The mode of transmission of electrical effects†

Marvin Charton*‡[a] and Barbara I. Charton[b]

[a] *Chemistry Department, Pratt Institute, Brooklyn, NY 11205, USA*
[b] *Library, St. Johns University, 8000 Utopia Parkway, Jamaica, NY 11439, USA*

Received (in Cambridge, UK) 23rd February 1999, Accepted 22nd July 1999

The dependence of substituent electrical effect transmission on substituent–reaction site distance and on the charge on reactant and product or transition state has been studied in the systems X–G–Y and X–Y where X is a variable substituent, Y a reaction site, and G a skeletal group. Reaction types studied were molecule–molecule (MM), molecule–ion (MI), and molecular ionization (Mi). MM reactions include proton transfer equilibria (pK_a's) of compounds with Y = CO_2H, OH, SO_2NH_2, NR_2H^+, azarenes, $PO_2(OH)^-$, and SH; gas phase ΔG_{acid} values for Y = CO_2H and OH, and proton affinities for NR_2H^+, proton transfer reaction rates for $XGCO_2H$ with Ph_2CN_2, and hydrogen bonding equilibria for XGCN (pK_{HB}). MI's include rates of base catalyzed ester hydrolysis, nucleophilic substitution of $PhCOCH_2Br$ by $XGCO_2^-$, and protodetritiation of T-substituted arenes. Mi reactions were solvolyses of XGCHLgMe (Lg is a leaving group) and $XGCMe_2Cl$. The measure of electrical effect magnitude used was *L*, the coefficient of the localized (field and/or inductive) effect obtained from correlation of appropriate data sets with linear free energy relationships. The substituent–reaction site distance was parameterized by *n*, the number of bonds between the substituent and the nearest atom of the reaction site undergoing bond change (Y^1). Correlations of L with $1/n^2$ and $1/n$; and of log $|L|$ with log *n* by simple linear regression analysis determined the dependence of *L* on *n*. Data sets with very large values of θ, the angle between the X–G bond and the line joining X and Y^1, were excluded. Data in aqueous–organic solvent mixtures can be combined into a single data set regardless of the solvent composition, probably due to preferential solvation by water. The results do not agree with the Kirkwood–Westheimer model for MI, Mi, and some MM reactions all of which show a dependence on $1/n$ rather than $1/n^2$. They support a modified field effect as the mode of transmission. This model differs from the Kirkwood–Westheimer model *m* seeming to depend on the charge difference between initial and final states.

Introduction

There has been considerable interest for many years in the manner of transmission of electrical effects. Electrical effect transmission may be defined as the process by which the electrical effect of a substituent X bonded to the *i*-th position of some skeletal group G reaches the reaction site Y bonded to the *j*-th position of G. Two modes of transmission had been proposed in the first half of this century, the inductive effect and the field effect. Both were first suggested by Derick.[1] The inductive effect involves the successive polarization of the bonds between X and Y. The initial electrical effect is decreased by passage through each bond by a fraction called the falloff factor, *f*, which is reported to have a value of 0.33–0.36.[1–3] The field effect is propagated directly through space, it is a function of the distance *r* between X and Y and of the cosine of the angle θ between the XG bond and the XY distance.[4,5] Classical electrostatics requires the dependence on cosine θ when the substituent is a dipole. Early attempts to distinguish between the two models were inconclusive.[3,6,7] Though there was a preference for the inductive effect as the major mode of transmission by some the field effect could not be excluded. Reynolds[8] reviewed the problem of the model of transmission some time ago. Bowden and Grubbs[9] have reviewed the evidence for the angular dependence of the electrical effect. This evidence supports the field effect as the mode of transmission. Exner and Friedl[10] have argued that overall the evidence is unclear and that electrical effect transmission may be accounted for by a modified form of the inductive effect. Bowden and Grubbs have recently summed up the evidence supporting the field effect model.[11] This model is based on the assumption that classical electrostatics is valid on the molecular scale. Its most successful variant is that of Kirkwood and Westheimer[5,12] (the KW model). According to the authors, "*Since the theory does not consider numerous factors (e.g. electrical saturation, electrostriction, the detailed structure of the solute molecules) it is necessarily only approximate and of restricted application*".[13] Bowden and his students have clearly demonstrated the importance of the angular dependence predicted by the KW model. One of us has described elsewhere[14] evidence that the inductive effect mode of transmission is not in accord with the observed dependence of electrical effect transmission on distance. It seemed to us that both classical models of transmission are flawed to some extent. Here we have investigated the transmission of the electrical effect for the widest range of reaction types for which we could find data reported. Our objective is twofold: 1. To determine the **actual** dependence of electrical effect transmission on distance. 2. To determine its dependence on the charge on reactants and on products.

Our first task is to choose an appropriate measure of the extent of transmission of electrical effects. In linear free energy relationships (LFER) for substituent effects the magnitude of the universal (polar/field/inductive/localized) electrical effect of a substituent X acting on an active site Y through an intervening skeletal group G on some measurable property *Q* is given by either ρ,[15] *C*, or *L* depending on the choice of the correlation

---

† Appendix I, supplementary Tables 1–7 and supplementary Figs. 1–7 are available as supplementary data. For direct electronic access see http//www.rsc.org/suppdata/p2/1999/2203, otherwise available from BLDSC (SUPPL. NO. 57618, pp. 30) or the RSC Library. See Instructions for Authors available *via* the RSC web page (http//www.rsc.org/authors).

‡ Parts of this work were presented at the Fourth Kyushu International Symposium on Physical Organic Chemistry, Fukuoka, 1993; and the Sixth Kyushu International Symposium on Physical Organic Chemistry, Fukuoka, 1995.

equation. ρ, C, or L are equivalent to each other. We have shown that it is possible to account for electrical effects in all types of systems in terms of the "pure" parameters σ_l, the localized (polar/field/inductive or universal) electrical effect constant; σ_d, the delocalized (resonance) electrical effect constant; and σ_e, the electronic demand electrical effect constant, the equation used is denoted the LDR eqn.[16] It is convenient to use as names of LFER the symbols representing the coefficients of the independent variables. L is the coefficient of σ_l in the L, LD and LDR equations. C is the coefficient of σ_C in the CR equation, ρ is the coefficient of the Hammett σ constants in the Hammett and the modified Yukawa–Tsuno equations. The σ_D, σ_C, and Hammett σ constants are composite parameters consisting of a combination of σ_d and σ_e; of σ_l and σ_d; and of σ_l and σ_d and σ_e respectively. All of these LFER can be derived from the LDR equation, the choice of correlation equation depends on the number of data points available in the data set. As set size decreases the number of independent variables in the correlation equation chosen must also decrease. All of these equations are described in Appendix I of the supplementary data.

The KW model categorizes reaction types as either monopole–dipole or monopole–monopole depending on the nature of the substituent. In the two types of reaction to which the model was applied, acid ionization[5] and basic ester hydrolysis,[12] the products are of the monopole–dipole type. We have preferred to classify the reaction types studied in this work in terms of the reactants, the categories are molecule–molecule (MM); molecule–ion (MI); and molecular ionization (Mi).

The dependence of *L* on distance

1. The dependence of *L* on distance for molecule–molecule reactions

The KW model predicts that the magnitude of L should be a linear function of $1/r^2$ for reactions between monopoles and dipoles. The ionization of substituted acids is considered to be of this type. It has been reported that ρ values can be estimated from a relationship derived from the Kirkwood–Westheimer equation and the Hammett equation.[17–19] The former is written as either eqn. (1*a*) or (1*b*), where e is the charge on an electron,

$$\log k_X/k_H = (e\mu_X \cos\theta)/(2.303RTD_E r^2) \qquad (1a)$$

$$\log k_X/k_H = (e\mu_X \cos\theta)/(2.303RTD_E r) \qquad (1b)$$

R the gas constant, T the temperature, D_E the effective relative permittivity, μ_X the bond moment of the X–G or X–Y bond, and r the distance from X to Y. Eqn. (1*a*) is derived for the case in which the substituent is a dipole and the active site in the product is a monopole, while eqn. (1*b*) is derived for the case in which both the substituent and the active site in the product are monopoles.

The equation derived for the calculation of ρ is[19] either eqn. (2*a*) or eqn. (2*b*) for the dipole–monopole and monopole–monopole cases respectively. The superscript ° designates values for the reference skeletal group (1,4-phenylene).

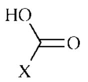
$$\rho_G = (\cos\theta_G {r^2}_{G^\circ}\rho_{G^\circ})/({r^2}_G \cos\theta_{G^\circ}) = (\cos\theta_G \hat{C})/{r^2}_G \qquad (2a)$$

$$\rho_G = (\cos\theta_G r_{G^\circ}\rho_{G^\circ})/(r_G \cos\theta_{G^\circ}) = (\cos\theta_G \hat{C}')/r_G \qquad (2b)$$

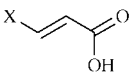
If we consider only skeletal groups G for which θ is less than about 45° then we can assume that $\cos\theta$ is approximately constant as it must lie between 0.707 and 1. Then eqns. (3) apply. As an appropriate approximate measure of distance we have taken n, the number of bonds on the shortest path between X and the nearest atom of Y which undergoes bond formation or cleavage (designated Y^1). Thus in the case of the carboxylic acids **1**, **2**, **3** and **4**, the n values are 2, 4, 3 and 3 respectively. For the azarenium ions, **5**, **6**, **7** and **8**, the n_r values are 2, 4, 8 and 10 respectively.

$$L = \hat{C}\,\overline{\cos}\,\theta/r^2 = \hat{C}'/r^2 \qquad (3a)$$

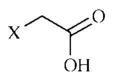
$$L = \hat{C}'\,\overline{\cos}\,\theta/r = \hat{C}'/r \qquad (3b)$$



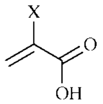
1



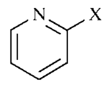
2



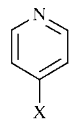
3



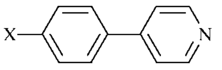
4



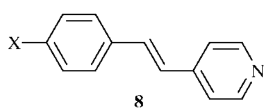
5



6



7



8

The classical inductive effect is given by eqn. (4), where Ç is a constant, f is the falloff factor, k is the number of paths between X and Y^1, and n is the number of bonds in each path.

$$L = \c{C}\sum_{i=1}^{k} f^n \qquad (4)$$

Values of L, C or ρ taken from the literature[16,20] or obtained from the appropriate correlations for carboxylic acid, hydroxy, *N*-hydroxy, ammonium, sulfonamide, azarenonium, and thiol data sets involving proton transfer equilibria in water, aqueous organic solvents, protic solvents, and the gas phase have been correlated with eqns. (5*a*) and (5*b*) by means of a simple linear regression analysis. Rate constants for carboxylic acid esterification with diphenyldiazomethane in ethanol, basic ester hydrolysis in water and in aqueous organic solvents, carbenium ion formation, reaction of carboxylate ion with phenacyl bromide, and protodetritiation were also correlated with these equations as were pK_{HB} values for nitriles. Further evidence can be obtained by writing eqns. (3*a*) and (3*b*) in logarithmic form, giving eqns. (6). Eqns. (6*a*) and (6*b*) require that the coefficient of $\log n$ be -2 and -1 respectively. Correlation of the L values with eqn. (7) in which b_o equals $\log\hat{C}$ is a method of determining the exponent m. The data generally fit eqns. (5*a*) and (5*b*) much better than they do eqn. (7). Two possible reasons for this behavior are: 1. A disadvantage of eqns. (5*a*) and (5*b*) is that L values for $n = 1$ or 2 have a disproportionate effect on the model. The data points for n greater than 5 tend to form a cluster. The coefficient $\hat{C}$ tends to be large, resulting in better fit. 2. The logarithmic nature of eqn. (7) results in a range for the independent variable of one order of magnitude when n varies from 1 to 10 whereas for the same range of n the

$$L = \hat{C}n^{-2} + a_o \qquad (5a)$$

$$L = \hat{C}n^{-1} + a_o \qquad (5b)$$

$$\log|L_{GY}| = -2\log n_{GY} + \log\hat{C} \qquad (6a)$$

$$\log|L_{GY}| = -\log n_{GY} + \log\hat{C} \qquad (6b)$$

$$\log|L| = m\log n + b_o \qquad (7)$$

Table 1 Examples of *L* value data sets used in correlations

CA1	*L*, $XGCO_2H$, pK_a, aq., 25 °C. G, *L*: 4-$PnOCH_2$, 0.317; C≡C, 2.40; *E*-2-Vn, 2.41; 4-Pn, 0.963; 4-$PnCH_2$, 0.557; none, 9.63; 4-Pn, 0.976; 4-$PnSCH_2$, 0.359; *E*-2-cPrn, 1.87; 4-$PnS(O)CH_2$, 0.221; 4-$PnS(O)_2CH_2$, 0.258; 4′-Pn-E-2-cPrn, 0.161; CH_2, 4.05; 1-Vn, 4.14; *E*-4-cHxn, 0.747; 5,2-Fn, 1.37; 5,2-Tn, 1.24; 5,2-Spn, 1.27; 5,2-Tpn, 1.18; 5.2-Prn, 1.40; *E*-4-Pn-2-Vn, 0.446; *E*-3-Pn-2-Vn, 0.439; 3-$PnOCH_2$, 0.273; 4-$PnOCH_2$, 0.352; 4-$PnSeCH_2$, 0.399; 3-$PnSCH_2$, 0.174; 3-$PnS(O)_2CH_2$, 0.243; 3-$PnSeCH_2$, 0.288; CHMe, 4.04; *E*-CX=CHPh, 4.25; *E*-CX=CHMe, 3.55; $CX{=}CMe_2$, 3.56; *E*-MeC=CXH, 2.58; *E*-CH=CXMe, 2.84; *E*-2-$VnCH_2$, 1.18
HA2	*L*, XGOH, aq., 25 °C. G, *L*: 4-Pn, 2.77; none, 16.4; 4-$PnCHCF_3$, 1.01; 4-$PnC(OH)CF_3$, 1.11; CH_2, 7.32; 4,1-Nn, 3.30; 3,1-Nn, 2.40; 4,2-Nn, 2.14; 2.3-Ttr, 3.65, 2-Pn, 4.06; 3-Pn, 2.30
BEH1	*L*, $XGCO_2Et$, log *k*, OH^-, aq. EtOH, 24–30 °C. G, *L*: 4,1-bc[2.2.2]Ocn, 2.19; 4,1-Nn, 2.37; 4′-Pn-4-Pn, 0.607; 5,2-Fn, 2.83; 5,2-Tn, 2.10; 5,2-Bfn, 1.52; 5,2-Idn, 1.08; 4-PnC≡C, 1.00; 3-Pn, 2.33; 4-$PnCH_2$, 1.14; 4-$PnCH_2$, 1.07; *E*-4′-Pn-2-Vn, 1.14; 4-$PnCH_2CH_2$, 0.642; none, 12.6; 5,3-Pyn, 2.23; *E*-4′-Pn-2-cHxn, 0.769; none, 10.5; 4-Pn, 2.31; 4-Pn, 2.33; 3-Pn, 2.39; 3,1-Nn, 2.18; 3-$PnCH_2CH_2$, 0.615; 4-$PnCH_2$, 1.04; 3-$PnCH_2$, 1.25; 7,2-Fln, 0.737
CF3	*L*, $XGCMe_2Cl$, log *k*, aq. MeAc, 25 °C. G, *L*: 3-Pn, 4.84; 4-Pn, 5.02; 4′-Pn-4-Pn, 1.69; 6,2-Nn, 3.52
CF4	*L*, G, CMe_2Cl, log *k*, 80% aq. EtOH, 25 °C. G, *L*: 6,2-Pyn, 4.07; 3-$PnCH_2$, 1.20, 4-$PnCH_2$, 1.22, CH_2CH_2, 3.16, none, 10.1

Abbreviations: Vn, vinylene; Pn, phenylene; Nn, naphthylene; Azn, Azulylene; Fln, fluorenylene; cPrn, cyclopropylene; cBun, cyclobutylene; cPen, cyclopentylene; cHxn, cyclohexylene; s[3.3]Hpn, spiro[3.3]heptylene; bc[2.2.1]Hpn, bicyclo[2.2.1]heptylene; bc[2.2.2]Oen, bicyclo[2.2.2]octenylene; bc[2.2.2]Ocn, bicyclo[2.2.2]octylene; Adn, adamantylene; Cbn, cubanylene; Fn, furylene; Tn, thiophenylene; Spn, selenophenylene; Tpn, tellurophenylene; Prn, pyrrolene; Pyn, pyridylene; Ozn, oxazolylene; Tzn, thiazolylene; Bfn, benzofurylene; Btn, benzothiophenylene; Pnn, pyrenylene; Idn, indolylene; Py, pyridyl; Qu, quinolyl; iQu, isoquinolyl. Absence of a skeletal group is indicated by "none". For all of the data sets studied see Table 1 of the supplementary data.

independent variable of eqn. (5) has a range of two orders of magnitude.

2. The dependence of *L* on distance for ion–molecule reactions

Westheimer and Shookhoff[12,13] have reported that for basic ester hydrolysis the field effect is given by eqn. (1*a*). Correlations were carried out with both eqns. (5*a*) and (5*b*).

3. The dependence of *L* on distance and solvent composition

In order to provide further information on the validity of eqns. (3*a*) and (3*b*) we have examined the use of eqns. (8*a*) and (8*b*).

$$L = \hat{C}n^{-2} + \dot{S}\varphi_m + a_o \qquad (8a)$$

$$L = \hat{C}n^{-1} + \dot{S}\varphi_m + a_o \qquad (8b)$$

These relationships permit the combination of data obtained in aqueous organic solvents varying over a wide range of compositions into a single data set. This results in data sets which are large enough to provide statistically significant results. They were obtained from eqns. (5*a*) and (5*b*) by the addition of a term in φ_m, the mole fraction of organic solvent in the mixture.[21] The φ_m values are given in Table 1 of the supplementary data.

Results

Examples of the available XGY data sets are given in Table 1, the rest are reported in Table 1 of the supplementary data. They have been correlated with eqns. (5*a*) and (5*b*), and (7); and also with (8*a*), and (8*b*) where applicable by means of either simple or multiple linear regression analysis as appropriate. Results of the best correlations with eqns. (5*a*) and (5*b*) are reported in Table 2, coefficients and their standard errors from correlations with eqn. (7) are reported in Table 3, the best results for correlations with eqns. (8*a*) and (8*b*) are given in Table 4. Complete results are reported in Tables 2, 3, and 4 of the supplementary data. Significant correlations have been obtained with eqns. (5*a*), (5*b*) and (7). Correlations with eqns. (8*a*) and (8*b*) show that there is generally no dependence on solvent composition. Plots of *L* against $1/n^2$ for set CA2 and against $1/n$ for sets BEH1 and CF1 are given in Figs. 1–3. Other plots are given in Figs. 1–6 of the supplementary data.

In order to determine whether the *L* values fit eqn. (9) and

$$L = \hat{C}n^m \qquad (9)$$

what the value of *m* is we have used the quantities defined below. For the comparison of correlations with eqns. (5*a*) and (5*b*) they are given by the quantities Δ^*, r_F, and Δ^o defined in eqns. (10), (11) and (12) respectively. Thus eqn. (10) applies,

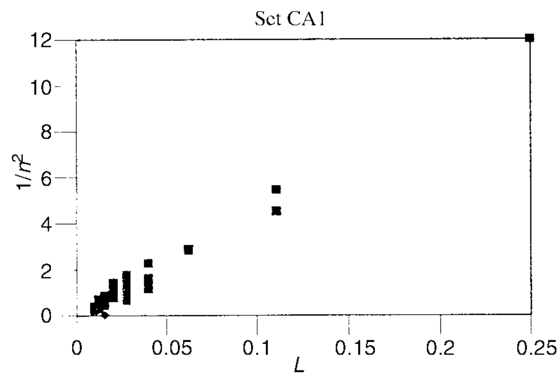


Fig. 1 *L* (abscissa) *vs.* $1/n^2$ (ordinate) for set CA1, carboxylic acids in water at 25 °C.

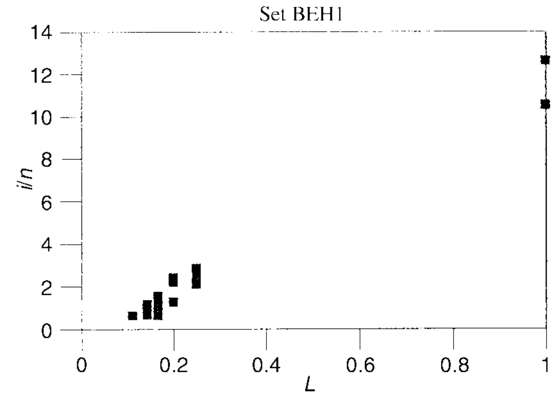


Fig. 2 *L* (abscissa) *vs.* *i*/*n* (ordinate) for set BEH1, the reaction of ethyl carboxylates with hydroxide ion in aqueous EtOH at 25 °C.

$$\Delta^* = \Delta 100r^2 \equiv (100r^2)_2 - (100r^2)_1 \qquad (10)$$

$$r_F = F_2/F_1 \qquad (11)$$

$$\Delta^o = \Delta S^o \equiv S_2{}^o - S_1{}^o \qquad (12)$$

where the subscripts 1 and 2 indicate the values obtained for correlations with eqns. (5*b*) and (5*a*) respectively. $100r^2$ is the percentage of the data accounted for by the regression equation. Then if $\Delta^* > 1.50$ $m = 2$, if $\Delta^* < -1.50$ $m = 1$, if $1.50 > \Delta^* > -1.50$ then *m* is uncertain.

Table 2 Results of best correlations with eqns. (5*a*) and (5*b*)

Set	Y	Solv.	$\hat{C}$	s_C	a^o	s_{ao}	$100r^2$	*F*	S_{est}	S^o	*N*
CA1	CO_2H	W	39.5	0.843	−0.319	0.0704	98.25	2195	0.24	0.25	35
CA2	CO_2H	AE	46.7	1.22	−0.123^n	0.107	96.91	1472	0.33	0.18	49
CA3	CO_2H	AMCS	46.4	1.45	0.0063^n	0.114	98.84	1019	0.34	0.117	14
CA4	CO_2H	AM	30.3	1.28	0.208^n	0.114	98.28	562.8	0.29	0.145	12
CA5	CO_2H	M	35.7	2.07	−0.018^n	0.2	97.69	296.4	0.44	0.172	9
CA6	CO_2H	E	46	4.64	−0.144^n	0.644	98	98.14	0.82	0.2	4
CA7	CO_2H	DMSO	69	7.86	−0.149^n	1.09	97.47	76.95	1.39	0.225	4
CA8	CO_2H	AD	41.8	2.11	−0.078^n	0.117	95.61	392	0.33	0.221	20
CA9	CO_2H	ABCS	*13.3*	*1.93*	*−0.935*	*0.301*	*92.19*	*47.25*	*0.19*	*0.342*	*6*
CA10	CO_2H	g	153	16.1	7.93	1.75	94.75	90.27	1.22	0.271	7
CA11	CO_2H	DMF	56	4.81	0.319^n	0.249	91.13	135.4	0.38	0.207	6
CA12	CO_2H	AN	48.9	2.15	0.384^n	0.268	99.42	515.3	0.42	0.098	5
HA1	OH	AE	*18.6*	*3.15*	*−0.633*^{n}	*0.529*	*68.45*	*34.71*	*0.76*	*0.596*	*18*
HA2	OH	W	*18.0*	*0.888*	*−1.74*	*0.359*	*97.85*	*409.5*	*0.68*	*0.162*	*11*
HA3	OH	g	38.7	10	23.7	4.77	83.18	14.84	7.88	0.529	5
HA4	CZ=NOH	W	47.1	3.63	−0.102^n	0.272	97.12	168.5	0.43	0.201	7
PA1	$PO(OH)O^-$	W	*7.73*	*0.078*	*−0.11*	*0.0446*	*99.98*	*9899*	*0.05*	*0.02*	*4*
AB1	NH_3^+	W	23.7	1.19	1.66	0.334	97.07	397.8	1.09	0.185	14
AB2	NH_3^+	AE	*31.5*	*0.566*	*−2.65*	*0.18*	*99.68*	*3111*	*0.48*	*0.062*	*12*
AB3	NH_2	g	*85.2*	*5.59*	*−8.85*	*3.03*	*98.72*	*232.3*	*3.63*	*0.146*	*5*
AB4	NMe_2H^+	W	22.7	1.66	3.14	0.772	98.42	187.4	1.34	0.162	5
			28.2	*2.42*	*−3.09*	*1.31*	*97.84*	*135.7*	*1.57*	*0.19*	*5*
SA1	SO_2NH_2	W	25	0.855	0.0083^n	0.0912	99.42	853.1	0.18	0.09	7
AzB1	=NH$^+$–	W	*24.2*	*1.36*	*−1.97*	*0.421*	*94.64*	*318*	*0.8*	*0.244*	*20*
TA1	SH	W	*19.9*	*1.5*	*−2.25*	*0.81*	*98.33*	*176.5*	*0.97*	*0.167*	*5*
HB1	CN	CCl_4	6.92	0.936	0.175^n	0.107	93.19	54.7	0.19	0.32	6
BEH1	CO_2Et	AE	*12.4*	*0.422*	*−0.771*	*0.143*	*97.39*	*859.9*	*0.47*	*0.168*	*25*
BEH2	CO_2Et	W	*12.0*	*1.09*	*−1.68*	*0.557*	*96.82*	*121.8*	*0.716*	*0.218*	*6*
BEH3	CO_2Et	AA	*6.80*	*1.54*	*0.805*^{n}	*0.674*	*79.63*	*19.55*	*1.11*	*0.534*	*7*
BEH4	CO_2Et	AD	*8.19*	*1.59*	*0.42*^{n}	*0.452*	*89.84*	*26.52*	*0.47*	*0.412*	*5*
BEH5	CO_2Me	AD	*14.8*	*3.33*	*−0.62*^{n}	*0.678*	*86.88*	*19.87*	*0.569*	*0.468*	*5*
BEH6	CO_2Me	AM	*7.04*	*2*	*0.472*^{n}	*0.454*	*67.45*	*12.43*	*0.21*	*0.659*	*8*
CD1	CO_2H	E	19.1	0.925	0.132^n	0.0723	97.27	427.7	0.21	0.178	14
			11.7	*0.494*	*−1.25*	*0.11*	*97.91*	*563.3*	*0.18*	*0.156*	*14*
CF1	CHLgMe	AE	*29.3*	*5.38*	*−0.416*^{n}	*1.18*	*83.16*	*29.63*	*0.54*	*0.474*	*8*
CF2	CHClMe	AA	*21.1*	*5.6*	*−0.196*^{n}	*1.13*	*66.98*	*14.2*	*0.79*	*0.652*	*9*
CF3	CMe_2Cl	AA	*22.2*	*8.07*	*−0.144*^{n}	*1.48*	*79.13*	*7.582*	*0.86*	*0.646*	*4*
CF4	CMe_2Cl	AE	*10.2*	*1.52*	*−0.034*^{n}	*0.755*	*93.81*	*45.44*	*1.05*	*0.321*	*5*
PD1	=CT −	TFA	28.3	6.49	2.54	1.11	79.1	18.93	1.73	0.541	7
			18.9	*4.26*	*0.0426*^{n}	*1.58*	*79.69*	*19.62*	*1.71*	*0.533*	*7*
NS1	CO_2^-	AA	*3.63*	*0.373*	*−0.139*^{n}	*0.0929*	*94.04*	*94.63*	*0.13*	*0.282*	*8*

Lg, leaving group. Values in italics are for correlations with eqn. (11), other values are for correlations with eqn. (9). Abbreviations. Set: CA, Carboxy acidity; HA, hydroxy acidity; PA, phosphonate acidity; AB, amine basicity; SA, sulfonamide acidity; AzB, azarene basicity; TA, thiol acidity; HB, hydrogen bonding; BEH, basic ester hydrolysis; CD, Carboxy–diphenyldiazomethane reactivity; CF, carbenium formation; PD, protodetritiation; NS, nucleophilic substitution. Solvent: W, water; AE, aq. ethanol; AMCS, aq. methyl Cellosolve; AM, aq. methanol; M, methanol; E, ethanol; DMSO, dimethyl sulfoxide; ABCS, aq. butyl Cellosolve; AA, aq. acetone; AD, aq. dioxane; g, gas phase; DMF, dimethylformamide; AN; acetonitrile; TFA, trifluoroacetic acid. Compositions of aqueous organic solvents are variable unless otherwise noted. Superscript n, not significant. Aqueous solvents are of varying composition unless otherwise noted. a. 26–29 mol% AE. b. 30.9 mol% AM. c. 23.6 mol% AE. d. 68.9–69.6 mol% AE. Statistics: $100r^2$, $100R^2$, percentage of the data accounted for by the regression equation; A$100R^2$, percentage of the data accounted for by the regression equation adjusted for the number of independent variables, *F*, a test for the significance of the regression coefficients and a measure of the goodness of fit; S_a, S_{ao}, standard errors of the regression coefficients; S_{est}, standard error of the estimate; S^o, standard error of the estimate divided by the root mean square of the data. Statistics for all correlations with eqns. (5*a*) and (5*b*) are given in Table 2 of the supplementary data.

In eqn. (11) the subscripts 1 and 2 have the same meaning as that in eqn. (10). *F* is a test for the significance of the regression coefficients. Then if $r_F > 5/4$ $m = 2$, if $r_F < 4/5$ $m = 1$, if $5/4 > r_F > 4/5$ *m* is uncertain.

In eqn. (12) again the subscripts 1 and 2 have the same meaning as that in eqn. (10). S^o is the standard error of the estimate divided by the root mean square of the data. Then if $\Delta^o > 0.0500$ $m = 1$, if $\Delta^o < -0.0500$ $m = 2$, if $0.0500 > \Delta^o > -0.500$ *m* is uncertain.

In interpreting the results of correlations with eqn. (7) we have made use of the quantity r_t, the ratio of the *t* tests for the difference between 2 and *m* and 1 and *m*. S_m is the standard error of the coefficient *m*. Thus:

$$t_2 = |2 - m|/S_m \qquad (13a)$$

$$t_1 = |1 - m|/S_m \qquad (13b)$$

$$r_t \equiv t_2/t_1 = |2 - m|/|1 - m| \qquad (14)$$

Then if $r_t > 1.50$ $m = 1$, if $r_t < 0.667$ $m = 2$, if $2.00 > r_t > 0.500$ *m* is uncertain.

A comparison of log $\hat{C}$ obtained from eqns. (5*a*) and (5*b*) with b_o from eqn. (7) can also be used as a test for the value of *m*. $\hat{C}$ is the coefficient of the independent variable in eqns. (5*a*) and (5*b*) while b_o is the intercept of eqn. (7). The ratio of the *t* tests for the difference between log $\hat{C}$ and b_o, r_{tc}, is given by eqns. (15) and (16).

$$t_2 = |\log \hat{C}_2 - b_o|/S_{bo} \qquad (15a)$$

$$t_1 = |\log \hat{C}_1 - b_o|/S_{bo} \qquad (15b)$$

$$r_{tc} = t_2/t_1 = |\log \hat{C}_2 - b_o|/|\log \hat{C}_1 - b_o| \qquad (16)$$

The subscripts 1 and 2 designate $\hat{C}$ values obtained from eqns. (5*b*) and (5*a*) respectively. If $r_{tc} < 0.2$ then *m* is −2; if $0.2 < r_{tc} < 5$, then *m* is indeterminate; if r_{tc} is > than 5, $m = -1$.

Table 3 Coefficients and their standard errors for correlations with eqn. (7)

Set	m	s_m	b^o	S_{bo}	Set	m	s_m	b^o	S_{bo}
CA1	−2.75	0.141	1.96	0.103	AB3	−1.36	0.166	1.9	0.0801
CA2	−2.15	0.129	1.72	0.101	AB4	−1.22	0.215	1.37	0.104
CA3	−2.17	0.178	1.79	0.144	SA1	−2.06	0.217	1.43	0.148
CA4	−2.1	0.149	1.6	0.112	A_ZB1	−1.65	0.108	1.55	0.0694
CA5	−1.79	0.333	1.42	0.209	TA1	−1.45	0.178	1.28	0.0859
CA6	−1.79	0.264	1.54	0.157	HB1	−1.3	0.297	0.595	0.2
CA7	−1.73	0.289	1.70	0.172	BEH1	−1.37	0.11	1.14	0.0782
CA8	−2.17	0.22	1.7	0.166	BEH2	−1.4	0.253	0.997	0.122
CA9	−2.05	0.352	1.68	0.298	BEH3	−2.23	0.646	1.76	0.42
CA10	−1.07	0.255	1.91	0.167	BEH4	−1.13	0.259	1.07	0.117
CA11	−1.91	0.271	1.75	0.206	BEH5	−1.67	0.559	1.49	0.422
CA12	−1.83	0.344	1.61	0.218	BEH6	−0.836	0.198	0.855	0.131
HA1	−1.83	0.228	1.70	0.187	CD1	−1.91	0.17	1.29	0.126
HA2	−1.4	0.171	1.25	0.101	CF1	−1.12	0.193	1.52	0.13
HA3	−0.703	0.135	1.79	0.0611	CF2	−1.06	0.233	1.34	0.17
HA4	−2.08	0.109	1.69	0.0785	CF3	−1.28	0.423	1.53	0.333
PA1	−1.05	0.0151	0.887	0.0082	CF4	−1.16	0.294	1.05	0.171
AB1	−1.7	0.261	1.49	0.164	PD1	−0.945	0.153	1.25	0.0871
AB2	−1.89	0.225	1.68	0.196	NS1	−1.4	0.18	0.705	0.18

Complete results of the correlations are given in Table 3 of the supplementary data.

Table 4 Best results for correlations with eqns. (8*a*) and (8*b*)

Set	Y	Solvent	C	S_C	$\acute{S}$	$S_{\acute{S}}$	a_o
CA2	CO_2H	AE	46	1.26	0.647	0.342	−0.286
CA4	CO_2H	AM	30	1.35	−0.356	0.385	0.355
			19	0.653	1.43	0.319	−2.59
CA8	CO_2H	AD	41.5	1.97	1.03	0.531	−0.297
HA1	OH	AE	21	3.96	−1.03	1.01	−0.642
AB2	NH_3^+	AE	31.5	0.573	1.27	1.41	−2.93
Set	S_{ao}	$100R^2$	$A100R^2$	F	S_{est}	S^o	N
CA2	0.107	97.13	97.07	778.1	0.321	0.175	49
CA4	0.196	98.41	98.25	277.8	0.294	0.146	12
	0.227	99.06	98.97	473.8	0.226	0.112	12
CA8	0.156	96.41	96.21	228.2	0.307	0.206	20
HA1	0.528	70.47	68.63	17.90	0.761	0.595	18
AB2	0.357	99.71	99.68	1527	0.483	0.0626	12

For abbreviations see Table 1. For complete results of correlations see Table 4 of the supplementary data.

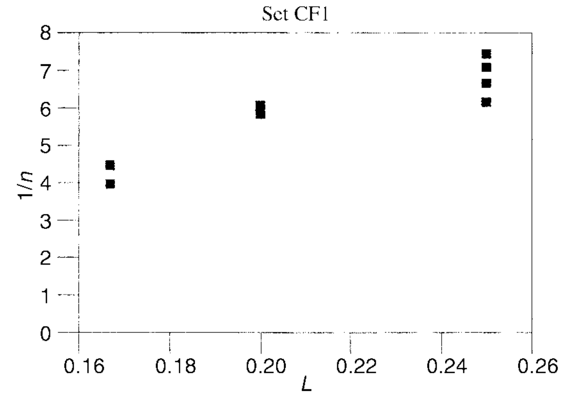


Fig. 3 L (abscissa) *vs.* $1/n$ (ordinate) for set CF1, the solvolysis of XGCHLgMe, Lg = Cl or OpNP, in 80% aqueous ethanol at 25 °C.

Values of log $\hat{C}_1$, log $\hat{C}_2$, b_o, r_{tc} and m are given in Table 5 of the supplementary data.

Finally, in evaluating the results it is also necessary to consider the range, Δn, of n values encompassed by the data set given by the difference between the maximum and minimum values, n_{mx} and n_{mn} respectively, of n. This should be as large as possible, certainly not less than four units, with an n_{mn} value of three or less. The number of data points, N, in the set is also an important factor as this determines the number of degrees of freedom in the regression analysis. In correlations with eqns. (5*a*), (5*b*) and (7) N should be at least 5, for correlations with eqns. (17*a*), (17*b*) and (18) it should be at least 8. Values of $\varDelta^*$, r_F, $\varDelta^o$, r_t, r_{tc}, Δn, n_{mn} and N are reported in Table 5.

Molecule–molecule reactions

1. Proton transfer equilibria. The KW model considers the ionization of substituted acids to be monopole–dipole reactions. We consider them to be molecule–molecule reactions. Data sets representing ionization constants of XGY with Y equal to CO_2H, OH, MZ=NOH, $P(OH)O_2^-$, MH_3^+, NMe_2H^+, SO_2NH_2, $-NH^+=$, and SH have been studied. The most extensively studied are the carboxylic acids (sets CA1–CA12) which include data determined in protic solvents including water, methanol and ethanol, and aqueous–organic solvents; dipolar aprotic solvents including dimethylformamide, acetonitrile, and dimethyl sulfoxide; and the gas phase. On the basis of the method of evaluation described above all but sets CA4 and CA9 are best fit by eqn. (5*b*) with $m = -2$, the value of m for these two sets is uncertain. As set CA9 has n_{mn} equal to 5 it can be disregarded. Even with the inclusion of all CA data sets the mean value of m for the ionization of carboxylic acids in various media (sets CA 1–12) is -1.96 ± 0.387. Clearly, $XGCO_2H$ data sets fit eqn. (5*a*) best, as expected from the KW model. The *N*-hydroxy and sulfonamido data sets (HA4 and SA1) also seem to obey eqn. (5*a*). For the *C*-hydroxy, amino and thiol data sets (sets HA1–3, AB1–4 and TA1) the best fit is either with eqn. (5*b*) with $m = -1$ or no discrimination between eqns. (5*a*) and (5*b*) is possible. With the inclusion of all of these data sets the mean value of m is

Table 5 Values of Δ^*, r_F, Δ^o, r_t, r_{tc}, n_{mn}, Δn, N, and assigned m

Set	Δ^*	r_F	Δ^o	r_t	r_{tc}	n_{mn}	Δn	N	m
CA1	4.00(2)	3.86(2)	−0.116(2)	0.422(2)	0.590(U)	2	6	31	−2
CA2	10.09(2)	5.09(2)	−0.192(2)	0.130(2)	0.118(2)	2	8	49	−2
CA3	5.34(2)	5.90(2)	−0.158(2)	0.145(2)	0.316(U)	2	10	14	−2
CA4	−0.78(U)	0.578(1)	0.033(U)	0.0909(2)	0.343(U)	2	10	12	U
CA5	6.25(2)	3.96(2)	−0.160(2)	0.370(2)	1.63(U)	2	6	9	−2
CA6	7.09(2)	3.74(2)	−0.177(2)	0.266(2)	2.00(U)	2	4	4	−2
CA7	5.51(2)	3.36(2)	−0.176(2)	0.370(2)	2.80(U)	2	4	4	−2
CA8	1.45(U)	1.35(2)	−0.034(U)	0.145(2)	0.195(2)	3	5	20	−2
CA9	−1.39(U)	0.833(U)	0.029(U)	0.0476(2)	0.0893(2)	5	5	6	−2
CA10	4.20(2)	4.16(2)	−0.103(2)	9.00(1)	3.50(U)	2	4	7	−2
CA11	9.10(2)	4.60(2)	−0.217(2)	0.0989(2)	0.01111(2)	3	5	6	−2
CA12	1.74(2)	4.08(2)	−0.0984(2)	0.205(2)	0.727(U)	2	5	5	−2
HA1	−4.78(1)	0.775(1)	0.043(U)	0.205(2)	0.116(2)	4	8	18	U
HA2	−1.98(1)	0.510(1)	0.063(1)	1.50(1)	7.00(1)	1	5	11	−1
HA3	−8.38(1)	0.667(1)	0.094(1)	3.87(1)	1.86(U)	1	4	6	−1
HA4	1.18(U)	1.43(2)	−0.037(U)	0.0741(2)	0.0116(2)	3	4	7	−2
PA1	−3.50(1)	0.006(1)	0.245(1)	19.0(1)	81.0(1)	2	4	4	−1
AB1	0.37(U)	0.885(U)	−0.011(U)	0.429(2)	2.20(U)	1	6	14	U
AB2	−1.69(1)	0.157(1)	0.093(1)	0.124(2)	1.28(U)	1	10	12	−1
AB3	−2.72(1)	0.306(1)	0.114(1)	1.78(1)	2.33(U)	1	4	5	−1
AB4	0.58(U)	0.725(1)	0.028(U)	3.55(1)	0.125(2)	1	4	5	−1
AzB1	−1.34(U)	0.787(1)	0.029(U)	0.538(2)	0.188(2)	2	8	20	−2
SA1	1.01(U)	2.75(2)	−0.059(2)	0.0566(2)	0.143(2)	1	5	7	−2
TA1	−4.55(1)	0.256(1)	0.155(1)	1.22(U)	4.50(U)	1	4	5	−1
HB1	6.26(2)	2.06(2)	−0.073(2)	1.43(U)	6.13(1)	2	7	6	−2
BEH1	−2.28(1)	0.521(1)	0.063(1)	1.70(1)	2.40(U)	1	8	25	−1
BEH2	−1.64(1)	0.649(1)	0.057(1)	1.50(U)	0.232(U)	1	5	6	−1
BEH3	−8.00(1)	0.645(1)	0.094(1)	49.0(1)	0.256(U)	3	3	6	−1
BEH4	−11.67(1)	0.405{1}	0.191(1)	6.69(1)	0.600(U)	2	7	5	−1
BEH5	−8.05(1)	0.562(1)	0.126(1)	0.818(U)	<0.001(2)	3	6	7	−1
BEH6	−5.40(1)	0.787(1)	0.052(1)	5.03(1)	51.4(1)	3	5	12	−1
CD1	0.64(U)	0.758(1)	0.022(U)	0.0989(2)	0.0455(2)	2	6	14	−2
PD1	−0.59(U)	0.962(U)	0.008(U)	24.0(1)	7.83(1)	2	6	8	−1
NS1	−7.07(1)	0.424(1)	0.135(1)	1.50(U)	0.386(U)	1	8	8	−1
CF11	1.03(U)	0.870(U)	−0.020(U)	15.7(1)	7.83(1)	4	5	16	−1
CF12	−6.78(1)	0.562(1)	0.112(1)	4.56(1)	3.50(U)	1	8	9	−1

Numbers in parentheses are assigned values of $-m$, U indicates uncertain. m values in the column headed m are probable.

-1.44 ± 0.396. *This result is not expected.* The KW equation predicts $m = -2$ for all of these reactions.

For the ionization of azaarene basicities m is likely to be -2.

2. Proton transfer rate. Rate constants for the reaction of diphenyldiazomethane with $XGCO_2H$ seem to have an m value of -2 in agreement with the KW model.

3. Hydrogen bonding equilibria. pK_{HB} values for hydrogen bond complex formation of 4-fluorophenol with XGCN have an m value of -2, again in agreement with the KW model.

Ion–molecule reactions

1. Base catalyzed ester hydrolysis. We have first correlated the subset of set BEH1 for which φ_m is 68.9 and 69.6 and then all of set BEH1 with eqns. (5*a*), (5*b*) and (7). As the results were about the same we then assumed that as was the case for proton transfer reactions L in aqueous organic solvents is independent of φ_m. The six data sets for base catalyzed hydrolysis of $XGCO_2Et$ (BEH1–4) and $XGCO_2Me$ (BEH5–6) are all best fit by eqn. (5*b*) (m is equal to -1). Although set BEH3 is unreliable due to its small range the remaining sets are reliable and the result is clear. With the inclusion of all these data sets (BEH1–6) the mean value of m from correlations with eqn. (7) is -1.42 ± 0.469. This result disagrees with the extension of the KW model by Westheimer and Shookhoff which considered these reactions to be of the monopole–dipole type for which m is -2. *Again, this result is unexpected.*

2. Other ion–molecule reactions. The reaction of $XGCO_2^-$ with phenacyl bromide has m equal to -1, it is best fit by eqn. (5*b*). This is also true of ionization constants for $XGP(OH)O_2^-$ (set PA1) but the result is unreliable due to the small size of the data set. The protodetritiation of XArT (set PD1) seems to give a value of m of -1. As the reaction involves rate determining transfer of T^+ from a positive ion it belongs in the ion–molecule category.

Molecular ionization

As the transition state for this type of reaction involves bond breaking but no bond forming it does not fit into the two categories above. The KW model would presumably classify these reactions as monopole–dipole types which would be expected to obey eqn. (5*a*). The reaction sites studied are CHLgMe where Lg indicates a leaving group (sets CF1, CF2); and CMe_2Cl (sets CF3, CF4). For sets CF1–4 the mean value of m obtained from correlations with eqn. (7) is -1.16 ± 0.0929. For each reaction site data sets are available in aqueous ethanol and in aqueous acetone. In order to enlarge the data sets and improve the reliability of the results we have combined the data sets for each reaction site into a single set by introducing an indicator variable for the effect of solvent. The correlation equations then become eqns. (17) and (18), in which

$$L = \hat{C}n^{-2} + sSv + a_o \qquad (17a)$$

$$L = \hat{C}n^{-1} + sSv + a_o \qquad (17b)$$

$$\log|L| = m \log n + sSv + b_o \qquad (18)$$

Sv takes the value 1 for one solvent and 0 for the other. Best results of correlations with eqns. (17*a*), (17*b*) and (18) are pre-

Table 6 Best results of correlations with eqns. (17*a*), (17*b*) and (18)

Set	$\hat{C}/m$	S_a/S_m	s	S_s	a_o/b_o	S_{ao}/S_{bo}	$100R^2$	$A100R^2$	F	S_{est}	S^o	N
CF11	60.9	6.88	1.77	0.238	1.24	0.327	92.8	92.28	83.76	0.467	0.298	16
	23.7	*2.9*	*1.72*	*0.256*	*−0.92*	*0.596*	*91.77*	*91.19*	*72.52*	*0.499*	*0.318*	*16*
	−1.06	**0.11**	**0.156**	**0.023**	**1.32**	**0.081**	**93.47**	**93.01**	**93.08**	**0.044**	**0.283**	**16**
CF12	*10.5*	*1.47*	*2.06*	*0.76*	*−0.14*	*0.736*	*89.46*	*87.96*	*25.47*	*1.03*	*0.398*	*9*
	−1.18	**0.215**	**0.395**	**0.112**	**1.06**	**0.127**	**83.68**	**81.35**	**15.39**	**0.144**	**0.495**	**9**

Results in ordinary type are for correlations with eqn. (17*a*), those in italics are for correlations with eqn. (17*b*), those in bold are for correlations with eqn. (18). For complete results see Table 7 of the supplementary data.

sented in Table 6, complete results in Table 6 of the supplementary data. The results for the XGCHLgMe indicate an m value of -1 though some uncertainty remains, perhaps due to n_{mn} having a value of 4 for the combined set (set CF11). The results for the XGCMe$_2$Cl show clearly that m is -1. The mean value of m for sets CF11–12 from correlations with eqn. (17*b*) is -1.12 ± 0.06. *Clearly, the results for molecular ionization are unexpected.* The KW model predicts m equal to -2 for all monopole–dipole reactions, thus it must also predict m equal to -2 for the Mi reactions studied here. It would be interesting to have results for XGCHLg with an ionic leaving group.

The validity of eqn. (7)

Eqn. (7) is based on the assumption that the mode of transmission of electrical effects is given by eqn. (9). If this is so then there should be no significant difference between b_o from eqns. (7) or (18) and log $\hat{C}$ from eqns. (5*a*), (5*b*), (8*a*), (8*b*), (17*a*) and (17*b*). This is indeed the case, providing additional justification for the validity of eqn. (7).

Discussion

The mode of transmission of electrical effects

The results we have obtained are in accord with a field effect model of some kind. They involve studies of a fairly wide range of reactions under a wide range of conditions. The classical inductive effect cannot possibly account for them. When our results are combined with the studies of Bowden and co-workers[9,11] on the effect of the angle, the NMR studies of Adcock and co-workers,[22] and our own previous results[14] we can only conclude that the mode of transmission of electrical effects is some form of the field effect other than the KW model. The original KW model considered only monopole–dipole reactions to which all proton transfer reactions of neutral acids and bases as well as those of esters with hydroxide ion were supposed to belong. Our results show that the proton transfer reaction of C-hydroxy, amino, and thiol compounds are usually best fit by eqn. (5*b*). The reactions studied and the value of m assigned to them are given in Table 7. Let us consider the magnitude and the degree of concentration of the charges on the initial and final reaction sites, Y_i and Y_f respectively, in the reactions studied. Y_i is the reaction site in the reactant and Y_f that in the product or transition state. We can classify charges as either concentrated or diffuse. A concentrated charge has more than half of the total positive or negative charge localized on a single atom of the reaction site. A diffuse charge has most of the charge divided between two or three atoms of the reaction site. The charge on the reaction site can further be qualified as large ($>0.50e$), intermediate ($0.25e < q < 0.50e$), or small ($<0.25e$); where e is the magnitude of the charge on the electron. We can use this system to assign empirically charges q_i and q_f to the reaction sites and crudely estimate the magnitude of Δq, the difference between the initial and final charges. The q_i, q_f and Δq values are also reported in Table 7. In Table 7 of the supplementary data are given the number of data sets having a particular value of Δq for each value of m. Those for which m is equal to -1 generally have a large concentrated charge difference on the reacting atom. Those for which m is equal to -2 have a diffuse charge difference as the result of delocalization. Those which have uncertain m values show a range of Δq values. Fig. 4 illustrates the point. As an alternate approach we have assigned numerical values to the Δq estimates with S, M and L equal to 1, 2 and 3 respectively for a concentrated and 0, 1 and 2 for a diffuse charge. $M - L$ is thus equal to 2.5 for a concentrated and 1.5 for a diffuse charge on this scale. Correlation of the observed m values for the 21 data sets that both met the requirements for reliability and gave results regarded as probably correct for m gave on exclusion of four outliers (sets, CA1, HB1, AB2, and BEH6) the eqn. (19), with $100r^2$, 64.49; F,

$$m = 0.869(\pm 0.167)\Delta q - 3.68(\pm 0.403) \qquad (19)$$

27.24; S_{est}, 0.277; S^o, 0.634; N, 17. The regression equation is statistically highly significant, particularly in view of the very approximate numerical values assigned to the charges S, M, L and ML. A plot of $-m$ against assigned Δq values is given in Fig. 10 of the supplementary data.

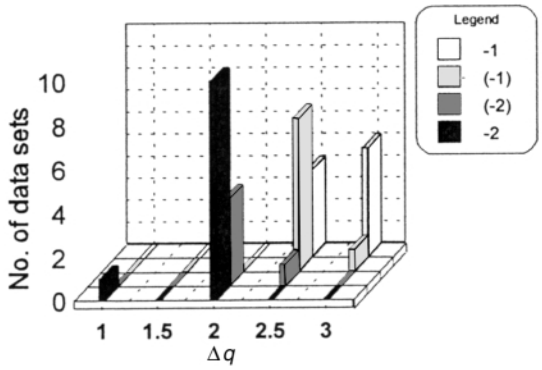


Fig. 4 The variation of N_{ds}, the number of data sets having a given m value, with Δq. N_{ds}, (abscissa) *vs.* Δq (ordinate). m values in parentheses are uncertain, others are probable.

The results support our proposal that m is dependent on Δq but are not sufficient to prove it conclusively. Furthermore, although we have assumed that m is either -1 or -2 we cannot exclude the possibility that m varies continually between 0.5 and 3.0, depending on the value of Δq which certainly varies from 0 to 1. We will refer to this model in which transmission depends on the reaction site difference in charge between reactant and product (or transition state) as the modified field effect (MFE) model.

A field effect is characterized by three properties: 1. A dependence on cos θ; 2. A dependence on r^{-m}; 3. A dependence on substituent and reaction site charge in product (KW) or charge difference between reactant and product/transition state (MFE). Bowden and his students have provided evidence for 1, in this work we have provided evidence for 2 and 3. In view of the fact that the classical inductive effect is independent of charge or charge difference it is difficult to see how it can account for these results.

Table 7 *q* and Δq as a function of Y

Sets	Y_i	Reagent	Y_f	Product	Q	q_i	q_f	Δq	*m*
HA2	OH	H_2O	O^-	H_3O^+	pK_a	0	c	L	−1
HA3	OH		O^-		Δg	0	c	L	−1
AB2	NH_2	H_2O	NH_3^+	OH^-	pK_{ca}	0	c	L	−1
AB3	NH_2		NH_3^+		PA	0	c	L	−1
TA1	SH	H_2O	S^-	H_3O^+	pK_a	0	c	L	−1
BEH1,2	CO_2Et		$C(OEt)(OH)^{\delta-}O^{\delta-}$		log *k*	c	d	M–L	−1
BEH6	CO_2Me		$C(OMe)(OH)^{\delta-}O^{\delta-}$		log *k*	c	d	M–L	−1
CF12	CMe_2Cl		$CMe_2^{\delta+}Cl^{\delta-}$		log *k*	0	c	M–L	−1
AB4	NMe_2	H_2O	NMe_2H^+	OH^-	pK_{ca}	0	c	L	*−1*
PA1	$P(OH)O_2^-$	H_2O	PO_3^{2-}	H_3O^+	pK_a	d	Ld	M–L	*−1*
BEH3	CO_2Et		$C(OEt)(OH)^{\delta-}O^{\delta-}$		log *k*	c	d	M–L	*−1*
BEH4	CO_2Et		$C(OEt)(OH)^{\delta-}O^{\delta-}$		log *k*	c	d	M–L	*−1*
BEH5	CO_2Me		$C(OMe)(OH)^{\delta-}O^{\delta-}$		log *k*	c	d	M–L	*−1*
NS1	CO_2^-	$BzCH_2Br$	$O^{\delta\delta-}COCH_2BzBr^-$		log *k*	d	c	M–L	*−1*
PD1	$(–CHT–)^+$	$CF_3CO_2^-$	$(–CH–)^{\delta+}TOC(CF_3)O^{\delta-}$		log *k*	sd	sd	M–L	*−1*
CF11	ArCHLgMe		$ArCHNe^{\delta+}Lg^{\delta-}$		log *k*	sc	sc	M–L	*−1*
HA1	OH	H_2O	O^-	H_3O^+	pK_a	c	c	L	U
AB1	NH_2	H_2O	NH_3^+	OH^-	pK_{ca}	c	c	L	U
CA4	CO_2H	H_2O	CO_2^-	H_3O^+	pK_a	d	d	M	U
CA6	CO_2H	HSv	CO_2^-	H_3Sv^+	pK_a	d	d	M	*−2*
CA7	CO_2H	Sv	CO_2^-	HSv^+	pK_a	d	d	M	*−2*
CA9	CO_2H	H_2O	CO_2^-	H_3O^+	pK_a	d	d	M	*−2*
CD1	CO_2H	Ph_2CN_2	$O^{\delta-}COHCPh_2N_2^{\delta+}$		log *k*	d	d	M	*−2*
AzB1	–N=	H_2O	$–NH^+=$	OH^-	pK_{ca}	d–c	d–c	M–L	*−2*
CA,1,2,3,8	CO_2H	H_2O	CO_2^-	H_3O^+	pK_a	d	d	M	−2
CA5	CO_2H	HSv	CO_2^-	H_2Sv^+	pK_a	d	d	M	−2
CA11,12	CO_2H	Sv	CO_2^-	HSv	pK_a	d	d	M	−2
CA10	CO_2H	(g)	CO_2^-		ΔG	d	d	M	−2
SA1	SO_2NH_2	H_2O	SO_2NH^-	H_3O^+	pK_a	d	d	M	−2
HA4	MZ=NOH	H_2O	$MZ=NO^-$	H_3O^+	pK_a	0	d	M	−2
HB1	CN	4-FPnOH	$–CN\cdots H\cdots O–$	—	pK_{KB}	0	sd	S	−2

Sets in italics are unreliable as they do not meet the requirements described in the text. Abbreviations: sd, small diffuse; d, diffuse; ld, large diffuse; sc, small concentrated; c, concentrated; Ld, large concentrated; L, large; M, medium; S, small. pK_{ca} = p*K* of the conjugate acid.

The effect of solvent

1. Solvent composition in mixed solvents. The correlations obtained with eqns. (8*a*) and (8*b*) clearly show that in general there is no dependence on solvent composition for either the CA sets or the BEH sets. These results can best be explained by preferential solvation of the carboxylic acid or ester by water. It must be noted however that $\hat{C}$ for carboxylic acid ionization in aqueous methanol (set CA4) seems to deviate somewhat from the mean value obtained for the other aqueous organic solvents and water of 43.5 ± 3.07 (sets CA1–3, 8, 9). For the most reliable BEH sets (BEH1, 2, 4, 5) the mean value of $\hat{C}$ obtained from correlation with eqn. (14) is 12.1 ± 2.88.

2. Variation with solvent. We have considered the correlation of the $\hat{C}$ values obtained for carboxylic acids in pure solvents with eqn. (20), where P_{Sv} is some solvent parameter. The use of

$$\hat{C} = a_1 P_{Sv} + a_o \qquad (20)$$

the E_T values of Reichart,[23] and the AN (acceptor number) and DN (donor number) parameters of Gutmann[24] as solvent parameters in eqn. (20) gave barely significant correlations, the use of the solvent relative permittivity *D* gave no relationship at all. Correlation with eqn. (21), where b_{OH}, the number of OH

$$\hat{C} = a_1 P_{Sv} + a_2 b_{OH} + a_o \qquad (21)$$

groups, takes the value 1 when the solvent has an OH group and 0 otherwise gave significant results when P_{Sv} is DN. The regression equation is given by eqn. (22), with $100R^2$, 88.18; $A100R^2$,

$$\hat{C} = 1.11(\pm 0.446)\mathrm{DN} - 12.3(\pm 4.75)b_{OH} + 31.9(\pm 10.9) \qquad (22)$$

85.23; *F*, 11.19; S_{est}, 5.28; S^o, 0.486; *N*, 6. Correlation with eqn. (23), where n_H is the number of OH bonds in the solvent

$$\hat{C} = a_1 \mathrm{DN} + a_2 n_H + a_o \qquad (23)$$

molecule gave much poorer results. All of the solvent parameters used with eqns. (20), (21) and (23) are given in Table 8.

Table 8 Solvent parameters used in correlations with eqns. (20), (21), and (23)

Solvent	E_T	D^{20}	AN	DN	b_{OH}	n_H
Water	63.1	80.1	54.8	18.0	1	3
MeOH	55.5	33.62	41.3	19	1	1
EtOH	51.9	25.00	37.1	20	1	1
MeCN	45.0	37.5	19.3	14.1	0	0
$HCONMe_2$	43.8	38.3	16.0	26.6	0	0
Me_2SO	45.0	48.9	19.3	29.8	0	0

It seems that the solvent effect on $\hat{C}$ for carboxylic acid ionization is real and depends on the hydrogen bonding capabilities of the solvent. Much further work is required before this can be regarded as certain.

The estimation of *ρ*, *L* and *C*

If values of these parameters are available it is often possible to estimate pK_as and log *k* values of interest. While values of *h* are also necessary pK_H values may be available for use as approximate *h* values. Alternatively, pK_H values for HGY can be estimated from the appropriate substituent constants for the HG group and the *L* or *C* values for XY (the set for which there is no skeletal group G). *ρ*, *L* and *C* can be estimated from the values for $\hat{C}$ and a_o when these are known for the reaction of interest.

Conclusions

Our results show clearly that substituent electrical effect dependence on distance for the chemical reactivity of a reaction site as measured by *L* (or *ρ* or *C*) is given by eqn. (9) when the

$$L = \hat{C}n^m \qquad (9)$$

angle θ is not greater than 45°. The value of m seems to be determined by the magnitude of the change in charge on the reacting atom, Δq. When Δq is equal to or less than medium in size m is -2, otherwise it is -1. These results are in accord with the MFE model for the transmission of electrical effects which depends on the difference in charge between initial and final states. More work is required to verify this model. When the angular effect is included the dependence of the electrical effect on distance and angle can be described by either eqn. (24) or

$$L = \check{C} \cos \theta \, n^m \qquad (24)$$

(25), where m is -1 or -2, $\check{C}$ is a constant, and n and θ have been defined above.

$$\log |L| = \log \cos \theta + m \log n + \log \check{C} \qquad (25)$$

The reactivities in aqueous organic solvents studied in this work generally show no dependence on solvent composition, this is presumably due to preferential solvation by water. Carboxylic acid ionization studied in pure solvents seems to show a dependence on the hydrogen bonding properties of the solvent and the solvent basicity.

The results of correlations with eqn. (9) can be used to estimate values of ρ, L and C that in turn can be used in conjunction with other coefficients to predict values of pK_a or log k for compounds of interest.

References

1 C. G. Derick, *J. Am. Chem. Soc.*, 1911, **33**, 1152.
2 G. K. Branch and M. Calvin, *The Theory of Organic Chemistry*, Prentice-Hall, New York, 1941, p. 211.
3 R. W. Taft, in *Steric Effects in Organic Chemistry*, M. S. Newman, Ed., Wiley, New York, 1956, p. 556.
4 J. G. Kirkwood, *J. Chem. Phys.*, 1934, **2**, 351.
5 J. G. Kirkwood and F. H. Westheimer, *J. Chem. Phys.*, 1938, **6**, 506; F. H. Westheimer and J. G. Kirkwood, *J. Chem. Phys.*, 1938, **6**, 513; F. H. Westheimer and M. W. Shookhoff, *J. Am. Chem. Soc.*, 1939, **61**, 555; F. H. Westheimer, *J. Am. Chem. Soc.*, 1939, **61**, 1977; J. N. Sarmousakis, *J. Chem. Phys.*, 1944, **12**, 277.
6 H. B. Watson, *Modern Theories of Organic Chemistry*, Oxford University Press, Oxford, 1937, p. 40.
7 A. E. Remick, *Electronic Interpretations of Organic Chemistry*, 2nd Edn., Wiley, New York, 1949, p. 93.
8 W. F. Reynolds, *Prog. Phys. Org. Chem.*, 1983, **14**, 165.
9 K. Bowden and E. J. Grubbs, *Prog. Phys. Org. Chem.*, 1992, **19**, 183.
10 O. Exner and Z. Friedl, *Prog. Phys. Org. Chem.*, 1992, **19**, 259.
11 K. Bowden and E. J. Grubbs, *Chem. Soc. Rev.*, 1996, **25**, 171.
12 F. H. Westheimer and M. W. Shookhoff, *J. Am. Chem. Soc.*, 1940, **62**, 269; F. H. Westheimer, *J. Am. Chem. Soc.*, 1940, **62**, 1892; F. H. Westheimer, W. A. Jones and R. A. Lad, *J. Chem. Phys.*, 1942, **10**, 478; D. Price and F. H. Westheimer, *J. Chem. Phys.*, 1943, **11**, 150.
13 F. H. Westheimer and J. G. Kirkwood, *Trans. Faraday Soc.*, 1947, **43**, 77.
14 M. Charton, *J. Phys. Org. Chem.*, 1999, **12**, 275.
15 L. P. Hammett, *Physical Organic Chemistry*, McGraw-Hill Book Co., New York, 1940, pp. 203–204.
16 M. Charton, *Prog. Phys. Org. Chem.*, 1987, **16**, 287.
17 M. Charton, Dissertation, Stevens Institute of Technology, Hoboken, NJ, 1962.
18 K. Bowden, *Can. J. Chem.*, 1963, **41**, 684; K. Bowden, *Can. J. Chem.*, 1963, **41**, 2781.
19 M. Charton, *CHEMTECH*, 1974, 502.
20 M. Charton, *Prog. Phys. Org. Chem.*, 1981, **13**, 119; M. Charton in *Correlation Analysis in Chemistry: Recent Advances*, N. B. Chapman and J. Shorter, Eds., Plenum, New York, 1978, pp. 175–268; M. Charton, *Prog. Phys. Org. Chem.*, 1971, **8**, 235; M. Charton, *Prog. Phys. Org. Chem.*, 1973, **10**, 81; M. Charton, in *The Chemistry of Dienes and Polyenes*, Z. Rappoport, Ed., Wiley, Chichester, 1997, pp. 683–732; D. D. Perrin, B. Dempsey and E. P. Serjeant, *pKa Prediction for Acids and Bases*, Chapman and Hall, London, 1981.
21 M. Charton, in *The Chemistry of Organic Germanium, Tin and Lead Compounds*, S. Patai, Ed., Wiley, Chichester, 1995, pp. 603–664.
22 W. Kitching, M. Bullpit, D. Gartshore, W. Adcock, T. C. Khor, D. Doddrell and I. D. Rae, *J. Org. Chem.*, 1977, **42**, 2411; W. Adcock and T. C. Khor, *J. Am. Chem. Soc.*, 1978, **100**, 7799; W. Adcock and D. P. Cox, *J. Org. Chem.*, 1979, **44**, 3004.
23 C. Reichart, *Solvent Effects in Organic Chemistry*, Verlag Chemie, Weinheim, 1979.
24 V. Gutmann, *The Donor-Acceptor Approach to Molecular Interactions*, Plenum Press, New York, 1978.

Paper 9/01495K