Structure–Reactivity Relationships

DOI: 10.1002/anie.201005816

Limitations of the $s(E+N)$ and Related Equations: Solvent Dependence of Electrophilicity**

*T. William Bentley**

electrophilicity · kinetics ·
linear free-energy relationships · nucleophilicity ·
solvent effects

Generic expressions that capture the essence of reactivity patterns still have important roles, despite the huge advances in computational chemistry for individual systems.[1,2] New generic equations are relatively rare, and critical scrutiny of the scope and reliability of two such equations [Eqs. (1) and (2)] is presented below.

$$\log k = s\,(E+N)\text{, or changing the symbol for } s \qquad \log k = s_{\mathrm{N}}\,(E+N) \tag{1}$$

$$\log k/s_{\mathrm{N}} = s_{\mathrm{E}}\,(E+N) \tag{2}$$

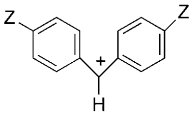
The $s(E+N)$ equation [Eq. (1)] has been developed empirically from constant selectivity relationships[3,4] to correlate logarithms of rate constants ($\log k$) at 20 °C for a huge range of reactions of electrophiles (electrophilicity E) and nucleophiles (nucleophilicity N); in Equation (1), s (or s_{N})[5–7] is referred to as a "nucleophile-specific" parameter,[7] and in Equation (2), s_{E} is an "electrophile-specific"' parameter.[7]

Initially Equation (1) was intended to be "semi-quantitative"(reliable to an order or two of magnitude),[3] but confidence has grown to such an extent that many values of E and N are quoted to an accuracy of four significant figures. Furthermore, it has recently been proposed[7] that Equation (2) is the basis of a general scale of nucleophilicity from which various other scales can be derived: for example, Equation (1) is derived from Equation (2) if $s_{\mathrm{E}} = 1$, as observed for many cation–anion recombinations.[8]

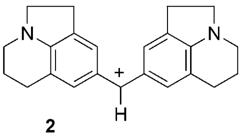
When an equation fits a huge amount of data, it is tempting to conclude that a fundamental "general relationship"[9] has been established. The situation is reminiscent of claims made in the early 1980s for solvent effects; after lively exchanges of strongly opposing views,[10–12] the consensus[13,14] is that linear free-energy relationships are local empirical rules or quantitative similarity models.[10,12] Far from behaving as a fundamental law, providing a challenge to theory,[9] Equation (2) is shown below (Table 1) to be incorrect.

The constants (independent variables) for typical correlations (e.g. Hammett σ) are usually defined directly from experimental data, and the slopes (e.g. ρ) are optimized, adjustable parameters.[14] In contrast, over 1000 values of the adjustable parameters E, N, or s_{N} have emerged from complex optimizations of data for benzhydrylium cations (e.g. **1** and **2**),[15] illustrated in Scheme 1 for the parts required for water as solvent (the complete scheme is much larger and is explained in detail elsewhere).[16]

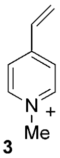
Preferably,[22] there should be a ratio of five data points per independent or explanatory variable, whereas in the



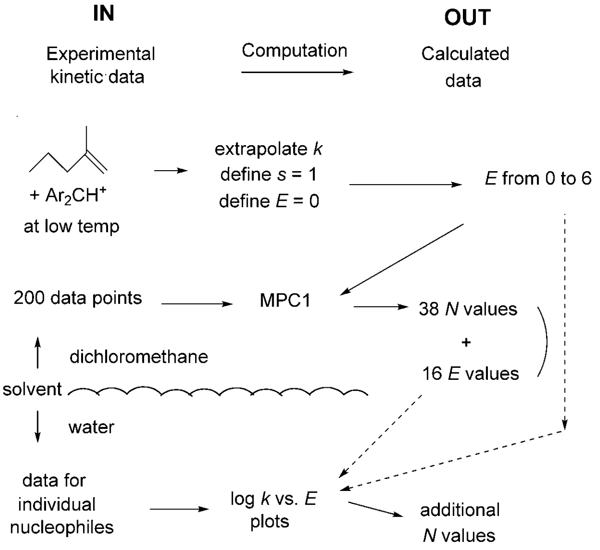
1



2



3



Scheme 1. Steps leading to values of E for cations and then to values of N for nucleophiles in water.[17–21]

---

[*] Dr. T. W. Bentley
Chemistry Unit, Grove Building, School of Medicine
Swansea University
Singleton Park, Swansea SA2 8PP, Wales (UK)
Fax: (+44) 179-229-5554
E-mail: t.w.bentley@swansea.ac.uk

[**] Helpful comments from D. Bethell are gratefully acknowledged. s = slope, E = electrophilicity, N = nucleophilicity.

multiparameter correlation (MPC1, Scheme 1) required for the initial values of E, 92 parameters are calculated from only 200 data points; a ratio of approximately 3:1 is typical for calculations of N and s_N for each nucleophile from plots of $\log k$ versus E ($N = s_N N/s_N$ or intercept/slope).[18–21,23]

Extrapolations are often required to obtain intercepts, thus introducing extrapolation errors arising from the complex substituent effects in benzhydryl substrates[24] and from cross interactions between the electrophile and the incoming nucleophile.[25] The E scale currently spans a range from $+6$ to -23,[15] and the lower region refers to neutral substrates.[4,23] Unlike other structure–reactivity correlations,[13,14] the reference substrate for the E scale of electrophilicity ($E = 0$ for **1**, $Z = OMe$) is located well away from the center of the E scale, so changing to **1** ($Z = NMe_2$) having $E = -7.02$[17] would reduce extrapolation errors.[16,26]

A key assumption in the design of Equation (1) is that solvent-independent values of E for benzhydrylium cations, reference electrophiles for Equations (1) and (2), can be obtained by a multiparameter correlation (MPC1, Scheme 1) of kinetic data for $\pi_{C=C}$ nucleophiles (e.g. alkenes) in dichloromethane;[17] it is considered[27,28] to be unlikely that the same E values would apply to the $\log k$ versus E plots[18–21] for other solvents such as water.

The limitations of Equations (1) and (2) will be illustrated below by correlations of kinetic data for reactions of benzhydrylium cations **1** and **2**[18–21] and for the methylvinylpyridinium cation **3**.[29] All of these reactions involve rate-determining nucleophilic attack in water for various steric environments, for example at secondary carbon atoms for **1** and **2** and at a primary carbon atom for **3** (Michael addition to the double bond[29]) by 35 primary, secondary, and α-effect amines, amino acids, and peptides.

When a fixed nucleophilicity scale is defined directly from $\log k$ for **1** ($Z = NMe_2$) or **2**, as suggested recently,[16,26] plots give slopes (s_E) and intercepts which refer to differences between E values of the two substrates (see Table 9 of Ref. [26]). There are excellent correlations[8] (Figure 1 and Figure 2) of $\log k$ for **2** or **3** versus $\log k$ for **1** ($Z = NMe_2$) for amines in water, but the intercept of Figure 1 of -1.87 ± 0.05 is not in agreement with the published value of -3.0 for the difference in E values.[17]

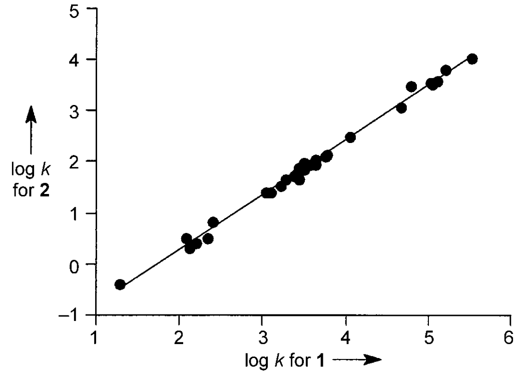


Figure 1. Correlation of $\log k$ for **2** vs. $\log k$ for **1** ($Z = NMe_2$) at 20 °C for 32 amines and hydroxide ion in water: slope $= 1.075 \pm 0.014$; intercept $= -1.87 \pm 0.05$ ($n = 33$, $R^2 = 0.995$, std error $= 0.08$); data from Refs. [18–21].

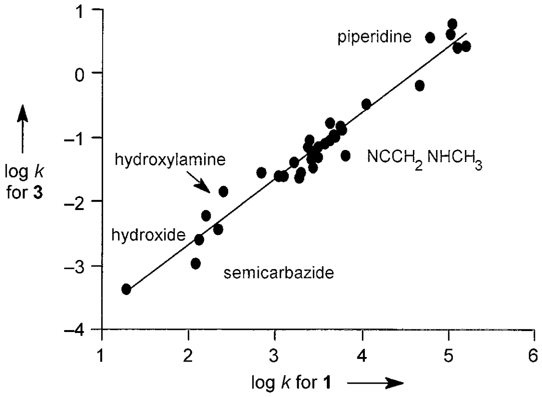


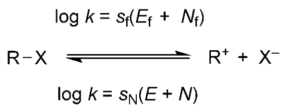
Figure 2. Plot of $\log k$ for **3** for amines in water at 25 °C vs. $\log k$ for **1** ($Z = NMe_2$) at 20 °C: slope $= 1.03 \pm 0.04$; intercept $= -4.74 \pm 0.14$ ($n = 35$; $R^2 = 0.954$; std error $= 0.21$); data for **3** from Ref. [29].

In contrast, the corresponding plot for 22 nucleophiles in dichloromethane gave a significantly more negative intercept of -2.66 ± 0.16.[26] Hence, it appears that the E scale is attenuated in water, presumably because water is a good electron-pair donor solvent.[30] The slope of Figure 1 ($s_E = 1.07$) in water is the same as that of the plot in dichloromethane,[26] thus suggesting that this parameter is of practical use (i.e., that it is consistent and predictable) and so of theoretical interest.[31]

When $s_E \approx 1.00$ (as in Figure 1 and Figure 2), a constant selectivity relationship is obeyed, and it can be expressed as $\log k = E + N$ (see Equations (30) and (31) in Ref. [32]). According to this equation, if correct values of E were used, a plot of $\log k$ versus E (for a fixed nucleophile) would have a slope close to unity, so $s_N \approx 1.00$ [Eq. (1)]. Consequently, published values of $s_N = 0.6$ or 0.7 in water[18] and DMSO[23] are too low, and the substituent effect s_N is not simply a 'nucleophile-specific parameter'. As illustrated below for hydrogen isotope effects, there is a distinction between substrate and solvent effects.

In the current design of Equations (1) and (2), all of the solvent effects are included in the N and s_N terms.[32] For cation–nucleophile recombinations in protic solvents, anion desolvation is a dominant process for both anions[33] and amines;[34] for cations $R_1R_2CH^+$, $k_H/k_D \approx 1.00$,[35] thus suggesting that there is little change in hybridization at the carbenium center in the transition state.[8] In contrast, reactions of alkenes in dichloromethane show $k_H/k_D \approx 0.80$, thus suggesting that "rehybridization of the carbenium ion center is far advanced in the transition state."[36] It is reasonable in principle to correlate these effects by a nucleophilicity parameter (e.g. N) and by a substituent-effect parameter (e.g. s_N) to allow for variations in electron demand.[37] A limitation is that N is calculated indirectly; the value of N depends on E, defined in dichloromethane by MPC1 (Scheme 1), even though reactions in dichloromethane differ significantly from reactions in protic solvents.

Further insights are available from extensive research on S_N1 reactions, which are the reverse of cation–anion recombinations (Scheme 2).[8,27] The corresponding terms (E_f for electrofugality of the cation or electrophile and N_f for



$$R{-}X \underset{\log k = s_N(E + N)}{\overset{\log k = s_f(E_f + N_f)}{\rightleftharpoons}} R^+ + X^-$$

Scheme 2. Equations for S_N1 reactions and their reverse (cation–anion recombinations).

nucleofugality of the anion or leaving group) can be calculated from data for benzhydryl substrates.[38] Consequently, the analogous equation ($\log k = s_f(E_f + N_f)$) for the reverse of cation–anion recombinations behaves as a benzhydryl similarity model. However, for other cations (e.g. *tert*-butyl or 1-phenylethyl), different values of N_f are needed if a fixed value of E_f is assumed,[38] or if N_f is fixed, then E_f depends on the solvent.[39] The results can be explained by differences in solvation between alkyl and more delocalized cations.[27,40–44] As shown below, the same applies to Equation (1), so it should be applied cautiously as an additivity rule.

The problems associated with Equation (2) are more severe. Despite the fact that other equations have been derived from Equation (2),[7] there is currently no satisfactory derivation of Equation (2).[45] A modification has already been suggested in which the $s_E E$ term in Equation (2) is replaced by E in Equation (3).[46,47] Arguably, one of the major problems associated with Equations (1)–(3) is that there is no defined standard reaction; instead of calculating relative rates ($\log k/k_0$), $\log k$ is obtained directly.

$$\log k/s_N = E + s_E N \tag{3}$$

For **3**, the more precise correlation is a direct plot of $\log k$ versus $\log k$ for **1** ($Z = NMe_2$; Figure 2). Hydroxide, a suitable reference nucleophile (k_0),[8] fits the correlation line, so the intercept would be negligible if relative rates ($\log k/k_0$) were plotted. The corresponding plot for both Equations (2) and (3) is of $\log k/s_N$ versus N (Figure 3), thus giving a slope (s_E) of 0.77 instead of 1.03 (Table 1). As expected because of differences in charge delocalization,[40–44] N values obtained from data for **1** or **2** are not appropriate for **3** (especially for semicarbazide as nucleophile).

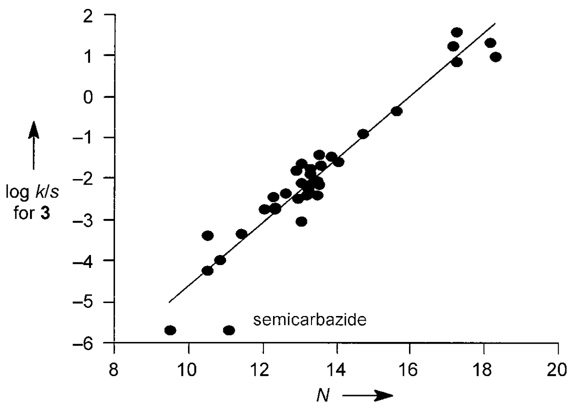


Figure 3. Plot [Eqs. (2) and (3)] of $\log k/s_N$ for **3** in water at 25 °C vs. N: slope $= 0.772 \pm 0.042$; intercept $= -12.33 \pm 0.57$ ($n = 35$; $R^2 = 0.912$; std error $= 0.51$); see also Figure 8 of reference [18].

Table 1: Comparisons of best estimates (Figure 2) with Figure 3 [Eqs. (2) and (3)] for reactions of **3** with amines, amino acids, and peptides.

Parameter	Figure 2	Equation (2)	Equation (3)
s_E	1.03	0.77[a]	0.77
E	−11.7[b]	−16.0[a]	−12.3
std error[c]	0.21	0.51	0.51

[a] Figure 8 of reference [18]) gives $s_E = 0.762$ and an intercept ($s_E E$) of −12.27. [b] See text. [c] Standard errors cannot be compared directly because Equations (2) and (3) require values of $\log k/s_N$, which are about twice as large as $\log k$.

An excellent correlation for **1** ($Z = NMe_2$; not shown) for Equations (2) and (3) gave slope $= 1.003 \pm 0.007$; intercept -7.09 ± 0.09; $n = 35$; $R^2 = 0.998$, standard error $= 0.085$. The intercept agrees with the published value of −7.02 for E.[17] So there is satisfactory agreement that $s_E = 1.0$ for **1** ($Z = NMe_2$), but only Figure 2 shows a unit slope for **3**. The satisfactory results arise because experimental data for **1** ($Z = NMe_2$) and very closely related cations are used in the same equation to calculate the parameters s_N, N, and E.

According to the correlation (Figure 2), $\log k$ for **3** (at 25 °C) is 4.7 less than that for **1** ($Z = NMe_2$), and this difference should be a satisfactory measure of the difference in electrophilicity (E), because the plot is of unit slope. In comparisons of the various equations (Table 1), a small temperature difference (20 °C for **1**, 25 °C for **3**) is ignored. Assuming that $E = -7.02$ for **1** ($Z = NMe_2$),[17] E for **3** is calculated from Figure 2 to be −11.7, in satisfactory agreement with $E = -12.3$, obtained from the intercept of Figure 3 before division by s_E. Consequently, Equation (2) is incorrect; values of E are more consistent with Equation (3)[46,47] than with Equation (2), but neither leads to the expected conclusion that $s_E = 1.0$ (Figure 2).[8]

For the tertiary amine imidazole ($\log k$ for **1** ($Z = NMe_2$) $= 1.51$, $s_N = 0.57$, $N = 9.63$),[48] predicted values for **3** are slightly faster than the observed $\log k$ for **3** of −3.51.[29] Figure 2 gives −3.18 and Figure 3 gives −2.80 (steric effects are known to complicate related calculations[29,35a,49]).

In general, caution is required when applying multiparameter correlations.[13] The complex data processing (Scheme 1) has obscured the 1:1 relationship[8] between $\log k$ for **1** ($Z = NMe_2$) and **3** (Figure 2). Inconsistent results for **3** show that Equations (1)–(3) are not robust. An underlying cause is the solvent dependence of electrophilicity (E), contrary to a key tacit assumption required to obtain N values using Equation (1).[18–21]

When $s_E \neq 1$ in Equation (2), E becomes based on a floating scale coexisting with the established E scale based on $s_E = 1$,[17] an unjustifiable situation.[45] When $s_E = 1$ in Equation (2), it is the same as the original Equation (1). Consequently, the recently proposed[7] general equation [Eq. (2)] should not be used; plots of $\log k/s_N$ versus N should be based on Equation (1) and so drawn with a slope of unity.

Alternative equations which include an s_E term and correspond to Figure 1 and Figure 2 can be used for predictions and also to gain additional insights;[16,26–28] these equations can also be used to correlate the same range of reactivity as Equation (1), and they show that N in Equation (1) has a floating scale[45] deviating by up to ten orders of magnitude from a fixed scale.[16]

Received: September 16, 2010

---

[1] R. A. Marcus, *Frontiers of Physical Organic Chemistry Faraday Discussion* **2010**, *145*, 555.
[2] J. Michl, *Frontiers of Physical Organic Chemistry Faraday Discussion* **2010**, *145*, 558.
[3] H. Mayr, M. Patz, *Angew. Chem.* **1994**, *106*, 990 – 1010; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 938 – 957.
[4] R. Lucius, H. Mayr, *Angew. Chem.* **2000**, *112*, 2086 – 2089; *Angew. Chem. Int. Ed.* **2000**, *39*, 1995 – 1997.
[5] As s is the symbol in the Swain–Scott equation,[6] s_N will be used in Equation (1).
[6] C. G. Swain, C. B. Scott, *J. Am. Chem. Soc.* **1953**, *75*, 141 – 147.
[7] T. B. Phan, M. Breugst, H. Mayr, *Angew. Chem.* **2006**, *118*, 3954 – 3959; *Angew. Chem. Int. Ed.* **2006**, *45*, 3869 – 3874.
[8] C. D. Ritchie, *Can. J. Chem.* **1986**, *64*, 2239 – 2250.
[9] H. Mayr, A. R. Ofial, *J. Phys. Org. Chem.* **2008**, *21*, 584 – 595.
[10] M. Sjöström, S. Wold, *Acta Chem. Scand. Ser. B* **1981**, *35*, 537 – 554.
[11] M. J. Kamlet, R. W. Taft, *Acta Chem. Scand. Ser. B* **1985**, *39*, 611 – 628.
[12] S. Wold, M. Sjöström, *Acta Chem. Scand. Ser. B* **1986**, *40*, 270 – 277.
[13] T. W. Bentley, M. S. Garley, *J. Phys. Org. Chem.* **2006**, *19*, 341 – 349.
[14] A. Williams, *Free Energy Relationships in Organic and Bioorganic Chemistry*, Royal Society of Chemistry, London, **2003**.
[15] http://www.cup.uni-muenchen.de/oc/mayr.
[16] T. W. Bentley, *J. Phys. Org. Chem.* **2011**, *24*, 282 – 291.
[17] H. Mayr, T. Bug, M. F. Gotta, N. Hering, B. Irrgang, B. Janker, B. Kempf, R. Loos, A. R. Ofial, G. Remennikov, H. Schimmel, *J. Am. Chem. Soc.* **2001**, *123*, 9500 – 9512.
[18] F. Brotzel, Y. C. Chu, H. Mayr, *J. Org. Chem.* **2007**, *72*, 3679 – 3688.
[19] F. Brotzel, B. Kempf, T. Singer, H. Zipse, H. Mayr, *Chem. Eur. J.* **2007**, *13*, 336 – 345.
[20] F. Brotzel, H. Mayr, *Org. Biomol. Chem.* **2007**, *5*, 3814 – 3820.
[21] S. Minegishi, H. Mayr, *J. Am. Chem. Soc.* **2003**, *125*, 286 – 295.
[22] M. H. Abraham, P. L. Grellier, J.-L. M. Abboud, R. H. Doherty, R. W. Taft, *Can. J. Chem.* **1988**, *66*, 2673 – 2686.
[23] R. Lucius, R. Loos, H. Mayr, *Angew. Chem.* **2002**, *114*, 97 – 102; *Angew. Chem. Int. Ed.* **2002**, *41*, 91 – 95.
[24] Y. Tsuno, M. Fujio, *Adv. Phys. Org. Chem.* **1999**, *32*, 267 – 385.
[25] a) I. Lee, *Adv. Phys. Org. Chem.* **1992**, *27*, 57 – 117; b) P. Denton, C. D. Johnson, *J. Chem. Soc. Perkin Trans. 2* **1995**, 477 – 481.
[26] T. W. Bentley, *J. Phys. Org. Chem.* **2010**, *23*, 836 – 844.
[27] D. N. Kevill, M. J. D'Souza, *J. Chem. Res.* **2008**, 61 – 66.
[28] T. W. Bentley, *J. Phys. Org. Chem.* **2010**, *23*, 30 – 36.
[29] J. W. Bunting, C. K. M. Heo, *J. Chem. Soc. Perkin Trans. 2* **1994**, 2279 – 2290.
[30] C. Reichardt, *Solvents and Solvent Effects in Organic Chemistry*, 3rd ed., Wiley-VCH, Weinheim, **2003**, Section 2.6.
[31] S. Hoz, D. Speizman, *J. Org. Chem.* **1983**, *48*, 2904 – 2910.
[32] R. A. More O'Ferrall, *Adv. Phys. Org. Chem.* **2010**, *44*, 19 – 122.
[33] C. D. Ritchie, P. O. I. Virtanen, *J. Am. Chem. Soc.* **1972**, *94*, 4966 – 4971.
[34] R. A. McClelland, V. M. Kanagasabapathy, N. S. Banait, S. Steenken, *J. Am. Chem. Soc.* **1992**, *114*, 1816 – 1823.
[35] a) C. A. Bunton, N. Carrasco, F. Davoudzadeh, W. E. Watts, *J. Chem. Soc. Perkin Trans. 2* **1980**, 1520 – 1526; b) T. Van Tham, R. A. McClelland, *Can. J. Chem.* **2001**, *79*, 1887 – 1897.
[36] H. Mayr, R. Schneider, C. Schade, J. Bartl, R. Bederke, *J. Am. Chem. Soc.* **1990**, *112*, 4446 – 4454.
[37] H. C. Brown, *The Non-Classical Ion Problem*, Plenum, New York, **1977**, chap. 10.
[38] T. W. Bentley, *Chem. Eur. J.* **2006**, *12*, 6514 – 6520.
[39] B. Denegri, A. R. Ofial, S. Jurić, A. Streiter, O. Kronja, H. Mayr, *Chem. Eur. J.* **2006**, *12*, 1657 – 1666.
[40] K. T. Liu, H. C. Sheu, H. I. Chen, P. F. Chiu, C. R. Hu, *Tetrahedron Lett.* **1990**, *31*, 3611 – 3614.
[41] T. W. Bentley, I. S. Koo, S. J. Norman, *J. Org. Chem.* **1991**, *56*, 1604 – 1609.
[42] D. N. Kevill, N. H. Ismail, M. J. D'Souza, *J. Org. Chem.* **1994**, *59*, 6303 – 6312.
[43] D. N. Kevill, M. J. D'Souza, *Curr. Org. Chem.* **2010**, *14*, 1037 – 1049.
[44] D. S. Kumar, K. P. Jayakumar, S. Balachandran, *J. Phys. Org. Chem.* **2010**, *23*, 783 – 788.
[45] There is a derivation of Equation (2) in footnote 47 of Ref. [7] that combines constants and variables in such a way that both E and N become based on floating scales (having less than the usual attachment to fixed reference points).
[46] V. Dichiarante, M. Fagnoni, A. Albini, *J. Org. Chem.* **2008**, *73*, 1282 – 1289.
[47] This alternative to Equation (2) was originally expressed as $\log k = s(E+eN)$.[46]
[48] M. Baidya, F. Brotzel, H. Mayr, *Org. Biomol. Chem.* **2010**, *8*, 1929 – 1935.
[49] P. Rodriguez-Dafonte, F. Terrier, S. Lakhdar, S. Kurbatov, R. Goumont, *J. Org. Chem.* **2009**, *74*, 3305 – 3315.