Reaction Mechanism from Structure-Energy Relations. 1. Base-Catalyzed Addition of Alcohols to Formaldehyde

Ernest Grunwald

Contribution from the Chemistry Department, Brandeis University, Waltham, Massachusetts 02254. *Received November 16, 1984*

Abstract: There is long-standing evidence that in the general base catalyzed addition of alcohols to formaldehyde, C···O bond formation and proton transfer occur simultaneously. Structure–energy relations for this reaction are complicated. The slopes of Brønsted plots vary greatly, even for structurally similar alcohols, and plots of log k vs. pK_a of the alcohols go through minima. The data thus provide good material for testing the author's recent theory of structure–energy relations for concerted reactions. The theoretical equations are nonlinear and depend specifically on the reaction mechanism. Of four mechanisms, only that favored by previous independent work fits the data well, even onto reproducing the rate-constant minima. The rate-determining step of this mechanism is $R'CH_2COO^- + RCH_2OH + H_2C{=}O \rightarrow R'CH_2COOH + RCH_2OCH_2O^-$. R ranges in electronegativity from CH_3 to CF_3, R' from H to CN. Assuming this mechanism to be correct, progress of C···O bond formation (u^*) and proton transfer (v^*) were deduced from the theoretical equations. The results show that mean progress at the transition state, $(u^* + v^*)/2$, varies only sightly with substitution, but that disparity of progress of the two reaction events, $(v^* - u^*)/2$, varies markedly both with R and R' and changes sign within the series. Rate-constant minima occur near points where $u^* = v^*$, thus proving the effectiveness of disparity at lowering the free energy of activation.

Structure–energy relations have long been used to help identify reaction mechanisms.[1,2] Recently I formulated a theory of structure–energy relations which is capable of giving a detailed description of *concerted* reaction mechanisms.[3] This theory allows for the effect of disparity of progress of concerted reaction events at the transition state, and for the variation of that disparity within a reaction series. The theory is an extension of Marcus rate–equilibrium theory[4] and its developments,[5-10] and it attempts a quantification of the use of More O'Ferrall diagrams,[11] following contributions by others.[12-19] In this and the following paper, I shall test the theory by applying it to a family of reactions with unusually complex structure–energy relationships, namely, the general acid and base catalyzed addition of alcohols to formaldehyde (eq 1).

$$RCH_2OH + H_2C{=}O + \text{catalyst} \underset{k_r}{\overset{k_f}{\rightleftharpoons}} RCH_2OCH_2OH + \text{catalyst} \quad (1)$$

It will be useful to begin with some definitions. A *concerted reaction mechanism* is one in which potentially consecutive reaction events (such as breaking an existing bond and forming a new one, or bond-breaking and electron delocalization, etc.) actually take place simultaneously. A *progress variable* is any property of a reacting molecular configuration which measures the relative progress of a reaction event; I shall use normalized progress variables which vary from zero for the reagents to unity for the products.

Let v and u denote the progress variables for a reaction in which there are two concerted reaction events. For any *single* reaction, v is a monotonic function of u. If this function is practically identical for all reactions in a series, then the reaction coordinate z for the series depends in effect on a single progress variable, which might be either u or v or some other monotonic normalized function of u. If, however, $v(u)$ varies for different reactions in the series, then z is a function of two progress variables; that is, $z = z(u,v)$. Other concerted mechanisms are conceivable in which z depends on more than two progress variables. Because theoretical structure–energy relationships and their specific parameters depend on the number and nature of the progress variables, a concerted mechanism is defined theoretically by its progress variables.

Rate Constants for Base-Catalyzed Reaction. I chose reaction 1 for the testing of theory because of the complicated phenomenology, which is supported by good data, and because of consistent evidence that the rate-determining addition of RCH_2OH to the carbonyl group is concerted with proton transfer to or from the catalyst.[12c,14,20] The reaction mechanism has been reviewed by Bell[20] and, more recently, by Funderburk, Aldwin, and Jencks (FAJ),[14] whose data I shall use. In this paper I shall consider the base-catalyzed reaction.

Relevant structure–energy plots for the base-catalyzed reaction are shown in Figures 1 and 2. The rate constants (k_r) refer to

(1) Hammett, L. P. "Physical Organic Chemistry", 1st ed.; McGraw-Hill: New York, 1940; 2nd ed., 1970.
(2) Leffler, J. E.; Grunwald, E. "Rates and Equilibria of Organic Reactions"; Wiley: New York, 1963.
(3) Grunwald, E. *J. Am. Chem. Soc.* **1985**, *107*, 125.
(4) (a) Marcus, R. A. *J. Phys. Chem.* **1968**, *72*, 891. (b) Cohen, A. O.; Marcus, R. A. *Ibid.* **1968**, *72*, 4249. (c) Marcus, R. A. *J. Am. Chem. Soc.* **1969**, *91*, 7224. (d) Marcus, R. A. *Faraday Symp. Chem. Soc.* **1975**, *10*, 60.
(5) Kurz, J. L. *Chem. Phys. Lett.* **1978**, *57*, 243.
(6) (a) Albery, W. J. *Annu. Rev. Phys. Chem.* **1980**, *31*, 227. (b) Albery, W. J. *Faraday Discuss. Chem. Soc.* **1982**, *74*, 245.
(7) (a) Murdoch, J. R. *J. Am. Chem. Soc.* **1983**, *105*, 2667. (b) Murdoch, J. R. *Ibid.* **1983**, *105*, 2159. (c) Murdoch, J. R.; Magnoli, D. E. *Ibid.* **1982**, *104*, 3792. (d) Magnoli, D. E.; Murdoch, J. R. *Ibid.* **1981**, *103*, 7465.
(8) (a) Kreevoy, M. M.; Oh, S.-W. *J. Am. Chem. Soc.* **1973**, *95*, 4805. (b) Albery, W. J.; Kreevoy, M. M. *Adv. Phys. Org. Chem.* **1978**, *15*, 87. (c) Kreevoy, M. M.; Lee, I. S. H. *J. Am. Chem. Soc.* **1984**, *106*, 2550. (d) Kreevoy, M. M.; Konasewich, D. E. *Adv. Chem. Phys.* **1971**, *21*, 243.
(9) (a) Kresge, A. J. *Chem. Soc. Rev.* **1973**, *2*, 475. (b) Kresge, A. J.; Sagatys, D. S.; Chen, H. L. *J. Am. Chem. Soc.* **1977**, *99*, 7228. (c) Kresge, A. J. *Ibid.* **1980**, *102*, 7797.
(10) (a) Lewis, E. S.; Kukes, S.; Slater, C. D. *J. Am. Chem. Soc.* **1980**, *102*, 1619. (b) Lewis, E. S.; Hu, D. D. *Ibid.* **1984**, *106*, 3292.
(11) (a) More O'Ferrall, R. A. *J. Chem. Soc. B* **1970**, 274. (b) Albery, W. J. *Prog. React. Kinet.* **1967**, *4*, 355.
(12) (a) Jencks, W. P.; *Chem. Rev.* **1972**, *72*, 705. (b) Young, P. R.; Jencks, W. P. *J. Am. Chem. Soc.* **1979**, *101*, 3288. (c) Palmer, J. L.; Jencks, W. P. *Ibid.* **1980**, *102*, 6472. (d) Cox, M. M.; Jencks, W. P. *Ibid.* **1981**, *103*, 580. (e) Gandler, J. R.; Jencks, W. P. *Ibid.* **1982**, *104*, 1937.
(13) Jencks, D. A.; Jencks, W. P. *J. Am. Chem. Soc.* **1977**, *99*, 7948.
(14) Funderburk, L. H.; Aldwin, L.; Jencks, W. P. *J. Am. Chem. Soc.* **1978**, *100*, 5444.
(15) Critchlow, J. E. *J. Chem. Soc., Faraday Trans. 1* **1972**, *68*, 1774.
(16) Murdoch, J. R. *J. Am. Chem. Soc.* **1983**, *105*, 2660.
(17) (a) Harris, J. C.; Kurz, J. L. *J. Am. Chem. Soc.* **1970**, *92*, 349. (b) Harris, J. M.; Shafer, S. G.; Moffatt, J. R.; Becker, A. R. *Ibid.* **1979**, *101*, 3295.
(18) (a) Gajewski, J. J.; Gilbert, K. E. *J. Org. Chem.* **1984**, *49*, 11. (b) Gajewski, J. J. *J. Am. Chem. Soc.* **1979**, *101*, 4393.
(19) (a) Guthrie, J. P. *J. Am. Chem. Soc.* **1980**, *102*, 5286. (b) Bernasconi, C. F.; Gandler, J. R. *Ibid.* **1978**, *100*, 8117.
(20) Bell, R. P. "The Proton in Chemistry", 2nd ed.; Cornell University Press: Ithaca, N.Y., 1973; pp 183-190.

0002-7863/85/1507-4710$01.50/0

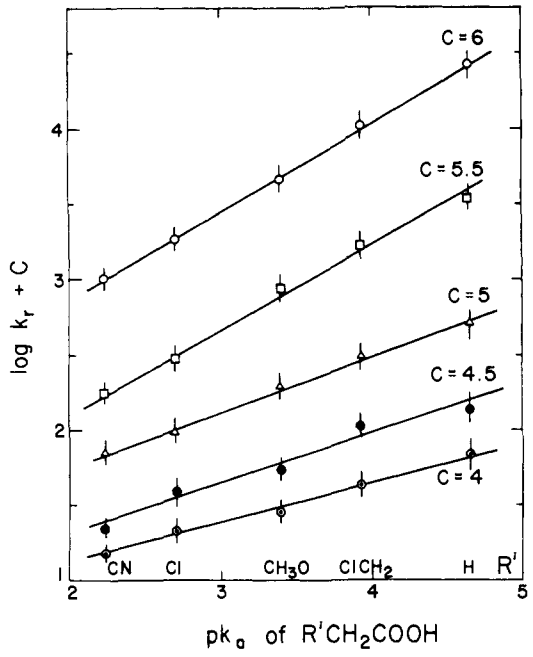

Figure 1. Brønsted plot for base-catalyzed of RCH_2OCH_2OH vs. pK_a of catalyst in water. *C*, R, slope of straight line: 4, CH_3, 0.267; 4.5, CH_3OCH_2, 0.330; 5, $ClCH_2$, 0.360; 5.5, Cl_2CH, 0.549; 6, F_3C, 0.586 (data from ref 14).

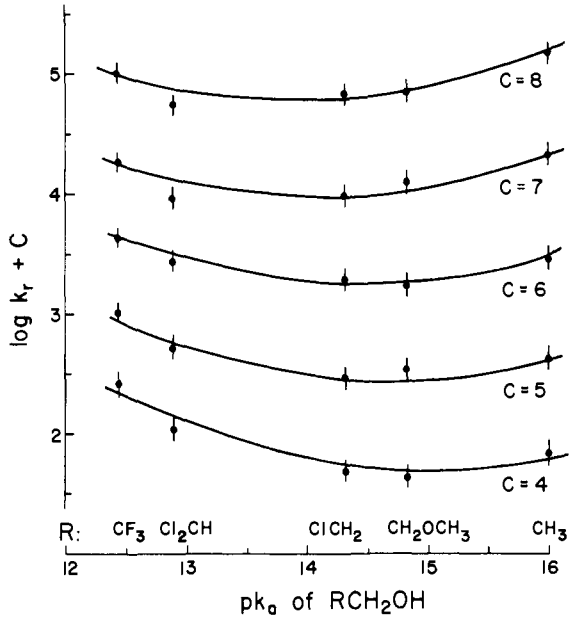

Figure 2. Logarithmic plot of rate constant for cleavage of RCH_2OCH_2OH catalyzed by $R'CH_2COO^-$ in water, vs. pK_a of the alcohol produced. C, R': 4, H; 5, $ClCH_2$; 6, CH_3O; 7, Cl; 8, CN (data from ref 14).

the reverse reaction in eq 1, which was made irreversible by trapping the formaldehyde. Figure 1 shows Brønsted plots of log k_r vs. pK_a of the conjugate of the base catalyst. These are straight lines, as expected, but the slopes are unusually variable.

Figure 2 shows plots of log k_r vs. pK_a of the *alcohol*. These all go through a minimum. Since nucleophilicity for a series of closely similar substrates, such as RCH_2OH, parallels basicity, one would have expected log k_r to increase with pK_a of the alcohol. Thus the minima beg to be explained. Rate-constant minima sometimes indicate a change of mechanism, but that is not expected here because the changes in alcohol structure are small and occur outside the reaction zone.

Theoretical Formulation. The theory of structure–energy relations[3] which I wish to test is nonlinear and permits rate-constant minima even though the reaction events themselves vary with

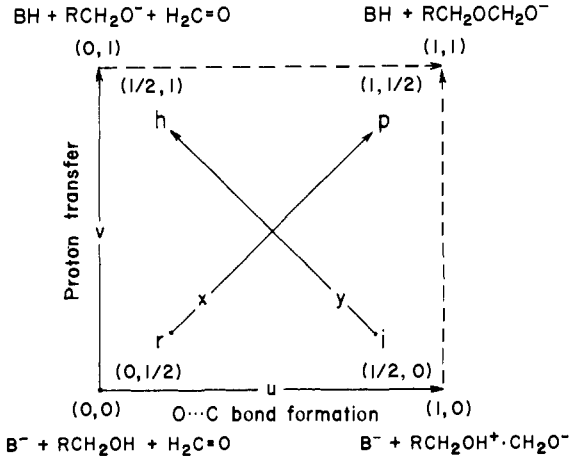

Figure 3. Progress variables and reaction events in the rate-determining step of mechanism 2 for the base-catalyzed addition of alcohols to formaldehyde.

structure in a normal manner. It turns out that if I adopt the two-step mechanism 2 favored by FAJ,[14] the theory reproduces

$$B^- + HOCH_2R + H_2C{=}O \underset{k_-}{\overset{k_+}{\rightleftharpoons}} BH + RCH_2OCH_2O^- \quad (\Delta G^\circ) \quad (2a)$$

$$RCH_2OCH_2O^- + HB \overset{\text{fast}}{\rightleftharpoons} RCH_2OCH_2OH + B^- \quad (\Delta G^{HB} - \Delta G^1) \quad (2b)$$

the minima, while other mechanisms that I have tested do not. This section will review the basic theory[3] and apply it to the mechanism of eq 2.

The rate-controlling step (2a) involves two reaction events: bond formation to >C=O, and proton transfer to B^-. In Figure 3 these events and their normalized progress variables u and v form the coordinates of a More O'Ferrall diagram. Upon rotation of the u,v axes by 45°, translation of the origin, and renormalization, one obtains an x,y coordinate system (included in Figure 3) whose x axis connects reagents r and products p, and whose y axis connects the two intermediate states that obtain when the reaction events occur stepwise (i when u is followed by v; h when v is followed by u). The u,v coordinates are related to the x,y coordinates by:[3]

$$x = (v + u)/2 \quad (3a)$$

$$y = (v - u)/2 + \tfrac{1}{2} \quad (3b)$$

Accordingly, x measures mean progress, and $y - {}^1/_2$ measures disparity of progress of the two reaction events. The process r → p is the main reaction. The process i → h has been called a *disparity reaction*. To apply Marcus theory it is convenient to represent the free energy G as a function of x and y.[3]

For many reaction series it is possible to transform the progress variables chosen initially so that $G(x,y)$ is a quadratic function.[3,5,7,13] In such cases, on applying the assumptions of Marcus theory and appropriate boundary conditions,[3] $G(x,y)$ can be expressed in the useful form:

$$G(x,y) = c + 4\gamma x(1 - x) + x\Delta G^\circ - 4\mu y(1 - y) + y\Delta G' \quad (4)$$

Here $\Delta G^\circ = G(p) - G(r)$, $\Delta G' = G(h) - G(i)$, and c is a constant depending on the zero point of the energy scale. By definition, both of the energy quantities γ and μ are positive. When $|\Delta G^\circ|$ is less than about 2γ, γ may be treated as a constant characteristic of the reaction family and is called the *intrinsic energy barrier*. Similarly, when $|\Delta G'|$ is less than about 2μ, μ may be treated as a constant and is called the *intrinsic energy well*.

According to eq 4, the coordinates of the transition state are given by (5). The activation free energy $\Delta G^\ddagger$ is defined by (6).

$$x^\ddagger = \tfrac{1}{2} + \Delta G^\circ/8\gamma \quad (5a)$$

$$y^\ddagger = \tfrac{1}{2} - \Delta G'/8\mu \quad (5b)$$

$$\Delta G^{\ddagger} = G^{\ddagger} - G(\mathrm{r}) = G(x^{\ddagger},y^{\ddagger}) - G(0,\tfrac{1}{2}) \quad (6)$$

On substituting in eq 4 and solving, one obtains eq 7.[3]

$$\Delta G^{\ddagger} = \gamma + \tfrac{1}{2}\Delta G^{\circ} + (\Delta G^{\circ})^2/16\gamma - (\Delta G')^2/16\mu \quad (7)$$

Note that eq 7 is a quadratic rather than a linear structure–energy relationship and thus can accommodate rate-constant minima. In many reaction series, ΔG° for the main reaction varies only slightly, while $\Delta G'$ for the disparity reaction varies substantially. If μ is sufficiently small, eq 7 simplifies to the approximate form (8). The operator δ in (8) denotes the change

$$\delta[\Delta G^{\ddagger}] \approx -\delta[(\Delta G')^2/16\mu] \quad (8a)$$

$$|\delta[\Delta G^{\ddagger}]| >> |\delta[\tfrac{1}{2}\Delta G^{\circ} + (\Delta G^{\circ})^2/16\gamma]| \quad (8b)$$

in the indicated free-energy function due to a change of reaction within the reaction series. When reactivity within a reaction series conforms to eq 8, it follows from the differential of (7) that $\Delta G^{\ddagger}$ goes through a maximum, and the rate constant goes through a minimum, when $\Delta G' \approx 0$.

Comparison of Theory and Observation. To test eq 7 one needs to know the reaction mechanism. For the base-catalyzed reaction, the mechanism of eq 2 is reasonably well established,[12c,14,20] even though part of the deduction makes use of ideas similar to those that lead to eq 7 and thus may not be truly independent of the present test. It is worth mentioning also that a powerful diagnostic tool has now been added, in that $\Delta G'$ must pass through zero near the rate-constant minimum. Of the mechanisms that I have considered, only that of eq 2 has $\Delta G'$ pass through zero. The rate-determining step of that mechanism is (9a). The corresponding disparity reaction (i → h, Figure 3) is (9b).

$$\mathrm{B^- + HOCH_2R + H_2C{=}O \rightarrow BH + RCH_2OCH_2O^-} \quad \Delta G^{\circ} \quad (9a)$$

$$\mathrm{B^- + RCH_2OH^+{\cdot}CH_2O^- \rightarrow BH + RCH_2O^- + H_2C{=}O} \quad \Delta G' \quad (9b)$$

ΔG° and $\Delta G'$ in (9) can be evaluated by combining appropriate free-energy changes defined in eq 10–16:

$$\mathrm{BH = H^+ + B^-} \quad \Delta G^{\mathrm{HB}} \quad (10)$$

$$\mathrm{RCH_2OH = H^+ + RCH_2O^-} \quad \Delta G^{\mathrm{HOR}} \quad (11)$$

$$\mathrm{RCH_2OCH_2OH = H^+ + RCH_2OCH_2O^-} \quad \Delta G^{1} \quad (12)$$

$$\mathrm{RCH_2OH^+{\cdot}CH_2OH = H^+ + RCH_2OCH_2OH} \quad \Delta G^{\mathrm{IH}} \quad (13)$$

$$\mathrm{RCH_2OCH_2OH = RCH_2OH^+{\cdot}CH_2O^-} \quad \Delta G^{\mathrm{Z}} \quad (14)$$

$$\mathrm{H_2C{=}OH^+ = H^+ + H_2C{=}O} \quad \Delta G^{\mathrm{K}} = -4.0 \text{ kcal} \quad (15)$$

$$\mathrm{RCH_2OH + H_2C{=}O = RCH_2OCH_2OH} \quad \Delta G^{\mathrm{hemi}} = -4.04 \text{ kcal} \quad (16)$$

Numerical values are listed in Table I. For reactions 10 and 11 they are the results of direct measurements, as summarized by FAJ.[14] For reactions 12–15 they are estimates based on data for model substrates, with substituent effects estimated by familiar linear free-energy methods. Except for the base value adopted for ΔG^{Z}, which was treated as a parameter of the least-squares fit,[21a] these values similarly were taken from the work of FAJ.[14,22a]

Table I. Standard Free-Energy Changes (kcal/mol) at 25 °C for Reactions 10–14[a] (alcohol = RCH_2OH; base catalyst $B^- = R'CH_2CO_2^-$)

	R =				
	CH_3	CH_3OCH_2	$ClCH_2$	Cl_2CH	CF_3
ΔG^{HOR}	21.82	20.22	19.52	17.58	16.96
ΔG^{1}	18.50[b]	18.18	18.03	17.65	17.52
ΔG^{IH}	−6.3[c]	−7.9	−8.6	−10.6	−11.2
ΔG^{Z}	15.4[d]	16.7	17.3	19.0	19.5
	R′ =				
	H	$ClCH_2$	CH_3O	Cl	CN
ΔG^{HB}	6.34	5.36	4.64	3.68	3.04

[a] See text for details. [b] $\Delta G^{1} = 14.1_3 + 0.200\Delta G^{\mathrm{HOR}}$. [c] $\Delta G^{\mathrm{IH}} = -28.1_6 + \Delta G^{\mathrm{HOR}}$. [d] $\Delta G^{\mathrm{Z}} = 33.9_3 - 0.851\Delta G^{\mathrm{HOR}}$.

Table II. Values (in kcal/mol) of ΔG°, $\Delta G'$, and $\Delta G_+^{\ddagger}$ for the Rate-Determining Step of Mechanism 2 (alcohol = RCH_2OH; base catalyst = $R'CH_2CO_2^-$)

R	R′	ΔG°	$\Delta G'$	$\Delta G_+^{\ddagger}$
CH_3	H	8.12	4.16	16.36
	$ClCH_2$	9.1	5.14	16.64
	CH_3O	9.82	5.86	16.87
	Cl	10.78	6.82	17.05
	CN	11.42	7.46	17.26
CH_3OCH_2	H	7.8	1.18	16.64
	$ClCH_2$	8.78	2.16	16.75
	CH_3O	9.5	2.88	17.18
	Cl	10.46	3.84	17.38
	CN	11.1	4.48	17.70
$ClCH_2$	H	7.65	−0.11	16.57
	$ClCH_2$	8.64	0.88	16.85
	CH_3O	9.36	1.6	17.11
	Cl	10.31	2.55	17.52
	CN	10.96	3.2	17.72
Cl_2CH	H	7.27	−3.7	16.09
	$ClCH_2$	8.26	−2.72	16.52
	CH_3O	8.98	−2.0	16.91
	Cl	9.93	−1.04	17.54
	CN	10.58	−0.4	17.85
CF_3	H	7.14	−4.86	15.57
	$ClCH_2$	8.13	−3.87	16.12
	CH_3O	8.85	−3.15	16.60
	Cl	9.8	−2.2	17.13
	CN	10.45	−1.55	17.49

The value of ΔG^{hemi} in (16) is based on direct measurements for R = CH_3. The same value was used for all alcohols in the present series, because preliminary results for R = CF_3 cited by FAJ[22b] and direct measurements for a similar series of hemithioacetals[23] suggest that substituent effects are likely to be small. In terms of the free-energy changes defined in eq 10–16, ΔG° and $\Delta G'$ for reactions 9a and 9b are given by:

$$\Delta G^{\circ} = -\Delta G^{\mathrm{HB}} + \Delta G^{1} + \Delta G^{\mathrm{hemi}} \quad (17a)$$

$$\Delta G' = -\Delta G^{\mathrm{HB}} + \Delta G^{\mathrm{HOR}} - \Delta G^{\mathrm{hemi}} - \Delta G^{\mathrm{Z}} \quad (17b)$$

Free energies of activation $\Delta G^{\ddagger}$ were calculated using the Eyring equation:

$$\text{rate constant} = (k_{\mathrm{B}}T/h)\exp(-\Delta G^{\ddagger}/RT) \quad (18)$$

The experimental rate constants k_r (eq 1) reported for the reverse reaction are formal quantities, independent of mechanism, and thus define a formal activation free energy $\Delta G_r^{\ddagger}$. The corre-

(21) (a) On the basis of model substrates, the value of ΔG^{Z} for $C_2H_5OCH_2OH$ is 16.4 kcal, instead of 15.4 kcal as listed in Table I. This estimation requires evaluation of $pK_1 - pK_2$ for the dibasic acid $C_2H_5{\cdot}OH^+{\cdot}CH_2OH$, but data for model substrates which have both the correct charge type and the correct O–C–O basic sites do not seem to exist. Theory and experience with acidity functions show that matching the basic site is important, while matching the charge type ν is less important because the electrostatic contribution to $pK_1 - pK_2$ is in first approximation proportional to $\nu^2 + (\nu - 2)^2 - 2(\nu - 1)^2 = 2$, which is independent of ν. The estimate of 16.4 kcal obtained for $C_2H_5OCH_2OH$ is based on essentially equal values of $pK_1 - pK_2$ reported for $(CF_3)_2C(OH)_2$ and $O{=}C(OH_2)$. (b) When ΔG^{Z} is not treated as an adjustable parameter but a base value of 16.4 kcal for $C_2H_5OCH_2OH$ is used, σ(fit) is 0.213 kcal and σ(data) is 0.273 kcal. Thus the fit of eq 7 is acceptable on this basis as well.

(22) (a) It is desirable to use one consistent set of free-energy values, even though some of the estimates may since have been improved. For instance, a value of −5.37 has been recommended for pK_a of formaldehyde. Cox, R. A.; Smith, C. R.; Yates, K. *Can. J. Chem.* **1979**, *57*, 2952. (b) Borcsok, E.; Kosiba, B. In ref 14, footnote 26.

(23) Gilbert, H. F.; Jencks, W. P. *J. Am. Chem. Soc.* **1977**, *99*, 7931.

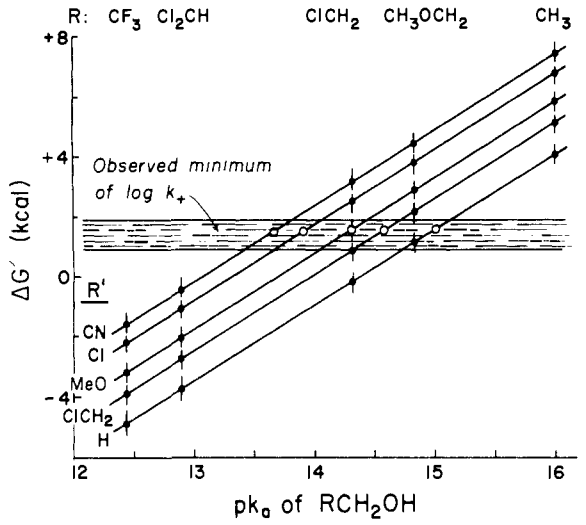

Figure 4. Plot of $\Delta G'$ for the rate-determining step of mechanism 2 vs. pK_a of the alcohol reagent. The base catalyst is $R'CH_2COO^-$. Log k_r goes through a minimum when $\Delta G' = 1.4 \pm 0.5$ kcal, as indicated by the horizontal band. Values of $\Delta G'$ at which log k_r is predicted to go through a minimum range from 1.49 to 1.57 kcal and are shown by open circles.

sponding formal quantity $\Delta G_f^{\ddagger}$ for the forward reaction is given by:

$$\Delta G_f^{\ddagger} = \Delta G_r^{\ddagger} + \Delta G^{hemi} \quad (19)$$

We are, however, interested in the mechanistically defined quantity $\Delta G_+^{\ddagger}$ based on the forward rate constant k_+ in eq 2a. Because reaction 2b in this mechanism is fast, the transition state for the overall formal reaction 1 is identical with that for reaction 2a. Thus $\Delta G_+^{\ddagger} = \Delta G_f^{\ddagger}$. Substitution in (19) then leads to;

$$\Delta G_+^{\ddagger} = \Delta G_r^{\ddagger} + \Delta G^{hemi} \quad (20)$$

Results thus obtained for $\Delta G°$, $\Delta G'$, and $\Delta G_+^{\ddagger}$ for an array of 25 reactions are listed in Table II. The values of $\Delta G°$ are all positive. Therefore in eq 7, the $\Delta G°$ terms alone cannot produce rate-constant minima. On the other hand, the values of $\Delta G'$ range from +7.5 to −4.9 kcal/mol. The $(\Delta G')^2$ terms in (7) go through a minimum whenever $\Delta G' = 0$ and thus will cause rate-constant minima.

Figure 4 shows a plot of $\Delta G'$ vs. pK_a of the alcohol reagent. While the positions of the minima vary with pK_a, within their 0.5-kcal/mol uncertainty they occur at practically equal values of $\Delta G'$, in good agreement with prediction based on eq 7 (open circles in Figure 4). The high quality of fit no doubt comes in part from treating ΔG^Z for $C_2H_5OCH_2OH$ as an adjustable parameter; see (17b). Yet this single parameter reproduces five minima. Moreover, the qualitative aspects of the fit would not be worsened by a possible 1–2 kcal/mol error in the base value adopted for ΔG^Z.[21b]

In the least-squares calculation of parameters for eq 7, I used the general method of Deming and the specific algorithm of Wentworth.[24] I used the data in Table II and the following estimates of standard errors: for $\Delta G_+^{\ddagger}$, 0.14 kcal (based on a standard error of 0.10 unit for log k_r); for $\Delta G°$ and $\Delta G'$, 0.4 kcal (based on the experience that equilibrium constants obtained by comparisons with chemical models are accurate to about a factor of 2). The following results were obtained:

$$\gamma = 12.10 \pm 0.08 \text{ kcal/mol}$$

$$\mu = 2.98 \pm 0.60 \text{ kcal/mol}$$

$$\Delta G^Z = 15.4 \pm 0.5 \text{ kcal/mol for } C_2H_5OCH_2OH$$

$$\sigma(\text{fit}) = 0.188 \text{ kcal/mol}$$

$$\sigma(\text{data}) = 0.273 \text{ kcal/mol}$$

In the above, σ(fit) is the standard error of fit of eq 7, and may be compared with σ(data), which expresses the inherent accuracy of the data. σ(data) is defined as follows. Let $y = f(x)$, and let $\{y_i,x_i\}$ denote the experimental data set, $i = 1$ to m. Then σ(data) is the average error of the difference function, $[y_i - f(x_i)]$, $i = 1$ to m. The latter is computed from standard errors in y_i and x_i by familiar formulas for the propagation of error. Note that $\sigma(\text{fit})/\sigma(\text{data}) < 1$, showing that the fit is better than required by the accuracy of the data. Note also that σ(fit) is only slightly greater than the standard error estimate for $\Delta G_+^{\ddagger}$, in spite of the fact that some of the free-energy changes (eq 10–16) which go into the theoretical prediction are relatively inaccurate.[21b]

Other Reaction Mechanisms. Equation 7 was applied to three other reaction mechanisms which cover a variety of rate-determining steps. The mechanism shown as (21) + (22) differs from (2) in that the rate-determining step (22) is acid-catalyzed addition of RCH_2O^- to the carbonyl group, rather than base-catalyzed addition of RCH_2OH. The formal rate law is one of general base catalysis, and the rate-determining step involves two progress variables:

$$B^- + RCH_2OH \overset{\text{fast}}{\rightleftharpoons} BH + RCH_2O^- \quad (21)$$

$$RCH_2O^- + H_2C{=}O + HB \underset{k_-'}{\overset{k_+'}{\rightleftharpoons}} RCH_2OCH_2OH + B^- \quad (22)$$

Mechanism 21 + 22 had previously been ruled out by FAJ, partly because some of the observed reactivities can be accommodated only by microscopic rate constants of physically impossible magnitudes.[14] In the present study, using free-energy data appropriate for this mechanism, the fit of eq 7 is decisively poor. σ(fit) is 1.1 kcal, four times greater than σ(data).

A brief description of the preceding calculation will explain the marked dependence on the reaction mechanism. For the rate-determining main reaction (22), the disparity reaction is $RCH_2OCH_2O^- + HB \rightarrow RCH_2O^- + H_2C{=}OH^+ + B^-$. $\Delta G° = \Delta G^{HB} - \Delta G^{HOR} + \Delta G^{hemi}$. $\Delta G' = \Delta G^{HB} + \Delta G^{HOR} - \Delta G^K - \Delta G^{hemi} - \Delta G^I$. $\Delta G_+^{\ddagger} = \Delta G_r^{\ddagger} + \Delta G°$. For mechanism 2, the corresponding equations are (9a,b), (17a,b), and (20).

Two other mechanisms, shown respectively as (23) + (24) and (23) + (25), involve simultaneous acid and base catalysis. Their rate-determining steps depend on three progress variables, labeled u, v, and w:

$$HO^- + HB \overset{\text{fast}}{\rightleftharpoons} HOH + B^- \quad (23)$$

$$B^- + HOCH_2R + H_2C{=}O + HOH \overset{\text{slow}}{\rightleftharpoons} BH + RCH_2OCH_2OH + OH^- \quad (24)$$

$$HO^- + HOCH_2R + H_2C{=}O + HB \overset{\text{slow}}{\rightleftharpoons} HOH + RCH_2OCH_2OH + B^- \quad (25)$$

These mechanisms are outside the scope of two-variable More O'Ferrall diagrams (and thus were not tested by FAJ), but can be examined by an extension of the present theory.[25] This has been done for (23) + (24). The fit, though not decisively poor, is significantly poorer than that of eq 2. σ(fit) is 0.36 kcal, 1.4 times greater than σ(data), and the least-squares equation does not predict the rate-constant minima. Acceptable intrinsic parameters for simultaneous acid and base catalysis are, however, available from studies of the acid-catalyzed reaction.[25] They predict that reaction according to (23) + (24) accounts for 0.03–3.4% of the total base-catalyzed rate, and that reaction according to (23) + (25) accounts for 10^{-3} to 10^{-4}%. It appears, therefore, that these reaction paths make only small contributions.

In the remainder of this paper I shall assume that eq 2 represents the dominant base-catalyzed reaction path.

Transition-State Coordinates. Using data presented earlier, transition-state coordinates $x^{\ddagger}$,$y^{\ddagger}$ for the rate-determining step

(24) (a) Deming, W. E. "Statistical Adjustment of Data"; Wiley: New York, 1943. (b) Wentworth, W. E. *J. Chem. Educ.* **1965**, *42*, 96.

(25) Grunwald, E., *J. Am. Chem. Soc.*, following paper in this issue.

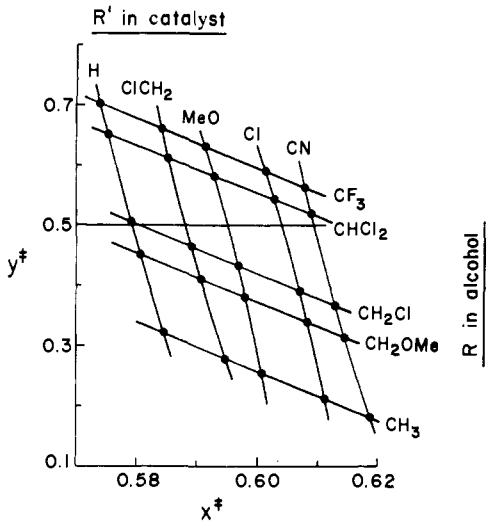

Figure 5. Transition-state coordinates projected from quadratic energy surfaces (eq 5) for the rate-determining step of mechanism 2. Note the difference in the scales of $x^{\ddagger}$ and $y^{\ddagger}$. When $y^{\ddagger} > 0.5$, proton transfer is ahead of C···O bond formation. When $y^{\ddagger} < 0.5$, C···O bond formation is ahead of proton transfer.

Table III. Transition-State Cordinates in the Rate-Determining Step of Mechanism 2

R[a]	R′	$x^{\ddagger}$	$y^{\ddagger}$	$u^{\ddagger}$[b]	$v^{\ddagger}$[c]
CH_3	H	0.584	0.325	0.758	0.409
	CH_3O	0.601	0.254	0.847	0.355
	CN	0.618	0.187	0.931	0.305
$ClCH_2$	H	0.579	0.505	0.574	0.584
	CH_3O	0.597	0.433	0.664	0.530
	CN	0.613	0.366	0.747	0.479
CH_3	H	0.574	0.704	0.370	0.778
	CH_3O	0.591	0.632	0.459	0.724
	CN	0.608	0.565	0.543	0.670

[a] Base catalyst = $R'CH_2COO^-$; alcohol = RCH_2OH. [b] O···C bond formation. [c] Proton transfer to $R'CH_2COO^-$.

(2a) of mechanism 2 were calculated from eq 5, and results obtained for the 25 reactions are mapped in Figure 5. Note that $x^{\ddagger}$ and $y^{\ddagger}$ are projected from quadratic energy surfaces. Thus $x^{\ddagger}$, the coordinate which measures mean progress, is not simply equal to a familiar progress variable such as change of bond order or change of electric charge; however, it varies monotonically with such variables. The disparity coordinate $y^{\ddagger}$ is defined in such a way (eq 3b) that the disparity ($v^{\ddagger} - u^{\ddagger}$) changes sign when $y^{\ddagger}$ crosses $^1/_2$.

Figure 5 shows that the variation of $x^{\ddagger}$ in the series is relatively small, and the dependence of $x^{\ddagger}$ on $\Delta G°$ conforms to Hammond's rule. On the other hand, the variation of $y^{\ddagger}$ is large; substituents in the alcohol reagent cause marked effects, and $y^{\ddagger}$ varies so much that the disparity between proton transfer and O···C bond formation even changes sign. When $y^{\ddagger} > 0.5$, proton transfer is ahead of O···C bond formation, at the transition state. When $y^{\ddagger} < 0.5$, O···C bond formation is ahead of proton transfer. These effects are clearly shown for representative reactions in Table III, which lists not only $x^{\ddagger}$ and $y^{\ddagger}$, but also the original progress variables $u^{\ddagger}$ and $v^{\ddagger}$; see Figure 3.

According to the present theory, all of the reactions plotted in Figure 5 proceed by the same mechanism. This is because the nature of the rate-determining reaction and of the disparity reaction remains constant, and because all rate constants are reproduced by eq 7 with the same pair of intrinsic parameters γ and μ. The change of sign of the disparity within the reaction series and the corresponding change in the nature of the transition states are fundamental, however, and may cause other criteria to indicate other than a constant reaction mechanism. For example, Figure 6 compares plots of log k_r for reaction 1 for a fixed series of alcohols but different pairs of catalysts. When the

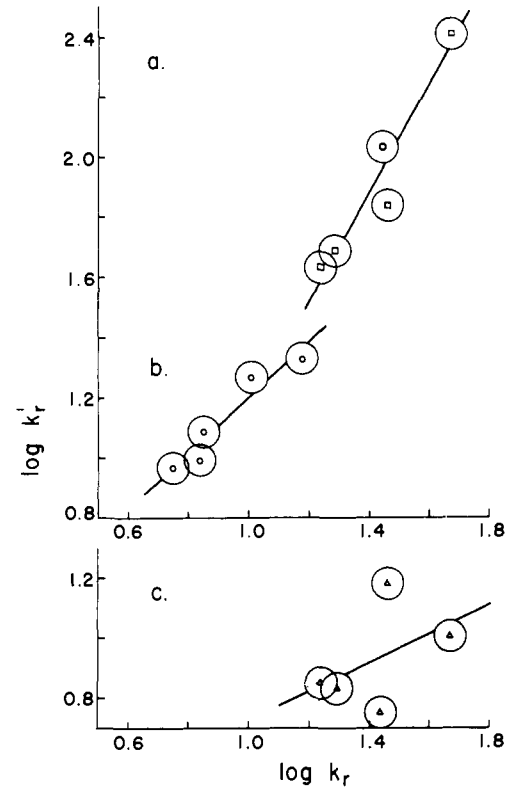

Figure 6. Plots of log k_r for reaction 1 for a fixed series of alcohols and the following pairs of base catalysts: (a) acetate (ordinate) and methoxyacetate; (b) chloroacetate (ordinate) and cyanoacetate; (c) cyanoacetate (ordinate) and methoxyacetate. Plots a and b are acceptably linear; plot c is a scatter diagram. Data are from ref 14.

catalysts are acetate and methoxyacetate, or chloroacetate and cyanoacetate, the plots are nearly linear, suggesting similarity of mechanism (Figure 6a,b). When the catalysts are cyanoacetate and methoxyacetate, however, the plot is a scatter diagram, suggesting a difference of mechanism (Figure 6c). The present theory avoids such ambiguity by treating reaction mechanism and disparity of reaction events as distinct variables.

Some Fitting Comments. So strong is the tradition of linear free-energy relations[1,2] that when a structure–energy plot can plausibly be fitted to a straight line, one tends to assume that the straight line is the true relationship. But there is no statistical necessity for this; the true relationship might be a gentle curve. This issue is relevant because relationships such as those plotted in Figure 1 have been identified as straight lines,[14] whereas the present theory requires that they be curves. In order to probe for possible inconsistency, I shall present a method for computing the signal/noise ratio (S/N) in the detection of curvature and show that, for the plots of Figure 1, the predicted curvature is too small to be detected.

Let $f(x)$ denote the theoretical curve, and let $\{y_i,x_i\}$ denote the experimental data set, $i = 1$ to m. As before, let σ(data) denote the average error of $[y_i - f(x_i)]$. Let $f_i = f(x_i)$, and compute the set $\{f_i,x_i\}$, $i = 1$ to m. We wish to find whether the set $\{f_i,x_i\}$, which represents the theoretical curve at the experimental points x_i, differs detectably, beyond the experimental noise level established by σ(data), from a straight line.

Let $L(x)$ be the equation of the straight line which gives best fit to $\{f_i,x_i\}$, and let $\sigma(L\backslash f)$ denote its standard error of fit. $\sigma(L\backslash f)$ measures the desired signal S, i.e., the deviation between the theoretical curve and the closest straight line, in the same sense that σ(data) measures the noise N. Since S/N for multiple experiments increases as $m^{1/2}$, where m denotes the number of experiments, the final expression for S/N in the detection of curvature is:

$$S/N = m^{1/2}\sigma(L\backslash f)/\sigma(\text{data}) \qquad (26)$$

Detectability depends on whether S/N is greater or less than unity.

As S/N decreases below unity, detection of curvature quickly becomes impossible.

For the data points plotted in Figure 1, $x = pK_a$ and $f(x) = \log k_r$. The latter was calculated from eq 7, using the parameters resulting from the least-squares fit for mechanism 2. Values of S/N computed in this way for the five plots shown in Figure 1 range from 0.6 to 0.7, the mean S/N being 0.66. This is well below unity. One may conclude, therefore, that the predicted curvature of the plots is too gentle to be detected.

A second fitting comment concerns solvation effects, or rather the number of adjustable parameters to be used. When the formation of a reactive encounter complex from the reagents needs to be treated as a separate prior equilibrium, Marcus[4] and others[6,8d,9,10b] have suggested modifications of the basic theory which would recast eq 7 in the form:

$$\Delta G^{\ddagger} = \gamma^{\circ} + \tfrac{1}{2}\Delta G^{\circ} + (\Delta G^{\circ})^2/16\mu^{\circ} - (\Delta G')^2/16\mu' \qquad (27)$$

There are now three adjustable parameters: γ°, μ°, and μ'. The original γ in eq 7 has been replaced by γ° and μ°: γ° gathers up all effects on reactivity that remain constant in the reaction series, while μ° gathers up all effects that vary as $(\Delta G^{\circ})^2$. Their values are different if certain physical effects, including solvation effects in the formation of reactive encounter complexes, become important. μ' corresponds to μ in eq 7.

The fit of the three-parameter eq 27 represents an improvement over that of eq 7. Using the same 25 data sets as before (Table II), least-squares results are: $\gamma^{\circ} = 12.66 \pm 0.2$ kcal; $\mu^{\circ} = 21.8 \pm 2.8$ kcal; $\mu' = 3.66 \pm 0.8$ kcal. σ(fit) = 0.137 kcal, which compares favorably with that of eq 7, 0.188 kcal, and is consistent with the 0.14-kcal experimental error of $\Delta G_+^{\ddagger}$. Because the formal change from (7) to (27) has a physical basis, the improvement of fit is not trivially due merely to the introduction of an additional parameter.

Despite this good result, I prefer eq 7 for the investigation of reaction mechanism. The physical basis of eq 7 is of broad scope, and although there are soft spots, the inherent approximations are understood. Using eq 7, transition-state coordinates can be simply deduced. Assuming that eq 2 represents the dominant mechanism, the fit of eq 7 is really quite good, and one ought not add to the pitfalls in deducing reaction mechanism by needlessly increasing the number of adjustable parameters.

Registry No. CH_2O, 50-00-0; CH_3CH_2OH, 64-17-5; $CH_3OCH_2CH_2OH$, 109-86-4; $ClCH_2CH_2OH$, 107-07-3; Cl_2CHCH_2OH, 598-38-9; CF_3CH_2OH, 75-89-8; CH_3COO^-, 71-50-1; $ClCH_2CH_2COO^-$, 5102-76-1; $CH_3OCH_2COO^-$, 20758-58-1; $ClCH_2COO^-$, 14526-03-5; $NCCH_2COO^-$, 23297-32-7.

Reaction Mechanism from Structure–Energy Relations. 2. Acid-Catalyzed Addition of Alcohols to Formaldehyde

Ernest Grunwald

Contribution from the Chemistry Department, Brandeis University, Waltham, Massachusetts 02254. *Received November 16, 1984*

Abstract: Previous theory of structure–energy relations is extended to mechanisms with three concerted reaction events. Data for 25 reactions (five alcohols, five acid catalysts) are examined on the basis of four different mechanisms. Only one of the mechanisms fits well. It involves concerted C···O bond formation, proton donation by the acid catalyst, and proton acceptance by a water molecule, according to $H_2O + HOCH_2R + H_2C{=}O + HOOCCH_2R' \rightarrow H_2OH^+ + RCH_2OCH_2OH + {}^-OOCCH_2R'$. The fit for this mechanism is decisively good in three practically independent tests. The contrast between the present mechanism and that which fits the base-catalyzed reaction (which does not involve HOH as a reagent) can be explained as due to peculiarities of relative acid–base properties of RCH_2OCH_2OH and H_2O. No inconsistency appears with other data. Transition-state coordinates for the acid-catalyzed reaction are tabulated. Progress of C···O bond formation, though variable within the reaction series, is well ahead of that of the proton transfers. Further analysis of the theoretical free-energy surfaces indicates that disparity of progress of the concerted reaction events reaches a maximum at the transition state.

This paper continues the use of structure–energy relations for deducing the mechanism of the addition of alcohols to formaldehyde. The preceding paper[1] (hereafter called part 1) examines the mechanism of general base catalysis, using a theory of reactivity[2] which applies when there is disparity of progress of two concerted reaction events. This theory, which develops Marcus rate–equilibrium theory[3] so as to quantify the use of More O'Ferrall diagrams,[4] predicts diagnostically different structure–energy relations for different reaction mechanisms. In the case of base catalysis, only one of the suggested mechanisms fits the data, and this mechanism agrees with that indicated by other methods.[1,5,6]

I shall now examine the mechanism of general acid catalysis. The specific reaction series is shown symbolically in eq 1, where R and R′ are variable substituents.

$$R'CH_2COOH + RCH_2OH + H_2C{=}O \underset{{}^{a}k_r}{\overset{{}^{a}k_f}{\rightleftharpoons}} RCH_2OCH_2OH + R'CH_2COOH \qquad (1)$$

I shall use a 5 × 5 matrix of data: five primary alcohols and five carboxylic acid catalysts, as reported by Funderburk, Aldwin, and Jencks (FAJ).[5] The experimental rate constants ${}^{a}k_r$ apply to the *reverse* reaction in eq 1, which was caused to go to completion by trapping the formaldehyde.

Brønsted plots of $\log {}^{a}k_r$ vs. pK_a of the acid catalyst are shown in Figure 1. Plots of $\log {}^{a}k_r$ vs. pK_a of the alcohol are shown in Figure 2. These figures also show best-fitting straight lines obtained by least squares. Most of the lines reproduce the data adequately. However, as pointed out before,[1] there is no statistical necessity for the real relationships to be straight lines. They may be gentle curves.

(1) Grunwald, E. *J. Am. Chem. Soc.*, preceding paper in this issue.
(2) Grunwald, E. *J. Am. Chem. Soc.* **1985**, *107*, 125.
(3) Marcus, R. A. *J. Phys. Chem.* **1968**, *72*, 891.
(4) More O'Ferrall, R. A. *J. Chem. Soc. B* **1970**, 274.
(5) Funderburk, L. H.; Aldwin, L.; Jencks, W. P. *J. Am. Chem. Soc.* **1978**, *100*, 5444.
(6) Palmer, J. L.; Jencks, W. P. *J. Am. Chem. Soc.* **1980**, *102*, 6472.

0002-7863/85/1507-4715$01.50/0