

Substitution Effects in Kinetically Controlled Reactions

By **Peter Luby**, Institute of Chemistry, Slovak Academy of Sciences, 809 33 Bratislava, Dúbravská cesta, Czechoslovakia

A simple method for determination of relative reactivity of functional groups of polyfunctional substances in the form of a kinetic rate constant ratio is shown to be defective especially for reactions carried out to a high degree of substitution. A general scheme of kinetic rate equations implying substitution effects by functional groups upon reactivity of unsubstituted ones is developed for an irreversible reaction of a trifunctional substance. The method is exemplified for Kuhn methylation of methyl β -D-xylopyranoside. Good agreement of the proposed model with experimental results is shown, which facilitates the expression of substitution effects by virtue of free activation energy changes. A convenient method of numerical analysis of experimental results is developed which simplifies complicated least squares procedures considerably.

QUANTITATIVE evaluation of relative reactivities of functional groups of a polyfunctional substance given by a simple relation of kinetic rate constants appears to be imperfect for several types of reactions. More additional parameters are necessary for the total description of kinetic behaviour of such a substance than a few rate constants whose number would be equal to that of the

functional groups. The conversion rate of a group often depends on whether or not the other groups of the same molecule have reacted. Kinetic equations should therefore be fitted by coefficients allowing for the influence of reacted functional groups of the same molecule,

Etherification of carbohydrates may be taken as an

example. For example the reactivity of the 3-hydroxy-group in methylation of xylan¹ and amylose² in the presence of aqueous alkali is doubled, once the 2-hydroxy-group has been methylated. A similar effect was observed for the reaction of some glucose derivatives with dimethyl sulphate² and with 2-(diethylamino)ethyl hydrogen sulphate.³ Even larger changes in activities of hydroxy-groups of glucosides caused by the individual substitution of the 2- and 3-hydroxy-groups have been found in the reaction with *NN*-dimethylaziridinium chloride.⁴

An effect of the change in free activation energy $\Delta G_{i,j}$ induced by conversion of other groups of the same polyfunctional molecule is usually called the substitution effect⁵ (s.e.) or neighbouring group effect.⁴

A special case of s.e., *viz.* when the change $\Delta G_{i,j}$ is proportional to the number of reacted functional groups and called the linear substitution effect, has been solved theoretically for a kinetically controlled reaction.⁵ A generalization of such a mathematical treatment consisted in admitting that $\Delta G_{i,j}$ need not be proportional to the number of reacted functional groups.⁶ The simplifying assumption of superposition of individual increments of $\Delta G_{i,j}$ as well as the principle of equal initial reactivity of functional groups still remained in the latter approach. Thus the theory is not applicable *e.g.* to kinetically controlled reactions of carbohydrates.

A kinetic free of the aforementioned simplifications describing the behaviour of a substance entering the reaction with as much as three functional groups is developed in the present paper. Comparison of the theory with experimental results⁶ gives the s.e. contributions to the standard free energy of forming any species of the reaction mixture for Kuhn methylation of β -D-methylxyloside. A modification of the least squares method for evaluation of experimental curves is suggested for this purpose.

THEORY

Regarding a substance with three functional groups a—c whose reactivity need not be mutually independent, one should consider nine distinct s.e. coefficients expressing this dependence, *viz.* $N_{a,b}$, $N_{a,c}$, $N_{b,a}$, $N_{b,c}$, $N_{c,a}$, $N_{c,b}$, N_{ab} , N_{ac} , N_{bc} , where the subscripts separated by commas refer to both the substituted functions and to those entering the reaction. They are defined by the respective change of free activation energy by relationships (1a and b) *etc.* $\Delta G_{a,b}$ Denotes an increment of the

$$N_{a,b} = \exp - (\Delta G_{a,b}/RT) \quad (1a)$$

$$N_{ab} = \exp - (\Delta G_{ab}/RT) \quad (1b)$$

free activation energy of substitution of a trifunctional

substance to the second degree, when the functional group a has already been substituted and a new substitution is occurring on function b. Similarly, ΔG_{ab} refers to the third substitution taking place on c. The choice of a reference value for the above increments follows from the definitions; it is the free activation energy $G_{0,i}$ of the corresponding conversion of function *i* which takes place on the unsubstituted species at the temperature of the reaction system. Hence equation (2)

$$\Delta G_{a,b} = G_{a,b} - G_{0,b} \quad (2)$$

applies where $G_{a,b}$ denotes the free activation energy of conversion of function b, taking place on a species in which a is already converted.

S.e. coefficients are taken as correction factors to the rate constants for first- or second-order equations, the latter taking the forms (3)—(6). k'_a , k'_b , and k'_c are

$$ds_0/dt = -(k'_a + k'_b + k'_c)s_0s_R \quad (3)$$

$$ds_a/dt = k'_a s_0 s_R - (k'_b N_{a,b} + k'_c N_{a,c})s_a s_R \quad (4)$$

$$ds_{ab}/dt = (k'_a N_{b,a} s_b + k'_b N_{a,b} s_a - k'_c N_{ab} s_{ab})s_R \quad (5)$$

$$ds_{abc}/dt = (k'_a N_{bc} s_{bc} + k'_b N_{ac} s_{ac} + k'_c N_{ab} s_{ab})s_R \quad (6)$$

the absolute rate constants of the conversion of functionalities a, b, and c, respectively, s_0 denotes the molar fraction of the parent (unsubstituted) substance, and s_R the molar fraction of reagent. Concentrations of the remaining substances s_a , s_{ab} , s_{abc} , *etc.* are indexed according to the respective functional groups reacted; all of them are related to the initial molar amount of the substrate, *i.e.* $s_0(t=0) = 1$, where *t* is time. Expressions for ds_b/dt , ds_c/dt , ds_{ac}/dt , and ds_{cb}/dt may be obtained by simple variation of the subscripts in equations (4) and (5).

The introduction of dimensionless magnitudes (7)—(9)

$$k_a = k'_a / (k'_a + k'_b + k'_c) \quad (7)$$

$$k_b = k'_b / (k'_a + k'_b + k'_c) \quad (8)$$

$$k_c = k'_c / (k'_a + k'_b + k'_c) \quad (9)$$

for the rate constants is more convenient for the following treatment.

A routine solution of the rate equations (see refs. 5 and 6) requires elimination of time by dividing them by equation (3) and subsequent integration. It stands to reason that whether the reaction order is first or second is immaterial because s_R cancels. Thus equations (10)—(16) are produced with the definitions (17)—(22).

This conception represents a generalization of the

¹ I. Croon and T. E. Timmel, *J. Amer. Chem. Soc.*, 1960, **82**, 247.

² I. Croon, *Acta Chem. Scand.*, 1959, **13**, 1235.

³ E. J. Roberts and S. P. Rowland, *Carbohydrate Res.*, 1967, **5**, 1.

⁴ E. J. Roberts, C. P. Wade, and S. P. Rowland, *Carbohydrate Res.*, 1972, **21**, 357.

⁵ M. Gordon and G. R. Scantlebury, *Trans. Faraday Soc.*, 1964, **60**, 604.

⁶ P. Luby, *J. Phys. Chem.*, 1974, **78**, 1083.

principles of a kinetically controlled s.e.⁵ in the following ways.

The initial reactivity of functional groups need not be equal, *i.e.* $k_a \neq k_b \neq k_c$. The values of the free activation energy changes need not be proportional to the

$$s_a = k_a s_0 (s_0^{x_a-1} - 1) / (1 - x_a) \quad (10)$$

$$s_b = k_b s_0 (s_0^{x_b-1} - 1) / (1 - x_b) \quad (11)$$

$$s_0 = k_c s_0 (s_0^{x_c-1} - 1) / (1 - x_c) \quad (12)$$

$$s_{ab}/k_a k_b = (s_0^{x_a} - s_0^{y_{ab}}) N_{a,b} / (1 - x_a)(y_{ab} - x_a) + (s_0^{x_b} - s_0^{y_{ab}}) N_{b,a} / (1 - x_b)(y_{ab} - x_b) + (s_0 - s_0^{y_{ab}}) [N_{a,b}(1 - x_b) + N_{b,a}(1 - x_a)] / (1 - x_a)(1 - x_b)(1 - y_{ab}) \quad (13)$$

$$s_{ac}/k_a k_c = (s_0^{x_a} - s_0^{y_{ac}}) N_{a,c} / (1 - x_a)(y_{ac} - x_a) + (s_0^{x_c} - s_0^{y_{ac}}) N_{c,a} / (1 - x_c)(y_{ac} - x_c) + (s_0 - s_0^{y_{ac}}) [N_{a,c}(1 - x_c) + N_{c,a}(1 - x_a)] / (1 - x_a)(1 - x_c)(1 - y_{ac}) \quad (14)$$

$$s_{bc}/k_b k_c = (s_0^{x_b} - s_0^{y_{bc}}) N_{b,c} / (1 - x_b)(y_{bc} - x_b) + (s_0^{x_c} - s_0^{y_{bc}}) N_{c,b} / (1 - x_c)(y_{bc} - x_c) + (s_0 - s_0^{y_{bc}}) [N_{b,c}(1 - x_c) + N_{c,b}(1 - x_b)] / (1 - x_b)(1 - x_c)(1 - y_{bc}) \quad (15)$$

$$s_{abc} = 1 - s_0 - s_a - s_b - s_c - s_{ab} - s_{ac} - s_{bc} \quad (16)$$

$$x_a = N_{a,b} k_b + N_{a,c} k_c \quad (17)$$

$$x_b = N_{b,a} k_a + N_{b,c} k_c \quad (18)$$

$$x_c = N_{c,a} k_a + N_{c,b} k_b \quad (19)$$

$$y_{bc} = N_{bc} k_a \quad (20)$$

$$y_{ac} = N_{ac} k_b \quad (21)$$

$$y_{ab} = N_{ab} k_c \quad (22)$$

number of substituted functional groups. Moreover, the principle of their superposition, *e.g.* $\Delta G_{ab} = \Delta G_{a,c} + \Delta G_{b,c}$, is abandoned.

Numerical Analysis.—An ordinary solution of the integrated rate equations for the rate constants seems to be the most obvious way for their calculation, as was done *e.g.* by Staněk *et al.*⁷ for a bifunctional system. However, such a method allows only calculation for one experimental point corresponding *e.g.* to the final distribution of the reaction products. High accuracy measurements are therefore required to assure coherence of the results with other points on the kinetic curves. Moreover, parallel experiments are necessary to obtain additional information on the intermediate conversion rate, so that the solution is unique. The availability of the intermediate products in amounts sufficient for high precision kinetic measurements is therefore unavoidable.

Another possibility consists in evaluation of complete kinetic curves by the least squares method. However, the complexity of the problem increases exponentially with the number of functional groups f . For example

the number of possible reaction components is a binomial function 2^f . The number of parameters describing the system at all increases even more quickly [equation (23)]. For instance a system of 12 equations

$$p = f2^f/2 \quad (23)$$

comprising 12 unknowns (three rate constants, nine s.e. coefficients) must be solved for a trifunctional system. This would be an unusually tedious task.

Fortunately, such a system can be split in two, comprising six and nine independent equations, respectively. Suffixes i and $j = 2, 3$, or 4 denoting functional groups of methyl β -D-xyloside (the substrate of Kuhn methylation, chosen as an example) will be used further. The simplification consists in a condition that normal equations are solved for k_2, k_3, k_4, x_2, x_3 , and x_4 so that the fit of curves s_2, s_3 , and s_4 is the best possible, instead of more rigorous requirement relating this condition to all values, s_{23}, s_{24} , and s_{34} included. Such a simplification causes negligible departures of the theoretical curves s_{23}, s_{24} , and s_{34} and agreement with experiment is good as shown later.

The first six normal equations take the general forms (24) and (25) where equation (26) applies, \bar{s}_{ni} being the

$$\sum_{n=1}^m \sum_{i=2}^4 \partial Q_{ni} / \partial k_j' = 0 \quad (24)$$

$$\sum_{n=1}^m \sum_{i=2}^4 \partial Q_{ni} / \partial x_j = 0 \quad (25)$$

$(j = 2, 3, \text{ or } 4)$

$$Q_{ni} = (\bar{s}_{ni} - s_{ni})^2 \quad (26)$$

experimental value for the n th measurement, s_{ni} the corresponding theoretical value [equations (10)—(15)], and m the total number of experimental points. Lest the expressions are too complicated we avoid double summations by solving the system for a particular case, $m = 1$.

Taking into account relationships (7)—(9), (10)—(12), and (26), the partial derivative in (24) can be expressed as in (27).

$$\partial Q_i / \partial k_j' = (\partial Q_i / \partial s_i) (\partial s_i / \partial k_i) (\partial k_i / \partial k_j') \quad (27)$$

Owing to the two degrees of freedom which follow from the arbitrary values of the N coefficients [equations (17)—(19)], x_i is independent of k_i and s_i may therefore be considered as a linear function of k_i . Performing

$$\partial s_i / \partial k_i = s_i / k_i \quad (28)$$

partial derivations of the right hand side terms of (27) we obtain (29) and (30) in accord with (7)—(9), (26), and (28). Insertion of (29) and (30) in (24) provides (32) after

$$(\partial Q_i / \partial k_j')_{i \neq j} = 2(\bar{s}_i - s_i) s_i / k' \quad (29)$$

$$\partial Q_j / \partial k_j' = -2(\bar{s}_j - s_j) s_j (1 - k_j) / k_j k' \quad (30)$$

$$k' = k_2' + k_3' + k_4' \quad (31)$$

⁷ J. Staněk, P. Chuchvalec, K. Čapek, K. Kefurt, and J. Jarý, *Carbohydrate Res.*, 1974, **36**, 273.

rearrangement. Apparently, the right-hand side of (32)

$$(\bar{s}_j - s_j)s_j/k_j = \sum_{i=2}^4 (s_i - s_i)s_i \quad (32)$$

is independent of j ; returning to a general case $m > 1$ for $j = 2, 3$, or 4 , we get equation (33). The general form of the next three normal equations follows from (10)—(12)

$$\begin{aligned} \sum_{n=1}^m (\bar{s}_{n2} - s_{n2})s_{n2}/k_2 &= \sum_{n=1}^m (\bar{s}_{n3} - s_{n3})s_{n3}/k_3 \\ &= \sum_{n=1}^m (\bar{s}_{n4} - s_{n4})s_{n4}/k_4 \quad (33) \end{aligned}$$

and (25). Two equations of type (33) and three of (34)

$$\sum_{n=1}^m (\bar{s}_j - s_{nj})(s_{nj} + k_j s_{n0} x_j \ln s_{n0}) = 0 \quad (34)$$

along with the trivial relation (35) complete the system of six independent equations with six unknowns.

$$k_2 + k_3 + k_4 = 1 \quad (35)$$

Normal equations for disubstituted species may be obtained by analogy. Substituting (37) in (36), taking

$$\begin{aligned} \sum_{n=1}^m (\partial Q_{ij}/\partial N_{i,j})_n &= \sum_{n=1}^m (\partial Q_{ij}/\partial N_{j,i})_n \\ &= \sum_{n=1}^m (\partial Q_{ij}/\partial N_{ij})_n = 0 \quad (36) \end{aligned}$$

$$Q_{ij} = (\bar{s}_{ij} - s_{ij})^2 \quad (37)$$

into account equations (13)—(15) we obtain (38)—(40).

$$\sum_{n=1}^m [(\bar{s}_{ij} - s_{ij})\partial s_{ij}/\partial N_{i,j}]_n = 0 \quad (38)$$

$$\sum_{n=1}^m [(\bar{s}_{ij} - s_{ij})\partial s_{ij}/\partial N_{j,i}]_n = 0 \quad (39)$$

$$\sum_{n=1}^m [(\bar{s}_{ij} - s_{ij})\partial s_{ij}/\partial N_{ij}]_n = 0 \quad (40)$$

Both (38) and (39) are linear relations between $N_{i,j}$ and $N_{j,i}$. Equation (40) represents N_{ij} as a quadratic function of $N_{i,j}$ or $N_{j,i}$.

Parameters determining relative reactivity of functional groups of methyl β -D-xylopyranoside during Kuhn methylation⁸

Hydroxy-group	Kinetic rate constant	S.e. coefficients		Increment of the free activation energy at 298 K (cal mol ⁻¹)	
		Second degree	Third degree	Second degree	Third degree
2-OH	0.4032	$N_{3,2}$ 0.8136	$N_{34} \infty$	$\Delta G_{3,2}$ 122	$\Delta G_{34} -\infty$
		$N_{4,2}$ 1.45		$\Delta G_{4,2}$ 220	
		$N_{2,3}$ 0.5656	N_{24} 0.5517	$\Delta G_{2,3}$ 337	ΔG_{24} 352
3-OH	0.261	$N_{4,3}$ 0.2316		$\Delta G_{4,3}$ 866	
		$N_{2,4}$ 1.2736	N_{23} 0.9643	$\Delta G_{2,4}$ -143	ΔG_{23} 21.5
		$N_{3,4}$ 0.4787		$\Delta G_{3,4}$ 436	

Solving the system of (38)—(40) we find three constants, say $N_{a,b}$, $N_{b,a}$, and N_{ab} . Following values $N_{a,c}$

⁸ Yu. S. Ovodov and E. V. Evtuschenko, *Carbohydrate Res.*, 1973, **27**, 169.

and $N_{b,c}$ are determined by (17) and (18). The remaining pair of the first degree s.e. coefficients $N_{c,a}$ and $N_{c,b}$ may be found from (19) and either (38) or (39). After all N_{ac} and N_{bc} are unambiguously given by equations of the type (40). Obviously, three equations of the type (40), two of (39), one (38) together with (17)—(19) form a system of nine independent relations, wherefrom all the s.e. coefficients may be calculated.

Application to Experimental Results.—Because of the comprehensive data on Kuhn methylation of methyl

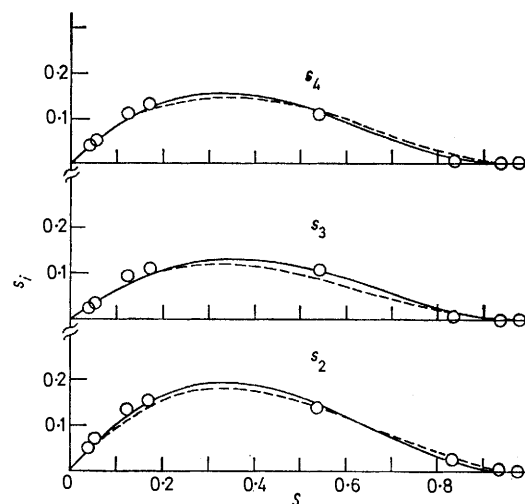


FIGURE 1 Kuhn methylation of methyl β -D-xyloside: molar fractions of monosubstituted species s_2 , s_3 , s_4 respectively, plotted against the mean conversion degree. Points, experimental results from Table III of ref. 8; full lines, values listed in the Table; dashed lines, values for the ratio $k_2 : k_3 : k_4 = 0.3907 : 0.2803 : 0.329$, s.e. disregarded

β -D-xylopyranoside (Table III of ref. 8) even for an advanced state of reaction, complete information concerning the character of the s.e. may be obtained for this reaction. Numerical analysis resulted in a set of rate constants and s.e. coefficients. Their values are listed in the Table along with the related values of free activation energy changes. Theoretical course of the reaction corresponding to these values is compared with experimental values by the full lines in Figures 1—3, where molar fractions of various species are plotted

against the mean substitution degree S given by equation (41). The best approximation made possible by ordinary

$$S = (s_2 + s_3 + s_4 + 2s_{23} + 2s_{24} + 2s_{34} + 3s_{234})/3 \quad (41)$$

rate equations (with the s.e. disregarded) is compared with the same figures on dashed lines corresponding to the ratio $k_2 : k_3 : k_4 = 0.3907 : 0.2803 : 0.3290$. A striking

deviation is shown by the curve s_{24} and also s_{34} . The latter should be identical with the ordinate axis.

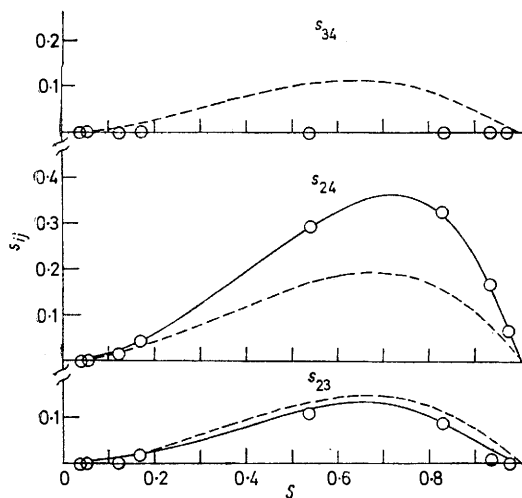


FIGURE 2 Kuhn methylation of methyl β -D-xyloside: molar fractions of disubstituted species s_{23} , s_{24} , s_{34} respectively, plotted against the mean conversion degree. Points, experimental results from Table III of ref. 8; full lines, values listed in the Table; dashed lines, values for the ratio $k_2 : k_3 : k_4 = 0.3907 : 0.2803 : 0.329$, s.e. disregarded

ing deviation is shown by the curve s_{24} and also s_{34} . The latter should be identical with the ordinate axis.

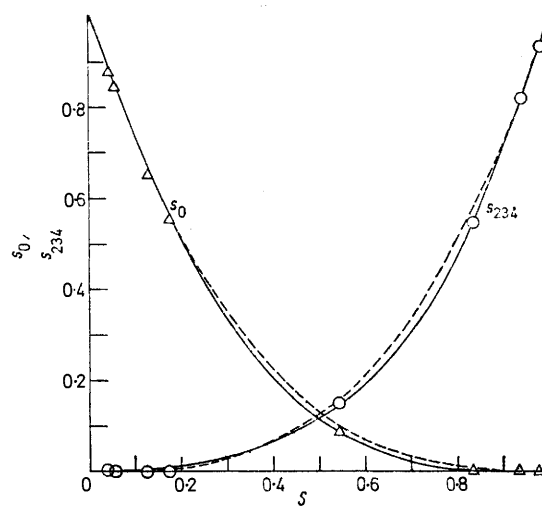


FIGURE 3 Kuhn methylation of methyl β -D-xyloside: molar fractions of parent s_0 and fully substituted species s_{234} respectively, plotted against the mean conversion degree. Points, experimental results from Table III of ref. 8; full lines, values listed in the Table; dashed lines, values for the ratio $k_2 : k_3 : k_4 = 0.3907 : 0.2803 : 0.329$, s.e. disregarded

rate equations which is insufficient especially at higher degrees of substitution.

[4/1738 Received, 19th August, 1974]