conditions, consisting of $(M-1)^+$ m/z 153 as well as m/z 137 which is the base peak in the spectrum[9]. Neither of these 2 ions (m/z 153, 137) are present in the methane CI spectrum of oct-2-enal. Similar mass-chromatography carried out under EI conditions gave supportive evidence for the presence of linalool in *D. superstitiosus* as the EI-spectrum of the sample from the cotton stainer (fig. 2B) is virtually superimposable on that of authentic linalool (fig. 2A).

Linalool was not detected in the dorsal abdominal scent glands of the larvae of *D. superstitiosus*. It was absent from any other parts of the larvae and adult insect when examined under our GC-MS conditions; its production is evidently confined exclusively to the adult metathoracic scent gland. It was not found in either the metathoracic or the 3rd abdominal scent gland of *Pyrrhocoris apterus* L., which indicates that it is not universally present in the species of Pyrrhocoridae. The larval glands of *D. superstitiosus* were found to contain 4-oxo-oct-2-enal, 4-oxo-hex-2-enal, oct-2-enal, tridecane, hex-2-enal, hexanal but no linalool.

Discussion. Although it is produced by the cotton plant[10], it is considered unlikely that linalool from a dietary source is taken up preformed and accumulated in the metathoracic scent gland of cotton stainers. It is most probably produced by a conventional isoprenoid pathway in the tissue of the insect. We suggest that linalool has some unique role to play in the ecology of cotton stainers of the genus *Dysdercus*.

Possible roles for the scent volatiles of *Dysdercus* (defence, alarm, cannibal, aggregating and sex attractant behaviours) have been suggested[6,11].

1 To whom correspondence should be addressed. We thank the British Inter Universities Council and Bayero University Research and Higher Degrees Committee for their financial support, Dr B.W. Staddon for his valuable assistance, Dr D.E. Games for arranging mass spectral facilities.

2 V.M. Dazzini and M. Pavan, Publ. Inst. Ent. Univ. Pavia *5*, 1 (1979).
3 B.W. Staddon, Adv. Insect Physiol. *14*, 351 (1979).
4 H. Remold, Nature *198*, 764 (1963).
5 D.H. Callam and G.C. Scott, J. Insect Physiol. *15*, 1695 (1969).
6 I.J. Everton, D.W. Knight and B.W. Staddon, Comp. Biochem. Physiol. *63B*, 157 (1979).
7 J.R. Aldrich, M.S. Blum, H.Lloyd and H.M. Fales, J. chem. Ecol. *4*, 161 (1978).
8 B.W. Staddon, I.J. Everton and D.E. Games, Comp. Biochem. Physiol. *62B*, 259 (1979).
9 D.W. Knight, M.E. Knight, M. Rossiter, A. Jones and D.E. Games, Adv. Mass-Spectrometry *8*, 1116 (1980).
10 P.A. Hedin, Envir. Ent. *5*, 1234 (1976).
11 D.H. Callam and A. Youdewei, J. Insect Physiol. *14*, 1147 (1968).

The composition of Taft E_s^o constants: An application of biased alternatives to least squares

H. Mager[1] and A. Barth[2]

Academy of Sciences of the GDR, Institute of Plant Biochemistry, DDR-4010 Halle/Weinberg (German Democratic Republic), and University of Halle, Department of Biochemistry, Section of Biosciences, DDR-4010 Halle/Weinberg (German Democratic Republic), 26 January 1981

Summary. A systematic analysis of the composition of E_s^o like rate constants clearly revealed that Taft E_s^o-values depend upon the size of the substituents. Further evidence in favor of this view is adduced even in a case where OLS led to the conclusion that E_s^o should be completely independent of the size of the substituents, since biased estimators (PCRA, LRRA) showed that this statement is not correct. Furthermore, it seems that the magnitude of the steric effect represented by E_s^o is a function of the thickness of the substituent along 2 directions perpendicular on its main axis and is not influenced by its length.

An essential prerequisite of every quantitative approach to biological structure-activity (QSAR) or chemical structure-reactivity (LFER) relationships is the development of reliable substituent constants. According to Verloop, Hoogenstraaten and Tipker[3], the development of steric substituent constants is still in its early stages. The steric substituent constants E_s^o have been defined by Taft[4] on the basis of the acidic hydrolysis and esterification of ortho substituted benzoates and benzoic acids, respectively. In 1969, Charton[5] reexamined E_s-values for aliphatic and aromatic systems and concluded that the aliphatic E_s-values are linear functions of the corresponding van der Waals radii r_v, whereas E_s^o should not in any way depend upon the substituent's size. However, it has been demonstrated recently that the interpretation of Taft E_s^o constants in terms of predominating inductive, mesomeric, and steric contributions depends at least partly upon the sample structure, although in nearly all cases about 80% of the variation of E_s^o could have been accounted for by the van der Waals radius[6]. Furthermore, the independent variables used in such correlations are often multicollinearly related, a fact from which a lot of problems may arise if an interpretation is intended. In order to reduce the uncertainty associated with the magnitudes and even the signs of the regression coefficients if multicollinearity is present in the data, the application of biased estimators[7-11] may be preferable. Our work was directed towards further clarifying the meaning of E_s^o in terms of inductive, mesomeric, and steric influences. In order to keep the problems of interpretation to a minimum, solvent effects onf E_s^o have been disregarded and, furthermore, instead of E_s^o the rate constants of the basic reactions: a) esterification of benzoic acids by methanol/HCl at 25 °C and 40 °C[12,13] and b) esterification of

Table 1. Some measures of the steric substituent effect[a]

Group	z_1	E_s(corr.)	E_s^o
H	0.2223	0.79	–
F	0.3300	0.42	0.49
I	−0.5054	−0.74	−0.20
OMe	0.9819	0.34	0.99
OEt	0.9146	0.29	0.90
Me	−0.3459	0.00	0.00
NO_2	−1.2447	−0.87	−0.75
Br	−0.2716	−0.42	0.00
Cl	−0.0811	−0.17	0.18

[a]For notations and references see text.

Table 2. Relationships between measures of the steric substituent effect and other substituent constants[a]

Set	Method	$\mathfrak{F}$	$\mathfrak{R}$	r_v	x_7	L	B_1	B_2	B_3	B_4	$B_{1,max}$[b]	R^2 [b]
1-1	OLS	0.119[c]	−2.111	−1.356	–	–	–	–	–	–	0.5740	0.9508
1-2	OLS	0.000	−2.133	−1.243	–	–	–	–	–	–	0.1818	0.9488
1-3	OLS	0.454	−0.272[c]	−1.300	0.765	–	–	–	–	–	0.9019	0.9926
1-4	OLS	0.500	0.000	−1.308	0.861	–	–	–	–	–	0.5734	0.9918
1-5	OLS	–	–	–	–	−6.260	16.693	−4.700	−13.025	12.214	0.9995	0.9922
1-6	OLS	0.213	–	–	0.783	−3.139	8.114	−2.671	− 6.183	5.907	>0.9999	1.0000
1-7	LRRA	0.474	–	–	1.062	0.000	−0.478	0.000	− 0.440	0.000	–[d]	0.9968
2-1	OLS	−0.562	−2.621	0.000	–	–	–	–	–	–	0.8450	0.9727
2-2	PCRA	0.000	−1.235	−1.230	–	–	–	–	–	–	–[d]	0.8971
3-1	OLS	−0.005[c]	−0.047[c]	−1.847	–	–	–	–	–	–	0.5740	0.9742
3-2	OLS	0.000	0.000	−1.888	–	–	–	–	–	–	–	0.9742

[a]A stroke indicates that this variable was not included in the basis set, while 0.000 denotes a nonsignificant variable (5% level) after the reduction step. All other regression coefficients are significant at the 5% level. [b]$B_{I,max}$ denotes the highest coefficient of internal determination in the set, which is a suitable measure of multicollinearity and R^2 is the squared multiple correlation coefficient. [c]Not significant at the 5% level. [d]Not computed, since biased estimators have been applied.

benzoic acids by cyclohexanol/HCl at 55 °C and 65 °C[14] have been used. Since there is a high correlation among the rate constants, a principal component analysis[7,15] was performed. The 1st eigenvalue of the correlation matrix explains more than 99% of the variability among the rate constants ($l_1 = 3.9638$, total standardized variability: 4.0000). Therefore, it seemed us to be reasonable to use the 1st principal component z_1 as an overall measure of the effects common to both reactions. Indeed, it has been shown that the results obtained using the 1st principal component of the corresponding variance-covariance matrix are very similar to those obtained using E_s^o only, but often the goodness of fit is improved[16]. The 1st principal component of the correlation matrix will be mainly determined by such factors which simultaneously influence all reactions. In order to evaluate the relative contributions of inductive, mesomeric, and steric parameters to the total substituent effect, equation (1) was applied

$$y = b_0 + b_1\mathfrak{F} + b_2\mathfrak{R} + b_3 r_v + b_4 x_7 + b_5 L + b_6 B_1 + b_7 B_2 + b_8 B_3 + b_9 B_4 \quad (1)$$

with

$y = z_1, E_s^o, E_s$ (corr.),

where $\mathfrak{F}$ and $\mathfrak{R}$ denote the substituent constants of Swain and Lupton[17], x_7 is a dummy variable accounting for 'special effects' of OMe, OEt and NO_2 groups (concerning the interpretation of x_7 see Mager[16,18]) indirectly related to their resonance capacities, and L, B_1 to B_4 are steric parameters defined elsewhere[3]. In order to reduce the probability of misspecification due to multicollinearity, additionally to ordinary least squares (OLS) the following estimators were applied: a) latent root regression analysis[7,8] (LRRA), b) principal component regression analysis[7,8] (PCRA). The values used for the dependent variables are displayed in table 1 and the results of our correlations are summarized in table 2 (in every case only a subset of eqn (1) was used as a starting equation).

From the results presented in table 2 the following conclusions may be drawn:

1. If the variable pool consists only of $\mathfrak{F}$, $\mathfrak{R}$, and r_v, the total variability of z_1 is best explained by a linear combination of $\mathfrak{R}$ and r_v, i.e., z_1 was found to be dependent upon steric 'buik' (sets 1-1, 1-2). Thus, by virtue of the same arguments as applied previously[16] E_s^o-values should not be independent of the size of the substituents.

2. Although at least F, Cl, and Br should exert remarkable mesomeric effects, $\mathfrak{R}$ may be replaced by the dummy variable x_7 with

$$x_7 = \begin{cases} +1 \text{ for OMe, OEt} \\ -1 \text{ for } NO_2 \\ 0 \text{ for the other substituents} \end{cases}$$

which simultaneously yields a much improved fit (sets 1-3, 1-4). Therefore, we suppose that the significance of the mesomeric substituent constant $\mathfrak{R}$ in sets 1-1 and 1-2, respectively, does not represent a pure resonance effect. It is concluded that OMe, OEt, and NO_2 groups act via some unknown mechanism. But this problem is clearly outside the range of questions which can be treated by simple application of LFER's, and speculations about hindered solvation of the excited state, 'tip over' mechanisms etc. should be avoided.

3. Nearly the total variability of z_1 may be accounted for by the steric substituent parameters L, B_1 to B_4 (set 1-5). However, this equation suffers from a very high degree of multicollinearity and there is no possibility of placing any significance on the point estimates. This view is supported by consideration of set 1-6, where the magnitudes of the regression coefficients for the steric parameters are reduced by a constant factor of about 0.5. Since we are dealing here with total 'subsets' selected on the basis of physical considerations and prior information rather than arising from a subset selection procedure, the probability of chance correlations as described by Topliss and Edwards[19] is minimized, apart from the fact that a properly applied backward elimination (including simultaneous tests, see Mager[7,20] and the references cited therein) should not suffer from that type of chance correlation. Furthermore, the values calculated for the chance correlation measures given by Rencher and Pun[21] are clearly inferior to the R^2's obtained.

4. In sharp contrast to the lack of consistency for equations computed using ordinary least squares, the application of latent root regression analysis as developed by Webster et al.[22] yields more reliable and interpretable results (set 1-7). Although the fit is not affected, the weightings for the inductive effect and for x_7 are now approximately in the same range as in set 1-4. Furthermore, the significant regression coefficients for the steric parameters have the correct signs. Set 1-7 may be considered as an 'independent' proof for the significance of x_7 and, therefore, for the non-uniform behavior of ortho-substituents in acidic esterifications, which might have been an important reason for the controversy concerning the relative performances of the various substituent properties in the reactions of ortho substituted aromatic compounds. Obviously, the steric effect in the reactions investigated and hence represented by z_1 and E_s^o, respectively, is independent of the length of the substituents and is probably only a function of the 2 shape parameters B_1 and B_3.

5. Even in the case where OLS gives a subset containing only $\mathfrak{F}$ and $\mathfrak{R}$ (set 2-1, NO_2 omitted), if E_s^o is regressed vs the corresponding variables of the pool ($\mathfrak{F}$, $\mathfrak{R}$, r_v), the

application of PCRA shows that the influence of the steric term was suppressed due to multicollinearity (set 2-2).
6. As is to be expected, the E_s (corr.) values given by Mager[16] are strongly related to the van der Waals radius r_v without any significant inductive or mesomeric contributions (set 3-1 and 3-2). This may be regarded as an additional proof for the usefulness of Charton's ν values[23] in both aromatic and aliphatic reactions.

1 Reprint requests should be addressed to H.M. DDR-4090 Halle-Neustadt, Bl. 411/3 (German Democratic Republic).
2 Acknowledgment. We would like to thank Prof. J.T. Webster, Department of Statistics, Southern Methodist University, Dallas, for sending us s computer program of latent root regression analysis. We also thank the referees whose insightful comments were most helpful in improving the presentation.
3 A. Verloop, W. Hoogenstraaten and J. Tipker, in: Drug Design, vol. 7, p. 165. Ed. E.J. Ariens. Academic Press, New York 1976.
4 R.W. Taft, in: Steric Effects in Organic Chemistry, p. 556. Ed. M.S. Newman. Wiley, New York 1956.
5 M. Charton, J. Am. chem. Soc. *91*, 615 (1969).
6 H. Mager, P.P. Mager and A. Barth, Tetrahedron *35*, 1953 (1979).
7 H. Mager, Moderne Regressionsanalyse in Chemie und Biologie. Sauerländer, Aarau, in press (1982).
8 R.R. Hocking, Biometrics *32*, 1 (1976).
9 H. Mager, Pharmazie, in press (1982).
10 H. Mager, Dissertation, Akademie der Wissenschaften der DDR, 1977.
11 P.P. Mager, H. Mager and A. Barth, Scientia pharm. *48*, 2 (1980).
12 R.J. Hartman and A.G. Gassman, J. Am. chem. Soc. *62*, 1559 (1940).
13 R.J. Hartman and A.M. Borders, J. Am. chem. Soc. *59*, 2107 (1937).
14 R.J. Hartman, H.M. Hoogsteen and G.A. Moede, J. Am. chem. Soc. *66*, 1714 (1944).
15 D.F. Morrison, in: Multivariate Statistical Methods, p. 221. McGraw-Hill, New York 1967.
16 H. Mager, Tetrahedron *37*, 509 (1981).
17 C.G. Swain and E.C. Lupton, J. Am. chem. Soc. *90*, 4328 (1968).
18 H. Mager, Tetrahedron *37*, 523 (1981).
19 J.G. Topliss and R.P. Edwards, J. med. Chem. *22*, 1238 (1979).
20 H. Mager. Scientia pharm. *45*, 75 (1977).
21 A.C. Rencher and F.C. Pun, Technometrics *22*, 49 (1980).
22 J.T. Webster, R.F. Gunst and R.L. Mason, Technometrics *16*, 513 (1974).
23 M. Charton, J. Am. chem. Soc. *97*, 1552 (1975).

Space-filling molecular models of oxiranes (epoxides)

U. Séquin[1]

Institut für Organische Chemie der Universität, St. Johanns-Ring 19, CH-4056 Basel (Switzerland), 31 August 1981

Summary. An oxirane unit was constructed from aluminium spheres to fit the popular CPK system of space-filling molecular models.

Molecular models are widely recognized as valuable tools for the study of stereochemical relationships in organic chemistry and biochemistry[2]. The 2 most popular and most frequently used sorts are the framework models of the Dreiding type[3] and the space-filling models of the CPK type[4]. Whereas the framework models are most suitable for the determination of distances between nonbonded atoms and of torsion angles, the space-filling type lends itself particularly well to the study of intra- and intermolecular steric interactions and to the estimation of the preferred conformations of molecules.
We recently started to investigate the conformation of a substituted bioxirane related to the anti-tumor antibiotic hedamycin[5], and wanted to use space-filling models for our studies. Since no CPK units are commercially available for 3-membered rings, we designed an oxirane unit compatible with the CPK models.

The model unit was devised to represent an average oxirane. Thus, the dimensions used for its construction were a compromise (fig. 1) between those measured for ethylene oxide itself[6], those of the standard Dreiding oxirane unit[7] and of the covalent and Van der Waals radii used in the CPK system[8].

The oxygen atom was cut from an aluminium sphere of radius 16.87 mm (corresponding to a Van der Waals radius of 1.35 Å at the scale of 12.5 mm/Å used in the CPK models). The covalent radii were adjusted to 0.66 Å. No provision was made for a socket, as the oxygen atom was to be glued to the carbons. The carbon atoms were made from aluminium spheres of radius 15.62 mm (1.25 Å), the size used for sp^3-hybridized carbons in the CPK models[8]. Covalent radii of 0.77 Å (towards oxygen), 0.735 Å (towards carbon) and 0.77 Å (towards hydrogen) were used.

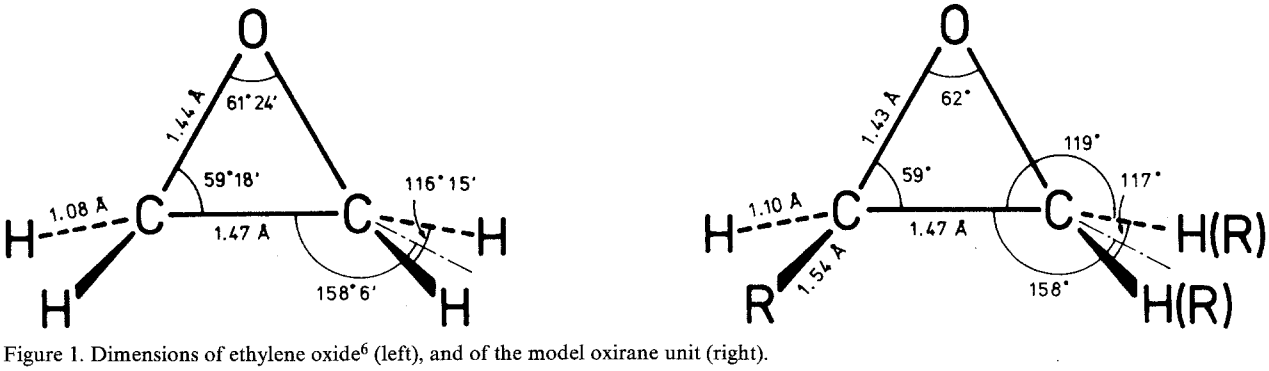

Figure 1. Dimensions of ethylene oxide[6] (left), and of the model oxirane unit (right).