PAPER CHROMATOGRAPHY AND STRUCTURAL RELATIONSHIPS OF ORGANIC ACIDS

J. R. HOWE
Low Temperature Research Station, Cambridge (Great Britain)
(Received July, 23rd 1959)

INTRODUCTION

An equation relating the partition coefficient of a solute between two liquid phases to its R_F value on paper was postulated by CONSDEN, GORDON AND MARTIN[1] in 1944:

$$\alpha = \frac{A_L}{A_S}\left(\frac{1}{R_F} - 1\right) \qquad (1)$$

where α = partition coefficient of solute between the stationary and moving phases, and A_L and A_S are the cross sectional areas of the moving and stationary phases respectively.

Following on from this in 1949, MARTIN[2] derived a relationship in which the free energy required to transport a molecule from one "ideal" liquid phase to another was proportional to the logarithm of the partition coefficient of the solute between the two phases. He then made two simplifying assumptions:

(1) that the total free energy of a molecule is composed of the sum of the free energies of the constituents composing the molecule; thereby implying that all isomers containing the same constituents would have the same partition coefficients;

(2) that the free energy required to transfer a given constituent (*e.g.* $-CH_2-$) of the molecule, from one solvent to another, is independent of the remainder of the molecule.

From this he was able to predict that the difference between the logarithms of the parition coefficients for adjacent members of a homologous series would be constant provided that the degree of ionisation of the members was the same. Therefore from equation (1) above, if A_L/A_S is constant over the whole paper, $\log(1/R_F - 1)$ should decrease by equal steps as the number of substituents is increased in a homologous series.

In order to test this proposition BATE-SMITH AND WESTALL[3] introduced the term "R_M value" such that

$$R_M = \log\left(\frac{1}{R_F} - 1\right) \qquad (2)$$

and found practically constant ΔR_M values for the introduction of hydroxyl and glucose groups into a number of aromatic compounds. The relationship was further investigated for organic acids by ISHERWOOD AND HANES[4] in a range of propanol–

ammonia solvents. They found ΔR_M constant for unit increase of n in the homologous dicarboxylic acid series $HOOC \cdot (CH_2)_n \cdot COOH$, but with the monocarboxylic acids $H \cdot (CH_2)_n \cdot COOH$ the ΔR_M values for unit addition of $-CH_2-$ decreased.

In the present paper, this type of investigation has been extended to a large number of organic acids in which several homologous series are included. The R_M values of the acids are compared, both in an alkaline solvent where the acids were ionised and also in an acidic solvent where the ionisation of the carboxyl groups of the acids was suppressed. By relating the R_M values from the two solvents it is possible to predict the number, type and configuration, of many of the substituent groups present in the acid molecule.

The literature on paper chromatography contains numerous examples of factors which affect the R_F value. These factors include temperature, purity of solvents, equilibration, volume of solvent relative to volume of chromatogram tank, ascending or descending solvent, distance travelled by solvent, type of paper, distance of point of application of spot from solvent level, pH of solution applied to the base-line, etc. Although the individual R_F values may vary with each of these factors, the overall chromatographic pattern of spots is likely to be similar no matter what conditions are chosen provided that a standardised procedure is adopted. Therefore the experimental conditions actually used in this paper are described in some detail in the next section and were always followed, very closely, for each set of determinations of R_F values.

MATERIALS AND METHOD

Chromatogram jar, dimensions: 29 × 19 × 45 cm.
Whatman No. 1 filter paper, dimensions: 23 × 45 cm.
Alkaline solvent: *n*-propanol–2 *N* aqueous ammonia (70:30).
Acid solvent: *n*-propanol–water saturated with sulphur dioxide (70:30).
Spray for acids: 15 ml Universal Indicator (B.D.H.) + 2–3 ml 0.1 *N* sodium hydroxide[5].
Spray for amino acids: 0.4 % ninhydrin + 0.2 % cobalt chloride + 5 % water in isopropyl alcohol[6].
Spray for neutral compounds: 0.2 *N* silver nitrate + 880 ammonia (6:1)[7].

Whatman No. 1 filter papers were washed prior to use with 2 *N* aqueous acetic acid, followed by distilled water and then 10 *N* aqueous ammonia solution as described by ISHERWOOD AND HANES[4]. All acids which were to be revealed later by a pH indicator were prepared as 0.1 or 0.2 *N* solutions and those to be revealed by ninhydrin were at 0.05 *M* strength. Within each homologous series all the individual members were of the same normality and all acids were dissolved in 2 *N* ammonium hydroxide solution. 4 μl spots of the solutions of the ammonium salts were applied on a starting line drawn 7.5 cm from one end of the paper and the papers were then hung in the jars for descending chromatography.

The reagents used in the preparation of solvents were the purest commercial grades available and no further purification was attempted. Solvents were always

prepared accurately as small variations, particularly in water content, had a significant effect on R_F value. This is important because the derived relationships depend on the fact that the final solvents were always of the same composition. The equilibration liquid, which consisted of 200 ml of either the acidic or the basic solvent, was splashed down the sides of the jar and the papers were gently flapped inside the jar for 2 h by the method described by HANES AND ISHERWOOD[8]. Solvent was then added to the trough and allowed to descend down the paper for about 10 h in which time the solvent front had advanced approximately 27 cm beyond the starting line. Chromatography was carried out at a constant temperature of 20° ± 0.5°.

When the papers were sprayed with Universal Indicator, the acid spots appeared immediately and their exact position was outlined because the contrast between the spot and background was usually greatest at this stage. These chromatograms can be retained as permanent records if stored away from acidic or basic vapours.

RESULTS AND DISCUSSION

Chromatography was carried out under the above conditions and the results are presented in Table I as R_F and R_M values for each solvent. In order to emphasise the structural relationship of these acids they are arranged in groups, in the Table, primarily according to whether they are mono-, di-, tri- or tetra-carboxylic acids. Each carboxyl grouping of acids is subdivided into the substituted groups, *e.g.* hydroxyl, amino etc. and finally the acids themselves are placed in homologous series, where possible, followed by any other acids of similar substitution in order of increasing carbon number. To assist in the comparison of related compounds a separate column is devoted to the number of carbon atoms in each acid. All the compounds in Table I have been numbered consecutively from 1 to 111 and these numbers are used in the figures and text.

Solvents

The solvents used were chosen as a result of the following considerations:

(a) In order to restrict the number of variables, systems of 1 organic component only were considered, *i.e.* 1 organic component + dilute aqueous acid (or alkali). Furthermore it was considered desirable to have the two solvents as similar in composition as possible, differing from one another only in the small amount of acid or base which had to be added.

(b) A large difference was required between the R_F values of each acid in the two solvents so that the effect of the –COOH group was accentuated. In a series of investigations with solvents containing alcohols it was found that this difference increases with the chain length of the alcohol but is limited by the low solubility of water in the higher alcohols. In solvents with low water contents the salts tend to streak and with the longer chain alcohols the di-, tri- and tetra-carboxylic acids in the alkaline solvent do not move from the base line and therefore cannot be distinguished from one another. This meant that propanol was the highest chain length

TABLE I

R_F AND R_M VALUES OF ACIDS AND SOME NEUTRAL COMPOUNDS

Code No.	Carbon No.	Compound	*n*-Propanol–2 N ammonia (70:30)		*n*-Propanol–water-sat. SO_2 (70:30)		(R_M) alkaline $-(R_M)$ acid
			R_F	R_M	R_F	R_M	
		MONOCARBOXYLIC ACIDS					
		Unsubstituted fatty acids					
1	1	Formic	0.47	+0.05	Acids too volatile to record		
2	2	Acetic	0.48	+0.03			
3	3	Propionic	0.56	−0.10			
4	4	*n*-Butyric	0.62	−0.21			
5	5	*n*-Valeric	0.67	−0.31			
6	6	Caproic	0.73	−0.43	0.90	−0.95	0.52
7	7	Heptanoic	0.76	−0.50	0.91	−1.00	0.50
8	8	Caprylic	0.79	−0.57	0.92	−1.06	0.49
9	9	Nonanoic	0.80	−0.60	0.92	−1.06	0.46
10	10	Decanoic	0.81	−0.63	0.93	−1.12	0.49
11	12	Lauric	0.81	−0.63	0.93	−1.12	0.49
12	14	Myristic	0.81	−0.63	0.93	−1.12	0.49
13	4	Isobutyric	0.61	−0.19	Acids too volatile to record		
14	5	Isovaleric	0.66	−0.29			
		Mono-hydroxy acids					
15	2	Glycollic	0.38	+0.21	0.69	−0.35	0.56
16	3	Lactic	0.47	+0.05	0.77	−0.52	0.57
17	4	2-Hydroxy-*n*-butyric	0.54	−0.07	0.82	−0.66	0.59
18	5	2-Hydroxy-*n*-valeric	0.62	−0.21	0.85	−0.75	0.54
19	6	2-Hydroxy-caproic	0.68	−0.33	0.88	−0.86	0.53
20	8	2-Hydroxy-caprylic	0.76	−0.50	0.90	−0.95	0.45
21	15	2-Hydroxy-pentadecanoic	0.78	−0.55	0.91	−1.01	0.46
22	4	3-Hydroxy-*n*-butyric	0.50	0.00	0.79	−0.57	0.57

(continued on p. 393)

TABLE I (*continued*)

Code No.	Carbon No.	Compound	*n*-Propanol–2 *N* ammonia (70:30)		*n*-Propanol–water–sat. SO_2 (70:30)		(R_M) alkaline $-(R_M)$ acid
			R_F	R_M	R_F	R_M	
	Di-hydroxy acids						
23	3	Glyceric	0.37	+ 0.23	0.62	— 0.21	0.44
	Mono-amino acids						
24	2	Glycine	0.31	+ 0.35	0.42	+ 0.14	0.21
25	3	α-Alanine	0.39	+ 0.19	0.56	— 0.10	0.29
26	4	α-Amino-*n*-butyric	0.47	+ 0.05	0.63	— 0.23	0.28
27	5	Norvaline	0.58	— 0.14	0.71	— 0.39	0.25
28	6	Norleucine	0.68	— 0.33	0.77	— 0.53	0.20
29	8	α-Amino-caprylic	0.75	— 0.48	0.82 tail	— 0.66	0.18
30	3	β-Alanine	0.34	+ 0.29	0.50	0.00	0.29
31	4	γ-Amino-butyric	0.36	+ 0.25	0.54	— 0.07	0.32
32	6	ε-Amino-caproic	0.46	+ 0.07	0.65	— 0.27	0.34
33	8	ω-Amino-caprylic	0.63	— 0.23	0.75	— 0.48	0.25
34	4	α-Amino-isobutyric	0.46	+ 0.07	0.65	— 0.27	0.34
35	5	Valine	0.55	— 0.09	0.68	— 0.33	0.24
36	6	Leucine	0.65	— 0.27	0.76	— 0.50	0.23
37	4	β-Amino-butyric	0.43	+ 0.12	0.60	— 0.18	0.30
38	5	γ-Amino-valeric	0.45	+ 0.09	0.62	— 0.21	0.30
39	6	Isoleucine	0.64	— 0.25	0.74	— 0.45	0.20
	Mono-amino, mono-amide acids						
40	4	Asparagine	0.28	+ 0.41	0.30	+ 0.37	0.04
41	5	Glutamine	0.29	+ 0.39	0.34	+ 0.29	0.10
	Mono-amino, mono-hydroxy acids						
42	3	Serine	0.35	+ 0.27	0.43	+ 0.12	0.15

(*continued on p. 394*)

TABLE I (*continued*)

Code No.	Carbon No.	Compound	*n*-Propanol–2 *N* ammonia (70:30)		*n*-Propanol–water-sat. SO_2 (70:30)		(R_M) alkaline $-(R_M)$ acid
			R_F	R_M	R_F	R_M	
		Di-amino acids					
43	5	Ornithine mono-HBr	0.27	+ 0.43	0.28 elong.	+ 0.41	0.02
44	6	Lysine mono-HCl	0.29	+ 0.39	0.30 elong.	+ 0.37	0.02
		Other amino acids					
45	5	Proline	0.45	+ 0.09	0.55	— 0.09	0.18
46	5	Hydroxyproline	0.37	+ 0.23	0.46	+ 0.07	0.16
47	6	Citrulline	0.32	+ 0.33	0.42	+ 0.14	0.19
48	9	Tyrosine	0.46	+ 0.07	0.63	— 0.23	0.30
49	11	Tryptophan	0.58	— 0.14	0.68 tail	— 0.33	0.19
50	9	Phenylalanine	0.64	— 0.25	0.71	— 0.39	0.14
51	5	Methionine	0.56	— 0.10	0.68	— 0.33	0.23
52	6	Arginine mono-HCl	0.25	+ 0.48	0.36 s. tail	+ 0.25	0.23
		Halogen-acids					
53	3	2-Bromo-propionic	0.63	— 0.23	0.87	— 0.83	0.60
54	4	2-Bromo-*n*-butyric	0.67	— 0.31	0.88	— 0.86	0.55
55	5	2-Bromo-*n*-valeric	0.73	— 0.43	0.90	— 0.95	0.52
56	6	2-Bromo-caproic	0.77	— 0.52	0.91	— 1.00	0.48
57	2	Trichloro-acetic	0.74	— 0.45	0.84	— 0.72	0.27
		Mono-phenyl acids					
58	7	Benzoic	0.64	— 0.25	0.91	— 1.00	0.75
59	8	Phenylacetic	0.66	— 0.29	0.91	— 1.00	0.71
60	9	3-Phenyl-propionic	0.71	— 0.39	0.92	— 1.06	0.67
61	11	5-Phenyl-*n*-valeric	0.76	— 0.50	0.94	— 1.19	0.69
62	9	Phenyl-lactic	0.68	— 0.33	0.91	— 1.00	0.67
63	9	Cinnamic	0.70	— 0.37	0.91	— 1.00	0.63

(*continued on p. 395*)

TABLE I (*continued*)

Code No.	Carbon No.	Compound	*n*-Propanol–2 *N* ammonia (70:30)		*n*-Propanol–water-sat. SO_2 (70:30)		(R_M) alkaline $-(R_M)$ acid
			R_F	R_M	R_F	R_M	
		DICARBOXYLIC ACIDS					
	Unsubstituted saturated acids						
64	2	Oxalic	0.17	+ 0.69	0.67	— 0.31	1.00
65	3	Malonic	0.20	+ 0.60	0.76	— 0.50	1.10
66	4	Succinic	0.23	+ 0.53	0.79	— 0.57	1.10
67	5	Glutaric	0.27	+ 0.43	0.82	— 0.66	1.09
68	6	Adipic	0.30	+ 0.37	0.84	— 0.72	1.09
69	7	Pimelic	0.35	+ 0.27	0.87	— 0.83	1.10
70	8	Suberic	0.40	+ 0.18	0.89	— 0.91	1.09
71	9	Azelaic	0.47	+ 0.05	0.90	— 0.95	1.00
72	10	Sebacic	0.53	— 0.05	0.91	— 1.00	0.95
73	5	Ethyl-malonic	0.29	+ 0.39	0.87	— 0.83	1.22
74	6	3-Methyl-glutaric	0.31	+ 0.35	0.86	— 0.79	1.14
75	7	2,2-Dimethyl-glutaric	0.35	+ 0.27	0.89	— 0.91	1.18
76	7	3,3-Dimethyl-glutaric	0.31	+ 0.35	0.88	— 0.86	1.21
77	8	3-Methyl-3-ethyl-glutaric	0.37	+ 0.23	0.89	— 0.91	1.14
	Mono-hydroxy acid						
78	4	Malic	0.20	+ 0.60	0.70	— 0.37	0.97
	Di-hydroxy acid						
79	4	Tartaric	0.18	+ 0.66	0.56	— 0.10	0.76
	Mono-amino acids						
80	4	Aspartic	0.19	+ 0.63	0.44	+ 0.10	0.53
81	5	Glutamic	0.20	+ 0.60	0.53	— 0.05	0.65
82	6	α-Amino-adipic	0.21	+ 0.57	0.57	— 0.12	0.69
83	7	α-Amino-pimelic	0.24	+ 0.50	0.63	— 0.23	0.73

(*continued on p. 396*)

TABLE I (*continued*)

Code No.	Carbon No.	Compound	*n*-Propanol–2 *N* ammonia (70:30)		*n*-Propanol–water-sat. SO_2 (70:30)		(R_M) alkaline $-(R_M)$ acid
			R_F	R_M	R_F	R_M	
		Aromatic acids					
84		Phenylsuccinic	0.39	+0.19	0.86	−0.79	0.98
85		Phthalic	0.33	+0.31	0.86	−0.79	1.10
		Unsubstituted double bonded acids					
86	4	Fumaric	0.27	+0.43	0.85	−0.75	1.18
87	4	Maleic	0.25	+0.48	0.78	−0.55	1.03
88	5	Mesaconic	0.27	+0.43	0.85	−0.75	1.18
89	5	Citraconic	0.28	+0.41	0.84	−0.72	1.13
90	5	Itaconic	0.29	+0.39	0.84	−0.72	1.11
		TRICARBOXYLIC ACIDS					
		Unsubstituted saturated acids					
91	6	Tricarballylic	0.13	+0.83	0.80	−0.60	1.43
		Hydroxy acids					
92	6	Citric	0.11	+0.91	0.69	−0.35	1.26
93	6	Isocitric	0.11	+0.91	0.69	−0.35	1.26
		Aromatic acid					
94	9	Trimesic	0.12	+0.86	0.87	−0.83	1.69
		Unsubstituted double bonded acids					
95	6	*cis*-Aconitic	0.14	+0.79	0.74	−0.45	1.24
96	6	*trans*-Aconitic	0.14	+0.79	0.86	−0.79	1.58

(*continued on p. 397*)

TABLE I (*continued*)

Code No.	*Carbon No.*	*Compound*	*n-Propanol–2 N ammonia (70: 30)*		*n-Propanol–water-sat. SO_2 (70: 30)*		*(R_M) alkaline $-(R_M)$ acid*
			R_F	R_M	R_F	R_M	
			TETRACARBOXYLIC ACIDS				
97	10	Prehnitic	0.04	+ 1.38	0.77	— 0.52	1.90
98	10	Pyromellitic	0.05	+ 1.28	0.81	— 0.63	1.91
			INORGANIC ACIDS				
99		Orthophosphoric	0.11	+ 0.91	0.62	— 0.21	1.12
100		Sulphuric	0.18	+ 0.66	0.31	+ 0.35	0.31
101		Hydrochloric	0.43	+ 0.12	0.51	— 0.02	0.14
102		Nitric	0.53	— 0.05	0.61	— 0.19	0.14
103		Perchloric	0.62	— 0.21	0.71	— 0.39	0.18
			NEUTRAL COMPOUNDS				
104	18	Raffinose	0.20	+ 0.60	0.20	+ 0.60	0.00
105	12	Sucrose	0.35	+ 0.27	0.36	+ 0.25	0.02
106	6	Glucose	0.40	+ 0.18	0.41	+ 0.16	0.02
107	6	Sorbitol	0.41	+ 0.16	0.42	+ 0.14	0.02
108	6	Dulcitol	0.43	+ 0.12	0.43	+ 0.12	0.00
109	6	Fructose	0.43	+ 0.12	0.45	+ 0.09	0.03
110	5	Ribose	0.50	0.00	0.50 s. tail	0.00	0.00
111	3	Glycerol	0.59	— 0.16	0.60	— 0.18	0.02

elong. = elongated spot; s. tail = short tail.

alcohol that could be used although its resolving power for closely similar acids was not as great as the higher alcohols.

The investigation also included a search for a solvent that would give a constant value of ΔR_M for the addition of $-CH_2-$ in a few of the homologous series of acids. This value was not always constant for all the series of acids examined in any one alcohol but here again the two propanol solvents chosen were one of the best combinations.

(c) Acids and bases added to the solvent must be volatile to allow for the subsequent operation of a pH indicator spray and they both must be strong enough to suppress the ionisation of the acidic or basic groups in the acid to be investigated.

Ammonia is a strong enough base to form salts with all the organic acids studied here and for many of the ampholytic amino acids but may not be strong enough to suppress the basic groups in amino acids such as arginine. However its volatility and purity made it superior to other bases. Paper chromatograms developed in alcohol–ammonia mixtures and then sprayed with Universal Indicator or silver nitrate always show a false front whose position is dependent on the nature of the alcohol and the amount of water (and/or ammonia) in the solvent. With the solvent described above, this front appeared at an R_F of approx. 0.40 as a line of demarcation behind which the paper background was faintly alkaline and beyond which it was slightly acid.

On the acidic side sulphurous acid was preferred to the more commonly used formic or acetic because of its greater strength (*cf.* VAS[9]). This is particularly necessary when dealing with a strong organic acid, such as oxalic acid, in the form of its ammonium salt as the method described depends largely on the fact that the acid was forced completely into the unionised state. The ammonium ion (from the ammonium salts of the acids applied to the starting line) interacts with sulphur dioxide in the solvent and results in an extra acid spot which has an R_F value of 0.46. This should not cause confusion with an unknown acid spot because it occurred in a part of the chromatogram away from all other "indicator-positive" organic acids studied, it was also slower in colouring up and less intense than the free acid spots.

A further advantage of sulphur dioxide is its volatility which allows the paper to be sprayed as soon as it is dry without any further treatment. This makes it possible to chromatograph semi-volatile acids which would otherwise be lost in the steaming, heating or leaving overnight required in the removal of formic or acetic acids from the paper.

The effect of the carboxyl group on the R_F value (Fig. 1)

The dominating effect of the carboxyl group on the R_F values of the compounds is illustrated in Fig. 1 which contains all the results of Table I in the form of a two-dimensional graph of R_F values for the two solvents. Arbitrary boundaries have been drawn to separate the acids according to their carboxyl number in such a way that each acid of Table I falls into its correct grouping.

Neutral compounds, *i.e.* those that contain no ionisable groups and which are chemically unaffected by the presence of ammonia or sulphur dioxide, have closely

similar R_F values in the two solvents. Ideally the R_F should be identical and the values irrespective of constitution should fall on a 45° slope line. In practice it was found that the values in the acid solvent were equal to or slightly higher than the alkaline solvent. There was no visible evidence on the chromatograms of inter-action between the sugars and sulphur dioxide, nor of hydrolysis of the polysaccharides.

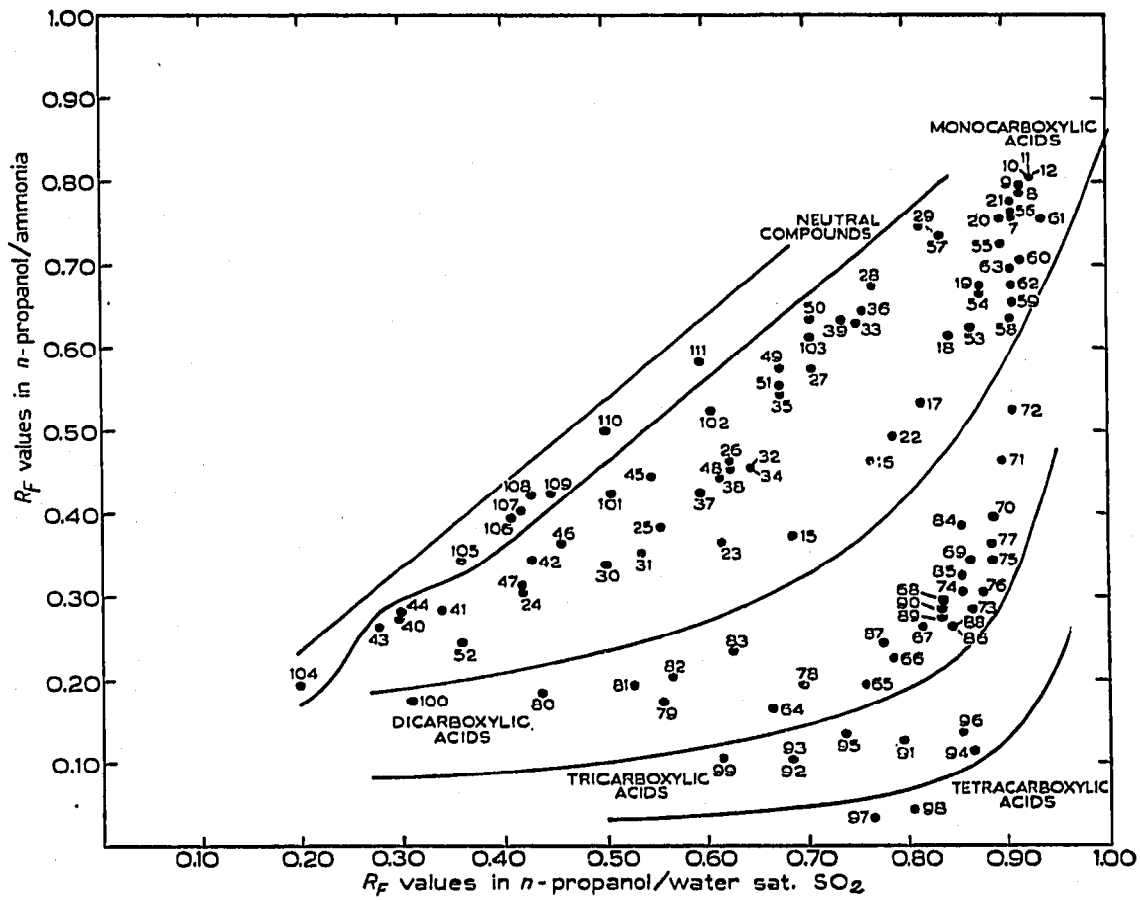

Fig. 1. R_F values in n-propanol–ammonia plotted against R_F values in n-propanol–water saturated with SO_2, for compounds listed in Table I.

The separation (in Fig. 1) between the slower running neutral compounds raffinose and sucrose (code Nos. 104 and 105) and the monocarboxylic amino aicds, asparagine, ornithine and lysine (code Nos. 40, 43 and 44) is achieved only by drawing an irregular boundary between them. This would be impossible if other neutral compounds, with R_F values intermediate between raffinose and sucrose, had been included. As the neutral compounds and the amino acids were identified by different sprays there was no difficulty in placing each in its proper group.

A few inorganic acids are also included in Fig. 1. The strong, monobasic, mineral acids—hydrochloric, nitric and perchloric (code Nos. 101, 102 and 103)—are in the monocarboxylic group whereas dibasic sulphuric acid (code No. 100) is just in the dicarboxylic group. This is an unexpected agreement between the behaviour of the

weak organic acids and the strong mineral acids because the sulphurous acid in the solvent is presumably not strong enough to affect the ionisation of the latter. On the other hand, tribasic orthophosphoric acid (code No. 99) is a much weaker acid and therefore there is some justification for its appearance along with the other tricarboxylic acids. A further investigation of inorganic acids was not attempted, so it is not possible to say, at this stage, whether there is any direct relationship between the basicity of inorganic acids and their R_F values.

Included in the monocarboxylic and dicarboxylic groups are a wide variety of different types of acids *e.g.* unsubstituted, halogen-, phenyl-, hydroxyl-, amino-, double bonded acids, etc. Each of these substituted groups is therefore of secondary importance, compared with the carboxyl group, in determining the R_F value of the acid in this class of solvent. This is well illustrated by the fact that aspartic and glutamic acids (code Nos. 80 and 81) are in the dicarboxylic group whereas the corresponding monoamides of these, asparagine and glutamine (code Nos. 40 and 41), occur with the monocarboxylic acids in fact nearly with the neutral compounds as described above.

The R_M values of homologues (Fig. 2)

If the different distances moved by the acids are primarily due to differences in their partition coefficients in these single phase solvents, then the predictions of MARTIN etc., outlined in the introduction, can be applied. R_M values from series of structurally related compounds, when plotted against unit additions of a substituted group in each series, should form sets of straight parallel lines such that the identity of unknown acid spots from a chromatogram could be obtained by extrapolation or interpolation. LEDERER[10] examined a number of homologous series in different solvents, from published lists of R_F values, and came to the conclusion that a linear relationship exists between plots of R_M values against the number of $-CH_2-$ groups in each series. The extent to which this ideal situation applies to the system described here was investigated by considering the addition of $-CH_2-$ to the molecule in nine similar homologous series.

The R_M values for these homologues are plotted against carbon number in Figs. 2A and 2B for the two solvents. These plots show that the introduction of –phenyl, –Br, and $-CH_2-$ into the molecule results in a decrease of R_M value (increase in R_F) whereas –OH, $-NH_2$, and –COOH increase the R_M value in both solvents. The values of the phenyl-substituted fatty acids (code Nos. 58–61, Table I) are plotted in these two figures against the number of carbon atoms in the fatty acid part of the molecule alone. Calculations on the R_F values of HASHMI AND CULLIS[5] show that the substitution of iodine into the fatty acid molecule also reduces the R_M value of the acid in a very similar propanol–ammonia solvent. The extent of the reduction is slightly greater than that resulting from the corresponding introduction of bromine into the molecule.

In both solvents the majority of series show a close approximation to linearity up to compounds containing 8 carbon atoms; above this value the straight-chain and hydroxy-monocarboxylic acids rapidly approach a limiting value as found by ISHERWOOD AND HANES[4] in similar alkaline solvents. Although straight lines or regular

curves have been drawn to represent the behaviour of the series, the individual members often show small departures from the line. By always comparing all members of any one series on the same paper, at the same time, these minor variations have been shown to be repeatable and in most cases are true expressions of the individuality of the separate acids. A good example of this is often shown by the first member of a homologous series.

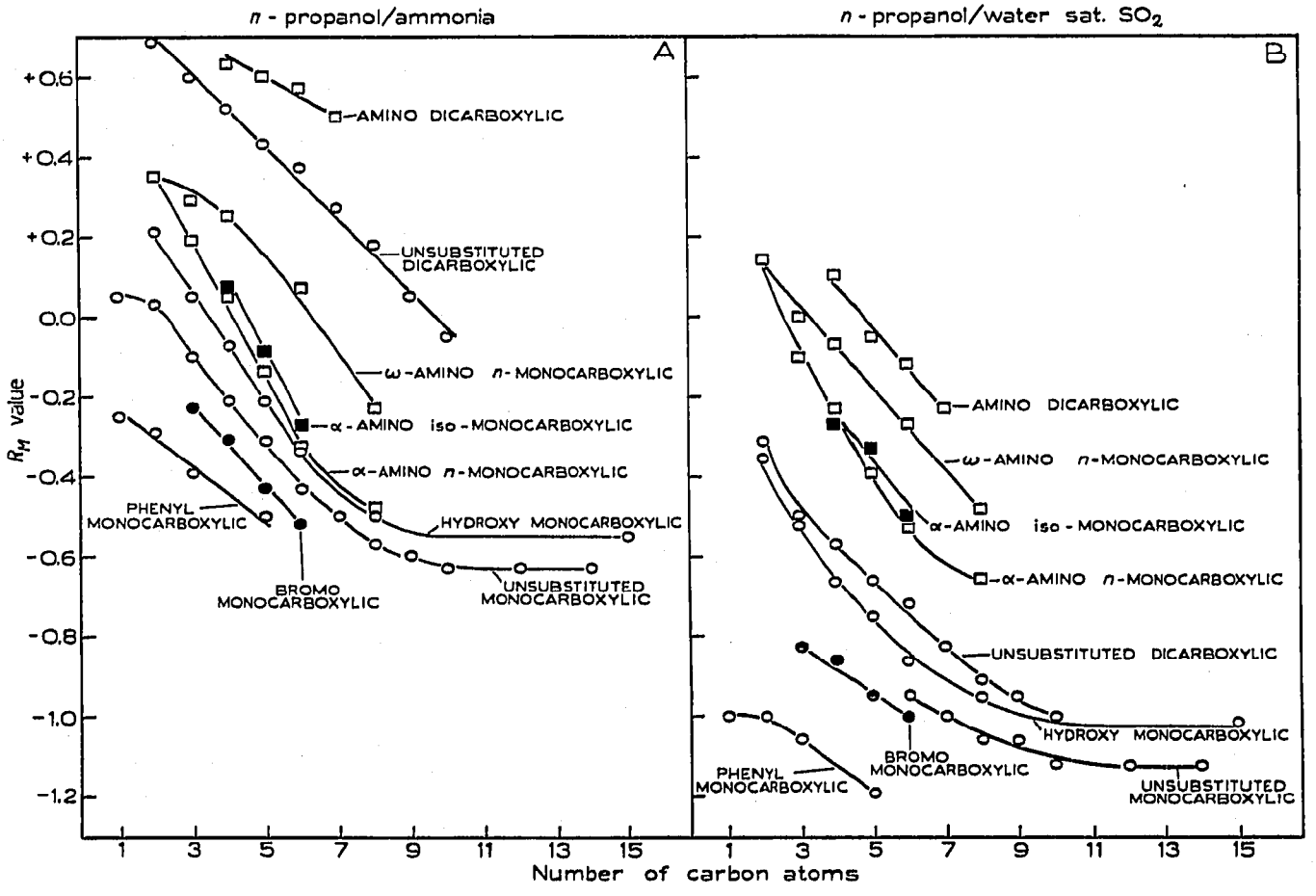

Fig. 2. Variation of R_M values with number of carbon atoms in homologous series. □ ■ amino acids, ○ ● non-amino acids.

The lines representing the series show a general similarity of slope but they are not strictly parallel to one another. In fact the divergencies from parallel are such that it is not possible to give a precise ΔR_M value for the addition of $-CH_2-$ in all the series; even over the linear parts of the graph the ΔR_M values of different lines vary, from 0.05 to 0.15 approximately, in both solvents.

The short homologous series of α-amino-iso-acids are very similar to the α-amino-n-acids a condition which corresponds to the lack of distinction of n- from iso- in the non-substituted acids butyric and valeric. On the other hand, as noted by SCHAUER AND BULIRSCH[11], the position of the amino group in the molecule has a relatively large influence on the final R_M value of the compound. An acid containing a terminal

(ω-) amino group has a higher R_M value (lower R_F) than the corresponding acid with an α-amino group. The two amino acids containing $-NH_2$ substituted part way down the chain (*i.e.* β-amino-*n*-butyric and γ-amino-*n*-valeric) adopt intermediate positions between these two extremes.

At least part of the explanation of these divergencies must lie in the assumptions made by MARTIN which were to a first approximation only. It must be assumed therefore that the partition coefficients of isomers are of the same order of magnitude only. Furthermore that the free energy required to transport $-CH_2-$ from one phase to another is partly dependent not only on the other groups present in the molecule but also on the manner in which they are arranged. This cannot of course be directly verified with the two single phase solvents used here. If these practical reservations are allowed for, the theoretical predictions prove to be of immense value in the determination of the structure of unknown acids. From the practical point of view the difference in R_F values shown by some isomers increases the use of the chromatographic method of identification.

No attempt has been made to follow the work of REICHL[13] or of SCHAUER AND BULIRSCH[11] who calculate average ΔR_M values for many substituent groups from the R_M values of relatively few compounds. In the present results the ΔR_M values for unit additions of $-CH_2-$ were not constant when a large number of acids were examined, therefore it was considered premature to attempt such calculations on the substituent groups.

The alkaline solvent—acid solvent R_M difference (Fig. 3)

REICHL[13] expressed his results as $R_M = \log [R_F/(1-R_F)]$ because this function increases as the R_F value increases. This value is the negative logarithm of the value defined by equation (2) above and therefore to avoid any further confusion in the literature, REICHL's values will be referred to as $(-R_M)$ values in this paper. He examined 36 acids in 2 pairs of solvents and plotted his results as a two-dimensional graph of the $(-R_M)$ values in one solvent against those in the other. He found that on drawing parallel lines across his graph he could separate the acids into groups according to the number of -COOH groups they contain. In one set of solvents the amino acids occurred in the wrong group and in both sets of solvents phthalic and maleic acids were anomalous, which he concluded was due to interaction between the adjacent carboxyl groups—an "*ortho* effect" (*cf.* BAKER[14], BATE-SMITH AND WESTALL[3]). In the present paper, no such anomalous behaviour was noticeable with these or other acids containing adjacent -COOH groups.

REICHL also noted that the difference in the $(-R_M)$ values for each acid, between an acidic and a neutral solvent, was a good index of the number of carboxyl groups in that acid. A similar R_M difference has been used by MACEK AND VEJDĚLEK[15] to determine the number of α-glycol groups in the veratrum alkaloids. In their method the R_M difference is between the values obtained for each alkaloid, chromatographed first on an ordinary paper and then, under similar conditions, on a paper impregnated with boric acid.

Values of the expression (R_M) alkaline solvent — (R_M) acid solvent are tabulated in Table I. With the larger number of acids considered here the range of values included in the monocarboxylic group (0.02 to 0.75) overlap somewhat the range covered by the dicarboxylic group (0.53 to 1.22). However, to illustrate the value of this expression, the homologous series and other compounds are graphed in Fig. 3 with the R_M

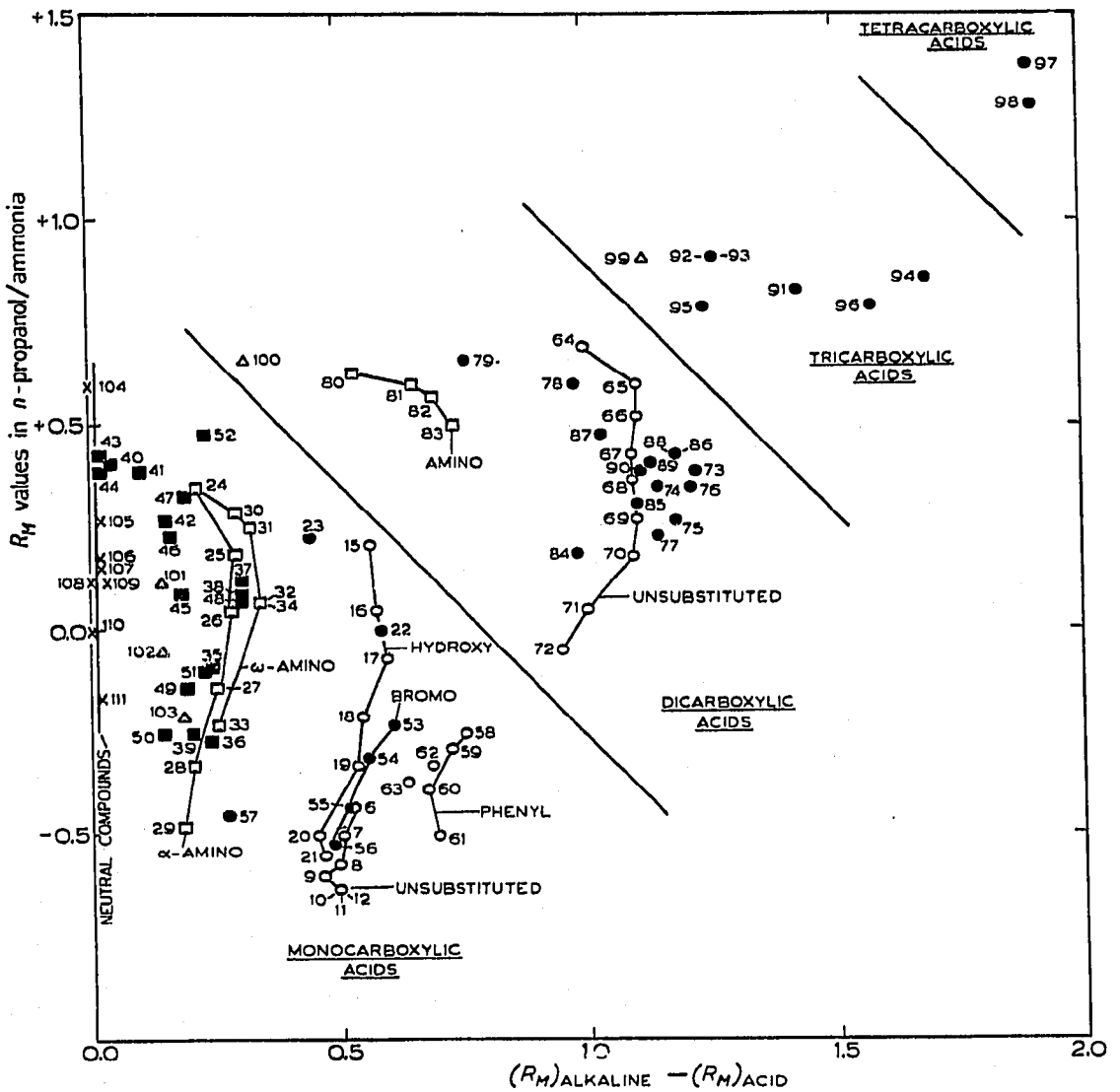

Fig. 3. R_M values in *n*-propanol–ammonia of compounds listed in Table I, plotted against the difference in R_M values between the two solvents. ×, neutral compounds; △, inorganic acids; □ ■, amino organic acids; ○ ●, non-amino organic acids.

difference between the two solvents as abscissa against the R_M in propanol–ammonia as ordinate. Plotted in this way the overlapping of values from one carboxyl group to another is avoided and it is possible to draw an equi-spaced series of parallel lines to separate the acids into their correct groups.

Neutral compounds occur on the extreme left of the graph and are represented by the vertical line at 0.01 R_M difference between the two solvents. The same overlapping

with a few of the amino acids, as previously explained for Fig. 1, is also noticeable here. The homologous series are shown connected point to point by full lines which are nearly vertical and the points composing them are in order of increasing carbon number downwards. A vertical line for these series would mean that the addition of $-CH_2-$ to an acid produced the same change in partition coefficient irrespective of whether the acid was ionised or unionised.

The homologous series of monocarboxylic α-amino-iso-acids (code Nos. 34–36) are not connected by lines in the figure because they would obscure part of the series of α-amino-n-acids (code Nos. 24–29) and ω-amino-n-acids (code Nos. 30–33) which are shown connected. These three homologous series together with β-aminobutyric, γ-aminovaleric and isoleucine (code Nos. 37–39) all occupy a very small sub-area of the monocarboxylic group. Therefore any unknown acid spot which occurs in this sub-area is likely to be an isomer of the amino-substituted fatty acids. Tyrosine (code No. 48), an aromatic amino acid, also falls inside this area and is therefore an exception to the statement.

Similar sub-groups and areas can be marked off for other derivatives and configurations. For example, all the amino acids of Table I occupy the left hand area of their respective carboxylic groupings in Fig. 3, whereas all the non-amino organic acids are quite separate to the right. The intervening space between these two sub-groups will probably be occupied by the mono-, di- or tri- etc. hydroxy-substituted organic acids. A start in this direction can be seen in the sequence succinic–malic–tartaric acid (code Nos. 66, 78, 79) which have 0, 1 and 2 hydroxyl groups respectively with the same number of carbon atoms in each. The addition of each hydroxyl group brings the acid position nearer to the amino acid sub-group. It is anticipated therefore that certain heavily hydroxylated acids will occur in with the amino acids.

The three homologous series of monocarboxylic acids *i.e.* unsubstituted, bromo- and hydroxy-substituted respectively, also occur very close together. But the homologous phenyl-substituted acids (code Nos. 58–61) together with phenyl-lactic and cinnamic acids (code Nos. 62, 63) are to the right of these and suggest the presence of a separate aromatic sub-group. Finally, it is worth mentioning that the few inorganic acids studied occupy positions to the left of each carboxylic group.

Recently it has been found that certain substituted aromatic acids such as syringic and protocatechuic acids do not fit into their correct carboxyl grouping. Detailed discussion of these and other inter-relationships contained in this paper will be postponed until more information concerning related compounds has been obtained.

CONCLUSION

As there is such a close and definite relationship between R_M value and molecular structure of an acid it is clear that there will be many more acids than those already studied which will fit into the carboxyl group areas drawn here.

The scheme outlined should prove useful not only in conjunction with selective spray reagents but also in the examination of unknown compounds detected by non-

specific reagents. For example, with radioactive isotopes revealed by autoradiographs it may be possible to show by this method

(a) whether the compound is neutral or not;
(b) the number of carboxyl groups in the molecule;
(c) the presence of $-NH_2$ groups;
(d) an indication of other substituted groups such as phenyl or hydroxyl.

ACKNOWLEDGEMENTS

I would like to thank Dr. E. C. BATE-SMITH, Dr. M. INGRAM, Dr. H. G. WAGER and Mr. F. A. E. PORTER for the help and guidance they have given throughout this investigation.

SUMMARY

An acidic and a basic solvent containing the same proportions of propanol and water have been used for the study of about 100 organic acids by paper chromatography. The acid and the base were solutions of the gases sulphur dioxide or ammonia in water and therefore were readily removed after chromatography. Sulphurous acid has the advantage of being a stronger acid than those normally used, *e.g.* formic and acetic acid. The R_M values of the organic acids in the two solvents can be arranged into clear-cut groups dependent on the number of carboxyl and other substituted groups present, in such a way that the number of carboxyl groups in an unknown acid can be confidently predicted. An indication can also be obtained of other groups such as alkyl, aryl, amino, hydroxy, bromo etc. which may be present, and the manner in which they are arranged. The results include nine homologous series whose members differ from one another only in the additon of $-CH_2-$ to the molecule.

REFERENCES

1 R. CONSDEN, A. H. GORDON AND A. J. P. MARTIN, *Biochem. J.*, 38 (1944) 224.
2 A. J. P. MARTIN, *Biochem. Soc. Symposia*, Cambridge University Press, London, No. 3, 1949, p. 4.
3 E. C. BATE-SMITH AND R. G. WESTALL, *Biochim. Biophys. Acta*, 4 (1950) 427.
4 F. A. ISHERWOOD AND C. S. HANES, *Biochem. J.*, 55 (1953) 824.
5 M. H. HASHMI AND C. F. CULLIS, *Anal. Chim. Acta*, 14 (1956) 336.
6 L. F. WIGGINS AND J. HOWARTH WILLIAMS, *Nature*, 170 (1952) 279.
7 S. M. PARTRIDGE, *Nature*, 158 (1946) 270.
8 C. S. HANES AND F. A. ISHERWOOD, *Nature*, 164 (1949) 1107.
9 K. VAS, *Acta Chim. Acad. Sci. Hung.*, 1 (1951) 335.
10 M. LEDERER, *Proc. 2nd Intern. Congr. of Surface Activity, London*, 1957, p. 506.
11 H. K. SCHAUER AND R. BULIRSCH, *Z. Naturforsch.*, 10b (1955) 683.
12 E. R. REICHL, *Monatsh. Chem.*, 86 (1955) 69.
13 E. R. REICHL, *Mikrochim. Acta*, (1956) 958.
14 W. BAKER, *Nature*, 137 (1936) 236.
15 K. MACEK AND Z. J. VEJDĚLEK, *Nature*, 176 (1955) 1173.