

CHROM. 7684

COMPUTER APPLICATIONS IN GAS-LIQUID CHROMATOGRAPHY

I. PREDICTION OF COMPONENT RESOLUTION FROM LIMITED INPUT DATA

C. A. STREULI and W. H. MULLER

Lederle Laboratories Division, American Cyanamid Company, Pearl River, N.Y. 10965 (U.S.A.)

and

M. ORLOFF*

Chemical Research Division, American Cyanamid Company, Stamford, Conn. 06904 (U.S.A.)

(First received January 18th, 1974; revised manuscript received June 11th, 1974)

SUMMARY

It is shown that the temperature at which resolution of mixtures will occur in gas-liquid chromatography can be accurately predicted by a computer using input data obtained at only three temperatures.

INTRODUCTION

In the past decade, the principle use of computers in gas-liquid chromatography (GLC) has been to obtain quantitative data on peak areas and to calculate the percentage compositions of mixtures from on-line data gathering. This limited application, however, has neglected an important area, and one that is intrinsic to physical science: the prediction of further results from limited experimental data. There has been some work on structure-retention correlation in the past 5 years¹⁻³ and recently on column selection⁴, but the principle use of computers in GLC seems to have been aimed in a different direction. However, accurate prediction of results can save a great deal of experimental time and is, therefore, of considerable value.

A number of mathematical equations have been developed in GLC that relate retention volume and peak spread to known physical parameters, and a computer is uniquely qualified to do the tedious calculations necessary to predict retention volume and peak spread at any temperature if given limited input data.

Consider first the necessary mathematical expressions. They are⁵:

$$\log V_{\theta} = \frac{-\Delta H_s}{2.3R} \cdot \frac{1}{T} + c_1 \quad (1)$$

* Present address: Organic Chemicals Division, American Cyanamid Company, Bound Brook, N.J. 08805, U.S.A.

$$V_g = F_c \cdot \frac{1}{w_l} \cdot \frac{273}{T} \cdot j \cdot (t_g - t_a) \quad (2)$$

$$j = \frac{2 p_a}{p_l + p_a} \quad (2a)$$

$$F_c = F \left(\frac{T}{T_a} \right) \cdot \left(\frac{p_a - p_w}{760} \right) \quad (2b)$$

$$N^{\pm} = \left[\left(\frac{t_g}{\frac{1}{2} \text{ wid.}} \right)^2 5.548 \right]^{\pm} \quad (3)$$

$$\text{Peak spread} = \frac{4 t_g}{N^{\pm}} \quad (4)$$

where V_g is the specific retention volume, ΔH_s the heat of solution of solute in substrate, R the gas constant, T the absolute temperature of the column ($^{\circ}\text{K}$), F the flow-rate, F_c the corrected flow-rate, w_l the weight of substrate on the support, t_g and t_a the retention time of the solute and the dead time of the column, respectively, p_a and p_l the outlet and inlet pressure of the column, respectively, p_w the vapor pressure of water at ambient temperature (T_a), N the number of theoretical plates and $\frac{1}{2}$ wid. the width of the peak at half-height. All of these terms are well known to gas chromatographers and are defined in ref. 4.

Eqns. 3 and 4 utilize the idea of theoretical plates to determine the number of plates in a column and then determine how fast a peak will spread at the base depending on its retention time. Eqn. 2a approximates a corrected flow-rate from a measured flow-rate, and eqn. 2b corrects for pressure drops in the column. The last two equations can be worked into eqn. 2 to give:

$$V_g = F \cdot \left(\frac{273}{T_a} \right) \cdot \left(\frac{1}{w_l} \right) \left(\frac{p_a - p_w}{760} \right) \left(\frac{p_a}{p_l + p_a} \right) (t_g - t_a) \quad (2c)$$

Thus V_g can readily be calculated from easily measured parameters, as can be N from eqn. 3. The most important equation is eqn. 1. Eqn. 1 is the equation of a straight line, $y = a \cdot x + b$, with $1/T$ as the x parameter, $\log V_g$ the y parameter, $-\Delta H/2.3 R$ the slope and c_1 the intercept. Two values of x and y are sufficient to determine the slope and intercept, but a better solution for a straight line from experimental data is a least-squares solution using more values of x and y . A least-squares solution requires at least three points. Once the slope and intercept have been determined, $\log V_g$ (and V_g) are predictable at any value of T and ultimately through a reworking of eqn. 2c, t_g can be obtained when the flow-rate, pressure conditions and w_l are known. This will, of course, apply only to a single liquid phase, as the slope and intercept change with the substrate. It also assumes that ΔH is independent of temperature.

The problems then considered were: how many experimental data points are necessary to obtain reliable values of slope and intercept in eqn. 1; can an average value of N^{\pm} (\bar{N}^{\pm}) from these data maintain a small deviation; and can the use of these limited data then be used to predict optimum conditions for a separation? A computer can easily carry out all the necessary mathematical procedures in a very short time and make the desired predictions.

EXPERIMENTAL

The gas chromatograph used in this work was an F & M Model 720 dual-column instrument equipped with thermal conductivity detector, column, injection block and detector temperatures which can be read to ± 1 °C. Column temperatures were calibrated from 70 to 230 °C. Flow-rates were measured in a soap-bubble flow meter. The carrier gas (helium) was held at a constant inlet pressure of 30 lb./in.². Ambient pressure and temperature were measured in the usual way and injections were 1 μ l for liquids or 3 μ l for solid samples dissolved in either dimethylformamide or chloroform; t_r , t_a and peak half-width were read from a strip-chart recorder, estimates being made to 0.05 min. These measurements constitute the input data for the computer. All columns used were 6 ft. long, 4 mm I.D. stainless steel packed with the appropriate liquid, adsorbed on Chromosorb W, 60–80 mesh, acid-washed. The liquids employed were a hydrocarbon* system, a silicon-containing carborane polymer** and a polyethylene glycol***, which represent non-polar, semi-polar and polar substrates, respectively. In this paper (and in Part II⁶), these three systems will be referred to as substrates 1, 2, and 3, respectively. The weights of liquid loads were determined to the nearest milligram and had an error of less than 0.1 %. This is about the order of magnitude for temperature and pressure measurements also, as temperatures are converted to T (°K) and pressures are measured in torr.

RESULTS

The first problem considered was variation of the values in the slope and intercept with the number of data points given to the computer for a least-squares solution of eqn. 1 and the value of \bar{N}^{\dagger} computed from eqn. 3. Results for isoquinoline on substrate 3 (weight 1.880 g) obtained from data taken at 10 °C intervals from 150 to 240 °C are listed in Table I. Experimental values for $\log V_r$ ranged from 2.9 to 1.8 over this interval considered.

TABLE I

VARIATION OF SLOPE, INTERCEPT, AND \bar{N}^{\dagger} WITH NUMBER OF DATA POINTS FOR LEAST-SQUARES SOLUTION

No. of points	Slope	Intercept	s^*	\bar{N}^{\dagger}
10	-2658.1	-3.4058	± 0.0072	33.8
5	-2673.7	-3.4435	± 0.0069	34.9
3	-2680.9	-3.4579	± 0.0115	34.5

* s is the standard deviation in $\log V_r$ between the values calculated from Table I and eqn. 1, and the measured values.

The slope value calculated using three points differs from that using ten points by 0.85%, the intercept by 1.5% and the \bar{N}^{\dagger} value by 2%.

A second test was run using *p*-xylene on the same substrate and comparing

* Available under the trade mark Apiezon M.

** Available under the trade mark Dexsil 300.

*** Available under the trade mark Carbowax 20M.

results using eight and three sets of data points. Here \bar{N}^{\pm} values were calculated and s values for the eight-point solution were 0.0196 and for the three-point solution 0.0397. Tentatively, it appeared that three points were sufficient to make credible predictions of t_g and Δ (peak spread).

Further work was carried out on other compounds and substrates and all the results are listed in Table II.

The results in Tables I and II indicate three sets of data (T , t_g and half peak width) and are sufficient for evaluation of slope, intercept and \bar{N}^{\pm} and give relatively small errors in back-calculating t_g and base peak width, the sought-for values.

The second group of problems involved replications of data from day to day, at different flow-rates on the same column, or from column to column of the same material.

TABLE II
RESULTS FOR VARIOUS SUBSTRATES AND COMPOUNDS

Column	Compound	No. of data points	Slope	Intercept	s^*	\bar{N}^{\pm}	Error	
							t_g	Δ
Substrate 3	<i>p</i> -Xylene	8	-1880.9	-3.0604	0.0196	32.8	± 0.21	± 0.06
		3	-1839.5	-2.9427	0.0397	32.8	± 0.21	± 0.09
Substrate 3	Pyrrole	8	-2481.5	-3.8354	0.0162	31.6	± 0.31	± 0.06
		3	-2442.0	-3.7529	0.0076	31.3	± 0.05	± 0.05
Substrate 1	Tetrahydro-naphthalene	7	-2322.9	-2.9798	0.0043	41.8	± 0.08	± 0.08
		3	-2325.0	-2.9831	0.0099	41.7	± 0.08	± 0.05
Substrate 2	Acetophenone	8	-2232.8	-3.1861	0.0210	33.0	± 0.17	± 0.07
		3	-2205.7	-3.1257	0.0057	33.2	± 0.03	± 0.10

* s is as defined in Table I, so the grand average for the standard deviation in $\log V_g$ is of the order of ± 0.015 .

Variation of flow-rate was done with toluene on 15% substrate 3. The original run had a corrected flow-rate of 96.45 ml/min and three data points generated slope, intercept and \bar{N}^{\pm} values of -1687.8, -2.7903 and 32.5, respectively. A second run was made 10 days later at a corrected flow-rate of 58.6 ml/min. The measured values at 80 °C for t_g and Δ were 5.15 and 0.68 min, respectively. The computer predicted values of 5.20 and 0.64 min, respectively. Other comparable runs made with naphthalene and benzyl alcohol on substrate 2 and tetrahydronaphthalene and 2-methylnaphthalene on substrate 1 gave average errors in t_g of ± 0.28 min and peak spread of ± 0.09 min. Time intervals between any two sets of runs were 1 day to 3 months.

The two column test was carried out on substrate 3 with weights (w_i) of 1.880 and 2.130 g. Constants calculated for *p*-xylene on the first column were used for predictions on the second. Errors in t_g and peak spread were ± 0.09 min and ± 0.04 min, respectively (average of three). In each case tested, the three data point original solution appears to be valid.

The final problem was prediction of resolution of compounds using data obtained from running components of the mixture separately. Three mixtures were studied on all three substrates: quinoline and isoquinoline; aniline, N-methylaniline and N,N-dimethylaniline; and the three isomeric xylenes.

TABLE III
SEPARATION OF QUINOLINE AND ISOQUINOLINE ON SUBSTRATE 3

Partial computer print-out						Experimental data					
$T (^{\circ}\text{K})$	$(t_R)_1$	$(t_R + \frac{1}{2}\Delta)_1$	$(t_R)_2$	$(t_R - \frac{1}{2}\Delta)_2$	Overlap	$T (^{\circ}\text{K})$	$(t_R)_1$	Error	$(t_R)_2$	Error	Overlap*
423	21.45	22.62	23.99	22.62	0.00	423	21.27	0.18	24.70	-0.71	r
433	15.60	16.45	17.31	16.32	-0.13	433	15.25	0.35	17.62	-0.31	dnrc
443	11.54	12.16	12.71	11.98	-0.18	453	8.52	0.15	9.80	-0.31	dnrc
453	8.67	9.14	9.49	8.95	-0.19	483	3.79	0.26	4.46	-0.19	dnrc
483	4.05	4.27	4.37	4.12	-0.15						

* r = resolves; dnrc = does not resolve completely.

The computer was given the appropriate slope, intercept and \bar{N}^{\dagger} values for each compound together with flow-rate, pressure data and T_a and was asked to predict at 5 °C intervals values for $(t_R)_1$, $(t_R + \frac{1}{2}\Delta)_1$, $(t_R)_2$, $(t_R - \frac{1}{2}\Delta)_2$ and overlap. A negative value for overlap would indicate that the mixture does not resolve completely. Limits of t_R were set between 4 and 25 min. A typical comparison of prediction and experimental results is shown in Table III.

It can be seen that the computer correctly predicted that the mixture would just resolve at 423 °K (150 °C) but not at any higher temperatures. A positive number in the last column would give greater assurance of resolution. On substrate 2, the computer predicted resolution from 393 up to 423 °K but this was not born out experimentally because of tailing of the first peak. On substrate 1 it correctly predicted no resolution between 413 and 453 °K.

With the isomeric xylenes, the computer correctly predicted reality at all temperatures tested, *i.e.*, the *ortho*-isomer could be resolved from the *meta*- and *para*-isomers, but the latter two could not be resolved within the time limits set. For the mixture of anilines, the computer predicted complete resolution at 373 °K for the three components, which is experimentally correct. The overlap values were 7.97 and 3.04. It also predicted positive values up to 403 °K, but here resolution failed because of tailing of N-methylaniline. It made correct predictions for substrate 1 and substrate 3, however. The principle problem appears to be a non-Gaussian peak when prediction fails, as t_R values are always within limits of error.

CONCLUSIONS

(1) Three data sets are sufficient to obtain a least-squares solution of eqn. 1 and to give an average \bar{N}^{\dagger} value. These values are valid for changes in flow-rate, ambient pressures and temperatures, and liquid load over a time period.

(2) Utilizing these values, the computer can predict with a high degree of accuracy the correct temperature for resolution of several components. The overlap value should probably be at least +1.0 and if tailing is severe this value should be higher.

(3) The method can be applied to unknown compounds or simple mixtures of unknown compounds, as it derives from experimental input data of limited amount.

REFERENCES

- 1 A. W. Landon and S. Sandler, *Anal. Chem.*, 45 (1973) 921.
- 2 L. Rohrschneider, *J. Chromatogr.*, 22 (1966) 6.
- 3 P. H. Weiner and D. G. Howery, *Anal. Chem.*, 44 (1972) 1189.
- 4 K. Lekova and M. Gerasimoff, *Chromatographia*, 7 (1974) 69.
- 5 S. Dal Nogare and R. S. Juvet, Jr., *Gas-Liquid Chromatography*, Interscience, New York, 1962, pp. 70-81.
- 6 C. A. Streuli and M. Orloff, *J. Chromatogr.*, 101 (1974) 23.