CHROM. 4538

Use of normalization techniques for establishing elution positions of chromatographic peaks

Chromatographic separation systems are now used in many analytical instruments. Assignment of identities to the resulting chromatographic peaks is a prerequisite to quantification of the compounds represented by these peaks and is usually made on the basis of elution position (position of maximum peak height). The elution positions must be precise to allow peak identification, especially for high-resolution systems in which the peaks may be eluted very closely together[1–4]. This is particularly true if computer techniques are to be used to evaluate the chromatogram[5].

In instances where chromatograms are evaluated manually, human judgment can sometimes compensate for shifts in elution position caused by variations of eluent flow rate, temperature changes, pH changes, etc., and allow exact peak identification to be established. In computer evaluation, corrections must be applied to the peak elution positions to compensate for these shifts. The use of normalization techniques to make such corrections is the subject of this paper. The data reported here were obtained from high-resolution carbohydrate automated analyses of body fluids[6].

Carbohydrate chromatograms

The automated carbohydrate analyzer utilizes strongly basic anion-exchange resin to chromatograph soluble sugars in borated physiological fluids, and sulfuric acid–phenol colorimetry to detect the eluted carbohydrates. The chromatograms subsequently generated are continuous plots of the absorbance of the eluate–acid–phenol reaction mixture at 480 and/or 490 mμ *versus* time (see Fig. 1)[3,6]. A large number of carbohydrate analyses of the urine from normal subjects have been made to establish "base-line" chromatograms or chromatographic patterns.

The chromatographic peaks were identified by their elution time or position relative to well-defined reference peaks, as shown in Fig. 1. The peaks were identified by name or by an alphanumeric label if unknown. For example, the large peak sucrose is followed by raffinose and several smaller unknown peaks that are designated Su+1, Su+2, etc.

Normalization of peak elution positions

Assignment of a name or a label to a chromatographic peak may require correction of the elution positions due to elution position shifts caused by variation of operating parameters. The correction method we have used is normalization of the elution position of a peak by comparison with one or two other well-defined large peaks. The simplest correction is to normalize to only one peak. For example, in the case of the carbohydrate chromatograms, glucose which is eluted at about 12.6 h (Fig. 1) was used for normalization. In this method it is assumed that all peak positions vary in a way proportional to variation in the reference peak. For example, if the reference peak elutes at 0.9 of its average position, then all peaks will elute at 0.9 of their average positions. Normalization to one peak requires that each sugar elution position (time) for a specific analysis be multiplied by a normalization

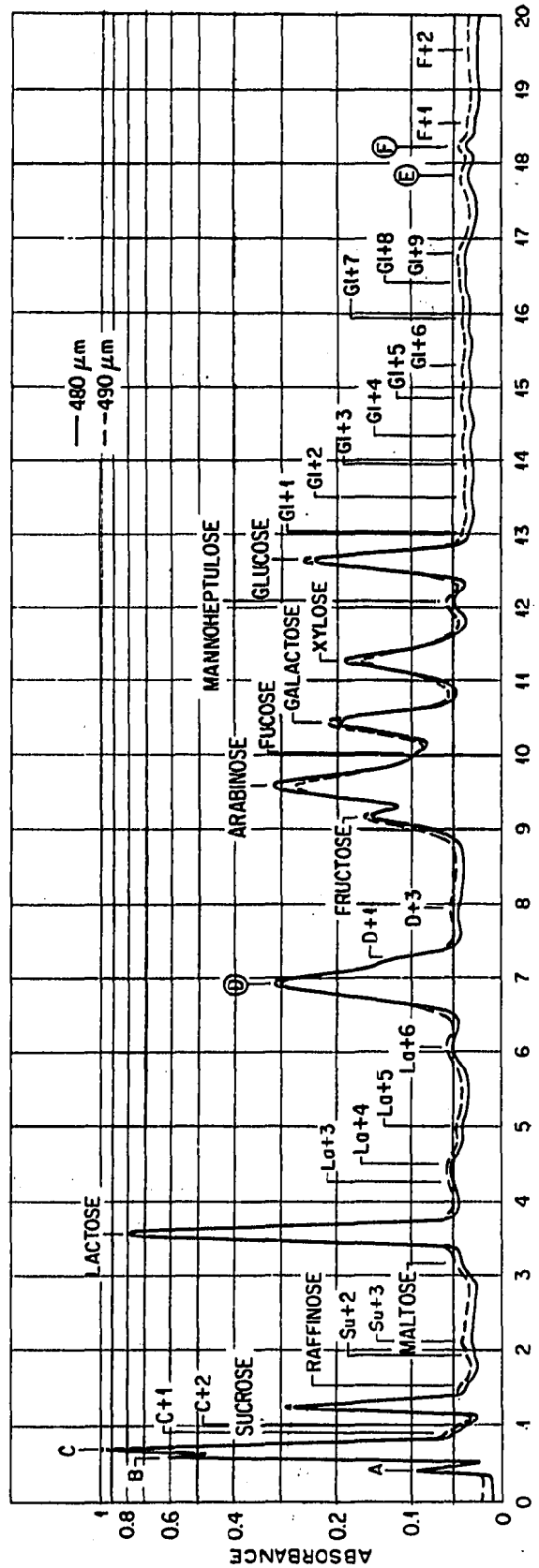

Fig. 1. Typical carbohydrate chromatogram of a urine composite showing peak identification. Run conditions are given in refs. 3 and 6.

factor, F, defined as follows:

$$F = \bar{t}_G/t_G' \quad (1)$$

in which $\bar{t}_G$ is the average elution time of the normalization peak, *e.g.*, glucose, in a series of previous chromatograms, and t_G is the elution time of the same peak determined for the specific chromatogram being analyzed. Thus, the normalized elution time for any sugar, t_X^*, is expressed by

$$t_X^* = F \cdot t_X \quad (2)$$

A second procedure, which involves normalization to two peaks, was also evaluated. This procedure is based on the assumption that each peak varies by some constant shift in elution time plus an amount proportional to the shifts of reference peaks. To compensate for these shifts it is necessary to normalize to two well-defined

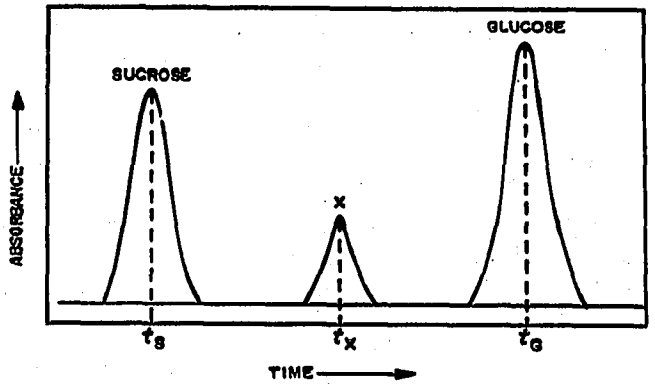

Fig. 2. Choise of reference peaks in the carbohydrate chromatogram for normalizing peak positions.

peaks, such as sucrose (elution time of 1.22 h) and glucose (elution time of 12.64 h) (see Fig. 2). The resulting mathematical model is then two coupled algebraic expressions:

$$F = \bar{t}_S/(\bar{t}_S + \Delta t), \quad (3)$$

$$F = \bar{t}_G/(t_G + \Delta t), \quad (4)$$

where the subscript S designates the second normalization peak, *e.g.*, sucrose, and Δt is the constant shift of elution time. It is assumed that Δt is the same for sucrose and glucose. Further defining

$$R = \bar{t}_G/\bar{t}_S', \quad (5)$$

and solving eqns. 3 and 4 simultaneously we obtain:

$$\Delta t = \frac{t_G - Rt_S}{R-1} \quad (6)$$

The second unknown, F, can then be determined by substitution of the numerical value of Δt into eqn. 3 or 4.

Normalization to two peaks requires that Δt and F be calculated for each specific chromatogram as indicated in the preceding mathematical derivation. Each sugar

elution time for that specific chromatogram is then normalized by addition of the Δt correction to the elution time and multiplying the resulting sum by the factor, F:

$$t_X^* = (t_X + \Delta t)F \quad (7)$$

RESULTS AND DISCUSSION

Normal human urine specimens were chromatographed using a routine 20-h elution period[6]. The average elution positions of major peaks determined from the urine chromatograms of seven normal subjects are given in Table I. A comparison is made of the average peak elution times obtained by normalization to glucose only and those obtained by normalization to both glucose and sucrose.

TABLE I

AVERAGE PEAK ELUTION TIMES AND AVERAGE ELUTION TIMES NORMALIZED TO ONE PEAK (GLUCOSE) AND TO TWO PEAKS (GLUCOSE AND SUCROSE) AS DETERMINED FOR CARBOHYDRATE CHROMATOGRAMS OF THE URINE OF SEVEN NORMAL HUMAN SUBJECTS

Peak	*Without normalization*		*Normalization to one peak*		*Normalization to two peaks*	
	Average elution time (h)	*Standard deviation (h)*	*Average elution time (h)*	*Standard deviation (h)*	*Average elution time (h)*	*Standard deviation (h)*
A	0.39	0.02	0.39	0.02	0.39	0.02
B	0.58	0.03	0.58	0.04	0.58	0.02
C	0.66	0.03	0.66	0.04	0.66	0.03
Sucrose	1.22	0.03	1.22	0.05	1.22	0.00
Raffinose	1.44	0.04	1.44	0.05	1.44	0.02
Maltose	3.06	0.10	3.07	0.12	3.07	0.09
Lactose	3.34	0.16	3.35	0.17	3.35	0.15
Ribose	3.54	0.05	3.60	0.04	3.58	0.02
D	6.70	0.22	6.70	0.16	6.70	0.16
Fructose	9.13	0.24	9.13	0.05	9.12	0.06
Arabinose	9.57	0.25	9.56	0.05	9.56	0.06
Fucose	10.00	0.27	9.97	0.06	9.56	0.07
Galactose	10.53	0.30	10.53	0.16	10.53	0.17
Xylose	11.35	0.25	11.35	0.07	11.35	0.07
Mannoheptulose	12.08	0.27	12.08	0.03	12.08	0.03
Glucose	12.64	0.29	12.64	0.00	12.64	0.00
E	17.58	0.34	17.57	0.15	17.58	0.17
F	17.92	0.38	17.92	0.24	17.92	0.25

The uniformity of the elution positions, *i.e.*, the smaller standard deviation of the average elution time, is improved considerably in the portion of the chromatogram from 6 to 18 h when the peak elution times are normalized to one peak. For example, the standard deviation of elution time for arabinose is decreased by a factor of 5. Prior to 6 h and subsequent to 18 h there is relatively little change in the standard deviation of the elution positions. Normalization to two peaks further improves the uniformity of elution positions in the front of the chromatogram, from 0.6 to 6 h; for example, the standard deviation of the average elution time of ribose is decreased by

a factor of 2. Subsequent to 6 h, however, normalization to two peaks yields approximately the same results as normalization to one peak.

In computer evaluation of the chromatographic data, identification of peaks would be dependent upon the normalized peak elution position falling within a preassigned time interval. From the above data it can be seen that normalization allows one to significantly reduce that time interval. For example, assignment of the identity of fructose to a chromatographic peak would be dependent upon the normalized peak elution position falling within the time interval 9.12 ± 0.06 h (if plus or minus one standard deviation is found to be practicable). This would not interfere with the arabinose peak with a normalized time interval of 9.56 ± 0.06 h. However, in the case of the unnormalized time intervals, there can be overlap between these two peaks within the one standard deviation limits. This would make computer evaluation of the chromatogram almost impossible.

For completely automated computer identification the computer must, of course, first search the chromatogram for and determine the elution positions of the reference peaks used for normalization. This requires that such peaks be readily recognizable either by very constant elution position or some distinctive characteristic. In this case the two normalizing peaks were chosen because they are always present as very large peaks.

Gratitude is expressed to Mr. E. SCHONFELD, who contributed many helpful suggestions and assisted in the development of the computer programs used to evaluate the data.

The research was supported by the National Institute of General Medical Sciences.

Oak Ridge National Laboratory, Oak Ridge, Tenn. 37830 (U.S.A.)*

ROBERT L. JOLLEY
CHARLES D. SCOTT

1 C. D. SCOTT, J. E. ATTRILL AND N. G. ANDERSON, *Proc. Soc. Exptl. Biol. Med.*, 125 (1967) 181.
2 C. D. SCOTT, *Clin. Chem.*, 14 (1968) 521.
3 R. L. JOLLEY AND M. L. FREEMAN, *Clin. Chem.*, 14 (1968) 538.
4 C. D. SCOTT, R. L. JOLLEY AND W. F. JOHNSON, Prototype systems for the automated, high-resolution analyses of UV-absorbing constituents and carbohydrates in body fluids, *First Annual Symposium on Automated, High-Resolution Analyses in the Clinical Laboratory, Oak Ridge National Laboratory, March 13–14, 1969*, to be published as a supplement to *Am. J. Clin. Pathol.*
5 C. D. SCOTT AND J. M. JANSEN, JR., A computer technique for on-line analyses of complex column chromatographic data, *First Annual Symposium on Automated, High-Resolution Analyses in the Clinical Laboratory, Oak Ridge National Laboratory, March 13–14, 1969*, to be published as a supplement to *Am. J. Clin. Pathol.*
6 R. L. JOLLEY, J. L. JAINCHILL, C. D. SCOTT, K. S. WARREN AND M. L. FREEMAN, Carbohydrates in normal urine and blood serum as determined by high-resolution column chromatography, *First Annual Symposium on Automated, High-Resolution Analyses in the Clinical Laboratory, Oak Ridge National Laboratory, March 13–14, 1969*, to be published as a supplement to *Am. J. Clin. Pathol.*

Received December 8th, 1969

* Operated for the U.S. Atomic Energy Commission by the Union Carbide Corporation.