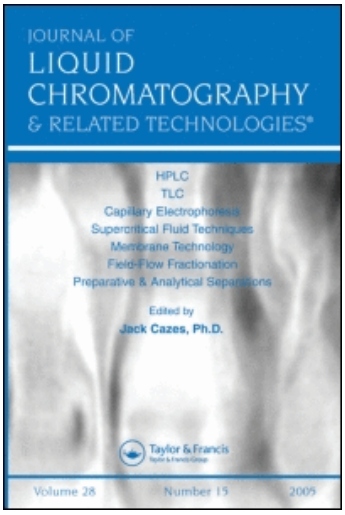

Journal of Liquid Chromatography & Related Technologies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

Data Acquisition and Processing for High Speed Liquid Chromatography

J. M. Anderson[a]
[a] Alltech Associates, Deerfield, IL

To cite this Article Anderson, J. M.(1983) 'Data Acquisition and Processing for High Speed Liquid Chromatography', Journal of Liquid Chromatography & Related Technologies, 6: 14, 2809 — 2828
To link to this Article: DOI: 10.1080/01483918308064949
URL: http://dx.doi.org/10.1080/01483918308064949

JOURNAL OF LIQUID CHROMATOGRAPHY, 6(14), 2809-2828 (1983)

DATA ACQUISITION AND PROCESSING FOR HIGH SPEED LIQUID CHROMATOGRAPHY

J. M. Anderson
Alltech Associates
2051 Waukegan Road
Deerfield, IL 60015

SUMMARY

The effect of sampling rate and digital smoothing on data acquired from high speed liquid chromatography (HSLC) is explored. The amount of data required per peak is determined from the precision of area, height, and retention time measurements. The correct sampling rate is related mathematically to column characteristics and operating conditions. The effect of a modified moving average digital smoothing routine on peak width and height is investigated. Digital smoothing functions are shown to behave similarly to analog noise filters. The merits of raw data storage and post analysis processing are discussed in light of the short analysis times in HSLC and the decreased cost of computer memory.

INTRODUCTION

Short HPLC columns filled with 3 micron particles, when operated at high flow rates, are capable of performing many separations 5 - 50 times faster than with conventional HPLC columns, with little or no loss of resolution. Specialized or modified HPLC equipment has been developed to take full advantage of this technique. The high efficiency and low volume of the

 0148-3919/83/0614-2809$3.50/0

peaks generated in high speed liquid chromatography (HSLC) require reduced volume in injectors, connecting tubing, and detector flow cells.[1] Detector electronics must also be modified to cope with the sharp, relatively high frequency peaks. Noise filtering networks must have low time constants so as not to distort peak shape and thereby decrease resolution.[1-3] In a similar fashion, the data handling system employed for HSLC must be carefully designed to accurately represent the chromatographic data in a digital format without significantly distorting peak shape.

This paper describes a data handling system for use with HSLC based on the Apple II computer. The effects of sampling rate and digital smoothing routines on chromatographic data are investigated. The use of low cost RAM memory for raw data retention is described.

EXPERIMENTAL

CHROMATOGRAPHY

A Gilson Model 303 Pump (Gilson Medical Electronics, Middleton, WI, USA) in conjunction with an LDC Mark III Pulse Dampener (LDC, Riviera Beach, FL, USA) was used as the solvent delivery system. Samples were injected with a Rheodyne 7410 internal loop injector fitted with either a 2 or 5 ul sample loop (Rheodyne, Cotati, CA, USA). Separations were performed on 100mm x 4.6mm columns packed with 3 micron RoSiL C18

DA (Alltech Associates, Deerfield, IL, USA). Tubing used to connect the column to the injector and detector was 1/16" OD x .004" ID (Alltech Associates, Deerfield, IL, USA). A Kratos 773 UV-Visible Detector (Kratos Analytical Instruments, Westwood, NJ, USA), equipped with a 0.5 ul flow cell, was used to monitor column effluent at 254 nm. The square wave rise time for the detector's noise filter was set at 100 msec. Samples and mobile phases were as noted with each figure.

DATA SYSTEM

An Apple II Plus with 48K RAM memory was used as the host processor. Peripherals consisted of a single floppy disk drive with controller and an Apple DOT Matrix Printer (All obtained from Alltech Associates, Deerfield, IL, USA). Analog to digital (A/D) conversion and data integration was done with an Analytical Computers' Chromcard (Analytical Computers, Elmhurst, IL, USA) supplied with an optional 128K RAM memory expansion card.

The Chromcard consists of a printed circuit board that fits into one of the empty slots in the back of the Apple and receives the analog signal from the detector's recorder output. The accompanying software digitizes incoming data with 12 bit precision at up to 20 Hz. The chromatogram is displayed in real time on the Apple's CRT, using the high resolution graphics mode.

The incoming data is smoothed using 3, 5 or 9 points and is fitted to a quadratic curve. The first and second derivatives are calculated to define the onset of the peak, the position of valleys between peaks, and the proper peak end point. User defined valves for slope sensitivity, minimum peak height and minimum peak area are used to discriminate against noise and other artifacts. Raw data points are stored in RAM memory for subsequent re-analysis or for transfer to disk for permanent storage.

Once all peaks have been defined and raw areas calculated, retention times are compared with expected values supplied by the operator. If experimental values match expected values within a user defined window, peak names are assigned. Raw peak areas are then normalized, compared with an internal standard (if desired), and multiplied by response factors prior to generation of the final report.

The final report lists all chromatographic conditions, together with operator name, date, and sample identification. Each peak is listed with retention time, area, area %, normalized area and concentration. A typical report is shown in Figure 1. The chromatogram may then be reproduced on the line printer in a dot-matrix format, as shown in Figure 2.

DATE 06/22/83 OPERATOR DSB SAMPLE EXPT1 SIZE 20 ULITERS
METHOD N1.M METHOD DESCRIPTION NUCLEOTIDES

COLUMN DESCRIPTION:
COLUMN ID 8 MM OD PACKING C 18 REVERSE TYPE RADIAL PAK
MOBILE PHASE 50 MM AMM/PHOS PH6 FLOW RATE 1 ML/MIN PRESSURE 500 PSI

OPERATING CONDITIONS:
INITIAL TEMP ENDING TEMP TEMP RATE
DETECTOR TYPE UV 258NM DETECTOR SENS 0.2 DETECTOR TEMP

PK#	COMPOUND NAME	TIME	HT	WIDTH	AREA	AREA%	NORM	CONC
1	GTP	5.18	7.48	14.38	9472	14.88		0.
2	GDP	5.67	10.90	20.88	17709	27.82		0.
3	GMP	7.51	17.79	23.06	36467	57.29		0.

FIGURE 1

DOT MATRIX REPRODUCTION OF CHROMATOGRAM

FIGURE 2

RESULTS AND DISCUSSION

The extremely narrow peaks generated by HSLC present a unique problem to chromatography data systems. These narrow peaks are a result of the high efficiency of HSLC columns, coupled with the speed at which peaks elute. The speed arises from the use of short columns (30 - 150 mm) and operation at high mobile phase velocities. A comparison of HSLC with conventional HPLC is shown in Figure 3. The conventional column (150 mm, 5 micron) separates seven components in just under 9 minutes. The HSLC column (100 mm, 3 micron) separates the same mixture in just under 50 seconds, with little or no loss of resolution. The amount of information available from both chromatograms is the same, but the time base has been decreased almost twelvefold. The data system, to maintain the same integrity of information, must digitize the chromatogram twelve times faster, requiring higher sampling rates. Simultaneously, the peaks shift to a higher frequency, increasing the effect of digital smoothing routines on the peak shape.

A comparison may be made between these digital effects and their analog counterparts. Faster sampling rates correspond to faster strip chart recorder speeds. Improved digital smoothing routines parallel analog noise filters with shorter time constants.

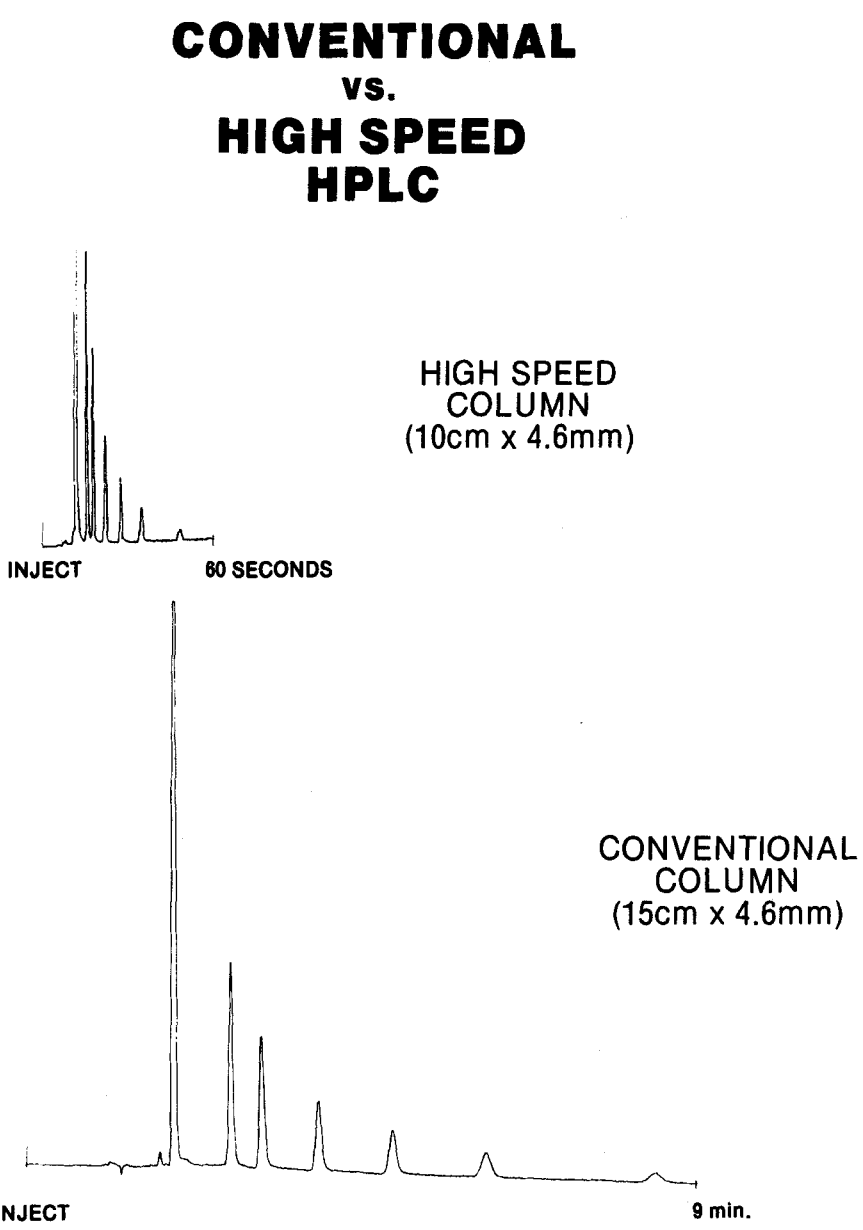

FIGURE 3 Conventional Column, 15cm x 4.6mm P/W RoSiL C18 HL 5 micron. High Speed Column 10cm x 4.6mm P/W RoSiL C18 HL 3 micron. Mobile Phase 75/25 Acetonitrile/Water. 1.0ml/min for Conventional Column, 3.0ml/min for High Speed Column. Sample - Anilides.

SAMPLING RATE

The sampling rate determines the interval between consecutive examinations of the analog signal. The speed of the analysis and the shape of the peaks determine the sampling rate required--the sharper the peaks, the greater the sampling rate. The sampling rate should be fast enough to provide an adequate digital representation of the peak profile. On the other hand, excessive sampling rates will generate large quantities of data. If raw data is being stored, this rapidly fills the computer's memory.

The sampling rate required for a given peak can be related to chromatographic parameters using Equation (1).

$$(1) \quad W = \frac{4\ Vo\ (1 + k^1)}{Q\ N^{\frac{1}{2}}}$$

W = Peak Width (at Base)

Vo = Column Void Volume

Q = Volume Flow Rate

N = Column Efficiency

k^1 = Capacity Factor

If the peak duration (W) is known, it is simple to calculate the sampling rate, using Equation (2).

$$(2) \quad F = \frac{C\ Q\ N^{\frac{1}{2}}}{4\ Vo\ (1 + k^1)}$$

F = Sampling Rate

C = The Number of Data Points Desired per Peak

Inspection of Equation (2) reveals that sampling rate increases with increased flow rate (Q), increased efficiency (N), and decreased column void volume (Vo). All three of these occur in HSLC. Smaller particles result in higher efficiencies; shorter columns have smaller void volumes, and shallow H/U curves for small particles allow operation at higher flow rates.

Choosing the correct value for C depends on the precision required from subsequent calculations. In most cases area, height, and retention time are extracted from raw peak data. Examining the precision of area, height, and retention time at various sampling rates may be used to determine the correct value for C.

The Chromcard system allows A/D conversion rates of 20 Hz, 10 Hz, 2 Hz, 1 Hz, and 0.5 Hz. For the high speed separation shown in Figure 4, A/D conversion rates of 20 Hz, 10 Hz, and 2 Hz were examined. Ten runs were collected at each rate. The earliest eluting peak (phenol) is the sharpest of the five, hence presenting the most demanding case to the data handling system. The mean values for area, height, and retention time, along with the relative standard deviation (RSD) for each quantity are shown in Table 1.

The mean values for retention time, peak height and peak area are nearly identical when calculated from data collected at 10 Hz and 20 Hz. As expected, the precision obtained from data collected at 20 Hz is slightly

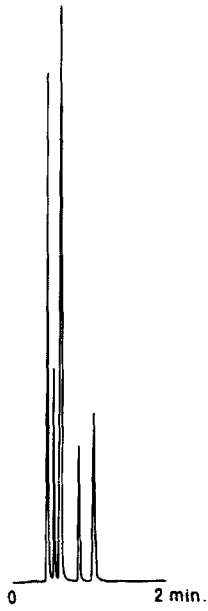

FIGURE 4 Column 10cm x 4.6mm P/W RoSiL C18 HL 3 micron. Mobile Phase 75/25 Acetonitrile/Water. Flow 3.0ml/min 3000 PSIG. Peak 1, Phenol; Peak 2, Benzaldehyde; Peak 3, N,N-Diethyl-m-Toulamide; Peak 4, Toluene; Peak 5, Ethyl Benzene.

TABLE 1

Integrator Precision at Various A/D Conversion Rates*

Sampling Rate (Points/Sec.)	Mean Value/% RSD Retention Time (Seconds)	Peak Area	Peak Height
20	21.83 (0.771)	1522.55 (1.329)	2712.108 (1.271)
10	21.84 (0.418)	1536.92 (1.512)	2622.113 (1.497)
2	21.95 (1.293)	1623.90 (16.38)	1901.200 (12.34)

*10 runs at each rate. Data for Peak 1, phenol.

better, although the maximum RSD for 10 Hz is only 1.512%. This is acceptable precision for most applications. When the sampling rate is dropped to 2 Hz, however, the mean values for peak height and peak area differ greatly from those obtained at 10 and 20 Hz. The precision at the 2 Hz rate also becomes unacceptable (RSD>12%).

This information suggests that the number of points collected at the 10 Hz and 20 Hz rates provides an adequate data base for integration with good precision. Table 2 displays the number of data points obtained experimentally at each sampling rate, along with the theoretical value obtained by solving Equation (2) for C.

TABLE 2

Number of Data Points Collected at Various A/D Conversion Rates

Sampling Rate (Points/Second)	Number of Data Points	
	Experimental*	Calculated**
20	19.6	16.4
10	10.1	8.2
2	3.2	1.6

* Average values for 10 runs. Data for Peak 1, phenol.

** Calculated using $C = \frac{4 F Vo(1 + k^1)}{Q N^{\frac{1}{2}}}$, from Eqn (2). N = 11,300 plates.

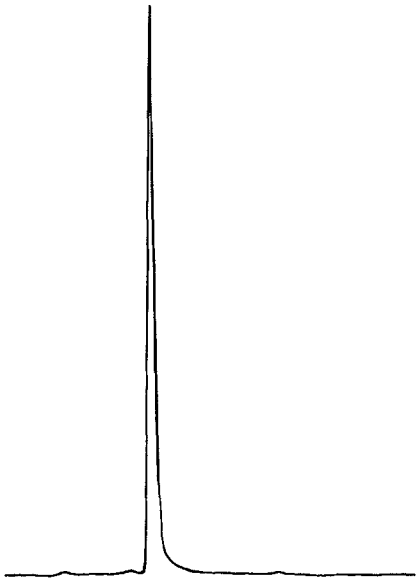

FIGURE 5 Phenol Peak Profile

The lack of correlation between the experimental and calculated values may be explained by the non-Gaussian shape of the phenol peak. The efficiency value used in calculating C by Equation (2) was obtained using bandwidth at 50% of peak height. This method assumes a Gaussian peak shape, and as Figure 5 reveals, the phenol peak has a significant tail. This leads to an inflated value for N and a reduced value for C. Nonetheless, Equation (2) provides a reasonable method for choosing sampling rate. This data indicates that a C value of 10 - 20 points/peak (minimum) will provide acceptable precision.

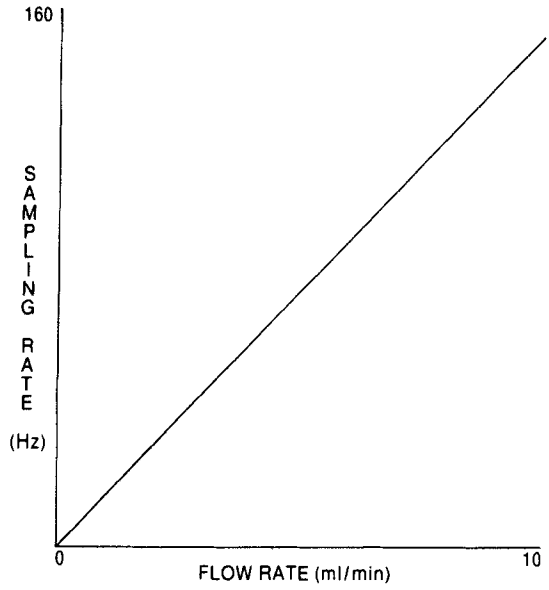

FIGURE 6 Sampling Rate vs. Flow Rate

For a typical 3 micron HSLC column 100 mm long with 12,000 plates and a Vo of 0.6 ml, then substituting in Equation (2) with C = 20, yields Equation (3).

$$(3) \quad F = \frac{15.22\ Q}{(1 + k^1)}$$

A peak eluting at the void volume ($k^1 \approx 0$) will require the fastest sampling rate. For HSLC with 4.6mm columns, flow rates of up to 6 ml/min may be encountered. Figure 6 displays the plot of sampling rate vs. flow rate, for a peak eluting at the void volume. For a flow rate of 6 ml/min, sampling rates of

close to 100 Hz may be required to generate 20 data points. Chromcard's current maximum sampling rate of 20 Hz should be adequate for most HSLC and for all conventional HPLC. However, to provide adequate precision and accuracy for the fastest analyses, a system with higher sampling rates would be desirable.

DIGITAL SMOOTHING

Most commercial LC detectors incorporate some type of analog noise filter as part of the signal electronics. The purpose of these filters is to eliminate noise at frequencies higher than that of the chromatographic peaks. Most of these filters can be characterized by their time constant. The time constant should be high enough to remove short term noise without distorting the peak shape. If a 5% increase in bandwidth due to noise filters is deemed acceptable, then the time constant should not exceed 32% of the peak's bandwidth.(3) Haddad, et al., quantifies the effects of RC filters on chromatographic efficiency and resolution.(2) As the peak's bandwidth decreases, the effect of the noise filters on the peak shape increases. Analog noise filters for HSLC have been designed with small time constants so as not to significantly degrade peak shape.

Chromatography data systems usually include software for smoothing digitized data. The purpose is similar to that of analog noise filters--to remove

spurious signals from chromatographic data. The simplest of these numeric noise suppression methods is the moving average. The values of several consecutive data points are summed and divided by the number of points collected. This average is used for the value of the middle point in the array. The oldest point is dropped; a new point is added, and the process is repeated for the new array. In this manner, a "window" moves through the data, assigning a value for the middle data point in the window based on the past and future points in the array. The number of points used in the average (the size of the window) determines the extent of the smoothing. The length of the smoothing array can be qualitatively compared with the magnitude of the time constant used in analog filters. The more points used in the smoothing function, the greater the distortion of peak shape. As the peak's bandwidth becomes smaller, the effect of the moving average on the peak shape becomes more significant.

The software used with Chromcard uses a modified moving average as described by Savitsky and Golay.(4) The software operates on a 3, 5, or 9 point window. Instead of calculating a simple average of the data points, the data in the array is fitted to a quadratic curve ($y = ax^2+bx+c$) using a least squares approximation. A quadratic curve fit is superior to a simple

moving average in that the quadratic function can be fitted more correctly to a rapidly changing signal such as a chromatographic peak. The result is better smoothing with less distortion.

The effect of the size of the window on peak shape is shown in Table 3. Using a 3, 5, or 9 point window, data from the phenol peak (Fig. 4) was acquired and smoothed with a quadratic curve fit. Peak height and peak bandwidth (at 50%) were measured for the smoothed data. Efficiency was calculated using the 1/2 height method.

Inspection of bandwidth and peak height values in Table 3 demonstrates that digital smoothing affects peak shape in a manner similar to analog noise filters. If the smoothing function begins to affect signals in the same frequency range as the chromatographic peaks, then the apparent bandwidth will increase, decreasing efficiency. Peak height will decrease, affecting

Table 3

Effect of Digital Smoothing on Peak Width and Height

Points Used in Smoothing Routine	Peak Height	Peak Width (50%)	Efficiency (Plates)
3	3529	6.750	12,100
5	3526	6.750	12,100
9	3385	7.125	11,000

sensitivity. In the case of the quadratic curve fit program applied in this work, no decrease in performance due to digital smoothing occurred for the 3 or 5 point windows. Using 9 points, however, reduced measured efficiency by 8.5% and peak height by 4.1%.

Unlike analog noise filters, whose characteristics are fixed by the value and nature of their electronic components, digital smoothing functions may take on almost any characteristic. The virtually unlimited ability to manipulate data through software makes this so. The Savitsky-Golay procedure, for example, may be modified to fit the raw data to virtually any mathematical function. In the future, it may be possible to dispense completely with analog signal processing in favor of more powerful, versatile digital methods.

When applying digital smoothing to chromatographic data, care must be taken to avoid distortion of peak shape. When the frequency of the peaks becomes closer to that of the noise to be eliminated, such as in HSLC, the design of the digital smoothing becomes increasingly important.

RAW DATA STORAGE

Chromatography data systems may be categorized according to the way they treat raw data. So called "on-the-fly" systems process data as it is received from the A/D converter, determining when peaks occur and

performing integration as the analysis takes place. Raw data values are discarded after they are processed. Small microprocessor based data systems usually have been limited to operation in this mode due to the high cost of solid state memory and the limited addressing capacity of 8 bit processors. In some cases, raw peak data can be stored in a small solid state memory buffer, then transferred to disc when the buffer is filled. The buffer is then filled with new data until the buffer once again requires "dumping" to disc. Unfortunately, transfer of data to a typical floppy disc can take up to one full second, during which time data acquisition must be interrupted. This places a severe limitation on sampling rate.

Larger computers with increased memory capacity have normally been employed for peak processing with retention of raw data. The major advantage of raw data retention is the ability to reanalyze the data under an alternate set of conditions. Peak processing parameters may be modified after the analysis to cope with unforeseen changes in the incoming data. For example, peak threshold and minimum area parameters can be adjusted to eliminate an unexpected peak from calculations.

In 1980 Reese predicted that data systems would be developed to operate with raw data retention as the cost of memory decreased.(5) Since that time, advances in

small personal computers have made fairly sophisticated computing power available at modest costs. Relatively inexpensive, high volume, solid state memory has also been developed making it possible to retain a fair amount of raw data with a microprocessor based system.

HSLC is an ideal candidate for systems of this nature since the amount of data generated is relatively small. Although the amount of data per unit time is high in HSLC due to the high sampling rates, the analyses themselves are typically quite short. In most cases, the entire separation takes place in less than ten minutes.

The data system employed in this work was equipped with a 128K RAM memory expansion card for the purpose of

Table 4

Data Storage Capacity at Various Sampling Rates*

Sampling Rate (Hz)	Approximate Storage Capacity (Minutes)
100	10
75	13.34
50	20
40	25
20	50
10	100

*Based on 128K system.

raw data storage. Table 4 shows the amount of raw chromatographic data (in minutes) that may be stored with 128K memory at various sampling rates. Even at sampling rates of 100 Hz, ten minutes of raw data may be stored in a 128K buffer. This is certainly adequate for even the most demanding HSLC.

CONCLUSION

With the advent of HSLC, the components of the LC system have been critically reviewed and modified to cope with the narrow, high speed peaks produced by HSLC columns. The data system must be viewed in the same light. Sampling rates should be adjusted to produce a minimum of 10 - 20 points per peak. Digital smoothing programs should be carefully evaluated to insure that no artificial decreases in performance are introduced. With the advent of low cost RAM memory to complement microcomputers, data processing with retention of raw data is possible.

REFERENCES

1. Cunico, R. L., Varian Instrument Applications, 16 (2), (1982) p. 9.

2. Haddad, P. R., Keating, R. W., and Low, G. K. C., J. of Liquid Chromatogr., 5, (1982) p. 853.

3. Kucera, P., J. of Chromatogr., 198, (1980), pg. 93.

4. Savitzky, A., and Golay, M., Anal. Chem., 36, (1964), pg. 1627.

5. Reese, C. E., J. of Chromatogr. Sci., 18, (1980), pg. 249.