Quantitative Comparison of Combined Gas Chromatographic/Mass Spectrometric Profiles of Complex Mixtures

Dennis H. Smith* and Michael Achenbach

Departments of Genetics and Chemistry, Stanford University, Stanford, California 94305

William J. Yeager, Patricia J. Anderson, William L. Fitch, and Thomas C. Rindfleisch

Department of Genetics, Stanford University, Stanford, California 94305

We describe a method, termed HISLIB, for qualitative and quantitative comparisons of complex mixtures of organic compounds. Our method compares combined gas chromatographic/mass spectrometric (GC/MS) profiles of new mixtures with historical libraries of GC/MS data on related mixtures. Co-occurrence of components is established by matching both retention indexes and mass spectra after background removal and resolution of overlapping GC components. Quantitation is achieved by comparing relative concentrations of components, calculated using internal standards. Uses include validation of analytical procedures, determination of variations among controls, and rapid detection of novel (in identity or amount) components in new mixtures.

Improvement of combined gas chromatographic/mass spectrometric (GC/MS) instrumentation, including automation of many previously manual procedures, has resulted in the routine generation of large volumes of data. There has been a steady progression of developments in computer-based procedures for analysis of these data. The first major development to assist scientists in data analysis was library search techniques (1, 2). These techniques are valuable not only in identification of previously observed components, but also in noting which components are not found in the library and are thus subjects for more sophisticated procedures for structure elucidation. More recently, relative retention indexes (RRI's) have been used to improve the specificity of identification in cases where related (e.g., isomeric) compounds exhibit similar mass spectra (3–5). To improve the quality of mass spectra obtained from GC/MS systems, which facilitates both library matching and interpretation of spectra of unknown compounds, computer programs have been developed to remove background and resolve overlapping GC components (6, 7).

We describe a logical synthesis and extension of the above procedures designed to automate the task of quantitative comparison of GC/MS results obtained on complex mixtures of organic compounds. A method for qualitative comparison of GC/MS profiles to detect anomalous compounds has been reported (8). But that method is limited to comparison of two sequential analyses and does not provide detailed quantitative information. Our developments remove these restrictions. We carry out quantitative comparisons which couple the specificity of the mass spectrum and RRI to identify each compound (3–5), with calculated relative concentrations to determine their relative amounts. This method, "HISLIB", is based on comparing new mixtures to "historical" libraries of previous results. It is capable of detecting new components and anomalous concentrations of previously encountered com-

ponents. Applications of HISLIB include: (a) validation of analytical procedures used to isolate complex mixtures; (b) development of historical libraries which might include complete summaries of all past observations, compilations of controls, or any other selected subset of results; (c) computation of average mass spectra and RRI's of known compounds to improve the quality of existing libraries of mass spectral data; and (d) rapid comparison of new data to previously compiled library(ies) to detect differences in kind and/or amounts of individual components.

During preparation of this report, a paper appeared (9) which addresses some of the same issues raised in our discussions. In fact, that paper utilizes concepts and earlier programs and data from our own laboratory. The work described in our report represents the results of a maturation of these concepts and programs and the development of new programs and GC/MS procedures designed specifically to obtain reliable, quantitative results. Indeed, several of our new developments are solutions to problems discussed by Blaisdell as deficiencies of his method (9).

EXPERIMENTAL

We routinely collect complete GC/MS runs, including repetitively scanned mass spectra. Any system capable of providing these data is potentially adequate; we employ a Finnigan Instrument Corp. model 1015 quadrupole mass spectrometer controlled by a Digital Equipment Corp. (DEC) PDP-11/20 computer (10). Subsequent data processing is done on a DEC PDP-11/45 with 28K words of core memory, a 5M word disc drive, teletype, printer, CRT, and Versatek printer/plotter. Unless otherwise noted, GC and GC/MS experiments were performed on a Finnigan Instruments Corp. model 9500 gas chromatograph, employing 6-foot U-shaped $^1/_8$-inch i.d. columns, packed with 10% OV-17 on 100/120 mesh Gas-Chrom Q. Initial temperatures (usually 80 °C) were maintained for 4 min followed by temperature programming at a rate of (usually) 4 deg/min.

For optimum use of HISLIB, it is desirable to apply a number of preprocessing steps to experimental data. Because library matching, determination of RRI's and, particularly, measurements of relative concentrations depend strongly on spectra free from background and overlapping components, we first process the GC/MS data with the CLEANUP (7) program. Next we determine RRI's for each detected component, and compute relative concentrations based on one or more internal standards. We then match each spectrum against an existing library of mass spectral data, in our case a library of compounds of biological interest (11). Finally, the resulting data are combined with previous results to update the historical library or are compared against an existing historical library. The flow of data through these steps is summarized in Figure 1.

The HISLIB system can be used without applying some of the processing steps above. However, without "clean" spectra, both library search results and RRI's are compromised, especially for components of low abundance. Accurate quantitation becomes effectively impossible. RRI's are very important to increase the
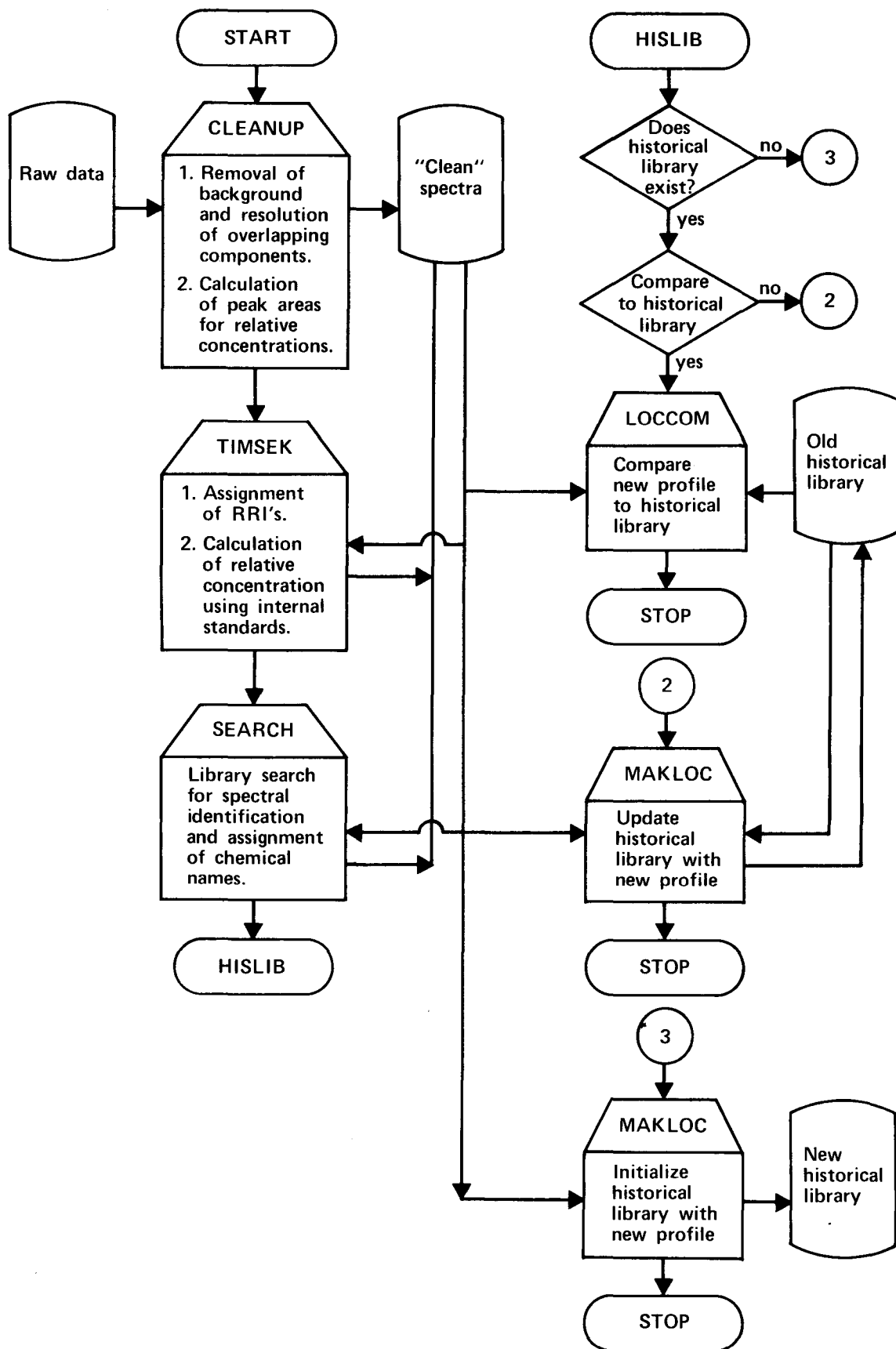
Figure 1. Major steps in processing a complete set of GC/MS data to establish and search an historical library

specificity of mass spectral identifications; the combination of both is highly effective in distinguishing materials with similar spectra (3-5). Matching spectra against a compendium of spectra of known compounds (as opposed to the historical libraries discussed here) is also not essential, but is useful in assigning

names of known compounds to their spectra as a guide in interpreting the results of comparisons to the historical library.

The following sections describe the details of our method, assuming these preprocessing steps are performed. Further information about the programs described below, including

availability, may be obtained by writing the authors.

Automatic Determination of Relative Retention Indexes (RRI's). We use an extension of the method proposed by Nau and Biemann (3, 4) for determination of RRI's. Our procedure is automatic and calculates reproducible RRI's under variable instrumental and experimental conditions including unavoidable changes in initial GC column temperature, carrier gas flow, or temperature programming rates. It requires only three internal hydrocarbon standards for the analysis of a GC/MS run.

As previously described (3, 4), each column is calibrated with a mixture of hydrocarbons (we use 1 μL of an approximately 1 μg/μL solution each of n-C_{10} through n-C_{26}, and n-C_{28}). This calibration yields a file of 18 data points relating carbon numbers and mass spectrometer scan numbers. Each subsequent GC/MS experimental run using that column is processed using this calibration file as a reference (assuming that conditions of temperature programming, initial temperatures, and flow rate are approximately the same—see Results and Discussion). Three of the hydrocarbons used in the calibration run are added to each experimental mixture. The CLEANUP program is run to isolate representative spectra and to assign scan numbers corresponding to elution times for each component. The TIMSEK program (Figure 1) then locates the three added standards by matching their known spectra in windows about the expected elution scan numbers and fits the three observed hydrocarbon scan numbers to those corresponding in the calibration run using a least-squares method. We assume that differences in conditions between a given experimental run and the calibration run can be accounted for by a linear transformation of the elution time scale as given in Equation 1a. We determine the linear coefficients A and b by maximizing the correspondence between the elution times of the three standards in the experimental and calibration runs; or equivalently minimizing the error function given in Equation 1b with respect to A and b.

$$S_{cal} = AS_{exp} + b \tag{1a}$$

where S_{cal} is a scan number in the elution time scale of the calibration file, S_{exp} is a scan number in the elution time scale of the experimental run, and A and b are the linear transformation coefficients.

$$E^2 = \sum_{i=1}^{n} [S_{i(cal)} - AS_{i(exp)} - b]^2 \tag{1b}$$

where $S_{i(cal)}$ is the scan number of the ith standard in the calibration file, $S_{i(exp)}$ is the scan number of the ith standard in the experiment file, and i indexes over the internal standards used (n = 3 in our case).

Once A and b are determined (Equation 1b), Equation 1a is used to determine the effective scan number for elutants in the experimental run as transformed to the calibration run time scale. These effective scan numbers are converted to RRI's by a linear interpolation or extrapolation using the nearest hydrocarbons measured in the calibration file (3, 4). (If the GC is operated isothermally, a logarithmic interpolation/extrapolation is used.)

This method differs from that of Nau and Biemann in that the least squares fitting procedure (Equation 1b) takes explicit account of both linear offsets and expansion or contraction of the scan number/retention index curve rather than simply optimizing about the midpoint of the range (3, 4).

Determination of Relative Concentrations. Relative concentrations are determined by TIMSEK (Figure 1) based on any one or combination of the internal standards selected by the user prior to obtaining GC/MS data. Ideally, standards should be chosen that reflect the kinds of compounds one wishes to quantitate, the variety of analytical procedures used to isolate mixtures to be analyzed, the sensitivity of spectra to changing MS conditions, and other considerations that affect accurate and reproducible quantitation using any analytical procedure. We wish only to point out that care must go into the selection and use of such standards. TIMSEK uses a pre-established library of spectra of standards together with their RRI's. The standard(s) selected is searched for in the GC/MS data by looking for the closest spectrum match (Equation 4 below) within a narrow retention index window (+/−0.2 methylene unit). This is similar to the method of Sweeley et al. (5). Having found the internal
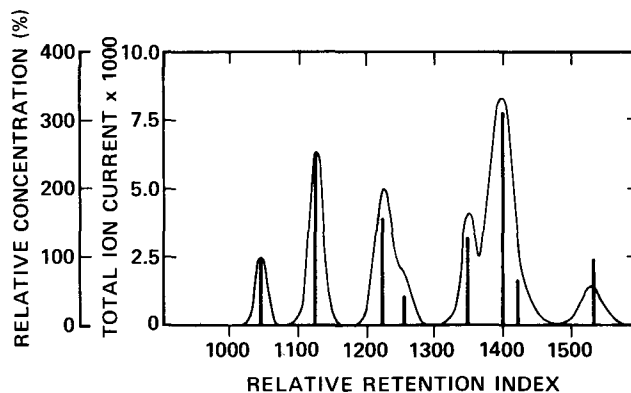
Figure 2. A schematic representation of a GC/MS profile, displaying relative concentrations of detected elutants vs. relative retention index, superimposed on the corresponding total ion current plot. For each detected elutant, the total ion current profile is effectively collapsed (Equations 2 and 3) into a line whose height represents relative concentration

standard(s), the relative concentration of the ith component is calculated according to Equation 2,

Rel Concentration = 100 $\times$

$$\frac{\text{Areal TIC of } i\text{th component}}{\text{Areal TIC of internal standard}} \tag{2}$$

The "areal" total ion current (TIC) measures the area of the GC peak of the ith component, not simply its height. The area for each GC peak is derived from the raw mass spectral data using the peak model determined for each spectrum during CLEANUP (7). The intensity (ion abundance expressed as peak height) of each mass in the spectrum of the ith component is determined by fitting the data-adaptive peak model to the intensity profile for each mass (fragmentogram) about the position of elution of the component (7). Simpson's Rule is used to determine the area of the model peak. The areal total ion current for the ith component is given by Equation 3,

$$\text{Areal TIC}_i = \frac{A_{i(model)}}{h_{i(model)}} \sum_m I_{im} \tag{3}$$

where $A_{i(model)}$ and $h_{i(model)}$ are the area and height of the peak model for the ith component and I_{im} is the ion abundance (peak height) at mass m in the mass spectrum of the ith component after processing by the CLEANUP program.

If more than one standard is used, the basis for relative concentrations is the average of the areal total ion currents for the standards. The inclusion of multiple standards provides the opportunity for a better statistical basis for computing relative concentrations since statistical fluctuations in measuring the areal TIC of one are reduced by averaging with the areal TIC's of the others. Depending on the relative quantities and reproducibility of the various standards included, a weighted average may be appropriate to account for different relative a priori uncertainties in the TIC's among them. In our case, these are comparable and a straightforward average is used. An improvement in quantitation standard reproducibility can be expected increasing approximately as the square root of the number of standards included. Results illustrating the advantages of multiple standards are presented in Results and Discussion.

Assembling an Historical Library of GC/MS Profiles. We define a "profile" for a GC/MS experiment as an assembly of data consisting of: (a) The (unnormalized) spectrum of each component after component detection, background removal and resolution of overlapping components; (b) the retention index of each component; (c) the relative concentration of each component; and (d) (optionally) a name for each component which may be a simple experiment code or a name associated with the component during routine library search (Figure 1). A GC/MS profile by this definition may be visualized as shown schematically in Figure 2. The relative concentrations are depicted as vertical bars at the appropriate elutant locations superimposed on a normal total ion current plot (total ion current vs. RRI). The height of each bar

corresponds to the areal total ion current or relative concentration through Equations 2 and 3. The relative heights of the bars will approximate the relative heights of the respective peaks in the total ion current plot. However, depending on the area/height ratio of the model peak for each elutant (Equation 2), the relative concentration can differ substantially from the peak height in the total ion current plot (e.g., compare the first and last peaks in Figure 2, both of which have relative concentrations of 100%).

An historical library is assembled by HISLIB by taking the GC/MS profile from an experiment and adding it to the library (Figure 1). If the library is initially empty, the profile becomes the library. If the library already contains at least one profile, the new profile is added as follows. Each spectrum in the new profile is compared to each spectrum in the library within a narrow retention index window (e.g., $+/-0.2$ methylene, or $+/20$ RRI, units for our work). A spectral match score, in this case a cross-correlation score, is calculated by Equation 4,

$$\text{Spectral Score} = 1000 \times$$

$$\frac{\left[\sum_m e_{m(\text{prof})} \; e_{m(\text{hist})} \right]^2}{\sum_m e_m^{\;2}{}_{(\text{prof})} \; \sum_m e_m^{\;2}{}_{(\text{hist})}} \qquad (4)$$

where spectra are reduced to the two most abundant ions every 14 amu (12) and the spectral intensities are encoded before matching. $e_{m(\text{prof})}$ and $e_{m(\text{hist})}$ are the encoded intensities at mass m for the new profile and the historical library, respectively. They are quantized to have values 0, 1, 2, or 3 corresponding to the relative intensity ranges 0–4, 5–16, 17–64, and 65–100% of base peak, respectively.

The definition in Equation 4 has several useful properties, based on Schwartz's inequality (13). The spectral match score calculated is independent of the order in which spectra are compared. If two ions of the same mass are present, a positive contribution to the score results. More abundant ions are weighted more heavily because of the squared term. The score is guaranteed to be between zero and 1000, 1000 representing a perfect match. Equation 4 is similar to the "degree of coincidence" score used by Jellum et al. (8), except that Equation 4 uses encoded peak heights rather than just the number of peaks.

The spectral match score and the proximity of the retention indexes are combined through an heuristic evaluation function (Equation 5a) which yields the final score. This final score is the spectral match score weighted by a trapezoidal function (Equation 5b) which penalizes for disparate RRI's. The weight is unity if the difference in RRI's is less than five units and decreases linearly to a threshold weight as the absolute difference in RRI's becomes greater than 5 units up to the empirical cutoff of 20 RRI units.

$$\text{Final Score} = \text{Spectral score} \times W(\Delta \text{RRI}) \qquad (5a)$$

Where $\Delta \text{RRI} = (\text{RRI}_{\text{exp}} - \text{RRI}_{\text{lib}})$ and RRI_{exp} and RRI_{lib} are the relative retention indexes for the experiment and library components respectively. The weighting function, $W(x)$, is defined by,

$$W(x) = 1 \qquad\qquad ; |x| < 5 \text{ RRI units}$$
$$= 1 - \frac{(\text{Maxscore} - \text{Minscore})}{15 \; \text{Maxscore}} x \quad ; 5 \leqslant |x| < 20$$
$$= 0 \qquad\qquad ; 20 \leqslant |x| \quad (5b)$$

where Maxscore = 1000 and Minscore = 400.

If this final score exceeds 400, the experiment compound is considered a potential match to the library compound. If there is more than one potential match between closely eluting experimental and library compounds, the ambiguity is resolved by a procedure (see below) that maximizes the overall correspondence between the pattern of experimental and library elutants. The Minscore value of 400 was derived empirically by examining the distribution of scores obtained by matching every nth spectrum in chemically related subsets of our library (11) with all spectra in that subset. A representative distribution of the number of matches with a given score as a function of score is shown in Figure 3. From a number of such curves a value of 400 was chosen as
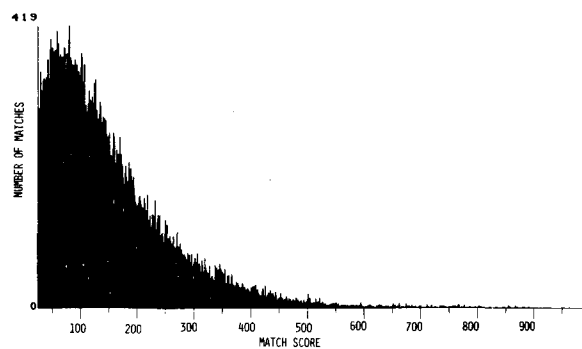

Figure 3. A histogram of the number of library comparisons with a given match score vs. score. Scores were obtained using Equation 4. Comparisons were made in two subsets of the library of Markey et al. (11) containing 750 compounds total. Every 5th spectrum was compared with the other spectra in each subset for a total of approximately 65 000 comparisons

a threshold for distinguishing matches and non-matches (this threshold will depend to some extent on the range of compounds included in the library, derivatization procedures, etc.). Although artifactual matches may occasionally yield scores higher than the threshold, the RRI weighting (Equation 5a) significantly reduces such occurrences.

Assignment of New Spectra to the Historical Library. The final step in correlating a new profile with an historical library involves selecting between alternative matchings of experimental and historical library spectra with similarly high final scores. This occurs frequently with isomeric compounds with similar spectra and retention indexes, and accidentally as, for example, with compounds whose spectra are similar due to domination of the fragmentation pattern by ions from a functionality added during derivatization.

We have implemented a pattern matching procedure to resolve such ambiguities. Briefly, the procedure attempts to maximize the consistency between a new experiment and the library, assuming they are derived from similar mixtures. In a region containing ambiguities, a matrix representing every possible assignment relating experiment and library spectra to one another is analyzed using an algorithm which can trace and rank all self-consistent "paths" through the matrix (14). Such paths include those which create new entries in the historical library, i.e., paths with some spectra in the new profile not being matched to any existing spectrum in the historical library. Consistency constraints on the assignments include: (a) the scoring threshold must be exceeded for a match to be considered, (b) RRI order must be preserved, and (c) a spectrum in either set can be assigned to at most one spectrum in the counterpart set. Finally, the "best" assignment is that which has the highest total score, where the total score is the summation of scores (Equation 5a) for each candidate pairing of spectra between the two sets (the score is not incremented for a spectrum found to be only in one set). This procedure is driven strongly toward maximum overlap between the two sets of spectra. This is justified when the minscore is high enough to reject dissimilar spectra and the GC/MS profiles are from related mixtures.

Once specific assignments have been made, spectra from the new profile are added to the library. New entries are created for components which scored less than minscore against library entries, or which were assigned as new entries by the above pattern-matching algorithm. When a pairing with an existing library entry is made, the new spectrum is averaged with the library spectrum for that entry, effectively weighting each contributing spectrum by its total ion current. At the same time, the new relative concentration and retention index are averaged with the previous values. Note that an important advantage of this approach is that components need not be identified by name, only by occurrence in terms of RRI and mass spectrum (9).

Comparing New Profiles to the Historical Library. Once a suitable historical library has been prepared, subsequent profiles can be compared to it to detect similarities and differences. In practice, we use the same program used to assemble the library to perform the comparisons, changing only a flag which prevents

using the new data to update the library and which causes a summary output to be produced indicating the results of comparison. Individual users may select different formats for such a summary. The one used in our laboratory (see Results and Disucssion) was chosen to focus the attention of the user on components observed in significantly different relative concentration and on new components present in the profile regardless of relative concentration.

Manual Method of Extraction of Urinary Organic Acids. To 3 mL of freshly thawed urine is added an aliquot of *m*-chlorophenylacetic acid solution (84 μg, 0.49 μmol, in H_2O) as an internal standard for quantitation. The urine is then acidified with six drops of 3 N hydrochloric acid and extracted three times with 1:1 ether–ethyl acetate (6 mL total). The combined organic extracts are dried (Na_2SO_4) and evaporated to dryness in vacuo. The resulting residue is dissolved in methanol–ethyl acetate (1:1, 3.0 mL) and a 1.0-mL portion of this is transferred to a Teflon-capped glass vial. The solvent is blown off with a stream of nitrogen.

DEAE-Sephadex Anion Exchange Method of Extraction of Urinary Organic Acids. As in the manual method, *m*-chlorophenylacetic acid (84 μg, 0.49 μmol) is added to 3.0 mL of urine in a 12-mL centrifuge tube. Barium hydroxide solution (0.1 M, 3.0 mL) is added, the mixture is quickly stirred and centrifuged for 15 s. The supernatant is removed and treated with hydroxylamine hydrochloride (50 mg, 0.7 mmol). This mixture is heated at 60 °C for 30 min, allowed to cool, and neutralized to pH 7–8 with dilute hydrochloric acid. This solution of organic acids and oximes of keto-acids is then placed on a DEAE-Sephadex A-25 column (1.0 cm × 5.0 cm) prepared as previously described (15). After the acid solution is passed onto the column, the resin is washed twice with distilled water (5.0 mL) to remove neutral and basic constituents. The organic acids are then eluted with 1.5 M pyridinium acetate solution (15 mL). An aliquot of this elutant (5.0 mL) is lyophilized to dryness at 5–10 μ pressure, the residue taken up in methanol–ethyl acetate (1:1, 2 mL) and transferred to a Teflon-capped glass vial. The solvent is blown down with a stream of nitrogen.

Trimethylsilylation. The urinary acids (from either of the above procedures) are treated with *N,O*-bis(trimethylsilyl)trifluoroacetamide ("BSTFA", 100 μL) and heated at 60 °C for 30 min. Before analysis, a solution of hydrocarbon standards (5 μL of a 5 μg/μL solution of dodecane, octadecane, and tetracosane) is added as a reference for RRI calculations.

RESULTS AND DISCUSSION

Relative Retention Index Calculations. Because of the increasingly important role of relative retention indexes in analysis of GC/MS profiles (this study and Ref. 3, 4, and 5), we have evaluated our method, TIMSEK, for calculation of RRI's (Experimental section) in several ways. We made radical changes to temperature programming rates, starting temperatures, and carrier gas flow rates for GC/MS runs subsequent to a calibration run. These changes simulate perturbations of the system far beyond what we expect in normal operation.

We present in Figure 4a plots of carbon number vs. scan number for three GC temperature programming rates, 4, 6, and 8 °C/min. Using the 4 °C/min GC/MS experiment as the calibration run, we used the method described previously (see Experimental, and Equations 1a and 1b) to compute transformed scan numbers and retention indexes of hydrocarbons in the "experimental" 6 and 8 °C/min runs. For these trials we assigned the three standards C_{12}, C_{18}, and C_{24} manually because of the large discrepancies in elution times compared to the 4 °C/min calibration. The results are presented in Figure 4b.

The curves (Figure 4b) are superimposable, indicating that the method has corrected for the considerable contraction in the carbon number vs. scan number scale (Figure 4a). A more accurate measure is the set of RRI's calculated for the hydrocarbons in the experimental runs, which are effectively unknowns. We present in Table I the average absolute error,

Table I. Average Absolute Error and Standard Deviations of RRI Measurements with Variation of GC Temperature Programming Rate, Based on a Four °C/min Calibration

GC programming rate, °C/min	Average absolute error, RRI units[a]	Std dev
6	4.5	4.9
8	4.4	3.4

[a] 100 times the value in methylene units.

Table II. Average Absolute Error and Standard Deviation of RRI Measurements with Variation of GC Starting Temperatures, Based on an Initial Temperature of 80 °C as the Calibration Run

GC initial temperature, °C	Average absolute error, RRI units	Std dev
60	3.1	2.6
100	21.2	36.

Table III. Average Absolute Error and Standard Deviation of RRI Measurements with Variation of GC Carrier Gas Flow Rate Based on 30 mL/min as a Calibration Run

| Carrier gas flow rate, mL/min	Average absolute error, RRI units	Std dev
25	1.5	1.1
35	1.6	0.5

in RRI units, and the standard deviation of the measurements from the expected values for each programming rate based on the experiment at 4 °C/min as the calibration. These results should be evaluated considering that the determination of component elution times by the CLEANUP program to the nearest spectral scan time leaves an uncertainty of a fraction ($^1/_3$ to $^1/_2$) of a scan. Under our experimental conditions, one scan represents approximately 0.03 methylene unit (3 RRI units, Figure 4) at a temperature programming rate of 4 °C/min.

We next performed a similar experiment, this time varying the starting temperature beginning with 80 °C (used as the calibration run), then using 60 and 100 °C starting temperatures as experimental data. Results are presented in Table II.

Finally, we evaluated the ability of the method to cope with variation in GC carrier gas flow. Because these changes cause less extreme contractions and expansions of the scan number, or RRI, scale than variations in temperature programming, rates, TIMSEK performs very well for flow rates of 25, 30, and 35 mL/min. Using 30 mL/min as a calibration run, results are summarized in Table III. The ability of Equations 1a and 1b to adjust for variations in carrier gas flow rate is reflected in Figures 4c and 4d.

Based on these data, the initial GC colummn temperature is the most critical parameter to control to ensure accurate RRI's and, fortunately, is the easiest to control precisely. The initial isothermal period (see Experimental) at higher initial temperature distorts the linearity of the RRI vs. scan number curve at low (approximately n-C_{12}) carbon numbers and is responsible for the large deviations noted at higher initial temperatures (Table II).

Method for Quantitation. Areas of gas chromatographic peaks are widely used for purposes of quantitation of ma-
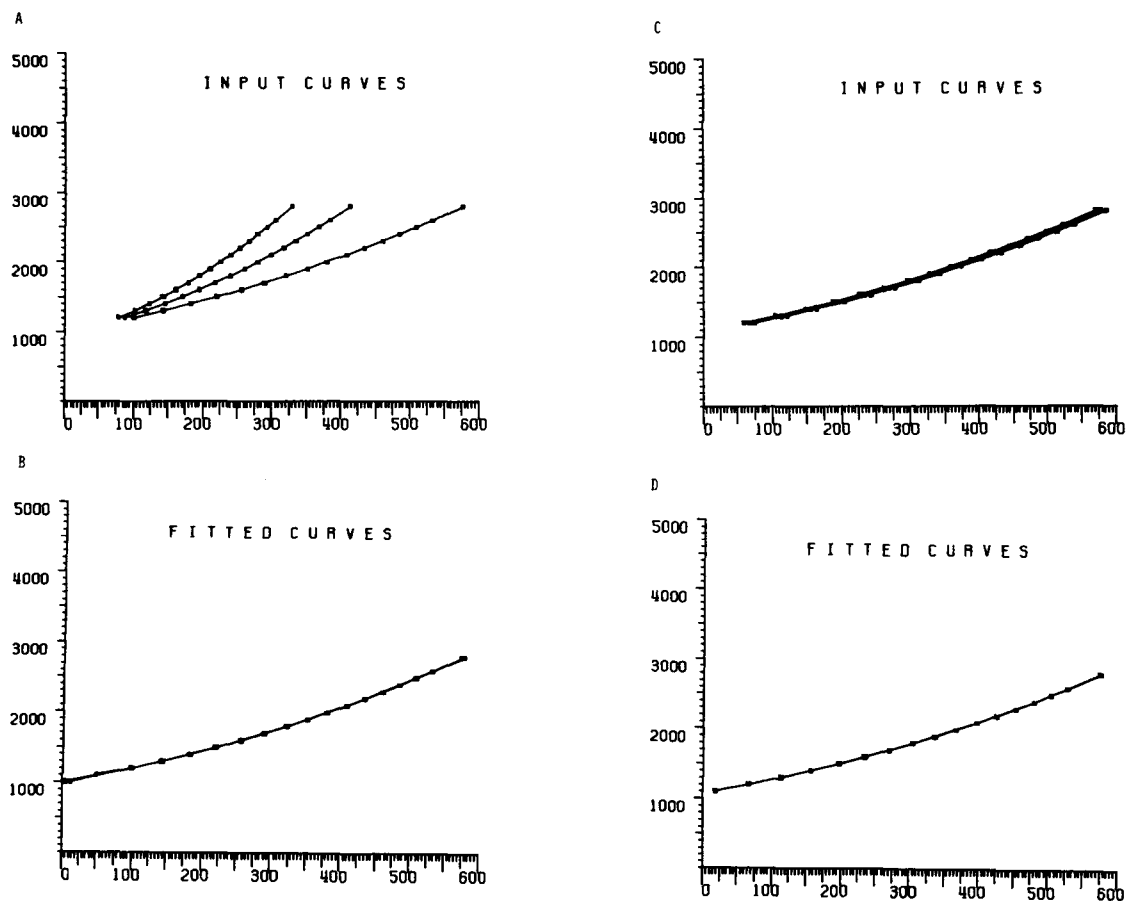
Figure 4. (a) Relative retention indexes vs. scan number for three hydrocarbon standard analyses (see text), at GC programming rates of 4, 6, and 8 °C/min. (b) Relative retention indexes vs. scan number for three analyses in Figure 4a, normalized to the 4 °C/min run using the linear transformation of Equation 1a. (c) Hydrocarbon standard analyses similar to those in Figure 4a, but with varying flow rates of 25, 30, and 35 mL/min. (d) Results of normalizing the runs in Figure 4c to the 30 mL/min run using the linear transformation as in Figure 4b

terials. When complex mixtures are analyzed by GC alone, however, there arise questions of homogeneity of GC peaks and identity of components among different analyses. RRI's and GC peak shapes are often insufficient to answer these questions, particularly when new components are observed in routine screening procedures. For these reasons GC/MS is now used extensively to analyze complex mixtures. Programs such as CLEANUP and library search techniques assist scientists in qualitative analysis of such mixtures. But much less progress has been made in obtaining quantitative results. Although workers in the petroleum industry have performed quantitative type analyses utilizing mass spectrometry for many years, such analyses depend on detailed knowledge of compound types present and careful calibration of the mass spectrometer with a suite of standards. These conditions are not met in most GC/MS analyses of mixtures.

We choose to use the areal total ion current of the internal standard(s) to compute relative concentrations of each component (Equations 2 and 3). These relative concentrations should be a better measure of the amount of material present than calculations based on single or selected ions (5) which are subject to greater statistical variation, given that the GC column resolution together with CLEANUP is able to remove other contributions to the spectrum. We stress that it is essential for accurate relative concentrations that the method chosen (e.g., CLEANUP) to remove background and overlapping components be able to apportion intensity of an ion common to overlapping components appropriately to the individual spectra (7), rather than assigning the ion to one spectrum or the other (6). Of course, measurement of relative concentrations provides a means for quantitative comparison of profiles but does not determine the actual amount of each

component. Auxiliary methods designed for quantitation of individual components, e.g., mass fragmentography or "selected ion monitoring" (16), can be used to establish relationships between relative concentrations and actual amounts of materials.

We evaluated the reproducibility of relative concentrations based on areal total ion current values by analyzing five GC/MS profiles of our mixture of hydrocarbon standards. Each profile was treated as an unknown. HISLIB was used to correlate and summarize the data. We determined relative concentrations in two ways. First, we employed n-C_{18} as the internal standard. Then we employed the average of the areal total ion currents for n-C_{12}, n-C_{18} and n-C_{24} as the basis for determining the relative concentration of each component. Results are summarized in Table IV.

The results in Table IV are a measure of the reproducibility of the data acquisition and analysis procedures. They indicate the variance to be expected using this method and a single internal standard. The results also indicate significant improvement in precision of relative concentrations when multiple standards are used to smooth out statistical fluctuations in the areal total ion current of a single internal standard. The deviations in relative concentration based on n-C_{18} alone average about 6.4% of relative concentration. Using n-$C_{12,18,24}$ together, the deviations are reduced to about 4.2%, consistent with the square root of three improvement to be expected a priori for three vs. one standard.

The results, however, do not measure variations in isolation or derivatization procedures or long-term variations in performance of the GC/MS system. Results presented below indicate variations observed in a complete analytical procedure in our laboratories. Other workers must evaluate their own

Table IV. Average Relative Concentrations and Standard Deviations for Five Analyses of a Mixture of n-Alkanes

Carbon No.	Relative retention index		Relative concentration based on $n\text{-}C_{18}$		Relative concentration based on $n\text{-}C_{12,18,24}$	
	Av.	S.D.	Av.	% S.D.	Av.	% S.D.
11	1101	1.3	30.4	7.6	13.4	5.2
12	1199	0.4	40.6	7.6	17.9	6.7
13	1300	0.6	48.0	6.2	21.2	3.8
14	1400	0.4	56.4	6.6	24.9	4.0
15	1499	1.2	62.9	6.2	27.8	4.0
16	1600	0.7	75.3	4.5	33.2	1.8
17	1698	1.1	80.3	3.0	35.5	1.4
18	1801	0.7	100.0	· · ·	44.2	3.4
19	1900	1.6	96.7	4.1	42.7	2.6
20	2000	1.3	102.0	5.3	45.0	2.9
21	2101	1.4	109.5	5.6	48.3	3.9
22	2198	0.0	109.0	5.7	48.1	3.1
23	2298	1.2	115.4	5.4	50.9	3.9
24	2399	0.7	85.8	7.9	37.8	5.0
25	2498	1.9	73.9	7.4	32.6	5.2
26	2596	2.7	75.7	7.3	33.4	4.5
28	2800	0.7	74.6	12.3	32.9	10.0

procedures similarly. One advantage of HISLIB is that such evaluations are greatly simplified.

Application Example—Comparison of Isolation Procedures. During a study of different isolation procedures for various organic fractions of human body fluids, we have used HISLIB as an aid to monitoring analytical procedures and intercomparing methods. We select as an illustrative example a comparison of two isolation procedures for the organic acid fraction of human urine, manual extraction and anion exchange (see Experimental). Data already exist in the literature for these isolation procedures, using GC methods for quantitation (15). To evaluate these procedures for both keto and hydroxy acids, we used aliquots of a 24-h urine sample of a patient previously diagnosed as having phenylketonuria (PKU). The patient was on a low phenylalanine diet at the time the urine was collected.

HISLIB was used to construct a library containing results from analysis of five aliquots of the above urine, using the manual extraction method. A representative total ion current plot for one of the analyses is shown in Figure 5a. The abundant phenylacetic, p-hydroxyphenylacetic, phenyllactic, and p-hydroxyphenyllactic acids (as TMS ethers/esters) (RRI's 1433, 1763, 1689, 1993, respectively) are notable characteristic compounds excreted in this disease; the abundance of phenylpyruvic acid is very low (in the baseline for the injected amount of the total mixture in this experiment) compared to the amount excreted prior to dietary control. The complete historical library is presented in Table V.

The reproducibility of relative concentrations is reduced relative to data presented in Table IV, because now all the variables of the isolation and derivatization procedure affect the results. However, the precision of our results is generally higher than that reported for similar analyses using a GC method for quantitation (15). One reason for the improved precision is, we feel, the addition of the quantitation standard at the beginning of the isolation procedure rather than just prior to derivatization (15). The only component with an inordinately large standard deviation is dioctylphthalate (RRI 2778). We attribute this artifact to sources other than the urine itself.

To the historical library in Table V, we compared data from an analysis of the same urine sample, but using anion exchange as the isolation method. Selected results of the comparison of a representative GC/MS profile (the total ion current plot is shown in Figure 5b) with the historical library are presented in Table VI.

As discussed previously (15), the two isolation procedures yield quite different GC/MS profiles (Figure 5). These differences are quantitated by the HISLIB output and can be quickly observed by scanning the "DISCREPANCY" column of Table VI (see footnote to the Table for explanation of terms). Some components, e.g., palmitic acid, are observed in nearly equal amounts in the two procedures. Other components, e.g., urea-diTMS and succinic acid-diTMS, are observed in significantly different quantities. There are

Table V. An Historical Library Containing Organic Acid Analyses of Five Aliquots of Urine from a PKU Patient under Dietary Control

Retention index	Std dev	No. Occ.	Rel Concn	% Std dev	Chemical name
1131	· · ·	1	9.2	· · ·	?
1200	0.0	5	97.8	4.9	C_{12} [a]
1359	0.9	5	46.9	17.7	Urea-diTMS
1366	· · ·	1	10.0	· · ·	Benzoic acid-TMS
1409	1.3	5	18.9	9.0	Succinic acid-di-TMS
1433	0.6	5	243.6	17.6	Phenylacetic acid-TMS
1560	· · ·	1	6.3	· · ·	Erythronic acid-tetra-TMS
1596	1.0	2	11.1	21.6	Threonic acid-tetra-TMS
1620	1.5	5	100.0	· · ·	m-Chlorophenylacetic acid-TMS [b]
1643	0.9	3	10.3	8.7	?
1664	1.7	3	13.9	15.8	?
1689	1.0	5	654.0	10.7	Phenyllactic acid-di-TMS
1763	1.0	5	78.7	10.4	p-Hydroxyphenylacetic acid-di-TMS
1798	1.0	5	343.6	7.7	C_{18} [a]
1847	1.4	5	8.1	16.0	Unknown phthalate
1887	1.1	4	17.9	16.2	Citric acid-tetra-TMS
1898	0.8	5	13.2	20.5	Unknown mixture
1991	1.2	5	76.0	3.0	p-Hydroxyphenyllactic acid-tri-TMS
2049	0.7	5	285.8	7.6	Unknown δ [c]
2093	1.0	5	39.1	28.1	Palmitic acid-TMS
2189	· · ·	1	4.8	· · ·	?
2293	0.7	5	27.3	12.8	Stearic acid-TMS
2401	0.7	5	516.0	5.2	C_{24} [a]
2778	3.8	5	83.2	86.5	Dioctylphthalate

[a] Internal RRI standard, added just prior to GC/MS analysis. [b] Internal quantitation standard, added to the urine prior to initial extraction. [c] Subsequently identified as N-acetylphenylalanine-TMS.
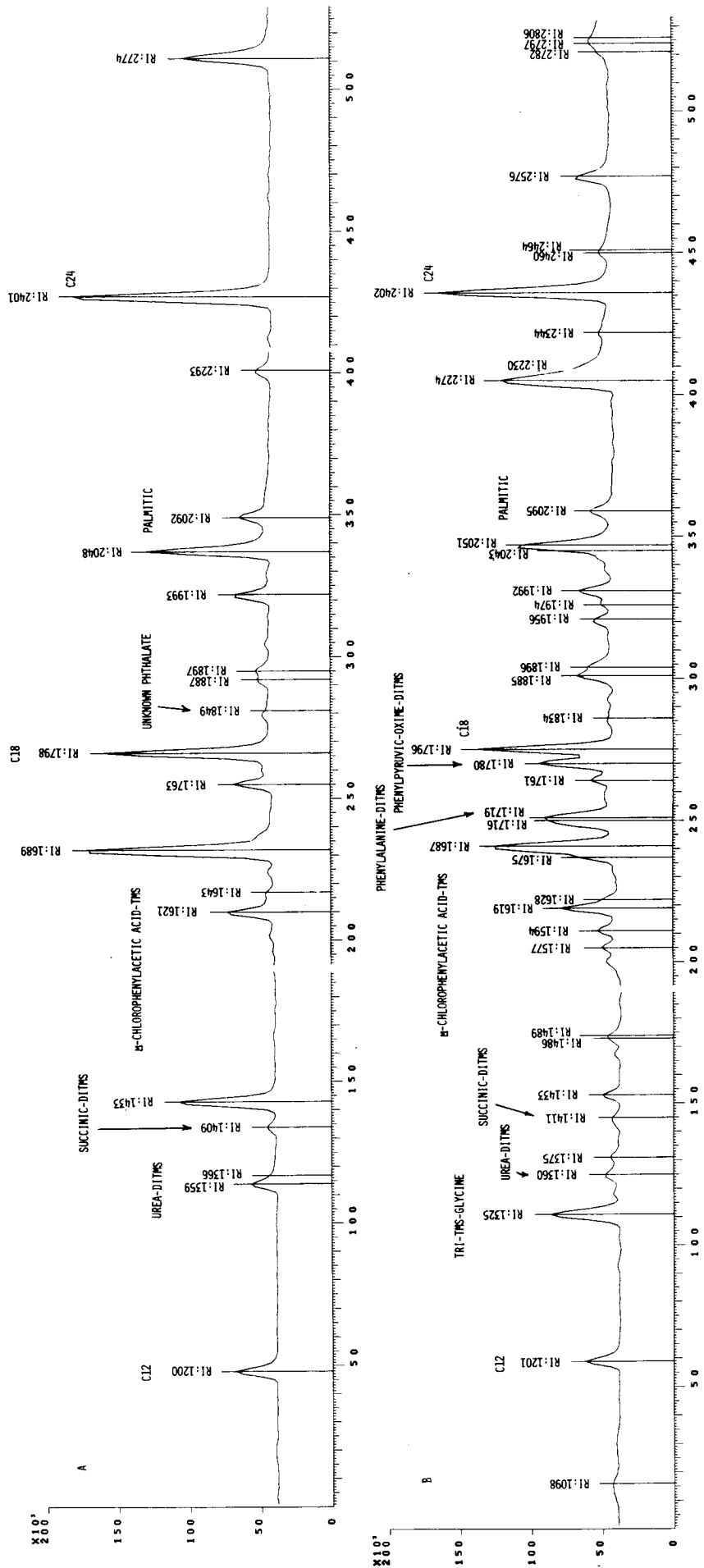
Figure 5. (a) Plot of total ion current vs. scan number for a representative GC/MS experiment analyzing a trimethylsilylated mixture of organic acids isolated by the manual method of extraction. (b) Total ion current plot of the same mixture of organic acids isolated by the anion-exchange method. *m*-Chlorophenylacetic acid–TMS is the internal quantitation standard. Minor discrepancies in RRI's between Figure 5a and Table VI are due to the fact that RRI's in Table VI are averages for five analyses

Table VI. Selected Results of Comparison of the GC/MS Profile of a Mixture of Urinary Organic Acids Isolated by Anion Exchange to the Historical Library of Table V. The Urine Sample was an Aliquot of the Same Urine Used to Construct the Library

Cmpd name[a]	RRI	Rel concn	HISLIB name	N[b]	RRI	Std dev	Rel concn	% Std dev	Discrepancy[c]
TRI-TMS-GLY	1325	157.5							NEW
UREA-DITMS	1360	18.3	UREA-DITMS	5	1359	0.9	46.9	17.7	-***
SUCCINIC	1411	9.3	SUCCINIC	5	1409	1.3	18.9	9.0	-******
PHENYLALA	1719	81.5							NEW
PHENYLPYRUV	1780	180.1							NEW
			?(phthalate)	5	1847	1.4	8.1	16.0	MISSING
STEARIC[d]	2095	34.9	PALMITIC	5	2093	1.0	39.1	28.1	-

[a] Names are truncated to conserve space. [b] N is the number of occurrences of the compound in the historical library. [c] A discrepancy of "NEW" indicates a component which was observed in the experimental GC/MS profile but is not present in the historical library. A discrepancy of "MISSING" indicates the reverse situation. The sign and number of asterisks indicate, for components present in both experiment and library, the sign and magnitude of the mismatch in relative concentration, one asterisk/standard deviation unit. [d] The best match to the library (10) is incorrect.

Table VII. Selected Results of Quantitative Comparison of a GC/MS Profile of Trimethylsilyl Derivatives of Urinary Organic Acids (72 h after Derivatization) with an Historical Library Composed of the Same Mixture Analyzed Repetitively at Earlier Times[a]

Cmpd name	RRI	Rel concn	HISLIB name	N	RRI	Std dev	Rel concn	% Std dev	Discrepancy
(b)	1323	162.3	2-AMINOETHANOL	5	1326	0.9	161.9	2.2	
GLYCERIC	1370	11.7	GLYCERIC	4	1376	0.9	12.6	23.0	-
PHENYLACETIC	1430	32.2	PHENYLACETIC	5	1433	0.6	33.6	6.5	-*
THREONIC	1592	30.8	THREONIC	5	1595	1.0	33.9	10.6	-*
PHENYLPYRUV	1781	172.2	PHENYLPYRUV	5	1781	1.1	179.5	16.4	-
4-OH-PHENYLL	1990	58.9	4-OH-PHENYLLAC	5	1992	0.8	52.1	7.3	**
PALMITIC	2093	55.7	PALMITIC	5	2095	0.7	34.8	9.2	*******

[a] See text for description. [b] Matched poorly to master library, but scored well (835) against average spectrum in historical library.

components missing in the new experiment, e.g., the unknown phthalate at RRI 1847 (Tables V and VI). There are new components, including, of course, phenylpyruvic acid–oxime–TMS (RRI 1780) because the manual extraction procedure did not include formation of oximes, and two amino acids, glycine and phenylalanine as the TMS derivatives (RRI 1325, 1719) which proved to be artifactual on a repeat analysis.

Application Example—Time Stability of Derivatives. We have also used HISLIB to monitor the long term stability of the trimethylsilyl derivatives of organic acid fractions, isolated by ion exchange, of human urine. Five samples were analyzed (from the same patient as above) 1, 2, 4, 8, and 24 h after derivative formation. The GC/MS profiles resembled each other closely enough to indicate decomposition was minimal after 24 h. After 72 h, a sixth analysis was made of the same mixture and this new GC/MS profile compared to an historical library composed of the first five experiments. We present in Table VII selected results from this comparison.

The GC/MS profile of the sixth experiment remained very similar to the previous profiles. This is established quantitatively by the strong similarity of the profiles reflected in the data of Table VII and summarized in the discrepancy column of the Table. We have no explanation for the observation of significantly greater amounts of palmitic acid–TMS and also succinic acid–diTMS; all other components compared very favorably.

Other Applications of HISLIB. From the preceding discussions, several other applications of HISLIB are suggested. We have presented examples of the use of HISLIB to check on the reproducibility of instrumentation and analytical procedures utilized to study complex mixtures. Clearly, the same technique can be used to explore other variables in an analytical scheme. HISLIB should facilitate detailed intercomparisons of complex mixtures, for example those encountered in diagnostic medicine where enhancements of GC/MS techniques are desirable (17).

Because the historical library can be updated at will, it is easy to maintain a long-term history of analysis of a particular type of mixture. Maintenance of several such libraries for different types of mixtures is a simple task. In fact, different historical libraries can be compared with one another, opening the possibility for comparison of results among laboratories engaged in similar research.

HISLIB averages spectra of the same compound. Thus, statistical variations in ion abundances are reduced as additional examples are encountered. The resulting average spectrum is frequently of much higher quality than a single spectrum in existing libraries. We have implemented a mechanism for adding averaged spectra to or replacing spectra in our primary library. This provides a mechanism for gradual improvement of spectral libraries with time. In addition, RRI's are included with the spectra, enabling us to improve the certainty with which subsequent spectra are matched to the primary library.

The method of comparing new profiles to an existing historical library quickly focuses attention on known materials present in abnormal quantities and on new components. The latter become subjects for more sophisticated structure elucidation procedures (18) which can now use the (high quality) mass spectral data directly to assist in solving the structures of unknowns (19).

Limitations. There are several limitations to the procedure which must be mentioned. We have not yet thoroughly investigated variations in relative concentrations with instrument operating parameters. The performance of any mass spectrometer may change as a function of time. Any change in performance which affects the ionization of the internal standard(s) relative to other mixture components will affect results of quantitation. This can be avoided in part by using several different standards in each run.

In the present implementation of the program, there are several deficiencies in the data analysis scheme. We have not

included a procedure for easily deleting selected old experiments from an historical library—the library must currently be recreated excluding undesired experiments. Also the spectrum averaging scheme makes no decisions about including ions of low abundance—all are included. Ions which occur infrequently are diminished in importance as additional spectra are averaged, but they are not rejected because we have not yet developed adequate heuristics for removing such ions.

ACKNOWLEDGMENT

We thank Edwin Blaisdell, Mark Stefik, Wilfred Pereira, and Alan Duffield for their contributions to initial discussions of the problem addressed by HISLIB.

LITERATURE CITED

(1) R. G. Ridley, in "Biochemical Applications of Mass Spectrometry", G. R. Waller, Ed., John Wiley and Sons, New York, N.Y., 1972, p 177.
(2) F. W. Karasek and J. Michnowicz, *Res./Dev.*, **38**, May 1976.
(3) H. Nau and K. Biemann, *Anal. Lett.*, **6**, 1071 (1973).
(4) H. Nau and K. Biemann, *Anal. Chem.*, **46**, 426 (1974).
(5) C. C. Sweeley, N. D. Young, J. F. Holland, and S. C. Gates, *J. Chromatogr.*, **99**, 507 (1974).
(6) J. E. Biller and K. Biemann, *Anal. Lett.*, **7**, 515 (1974).
(7) R. G. Dromey, M. J. Stefik, T. C. Rindfleisch, and A. M. Duffield, *Anal. Chem.*, **48**, 1368 (1976).
(8) E. Jellum, P. Helland, L. Eldjarn, U. Markwardt, and J. Marhofer, *J. Chromatogr.*, **112**, 573 (1975).
(9) B. E. Blaisdell, *Anal. Chem.*, **49**, 180 (1977).
(10) W. E. Reynolds, in "Biochemical Applications of Mass Spectrometry", G. R. Waller, Ed., John Wiley and Sons, New York, N.Y., 1972, p 109.
(11) S. P. Markey, W. G. Urban, and S. P. Levine, "Mass Spectra of Compounds of Biological Interest", *U.S. At. Energy Comm. Rep.*, **No. TID-26553**, National Technical Information Service, U.S. Dept. of Commerce, Springfield, Va. 22161.
(12) H. S. Hertz, R. A. Hites, and K. Biemann, *Anal. Chem.*, **43**, 681 (1971).
(13) See for example, A. N. Kolmogorov and S. V. Fomin, "Elements of the Theory of Functions and Functional Analysis, Volume 1: Metric and Normed Spaces", Graylock Press, Rochester, N.Y., 1957, p 16.
(14) E. W. Dijkstra, *Numerische Math.*, **1**, 269 (1959).
(15) J. A. Thompson and S. P. Markey, *Anal. Chem.*, **47**, 1313 (1975).
(16) C. G. Hammer, B. Holmstedt, and R. Ryhage, *Anal. Biochem.*, **25**, 532 (1968).
(17) O. Stokke, *Biomed. Mass Spectrom.*, **3**, 97 (1976).
(18) R. E. Carhart, D. H. Smith, H. Brown, and C. Djerassi, *J. Am. Chem. Soc.*, **97**, 5755 (1975).
(19) D. H. Smith and R. E. Carhart, in "Chemical Applications of High Performance Spectrometry", M. L. Gross, Ed., Proceedings in press.