Journal of Chromatography, 379 (1986) 3—26
Biomedical Applications
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMBIO. 2993

REVIEW

ADVANCED INSTRUMENTATION AND STRATEGIES FOR METABOLIC PROFILING

JOHN F. HOLLAND

Department of Biochemistry, Michigan State University, East Lansing, MI 48824 (U.S.A.)

JOHN J. LEARY

Department of Chemistry, Michigan State University, East Lansing, MI 48824 (U.S.A.)

and

CHARLES C. SWEELEY*

Department of Biochemistry, Michigan State University, East Lansing, MI 48824 (U.S.A.)

(Received January 2nd, 1986)

CONTENTS

0378-4347/86/$03.50

1. INTRODUCTION

Two-dimensional graphic representations of information about the metabolic status of an individual were first developed by Williams [1]. By employing paper chromatography, taste thresholds, and other analyses, several related physiological compounds were analyzed for each subject under study. Williams [1] was able to graphically demonstrate the results for each individual by utilizing vector angles to represent the various analytes and vector distances to represent the analyte concentrations. As shown in Fig. 1, these graphs revealed an individuality of the pattern for each subject. A genetic basis for this individuality was inferred when differences in the patterns normally observed from one person to another nearly disappeared when identical twins were analyzed, as shown in Fig. 2. These patterns led to the concept of a unique metabolic state that may be useful as a sort of fingerprint in revealing individual metabolic idiosyncrasies and evaluating conditions of health.

The term metabolic profile was coined by Horning and Horning [2] and applied to the graphic results from gas chromatographic (GC) analyses of specific types of metabolites from various body fluids. They established the methodology for several different classes of metabolites, including the derivati-

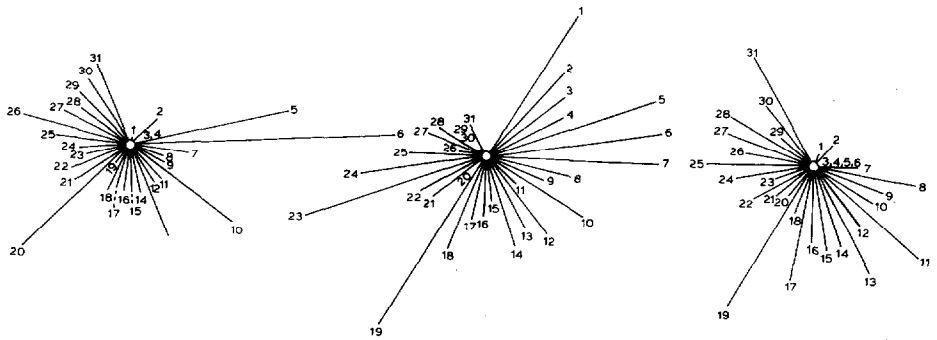


Fig. 1. The metabolic patterns of three unrelated subjects. Vectors 1–17 represent taste thresholds, and vectors 18—31 indicate the concentration of urinary metabolites, measured by paper chromatography. Vector identification: 1, creatinine; 2, sucrose; 3, potassium chloride; 4, sodium chloride; 5, hydrochloric acid; 6, uric acid; 7, glucose; 8, leucine; 9, valine; 10, citrulline; 11, alanine; 12, lysine; 13, taurine; 14, glycine; 15, serine; 16, glutamic acid; 17, aspartic acid; 18, citrate; 19 and 20, unidentified; 21, gonadotropic hormone; 22, pH; 23, pigment/creatinine; 24, chloride/creatinine; 25, hippuric acid/creatinine; 26, creatinine; 27, taurine; 28, glycine; 29, serine; 30, citrulline; 31, alanine. In the vector representation of taste threshold and urinary metabolites identical compounds may be represented by two different numbers. (Taken from ref. 1 with the permission of author and publisher.)

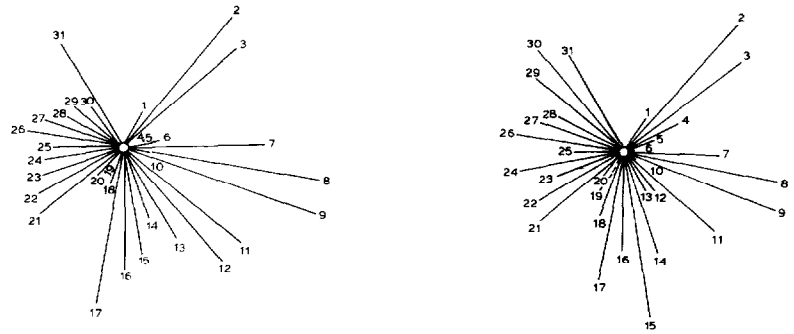


Fig. 2. The metabolic patterns of urinary constituents of a pair of identical twins. For vector identification see legend to Fig. 1. (Taken from ref. 1 with the permission of the author and publisher.)

zation techniques necessary to volatilize and analyze a great number of compounds from a single sample by means of a single GC analysis. Metabolic profiles were defined to be multi-component GC analyses that produced patterns for a group of metabolically or analytically related metabolites.

The term profile was used to indicate that each individual would have a specific GC pattern or profile of these compounds. The term itself means a two-dimensional representation of a three-dimensional object, in effect a cross section. By direct analogy, a metabolic profile is a two-dimensional cross section of a complex multi-dimensional physiological state delineated by the sample source, chemical processing, and method of analysis. Although defined by Horning and Horning [2] to be a gas chromatogram, in reality a profile may be generated by any one of a number of instrumental techniques. A general profile is shown in Fig. 3. Although immediate changes in an individual's profile can be assumed to indicate changes in the state of that person's health, it is possible that, over time, even the presence of insipient disease might be revealed by subtle changes in the metabolic profile.

A distinction must be made between screening and profiling analyses. In screening, or multiple analyte analysis, several compounds are measured, sometimes simultaneously, but in all situations the amounts of these analytes are compared to the reference ranges in which they are normally found. As an illustration, clinical laboratories now employ automated analyzers that provide

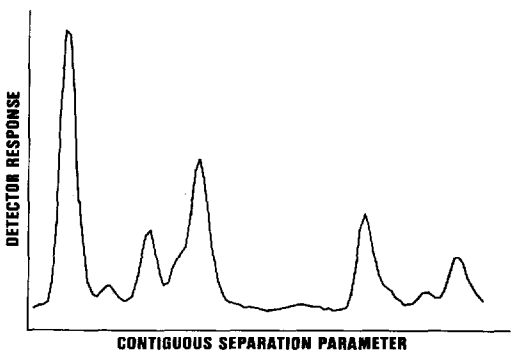


Fig. 3. An example of a metabolic profile in which all the detected components contribute to the final appearance.

quantitative results for up to 24 or more different components of serum or urine. The concentration of each analyte is to be considered as a single variable and used in the diagnosis of disease or the evaluation of the state-of-health based on the clinically-determined relevance of that particular analyte. In the metabolic profile, however, the absolute values of particular analytes are not the only major concern, but rather, the relative relationship of these analytes to one another is an equally important consideration. Another major difference between typical multi-component analysis and metabolic profiling is that in the profile, the various analytes are usually metabolically or chemically related. The metabolic profile generally arises from a single analytical procedure, which imposes limitations on the chemical types of analytes that can be determined in a given analysis. For example, profiles may be obtained for steroids, organic acids, prostaglandins, peptides, etc., but not usually mixtures of these different kinds of analytes. The third major difference between multiple analyte analysis and profiling lies in the fact that the profile utilizes information from unknowns as well as knowns. Lastly the profile gathers information on all of the substances present in the sample without need for prior target compound designation. It follows that the profile is more uniquely able to serve in distinguishing metabolic differences as the results of analyses are studied.

The original concept of metabolic patterns emphasized a characteristic and somewhat reproducible profile for individual subjects. That is, the knowledge of a person's unique profile could be used in evaluating the analysis of subsequent samples from the same individual. Changes in the profile hopefully could be related to changes in the state of the person's health. This approach to profiling has not been widely employed because of its prohibitive cost in resources, the perturbance of the profile by diet and drugs, and difficulty in maintaining continuing availablility of the subjects under study over a prolonged time.

An extension to the meaning of the term metabolic profile in most applications of the technique has been to define the characteristics of a reference set against which individual profiles can be compared. Here, rather than observing changes of an individual profile as a function of time, the individual profile is compared to the accumulated profile of a designated reference set. Adult diabetics, elderly schizophrenics, and children with neuroblastoma, are examples of reference sets, as are healthy adult males, healthy institutionalized children, etc. By this comparison, it can be determined whether or not the specific subject under study is a member of the set being referenced. These comparisons can work in both directions. For instance, it may be of interest to determine whether a particular subject has a profile which would statistically place it in a particular reference set. Conversely, a particular subject profile may be determined to be significantly different from the reference set being compared.

Although the basic objective of metabolic profiling is to obtain a definitive pattern, in reality, the profile generates quantitative analyses of many compounds. Quite often an outstanding variation of a particular component against the normal of the reference set can lead to the differential diagnosis of a specific health problem. This type of diagnostic assistance is a bonus of profiling analyses. Because of the ease of obtaining such information, profiling of body fluids has become a major contributor to the detection of specific

TABLE 1

ADVANTAGES AND DISADVANTAGES OF METABOLIC PROFILING

Advantages

(1) All of the information generated by multiple analyte analyses is available for evaluation of each individual component in the mixture.
(2) Relationships between components are revealed as well as quantitative amounts.
(3) Unknown as well as known components are used in establishing patterns.
(4) All the components present are quantitated in each analysis. Hence, there is no need for prior identification of target compounds.

Disadvantages

(1) Only one method of analysis can be employed in generating a single profile. Thus, the scope of the analysis will be both instrumental method- and compound-type-specific.
(2) Sample preparation and analysis of non-differentiating multiple analytes can be difficult.
(3) The instrumentation required is often very expensive to buy and maintain and the analyses can be time- and resource-consuming.

metabolic human disorders. This favorable byproduct should be differentiated, however, from the long-term objective of metabolic profiling, which is to characterize the state of a person's health by observation of the relationships observed in the profile. The general advantages and disadvantages inherent in profiling types of analysis are listed in Table 1.

2. GENERAL REQUIREMENTS FOR PROFILING

One of the major analytical requirements for profiling is that the instrumental and chemical methods chosen must be appropriate for the analysis of as many as possible of the very large number of components of a certain class in the biological sample. Because of their numbers, these components must be analyzed with a high degree of selectivity. The method must incorporate some means of separating and individually identifying the various components. Another requirement is that the various analytes must be measured with a detection system that does not introduce a severe bias to the quantitation of the different compounds. Further, there is an extremely large dynamic range with which the measurements must be made. Metabolites and compounds found in biological mixtures such as blood or urine vary widely in their concentrations. One must be able to detect and accurately measure one component of a mixture that may be 5000 to 10 000 times more dilute than another component of the same mixture. Both must be measured within the limits of clinical accuracy and from a single preparation of the sample.

The methodology, both instrumental and chemical, becomes the major determining factor in the applicability and the success of various profiling investigations. The following sections will illustrate various kinds of profiling by selected examples. The major divisions for this section have been made on the basis of the instrumental techniques employed. In general, subject areas chosen were limited to those applications where pattern recognition is either used directly or made available by the technique. Additionally, the analysis must meet the criteria of using a single method for all of the compounds of a related type and must be analyzed in a single run by that method.

3. PROFILING BY GAS CHROMATOGRAPHY

GC was the method with which Horning and Horning [2] developed and defined the concept of metabolic profiling. There are several instrumental reasons encouraging this choice. With the advent of flame ionization detection (FID), GC became an analytical procedure that provided the largest dynamic range of physical measurement ever available in a wide-spread instrument system. This is exceedingly important, since variations in the concentration of metabolites encompass several orders of magnitude. Six-to-seven orders of linear attenuation are typically available by FID, providing an opportunity to more effectively meet the needs of profiling. In addition, the response of the flame ionization detector is relatively constant for similar chemical species encountered in a profiling sample. This feature allows for an acceptable quantitative accuracy in multi-component analysis. The wide choice of stationary phases and column types and dimensions, along with variable operating parameters such as temperature and programming rate, and carrier gas type and flow-rate allow an optimization of the capacity of the method to separate the various components of the complex mixtures, greatly aiding in the qualitative aspects of the analysis. Lastly, and by no means of the least significance, the common practice of utilizing a strip chart recorder for output of the time versus intensity curve (gas chromatogram) provided a convenient visual representation of all of the components and their relative abundances. The gas chromatogram became the metabolic profile. In their original work, Horning and Horning [2] utilized packed column GC to obtain individual profiles, whilst GC with mass spectrometry (MS) was used as the arbitrator for qualitative analyses.

3.1. Typical areas of applications

Two typical examples of profiling by GC are briefly discussed here to illustrate the applicability and complexity of the problems one may meet in using the profiling approach. For more detailed information particularly on steroids the reader is referred to the specialized chapter of this volume (see pp. 91—156).

Important differences in the *steroid profiles* of adults, newborns, children, and pregnant women have been observed. Evidence of clinical disorders such as Cushing's disease, Addison's disease, hyperaldosteronism, and deficiencies in steroid enzymatic pathways can be detected in steroid profiles by noting the elevation or diminution of particular steroid metabolites [3].

Steroid profiling by GC has been performed on numerous complex biological matrices including urine, plasma, saliva, cerebral spinal fluid, amniotic fluid, and various tissues. The general sample preparation scheme for these compounds includes extraction and isolation of steroids from the physiological matrix, separation of free steroids from steroid conjugates, and derivatization. Sample purification is the rate-limiting step in the production of steroid profiles.

Urinary steroids are excreted in three forms: free steroids, glucuronide conjugates, and sulfate conjugates. Both free and conjugated steroids must be

extracted in uniformly high yields and with minimal structural-based discrimination. These important criteria increase the complexity of the isolation technique. Normally, the conjugates are cleaved by solvolysis or enzyme hydrolysis to yield a sample containing only free steroids. Alternately, steroid conjugates may be analyzed intact using techniques such as fast atom bombardment (FAB)-MS. The conjugates may also be derivatized and analyzed by GC. Extremely thermostable capillary columns must be employed for the separation of these derivatized conjugates.

Derivatives of free steroids are prepared in order to promote volatility and to protect against thermal degradation during GC analysis. Methyloxime–trimethylsilyl derivatives are usually preferred. Methoxylamine hydrochloride is first used to protect the keto groups from enolization during silylation by forming methyloximes. Steroid hydroxyl functions then undergo trimethylsilylation. The derivatized steroids may then be chromatographed.

Urinary *organic acids*, the end-products of the intermediary metabolism of carbohydrates, amino acids, lipids, and nucleotides represent another illustrative example of GC profiling. They are a very complex mixture of polar and non-polar substances that range in molecular weight from 60 (acetic acid) up to about 300 (stearic acid and sialic acid). The abnormally high excretion of particular organic acids has been correlated with genetic disorders called organic acidurias, and it is now possible to utilize the metabolic profile of organic acids to assist in the differential diagnosis of many inherited metabolic disorders [4]. A typical profile of urinary organic acids, separated on a 3.65 m × 4 mm glass column packed with 5% SE-30, is shown in Fig. 4. A mixture of methyl esters and O-trimethylsilyl methyl esters was used in this analysis [2].

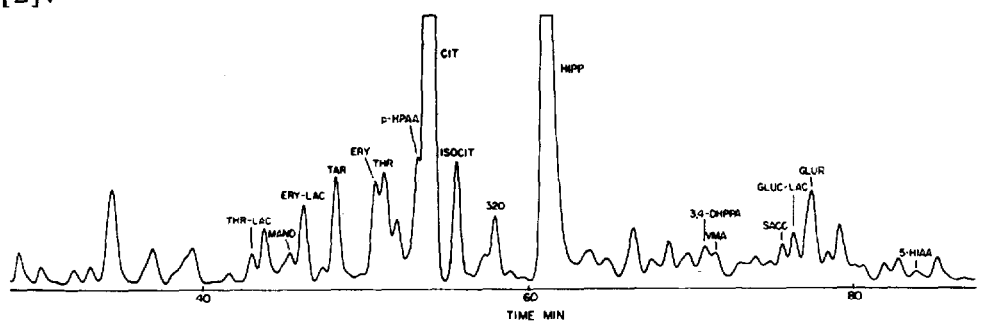


Fig. 4. Metabolic profile of human urinary acids (24-h collection, female subject) separated by packed column GC. Peaks: THR-LAC = threonic acid lactone; MAND = mandelic acid; ERY-LAC = erythreonic acid lactone; TAR = tartaric acid; ERY = erythreonic acid; THR = threonic acid; p-HPAA = *p*-hydroxyphenylacetic acid; CIT = citric acid; ISOCIT = isocitric acid; 320 = unknown metabolite (mol. wt. 320); HIPP = hippuric acid; 3,4-DHPPA = 3,4-dihydroxyphenylpropionic acid; VMA = vanilmandelic acid; SACC = saccharic acid; GLUC-LAC = gluconic acid lactone; GLUR = glucuronic acid; 5-HIAA = 5-hydroxyindole acetic acid. (Taken from ref. 2 with permission of the authors and publisher.)

4. PROFILING BY GAS CHROMATOGRAPHY—DATA SYSTEM

The development of bonded-phase fused silica GC columns has significantly increased the stability and reproducibility of capillary GC and greatly reduced column bleed and other detrimental variables. The excellent resolution

of these capillary columns provides an opportunity to resolve the majority of the components of complex mixtures in the time domain. This capacity is encouraging the development of GC-only profiling. This technique requires the use of a computer system which, for maximum convenience and utility, should be on-line and dedicated. The advantages of profiling by GC—data system (DS) lie in the relatively low cost of the instrumentation required when compared to other profiling techniques, and a wide spectrum of applications. Since the methodology utilized by this approach historically resides in the art of GC—MS—DS, concern for the absence of a mass axis is evidenced in the original applications. This can be remedied by using a sequential protocol with all samples being run by GC—DS and samples statistically aberrant or with identification ambiguities being re-run by GC—MS—DS. In fact, the availability of megabore column technology supports this new dual approach by making possible a congruence in the order of elution between both methods. This is essential for accurate, effective cross-referencing between the data files for the GC—DS and GC—MS—DS runs. Otherwise, it would be very difficult to match the mass spectra with corresponding peaks on the GC profile.

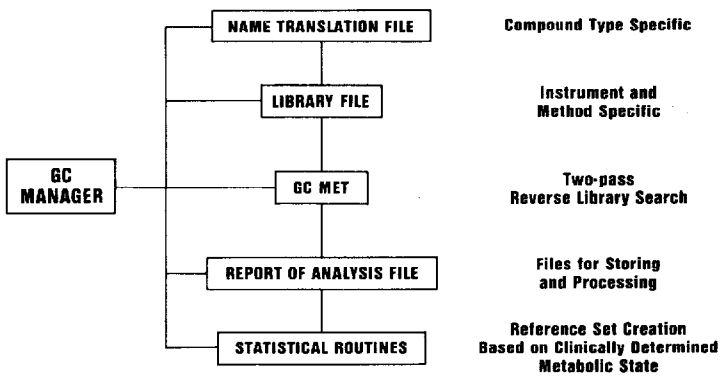


Fig. 5. Flow-chart showing GC MET routines.

In an effort to maximize the information obtained by GC-only profiling, a computer system called GC MET (GC Metabolites) has been developed that attaches directly to the output of the integrator of a commercial gas chromatograph. At the end of each run, the data are transferred directly to an IBM-PC for storage and analysis. An integrated software system utilizing a reverse library search has been written for profiling analysis, as diagrammed in Fig. 5. Since, for any profiling analysis, a single type of body tissue or fluid is employed, the composition of the samples will be relatively similar. There is a fixed number of components that can be present and for the most part they must exist in relatively reproducible relationships because the samples originate from living, metabolically-functioning humans. This permits application of a closed set principle with the actual members of the set being used as retention index standards. This removes the need for co-injected standards that can obscure sample peaks in the region of their appearance, and often are chemically different from the members of the set, and thereby may introduce anomalous elution behavior.

Maximum retention index accuracy can be obtained by using selected members of the set as standards. This approach necessitates two passes through

the data file during computer analysis of profiling data. The first pass locates those members of the set that are the most recognizable in that sample. These compounds are then used as time standards for the conversion of retention time to retention index for all components during the second pass. Automatic diagnostic output permits instant operator evaluation of the quality of the identification assignments.

All peaks found in the spectrum are identified as being in the library (knowns and unknowns) or not in the library. Prominent peaks not in the library are highlighted in the output routine for further consideration. Although the retention index is the major analytical determinant, other factors such as resolution of the peak, peak area, relativity to nearest neighbors, and other pattern factors are also considered. The following equation documents the calculation of the confidence factor (CF) which is used to evaluate the accuracy of the component identification:

$$CF = 100 \cdot \frac{10 - |\Delta^{RI}|}{10} \cdot f_r \cdot f_p \cdot f_c$$

where

RI = library retention index;
f_r = resolution factor, determined by the method used for area calculation;
f_p = pattern factor, determined by known linkages to other metabolites;
f_c = concentration factor, determined by comparison of measured area to expected area (this factor is only used in the first pass confidence calculation, not the final calculation).

The output of this routine is a "result of an analysis file" that contains compound number, amount, and confidence factor. These files are submitted to statistical routines that create reference sets and compare single files to designated reference sets or one set to another set. Statistical variations are detected by this means as are conformity of profiles to specific sets.

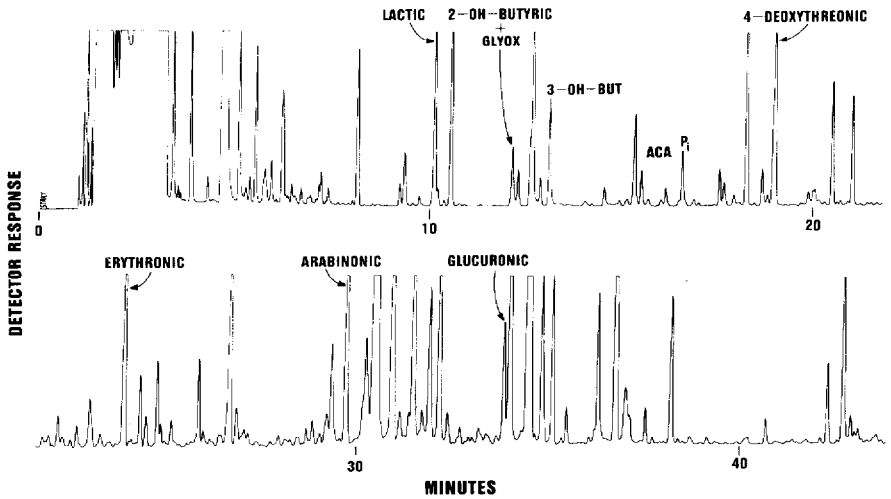


Fig. 6. Urinary organic acid profile generated on a 25-m Ultra II capillary column. Temperature programme, 60–275°C at 5°C/min; starting flow-rate of helium, 42.5 cm/s. See also Table 2.

TABLE 2

TABULAR REPORT OF URINARY ORGANIC ACIDS IDENTIFIED BY GC—DS PROFILING

This sample also contained 120 unknown metabolites, a few of which are artifacts of the chemical isolation procedure and derivatization reagents.

Metabolite	Concentration (μg/mg creatinine)
Phenol	2.9
Lactic	48.0
2-Hydroxyisobutyric	3.9
Glycolic	33.0
Glyoxylic oxime, 2-hydroxybutyric	24.0
Oxalic	14.0
p-Cresol	110.0
Pyruvic oxime	9.0
3-Hydroxybutyric, sulfuric	45.0
2-Hydroxyisovaleric	0.83
2-Ketobutyric oxime	0.69
3-Hydroxyisovaleric	16.0
2-Ketoisovaleric oxime (I)	0.93
2-Ketoisovaleric oxime (II)	0.75
Urea, 2-ketovaleric oxime, 2-hydroxy-isocaproic, acetoacetic oxime (I)	31.0
Benzoic	36.0
2-Ketocaproic oxime	0.43
Caprylic	2.7
Acetoacetic oxime (II)	7.6
Phosphoric	30.0
2-Keto-3-methylvaleric oxime, cyclohexylacetic, 2-ketovaleric oxime	4.2
Phenylacetic	1.4
d-Glyceric, maleic	0.92
Succinic	7.1
Methylsuccinic	4.60
l-Glyceric	9.40
Fumaric	3.4
4-Deoxyerythronic	17.0
4-Deoxythreonic	83.0
Glutaric	7.0
Hydrocinnamic, 3-hydroxypyruvic oxime	1.4
3-Deoxytetronic	8.8
3-Methylglutaric	0.81
2-Deoxytetronic	23.0
o-Methoxybenzoic	0.6
Citramalic	3.3
Malic	8.1
2-Ketooctanoic oxime	1.2
Adipic	3.6
Salicylic	37.0
Pyroglutamic	24.0
Methyladipic	7.9
4-Methoxyphenylacetic	5.2
Hydroxymethylfuroic	11.0
Erythronic	71.0
m-Hydroxybenzoic	13.0
Threonic	18.0

TABLE 2 (continued)

Metabolite	Concentration (μg/mg creatinine)
2-Hydroxyglutaric, phthalic	18.0
3-Phenyllactic	1.8
Pimelic	6.9
3-Hydroxy-3-methylglutaric, 3-hydroxyphenylacetic	6.0
4-Hydroxybenzoic, glutamic	2.5
2-Ketoglutaric oxime, oxalosuccinic oxime	32.0
4-Hydroxyphenylacetic	12.0
Lauric, phenylpyruvic, furandicarboxylic	9.3
2-Deoxyribonic	51.0
3,4-Dimethoxyphenylacetic	15.0
Suberic	2.0
Homoveratric	2.3
m-Hydroxyphenylpropionic	14.0
Quinolinic	4.5
Aconitic, orotic	1.4
3-Hydroxyanthranilic	16.0
Vanillic	6.6
Homovanillic	7.5
Gentisic, 3-methoxycinnamic	16.0
Terephthalic, 4-hydroxymandelic, ribonic	18.0
Arabinonic	59.0
Protocatechuic, *p*-methoxycinnamic	48.0
Hippuric (I), citric, isocitric, 3,4-dihydroxyphenylacetic	440.0
m-Hydroxyphenylhydracrylic	160.0
3-Deoxyribohexonic	15.0
2-Ketogluconic oxime, dehydroascorbic	120.0
Vanilmandelic, sebacic, 3-deoxyhexonic	5.1
Gluconolactone, galactonolactone (I)	81.0
Glucose (I)	160.0
Galacturonic oxime, *p*-coumaric	2.4
Galactonolactone (II)	4.0
Ascorbic, indoleacetic, gallic	23.0
Glucuronic	50.0
Glucose (II)	230.0
Gluconic	140.0
Kynurenic, salicyluric	480.0
Galactaric	17.0
1,12-Dodecanedioic	0.42
Uric	240.0
Stearic	1.7
N-Acetylneuraminic	24.0

A typical profile of urinary organic acids, separated as the per-trimethylsilyl (oxime) derivatives, is shown in Fig. 6. The separation was carried out on a 25-m Ultra II column (Hewlett-Packard), programmed from 60 to 275°C at 5°C/min and with a starting linear velocity of helium of 42.5 cm/s. Table 2 summarizes the information obtained from this profile by computer-assisted identification of components by a library search algorithm (described above). The library for this application contains over 300 compounds. Only the positively identified components are included in the summary; information about unknown substances is reported by the computer in a separate tabular output.

Profiling by the dual approach is utilized by Meridian Labs. in their profiling services. An IBM-PC attached to a Hewlett-Packard Model 5890A gas chromatograph with a printing integrator (HP 3392A) provides the initial GC—DS analysis. Samples with documented abnormalities are then run by GC—MS—DS with a compatable megabore GC column.

5. PROFILING BY GAS CHROMATOGRAPHY—MASS SPECTROMETRY

The attachment of the gas chromatograph to the mass spectrometer occurred in the late 1950's and early 1960's. This hybrid instrument can be envisioned in two ways. The gas chromatograph can be considered as an exotic inlet for the mass spectrometer, thereby greatly expanding the versatility and applications of the mass spectrometer into many different areas of organic chemistry and biochemistry. On the other hand, the mass spectrometer may be considered an exotic detector for the gas chromatograph, producing a multidimensional detecting system for GC. The latter sense best illustrates its role in profiling. Three-dimensional data files are readily obtained by this instrument system involving chromatographic time, ion intensity, and ion mass. Full utility of GC—MS was not realized, however, until the attachment of the minicomputer data system to the mass spectrometer creating the technique of GC—MS—DS. This instrumentation rapidly became the most outstanding and versatile analytical system available to the scientist. Quite naturally, an analysis protocol as exacting as that for profiling would gravitate toward GC—MS—DS and a great deal of the activity in the area of profiling has centered on this instrumentation for the past fifteen years.

Horning and Horning [2] employed GC—MS—DS with chemical ionization (CI) to obtain a profile of neutral drug metabolites as early as 1971. However, in this application the mass axis was not directly utilized, but rather the total ion current was summed for each scan to provide the ordinate for the profile. In this mode the mass spectrometer serves in an analogous manner to the flame ionization detector. Later approaches exploit the mass information as well as the time domain. Today GC—MS—DS remains the dominant instrumental vehicle utilized in metabolic profiling research throughout the world. For the most part, magnetic sector mass spectrometers have been used almost exclusively for these applications because of the reproducibility of fragmentation pattern ratios, the accuracy of the ion current measurement, and the wide dynamic range with which ion current measurements can be made. However, many investigations have also been made with quadrupole mass spectrometers, the principal advantage of which lies in the ability to scan the mass axis at a faster rate, making the mass spectrometer more amenable to high resolution capillary GC.

Gates et al. [5] employed an LKB-9000 gas chromatograph—mass spectrometer for metabolic profiling analyses of organic acids and developed a disk-based data system utilizing the DEC PDP-8e computer system. Ion current measurements can be taken with a range from 0 to 500 000 providing a linear dynamic range of greater than three orders of magnitude. The mass-related values are obtained originally in the output of a Hall probe field sensor. This output is squared by analog means and fed into the computer for digitalization

and storage utilizing two analog channels which provide total resolution of thirteen bits sufficient to ensure unit mass assignments in the range up to 800 a.m.u. Scans are repetitively made from the beginning to the end of the GC elution. The scans are initiated by the computer as a function of preset time intervals. The data from successive scans are collected and stored as mass—intensity pairs for each scan. Once the data have been collected specific programs for profiling analyses can be executed.

Profiling software can be divided into three general routines: library builder, data analysis, and statistical package. The heart of the software is the analysis program which is based upon a reverse library search. In contrast to a forward library search where a particular mass spectrum is taken to a library to determine the possible entries in that library that are closest to a match with the mass spectrum, a reverse library search involves taking members of the library to the raw data file in order to determine if any library member may be contained in the acquired data. Reverse library searching was chosen for profiling because the number of components that can be analyzed in any sample is limited and the nature of these components is restricted by the method of analysis. An advantage of the reverse library search is that it allows the analysis of two or more compounds that are eluting together. Furthermore, it is a very fast method of analysis. It identifies the presence of a library entry and subtracts it from the data file. This results in a residual that may contain one or more other entries in the library, or may contain one or more compounds not in the library. Iterations of this logic may reveal several co-eluting compounds.

The major disadvantage of the reverse library search is that it cannot locate and identify compounds that are not in the library. Library builder or generator-type programs must be employed that will take the residual subunits that may remain after all pertinent library entries have been subtracted from the data file and utilize this information to identify new compounds that may then be entered into the library. The heart of the reliability of any profiling analysis is in the library. Building the library is a critical and ever-continuing process. Information pertinent to the identification of any compound and to its quantitation must be incorporated into the library prior to analysis.

In the analysis program, a retention index scheme similar to that of Kovats is used. In order to convert retention time to retention index, several hydrocarbon standards from C_{10} to C_{26} are co-injected with the sample. All of the library entries have an identifying retention index which has been determined by rigorous inspection of the chromatographic performance of various standards and samples. This retention index forms the center of a search range which the program will employ by assuming that if a specific compound is present in a particular sample it must elute within a specific time interval or window which is set sufficiently large to incorporate variations in the normal chromatographic behavior. In the execution of the program for each library entry, the indicated window is searched for the presence of a designate ion that peaks within the limits of the window. The designate ion is chosen to be most indicative of that particular compound in the milieu with which it will co-elute. Upon successful identification of a designate ion peak, a confirming ion set of

two to five ions is sought. The ion intensity relationship of the confirming ion set, linked to that of the designate ion and to the variation of the retention index library value from the observed value, are all utilized in the calculation of a probability factor which indicates the confidence with which that particular compound has been identified. This value is kept in the results file and may be used to place restrictions on subsequent statistical routines. The confidence factor (CF) is dependent upon a designate ion factor, an ion match factor (f_{IM}), and a retention index factor (f_{RI}). If the designate ion peak is not found, then CF = 0; if the designate ion peak is found, then CF = $f_{IM} \times f_{RI}$.

$$f_{IM} = 100 \times \left(1 - \frac{\sum_{1}^{i} X_n}{\sum_{1}^{i} X_d}\right)$$

where

X_n = library ratio — measured ratio;

X_d = library ratio + measured ratio, where ratio = $\dfrac{\text{confirming ion intensity}}{\text{designate ion intensity}}$

If expected ratio is less than 0.25, $X = 0.75X_n$. If expected ratio is less than 0.15, $X = (0.75)^2X_n$. If ion intensity is less than 10 000, $X = 0.75X_n$. If ion intensity is less than 4000, $X = (0.75)^2X_n$.

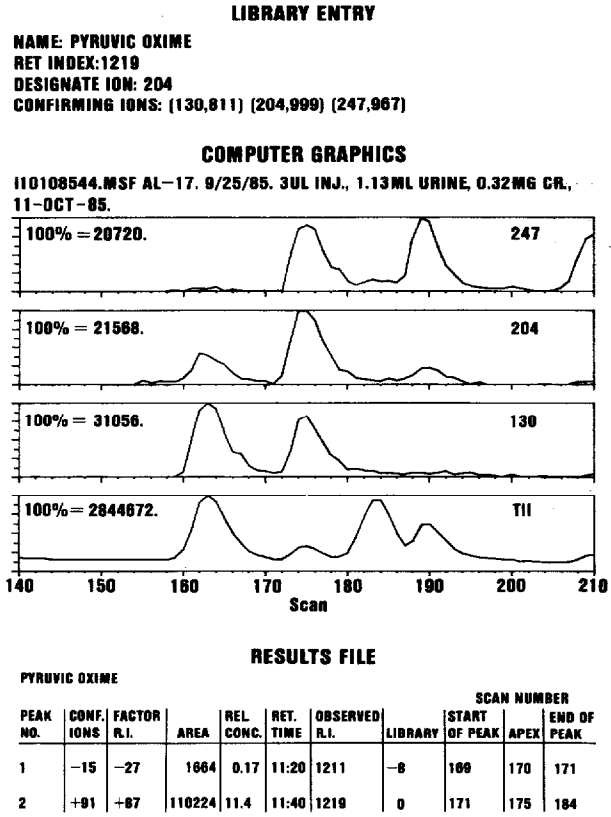
$$f_{RI} = 10 - \left|\frac{\Delta RI}{20}\right|$$

Here RI = library retention index.

In general, if CF is from 70 to 100, the identification is quite certain; if CF is from 50 to 70, the identification may be correct, but if CF is from 0 to 50, the compound is not present.

The initial procedure of the analysis program, which is called MSSMET, is a search for the internal hydrocarbon standards. Once located, these standards are used to construct the relationship between the retention time and the retention index for all other measurements in the run. The library entries are searched in the manner described above within progressively increasing window centers throughout the duration of the run file. A co-injected quantitative standard is located and the area of its designate ion peak calculated and used to quantitate the other peaks. An additional analytical factor is used which compensates for the proportionality with which a particular ion current may reflect the total concentration of the original compound in the mixture.

Examples of library entries and results of analyses output are shown in Fig. 7. All members of the library are searched for in the data file for each GC—MS run. Large peaks not in the library are keyed by a diagnostic routine in the output files. Since the profile employs a relative relationship it is important to note that unknown compounds which had been identified as being part of the profile are utilized and measured as well as known compounds. The results of the reverse library search are condensed into a file containing only the identification of the compound, the amount of the com-



RESULTS FILE

PYRUVIC OXIME

PEAK NO.	CONF. IONS	FACTOR R.I.	AREA	REL. CONC.	RET. TIME	OBSERVED R.I.	LIBRARY	SCAN NUMBER START OF PEAK	APEX	END OF PEAK
1	-15	-27	1664	0.17	11:20	1211	-8	169	170	171
2	+91	+87	110224	11.4	11:40	1219	0	171	175	184

Fig. 7. MSSMET analysis of urinary organic acids of a patient with Alzheimer's disease. The total ion intensity shows four peaks between scan 160 and scan 200. The MSSMET identifications were C_{12} hydrocarbon plus glyoxylic oxime (scan 163), pyruvic oxime (scan 175), cresol (scan 184), and 3-hydroxyisovaleric (scan 190). Using the library entry for pyruvic oxime, the computer found two peaks within a window of ± 60 seconds (± 15 scans) containing the designate ion at *m/z* 204. Only the second peak, at scan 175, fits the library set of confirming ions and retention index.

pound, and the confidence with which the compound has been identified. These results of analyses files may then be utilized to create reference sets whose membership is determined by conditions placed upon the sample subjects. Once a meaningful reference set has been created, the results of a single analysis may be compared against this set in any one of a number of statistical routines.

5.1. Gas chromatography—chemical ionization mass spectrometry

Profiling by GC—CIMS was first investigated by Horning and Horning [2], as cited above. An advantage of CI lies in the relative simplicity of the spectra produced. Very little fragmentation occurs, and the dominant ion current for each analyte appears as a protonated or adduct molecule ion species, or as a product of a simple elimination from one of these species. The resulting simplified spectra can often be useful in the resolution of complex mixtures

and offers an increase in sensitivity and a reduction in background interference. This technique, however, has not enjoyed widespread application for profiling analysis because in many cases the simplified spectra present too few ions to resolve ambiguities of identification and, more importantly, the processes producing the ions are affected by the composition of the sample in the source and the source pressure, both of which vary during sample elution. Accurate quantitation suffers under dynamic CI conditions.

5.2. Application of the gas chromatography—mass spectrometry—data system

Analyses of urinary *organic acids* by GC—MS—DS have traditionally involved visual inspection of the profile for abnormally high levels of a particular substance, and a forward library search routine to compare its mass spectrum

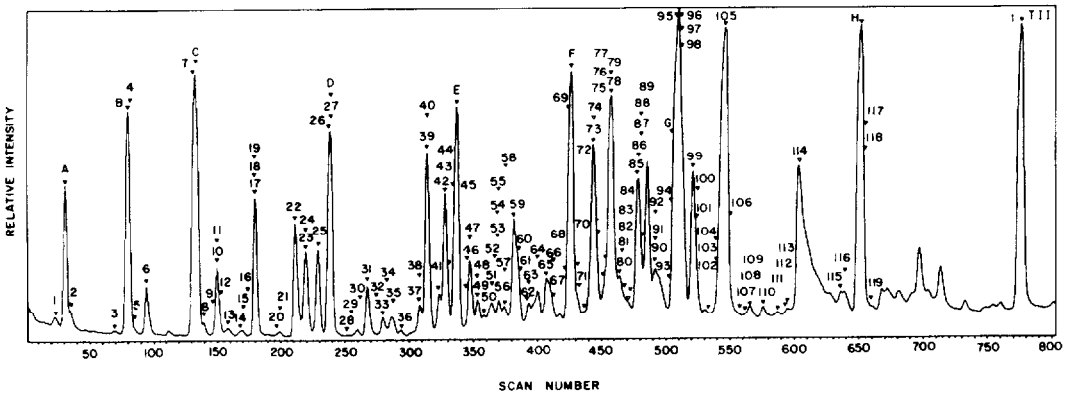


Fig. 8. Urinary acids identified by MSSMET in a urine sample from a healthy adult subject. Peaks labeled A through I are straight-chain hydrocarbons with 10, 11, 12, 14, 16, 18, 20, 24, and 28 carbon atoms per molecule, respectively. Other peaks: (1) U-1; (2) U-50; (3) α-hydroxyisobutyric; (4) lactic; (5) U-2; (6) glycolic; (7) β-hydroxybutyric; (8) U-4 (pyruvic oxime); (9) U-79; (10) U-5 (cresol); (11) oxalic; (12) U-6; (13) U-51; (14) U-7; (15) glycerol; (16) levulinic; (17) malonic; (18) methylmalonic; (19) U-RA 183; (20) U-OXB1; (21) U-9 (2-methylglyceric); (22) phosphoric; (23) U-10 (deoxyerythronic); (24) benzoic; (25) U-11 (deoxythreonic); (26) succinic; (27) fumaric; (28) phenylacetic; (29) nicotinic; (30) citramalic; (37) malic; (38) U-58 (3-methylglutaconic-peak 1); (39) U-16 (erythronic); (40) U-80 (3-methylglutaconic-peak 2); (41) U-59 (threonolactone); (42) U-17 (threonic); (43) mandelic; (44) adipic; (45) 3-methyladipic; (46) *o*-hydroxybenzoic; (47) U-60; (48) α-hydroxyglutaric; (49) U-61; (50) β-hydroxy-β-methylglutaric; (51) U-21; (52) U-82; (53) *m*-hydroxybenzoic; (54) pyroglutamic; (55) U-83 (hydroxymethylfuroic); (56) U-22; (57) *o*-hydroxyphenylacetic; (58) U-84; (59) tropic (internal standard); (60) arabonolactone; (61) α-ketoglutaric oxime; (62) *p*-hydroxybenzoic; (63) *m*-hydroxyphenylacetic; (64) U-24; (65) *p*-hydroxyphenylacetic; (66) ribonolactone; (67) arabonic; (68) suberic; (69) β-glycerophosphoric; (70) U-64; (71) U-87; (72) U-65; (73) α-glycerophosphoric; (74) U-26; (75) *cis*-aconitic; (76) U-66; (77) U-67; (78) U-68; (79) citric; (80) azelaic; (81) terephthalic; (82) vanillic; (83) U-89; (84) U-29; (85) homovanillic; (86) galactono-1,4-lactone; (87) *m*-hydroxyphenylhydracrylic; (88) veratric; (89) U-30; (90) *o*-coumaric; (91) hexuronic; (92) gluconic; (93) *p*-hydroxyphenyllactic; (94) U-72; (95) vanilmandelic; (96) ascorbic; (97) U-91; (98) hexuronic; (99) hexuronic; (100) hydrocaffeic; (101) U-74; (102) U-2071; (103) palmitic; (104) U-75; (105) hippuric; (106) caffeic-peak 1; (107) U-76 (hydroxydecanedioic); (108) U-77; (109) U-37; (110) indoleacetic; (111) U-NE8; (112) caffeic-peak 2; (113) urocanic; (114) uric; (115) U-41; (116) *m*-hydroxyhippuric; (117) U-42; (118) 3,4,5-trimethoxycinnamic; (119) 5-hydroxyindoleacetic. Unknowns are identified by a number containing U; numbers are given chronologically to peaks in the chromatogram. In some cases unknowns are tentatively identified. (Taken from ref. 5 with the permission of the authors and publisher.)

with those of a large number of reference compounds [6]. Thus, only a few peaks in a profile are identified, and they are chosen on the basis of being excreted in amounts substantially greater than normal. The computer system described above has the advantage that most, if not all, of the metabolites are identified and quantitated. The profile illustrated in Fig. 8, for example, contained more than 100 identified organic acids and numerous unknowns, the levels of each of which were outputted in a tabular form.

As with organic acids, the urinary *steroid profiles* obtained from GC—MS—DS are best analyzed using a reverse library search algorithm. Historically, obtaining both mass spectral and retention index information insured the positive identification of the urinary steroids. Structural epimers of urinary metabolites such as androsterone and etiocholanolone produce similar mass spectra, and one therefore needs to know the order of elution in order to characterize each compound. On the other hand, steroids with differing structures may exhibit incomplete chromatographic separation and must be identified on the basis of MS information.

When used in conjunction with other methods, GC—CIMS can be extremely helpful as evidenced by Szczepanik et al. [7] where this technique was employed with GC—electron impact (EI) MS and selected ion monitoring (SIM) in the analysis of 48 *bile acids* and related compounds. The acids were chromatographed as their methyl esters derivatives using 0.5% SP-525 as the stationary phase on 100—200 mesh Gas-Chrom Q in a glass column (103 cm × 1 mm). The combination of the relative retention time, the EI spectra, and the CI data formed a balanced and integrated system for the characterization of bile acid mixtures.

6. PROFILING BY PLANAR CHROMATOGRAPHY

6.1. Paper chromatography

Although Williams [1] published his patterns years before the definition of the metabolic profile, his outstanding work meets all of the criteria of a profile except that the analyses were not carried out by a single method and with a single sample. Williams [1] used data from several analyses, mostly paper chromatography, to create the "individual man" metabolic patterns as opposed to the non-existent "normal man" patterns. Several different solvents and color reagents were used to analyze saliva and urine for a variety of biochemical compounds including amino acids and their derivatives, sugars and sugar derivatives, vitamins and vitamin derivatives, purines, pyridines, carboxylic acids, and various other organic compounds as well as some inorganic ions. In order to expedite the very large number of analyses in this study, only the simplest procedures were used, i.e. visual comparison of color and quantitation by measurement of spot area. The large variations of the patterns between individuals appeared to justify the casual procedures.

In another application of paper chromatography, Armstrong et al. [8] analyzed the phenolic acids in human urine. After development of two dimensional paper chromatograms with reagents such as diazotized sulfanilic acid, visual patterns served as the basis for definition of profiles of these

metabolites. An interesting application was the search for a biochemical basis for schizophrenia, which resulted in the identification of *m*-hydroxyhippuric and *m*-hydroxyphenylhydracrylic acids as being typically over-excreted by most patients with schizophrenia but which were actually found to be derived from coffee rather than an abnormal metabolic pathway [9].

6.2. Thin-layer chromatography

Thin-layer chromatography (TLC) offers several unique features in its application to profiling analysis. Very small amounts of compounds can be processed and analyzed without loss and in many situations, with less extensive preparation than that needed for paper chromatography. Additionally, the spots and bands diffuse less during chromatographic development. The developed TLC plate is also amenable to direct staining and quantitative scanning (densitometry) by photoelectric means. Two types of profiles can be obtained by this technique. In one case, the stained plate can be photographed, with the print itself serving as the template for future comparison. However, a more visually interpretable profile can be obtained from the densitometer recordings from linear scans of the plate.

In an excellent early application of this technique, Shackleton and Mitchell utilized silica gel plates for the measurement of 3β-hydroxy-Δ^5-steroids, using antimony trichloride for color development. A reflection scanner was used for quantitative analysis by comparison with developed standards. Other thin-layer techniques and GC—MS were used to confirm the identity of the steroids.

7. PROFILING BY COLUMN LIQUID CHROMATOGRAPHY

7.1. Classical column liquid chromatography

Since the elution process of column liquid chromatography separates in the time domain the various components of a mixture, a means for serial on-line detection is normally employed with the resulting recorder tracing creating a two-dimensional chromatogram in a manner analogous to that of GC. Again, these chromatograms are profiles whenever the other analysis criteria are met. Detection means such as refractive index and absorption spectrometry are very accurate but are somewhat inferior in terms of dynamic range and sensitivity. Fluorescence detection greatly increases the sensitivity with a somewhat greater working range. Newer techniques such as the thermal lens effect have not been employed to date, but have the potential to increase the analytical scope of column chromatographic analyses for profiling applications.

Ion-exchange chromatography has been utilized by many investigators for amino acid analyses. The application of this technique in situations which require isolation, purification, and digestion of polypeptides prior to analyses does not fall within the realm of profiling. In these cases, the profile establishes the amino acid composition of the polypeptide but gives little or no information regarding a metabolic state or condition. Nevertheless a lot of methodological experience applicable today in the profiling procedures stems from the studies about the primary structure of proteins. There are many

applications, however, where amino acid profiles from samples of urine, blood, or tissue extracts is important in clinical diagnosis [11]. For more detailed information the reader is directed to p. 177 of this volume.

Jolley and Freeman [12] utilized anion-exchange chromatography with a colorimetric detection system for the analysis of carbohydrates in normal and diabetic urine and serum. Monosaccharides and small oligosaccharides were separated on columns (1.5 m × 0.62 cm of Dowex 1 with gradients of various borate buffers) which required very long analysis times (up to 24 h). An ultraviolet (UV) detector was later employed for the profiling analysis of a variety of UV-absorbing metabolites. This application, which is one of the few examples where biochemically divergent metabolites have been analyzed in the same run, typically required almost two days to complete the chromatogram [13].

Barness et al. [14] used silicic columns and colorimetric detection for the analysis of organic acids from blood plasma, red cells, and cerebrospinal fluid. The method was nearly three orders of magnitude less sensitive than GC but had the advantage that urine samples could be directly analyzed without a chemical pre-purification step.

7.2. High-performance liquid chromatography

The technique with the greatest potential to challenge GC for profiling analysis is high-performance liquid chromatography (HPLC). The sometimes extensive sample clean-up and derivatization steps necessary prior to GC are greatly abbreviated or completely omitted in applications of HPLC to profiling analysis. HPLC separations can isolate, clean-up, and analyze, all in the same run. Presently, the characteristics of the detection systems are the limiting factors in the increased utility of this technique. HPLC is more powerful than classical column liquid chromatography both in column efficiency and analysis speed, but it suffers the same restrictions in dynamic range and sensitivity since both methods commonly employ similar kinds of detection. The versatility of this technique, however, has encouraged the development of more exotic detection systems such as MS and Fourier-transform infrared spectroscopy and its application to profiling will undoubtedly increase in the future.

The applicability of modern HPLC procedures can be demonstrated on the analyses of amino acids from plasma or urine samples where the results were achieved in less than 1 hour on reversed-phase C_{18} columns [11, 15]. This topic is discussed in more detail on p. 177.

8. PROFILING BY COLUMN LIQUID CHROMATOGRAPHY–MASS SPECTROMETRY

The introduction of analytes in a liquid chromatography effluent stream into the mass spectrometer has been a difficult analytical problem, which has limited the proliferation of this technique. Several approaches have been developed, most notable of which are the moving belt, CI reagent gas transfer, and more recently, the thermospray. Problems inherent in each of these transfer techniques may place limitations on the compounds that can be analyzed, as well as sensitivity, reproducibility, and dynamic range of the

system, making them relatively unattractive for general profiling analysis.

One instrumental approach, however, has been developed by Horning et al. [16] with the potential for profiling application. This system is called atmospheric pressure ionization (API) MS. The entire effluent from the liquid chromatograph is vaporized in an API chamber. Ions are generated external to the low pressure source of the mass spectrometer by a corona discharge. Very high ion densities are formed by this method, some of which pass into the mass spectrometer for analysis. This method of ionization significantly increases both the sensitivity and the range of the subsequent MS analysis.

9. PROFILING BY ELECTROPHORESIS

The classical blood protein analysis by electrophoresis has for many years produced profiles of the circulating proteins. Although limited in resolution, sensitivity, and dynamic range, these techniques, with densitometer tracing, fit all of the criteria of profiling and, in fact, may be viewed as the most senior of the methodologies practising this art. In a single typical procedure, all of the 150 or so soluble proteins are measured. In fact, all but ten to twenty of these proteins are usually present at too low a concentration to be detected. However, any significant positive variation could affect the tracing. Presently, the method is sufficiently mature that nearly all of the components of the profile amenable to any particular procedure are now known. It is important to recall, however, that in the earliest reported use of this technique, the profile contained mostly unknown substances. Identifications were sometimes based upon gel migration (molecular weight) relative to reference standards. From the start, pattern recognition and pattern analysis has been of paramount importance in the determination of the clinical significance of each analysis. Indeed, the development of this art may serve as a model for all of profiling analysis, clearly emphasizing the utility of unknowns as well as knowns and the importance of interrelationships as well as absolute quantitative amounts.

In recent years, a more esoteric version of electrophoresis has been developed, two-dimensional gel electrophoresis. In the prominent manifestations of this technique, gels are cast such that one axis, usually the ordinant, resolves on the basis of sodium dodecyl sulfate (SDS) molecular weight, while the other axis resolves by isoelectric focussing. The resulting two-dimensional map can be photographed or otherwise reproduced and/or analyzed by computer means. At this stage of the development of this technique, most of the proteins visualized by Coomassie blue or silver staining are still unidentified. All may be given a number or a name based upon the field coordinates where they are found. Visual pattern recognition is the dominant mode of profile analysis. The establishment of a library or a protein index classification has been proposed based upon the specific fluid or tissue used as the sample source. Significant variations in the profiles are sought and where observed, attempts are made to correlate with the metabolic state or other conditions of health.

A complete characterization of the proteins of a single cell may be the most difficult of the profiling problems attempted to date. The two-dimensional electrophoresis method, as developed by O'Farrell [17] and Anderson and

Anderson [18] has led to an ability to separate more than 1300 polypeptides from asynchronous Hela cells [19]. The problems associated with this approach are (1) to have sufficient resolution to separate up to 2000 or more proteins; (2) to cope with both basic and acidic proteins in the same mixture; (3) to visualize components that may be as much as 10 000-fold different in concentration; and (4) to carry out these analyses with sufficient reproducibility in that any component can be identified from its map locus.

10. PROFILING BY FAST ATOM BOMBARDMENT-MASS SPECTROMETRY

The development of various desorption/ionization techniques in the 1970's extended the ability of MS to analyze polar non-volatile compounds of medium to high mass. In particular, the development of FAB in 1981 enabled new approaches for the analysis of intractable molecules. In this technique, the sample is dissolved in a viscous solvent such as glycerol and placed on a probe external to the MS ion source. After insertion of the sample-laden probe into the vacuum of the source, the probe is bombarded with a beam of energetic (6—10 keV) atoms, usually argon or xenon.

Both positive and negative analyte ions are created by this technique and are measured with little or no fragmentation. The major advantage of this technique is that it extends the scope of MS analysis to many types of biomolecules without the need for cumbersome and sometimes impossible derivatizations.

The limitations of FAB-MS for general profiling applications lie in the time-consuming sample insertion processes, the one-dimensional nature of the analysis (searches by mass only) and variations in the ion currents which made quantitation difficult. The use of a neutral beam is not essential as comparable data can be obtained using energetic ion beams in a technique called secondary ion mass spectrometry (SIMS). In applications to profiling FAB and SIMS can be considered as analogous.

Shackleton and Straub [20] utilized both FAB and SIMS for the direct profiling analysis of steroid conjugates in urine. The steroids and steroid conjugates were separated from the sample by either Sephadex LH-20 chromatography or diethylaminopropyl-LH-20 chromatography. The samples were analyzed with a Kratos MS-50 mass spectrometer with a FAB accessory. The negative ion spectra were found to be more informative with the exception of the sulfates. The advantage of profiling by this technique was that it enabled mixtures of steroid conjugates to be resolved without major interferences from fragment ions. The major disadvantage was the loss of the fragmentation information needed to distinguish isomeric compounds. The potential of this technique for specific types of profiling analysis is obvious and applications should grow as this art is extended.

11. PROFILING BY MASS SPECTROMETRY—MASS SPECTROMETRY

The development of tandem mass analyzers has been of great interest to practitioners of the art of profiling. These complex mass spectrometers use selected magnetic (B), electric (E) and quadrupole (Q) mass separators in

Fig. 9. Instrumental configuration used for profiling by MS—MS.

various serial combinations to achieve a second order of mass separation and measurement. In a typical manifestation of this art, as shown in Fig. 9, a sample is ionized and resolved by the first mass separator. The ion selected for transmission (parent ion) is directed through a reaction chamber where it interacts with a gaseous molecular species. Energy transferred to the ion produces dissociations [collisional activated dissociation (CAD) or collisionally induced dissociation (CID)] which yield neutral products and smaller ionic species (daughter ions). These ion products are resolved by the second mass analyzer and measured directly or under various protocols defining relationships between the two mass separators (linked scanning modes).

The special interest of this type of MS instrumentation for profiling lies in the ability of the first mass separator to isolate from a complex sample mixture a single beam of one mass-to-charge ratio. This ion beam may contain one or more ionic species but is unquestionably less complex than the original mixture. The daughter ions from this parent ion beam can reveal the presence of more than one species and provide information for both qualitative and quantitative analysis. Using the first separator to isolate the components under study and the second separator to analyze them has the potential for profiling applications performed without extensive clean-up procedures. The reduction in the background for the final analysis furnishes a favorable signal-to-noise for the measurement which may increase both sensitivity and dynamic range. Obviating the need for extraction, derivatization, and chromatography greatly reduced the total analysis time.

Hunt et al. [21] cite total analysis time in the vicinity of 15 min for the profiling of carboxylic acids in lyophilized urine samples. Using the technique of negative chemical ionization, three different scan modes were used: daughter ion, parent ion, and neutral loss scans.

12. PROFILING BY FIELD IONIZATION MASS SPECTROMETRY

In the presence of a large field, gradient ions can be formed with very little internal energy and, hence, very little fragmentation. The large field gradients are produced in the mass spectrometer source by use of various materials with very sharp points or edges. These range from activated emitters (usually a tungsten wire with amorphous carbon dendrites), to razor blades, broken rods, and pre-formed ridges or orifices.

In this technique, the samples are volatilized by thermal means, and directed to the field emitter. Ions are formed on or near the emitter surface and are accelerated into the mass analyzer by the high fields.

The advantage of field ionization (FI) for profiling lies only in the soft ionization which yields very simple spectra approaching the one ion-one analyte ideal. This makes identification relatively simple in situations where a limited number of analytes are present. The drawbacks of FI for profiling lie in the necessity for prior volatilization and the one-dimensional nature of the

resulting resolution. The very high fields essential to the technique produce a susceptibility for unwanted spark discharges and non-uniform gradient distributions. These artifacts, coupled with the condensing/desorption processes on the emitter itself, result in relatively poor quantitative measurements. These disadvantages have severely limited profiling applications of this technique.

13. CONCLUDING REMARKS

Do profiling techniques that are expensive and time- and resource-consuming have a growth outlook for the future? Even assuming the present level of instrumentation technology, there will continue to be an expansion of the applications of profiling. The value of pattern recognition and the correlation of pattern factors to metabolic state have clearly been documented by the utility of electrophoresis patterns for blood proteins and isoenzyme analysis. Extensions of the art to subjects other than humans is also occurring, as evidenced by e.g. the phenotyping of bacteria by the lipid composition profile. A commercial system is now available for this analysis. Although this is not a profile of a metabolic state, the strategy is identical to that of profiling.

A direct result of profiling analysis has been the identification of numerous metabolites, many of which are novel, as the investigators build their libraries. This number of known metabolites will increase as various applications mature, giving profiling a scientific justification that could stand alone as sufficient rationale for the effort and expense. In summary, the present state-of-the-art is healthy, and growing. The future will be even better.

It can be anticipated, however, that there will be major growth in three areas germaine to profiling. The first area is in the analysis of the profile. The mostly human visualization of profiles, as used today, will be replaced by mostly computer visualizations in the future. Not only will the obvious and trivial associations be made automatically but, in addition, sophisticated statistical packages including modelling, pattern recognition, and artificial intelligence will become common practice. Computer-assisted pattern recognition has been the objective of several investigators and great progress is apparent. On this foundation, a new art will be established in the future. The second area of profiling that will be altered in future practice lies in the use of robots to automate the time-consuming, repetitive, and exacting separations, derivatizations, and chemical processing routines where they are encountered. Automation of the analysis itself will also become commonplace. Although the initial cost of this automation will be high, the cost per analysis will fall rapidly after a period of time using this mode of analysis. The third area of major change in profiling will concern the scope of the analysis. The one-dimensional profiles that were essential for the human cognitive processes will yield to multidimensional analyses under the eye of the computer. Several analyses will be combined under linked logical relationships to increase the qualitative ability and enhance the sensitivity and range of profiling applications. For example, a single sample may be analyzed utilizing a split-stream and two different chromatographic columns operating under two different experimental protocols. The results of these two profiles will then be combined by computer means and both used in the decoding of the information contained. In addition to multi-

dimensional analyses, new and specific detection systems are becoming widespread. In automated analytical systems several of these detectors may be incorporated, each utilizing its specificity and sensitivity in an overall network that yields non-discriminating quantitative analysis of very large numbers of similar metabolites. It is obvious as we move into the future that profiling done on the basis of a single analysis will be superceded by profiling analyses that are automated, computerized, and broadened in scope, yielding multi-dimensional pictures of health.

14. SUMMARY

This review summarizes the instrumental techniques applicable for metabolic profiling analysis. Since its origin with the work of Roger Williams, use has been made of planar chromatography, column liquid chromatography, electrophoresis, gas–liquid chromatography, and mass spectrometry to obtain analyte patterns. Strategies for the analysis of various metabolites including organic acids, steroids, drugs, bile acids, carbohydrates, and proteins are discussed.

15. ACKNOWLEDGEMENTS

Research reported herein and preparation of the manuscript were supported in part by a research grant (RR00480) from the Divison of Research Resources, National Institutes of Health. The authors are grateful to B. Musselman and M. Micke for assistance in preparing graphic materials, J.T. Watson for helpful suggestions, and V. McPharlin for preparing the manuscript.

REFERENCES

1 R.J. Williams, Biochemical Institute Studies IV. Individual Metabolic Patterns and Human Disease: An Exploratory Study Utilizing Predominantly Paper Chromatographic Methods, University of Texas Publication No. 5109, University of Texas, Austin, TX, 1951, Ch. 1.
2 E.C. Horning and M.G. Horning, Clin. Chem., 17 (1971) 802.
3 C.H.L. Shackleton and W. Chai, Endocrine Rev., 6 (1985) 441.
4 M. Tuchman and R.A. Ulstrom, Adv. Pediatr., 32 (1985) 469.
5 S.C. Gates, M.L. Smisko, C.L. Ashendel, N.D. Young, J.F. Holland and C.C. Sweeley, Anal. Chem., 50 (1978) 433.
6 E. Jellum, J. Chromatogr., 143 (1977) 427.
7 P.A. Szczepanik, D.L. Hachey and P.D. Klein, J. Lipid Res., 17 (1976) 314.
8 M.D. Armstrong, K.N.F. Shaw and P.E. Wall, J. Biol. Chem., 218 (1956) 293.
9 M.D. Armstrong and K.N.F. Shaw, J. Biol. Chem., 225 (1956) 269.
10 G.H.L. Shackleton and F.L. Mitchell, Steroids, 10 (1967) 359.
11 G.A. Qureshi, L. Fohlin and J. Bergström, J. Chromatogr., 297 (1984) 91.
12 R.L. Jolley and M.L. Freeman, Clin. Chem., 14 (1968) 538.
13 R.L. Jolley and C.D. Scott, Clin. Chem., 16 (1970) 687.
14 L.A. Barness, G. Morrow, R.E. Nocho and R.A. Maresca, Clin. Chem., 16 (1970) 20.
15 H. Godel, T. Graser, P. Földi, P. Pfaender and P. Fürst, J. Chromatogr., 297 (1984) 49.
16 E.C. Horning, D.I. Carroll, I. Dzidic, K.D. Haegele, M.G. Horning and R.N. Stillwell, J. Chromatogr., 99 (1974) 13.
17 P.H. O'Farrell, J. Biol. Chem., 250 (1975) 4007.
18 N.G. Anderson and L. Anderson, Clin. Chem., 28 (1982) 739.
19 R. Bravo and J.E. Celis, Clin. Chem., 28 (1982) 766.
20 C.H.L. Shackleton and K.M. Straub, Steroids, 40 (1982) 35.
21 D.F. Hunt, A.B. Giordani, G. Rhodes and D.A. Herold, Clin. Chem., 28 (1982) 2387.