Automated Metabolic Profiling of Organic Acids in Human Urine.
I. Description of Methods

Stephen C. Gates,¹ Nancy Dendramis,² and Charles C. Sweeney²,³

We describe a complete procedure for separation and mechanized analysis of organic acids in human urine. The acid fraction of urine is separated by anion-exchange chromatography on diethylaminoethyl-Sephadex. Individual acids are identified and measured by use of a gas-chromatography/mass spectrometer/computer system that can clearly distinguish contributions from at least 150 substances in a single sample. We discuss analytical recoveries, contributions from the sample separation process, stability of stored samples before and after processing, and reproducibility of the extraction procedure.

Additional Keyphrases: separating data on peaks in a multicomponent chromatogram by computer techniques • gas-chromatography/mass spectrometer/computer • components of urine

A ubiquitous “problem” encountered by the clinical chemist is the presence of “uninteresting” compounds in a sample of biological material being assayed for a particular substance or group of substances. Regardless of whether these substances are detected as background, matrix effects, or interfering substances, elaborate separation schemes must often be implemented to eliminate them sufficiently so that the analytes of interest can be measured. Alternatively, the data can be separated after they have been collected. That is to say, the multicomponent sample is analyzed with use of a detector that responds equally well to a wide variety of components; after the data-collection process, a sophisticated computer algorithm, using pattern recognition, factor analysis, principal component analysis, or other such techniques, is applied to separate the data for the compound of interest.

A potential advantage of a data-separation technique, rather than a chemical-separation technique, is the possibility that sample processing and assay development time can be reduced, because a specific method need not be discovered for each component of interest. Data-processing techniques also provide the analyst with the capability to measure tens or hundreds of components simultaneously, yielding data that are commonly referred to as “metabolic patterns” (1) or “metabolic profiles” (2, 3).

In most systems developed thus far for metabolic profiling, gas-chromatographic (GC)⁴ peak areas are used to evaluate compounds; thus, for example, three groups (4-10) have utilized peak areas obtained from low-resolution (packed) GC column separations of the organic acids in human urine. This approach, however, is incapable of accurately separating peak area contributions from minor or incompletely resolved peaks. A more satisfactory approach (11, 12) has been used for some types of samples: separation of volatiles from urine by capillary GC and qualitative analysis by computerized pattern-recognition techniques. GC peak areas (or heights) are still used for quantitation, but many components are not completely separated even by the capillary columns, and so quantitation of some compounds may still be inaccurate.

Two mass spectrometry systems have recently been reported. In one (13, 14), areas or heights of individual mass chromatogram (15) peaks are used. In the second, contributions of more than one component of unresolved GC peaks are resolved by a tabular peak-modeling technique applied to mass chromatograms (16, 17). In both systems, gas chromatography and mass spectrometry are used, with computer acquisition and analysis of the data.

We describe here our development of a completely mechanized system that can accurately and precisely quantitate complex biological mixtures. Our experience (18-23) has indicated that the relatively simple procedures that depend on the determination of GC peak areas work well with mixtures of pure compounds but are inadequate when applied to biological samples. In the mass spectral metabolite (MSSMET) system that we have developed (23), algorithms that make use of the intensities of selected ions for quantitation make possible the measurement of components that are completely unresolved by the GC. During the past two years, we have repeatedly tested this system with mixtures of organic acids from human urine.

Methods

A modified version of the procedure of Thompson and Markey (5) is used to obtain the organic acids from urine. DEAE-Sephadex A-25 (Pharmacia Fine Chemicals, Piscataway, N.J. 08854) is swollen for 48 h in 1.0 mol/liter pyridinium acetate, which is followed by treatment with 0.5 mol/liter pyridinium acetate. Both solutions are changed several times. A slurry of the DEAE-Sephadex is added to a glass column (1 cm i.d.) with a 200-ml reservoir (Kontes, Vineland, N.J. 08360) containing a lightly-packed plug of glass wool. When the column packing is 8 cm high, 21 ml of 0.5 mol/liter pyridinium acetate is added and allowed to drain through the column. The column is left at this stage until the sample is ready.

Urine, previously stored at −80 °C, is brought to room temperature in a bath of warm water. A 1.0 g/liter solution of the internal standard (tropic acid) in methanol, also stored at −80 °C, is likewise warmed and 50 µl of it is placed in a si-
Fig. 1. Typical library entry for MSSMET analysis of urine samples, based on a reverse library search using GC retention indices and selected mass chromatograms. The MSSMET library includes "options" (entries preceded by a slash), which automatically set peak-detection parameters and compound entries (preceded by an asterisk). For each compound, the library entry includes an identifying number and IUPAC and common names; GC retention index; m/z of the "designate" ion and value of the "k-factor" (23); and the masses and relative intensities of the confirming ions.

lanized 13 x 100 ml test tube. The methanol is removed with a stream of dry nitrogen. The urine is stirred to homogeneity and an aliquot calculated to contain approximately 1.5 mg of creatinine (6) is placed in the test tube containing the tropic acid. The contents of the test tube are sonicated to assure complete mixing.

To each solution thus obtained is added 3 ml of aqueous 0.1 mol/liter Ba(OH)₂ (Fisher Scientific Co., Pittsburgh, Pa. 15219). The contents of each tube are mixed and then centrifuged for 30 s at maximum speed on a small table-top centrifuge. The supernatant solution is removed by pipette and the precipitate washed and centrifuged twice, an additional 1.0 ml of 0.1 mol/liter barium hydroxide solution being added after each centrifugation.

Oxime derivatives of oxo-acids in the mixture are prepared by adding to it 200 µl of a 75 g/liter solution of hydroxylamine hydrochloride (J.T. Baker Chemical Co., Phillipsburg, N.J. 08865) and then heating at 80 °C for 20 min. The samples are then cooled to room temperature in an ice bath and 2 mol/liter hydrochloric acid or 2 mol/liter acetic acid solution is added until the pH is 7 to 8.

The sample so prepared is applied to the DEAE-Sephadex column and cations and neutral substances are eluted with 50 ml of redistilled water. This eluate is discarded. Next, the anionic (acidic) metabolites are eluted with 40 ml of 1.5 mol/liter pyridinium acetate into a clear, silanized, 250-ml round-bottom flask. The sample is frozen in a solid CO₂-acetone bath, with rotation of the flask to evenly coat the sample on the inside of the flask. If necessary, the flask is stored at -80 °C until the lyophilizer is available. The sample is dried on a lyophilizer (Virtis, Gardiner, N.Y. 12525) at a pressure of 13.3 Pa (0.1 Torr) or less until the sample is about 5 ml. The sample is then melted and transferred to a silanized 50-ml conical centrifuge tube with a ground-glass stopper. The sample is resubmerged by submerging the stopped tube in a solid CO₂-acetone slurry. The stoppers are then removed and the tube mouths covered with several layers of coarse-mesh cheesecloth, secured with a rubber band, to prevent losses of sample material. The centrifuge tubes are loaded into a 1-liter lyophilizer flask (Virtis) and the contents lyophilized to dryness. Tubes are removed as soon as the material is dry and powdery. The dried samples are silylated in the same tubes by adding 250 µl of a silylating mixture containing bis(tri-methylsilyl)trifluoroacetamide and trimethylchlorosilane (both from Pierce Chemical Co., Rockford, Ill. 61051) and redistilled dry pyridine (200/2.50 by vol). The stopped tubes are heated at 80 °C for 1 h; an inverted test tube rack is placed over the stoppers to hold them in place. The tubes are shaken at least once during the heating process, to ensure complete silylation.

The silylated samples are immediately transferred to silanized glass capillaries, prepared from 100-µl disposable micropipettes (Dade, Miami, Fla. 33152) that are broken in half and flame-sealed at one end. Each capillary is filled about half full with a 500-µl syringe, then flame-sealed until a small bubble begins to form in the glass at the heated end. Sealed capillaries are stored at 4 °C until used. Between each use, the 500-µl syringe is cleaned with dry redistilled methanol, followed by redistilled hexanes (residual methanol will react with the derivatizing agents).

Gas-Chromatographic-Mass-Spectrometric Analysis Procedure

Derivatized samples are analyzed with an LKB 9000 gas chromatograph-mass spectrometer (GC-MS) as previously described (23). In brief, the analysis consists of separation of the sample on a 310 cm x 2 mm (i.d.) column of 5% OV-17 on 80/100 mesh Supelcoport (Supelco, Bellefonte, Pa. 16823). The column oven is temperature programmed from 60 to 280 °C at 4 °C/min while the magnet of the mass spectrometer is scanned repetitively over field strengths corresponding to m/z 50 to 550 every 4 s. Data are collected with a PDP 8/e computer system, then transferred to a PDP 11/40 computer.
Fig. 3. Mass spectral metabolite system output
Output is for data shown in Fig. 2; peaks at scans 382 and 407 are from the trimethylsilyl derivatives of tropic acid and p-hydroxyphenylacetic acid, respectively. In addition to the compound number and name, the following information is printed for each peak of the designated ion within the window: a sequence number; the match category (+, 7 or -) designates ion peak area and substance concentration, expressed in exponential notation. Below each of these is the corresponding value based on peak height instead of peak area. In addition, retention data are printed for each peak, including the observed retention time in minutes and seconds; deviation of the retention time from the value expected; observed retention index; deviation of the retention index from the value expected; and the scan numbers corresponding to the beginning, apex, and end of the peak of the designated ion

Fig. 4. Reproducibility of the analytical procedure
Two aliquots of the same urine sample were each separated on DEAE-Sephadex columns. The resulting organic acid fractions were trimethylsilylated, derivitized, and then analyzed on separate days by GC-MS. Hydrocarbon standards were co-injected with each sample; hence m/z 73 is plotted to show only the trimethylsilyl derivatives present in each mixture.

System for analysis by MSSMET (19–23). MSSMET is used to convert the approximately 5 × 10^5 words of data generated from the GC-MS analysis of each sample to a set of 100 to 200 compound names and uncorrected or relative concentrations. It uses GC retention indices in a reverse library search of selected ion mass chromatograms. Its operation is virtually completely automatic. This process is illustrated in Figures 1–3.

Results and Discussion
Separation of Organic Acids
Since quantitative as well as qualitative accuracy is desired, the chemical separation procedure must be reproducible, so that contributions of the procedure to total variability will be as small as possible. Unfortunately, diethyl ether–ethyl acetate extraction and other extraction methods commonly used to separate organic acids from urine are not very reproducible, especially with polar substances. The original DEAE-Sephadex procedure described by the Hornings (24) has not been satisfactory, in our experience, because the eluate from the column often contains so much of some unidentified residue that it cannot be lyophilized to dryness. Hence, large quantities of silylating reagent must be used, producing solutions that are too dilute for satisfactory GC-MS analysis.

A somewhat more satisfactory procedure is the modified DEAE-Sephadex separation proposed by Thompson and Markay (5), in which barium hydroxide treatment before separation of the sample on the DEAE-Sephadex removes much of the phosphates and sulfates that normally are present in urine. In most other respects the procedure follows that proposed by Chalmers and Watta (7, 8), who made a very careful study of the optimum conditions for the DEAE-Sephadex procedure and suggested O-ethyloxime formation of oxo acids before silylation.

The anion-exchange portion of the Thompson and Markay procedure gave poor reproducibility in our hands. Modifications to improve the method have included an increase in the quantities of eluents (water and 1.5 mol/liter pyridinium acetate), the addition of pyridine to the silylating mixture, the use of sealed silanized glass capillaries to store silylated samples, elimination of a step involving transfer of the acids in methanol, and the use of creatinine concentration as an index to the amount of sample to be applied to the column. It should be noted that the ratio of eluents in this modification corresponds very closely to that used by the Hornings (24) and by Chalmers and Watta (7). The use of silanized capillaries for storing silylated samples results in long-term stability of the samples, even at room temperature. Multiple peaks that might be produced from nitrogen-containing compounds, particularly indolic acids, are eliminated by the added pyridine in the silylating mixture. The barium hydroxide step leads to a much dryer product. Domino and Mathews (University of Michigan, private communication) have recently suggested that the residue after lyophilization can be decreased even further if acetic acid is used instead of hydrochloric acid to adjust the pH of the urine before the sample is placed on the DEAE-Sephadex column. Furthermore, the use of a creatinine standard helps ensure that roughly the same total quantity of organic acids is separated from each urine sample, which is not the case if a fixed volume of sample is added to the column each time.

With these changes in the procedure, the method has been reproducible (Figure 4) and analytical recoveries of many substances exceeded 90% (Table 1), while others were not recovered in good yields, perhaps because of the barium precipitation step (25, 26). Coupled with reasonable stability of stored urine samples (Figure 5) and a reasonably clean procedural blank (Table 2), the method provides acceptable quantitative results for most substances studied.

Table 1. Analytical Recoveries of Organic Acids by Use of the Barium Hydroxide–DEAE-Sephadex Method

| Compound added | Mean recovery, % *
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Succinic acid</td>
<td>92</td>
</tr>
<tr>
<td>α-hydroxybenzoic acid</td>
<td>67</td>
</tr>
<tr>
<td>β-Hydroxy-β-methylglutaric acid</td>
<td>98</td>
</tr>
<tr>
<td>Tropic acid</td>
<td>91</td>
</tr>
<tr>
<td>α-Glycerophosphoric acid</td>
<td>87</td>
</tr>
<tr>
<td>Citric acid</td>
<td>16</td>
</tr>
<tr>
<td>Homovanillic acid</td>
<td>94</td>
</tr>
<tr>
<td>Vanilmandelic acid</td>
<td>98</td>
</tr>
<tr>
<td>Ascorbic acid</td>
<td>10</td>
</tr>
<tr>
<td>Indoleacetic acid</td>
<td>92</td>
</tr>
</tbody>
</table>

* Based on analysis of duplicate samples at each of two different concentrations of the added compound in urine.
Collection and Analysis of GC-MS Data

For maximum sensitivity and precision, the mass spectrometer is set at the fastest magnet scan speed (4 s per scan cycle from m/z 50 to 550 and back) and highest multiplier gain (gain 8 on the LKB-9000 instrument) that can be used without severely impairing the quality of the data. At these settings, the mass spectrometer collects data so rapidly and over such a wide dynamic range that a very sophisticated data system (23) must be used, and the mass spectrometer itself must be kept at optimum operating conditions. In addition, the GC columns must be tested daily for signs of degradation and replaced frequently. Ion source sensitivity is extremely critical, because many compounds are present in only very low concentrations. Therefore, it may be necessary to spend a great deal of time in cleaning the ion source and separator parts, fine-focusing the ion source, and eliminating sporadic sources of electronic “noise.”

After transfer to a larger computer (a PDP 11/40), the GC-MS data are analyzed by MSSMET. Because MSSMET uses a reverse library search procedure, in which the multiple mass spectra obtained during the GC run of the urinary organic acids are searched for matches to a small library of spectra of substances expected to be present, it is important to select a suitable library of organic acids. Many of the reference spectra were obtained by analysis of reasonably pure reference compounds, but more were obtained by inspection of individual urine-sample GC-MS data sets for spectra not in the library. These additional spectra were located by using a manual inspection of the data and two automated methods designed for this purpose (27). Specific ions differentiating these compounds ["designate and confirming ions" (20)] were selected by an automated method (23) or manually. Several different libraries were developed as this process continued (27); the most recent one is available from the editorial office of this journal. Substances identified with this library are shown for a typical urine in Figure 6.

The single most important item of information in the library entry for each compound is the GC retention index, because intralaboratory precision for this datum exceeds 0.2% (23)—although it may be considerably poorer than this because of the variability in both the liquid and solid phases of

---

5 Blaisdell, B. E., and Sweeney, C. C., Determination in gas chromatography-mass spectrometry data of mass spectra free of background and neighboring substance contributions, manuscript in preparation.
Fig. 6. Organic acids from urine, identified by MSSHMET as follows.

Added hydrocarbons with 10, 11, 12, 14, 16, 18, 20, and 24 carbon atoms are identified as A–H, respectively. Substances not yet positively matched to pure reference compounds are designated "UN," with tentative identification in parentheses. The urinary acids are (1) UN1, (2) α-hydroxysebutyric, (3) lactic, (4) UN43, (5) UN4, (6) glycolic, (7) UN6, (8) β-hydroxybutyric, (9) UN7 (pyruvic oxide), (10) UN8, (11) cresol, (12) UN10, (13) UN11, (14) glycerol, (15) methylmalonic, (16) 2-methylglyceric, (17) phosphoric, (18) 4-deoxyerythronic, (19) benzolic, (20) 4-deoxythronic, (21) UN18, (22) succinic, (23) fumaron, (24) phenylacetic, (25) nicotinic, (26) UN20, (27) 2-deoxytreonic, (28) UN21, (29) UN23, (30) glutaric, (31) citramalic, (32) UN24, (33) erythronic, (34) UN27 (3-methylglutaric), (35) threonic, (36) adipic, (37) UN32, (38) 3-methyladipic, (39) α-hydroxybenzolic, (40) UN34, (41) α-hydroxyglutaric, (42) UN34, (43) UN35, (44) α-hydroxy-β-methylglutaric, (45) UN38, (46) UN37, (47) γ-hydroxybenzolic, (48) pyroglyutaric, (49) UN38 (hydroxymethylfuroic), (50) UN39, (51) UN40, (52) α-hydroxyphenylacetic, (53) pimelic, (54) topic internal standard, (55) α-ketoglutartic oxide, (56) p-hydroxybenzolic, (57) m-hydroxyphenylacetic, (58) UN42, (59) p-hydroxyphenylacetic, (60) arabonic, (61) suberic, (62) β-glycerophosphoric, (63) UN45, (64) α-glycerophosphoric, (65) UN47, (66) cis-aconitic, (67) UN49, (68) UN50, (69) citric, (70) UN52, (71) azelaic, (72) terahyphalic, (73) vanillic, (74) UN53, (75) UN54, (76) homovanillic, (77) m-hydroxyphenylacetic, (78) vanitic, (79) UN56 (80) UN57, (81) o-umaric, (82) unknown hexuronic, (83) gluconic, (85) UN58, (86) sebacic, (87) p-hydroxyphenylacetic, (88) vanilmandelic, (89) ascorbic, (90) UN61, (91) UN62, (92) UN63, (93) UN68, (94) fureric, (95) hippuric, (96) UN67, (97) UN69, (98) UN70, (99) 3-indolyacetic, (100) UN71, (101) cinnamic, (102) UN72, (103) furelic, (104) uric, (105) UN73, (106) UN74, (107) UN75, (108) m-hydroxyhippuric, (109) UN76, (110) 3,4,5-trimethoxycinnamic, (111) 3-(5-hydroxindolyl)-acetic, and (112) UN 77.

The GC packing material. Thus, for example, we have noticed that retention indices on OV-17 can vary from batch to batch by as much as 0.5%. Furthermore, we have encountered one batch of OV-17 for which retention indices varied from the other batches by as much as 1%, and we have also noticed considerable differences depending upon the solid support chosen. For these reasons, we recommend purchase of large batches of packing material and careful testing of each batch to ensure uniform retention indices.

We conclude that MSSHMET is suitable for metabolic profiling of extremely complex biological mixtures. The use of GC-MS repetitive scanning data currently appears to be satisfactory for both qualitative and quantitative analyses of complex biological fluids. Sophisticated high-resolution GC-computer systems [of the type developed by Robinson and Pauling (11, 12)] are generally cheaper and more reliable than GC-MS systems, but as yet they cannot fully resolve all components and provide quantitative analyses for each compound. Thus, the two methods are probably best seen as complementary, the capillary technology being the more suitable for screening large numbers of samples. For the most interesting samples can then be examined by use of MSSHMET-type techniques. In any case, the challenge then becomes the development of better separation procedures, testing on appropriate samples (28), the design of appropriate statistical techniques to analyze the data and, perhaps most important in the long run, the selection of appropriate clinical and biochemical questions to answer with this powerful new technology.

We gratefully acknowledge the technical assistance of J. Harten and the assistance of D. Byrne in preparing the manuscript. We also appreciate the use of a set of retention indexes loaned to us for comparison purposes by Dr. Dennis Smith (Department of Genetics, Stanford University). This work was supported in part by a research grant (RR-00480) from the Biotechnology Resources Branch, Division of Research Resources, NIH.

Table 2. Analysis of Procedural Blank

<table>
<thead>
<tr>
<th>Compound</th>
<th>Retention index</th>
<th>Relative area in blank</th>
<th>Relative area in adult urine samples</th>
<th>Mean relative area in adult urine samples ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN#</td>
<td>1011</td>
<td>7.8</td>
<td>1.2 ± 1.0</td>
<td></td>
</tr>
<tr>
<td>Lactic acid</td>
<td>1107</td>
<td>0.9</td>
<td>9.3 ± 5.5</td>
<td></td>
</tr>
<tr>
<td>UN3</td>
<td>1141</td>
<td>1.1</td>
<td>1.8 ± 1.3</td>
<td></td>
</tr>
<tr>
<td>Glycerol</td>
<td>1284</td>
<td>0.72</td>
<td>2.3 ± 0.8</td>
<td></td>
</tr>
<tr>
<td>Palmitic acid</td>
<td>2102</td>
<td>0.3</td>
<td>1.2 ± 1.1</td>
<td></td>
</tr>
<tr>
<td>UN70</td>
<td>2166</td>
<td>7.9</td>
<td>8.5 ± 5.5</td>
<td></td>
</tr>
<tr>
<td>3,4,5-Trihydroxycinnamic acid—peak 2</td>
<td>2443</td>
<td>0.9</td>
<td>0.3 ± 0.3</td>
<td></td>
</tr>
</tbody>
</table>

* UN, unknown in Fig. 6.