ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Chromatography B, 789 (2003) 283–301

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Method development in high-performance liquid chromatography for high-throughput profiling and metabonomic studies of biofluid samples

Hai Pham-Tuan, Lefteris Kaskavelis, Clare A. Daykin, Hans-Gerd Janssen*

Central Analytical Science Unit, Unilever R&D Vlaardingen, P.O. Box 114, 3130 AC Vlaardingen, The Netherlands

Received 21 October 2002; received in revised form 6 January 2003; accepted 13 January 2003

Abstract

''Metabonomics'' has in the past decade demonstrated enormous potential in furthering the understanding of, for example, disease processes, toxicological mechanisms, and biomarker discovery. The same principles can also provide a systematic and comprehensive approach to the study of food ingredient impact on consumer health. However, ''metabonomic'' methodology requires the development of rapid, advanced analytical tools to comprehensively profile biofluid metabolites within consumers. Until now, NMR spectroscopy has been used for this purpose almost exclusively. Chromatographic techniques and in particular HPLC, have not been exploited accordingly. The main drawbacks of chromatography are the long analysis time, instabilities in the sample fingerprint and the rigorous sample preparation required. This contribution addresses these problems in the quest to develop generic methods for high-throughput profiling using HPLC. After a careful optimization process, stable fingerprints of biofluid samples can be obtained using standard HPLC equipment. A method using a short monolithic column and a rapid gradient with a high flow-rate has been developed that allowed rapid and detailed profiling of larger numbers of urine samples. The method can be easily translated into a slow, shallow-gradient high-resolution method for identification of interesting peaks by LC–MS/NMR. A similar approach has been applied for cell culture media samples. Due to the much higher protein content of such samples non-porous polymer-based small particle columns yielded the best results. The study clearly shows that HPLC can be used in metabonomic fingerprinting studies.
© 2003 Elsevier Science B.V. All rights reserved.

Keywords: High-throughput profiling; Metabonomics; Biological fluids; Monolithic columns

1. Introduction

The vast amount of food ingredients with potential positive health impacts and the increasing interest in the assessment of product–consumer interactions require rapid, advanced analytical tools to comprehensively analyze measurable parameters within consumers, such as biofluid metabolites. Conventional analytical approaches that use information from single pre-identified components with known properties present a number of disadvantages which restrict their usefulness. These include: (i) laborious and often involving elaborate sample preparation, (ii) multiple procedures, each limited to the analysis of only a few compounds, (iii) artifacts such as metabolite breakdown, which contribute to variable or poor

*Corresponding author. Fax: +31-10-460-5310.
E-mail address: hans-gerd.janssen@unilever.com (H.-G. Janssen).

1570-0232/03/$ – see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S1570-0232(03)00077-1

recovery, (iv) failure to meet the required sample throughput rate and (v) inability to detect synergistic effects. Therefore, in order to further the understanding of the relationships between consumer health and food ingredients it is critical that novel analytical concepts are developed and implemented.

"Metabonomics", which has previously been defined as "the quantitative measurement of multiparametric response of living systems to pathophysiological stimuli or genetic modification" [1] provides the potential to circumvent many of these problems. By profiling biofluids, non-selective, "information-rich" metabolite profiles of complex biological samples can be obtained with minimal sample preparation. The key point of this methodology is to have in disposition a generic analytical method for rapid biofluid sample profiling together with a chemometric method for data evaluation. Whilst NMR spectroscopy is a mature technique in this field [2–4], chromatography is more or less still in its infancy, a notable exception being the work by Plumb et al. [5]. Chromatography is used abundantly in biofluid analysis [6], but almost exclusively for target component analysis and not for whole-sample fingerprinting combined with chemometrics. The high separation power of chromatographic techniques and their ability to achieve high sensitivity are strong incentives for the consideration of their use in biofluid fingerprinting as well. If successfully developed, chromatography would provide additional and complementary information to that achieved with NMR, especially for low abundance metabolites. Together, the two techniques would give a more comprehensive picture of the sample profile.

At this point it is appropriate to provide definitions of the words "target compound analysis", "screening", and "profiling". In target compound analysis, one or more known compounds are quantified, generally with a high degree of accuracy. Target compound analysis usually requires extensive sample preparation and/or the use of sensitive and selective detection and identification devices. Screening is (more or less) a faster and less accurate mode of target compound analysis. Samples are rapidly monitored (screened) for the presence of (a group of) *known* compounds. If detected, the levels of the analytes are estimated, or it is decided to perform a more reliable and accurate target compound analysis. Profiling or fingerprinting, unlike screening and target compound analysis, does not look to specific, pre-identified compounds but looks at the *entire* chromatogram as a fingerprint of the sample. In this approach, all constituents of the sample can be considered as "analytes" and the fingerprint is optimised to provide information relating to as many as possible of these constituents. Here it is also appropriate to emphasise the difference between our method of whole-sample fingerprinting and "classical" metabolite profiling or compound screening in the pharmaceutical industry. In the metabonomics approaches as we discuss here, chromatograms of two groups of individuals, e.g. control and intervention groups, generally have only subtle differences superimposed on a relatively large within-group variability. Simple strategies such as, for example, looking for new peaks that show up, or detecting compounds that resemble the molecular structure of the starting compound, are of little use here.

A limiting factor in understanding the biochemical information from HPLC profiles of biofluid samples is their complexity. Biofluid samples contain many thousands of metabolites and due to the complex nature of biofluid profiles, subtle changes in metabolites can be overlooked when examined by eye. The use of chemometric pattern recognition techniques such as, for example, principal component analysis (PCA) for interpretation of the data generated is therefore mandatory [7–10]. This approach provides an efficient, non-selective procedure for analysing biological samples and allows for the correlation of metabolic responses and health effects of "functional food" ingredients. Before chromatography can be used for high throughput sample profiling/screening, a number of practical issues inherent to any of the chromatographic techniques have to be resolved. These problems can be summarized as follows: (i) poor system stability and profile reproducibility, (ii) limited speed of analysis, (iii) limited range of compounds covered in a single analysis, (iv) interference of the sample matrix, and (v) sample preparation complexity.

In target compound analysis using liquid chromatography, the aim is to find, identify and quantify specific target compounds. As a consequence, the usual route for method development is to find the best chromatographic conditions to separate the

expected solutes from each other and, probably even more importantly, from matrix interferences. In comprehensive profiling studies, in contrast, there is no longer a clear distinction between target compounds and matrix, but rather, any component present in the sample is a potential target. This requirement has laid a tremendous weight on the separation power of the chromatographic system to be applied. The analysis time must be shortened by developing rapid LC methods [11,12]. The peak capacity can be improved via development of a comprehensive LC set-up [13]. Last but not least, the stability of the sample profiles should be improved by investing in a stable and robust LC system. This last point can then be aided by rigorous post-analysis data treatment procedures [14]. Curiously, stationary phase selectivity is not really a parameter of concern: Fingerprints on other phases are likely to look different, but are not expected to have a higher information content.

This contribution aims to prove that with a thorough and careful optimization process, HPLC can be used successfully for biofluid profiling in metabonomic studies. This will include investigations as to whether or not rapid profiling methods, using LC–DAD-FD systems, can provide stable sample fingerprints that can be used in pattern recognition processes. Data analysis will be used to identify regions in the chromatographic fingerprints where metabolic changes occur. The possibility of identifying peaks by converting the fingerprint separation to slower hyphenated methods such as, for example, LC–MS(–MS)/NMR [15] will be studied. The study will focus on urine and cell culture samples. Preliminary results with saliva and plasma samples will also be presented.

2. Experimental

Several HPLC systems, including two low-pressure mixing systems (LPM) as well as a high-pressure mixing system (HPM) have been evaluated during the development process.

The HPM system consisted of two Gilson 305 pumps (Gilson International, Den Haag, The Netherlands) equipped with analytical 5SC pump heads (50 μl stroke volume), a Gilson 805 manometric unit and a Gilson 811C dynamic mixer. The standard 1.5-ml volume of the mixing chamber was reduced to 0.46 ml by insertion of a laboratory-made PTFE plug into the void volume of the chamber.

Two low-pressure mixing systems were used. The first was a Gynkotek 480P quaternary pump (Gynkotek, Germering bei Munich, Germany). The second instrument was a HP 1100 quaternary system (Agilent Technologies, Palo Alto, CA, USA). Both systems have mixing chambers with a volume of approximately 1.0 ml.

Samples were delivered either with a carousel-type autosampler (Marathon, Spark Holland, Emmen, The Netherlands) connected to the Gilson or Gynkotek pump systems, or with the *xyz*-type autosampler G1313A integrated in the HP 1100 system.

A system of two detectors consisting of a fluorescence detector (FD) and a diode-array detector (DAD) arranged in series, was used to record sample profiles. The FD was a Jasco FP-1520 Intelligent Fluorescence detector (JASCO, Tokyo, Japan). Throughout the work, an excitation wavelength of 280 nm and an emission wavelength of 400 nm was used. The DAD was either a HP 1040M (Agilent Technologies) for the Gilson or Gynkotek pump system, or a HP G1316A integrated in the HP 1100 system. TurboChrom data acquisition software (Perkin-Elmer, Shelton, CT, USA) was used for the fluorescence signal. An HP ChemStation system (Agilent Technologies) was used for DAD control and DAD signal acquisition.

All organic solvents, i.e. methanol (MeOH) and acetonitrile (ACN), both Lichrosolv grade, as well as the mobile phase additives formic acid (HFm) and trifluoroacetic acid (TFA), were purchased from Merck (Darmstadt, Germany). Water was filtered through the Milli-Q Plus system (Millipore, Etten-Leur, The Netherlands). Solvents were degassed by helium sparging and on-line degassing. The details of the columns used are summarised in Table 1.

Prior to injection the samples were centrifuged at 3500 rpm for 3–4 min or filtered through 0.45-μm GHP membrane Acrodisk 13-mm disk filters (Gelman Sciences, MI, USA) to eliminate particulates. Furthermore, an in-line filter (0.5 μm) was incorporated in the flow path between the injector and column to protect the column from blockage.

Table 1
Columns investigated for rapid fingerprinting of urine samples

Column type	Stationary phase	Manufacturer	Length (mm)	I.D. (mm)	Particle size (μm)
Small-particle non-porous polymer-based	Chromspher C_{18}	Chrompack–Varian (Middelburg, The Netherlands)	30	4.6	1.5
Short-length porous silica-based	Chromspher ODS	Chrompack–Varian	30	4.6	5
‘‘Normal-length’’ porous silica-based	Chromspher C_{18}	Chrompack–Varian	250	4.6	5
‘‘Small-bore’’ porous silica-based	Chromspher C_{18} glass	Chrompack–Varian	100	3.0	5
Monolith	Chromolith SpeedROD RP-18e	Merck (Darmstadt, Germany)	50	4.6	–

3. Results and discussion

High throughput profiling requires the development of a rapid LC method allowing highly detailed fingerprinting of complex samples containing a wide range of compounds. The requirements for this method are as follows:

- Short analysis time and high resolution for rapid and detailed sample profiling.
- Excellent run-to-run repeatability of retention times because chromatograms are compared as fingerprints of the sample. For this reason, alternative approaches for faster analysis such as parallel analysis [16] or column switching systems with simultaneous analysis and re-equilibration [17,18] are not included in the present study. Such systems, although extremely fast, will unavoidably result in larger retention time fluctuations.
- Based on reversed-phase liquid chromatography because this technique offers the best coverage for metabolites in urine and blood (plasma) samples.
- Due to the very wide polarity and molecular mass range of the components present in the biofluid samples, rapid gradients need to be run from a very low percentage of modifier to the maximum, usually 100%. This wide range and high programming rate necessitates the use of a simple binary solvent system for stable and robust gradient operation.
- The total analysis cycle, including column re-equilibration, should be less than ca. 10 min, i.e. comparable to ^{1}H NMR-based studies. Due to the compulsory postrun re-equilibration, the actual analysis should be completed in 5 min.

3.1. Column selection

Column selection depends strongly on prior knowledge of the samples to be analyzed, i.e. the sample matrix, the range of target compounds, etc. For example, urine samples from healthy individuals contain almost no high molecular mass compounds such as proteins due to filtration via the renal pathway. The compounds of interest are therefore predominantly low molecular mass, very hydrophilic compounds. Hence, reversed-phase columns with strong retention power, can and should be used for fingerprinting purposes.

Another important criteria for the development of the generic rapid LC method is its transferability. Once ‘‘interesting regions’’ have been found in the fingerprints, it should be possible to run the method on hyphenated LC–NMR and LC–MS instruments without excessive modifications. Columns with normal inner diameters, i.e. in the range of 3–4.6 mm I.D., were used to facilitate later coupling to NMR.

The column length is another parameter to consider. A conventional 25-cm column provides an excellent separation power, but on such a column, a total cycle time of 10 min will be difficult to meet. The elution profiles of urine obtained for the shorter columns are given in Fig. 1. From this figure, it can be seen that the retention power of the non-porous, small-particle polymer-based column is too weak to

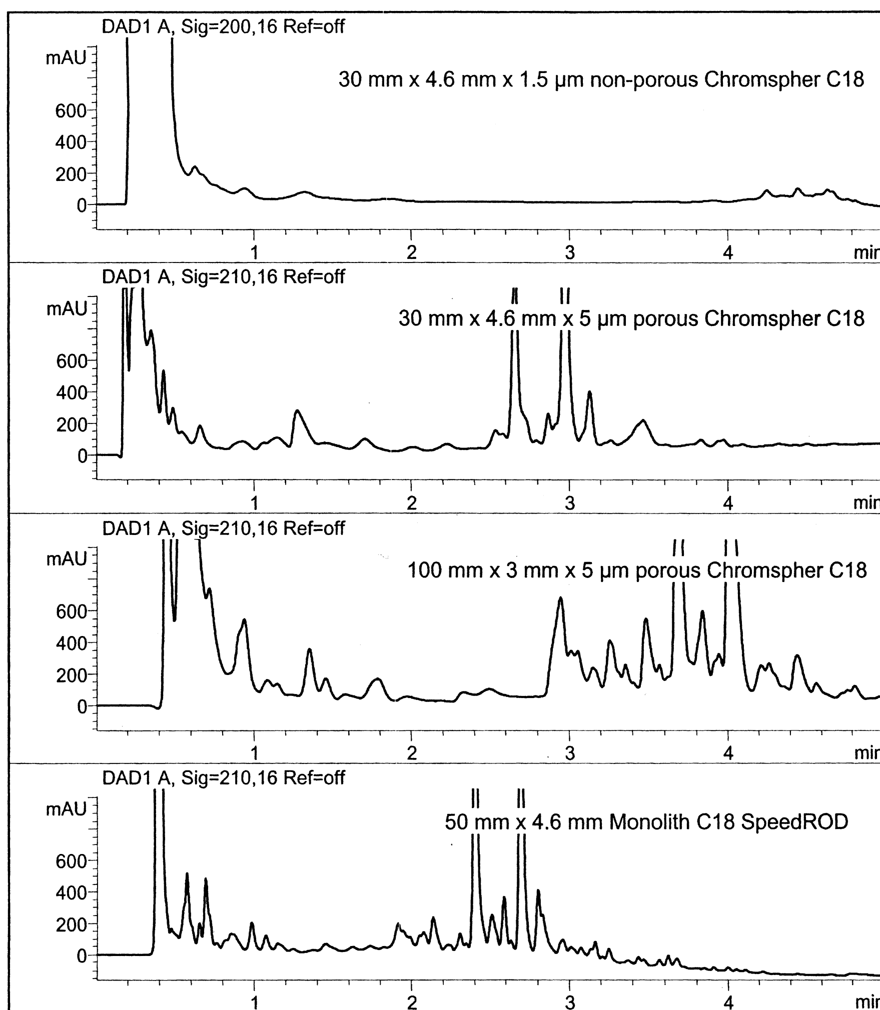

Fig. 1. Urine profile on different reversed-phase HPLC columns. Injection volume 10 µl. Solvents: A=0.1% HFm in Milli-Q water, B=0.1% HFm in MeOH. Gradient 2% B (0–0.5 min) to 90% B (4.5–5 min), flow-rate 1 ml/min.

retain very hydrophilic species. Therefore, these elute as a large peak in the void volume. Furthermore, the loadability of this column is very low due to the low surface area. Column overloading most likely is the cause for the observed band broadening for the later eluting, more hydrophobic species. Porous silica-based columns are more retentive for hydrophilic metabolites. The short 30-mm column packed with normal size (5 µm) particles gives a rather low plate number. On the other hand, when using the 100 mm, 3 mm I.D. small-bore column, elution is not complete within the 5-min time limit. The monolith column gives the most "detailed" profile for urine samples. Another advantage of the monolith column is the high flow-rate that can be used without excessive pressure drops. This allows faster post-run re-equilibration.

In contrast to urine, for samples that contain high levels of proteins, i.e. blood plasma or cell culture media, columns with weaker retention based on either non-porous particles or large pore materials have to be used to provide a complete screening profile and to avoid column clogging. Profiles for cell culture media samples obtained on the non-

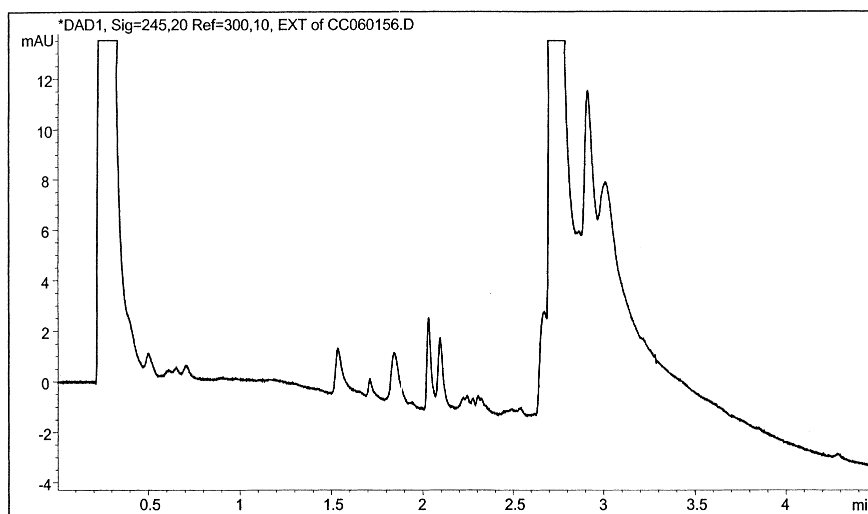

Fig. 2. HPLC profiles of cell culture media on non-porous polymer-based 30 mm×4.6 mm×1.5 μm Chromspher C_{18} column. Gradient: 5% B (0 min) to 90% B (4–5 min) at 1 ml/min, A=0.1% HFm in Milli-Q water, B=0.1% HFm in ACN.

porous, small particle column are given in Fig. 2. Albumin, one of the main constituents in the cell culture media, elutes as a large peak in the middle of the chromatogram (around 2.7 min). Very hydrophilic species are not retained on this column and elute in the void volume. Separation of these components can be improved using a more retentive porous silica-based column. Large molecular mass species, however, are completely retained even on a very short (3 cm) column after extended washing with a high organic content mobile phase [19].

3.2. Solvent selection

With regard to solvent selection, more or less similar criteria as in the selection of the column apply. For the robustness of the method and to allow future coupling with MS and NMR spectroscopy, a binary solvent system was chosen. For combination with NMR, the selection of the solvent is limited to methanol or acetonitrile as the organic modifier. Due to the fact that numerous metabolites present in urine are highly hydrophilic, a weaker organic solvent such as, for example, methanol, is preferable in aiding the analysis of urine samples (Fig. 3). On the other hand, if the sample also contains strongly retained hydrophobic compounds, e.g. blood plasma, a "stronger" solvent such as acetonitrile may provide a more comprehensive profile.

3.3. Selection of buffer system

To obtain stable and reproducible sample fingerprints, the pH of the mobile phase must be strictly controlled. In order to be readily applicable for hyphenated systems such as LC–MS and LC–NMR, volatile and less-protonated buffer systems are preferred. For this reason, 0.1% formic acid (HFm) or 0.1% trifluoroacetic acid (TFA) were used exclusively. For sample profiling using the DAD, low wavelengths are generally more informative. TFA has a distinct advantage over formic acid (HFm). Due to its lack of protons, it is more compatible with NMR spectroscopy than HFm. On the other hand, TFA is known to have strong ionization suppression effects towards several types of components in LC–ESI-MS [20]. Fortunately, post column addition of "TFA fix", i.e. 75% propionic acid in isopropanol, can be used to restore the MS signal.

3.4. Suppression of retention time shift

The main obstacle for using chromatographic data for sample fingerprinting is the fluctuation of retention times in the chromatograms. Even only minor

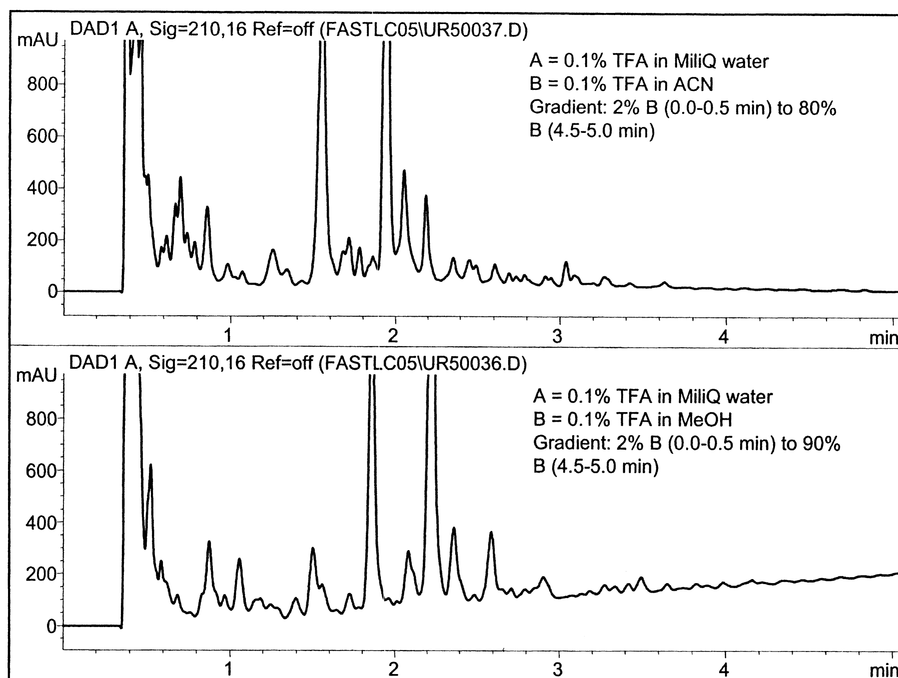

Fig. 3. Urine profiles on the monolith column with ACN (upper trace) and MeOH (lower trace) as organic eluent.

time shifts of the order of one peak width or less can affect the results dramatically. Although even a relatively inexperienced chromatographer can intuitively recognize the patterns being shifted in different chromatograms, for data-analysis tools such as principal component analysis (PCA), each component should elute at exactly the same retention time in order to minimize the inherent variation in the data set. A data pre-treatment step can adjust the variation in retention time to a certain extent, but the "physical" retention time has to be as accurate as possible.

The main causes for the irreproducibility of the retention times are: instability in the mobile phase composition or pH, fluctuation of the column temperature, column aging and/or contamination, and inadequate column post-run re-equilibration. Fluctuation of the temperature can be reduced with the use of a column thermostat. The pH of the eluent can be kept constant by adding the buffer to both the A and B solvent reservoirs. Because the pH regulators used (TFA or HFm) also have strong ion-pairing characteristics, it is very important to have them at equal percentages in both solvents to avoid disrupting the equilibration of the stationary phase surface during rapid gradient programming. This also helps to minimize the baseline noise.

The most important parameter to control if retention time shifts have to be minimized is solvent mixing during the gradient programming. The use of a pre-mixed starting solvent is recommended to ensure a homogeneous solvent mixture and a properly equilibrated column at the start of the gradient. The pump system is the main factor in providing a stable and robust gradient. Comparing five consecutive urine injections on the low-pressure mixing Gynkotek M480 and high-pressure mixing Gilson 305 (Fig. 4), it can clearly be seen that the high-pressure mixing system provides a more stable gradient. In the low-pressure mixing system, the frequency of the proportional valve determines the composition of the solvent mixture in the mixing chamber. At very steep gradients, the required valve switching frequencies will exceed the maximum switching frequency of the proportioning valve. Moreover, air bubble formation is less likely to occur in LC systems employing high-pressure mixing.

Fig. 4 (insert) also shows that even with the

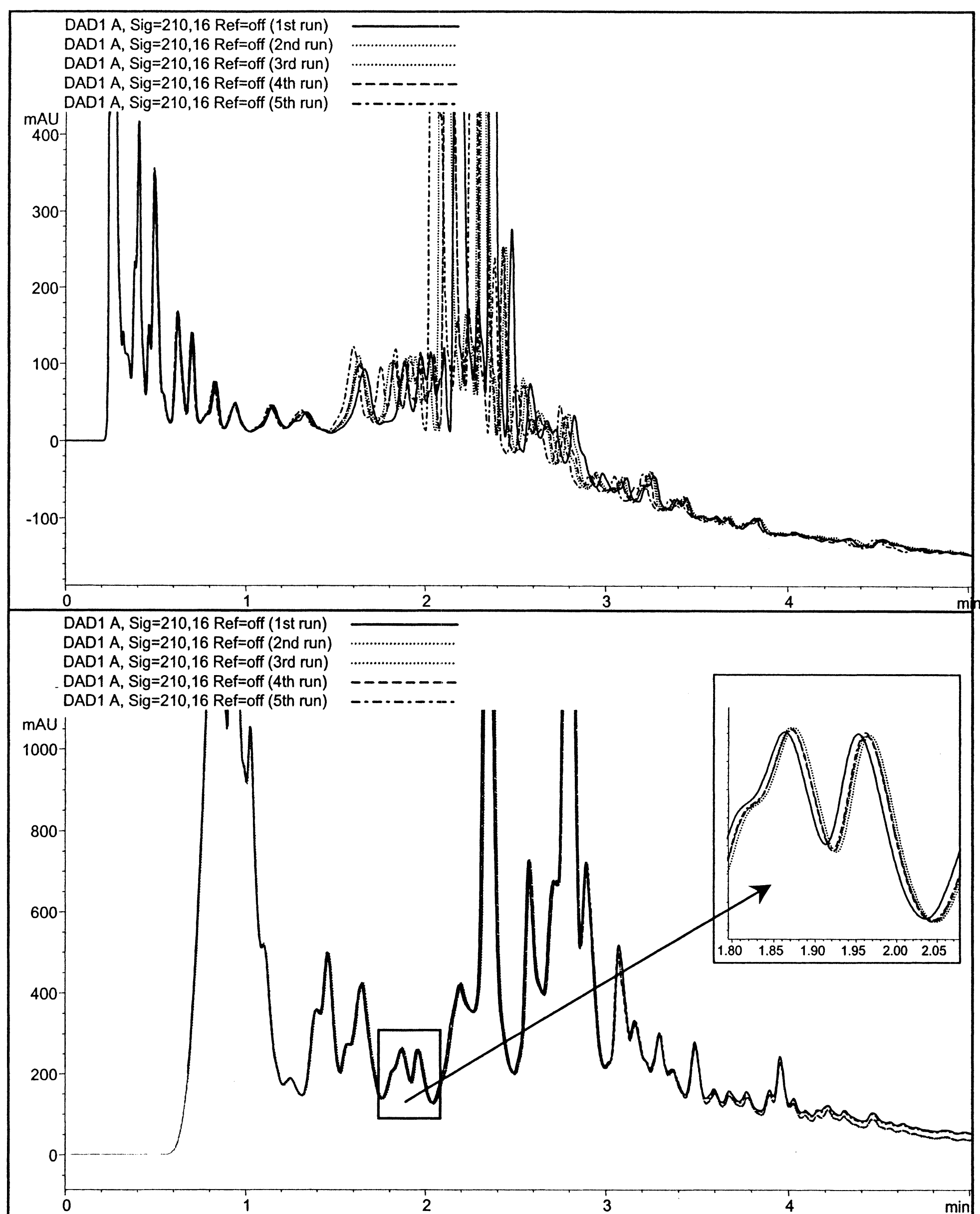

Fig. 4. Five consecutive injections of a urine sample on the LPM pump system Gynkotek M480 (upper trace) and on the HPM pump system Gilson 305 (lower trace). Column: 50 mm×4.6 mm RP-18 monolith SpeedROD (Merck); gradient: 0% B (0 min) to 90% B (4–5 min); A=0.1% TFA in MeOH/Milli-Q water (5:95 w/w); B=0.1% TFA in MeOH, flow-rate 2 ml/min (upper trace) and 1 ml/min (lower trace), respectively.

high-pressure mixing system, the very first run (solid line) gives a noticeably shifted profile compared to subsequent runs. This is most probably caused by the fact that the column was at a different equilibrated state when the first run of the sequence was started. The time that the column is kept under the starting conditions of the gradient is different. This observation leads to the conclusion that for a sample

sequence run, it is necessary to keep the time interval between the runs, i.e. total run time including post-run re-equilibration and the injection time for the subsequent run, perfectly constant. On the carousel-type autosamplers such as the Marathon, this requirement can only be met when the washing steps between the injections are eliminated. The *xyz*-type autosampler can also be used with the ''overlapping injection'' function, i.e. the sample is drawn into the loop or injection needle, ready to be injected before the post-run time of the previous run has elapsed.

Another advantage of high-pressure mixing systems is that efficient solvent mixing can be obtained at a much smaller volume of the mixing chamber. This reduces the delay time of the gradient. Moreover, the true gradient profile will show a better agreement with the set gradient. To minimize the volume effect, the standard 1.5-ml mixing chamber of the Gilson 811C dynamic mixer was partially filled with a laboratory-made PTFE plug giving an effective delay volume of 0.46 ml. This modification improved the gradient profile significantly (Fig. 5).

Another parameter that requires attention is the programming range of the organic modifier. The wide range of component polarities would require the gradient to be run from the lowest possible percentage of organic modifier to the highest value. A steep gradient would, however, reduce the separation power of the system. Moreover, for normal reversed-phase columns, the use of a 100% water mobile phase should be discouraged to avoid phase collapsing or slow re-equilibration. This problem will be discussed in detail in Section 3.5.

In conclusion, the following precautions are recommended:

- Use column thermostat;
- Use pre-mixed solvents;
- Avoid 100% water mobile phase;
- Use high-pressure mixing with a low mixing chamber volume;
- Maximize post-run column equilibration volume using higher a flow-rate; and
- Ensure a precise time interval between injections.

3.5. Sequence run arrangement

As stated several times throughout this article, the most important feature of a rapid HPLC profiling

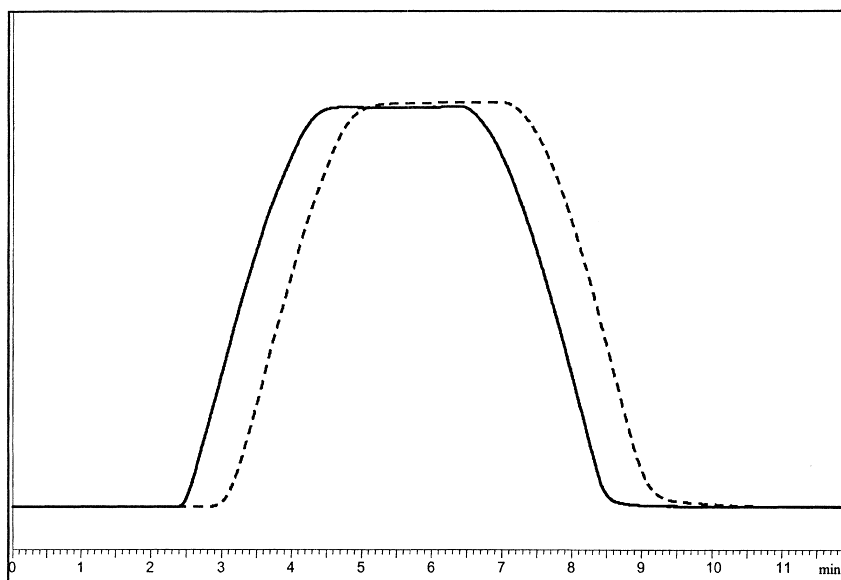

Fig. 5. Gradient profile of the Gilson 305 pump system with standard 1.5-ml mixing chamber (dotted line) and modified 0.46-ml mixing chamber (solid line). Solvents: A=Milli-Q water; B=0.1% acetone in MeOH. Gradient: 0% B (0–2 min) to 100% B (4–6 min) to 0% B (8 min) at 2 ml/min. UV setting: 254 nm.

method is the stability of the sample profile during the entire batch of samples. During an analysis sequence, the system performance can deteriorate due to a number of reasons such as for example mechanical failure, leakage, blockage (injector, flow path, tubing, column), column aging, etc. The most sensitive part of the HPLC system in this respect is the column itself. The column can lose its resolution due to stationary phase degradation and/or accumulation of impurities. An example of column degeneration can be seen in Fig. 6. Most likely, the C_{18} chains of the stationary phase collapsed due to the low percentage of organic solvent, 2% MeOH in this case, in the initial part of the gradient. Normal C_{18} stationary phases can tolerate a few short runs at low organic modifier content (less than 5%), but the column eventually degenerates when a large volume of a high water content solvent has passed through. New types of reversed-phase HPLC columns that tolerate prolonged use with low percentages of modifier have recently become commercially available. They are, however, not fully compatible with the strong acids (TFA and HFm), selected to be used as mobile phase pH regulators in this experiment. Under these conditions, the stationary phase selectivity changed gradually during prolonged operation with a 100% water mobile phase (Fig. 7).

Another possible cause of column aging is build-up of residual proteins/contaminants present in the urine samples on the surface of the stationary phase. If the post-run washing period is not long enough or does not go to a high enough organic content, band broadening occurs due to surface contamination.

Any time when a disruption in the system performance occurs, a certain number of samples have to be reanalyzed. Under optimum conditions, a HPLC column can typically handle up to 1000 samples. Care should be taken so that the total number of analyses for any study, i.e. number of rapid profiling runs, number of eventual re-runs and number of identification runs, does not exceed this limit. This is to avoid possible batch-to-batch variability in column performance.

The system stability and the total number of analyses can be closely controlled, and if necessary corrected, by incorporating reference sample runs into the sequence at regular intervals between actual sample runs. The reference runs are carried out using one sample, the reference sample. Differences in the profiles from the reference runs can be used to

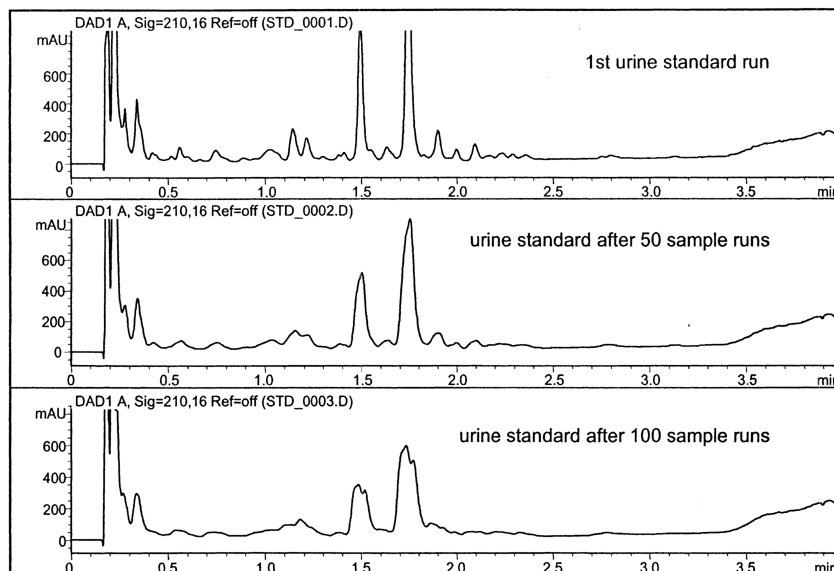

Fig. 6. Column performance degradation during rapid HPLC profiling of urine samples. Gradient: 0% B (0–0.5 min) to 40% B (3.5–4 min) at 5 ml/min, post-run time 1 min; A=0.1% TFA in Milli-Q water, B=0.1% TFA in MeOH.

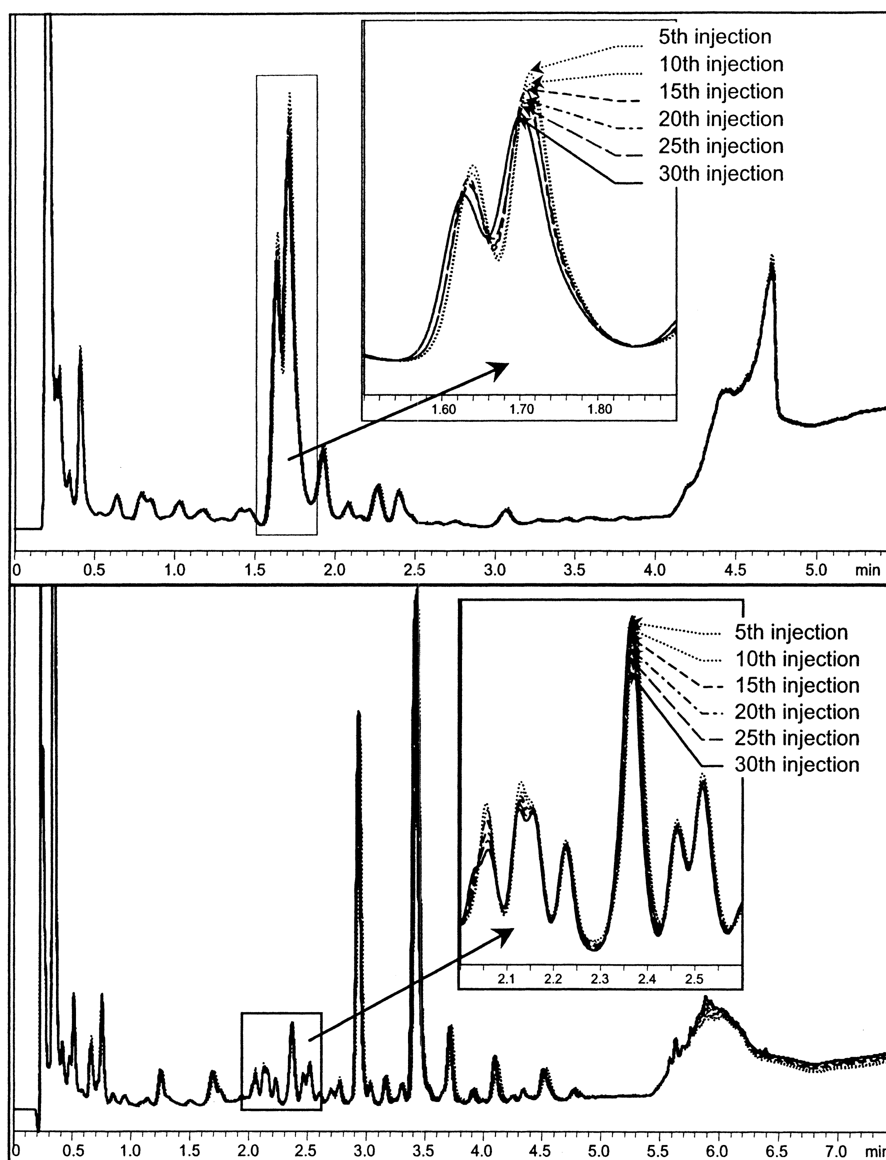

Fig. 7. Repetitive injections of a standard urine sample on the 50 mm×4 mm×5 μm Discovery Amid C_{16} column (upper traces) and 50 mm×4.6 mm×3 μm Polarity dC_{18} column (lower traces). Gradients: 0% B (0–0.62 min) to 30% B (3.75 min) to 100% B (4.37–5.62 min) at 4 ml/min, postrun time: 2.5 min for the Discovery column and 0% B (0–0.83 min) to 30% B (5 min) to 100% B (5.83–7.5 min) at 3 ml/min, postrun time 2.5 min. Solvents A=0.1% TFA in Milli-Q water, B=0.1% TFA in MeOH. UV wavelength: 210 nm.

correct retention time shifts in the chromatograms and also to obtain a general impression about the accuracy of the measurements. Occasionally, the difference between successive reference runs can be unacceptable. In such a case, the runs between the last acceptable reference run and the failed reference run need to be repeated.

How often the reference and the blank have to be run is a careful balance between the risk of having to re-run a large number of samples and the total

number of analyses possible on a single column. The detailed mathematical formulation of this problem will be reported in a separate publication [21]. As a rule of thumb, a reference run for every 15–20 sample runs will be the most economical choice.

3.6. Method transferability

A generic rapid HPLC profiling method has to be easily transferable to a ''slow'', high resolution method for identification of the compounds responsible for differences in the fingerprints. In the rapid screening method, resolution is sacrificed for speed. In a metabonomic, dietary intervention study with human volunteers, a large number of biofluid samples are profiled and the data evaluated to examine the effect of the intervention on the volunteers. Regions of the chromatograms that contain information relating to metabolic changes in the volunteers can then be observed via study of so-called PCA ''loadings'' plots. A logical question then arises: which components are responsible for the effect? A subset of typical samples that clearly show the metabolic effect of interest will then be selected and re-analyzed with hyphenated techniques such as LC–MS, LC–NMR and/or LC–NMR-MS [22–24] for identification. It is, therefore, very important that the rapid-profiling method can be easily converted to a slow, shallow-gradient high-resolution HPLC–MS/NMR method. Speed of analysis no longer has the highest priority. Now the resolution of the chromatographic run becomes the critical factor.

Hyphenated techniques, such as for example LC–MS/NMR, operate under rather strict conditions compared to universal HPLC–DAD/FD runs. As an example, both NMR and MS detectors have strict requirements in terms of the solvents and buffers to be used. Moreover, unlike MS which is very compatible with rapid gradients [25], NMR spectrometers prefer low LC flow-rates and shallow gradients due to the rather low sensitivity and low scan speed.

A fast gradient method can be translated straightforwardly into a slower, higher resolution method by applying a shallower gradient and a proportionally reduced flow-rate. Here the flow-rate was kept constant for both methods at 1 ml/min to avoid arriving at a sub-optimal velocity. Due to the high back-pressure of the column, the rapid gradient method could not be run at a higher linear velocity. This strategy was applied for the profiling and identification of cell culture media samples using the non-porous small particle column (Fig. 8). Although the overall profile can be easily ''translated'', a few peaks are seen to elute at different relative positions. This is most likely due to the slightly different composition of the starting eluent in the two methods. A lower organic solvent percentage, 2% instead of 5%, was used in the slow gradient method in order to increase the resolution of hydrophilic components.

The monolithic column, which has been selected for the urine samples, on the other hand, tolerates higher flow-rates without excessive back-pressures. Here, a rapid profiling method can be created by combining high flow-rates with gradients adjusted accordingly. The elution order of the peaks will then be preserved. Resolution, however, is generally reduced at higher flow-rates. Van Deemter curves of components with different molecular masses and polarities are shown in Fig. 9. From this figure, it can be seen that for hydrophilic compounds, the optimal condition lies at lower flow-rates while it moves towards higher values for more hydrophobic compounds. This discrepancy with chromatographic theory is most likely caused by measurement difficulties where peaks show tailing. For the hydrophilic components, the column efficiency is not reduced significantly when the flow-rate is increased from 1 to 5 ml/min, i.e. less than a factor of two in the case of hippuric acid. This is acceptable for a gain in analysis speed of a factor of five.

Fig. 10 shows a typical profile of urine on a monolithic column under different flow-rates with correspondingly adjusted gradients. When the chromatograms are aligned, it can be seen that the elution order of the chromatograms is preserved (Fig. 11).

The preservation of elution order during the transfer from a rapid profiling method to a slow HPLC–DAD-MS identification method proved to be very useful. An example can be found in a study of the effect of tea consumption [26]. PCA of rapid profiling of over 500 urine samples on the fluorescence detector shows a clear effect of the tea intervention (Fig. 12). The loading plot of the PC analysis represents the level of contribution of each peak in the chromatogram to the overall effect seen

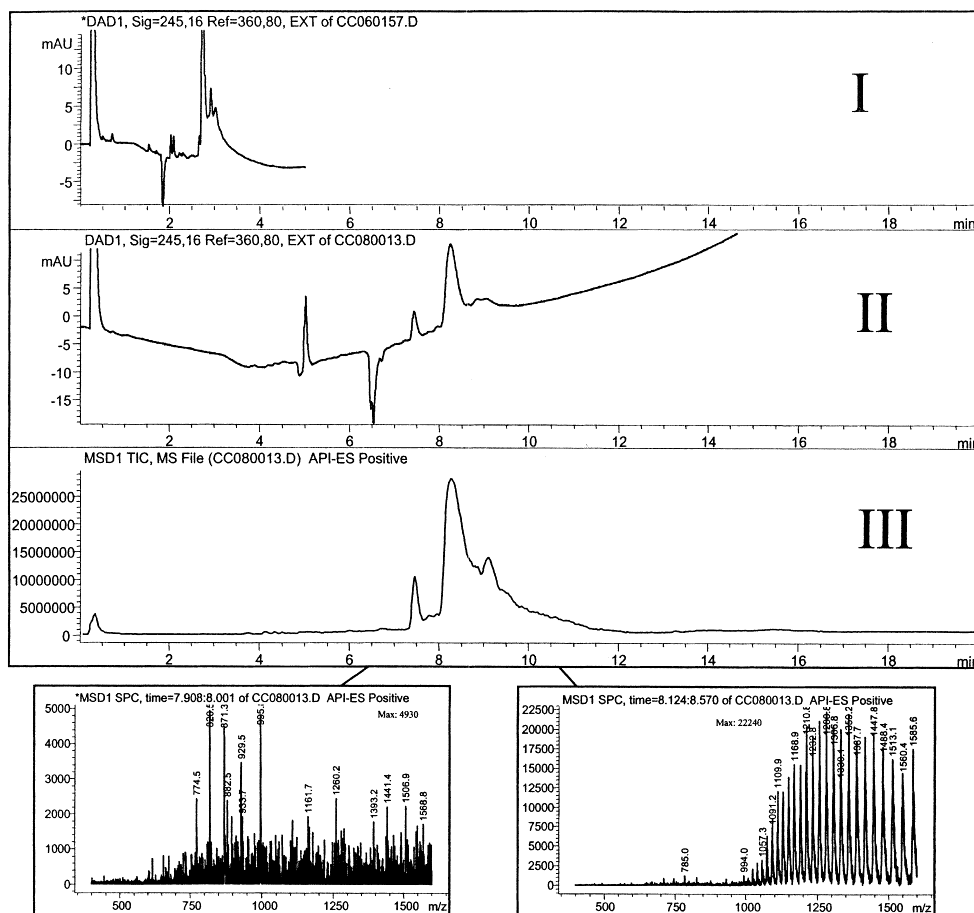

Fig. 8. "Translation" of a rapid HPLC–DAD method (I) for profiling to a slow HPLC–DAD (II)–MS (III) method for identification. Inserts: MS spectra collected at the indicated times; left: a protein with molecular mass of 13 946±20 Da; right: albumin with a molecular mass of ca. 60 000 Da. Cell culture media samples on non-porous polymer based 30 mm×4.6 mm×1.5 μm Chromspher C_{18} column. Solvents: A=0.1% HFm in Milli-Q water, B=0.1% HFm in ACN. Gradient: I—5% B (0 min) to 90% B (4–5 min); II, III—2% B (0–2 min) to 90% B (15–20 min).

in the score plot. A high peak in the loading plot indicates that the corresponding peak in the LC sample profile has a larger contribution to the overall effect (Fig. 13). LC–MS experiments are then carried out to identify the peaks eluting in that particular region of the chromatogram. This is preferably done in a "higher chromatographic resolution mode" to facilitate unambiguous peak identification". LC–MS runs are therefore carried out at a lower flow-rate and a shallower gradient. To "relocate" the peaks of interest in the TIC chromatogram, the short, rapid DAD-FD chromatogram has to be "stretched-out" to align with the long, slow DAD-MS chromatogram. Because the FD and MS chromatograms do not have the same appearance, direct peak assignment from the FD trace to MS trace is not always possible. A cross-reference or a "bridge" chromatogram is needed. Peak alignment (Fig. 13) from FD to MSD is possible using the DAD signal, which is present on both systems. When the peak in question is identified, selected ion chromatograms can be extracted and PCA is carried out again to confirm the effect. For complete structure elucidation, however, more elaborate work is needed. This might involve multi-step sample clean-up, separation to isolate and concentrate the com-

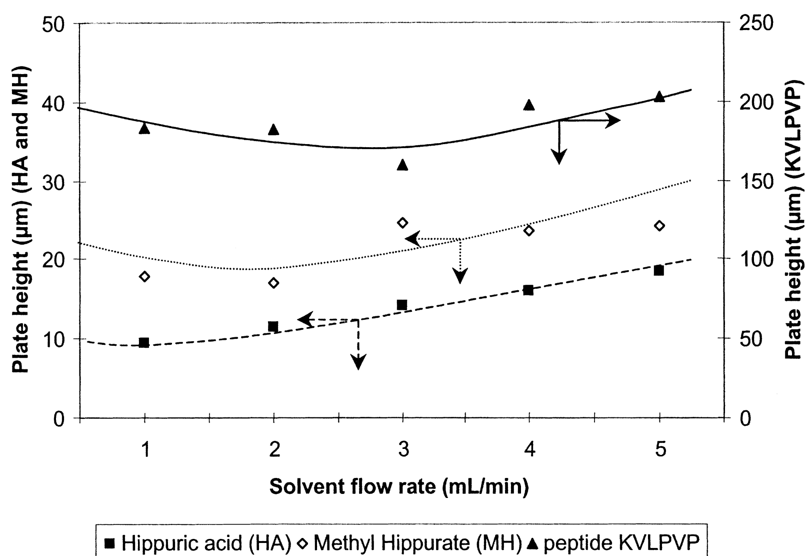

Fig. 9. Van Deemter curves for different components on a monolithic column (50×4.6 mm RP-18 SpeedROD Chromlith, Merck). Solvent: isocratic 0.1% TFA in MeOH/Milli-Q water (5:95 w/w), UV detector 210 nm, 16-nm bandwidth without reference.

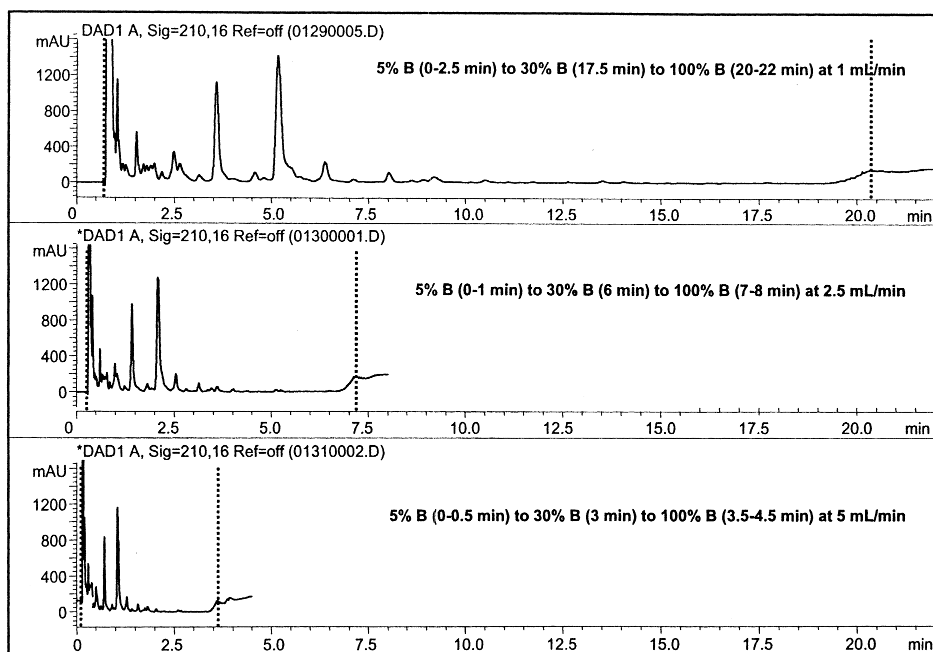

Fig. 10. Profile of a standard urine sample on the monolithic column at different flow-rates and correspondingly adjusted gradients. Non-aligned chromatograms.

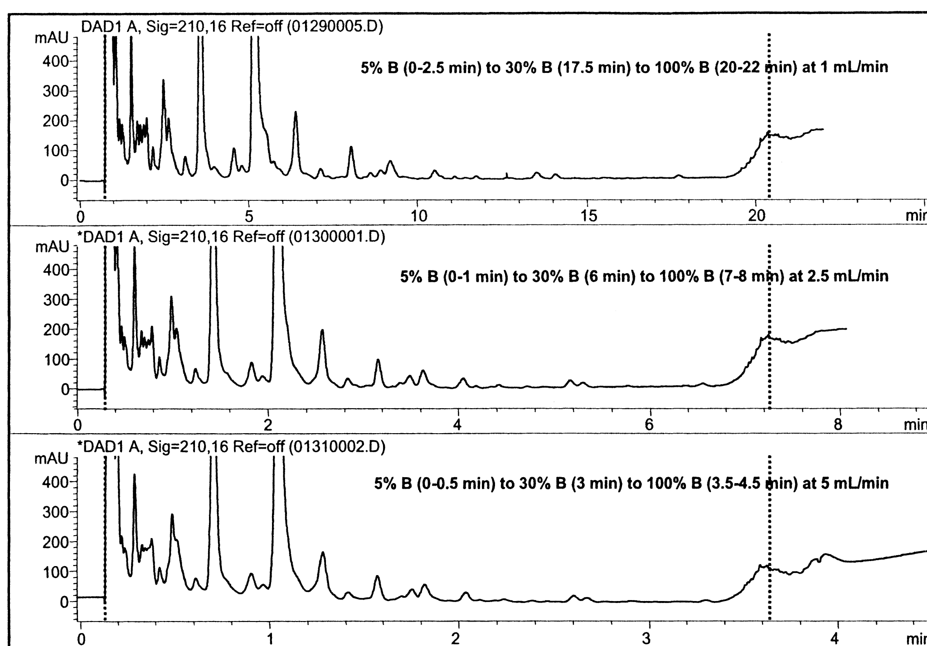

Fig. 11. Profile of a standard urine sample on the monolithic column at different flow-rates and correspondingly adjusted gradients. Aligned chromatograms.

ponents of interest and analysis using MS–MS and NMR.

3.7. Rapid profiling of serum and saliva

So far, the monolith column has proven to be very useful for rapid profiling of urine samples. Because of the high retention power for the low molecular mass metabolites, which are the most relevant for the metabonomic studies, it is of interest to evaluate its performance on other types of biofluid samples, i.e. serum and saliva.

Serum or blood plasma samples generally contain large amounts of proteins, which impose a great danger for reversed-phase columns because of protein precipitation at higher organic content of the mobile phase. Serum or plasma proteins, therefore, have to be removed beforehand. An effective and rapid technique for protein removal is organic solvent precipitation [27].

Fig. 14 shows the rapid profiling of porcine serum on the monolithic column after protein removal with ACN (ratio sample/ACN=1:1) and MeOH (ratio sample/MeOH=1:2). The chromatograms are reconstructed with baseline subtraction from a blank Milli-Q water injection. The profiles are quite different depending on the organic solvent used for protein precipitation, although several common peaks can be seen. To minimize deleterious effects of residual proteins present in the samples, ACN was selected as organic modifier. A combination of high flow-rate and high ACN content was used during the washing period of the gradient.

A profile of human saliva was recorded without protein precipitation (Fig. 15). The sample was filtered through a 0.45-μm Acrodisk membrane filter. A similar instrument as for the porcine serum experiment was used. This is, however, only the profile of a single sample. The method has to be evaluated on larger numbers of samples to establish the stability of the profile. The accumulation of proteins from the saliva is expected to impose a great danger to the column performance.

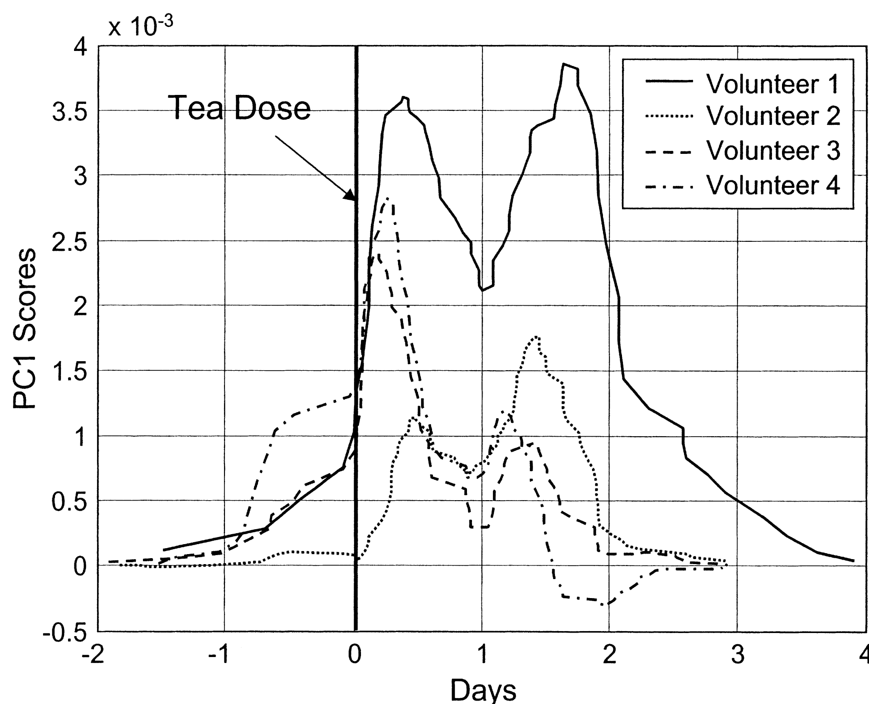

Fig. 12. Second derivative of the PC score plot of urine sample profiles (fluorescence detector signal) denoting the dynamic effect of the tea intervention for different volunteers. The *x*-axis shows the times urine samples were taken, relative to the administration of tea doses.

Preliminary results shown above clearly proved the applicability of the monolithic column in rapid profiling of biofluid samples with higher levels of proteins. A combination of high flow-rates and rapid gradients provides a detailed profile of low molecular mass compounds in these samples and avoids the accumulation of residual high molecular mass species on the column.

4. Conclusions

HPLC has been successfully used for fingerprinting analysis of biofluid samples. Generic methods for rapid profiling of biofluid samples have been developed. These can be successfully run on standard equipment if proper precautions are taken. For samples with a low protein-content such as urine and deproteinised serum/plasma, the optimum method uses a short-length monolithic C_{18} reversed-phase column (50×4.6 mm RP-18 SpeedROD Chromolith, Merck). High speed profiling is guaranteed by a combination of high flow-rates and rapid gradients. This high flow-rate, rapid gradient method is easily converted to a high-resolution slow-gradient method such as, for example, LC–MS/NMR for identification purposes. The method developed has been applied successfully for fingerprinting large numbers of urine samples from different studies.

For samples containing high levels of proteins such as cell culture media samples, a method based on the use of a stationary phase of weaker, non-porous small polymer particles (30×4.6 mm×1.5 μm C_{18} Chromspher, Varian-Chrompack) was developed. Some 40 samples were analysed and stable profiles were obtained. Different compounds, low molecular mass as well as proteins, have been identified as being responsible for the differences in sample fingerprints.

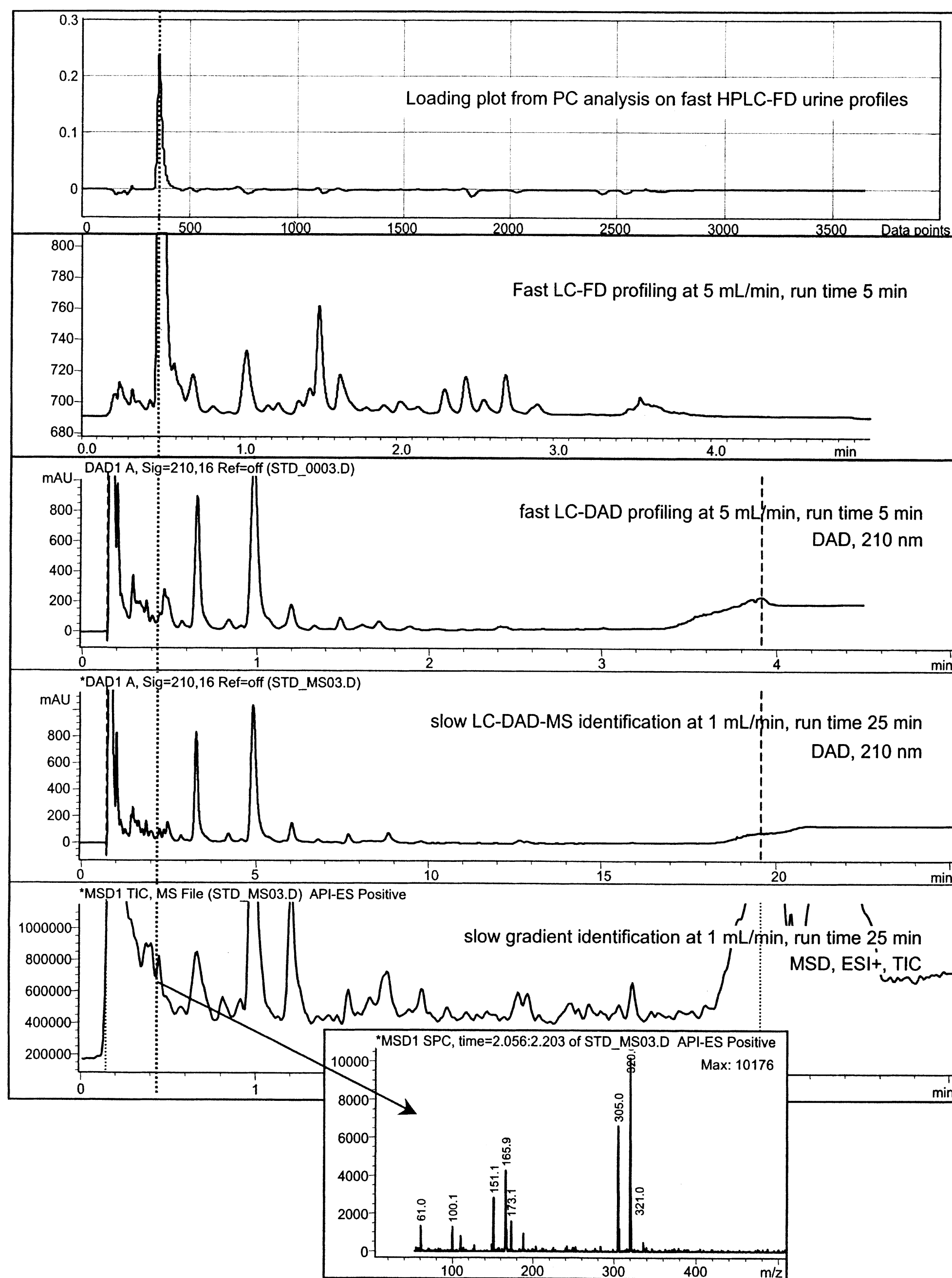

Fig. 13. "Translation" from rapid HPLC–DAD-FD profiling method to slow HPLC–DAD-MS identification method.

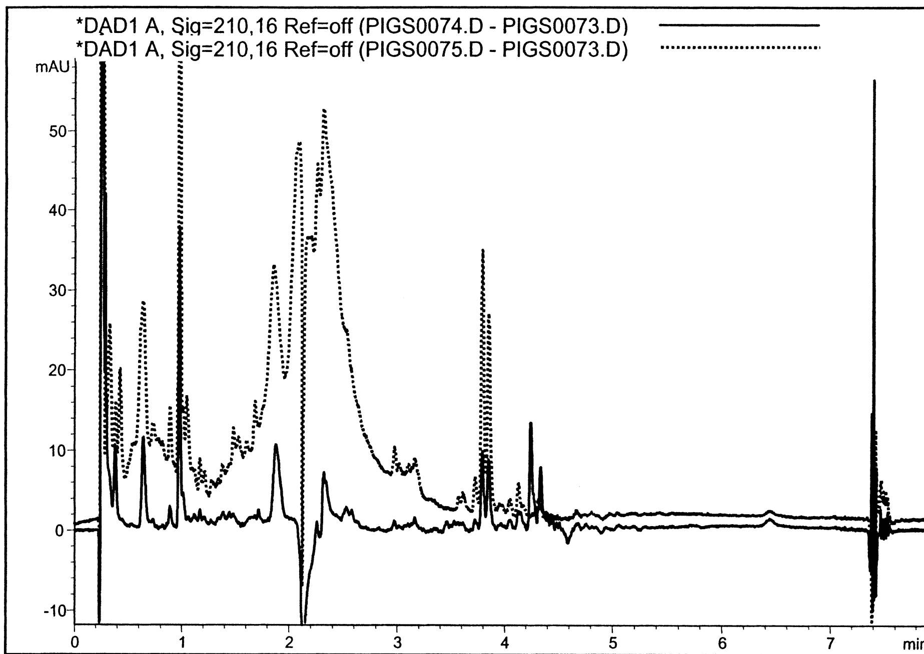

Fig. 14. Porcine serum profile on the monolithic column after protein precipitation with MeOH (ratio 1:2, solid trace) and ACN (ratio 1:1, dotted trace). Column: 50 mm×4.6 mm RP-18 SpeedROD Chromolith, Merck. Solvents: A=0.1% TFA in ACN/Milli-Q water (5:95 w/w), B=0.1% TFA in ACN. Gradient: 0% B (0 min) to 100% B (4–7 min) to 0% B (7.1 min); flow program: 3 to 5 ml/min (5.1–9 min) to 3 ml/min (9.1 min). Injection volume 5 μl.

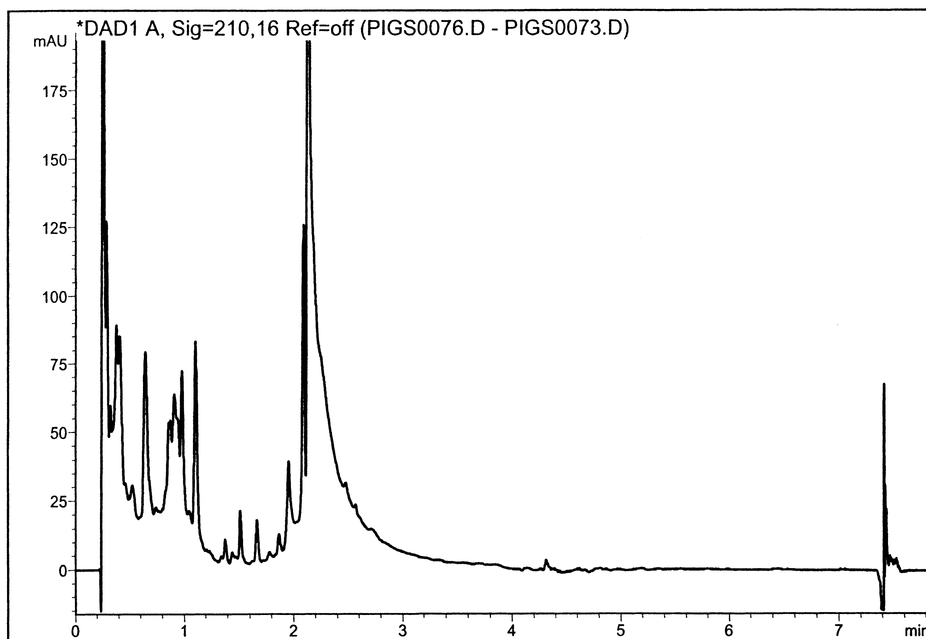

Fig. 15. Human saliva profile on the monolithic column without protein precipitation. Chromatographic conditions: see Fig. 14.

References

[1] J.K. Nicholson, J.C. Lindon, E. Holmes, Xenobiotica 29 (1999) 1181.
[2] T.W.M. Fan, Prog. NMR Spectrosc. 28 (1996) 161.
[3] J.C. Lindon, J.K. Nicholson, J.R. Everett, Annu. Rep. NMR Spectrosc. 38 (1999) 1.
[4] J.K. Nicholson, I.D. Wilson, Prog. NMR Spectrosc. 21 (1989) 449.
[5] R.S. Plumb, C.L. Stumpf, M.V. Gorenstein, J.M. Castro-Perez, G.J. Dear, M. Anthony, B.C. Sweatman, S.C. Connor, J.N. Haselden, Rapid Commun. Mass Spectrom. 16 (2002) 1991.
[6] G.B. Scarfe, J.C. Lindon, J.K. Nicholson, P. Martin, B. Wright, S. Taylor, E. Lenz, I.D. Wilson, Xenobiotica 30 (2000) 717.
[7] L. Eriksson, E. Johansson, N. Kettaneh-Wold, S. Wold, Introduction to multi and megavariate data analysis using projection methods (PCA & PLS), Umetrics, 1999.
[8] K.P.R. Gartland, C.R. Beddell, J.C. Lindon, J.K. Nicholson, Mol. Pharmacol. 39 (1991) 629.
[9] J.C. Lindon, E. Holmes, J.K. Nicholson, Prog. NMR Spectrosc. 39 (2001) 1.
[10] T.M. Alam, K. Alam, Spectroscopy 16 (2001) 18.
[11] J. Zweigenbaum, J. Henion, Anal. Chem. 72 (2000) 2446.
[12] J.F. Banks, E.E. Guiciek, Anal. Chem. 69 (1997) 3973.
[13] D.B. Wall, M.T. Kachman, S. Gong, R. Hinderer, S. Parus, D.E. Misek, S.M. Hanash, D.M. Lubman, Anal. Chem. 72 (2000) 1099.
[14] D. Bylund, R. Danielsson, G. Malmquist, K.E. Markides, J. Chromatogr. A 961 (2002) 237.
[15] J.C. Lindon, J.K. Nicholson, I.D. Wilson, J. Chromatogr. B 748 (2000) 233.
[16] M. Jemal, M. Huang, Y. Mao, D. Whigan, M.L. Powell, Rapid Commun. Mass Spectrom. 15 (2001) 994.
[17] S.R. Needham, LCGC North Am. 18 (2000) 1156.
[18] J.S. Janiszewski, K.J. Rogers, K.M. Whalen, M.J. Cole, T.E. Liston, E. Duchoslav, F.G. Fouda, Anal. Chem. 73 (2001) 1495.
[19] G. Rozing, M. Serwe, H. Weissgerber, B. Glatz, LCGC North Am. 19 (2001) 26.
[20] F.E. Kuhlmann, A. Apffel, S.M. Fischer, G. Goldberg, P.C. Goodley, J. Am. Soc. Mass Spectrom. 6 (1995) 1221.
[21] L. Kaskavelis, H. Pham-Tuan, Optimization of the total number of experimental runs, J. Chromatogr. A, in preparation.
[22] J.C. Lindon, J.K. Nicholson, I.D. Wilson, in: K. Albert (Ed.), On-line LC–NMR and Related Techniques, Wiley, Chichester, UK, 2002.
[23] M. Dachtler, T. Glaser, K. Kohler, K. Albert, Anal. Chem. 73 (2001) 667.
[24] M. Sandvoss, A. Weltring, A. Preiss, K. Levsen, J. Chromatogr. A 917 (2001) 75.
[25] J.M. Lazar, E.D. Lee, J.C.H. Sin, A.L. Rockwood, K.D. Onuska, M.L. Lee, LCGC North Am. 18 (2000) 110.
[26] L. Kaskavelis, H. Pham-Tuan, C.A. Daykin, Modelling the dynamic effect of tea in the human body using a novel metabonomic study design—An exploratory study, Chemometr. Intell. Lab. Syst., submitted.
[27] C.A. Daykin, P.J.D. Foxall, S.C. Connor, J.C. Lindon, J.K. Nicholson, Anal. Biochem. 304 (2002) 220.