A Rapid Simple Approach to Screening Pharmaceutical Products Using Ultra-Performance LC Coupled to Time-of-Flight Mass Spectrometry and Pattern Recognition

Robert S. Plumb[1,*], Michael D. Jones[2], Paul D. Rainville[2], and **Jeremy K. Nicholson[1]**

[1]Imperial College, Faculty of Medicine, Sir Alexander Fleming Building, South Kensington, London, UK, SW7 and [2]Waters Corporation, 34 Maple Street, Milford, MA, 01757

Abstract

The comparison of batches of pharmaceutical product or raw active pharmaceutical ingredients (API) for product release can be time consuming and tedious process. It often requires long analysis times and potentially several liquid chromatography–tandem mass spectrometry (LC–MS–MS) analytical runs to determine the identity of the impurities and their relationship to the active pharmaceutical ingredient. The combination of a high resolution (sub 2 µm porous particle) LC coupled to exact mass MS, principal components analysis (PCA) allowed for the rapid classification of batches of Simvastatin tablets according to their impurity profile. Evaluating the ultra-performance LC–MS exact mass data with PCA allowed for the impurities of Simvastatin to be easily detected and identified. This approach to impurity batch analysis should be applicable to many other forms of batch analysis, fermentation broths, food production, and API manufacturing.

Introduction

The production of pharmaceutical bulk raw materials and finished product is a high value, capital-intense process. With batches of material valued at hundreds of thousands of dollars, any delay in the delivery of raw materials to production, and of pharmaceutical medicine to distributors can be extremely costly. Hence, the ability to accurately measure the impurity profile of an active pharmaceutical ingredients (API) or product is critical to the manufacturing process. It is a regulatory requirement for a pharmaceutical manufacturer to have a specific, accurate, reliable assay for the acceptance of raw materials and the testing of a finished product (1). This is often achieved by the use of techniques such as high-performance liquid chromatography (HPLC) with UV detection. These assays are often slow and monitor only the known impurities/degradents. The impurity profile of a finished product is regularly used to check for counterfeit products, this is often achieved by the combination of liquid chromatography with mass spectrometry (MS), especially MS–MS (2). The small, minor impurities present in the sample are diagnostic of the route of synthesis employed in the manufacture of the API and thus can be used to monitor for likeness to the innovator's process and identify illegally copies of valuable products. The data reduction and analysis process is normally achieved by employing time consuming manual analysis of the data and comparison to authentic standards, which requires the synthesis of these standards. Therefore the comparison of samples from different batches of production requires a significant amount of manual analysis, tabulation, and quantitation and is limited by the fact that it only addresses known impurities (3).

Reversed-phase LC with UV or photodiode array detection has become the technique of choice for this operation, due to its compatibility with the samples, resolution, specificity, and sensitivity. The need to comprehensively separate all of the impurities in the sample often results in relatively long analysis times, typically 30–50 min (3). This is due to the moderate resolution developed by the traditional 3.5 and 5 µm particles used in the separation process. More recently, the introduction of sub-2 µm porous LC packing materials (4) has allowed for extremely high resolution chromatograms to be generated in just a few minutes, allowing analysis times for impurity analysis to be significantly reduced (5). The extra resolution and sensitivity of these sub-2 µm chromatography particles has attracted interest of researchers faced with the analysis of complex samples. Nielen et al. (6) employed the combination of ultra-performance LC (UPLC) and hybrid quadrupole TOF MS (Q-TOF) for the detection of designer steroids in urine samples; Wren and Tchelitcheff (7) employed these small particles with MS detection for the detection of a series of beta blockers, reducing the analysis time from 10 min to just 3.5 min; and Haynes et al. showed how analysis times and sensitivity could be significantly improved in bioanalysis LC–MS–MS by utilizing the enhanced chromatographic performance of the sub-2 µm chromatography (8).

The rapid evaluation of batches for product is essential to the timely release of product. The acceptance or rejection of batches relies on the comparison of the batch under test with a known set of parameters or acceptance criteria. The evaluation of pharmaceutical product is usually achieved by the comparison of the batch under test with a control standard against a standard oper-

* Author to whom correspondence should be addressed: email R.Plumb@ic.ac.uk.

ating procedure driven acceptance level for impurities. This approach, however, is limited to the detection of components that are visible by the current detection methodology, and thus the presence of a new impurity that is not visible, by UV for instance, would be missed.

The statistical analysis of complex biological data sets using proton nuclear magnetic resonance (9), HPLC–MS (10), and UPLC–MS (11) has been employed in metabonomic studies to detect and visualize the differences between mammalian samples either following the administration of a candidate pharmaceutical (12) or as a result of disease state progression. This is achieved by employing simple, unbiased, statistical tools, such as principal components analysis (PCA) and partial least squares (PLS) analysis or more powerful applications such as partial least squares discriminat analysis (PLS-DA) or more advanced techniques such as statistical total correlation spectroscopy (13), which allow large data sets to be visualized and the cross correlation of data to be achieved. In this paper, we present the use of a rapid high resolution LC–MS screening approach combined with simple statistical analysis, to classify Simvastatin tablets from different manufactures and identify the impurities in each batch of sample.

Experimental

Chemicals

Ammonium acetate and acetic acid were obtained from Sigma-Aldrich Chemicals (St. Louis, MO). Acetonitrile was obtained from Fisher Scientific (ThermoFisher, Waltham, MA), and the distilled water was produced in-house using a Millipore MilliQ system (MilliPore, Billerica, MA). The Simvastatin (S) standard compound was purchased from the United States Pharmacopoeia (Rockville, MD), and the Simvastatin tablets were obtained from four different manufactures (Lupin Limited Mumbai India, Merck Sharp & Dohme Pty Limited, Granville NSW, Australia, USV Limited Mumbai India, Merck & Co Inc., Rahway, NJ).

Chromatography

The chromatographic separations were performed on a Waters ACQUITY BEH C18 column (2.1 × 100 mm, 1.7 µm) (Waters Corporation, Milford, MA), the column was operated at 40°C and eluted with a linear gradient of 50–100% acetonitrile versus ammonium acetate (pH 5) over 5 min followed by a hold at 100% acetonitrile for 0.5 min before returning to original starting conditions over 0.1 min. The separations were performed on Waters ACQUITY Ultra Performance LC system with a mobile phase flow rate of 800 µL/min, generating a column back-pressure of up to 10,500 psi.

Mass spectrometry

MS was performed on a Waters Q-TOF Premier (Waters MS Technologies, Manchester, UK) orthogonal acceleration time of flight mass spectrometer operating in positive ion mode. The nebulization gas was set to 800 L/h at a temperature of 350°C, the cone gas was set to 10 L/h, and the source temperature set to 120°C. A capillary voltage and a cone voltage were set to 3200 V and 60 V, respectively. The Q-TOF Premier was operated in V optics mode with 10,000 resolution (FWHM). The data acquisition rate was set to 0.095 s, with a 0.05 s inter-scan delay. Data were collected for 10 min, using alternating collision energies of 5 eV and 25 eV to provide precursor and fragment ion information. All analyses were acquired using the lockspray to introduce a reference compound via an indexed auxillary sprayer to ensure accuracy and reproducibility; leucine-enkephalin was used as the lock mass (*m/z* 556.2771) at a concentration of 300 pg/µL and flow rate of 30 µL/min. Data was collected in centroid mode from 100–1000 *m/z* with a lockspray frequency of 11 s, and data averaging over 10 scans. The instrument is operated in a wide band rf mode in which alternating parallel scans are utilized. The low energy scan provided intact *m/z* information, while the high-energy scan provided fragment ion information allowing for the comprehensive generation of mass spectrometry data in one analysis. Following data collection, the data were processed in many different ways to reveal common precursor ions, common fragment ions, and constant neutral loss data without the need to perform additional experiments to obtain the requisite data. The first quadrupole is operated in a wide band rf-mode for both the precursor and fragment ions collected in accurate mode. This facilitates the determination of the elemental composition of the fragment ions during structural elucidation.

The statistical analysis was performed using the Waters MassLynx software and MarkerLynx application manager. The peaks were detected and integrated with the ApexTrak software; the data was then aligned using the proprietary algorithm within the software to produce an aligned data table. This table was then automatically reintegration to ensure all detected peaks were correctly assigned in each sample. The data were then subjected to PCA analysis with mean centering and pareto scaling.

Results and Discussion

Liquid chromatography has long been the technique of choice for impurity profiling and product analysis in the pharmaceutical industry. This technique relies on the complete resolution of the known products in a reliable reproducible manner. The ability to resolve all of the components in the sample relies on efficiency of the chromatographic column, the duration of the analysis and the selectivity of the mobile phase/stationary phase combination. The efficiency of the chromatography column is dependent upon the efficiency of the column packing and in turn the particle size. Chromatographic theory dictates that smaller particles generate higher resolution separation, but also demand/require higher mobile phase linear velocities (flow rates) for optimal performance (14). Thus, conveniently, with smaller particles superior performance is also accompanied by faster analysis. The only drawbacks to the use of these smaller particles for liquid chromatography are there requirement for higher operating pressures and LC systems with low delay volumes, both of these have been addressed by modern instrument design (15). The data displayed in Figure 1 shows the positive ion LC–MS analysis of a standard solution of Simvastatin at a con-

centration of 10 µg/mL. The column was eluted under gradient conditions from 50% to 100% acetonitrile over 5 min with ammonium acetate as the aqueous modifier. The data generated shows a high resolution chromatogram with major peak eluting at 3.13 min, the mass spectra of this peak corresponded to that of Simvastatin with a *m/z* value of 419. The chromatography system demonstrated excellent retention time and peak shape reproducible over the duration of the study.

Samples from 4 separate batches of Simvastatin tablets were analyzed using the LC–MS system, previous with-in batch analysis revealed no significant differences between tablets for each manufacturer. An example of the chromatograms obtained from two separate batches of Simvastatin tablets is shown in Figure 2. A careful review of the TIC obtained for both systems showed significant differences between the two samples, with the sample in chromatogram B exhibiting extra peaks with retention times of 3.70, 3.99, and 4.13 min compared with those obtained in the sample illustrated in chromatogram A. The accurate mass MS and MS–MS data allowed the two peaks eluting with a retention time of 3.99 and 4.13 to be identified as the Simvastatin-acetate and anhydro-Simvastatin impurities, respectively. The Simvastatin acetate impurity MS spectrum showed a dominant ion with a mass of 401.2648, corresponding to an elemental composition of $C_{25}H_{37}O_4$ for the MH+ ion, with a mass accuracy of 3 ppm. The fragment ion generated from this peak gave a diagnostic fragment ions of *m/z* = 285 and 199, suggesting that the lower part of the molecule remained unchanged, which in turn suggested that the change in structure occurred in the upper section of the molecule. The reduction in mass of 18 compared to the Simvastatin molecule suggested the loss of water, this information combined with the fragmentation pattern allowed this molecule to be identified as a Simvastatin anhydro impurity. The peak, eluting with a retention time of 3.99 min, produced a MH^+ ion with an *m/z* value of 461.2917. This *m/z* = 461 ion generated an isotope fit value of 1.5 for the elemental composition $C_{27}H_{41}O_6$. This information, combined with the fragmentation ions 423, 307, and 199 produced in the MS–MS analysis allowed the analyte to be identified as the Simvastatin acetate impurity.

A simple way to compare the data from each manufactur's tablets would be to compare the ratio of the peaks of interest; the data in Table I shows the ratio of the major impurities as a percentage of the major peak intensity. We can see from this data that this approach does not really give the scientist any useful insight into the relative impurity levels of the samples. Many researchers have reported the successful use of LC–MS with

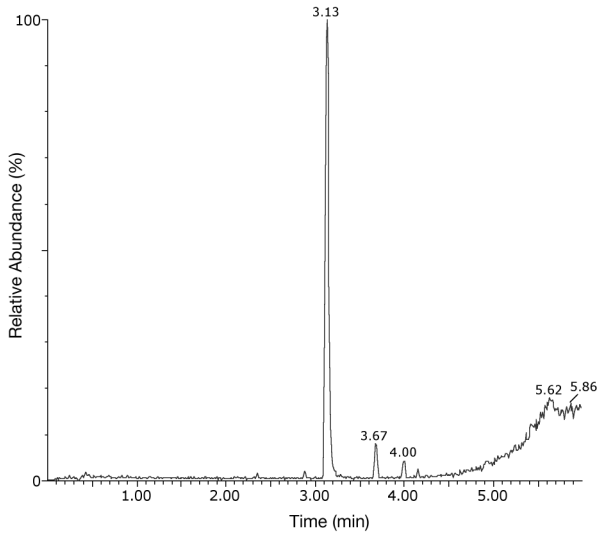

Figure 1. Positive ion electrospray TIC LC–MS trace from the gradient elution of Simvastatin standard.

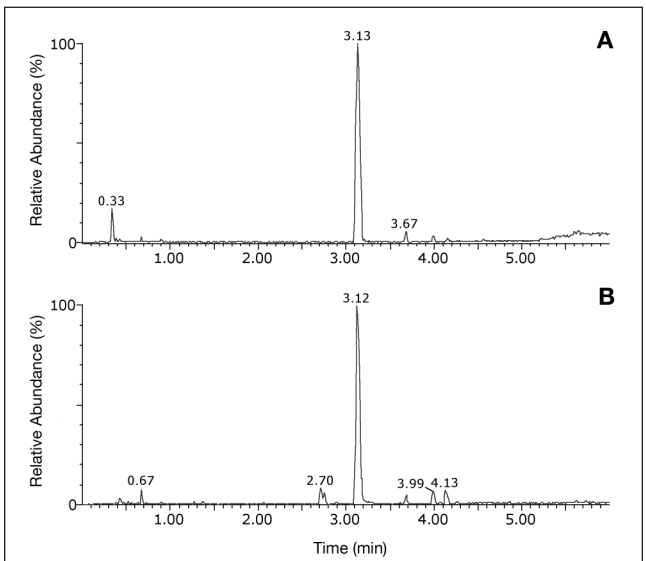

Figure 2. Representative positive ion LC–MS trace from samples I and II.

Table I. Relative Intensities of the Acid, Anhydro, and Acetate Impurities in the Simvastatin Tablet Samples I, II, III, and IV

	Simvastatin Acid	Simvastatin Acetate	Simvastatin Anhydro
Sample 1	0.86	1.57	1.74
Sample 2	0.94	1.64	2.0
Sample 3	1.36	3.38	5.10
Sample 4	1.91	1.22	2.66

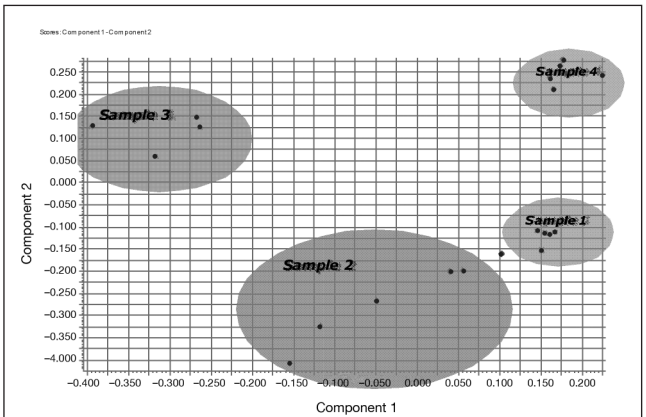

Figure 3. PCA scores plot for the positive ion LC–MS data showing P1 vs. P2.

statistical analysis to facilitate the identification of groups and classes in metabolomics/metabonomics studies (16–18). In these studies, the researchers have employed simple non-biased statistical analysis such as PCA and PLS as well as more complex analysis such as PLS-DA to group samples in similar and dis-similar clusters and identify the analytes responsible for the observed clustering. The positive ion data collected from the six replicate injections of each of the four samples was evaluated by principle components analysis, using mean centering and pareto scaling. The P1/P2 PCA score plot produced is displayed in Figure 3. Here we can see that the PCA analysis grouped the samples into four discrete clusters. The loadings plot, which indicated the peaks contributing most significantly to the variation observed in the data, for the 4 Simvastatin tablets groups is shown in Figure 4. For simplicity only, the *m/z* value is given, although the peaks are actually described by a mass retention time pair. In this data, we can see that there is a strong signal from the Simvastatin moiety (*m/z* = 419).

The data was also simplified using the MarkerLynx software such that the Simvastatin signal and any fragment ions relating to a particular impurity were removed, prior to further PCA statistical analysis (Figure 5). The data was further simplified by employing a mass deficiency filter of –50 mDa to +20 mDa below/above the accurate mass of Simvastatin. The resulting PCA data generated then contained only the peaks related to the Simvastatin molecule. This eliminated the effects of tablet excipient differences between the batches of tablets; whilst these may be an important difference to the efficacy of a tablet formulation effecting its dissolution rate and hence overall bioavailability, it is not the purpose of this paper to illustrate the effect of these components. In Figure 5, we can see that the major ions contributing to the positioning of the Sample I tables are the *m/z* = 441 and *m/z* = 457 ion. The ions that contribute to the variance observed in the data are summarized in Table II. The data in Table II illustrates that sample 2 has the greatest number of impurity peaks contributing to its position in statistical space. This fact can be illustrated by the comparison of representative base peak intensity (BPI) chromatogram for samples 1 and 3 (Figure 6). Here we can see that there are significantly more impurity peaks in the sample 3 (Figure 6A) chromatogram compared to the sample 1 chromatogram (Figure 6B), thus giving confirmatory evidence to the information generated in the PCA analysis.

The low collision energy MS data was used to provide precursor ion data from which the elemental composition data

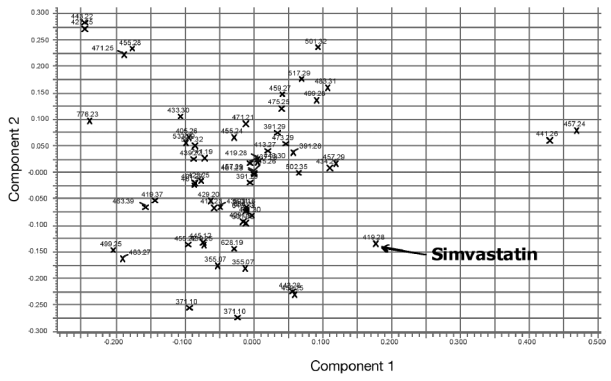

Figure 4. PCA loadings plot of P1 vs. P2 with Simvastatin signal included.

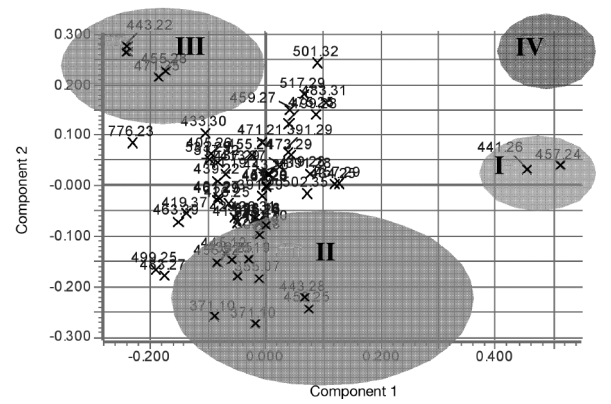

Figure 5. PCA loadings plot of P1 vs. P2 with the Simvastatin signal excluded; the overlaid PCA scores plot positions for samples I–IV.

Table II. Impurities Responsible for the Relative Position of the Samples in the PCA Plot

Ions Contributing to clustering	Sample I	Sample II	Sample III	Sample IV
371.10		X		
355.07		X		
443.28		X		
459.25		X		
457.24	X			
441.26	X			
455.28			X	
471.25			X	
443.22			X	

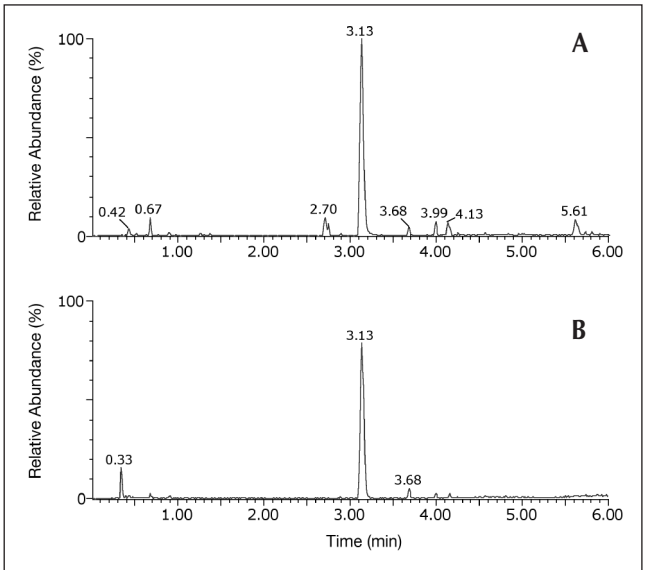

Figure 6. Comparison of positive ion TIC LC–MS traces for sample I (trace B) and sample II (trace A).

could be obtained for each peak. The simultaneously acquired high collision energy data was used to provide fragment information and accurate mass data on each peak produced. This approach has previously been reported by Bateman et al. (19) for the analysis of peptides also by Johnson and Plumb (20) for the analysis of acetaminophen metabolites and by Wrona et al. (21) for the analysis of in vitro metabolites. This approach was utilized to confirm the identity of the impurities detected in this study and those contributing to the observed group clustering. The impurity identification data was used to annotate the PCA loadings plot, highlighting which impurity was responsible for the position of the individual samples on the PCA scores plot (Figure 7). As can be seen from this plot, the sample II tablets contained considerably more impurities than the other samples. The impurities identified were consistent with those already reported (22). These impurities differences are most likely due subtle differences in the manufacturing process or the solvents used in the process. It is not the purpose of this communication to identify the reasons why these batches are different, this example has been used to illustrate the power of this technique to evaluate the similarities or differences between batches and samples.

Conclusion

The results given in this study demonstrate how the combination of a fast LC–MS methodology with simple unbiased statistical analysis can be used to identify the differences between individual batches of pharmaceutical products. The statistical analysis allowed the rapid facile analysis of the samples allowing them to be grouped into similar and dissimilar groups. The statistical process facilitated the detection of all of the drug substance impurities responsible for the observed group clustering. The use of a hybrid-quadrupole TOF-MS facilitated the collection of accurate mass LC–MS data, allowing the information to be simplified using a mass deficiency filter; constraining the statistical processing to only those signals relating to the impurities of the active drug product. The simultaneous acquisition of the low and high collision energy data allowed for the confirmation of the identity of the impurities via the accurate mass value of the intact molecule and the fragmentation pattern produced in the high collision energy experiment.

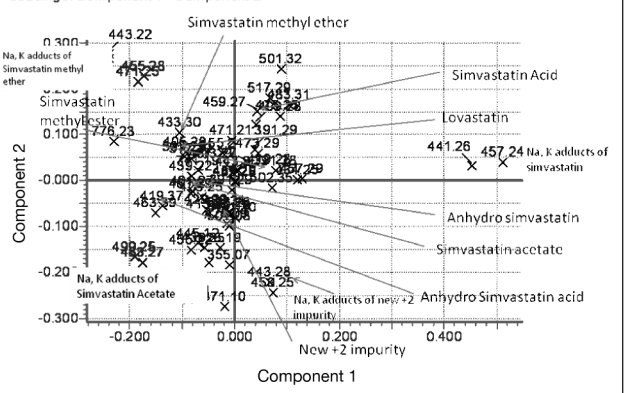

Figure 7. PCA loadings plot P1 vs P2 with the impurity peaks labeled.

Whilst in this example we employed LC–MS and PCA to determine and illustrate the differences between tablets made by different manufactures, it could also be used to identify the differences between batches of samples produced by the same manufacturer. This approach would allow for a simple batch control process to be developed, without the need to identify every peak in the sample allowing for a rapid decision to be made on product quality and batch release. There is a further advantage to this process compared to traditional impurity monitoring, where known impurities are monitored and unknown impurities may be ignored as it takes into account all detected analytes in the sample. As these "new" analytes/impurities could, potentially be, more toxic than those already known, thus this approach allows a more comprehensive, faster approach to the monitoring of pharmaceutical products.

References

1. Guidance for Industry Q7A Good Manufacturing Practice. Guidance for Active Pharmaceutical Ingredients. \\CDS018\CDERGUID\4286fnl.doc.
2. C. Eckers, A.M. Laures, K. Giles, H. Major, and S. Pringle. Evaluating the utility of ion mobility separation in combination with high-pressure liquid chromatography/mass spectrometry to facilitate detection of trace impurities in formulated drug products. *Rapid Commun. Mass Spectrom.* **21(7):** 1255–63 (2007).
3. C. Bharathi, D.K. Chary, M.S. Kumar, R. Shankar, V.K. Handa, R. Dandala, and A. Naidu. Identification, isolation and characterization of potential degradation product in risperidone tablets. *J. Pharm. Biomed. Anal.* **7;46(1):** 165–9 (2007).
4. J. Mazzeo, U. Neue, M. Kele, and R. Plumb. Advancing LC performance with smaller particles and higher pressures. *Anal. Chem.* **77(23):** 460A–7A (2005).
5. R.S. Plumb, M.D. Jones, P. Rainville, and J.M. Castro-Perez. The rapid detection and identification of the impurities of simvastatin using high resolution sub 2 microm particle LC coupled to hybrid quadrupole time of flight MS operating with alternating high-low collision energy. *J. Sep. Sci.* **17;30(16):** 2666–2675 (2007).
6. M.E. Touber, M.C. van Engelen, C. Georgakopoulus, J.A. van Rhijn, and M.W. Nielen. Multi-detection of corticosteroids in sports doping and veterinary control using high-resolution liquid chromatography/time-of-flight mass spectrometry. *Anal. Chim. Acta.* **14;586(1–2):** 137–46 (2007).
7. S.A. Wren and P. Tchelitcheff. Use of ultra-performance liquid chromatography in pharmaceutical development. *J. Chromatogr., A* **1119(1–2):** 140–6 (2006).
8. J.X. Shen, H. Wang, S. Tadros, and R.N. Hayes. Simultaneous determination of desloratadine and pseudoephedrine in human plasma using micro solid-phase extraction tips and aqueous normal-phase liquid chromatography/tandem mass spectrometry. *J. Pharm. Biomed. Anal.* **40(3):** 689–706 (2006).
9. O. Beckonert, H.C. Keun, T.M. Ebbels, J. Bundy, E. Holmes, J.C. Lindon, and J.K. Nicholson. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts**..** *Nat. Protoc.* **2(11):** 2692–703 (2006).
10. I.D. Wilson, R. Plumb, J. Granger, H. Major, R. Williams, and E.M. Lenz. HPLC-MS-based methods for the study of metabonomics. *J. Chromatogr., B* **5;817(1):** 67–76 (2005).
11. I.D. Wilson, J.K. Nicholson, J. Castro-Perez, J.H. Granger,

K.A. Johnson, B.W. Smith, and R.S. Plumb. *J. Proteome Res.* **4(2):** 591–8 (2005).
12. S. Rezzi, Z. Ramadan, F.P. Martin, L.B. Fay, P.V. Bladeren, J.C. Lindon, J.K. Nicholson, and S. Kochhar. Human metabolic phenotypes link directly to specific dietary preferences in healthy individuals. *J. Proteome Res.* **2;6(11):** 4469–4477 (2007).
13. O. Cloarec, M.E. Dumas, A. Craig, R.H. Barton, J. Trygg, J. Hudson, C. Blancher, D. Gauguier, J.C. Lindon, E. Holmes, and J. Nicholson. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Anal. Chem.* **1;77(5):** 1282–9 (2005).
14. R.S. Plumb, P. Rainville, B.W. Smith, K.A. Johnson, J. Castro-Perez, I.D. Wilson, and J.K. Nicholson. Generation of ultrahigh peak capacity LC separations via elevated temperatures and high linear mobile-phase velocities. *Anal. Chem.* **15;78(20):** 7278–83 (2006).
15. M.E. Swartz. UPLC™ : An introduction and review**.** *J. Liq. Chrom. & Rel. Technol.* **28(7-8):** 1253–63 (2005).
16. M. Ala-Korpela. Potential role of body fluid 1H NMR metabonomics as a prognostic and diagnostic tool. *Expert Rev. Mol. Diagn.* **7(6):** 761–73 (2007).
17. Z.D. Zeng, Y.Z. Liang, F.T. Chau, S. Chen, M.K. Daniel, and C.O. Chan. Mass spectral profiling: an effective tool for quality control of herbal medicines. *Anal. Chim. Acta.* **5;604(2):** 89–98 (2007).
18. L.M. Smith, A.D. Maher, O. Cloarec, M. Rantalainen, H. Tang, P. Elliott, J. Stamler, J.C. Lindon, E. Holmes, and J.K. Nicholson. Statistical correlation and projection methods for improved information recovery from diffusion-edited NMR spectra of biological samples. *Anal. Chem.* **1;79(15):** 5682–9 (2007).
19. R.H. Bateman, R. Carruthers, J.B. Hoyes, C. Jones, J.I. Langridge, A. Millar, and J.P. Vissers. A novel precursor ion discovery method on a hybrid quadrupole orthogonal acceleration time-of-flight (Q-TOF) mass spectrometer for studying protein phosphorylation. *J. Am. Soc. Mass Spectrom.* **13:** 792–803 (2002).
20. K.A. Johnson and R. Plumb. Investigating the human metabolism of acetominophen using UPLC and exact mass oa-TOF MS. *J. Pharm. Biomed. Anal.* **39:** 805–810 (2005).
21. R.S. Plumb, M.D. Jones, P. Rainville, and J.M. Castro-Perez. The rapid detection and identification of the impurities of simvastatin using high resolution sub 2 microm particle LC coupled to hybrid quadrupole time of flight MS operating with alternating high-low collision energy. *J. Sep. Sci.* **17;30(16):** 2666–2675 (2007).

Manuscript received November 21, 2007;
revision received December 17, 2007.