Technique review

Development of high throughput dispersive LC–ion mobility–TOFMS techniques for analysing the human plasma proteome

Xiaoyun Liu, Manolo Plasencia, Susanne Ragg, Stephen J. Valentine and David E. Clemmer

Date received: 28th May 2004

Xiaoyun Liu
obtained his BS from Jilin University in China and his MS from Michigan State University. Currently he is a second-year graduate student at Indiana University.

Manolo Plasencia
obtained his BS and MS degrees from Florida International University. Currently he is a third-year graduate student at Indiana University.

Susanne Ragg
is an assistant professor of pediatrics hematology/oncology at the Indiana University Medical School. She received her MD from Heidelberg University and her PhD from the institute for Human Genetics and Anthropology at Heidelberg University.

Stephen J. Valentine
received a BA in chemistry from Adams State College (1995) and a PhD in analytical chemistry from Indiana University in 2000. He worked for three years as a research scientist at Beyound genomics. He is currently working as a research scientist at Indiana University.

David E. Clemmer
received a BS degree from Adams State College and then a PhD from the University of Utah (1992). He joined the faculty at Indiana University as an assistant professor of analytical chemistry in 1995, and is currently a Robert and Marjorie Mann Chair and the Chairman of the Department of Chemistry at Indiana University. He is a former guitarist for the Yard Dogs.

Keywords: *biomarker, ion mobility spectrometry, mass spectrometry, plasma, proteomics*

David E. Clemmer,
Department of Chemistry,
Indiana University,
Bloomington IN, 47405, USA

Tel: +1 812 855 8259
Fax: +1 812 855 8300
E-mail: clemmer@indiana.edu

Abstract

A technique that combines ion mobility spectrometry (IMS) with reversed-phase liquid chromatography (LC), collision-induced dissociation (CID) and mass spectrometry (MS) has been developed. The approach is described as a high throughput means of analysing complex mixtures of peptides that arise from enzymatic digestion of protein mixtures. In this approach, peptides are separated by LC and, as they elute from the column, they are introduced into the gas phase and ionised by electrospray ionisation. The beam of ions is accumulated in an ion trap and then the concentrated ion packet is injected into a drift tube where the ions are separated again in the gas phase by IMS, a technique that differentiates ions based on their mobilities through a buffer gas. As ions exit the drift tube, they can be subjected to collisional activation to produce fragments prior to being introduced into a mass spectrometer for detection. The IMS separation can be carried out in only a few milliseconds and offers a number of advantages compared with LC–MS alone. An example of a single 21-minute LC–IMS–(CID)–MS analysis of the human plasma proteome reveals ~20,000 parent ions and ~600,000 fragment ions and evidence for 227 unique protein assignments.

INTRODUCTION

The problem with high throughput LC–MS methods for proteomics

The combination of mass spectrometry (MS) and collision-induced dissociation (CID) with multiple dimensions of condensed-phase separation strategies — such as strong cation exchange chromatography (SCX) with reversed-phase liquid chromatography (LC) — has become a standard platform for the identification of peptides and proteins in complex proteomics mixtures.[1–5] While it is difficult to overstate the impact that these strategies are having (literally revolutionising the analysis of complex systems), this platform is not without limitations. Among these is the fact that the combination of SCX with LC is time consuming — a typical two-dimensional experiment requires many hours to complete. Additionally, MS/MS strategies for peptide and protein identification on such complex systems are subject to problems with sampling — especially when coupled to LC. In a complex mixture, any two experiments (including back-to-back analyses of the same sample) will differ in which components are selected for MS/MS experiments and subsequently identified — a result of

slight variations in the positions of components (and intensities of parent ions) within the two-dimensional separation space. Ultimately, these sampling issues come about because the analysis requires that one precursor ion is initially selected by the MS; also, during the time that the precursor ion is selected, exposed to energising collisions to produce fragment ions (ie CID), and subsequently analysed again by MS to detect fragments (the MS/MS process may take a few seconds for each peptide), many other components of the complex mixture may elute from the column and thus be missed by the analysis.

Sampling issues become even more significant as the 'complexity' of a mixture increases. By 'complexity', is meant the number of different components and the extent to which they vary in concentration. While it is relatively straightforward to write data-dependent software that allows the instrument to find and select the most intense parent ions (at any given time) for MS/MS experiments, it is difficult to identify and select relatively low concentration components (unless it is already known that they exist). Arguably, low-abundance species, such as those associated with membrane proteins or post-translational modifications, may be of most interest from a functional point of view. What is needed is a high throughput technology that is capable of analysing all of the ions all of the time.

Dispersing components and nesting measurements: development of LC–IMS–TOF

Over the past several years, researchers in Professor Clemmer's group have attempted to develop such an approach by combining dispersive separation techniques with nested measurements — a combination of LC, ion mobility spectrometry (IMS) and MS. Ion mobility techniques have a long history in analytical chemistry as a means of detecting analytes (eg explosives and chemical warfare agents[6,7]). These applications are suitable because IMS is selective, relatively easy to implement and can be used at high analysis speeds. Ion selectivity is based on the uniqueness of the mobility of an ion as it is pulled through a buffer gas under the influence of a weak electric field — conceptually similar to well-known electrophoresis techniques. In the gas phase, an ion's electrophoretic mobility is well defined and depends only on the overall shape and charge of the ion (and experimental parameters that can be normalised to standard conditions). Thus, unlike gel measurements, gas-phase ion mobilities are highly reproducible — any two measurements typically agree (from run to run and laboratory to laboratory) to within 1 per cent (relative uncertainty). Because the density of a gas is $\sim 10^3$ times less than that of a liquid, separation times are of the order of 10–100 milliseconds (ms) rather than minutes to hours, as is the case for other condensed-phase separation techniques. Additionally, the combination of LC with IMS involves two dispersive approaches and the difference in time required for each dimension allows measurements to be recorded in a 'nested' fashion (see below). In this approach, no ions are intentionally selected (and thus none is intentionally discarded). When combined with a dispersive MS approach such as time-of-flight (TOF) MS, it is possible to examine large mixtures of ions and obtain reproducible representations of complex mixtures from run to run.

Here, a brief technical review of the early development of this technique for use in the analysis of complex peptide mixtures will be provided. The method will be illustrated by showing some very preliminary attempts to analyse rapidly the human plasma proteome. The instrumentation and methods are at an early developmental stage and (as discussed) there are many limitations associated with applications involving proteomics. Nevertheless, it appears that these methods may already offer some advantages in terms of speed,

reproducibility and coverage that will encourage further development.

TECHNIQUES

Mobility measurements

A number of authoritative reviews provide detailed descriptions about IMS background, instrumentation, theory and use as a structural probe.[8–11] The time required for an ion to drift through a buffer gas under the influence of a uniform electric field (E) depends on its mobility ($K = v_D \cdot E^{-1}$, where v_D is the drift velocity of the ion).[12] It is useful to report values as reduced mobilities, as these allow any two measurements to be compared. The reduced mobility, K_0, is calculated from the following equation:[12]

$$K_0 = \frac{L^2}{t_D \cdot V} \times \frac{273.2}{T} \times \frac{P}{760}$$

where the parameters L, V, T, P *and* t_D refer to the drift tube length, drift voltage, buffer gas temperature and pressure and measured drift time, respectively. The ability to separate peaks depends on the instrumental resolving power ($t_D/\Delta t$, where Δt corresponds to the full width at half maximum of the peak) as given by:

$$\frac{t_D}{\Delta t} = \left(\frac{LEze}{16 k_B T \ln 2}\right)^{1/2}$$

where k_B, z and e are Boltzmann's constant, the ion charge state and the electron charge, respectively.[8] Typically, the resolving power of peaks in measurements is of the order of 18–40; as can be seen from the equation, higher resolving powers can be obtained using higher drift fields. Resolving powers as high as $t_D/\Delta t = 240$ have been reported.[13]

Nesting measurements: timing is everything

The term 'nesting' was taken by analogy with nested Do/Continue loops of FORTRAN code; that is, an inner loop runs within outer loops. The LC–IMS–MS analysis is a nested measurement in the sense that dispersion into each new dimension occurs in much shorter times than the dispersion time associated with the previous dimension. The shortest measurement (innermost loop) is associated with dispersion of ions in the TOFMS based on mass to charge (m/z), a process that requires from ~3 to 30 μs in the authors' instruments; hundreds (to thousands) of TOFMS measurements can be carried out during the time required for the IMS separation, where ion drift times are typically 3–30 ms. The two-dimensional IMS–MS measurement (middle loop) can be carried out thousands of times during the LC separation (which typically requires minutes to hours).

Making a measurement

Figure 1 shows a schematic diagram of one experimental apparatus that has been constructed and which is used widely in the authors' laboratory, the LC–IMS–MS configuration. Briefly, LC–IMS–MS analysis is carried out as follows. LC is performed using standard capillary LC procedures. Peptides elute from a pulled tip capillary column and are electrosprayed into the source region of the IMS–MS instrument. Peptide ions produced from the continuous electrospray ionisation (ESI) beam are focused into a RF-focusing device (in this case, a linear octopole ion trap; in other cases, Paul geometry traps[14] and ion funnel traps[15] have been used) and accumulated to create a short concentrated packet of ions. Every 10 ms, this accumulated ion packet is ejected from the trap into the drift tube as a 100 μs-wide ion packet, initiating the IMS separation. Inside the drift tube, the packet of ions migrates under the influence of a weak electric field (imposed by a series of cylindrical electrodes) and different components separate due to differences in their mobilities through the buffer gas. Ions with compact structures (which undergo fewer collisions) have higher mobilities than ions with more open conformations. Also, high charge states typically have higher mobilities than

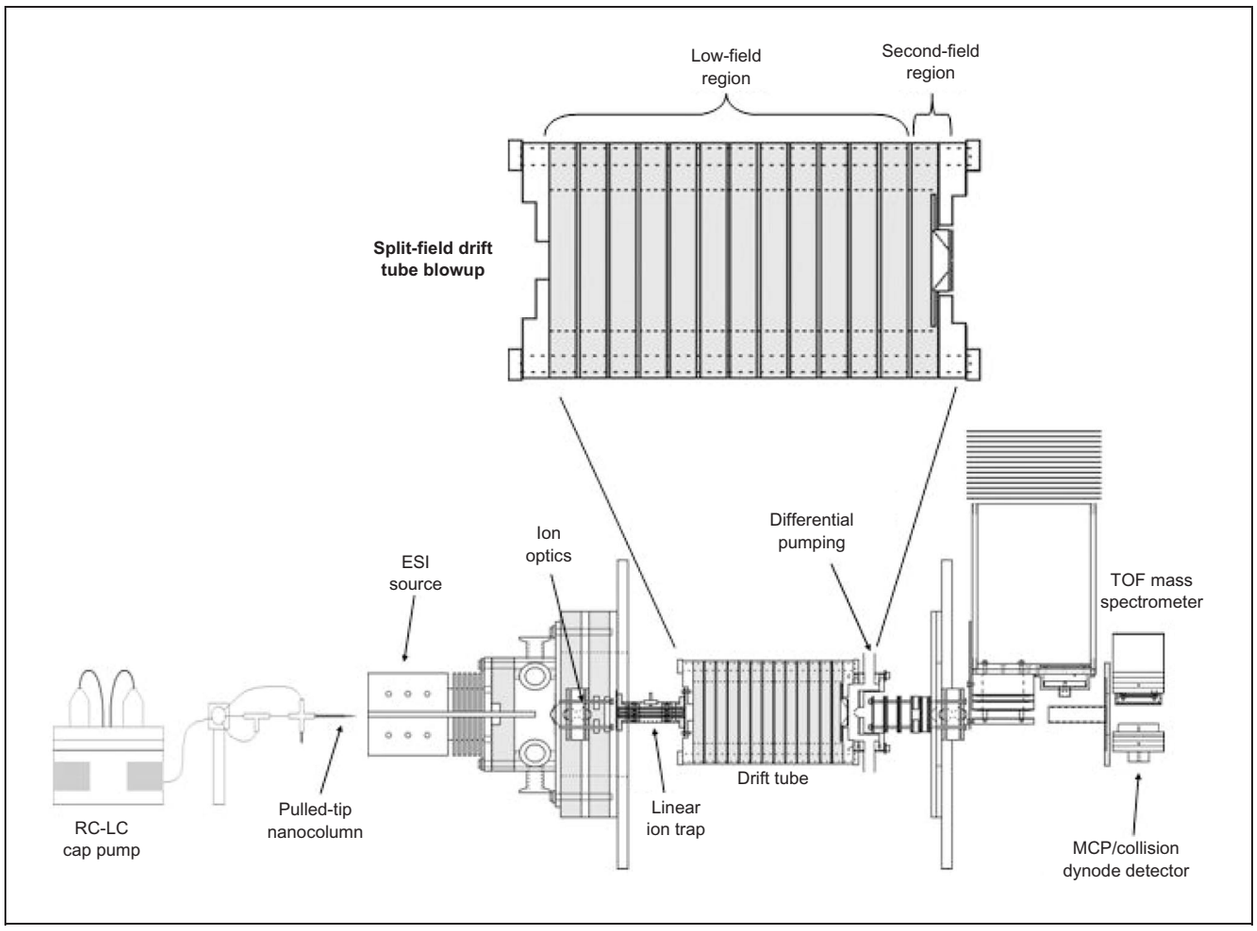

Figure 1: Schematic ion mobility/time-of-flight instrument with a split-field drift tube (authors' unpublished data)

low charge states because they experience a greater drift force. On exiting the drift tube, ions enter the source region of a TOF mass spectrometer. The pulse sequence of the TOF measurement is synchronised with the pulse that is used to inject ions into the drift tube.

Example IMS–TOFMS data

It is useful to illustrate just the IMS–TOF portion of an analysis with an example dataset. Figure 2 shows a two-dimensional plot of a nested drift-flight time measurement of peptides obtained from the digestion of 14 proteins.[16] In this case, the mixture was simple so that no LC separation was required and data were acquired by electrospraying the mixture directly into the IMS–TOFMS instrument. Measurements were recorded using a high-resolution drift tube that has a typical resolving power for tryptic peptides of ~50–90. Clearly, many peaks are observed across a range of drift times and flight times and most of these peaks can be assigned to tryptic peptides that are expected upon digestion of the mixture. The dataset is divided into several regions: region (a) encompasses the $[M+2H]^{2+}$ and $[M+3H]^{3+}$ precursor ions; regions (b) and (c) show narrow subsets (associated with relatively high and relatively low mobility ions) within this region; region (d) encompasses the $[M+H]^{+}$ precursors; and regions (e) and (f) show subsets within region (d). The sum of the signals over each of these ranges, shown in Figure 3, illustrates an important advantage of IMS–MS techniques. Summing the ion signal over broad ranges

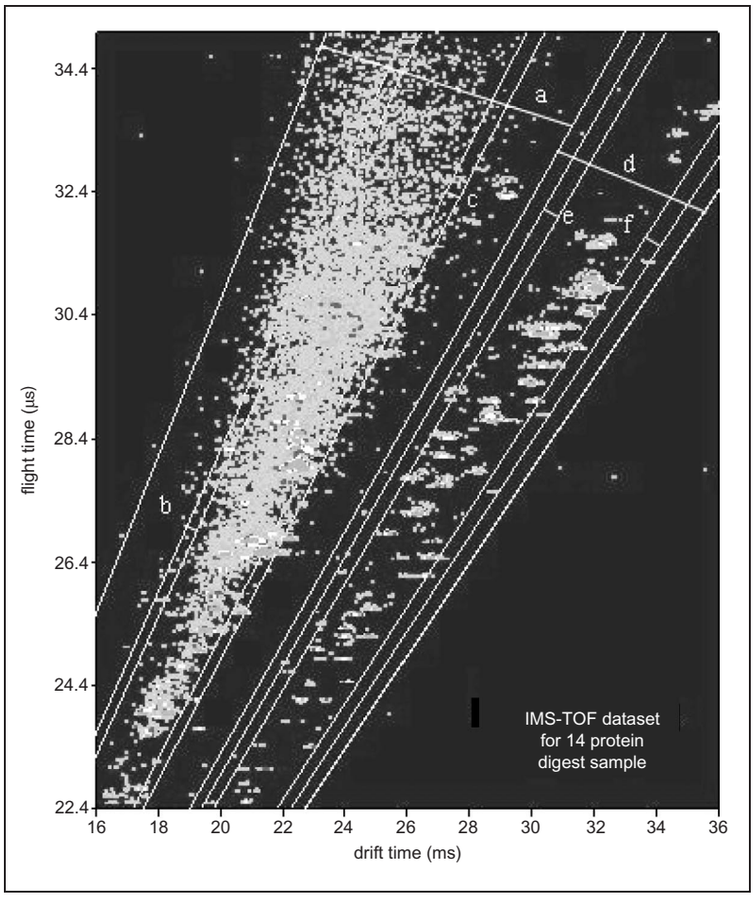

Figure 2: This figure is a modified form of the data shown in ref. 16. The plot shows a two-dimensional dataset for a mixture of digests from 14 proteins. Boundary regions for diagonal slices that were used to obtain summed mass spectra across the charge state families are indicated by lines. The solid bold lines denote boundaries used to obtain summed mass spectra over the $[M+H]^+$ or $[M+2H]^{2+}$ and $[M+3H]^{3+}$ charge state families, labelled (a) and (d), respectively. The normal solid lines (slices b and e) designate the regions used to obtain high-mobility (low drift time) summed mass spectra. The normal solid lines (slices c and f) designate the regions used to obtain low-mobility (high drift time) summed mass spectra. The proteins used for this study were albumin (bovine, dog, horse, pig and sheep), aldolase (rabbit), β-casein (bovine), cytochrome c (horse), β-lactoglobulin (bovine), myoglobin (horse) and haemoglobin (human, pig, rabbit and sheep). One hundred and sixty-eight peaks were assigned to tryptic fragments expected from the proteins and the number of tryptic peptide ions identified for each protein was $\geqslant$4

of the drift time — such as regions (a) and (d) — leads to spectra with high levels of background signal. Normally, in an MS measurement, this would be referred to as the background or noise level of the spectrum and any information about peaks of interest would need to exist as a signal that is substantially above this noise level in order to be useful. The ability to examine peaks across a narrow range of drift time windows effectively allows many peaks that would be buried under much larger features to be clearly observed. For example, most of the peaks associated with the spectra labelled (b), (c), (e) and (f) would not be detectable with MS alone. Thus, perhaps the most obvious advantage of the IMS dimension is that it allows many small signals to be detected in the presence of large signals (as long as the signals correspond to ions with different mobilities).

A modulated split-field drift tube for the acquisition of precursor and fragment ion information

As mentioned above, the combination of LC separations with MS/MS presents challenging problems associated with sampling. The authors' solution is to eliminate the initial process of selecting a single precursor ion and, instead, to simplify the mixture by the combination of LC and IMS separation prior to CID and MS detection. To do this, an LC split-field drift tube device has been developed (Figure 1). In this configuration, LC-separated peptides are electrosprayed to produce ions that are separated again in a relatively long first-field region of the drift tube that is operated at low fields (usually 3–10 $V \cdot cm^{-1}$); as the LC and mobility-dispersed species exit the first-field region, they enter a much shorter second-field region that can be operated at low fields (to pass the intact precursor ions into the MS) or at substantially higher fields (in order collisionally to activate the parent ions and produce fragments, which then enter the MS). By modulating the second-field region (between precursor and fragment ion conditions), it is possible to generate detailed profiles for most $[M+2H]^{2+}$ and $[M+3H]^{3+}$ precursor ions — and many more ions are sampled. In this approach, no single ion is rigorously selected by its m/z value for the fragmentation experiment; however,

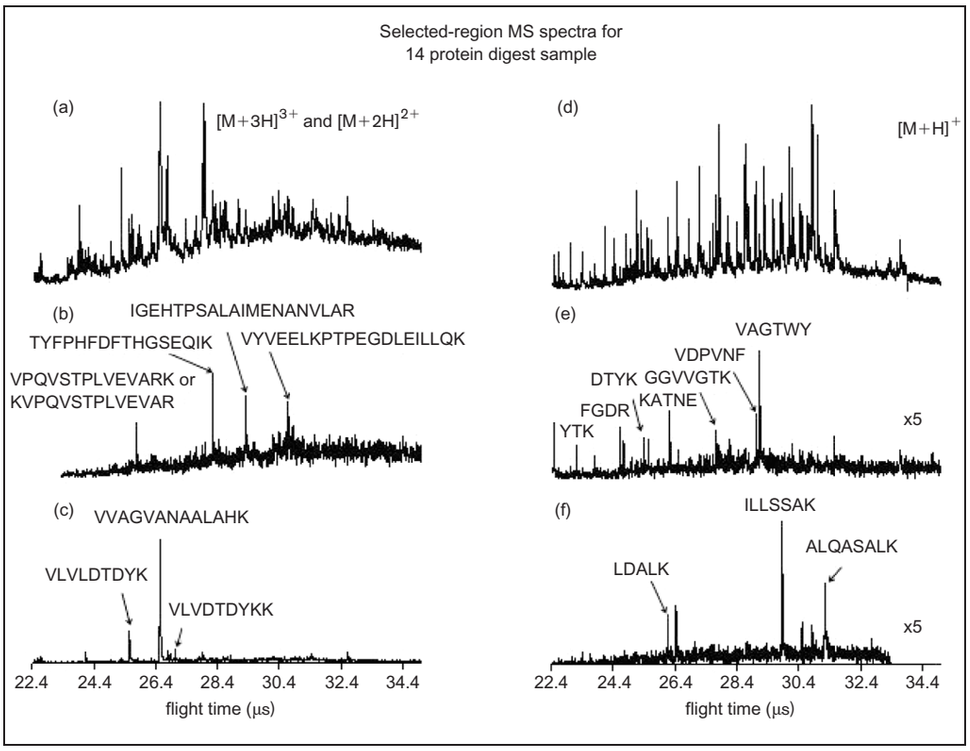

Figure 3: This figure is modified from its original form in ref. 16. The mass spectral plots correspond to the regions shown in Figure 2. Summed mass spectra for the (a) $[M+3H]^{3+}$ and $[M+2H]^{2+}$ charge state families; (b) high mobility $[M+3H]^{3+}$ and $[M+2H]^{2+}$ ions; (c) low mobility $[M+3H]^{3+}$ and $[M+2H]^{2+}$ regions; (d) $[M+H]^{+}$ charge state family; (e) high mobility $[M+H]^{+}$ charge state family; and (f) low mobility region of the $[M+H]^{+}$ charge state family. Note that a primary advantage of the IMS separation is the ability to separate ions such that many small features can be observed

because fragments and parents are coincident in LC time and drift time (through the first-field region), it is possible to group fragments along with the antecedents (precursor ions from which fragments came) in order to produce precursor MS and CID–MS fragment spectra. In many cases, only a single precursor (and its fragments) is found at a specific retention time and drift time combination, and often these data can be used to identify a peptide that is unique to a protein (eg by a database searching approach such as MASCOT).[17]

Example LC–IMS–(CID)–MS proteome analysis

Figure 4 illustrates the separation capabilities of the LC–IMS combination. This example dataset was acquired for a sample of tryptic peptides that are obtained upon digestion of soluble proteins extracted from human urine.[18] Here, many of the features separated along the drift time axis at given retention times are species that are not separated by LC alone. Figure 4 also shows an example of the IMS–(CID)–MS data near frame 300. A MASCOT search provides evidence for the xanthine dehydrogenase/oxidase, uromodulin and albumin proteins (from this single frame).

Using this approach, the human plasma proteome has been examined. Prior to 2002, only ~289 proteins had been identified in human plasma.[5,19] Recently, Smith and co-workers[5] reported that they had evidence for between ~800 and 1,700 plasma proteins. A brief summary of their results is provided in Table 1.

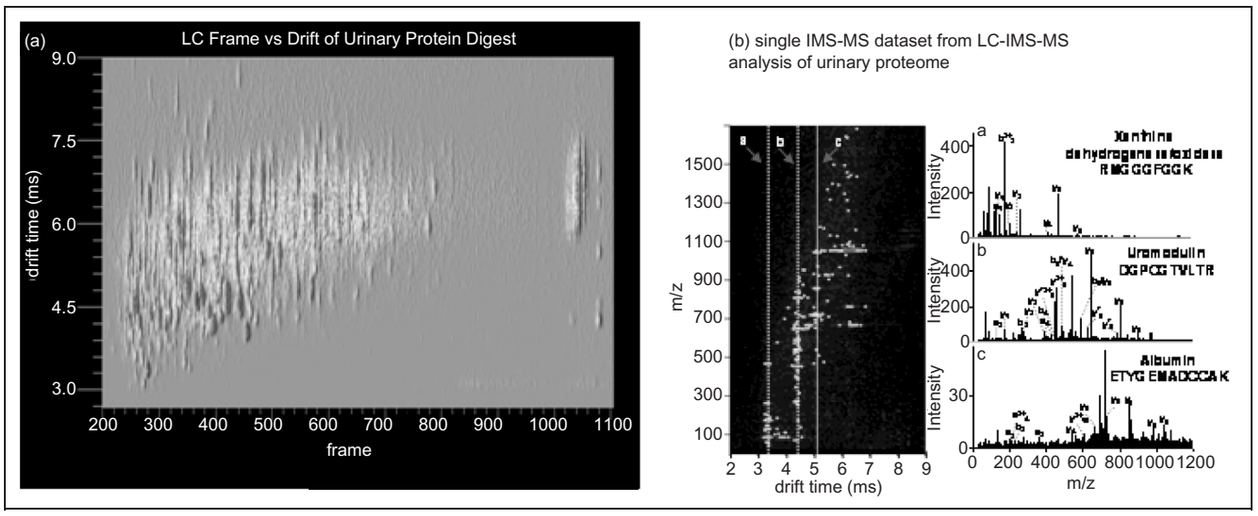

Figure 4: The plots shown here are modified from their original form in ref. 18. Part (a) (left) shows a plot of the two-dimensional separation of a tryptic digestion of the soluble proteins that are found in human urine. Note that during the preparation (before digestion), the removal of albumin has been attempted. This separation is obtained by microcolumn LC combined with ion mobility separations. Each frame corresponds to a complete two-dimensional mobility/time-of-flight dataset under parallel collision-induce dissociation (CID) conditions. In this dataset, each two-dimensional frame was integrated for 500 experimental cycles (a total acquisition time of 5 s for each frame; acquisition times for two-dimensional mobility/time-of-flight data are typically 0.5–5 s.) Part (b) (right) shows a single two-dimensional mobility/time-of-flight frame from this dataset. Three example CID/MS spectra for individual peptides can be obtained by summing the ion signal that is found at a single drift time. Several aspects of this plot are important. First, mobility and LC retention time are relatively orthogonal properties. The authors estimate the two-dimensional peak capacity (using a FWHM definition) of this dataset to be ~7,000–10,500 (this assumes a peak capacity of 350 for the LC separation and ~20–30 for the ion mobility separation). They also note that these data have been cut off so that small peaks are not observed. A single complete LC–IMS–MS dataset is typically 0.2–20 gigabytes in size. The urine sample was filtered with a regenerated cellulose membrane filter (pore size of 10 kDa). It was then diluted with 2.0 M urea solution in 50 mM phosphate buffer (pH = 7.5). Proteins were precipitated with 85 per cent ammonium sulphate solution and pelleted with centrifugation (15,000 rpm, 45 min). Protein pellets were resuspended in 0.5 ml deionised water and subsequently desalted online using a 5.0 ml HiTrap desalt column with Sephadex G-10 agarose gel (Amersham Biosciences Corporation, Piscataway, NJ). This created an exclusion limit of 2 kD. The extracted proteins were lyophilised. A HiTrap Blue affinity column was subsequently used to remove abundant proteins. The protein solution was desalted, lyophilised and digested. Prior to digestion, disulphide bonds were reduced and alkylated by the addition of dithiothreitol (10 mM) in an 8 M urea solution. After incubating for 2 hours (37°C), iodoacetamide (20 mM total concentration) was added, and the sample was incubated at 0°C (in darkness) for 2 hours. Excess (40×) cysteine was added to quench the reaction. The alkylated protein solution was diluted to ~1 M urea with phosphate buffer. Then, TPCK-treated trypsin was added to the solution at a 1:50 (trypsin:total protein) ratio. Samples were incubated at 37°C for 24 hours. Tryptic peptides were then desalted using a C18 sep pak cartridge. The peptide solution was subsequently lyophilised

Tryptic peptides were analysed by combinations of LC–MS/MS and SCX–LC–MS/MS. In direct LC–MS/MS measurements (using a 300-min LC gradient), an average of 475 peptides were identified by a database search, corresponding to an average of 126 proteins. The authors stressed sampling issues and noted that upon ten separate 300-min LC–MS runs, they found evidence for 902 peptides and 300 proteins. The SCX–LC–MS/MS combination yielded 881 proteins from 39 300-min LC runs. Although the increased coverage is a significant breakthrough, these methods illustrate timescale and sampling obstacles.

Table 1 also summarises some of the authors' early attempts to characterise plasma with LC–IMS–(CID)–MS methods. Using a 260-min gradient (similar to ref. 5, with a similar database search [MASCOT] for identification), evidence was found for 39,000 precursor

Table 1: Numbers of peptides and proteins identified from LC–MS and LC–IMS–MS analysis of human plasma

Technique	Time (minutes)[e]	Number of peptides identified[f]	Number of unique proteins[f]
LC–MS[a]	300	475[b]	126[b]
LC–MS (10 runs, cumulative)[a]	3,000	902	300
SCX–LC–RPLC–MS (39 runs, cumulative)[a]	11,700	Not reported	881
LC–MS[c]	21	99	32
LC–IMS–CID–MS[d] (one pedestal)	21	114	67
LC–IMS–CID–MS (three pedestals)	21	270	227
LC–MS	260	347	90
LC–IMS–CID–MS (one pedestal)	260	522	281

[a] Values were obtained from ref. 5. High-efficiency liquid chromatography (LC) separations were coupled with ion trap mass spectrometry.
[b] These values correspond to the average values obtained from those reported for ten reversed phase liquid chromatography (RPLC) runs in ref. 5.
[c] A commercial LC-Q DECA XP ion trap mass spectrometer (ThermoElectron Inc., Waltham, MA) was coupled with a nanoflow LC system (Dionex, Inc. Sunnyvale, CA) for the LC–mass spectrometry (MS) experiments. LC separation was achieved on a pulled tip capillary LC column packed in-house (15 cm×75 μm i.d. packed with 5 μm, 100 Å Magic C18AQ).
[d] A LC–IMS–time-of-flight (TOF) instrument (constructed in-house) was utilised for the ion mobility spectrometry (IMS) studies. For the one pedestal studies, parent ion datasets were collected and, in a separate experiment, collision-induced dissociation (CID) datasets were recorded. For the three pedestal dataset, parent ion and fragment ion retention time windows were recorded in a single experiment. The same LC column was employed for all condensed-phase separations.
[e] The time refers to the elution time of the experiment (roughly the gradient time). For the experiments reported in ref. 5, a cumulative time was reported where multiple runs resulted in a list of cumulative protein identifications. The solvent system utilised in these studies consisted of solvents A (96.95% water, 2.95% acetonitrile and 0.1% formic acid) and B (99.9% acetonitrile and 0.1% formic acid). The gradients (expressed as percentage solvent B) employed were: 10–20 per cent over 10 minutes; 20–30 per cent over 4 minutes; 30–38 per cent over 7 minutes; and 0–55 per cent over 260 minutes for the 21- and 260-minute elution time experiments, respectively. The gradients were the same for both instrumentation platforms.
[f] The number of unique peptides identified includes all unique peptides with scores above the homology score provided by MASCOT[17] (indicating the peptide match has <5 per cent chance of occurring randomly). Here, redundant peptide IDs have been removed. The number of unique proteins was obtained from these peptides. False positives were removed by visually inspecting MS/MS data for peptides with scores at or slightly above homology and, in some cases, by visually inspecting all assignments.

ion peaks, corresponding to 522 peptides and 281 proteins (347 peptides and 90 proteins were found from analysis of the same sample by LC–MS/MS; Table 1). Reducing the LC separation time by using a 21-min gradient, 114 peptides and 67 proteins were detected. Figure 5 depicts these data in terms of the number of unique peaks that were picked by the authors' in-house software (~20,000 parent ion peaks and ~600,000 fragment ion peaks using a relatively high count threshold; the actual number of peaks varies depending on the threshold). An analogous LC–MS/MS experiment on this timescale yielded 99 peptides and 32 proteins. Examination of these data reveals that many precursors did not undergo efficient fragmentation under the single second-field voltage setting. Thus, a multi-stepped second-field region was developed, where a single parent ion distribution is correlated with three different second-field fragmentation conditions. New data (a different plasma sample) yielded 270 peptides corresponding to 227 proteins. Many classical plasma proteins and tissue leakage proteins were detected. A very high dynamic range from this approach is

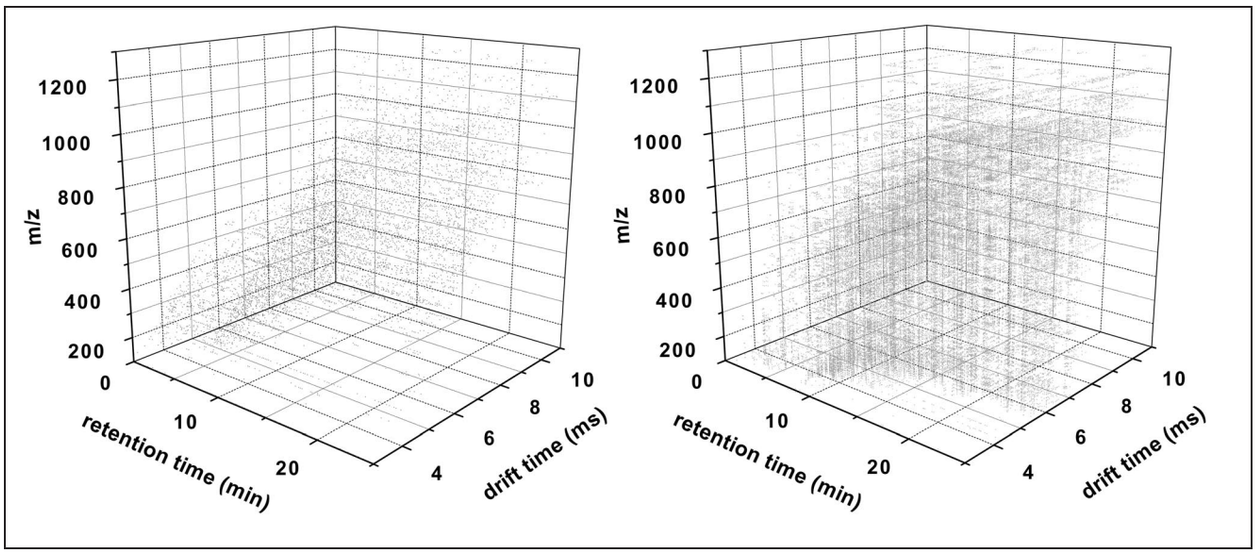

Figure 5: Three-dimensional plots of peak positions from parent-ion (left) and CID (right) data from analysis of tryptic peptides from digestion of proteins isolated from human plasma. Note that these are peak positions from processed data (not raw ion signals). This sample was obtained by digesting plasma that had been desalted on a PD-10 desalting column (Amersham Pharmacia Biotech, Uppsala, Sweden). Disulphide bonds were then reduced and alkylated as described in the caption to Figure 4. The alkylated protein solution was then diluted with a phosphate buffered saline (PBS) solution (pH = 7.5) to a final concentration of 2 M urea. Then, tryptic digestion proceeded as outlined in the caption to Figure 4. Tryptic peptides were then desalted using C4 Extraction Columns (J. T. Baker, Inc., Phillipsburg, NJ) and dried on the speed vac. Six microlitres of a $\sim$0.040 μg/μl solution of the digested peptides (determined by Bradford assay) were loaded onto the LC column for analysis

anticipated, but a detailed analysis has not been carried out. At this point, a dynamic range of $>10^6$ is certain.

Limitations: data storage, computational analysis and informatics

These preliminary methods are highly limited. Most significant is that a single file size (in a compressed form) ranges from 5 to >20 gigabytes. The authors aim to analyse entire datasets with a single algorithm; this is particularly attractive for dispersive techniques because, upon calibration, it should be possible to signal average complete datasets to improve the signal-to-noise levels of low-abundance components. There is currently no commercially available software that allows peaks to be quickly located and collated into groups of parent and fragment spectra. Data analysis is rate limiting. One glimmer of hope comes from the reproducibility of peak positions. Once a peptide has been identified, it can be directly mapped to a location, allowing a quick check of existence against new precursor and fragment spectra.

The amount of information that is considered is also limited. Current software does not include information about the ion charge state (apparent from both isotopic and mobility resolution), and the mass determination that is used is far below the actual measurement. Additionally, the mobility of an ion contains information about its shape — and different sequences have different shapes. Such information should be valuable as a means of further corroborating database assignments or as a constraint that helps eliminate false positives.

Acknowledgments

This work is supported in part by the NIH (GM59148), NSF (CHE0078737) and the Indiana Genomics Initiative (INGEN). We thank our

insightful colleagues (especially Professors Fred Regnier and Steve Naylor) and critical reviewers of our manuscripts.

References

1. Washburn, M. P., Wolters, D. and Yates, J. R. (2001), 'Large-scale analysis of the yeast proteome by multidimensional protein identification technology', *Nat. Biotechnol.*, Vol. 19, pp. 242–247.
2. Wolters, D. A., Washburn, M. P. and Yates, J. R. (2001), 'An automated multidimensional protein identification technology for shotgun proteomics', *Anal. Chem.*, Vol. 73, pp. 5683–5690.
3. Peng, J., Elias, J. E., Thoreen, C. C. *et al.* (2003), 'Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC–MS/MS) for large-scale protein analysis: The yeast proteome', *J. Proteome Res.*, Vol. 2, pp. 43–50.
4. Chen, J., Balgleg, B. M., DeVoe, D. L. and Lee, C. S. (2003), 'Capilliary isoelectric focusing-based multidimensional concentration/separation platform for proteome analysis', *Anal. Chem.*, Vol. 75, pp. 3145–3152.
5. Shen, Y., Jacobs, J. M., Camp, D. G. II *et al.* (2004), 'Ultra-high-efficiency strong cation exchange LC/RPLC/MS/MS for high dynamic range characterization of the human plasma proteome', *Anal. Chem.,* Vol. 76, pp. 1134–1144.
6. St. Louis, R. H. and Hill, H. H. (1990), 'Ion mobility spectrometry in analytical chemistry', *Crit. Rev. Anal. Chem.*, Vol. 21, pp. 321–355.
7. Collins, D. C. and Lee, M. L. (2002), 'Developments in ion mobility spectrometry-mass spectrometry', *Anal. Bioanal. Chem.*, Vol. 372, pp. 66–73.
8. Revercomb, H. E. and Mason, E. A. (1975), 'Theory of plasma chromatography/gaseous electrophoresis — A review', *Anal. Chem.*, Vol. 47, pp. 970–983.
9. Clemmer, D. E. and Jarrold, M. F. (1997), 'Ion mobility measurements and their applications to clusters and biomolecules', *J. Mass Spectrom.*, Vol. 32, pp. 577–592.
10. Hoaglund Hyzer, C. S., Counterman, A. E. and Clemmer, D. E. (1999), 'Anhydrous protein ions', *Chem. Rev.*, Vol. 99, pp. 3037–3079.
11. Wyttenbach, T. and Bowers, M. T. (2003), 'Gas phase conformations: The ion mobility/ion chromatography method', *Mod. Mass Spectrom. Top. Curr. Chem.*, Vol. 225, pp. 207–232.
12. Mason, E. A. and McDaniel, E. W. (1988), '*Transport Properties of Ions in Gases*', Wiley, New York, NY.
13. Srebalus, C. A., Li, J., Marshall, W. S. *et al.* (1999), 'Gas phrase separations of electrosprayed peptide libraries', *Anal. Chem.*, Vol. 71, pp. 3918–3927.
14. Hoaglund, C. S., Valentine, S. J. and Clemmer, D. E. (1997), 'An ion trap interface for ESI-ion mobility experiments', *Anal. Chem.*, Vol. 69, pp. 4156–4161.
15. Wyttenbach, T., Kemper, P. R. and Bowers, M. T. (2001), 'Design of a new electrospray ion mobility mass spectrometer', *Int. J. Mass Spectrom.*, Vol. 212, pp. 13–23.
16. Taraszka, J. A., Counterman, A. E. and Clemmer, D. E. (2001), 'Gas phase separations of complex tryptic peptide mixtures', *Fresenius J. Anal. Chem.*, Vol. 369, pp. 234–245.
17. URL: http://www.matrixscience.com
18. Moon, M. H., Myung, S., Plasencia, M. *et al.* (2003), 'Nanoflow LC/ion mobility/CID/TOF for proteomics: Analysis of a human urinary proteome', *J. Proteome Res.*, Vol. 2, pp. 589–597.
19. Anderson, N. L. and Anderson, N. G. (2002), 'The human plasma proteome — History, character and diagnostic prospects' *Mol. Cell Proteomics*, Vol. 1, pp. 845–867.