Available online at www.sciencedirect.com
ScienceDirect
International Journal of Mass Spectrometry 268 (2007) 158–167
International Journal of Mass Spectrometry
www.elsevier.com/locate/ijms

Nanoelectrospray peptide mapping revisited: Composite survey spectra allow high dynamic range protein characterization without LCMS on an orbitrap mass spectrometer

Aiping Lu [a], Leonie F. Waanders [a], Reinaldo Almeida [b], Guoqing Li [a], Mark Allen [b], Jürgen Cox [a], Jesper V. Olsen [a], Tiziana Bonaldi [a], Matthias Mann [a,*]

[a] *Department for Proteomics and Signal Transduction, Max-Planck-Institute for Biochemistry, Am Klopferspitz 18, D-82131 Martinsried, Germany*
[b] *Advion Biosciences Limited, Rowan House, 26 Queens Road, Norfolk NR9 3DB, UK*

Received 19 February 2007; received in revised form 9 May 2007; accepted 14 May 2007
Available online 21 May 2007

Abstract

Mass spectrometric (MS) determination of the primary structure of proteins, including post-translational modifications, remains a challenging task. Proteins are usually digested to tryptic peptides that are measured either by MALDI peptide mapping or by liquid chromatography online coupled to tandem MS (LC–MS/MS). Here we instead analyze peptides by a chip implementation of nanoelectrospray (TriVersa Nanomate, Advion Biosciences), coupled to a linear ion-trap–orbitrap hybrid instrument (LTQ-Orbitrap, Thermo Fisher). The C-trap connecting the linear ion-trap and orbitrap is filled repeatedly in different *m/z* ranges with up to a million charges. Each range is analyzed in the orbitrap repeatedly and separately, creating a survey spectrum composed of hundreds of single spectra. The composite spectrum is inherently normalized for different *m/z* ranges due to their different fill times and retains information on the variability of mass measurement and intensity. Nanoelectrospray offers analysis times of more than 30 min/μl of peptide mixture, sufficient for in-depth peptide characterization by high resolution C-trap fragmentation in addition to high sensitivity ion-trap fragment analysis. We obtain over 6000-fold dynamic range and subfemtomole sensitivity. Automated analysis of digested BSA resulted in sequence coverage above 80% in low femtomole amounts. We also demonstrate identification of seven modified peptides for a purified histone H3 sample. Static spray allows relative quantitation of the same peptide with different modifications. Chip-based nanoelectrospray on an orbitrap instrument thus allows very high confidence protein identification and modification mapping and is an alternative to MALDI peptide mapping and LC–MS/MS.

Keywords: LTQ-Orbitrap; Peptide mass mapping; Peptide sequencing; Dynamic range; Protein modification

1. Introduction

During the last few years, efforts in mass spectrometry-based proteomics [1] have concentrated on the qualitative and quantitative analysis of complex protein mixtures. However, most biological mechanisms involve protein modifications, which are not easily or comprehensively picked up in these large-scale experiments [2,3]. In contrast to the few peptides required for identification by MS, the analysis of post-translational modifications (PTMs) in principle requires peptides covering every part of the protein (100% sequence coverage). Furthermore, some modifications may be sub-stoichiometric, even in the purified protein of interest, requiring the analysis of several peptides covering the same sequence stretch.

MALDI Time-Of-Flight (TOF) and MALDI-TOF/TOF are popular methods to identify gel-separated proteins. MALDI sample preparation has been optimized and is rapid and convenient [4]. MALDI-TOF/TOF has been increasingly automated and now allows large number of gel spots to be identified, i.e., in combination with 2D gel electrophoresis. Nevertheless, the trend towards mixture analysis and quantitative proteomics have made LC–MS/MS 'shotgun' methods increasingly popular [5–7]. In particular, the quality of MS/MS data in LC–MS/MS often makes protein identifications much more specific than with the MALDI method [8]. Further advantages of LC–MS/MS are its high sensitivity as peptides are concentrated into very small

* Corresponding author. Tel.: +49 89 8578 2778; fax: +49 89 8578 2219.
E-mail address: mmann@biochem.mpg.de (M. Mann).

1387-3806/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2007.05.006

peak volumes and the extra information contained in the chromatographic retention time of each peptide. Disadvantages are the dynamic nature of LC–MS/MS, which makes it difficult to do repeat measurements of the same peak as well as to apply several fragmentation techniques during the elution time of typically less than 30 s.

In theory, nanoelectrospray [9,10] which is static and allows directed measurements offers a compromise allowing both ready identification of proteins without LC separation while still offering extremely high accuracy protein identification and mapping of post-translational modifications. The original 'manual' nanoelectrospray has now largely fallen out of favor, mainly because of its low throughput. However, recently nanoelectrospray has been revived in a chip-based form, commercially in the form of the Advion TriVersa Nanomate. Here we investigate the combination of this automated nanoelectrospray with a powerful new mass spectrometer, the hybrid linear ion-trap–orbitrap [11].

2. Experimental

2.1. Sample preparation for protein standards

Unless otherwise specified, chemicals were from Sigma Aldrich. Bovine serum albumin (BSA, 2 mg/ml Bio-Rad) was diluted to a concentration of 4 pmol/μl with 6 M urea/2 M thiourea, incubated in 1 mM DTT (final concentration) for 45 min at 56 °C for protein reduction and subsequently in 5.5 mM iodoacetamide (final concentration) at room temperature in the dark for 30 min for alkylation. The solution was digested with 1:50 (w/w) protein amount of endoproteinase Lys-C (Wako) for 4 h at room temperature, then diluted 4× with 50 mM NH_4CO_3 and digested further with 1:50 (w/w) protein amount of trypsin (Promega) overnight at 37 °C. The digestion was stopped by adding 1% (v/v) of absolute TFA. BSA peptides were desalted and stored on RP-C_{18} StageTip columns [12] and eluted right before mass spectrometric analysis with 50% methanol/0.5% formic acid.

2.2. Histone H3 sample preparation

Complete™ proteases inhibitors (Tablet, Roche) were added to all buffers below and the solutions were cooled to 4 °C before use. Semi-confluent HeLa cells were collected and resuspended in Buffer-N (15 mM Hepes–KOH pH 7.6, 60 mM KCl, 15 mM NaCl, 0.5 mM EGTA, 10% Sucrose). Lysis was performed by adding 0.2% NP40 and rotating the cell suspension for 10 min at 4 °C. Cell lysates were carefully poured on 20 ml sucrose cushions (20% sucrose in Buffer-N). Nuclear pellets were fractionated upon centrifugation (4000 rpm, 15 min, 4 °C) and washed in PBS. Core histones, together with linker histones and high mobility group proteins (HMG) were then extracted by adding a half volume of ice-cold HCl (0.8 M) overnight with continuous rotation at 4 °C. The sample was centrifuged for 10 min at 12,000 × *g*, and histones and the other acid-soluble proteins remained in the supernatant. Residual histones were re-extracted for 3–4 h in 0.4 M ice-cold HCl, the supernatants derived from the two extractions were pooled and dialyzed against 100 mM ice-cold acetic acid. The dialyzed sample was aliquoted, lyophilized, and evaluated for purity and concentration by resuspension in H_2O and by performing SDS-PAGE (18%).

About 100 μg histone sample was resuspended in 100 μl 0.1% TFA, 2% ACN and directly loaded onto a reverse phase HPLC column (Jupiter C_{18}, 250 × 4.60, 5 μm, 300 Å) (Phenomenex) connected to an Aekta LC-system (Amersham). Individual histones were separated by applying a gradient from 20% to 80% ACN in 0.1%TFA.

The total amount of histone H3 was estimated by SDS-PAGE. A fraction containing 1.5 μg of histone H3 was dried down and redissolved in a buffer composed of 100 mM Tris–HCl, 10 mM $CaCl_2$, pH 7.6 for overnight Arg-C (1:50, w/w) digestion at 37 °C. One half of the peptide solution was desalted and stored using RP-C_{18} StageTip columns, while the other half was desalted and stored using SCX (Strong Cation Exchange) StageTip columns. Peptides on the RP-C_{18} column were eluted by 50 μl 80% acetonitrile/0.5% acetic acid, and the peptides on the SCX column were eluted by 50 μl 5% ammonium hydroxide/30% methanol. Both eluates were combined, dried down, and redissolved in 50% methanol/0.5% formic acid (1 pmol/μl or 15 ng/μl) for nanoelectrospray.

2.3. Mass spectrometric analysis

All experiments were performed using a linear ion-trap–orbitrap hybrid mass spectrometer (LTQ-Orbitrap, Thermo Fisher Scientific, Bremen, Germany) with a TriVersa Nanomate (Advion Biosciences, Ithaca, USA) as ion source. A positive voltage (1.5 kV) is applied on the chip while the mass spectrometer sample orifice remains at 0 kV. The electrostatic field between the chip and the orifice drives the positive ions towards the mass spectrometer. The flow rate is dependent on the chip diameter. When not mentioned otherwise, all results were acquired with the low flow rate chip (i.d. 2.5 μm) from Advion, providing a flow rate of 20 nl/min. At this flow rate, 1 μl of sample provided stable static electrospray longer than 30 min, just like in 'classical' nanoelectrospray.

Every sample, consisting of 1 μl of solution was sprayed twice and MS spectra were acquired either by full range acquisition (full scan) or multiple overlapping segmented range acquisition (selected ion monitoring, or SIM scans). For the BSA sample, four segmented SIM mass ranges (300–500, 450–650, 600–800, 750–1350) were recorded. For the histone H3 sample, the four SIM segments were chosen as 300–550, 500–650, 600–750, and 700–900 *m/z*. MS/MS fragmentation was performed by data-dependent selection of the five most intense peaks in the segmented mass range. 'Dynamic exclusion' was set to 150 s, longer than the acquisition time per two overlapping segments.

2.4. Data analysis

The Mascot engine was used for mass spectrometry data identification (Matrix Science, London, UK). BSA peaks were searched in IPI_Human_v313 to which the BSA sequence had

been added, using 5 ppm maximum mass deviation (MMD [13]) for precursor ions, 0.5 Da MMD for fragment ions, carbamidomethylation (C) as fixed modification, and oxidation (M), *N*-acetylation, deamidation (NQ), pyro-glutamate (N-term QC) as variable modifications. Up to three missed cleavages were allowed and every fully tryptic, unique peptide ('bold red' in the MASCOT report) without a second protein match was accepted as a hit.

Histone H3 peaks were searched in a histone database (276 non-redundant sequences, including different histone proteins/variants, keratins and the proteases used), using 5 ppm MMD for precursor ion, 0.5 Da MMD for ion-trap fragmentation, and 0.01 Da mass tolerance for C-trap fragmentation (minimum possible in Mascot), and seven variable modifications, including methionine oxidation, N-terminal acetylation, mono-, and dimethylation of lysines and arginines and lysine trimethylation and acetylation. Up to two missed cleavages were allowed and peptides with a score higher than that corresponding to a significance value of $p = 0.05$ were accepted.

3. Results and discussion

3.1. Automated nanoelectrospray coupled to the LTQ-Orbitrap

'Classical' nanoelectrospray requires handling of fragile pulled needles, which is both time consuming and a skill demanding considerable dexterity. In contrast, the TriVersa achieves the same low flow rates and thereby sensitivity using a micro-machined chip that is operated completely automatically. Here we describe operation of the automated nanoelectrospray combined with a high accuracy mass spectrometer, the LTQ-Orbitrap. The TriVersa automatically takes a tip, aspirates the sample, and transfers it to the nozzle of the chip, located in front of the mass spectrometer (Fig. 1). As can be seen in the figure, the LTQ-Orbitrap contains a C-trap, which functions as a container for ions transferred from the ion-trap and waiting to be ejected into the high-resolution analyzer—the orbitrap. Importantly, the instrument allows any ion population isolated in the ion-trap to be accumulated in the C-trap for final high-resolution analysis in the orbitrap. This high-resolution scan in the orbitrap takes 0.25–1 s, depending on the resolution chosen.

While the instrument is extremely sensitive, its duty cycle is limited by the fact that the C-trap only accommodates 10^6 ions, which is often achieved with ion accumulation for just a few milliseconds. Secondly, the dynamic range is also limited by dominant ions (typically in the low to middle *m/z* range), which can make up a large fraction of the total ion population. We reasoned that the combination of nanoelectrospray and LTQ-Orbitrap should allow us to ameliorate both problems. Instead of acquiring a single full scan spectrum, we decided to acquire a large number of spectra by filling up the C-trap to capacity for each of a number of segmented mass ranges. This should lead to a 'normalized' mass spectrum consisting of a 'matrix' of individual spectra for several mass segments and averaged over many scans. This 'composite' spectrum should have a much larger dynamic range and peptide mass measurement accuracy than a single full scan spectrum or averaged full scan spectra.

Furthermore, the long spray time allows directed and iterative peptide fragmentation experiments. Peptides can be identified by peptide mass fingerprinting (PMF), ion-trap fragmentation with read out in the ion-trap or in the orbitrap, fragmentation in the C-trap or any combination of these. We therefore sought to devise efficient MS/MS schemes to characterize the maximum number of peptide peaks.

3.2. Acquisition methods for the composite spectrum and MS/MS acquisition

We found that a three step procedure, encompassing peptide mapping, data-dependent sequencing and directed sequencing of 'missing' peaks, was optimal for protein characterization (Fig. 2). In the first step the full mass range is divided by SIM scans into multiple overlapping segments (several hundred *m/z* units wide), which were acquired in the orbitrap. The segmented mass ranges, shown in Fig. 2A, were chosen so that the accu-

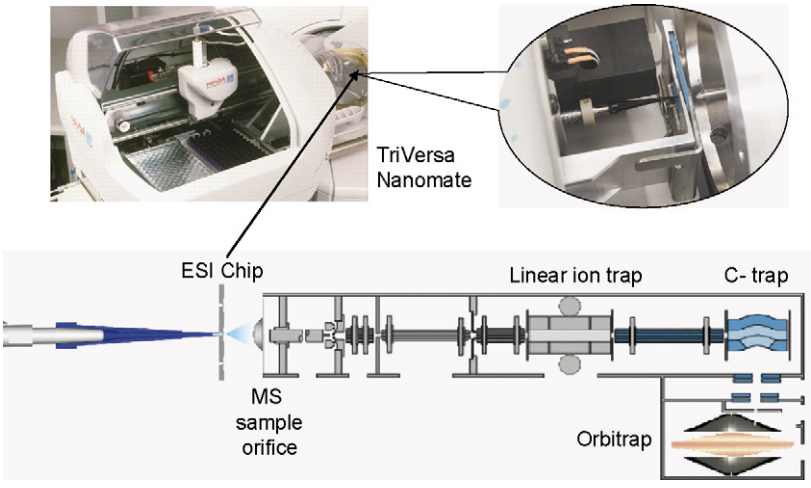


Fig. 1. Schematics of the TriVersa Nanomate coupled to the LTQ-Orbitrap. Samples were applied by coated tips to the nozzle of the electrospray chip in front of the TriVersa instrument. The low flow rate chip (i.d. 2.5 μm) provided a stable flow rate of 20 nl/min, which is in the 'true' nanoelectrospray range.

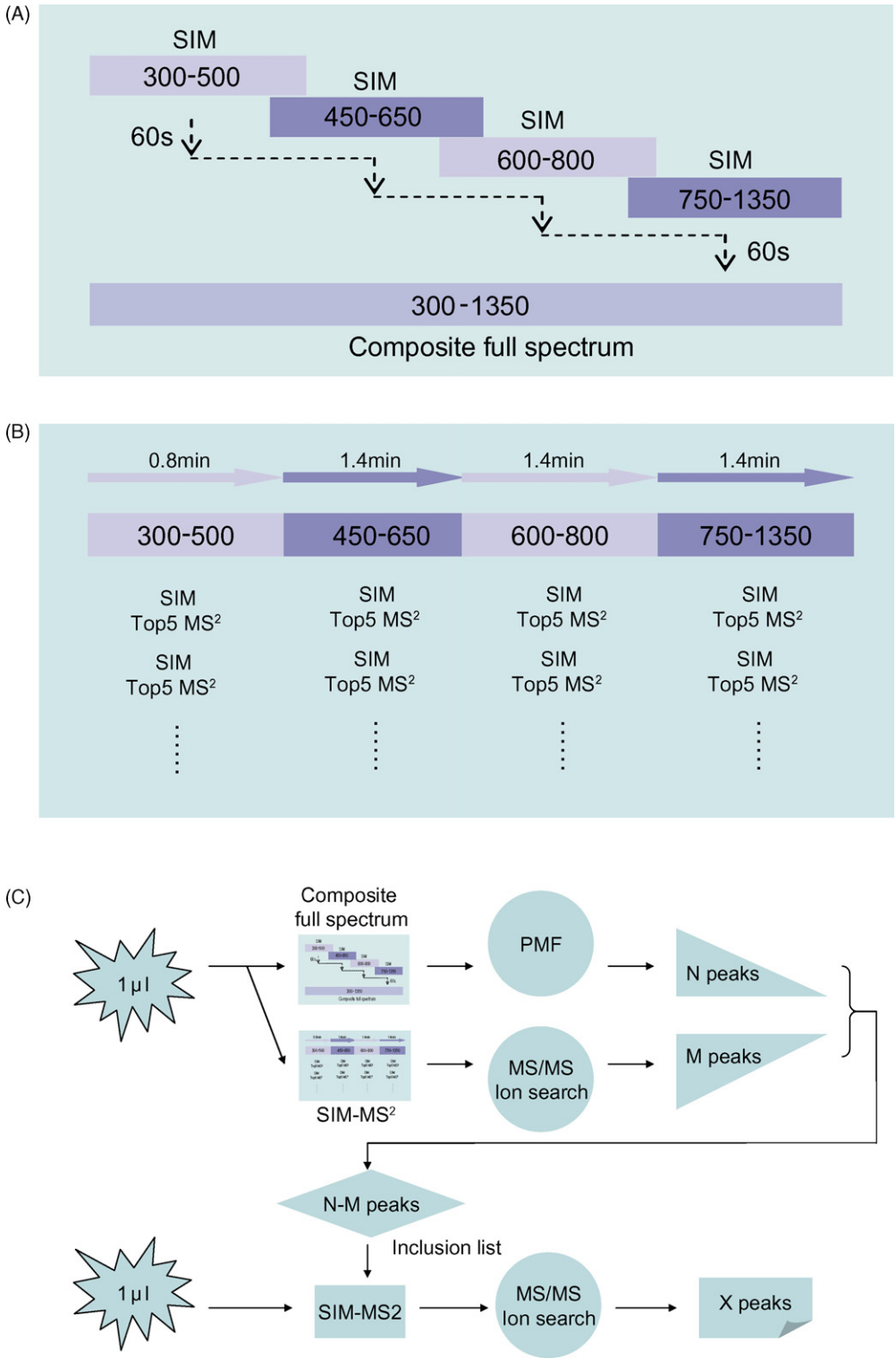


Fig. 2. Schematic description of the acquisition strategy as performed for protein characterization of tryptic digested BSA. (A) Selected ion monitoring (SIM) scans of multiple mass segments were repeatedly analyzed in the orbitrap mass spectrometer and these multiple SIM acquisitions were combined in one 'composite' spectrum. (B) After each SIM scan, the five most intense ions were data-dependently selected for MS/MS fragmentation in either ion-trap or orbitrap. (C) Overview of the complete protein characterization method, including the acquisition of the composite spectrum and the data-dependent SIM–MS^2 mode. These two methods are followed by a third strategy that aims to identify peptides that were not sequenced yet. In this directed SIM–MS^2 mode all precursor masses that were not fragmented so far are placed on an inclusion list. The entire experiment can be carried out with only 2 μl of very diluted sample solution.

Table 1
BSA sequence as identified by the three-step method introduced in this paper. In total 2 μl of trypsin digested BSA sample was used for every concentration. One microliter BSA was used to generate the composite SIM spectra and to carry out the data-dependent SIM–MS^2 method. Both methods were performed three times to assess their reproducibility. The second microliter was used specifically to characterize the peptides only identified by PMF but not sequenced yet. When all identification methods were combined, a sequence coverage of 66% was reached even for 500 attomole/μl

BSA concentrations	Acquisition methods	Sequence identified		
		per Individual Run (%)	per Acquisition Method (%)	Overall (%)
25 fmol/μl	Composite SIM	87	88	88
		79		
		81		
	SIM–MS^2	63	75	
		73		
		72		
	Inclusion list MS^2	1.4	1.4	
5 fmol/μl	Composite SIM	73	82	83
		77		
		78		
	SIM–MS^2	54	65	
		60		
		60		
	Inclusion list MS^2	10.1	10.1	
500 amol/μl	Composite SIM	45	60	66
		42		
		49		
		24	35	
	SIM–MS^2	20		
		24		
	Inclusion list MS^2	10.5	10.5	

mulation time for every segment would be similar. Each mass segment window was measured many times to gain sensitivity and precision. If several minutes are allocated to acquisition of the composite survey spectrum, then each mass segment is typically acquired more than 100 times.

First identification is based on peptide mass fingerprint (PMF) analysis. Because of the high mass accuracy of the orbitrap, particularly when including an internal mass standard in each spectrum (see below), dominant proteins in the sample are readily identified at this stage. In the second step data-dependent fragmentation is performed in each segmented range. Again SIM survey scans are recorded for each *m*/*z* range but now they are followed by ion-trap MS/MS spectra of the five most intense peaks. For each *m*/*z* range, the SIM–MS^2 cycle is repeated for a time adapted to the number of MS/MS candidates (Fig. 2B). Peptide identification is performed in the MS/MS ion search mode and peaks identified by PMF are confirmed by the MS/MS ion search. Since MS/MS spectra contain more information than the peptide mass alone, many peaks that cannot be identified only by the precursor mass are identified at this stage. This analysis still leaves some peptide peaks unfragmented—mainly because of their low signal, which may mean that they do not appear in every scan. These peaks are then targeted by a so-called 'inclusion list' in the third part of the measurement sequence. Fig. 2C presents an overview of the three-step sequence.

Since the acquisition methods for composite full spectra and SIM–MS^2 take only 4 and 5 min, respectively, 1 μl of sample sprays long enough to record both steps three times. A second microliter is used for step three in which we specifically target peaks not fragmented yet. Several microscans are applied for both MS (SIM) and MS/MS acquisition to boost sensitivity and data quality.

3.3. Subfemtomole sensitivity

Having established an efficient protocol for comprehensive protein characterization, we wanted to assess its sensitivity on a model protein. Using the strategy as described in Fig. 2C, we obtained a sequence coverage of more than 80% for BSA. The missing peptides were generally very short and some of them did appear under different digestion conditions as 'missed cleavage' peptides. We found that the BSA concentration could be diluted to 25 fmol/μl without losing protein sequence coverage (data not shown). Illustrating the excellent sensitivity of the set up described in this paper, more than 60% of the BSA sequence was still identified when the protein was diluted to 500 amol/μl (see Table 1). As shown in the table, the inclusion list SIM–MS^2 method turned out to be particularly advantageous for lower protein concentrations.

3.4. Extremely high mass precision in the composite full spectra

The orbitrap detection is based on inherently very precise frequency measurement and is, in our experience, much less affected by space charge than a Fourier-transform ion cyclotron

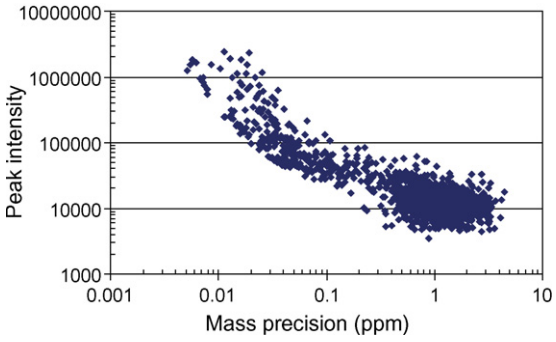


Fig. 3. Mass precision correlated to ion intensity. Plot of the intensity of different ions extracted from the composite spectrum vs. their precision, from thousands of scans depicted on a double logarithmic scale. Note that for most peaks we achieve sub ppm mass precision.

resonance (FT-ICR) instrument. When combined with 'lock mass' injection in every spectrum, the orbitrap is capable of achieving low to sub-ppm mass accuracy [14]. We reasoned that, by measuring each SIM mass range multiple times, the mass values of peaks extracted from the composite full spectra should become even more precise, since the standard deviation of a population is inversely dependent on the square root of the number of measurements. In our case, by repeated measurement of the same spectrum for 100 times, the standard deviation of the mass value should be decreased by a factor of 10. In Fig. 3, the peak intensity is plotted versus peak precision on a double logarithmic scale, and the data show that more intense ions yield better precision within the same measuring time. As can be seen in the figure, a large percentage of the peptides have a precision well below 1 ppm, with some peptides even reaching 100 ppb or less. Thus we conclude that the TriVersa–LTQ-Orbitrap combination is capable of extremely high mass accuracy, comparable or superior to any other platform currently used in proteomics.

3.5. More than 6000-fold dynamic range in the composite full spectra

Summing up multiple spectra filters out noise but boosts low intense ions that cannot be distinguished from background peaks in single spectra. We demonstrate further improvement of the dynamic range by collecting segmented *m/z* ranges instead of one full spectrum. Often a few, very intense ions comprise 90% of the total ion number of a spectrum, whereas in the segmented mass range regions with low intense signals are accumulated for a longer time in order to reach the same specified injection target value.

As depicted in Fig. 4A, the composite spectrum is much more feature rich in the higher mass range compared to the spectrum consisting of averaged full mass range spectra with the same total acquisition time. Both the composite spectra and the summed-up full spectra were acquired in the same experiment. As is apparent from the figure, the S/N was much better (see arrows). With the 25 fmol/μl tryptic-digested BSA sample, we assigned BSA peaks with a signal intensity difference of up to 6700 (Fig. 4B). While the dynamic range of the orbitrap is specified at 10^4, this value applies only to a simple two-component mixture. In our experience, dynamic range in complex peptide mixture analysis is around 10^3 in LC–MS experiments, so the composite spectra exhibit a comparable or superior dynamic range for complex samples to online experiments.

3.6. Sequence coverage comparable to LC–MS

We compared the sequence coverage obtained after 5 min of SIM–MS^2 acquisition with a conventional LC–MS run, both times using 50 fmol of BSA and the same parameter settings on the LTQ-Orbitrap. For this experiment, a higher flow rate chip (i.d. 5 μm) was used for nanoelectrospray, resulting in a flow rate of 200 nl/min. The LC–MS run took in total 53 min, of which the actual gradient lasted for 28 min. Both methods were performed twice. The sequence coverage obtained by these two methods was very comparable, 83.7% for nanoelectrospray SIM–MS^2 and 78.5% for LC–MS/MS. Detailed identification information for each tryptic peptide in the BSA sequence is shown in Fig. 5. Most of the peaks were identified by both methods, but some low intensity peaks were only sequenced in LC–MS. This may be due to the concentration effect of chromatography, where each ion elutes in a very short time span in contrast to the long but 'diluted' duration in nanoelectrospray. On the other hand, there were a few peptides that co-eluted with others and disappeared before having a chance to be fragmented in LC–MS but those were sequenced in the SIM–MS^2 run.

3.7. Characterization of Histone H3 post-translational modifications

Histones are the protein constituents of nucleosomes around which DNA is wound in eukaryotic cells. Histone tails on the nucleosome are subject to enzyme-mediated post-translational modifications (PTMs) of selected amino acids, such as lysine acetylation, lysine and arginine methylation, serine phosphorylation and attachment of ubiquitin [15,16]. These modifications, singly or in combination, are thought to generate an epigenetic code that specifies different patterns of gene expression and silencing [17]. Characterization of post-translational modifications on bulk histones by mass spectrometric approaches has proven to be very successful as recently reviewed by Hunt and co-workers [18]. Here we investigate the suitability of the nanoelectrospray–orbitrap combination to characterize modifications on histone H3 purified from human HeLa cells, separated from other histone molecules by RP-chromatography and in-solution digested with Arg-C protease. In order to distinguish between several modifications present on such molecules, many of them only differing in single methyl or acetyl groups, we also employed high resolution read out of MS/MS spectra in the orbitrap. Furthermore, fragmentation was performed by higher energy injection into the C-trap [19], which leads to triple-quadrupole like behavior and preservation of the full mass range in the MS/MS spectra. Fragmentation spectra were acquired at a resolution of 30,000 and the mass accuracy was in the low ppm range for these fragmentation spectra.

More than 500 peaks were extracted from the composite full spectra. Spectra recorded in SIM–MS^2 mode were searched in

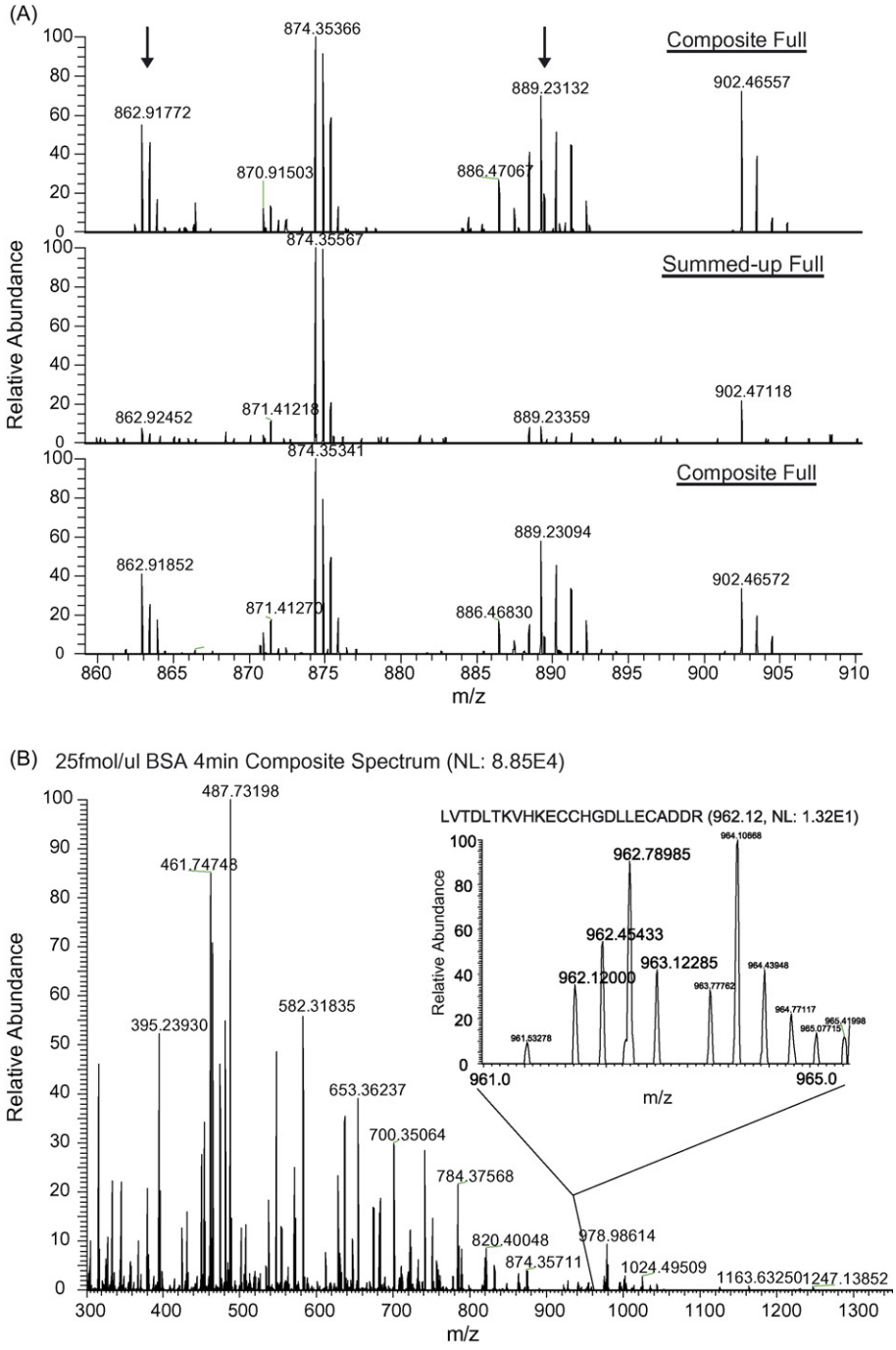


Fig. 4. Advantages of composite spectra. (A) In the composite spectrum ions of low intensity were boosted in comparison to full spectra summed-up for the same time span. 500 attomole/μl BSA sample was sprayed, and the composite spectra and the summed-up full spectra were acquired directly after one another. Whereas for intense ions there is no large visible difference, the S/N ratio for low intense ions increases dramatically. (B) A dynamic range of over 6000-fold was obtained in the composite spectrum. 25 fmol/μl BSA resulted in a composite spectrum, in which the most intense peptide with *m/z* of 487.73 (2+, DLGEEHFK) had an intensity of 8.85E4, while the peptide at *m/z* 962.12 (3+, LVTDLTKVHKECCHGDLLECADDR) was observed with an intensity of 13.2 (inset), resulting in a dynamic range of 6700.

a histone database and 46% of the sequence of histone H3 was identified (based on peptides with a score higher than $p < 0.05$). The heavily modified N-terminal sequence was completely covered and seven differently modified peptides were detected (Table 2). In six of them the modified residues were unambiguously determined. In particular, the high mass accuracy of the orbitrap allowed easy distinction between trimethylation and acetylation, both of which are important histone modifications that have the same nominal mass. In the seventh peptide, the trimethylated and acetylated peptide KSTGGKAPR, the modified sites could not unambiguously be assigned to the sequence since the fragmentation was performed in the LTQ and the mass difference between these modifications (0.0364 Da) is far less than what the LTQ can distinguish. Therefore, a second

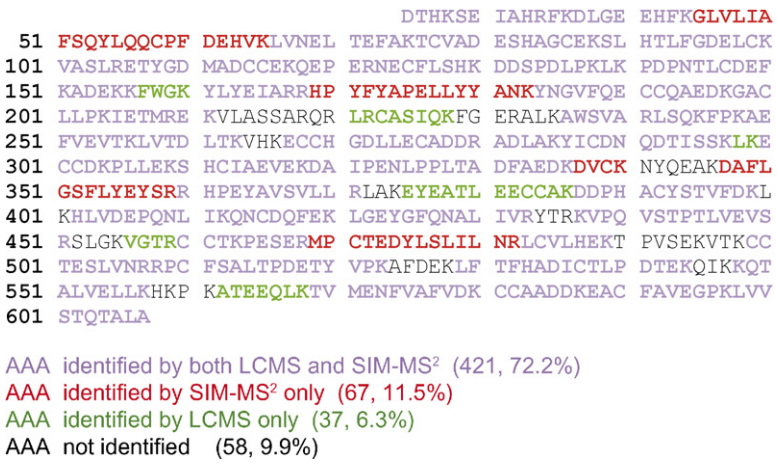


Fig. 5. Comparison of the BSA sequence coverage obtained in SIM–MS^2 vs. LC–MS/MS mode. In both experiments 50 fmol of BSA solution were used and both measurements were performed with the same instrument settings. Peptides in blue are identified in both methods, red peptides only in SIM–MS^2 and green peptides only in LC–MS mode. Black peptides were not identified.

Table 2
Histone H3 modifications (combinations) identified by the SIM–MS^2 method

Modified amino acids	Sequence
K4 monomethylation	TKQTAR
K9 dimethylation + K14 acetylation K9 trimethylation + K14 acetylation	KSTGGKAPR
K23 acetylation K18 acetylation + K23 acetylation	KQLATKAAR
K79 monomethylation K79 dimethylation	EIAQDFKTDLR

microliter of the sample was sprayed and C-trap fragmentation was performed combined with recording of the MS/MS spectra in the orbitrap. With the resulting high mass accuracy, both types of modifications were confirmed and trimethylation was assigned to K9 and acetylation to K14.

The reported seven modified peptides are relatively abundant in the bulk preparation histone sample, and have already been reported by either top-down [20] or LC-based bottom-up [18] mass spectrometric method. Our approach provides an alternative way to characterize histone by bottom up mass spectrometric analysis without online LC separation. Preliminary work furthermore indicates that modifications on short peptides that escape detection by LC–MS/MS can be detected by automated nanoelectrospray (data not shown).

3.8. Relative quantitation by deconvoluted peak intensity

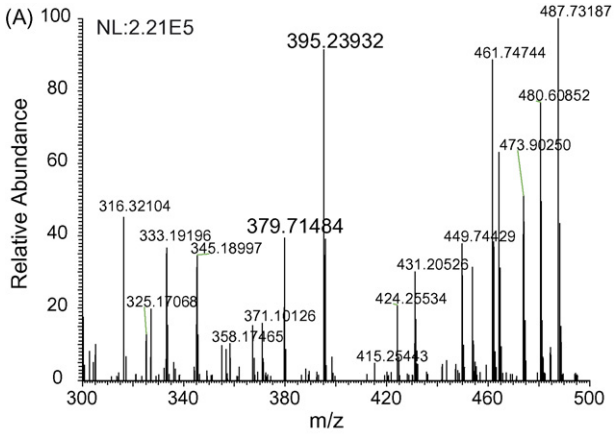
As mentioned above, quantification of peptides and proteins is becoming more and more important. An advantage of the acquisition of a large number of spectra is the increasing precision, not only of the mass value but also of the intensity ratio between ions. This will be illustrated with the ratio between two different BSA peptides as well as with the ratio between a methylated and non-methylated histone peptide.

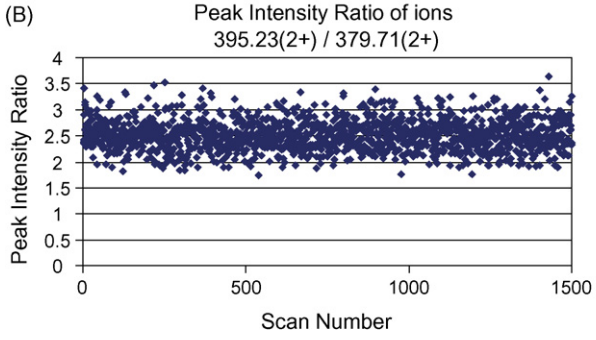
Fig. 6 illustrates the relative quantitation of BSA peptides. Panel A shows the *m*/*z* segment 300–500. Fig. 6B reveals that the ratio of the relative intensities of the peptides with *m*/*z* 395.239 (2+) and 379.715 (2+) varied per spectrum between 1.7 and 3.5. However, with increasing number of accumulated spectra quantitation becomes more and more precise. As shown in Fig. 6C, the 99% confidence interval for quantitation decreases from 13% after accumulating 10 scans to 0.9% after accumulation of 1500 scans (15 min acquisition).

In the case of histone H3 we investigated quantitation of the normal peptide against a slightly modified form. This is possible in nanoelectrospray since both peptides are present in the same scan. In order to avoid inaccuracy due to transmission 'edge effects' in the SIM windows, we chose to quantify based on the full spectrum. The relative quantitation values for identified histone H3 peptide pairs are listed in Table 3. The amounts of several methylated peptides were about 10-fold less than those of the unmodified peptides. Note that this value gives a general idea

Table 3
Quantitation values of identified peptide pairs with the same sequence but different modifications

Identified peptide pairs	Quantitation value with 99% confidence interval
TKQTAR/TK$_{(methyl)}$QTARK	11.791 ± 0.916
K$_{(dimethyl)}$STGGK$_{(acetyl)}$APR/K$_{(trimethyl)}$STGGK$_{(acetyl)}$APR	9.849 ± 0.826
KQLATK$_{(acetyl)}$AAR/KQLATK$_{(diacetyl)}$AAR	14.651 ± 1.307
EIAQDFKTDLR/EIAQDFK$_{(methyl)}$TDLR/ EIAQDFK$_{(methyl)}$TDLR/EIAQDFK$_{(dimethyl)}$TDLR	11.391 ± 0.868 8.206 ± 0.990





(C)

Number of scans	Average of peak intensity ratios	quantitation ratio with 99% confidence interval
10	2.748	2.748 ± 0.354
50	2.623	2.623 ± 0.124
100	2.570	2.570 ± 0.086
200	2.546	2.546 ± 0.058
300	2.524	2.524 ± 0.051
400	2.526	2.526 ± 0.047
500	2.513	2.513 ± 0.039
1000	2.511	2.511 ± 0.027
1500	2.513	2.513 ± 0.022

Fig. 6. Highly accurate relative quantitation by the ratio of ion intensities within one segmented spectrum. Segmented SIM scans from *m/z* 300 to 500 were acquired for 15 min with 25 fmol/μl tryptic BSA. (A) A single SIM scan including the ions with *m/z* 395.239 (2+) and 379.715 (2+) of which the ratios of intensities were quantified. Per single scan the ratio varies from 1.5 to 3.5 (B). With an increasing number of accumulated spectra, the precision of the quantitation ratio increases, as shown in (C).

of the absolute stoichiometry of this methylation site but that it needs to be corrected for the different ionization efficiencies of the modified versus the unmodified peptides [21].

4. Conclusions and perspectives

In this paper, we have endeavored to revive nanoelectrospray, an 'old' protein mapping method using no chromatographic peptide separation, as an alternative to MALDI peptide mapping. Using the advantages of a stable spray in combination with the LTQ-Orbitrap mass spectrometer, we have introduced the concept of 'composite spectra', which are spectra composed of a high number of segmented SIM scans. These composite spectra allow very high sensitivity, accuracy and dynamic range due to optimized C-trap fill times for each mass segment. In the automated format of the TriVersa Nanomate, nanoelectrospray measurements are robust, user-friendly and easily amenable for different protein samples while using very low amounts of spraying solution. Since the instrument can readily switch between MS and MS/MS, and between fragmentation in the linear ion-trap and the C-trap, it offers a large number of complementary methods and is very flexible and user-friendly.

The LTQ-Orbitrap mass spectrometer is a dual instrument with two independent detection systems (orbitrap and LTQ), which can be operated simultaneously, thus the ideal combination would be to fragment high intense ions in the orbitrap while simultaneously performing ion-trap fragmentation of low intensity ions. This would increase the duty cycle and analysis speed. However, this requires direct access to the LTQ-Orbitrap acquisition software, which we are currently lacking. At the other extreme of acquisition sophistication is the so-called 'ion-mapping' technique. In this method, the whole mass range is scanned step by step by SIM-scans of for example 6 Da windows with and without applying collision energy to fragment the ions in this small window. Especially for complex mixtures this method could in principle be very valuable in a nanoelectrospray setup since the dynamic range is expected to be further increased by these small segments. However, when we tried this method we found that it allocates too much time to 'empty' regions and is thus overall less efficient than the method described here.

In order to further improve protein characterization, we plan to access to the LTQ-Orbitrap acquisition software directly and perform genuine 'real time' data acquisition. Fragmentation (MS^2 or MS^3) will focus on the peaks recognized in the survey scan but not identified in the search with expected variable modifications. This will allow identification of new peptide sequences, variant alleles or unexpected modifications. Of course, digestion with multiple enzymes is also an obvious next step for even more for more in-depth characterization of modified peptides.

Acknowledgments

The authors thank other members of the department for Proteomics and Signal Transduction and of Advion Biosciences for constructive comments and discussion. This work was partially supported by the EU research directorate in the 'Interaction Proteome' grant (LSHG-CT-2003-505520) and in the HEROIC grant (LSHG-CT-2005-018883).

References

[1] R. Aebersold, M. Mann, Nature 422 (2003) 198.
[2] M. Mann, O.N. Jensen, Nat. Biotechnol. 21 (2003) 255.
[3] O.N. Jensen, Nat. Rev. Mol. Cell. Biol. 7 (2006) 391.
[4] M. Kussmann, P. Roepstorff, Meth. Mol. Biol. 146 (2000) 405.
[5] J. Yates, A. McCormack, D. Schieltz, E. Carmack, A. Link, J. Protein Chem. 16 (1997) 495.
[6] J. Peng, S.P. Gygi, J. Mass Spectrom. 36 (2001) 1083.
[7] S.E. Ong, M. Mann, Nat. Chem. Biol. 1 (2005) 252.

[8] H. Steen, M. Mann, Nat. Rev. Mol. Cell. Biol. 5 (2004) 699.
[9] M. Wilm, M. Mann, Anal. Chem. 68 (1996) 1.
[10] M. Wilm, A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, M. Mann, Nature 379 (1996) 466.
[11] A. Makarov, E. Denisov, A. Kholomeev, W. Balschun, O. Lange, K. Strupat, S. Horning, Anal. Chem. 78 (2006) 2113.
[12] J. Rappsilber, Y. Ishihama, M. Mann, Anal. Chem. 75 (2003) 663.
[13] R. Zubarev, M. Mann, Mol. Cell. Proteomics 6 (2007) 377.
[14] J.V. Olsen, L.M. de Godoy, G. Li, B. Macek, P. Mortensen, R. Pesch, A. Makarov, O. Lange, S. Horning, M. Mann, Mol. Cell. Proteomics 4 (2005) 2010.
[15] H.T. Spotswood, B.M. Turner, J. Clin. Invest. 110 (2002) 577.
[16] T. Kouzarides, Cell 128 (2007) 693.
[17] T. Jenuwein, C.D. Allis, Science 293 (2001) 1074.
[18] B.A. Garcia, S.B. Hake, R.L. Diaz, M. Kauer, S.A. Morris, J. Recht, J. Shabanowitz, N. Mishra, B.D. Strahl, C.D. Allis, D.F. Hunt, J. Biol. Chem. 282 (2007) 7641.
[19] J.V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, M. Mann, Nat. Meth., in press.
[20] C.E. Thomas, N.L. Kelleher, C.A. Mizzen, J. Proteome Res. 5 (2006) 240.
[21] H. Steen, J.A. Jebanathirajah, J. Rush, N. Morrice, M.W. Kirschner, Mol. Cell. Proteomics 5 (2006) 172.