RESEARCH ARTICLE

Increased proteome coverage by combining PAGE and peptide isoelectric focusing: Comparative study of gel-based separation approaches

Ilian Atanassov[1] *and Henning Urlaub*[1,2]

[1] Bioanalytical Mass Spectrometry Group, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany
[2] Bioanalytics, Department of Clinical Chemistry, University Medical Center, Göttingen, Germany

The in-depth analysis of complex proteome samples requires fractionation of the sample into subsamples prior to LC-MS/MS in shotgun proteomics experiments. We have established a 3D workflow for shotgun proteomics that relies on protein separation by 1D PAGE, gel fractionation, trypsin digestion, and peptide separation by in-gel IEF, prior to RP-HPLC-MS/MS. Our results show that applying peptide IEF can significantly increase the number of proteins identified from PAGE subfractionation. This method delivers deeper proteome coverage and provides a large degree of flexibility in experimentally approaching highly complex mixtures by still relying on protein separation according to molecular weight in the first dimension.

Received: January 22, 2013
Revised: June 7, 2013
Accepted: June 17, 2013

Keywords:
IPGs / PAGE / Peptide IEF / Shotgun proteomics / Technology

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Large-scale proteome profiling based on shotgun MS analysis has become a widely used and indispensable tool for modern proteomics research. In typical shotgun proteomic workflows, proteins isolated directly from biological samples are hydrolyzed to peptides with proteases, and peptides are then identified by RP-HPLC coupled to MS/MS. The success of shotgun proteomic studies largely depends on the resolution of the peptide separation step prior to MS analysis and on the MS peak capacity. Peptide separation by RP-HPLC alone does not have the resolution power to separate all peptides derived from complex proteome samples. Moreover, even current, state-of-the-art MS instruments can handle only a limited number of peptides simultaneously, in terms of precursor selection and subsequent fragmentation. Thus, to further increase the overall number of identified proteins and proteome coverage, experimental samples must be partitioned into subsamples with lower complexity prior to LC-MS/MS analysis. Accordingly, several strategies operating at the level of proteins or peptides have been developed to reduce sample complexity, with the aim of obtaining the "deepest" possible proteome coverage. Reducing sample complexity at the protein level usually employs the so-called GeLC-MS/MS technique, which involves 1D SDS-PAGE separation of proteins and gel fractionation prior to protein digestion and LC-MS/MS. The "peptide strategy" employs various peptide separation techniques, including strong cation exchange chromatography that is performed either on-line (MudPit) [1] or off-line, IEF of peptides (pIEF) on IPG strips [2, 3], and anion exchange chromatography of peptides [4]. These studies make it evident that, in order to increase proteome coverage, a substantial reduction of sample complexity is required; this can be obtained for example by subfractionation of samples.

Here, we demonstrate how to increase the depth of shotgun proteome analyses by using a 3D approach combining 1D PAGE of proteins with pIEF and LC-MS/MS (PAGE-pIEF-LC-MS/MS) in order to reduce sample complexity. We compare the results obtained from this method with those from

Correspondence: Professor Henning Urlaub, Bioanalytical Mass Spectrometry Group, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany
E-mail: henning.urlaub@mpibpc.mpg.de
Fax: +49-551-201-1197

Abbreviations: FA, formic acid; **MW**, molecular weight; **NE**, nuclear extract; **pIEF**, peptide IEF

methods of either GeLC-MS/MS or pIEF-LC-MS/MS. Importantly, the PAGE-pIEF-LC-MS/MS method presented here led to a significant increase in the proteome coverage and allowed us to successfully identify low-abundance proteins. Moreover, it provides the unique advantage that not necessarily the entire proteome has to be investigated since the use of prefractionation according to the molecular weight (MW) (e.g. by PAGE) allows investigation of only a subset of proteins, e.g. low MW proteins. We analyzed a complex proteomic sample of nuclear extract (NE) from HeLa cells. First, we employed two common separation strategies, of PAGE and pIEF with LC-MS/MS, for in-depth profiling. Next, we coupled both separation techniques in such a manner that first the complexity of the protein mixture was reduced by 1D PAGE, then the peptides derived from each PAGE slice after in-gel digestion were separated by pIEF, extracted from the pIEF strip and, finally analyzed by LC-MS/MS (PAGE-pIEF-LC-MS/MS). We also examined the reproducibility of GeLC-MS, pIEF-LC-MS/MS, and PAGE-pIEF-LC-MS/MS.

2 Materials and methods

2.1 Sample preparation and GeLC-MS/MS analysis

HeLa cell NE was prepared according to Dignam et al. [5]. NE mixture corresponding to 70 μg of protein was precipitated with ethanol [6], and the resulting pellet was dissolved in 30 μL 1× LDS sample buffer (Invitrogen) and separated on a precast 4–12% NuPAGE Novex Bis-Tris gel (Invitrogen). Colloidal Coomassie-stained gels were fractioned into 23 slices using a tool developed in-house [7], and all slices that showed Coomassie staining (slices 2–22) were reduced, alkylated, and digested overnight with trypsin (Roche) as described previously [7]. Tryptic peptides were extracted, dried and redissolved in 20 μL 5% ACN with 0.1% formic acid (FA); triplicate injections of 5 μL each of the peptide solution were then analyzed by LC-MS/MS.

2.2 pIEF-LC-MS/MS analysis

For the pIEF workflow, 70 μg of ethanol-precipitated, HeLa NE was reduced, alkylated, and digested in-solution according to [6]. Briefly, the protein pellet was dissolved in 8 M urea and digested with Lys-C (Roche; 1:50 protease to protein ratio) for 3 h. The sample was then diluted with 100 mM ammonium bicarbonate solution to 2 M urea and digested with trypsin (Promega; 1:50) overnight at 25°C. The resulting peptides were desalted using STAGE tips [8]. For IEF, dried peptides were dissolved in 350 μL of 8 M urea, 0.2% IPG buffer (GE Healthcare), and then applied to ceramic strip holders. We used 18 cm IPG strip (pH 3–10; GE Healthcare). After pipetting the sample on top of the strip holder and wetting the 18 cm IPG strip by lowering and raising it three times, the strip was positioned on the incline at the cathodic end of the strip holder. The strip was then pushed with forceps, sliding it down the incline and moving it left towards the anode until the IPG gel reached the bottom of the strip holder (Supporting Information Fig. 1). Next, the IPG strip was covered with around 1.5 mL cover fluid (Pharmacia Biotech) and rehydrated overnight. Peptides were separated on an IPGphor (Pharmacia Amersham) for a total of 30 000 Vh at max 50 μA per strip using the following parameters: overnight rehydration at 0 V then 500 Vh at 500 V, then 1750 Vh at gradient 500 V to 3000 V, then 27750 Vh at 8000 V all steps performed at 20°C. After pIEF, IPG strips were cleaned of the cover fluid by immersing them for 10 s in *n*-hexane (Merck) and manually sliced by putting the strip on top of graph paper and striping the gel from the plastic into 36 pieces (0.5 cm each). Peptides were then extracted by sequentially incubating the gel slices for 30 min each with 1% FA; 50% ACN, 1% FA; and 99% ACN, 1% FA. Extracted peptides were desalted over STAGE tips. Peptides were finally dissolved in 20 μL 5% ACN with 0.1% FA; triplicate injections of 5 μL each of the peptide solution were then analyzed by LC-MS/MS.

2.3 PAGE-pIEF-LC-MS/MS analysis

A total of 70 μg HeLa NE was fractioned by 1D PAGE and analyzed in parallel with the GeLC-MS/MS approach described above. All steps were identical, except that peptides extracted from the PAGE slices were subsequently separated on a 7 cm (instead of a 18 cm, see above) IPG strip (pH 3–10, GE Healthcare) prior to LC-MS/MS. Tryptic peptides were extracted from PAGE slices, dried, dissolved in 150 μL solution of 8 M urea, 0.2% IPG buffer and separated on IPGphor for a total of 9950 Vh [9]. IEF was performed using the following parameters: max 50 μA per strip, overnight rehydration at 0 V then 250 Vh at 500 V, then 6500 Vh at gradient 500 V to 6000 V, and then 3200 Vh at 6000 V. All steps were performed at 20°C. Following pIEF, the IPG gel was fractioned into 13 slices (0.5 cm each), and peptides were extracted and desalted as described in the previous section. Dried peptides were dissolved in 16 μL 5% ACN with 0.1% FA; 8 μL of this was analyzed by LC-MS/MS. For the pIEF reproducibility studies peptides were dissolved in 20 μL 5% ACN with 0.1% FA from which 5 μL was injected and analyzed by LC-MS/MS

2.4 LC-MS/MS analysis

LC separation was performed on an Agilent 1100-LC system (Agilent Technologies). Peptides were loaded onto a trap column packed in-house (2 cm, 360 μm od, 150 μm id; ReproSil-Pur, C18, AQ 5 μm, Dr. Maisch HPLC GmbH) and separated at a flow rate of 300 nL/min on a 12 cm analytical column (75 μm id packed with ReproSil-Pur C18 AQ 5 μm resin; Dr. Maisch, Germany) prepared in-house. Separation was performed using a gradient mixture of 0.1% FA in water (buffer A) and 95% ACN, 0.1% FA (buffer B). A 33-min,

5–38% buffer B gradient was used. Eluting peptides were analyzed on a hybrid LTQ-Orbitrap-XL (ThermoScientific), operating in a data-dependent mode. Survey full scan MS spectra were acquired in the orbitrap (m/z 350−1600) with a resolution of 30 000 at m/z 400 and an automatic gain control target at 10^6. The five most intense ions with were selected for CID MS/MS fragmentation in the linear ion trap at an automatic gain control target of 30 000. Ions with unrecognized charge state and with charge state one were excluded. Detection in the linear ion trap of previously selected ions was dynamically excluded for 60 s. Internal calibration of the orbitrap was performed using the lock mass option [10].

2.5 Data analysis

Proteins were identified with MaxQuant [11] (version 1.0.13.13), using the Mascot search engine (version 2.3.03, Matrix Science) with the IPI human protein sequence database (version 3.72). The database was concatenated with typical contaminants and reversed sequences using the SequenceReverser.exe tool from the MaxQuant package. Default parameters were used for MaxQuant (7 ppm MS tolerance, 0.5 Da MS/MS tolerance, maximum two missed cleavages, 1% protein and peptide false discovery rate), except that the option "keep low scoring version of identified peptides" was turned off. Methionine oxidation and cysteine carbamidomethylation were set as variable and fixed modifications, respectively. The raw files from the GeLC-MS/MS reproducibility experiment were analyzed with MaxQuant (version 1.3.0.5) using the Andromeda search engine [12] again using the IPI human protein sequence database (version 3.72). Same (default) parameters were used as for MaxQuant version 1.0.13.13.

3 Results and discussion

3.1 Performance of GeLC-MS/MS versus pIEF-LC-MS/MS

We initially evaluated the performance of GeLC-MS/MS and pIEF-LC-MS/MS. Overall, 3540 proteins were identified based on 38 657 peptides (GeLC-MS/MS) compared to 3945 proteins and 31 113 peptides (pIEF-LC-MS/MS). We also examined the separation efficiency of both methods. For pIEF-LC-MS/MS, 79% of all identified peptides in all 36 pIEF slices were unique and found in one pIEF slice only, demonstrating the high resolving power of tryptic peptides separated by pIEF. For the GeLC-MS/MS approach, 66% of all identified proteins were unique and identified in a single PAGE slice only, while 85% of all identified peptides in all PAGE slices were unique and identified in a single PAGE slice (Supporting Information Fig. 2). We could demonstrate that the PAGE fractionation is efficient for protein separation and results in peptides that are nonredundant between PAGE slices, and that pIEF can indeed efficiently separate tryptic peptides derived from a complex sample in the pH 3–10 region.

3.2 pIEF of peptides from a single PAGE subfraction

Next, we coupled PAGE and pIEF in such a manner that first the complexity of the protein mixture was reduced by 1D PAGE, then the peptides derived from each PAGE slice, after in-gel digestion, were separated by pIEF, extracted from the pIEF strip and, finally analyzed by LC-MS/MS (Fig. 1A). This combined PAGE-pIEF-LC-MS/MS approach significantly increased the proteome coverage, as evidenced by comparing the protein and peptide sets identified from, for instance, PAGE slice #10 (MW region ∼75 kDa) either analyzed directly by LC-MS/MS (PAGE#10-LC-MS/MS) or with an additional peptide separation by pIEF prior to LC-MS/MS (PAGE#10-pIEF-LC-MS/MS; Fig. 1A). Including the pIEF step more than doubled the total number of proteins identified, and these included almost all of the proteins previously identified as well as a large number of additional proteins that had escaped identification before (Fig. 1B). MS intensity has previously been shown to be a good proxy for estimating protein abundance [13]. Intriguingly, proteins unique to our PAGE-pIEF-LC-MS/MS approach had lower intensity (summed extracted ion chromatograms) distributions as compared with those identified by both methods (Fig. 1B), indicating that the pIEF step subsequent to PAGE helped to detect lower abundance proteins. While only around 1200 peptides were identified by LC-MS/MS, performing pIEF doubled the total number of identified peptides (Fig. 1C), which resulted in an increase of the median sequence coverage of all proteins common to PAGE#10-LC-MS/MS and PAGE#10-pIEF-LC-MS/MS from 12 to 23% (Supporting Information Fig. 3). Peptides identified in PAGE#10 after pIEF-LC-MS/MS were compared with those identified by LC-MS/MS only and divided into three groups, based on the proteins to which they mapped (Fig. 1B): (i) common peptides mapping to common proteins i.e., those peptides identified by both PAGE#10-LC-MS/MS and PAGE#10-pIEF-LC-MS/MS (1861 peptides matching to the 291 common proteins); (ii) newly identified peptides identified by PAGE#10-pIEF-LC-MS/MS only (2484) that mapped to one of the 291 common proteins; and (iii) newly identified peptides (1139) that did not map to one of 291 common proteins but match to one of the 359 proteins that were identified, by PAGE-pIEF-LC-MS/MS only. All three peptide groups showed similar orthogonality plots of their LC retention time and IPG gel slice coordinates (Supporting Information Fig. 4) [14]. Thus, extensive prefractionation provided the mass spectrometer with enough time to sequence peptides coming from lower abundance proteins, while assuring that these peptides were efficiently separated from those derived from higher abundance proteins.

We also tested whether performing a longer LC gradient, a standard step when one is aiming at maximizing protein identifications from a sample/subfraction, would deliver a larger

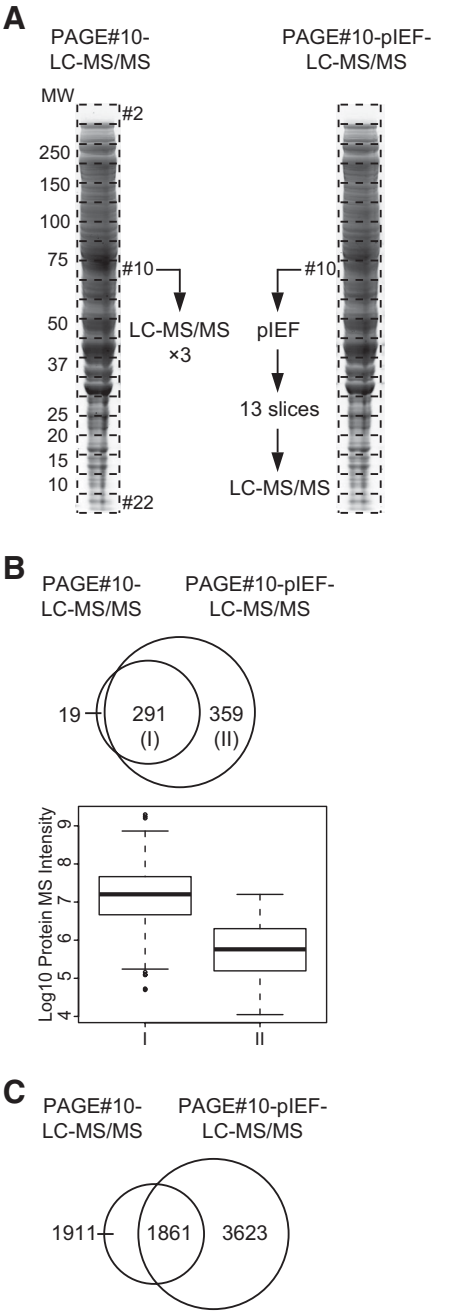

Figure 1. Proteome profiling by PAGE-pIEF-LC-MS/MS. (A) Schematic representation of the workflow. Proteins were first separated according to MW by 1D PAGE. Gels were fractionated, and proteins from gel slice #10 of the 1D PAGE were digested in gel and then analyzed by LC-MS/MS (PAGE#10-LC-MS/MS) or, after additional separation of the extracted peptides, by pIEF prior to LC-MS/MS (PAGE#10-pIEF-LC-MS/MS). (B) Overlap between proteins identified in PAGE#10-LC-MS/MS and PAGE#10-pIEF-LC-MS/MS and box plots showing the MS intensity distribution of "common" proteins, i.e. those identified by both PAGE#10-LC-MS/MS and PAGE#10-pIEF-LC-MS/MS (I) and of the "newly identified" proteins, i.e. those identified only when pIEF was also performed (II). (C) Overlap of identified peptides between PAGE#10-LC-MS/MS and PAGE#10-pIEF-LC-MS/MS.

number of protein/peptide identifications as compared with 7 cm pIEF-LC-MS/MS. For this purpose, we separated two samples of 70 μg HeLa NE by PAGE in separate lanes. Peptides extracted again from PAGE slice #10 were dried and redissolved in 22 μL 5% ACN with 0.1% FA; two injections of 5 μL each of the peptide solution were then analyzed using a 33 min LC gradient and two injections of 5 μL each using 221 min LC gradient; 5–38% buffer B. In parallel, total peptides derived from PAGE slice #10 from the second lane were separated using pIEF. Tryptic peptides were extracted from the pIEF strip and desalted. Dried peptides were redissolved in 22 μL 5% ACN with 0.1% FA; two injections of 5 μL each of the peptide solution were then analyzed by LC-MS/MS. Prolonging the LC gradient increased the number of identified proteins (425 compared with 238) and peptides (3792 compared with 2141). Importantly, pIEF of the same amount of extracted peptides clearly outperformed both separation steps in terms of identified proteins (741) and peptides (8606) (Supporting Information Fig. 5). While the analytical time for the pIEF approach was longer (around 11 h including column equilibration time) than employing both LC gradients (approximately 4 h for the longer gradient), the method clearly delivered a larger number of protein/peptide identifications.

3.3 Comparison between GeLC-MS/MS, pIEF-LC-MS/MS, and PAGE-pIEF-LC-MS/MS

We next performed pIEF of the peptides derived from the remaining PAGE slices and compared the results between all three proteomics workflows, namely, GeLC-MS/MS, pIEF-LC-MS/MS, and the novel PAGE-pIEF-LC-MS/MS. The number of proteins and peptides identified from each of these methods is shown in the schematic representation of the workflows (Fig. 2A). The differences in the number of identifications between the three approaches reflect the depth of the analysis and thus the proteome coverage. Approximately 3540 and 3940 proteins were identified by GeLC-MS/MS and pIEF-LC-MS/MS, respectively, and 5260 proteins by our PAGE-pIEF-LC-MS/MS method (Supporting Information Table 1); thus, there is a clear improvement provided by the additional dimension of separation. Due to the higher number of identified peptides, the median protein sequence coverage for all proteins common between GeLC-MS/MS and PAGE-pIEF-LC-MS/MS increased from 17% (GeLC-MS/MS) to 26% for PAGE-pIEF-LC-MS/MS. While there was a large overlap in the proteins identified by all three approaches (of around 3050 proteins), more than 1100 proteins were identified only after performing PAGE-pIEF-LC-MS/MS (Fig. 2B). Note that those 1100 proteins had a much lower MS intensity and emPAI [15] distribution as compared with those identified in all three workflows (Fig. 2C, Supporting Information Fig. 6), emphasizing that lower abundance proteins are more readily identified by PAGE-pIEF-LC-MS/MS approach than by either of the two other approaches. There were a number of proteins that could not be identified by our

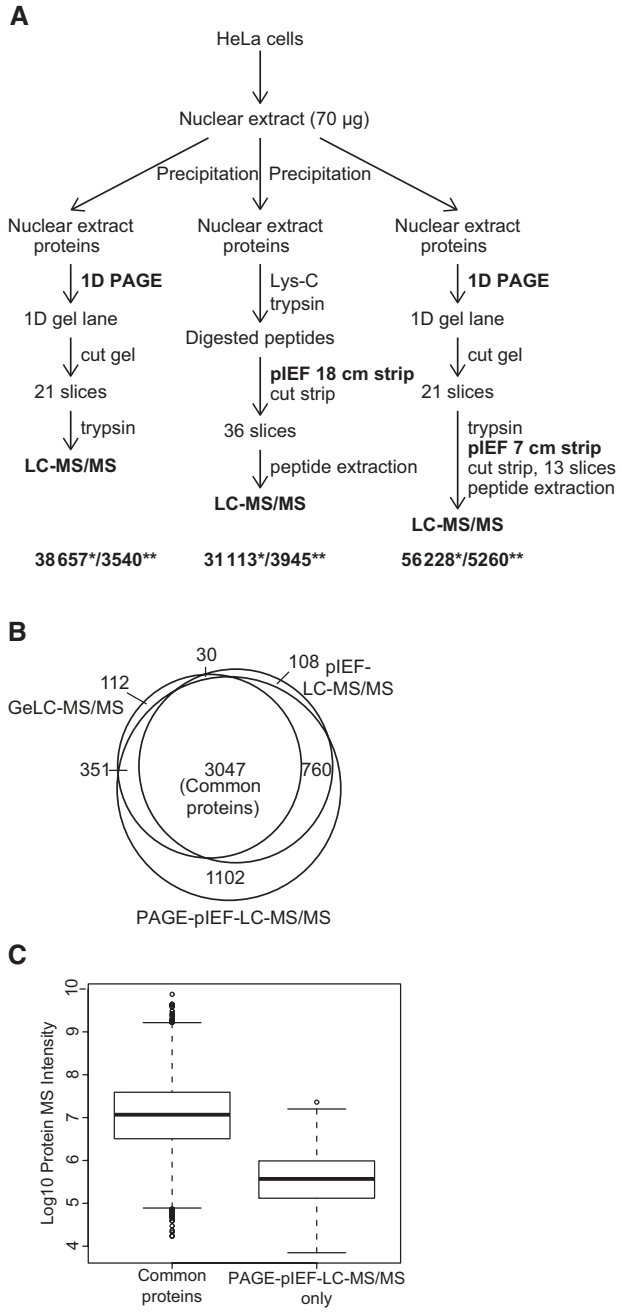

Figure 2. (A) Schematic presentation of the three workflows employed in the study, showing the number of identified peptides (*) and proteins (**). (B) Overlap of the proteins identified by GeLC-MS/MS, pIEF-LC-MS/MS, or PAGE-pIEF-LC-MS/MS. (C) Intensity distributions of proteins identified in all three workflows (common proteins) and those identified only by PAGE-pIEF-LC-MS/MS.

PAGE-pIEF-LC-MS/MS method. We therefore calculated the number of peptides that map to a protein in each of the protein groups not identified by any of the other two separation methods. We found that more than 90% of the 112 proteins identified only by GeLC-MS/MS and of the 108 proteins identified only by pIEF-LC-MS/MS were based on a single peptide. In comparison, of the 1100 proteins identified only by PAGE-pIEF-LC-MS/MS, 41% were identified by one peptide only, while the rest were identified by at least two peptides. This showed that, while more proteins of lower abundance were identified using our PAGE-pIEF-LC-MS/MS approach, most of those identifications were of higher confidence and based on more than one peptide.

3.4 Overall separation efficiency of PAGE and IEF

The success of the PAGE-pIEF-LC-MS/MS method depends on its performance in each dimension. For all except one of the 7 cm pIEF runs, more than 70% of the peptides were unique and identified in a single IPG gel slice only (Supporting Information Table 2), indicating that the peptide separation by pIEF was very efficient. In a PAGE slice-by-slice comparison of GeLC-MS/MS with PAGE-pIEF-LC-MS/MS, we saw a significant increase in the number of identified peptides and proteins, after applying pIEF, from almost all PAGE slices and protein MW regions (Supporting Information Fig. 7 and Supporting Information Table 3). We also noted that the peptides from the proteins newly identified by PAGE-pIEF-LC-MS/MS were spread throughout the pH range of the 7 cm IPG strips and the LC gradients, indicating efficient use of the separation space (Supporting Information Figs. 8 and 9). Regarding the protein separation in the first dimension (PAGE), 53% of the proteins and 73% of the peptides were uniquely identified in a single PAGE slice only (Supporting Information Fig. 10). These identifications showed greater redundancy when compared to GeLC-MS/MS (with 62% of the proteins and 83% of the peptides identified in a single PAGE slice only, see Supporting Information Fig. 2). In order to evaluate the increase of redundancy, we examined the distribution of the MS intensities of three proteins that were identified in the highest number of PAGE slices using PAGE-pIEF-LC-MS/MS. In both GeLC-MS/MS and PAGE-pIEF-LC-MS/MS approaches, these proteins had very high intensities in one main slice and low intensities in the bordering slices (Supporting Information Fig. 11). We noted the same when plotting the MS intensities of all proteins identified using the PAGE-pIEF-LC-MS/MS approach (Supporting Information Fig. 12 and Supporting Information Table 4). These results indicated that high resolution in protein separation was indeed achieved, with proteins concentrated mainly in a single PAGE slice and spread only in low amounts to the other PAGE slices. Due to increased sensitivity, the lower amounts in the additional gel slices were detected only after performing pIEF prior to LC-MS/MS, leading to an increase in redundancy.

3.5 Technical reproducibility of pIEF

Large-scale shotgun proteomic studies usually aim at the elucidation of complex biological phenomena. In order to

achieve biological significance in comparative studies, the experimental and technical variability must be measured and documented. As our method relies on two separation steps prior to LC-MS/MS analysis, we separately examined the technical reproducibility of pIEF and PAGE-pIEF. In order to evaluate the reproducibility of pIEF, we performed two replicate pIEF analyses of tryptic peptides from three different PAGE slices. We separated 70 μg HeLa NE by PAGE, cut the lane into 23 slices, and selected three slices corresponding to the high, middle, and low protein MW region (PAGE#05, MW ~150 kDa, PAGE#10, MW ~ 70 kDa, and PAGE#15, MW ~40 kDa). Tryptic peptides from each of the three slices were dissolved in 250 μL 8 M urea, 0.2% IPG buffer which solution was then split into two and used to run two technical replicate pIEF analyses (e.g. PAGE#05-pIEF-r1, PAGE#05-pIEF-r2, etc.). After IEF, IPG strips were fractioned and peptides were extracted and analyzed by LC-MS/MS. First, we examined the numbers of identified proteins and peptides and the overlap of the identified proteins in the replicate pIEF analyses and then, the distribution of identified peptides across the IPG strip. The numbers of identified proteins and peptides in both pIEF replicates was highly similar (Table 1). In addition, we observed a large overlap of the identified proteins in each replicate analysis (Fig. 3A–C). Interestingly, proteins

Table 1. Numbers of identified peptides and proteins in pIEF replicates

PAGE slice	PAGE#05		PAGE#10		PAGE#15	
pIEF replicate	r1	r2	r1	r2	r1	r2
Identified peptides	4803	5011	4505	4467	3237	3238
Identified proteins	376	368	533	529	528	516

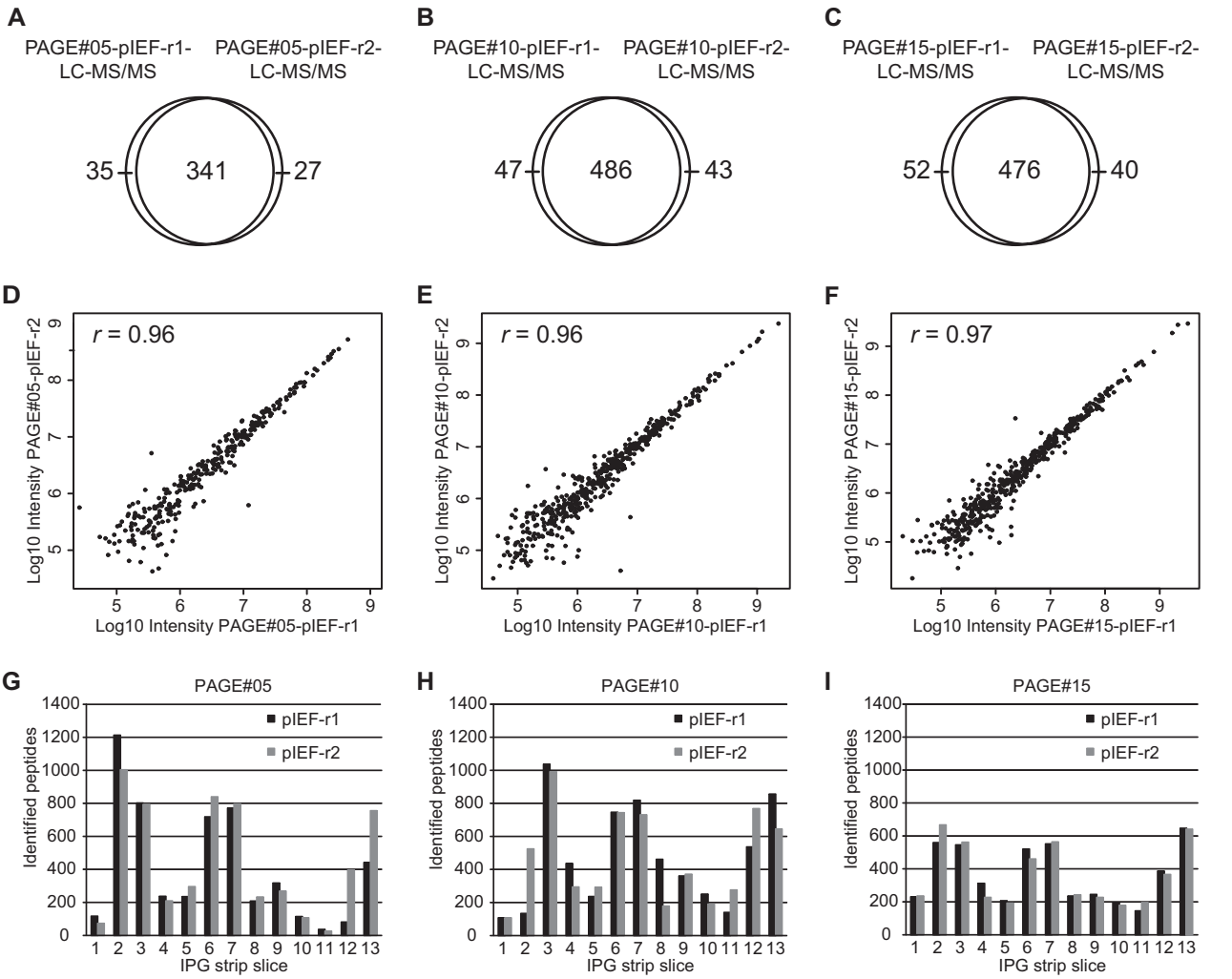

Figure 3. Overlap of identified proteins across replicate pIEF analyses of tryptic peptides from the same PAGE subfraction (A–C), correlation of intensity of proteins identified in both pIEF replicate analyses (D–F) and distribution of identified peptides on the IPG strips (G–I).

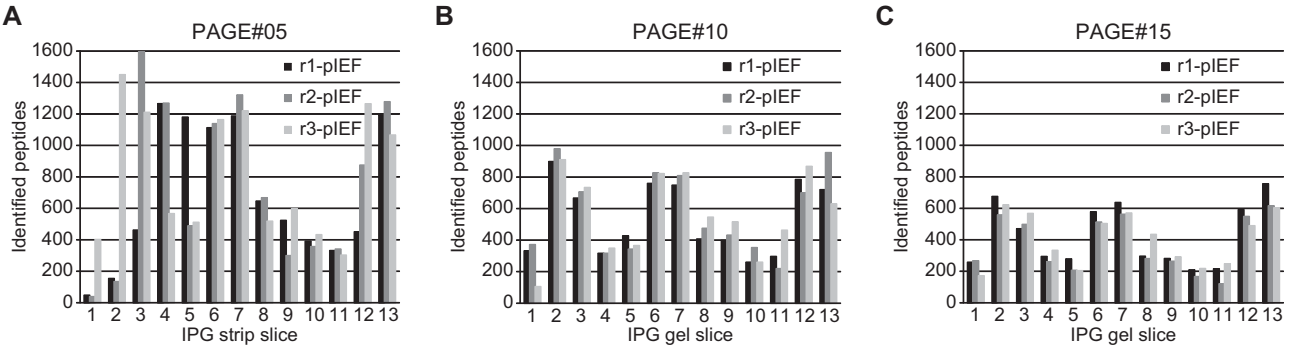

Figure 4. Distribution of identified peptides across the IPG strip from pIEF analyses of tryptic peptides from the same PAGE subfraction of three replicate PAGE lanes.

identified across pIEF replicates showed very high correlation of their MS intensity as the Pearson's correlation coefficient (*r*) between pIEF replicates was in the range 0.96–0.97 (Fig. 3D–F). Such high correlation values are characteristic of replicate LC-MS/MS analyses of the same sample, as shown by repeated injections and LC-MS/MS analysis of PAGE#05-pIEF-r1 and PAGE#05-pIEF-r2 (Supporting Information Fig. 13). We also observed very similar distribution of the identified peptides across the IPG strips (Fig. 3G–I). While there were differences in the numbers of identified peptides in some IPG strip slices, in all three experiments more than 99% of the peptides identified in both pIEF replicates were found in the same or within ±1 IPG gel slice. Thus, while each pIEF analysis includes several steps that can induce variability such as the IPG strip overnight rehydration, IEF, manual fractionation, extraction and desalting, we could show that upon proper handling the separation process and the protein/peptide identifications could be very reproducible.

3.6 Technical reproducibility of GeLC-MS/MS and PAGE-pIEF-LC-MS/MS

After showing the high technical reproducibility of pIEF-LC-MS/MS after PAGE we set to examine the reproducibility of the combined PAGE-pIEF-LC-MS/MS approach. As this method relies heavily on PAGE separation in the first dimension, we performed two separate experiments using PAGE in which peptides extracted from a PAGE slice were either (i) analyzed directly by LC-MS/MS separation without pIEF (GeLC-MS/MS) or (ii) separated by pIEF prior to LC-MS/MS analysis (PAGE-pIEF-LC-MS/MS). In both experiments we used three parallel replicate PAGE lanes of 70 μg HeLa NE. PAGE lanes were fractioned, and for further analysis we selected the same set of three PAGE slices (PAGE#05, PAGE#10, and PAGE#15) as in the pIEF experiment described in the previous section (Supporting Information Fig. 14). We assigned each parallel slice as a PAGE replicate (e.g. PAGE#05-r1, PAGE#05-r2, and PAGE#05-r3, etc.). First, we wanted to determine whether pIEF-LC-MS/MS analysis of peptides from replicate PAGE slices is reproducible. Therefore, we examined the distribution and numbers of identified peptides across the IPG strip in the PAGE-pIEF-LC-MS/MS approach. Similarly to the results described in the previous section, we observed a matching distribution of the identified peptides across the IPG strip from the pIEF analysis of replicate PAGE slices (Fig. 4), showing that the pIEF-LC-MS/MS separation was reproducible. The only difference that we found was in the acidic end (slices 1–5) in the pIEF analyses of PAGE#05-r1 and PAGE#05-r2 (Fig. 4A). We contributed this to improper rehydration at the acidic end of the IPG strip and underperformance of the IEF separation. Despite showing different peptide distribution, these five slices resulted in the identification of similar sets of peptides (data not shown). Next, we evaluated the overall reproducibility of GeLC-MS/MS and PAGE-pIEF-LC-MS/MS by looking at the numbers of identified proteins and peptides in a PAGE lane replicate. In both the GeLC-MS/MS and the PAGE-pIEF-LC-MS/MS approach there was a high consistency in the numbers of peptides and proteins identified (Table 2, Peptides, Proteins). The only exception was the pIEF analysis of PAGE#05 in which 6557, 7513, and 8569 peptides were identified in the three PAGE replicates. This difference was most probably a consequence of improper IPG strip rehydration, as discussed above. We further examined the overlap of protein identifications in all three PAGE replicates. As expected, in both experiments, the number of total proteins, i.e. those identified together in all three replicates, (Table 2, Total proteins) was higher compared with the number of identified proteins in a single replicate analysis (Table 2, Proteins). Meanwhile, the percentage of the total proteins that were also identified across all three replicates was higher for GeLC-MS/MS (65–67%) than for PAGE-pIEF-LC-MS/MS (52–63%). This showed that PAGE-pIEF-LC-MS/MS was slightly less reproducible than GeLC-MS/MS. Finally, we examined the correlation of protein intensity across PAGE replicates (Table 2, correlation protein intensity). Overall GeLC-MS/MS showed higher Pearson's correlation values ranging from 0.75 to 0.98, compared with PAGE-pIEF-LC-MS/MS with values between 0.67 and 0.97, further supporting the evidence that

Table 2. Numbers of identified peptides and proteins, correlation of protein intensity, and percentage of total proteins that are identified in all PAGE slice replicates using either GeLC-MS/MS or PAGE-pIEF-LC-MS/MS. Shown correlation is between PAGE replicates 1 and 2 (r1), 2 and 3 (r2), and 3 and 1 (r3). Total proteins indicate the number of non-redundant proteins identified together in all three replicate analyses.

	Experiment	GeLC-MS/MS			PAGE-pIEF-LC-MS/MS		
	PAGE replicate	r1	r2	r3	r1	r2	r3
PAGE#05	Peptides	3188	3241	3120	6557	751 3	8569
	Proteins	256	244	250	439	437	473
	Correlation protein intensity	0.89	0.95	0.89	0.67	0.68	0.97
	Total proteins		303			571	
	In all replicates [%]		67.00			58.14	
PAGE#10	Peptides	2443	2417	2408	5111	5629	5402
	Proteins	31 5	303	311	551	597	585
	Correlation protein intensity	0.84	0.98	0.83	0.80	0.70	0.95
	Total proteins		379			774	
	In all replicates [%]		65.1 7			52.71	
PAGE#15	Peptides	1538	1 705	1723	3583	3470	3512
	Proteins	218	233	237	585	593	581
	Correlation protein intensity	0.75	0.98	0.79	0.93	0.94	0.84
	Total proteins		271			716	
	In all replicates [%]		66.42			63.27	

PAGE-pIEF-LC-MS/MS was less reproducible than GeLC-MS/MS. Nevertheless, we observed that in both approaches, for all PAGE slices, two out of the three replicates showed very high correlation coefficients of the protein intensity: PAGE-r2/PAGE-r3 in GeLC-MS/MS and PAGE-r3/PAGE-r1 (slices #05 and #10) and PAGE-r1/PAGE-r2 and PAGE-r2/PAGE-r3 (slice #15) in PAGE-pIEF-LC-MS/MS (Table 2). These results indicated that very high reproducibility of PAGE-pIEF-LC-MS/MS can indeed be achieved.

Acknowledging that performing pIEF prior to LC-MS/MS is the methodological difference between PAGE-pIEF-LC-MS/MS and GeLC-MS/MS, it appears that including pIEF, which we showed to be highly reproducible, enhances shortcomings of the PAGE separation as indicated by slightly lower correlation of the protein intensity and overlap of identified proteins across replicates. This is expected since additional separation dimensions add sources of technical variation, which can decrease the overall reproducibility. Our findings are in agreement with other studies that compare 2D and 3D separation methods [16]. At the same time replicate PAGE-pIEF-LC-MS/MS analyses produced consistent numbers of peptide and protein identifications, which were higher than the corresponding GeLC-MS/MS analyses. This indicated that by introducing pIEF separation prior to LC-MS/MS we could increase the sensitivity of GeLC-MS/MS and achieve deeper coverage of the proteome.

4 Concluding remarks

In conclusion, we have established a PAGE-pIEF-LC-MS/MS approach that utilizes a large separation space and gives a deeper coverage of a proteome sample than either the GeLC-MS/MS or the pIEF-LC-MS/MS method used in this study. Methodologically, proteins are initially separated by 1D PAGE, the entire gel lane is fractionated, proteins are digested by trypsin in-gel, and extracted peptides are separated by pIEF prior to applying to LC-MS/MS. The PAGE and pIEF steps are easily interfaced without the need to exchange buffers, although pIEF might not tolerate high concentration of Ca^{2+}, commonly added to aid trypsin digestion. Our results show that pIEF is very reproducible and that adding it as a second separation dimension can significantly increase protein sequence coverage and the number of identified proteins from PAGE subfractionation. Moreover, pIEF allowed the identification of more proteins in comparison to a long LC gradient from a single PAGE slice. Our results demonstrated that our PAGE-pIEF-LC-MS/MS approach can be very successfully applied to proteomic analysis of specific PAGE regions. This would be especially advantageous if a proteomic sample has a high complexity with a specific MW bias. For our experiments, we used a fractionation tool that cuts the whole PAGE lane into a specific number of slices, but the gel lane can be fractioned into any number of slices prior to trypsin digestion and pIEF. Alternatively, the first dimension step can be substituted with another separation technique. This would allow for greater flexibility in custom-designing proteomics experiments to fit the specific needs.

We thank V.A. Raker for her help in preparation of this manuscript and Christof Lenz and Miroslav Nikolov for critical comments on the manuscript. This work was supported by the Max Planck Society.

The authors have declared no conflict of interest.

5 References

[1] Wolters, D. A., Washburn, M. P., Yates, J. R., An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* 2001, *73*, 5683–5690.

[2] Cargile, B. J., Bundy, J. L., Freeman, T. W., Stephenson, J. L., Gel based isoelectric focusing of peptides and the utility of isoelectric point in protein identification. *J. Proteome Res.* 2004, *3*, 112–119.

[3] Horth, P., Miller, C. A., Preckel, T., Wenz, C., Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis. *Mol. Cell. Proteomics* 2006, *5*, 1968–1974.

[4] Wisniewski, J. R., Zougman, A., Mann, M., Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. *J. Proteome Res.* 2009, *8*, 5674–5678.

[5] Dignam, J. D., Lebovitz, R. M., Roeder, R. G., Accurate Transcription initiation by RNA polymerase-II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 1983, *11*, 1475–1489.

[6] Schmidt, C., Lenz, C., Grote, M., Luhrmann, R., Urlaub, H., Determination of protein stoichiometry within protein complexes using absolute quantification and multiple reaction monitoring. *Anal. Chem.* 2010, *82*, 2784–2796.

[7] Schmidt, C., Urlaub, H., iTRAQ-labeling of in-gel digested proteins for relative quantification. *Methods Mol. Biol.* 2009, *564*, 207–226.

[8] Rappsilber, J., Ishihama, Y., Mann, M., Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* 2003, *75*, 663–670.

[9] Vaezzadeh, A. R., Hernandez, C., Vadas, O., Deshusses, J. J. M. et al., pICarver: a software tool and strategy for peptides isoelectric focusing. *J. Proteome Res.* 2008, *7*, 4336–4345.

[10] Olsen, J. V., de Godoy, L. M. F., Li, G. Q., Macek, B. et al., Parts per million mass accuracy on an orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* 2005, *4*, 2010–2021.

[11] Cox, J., Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 2008, *26*, 1367–1372.

[12] Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A. et al., Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* 2011, *10*, 1794–1805.

[13] de Godoy, L. M. F., Olsen, J. V., Cox, J., Nielsen, M. L. et al., Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* 2008, *455*, U1251-U1260.

[14] Manadas, B., English, J. A., Wynne, K. J., Cotter, D. R., Dunn, M. J., Comparative analysis of OFFGel, strong cation exchange with pH gradient, and RP at high pH for first-dimensional separation of peptides from a membrane-enriched protein fraction. *Proteomics* 2009, *9*, 5194–5198.

[15] Ishihama, Y., Oda, Y., Tabata, T., Sato, T. et al., Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* 2005, *4*, 1265–1272.

[16] Zgoda, V. G., Moshkovskii, S. A., Ponomarenko, E. A., Andreewski, T. V. et al., Proteomics of mouse liver microsomes: Performance of different protein separation workflows for LC-MS/MS. *Proteomics* 2009, *9*, 4102–4105.

