Martin Gilar[1]
Petra Olivova[1]
Asish B. Chakraborty[1]
Aleksander Jaworski[2]
Scott J. Geromanos[1]
John C. Gebler[1]

[1]Waters Corporation, Milford, MA, USA
[2]Franklin, MA, USA

Received September 24, 2008
Revised October 22, 2008
Accepted October 27, 2008

Research Article

Comparison of 1-D and 2-D LC MS/MS methods for proteomic analysis of human serum

1-D and 2-D LC methods were utilized for proteome analysis of undepleted human serum. Separation of peptides in 2-D LC was performed either with strong cation exchange (SCX)-RP chromatography or with an RP–RP 2-D LC approach. Peptides were identified by MS/MS using a data-independent acquisition approach. A peptide retention prediction model was used to highlight the potential false-positive peptide identifications. When applying selected data filtration, we identified 52 proteins based on 316 peptides in serum in 1-D LC setup. One hundred and eighty-four proteins/1036 peptides and 142 proteins/905 peptides were identified in RP–RP and SCX-RP 2-D LC, respectively. The performance of both 2-D LC methods for proteomic analysis is critically compared.

Keywords:
2-D / LC / Peptide / Proteomic / Retention prediction
DOI 10.1002/elps.200800630

1 Introduction

The majority of proteomic samples are highly complex mixtures of proteins/peptides [1–6]. Adequate separation of complex mixtures improves the success of MS-based proteomic analysis. Either 1-D or 2-D LC techniques are often employed [7–9]. The latter is expected to offer greater peptide resolution, in terms of chromatographic peak capacity [10].

Advanced separation in 2-D LC has been shown to allow for the identification of large numbers of peptides and proteins [3–8, 11–24]. The most common 2-D LC approach (for peptides) relies on a combination of strong cation exchange (SCX) and RP chromatographic modes [2, 5, 18, 25]. Recently, two groups have independently pioneered a new 2-D LC approach based on RP separation mode in both dimensions [14, 15, 22]. The different separation selectivity in first and second RP LC dimension is achieved by using a significantly different mobile phase pH. Typically, the first dimension is operated at pH 10, whereas the second at pH ~2.6. Because the eluent in first RP LC dimension contains ACN, the RP–RP 2-D LC was operated in an off-line mode with an intermediate sample concentration/evaporation step [14, 15, 22, 26, 27].

In a previous work, we proposed several alternative separation modes for 2-D LC separation of peptides and evaluated their mutual separation orthogonality [14, 15]. It has been shown that the separation orthogonality of SCX-RP and RP–RP 2-D LC approaches is comparable, at least for moderately complex sample. Delmotte *et al.* [28] compared SCX-RP and RP–RP 2-D LC methods for analysis of highly complex *C. glutamicum* proteomic sample. The authors concluded that SCX-RP method exhibits a greater degree of orthogonality compared with RP–RP 2-D LC. More recently, Dowell *et al.* compared several 2-D methods including RP–RP and SCX-RP 2-D LC using *Escherichia coli* as a test sample [27]. Nakamura *et al.* [29] have performed similar comparison for strong anion exchange-RP, SCX-RP, and RP–RP 2-D LC techniques. The authors of the reports found RP–RP 2-D LC to be suitable for proteomic analysis. The method provided a moderately higher number of identified peptides and proteins compared with SCX-RP 2-D LC. Although these results were supplemented by additional reports [30–32], the reasons remain unclear.

Several laboratories suggested using peptide retention prediction algorithms as effective filters for reducing false-positive (FP) peptide identifications, arguing that the retention time (RT) outliers are incorrect identifications and should be excluded from the data set ([19, 33–36],

Correspondence: Dr. Martin Gilar, Waters Corporation, 34 Maple Street, Milford, MA 01757, USA
E-mail: martin_gilar@waters.com
Fax: +1-508-482-3373

Abbreviations: ABRF, association of biomolecular resource facilities; **DDA,** data-dependent acquisition; **FA,** formic acid; **FP,** false positive; **MS^E,** alternate scanning MS and MS/MS data acquisition approach; **NH_4FA,** ammonium formate; **RT,** retention time; **SCX,** strong cation exchange

Krokhin *et al.*, http://hs2.proteome.ca/ssrcalc/ssrcalc.html, 2006). The majority of RT calculators are based on the ability to predict the peptide retention in RP LC from their amino acid composition using training set [31, 36, 37, 38]. This approach was recently extended into RP–RP 2-D LC, demonstrating that suitable retention models can be constructed for both low- and high-pH RP LC experiments [31, 32, 39]. Dwivedi *et al.* [40] have implemented the RP–RP 2-D HPLC method and shown an impressive accuracy of peptide retention prediction at both RP separation dimensions. In this report, we utilized peptide retention algorithm developed earlier [31] to visualize the falsely identified peptides.

Serum and plasma represent the most challenging proteomic samples. It is believed that as many as 10 000 proteins are present in serum, most of them at concentration levels undetectable with the current analytical methods. The protein dynamic concentration range spans over eleven orders of magnitude [3, 11, 12, 41, 42]. The concentration and enrichment of low abundant proteins are desirable [17, 24, 43, 44]. Removal of 6 or 20 major proteins reduces the serum protein content by 85 or 98%, respectively [13]. Notwithstanding the usefulness of protein depletion [17, 24, 45], the detection limit of low abundant proteins remains restricted by the mass load capacity of chromatographic column. Very large volumes of serum/plasma need to be processed to attain the limits of detection of low abundant proteins [43, 44].

Besides protein depletion, the fractionation on protein level is another strategy used to improve the dynamic range of proteomic analysis [44]. Although successful, this approach is laborious and is limited to soluble proteins and those that can be fractionated with sufficient recovery.

Limited success was reported when using 1-D LC-MS for analysis of undepleted plasma/serum [45]. Approximately 30–50 proteins were identified in the experiments described by Kapp *et al.* (only IgGs were depleted) [11, 12]. Majority of the identified peptides belonged to the abundant proteins, in particular human serum albumin.

The goal of the current work was to compare the success of 1-D LC and 2-D LC methods for analysis of serum proteome. Keeping in mind the ultimate goal of biomarker discovery, we selected human serum as the relevant sample, rather than using test samples such as yeast or *E. coli*. Because of danger of the unintentional protein losses during serum depletion (loss of non-targeted proteins), we chose to analyze an undepleted serum, accepting the limitations of lower dynamic range. SCX-RP and RP–RP methods were evaluated side-by-side using an identical sample with the aim of critically assessing the utility of novel RP–RP 2-D LC method for proteomic analysis. The aim was to compare the sets of confidently identified peptides and elucidate whether there are significant differences between the peptides identified in the RP–RP and SCX-RP 2-D LC methods.

2 Materials and methods

2.1 Chemicals

Formic acid (FA), TFA, ammonium hydroxide, and ammonium bicarbonate were purchased from Sigma-Aldrich (St. Louis, MO, USA). HPLC grade ACN was purchased from J.T. Baker (Phillipsburg, NJ, USA). A Milli-Q system (Millipore, Bedford, MA, USA) was used to prepare deionized water (18 MΩ cm) for HPLC mobile phases. MassPREP^{TM} protein digestion standards of enolase, alcohol dehydrogenase, phosphorylase b, hemoglobin, bovine serum albumin, and RapiGest^{TM} were obtained from Waters (Milford, MA, USA).

2.2 Human serum digest

Human serum sample was digested using the protocol published earlier [31, 32]. Briefly, 5 μL of human serum purchased from Gemini Bio-Products (Woodland, CA, USA), representing approximately 330 μg of total protein content, was transferred to an eppendorf tube, and combined with 20 μL of 100 mM ammonium bicarbonate (pH 8.5), and 25 μL of 0.2% aqueous solution of w/v RapiGest^{TM}. After protein denaturation (80°C for 15 min), 2.5 μL of 100 mM dithiothreitol was added to reduce disulfide bonds (60°C for 30 min). The proteins were alkylated by adding 2.5 μL of 200 mM iodoacetamide (room temperature, 30 min). Digestion was performed with Promega sequencing grade trypsin (100:1; total protein:-trypsin; w:w) at 37°C for overnight. The digestion mixture was acidified with 1 μL of 99% FA to break down the RapiGest^{TM} detergent prior to LC/MS analysis [46]. The final volume of the sample was ~100 μL; 2 μL was injected onto column for 1-D LC/MS analysis (this represents ~6.6 μg mass load, or 0.1 μL undiluted serum volume equivalent).

2.3 LC fractionation of serum digest

The fractionation of serum digest was performed using Alliance Bioseparations 2796 HPLC system equipped with 2996 photodiode array detector; 280 nm was used for UV detection. An external Rheodyne model 7725 injector (Rohnert Park, CA, USA) was used to inject 1 mL of serum digest on column (~3.3 mg total protein content) corresponding to 50 μL of human serum volume equivalent. Fractions were collected manually each 5 min, and evaporated to a final volume of 50 μL. The amount of sample injected onto second dimension RP analysis (1 μL) was the same as in both 2-D schemes and represented ~1 μL serum volume equivalent. This is a tenfold greater molar load for peptides compared with 1-D LC experiment, but not necessarily a greater mass load.

2.4 SCX chromatography

SCX separation was performed using PolySULFOETHYL A column (50 × 4.6 mm, 5 μm 300 Å, PolyLC, Columbia, MD, USA), which provided the best results in an earlier method optimization [15]. The digested serum sample was desalted prior to SCX fractionation using 3 cc Oasis HLB cartridge (60 mg). Briefly, the Oasis HLB SPE cartridge was conditioned with 1 mL of ACN followed by 1 mL of 0.1% aqueous TFA solution. The serum digest was spiked with TFA to a final concentration of 0.1%, loaded on cartridge, and subsequently washed with 1 mL of 0.1% TFA. Peptides were eluted with 1 mL of 70% ACN with 0.1% TFA and evaporated to dryness. Peptides were reconstituted in 1 mL of 0.1% FA prior to SCX fractionation. SCX separation temperature was 30°C; flow rate was 0.2 mL/min. Mobile phase A was water, B ACN, and C 400 mM ammonium formate (NH_4FA) aqueous buffer, pH 3.25. This buffer was prepared by titrating 5% FA aqueous solution (1.3 M FA) with concentrated ammonium hydroxide to pH 3.25. The NH_4^+ concentration calculated from the consumption of ammonium hydroxide was 400 mM. After the injection, the conditions were kept isocratic at 5 mM NH_4FA for 5 min. A shallow linear gradient from 5 to 160 mM was performed between 5 and 61 min, followed by a sharp gradient in the next 4 min from 160 to 300 mM NH_4FA. The concentration of ACN was maintained constant (25%) throughout the entire SCX chromatographic run.

2.5 RP chromatography

RP experiments at high pH were carried out using hybrid silica 150 × 2.1 mm, 3.5 μm, 130 Å column XTerra MS C18 (Waters). Mobile phase A was water, B ACN, and C 200 mM NH_4FA aqueous buffer; pH 10 (the NH_4FA buffer was prepared by diluting 12.5 g of concentrated ammonium hydroxide, 28%, into 900 mL of water and adding 1.62 mL of 99% FA; the pH was adjusted to 10, and the volume was brought to 1 L). Flow rate was 0.2 mL/min, separation temperature was 40°C. The B was maintained 0% for first 5 min of analysis, gradient was 0–60 % B from 5 to 80 min; the concentration of NH_4FA was maintained constant throughout the run.

2.6 Capillary LC analysis

Capillary LC of tryptic peptides was performed with a Waters nanoAcquity equipped with a Waters NanoEase™ Atlantis™ C_{18}, 300 μm × 15 cm RP column. The column temperature was set to 35°C. The 1-D LC experiment data were acquired using 0.1% aqueous FA as the mobile phase A and 0.1% FA in ACN as the mobile phase B. Gradient was 5–40% ACN in 90 min followed by 40–90% ACN wash in 10 min. The 2-D LC experiment analysis was performed with 0.1% aqueous FA as the mobile phase A and 0.1% FA in 80% ACN and 20% water as the mobile phase B. Gradient was 0–45% ACN in 90 min. The flow rate was 5 μL/min in all cases. The column was re-equilibrated at initial conditions (0% mobile phase B for 20 min). The lock mass, Glu-fibrinopeptide at 100 fmol/μL, was delivered by the auxiliary pump.

In 2-D LC analysis, the molar amount of peptides loaded on the column was approximately tenfold greater compared with 1-D LC. Because of an overload of the abundant peptides, peak broadening was observed. Reducing the mass load would improve the peak shape and mass accuracy; however, the prospect of detecting less abundant peptides will decrease. This problem is a typical scenario in proteomic research.

2.7 Mass spectrometer configuration

Capillary LC was directly connected to Waters Micromass Q-Tof Premier MS instrument. The mass spectrometer was operated in V-mode; the spectrum integration time was 1.0 s with an interscan delay time of 0.1 s. All analyses were performed using positive mode ESI using a NanoLock-Spray™ source. The lock mass channel was sampled every 30 s. Exact mass LC/MS and LC/alternate scanning MS and MS/MS data acquisition approach (MS^E) data were collected using 4 eV for MS and 15–40 eV for MS^E acquisition, respectively. Therefore, one cycle consisting of MS and MS^E scans was acquired every 2.0 s. The LC/MS data were acquired from m/z 50–2000.

2.8 Database search

The raw MS^E data were processed using Protein Lynx Global Server (PLGS 2.3 version 23) Identity^E software and searched against Swiss-Prot human database (20070206UniSwiss_HUMAN_plusBIC_Candidates-Cut2.fas, downloaded from the Association of Biomolecular Resource Facilities (ABRF) group web site). The database was concatenated with equal number of randomized sequences. The database was indexed using strict trypsin cleavage rules, one missed cleavage was allowed. Fixed modification of cysteine (carboxyamidomethylation) was selected. Tolerance used for peptides identification in 1-D LC experiment was set to auto (typically ±10 ppm for precursors and ±25 ppm for MS/MS fragments). The tolerance was relaxed to ±50 ppm for precursors and ±25 ppm for MS/MS fragments for 2-D LC experiment (because of column overloading and MS saturation, affecting the mass accuracy for the most abundant peptides).

2.9 Peptide retention prediction and outliers exclusion

The peptide RT prediction algorithm was described in a previous publication [31]. The predicted RT is calculated

from Eq. (1), considering the number of each amino acid (AA_i) in the peptide sequence and the amino acid retention factor b_i. The overall retention also depends on the peptide length L. The constants b_i, b_0, and c for a given chromatographic system were calculated from the regression using a training peptide set [31, 32]:

$$RT = (1 - c \cdot \ln L) \cdot \left(\sum b_i \cdot AA_i + b_0 \right) \quad (1)$$

Ideally, the correlation between experimental and predicted peptide retention is linear with a limited data scatter. Potential outliers can be excluded using Eq. (2); the exclusion criterion limit $E = 350$ was chosen arbitrarily. RT_{pred} and RT_{exp} are peptide predicted and experimental RTs, respectively:

$$E = |RT_{exp} - RT_{pred}| \cdot 100 \cdot \frac{|RT_{exp} - RT_{pred}|}{RT_{exp}} \quad (2)$$

3 Results and discussion

3.1 LC MS^E analysis

MS^E was used for all 1-D and 2-D LC experiments in this paper. A thorough description of principles of alternate scanning MS^E data acquisition has been given elsewhere [20, 47, 48]. The alternate scanning mode allows for data-independent acquisition [20, 21, 48], as opposed to a conventional data-dependent acquisition (DDA). Data-independent acquisition is well suited for simultaneous qualitative and quantitative proteomic analysis [21].

Briefly, the MS^E alternate scanning method relies on rapid, parallel data acquisition of two sets of data. The first set is a compilation of MS scans with low collision cell energy setting; the data set essentially comprises unfragmented peptides (precursors). The second set includes all MS scans performed at elevated collision cell energy setting. Resulting mass chromatograms contain all MS/MS ions originating from indiscriminately fragmented precursors. Two mass chromatograms are shown in Figs. 1A and B for a 1-D LC MS^E analysis of human serum digest. As a result of alternate scanning, the time scale of the mass chromatograms is offset by 1 s (scanning frequency used in this particular experiment). However, the precursors and their fragments are mutually linked *via* a chromatographic time axis, as both MS and MS/MS chromatograms are acquired in a single LC experiment.

Figure 1C illustrates the MS scan for a selected RT (indicated by an arrow). The dominant signal belongs to the peptide QNCELFEQLGEYK originating from human serum albumin. Figure 1D shows the high-energy scan acquired 1 s later, featuring the composite MS/MS spectra of all precursors eluting from the column at given chromatographic time. Due to the presence of a dominant peptide, most of the fragments in Fig. 1D are *b* and *y* ions belonging to QNCELFEQLGEYK. However, the MS/MS spectra of other minor components can also be reconstructed (appropriate *b* and *y* ions assigned to corresponding precursors). The basic assumption used for reconstruction is that detected precursors must appear as a peak at the identical time as their MS/MS fragments (extracted from MS/MS chromatogram). The high accuracy and resolution of QTOF MS plays an essential role in the MS/MS spectra deconvolution. In special cases when two or more precursors cannot

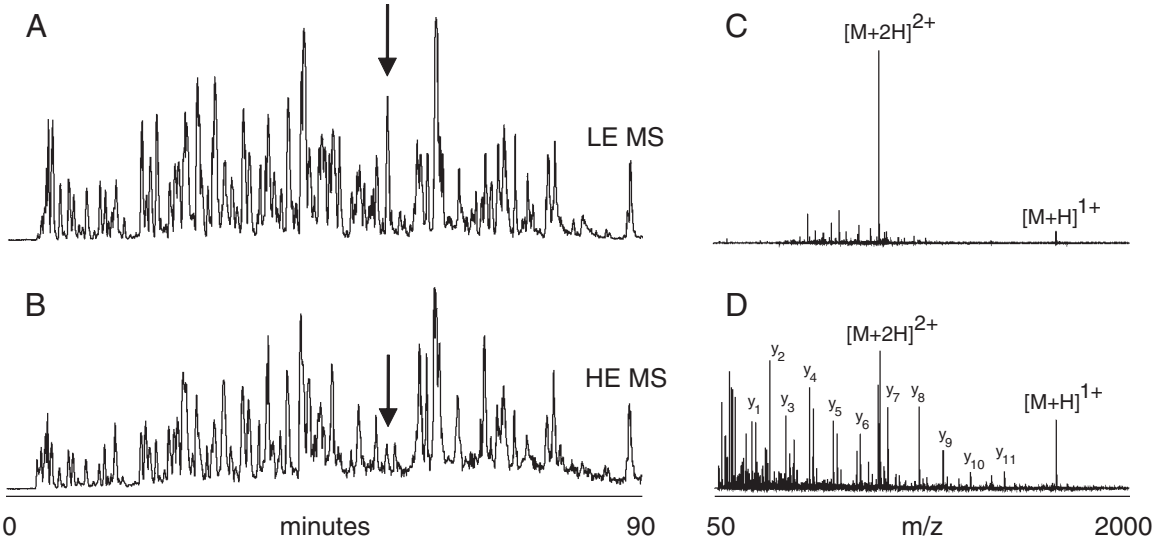

Figure 1. MS^E analysis of human serum tryptic digest in 1-D LC setup. *Via* rapid alternate scanning, two mass spectrum chromatograms are collected. Chromatogram (A) contains a compilation of MS scans at low-energy collision cell setting. Chromatogram (B) contains a compilation of MS/MS data; all precursors have been indiscriminately fragmented in the collision cell. Arrows point to a peak in chromatogram. Its MS spectrum (C) for 1 s MS scan shows ions for one dominant peptide along with many minor ones. MS/MS spectrum (D) acquired 1 s later contains a sum of all MS/MS fragments of all peptides eluting at the particular chromatographic time.

be chromatographically distinguished (ideal coelution), the reconstruction of MS/MS spectra relies on *in silico* depletion algorithm. The detailed explanation of the software algorithms is outside the scope of this manuscript. All reconstructed MS^E spectra were submitted to a database search.

3.2 2-D LC MS^E analyses

Two methods have been used for fractionation of human serum digest in first LC dimension. Either SCX or RP columns were utilized to separate peptides into distinct fractions (Fig. 2). Collected fractions were concentrated and analyzed using a capillary RP LC system with MS^E detection.

Most of the researchers using 2-D LC for proteome research chose to collect 5–15 fractions in first dimension [5, 17, 49–51]. This rather low number is driven by the practicality of experiment, namely by its length and complex data analysis. Another factor for consideration is the overlap in fraction content. A fraction content overlap exceeds 50% when the fractionation interval approaches the average peak width in the first separation dimension [15]. With those aspects in mind, we selected linear gradient and fractionation interval such that 10–14 fractions are generated in both RP and SCX experiments.

Figure 2 illustrates the performance of both 2-D LC methods. The RP–RP method shows good peak spread in second dimension, although some unused separation space could be observed. Interestingly, first RP dimension shows relatively evenly distributed peaks across the linear gradient, whereas in SCX chromatogram, the majority of peptides elute within fractions 4–6. Although it is possible to use a shallower gradient to achieve better peptides distribution between more fractions, the peaks would become proportionally broader. It has been found earlier that long gradients do not significantly improve the peak capacity of the SCX method [15]. Alternatively, a more frequent fractionation can be used to divide the fraction 4–6 into subsets [16, 25, 50]. The disadvantage of this approach is a greater peptide overlap between adjacent fractions [15].

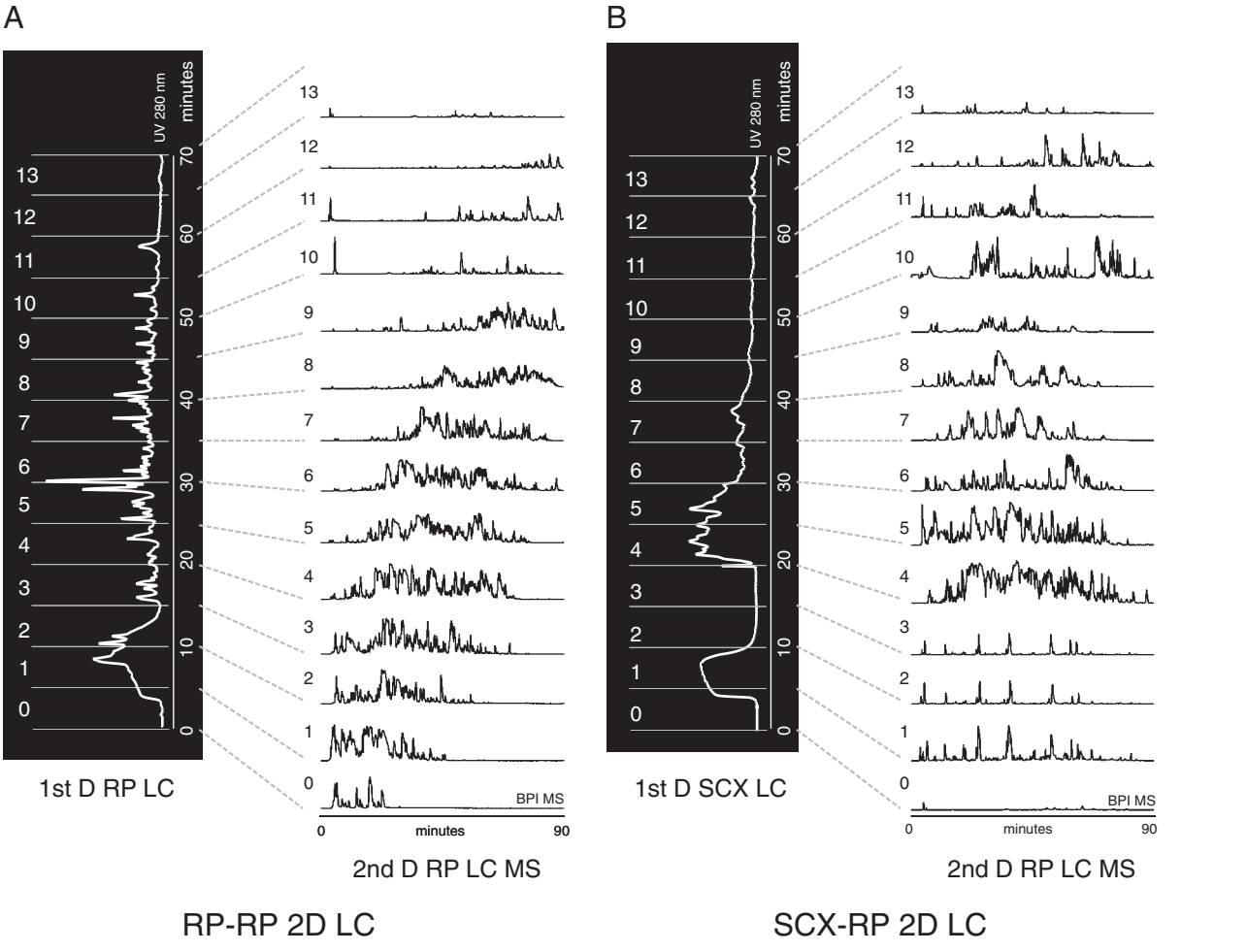

Figure 2. Separation of human serum tryptic digest in (A) RP–RP 2-D LC and (B) in SCX-RP 2-D LC setup. The fractions in first dimension were collected each 5 min.

Both first dimension chromatograms contain large peaks eluting soon after the sample introduction on column. These are signals of alkylation and reduction agents present in the sample. In RP experiment, some small hydrophilic peptides elute at the beginning of the chromatogram close to column void volume (see Fig. 2A). In the SCX experiment, we detected a minor breakthrough of peptides, evident as peaks present in second dimension LC-MS chromatograms, fractions 1–3. The pattern of eluting peaks is similar; these peaks are albumin-related peptides, the most abundant species in the sample.

3.3 Database search and data filtering

In order to compare the effectiveness of 2-D LC methods, the MS^E data have to be converted to peptide and protein identifications. The ABRF human database was concatenated with equal number of randomized decoy proteins in order to estimate the protein false discovery rate [5, 12, 16, 19, 25, 52]. Identity^E software search criteria are typically set to maintain low FP rate (*e.g.* 4% of random proteins pass the validation criteria). We have purposefully relaxed this criterion and accepted the protein identification up to ~25 % FP rate (25% random protein identification in the final set). Resulting data were exported and processed in Microsoft Excel. Additional data filters were applied and their efficiency was monitored as their ability to eliminate random protein identifications.

The data filters included identification repeatability and a RT prediction algorithm. Plots of experimental *versus* predicted RT are shown in Fig. 3 for RP–RP 2-D LC experiment. The graphs allow for visual appreciation of the numbers of RT outliers, which are indicative of FP peptide identifications [31, 34, 36]. 1-D LC experiment and SCX-RP 2-D LC data exhibit similar trends to those in Fig. 3 (data not shown).

Figure 3A shows 4953 so-called pass1 peptides, that is peptides that passed the validation criteria, such as mass accuracy, fragmentation pattern and contained at least three *b* or *y* ions or *b*/*y* ions combined. All peptides identified in combined triplicate analysis of all 14 RP–RP 2-D LC fractions are shown. The number of RT outliers declines sharply in Fig. 3B, where only repeatable peptide identifications are accepted (peptides identified at least twice in three repetitive analyses of the sample). This observation suggests that non-repeatable identifications often represent FP identifications. The proteins identified by a single peptide only (''one-hit-wonders'') are shown in Fig. 3B as open circles. Despite the fact that these identifications are repeatable, about 30% of these are questionable. Figure 3C shows the data set with applied RT exclusion filter (for details, see Section 2). RT exclusion filter successfully reduces the false discovery rate (measured as random peptides/proteins in the data; see the Supporting Information).

The usefulness of the RT prediction is illustrated in Fig. 3D, which plots the random peptide identifications only. Experimental RTs of randomly identified peptides do not show correlation with RT values predicted from their sequence. This finding confirms that RT filter is indeed useful for excluding peptides that were randomly or artificially identified in the data set. Many outliers shown in Fig. 3B are indeed random peptides and can be removed by RT filtering.

The sole purpose of the data filtration was to obtain highly confident set of peptides. Naturally, when accepting only repeatable identifications, and excluding the RT outliers, the final numbers of peptides and proteins are greatly reduced. Nevertheless, the resulting peptides represent more valid set suitable for comparison of RP–RP and SCX-RP 2-D LC methods.

3.4 Comparison of 1-D and 2-D LC methods

Using the criteria described in Section 3.3 (repeatable identifications, three *b* and *y* MS/MS fragments), we identified 316 unique peptides belonging to 52 proteins in 1-D LC experiment. Because 2-D LC provides for greater peak capacity, and the molar amount of peptides was ~tenfold higher compared with 1-D LC, one would expect that 2-D LC methods provide more extensive list of peptides and proteins. Indeed, RP–RP 2-D LC identified 1083 peptides and 191 proteins, whereas SCX-RP 2-D LC 958 proteins and 152 proteins.

When considering only proteins based on two or more peptides, 106 proteins were identified in RP–RP and 88 in SCX-RP 2-D LC experiments. The overlap in proteins between these two experiments was 83% (73 proteins). One hundred and twenty-nine unique proteins (150 when including possible isoforms) were identified in combined 2-D LC experiments (identified by two or more peptides). The list of peptides with their RT is provided in Supporting Information Table 2. The overlap between the 2-D LC experiments indicates that the analytical methods are not complementary to a degree as suggested in literature [3, 12, 16]. Significant variability in peptide/protein identification is inherent to the DDA MS/MS method [28, 53, 54] and can be mistakenly interpreted as complementarities between different separation methods. The data-independent analysis used in this work has been shown to provide highly repeatable and reproducible results [55, 56], although some variability remains for peptides approaching the detection (identification) limit.

One assumes that peptides identified in 1-D LC should be confirmed by data from 2-D LC experiment. Indeed, 95% of the peptides detected in 1-D LC were also found in combined 2-D LC experiment. Five proteins (out of 52 found in 1-D LC) were not observed in 2-D LC. Those five protein identifications were based on a single peptide only and are most likely FP identifications. When considering only identifications based on two or more peptides, 100% 1-D LC proteins were confirmed by 2-D LC experiment.

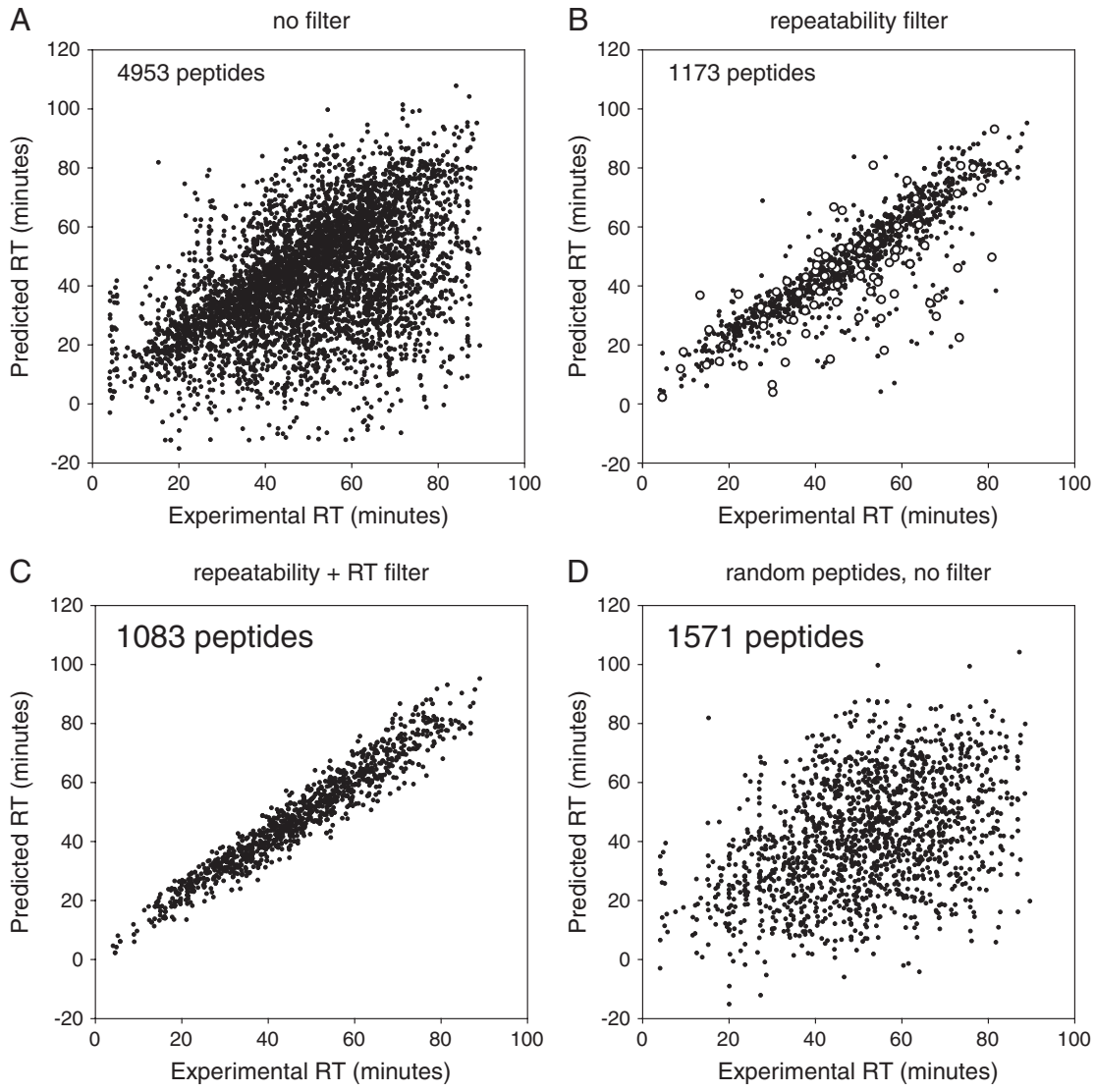

Figure 3. Correlation between experimentally measured and predicted RTs of peptides identified *via* a database search using various data filters. (A) All peptides identified in RP–RP 2-D LC experiment with minimum three *b*/*y* matched ions in precursor MS/MS spectrum. (B) Repeatable peptide identifications (peptide identified at least in two out of three repetitive experiments) with minimum three *b*/*y* ions matched. (C) Repeatable peptide identifications with three or more *b*/*y* ions matched; RT outliers were excluded. (D) Random peptides identified with three or more *b*/*y* ions. No repeatability filter was applied.

The RTs of peptides identified in 1-D LC experiment were plotted against the same peptides detected in 2-D LC (Fig. 4). The RTs show a strong correlation, despite the fact that the peptides were identified in separate experiments using slightly different gradients. The observed correlation increases our confidence in the peptide identification validity.

The numbers of proteins identified in serum or plasma are typically on order of hundred for 1-D LC experiment or several hundred for multi-dimensional setup [43, 45, 49, 50]. The numbers are lower compared with other type of proteomic samples (*e.g.* yeast [5, 25]). This is due to a great dynamic range of plasma/serum proteins.

The 2-D LC experiment presented here significantly improved the sequence coverage for proteins found in 1-D LC, and identified additional proteins. However, tenfold greater mass load in 2-D LC experiments resulted in only ~threefold greater numbers of identified peptides and proteins. This is expected, because the concentration levels of proteins in serum and other samples decrease exponentially [21, 48].

3.5 RP–RP versus SCX-RP 2-D LC method performance

Since the introduction of low/high-pH RP–RP 2-D LC concept in 2005 [14, 15, 22], several laboratories have investigated the method for proteomic applications [26, 30]. During the preparation of this manuscript, additional

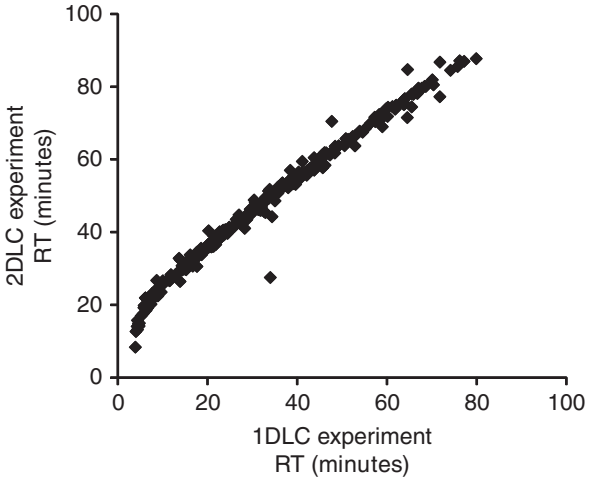

Figure 4. Correlation between experimental RTs of peptides identified in 1-D and 2-D LC experiments. Good correlation suggests that peptides were confidently identified and few FP peptide assignments were made.

reports compared RP–RP with SCX-RP and other 2-D LC methods [27, 29]. The results suggest that RP–RP 2-D LC is a useful method for proteomic analysis, and complements the state-of-the-art SCX-RP 2-D LC. Although it appears that the RP–RP method provides for moderately greater numbers of protein and peptides than SCX-RP 2-D LC, no explanation for this observation was provided [27–30]. We will investigate this question in more detail.

First, we will compare the apparent orthogonality of separation. Figure 2 illustrates that RP–RP 2-D LC provides and interesting orthogonality, despite the fact that both first and second separation dimension columns are essentially packed with the same type RP C18 sorbent. Figure 5 shows another representation of orthogonality, the 2-D surface coverage plots for RP–RP and SCX-RP 2-D LC [14]. The retention plots are similar to those published recently by Delmotte *et al.* [28] and also closely match the data obtained for less complex peptide mixtures [14, 15, 57].

Separation orthogonality is closely related to separation space coverage [14]. Fully orthogonal separation system will have the 2-D surface coverage plot randomly covered with the eluting peaks. Figure 5 indicates that some unused separation space exists for both investigated 2-D LC systems (see, *e.g.* sparse utilization of SCX-RP space for 11–13 fractions). We conclude that the in term of surface coverage, the RP–RP is similar to SCX-RP 2-D LC, as suggested in [14]. When collecting larger number of fractions, SCX-RP may exceed the orthogonality of RP–RP 2-D LC [28].

Figure 6 plots the number of peptides identified in each fraction for both separation methods. It appears that RP first separation dimension resolves peptides more evenly than SCX LC (see also Fig. 2). Large numbers of peptides in SCX separation mode elute in fraction 4; the numbers of peptides decline steeply in later fractions. Because the SCX separation selectivity depends on peptide charge, the tryptic

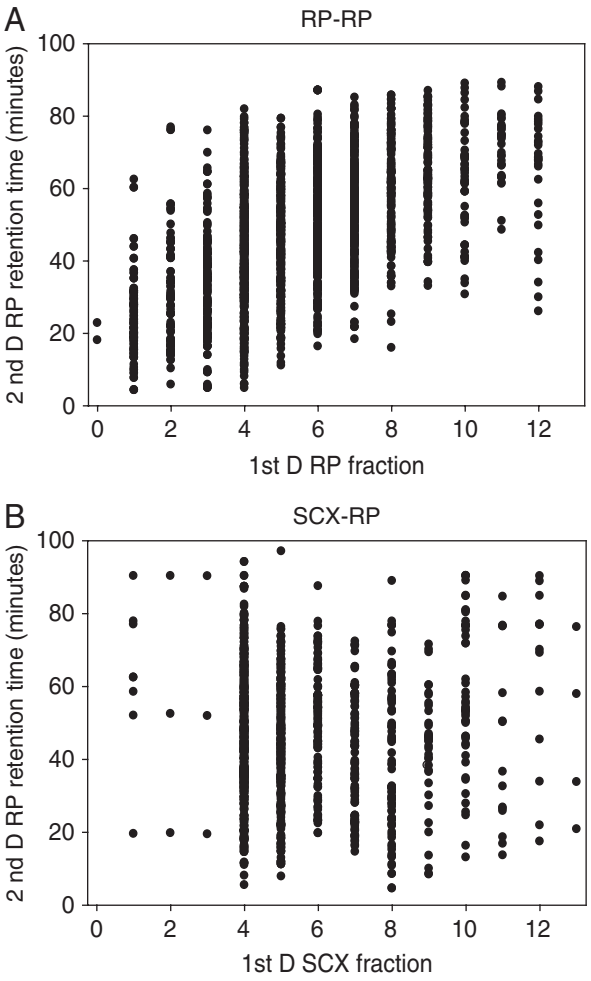

Figure 5. Orthogonality maps for two alternative 2-D LC methods. (A) RP–RP and (B) in SCX-RP 2-D LC.

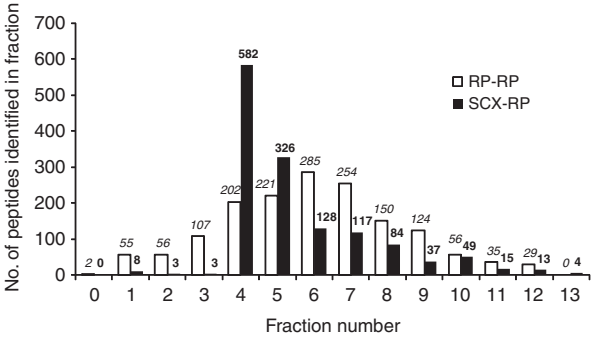

Figure 6. Number of peptides found in first LC dimension fractions for two alternative 2-D LC approaches (repeatable identifications with minimum three *b*/*y* ions were accepted; single peptide identifications were excluded). Peptides are more evenly distributed in RP–RP 2-D LC. For corresponding chromatograms, see Fig. 2.

peptides carrying typically 2^+ and 3^+ charge tend to elute in clusters [15, 25, 26, 28]. The sample complexity in selected SCX fractions is greater than one may expect by evenly

dividing the number of peptides in original sample by number of fractions.

More even distribution of peptides among the fractions in SCX first dimension can be achieved by collecting fractions more frequently during the initial phase of experiment. This would, however, increase the overlap in the fraction peptide content, not a favorable situation for DDA MS/MS analysis [53]. The average peptide overlap found between neighboring fractions in our experiment was 12% for RP–RP and 20% for SCX-RP 2-D LC. The overlap between fractions 4 and 5 in SCX mode was 39%. This observation is consistent with the broader peaks detected in SCX (Fig. 2).

Delmotte *et al.* [28] observed more even distribution of peptides between SCX LC fractions; earlier part of separation space was utilized. This is due to different column and gradient utilized in the authors' report [28]. It appears that the PolyLC PolySULFOETHYL^{TM} aspartamide column used in our work is more retentive than ProPac SCX-10. The degree of fraction overlap was not reported.

Dowell *et al.* [27] utilized various parameters such as p*I*, molecular weight, copy *per* cell number, and hydrophobicity of identified proteins to highlight some differences between 2-D separation strategies. Because the separation is performed on a peptide level, we investigated the peptide properties instead. Table 1 lists average p*I* and charge for peptides for both 2-D LC methods. As one would expect, the retention of peptides (fraction number) in SCX LC increases with the average peptide charge. The RP LC retention *versus* charge does not show a clear trend. Some p*I* trends are observed in Table 1, suggesting that basic peptides are more retained in both SCX and RP first dimension methods. However, when calculating the average charge or p*I* for all peptides detected in RP–RP and SCX-RP 2-D LC experiments, the values are nearly identical (charge 2.2 or 2.3 and p*I* 4.81 or 4.85, respectively). No distinction could be made between the methods, neither one seems to be better suitable for analysis of acidic or basic peptides.

More interesting results were obtained when calculating an average length of peptides detected in the fractions. Not surprisingly, the longer peptides (more hydrophobic) are better retained in first RP LC dimension. Average peptide length in the most populated SCX LC fractions 4 and 5 was relatively low. The average peptide lengths for the entire SCX-RP and RP–RP 2-D LC experiments were 14.0 and 15.5 amino acids. The overall average retention of the identified peptides in our experiment was 44.8 min for SCX-RP and 47.7 min for RP–RP 2-D LC method.

Although more factors can certainly play a role, the ability of RP–RP 2-D LC method to successfully identify longer (more hydrophobic) peptides emerged from further data analysis. Figure 7 indicates that SCX-RP method offered relatively greater number of identifications for 7–13 amino acids long peptides. The trend is opposite for RP–RP method, where noticeably larger numbers of peptides longer than 20 amino acids were identified. This observation may

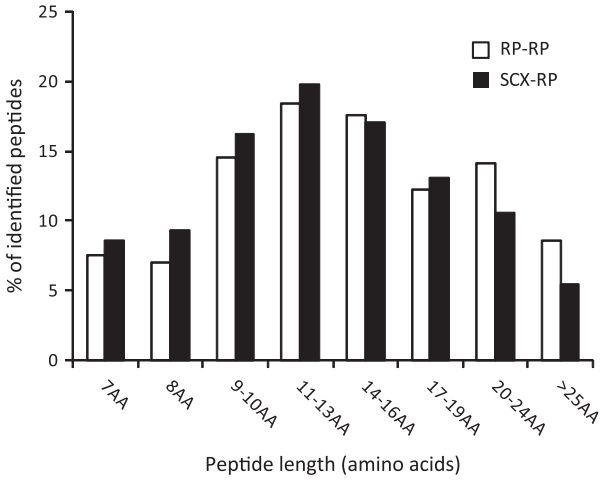

Figure 7. Relative frequency of peptides identified in RP–RP and SCX-RP 2-D LC experiments divided into groups according to their length. RP–RP method is more successful for peptides longer than 20 amino acids. The relative values for peptides were normalized to total number of unique identified peptides in RP–RP and SCX-RP 2-D LC (1033 and 888, respectively; repeatable identifications with minimum three *b*/*y* ions were accepted; single peptide identifications were excluded).

Table 1. Peptides average properties in RP or SCX first dimension fractions.

1st D	Fr. 0	Fr. 1	Fr. 2	Fr. 3	Fr. 4	Fr. 5	Fr. 6	Fr. 7	Fr. 8	Fr. 9	Fr. 10	Fr. 11	Fr. 12	Fr. 13
						Average charge of peptides in fraction[a)]								
RP	7.0	2.2	2.3	2.4	2.5	2.6	2.6	2.7	2.7	2.7	2.7	2.6	2.7	n.a.
SCX[b)]	*n.a.*	*1.9*	*1.4*	*1.4*	2.1	2.2	2.9	2.9	3.0	3.4	3.3	3.8	3.3	4.0
						Average pI of peptides in fraction[c)]								
RP	4.0	4.2	4.5	4.4	4.6	4.6	4.8	5.2	5.6	6.3	7.5	7.2	6.9	n.a.
SCX	*n.a.*	*4.3*	*4.9*	*4.9*	4.5	5.1	4.9	5.7	6.4	6.2	5.0	6.4	5.3	6.2
						Average length of peptides in fraction[b)]								
RP	12.7	11.0	10.7	11.9	14.0	15.3	16.4	16.6	17.8	17.9	17.3	18.6	16.4	n.a.
SCX	*n.a.*	*16.4*	*11.0*	*11.0*	13.6	12.4	17.0	13.6	13.4	14.6	19.1	15.7	17.7	17.8

a) The charge was calculated as a sum of *K*, *R*, and *H* in the peptide sequence plus 1 (terminal NH_2 group).
b) SCX fractions 0–3 contain only a few breakthrough peptides; values are in italics.
c) The p*I* values were calculated according to Shimura *et al.* [58].

be related to the selective loss of hydrophobic peptides in SCX LC first dimension separation. The loss of peptides is observed despite our effort to minimize it by selecting the SCX column with low residual hydrophobicity [59] and including 25% ACN in the mobile phase.

It is known that short peptides (7–8 amino acids) are less informative for protein identification (greater homology) than longer ones. The detection of more hydrophobic peptides can potentially explain the greater success of RP–RP method compared with SCX-RP 2-D LC reported in recent reports [26–30]. Nakamura *et al.* arrived to a similar conclusion when comparing several different RP and SCX columns for first dimension LC fractionation [29].

3.6 Prospects of RP–RP method for on-line 2-D LC proteomic analysis

Although the RP–RP 2-D LC method is well suited for off-line proteomic analysis, one may encounter difficulties when scaling down this method to nanoLC. Because RP fractions collected in first dimension contain high-pH buffer and organic solvent, the retention and concentration of the peptides on the trapping column, typically used in on-line 2-D nanoLC setup, could be compromised. This appears to limit the practicality of the proposed RP–RP 2-D LC method. One can envision that at-line sample dilution/neutralization by an aqueous acidic solvent can be performed prior to on-line sample concentration on a trapping column in order to eliminate the peptide breakthrough. Indeed, such on-line RP–RP 2-D nanoLC setup was recently implemented for analysis of complex peptide mixtures [60].

4 Concluding remarks

The comparison of two selected 2-D LC MS^E methods for proteomic analysis of undepleted human serum supports several conclusions: (i) Both 2-D LC methods are suitable for reducing sample complexity prior to MS analysis. (ii) The two investigated 2-D LC method identified similar proteins. Approximately 73% proteins identified by at least two proteins in SCX-RP were also found in RP–RP 2-D LC experiment. The methods are not strongly complementary. (iii) Retention prediction is a useful tool for proteomic data validation. One-hit-wonders are often highlighted as RT outliers, which are indicative of FP peptide identification. (iv) RP–RP is more suitable for analysis of hydrophobic peptides than SCX-RP 2-D LC. The loss of longer (hydrophobic) peptides in SCX-RP setup is probably responsible for moderately greater numbers of peptides and proteins identified in RP–RP 2-D LC.

The authors thank Dr. Martha Stapels and Dr. Sean M. McCarthy for their helpful comments to the manuscript.

The authors have declared no conflict of interest.

5 References

[1] Wehr, T., *LCGC North America* 2002, *20*, 954–962.

[2] Wolters, D. A., Washburn, M. P., Yates, J. R., III, *Anal. Chem.* 2001, *73*, 5683–5690.

[3] Adkins, J. N., Monroe, M. E., Auberry, K. J., Shen, Y., Jacobs, J. M., Camp, D. G., 2nd., Vitzthum, F. *et al.*, *Proteomics* 2005, *5*, 3454–3466.

[4] Aebersold, R., Anderson, L., Caprioli, R., Druker, B., Hartwell, L., Smith, R., *J. Proteome Res.* 2005, *4*, 1104–1109.

[5] Elias, J. E., Haas, W., Faherty, B. K., Gygi, S. P., *Nat. Methods.* 2005, *2*, 667–675.

[6] Eriksson, J., Fenyo, D., *J. Proteome Res.* 2005, *4*, 387–393.

[7] Shen, Y., Strittmatter, E. F., Zhang, R., Metz, T. O., Moore, R. J., Li, F., Udseth, H. R. *et al.*, *Anal. Chem.* 2005, *77*, 7763–7773.

[8] Shen, Y., Zhang, R., Moore, R. J., Kim, J., Metz, T. O., Hixson, K. K., Zhao, R. *et al.*, *Anal. Chem.* 2005, *77*, 3090–3100.

[9] Tolley, L., Jorgenson, J. W., Moseley, M. A., *Anal. Chem.* 2001, *73*, 2985–2991.

[10] Gilar, M., Daly, A. E., Kele, M., Neue, U. D., Gebler, J. C., *J. Chromatogr. A* 2004, *1061*, 183–192.

[11] Kapp, E. A., Schutz, F., Connolly, L. M., Chakel, J. A., Meza, J. E., Miller, C. A., Fenyo, D. *et al.*, *Proteomics* 2005, *5*, 3475–3490.

[12] Omenn, G. S., States, D. J., Adamski, M., Blackwell, T. W., Menon, R., Hermjakob, H., Apweiler, R. *et al.*, *Proteomics* 2005, *5*, 3226–3245.

[13] Echan, L. A., Tang, H. Y., Ali-Khan, N., Lee, K., Speicher, D. W., *Proteomics* 2005, *5*, 3292–3303.

[14] Gilar, M., Olivova, P., Daly, A. E., Gebler, J. C., *Anal. Chem.* 2005, *77*, 6426–6434.

[15] Gilar, M., Olivova, P., Daly, A. E., Gebler, J. C., *J. Sep. Sci.* 2005, *28*, 1694–1703.

[16] Li, X., Gong, Y., Wang, Y., Wu, S., Cai, Y., He, P., Lu, Z. *et al.*, *Proteomics* 2005, *5*, 3423–3441.

[17] Martosella, J., Zolotarjova, N., Liu, H., Nicol, G., Boyes, B. E., *J. Proteome Res.* 2005, *4*, 1522–1537.

[18] Masuda, J., Maynard, D. M., Nishimura, M., Ueda, T., Kowalak, J. A., Markey, S. P., *J. Chromatogr. A* 2005, *1063*, 57–69.

[19] Qian, W. J., Liu, T., Monroe, M. E., Strittmatter, E. F., Jacobs, J. M., Kangas, L. J., Petritis, K. *et al.*, *J. Proteome Res.* 2005, *4*, 53–62.

[20] Silva, J. C., Denny, R., Dorschel, C. A., Gorenstein, M., Kass, I. J., Li, G. Z., McKenna, T. *et al.*, *Anal. Chem.* 2005, *77*, 2187–2200.

[21] Silva, J. C., Gorenstein, M. V., Li, G. Z., Vissers, J. P., Geromanos, S. J., *Mol. Cell. Proteomics* 2006, *5*, 144–156.

[22] Toll, H., Oberacher, H., Swart, R., Huber, C. G., *J. Chromatogr. A* 2005, *1079*, 274–286.

[23] Weatherly, D. B., Atwood, J. A., III, Minning, T. A., Cavola, C., Tarleton, R. L., Orlando, R., *Mol. Cell. Proteomics* 2005, *4*, 762–772.

[24] Zolotarjova, N., Martosella, J., Nicol, G., Bailey, J., Boyes, B. E., Barrett, W. C., *Proteomics* 2005, *5*, 3304–3313.

[25] Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., Gygi, S. P., *J. Proteome Res.* 2003, *2*, 43–50.

[26] Schley, C., Altmeyer, M. O., Swart, R., Muller, R., Huber, C. G., *J. Proteome Res.* 2006, *5*, 2760–2768.

[27] Dowell, J. A., Frost, D. C., Zhang, J., Li, L., *Anal. Chem.* 2008, *80*, 6715–6723.

[28] Delmotte, N., Lasaosa, M., Tholey, A., Heinzle, E., Huber, C. G., *J. Proteome Res.* 2007, *6*, 4363–4373.

[29] Nakamura, T., Kuromitsu, J., Oda, Y., *J. Proteome Res.* 2008, *7*, 1007–1011.

[30] Kilpatrick, L., Vizcaino, M., Morris, P. J., Stein, S. T., Poster 13, ISPPP 2007, Orlando 2007.

[31] Gilar, M., Jaworski, A., Olivova, P., Gebler, J. C., *Rapid Commun. Mass Spectrom.* 2007, *21*, 2813–2821.

[32] Gilar, M., Neue, U. D., *J. Chromatogr. A* 2007, *1169*, 139–150.

[33] Petritis, K., Kangas, L. J., Ferguson, P. L., Anderson, G. A., Pasa-Tolic, L., Lipton, M. S., Auberry, K. J. *et al.*, *Anal. Chem.* 2003, *75*, 1039–1048.

[34] Petritis, K., Kangas, L. J., Yan, B., Monroe, M. E., Strittmatter, E. F., Qian, W. J., Adkins, J. N. *et al.*, *Anal. Chem.* 2006, *78*, 5026–5039.

[35] Krokhin, O. V., Craig, R., Spicer, V., Ens, W., Standing, K. G., Beavis, R. C., Wilkins, J. A., *Mol. Cell. Proteomics* 2004, *3*, 908–919.

[36] Krokhin, O. V., *Anal. Chem.* 2006, *78*, 7785–7795.

[37] Guo, D. C., Mant, C. T., Hodges, R. S., *J Chromatogr.* 1987, *386*, 205–222.

[38] Meek, J. L., *Proc. Natl. Acad. Sci. USA.* 1980, *77*, 1632–1636.

[39] Spicer, V., Yamchuk, A., Cortens, J., Sousa, S., Ens, W., Standing, K. G., Wilkins, J. A. *et al.*, *Anal. Chem.* 2007, *79*, 8762–8768.

[40] Dwivedi, R. C., Spicer, V., Harder, M., Antonovici, M., Ens, W., Standing, K. G., Wilkins, J. A. *et al.*, *Anal. Chem.* 2008, *80*, 7036–7042.

[41] Anderson, N. L., Anderson, N. G., *Mol. Cell. Proteomics* 2002, *1*, 845–867.

[42] Anderson, N. L., Polanski, M., Pieper, R., Gatlin, T., Tirumalai, R. S., Conrads, T. P., Veenstra, T. D. *et al.*, *Mol. Cell. Proteomics* 2004, *3*, 311–326.

[43] Rose, K., Bougueleret, L., Baussant, T., Bohm, G., Botti, P., Colinge, J., Cusin, I. *et al.*, *Proteomics* 2004, *4*, 2125–2150.

[44] Faca, V., Pitteri, S. J., Newcomb, L., Glukhova, V., Phanstiel, D., Krasnoselsky, A., Zhang, Q. *et al.*, *J. Proteome Res.* 2007, *6*, 3558–3565.

[45] Schenk, S., Schoenhals, G. J., de Souza, G., Mann, M., *BMC Med. Genomics* 2008, *1*, 41–69.

[46] Yu, Y. Q., Gilar, M., Lee, P. J., Bouvier, E. S., Gebler, J. C., *Anal. Chem.* 2003, *75*, 6023–6028.

[47] Chakraborty, A. B., Berger, S. J., Gebler, J. C., *Rapid Commun. Mass Spectrom.* 2007, *21*, 730–744.

[48] Silva, J. C., Denny, R., Dorschel, C., Gorenstein, M. V., Li, G. Z., Richardson, K., Wall, D. *et al.*, *Mol. Cell. Proteomics* 2006, *5*, 589–607.

[49] Shen, Y., Kim, J., Strittmatter, E. F., Jacobs, J. M., Camp, D. G., II, Fang, R., Tolie, J. N. *et al.*, *Proteomics* 2005, *5*, 4034–4045.

[50] Fujii, K., Nakano, T., Kawamura, T., Usui, F., Bando, Y., Wang, R., Nishimura, T., *J. Proteome Res.* 2004, *3*, 712–718.

[51] Zhou, H., Dai, J., Sheng, Q. H., Li, R. X., Shieh, C. H., Guttman, A., Zeng, R., *Electrophoresis* 2007, *28*, 4311–4319.

[52] Cargile, B. J., Bundy, J. L., Stephenson, J. L., Jr., *J. Proteome Res.* 2004, *3*, 1082–1085.

[53] Liu, H., Sadygov, R. G., Yates, J. R., III, *Anal. Chem.* 2004, *76*, 4193–4201.

[54] Chen, H. S., Rejtar, T., Andreev, V., Moskovets, E., Karger, B. L., *Anal. Chem.* 2005, *77*, 7816–7825.

[55] Fadgen, K. E., Murphy, J. P., Stapels, M. D., poster, Waters Library Number: PSTR10070802, ASMS 2008.

[56] Murphy, J. P., Stapels, M. D., Fadgen, K. E., Geromanos, S., poster, Waters Library Number: PSTR10070805, ASMS 2008.

[57] Gilar, M., Olivova, P., Daly, A. E., Gebler, J. C., in: Cohen, A. S., Schure, M. R. (Eds.), *Multidimensional Liquid Chromatography: Theory and Applications in Industrial Chemistry and the Life Sciences*, Wiley, Hoboken, New Jersey 2008, pp. 261–284.

[58] Shimura, K., Kamiya, K., Matsumoto, H., Kasai, K., *Anal. Chem.* 2002, *74*, 1046–1053.

[59] Alpert, A. J., Andrews, P. C., *J. Chromatogr.* 1988, *443*, 85–96.

[60] Liu, H., Li, G., Finch, J. W., Geromanos, S. J., Gebler, J. C., poster P57–T ABRF 2007, Tampa, Florida 2007.

