UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Development of comprehensive two-dimensional low-flow liquid-chromatography setup coupled to high-resolution mass spectrometry for shotgun proteomics

Roca, L.S.; Gargano, A.F.G.; Schoenmakers, P.J.

[Link to publication]

Development of comprehensive two-dimensional low-flow liquid-chromatography setup coupled to high-resolution mass spectrometry for shotgun proteomics

Liana S. Roca [a, b, *], Andrea F.G. Gargano [a, b], Peter J. Schoenmakers [a, b]

[a] van 't Hoff Institute for Molecular Science, University of Amsterdam, Amsterdam, the Netherlands
[b] Centre for Analytical Sciences Amsterdam, Amsterdam, the Netherlands

HIGHLIGHTS

- Increased LC × LC peak capacity versus 1D-LC without an increase in analysis time.
- Complete sample transfer between HILIC and RPLC dimensions.
- Sample focusing on trap column to overcome dilution and solvent incompatibility.
- Increased protein identification rate.
- Application of HILIC × RPLC to kidney tissue and IMR-90 cell line.

GRAPHICAL ABSTRACT
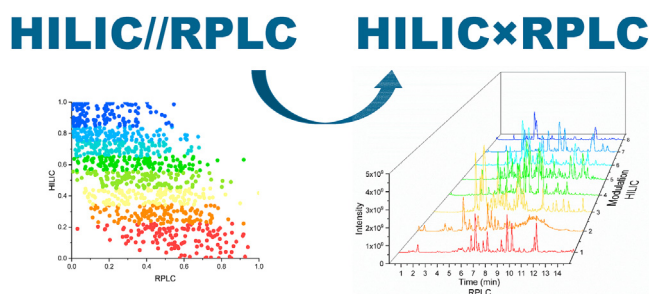
ARTICLE INFO

ABSTRACT

Bottom-up proteomics provides often small amounts of highly complex samples that cannot be analysed by direct mass spectrometry (MS). To gain a better insight in the sample composition, liquid chromatography (LC) and (comprehensive) two-dimensional liquid chromatography (2D-LC or LC × LC) can be coupled to the MS. Low-flow separations are attractive for HRMS analysis, but they tend to be lengthy. In this work, a low-flow, online, actively modulated LC × LC system, based on hydrophilic-interaction liquid chromatography (HILIC) in the first dimension and reversed-phase liquid chromatography (RPLC) in the second dimension, was developed to separate complex mixtures of peptides. Miniaturization permitted the analysis of small sample amounts (1−5 μg) and direct coupling with micro-ESI MS (1 μL min^{-1}). All components were focused and automatically transferred from HILIC to RPLC using stationary-phase-assisted active modulation (C18 traps) to deal with solvent-incompatibility or dilution issues. Optimization of the setup was performed for the HILIC columns and the RPLC columns to provide a more efficient separation and higher identification rates than obtained using one-dimensional (1D) LC. A 60% increase in peak capacity was obtained with the 2D setup compared to a 1D-RPLC separation and a 17

−34% increase in the number of proteins identified was achieved for the samples analysed (2D-yeast-8280 peptides and 2D-kidney tissue-8843 peptides), without increasing the analysis time (2 h).

1. Introduction

Liquid chromatography (LC) coupled to high-resolution mass spectrometry (HRMS) is commonly used to study protein expression and its modification in biological samples (*e.g.* cell cultures, tissues) [1]. Due to the difficulties associated with the liquid separation, ionization and fragmentation of proteins in the gas phase and the detection of ions with high mass-to-charge (m/z) ratios in MS [2], proteins are typically digested using enzymes (*e.g.* trypsin) and the resulting peptides are subsequently analysed using reversed-phase liquid chromatography (RPLC) coupled to tandem MS (MS/MS). To allow for the analysis of samples that are limited in quantity (and concentration), low-flow separations (*e.g.* 300 nL min^{-1}) with nano-electrospray ionization (nESI) sources are used [3,4]. The MS and MS/MS data are processed using software (*e.g.* MaxQuant [5], Mascot [6], PeptideShaker [7]) to identify peptides and to infer the presence of proteins in the sample based on characteristic peptides (unique proteotypic peptides). This analytical approach is usually described as bottom-up proteomics [8,9].

Tens of thousands of proteins are present in biological samples of complex organisms in a wide range of different concentrations (estimated at 10 orders of magnitude [10]). Furthermore, biological processes can induce several types of post-translational modifications (PTMs, *e.g.* phosphorylation, acetylation, glycosylation) [11,12]. The digestion process increases the complexity of the sample by generating up to a hundred peptides per protein. Therefore, achieving high-coverage proteome analysis of a complex cell system is a challenging endeavour, crucially relying on fast and high-resolution-separation and mass-spectrometry methods. Multidimensional separations can provide a large increase in the peak capacity in comparison with one-dimensional LC methods, provided that the separation dimensions are very different ("orthogonal") [13]. This will help reduce the coelution of analyte components and makes multidimensional methods attractive for proteomics analysis [14–17]. Orthogonality metrics asses the fraction of the separation space that is utilized for the separation of the components in a sample. The orthogonality can be assessed in several ways, for example focusing on the coverage of the separation space (global orthogonality) or on the uniformity of the space coverage (local orthogonality) [18].

An online multidimensional LC method often employed in proteomics analysis is the multidimensional protein identification technology (MudPIT) [19–21]. A column containing ion-exchange (IEC) particles in the first part and RPLC particles in the remainder is used in a MudPIT setup. Proteins or peptides are eluted using salt pulses from a strong-cation-exchange (SCX) resin or a mixture of weak-anion-exchange and SCX resins [22], focused on the RPLC column and separated using water to ACN gradients (typically containing 0.1% formic acid). Using a single column approach avoids dead volumes between separation dimensions and does not necessitate a valve system to couple the two columns. Separation methods are used that are mutually compatible and the eluting solvent in one dimension does not (significantly) influence the migration of analytes in the other dimension.

From a chromatographic perspective, this setup present limitations in terms of separation capacity and implementation, as it can be used only to couple IEC and RPLC separations (resulting in limited orthogonality) and employs pulsed elution in the first separation, leading to "undersampling" [23] and reducing the separation efficiency of the first separation dimension.

Other approaches have been investigated, which use online and offline comprehensive two-dimensional liquid chromatography (LC × LC) [*e.g.* Ref.[16].[24]. [25]. [26]. The advantage of an online setup is the automated sample transfer. It has been shown that RPLC (at a different pH), IEC and HILIC are good choices in combination with RPLC in a two-dimensional separation system.

In terms of selectivity, hydrophilic-interaction LC (HILIC) has emerged as a favourable alternative to IEC. The mechanism proposed for HILIC is based on partitioning between a water layer formed on the surface of the stationary phase and the bulk mobile-phase, with additional effects of electrostatic interactions and hydrogen bonding. Analytes are separated mainly based on their hydrophilicity, whereas in IEC separation is based on charge (peptides of interest generally carry two to four positive charges). HILIC provides a higher degree of orthogonality with RPLC [25,27] and can be easily coupled to MS when volatile additives are used. However, in HILIC, high concentrations of acetonitrile (ACN) are used, which creates a solvent mismatch with RPLC [28]. To overcome the solvent incompatibility, two-dimensional (2D) separations have been performed off-line, with the solvent being evaporated in between the two separations and replaced by a favourable solvent for the second-dimension (^{2}D) separation. If a 2D-LC separation is performed on-line, very small amounts can be transferred with the aid of loops. However, this may lead to sensitivity issues.

With the implementation of stationary-phase-assisted modulation (SPAM), the solvent mismatch can be overcome and narrower columns can be used in the second dimension [29]. In this case the simple loops between the two dimensions are replaced by small columns (traps) that allows concentration of the analytes in a smaller volume. The ^{1}D effluent is not transferred directly and, thus, the flow rate is not a limiting factor. The ^{1}D effluent can be diluted with a make-up flow, which will aid in solving the solvent incompatibilities. A similar type of set-up is already implemented for large volume injections in nano-HPLC (trap-and elute setup) [30]. To avoid volume overload on an analytical column, the sample can be loaded onto a pre-column (trap) at higher flowrates, using a loading pump and then a valve is switched to send the sample to the column using a solvent gradient.

The objectives of the present study were to develop a high-resolution LC × LC method to increase the number of analytes identified in bottom-up-proteomics studies and to enhance the coverage of the proteome. We aimed to combine HILIC with RPLC in capillary column format, using active modulation with dilution during sample transfer to circumvent solvent-incompatibility issues, all without increasing the required analysis time compared to a 1D separation.

2. Materials and methods

2.1. Chemicals

Acetonitrile (ACN, MS grade), 2-propanol (IPA, HPLC grade) and water (MS grade) for all MS measurements were purchased from Biosolve Chimie (Dieuze, France). Milli-Q water (18.2 mΩ) was obtained from a purification system (Millipore, Bedford, MA, USA). Bovine serum albumin (BSA, ≥96%) and formic acid (FA, analytical grade; 98%), were purchased from Sigma-Aldrich (Darmstadt, Germany). Yeast cell lysate was obtained from Promega (Madison, WI, USA).

Fused-silica capillaries (20, 50, 75, 100, 150 μm internal diameter (ID); all 360 μm outer diameter (OD)) (CM scientific, Silsden, United Kingdom), PEEK tubing (IDEX, Lake Forest, IL, USA), Next-advance frit kit (kasil 1624 and formamide; Troy, NY, USA), ferules, nuts and unions (Vici-Valco, Houston, TX, USA) were used to prepare the columns and connections.

C18 particles (3 μm diameter, 100 Å pore size) for the reversed-phase columns were obtained from NanoLCMS Solutions (Magic C18, Oroville, CA, USA). HILIC columns were packed in our laboratory with hydroxyethyl A particles (5 μm, 300 Å) from PolyLC (Columbia, MD, USA), Acquity UPLC BEH Amide (1.7 μm, 130 Å) from Waters (Milford, MA, USA), and ZIC-HILIC (3.5 μm, 100 Å) from Merck (Darmstadt, Germany).

2.2. Instrumentation

For the 1D and 2D experiments an Ultimate 3000 (ThermoFisher Scientific, Breda, The Netherlands) system was used, equipped with nano-pump (0.1–1.5 μL/min, NCP 3200RS) and nano-pump/loading-pump (0.1–1.5 μL/min nano pump, 1–100 μL/min loading pump, NCS 3500RS) modules and an auto-sampler (WPS-3000TPL RS). ProFlow nano selectors (ThermoFisher Scientific) were used as flow selectors for the nano pump. For the MS measurements an Orbitrap Q Exactive plus (QE) from ThermoFisher Scientific was used with a nanospray Flex source. The sampler was equipped with a 20-μL loop for the RPLC measurements, a 1-μL loop for HILIC measurements, or a 5-μL loop for the 2D measurements.

2.3. Sample preparation

The protein digestion was performed as described previously [31]. The peptide samples were dissolved either in 98% water, 2% ACN, 0.1% FA for RPLC measurements or 80% ACN, 20% 10 mM ammonium formate in water, pH 3 for HILIC and 2D measurements.

Human kidney tissue was obtained in homogenized fresh frozen form (Jasper Kers, Amsterdam UMC, Amsterdam, The Netherlands). The tissue aliquot (5 mg, 0.2 g mL^{-1} in 100-mM ammonium bicarbonate solution) was assumed to contain 1 mg of protein. The tissue was sonicated with 100 μL of 6 M urea solution, after which the same in-solution digestion procedure was followed.

Human IMR90 lung fibroblast cells (ATCC CCL-186) were grown in Dulbecco's modified Eagle media (DMEM) supplemented with 10% heat-inactivated, sterile-filtered fetal bovine serum (Life Technologies) and prepared according to what was described in Ref. [32]. The extracted proteins were precipitated with cold acetone (−20C). The acetone was then removed and the palette was suspended in 100 μL of 6 M urea solution. The same in-solution digestion was performed as described above.

2.4. Column packing

Firstly, frits were created in the fused silica capillaries. Kasil 1624 (60 μL) was mixed gently with formamide (20 μL). The capillaries

cut to the desired column length plus 30 mm were dipped in the solution for 1–2 s [33]. Then the capillaries were placed in an oven at 100C overnight. Before packing, the frit was cut to 1–2 mm and the capillary to the desired column length plus 5 mm.

The RPLC columns were packed by preparing a slurry of 0.1 g mL^{-1} particles in the packing solvent (methanol:isopropanol 50:50). The slurry was inserted in an empty column (4.6 mm × 50 mm) with one end drilled to be able to connect the capillary. The capillary was inserted 5 mm inside the packing column. A flow of 0.1 mL min^{-1} of the packing solvent was set. The column was kept under flow for 30 min after the packing was complete. The slurry left was collected and the packed column was removed. The top 5 mm were cut and the capillary was placed at the same level as the sleeve before making the final connection.

The same procedure was followed with the HILIC stationary phases, but the packing solvent was changed to 97% ACN, 3% 10-mM ammonium formate in water, due to the nature of the particles.

2.5. Chromatographic conditions

2.5.1. HILIC-MS separation conditions

Three different HILIC stationary phases were tested for the ^{1}D separation (ZIC HILIC, amide and hydroxyethyl). All were packed in-house and all columns were made to the same length (200 mm) and internal diameter (200 μm).

In this setup only one pump is used. The column is connected to the auto-sampler valve and no traps are used. The flow rate was set to 1 μL min^{-1} and 1 μg of yeast digest was loaded on the column. Mobile phase A was 10-mM ammonium formate in water, adjusted to pH 3, and mobile phase B was 97% ACN and 3% 10-mM ammonium formate, pH 3. A multi-segment gradient was used as follows: 95% B for 1 min, 95-85% B in 2 min, 85-75% B in 59 min, 75-65% B in 39 min, 65-50% B in 1 min and then back to 95% B in 1 min. The column was equilibrated with 95% B for 30 min.

2.5.2. RPLC-MS conditions

More attention was paid to the second dimension, since this has a greater impact on the final separation. The aspects considered were the efficiency of the separation (section 2.5.2A and B) and the detection sensitivity (section 2.5.2C). The mobile phases used were water with 2% ACN and 0.1% formic acid for channel A and the loading pump, and 80% ACN, 20% water, 0.1% formic acid for channel B.

For the RPLC optimization, all measurements were carried out in trap-and-elute mode using a 5 mm length, 300 μm ID trap column (C18, ThermoFisher Scientific). The loading pump (20 μL min^{-1}) was used to transfer the sample from the injector to the trap in 3 min. The column was kept in a column oven at 45C. A gradient from 5% B to 35% B was used for all measurements. For the peak integration and the determination of the peak width at half height MZmine (version 2.40.1) was used. The processing method will be detailed in section 2.7. Bovine-serum-albumin digest was chosen to represent a real proteomics sample and was used to characterize the ^{2}D columns. The sample was dissolved in water containing 2% ACN and 0.1% FA in concentrations of 0.2 μg μL^{-1} and 0.05 μg μL^{-1}.

2.5.2.1. RPLC-MS experiments to determine the optimal linear flow velocity. BSA digest 0.25 μL (50 ng, 0.2 μg μL^{-1}, quantified at protein level) was loaded on an RPLC column (C18, 10 mm length, 100 μm ID) to determine the effect of linear flow velocity on the peak capacity. The flow rates employed were 0.3, 0.5, 0.7, 0.9, 1.2, and 1.5 μL min^{-1}. The gradient time was varied in order to keep the same volumetric gradient (t_g/t_0) (see Table S1.1). The identified peptides from all the measurements were compared. From the common list of peptides, the most intense 35 peaks were chosen to

perform the peak capacity calculations (supplementary material Table S1.2). Equation (1) was used to calculate peak capacity [34] (derived in supplementary material *Equations S1.1, S1.2, S1.3*) and equations 2 and 3 were used to convert from volumetric flow rate to linear flow velocity.

$$n_c = \frac{t_g}{1.7 \times w_{\frac{1}{2}h}} + 1 \qquad (1)$$

Where n_c is the peak capacity, t_g the gradient time (min) and $w_{\frac{1}{2}h}$ is the peak width at half height (min).

$$\mu_s = \frac{F}{\pi \times r^2} \qquad (2)$$

Where μ_s is the linear flow velocity in an empty cylinder ("superficial velocity") (mm s^{-1}), F is the volumetric flow rate (mm^3 s^{-1}) and r is the column radius (mm)

$$\mu_t = \mu_s / \varepsilon_t \qquad (3)$$

Where μ_t is the mobile-phase velocity and ε_t is the total porosity of the column. We assumed a medium porosity column which would give a value of about 0.6 for ε_t [35].

The optimal mobile phase velocity was determined to be approximately 1.8 mm s^{-1} (Figure S1.1 supplementary material).

2.5.2.2. RPLC-MS effect of gradient steepness on peak capacity.

BSA digest (50 ng injected on column for all experiments) was used as the peptide sample. Three RPLC column dimensions were tested (100 mm length; 75, 100 and 150 μm ID) at an optimal linear flow velocity of 1.8 mm s^{-1} (see section 2.5.2A), which corresponded to 300, 533 and 1200 nL min^{-1}, respectively. The gradient times used were 3, 4, 5, 7, 10 and 15 min. The widths of 15 peptide peaks were used to calculate the peak capacity of each separation (for a list see supplementary material, Table S2.1).

2.5.2.3. RPLC-MS effect of flow rate on sensitivity.

RPLC columns with the same dimensions as those described in section 2.5.2B were used at their optimal linear flow velocity. The gradient time was kept constant (10 min) and the quantity of BSA digest loaded was varied (5, 10, 20 and 50 ng). The areas of extracted ion currents (EIC) for selected peptides (same list as in 2.5.2B) were considered to compare the MS sensitivity at different flow rates.
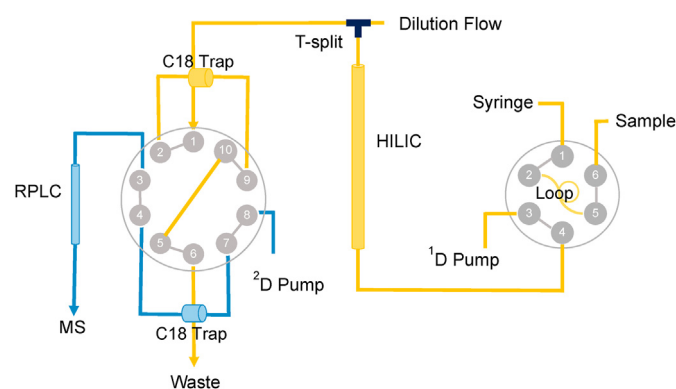
Fig. 1. Schematic of the 2D-LC setup used. The 6-port valve represents the injector valve and the 10-port valve is the modulation unit allowing to collect fractions of the sample from the first dimension and inject these in the second dimension. The 10-port valve and the RPLC column were kept in an oven at 45C.

2.5.3. HILIC × RPLC-MS optimization

A schematic of the HILIC × RPLC-MS setup is shown in Fig. 1. A HILIC ^{1}D column was run at 1 μL min^{-1}. The ^{2}D RPLC column was run at 1.2 μL min^{-1}, either with the 100-μm ID column or with the 150-μm ID column. A dilution flow of water containing 0.1% FA (9 μL min^{-1}) was provided by a loading pump, which was connected to the outlet of the HILIC column using a VHP MicroTee Assembly for 360 μm OD (P/N UH-750, Idex, Lake Forest, IL, USA). The tenfold diluted flow ^{1}D effluent containing the sample was retained on C18 traps (5 mm × 300 μm ID). This allows for the entire sample (peptides) to be transferred from the first to the second dimension in a solvent favourable for RPLC separation. The traps were attached to the valve using two Viper connections (10 mm × 30 μm ID; ThermoFisher Scientific). The other connections were made with fused-silica capillaries (250 mm × 20 μm ID, except 300 mm × 50 μm ID for the connection from the T-split to the valve, Fig. 1).

The gradient and valve switches were programmed using Xcalibur. A long step gradient was programmed for the first dimension and repeated short gradients for the second dimension. The ^{2}D gradient duration was set equal to the modulation time minus 1 min. This latter minute was used to wash the RPLC column increasing the % B to 80% in 0.1 min. Thereafter, the composition was programmed to return to the initial conditions in 0.4 min and the column was then equilibrated for 0.5 min. The first ^{2}D gradient ran from 5% B to 60% B. The gradient was then changed gradually, with the initial %B increasing linearly to 15% and the final %B decreasing linearly to 35%. This was because the more-hydrophobic components eluted in the first modulations from the HILIC 1D column required a higher % of ACN to be eluted from the ^{2}D RPLC column than the more-hydrophobic components eluting in the later modulations.

Multiple modulation times were considered (5, 10, 15 and 30 min) to determine a compromise between the sampling of the ^{1}D chromatogram and the resolution of the ^{2}D separation. A complication factor is that increasing the number of modulations implies that the fraction of the time lost due to sample loading and column equilibration is increased.

In our setup, the volume of the trap and the valve ports used were calculated to be about 640 nL. At a flow rate of 300 nL min^{-1} this will result in a delay in excess of 2 min. This may be quite significant. For example, in case of a 10 min modulation a 2 min delay represents a 20% reduction of the ^{2}D separation time. By increasing the flow rate to 1200 nL min^{-1} we can reduce the effect of the system volume to 0.5 min. However, we must still consider the dead volume of the ^{2}D column. A 100 mm × 75 μm ID column (RP75) run at 300 nL min^{-1} will have a dead time of about 1 min. If we increase the column ID, while keeping the length and linear flow velocity constant, the column dead time will stay the same. Running the separation at a higher than optimal linear flow velocity would reduce the dead time and the system dwell time simultaneously. The ^{2}D setup was therefore operated with the 150-μm or 100-μm ID columns (RP 150, RP100) at 1.2 μL min^{-1}.

2.6. Mass-spectrometry conditions

All separations were performed with direct injection into the nano-LC setup coupled, using a nanospray Flex source, to a Q-Exactive Plus (Thermo Fisher, Bremen, DE).

The tune method was set with the following parameters: capillary temperature 275C, S-lens RF level 55, and spray voltage between 1.85 and 2.1 kV, depending on the status of the emitter. The MS1 conditions were set to a resolution of 70,000, automatic gain control (AGC) target 3·10^6, maximum IT 60 ms, and a scan range from $m/z = 375$ to 1575. The MS2 parameters were:

resolution 17,500, AGC target 10^5, maximum IT 100 ms, isolation window ($\Delta m/z$) 1.8, fixed first mass $m/z = 110$, normalized collision energy (NCE) 27, number of MS/MS (loop count) 10. The data-dependent settings were a minimum AGC target of $2\cdot10^3$, intensity threshold $2\cdot10^4$, and charge exclusion of unassigned species, charge of 1, 7, 8 and larger than 8. The dynamic exclusion was set to 20 s and the chromatographic peak width was set to 15 s for the 1D experiments and 10 s for the 2D experiments.

2.7. Data processing

2.7.1. Bottom-up MS/MS database searches

MaxQuant (open-source computational platform, version 1.6.5.0) and Proteome Discoverer (Thermo Scientific, version 2.4.1.15) were used to identify the peptides in all samples. Carbamidomethyl (C) was used as a fixed modification and the variable modifications were set to oxidation (M) and acetylation (protein N-Terminal). Trypsin was specified as the enzyme, with a maximum of two missed cleavages. The false discovery rate (FDR) for the peptide identification was set to 1% [36].

FASTA files [37] with reviewed peptide sequences were downloaded from uniprot.org [38]. For the yeast measurements, baker's yeast was selected and only the reviewed sequences were downloaded in uncompressed form (created on July 05, 2019). For the kidney-tissue samples and the IMR90 cell-line sample, reviewed sequences of the complete human proteome were downloaded (created on February 12, 2019). To speed up the identification, for BSA only the reviewed sequences for the protein were downloaded (created on June 13, 2019) and not the bovine proteome. Information on the peptides, proteins, peptide-to-spectrum matches (PSMs), and retention times was gathered from both MaxQuant and Proteome Discoverer.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [39] partner repository with the dataset identifier PXD022330 and 10.6019/PXD022330.

2.7.2. Extraction of chromatographic characteristics

Identified sequences from MaxQuant or Proteome Discoverer were used to create csv files for targeted peak detection in MZmine (the peptide lists can be found in supporting materials S1 and S2). For each type of experiment an intermediate gradient duration was chosen for creating the csv file containing m/z values, retention times and sequences identified. The assignment was performed using the targeted-peak-detection function with the correct peak list, intensity tolerance 50%, noise level $6.0\cdot10^3$, m/z tolerance 0.02 Da or 10 ppm. The retention time tolerance was set to an arbitrary high value of 15 min due to the very different conditions between some of the measurements. The deconvolution function was used on the obtained peak lists to determine the charge state of the species. The peak integration was assessed visually and all peptides with poor integration were removed from further consideration. The retention time, peak width at half height and charge state were extracted for each peptide analysed.

2.7.3. Orthogonality calculations

Scatter plots between the HILIC and RPLC data were created (see Fig. 2), using the retention times of common peptides. If one sequence was identified multiple times the sequence with the highest identification score was chosen. The retention times of the peptides were normalized using equation 4.

$$t_{i \text{ (normalised)}} = \frac{t_i - t_{min}}{t_{max} - t_{min}} \tag{4}$$

Where t_{min} and t_{max} represent the first and the latest eluting peak in the chosen series, respectively, and t_i is the retention time of the chosen peak. The orthogonality scores were calculated with a modified Matlab script (courtesy of John Mommers, DSM Material Science Centre, Geleen, The Netherlands) [40], in which the bin-counting method [27] and the equations for the Asterisk approach [41] were implemented. The results from the bin-counting method is dependent of the number of analytes in the sample. It determines the space coverage as the percentage of occupied bins in the 2D separation space. Ideally, each analyte will occupy one bin. Thus, the number of bins should match the number of analytes to be able to make a fair conclusion. Therefore, the bin-counting method was modified so as to have a number of bins similar to the number of analytes included in the scatter plots. For the Asterisk method the same equations were applied, independent of the number of analytes included in the calculations. The Asterisk metric calculates the orthogonality from the distances of each peak to four axes, the vertical axes, the horizontal axes and the two diagonals. The final orthogonality is given as a percentage.

3. Results and discussion

Online coupling of HILIC $\times$ RPLC for the separation of peptides shows great promise in increasing the peak capacity without a great increase in analysis time. This is expected to lead to a better analysis of complex proteomics samples. In this project a 2D setup was built with micro flow rates in both dimensions. The effects of an increase in flow rate in the second dimension on peak capacity and MS sensitivity were investigated. Three samples with different levels of complexity were used, *i.e.* bovine serum albumin digest (one protein), yeast-proteome digest (about 6700 proteins) and human tissue (about 20,000 proteins).

3.1. First-dimension HILIC - orthogonality with respect to RPLC

We tested three types of HILIC stationary-phase chemistries, *viz.* Two neutral phases (hydroxyethyl and amide) and one charged type (sulfobetaine, ZIC-HILIC). The method was developed on the basis of previously described research [31]. HILIC-MS methods were optimized in order to spread the yeast peptides across the chromatogram. A total of 8908 unique peptides were identified from the three HILIC columns and the RPLC column. The highest number was obtained with the RPLC separation (7904 peptides, 1497 proteins) followed by ZIC HILIC (3098 peptides, 695 proteins), amide HILIC (2432 peptides, 584 proteins) and hydroxyethyl HILIC (2389 peptides, 556 proteins). The larger number of identified peptides can be explained by a more-efficient separation in RPLC compared with HILIC, augmented by the lower internal diameter and flow rate used for RPLC, while keeping the amount injected constant. The calculated peak capacity for the RPLC separation was 398. For the HILIC columns the calculated peak capacity was 158 for ZIC HILIC, 84 for the amide column and 141 for the hydroxyethyl column.

The identified peptides from the separations on the three HILIC columns and the RPLC column were compared and a common list was obtained (1064 peptides). To select the HILIC column with the highest separation power and orthogonality with RPLC, the retention times were normalized and scatter plots were created (Fig. 2). To assess the orthogonality both the asterisk equation [41] and the bin-counting method [27] were adopted.

The column showing the least correlation and the highest
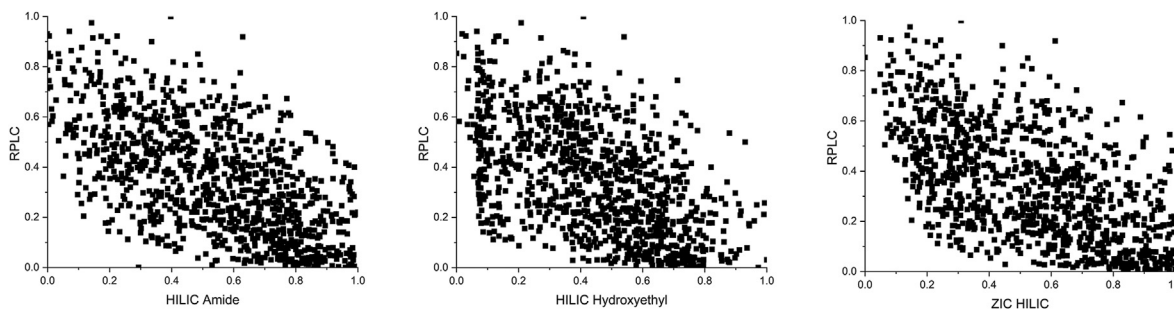
Fig. 2. Scatter plots of retention data obtained from one-dimensional separations using RPLC and three different HILIC columns. A common list of peptides between the four measurements was considered. Axes show normalized retention times.

Table 1
Orthogonality considerations of the three HILIC column compared to RPLC separation. 1064 common peaks were considered and the number of bins was 1024 (32×32).

Column	Bin %	Asterisk
HILIC Amide	48	59
HILIC Hydroxyethyl	42	51
HILIC ZIC	46	60

orthogonality based on the Asterisk metric was the ZIC HILIC column, followed by amide and then the hydroxyethyl. When considering the bin-counting method, the amide columns was found to outperform the ZIC HILIC column by 2% (see Table 1). Another observation was a clustering of peaks at the beginning of the separation on the hydroxyethyl column. If this were to be improved the orthogonality would be increased. It can be concluded that any of the three column chemistries would provide a largely orthogonal separation when coupled with RPLC. Irrespective of the column chemistry, the peptide retention order on the HILIC columns is quite similar (Figure S3.1, supporting information). Due to its higher peak capacity, higher number of proteins identified and higher orthogonality to RPLC, the ZIC HILIC column was selected for use in the 2D setup.

3.2. Optimization of second-dimension LC conditions: influence of column dimensions on peak capacity and detection

The second dimension was developed to provide a fast, efficient separation, while maintaining the sensitivity of low-flow LC-MS analysis. The parameters investigated were column internal diameter, flow rate and gradient duration, to establish a balance between speed of analysis and efficiency.

Potential loss in peak capacity, when using fast ^{2}D gradients to improve the first-dimension sampling rate, was investigated. Theoretically the peak capacity of the 2D-sysytem will be the product of the peak capacities of the two dimensions used. However, since in a low-flow setup adequate sampling of each ^{1}D peak is not possible, the total peak capacity will be better approximated as the peak capacity of the second dimension multiplied by the number of modulations.

In all of the three columns (with different dimensions) investigated an increase of the peak capacity approximately with the square root of the gradient duration was observed. As seen in Fig. 3-A, the 75-μm ID column presented peak capacities between 26 and 59 for the tested gradient lengths. The 100-μm ID column showed peak capacities from 32 to 89 and the 150-μm ID column exhibited a peak capacity of 43−117. Larger columns, running at higher flow rates, resulted in more-efficient separations.

Another consideration was the influence on the MS response,

due to changes in sample dilution. Introduction of miniaturized ESI sources have shown great promise for increasing the sensitivity of detection and for allowing smaller sample quantities to be used. The standard setup for contemporary proteomics features a 75-μm ID column packed with C18 particles, run at 300 nL min^{-1} [42]. This has been a compromise between sensitivity, enhanced by lower flow rates, and robustness of the setup. Lately, the advantages of micro-flow liquid chromatography (columns of around 1 mm ID) for proteomics have been discussed [43]. The separation efficiency improves by having sharper peaks and avoiding column overloading, but the MS sensitivity suffers due to sample dilution, hence an ion-suppression effect.

In this work, we explored the use of second-dimension columns with slightly larger diameters than the standard 75-μm ID column, to improve the sample loading and to obtain sharper peaks (higher peak capacity). In Fig. 3-B a 1.3-fold decrease in sensitivity is observed when going to the 100-μm ID column and a 2.3-fold decrease when using a 150-μm ID column.

Clearly, we must compromise the MS sensitivity to have a faster and more efficient separation. However, by increasing the quantity of peptides loaded we could overcome the decreased sensitivity, while maintaining an increased peak capacity.

3.3. 2D platform − optimization and application

The low flow 2D-LC setup was used to analyze yeast-proteome digest as a standard test sample and finally was applied for the separation of fresh-frozen human-kidney-tissue homogenate and an IMR90 cell-line lysate. The protein digestion was performed in-house as descried previously [31]. Stationary-phase-assisted modulation was used to transfer the fractions from the first dimension to the second dimension. In this way, the solvent incompatibility issues between HILIC and RPLC were solved and all sample could be transferred automatically. The RPLC was coupled to ESI-MS for the analysis of the sample.

3.3.1. Modulation time

When developing a miniaturized LC × LC setup some of the critical aspects to consider are the (column) dead volume and the (system) dwell volume and their influence on the second-dimension separation. In analytical-scale setups, system dwell volumes are small with respect to the flow rates used (e.g. 100 μL of dwell volumes at flow rates of 1 mL min^{-1} result in a 6-s dwell time), allowing fast second dimension-separations (e.g. below 1 min) and frequent sampling of the first-dimension separation, significantly increasing the peak capacity of the 2DLC separation system [44]. However, in nano-LC the dwell volumes tend to be relatively large with respect to the operative flow rates, introducing significant time gaps between second-dimension runs. This can be
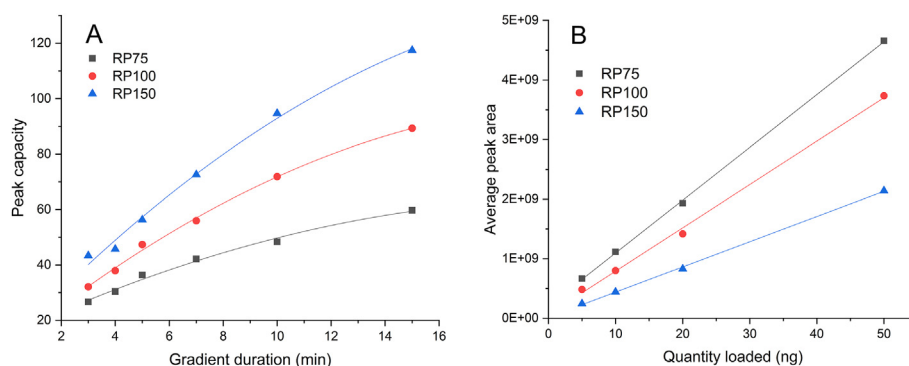
Fig. 3. (A) Peak capacity of three RPLC column with different dimensions as a function of gradient duration. Linear flow velocity was kept at 1.13 mm s-1. (B) MS sensitivity of three RPLC columns with different dimensions determined from the average peak area of 14 peptides. Black (□) 75 μm ID, red (●) 100 μm ID, blue (Δ) 150 μm ID columns. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

clearly observed in a series of modulations as analysis gaps in which no peptide is observed by the mass spectrometer.

We investigated the effect of the modulation time and the number of modulations per run. The duration of the HILIC gradient was kept constant at 2 h, independent of the modulation time tested. The total run time was expanded by up to 20 min to accommodate the final modulation. We expected more-efficient ^{2}D separations for longer modulation times. However, the ^{1}D separation will be jeopardized by undersampling. The modulation times tested were 5, 10, 15, and 30 min. The resulting numbers of modulations were 24, 12, 8, 4 respectively. A table documenting the peptide and protein identification in each case is included in supplementary material (Table S4.1).

When considering the 150-μm ID ^{2}D column run at the optimal linear flow velocity (1.2 μL min^{-1}; see Supplementary Material *S1*), the dwell volume of 0.64 μL of our system will result in a dwell time (^2t_D) of 0.5 min, while the column dead ($^2t_0$) time is 1 min, resulting in a total time loss per modulation ($^2t_D + {}^2t_0$) of 1.5 min. This implies that 30% of the time will be lost in case of 5-min modulations, 15% for 10-min modulations, 10% for the 15 min modulations, and only 5% for the 30-min modulations. If we replace the second-dimension column with a 100-μm ID one and we keep the flow rate the same, the $^2t_0$ is reduced to 0.5 min, and $^2t_D + {}^2t_0$ becomes 1 min. This implies that we regain 10% of the separation space for the 5-min modulations and 5%, 3.3% and 1.7% for the 10-min, 15-min and 30-min modulations, respectively (see supplementary material Figure S4.1).

However, the separation efficiency should be lower when using a smaller column run at a much higher linear flow velocity. The peak capacities for the two setups were calculated for one of the modulations, assuming that for all modulations in one run the second-dimension peak capacity (^2n) should be similar. One 10 min modulation run was selected for both the RP100 and RP150 columns. 15 peptide peaks were selected from modulation #7, 15 values for peak capacity were estimated based on the band widths of each of the peptides, and these 15 values were then averaged (for a list of peptides see supplementary material, Table S4.2). The RP150 showed a peak capacity of 83 and the RP100 setup one of 81. Assuming that the total peak capacity of the system equals the peak capacity of the second dimension multiplied by the number of modulations ($n_{mod} = 12$). Therefore, the total peak capacity of the system will be about 1000. This is a great improvement when considering the peak capacity of the 1D RP75 separation of the same 2-h duration (about 400).

Due to the very small difference in the obtained peak capacities of the two setups the 100 μm ID column was thought to provide a

better option due to the lower dead time ($^2t_D + {}^2t_0$). The longer modulations were determined to be more suitable for our setup, providing enough fractionation of the sample in the first dimension and a more limited time lost due to the dead time system.

3.3.2. Method optimization and repeatability

Triplicate measurements were performed to check the repeatability of the 2D-setup. The RPLC column was kept in a column oven at 45C to minimize the effects of fluctuations in the room temperature and to decrease the backpressure.

The identified species (about 4800 peptides per run) were not considered a good indication of the repeatability, due to the variability of the sample, ionization efficiency and precursor selection for MS/MS. Therefore, the retention times of a selection of the identified peptides were considered to determine the degree of variation between the triplicate measurements. We considered the common peptides without post-translational modifications from the three measurements (2742 peptides). We found an average variability of 1.7% (RSD) in the retention times between the triplicate measurements. Of the common peptides 192 showed a variability above 1%. If these were to be removed the average variability for the remaining 2550 peptides would drop to 0.01%. The observed variability may also be influenced by an incorrect assignment of an m/z value to a peptide sequence. We observed that multiple retention times (peaks) were assigned to some of the peptides. We used the identification score in Proteome Discoverer (xcorr) to filter out the duplicate assignments. During this filtering the wrong retention time may be removed, leading to an increased variability in the retention time for the triplicate measurements. Due to the large number of common peptides manual verification was not feasible. For an example of three overlaid extracted-ion-current chromatograms see supplementary material (Figure S4.2).

After the system was deemed repeatable, further considerations were made to optimize the method. The optimization of a 2D-LC system can be difficult and time consuming, due to the many parameters that can influence the efficiency of both the ^{1}D and ^{2}D separations, such as the linear flow velocity, the column temperature, the mobile-phase composition and the gradient program [44]. In this work, the parameters investigated were the linear flow velocity using two different column diameters in ^{2}D, the mobile-phase composition and the gradient duration. The latter factor is closely related to the modulation time. The optimization was performed by visual assessment of the obtained chromatograms.

Initially, it was observed that the first ^{2}D RPLC modulations needed a higher concentration of ACN to elute the peptides as compared to later modulations. This was expected, since more-

hydrophobic peptides will elute first in HILIC (^{1}D) and will require higher concentrations of ACN for elution in RPLC (^{2}D). A variable ("shifting") gradient was programmed, where the initial concentration of B (80% ACN, 20% water, 0.1% formic acid) was decreased linearly from 15% to 5% and the final concentration of B from 60% to 35% from the first to last modulation. A visual representation of the effect of the variable gradient compared to an identical repeated gradient can be found in supplementary materials (Figure S4.4).

The HILIC gradient was also modified to improve the spreading of the peaks. Initially the gradient was started at 95% B and held for 5 min, followed by a first step from 95% to 80% B in 40 min, then to 55% B in 72 min. The column was washed by going to 40% B in 1 min, to 70% B in 2 min, to 40% B in 1 min and back to 95% B in 1 min. The column was equilibrated for 28 min. With this gradient program it was observed that the middle fractions contained larger numbers of peaks. The elution of peptides earlier in the HILIC gradient was promoted by starting the gradient at 90% B and adding a step to 85% B in 5 min. The next step was from 85% to 70% B in 60 min, followed by a decrease to 50% B in 52 min. The washing steps were the same as in the previous gradient program. The column was equilibrated at 90% B for 28 min.

After the optimization of the HILIC gradient, the RPLC gradient for the 15 min modulation was optimized for each modulation individually (see Fig. 4). In supplementary materials Figure S4.5 a clear improvement in the separation can be observed for the manually adjusted gradient for each modulation versus the shifting linear gradient described above (see supplementary material figure S4.6 and S4.7). The improvement was also noticed in the number of peaks that could be subjected to MS/MS analysis ("sequenced peaks") and the number of peptides identified. For a generic shifting gradient, the number of peaks was 39,107 and 4529 peptides were identified, while for the manually optimized variable gradients the number of peaks increased to 49,194 and the number of peptides identified was 6606.

3.3.3. Application of the 2D setup and comparison with 1D RPLC

The same yeast sample was also run with 1D-RPLC with two column dimensions, *i.e.* 75 μm ID (300 nL min^{-1}) and 150 μm ID (1.2 μL min^{-1}). The gradient duration was equal to that of the ^{1}D HILIC gradient in the 2D measurements (120 min). Dilution in the column is proportional to the square of the column diameter [42].
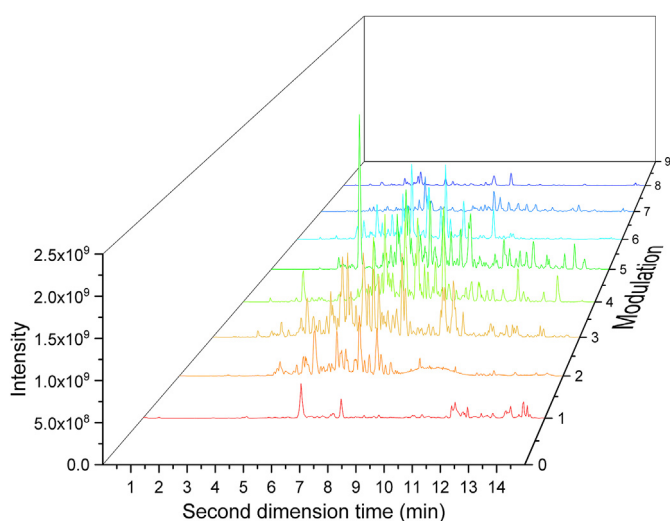
Therefore, a 4-fold greater dilution was expected when increasing the column diameter from 75 to 150 μm. This was also observed experimentally. The narrower column performed better when the same sample quantity was loaded (1 μg). However, a five-time increase in the sample quantity injected on the larger-diameter column led to more peptides being identified than with the 75-μm ID column. This beneficial effect is expected also for the 2D setup, since larger column diameters were used. Both 1 μg and 5 μg sample amounts were loaded during the 2D experiments to determine the effect of the sample concentration and to allow a comparison with the 1D runs.

The numbers of peptides and proteins identified are listed in Table 2. When 1 μg of sample was separated the RP75 column (1D) yielded the highest number of peptides identified, followed by the RP150 column and then the 2D separation (15 min modulation). However, for the number of proteins identified, the RP75 and 2D separations yielded similar values (nearly 1500), while about 300 fewer proteins were found with the RP150 column. When 5 μg were loaded, slightly more peptides were identified with the 2D and RP150 separations compared to the RP75 column (with 1 μg injected). The number of proteins found was comparable between RP150 (5 μg) and RP75 (1 μg). The 2D separation of 5 μg yeast-proteome digest yielded the highest number of proteins identified, with between 400 and 500 more proteins identified than on either 1D system (RP75, 1 μg or RP150, 5 μg). It can be concluded that dilution has a large effect on the identification of peptides in a sample as it decreases the intensity of the peaks present and consequently increase MS1 and MS/MS ion injection time, reducing the number of MS/MS scans that can be performed per analysis.

Finally, the setup was used for the characterization of proteomics samples from complex organisms: human-kidney-tissue proteome and IMR90 cells proteome. The same trends as with yeast were observed, but a larger number of proteins were identified, likely due to the more complex proteome.

For the IMR90 cells proteome we observed the highest number of identifications. We observed similar numbers of identified peptides and a slightly (5%) higher number of identified proteins using the 2D setup (15 min modulation) in comparison with the 1D separation on the RP75 column, when loading the same amount of sample (1 μg) on the two systems.

For the kidney-tissue-proteome, loading 1 μg of sample yielded similar results in terms of protein identification for the 1D (RP75) and 2D (15 min modulation) systems. When 5 μg of sample were loaded on the 2D system, the number of proteins identified

Fig. 4. 2D separation of yeast-proteome digest using 15-min modulations (120 min separation). ZIC HILIC as first dimension and C18 (100 μm ID) as second dimension. For conditions see Section 3.5.2 and supplementary materials figure S4.6.

Table 2
Comparison of results obtained with 1D 1D-RPLC compared to 2Dand 2D HILIC-RPLC separations of the yeast proteome, kidney-tissue proteome and IMR90 cells proteome.

File	Quantity loaded	Sample	Peptides	Proteins
1D RP75	1 μg	Yeast	7904	1497
1D RP150	1 μg	Yeast	6240	1130
1D RP150	5 μg	Yeast	8436	1457
2D (15 min modulation)	1 μg	Yeast	5905	1469
2D (15 min modulation)	5 μg	Yeast	8280	1931
1D RP75	1 μg	Kidney tissue	6629	1803
1D RP150	1 μg	Kidney tissue	5574	1351
2D (15 min modulation)	1 μg	Kidney tissue	5944	1764
2D (15 min modulation)	5 μg	Kidney tissue	8843	2750
2D (30 min modulation)	1 μg	Kidney tissue	5935	1868
1D RP75	1 μg	IMR90	10,005	2495
2D (15 min modulation)	1 μg	IMR90	9993	2624

increased by about 1000. This represents a 34% increase in the number of proteins identified using the 2D platform, without any increase in measurement time. The higher protein coverage available from the 2DLC-MS setup in the analysis of kidney samples suggest that 2D approaches could significantly help in the analysis samples from complex matrices.

Moreover, when comparing the chromatographic performance of the 1DLC analysis respect to the 2DLC analysis the average peak width of the peptides identified was 2.5 times higher. It is likely that the cycle time for the MS experiments in the MS system used in this study (between 0.9 and 1.3 s) does not allow to fully exploit the gain in chromatographic performance from the 2DLC-MS. We suggest that 2DLC methods in combination with the latest generation of orbitrap and or ion mobility -time of flight mass spectrometry instruments would allow to capture to a greater extent the chromatographic gains in terms of peak width reduction (and hence peak capacity) that 2DLC analysis allows for.

4. Conclusion

In this work we developed an online low-flow 2DLC setup for the separation of complex protein digest. The two separation dimensions were developed separately, keeping in mind their final application in the 2D-LC setup. The first dimension was chosen to be HILIC, due to its high orthogonality with RPLC. This was demonstrated with three types of column chemistries tested, which all showed a surface coverage in excess of 42% and asterisk orthogonality indices above 51%. The best performing column under the conditions tested was a ZIC HILIC column, which yielded a surface coverage of 46% and an asterisk orthogonality index of 60%. Therefore, this column was used in the 2D setup.

For the second-dimension separation a short run time, high peak capacity and good MS sensitivity are needed. For this purpose, three columns were investigated with different internal diameters (75, 100, and 150 μm ID) containing the same C18 particles (3 μm diameter) and having the same length (100 mm). The optimal linear flow velocity was determined to be about 1.8 mm s^{-1}, corresponding to volumetric flow rates of 300, 533 and 1200 nL min^{-1} in the three columns. The peak capacity increased, but the MS sensitivity decreased with increasing column diameter. The reduction in sensitivity was overcome by loading larger quantities of peptides.

Considering the implementation of the two-dimensional system, higher flow rates were needed to minimize the detrimental effects of the dwell volume and the column dead volume. Therefore, the 100-μm and 150-μm ID columns were used in the second dimension at a flow rate of 1.2 μL min^{-1}. The repeatability of the 2D setup was determined by considering the retention time variation. An average relative standard deviation of 1.7% was found. Manual, sample-dependent optimization of both dimensions (including variable gradients for each modulation) was needed to achieve the best result. In the future, computer-aided optimization would be highly attractive for an easier implementation of such a setup.

The 2D setup performed better than a one-dimensional separation with the 150-μm ID RPLC column. When the quantity of peptides was adapted the 2D setup also performed better than the 75-μm ID RPLC column. With the 2D setup we could obtain a much higher peak capacity in the same separation time (about 1000, as compared to 400 in 1D), which explains the better separation and higher identification capacity. A 17% increase in the number of proteins identified was obtained for the yeast proteome and a 34% for the human kidney-tissue-proteome with the 2D-LC setup compared to 1D RPLC (75 μm ID) when loading a five-fold higher quantity in the 2D setup.

CRediT authorship contribution statement

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2021.338349.

References

[1] M. Larance, A.I. Lamond, Multidimensional proteomics for cell biology, Nat. Rev. Mol. Cell Biol. 16 (2015) 269−280, https://doi.org/10.1038/nrm3970.

[2] Y. Zhang, B.R. Fonslow, B. Shan, M.-C. Baek, J.R. Yates III, J.R. Yates, Protein analysis by shotgun/bottom-up proteomics, Chem. Rev. 113 (2014) 2343−2394, https://doi.org/10.1021/cr3003533.Protein.

[3] M. Wilm, M. Mann, Analytical properties of the nanoelectrospray ion source, Anal. Chem. 68 (1996) 1−8, https://doi.org/10.1021/ac9509519.

[4] M.R. Emmett, R.M. Caprioli, Micro-electrospray mass spectrometry: ultra-high-sensitivity analysis of peptides and proteins, J. Am. Soc. Mass Spectrom. 5 (1994) 605−613.

[5] J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification, Nat. Biotechnol. 26 (2008) 1367−1372, https://doi.org/10.1038/nbt.1511.

[6] D.J.C. Pappin, D.M. Creasy, J.S. Cottrell, Probability-based Protein Identification by Searching Sequence Databases Using Mass Spectrometry Data Proteomics and 2-DE, 1999.

[7] M. Vaudel, J.M. Burkhart, R.P. Zahedi, E. Oveland, F.S. Berven, A. Sickmann, L. Martens, H. Barsnes, PeptideShaker enables reanalysis of MS-derived proteomics data sets: to the editor, Nat. Biotechnol. 33 (2015) 22−24, https://doi.org/10.1038/nbt.3109.

[8] B. Zhan, J.R. Yates, M.-C. Baek, Y. Zhang, B.R. Fonslow, Protein analysis by shotgun/bottom-up proteomics, Chem. Rev. 113 (2013) 2343−2394, https://doi.org/10.1021/cr3003533.

[9] S.S. Thakur, T. Geiger, B. Chatterjee, P. Bandilla, F. Fröhlich, J. Cox, M. Mann, Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation, Mol. Cell. Proteomics 10 (2011) 1−9, https://doi.org/10.1074/mcp.M110.003699.

[10] S. Eeltink, S. Dolman, R. Swart, M. Ursem, P.J. Schoenmakers, Optimizing the peak capacity per unit time in one-dimensional and off-line two-dimensional liquid chromatography for the separation of complex peptide samples,

J. Chromatogr., A 1216 (2009) 7368−7374, https://doi.org/10.1016/j.chroma.2009.02.075.

[11] M. den Ridder, P. Daran-Lapujade, M. Pabst, Shot-gun proteomics: why thousands of unidentified signals matter, FEMS Yeast Res. 20 (2020) 1−9, https://doi.org/10.1093/femsyr/foz088.

[12] M. Ke, H. Shen, L. Wang, S. Luo, L. Lin, J. Yang, R. Tian, Identification, quantification, and site localization of protein posttranslational modifications via mass spectrometry-based proteomics, Adv. Exp. Med. Biol. 919 (2016) 345−382, https://doi.org/10.1007/978-3-319-41448-5_17.

[13] P. Dugo, F. Cacciola, T. Kumm, G. Dugo, L. Mondello, Comprehensive multidimensional liquid chromatography: theory and applications, J. Chromatogr., A 1184 (2008) 353−368, https://doi.org/10.1016/j.chroma.2007.06.074.

[14] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions, J. Separ. Sci. 28 (2005) 1694−1703, https://doi.org/10.1002/jssc.200500116.

[15] V.-A. Duong, J.-M. Park, H. Lee, Review of three-dimensional liquid chromatography platforms for bottom-up proteomics, Int. J. Mol. Sci. 21 (2020) 1524, https://doi.org/10.3390/ijms21041524.

[16] G. Vanhoenacker, I. Vandenheede, F. David, P. Sandra, K. Sandra, Comprehensive two-dimensional liquid chromatography of therapeutic monoclonal antibody digests, Anal. Bioanal. Chem. 407 (2015) 355−366, https://doi.org/10.1007/s00216-014-8299-1.

[17] M. Kwiatkowski, D. Krösser, M. Wurlitzer, P. Steffen, A. Barcaru, C. Krisp, P. Horvatovich, R. Bischoff, H. Schlüter, Application of displacement chromatography to online two-dimensional liquid chromatography coupled to tandem mass spectrometry improves peptide separation efficiency and detectability for the analysis of complex proteomes, Anal. Chem. 90 (2018) 9951−9958, https://doi.org/10.1021/acs.analchem.8b02189.

[18] J. Jáčová, A. Gardlo, J.M.D. Dimandja, T. Adam, D. Friedecký, Impact of sample dimensionality on orthogonality metrics in comprehensive two-dimensional separations, Anal. Chim. Acta 1064 (2019) 138−149, https://doi.org/10.1016/j.aca.2019.03.018.

[19] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R.I. Yates, Direct analysis of protein complexes using mass spectrometry, Nat. Biotechnol. 17 (1999) 676−682, https://doi.org/10.1038/10890.

[20] M.P. Washburn, D. Wolters, J.R. Yates, Large-scale analysis of the yeast proteome by multidimensional protein identification technology, Nat. Biotechnol. 19 (2001) 242−247, https://doi.org/10.1038/85686.

[21] D.A. Wolters, M.P. Washburn, J.R. Yates, An automated multidimensional protein identification technology for shotgun proteomics, Anal. Chem. 73 (2001) 5683−5690, https://doi.org/10.1021/ac010617e.

[22] K.J. Webb, T. Xu, S.K. Park, J.R. Yates, Modified MuDPIT separation identified 4488 proteins in a system-wide analysis of quiescence in yeast, J. Proteome Res. 12 (2013) 2177−2184, https://doi.org/10.1021/pr400027m.

[23] G. Vivó-Truyols, S. Van Der Wal, P.J. Schoenmakers, Comprehensive study on the optimization of online two-dimensional liquid chromatographic systems considering losses in theoretical peak capacity in first- and second-dimensions: a pareto-optimality approach, Anal. Chem. 82 (2010) 8525−8536, https://doi.org/10.1021/ac101420f.

[24] E. Nägele, M. Vollmer, P. Hörth, C. Vad, 2D-LC/MS techniques for the identification of proteins in highly complex mixtures, Expert Rev. Proteomics 1 (2004) 37−46, https://doi.org/10.1586/14789450.1.1.37.

[25] P.J. Boersema, N. Divecha, A.J.R. Heck, S. Mohammed, Evaluation and optimization of ZIC-HILIC-RP as an alternative MudPIT strategy, J. Proteome Res. 6 (2007) 937−946, https://doi.org/10.1021/pr060589m.

[26] Y. Zhao, R.P.W. Kong, G. Li, M.P.Y. Lam, C.H. Law, S.M.Y. Lee, H.C. Lam, I.K. Chu, Fully automatable two-dimensional hydrophilic interaction liquid chromatography-reversed phase liquid chromatography with online tandem mass spectrometry for shotgun proteomics, J. Separ. Sci. 35 (2012) 1755−1763, https://doi.org/10.1002/jssc.201200054.

[27] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography, Anal. Chem. 77 (2005) 6426−6434, https://doi.org/10.1021/ac050923i.

[28] G. Groeneveld, M.N. Dunkle, M. Rinken, A.F.G. Gargano, A. de Niet, M. Pursch, E.P.C. Mes, P.J. Schoenmakers, Characterization of complex polyether polyols using comprehensive two-dimensional liquid chromatography hyphenated to high-resolution mass spectrometry, J. Chromatogr., A 1569 (2018) 128−138, https://doi.org/10.1016/j.chroma.2018.07.054.

[29] R.J. Vonk, A.F.G. Gargano, E. Davydova, H.L. Dekker, S. Eeltink, L.J. de Koning, P.J. Schoenmakers, Comprehensive two-dimensional liquid chromatography with stationary-phase-assisted modulation coupled to high-resolution mass spectrometry applied to proteome analysis of *Saccharomyces cerevisiae*, Anal. Chem. (2015), https://doi.org/10.1021/acs.analchem.5b00708.

[30] G. Mitulović, M. Smoluch, J.P. Chervet, I. Steinmacher, A. Kungl, K. Mechtler, An improved method for tracking and reducing the void volume in nano HPLC-MS with micro trapping columns, Anal. Bioanal. Chem. 376 (2003) 946−951, https://doi.org/10.1007/s00216-003-2047-2.

[31] L.S. Roca, S.E. Schoemaker, B.W.J. Pirok, A.F.G. Gargano, P.J. Schoenmakers, Accurate modelling of the retention behaviour of peptides in gradient-elution hydrophilic interaction liquid chromatography, J. Chromatogr., A (2019) 460650, https://doi.org/10.1016/j.chroma.2019.460650.

[32] P.F. Doubleday, L. Fornelli, N.L. Kelleher, Elucidating proteoform dynamics underlying the senescence associated secretory phenotype, J. Proteome Res. 19 (2020) 938−948, https://doi.org/10.1021/acs.jproteome.9b00739.

[33] M.F. Wahab, D.C. Patel, R.M. Wimalasinghe, D.W. Armstrong, Fundamental and practical insights on the packing of modern high-efficiency analytical and capillary columns, Anal. Chem. (2017), https://doi.org/10.1021/acs.analchem.7b00931.

[34] X. Li, D.R. Stoll, P.W. Carr, A simple and accurate equation for peak capacity estimation in two dimensional liquid chromatography, Anal. Chem. 81 (2009) 845−850, https://doi.org/10.1021/ac801772u.

[35] H.M. Quinn, A reconciliation of packed column permeability data: column permeability as a function of particle porosity, J. Mater. 2014 (2014) 1−22, https://doi.org/10.1155/2014/636507.

[36] D. Wang, B. Eraslan, T. Wieland, B. Hallström, T. Hopf, D.P. Zolg, J. Zecha, A. Asplund, L. Li, C. Meng, M. Frejno, T. Schmidt, K. Schnatbaum, M. Wilhelm, F. Ponten, M. Uhlen, H. Gagneur, H. Hahne, B. Kuster, A deep proteome and transcriptome abundance atlas of 29 healthy human tissues, Mol. Syst. Biol. 15 (2019) 1−16, https://doi.org/10.15252/msb.20188503.

[37] W.R. Pearson, D.J. Lipman, Improved tools for biological sequence comparison, Proc. Natl. Acad. Sci. U.S.A. 85 (1988) 2444−2448, https://doi.org/10.1073/pnas.85.8.2444.

[38] The UniProt Consortium, UniProt, A worldwide hub of protein knowledge, Nucleic Acids Res. 47 (2019) D506−D515, https://doi.org/10.1093/nar/gky1049.

[39] Y. Perez-Riverol, A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D.J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Pérez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg, Ş. Yilmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A.F. Jarnuczak, T. Ternent, A. Brazma, J.A. Vizcaíno, The PRIDE database and related tools and resources in 2019: improving support for quantification data, Nucleic Acids Res. 47 (2019) D442−D450, https://doi.org/10.1093/nar/gky1106.

[40] J. Mommers, S. van der Wal, Two metrics for measuring orthogonality for two-dimensional chromatography, J. Chromatogr., A 1586 (2019) 101−105, https://doi.org/10.1016/j.chroma.2018.11.081.

[41] M. Camenzuli, P.J. Schoenmakers, A new measure of orthogonality for multi-dimensional chromatography, Anal. Chim. Acta 838 (2014) 93−101, https://doi.org/10.1016/j.aca.2014.05.048.

[42] S.R. Wilson, T. Vehus, H.S. Berg, E. Lundanes, Nano-LC in proteomics: recent advances and approaches, Bioanalysis 7 (2015) 1799−1815, https://doi.org/10.4155/bio.15.92.

[43] Y. Bian, R. Zheng, F.P. Bayer, C. Wong, Y.C. Chang, C. Meng, D.P. Zolg, M. Reinecke, J. Zecha, S. Wiechmann, S. Heinzlmeir, J. Scherr, B. Hemmer, M. Baynham, A.C. Gingras, O. Boychenko, B. Kuster, Robust, reproducible and quantitative analysis of thousands of proteomes by micro-flow LC−MS/MS, Nat. Commun. 11 (2020) 1−12, https://doi.org/10.1038/s41467-019-13973-x.

[44] D.R. Stoll, X. Li, X. Wang, P.W. Carr, S.E.G. Porter, S.C. Rutan, Fast, comprehensive two-dimensional liquid chromatography, J. Chromatogr., A 1168 (2007) 3−43, https://doi.org/10.1016/j.chroma.2007.08.054.