Published in final edited form as:

Anal Chem. 2011 September 15; 83(18): 6996–7005. doi:10.1021/ac200639v.

An Online Nanoflow RP-SAX-RP LC-MS/MS Platform for Efficient and In-depth Proteome Sequence Analysis of Complex Organisms

Feng Zhou[†], Timothy W. Sikorski[‡], Scott B. Ficarro^{†,‡}, James T. Webber[†], and Jarrod A. Marto^{†,‡,§}

[†]Department of Cancer Biology and Blais Proteomics Center, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115-6084

[‡]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115-6084

Abstract

The dynamic range of protein expression in complex organisms coupled with the stochastic nature of discovery-driven MS/MS analysis continues to impede comprehensive sequence analysis and often provides only limited information for low-abundance proteins. High performance fractionation of proteins or peptides prior to mass spectrometry analysis can mitigate these effects, though achieving an optimal combination of automation, reproducibility, separation peak capacity, and sample yield remains a significant challenge. Here we demonstrate an automated nanoflow 3D LC-MS/MS platform based on high-pH reversed phase (RP), strong anion exchange (SAX), and low-pH reversed phase (RP) separation stages for analysis of complex proteomes. We observed that RP-SAX-RP outperformed RP-RP for analysis of tryptic peptides derived from *Escherichia coli*, and enabled identification of proteins present at a level of 50 copies per cell in *Saccharomyces cerevisiae*, corresponding to an estimated detection limit of 500 amol, from 40 µg total lysate on a low-resolution 3-D ion trap mass spectrometer. A similar study performed on a LTQ-Orbitrap yielded over 4000 unique proteins from 5 µg total yeast lysate analyzed in a single, 101 fraction RP-SAX-RP LC-MS/MS acquisition, providing an estimated detection limit of 65 amol for proteins expressed at 50 copies per cell.

Keywords

multidimensional separation; proteome profiling; nanoflow HPLC; RP-RP; reversed phase-reversed phase; RP-SAX-RP; reversed phase-strong anion exchange-reversed phase; online fractionation; two dimensional separation; three dimensional separation; LC-MS/MS

INTRODUCTION

Mass spectrometry-based proteomics is now well-established for characterization of proteins across a wide range of experimental contexts: recombinant protein expression systems, analysis of proteins resolved via SDS-PAGE,¹ multi-component protein complexes isolated through affinity purification,² post-translational modifications (phosphorylation,³⁻⁴ glycosylation,^{5,6-8} sulfation,⁹ sumoylation,^{10,11} and ubiquitination^{12,13}), protein sub-domains

[§]To whom correspondence should be addressed: Department of Cancer Biology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Smith 1158A, Boston, MA, 02215-5450, USA. Phone: (617) 632-3150 (office), Fax: (617) 582-4471, jarrod_marto@dfci.harvard.edu..

and termini,¹⁴⁻¹⁵ sub-cellular compartments,^{16,17} and finally whole proteomes.¹⁸⁻¹⁹ While advances in instrumentation continue to improve the discovery power of mass spectrometry in biomedical applications,²⁰⁻²⁶ it is nevertheless true that data quality often varies inversely with sample complexity. The deleterious effects observed during analysis of peptides present in complex matrices manifest in diminished performance of both MS and MS/MS scans and include: signal suppression of low-basicity peptides²⁷ and undersampling of low-abundance peptides during MS/MS analysis.²⁸ In addition, overlapping isotope profiles of co-eluting peptides can lead to errors in mass assignment, quantification and precursor selection.^{29, 30} Collectively these phenomena diminish the reproducibility of discovery-based experiments and inhibit characterization of low-abundance proteins.³¹⁻³² Biochemical fractionation of sub-cellular compartments and organelles, or affinity purification based on sequence tags or specific posttranslational modifications are effective strategies for analyte enrichment and reduction of sample complexity prior to LC-MS/MS analysis.^{16, 33-37} Although these techniques have proven reliable and robust, the protein and peptide concentration dynamic range of enriched samples often exceeds the analytical capabilities of current mass spectrometry platforms. In these cases, chromatographic pre-fractionation of proteins or peptides can significantly improve LC-MS/MS analysis. This is perhaps best exemplified by the traditional approach of gel-based protein separation, followed by in-gel digestion, peptide extraction, and LC-MS/MS analysis.³⁸⁻³⁹ Although these “Gel-LC” methods have been refined,^{18, 40-41} they are typically labor-intensive and low-throughput. These limitations catalyzed efforts to implement orthogonal peptide-based fractionation strategies, beginning with the combination of strong cation exchange (SCX) and reversed phase (RP) chromatography, termed “Multidimensional Protein Identification Technology (MudPIT).⁴²⁻⁴³ The success of the MudPIT approach has led to a plethora of alternative fractionation geometries, with the overall goal of improving separation power and proteome sequence coverage: Offline/online SCX-RP,⁴⁴⁻⁴⁵ offline immobilized pH gradient gels (IPG) coupled with RP-LC-MS/MS,⁴⁶⁻⁴⁷ offline solution-based IEF (Offgel) coupled with RP-LC-MS/MS,⁴⁸⁻⁴⁹ offline high-pH RP coupled with low-pH RP-LC-MS/MS,^{45, 50-52} offline mixed-mode high pH RP-SAX coupled with low pH RP-LC-MS/MS,⁵³ and online HILIC-SCX-RP.⁵⁴ Generally speaking the offline formats can accommodate large sample quantities (>100 µg), but are susceptible to losses resulting from additional sample handling, lyophilization, and non-specific adsorption to tube or other surfaces.^{45, 46} Online systems offer the advantage of efficient peptide recovery and transfer between all separation stages, but are typically limited in total sample capacity as compared to larger-scale offline formats. For example, Dowell and colleagues reported that online SCX-RP detected twice the number of peptides as compared to an equivalent offline SCX-RP format.⁴⁵ Similarly, work by Slebos et al.,⁴⁶ suggested that although the separation power of gel-based IEF fractionation of peptides as a first dimension is largely independent of sample quantity, the observed recovery and subsequent identification by LC-MS/MS varied inversely with total starting material. We recently demonstrated an automated, online multi-dimension fractionation platform that provided for direct comparison of SCX-RP and RP-RP configurations coupled to true nanoflow LC-MS/MS.⁵⁵ Although RP-RP exhibited superior analytical figures of merit as compared to SCX-RP under all conditions tested, we observed that fraction-to-fraction overlap of peptides began to increase beyond an analysis depth of 40, 1st dimension fractions. These results are consistent with fundamental descriptions of peak capacity for multi-dimension separations in general⁵⁶ and previous reports of RP-RP fractionation in particular.^{45, 50-51, 57} Collectively these observations and results ultimately limit the separation peak capacity and total proteome sequence coverage that can be obtained from 2-D platforms.

Here we expand our automated nanoflow RP-RP-LC-MS/MS platform⁵⁵ to include a stage of high pH SAX separation providing an integrated 3-D (RP-SAX-RP) fractionation system. Based on analysis of peptides derived from an *escherichia coli* standard we found that RP-

SAX-RP provided for significantly improved sequence identification at a fractionation depth beyond which RP-RP previously exhibited diminished peak capacity. Using biochemical data from a recent, large-scale protein expression study in *saccharomyces cerevisiae* as a benchmark, we found that RP-SAX-RP provided for MS/MS-based identification of peptides and proteins across a wide abundance range on both low resolution 3D ion trap and high performance Orbitrap instruments. Interestingly our data suggest an approximate 100-fold performance difference between previous generation and state-of-the-art mass spectrometers. Moreover, our results indicate that efficient, online fractionation strategies provide large-scale proteome sequence coverage from only a few micrograms of tryptic peptides. More generally, our data provide compelling evidence that improved capillary-based separation systems can augment the rapid pace of improvements in state-of-the-art mass spectrometry platforms.

EXPERIMENTAL SECTION

Due to space considerations experimental methods related to cell culture, sample preparation, mass spectrometry and data analysis are provided in Supplementary Materials.

Data Availability

Lists of proteins and peptides associated with Figures 2, 3, S1, and S2 are provided in Supplementary File S1. In addition, all native mass spectrometry files and mzResults files which include annotated MS/MS spectra are available for download from the Proteocommons website based on the following hash: hash link will be provided prior to publication.

3D RP-SAX-RP system

The 3D RP-SAX-RP platform (Figure 1) consisted of Waters UHPLC binary and isocratic pumps, an autosampler (Waters Corp., Milford, MA) and additional 6-port, 2-position valve (Valco Inc., Austin, TX). The 1st dimension reversed phase (RP) column consisted of a 250 μ m I.D. capillary packed with 15 cm of 5 μ m diameter XBridge C18 resin (Waters Corp., Milford, MA). An anion exchange (SAX) column (250 μ m I.D. $\times$ 15 cm of 10 μ m dia. POROS10HQ [AB Sciex, Foster City, CA] resin) was connected directly to the outlet of the 1st dimension RP column with a union. The 3rd dimension was constructed as described previously.⁵⁸⁻⁵⁹ The isocratic pump delivered either peptide samples or 1st (acetonitrile in 20 mM ammonium formate, pH 10) and 2nd (KCl in 20 mM ammonium formate, pH 10) dimension eluents at 2 μ L/min. through the sample loop. Discrete eluent concentrations used for all experiments are provided in Supplementary Table S1. The binary pump delivered 0.1% formic acid at 8 μ L/min. to dilute the organic content and acidify the 1st/2nd dimension effluent prior to the 3rd dimension pre-column, or provided for gradient elution (2-30% B in 45 minutes, A = 0.2 M acetic acid, B = acetonitrile with 0.2 M acetic acid) of peptides from the 3rd dimension reversed phase columns for LC-MS/MS analysis at a flow rate of $\approx$ 10 nL/min.⁶⁰ A Digital PicoView electrospray source platform (New Objective, Woburn, MA; model DPV-550 for the Orbitrap XL) was used to automatically position the emitter tip at the heated metal capillary inlet during LC-MS/MS acquisition or beneath a gravity-driven drip station during injection of peptide samples or 1st/2nd dimension eluents.

SAFETY CONSIDERATIONS

Safety glasses should be worn at all times during construction and use of fused silica based capillaries. In addition, a lab coat and gloves should be worn when handling other volatile organics in a chemical fume hood.

RESULTS AND DISCUSSION

We recently performed a detailed comparison of automated, online 2D RP-RP and SCX-RP fractionation platforms for analysis of peptides derived from complex cell lysates and affinity purified multi-component protein complexes.⁵⁵ Given the high degree of orthogonality that we observed between reversed phase separations performed at high- and low-pH, respectively, we asked whether addition of a third dimension would provide significantly improved fractionation power. Based on the buffer conditions used in the 1st dimension of our RP-RP platform (20 mM ammonium formate, pH 10) we chose strong anion exchange (SAX) as the 2nd dimension in a (now) 3D RP-SAX-RP configuration (Figure 1). Briefly, peptides are loaded through the autosampler (Figure 1, dashed box) in ammonium formate buffer (pH 10) and captured on the 1st dimension RP column. Peptides poorly retained on reversed phase at high pH are trapped on either the 2nd (SAX, pH 10) or 3rd (RP, pH 3.0) dimension columns. The low pH pre- and analytical columns constitute the 3rd dimension separation and are assembled in a vented⁵⁸⁻⁶⁰ configuration to facilitate loading of samples or 1st (acetonitrile, 20 mM ammonium formate, pH 10) and 2nd (KCl in 20 mM ammonium formate, pH 10) dimension eluents, respectively. The system provides complete, online fractionation, meaning that once peptides are injected from the autosampler well onto the 1st dimension column, they are automatically transferred to the 2nd and 3rd dimensions and are only exposed to the fused silica tubing, PEEK LC fittings, and flow paths of the 6-port valves.

Our previous analytical characterization studies⁵⁵ suggested that for LC-MS/MS analysis of *E. coli* peptides in a complex background of whole cell lysate, RP-RP performance appeared to peak at a depth of between 30-40 fractions. Based on this observation we asked whether RP-SAX-RP would provide for improved performance at a fractionation depth beyond which RP-RP previously exhibited diminishing returns. Towards this end we performed replicate RP-RP (40 fraction) and RP-SAX-RP (37 fraction) analyses of *E. coli* tryptic peptides. Figure 2A shows that while each fractionation technique provided reproducible data, the RP-SAX-RP platform identified 44% and 31% more peptides and proteins, respectively, as compared to RP-RP. As a surrogate analysis for separation peak capacity, we next asked whether the number of fractions across which discrete peptides were identified varied between the two separation platforms. Indeed, we observed that 75% of all peptide identifications spanned a single RP-SAX-RP fraction, while in the case of RP-RP (Figure 2B) this figure of merit was reduced to 60%. Consistent with these data we observed that 34 of 37 RP-SAX-RP fractions contained between 100-300 unique peptide identifications (Figure 2C), suggesting that this 3-D platform provided for uniform fractionation of *E. coli* peptides across the entire separation space.

Given the results above, we next asked whether RP-SAX-RP could provide sufficient fractionation depth to identify proteins expressed across a wide concentration dynamic range in the context of a shotgun, data-dependent type acquisition. In order to facilitate comparison of our results with other large-scale proteomic efforts, we switched to yeast as our model system. *S. cerevisiae* has a well-annotated genome and has been extensively studied at the whole proteome level using biochemical⁶¹ as well as shotgun^{18,43-44} and targeted⁶² mass spectrometry-based approaches. With these studies as a reference point, we next sought to establish an extreme limit for automated, online RP-SAX-RP fractionation. Towards this end we loaded 40 µg of peptides derived from yeast whole cell lysate and acquired 236 LC-MS/MS fractions on a LCQ Deca XP ion trap mass spectrometer (Supplementary Figure S1). This experiment required nearly 18 days of uninterrupted acquisition time and yielded 18,359 unique peptides (1% FDR) that mapped to 2902 proteins (see Supplemental Methods for details on protein inference). Although these data are based on somewhat dated mass spectrometry technology, it is nonetheless interesting to

note that our results compare very favorably to a previous study that utilized 80 fractions of offline SCX fractionation coupled to LC-MS/MS on the same mass spectrometry platform (1504 proteins identified from 1 mg protein lysate⁴⁴). This comparison suggests that continued improvements in fractionation can compensate for limitations in the stochastic nature of data-dependent MS/MS acquisition. Consistent with this hypothesis, we observed a surprisingly uniform distribution of unique peptide identifications across all 236 RP-SAX-RP fractions.

Several proteins identified in this analysis have been previously characterized with expression levels at or below 128 copies per cell (Table 1), including YOR093C (41copies/cell), YBL063W (49 copies/cell), and YJL084C (57 copies/cell).⁶¹ Annotated MS/MS spectra for representative peptides from these proteins are provided on our mzServer⁶³ via embedded links in Table 1. For comparison, Table 1 also lists low abundance proteins ($\leq$ 128 copies per cell) recently characterized by triple quadrupole-based MRM assays⁶². Interestingly we identified 8 out of these 15 proteins in our 236 fraction RP-SAX-RP analysis, suggesting that efficient fractionation can offset, in-part, the discrepancies in detection between discovery- and targeted-mode MS/MS analysis. Based on these data we estimate a detection limit of 500 amol ($\approx$ 50 protein copies per cell, 7E6 cells analyzed) and a dynamic range of 3.2E4 (YKL096W-a 1.59E6 copies/cell, 7E6 cells analyzed) for data-dependent analysis of total tryptic peptides on a 3-D ion trap instrument.

As an extension of the studies above, we next coupled our RP-SAX-RP separation system to a hybrid linear ion trap orbitrap (Orbitrap XL) mass spectrometer. As we expected significantly improved performance as compared to the 3-D ion trap instrument, we first reduced the sample quantity by nearly an order of magnitude (5 μ g versus 40 μ g), and then sought to establish the number of RP-SAX-RP fractions required to obtain yeast proteome coverage roughly equivalent to that observed with the LCQ platform (236 fractions, 18,359 unique peptides, 2902 proteins). We found that 19 fractions yielded 2809 protein identifications based on 8929 unique peptide sequences (Figure 3A, blue). Acquisition of these data required approximately 2 days of LC-MS/MS time. We next performed a second 19 fraction RP-SAX-RP analysis (Figure 3A, yellow) to establish a baseline for reproducibility of our separation platform. Here we identified 2856 proteins from 9021 peptides. Across both analyses we identified 2422 proteins in common ($\approx$ 75% reproducibility) and 3243 proteins in total. We also observed good agreement between signals of individual peptides identified in both runs (Figure 3B), with $\approx$ 80% of all intensity ratios within $\pm$ w2-fold of the mean value. We next tested whether the peptide content of individual fractions reproduced across these replicate runs by first requiring that peptide sequences were identified in the same fraction across both analyses; under these conditions, we observed that $\approx$ 48% of all peptides were reproducibly identified across discrete fractions in these two RP-SAX-RP analyses. However, the stochastic nature of MS/MS likely makes this result the lower bound of “reproducibility.” Hence we next used a more permissive approach which included as “reproduced” those MS-level precursor peaks that corresponded in mass ($\pm$ 10 ppm) and (third dimension) retention time ($\pm$ 0.5 minute) to peptide sequences identified in only one of the two analyses. Based on these criteria we found that potentially 98% of all identified peptides were reproducibly detected.

We next asked whether additional fractionation depth would provide higher proteome sequence coverage. In independent experiments we loaded 5 μ g of yeast tryptic peptides and acquired MS/MS data from 19, 51, and 101 RP-SAX-RP fractions, respectively. Encouragingly, at each fractionation depth we obtained data that constituted a significant superset of the previous experiment, with overlaps of at least 70% and 92% at the peptide (Supplementary Figure S2A) and protein (Supplementary Figure 2SB) level, respectively. Moreover we observed that at least 62% of unique peptide identifications were confined to a

single RP-SAX-RP fraction, even for the most in-depth analysis (101 fractions, Supplementary Figure S2C). As with the experiments performed on the low resolution instrument (Supplementary Figure S1) we readily detected low abundance proteins (≈ 50 copies per cell) at each fractionation depth, albeit at total acquisition times (e.g., number of fractions) far less than were required on the 3D ion trap (see embedded links in Table 1 for MS/MS spectra of representative peptides). Based on the sample quantity consumed per analysis (5 μg) we estimate a detection limit of 65 amol and a dynamic range of $3.2\text{E}4$ (YKL096W-a $1.59\text{E}6$ copies/cell, $8\text{E}5$ cells analyzed) for data-dependent analysis of total yeast tryptic peptides on a LTQ-Orbitrap XL. Collectively these results suggest that our 3D RP-SAX-RP platform provides for reproducible separation and sufficient peak capacity to enable identification of low abundance proteins in complex mixtures on low- and high-performance mass spectrometers.

For the sample quantity analyzed on the LTQ-Orbitrap (5 μg) it appeared that our 3D separation platform reached a point of diminishing returns between 50-100 fractions, as we only identified an additional 184 unique proteins (1377 unique peptides) by extending the analysis to 101 fractions. To determine whether additional fractionation primarily impacted protein sequence coverage (e.g., more peptides identified per protein) or protein dynamic range (e.g., identification of proteins at low expression levels), we next plotted our fractionation data from Supplementary Figures S1 and S2 along with that from a recent large-scale biochemical analysis⁶¹ of protein expression in yeast (Figure 4). We also included the combined set of unique protein identifications resulting from our 19, 51, and 101 fraction experiments. For each expression level bin, we used the quantitative western blot data as a reference and then plotted the subset of these proteins identified in our RP-SAX-RP analyses. Between a range of ≈ 60 and $\approx 4,000$ protein copies per cell, we observed a consistent trend in which the number of protein identifications increased systematically as a function of mass spectrometry technology (LCQ vs LTQ-Orbitrap) and fractionation depth (19, 51, and 101 fractions, or the combination thereof). However at the lowest expression level (Figure 4B, left, ≈ 30 -60 copies per cell) the mass spectrometry data appeared somewhat stochastic, suggesting that either this range represented the practical detection limit under the conditions tested, and hence fractionation provided little benefit, or the number of proteins ($n=9$) quantified by western blot was too small to serve as a reference. To explore this question further we next plotted two other expression categories from Ghaemmaghami et al.,⁶¹ “extremely low signal (<50 copies per cell)” and “no expression detected,” along with the corresponding proteins detected in our RP-SAX-RP analyses (Figure 4C). With larger sets of reference proteins (234 and 1982 for each category, respectively) we again observed a systematic increase in the number of identified proteins as a function of mass spectrometry technology and fractionation depth. Interestingly, we observed that the number of protein identifications more than doubled (from 62 to 134) for the “extremely low signal (<50 copies per cell)” category moving from the LCQ (236 fractions) to the LTQ-Orbitrap (101 fractions).

Finally in order to identify overall bias in protein identification as a function of expression level, we re-plotted the data in Figure 4A as a relative percentage of all proteins identified by each approach (Figure 5). Not surprisingly, both the LCQ (236 fractions) and Orbitrap (19 fractions) data exhibited a bias towards highly expressed proteins. However the combined data set (Orbitrap 19, 51, and 101 fractions) correlated very well with the western blot data across the majority of expression bins. Collectively our results suggest that increased fractionation provided for improved proteome coverage across the entire range of protein expression in yeast.

While cross-lab comparisons based on protein lists are admittedly wrought with caveats, it is nonetheless informative to compare our results with other analytically-focused studies

designed to maximize the discovery capacity of LC-MS/MS. For example, one recent report identified 3313 unique proteins in the context of a triplicate, 12 fraction off-line SCX-RP experiments from 2 mg yeast lysate.¹⁹ Similarly a combination of chromatographic and SDS-PAGE protein fractionation, followed by LC-MS/MS (off-line 3D), generated 75 fractions in total and identified 3639 proteins cumulatively across triplicate analyses from 1.5mg cell lysate.¹⁸ Finally, a triplicate, 24 fraction experiment based on offline isoelectric focusing peptide fractionation (OFFGEL) followed by LC-MS/MS (off-line IEF-RP) identified 3987 proteins from 0.3mg cell lysate.¹⁸ When compared to our data, in which we detected over 4000 unique yeast proteins from only 10% of equivalent input material, these results strongly suggest that high performance fractionation can significantly augment the performance capabilities of state-of-the-art mass spectrometry instrumentation. This observation is further corroborated by the data in Table 1, which compares the ability of our RP-SAX-RP platform to detect a set of low abundance proteins (<128 copies per cells) that were recently quantified by targeted MRM-type mass spectrometry assays. Cumulatively across our 3 analyses on the Orbitrap (19, 51, and 101 fractions) we detected 12 out of the 15 proteins, from approximately 2% of the input material as compared to the work of Picotti, et al.⁶²

Although the above comparisons provide useful insight, a recent, multi-lab report⁶⁴ illustrated the utility of considering numerous metrics when evaluating discovery-based proteomic methods. For example, it is informative to consider the overall analysis efficiency or rate of data production along with the total number of peptide and protein identifications. Towards this end, Table 2 provides detection efficiencies for the 2D/3D experiments on the Orbitrap (1, 19, 51, and 101 fractions) and the LCQ (37, 40, and 236 fractions) instruments, respectively. These data demonstrate the inevitable compromise between experiment time and the depth of proteome coverage, but nonetheless provide useful benchmarks for evaluation of overall system performance. It is also worth noting that while the capillary format and online geometry of our nanoflow RP-SAX-RP fractionation platform provides for fully automated LC-MS/MS analyses, there is a practical trade-off in terms of limited sample loading capacity. While it is true that offline configurations provide for analysis of larger analyte quantities, recent reports have noted that sample handling, lyophilization, and non-specific adsorption of peptides to tube surfaces can often reduce sample recovery in these schemes. In addition, these deleterious effects tend to scale inversely with total input, meaning that the performance for offline systems can be limited for sample quantities below ≈ 50 -100 μg .⁴⁵⁻⁴⁶ In contrast, we recently demonstrated that electrospray ionization efficiency compensates for chromatographic peak broadening at effluent flow rates below empirical Van Deemter minima,⁶⁰ and suggested that a renewed focus on small diameter ($\leq 25 \mu\text{m}$ I.D.) LC-MS assemblies operated at low nL/min flowrates would provide significantly improved analytical figures of merit and compliment the continued performance improvements available on state-of-the-art mass spectrometry platforms. Here we successfully coupled our nanoflow 1D columns with capillary-format high-pH RP and SAX to provide an automated, 3-D separation system for analysis of complex proteomes. The capacity of our RP-SAX-RP platform to provide improved detection and dynamic range is illustrated in Figure 4 in which we observed a systematic improvement in detection of low abundance proteins as a function of mass spectrometry technology and fractionation depth.

Consistent with our experience using 25 μm I.D. analytical columns with integrated electrospray emitters in a 1-D nanoflow configuration⁶⁰ or coupled with high-pH RP in an automated RP-RP configuration,⁵⁵ our RP-SAX-RP platform has proven to be very robust, providing uninterrupted operation for more than 100 injections spanning several months of analysis time without a major system failure; in fact the majority of data described herein were acquired using one column set over a span of ≈ 3 months.

CONCLUSIONS

Despite the growing use of mass spectrometry in biomedical research, limited dynamic range continues to hinder the characterization of low abundance proteins present in complex biological matrices. Multidimensional peptide fractionation⁴²⁻⁴³ improves data quality and facilitates analysis of proteomic results in the context of hypotheses of biological function, although achieving an optimal combination of analytical figures of merit remains a significant challenge. Herein we build upon our recent work in high performance 1-D⁶⁰ and 2-D⁵⁵ separations and describe an automated platform for 3-D RP-SAX-RP fractionation of tryptic peptides. Our observations are thematically aligned with other recent efforts⁶⁵⁻⁷¹ and provide compelling evidence that capillary-based fractionation formats offer a powerful combination of automation, reproducibility, separation peak capacity, and sample yield.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Eric Smith for assistance with the figures. Generous support for this work was provided by the Dana-Farber Cancer Institute and the National Institutes of Health, NHGRI (P50HG004233), and NINDS (P01NS047572).

REFERENCES

- (1). Shevchenko A, Wilm M, Vorm O, Mann M. *Analytical Chemistry*. 1996; 68:850–858. [PubMed: 8779443]
- (2). Gingras AC, Gstaiger M, Raught B, Aebersold R. *Nat. Rev. Mol. Cell Biol.* 2007; 8:645–654. [PubMed: 17593931]
- (3). Choudhary C, Mann M. *Nat. Rev. Mol. Cell Biol.* 2010; 11:427–439. [PubMed: 20461098]
- (4). Olsen JV, Blagoev B, Gnäd F, Macek B, Kumar C, Mortensen P, Mann M. *Cell*. 2006; 127:635–648. [PubMed: 17081983]
- (5). Patwa T, Li C, Simeone DM, Lubman DM. *Mass Spectrom. Rev.* 2010; 29:830–844. [PubMed: 20077480]
- (6). Kim YJ, Freas A, Fenselau C. *Analytical Chemistry*. 2001; 73:1544–1548. [PubMed: 11321307]
- (7). Madera M, Mechref Y, Klouckova I, Novotny MV. *J. Proteome Res.* 2006; 5:2348–2363. [PubMed: 16944947]
- (8). Yang ZP, Harris LE, Palmer-Toy DE, Hancock WS. *Clin. Chem.* 2006; 52:1897–1905. [PubMed: 16916992]
- (9). Yu YH, Hoffhines AJ, Moore KL, Leary JA. *Nat. Methods*. 2007; 4:583–588. [PubMed: 17558413]
- (10). Miller MJ, Barrett-Wilt GA, Hua ZH, Vierstra RD. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107:16512–16517. [PubMed: 20813957]
- (11). Blomster HA, Imanishi SY, Siimes J, Kastu J, Morrice NA, Eriksson JE, Sistonen L. *J. Biol. Chem.* 2010; 285:19324–19329. [PubMed: 20388717]
- (12). Xu P, Peng JM. *BBA-Proteins Proteomics*. 2006; 1764:1940–1947.
- (13). Peng JM, Schwartz D, Elias JE, Thoreen CC, Cheng DM, Marsischky G, Roelofs J, Finley D, Gygi SP. *Nat. Biotechnol.* 2003; 21:921–926. [PubMed: 12872131]
- (14). Schilling O, Barre O, Huesgen PF, Overall CM. *Nat. Methods*. 2010; 7:508–U33. [PubMed: 20526347]
- (15). Kleifeld O, Doucet A, Keller UAD, Prudova A, Schilling O, Kainthan RK, Starr AE, Foster LJ, Kizhakkedathu JN, Overall CM. *Nat. Biotechnol.* 2010; 28:281–U144. [PubMed: 20208520]
- (16). Wiederhold E, Veenhoff LM, Poolman B, Slotboom DJ. *Mol. Cell. Proteomics*. 2010; 9:431–445. [PubMed: 19955081]

- (17). Andersen JS, Lam YW, Leung AKL, Ong SE, Lyon CE, Lamond AI, Mann M. *Nature*. 2005; 433:77–83. [PubMed: 15635413]
- (18). de Godoy LMF, Olsen JV, Cox J, Nielsen ML, Hubner NC, Frohlich F, Walther TC, Mann M. *Nature*. 2008; 455:1251–U60. [PubMed: 18820680]
- (19). Swaney DL, Wenger CD, Coon JJ. *J. Proteome Res.* 2010; 9:1323–1329. [PubMed: 20113005]
- (20). Nevedomskaya E, Derks R, Deelder AM, Mayboroda OA, Palmblad M. *Anal. Bioanal. Chem.* 2009; 395:2527–2533. [PubMed: 19826795]
- (21). Wallace A, Millar A, Langridge J. *Nat. Methods*. 2007:AN12–AN13.
- (22). Shvartsburg AA, Smith RD. *Analytical Chemistry*. 2008; 80:9689–9699. [PubMed: 18986171]
- (23). Giannakopoulos AE, Thomas B, Colburn AW, Reynolds DJ, Raptakis EN, Makarov AA, Derrick PJ. *Rev. Sci. Instrum.* 2002; 73:2115–2123.
- (24). Olsen JV, Nielsen ML, Damoc NE, Griep-Raming J, Moehring T, Makarov A, Schwartz J, Horning S, Mann M. *Mol. Cell. Proteomics*. 2009:2759–2769. [PubMed: 19828875]
- (25). Second TP, Blethrow JD, Schwartz JC, Merrihew GE, MacCoss MJ, Swaney DL, Russell JD, Coon JJ, Zabrouskov V. *Analytical Chemistry*. 2009; 81:7757–7765. [PubMed: 19689114]
- (26). Moehring T, Strupat K, Delanghe B, Muenster H. *Mol. Cell. Proteomics*. 2005; 4:S308–S308.
- (27). Sterner JL, Johnston MV, Nicol GR, Ridge DP. *J. Mass Spectrom.* 2000; 35:385–391. [PubMed: 10767768]
- (28). Nilsson T, Mann M, Aebersold R, Yates JR, Bairoch A, Bergeron JJM. *Nat. Methods*. 2010; 7:681–685. [PubMed: 20805795]
- (29). Gouw JW, Tops BBJ, Mortensen P, Heck AJR, Krijgsveld J. *Analytical Chemistry*. 2008; 80:7796–7803. [PubMed: 18808151]
- (30). Kellmann M, Muenster H, Zomer P, Mol H. *J. Am. Soc. Mass Spectrom.* 2009; 20:1464–1476.
- (31). Delmotte N, Lasaosa M, Tholey A, Heinzel E, van Dorsselaer A, Huber CG. *J. Sep. Sci.* 2009; 32:1156–1164. [PubMed: 19360783]
- (32). Wang H, Chang-Wong T, Tang HY, Speicher DW. *J. Proteome Res.* 2010; 9:1032–1040. [PubMed: 20014860]
- (33). Ficarro SB, McClelland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM. *Nat. Biotechnol.* 2002; 20:301–305. [PubMed: 11875433]
- (34). Zhang Y, Wolf-Yadlin A, Ross PL, Pappin DJ, Rush J, Lauffenburger DA, White FM. *Mol Cell Proteomics*. 2005; 4:1240–50. [PubMed: 15951569]
- (35). Andersen JS, Mann M. *EMBO Rep.* 2006; 7:874–879. [PubMed: 16953200]
- (36). Burckstummer T, Bennett KL, Preradovic A, Schutze G, Hantschel O, Superti-Furga G, Bauch A. *Nat. Methods*. 2006; 3:1013–1019. [PubMed: 17060908]
- (37). Ficarro SB, Adelmant G, Tomar MN, Zhang Y, Cheng VJ, Marto JA. *Anal Chem.* 2009; 81:4566–75. [PubMed: 19408940]
- (38). Wilm M, Mann M. *Analytical Chemistry*. 1996; 68:1–8. [PubMed: 8779426]
- (39). Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T, Mann M. *Nature*. 1996; 379:466–469. [PubMed: 8559255]
- (40). Piersma SR, Fiedler U, Span S, Lingnau A, Pham TV, Hoffmann S, Kubbutat MHG, Jimenez CR. *J. Proteome Res.* 2010; 9:1913–1922. [PubMed: 20085282]
- (41). Fang Y, Robinson DP, Foster LJ. *J. Proteome Res.* 2010; 9:1902–1912. [PubMed: 20078137]
- (42). Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR. *Nat. Biotechnol.* 1999; 17:676–682. [PubMed: 10404161]
- (43). Washburn MP, Wolters D, Yates JR. *Nat. Biotechnol.* 2001; 19:242–247. [PubMed: 11231557]
- (44). Peng JM, Elias JE, Thoreen CC, Licklider LJ, Gygi SP. *J. Proteome Res.* 2003; 2:43–50. [PubMed: 12643542]
- (45). Dowell JA, Frost DC, Zhang J, Li LJ. *Analytical Chemistry*. 2008; 80:6715–6723. [PubMed: 18680313]
- (46). Slebos RJC, Brock JWC, Winters NF, Stuart SR, Martinez MA, Li M, Chambers MC, Zimmerman LJ, Ham AJ, Tabb DL, Liebler DC. *J. Proteome Res.* 2008; 7:5286–5294. [PubMed: 18939861]

- (47). Cargile BJ, Talley DL, Stephenson JL. Electrophoresis. 2004; 25:936–945. [PubMed: 15004858]
- (48). Balgley BM, Wang WJ, Song T, Fang XP, Yang L, Lee CS. Electrophoresis. 2008; 29:3047–3054. [PubMed: 18655040]
- (49). Hubner NC, Ren S, Mann M. Proteomics. 2008; 8:4862–4872. [PubMed: 19003865]
- (50). Gilar M, Olivova P, Daly AE, Gebler JC. J. Sep. Sci. 2005; 28:1694–1703. [PubMed: 16224963]
- (51). Song CX, Ye ML, Han GH, Jiang XN, Wang FJ, Yu ZY, Chen R, Zou HF. Analytical Chemistry. 2010; 82:53–56. [PubMed: 19950968]
- (52). Manadas B, English JA, Wynne KJ, Cotter DR, Dunn MJ. Proteomics. 2009; 9:5194–5198. [PubMed: 19771557]
- (53). Phillips HL, Williamson JC, van Elburg KA, Snijders APL, Wright PC, Dickman M. J. Proteomics. 2010; 10:2950–2960.
- (54). Wu CC, MacCoss MJ, Howell KE, Yates JR. Nat. Biotechnol. 2003; 21:532–538. [PubMed: 12692561]
- (55). Zhou F, Cardoza JD, Ficarro SB, Adelmant GO, Lazaro JB, Marto JA. J. Proteome Res. 2010; 9:6242–6255. [PubMed: 20873769]
- (56). Giddings JC. Analytical Chemistry. 1984; 56:1258.
- (57). Gilar M, Olivova P, Daly AE, Gebler JC. Analytical Chemistry. 2005; 77:6426–6434. [PubMed: 16194109]
- (58). van der Heeft E, ten Hove GJ, Herberts CA, Meiring HD, van Els C, de Jong A. Analytical Chemistry. 1998; 70:3742–3751. [PubMed: 9751018]
- (59). Licklider LJ, Thoreen CC, Peng JM, Gygi SP. Analytical Chemistry. 2002; 74:3076–3083. [PubMed: 12141667]
- (60). Ficarro SB, Zhang Y, Lu Y, Moghimi AR, Askenazi M, Hyatt E, Smith ED, Boyer L, Schlaeger TM, Luckey CJ, Marto JA. Analytical Chemistry. 2009; 81:3440–3447. [PubMed: 19331382]
- (61). Ghaemmghami S, Huh W, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. Nature. 2003; 425:737–741. [PubMed: 14562106]
- (62). Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R. Cell. 2009; 138:795–806. [PubMed: 19664813]
- (63). Askenazi M, Webber JT, Marto JA. Molecular & cellular proteomics: MCP. 2011; 10:7.
- (64). Rudnick PA, Clauser KR, Kilpatrick LE, Tchekhovskoi DV, Neta P, Blonder N, Billheimer DD, Blackman RK, et al. Molecular & cellular proteomics: MCP. 2010; 9:225–41.
- (65). Thakur SS, Geiger T, Chatterjee B, Bandilla P, Froehlich F, Cox J, Mann M. Mol. Cell. Proteomics. 2011 In Press.
- (66). Shen YF, Zhang R, Moore RJ, Kim J, Metz TO, Hixson KK, Zhao R, Livesay EA, Udseth HR, Smith RD. Analytical Chemistry. 2005; 77:3090–3100. [PubMed: 15889897]
- (67). Di Palma S, Stange D, De Wetering MV, Clevers H, Heck AJR, Mohammed S. J. Proteome Res. 2011 In press.
- (68). Di Palma S, Boersema PJ, Heck AJR, Mohammed S. Analytical Chemistry. 2011; 83:3440–3447. [PubMed: 21443167]
- (69). Shen YF, Jacobs JM, Camp DG, Fang RH, Moore RJ, Smith RD, Xiao WZ, Davis RW, Tompkins RG. Analytical Chemistry. 2004; 76:1134–1144. [PubMed: 14961748]
- (70). Kong RPW, Siu SO, Lee SSM, Lo C, Chu IK. J. Chromatogr. A. 2011; 1218:3681–3688. [PubMed: 21531424]
- (71). Luo Q, Gu Y, Wu SL, Rejtar T, Karger BL. Electrophoresis. 2008; 29:1604–11. [PubMed: 18383016]

Figure 1.

Schematic diagram of automated, online nanoflow RP-SAX-RP platform. **(A)** The autosampler first loads sample and then is used to inject first- (acetonitrile, orange) or second- (KCl, blue) dimension eluents, respectively. **(B, C)** An additional 6-port, 2-position valve provides a vented third dimension column configuration and ensures efficient capture of peptides on the low-pH RP precolumn. **(B)** Second dimension fractions are diluted (4:1) and acidified with reversed phase solvent A (0.1% formic acid, 3% acetonitrile) introduced by an ultra-high pressure binary pump. **(C)** An organic gradient is delivered by the binary pump to elute peptides from the third dimension column for MS/MS analysis. The vent valve generates a pre-column effluent split of $\approx 1000:1$ to provide a stable column/ESI flow rate of ≈ 5 -10 nL/min. Active solvent flow paths are highlighted in green. **(D)** A computer-controlled positioning platform (Digital PicoView) automatically moves the emitter tip between “electrospray” and “wash” positions during data acquisition and sample loading, respectively.

Figure 2.

(A) Replicate analyses of tryptic peptides derived from *e.coli* lysate by RP-RP (left, red) and RP-SAX-RP (right, blue) fractionation on a LCQ Deca mass spectrometer. RP-SAX-RP provided for higher numbers of identified peptides and proteins as compared to RP-RP for independent experiments spanning 37 and 40 fractions, respectively. (B) Peptide elution profiles, defined as the number of first (high-pH RP) or first and second (high-pH RP and SAX), fractions spanned by peptides detected in (A). (C) Distribution of unique peptide identifications in a 37 fraction RP-SAX-RP experiment.

Figure 3.

Technical replicate 19 fraction RP-SAX-RP analyses of *s.cerevisiae* tryptic peptides on a LTQ-Orbitrap mass spectrometer. **(A)** Venn diagrams illustrate the reproducibility of peptide ($\approx 55\%$) and protein ($\approx 75\%$) identifications. **(B)** Histogram distribution of Log_2 ratios for peptide precursors matched by unique amino acid sequence or a combination of intact peptide mass (± 10 ppm) and third dimension LC elution time (± 0.5 min).

Figure 4.

(**A, B**) Proteome coverage for *S. cerevisiae* as a function fractionation depth (19, 51, 101, 236 fractions) and mass spectrometry platform (LCQ Deca or LTQ-Orbitrap XL) as compared to biochemical-based quantification of protein expression.⁶¹ Western blot data (orange) are used as a reference set within each expression level bin. The count of reference proteins detected in each RP-SAX-RP experiment is plotted with the color scheme as indicated in the legend. The x-axis is labeled based on absolute expression level (copies per cell) as determined from TAP-GFP data⁶¹ or relative protein quantity calculated from absolute expression level⁶¹ and total lysate analyzed by each RP-SAX-RP experiment (LCQ – 40 µg, LTQ-Orbitrap – 5 µg). (**C**) Expression histograms plotted as in (**A, B**) for protein groups from Ghaemmaghani, et al, “Extremely low signal (< 50 copies/cell)” and “No expression detected,”⁶¹ not amenable to absolute quantification by western blot.

Figure 5.

Expression histogram as in Figure 4 (A), plotted as a percentage of the total protein count detected in each experiment. TAP-GFP data (orange) serves as the reference protein set in each expression bin. The x-axis is labeled based on absolute expression level (copies per cell as determined from TAP-GFP data⁶¹) or relative protein quantity calculated from absolute expression level and total lysate analyzed by each RP-SAX-RP experiment (LCQ – 40 µg, LTQ-Orbitrap – 5 µg).

Table 1

Detection of low expression level proteins. Each "Y" entry provides an embedded link to an annotated MS/MS spectrum via mzServer ⁶³.

Low copy number proteins (<128 copies/cell)	Detected in 19 fraction RP-SAX-RP (Orbitrap)	Detected in 51 fraction RP-SAX-RP (Orbitrap)	Detected in 101 fraction RP-SAX-RP (Orbitrap)	Detected in 236 fraction RP-SAX-RP (LCQ)
YLL040C	Y	Y	Y	Y
YNL014W	N	Y	Y	Y
YML109W	N	Y	Y	Y
YIL092W	N	Y	Y	N
YIL084C	Y	Y	Y	N
YKL145W	Y	Y	Y	Y
YKL075C	N	Y	Y	N
YIL002C	Y	Y	N	N
YHR015W	N	N	N	N
YPL008W	N	N	Y	N
YGL006W	N	Y	Y	Y
YKR031C	Y	Y	N	Y
YLR035C	N	N	N	N
YNR067C	Y	Y	Y	Y
YOR093C	Y	Y	Y	Y
<i>Success rate</i>	<i>7/15</i>	<i>12/15</i>	<i>11/15</i>	<i>8/15</i>

Table 2

Performance metrics for 2D (RP-RP) and 3D (RP-SAX-RP) LC-MS/MS platforms.

Organism	Platform	Instrument	# Fractions	Time (hr)	µg	Unique Peptides	Unique Proteins	Peptides /min	Proteins /min
<i>S. cerevisiae</i>	RP-SAX-RP	Orbitrap	1 fraction	2	5	2125	981	17.71	8.18
<i>S. cerevisiae</i>	RP-SAX-RP	Orbitrap	19 fraction	38	5	14576	3111	6.39	1.36
<i>S. cerevisiae</i>	RP-SAX-RP	Orbitrap	51 fraction	102	5	25091	3821	4.10	0.62
<i>S. cerevisiae</i>	RP-SAX-RP	Orbitrap	101 fraction	202	5	26468	4004	2.18	0.33
<i>S. cerevisiae</i>	RP-SAX-RP	LCQ	236 fraction	572	40	18359	2902	0.45	0.08
<i>E.coli.</i>	RP-SAX-RP	LCQ	37 fraction	74	20	4254	923	0.96	0.21
<i>E.coli.</i>	RP-RP	LCQ	40 fraction	80	20	2923	702	0.61	0.15