Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography B, 847 (2007) 54–61
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Improved proteomic discovery by sample pre-fractionation using dual-column ion-exchange high performance liquid chromatography ☆

Pierre C. Havugimana, Peter Wong, Andrew Emili *

Banting and Best Department of Medical Research, Department of Medical Genetics and Microbiology, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario, Canada M5S 3E1

Received 5 July 2006; accepted 23 October 2006
Available online 30 November 2006

Abstract

Clinically relevant biomarkers are urgently needed for improving patient diagnosis, risk stratification, prognosis and therapeutic treatments. There is a particularly compelling motivation for identifying protein-based indicators of early-stage disease for more effective interventions. Despite recent progress, the proteomic discovery process remains a daunting challenge due to the sheer heterogeneity and skewed protein abundances in biofluids. Even the most advanced mass spectrometry systems exhibit limiting overall dynamic ranges and sensitivities relative to the needs of modern biomedical applications. To this end, we report the development of a robust, rapid, and reproducible high performance ion-exchange liquid chromatography pre-fractionation method that allows for improved proteomic detection coverage of complex biological specimens using basic tandem mass spectrometry screening procedures. This form of sample simplification prior to global proteomic profiling, which we refer to collectively as 'fractionomics', increases the number and diversity of proteins that can be confidently identified in tissue and cell lysates as compared to the straight analysis of unfractionated crude extracts.

Keywords: Proteomics; Biomarker; Sample preparation; Pre-fractionation; HPLC; Ion exchange chromatography; LC–MS; Shotgun sequencing; Tandem mass spectrometry

1. Introduction

Proteomics is commonly viewed as the comprehensive study of the expression patterns, molecular interactions and functional states of the proteins present in a given cell, organ or intact organism of interest. The advent of global proteomic profiling strategies, combining the availability of genomic sequences with powerful new high-throughput screening technologies, increasingly makes it possible to define comprehensively the identities, relative abundance levels and modification states of proteins across various tissues and cellular fractions in a systematic genome-scale manner (reviewed by [1]). Elucidation of large-scale perturbations to such patterns as a result of pathological processes using comparative proteomic approaches is expected to provide insight into the fundamental basis of pathologies. It also offers the potential to reveal specific clinically-relevant protein biomarkers of disease states that have substantive diagnostic, prognostic and/or therapeutic value. The systematic monitoring of pathophysiological mal-adaptations at the proteomic level using validated animal-based models of disease, in particular, offers a particularly tractable experimental platform for comprehensive biomarker discovery. Given the tremendous complexity of clinical specimens like serum, urine, or biopsies, however, the critical challenge still remains the ability to identify and quantify the myriad of proteins present with high accuracy, sensitivity and dynamic range. This represents an ongoing, but still largely elusive, goal for the proteomic research community.

While gel electrophoresis represents a useful sample simplification technique, it suffers from severe dynamic range limitations and systematic bias. Conversely, gel-free methods of high-throughput liquid chromatography coupled to tandem mass spectrometry (LC–MS) have emerged as a more promising analytical platform for a more in-depth qualitative and quantitative proteomic analyses of complex biological systems (reviewed in [2–4]). Recent developments in automated shotgun sequencing

☆ This paper was presented at Biomarker Discovery by Mass Spectrometry, Amsterdam, The Netherlands, 18–19 May 2006.

* Corresponding author at: DCCBR, 160 College Street, Rm 914, Toronto, Ontario, Canada M5S 3E1. Tel.: +1 416 946 7281; fax: +1 416 978 7437.
E-mail address: andrew.emili@utoronto.ca (A. Emili).

1570-0232/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2006.10.075

now allow for the routine identification and quantitative monitoring of hundreds to thousands of proteins in a single analysis [5,6]. In particular, the ubiquitous MudPIT method (acronym for Multidimensional Protein Identification Technology), pioneered by Yates and associates [7,8], serves as an effective experimental approach for the exhaustive bottom-up characterization of the peptide products of proteolytic digests of complex biological mixtures. Yet, even today, global proteomic screens still typically result in the preferential identification of only higher abundance proteins, such as metabolic housekeeping enzymes, rather than signalling and regulatory proteins that lie closer to disease causality [9–11].

Detection limits by LC–MS can often be dramatically improved by applying sample simplification in advance of the shotgun analysis. Indeed, several recent notable large-scale proteomic studies [12–16] have combined sub-cellular fractionation as a means of enhancing detection coverage of organelle-specific factors localized to particular cellular compartments, wherein they perform their native biological and molecular functions. Affinity chromatography also serves as a convenient away to selectively isolate target proteins of special interest to high relative purity [17] assuming a biospecific ligand is available. Despite recent notable global scale interaction studies [18], this approach is not generally suitable for the routine study of whole proteomes. Surprisingly, other more traditional, but well trusted methods for biochemical fractionation of complex protein mixtures, such as high performance liquid chromatography (HPLC), gradient centrifugation, capillary electrophoresis and chromatofocusing, have not yet received as much attention to date, despite the fact that they offer the potential for further enhanced protein pre-separations that should lead to a corresponding increase in overall detection coverage.

HPLC offers a conceptually attractive pre-fractionation technique to alleviate protein complexity prior to the analysis by mass spectrometry [19–22]. Size exclusion chromatography, reverse-phase chromatography, affinity chromatography, and anion exchange chromatography represent proven, readily executed protein pre-fractionation platforms that amenable to high-throughput proteomic scale-up [23–26]. The different HPLC methods developed to date can be broadly characterized in terms of their empirical resolution, and whether the method is denaturing or not. Reverse-phase HPLC, a popular and powerful separation method that sorts proteins according to hydrophobicity, is typically used when sample denaturation and subsequent loss in protein activity is not an important consideration [27]. Alternatively, size exclusion chromatography, which separates proteins according to molecular size, is an excellent method for recovering proteins in an intact native conformation, but unfortunately offers relatively poor resolution.

Perhaps the most flexible and widely used method for protein fractionation in biochemical studies is ion exchange chromatography (IEX-HPLC). This method allows for the collection of proteins in relative unperturbed conformations without significant sample loss, while maintaining good resolution suitable for the separations of complex biological samples such as whole cell lysates prior to proteomic characterization (recently reviewed by [28,29]). In IEX-HPLC, the resolution of a mixture of proteins is achieved based on the differential retention of distinct proteins to the charged surfaces of the stationary phase adsorbent, due to differences in the local surface charge properties of proteins in solution [28]. Effective separations hinge on the selective sequential elution of the bound proteins after application of a salt gradient to the mobile phase. The eluent can then be monitored and individual fractions collected in a timed or targeted manner, with the differentially separated components subsequently subject to shotgun sequencing using basic LC–MS procedures. The protein enrichment and sample simplification achieved using this method should enhance the sensitivity of detection by tandem mass spectrometry since there is less chance for ion interference.

A diverse range of suitable bulk materials or pre-packed columns is commercially available for IEX-HPLC, which can potentially resolve even the most complex protein mixtures over a reasonable time frame. Anion exchange columns, which consist of resin bearing positively charged functional groups, can be used to separate negatively charged proteins (which represent the vast majority of soluble protein species typically expressed by a cell), while cation exchange columns resolve positively charged (e.g. histidine, lysine and arginine rich) proteins. In principle, the development of a method that combines the two types of columns in unison should provide an even more effective analytical separation tool prior to proteomic analysis.

Here, we evaluate a simple, rapid, and reproducible dual-column HPLC method for sample pre-fractionation that can be applied to many different types of biological samples as a means of enhancing detection coverage by LC–MS. Our approach involves one-step sample preparation using an aqueous low salt buffer at near physiological pH, followed by mixed-bed ion exchange gradient separations performed in tandem. Unlike a previous study [30], the collected protein fractions were obtained without column disconnection. The proteins are denatured and concentrated via acid precipitation, and digested extensively with trypsin. The resulting peptide mixtures are then analyzed sequentially by standard single-dimension reverse-phase capillary-scale electrospray LC–MS using automated data-dependent ion-trap tandem mass spectrometry. Finally, the peptide fragmentation patterns are interpreted by searching the acquired spectra against a reference protein sequence database. Although the procedure is straight-forward to implement in a generic analytical setting, we outline several key steps for optimizing sample fractionations using low salt mouse heart cytosolic extract and a high salt HeLa nuclear extract as illustrative test case examples.

2. Experimental methods

2.1. Chemicals and biologicals

Ultra-pure tris(hydroxymethyl)aminomethane (TRIS base), sodium chloride (NaCl), trichloroacetic acid (TCA), sodium azide (NaN_3), ammonium bicarbonate (NH_4HCO_3), calcium chloride ($CaCl_2$), HPLC grade water, HPLC grade acetonitrile (ACN), and HPLC grade glacial acetic acetic (AA) were

purchased from Fischer Scientific (Whitby, ON, Canada). Hydrochloric acid (6N HCL) was obtained from VWR International (Mississauga, ON, Canada), heptafluorobutyric acid (HFBA) from BioLynx (Brockville, ON, Canada), HPLC grade acetone and tissue culture media from Sigma–Aldrich (Oakville, ON, Canada), and protease cocktail inhibitor from Roche Diagnostics (Laval, QC, Canada). An equal mixture of Poroshell bulk immobilized trypsin (Applied Biosystems; Streetsville, ON, Canada) and TPCK-treated trypsin beads (Pierce Biotechnology; Rockford, IL, USA) were used for the tryptic digests.

2.2. Sample preparation

Mouse heart cytosol [3 mg ml^{-1} protein] was prepared in cardiac lysis buffer (250 mM sucrose, 50 mM Tris–HCl pH 7.6, 5 mM MgCl, 1 mM Dithiothreitol (DTT), and 1 mM Phenylmethylsulfonylfluoride) as described previously [31]. A 2 ml aliquot of frozen extract was thawed and briefly clarified by centrifugation at 14,000 rpm for 10 min in a pre-cooled Eppendorf centrifuge immediately prior to fractionation.

Loosely adherent HeLa cells (kindly provided by Ben Blencowe, Toronto, ON, Canada) were propagated in T-175 flask in a humidified incubator (37 °C, 5% CO_2) in DMEM high glucose media supplemented with 10% Fetal Bovine Serum. Cells were grown at 90% confluency, harvested, washed with phosphate buffer saline, trypsinized and pelleted by centrifugation (10 min, 1000 × *g*, room temperature). The supernatant was discarded and the cells stored at −80 °C. Cells were thawed, nuclei isolated and a nuclear extract (10 μg μl^{-1}) prepared in high salt buffer (20 mM HEPES, pH 7.9, 1.5 mM $MgCl_2$, 420 mM NaCl, 0.2 mM EDTA and 25% Glycerol) as reported previously [32]. The extract was stored at −80 °C prior to fractionation.

2.3. HPLC sample pre-fractionation

An integrated Agilent 1100 HPLC chromatograph system (Agilent Technologies, Mississauga, ON, Canada) consisting of a vacuum degasser, binary pump, refrigerated autosampler with 100 μl injector loop, thermostatted two column compartment, autocollection fraction module and multi-wavelength detector, was used in this study. Chromatogram acquisition and processing were performed according to basic procedures using the ChemStation software. The dual-column set-up, comprising a tandem WAX-CAT (PolyWAX LP, 200 × 2.1 mm i.d, 5 μm, 1000 Å; PolyCAT A, 50 × 2.1 mm i.d, 5 μm, 1000 Å) two-stage set-up. The system was protected from clogging with a weak anion exchange pre-column guard cartridge. All columns and guards were obtained from Canadian Life Science (Peterborough, ON, Canada). The column compartment was cooled to 17 °C while the other bays were chilled to 4 °C to minimize sample degradation.

Mobile phase Buffer A was prepared by dissolving 1.8 g of Tris in 1 liter of 10 mM HCl solution (pH 7.8 obtained without adjustment). Buffer B was prepared by dissolving 35 g of NaCl in Buffer A solution. A small amount (final concentration of 3 mM) of NaN_3 was added to minimize microbial outgrowth to each solution, which were also filtered using a 0.22 μm disposable membrane cartridge (Millipore, Cambridge, Canada) before use. Prior to sample injection into the HPLC system, the extracts were briefly clarified by centrifugation and stored in a pre-cooled autoloader. Injections were typically ∼120–150 μg total protein loaded per run. Elution was achieved using multi-step gradient, consisting of six transitions with increasing proportions of buffer B: (step 1; equilibration) 0%B, 0–8 min; (step 2; salt gradient) 0–45%B, 8–38 min; (step 3; high salt rinse) 45–100%B, 38–58 min; (step 4; high salt wash) 100%B, 58 − 66 min; (step 5; restoration) 100 − 0%B, 66–68 min; and, lastly, 0%B (step 6; re-equilibration) from 68 to 76 min. The column mobile phase flow rate was fixed at 250 μl min^{-1}. The chromatograms were monitored at 280 nm and timed fractions collected using an automated fraction collector cooled to 4 °C. A total of 26 fractions were collected per run, with one fraction isolated per ∼2.9 min (each ∼0.7 ml in volume for two successive injections).

2.4. Proteolysis and LC–MS sample preparation

The protein content of each HPLC fraction was precipitated by adding ice-cold neat TCA solution to a 10% (v/v) final concentration. Following overnight incubation at 4 °C, the samples were centrifuged at 14,000 rpm for 30 min in a cooled Eppendorf centrifuge. The supernatant was carefully aspirated and 600 μl ice-cold HPLC grade acetone gently added and removed to pH neutralize the protein pellets. The samples were air-dried in the fume-hood for 10 min, and resuspended in digest solution (50 mM NH_4HCO_3—1 mM $CaCl_2$) containing a 1:1 slurry of immobilized trypsin beads in a final volume of 60 μl. After incubation for two days at 30 °C with rotation, a 20 μl aliquot was withdrawn, mixed with an equal volume of LC–MS equilibration buffer (5% ACN, 0.5% AA, and 0.02% HFBA in HPLC grade water), and analyzed directly by LC–MS.

2.5. LC–MS analysis

For the proteomic detection, we used single-dimension reverse-phase chromatography coupled online to ion trap tandem mass spectrometry using standard conditions typically applied to samples of moderate complexity. Briefly, the digested peptide mixtures generated for each HPLC fraction were chromatographically resolved on a 150 μm inner diameter fused silica capillary micro-column (Polymicro Technologies, Phoenix, AZ, USA) bearing a fine nozzle created with a laser puller (P-2000, Sutter Instruments; Novato, CA, USA) that was connected in-line to an LTQ linear ion trap tandem mass spectrometer (Thermo Finnigan; San Jose, CA, USA). The columns were pre-packed with ∼8 cm of 5 μm Zorbax 300SB C_{18} resin (Agilent Technologies; Mississauga, ON, Canada) and were interfaced to a custom electrospray ion source. A Surveyor quaternary HPLC pump (Thermo Finnigan) was used to deliver a stable tip flow rate of ∼0.250 μl min^{-1} during the peptide separations. Elution of the peptides was achieved using a 90 min gradient (0 to 30% ACN in 60 min, then 30 to 80% ACN over

15 min, and returning to 0% buffer B in 15 min). Data-driven precursor ion selection and fragmentation was achieved in a fully automated data-dependent manner with dynamic target exclusion setting enabled.

2.6. Protein identification

The SEQUEST database search algorithm [33] was used to match up all the acquired spectra to peptide sequences encoded in a protein database downloaded from the European Bioinformatics Institute. The probabilistic STATQUEST model [34] was used to evaluate and assign confidence scores to all putative matches. Proteins were considered positively identified if detected with two or more high confidence (p-value <0.10) peptide candidates, each passing a minimum likelihood threshold corresponding to a 90% or greater probability of being correct. The proteomic patterns of the HPLC fractions were compared using the CONTRAST software tool [35]. Hierarchical clustering and visualization of the data in heat map format was performed as previously described [36].

3. Results and discussion

The objective of this study was to design and rigorously evaluate the effectiveness of a robust and generalizable IEX-HPLC pre-fractionation technique for improving routine proteomic detection coverage using standard gel-free tandem mass spectrometry screening procedures. The key objective was to retain and efficiently separate virtually all of the proteins present in a complex biological sample, excepting for a few non-binders and some limited residual bleed-through. For these reasons, we opted to develop a more efficient dual column mixed-bed ion exchange chromatography system, comprising both a terminal weak cation exchange (CAT) column (PolyCAT A, 50 × 2.1 mm i.d, 5 μm, 1000 Å) preceded by a weak ion exchange (WAX) column (PolyWAX LP, 200 × 2.1 mm i.d, 5 μm, 1000 Å). To minimize sample loss due to degradation and non-specific adsorption, sample loading and fractionation was completed quickly (within 76 min), with the entire system operated at sub-ambient temperatures (4–17 °C).

As representative biological test case samples, we evaluated the entire experimental system using first a low-salt ventricular muscle cytosolic extract obtained from mouse heart myocardium [14], and then a high-salt containing nuclear extract prepared from human-derived HeLa cancer cells [32]. By keeping the CAT column relatively short (50 mm length) and by applying a high flow rate (250 μl min^{-1}), we aimed to minimize the overall elution time, thereby avoiding undesirable (e.g. denaturation) and unpredictable (e.g. non-reversible absorbance) interactions of the proteins with the matrix, which may occur due to the presence of the buffering species (H-Tris$^+$) with the opposite charge to the CAT column packing material. The extracts were loaded and resolved without further pre-treatment as no further desalting or buffer exchange (even with high salt in the nuclear extract) was found to be required using conventional HPLC operating procedures.

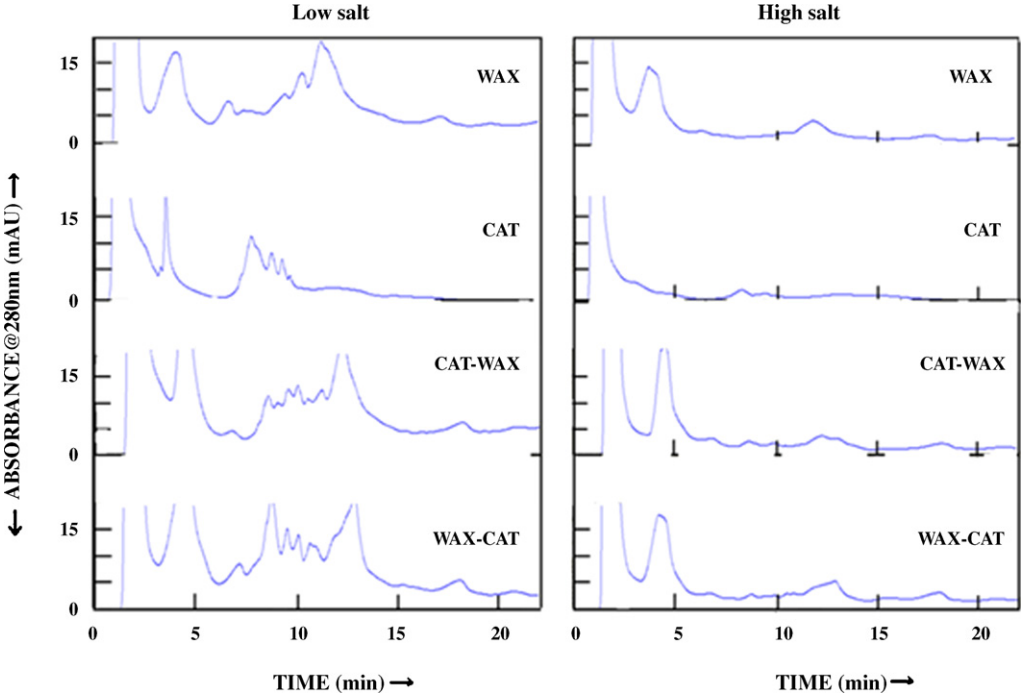

Fig. 1. Enhanced resolution using a tandem WAX-CAT chromatography system. Representative chromatograms showing the protein fractionations achieved with a mouse heart cytosolic protein extract containing either low salt (native, left panel) or high salt concentration (sample spiked with 420 mM NaCl; right panel) on single WAX, single CAT, tandem CAT-WAX, and tandem WAX-CAT column operations (top-to-bottom, respectively). Columns, 50 × 2.1 mm id, PolyWAX LP, and PolyCAT A; 5 μm particle size; 1000 Å pore size; flow-rate, 0.250 μl min^{-1}; temperature, 17 °C; linear gradient in 20 min from 0 to 0.6 M NaCl in Buffer A (see Section 2) followed by a 2.5 min washing at 0.6 M NaCl. Fifty microlitres of the sample (~150 μg total protein) was loaded.

As demonstrated in Fig. 1, single column configurations exhibited reduced retention and separation efficiencies. In terms of resolution, while it is obvious that a single WAX column outperformed a single CAT, a tandem arrangement exhibited markedly better resolving power than the WAX alone. This was true both for low and high-salt containing samples (*cf.* left and right panels). The WAX-CAT order showed moderately higher resolution than a CAT-WAX setup (see, for example, the degree of separation between the main eluting peaks from the void volume). Hence, it was determined that the WAX-CAT column geometry, together with an empirically optimized execution of the salt gradient (P.H.; data not shown), leads to the highest possible resolution of the sample along the entire chromatogram. This result concords with pioneering observations reported in the 1980's [37,38] on the mixed-bed IEX-HPLC of standard proteins mixtures (the only published papers on the subject, to our knowledge). The poor performance of single CAT alone for the high salt concentration was also anticipated. This material is designed to resolve basic proteins, which are not retained on WAX column, either by decreasing the mobile phase pH (starting at pH ∼7.0) or by increasing the salt concentration. Our relatively sub-optimal experimental conditions (elution buffer at pH 7.8 and especially the 420 mM NaCl in the high salt sample) negatively affect performance. Interestingly, a combination of WAX and CAT columns improved the separation of the salt-insensitive proteins (that is, a subset of proteins displayed stabilized binding at higher salt concentrations in the eluent) and likewise reduced the overall degree of unbound flow-through (which can be easily distributed uniformly in 8–10 fractions to facilitate identification).

Fig. 2, panel A shows a representative chromatographic profile, as recorded by UV-trace (absorbance at 280 nm), routinely achieved during the fractionation of the soluble heart cytosolic components under standardized chromatographic conditions. The procedure was also highly reproducible, as successive repeat injections of the same amount of sample material led to

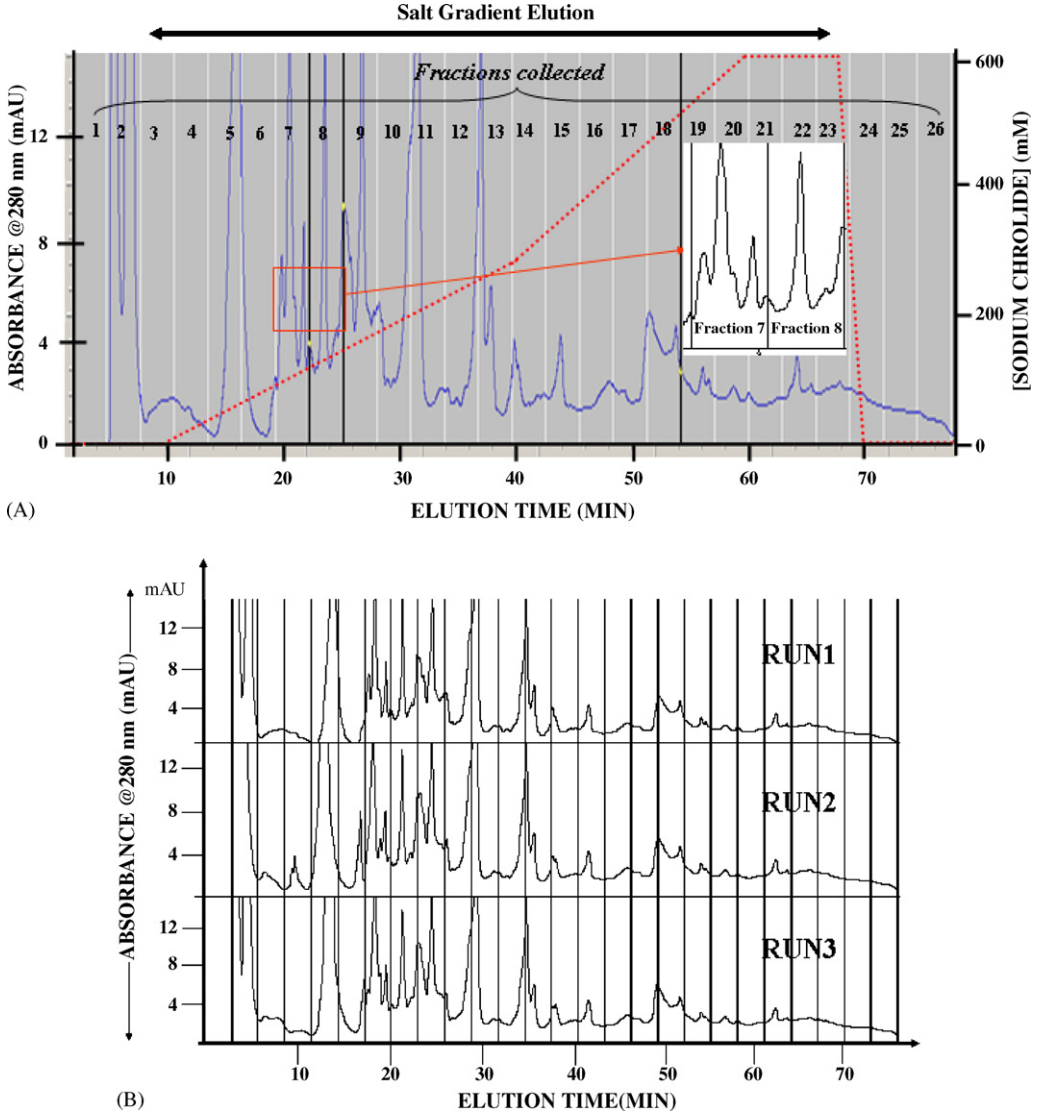

Fig. 2. Efficient IEX-HPLC separation of heart cytosolic proteins. Panel A provides a summary schematic of the sample pre-fractionation procedure superimposed over a representative chromatogram obtained with soluble mouse ventricle cytosol. Panel B shows the reproducibility of the IEX-HPLC method, indicating the UV-absorbance (280 nm) traces recorded for three successive HPLC runs after injection of the low salt heart cell extract.

essentially super-imposable protein chromatograms (again, as monitored by UV-absorbance), as illustrated in Fig. 2, panel B.

To determine the actual selectivity of the HPLC sorting method, the identities of the various proteins present in each fraction were identified by shotgun sequencing using automated data-dependent fragmentation of precursor peptide ions. To this end, the proteins in the disparate fractions were precipitated and digested extensively using trypsin. The resulting peptide mixtures were then chromatographically resolved and subsequently detected by means of microcapillary reverse-phase HPLC combined with online electrospray ionization into an attendant linear ion trap tandem mass spectrometer. All of the mass spectra generated by the instrument were then searched exhaustively against a suitable reference protein sequence database using the SEQUEST algorithm [33] and a list of highly confident (>90%) candidates deduced using the STATQUEST statistical model [34]. To examine and compare the proteomic patterns obtained, the spectral counts assigned to each putatively identified protein across the entire set of HPLC-derived samples in each of two separate experiments was deduced as a semi-quantitative measure [14]. For ease of inspection, the values were then clustered and visualized in a heat map format (Fig. 3, panel A), which showed the generally good reproducibility of the LC–MS screening procedure. Most of the proteins identified by LC–MS were detected, and hence presumably eluted, in a small subset of the fractions collected during the shallow linear gradient phase of the elution (e.g. fractions 4 to 20). Moreover, only a modest number of distinct proteins were identified in the flow-through fractions, indicating the high efficiency of the dual column set-up as a means for retaining the bulk of the complex sample material. This latter result can be explained by the fact that a near optimal interaction of the proteins (at neutral physiological pH) with the matrices is achieved due to low-salt concentration of the sample, while usage of a CAT column immediately downstream of the WAX column served to retain and separate loosely bound positively charged proteins, hence minimizing sample loss.

The maximum spectral counts recorded in one experiment was also calculated across the entire set of HPLC-derived samples during one representative analysis and compared to the corresponding value obtained in a straight analysis of an unfractionated reference cytosolic sample performed in parallel (Fig. 3, panel B). Collectively, the data indicated that the HPLC fractionation allowed for both the identification of far more proteins and the assignment of greater supporting spectral evidence as compared to a direct sample analysis.

By using longer columns, even more impressive separation results can be achieved (data not shown), albeit at the cost of a lengthened overall analysis time. However, a major drawback is that the eluting proteins would subsequently be collected in larger volume of buffer, which is a significant concern when detection of dilute low-abundance proteins is a priority. In this case, even pooling fractions from consecutive HPLC runs would likely not overcome the problem of sample concentration prior to LC–MS detection. Hence, to enhance the discovery of lower abundance nuclear proteins, such as transcription factors, two permutations to the basic fractionation procedure can be envisaged. First, one can generate smaller sample volumes by collecting more fractions (for instance, one fraction vial every minute, resulting in >50–100 fractions over the entire analysis). However, although this would yield a significant improvement in protein detection sensitivity due to improved sample recovery during the precipitation and digestion stages, it is not advisable under most circumstances since it leads to an increased need for time-consuming LC–MS. Alternatively, one could reduce the

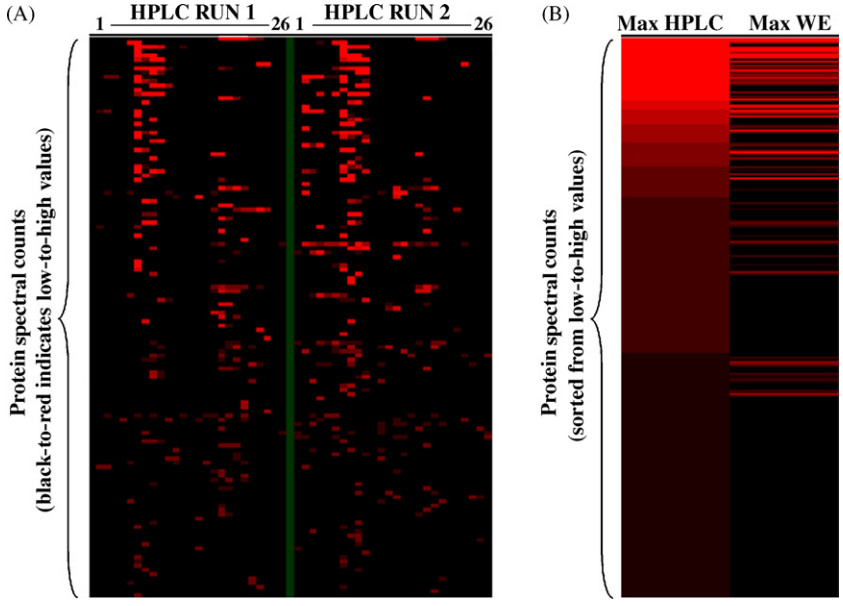

Fig. 3. Enhanced detection coverage by subsequent LC–MS analysis. Panel A outlines the proteomic profiles obtained by LC–MS analysis of the heart protein fractions recorded for two independent IEX-HPLC experiments (runs 1 & 2). Panel B shows the significant differences in overall detection coverage by LC–MS (indicated by the number of high confidence matching spectra recorded per protein) obtained by sample pre-fractionation with HPLC vs. straight analysis of the starting heart whole cell extract (WE).

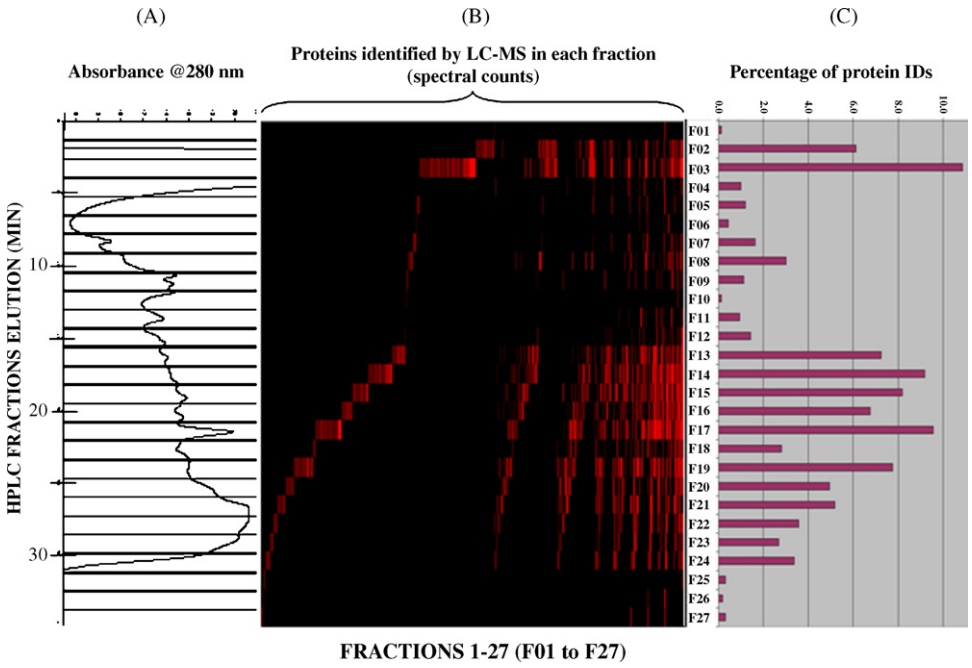

Fig. 4. HeLa nuclear extract HPLC pre-fractionation and LC–MS analysis. Panel A shows a UV-trace recording generated during IEX-HPLC fractionation of HeLa nuclear extract, as monitored at 280 nm. Panel B indicates the profiles of each of the proteins identified by LC–MS after IEX-HPLC fractionation in a heat map cluster format. Panel C shows a plot of the total number of proteins identified in each of the 27 individual fractions collected by IEX-HPLC fractionation.

column length, while maintaining a high flow rate and salt gradient, thereby decreasing the overall analysis time while allowing one to collect a reasonable number of fractions without much sample dilution. The pooling of adjacent fractions obtained from successive HPLC runs would also be of help in this case.

We explored this latter avenue by separating a HeLa cell derived nuclear extract using a 50 × 2.1 mm WAX column connected upstream of the 50 × 2.1 mm CAT column. The samples were clarified by centrifugation before injection. The columns were equilibrated in the starting buffer (10 mM Tris–HCl, pH 7.8). Two HPLC runs were executed in succession, with 40 μl of sample per injection, and fractions collected at every ∼1.3 min using a multi-well collection plate. After injection, the NaCl concentration of the mobile phase was linearly increased to up to 600 mM NaCl in 35 min, with the flow rate maintained at 250 μl min^{-1}. As shown in Fig. 4, despite the high salt concentration of the initial nuclear extraction buffer solution, a considerable number of different proteins were identified in almost each HPLC fractions. Again, these results indicate that pre-fractionation allows for both improved sensitivity (more proteins identified) and enhanced confidence (more spectra and peptides identified per protein) of detection across a broader range of proteins as compared with the more standard methods of direct LC–MS analysis of unfractionated cell lysates. This increased number of identified proteins presumably represents enhanced detection of lower-abundance proteins. In most cases, the eluting proteins were uniquely identified in a single or a few adjacent fractions, which reflects the overall high efficiency of the two column system to both retain and separate a complex mixture of proteins even in the presence of possible ionic competitors in the starting biological specimen.

4. Conclusions

Although a long established method, HPLC has the exciting potential to be used in synergistic conjunction with proteomic analyses. In this study, we used a dual column ion exchange HPLC system as a simple, reproducible and effective means of simplifying a complex proteome sample. Our preliminary results suggest that IEX-HPLC fractionation should be a useful method for global proteome analysis. The method was evaluated and validated by analyzing mouse heart cytosol and HeLa cell nuclear extracts as equally challenging test mixtures. Two sets of pilot profiling data indicated the effectiveness, rapidity, reproducibility and robustness of the procedure.

The approach described here is flexible, in that different fractionation regimes can be readily executed by a relatively simple modification of the salt gradient and the mode of sample collection. Hence, a proteome can be divided up among any number of HPLC fractions so as to optimize downstream analysis. The dual column set-up was characterized by impressive resolution, with minimal flow-through. Not surprisingly, we observed markedly improved subsequent proteomic detection coverage by LC–MS as indicated by a significant improvement in the number and quality of identifications made by tandem mass spectrometry. In effect, fractionation allowed for the detection of both more and different proteins (including presumably lower abundance factors). Furthermore, it enhanced the reliability of tentative candidate identifications by allowing for the recording of more supporting spectra.

An additional advantage of the technique is that, at least in principal, the many proteins present in a complex biological sample can be selectively enriched via consecutive HPLC injections,

by collecting and subsequently pooling the appropriate fractions prior to LC–MS analysis. These same qualities suggest a broad potential utility for improving biomarker discovery in a clinical proteomic setting. For these reasons, this method is now being routinely used in our laboratory as a standard sample preparation protocol for many of our proteomic projects, especially those aimed at biomarker discovery in biofluids. A potential caveat to such tandem arrangements is increased non-specific absorption and therefore sample loss and subsequent coverage bias by LC–MS. However, the use of long, thin IEX columns as a means to fractionate the protein sample prior to LC–MS analysis could be considered if further improvement in resolution (while minimizing loss) is a major consideration (e.g., to isolate a select group of low abundance proteins). In this case, a time window can be pre-defined so as to concentrate fraction volume. Conversely, if total analysis time is a concern, a combination of two short IEX columns might also well be considered.

Acknowledgements

We thank Vincent Fong for his expert assistance with the computational data analysis, Andy Alpert (PolyLC, Inc.) for helpful chromatography advice, and Dr. Thomas Kislinger for generous access to the heart tissue sample. This work was supported by operating grants to A.E. from the McLaughlin Centre for Molecular Medicine, Genome Canada and the Ontario Genomics Institute (OGI) and the Natural Science and Engineering Council of Canada (NSERC).

References

[1] T. Kislinger, A. Emili, Curr. Opin. Mol. Ther. 5 (2003) 285.
[2] B. Domon, R. Aebersold, Science 312 (2006) 212.
[3] T. Kislinger, A. Emili, Expert Rev. Proteomics 2 (2005) 27.
[4] R.A. Meyers, Encyclopedia of Molecular Cell Biology and Molecular Medicine, Wiley-VCH Verlag, Weinheim, 2005.
[5] T. Kislinger, A.O. Gramolini, Y. Pan, K. Rahman, D.H. MacLennan, A. Emili, Mol. Cell Proteomics 4 (2005) 887.
[6] M. Maziarz, C. Chung, D.J. Drucker, A. Emili, Mol. Cell Proteomics 4 (2005) 458.
[7] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R. Yates 3rd, Nat. Biotechnol. 17 (1999) 676.
[8] M.P. Washburn, D. Wolters, J.R. Yates 3rd, Nat. Biotechnol. 19 (2001) 242.
[9] T.J. Griffin, S.P. Gygi, T. Ideker, B. Rist, J. Eng, L. Hood, R. Aebersold, Mol. Cell Proteomics 1 (2002) 323.
[10] S.P. Gygi, Y. Rochon, B.R. Franza, R. Aebersold, Mol. Cell Biol. 19 (1999) 1720.
[11] K.L. Johnson, C.J. Mason, D.C. Muddiman, J.E. Eckel, Anal. Chem. 76 (2004) 5097.
[12] M. Dreger, L. Bengtsson, T. Schoneberg, H. Otto, F. Hucho, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 11943.
[13] L. Johnson, S. Mollah, B.A. Garcia, T.L. Muratore, J. Shabanowitz, D.F. Hunt, S.E. Jacobsen, Nucleic Acids Res. 32 (2004) 6511.
[14] T. Kislinger, B. Cox, A. Kannan, C. Chung, P. Hu, A. Ignatchenko, M.S. Scott, A.O. Gramolini, Q. Morris, M.T. Hallett, J. Rossant, T.R. Hughes, B. Frey, A. Emili, Cell 125 (2006) 173.
[15] V.K. Mootha, J. Bunkenborg, J.V. Olsen, M. Hjerrild, J.R. Wisniewski, E. Stahl, M.S. Bolouri, H.N. Ray, S. Sihag, M. Kamal, N. Patterson, E.S. Lander, M. Mann, Cell 115 (2003) 629.
[16] C.M. Taylor, S.E. Pfeiffer, Proteomics 3 (2003) 1303.
[17] G. Rigaut, A. Shevchenko, B. Rutz, M. Wilm, M. Mann, B. Seraphin, Nat. Biotechnol. 17 (1999) 1030.
[18] N.J. Krogan, G. Cagney, H. Yu, G. Zhong, X. Guo, A. Ignatchenko, J. Li, S. Pu, N. Datta, A.P. Tikuisis, T. Punna, J.M. Peregrin-Alvarez, M. Shales, X. Zhang, M. Davey, M.D. Robinson, A. Paccanaro, J.E. Bray, A. Sheung, B. Beattie, D.P. Richards, V. Canadien, A. Lalev, F. Mena, P. Wong, A. Starostine, M.M. Canete, J. Vlasblom, S. Wu, C. Orsi, S.R. Collins, S. Chandran, R. Haw, J.J. Rilstone, K. Gandi, N.J. Thompson, G. Musso, P. St Onge, S. Ghanny, M.H. Lam, G. Butland, A.M. Altaf-Ul, S. Kanaya, A. Shilatifard, E. O'Shea, J.S. Weissman, C.J. Ingles, T.R. Hughes, J. Parkinson, M. Gerstein, S.J. Wodak, A. Emili, J.F. Greenblatt, Nature 440 (2006) 637.
[19] V. Badock, U. Steinhusen, K. Bommert, A. Otto, Electrophoresis 22 (2001) 2856.
[20] M.-x. Gao, J. Hong, P.-y. Yang, X.-m. Zhang, Anal. Chim. Acta 553 (2005) 83.
[21] B.O. Keller, Z. Wang, L. Li, J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 782 (2002) 317.
[22] G. Van den Bergh, S. Clerens, F. Vandesande, L. Arckens, Electrophoresis 24 (2003) 1471.
[23] E. Barnea, R. Sorkin, T. Ziv, I. Beer, A. Admon, Proteomics 5 (2005) 3367.
[24] X. Li, Y. Gong, Y. Wang, S. Wu, Y. Cai, P. He, Z. Lu, W. Ying, Y. Zhang, L. Jiao, H. He, Z. Zhang, F. He, X. Zhao, X. Qian, Proteomics 5 (2005) 3423.
[25] J. Martosella, N. Zolotarjova, H. Liu, G. Nicol, B.E. Boyes, J. Proteome Res. 4 (2005) 1522.
[26] Z. Zhang, D.L. Smith, J.B. Smith, Proteomics 1 (2001) 1001.
[27] L.R. Snyder, J.L. Glajch, J.J. Kirkland, Practical HPLC Method Development, John Wiley & Sons, New York, 1997.
[28] P. Lescuyer, D.F. Hochstrasser, J.C. Sanchez, Electrophoresis 25 (2004) 1125.
[29] P.G. Righetti, A. Castagna, B. Herbert, F. Reymond, J.S. Rossier, Proteomics 3 (2003) 1397.
[30] B.L. Hood, M. Zhou, K.C. Chan, D.A. Lucas, G.J. Kim, H.J. Issaq, T.D. Veenstra, T.P. Conrads, J. Proteome Res. 4 (2005) 1561.
[31] T. Kislinger, A.O. Gramolini, D.H. MacLennan, A. Emili, J. Am. Soc. Mass. Spectrom. 16 (2005) 1207.
[32] C. Sandhu, M. Connor, T. Kislinger, J. Slingerland, A. Emili, J. Proteome Res. 4 (2005) 674.
[33] J.K. Eng, A.L. McCormack, I.I.I.J.R. Yates, J. Am. Soc. Mass. Spectrom. 5 (1994) 976.
[34] T. Kislinger, K. Rahman, D. Radulovic, B. Cox, J. Rossant, A. Emili, Mol. Cell Proteomics 2 (2003) 96.
[35] D.L. Tabb, W.H. McDonald, J.R. Yates 3rd, J. Proteome Res. 1 (2002) 21.
[36] B. Cox, T. Kislinger, A. Emili, Methods 35 (2005) 303.
[37] Y.-F. Maa, F.D. Antia, Z. El Rassi, C. Horvath, J. Chromatogr. 452 (1988) 331.
[38] Z. El Rassi, C. Horvath, J. Chromatogr. 359 (1986) 255.