Published in final edited form as:

Electrophoresis. 2012 January ; 33(1): 74–88. doi:10.1002/elps.201100431.

Liquid-phase-based separation systems for depletion, prefractionation and enrichment of proteins in biological fluids and matrices for in-depth proteomics analysis – An update covering the period 2008 – 2011

Subhashini Selvaraju and Ziad El Rassi*

Department of Chemistry, Oklahoma State University, Stillwater, OK 74078-3071

Abstract

This review article expands on the previous one (Y. Jmeian and Z. El Rassi, *Electrophoresis* 2009, 30, 249–261) by reviewing pertinent literature in the period extending from early 2008 to present. As the previous review article, the present one is concerned with proteomic sample preparation (e.g., depletion of high abundance proteins, reduction of the protein dynamic concentration range, enrichment of a particular sub-proteome), and the subsequent chromatographic and/or electrophoretic pre-fractionation prior to peptide separation and identification by LC-MS/MS. This review article is distinguished from its first version published in *Electrophoresis* 2009, 30, 249–261 by expanding on capturing/enriching sub-glycoproteomics by lectin affinity chromatography. Ninety-eight papers published in the period extending from early 2008 to the present have been reviewed. By no means this review article is exhaustive, giving the fact that its aim is to give a concise treatment of the latest developments in the field.

Keywords

Proteomics; Immunodepletion; ProteoMiner; Protein fractionation; Lectin affinity chromatography

1 Introduction

Proteomics analysis and profiling is most often hampered by the vast complexity of the biological samples such as body fluids, bio-extracts, tissues and cells. Not only that the dynamic concentration range of proteins in biological samples extends over several orders of magnitude, but also the very large number of the protein constituents make proteomics a sample preparation problem par excellence. These aspects of proteomics samples have necessitated the development of the most elaborate fractionation approaches to narrow down the dynamic concentration range and to reduce the complexity of the proteomics samples by fractionation. Further progress in proteomic sample preparation and fractionation is expected to allow the exploitation of the full potentials of the currently sophisticated and advanced mass spectrometry (MS) instrumentation for in-depth proteomics analysis.

This review article is divided into four major parts, including (i) depletion methods for the removal of high abundance proteins (HAPs), (ii) protein equalizer technology to reduce the dynamic protein concentration range, (iii) chromatographic and electrophoretic fractionation prior to LC-MS/MS to reduce proteomics sample complexity and (iv) capturing of specific

*Corresponding author: Prof. Ziad El Rassi, Department of Chemistry, Oklahoma State University, Stillwater, OK, 74078-3071; tel: 405-744-5931; fax: 405-744-6007, elrassi@okstate.edu.

sub-glycoproteome by lectin affinity chromatography (LAC) to focus on the sub-proteome that may have the biological information of interest. In most proteomics profiling, either step (i) or (ii) is performed first on the sample followed by step (iii) or (iv).

Published research articles on topics (i), (ii) and (iii) over the period extending from early 2008 to present have been reviewed here in order to expand on the previous review article by Jmeian and El Rassi [1] that discussed the liquid-based separation system employed for in-depth proteomic analysis for the time period 2002-early 2008. Since topic (iv) on the selective capturing of sub-glycoproteomics by LAC is a new addition to this review article, literature published over the last 5 years (2006 to the present) has been reviewed.

Other typical review articles are worth mentioning here because of the wealth of information they may provide to the readers of this Journal and also they are very related or complement the present review article. One must cite the very recent review article on protein equalizer technology by Righetti *et al.* [2], as well as the review articles on sample preparation and fractionation [3–6] and glycoproteomics and glycans capturing by LAC [7–9].

2 Different techniques currently used to reduce the complexity of proteomics samples – Sample treatment, pre-fractionation and selective capturing of sub-proteomics

2.1 Depletion methods

Depletion is one of the commonly used methods to reduce the complexity of a given proteomics sample as well as to narrow down the dynamic concentration range of its protein constituents despite reports that some of the depletion methods result in co-depletion of many other clinically important low abundance proteins [10].

2.1.1 Solvent solubilization/precipitation methods—Some of the low molecular weight (LMW) proteins that are present in plasma/serum reflect the patho-physiological state of a given subject and consequently could serve as potential biomarkers [11]. Thus, the identification of these LMW proteins gains importance in proteomics analysis. Kawashima *et al.* [12] introduced a differential solubilization (DS) method to extract/concentrate LMW proteins/peptides in serum with good reproducibility and yield. In this DS method, the LMW proteins were isolated by first diluting the serum with denaturing solution containing urea, thiourea and dithiothreitol (DTT), then slowly dropping the diluted serum into ice-cold acetone and immediately stirring at 4 C for 1h. This was followed by centrifugation, and the precipitate thus obtained was taken up with acidified acetonitrile. After mixing at 4 C for 1h, the mixture was centrifuged again for 15 min at 4 C. The LMW proteins/peptides were extracted in the supernatant. Using this DS method, the investigators were able to perform high-quality comparative analysis of more than 1500 LMW proteins/peptides from 1 μ L of serum by combining the DS method with reversed-phase chromatography (RPC) separation followed by matrix assisted laser desorption/ionization (MALDI)-time of flight (TOF)-MS. This allowed the detection of low abundance peptides in the subnanomolar range and containing many peptides bound to carrier proteins such as albumin. The authors reported the application of this method to the discovery of four new biomarker candidates of colon cancer with one of them is a fragment of the protein zyxin that possibly originated from tumor cells.

Warder *et al.* [13], described a protocol to precipitate the high abundance proteins (HAPs) which included albumin and transferrin. In brief, the serum was incubated with either DTT or tris(2-carboxyethyl)phosphine, and centrifuged to obtain the pellet which was rich in albumin. Analysis of the supernatant by MALDI-TOF-MS did not show any intact albumin

ions indicating that this protein was completely precipitated. This reproducible method was compatible with plasma from other species and can also be scaled up to large milliliter quantities.

2.1.2 Immunoaffinity depletion methods—Immunoaffinity depletion methods involve the use of immobilized antibodies to capture one or a group of HAPs. There are many commercially available immunodepletion columns to deplete a number of top HAPs. In a recent study by Tu *et al.* [14], the performance of the commercially available Agilent Technologies MARS-7 and MARS-14 columns to deplete 7 or 14 HAPs was evaluated in the shotgun proteomic analysis of human plasma. It was observed that the depletion columns were highly reproducible, but a few non-targeted proteins were also captured by the depletion columns. Immunodepletion of top 7 or 14 proteins resulted in 25% increase in identified proteins as compared to unfractionated plasma. Twenty-three low abundance plasma proteins (< 10 ng/mL) were detected, and this represented 5–6% of total identified proteins in the immunodepleted plasma.

In a recent study, a one-step process was reported to concentrate, purify and deplete albumin from urine [15]. Briefly, the urine proteins were first reduced, alkylated and transferred to a spin filter, where it was treated with anti-human serum albumin. The proteins were incubated and spun down to remove the waste. The supernatant solution containing the proteins were further tryptically digested, fractionated either using OFF-gel electrophoresis (OGE) or high pH RPC separation and analyzed by LC-MS/MS. By the OG fractionation, 703 proteins were identified whereas using the RPC separation 499 proteins were identified. This simple, efficient and reproducible method is compatible with diverse down-stream applications and is also a potential method to study other complex body fluids [15].

A comparative study was made to see the outcome when one, six, twelve or twenty major proteins were depleted using Beckman Coulter ProteomLab IgY-HAS, Agilent Technologies Hu-6HC and Beckman Coulter ProteomLab IgY-12, respectively [16]. When the number of depleted proteins was increased from twelve to twenty proteins the benefits were limited, whereas when 6 proteins were depleted many low abundance proteins were detected [16]. In another study [17], 12 highly and 77 moderately abundant proteins present in serum were depleted using GenWay Biotech's Seppro microbeads IgY12 and SuperMix columns, respectively. The authors reported the identification of 222 and 71 proteins when 89 ($77 + 12 = 89$) and 12 proteins were depleted, respectively. Clearly, this indicates that the simultaneous depletion of the high and the moderately abundant proteins would increase the number of identified low-abundance proteins. These findings were in good agreement with the results presented by Qian *et al.* [18] in which the same approach was followed to reduce the complexity of the proteome. The authors were able to establish the above-mentioned approach to identify the differentially expressed proteins in ovarian cancer sera. It has to be noted that despite a simultaneous depletion of both high and moderate abundance proteins was performed on the sera, the flow-through fractions still contained some of the moderately abundant proteins, indicating that the depletion was not complete [17, 18]. Also, depleting the 77 moderately abundant proteins might have a disadvantage that it can result in co-depletion of some of the clinically important proteins.

An assessment of the manufacturer's protocol to remove the bound albumin from an immunoaffinity column was described by Gundry *et al.* [19]. Beckman Coulter IgY-anti-human serum albumin spin column was evaluated in this study. It was observed that even after stripping the column for six consecutive times, albumin was still bound to the spin column. The authors suggested that either an alternative composition of the stripping buffer or a different protocol is necessary for effective removal of albumin, which would result in efficient regeneration of the spin column.

The performance of six commercially available depletion columns (Bio-Rad Aurum™ Affi-Gel® Blue mini kit, Sartorius Stedim Biotech Vivapure® anti-HSA/IgG kit, Qiagen Qproteome albumin/IgG depletion kit, Agilent Technologies MARC column - human 6, GenWay Biotech Seppro® MIXED12-LC20 column and Sigma-Aldrich ProteoPrep® 20 plasma immunodepletion kit) were evaluated based on efficiency of removal of HAPs, non-specific binding and total number of proteins spots detected in 2-DE [20]. Higher number of protein spots was detected in the flow through fractions of multiple affinity removal column (MARC) and Seppro IgY columns. It was concluded that the Seppro column showed overall best performance as compared to other five columns, whereas MARC column also provided notable performance and could be a choice in terms of economical conditions.

The performance of three different affinity and immunoaffinity subtraction spin columns (Montage albumin deletion kit, ProteomeLab IgY-12 partitioning kit and ProteoPrep 20 plasma immunodepletion kit) was compared using a pool of human cerebrospinal fluid (CSF) [21]. The bound and the unbound fractions of the CSF from the spin columns were analyzed by 1D gel electrophoresis and nanoLC-MALDI-TOF/TOF-MS and a comparison of the MASCOT MS/MS ion scores was also made. In the 1D gel electrophoresis, increased number of protein bands was observed in the flow through fractions from the three columns as compared to the intact CSF. The co-depletion of proteins that were bound to the HAPs by the so-called “sponge effect” was the largest on the purely affinity based Montage column. It was observed that all the three columns showed an increase in the MASCOT ion scores for the medium and low abundance proteins. It was also observed that the HAPs removal with the ProteomeLab IgY-12 spin column yielded reproducible results. Overall, 173 unique proteins were identified in the human CSF, which included 91, 128, 123 and 104 unique proteins from the intact CSF, Montage, ProteomeLab IgY-12 and ProteoPrep 20 spin columns, respectively.

2.1.3 Depletion by hydrophobic interaction chromatography—A method to deplete the high abundance proteins was proposed based on hydrophobic interaction chromatography (HIC) [22]. In this work, the hydrophobicity of 56 main plasma proteins was calculated based on their amino acid composition. These 56 proteins were classified as low, medium and highly hydrophobic through a cluster analysis. Some of the highly abundant proteins (albumin, Ig heavy chain γ -1, Ig heavy chain γ -2, Ig heavy chain γ -3, Ig δ , Ig heavy chain μ , serotransferrin, haptoglobin and fibrinogen γ) fell in the category of medium hydrophobicity proteins (i.e., cluster 2), which represented 70% of the highly abundant proteins. This HIC-based method was evaluated by 2-DE and the results were compared to those obtained by the immuno-affinity depletion of albumin. While both methods showed similar reproducibility, HIC allowed only the partial depletion of both α -1-antitrypsin and albumin. Despite the partial depletion of albumin, the HIC permitted the detection of twice the number of spot than the immuno-depletion method. However, realizing the incomplete depletion of albumin, the authors suggested that the HIC method would require further optimization in order to be used as complementary or alternative method to the immuno-depletion method.

2.2 Protein equalizer approach including its comparison with other sample treatment methods

In principle, combinatorial peptide ligand library (CPLL) reduces significantly (i.e., almost equalizes) the dynamic concentration range of proteins present in a given complex biological sample. In fact, under overloading conditions, and when a given proteomics sample is treated with the CPLL beads, HAPs immediately saturate the corresponding ligands whereas the proteins that are present at very low level gets selectively enriched by the corresponding ligands [23]. The CPLL beads were combined with other fractionation

methods such as differential gel electrophoresis (DIGE) and OGE fractionation [24, 25] in order to identify a high number of low abundance proteins. Also, the protein equalizer technology has allowed in-depth proteomics analysis of a complex mixture of human proteome [26] as well as of animal plasma proteome as it does not require any antibodies that are usually specific for human plasma proteins [27].

In a novel method, the equalized proteins were further fractionated based on their differences in isoelectric points using solid-state-buffers (SSB) associated with cation exchangers [24]. In earlier methods, the equalized proteins, which resulted from treating a given serum sample with the CPLL beads, were further fractionated by 1D gel analysis. This procedure yielded a number of fractions to be analyzed. To avoid this cumbersome process, protein equalizer in combination with SSB method was proposed, whereby the equalized proteins are allowed to adsorb on a solid phase where their net charge is opposite to that of the ion exchange column. This proposed method, which reduced the number of fractions to be analyzed, was compared to the performance of the classical anion exchange chromatography. It was observed that the eluted fractions from the SSB-based process had different ranges of isoelectric points, while the anion exchange chromatography did not show a good discrimination of the isoelectric points. When this method was compared to the performance of OGE fractionation (pI-based fractionation in off-gel format) after treating the serum with CPLL beads, it was observed that the SSB method detected more protein spots in 2-DE, although the OGE method resulted in better pI discrimination.

In another study [28], three different methods including HAPs precipitation, restricted access materials (RAM) combined with immobilized metal affinity chromatography (IMAC) and CPLL beads were evaluated to see which of these methods serves as a best fractionation step for the analysis of LMW proteins. The evaluation was performed based on the peptide/protein peaks generated from surface enhanced laser desorption/ionization (SELDI)-TOF-MS analysis and on the reproducibility of these methods. Even though the authors concluded that all the three methods were complementary, the CPLL beads efficiently captured the high molecular weight (HMW) proteins. The IMAC-RAM method identified some additional LMW protein peptides whereas the precipitation method using organic solvents did not give any new information on the peptide/protein peaks.

In a comparative study [29], three different commercially available protein enrichment kits were evaluated in terms of their effectiveness in detecting/accessing the low abundance proteins. The three different methods were immunodepletion using Seppro IgY14 that contained polyclonal antibodies raised against the 14 highest abundance proteins, a two-step immunodepletion process using Seppro IgY14 and Seppro IgY-Supremix system that contained a mixture of antibodies raised against the proteins present in the flow-through fraction of IgY12 and the third strategy involved the use of the CPLL beads. When the bound fractions were analyzed using 2-DE, as expected differences in the protein patterns were observed for the three methods. In another experiment by the authors, the flow through from IgY14 column was treated with CPLL beads. In principle, combining two different fractionation methods should allow for the identification of more number of proteins. In contrast, it was observed that the multi-step fractionation showed only slight increase in the sensitivity as compared to the onestep fractionation. It was observed that the one-step fractionation using the CPLL beads and the IgY14 fractionation increased the number of genes in 2-DE as compared to the unprocessed plasma. The author concluded that the combination of the two different fractionation methods did not show any significant increase in the number of genes in 2-DE, but it only made the whole process very expensive leading to few thousand dollars per sample.

In a very recent contribution by Di Girolamo *et al.* [30], the CPLL technology was further exploited in serum sample preparation, where it was strongly recommended to use boiling 4% SDS and 25 mM DTT for maximum release of captured proteins from the CPLL beads. This was realized by investigating four elution protocols consisting of (i) 4M urea and 1% CHAPS, (ii) 4M urea, 1% CHAPS and 5% acetic acid, (iii) 8M urea, 2% CHAPS and 5% acetic acid and (iv) boiling 4% SDS and 25 mM DTT. In all cases, 1 mL of serum was treated with 50 μ L of CPLL beads, which were then eluted with the 4 eluents just mentioned. With the first 3 eluents and after the first elution, the beads were re-eluted with eluent (iv) knowing that it offered maximal release from previous work by the authors [31]. It was found that eluent (i) released only ca. 20% of the proteins adsorbed to the beads while eluent (ii) released ca. 60% and eluent (iii) ca. 80%. On this basis, it was recommended that eluent (iv) is perhaps the only cocktail that would allow the release of the majority of low abundance proteins in pursuit of biomarker discovery. In fact, and as just mentioned Candiano *et al.* [31] have suggested the use of a highly effective, single elution similar to eluent (iv) to reduce the workload that was associated with the traditional three- to four-step sequential elution system consisting of various agents which generated many distinct fractions that added to the burden of having to analyze all of them.

Not only the release of captured proteins is an important operational aspect to pay attention to, but also the binding conditions are of equal importance. Fasoli *et al.* [32] investigated the pH involvement for a larger capture of protein species in sera and other proteomes. The traditional physiological saline of pH 7.2 seems to yield some loss of species initially present in the untreated sample. On this basis, the authors investigated two additional pH at acidic (pH 4.0) and alkaline (pH 9.3) pH values. It was found that the capture process is more extensive with a number of additional species captured at the two-pH extremes. That is, at pH 4.0 newly detected proteins were mostly acidic while at the alkaline pH additional protein species were more evenly distributed throughout the pI range towards the alkaline area. The role of pH was illustrated using snake venom and leaf protein extract. It was suggested that for reproducibility and comparative proteomics profiling, the operational pH should be kept under tight control.

In a study by Beseme *et al.* [33], CPLL was evaluated in combination with 2-DE in a restricted pH range (4–7). Also a comparison was made between the native, CPLL-treated and albumin and IgG depleted plasma samples to see which approach allowed access to the low-abundance proteins. In the albumin and IgG depleted plasma 427 spots were detected which reflected a 271% increase as compared to the native plasma, whereas in the CPLL-treated plasma 557 spots were detected corresponding to 355% and 130% increase as compared to native and the depleted plasma, respectively. The authors concluded that the CPLL technology was superior to the albumin and IgG depletion approach but the CPLL approach has the limitation that 1 mL of serum is needed in order to accumulate on the CPLL beads sufficient and detectable amount of proteins.

The reproducibility of CPLL beads in the quantitative and comparative serum proteomic analysis was recently investigated [34]. Also, a comparison was made with IgY-based affinity depletion approach in which 12 most abundant serum proteins were removed. Serum was divided into three portions and they were treated with three different ProteoMiner columns. It was observed that for the different ProteoMiner columns the quantitative and qualitative aspects were similar. The authors concluded that there could be up to 2-fold variation in the relative efficiencies of capturing proteins, a factor that might affect differential protein analysis. In the comparison process, the authors suggested that the ProteoMiner columns were at least as effective as the IgY approach in reducing the protein levels. However, the ProteoMiner had the limitation that it needs 1 mL of serum whereas the IgY approach required only 20 μ L. While this may represent a limiting factor for the

analysis of some clinical samples because of restricted availability, the larger volumes may be a contributing factor to increasing the ability to isolate lower abundance protein species. Indeed, the CPLL beads showed the ability to capture low abundance proteins such as cytokines.

In a study by Ernoul *et al.* [35], two different approaches, namely immuno-depletion and CPLL technology, were evaluated based on the number of identified proteins, presence of most abundant proteins and identification of medium and low concentration proteins in human plasma. First, either an immuno-depletion (using MARS-Hu 14 to remove 14 most abundant proteins) or CPLL treatment was performed on plasma and then the proteins were tryptically digested, labeled using isobaric tag for relative and absolute quantification (iTRAQ) and fractionated into 24 fractions by an OGE process, followed by analysis by nano-RPC-MALDI MS/MS. A total of 313 proteins were identified by both the immuno-depletion and the CPLL, in which 158 proteins were common to both methods. There were 85 and 70 proteins unique to the immuno-depletion step and the CPLL strategy, respectively. It was shown that both methods were equivalent in terms of estimated protein concentrations of the most abundant proteins using MS/MS spectral count. They also resulted in the identification of many medium and low abundance proteins. Since more than 1000 proteins were identified from cellular extracts using similar iTRAQ technique by the same research group [36], the authors concluded that the presence of HAPs in human plasma prevents the MALDI ionization of weak concentrated peptides and thus a search for new strategies to deplete HAPs or enrich low abundant proteins is necessary for the discovery of biomarkers.

The CPLL technology has been very effective in mining other proteomes. Farinazzo *et al.* [37] reported a large number of previously unreported egg yolk proteins. The use of CPLL has doubled the number of proteins from 115 unique gene products [38] to 255 unique proteins species. Also, CPLL was reported by D'Amato *et al.* [39] to allow the expansion of the protein list of whey proteins from as little as 17 unique gene products to a total of 149 unique protein species 100 of which were not described in previous proteomics investigations. In this work, for instance, a polymorphic alkaline protein, which was identified as an immunoglobulin (a minor allergen) was only found after treatment with CPLL. Similarly, D'Ambrosio *et al.* [40] investigated two types of peptide ligand libraries (PLL), containing hexapeptides terminating either with a primary amine or modified with a terminal carboxyl group in mining the chicken egg while proteome whereby they were the first to report the most comprehensive list of proteins. They reported 148 unique protein species upon treating the egg white with the two PLL; 35 proteins were captured specifically by the amino-terminus PLL while 33 were specifically enriched by the carboxy-terminus PLL. Furthermore, Mouton-Barbosa *et al.* [41] demonstrated CPLL in the reduction of the dynamic range of protein concentration in CSF and unmasking previously undetected proteins by nano-LC-MS/MS analysis on an LTQ-Orbitrap MS. Using a large pool of CSF from different sources in order to better characterize the protein content of this fluid, the authors were able to identify 1212 proteins in which 745 species were only detected after CPLL treatment. Due to the limited sample volume of CSF and low protein concentration, the conventional use of CPLL with large volume columns for treatment of patient samples is definitely precluded. Therefore, the CPLL-based method was optimized by the authors to be compatible with low volume samples. Under the miniaturized protocol, the CPLL treatment was still efficient as far as the reduction of the dynamic range of protein concentration was concerned and led to an increase of more than 100% of the number of identified proteins in one LC-MS/MS run. Interestingly, the CPLL approach was extended to the extensive analysis of the cytoplasmic proteome of human erythrocytes in combination with advanced MS by Roux-Dalvai *et al.* [42]. In this work, the erythrocyte cytoplasmic proteome which is composed essentially from 98% hemoglobin and 2% of many other proteins the identity of

which are largely unexplored, was treated by CPLL that allowed simultaneously to lower the signal of hemoglobin and amplify the signal of the low to very low abundance proteins in the cytoplasm of human red blood cells. This double achievement by the CPLL was fully demonstrated *via* 2-DE maps and nano-LC-MS/MS analysis. The authors reported with high confidence 1578 proteins in the cytoplasmic fraction of a highly purified preparation of RBCs, thus allowing a deep exploration of the classical RBC pathways as well as the identification of unexpected minor proteins. This study also reported the identification of eight different hemoglobin chains including embryonic and newly discovered globin chains. An interesting proteomic study based on the use of the CPLL technology has involved the deep cytoplasmic proteome of spinach leaves [43] where the investigators identified 332 unique proteins of which 114 are in common with the control, untreated CPLL samples, 18 are present only in the control and 190 represent the new species representing low abundance proteins. Not to overlook, CPLL was investigated in the comprehensive identification of proteins in whole human saliva due to their critical diagnostic potentials [44]. The CPLL technology was coupled to a three-dimensional peptide fractionation process involving sequential steps of preparative IEF, SCX, and capillary RPC. This coupling allowed the identification of 2340 proteins thus representing the largest saliva proteomic dataset generated using a single analysis platform. Approximately 20% of total salivary proteins are also found in plasma, and proteins in both body fluids show comparable functional diversity and disease-linkage. Incubation of supernatants from cell cultures (also called conditioned media, CMs) with CPLL allowed to reduce the dynamic range of the samples and led to the identification of 3 times more proteins than in untreated CM samples [45]. This was facilitated by combining the protein equalization *via* CPLL treatment and metabolic labeling with deuterated amino acid. The metabolic labeling was essential to distinguish between cellular proteins and homologous bovine proteins contained in the medium. Furthermore, CPLL technology proved very effective in profiling human serum proteome (disease-free serum as well as serum samples with different pancreatic diseases) when combined with SELDI-TOF-MS [46, 47]. Finally, the comprehensive analysis of low-abundance proteins in human urinary exosomes using CPLL technology in combination with peptide OGE fractionation and nanoLC-chip-MS/MS has been reported by Zhang *et al.* [48]. The human exosomes are 30 – 100 nm vesicles that originate as the internal vesicles in multivesicular bodies from every renal epithelial cell type facing the urinary track. They are thought as a suitable noninvasive starting material for biomarker discovery relevant to a variety of renal disease. After analysis by nanoLC-chip-MS/MS, 512 proteins were identified including a large number of proteins with extreme molecular weight or extreme *pI* value, which could not be well profiled by 2-DE alone.

Interestingly, the CPLL technology proved useful in enriching and detecting proteins in alcoholic drink matrices such as wine [49, 50] and beers [51]. Protein additives in wine such as casein was enriched and detected at the level of 1 µg/L of wine [49, 50] while plenty of residual barely proteins (>20) and as many as 40 trace proteins of *Saccharomyces cerevisiae* were readily detected in beers [51].

2.3 Chromatographic and electrophoretic enrichment/pre-fractionation methods prior to LC-MS/MS

This section describes the chromatographic and electrophoretic fractionations used prior to LC-MS/MS. As such, the fractionation step is treated independently as a one-dimensional separation process whenever it involves the offline use of one chromatographic column or multiple columns not used in serial or sequential fashion. When the offline fractionation step involves passing the proteomic sample on two or more columns in serial or in tandem format, then the process is referred to as a multidimensional fractionation set up. Whenever both electrophoretic and chromatographic fractionations are compared in a given

investigation, the classification as offline chromatographic or electrophoretic fractionation is based on the most accomplishing or performing fractionation step whether chromatographic or electrophoretic.

2.3.1 Chromatographic pre-fractionation using one dimension—A hydrophobic method based on an octadecyl (C₁₈) absorbent to enrich the LMW proteins present in serum was introduced by Wu *et al.* [52]. Serum was treated with a C₁₈ absorbent having an average pore size of 100 Å under urea and DTT denaturing conditions. After incubation and washing, the LMW proteins were eluted using 60% v/v acetonitrile solution. The eluted LMW proteins, which were enriched more than 100 fold, were subjected to SDS-PAGE analysis, 2-DE and isotope-coded affinity tags (ICAT) labeling quantitation. Since, the procedure does not involve any addition of salts, desalting steps were avoided before analyzing the samples by 2-DE. Serum proteins at the ng/mL levels such as myeloid-related proteins could be identified directly from 2-DE coupled with MALDI-TOF/TOF MS and LTQ-Orbitrap MS. Furthermore, in the presence of the denaturing conditions, the protein-protein interactions are broken, thus allowing the identification of the proteins that were bound to the HMW proteins. Also, more than 110 LMW proteins were identified from the serum and the eluted LMW fraction contained only 5% of the HMW proteins. In addition to serum proteins, this method was also demonstrated in analyzing proteins in cell and tissue extracts [52].

In an LC based one dimensional fractionation prior to 2-D LC-MS/MS (shotgun) proteomic analysis of trypsin-digested human MCF10A cell sample, a concatenated low pH (pH 3) and high pH (pH 10) RPC fractionation approach of the tryptic peptides was performed as an off-line first dimension separation [53]. From the first dimension 60 fractions were collected using either high-pH or low-pH RPC. The concatenation was done by pooling early, middle and late RPC fractions of the 60 total collected fractions whereby for instance the fractions nos. 1, 16, 31 and 46 were combined together, and the fractions nos. 2, 17, 32 and 47 were also similarly combined and so on. When compared to the traditional SCX chromatography, the high pH concatenation resulted in an increase of the number of peptide and protein identified by 1.8 and 1.6 fold, respectively, whereas the low pH concatenation yielded results that were comparable to the SCX results. This indicates that concatenation strategy using high-pH RPC in the first dimension is a better alternative to the traditionally used SCX separation [53].

Using a rather very elaborate LC based one dimension preparative fractionation involving multiple columns, Tucholska *et al.* [54] attempted to identify the low abundance human serum proteins. Various types of chromatography resins including propyl sulfate, quaternary amine, diethylaminoethanol, cibachron blue, phenol Sepharose, carboxy methyl sepharose, hydroxyl apatite, heparin, concanavalin A (Con A) and protein G were used to fractionate the serum proteins prior to LC-ESI-MS/MS. The MS/MS spectra were correlated to some 4396 distinct sequences of the human forward RefSeq by X!TANDEM of which 61% have been detected by other algorithms, but 3129 (73%) were never previously reported from blood by X!TANDEM. The authors reported that the depletion of albumin or IgG was not necessary when using the multiple forms of protein chromatography prior to LC-ESI-MS/MS analysis.

An offline one dimensional separation step involving the use of electrostatic repulsion-hydrophilic interaction chromatography (ERHIC) in shotgun proteomics for the comprehensive profiling of the rat kidney proteome was reported by Hao *et al.* [55]. In the ERHIC separation step, the peptides were fractionated into forty-six fractions on an anion exchange column and then analyzed by RPC-MS/MS. This offline separation mode was compared to the traditionally practiced SCX-RPC-MS/MS. In terms of unique peptide

identification, it was observed that both methods were complementary to each other as they separate peptides based on different properties. But in terms of number of proteins identified, ERLIC-cum-RPC-MS/MS showed higher number of proteins as compared to SCX-RPC-MS/MS. Also, a higher number of basic and hydrophobic peptides were observed in the former than in the later approach. In total, 5499 proteins and 35 847 peptides of rat kidney tissues were characterized.

In a recent report [56], glyco- and phosphopeptides from mouse brain membrane digest were enriched simultaneously by an optimized protocol using ERHIC. This ERHIC protocol resulted in the identification of 544 unique glycoproteins and 922 glycosylation sites whereas the previously reported hydrazide chemistry method resulted in the identification of only 192 glycoproteins and 345 glycosylation sites. Also, 383 phosphoproteins and 915 phosphorylation sites were recovered from the sample subjected to ERHIC. It was concluded that this efficient method, which simultaneously enrich glyco- and phosphopeptides can be applied to other samples such as body fluids, cells or tissues.

A comparison between two fractionation strategies (SDS-PAGE and SCX) of protein complexes prior to LC-MS/MS analysis was reported by Das *et al.* [57]. The evaluation was based on the protein coverage obtained *via* LC-MS/MS analysis of the two complexes consisting of the nuclear proteins Bmi-1 and GATA3 that were expressed with the cells at low and high levels, respectively. Immuno-precipitated proteins from the two different nuclear proteins were either separated by SDS-PAGE at the protein level or separated by SCX chromatography on the peptide level. The SCX separation resulted in 3-fold increase in terms of number of proteins identified by LC-MS/MS when compared to the number of proteins using SDS-PAGE fractionation. SCX fractionation also showed an increase in the individual protein coverage especially for the proteins expressed at low levels.

As expected the nature of the first dimension of separation prior to LC-MS/MS has an important impact on the depth of proteomics profiling. This was demonstrated in a comparison of RPC, SDS-PAGE and SCX [58] in the fractionation of microgram quantities of *E. coli* protein extract. The overall performance and relative fractionation efficiencies of each technique were assessed by comparing the total number of proteins identified by each method. It was found that the protein-level RPC and the high pH RPC peptide-level separations performed the best, identifying 281 and 266 proteins, respectively. The online pH variance SCX and the SDS-PAGE yielded modest performances with 178 and 138 proteins identified, respectively. The offline SCX had the worst performance with 81 proteins identified.

In a study by Manadas *et al.* [59], comparison of three different fractionation steps for first dimensional separation of peptides obtained from the membrane fraction of mouse cortical brain tissue was performed. They were RPC at pH 10 (RP-Hp), SCX with a pH gradient (SCX-pG) and OGE fractionation. The fractions collected from the three separation methods were analyzed by RP-low pH-LC-MS/MS. The RP-Hp approach resulted in the identification of 308 proteins whereas SCX-pG and OGE resulted in 219 and 256 proteins, respectively. Gene ontology analysis indicated that RP-Hp approach was more suitable for membrane protein analysis when compared to the other two methods.

Dong *et al.* [60] reported a novel method that could detect phosphopeptides present in attomole concentration from a protein digest. The method was based on the enrichment of phosphopeptides from β -casein digest using a strong anion exchange (SAX) capillary monolithic column. The bound phosphopeptides from the column were directly eluted into the MALDI target using 5% formic acid and analyzed by MALDI-TOF MS. When

compared to IMAC and metal oxide affinity chromatography this strategy is more compatible with MALDI MS.

2.3.2 Multidimensional chromatographic pre-fractionation—In a study by Cellar *et al.* [61], two-dimensional LC separation was carried out for sample enrichment and fractionation of mammalian proteomics. In the first dimension, IgY immunodepletion column was used to deplete the HAPs from the sample. The HAPs deprived sample was then passed through a trap cartridge (two 3 mm × 8mm, 20 μm, C4, 4000 Å), which serves as an injection loop for the on-line second dimension that consisted of an analytical C18 column with large pores. This on-line arrangement facilitated (i) on-line desalting, (ii) automatic buffer exchange, (iii) facile concentration and (iv) the fractionation of the proteins based on their polarity. This advantageous method offers a convenient on-line proteomic approach as compared to the traditional immunodepletion process, which results in dilution of the depleted protein fractions.

In a different study by Jmeian and El Rassi [62], an integrated fluidic platform was introduced to deplete the HAPs from human serum and to fractionate/concentrate the medium and low abundance proteins. The HAPs were depleted using monolithic protein A, protein G' and antibody columns, whereas the fractionation/concentration was achieved by tandem IMAC/RPC columns. Briefly, the serum sample deprived of albumin was injected into the tandem series of protein G', protein A, and antibody columns (connected in the order: protein G', protein A, antihuman α1-antitrypsin, anti-human transferrin, anti-human haptoglobin and anti-human α2-macroglobulin) in order to deplete the top seven HAPs. This was followed by online concentration by IMAC, which consisted of monolithic stationary phases having surface bound iminodiacetic acid (IDA) chelated with Zn²⁺, Ni²⁺ and Cu²⁺. The IDA-metal columns as well as the RPC column were arranged in the order of decreasing specificity. That is in the order of IDA-Zn²⁺ → IDA-Ni²⁺ → IDA-Cu²⁺ → RPC. While IDA-Zn²⁺ requires 2 vicinal histidine (His) residues on the protein surface for binding, IDA-Ni²⁺ necessitates at least 2 His residues, and IDA-Cu²⁺ one His residue suffices for protein binding. The proteins that are not retained on the IDA-metal columns are then captured by the RPC column, which is the last column in the tandem series. The eluted fractions from the metal-chelate columns and the RPC column were further analyzed using 2-DE. 1450 protein spots were detected from the gel analysis using SYPRO fluorescent stain. Analysis of the detected gel spots stained by Coomassie resulted in the identification of 295 proteins by LC-MS/MS and MALDI-TOF analysis.

The above-discussed IMAC-RPC platform was further expanded by substituting the immuno-depletion with the protein equalizer technology to reduce the dynamic protein concentration range [63]. In this set up, the equalized proteins were further fractionated on IMAC columns and the RPC column arranged in tandem in the order shown in the preceding section. The analysis of the bound fractions using LC-MS/MS resulted in the identification of 82 non-redundant proteins. When compared to the immunodepletion-based platform described above, the protein equalizer technology based platform resulted in the identification of a larger number of proteins. Also, the strategy of subsequent fractionation on the 4 tandem columns (IMAC and RPC columns) after equalization allowed the identification of more proteins than simply using the equalization by ProteoMiner™. While the ProteoMiner™ technology is superior in terms of the overall number of captured proteins it only complements the immuno-subtraction approach since the latter can capture the proteins that were not captured by the former.

In a different report [64], phosphopeptide enrichment was done using SAX column in the first step and the phosphopeptides present in the flow-through were further enriched by offline titanium dioxide (TiO₂) chromatography. In this method the tryptic digests from

HeLa cells were analyzed. The SAX method alone had bias towards the acidic phosphopeptides, whereas the TiO₂ or SCX-TiO₂ methods had bias towards basic phosphopeptides. The authors concluded that the less biased SAX-TiO₂ method could be applied to separate and detect phosphopeptides from complex cell proteome.

An integrated platform for proteomic profiling was established by Yuan *et al.* [65], in which proteins were first separated by a microcolumn packed with mixed weak anion exchange and weak cation exchange (WAX/WCX) microparticles using a series of salt steps followed by online digestion by trypsin immobilized microenzymatic reactor. The peptide digests were first trapped and desalted by two parallel C8 precolumns and then separated and identified by online μ RPLC-ESI-MS/MS. When compared to an off-line analysis under the same conditions, the online platform had higher sequence coverage, good reproducibility, high throughput and less analysis time for a four-protein mixture consisting of myoglobin, cytochrome c, bovine serum albumin and α -casein. The integrated platform was applied to proteins extracted from human lung cancer cells, yielding 30% more identified proteins than that obtained by the shotgun approach while cutting the analysis time by a half.

2.3.3 Electrophoretic fractionation methods—The electrophoretic fractionation method traditionally used in proteomics has been 2-DE. Some of the drawbacks of 2-DE (e.g., difficulty to separate low abundance proteins, acidic and basic proteins, very large and very small proteins and hydrophobic proteins) have favored the use of other electrophoretic methods e.g., liquid-phase isoelectric focusing, free-flow electrophoresis (FFE) and OGE.

In a recent study by Wang *et al.* [66], two different approaches were compared using human cancer cell lysates. The first approach was GeLC-MS/MS approach, in which proteins were first separated using SDS-PAGE and analyzed using LC-MS/MS. The second approach consisted of micro-scale solution (Micro-sol) isoelectric focusing of the proteins, which was followed by SDS-PAGE separation and finally the gel slices from SDS-PAGE were analyzed using LC-MS/MS. In Micro-sol liquid phase isoelectric focusing, protein samples are divided into multiple tandem electrode chambers using isoelectric membranes. As expected, 90% of the proteins identified in the first approach were also found in the second approach. Also, the second approach resulted in 22% increase in the number of identified proteins indicating that the Micro-sol isoelectric focusing is an effective fractionation step before SDS-PAGE in the proteomics workflow.

Unlike other liquid-based fractionation methods, which involve a solid support, FFE is done in the absence of any solid support such as a gel. In FFE, an electric field is applied perpendicularly to the flow, which results in the separation of charged particles according to their electrophoretic mobility or isoelectric point. As an application of FFE, it was applied to identify the proteins present in urine sample [67]. The intact urine proteins were concentrated using ultrafiltration and isoelectric focusing *via* FFE to obtain approximately 50 fractions. These fractions were tryptically digested and analyzed using RPC-MS. This strategy of including the FFE step before MS analysis facilitated the in-depth analysis of urine proteomics.

In a very recent investigation by Krishnan *et al.* [68], the first dimension separation was based on OGE fractionation which allowed them to identify 1373 proteins from human platelet proteome. The second dimension analysis was achieved by nano-LC-MS/MS. In the OGE, the peptides migrate according to their *pI* values. After focusing, the samples were recovered from the sample wells and then loaded into the RPC column for MS/MS detection. This shotgun proteome analysis by OGE fractionation and RPC fractionation in the first and second dimension, respectively, allowed in-depth analysis of the human platelet proteome. Also, OGE was used in fractionating the iTRAQ labeled peptides resulting from

the tryptic digest of human plasma proteome, which was first either immunodepleted by MARS or equalized by CPLL [69]. Each of the 24 OGE fractions was then analyzed by nano-LC on a C18 column and subsequently identified by MALDI-MS/MS. The immunodepletion and CPLL were found complementary with regards to the number of identified proteins. The iTRAQ labeling in combination with OGE fractionation allowed more than 300 different proteins to be identified from 400 μ g plasma proteins.

Since isoelectric focusing (IEF), which is the traditional first step for 2-DE fractionation, has some limitations, a microfluidic free flow (FF)-IEF device was developed for the continuous separation of proteins into 24 fractions [70]. The technique resulted in identification and isolation of low abundance proteins since it was able to concentrate the proteins up to 10–20 fold. When compared with the standard 2-DE the FF-IEF device showed higher separation resolution and reduced experimental time. Also, the device has the advantage that it can deplete albumin or hemoglobin with high reproducibility and retain HMW proteins and provide higher yield of proteins in a broader pI range.

Free-flow zone electrophoresis of peptides and proteins was performed on polydimethylsiloxane (PDMS) chip [71]. A single and a two-step free-flow zone electrophoresis was developed which was an ampholyte-free process that can isolate molecules within a narrow pI range from a small sample volume. The sorting of the peptides/proteins were validated by both MALDI and electrospray injection MS. In the single-step process molecules were separated into neutral, high or low pI values at a given buffer pH value. In the two-step process, two sorting steps were coupled *via* off-line titration to isolate peptides within a specific pI range. This process had the advantages that by changing the buffer pH, the pI range of the fractionation can be adjusted. The authors concluded that further development of this device would significantly impact MS-based bioanalysis.

Tryptic peptides obtained from two distinct ovarian endometrioid tumor-derived cell lines (MDAH-2774 and TOV-112D) were first separated into 40 fractions *per* run by off-line capillary IEF (cIEF) [72]. Each fraction was further subjected to second dimensional separation by RP-HPLC coupled to a nanoESI-ion trap. It was observed that the cIEF had high separation resolution with little overlap between adjacent fractions. A total of 1749 proteins were identified from the two cell lines with an overlap of 1092 proteins. To obtain the dynamic changes between the two cell lines, out of 828 differentially expressed proteins 532 proteins were used to retrieve functions/pathways and some new pathways were uncovered. The authors concluded that the above-mentioned strategy could define the pathways involved in the development of cancer.

Selective enrichment of cysteine (Cys)-containing peptides from complex proteomes (e.g., yeast *Saccharomyces cerevisiae* and human colon carcinoma RKO cells) was demonstrated by Lin *et al.* [73]. A reactive biotinylated probe was used to enrich the Cys-peptides followed by capture with streptavidin capture and mild alkaline release, which yielded a highly purified population of Cys-peptides with a residual S-carboxymethyl tag. The Cys-peptides were then fractionated by isoelectric focusing (IEF) and analyzed by LC-MS/MS. The degree of enrichment typically was 2 – 8 fold but ranged up to almost 20-fold for a few proteins. The authors concluded that the selective enrichment of Cys-peptides could enhance the proteome coverage, especially Cys-rich lower abundance proteins.

Secreted proteins from CAPAN-2 human pancreatic cancer derived cells were collected after performing stable isotope labeling by amino acids in cell culture (SILAC). The resulting stable isotope labeled proteome (SILAP) standard contained < 0.5% of individual unlabeled proteins. To each sample of pooled sera from patients with early stage pancreatic

cancer or controls, an equal amount of the SILAP standard was added. Proteins were separated by IEF prior to 2D-LC-MS/MS analysis [74]. Using this SILAC approach relative quantification of proteins was done which resulted in 121 candidate biomarkers from a total of 1065 identified proteins.

2.4 Capturing/targeting sub-glycoproteomics by lectin affinity chromatography

The following sections discuss the different strategies used in capturing specific sub-glycoproteomics by LAC. The discussion cover the literature published in the last five years (2006 – present). Progress made before 2006 in isolating glycoproteins using lectin capture technology has been reviewed in ref. [75].

Alterations in the glycosylation patterns of proteins are very common in diseases. For example, alterations in sialylation, fucosylation, high mannose-type, sialyl Lewis x (sLe^x) structures and changes in the degree of glycan branching have been reported in cancer patients [76–78]. Therefore, the identification and determination of structures and functions of glycoproteins are of great values in clinical studies. Recent reviews on the clinical applications of glycoproteomics can be found in refs. [9, 79], which describe the different glycomic approaches used for the identification of cancer biomarkers.

In the following discussion, the term single lectin column refers to analyzing an aliquot of a given sample on each lectin column separately regardless of the number of lectin columns involved. Whereas serial lectin columns refers to the use of LAC in a sequential or serial fashion where the pass through fraction from one lectin column is analyzed on a second lectin column and so forth.

2.4.1 Single or serial lectin columns—In a study aiming at targeting sialic acid-rich glycoproteins in pancreatic cancer serum [80], three different lectins were used namely wheat germ agglutinin (WGA), *Sambucus nigra* agglutinin (SNA) and *Maackia amurensis* lectin (MAL). MAL column could select glycoproteins containing NeuAc-Gal-GlcNAc with sialic acid (NeuAc) at the 3 position of galactose (Gal) while SNA column binds preferentially to sialic acid attached to terminal Gal in α -2,6 and to a lesser extent α -2,3 linkage and WGA column can interact with some glycoproteins *via* sialic acid residues (e.g., glycoproteins containing sialylated bisected biantennary complex type glycans) as well as glycoproteins containing bisected hybrid type glycans. The serum sample was first depleted of the 12 HAPs using an IgY-12 antibody column, and the sialylated glycoproteins were subsequently enriched using the three immobilized lectins operated as single columns. The enriched fractions were further fractionated by RPC using non-polar nonporous silica stationary phases. To achieve enhanced speed, resolution and reproducibility, the RPC column was heated at 60°C. The fractions from the RPC column were further subjected to SDS-PAGE analysis, and after in gel digestion with trypsin the proteins were identified by μ LC-MS/MS. In a different route, after RPC fractionation, the fractions were tryptically digested, and analyzed by μ LC-ESI-TOF in order to generate a glycopeptide map. In a third route, the RPC fractions were tryptically digested and the *N*-glycans were cleaved by the endoglycosidase PNGase F, and subjected thereafter to MALDI-QIT (quadrupole ion trap – TOF). These concerted efforts permitted the finding that plasma protease C1 inhibitor and the N83 glycosylation of α -1-antitrypsin were down regulated in pancreatic cancer serum with respect to disease-free serum.

In another application [81] of LAC, fucosylated serum glycoproteins were targeted using *Aleuria aurantia* lectin (AAL) to identify biomarkers of primary hepatocellular carcinoma (HCC). Prior to LAC with agarose-bound AAL, the serum was depleted of the top 12 most HAPs using commercially available microbeads comprising of 12 types of IgY polyclonal antibodies. It was observed that a core fucosylated biantennary glycan (FcA2G2) increased

from 5.8 % in disease-free patients to 10% in HCC sera and to 8.5% in cirrhotic serum. Increases were also observed in the level of a core fucosylated biantennary glycan with two sites of outer-arm fucosylation (FcA2G2F2). This structure increased from 5.5% in the pooled disease-free sera to 6.2% in the pooled cirrhotic sera and to 9.6% in the pooled HCC sera. As revealed by the authors, these changes were accompanied by decreases in the level of the biantennary glycan. In summary, the presence of four fucosylated proteins in the sera of patients with HCC namely hemopexin, α -2-HS-glycoprotein, anti-1-antichymotrypsin and transferrin were further validated/confirmed using lectin fluorophore-linked immunosorbent assay.

Abbott *et al.* [82] demonstrated the enrichment of β (1,6)-branched *N*-linked glycan structures using L-phytohemagglutinin (L-PHA) lectin which has affinity towards the same. The β (1,6)-branched *N*-linked glycans have been reported to serve as marker in the detection of tumor progression [83]. Disease-free and breast cancer tissues were profiled by LAC using an L-PHA column. In brief, the delipidated tissue samples were treated with biotinylated L-PHA lectin to enrich the β (1,6)-branched *N*-linked glycoproteins. The L-PHA bound glycoproteins were then captured using streptavidin particles and the captured proteins were eluted with a mobile phase consisting of urea/DTT/ammonium bicarbonate. The enriched glycoproteins were then subjected to nanoelectrospray-MS/MS, which allowed the identification of 34 proteins that were found to be elevated in breast cancer tissue when compared to the diseased-free tissue.

In an effort to identify putative altered glycoprotein biomarkers for lung adenocarcinoma, Hongsachart *et al.* [84] developed a method in which initially the healthy and the lung cancer serum were screened using seven fluorescein isothiocyanate (FITC) labeled lectins for specific glycoprotein profile of the sera. Based on the screening results, WGA, which showed highest specific binding with the glycoproteins, was selected for the enrichment of glycoproteins from the sera. Following this, a co-immunoprecipitation of haptoglobin using anti-haptoglobin was performed. The removal of haptoglobin from the WGA bound sample allowed the identification of increased number of differentially expressed proteins, which otherwise would have been masked by haptoglobin. Further analysis was done using 2-DIGE, and it was noted that the unbound fraction from the WGA column mostly contained the HAPs such as albumin and IgG. Using this strategy, three up-regulated (adiponectin, ceruloplasmin and glycosylphosphatidylinositol-80) and two down-regulated glycoproteins (cyclin H and Fyn) in lung cancer serum relative to healthy serum were identified and they were further validated by Western blot analysis.

A tandem affinity approach, which combines affinity fractionation and immunoaffinity depletion was reported to identify low-abundance proteins in human plasma [85]. In this approach, the glycoproteins present in plasma were first enriched by LAC using WGA. In the second step, immunoaffinity depletion was carried out using the heavy-chain fraction of the antibodies that were raised in llama (*Lama glama*) against the proteins that were captured in the lectin enrichment step. By using this strategy, the authors were able to selectively enrich carcinoembryonic antigen that was spiked in disease-free serum by a factor of 600–800 fold.

In a study by Jung *et al.* [86], LAC was performed using lectins with broad and narrow specificity such as Con A, *Helix pomatia* agglutinin (HPA), *Lycopersicon esculentum* lectin (LEL), AAL and *Lens culinaris* agglutinin (LCA) to analyze the changes in protein concentration of breast cancer plasma in comparison to disease-free plasma. In this study, the quantification of the protein concentration was achieved with stable isotope coding. The glycoproteins that were enriched by each LAC columns were tryptically digested, fractionated on an RPC column and analyzed using MALDI-MS/MS. It was observed that

small groups of proteins increased in concentration by 3-fold or more in the breast cancer as compared to the disease-free plasma.

In another study by Mann *et al.* [87], a label-free quantitative analysis of fucosylated serum glycoproteins was carried out using the lectins AAL and *Lotus tetragonolobus* agglutinin (LTA). First, seven HAPs namely, albumin, IgG, α -1-antitrypsin, IgA, transferrin, haptoglobin and fibrinogen were depleted from the serum samples. Thereafter, a serial LAC was performed using agarose-bound AAL and LTA lectins. The glycoproteins thus enriched by LAC were combined and further fractionated using an RPC column. The RPC fractionated proteins were tryptically digested and analyzed by nano-LC-ESI-MS/MS. To ensure that the same amount of proteins was subjected to RPC fractionation, a bicinchoninic acid assay was performed. The quantification of the proteins was done by the summation of the peak areas of the identified peptides. The utility of this strategy has been demonstrated in identifying several potentially interesting glycoproteins as good candidates for future studies of serum markers for esophageal adenocarcinoma progression.

In a study by Durham and Regnier [88], serial LAC was performed to selectively capture *O*-glycosylated peptides from the tryptic digests of human serum. The peptides were passed through Con A affinity column to remove the high mannose, hybrid type and biantennary complex type *N*-glycans. By doing so the high mannose glycans, which have affinity towards jacalin (JAC) would be completely removed. Then the flow through was passed through a JAC affinity column to capture the *O*-glycosylated peptides alone. The parent glycoprotein of the glycopeptides was identified by deglycosylation of *O*-glycopeptides by oxidative elimination. These peptides were further fractionated by RPC and analyzed using ESI-MS/MS. The approach was applied to fetuin as a model protein and all the *O*-glycosylated sites present in it were identified. Also, the *O*-glycosylated sites in over thirty glycoproteins from human serum were identified. It was concluded that this strategy allowed for enrichment of glycopeptides from complex samples and identification of *O*-glycosylation sites in proteins. However, the authors recognized that the removal of the glycans from peptides might result in loss of structural information that might be of biological significance.

Drake *et al.* [89] performed LAC on the tryptic digests of Multiple Affinity Removal System (MARS) 14-depleted human plasma. The lectins SNA and AAL were used to specifically capture the sialylated and the fucosylated glycopeptides as elevated levels of sialylation and fucosylation are observed in cancer plasma. Human lactoferrin served as positive control for fucosylated and sialylated glycopeptides while the high-mannose structures from yeast invertase served as negative controls. These two protein standards were spiked into the depleted human plasma. The human plasma peptides were then treated with *N*-glycosidase F to remove the *N*-linked glycans, and the resulting peptides were analyzed by ESI HPLC-MS/MS. The total number of glycoproteins identified was 122 containing 247 unique glycosites. Several of the identified glycoproteins were at the levels of ng/mL of plasma. It was concluded that the combination of LAC and MS constitutes a workflow that could be utilized into cancer biomarker discovery pipelines.

Serial LAC consisting of JAC, PNA and PHA immobilized lectin columns proved effective in capturing *O*-glycoproteome from *Drosophila melanogaster* S2 cells [90]. In this work, mucin-type *O*-glycosylation in S2 cells grown under serum-free conditions was shown to be restricted to the Tn-antigen (GalNAc α -Ser/Thr) and the T-antigen ((Gal β 1-3GalNAc α -Ser/Thr), a structural homogeneity that enables unique glycoproteomics strategies. This investigation involved the use of a label-free strategy for the isolation, profiling and analysis of *O*-glycosylated and consisted of serial LAC capture, 2-DE-based glycoprotein analysis by *O*-glycan specific mAbs and protein identification by MALDI-MS. Protein identity and *O*-

glycosylation was confirmed by ESI-MS/MS with detection of diagnostic sugar oxonium-ion fragments. 2-D reference maps and identified 21 secreted and intracellular mucin-type *O*-glycoproteins were established by the authors. The results showed that *Drosophila* S2 cells express *O*-glycoproteins involved in a wide range of biological functions including proteins of the extracellular matrix (laminin γ -chain, peroxidasin and glutactin), pathogen recognition proteins (Gnbp1), stress response proteins (glycoprotein 93), secreted proteases (matrix-metalloprotease 1 and various trypsin-like serine proteases), protease inhibitors (serpin 27A) and proteins of unknown functions.

2.4.2 Multi-lectin affinity chromatography (M-LAC)—In M-LAC, mixture of lectins having complementary specificities for different glycosylation are immobilized in a given column. After loading the sample onto the column, the M-LAC column is eluted sequentially using specific displacer for each lectin, i.e., the haptenic sugar. Due to the high complexity of serum that is usually amplified by the presence of HAPs, some of the approaches discussed below integrated the process of depletion of HAPs in order to facilitate the identification of the glycoproteins present in serum.

Pursuant to the initial studies on M-LAC, which involved lectins crosslinked to soft agarose support [91, 92], Hancock and co-workers reported an improved M-LAC the so-called high performance HP-M-LAC [93] for fast chromatographic affinity separations, which involved the immobilization of lectins to a styrene-divinylbenzene support matrix coated with a crosslinked polyhydroxylated polymer (POROS) with an active aldehyde functional group. A high lectin density (15 mg/mL of beads) was found to be optimal for the binding of glycoproteins from human plasma. The HP-M-LAC, which consisted of Con A, WGA and JAC was evaluated in plasma proteomics/glycoproteomics. Approximately 120 proteins were identified from the crude plasma and of these identified proteins 75% were glycoproteins. The affinity adsorbent had comparable binding properties and similar oligosaccharide specificity when compared to the previous study [91], but with better protein recoveries.

In another study by Yang *et al.* [94], glycoproteins were selectively captured from breast cancer and disease-free sera by M-LAC. The lectins used in this study were agarose-bound Con A, WGA and JAC. The proteins from the lectin bound fractions were identified by MS/MS analysis. The low-abundant breast cancer biomarker HER-2, which was not detected without the M-LAC step, was identified multiple times by MS/MS sequencing in 10 non-pooled breast cancer samples. An enzyme linked immunosorbent assay (ELISA) was performed to confirm the identification of HER-2. Gene ontology program was used for characterization of a large panel of candidate markers.

In a study by Plavina *et al.* [95], albumin and IgG were first depleted from the plasma sample followed by the enrichment of the glycoproteins using M-LAC (i.e., combining agarose-bound lectins in a 1:1:1 [v/v/v] ratio including Con A, WGA and JAC) and the identification of proteins using nano-LC-MS/MS. Also, a comparison of M-LAC with and without the depletion step was made, and it was shown that the total number of identified proteins (in bound and unbound fractions) increased from 120 to 191 by including the depletion step. To demonstrate the ability of this method, it was applied to biomarker discovery from psoriasis samples. It was observed that 11 proteins had different concentrations between the control and psoriasis plasma samples, and the protein galectin-3 binding protein was further validated using ELISA. The authors concluded that the combination of depletion of HAPs with M-LAC allowed the in-depth analysis of those proteins, which had concentrations of 10–100 ng/mL.

In another application of M-LAC by Hancock and co-workers [96], changes in breast cancer serum were identified using the three lectins mentioned above namely Con A, WGA and JAC, which were mixed in a ratio of 1:1:1 and packed into a column. Some of the HAPs like albumin, IgM, IgA, and IgG were depleted while other HAPs such as α -1-antitrypsin, transferrin and haptoglobin were not depleted as they might be involved in cancer related changes. The depleted serum was then subjected to M-LAC and the bound fractions were further analyzed using three different orthogonal analytical platforms to identify glycoproteins that had either concentration or glycan structure changes due to the breast cancer. In the first platform, 1D SDS-PAGE analysis was performed and three different detection methods namely, Coomassie blue staining, fluorescent staining of the glycoproteins and lectin blotting with biotinylated SNA were carried out. In the second platform the proteins were fractionated based on their pI values using digital ProteomeChip (dPC) to identify the breast cancer proteins, which show a change in their pI values. In the third platform, a lectin-antibody sandwich microarray was performed using AAL to detect the neutral glycan structure changes in the breast cancer serum. By all the three platforms the authors identified complement C3 beta chain, α -1-antitrypsin, transferrin and α -1B-glycoprotein as potential glycoproteins for further studies in breast cancer human serum.

In a very recent report from the same laboratory as above [97], an automated HPLC platform was introduced to remove HAPs and to fractionate glycoproteins using immuno-affinity depletion and M-LAC, respectively, in order to facilitate the identification of the breast cancer associated serum biomarkers. The collected glycoproteins obtained from M-LAC were further subjected to isoelectric focusing separation using a dPC, which had the operating range of pH 4.20~6.20 and 6.00~8.00. The gel plugs from the dPC were combined to get 10 fractions and these were subjected to in-gel digestion and LC-MS analysis. It was concluded that the inclusion of the isoelectric focusing using dPC fractionation after M-LAC extended the dynamic range of serum proteome and also resulted in the identification of low abundance proteins with higher sequence coverage. Finally, the proteins thrombospondin-1 and 5, α -1B-glycoprotein, serum amyloid P-component and tenascin-X were selected as promising candidates to analyze breast cancer serum.

Madera *et al.* [98] used 4 different lectins immobilized on macroporous silica for the enrichment of glycoproteins from serum. The immobilized lectins were Con A, SNA, *Ulex europaeus* agglutinin-I (UEA-I) and *Phaseolus vulgaris* agglutinin-L (PHA-L). Two different enrichment strategies were evaluated: serial LAC and M-LAC. In both cases, the authors performed off-line RPC pre-fractionation and subsequent protein identification by LC-MS/MS. While serial LAC permitted the identification of 108 proteins in the lectin bound fractions spanning a concentration dynamic range of 7–10 orders of magnitude, the M-LAC approach allowed the identification of only 67 proteins. This is not surprising since M-LAC, which is a mixed bed approach may introduce a lowering in the enrichment capacity of each lectin due to the fact that lectins are glycoproteins (except WGA, Con A and PNA) and may undergo bead to bead interactions among the various immobilized lectins.

3 Concluding remarks

This review article has summarized the relevant advances in the field of sample preparation and fractionation of proteomics samples. Despite the major progress that has been made over the last 3 years in this area of research, improved and automated platforms for sample preparation and fractionation of complex proteomics samples will be anticipated in order to realize more reliable and reproducible sample preparation and fractionation that would facilitate deeper proteomics profiling.

Acknowledgments

We gratefully acknowledge the financial support by the grant no. IR15GM096286-01 from the Department of Health and Human Services at the National Institute of Health.

Nonstandard Abbreviations

AAL	<i>Aleuria aurantia</i> lectin
Con A	concanavalin A
CPLL	combinatorial peptide ligand library
DIGE	differential gel electrophoresis
HAPs	high abundance proteins
HMW	high molecular weight
HPA	<i>Helix pomatia</i> agglutinin
LMW	low molecular weight
IMAC	immobilized metal affinity chromatography
JAC	jacalin
LAC	lectin affinity chromatography
LCA	<i>Lens culinaris</i> agglutinin
LEL	<i>Lycopersicon esculentum</i> lectin
LTA	<i>Lotus tetragonolobus</i> agglutinin
MAL	<i>Maackia amurensis</i> lectin
OGE	OFFgel electrophoresis
L-PHA	L-phytohemagglutinin lectin
PHA-L	<i>Phaseolus vulgaris</i> agglutinin-L
PNA	peanut agglutinin
SAX	strong anion exchange
SCX	strong cation exchange
SNA	elderberry lectin or <i>Sambucus nigra</i> agglutinin
UEA-I	<i>Ulex europaeus</i> agglutinin-I

References

1. Jmeian Y, El Rassi Z. Electrophoresis. 2009; 30:249–261. [PubMed: 19101934]
2. Righetti PG, Fasoli E, Boschetti E. Electrophoresis. 2011; 32:960–966. [PubMed: 21449065]
3. Ahmed FE. J Sep Sc. 2009; 32:771–7798. [PubMed: 19219839]
4. Fang X, Zhang WW. J Proteomics. 2008; 71:284–303. [PubMed: 18619565]
5. Ly L, Wasinger VC. Proteomics. 2011; 11:513–534. [PubMed: 21241016]
6. Matt P, Fu Z, Fu Q, Van Eyk JE. Physiol Genomics. 2008; 33:12–17. [PubMed: 18162500]
7. Hirabayashi J, Kuno A, Tateno H. Electrophoresis. 2011; 32:1118–1128. [PubMed: 21544837]
8. Kim YS, Yoo HS, Ko JH. Protein & Peptide Lett. 2009; 16:499–507.
9. Taylor AD, Hancock WS, Hincapie M, Taniguchi N, Hanash SM. Genome Med. 2009; 1:article no. 57.

10. Granger J, Siddiqui J, Copeland S, Remick D. *Proteomics*. 2005; 5:4713–4718. [PubMed: 16281180]
11. de Roos B. *Exp Rev Proteomics*. 2008; 5:819–826.
12. Kawashima Y, Fukutomi T, Tomonaga T, Takahashi H, Nomura F, Maeda T, Koda Y. *J Proteome Res*. 2010; 9:1694–1705. [PubMed: 20184378]
13. Warder SE, Tucker LA, Strelitzer TJ, McKeegan EM, Meuth JL, Jung PM, Saraf A, Singh B, Lai-Zhang J, Gagne G, Rogers JC. *Anal Biochem*. 2009; 387:184–193. [PubMed: 19454248]
14. Tu C, Rudnick PA, Martinez MY, Cheek KL, Stein SE, Slebos RJ, Liebler DC. *J Proteome Res*. 2010; 9:4982–4991. [PubMed: 20677825]
15. Vaezzadeh AR, Briscoe AC, Steen H, Lee RS. *J Proteome Res*. 2010; 9:6082–6092. [PubMed: 20923230]
16. Roche S, Tiers L, Provansal M, Seveno M, Piva MT, Jouin P, Lehmann S. *J Proteomics*. 2009; 72:945–951. [PubMed: 19341827]
17. Lin B, White JT, Wu J, Lele S, Old LJ, Hood L, Odunsi K. *Proteomics Clin Appl*. 2009; 3:853–861. [PubMed: 20559449]
18. Qian WJ, Kaleta DT, Petritis BO, Jiang H, Liu T, Zhang X, Mottaz HM, Varnum SM, Camp DG 2nd, Huang L, Fang X, Zhang WW, Smith RD. *Mol Cell Proteomics*. 2008; 7:1963–1973. [PubMed: 18632595]
19. Gundry RL, White MY, Noguee J, Tchernyshyov I, Van Eyk JE. *Proteomics*. 2009; 9:2021–2028. [PubMed: 19294703]
20. Polaskova V, Kapur A, Khan A, Molloy MP, Baker MS. *Electrophoresis*. 2010; 31:471–482. [PubMed: 20119956]
21. Wetterhall M, Zuberovic A, Hanrieder J, Bergquist J. *J Chromatogr B*. 2010; 878:1519–1530.
22. Mahn A, Reyes A, Zamorano M, Cifuentes W, Ismail M. *J Chromatogr B*. 2010; 878:1038–1044.
23. Simo C, Bachi A, Cattaneo A, Guerrier L, Fortis F, Boschetti E, Podtelejnikov A, Righetti PG. *Anal Chem*. 2008; 80:3547–3556. [PubMed: 18399644]
24. Restuccia U, Boschetti E, Fasoli E, Fortis F, Guerrier L, Bachi A, Kravchuk AV, Righetti PG. *J Proteomics*. 2009; 72:1061–1070. [PubMed: 19580890]
25. Sihlbom C, Kanmert I, Bahr Hv, Davidsson P. *J Proteome Res*. 2008; 7:4191–4198. [PubMed: 18690747]
26. Léger T, Lavigne D, Le Caër JP, Guerrier L, Boschetti E, Fareh J, Feldman L, Laprévote O, Meilhac O. *Clinica Chimica Acta*. 2011; 412:740–747.
27. Tu C, Li J, Young R, Page BJ, Engler F, Halfon MS, Canty JM, Qu J. *Anal Chem*. 2011; 83:4802–4813. [PubMed: 21491903]
28. De Bock M, de Seny D, Meuwis MA, Servais AC, Minh TQ, Closset J, Chapelle JP, Louis E, Malaise M, Merville MP, Fillet M. *Talanta*. 2010; 82:245–254. [PubMed: 20685463]
29. Bandow JE. *Proteomics*. 2010; 10:1416–1425. [PubMed: 20127685]
30. Di Girolamo F, Bala K, Chung MCM, Righetti PG. *Electrophoresis*. 2011; 32:976–980. [PubMed: 21449074]
31. Candiano G, Dimuccio V, Bruschi M, Santucci L, Gusmano R, Boschetti E, Righetti PG, Ghigerri GM. *Electrophoresis*. 2009; 30:2405–2411. [PubMed: 19593750]
32. Fasoli E, Farinazzo A, Sun CJ, Kravchuk AV, Guerrier L, Fortis F, Boschetti E, Righetti PG. *J Proteomics*. 2010; 73:733–742. [PubMed: 19879985]
33. Beseme O, Fertin M, Drobecq H, Amouyel P, Pinet F. *Electrophoresis*. 2010; 31:2697–2704. [PubMed: 20665525]
34. Dwivedi RC, Krokhlm OV, Cortens JP, Wilkins JA. *J Proteome Res*. 2010; 9:1144–1149. [PubMed: 20020779]
35. Ernault E, Bourreau A, Gamelin E, Guette C. *J Biomed Biotechnol*. 2010:927917. [PubMed: 19888438]
36. Ernault E, Gamelin E, Guette C. *Proteome Sci*. 2008; 6:article 27.
37. Farinazzo A, Restuccia U, Bachi A, Guerrier L, Fortis F, Boschetti E, Fasoli E, Citterio A, Righetti PG. *J Chromatogr A*. 2009; 1216:1241–1252. [PubMed: 19081572]

38. Mann K, Mann M. *Proteomics*. 2008; 8:178–191. [PubMed: 18046696]
39. D'Amato A, Bachi A, Fasoli E, Boschetti E, Peltre G, Sénéchal H, Righetti PG. *J Proteome Res*. 2009; 8:3925–3936. [PubMed: 19499900]
40. D'Ambrosio C, Arena S, Scaloni A, Guerrier L, Boschetti E, Mendieta ME, Citterio A, Righetti PG. *J Proteome Res*. 2008; 7:3461–3474. [PubMed: 18570458]
41. Mouton-Barbosa E, Roux-Dalvai F, Bouyssié D, Berger F, Schmidt E, Righetti PG, Guerrier L, Boschetti E, Burlet-Schiltz O, Monsarrat B, de Peredo AG. *Mol Cell Proteomics*. 2010; 9:1006–1021. [PubMed: 20093276]
42. Roux-Dalvai F, de Peredo AG, Simo C, Guerrier L, Bouyssié D, Zanella A, Citterio A, Burlet-Schiltz O, Boschetti E, Righetti PG, Monsarrat B. *Mol Cell Proteomics*. 2008; 7:2254–2269. [PubMed: 18614565]
43. Fasoli E, D'Amato A, Kravchuk AV, Boschetti E, Bachi A, Righetti PG. *J Proteomics*. 2011; 74:127–136. [PubMed: 21056706]
44. Bandhakavi S, Stone MD, Onsongo G, Van Riper SK, Griffin TJ. *J Proteome Res*. 2009; 8:5590–5600. [PubMed: 19813771]
45. Colzani M, Waridel P, Laurent J, Faes E, Rüegg C, Quadroni M. *J Proteome Res*. 2009; 8:4779–4788. [PubMed: 19711906]
46. Fakelman F, Felix K, Büchler MW, Werner J. *Arch Physiol Biochem*. 2010; 116:208–217. [PubMed: 20815689]
47. Fröbel J, Hartwig S, Paßlack W, Eckel J, Hass R, Czibere A, Lehr S. *Arch Physiol Biochem*. 2010; 116:174–180. [PubMed: 20662595]
48. Zhang Y, Li Y, Qiu F, Qiu Z. *Electrophoresis*. 2010; 31:3797–3807. [PubMed: 21082674]
49. Cereda A, Kravchuk AV, D'Amato A, Bachi A, Righetti PG. *J Proteomics*. 2010; 73:1732–1739. [PubMed: 20510390]
50. D'Amato A, Kravchuk AV, Bachi A, Righetti PG. *J Proteomics*. 2010; 73:2370–2377. [PubMed: 20813213]
51. Fasoli E, Aldini G, Regazzoni L, Kravchuk AV, Citterio A, Righetti PG. *J Proteome Res*. 2010; 9:5262–5269. [PubMed: 20722451]
52. Wu J, An Y, Pu H, Shan Y, Ren X, An M, Wang Q, Wei S, Ji J. *Anal Biochem*. 2010; 398:34–44. [PubMed: 19891953]
53. Wang Y, Yang F, Gritsenko MA, Clauss T, Liu T, Shen Y, Monroe ME, Lopez-Ferrer D, Reno T, Moore RJ, Klemke RL, Camp DG 2nd, Smith RD. *Proteomics*. 2011; 11:2019–2026. [PubMed: 21500348]
54. Tucholska M, Bowden P, Jacks K, Zhu P, Furesz S, Dumbrovsky M, Marshall J. *J Proteome Res*. 2009; 8:1143–1155. [PubMed: 19265436]
55. Hao P, Guo T, Li X, Adav SS, Yang J, Wei M, Sze SK. *J Proteome Res*. 2010; 9:3520–3526. [PubMed: 20450224]
56. Zhang H, Guo T, Li X, Datta A, Park JE, Yang J, Lim SK, Tam JP, Sze SK. *Mol Cell Proteomics*. 2010; 9:635–647. [PubMed: 20047950]
57. Das S, Bosley AD, Ye X, Chan KC, Chu I, Green JE, Issaq HJ, Veenstra TD, Andersson T. *J Proteome Res*. 2010; 9:6696–6704. [PubMed: 20968308]
58. Dowell JA, Frost DC, Zhang J, Li L. *Anal Chem*. 2008; 80:6715–6723. [PubMed: 18680313]
59. Manadas B, English JA, Wynne KJ, Cotter DR, Dunn MJ. *Proteomics*. 2009; 9:5194–5198. [PubMed: 19771557]
60. Dong M, Wu M, Wang F, Qin H, Han G, Dong J, Wu R, Ye M, Liu Z, Zou H. *Anal Chem*. 2010; 82:2907–2915. [PubMed: 20199055]
61. Cellar NA, Karnoup AS, Albers DR, Langhorst ML, Young SA. *J Chromatogr B*. 2009; 877:79–85.
62. Jmeian Y, El Rassi Z. *J Proteome Res*. 2009; 8:4592–4603. [PubMed: 19670910]
63. Selvaraju S, El Rassi Z. *Electrophoresis*. 2010; 32:674–685. [PubMed: 21365658]
64. Nie S, Dai J, Ning ZB, Cao XJ, Sheng QH, Zeng R. *J Proteome Res*. 2010; 9:4585–4594. [PubMed: 20681634]

65. Yuan H, Zhang L, Hou C, Zhu G, Tao D, Liang Z, Zhang Y. *Anal Chem.* 2009; 81:8708–8714. [PubMed: 19788244]
66. Wang H, Chang-Wong T, Tang HY, Speicher DW. *J Proteome Res.* 2010; 9:1032–1040. [PubMed: 20014860]
67. Foucher AL, Craft DR, Gelfand CA. *Methods Mol Biol.* 2010; 641:27–45. [PubMed: 20407940]
68. Krishnan S, Gaspari M, Della Corte A, Bianchi P, Crescente M, Cerletti C, Torella D, Indolfi C, de Gaetano G, Donati MB, Rotilio D, Cuda G. *Electrophoresis.* 2011; 32:686–695. [PubMed: 21337587]
69. Ernoul E, Bourreau A, Gamelin E, Guette C. *J Biomed Biotechnol.* 2009;927917. [PubMed: 19888438]
70. Wen J, Wilker EW, Yaffe MB, Jensen KF. *Anal Chem.* 2010; 82:1253–1260. [PubMed: 20092256]
71. Song YA, Chan M, Celio C, Tannenbaum SR, Wishnok JS, Han J. *Anal Chem.* 2010; 82:2317–2325. [PubMed: 20163146]
72. Dai L, Li C, Shedden KA, Misek DE, Lubman DM. *Electrophoresis.* 2009; 30:1119–1131. [PubMed: 19288585]
73. Lin D, Li J, Slebos RC, Liebler DC. *J Proteome Res.* 2010; 9:5461–5472. [PubMed: 20731415]
74. Yu KH, Barry CG, Austin D, Busch CM, Sangar V, Rustgi AK, Blair IA. *J Proteome Res.* 2009; 8:1565–1576. [PubMed: 19199705]
75. Morelle W, Canis K, Chirat F, Faid V, Michalski JC. *Proteomics.* 2006; 6:3993–4015. [PubMed: 16786490]
76. Brinkman-van der Linden EC, de Haan PF, Havenaar EC, van Dijk W. *Glycoconj J.* 1998; 15:177–182. [PubMed: 9557878]
77. Dage JL, Ackermann BL, Halsall HB. *Glycobiology.* 1998; 8:755–760. [PubMed: 9639536]
78. Saldova R, Royle L, Radcliffe CM, Abd Hamid UM, Evans R, Arnold JN, Banks RE, Hutson R, Harvey DJ, Antrobus R, Petrescu SM, Dwek RA, Rudd PM. *Glycobiology.* 2007; 17:1344–1356. [PubMed: 17884841]
79. Tian Y, Zhang H. *Proteomics Clin Appl.* 2010; 4:124–132. [PubMed: 21137038]
80. Zhao J, Simeone DM, Heidt D, Anderson MA, Lubman DM. *J Proteome Res.* 2006; 5:1792–1802. [PubMed: 16823988]
81. Comunale MA, Wang M, Hafner J, Krakover J, Rodemich L, Kopenhaver B, Long RE, Junaidi O, Bisceglie AM, Block TM, Mehta AS. *J Proteome Res.* 2009; 8:595–602. [PubMed: 19099421]
82. Abbott KL, Aoki K, Lim JM, Porterfield M, Johnson R, O'Regan RM, Wells L, Tiemeyer M, Pierce M. *J Proteome Res.* 2008; 7:1470–1480. [PubMed: 18271524]
83. Fernandes B, Sagman U, Auger M, Demetrio M, Dennis JW. *Cancer Res.* 1991; 51:718–723. [PubMed: 1985789]
84. Hongsachart P, Huang-Liu R, Sinchaikul S, Pan FM, Phutrakul S, Chuang YM, Yu CJ, Chen ST. *Electrophoresis.* 2009; 30:1206–1220. [PubMed: 19294700]
85. Mortezaei N, Harder S, Schnabel C, Moors E, Gauly M, Schluter H, Wagener C, Buck F. *J Proteome Res.* 2010; 9:6126–6134. [PubMed: 20839810]
86. Jung K, Cho W, Regnier FE. *J Proteome Res.* 2009; 8:643–650. [PubMed: 19099503]
87. Mann B, Madera M, Klouckova I, Mechref Y, Dobrolecki LE, Hickey RJ, Hammoud ZT, Novotny MV. *Electrophoresis.* 2010; 31:1833–1841. [PubMed: 20446296]
88. Durham M, Regnier FE. *J Chromatogr A.* 2006; 1132:165–173. [PubMed: 16919642]
89. Drake PM, Schilling B, Niles RK, Braten M, Johansen E, Liu H, Lerch M, Sorensen DJ, Li B, Allen S, Hall SC, Witkowska HE, Regnier FE, Gibson BW, Fisher SJ. *Anal Biochem.* 2011; 408:71–85. [PubMed: 20705048]
90. Schwientek T, Mandel U, Roth U, Müller S, Hanisch FG. *Proteomics.* 2007; 7:3264–3277. [PubMed: 17708590]
91. Yang Z, Hancock WS. *J Chromatogr A.* 2004; 1053:79–88. [PubMed: 15543974]
92. Yang Z, Hancock WS. *J Chromatogr A.* 2005; 1070:57–64. [PubMed: 15861788]
93. Kullolli M, Hancock WS, Hincapie M. *J Sep Sc.* 2008; 31:2733–2739. [PubMed: 18693314]

94. Yang Z, Harris LE, Palmer-Toy DE, Hancock WS. Clin Chem. 2006; 52:1897–1905. [PubMed: 16916992]
95. Plavina T, Wakshull E, Hancock WS, Hincapie M. J Proteome Res. 2007; 6:662–671. [PubMed: 17269723]
96. Zeng Z, Hincapie M, Haab BB, Hanash S, Pitteri SJ, Kluck S, Hogan JM, Kennedy J, Hancock WS. J Chromatogr A. 2010; 1217:3307–3315. [PubMed: 19782370]
97. Zeng Z, Hincapie M, Pitteri SJ, Hanash S, Schalkwijk J, Hogan JM, Wang H, Hancock WS. Anal Chem. 2011; 83:4845–4854. [PubMed: 21513341]
98. Madera M, Mechref Y, Klouckova I, Novotny MV. J Chromatogr B. 2007; 845:121–137.