Multidimensional Separations-Based Shotgun Proteomics

Marjorie L. Fournier,† Joshua M. Gilmore,† Skylar A. Martin-Brown,† and Michael P. Washburn*

Stowers Institute for Medical Research, Kansas City, Missouri 64110

Received January 25, 2007

Contents

* Address correspondence to this author at the Stowers Institute for Medical Research, 1000 E. 50th St., Kansas City, MO 64110 [telephone (816) 926-4457; fax (816) 926-4694; e-mail mpw@stowers-institute.org].
† These authors contributed equally to this work.

1. Introduction

The proteome refers to the collection of proteins in a given biological organism or system under a particular set of environmental conditions. The study of proteins, referred to as proteomics, is performed to identify the components of a particular proteome and analyze global changes in protein expression in response to different stimuli. This leads to an understanding of physiological and pathological states of an organism through a comprehensive analysis of biological processes. From a protein complex to whole cell, proteome analysis deals with highly complex mixtures, requiring more than one analytical dimension to achieve the high resolving power necessary for reliable analysis.

In 1975, O'Farrell and Klose described two-dimensional polyacrylamide gel electrophoresis (2D-PAGE)[1,2] that could resolve complex protein mixtures into thousands of spots. Years later, upon the development of matrix-assisted laser desorption ionization (MALDI)[3] and electrospray ionization (ESI),[4,5] combined with database searching,[6−9] the field of proteomics began to grow dramatically. Researchers were able to characterize complex mixtures of proteins and gain novel biological insights. Despite the longstanding success of 2D-PAGE coupled with mass spectrometry, several fundamental issues with the technology, including the challenges of identifying low-abundance proteins,[10−12] membrane proteins,[13] and proteins with extremes in isoelectric point (p*I*) and molecular weight (MW),[14,15] drove researchers to develop alternative approaches for the separation of complex mixtures. This led to the emergence of shotgun proteomics based on the coupling of high-performance liquid chromatography (HPLC) and mass spectrometry (MS). Similar to the shotgun genomic sequencing approach in which DNA is broken into smaller pieces prior to sequencing and reassembled in silico, proteins are first digested into peptides and then analyzed by multidimensional chromatography coupled to tandem mass spectrometry (MS/MS). Thousands of tandem mass spectra are then compared to theoretical tandem mass spectra using database searching algorithms for the identification of proteins in the sample (Figure 1b).

The inability of one-dimensional (1D) separation techniques to resolve complex biological samples for shotgun proteomics has required the development of multidimensional separation methods. A multidimensional separation includes two or more independent separation techniques (i.e., ion exchange, size exclusion, reversed phase, and affinity) coupled together for the analysis of a single sample.

10.1021/cr068279a CCC: $65.00
Published on Web 07/25/2007

Marjorie L. Fournier (far right) earned her Ph.D. degree in 2004 in Molecular Microbiology and Biotechnology from the Aix-Marseille University, France. Her Ph.D. research, conducted under the direction of Dr. Alain Dolla, was focused on the functional analysis of oxidative stress in anaerobic microorganisms by genetics and mass spectrometry-based proteomics. She is currently a postdoctoral research associate in Dr. Michael Washburn's laboratory at the Stowers Institute for Medical Research (Kansas City, MO). Her research is focused on using multidimensional separation-based quantitative proteomics to elucidate drug molecular mechanisms via dynamics of protein complexes and large-scale temporal analyses.

Joshua Gilmore (second from right) earned his B.S. degree in genetics from the University of Kansas in 2000 and his Ph.D. degree in biochemistry/biophysics from the University of Kansas in 2005. His Ph.D. research was conducted with Professor Jeffrey Urbauer and focused on the structural and functional characterization of bacterial anti-sigma factors. Prior to earning his Ph.D., he worked as a research assistant focusing on overexpression, purification, and biochemical analyses of proteins. He is currently a postdoctoral research associate in Dr. Michael Washburn's laboratory at the Stowers Institute for Medical Research. His research focuses on investigating protein–protein interactions and identifying post-translational modifications using multidimensional protein identification technology.

Skylar Martin-Brown (second from left) received her B.S. degree in chemistry from Truman State University in 2000. She earned her M.S. degree in chemistry from Kansas State University in 2004 under Dr. Maryanne M. Collinson, where her research focused on the examination of diffusion in sol–gel-derived thin films by single-molecule spectroscopy. After a brief stint as a chemistry instructor at Cloud County Community College in Concordia, KS, Skylar joined the Kansas Department of Health and Environment laboratories in 2004. Her work focused on the determination of lead concentration by inductively coupled plasma mass spectrometry in blood samples for the Kansas Childhood Lead Poisoning Prevention Program. She is currently a research technician in the proteomics core facility at the Stowers Institute for Medical Research in Kansas City, MO.

Michael Washburn (far left) received his B.A. degree in chemistry from Grinnell College in Iowa in 1992. He earned his Ph.D. degree in biochemistry/environmental toxicology at Michigan State University in 1998 under the direction of Dr. William W. Wells. He joined the laboratory of Dr. John R. Yates at the University of Washington in 1999 for postdoctoral research in strong cation exchange and reversed phased chromatography-based proteomics. In 2000, he joined the proteomics group at the Torrey Mesa Research Institute in San Diego, CA, and primarily developed stable isotope labeling quantitative proteomics approaches and investigated the relationship between mRNA and protein expression. He joined the Stowers Institute for Medical Research in 2003, where he is the Director of Proteomics. He co-leads a research team focused on developing and applying multidimensional chromatography-based quantitative proteomics approaches for investigating the dynamics of multiprotein complexes and biological systems.

Mainstream multidimensional separation techniques to date consist of two dimensions, which allow for enhanced resolution and peak capacity. Multidimensional chromatography coupled to mass spectrometry has rapidly grown in use and is now routinely part of the shotgun proteomics approach (Figure 1).

The amount of mass spectra generated from such complex protein mixtures is very difficult to interpret, which is then exacerbated by the large number of peptides produced from enzymatic digestion by proteases such as trypsin. Tens of thousands of peptides across a wide dynamic range of concentrations are generated. These peptides will have very similar mass to charge (m/z) ratios and different levels of abundance. This creates a challenge for the mass spectrometer with respect to detection and identification of low-abundance proteins. Therefore, powerful separation techniques are required to maximize the number of peptides for analysis by mass spectrometry.

This review will focus on the development of multidimensional chromatography coupled to tandem mass spectrometry for shotgun proteomics analysis. With a brief description of selected historical references regarding two-dimensional (2D) chromatography, the review will comprehensively cover the development of shotgun proteomics beginning with the theoretical groundwork laid by J. C.

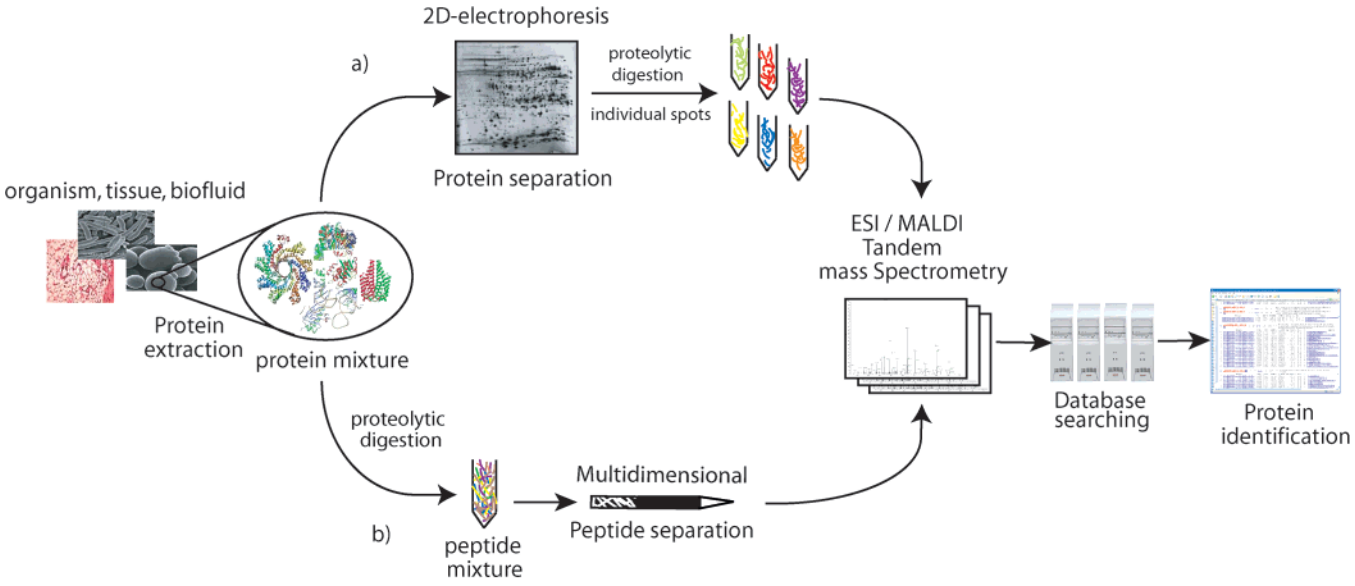

Figure 1. Proteomics analysis by (a) gel-based and (b) gel-free approaches. By the gel approach, the protein mixture is separated by two-dimensional electrophoresis, first by isoelectric focusing followed by SDS-PAGE. After spot visualization, proteins are extracted from the gel, digested, and analyzed by mass spectrometry for further identification by database searching. By the gel-free approach, the protein mixture is directly digested into a peptide mixture separated by multidimensional separation methods. Peptides are next analyzed by mass spectrometry. Proteins are identified from the generated mass spectra using database searching.

Giddings and colleagues. Finally, the review will contain a focused, but not comprehensive, description of informatics tools used to evaluate and assemble shotgun proteomics datasets for biological discovery.

2. Early Multidimensional Separation Theory

2.1. Peak Capacity and Orthogonality

Multidimensional separation couples two or more different separation methods by which analytes are first separated by one method and then by one or more additional independent separation methods. Greater chromatographic resolution obtained by multidimensional separations methods can be achieved by taking into consideration two fundamental criteria established by Giddings. First, the samples must be subjected to two or more independent separations. Frequently it is said that the separation techniques must be orthogonal, a term stemming from 2D thin layer chromatography (TLC) experiments in which the plate was developed in one direction and then rotated 90° and developed in a second direction.[16] The second criterion is that any components resolved in the first separation must maintain that resolution throughout the entire process.[17] The second requirement rules out the simple combination of any random separation techniques because it is possible to lose some or all of the resolution gained in the first separation during subsequent separations.[17] According to Giddings, the best way to ensure that resolution is not lost is to ensure that the sample pulse from the first technique is sufficiently small so that it can be reasonably resolved before being fed to subsequent techniques by which additional magnification may take place.[17]

The resolving power of any chromatographic technique is measured by the peak capacity. Giddings defined peak capacity as the "maximum number of peaks to be separated on a given column".[18] All peaks have a finite width as determined by the height of the theoretical plates and as such, only a limited number of peaks can fit into an allotted space. Giddings derived the peak capacity (n) to be

$$n = 1 + \frac{N^{1/2}}{m} \ln\left(\frac{V_n}{V_1}\right) \quad (1)$$

where $N^{1/2}/m$ is an approximation of peak width, V_1 is the volume at which the first peak elutes, and V_n is the volume at which the nth and final peak elutes.[18] The peak capacity for 2D separations is derived from 2D TLC experiments and is best described as[16]

$$n_{2D} = n_1 n_2 \quad (2)$$

and by extension the peak capacity for nth-order separations would be

$$\prod_{i=1}^{n} n_i \quad (3)$$

The actual peak capacities will be less than the calculated values due to additional broadening of the components in the second dimension[17] (Figure 2).

Although no researcher would take issue with improved resolution, one may question the necessity of it. Davis and Giddings published a series of reports on the statistical analysis of overlapping peaks in complex 1D chromatograms.[19−21] For a column with a peak capacity of 100, only 18 peaks will contain a single component. That constitutes 36% of the total sample components.[19] To recover 90% of the sample components as single peaks, the peak capacity must be at least 19 times larger than the number of components present. In that instance only 5.3% of the theoretical peak capacity would be utilized. The chromatograms would be practically empty.[19] Davis and Giddings also devised a statistical method for the estimation of the number of components in complex chromatograms[20] and tested it on actual chromatograms.[21]

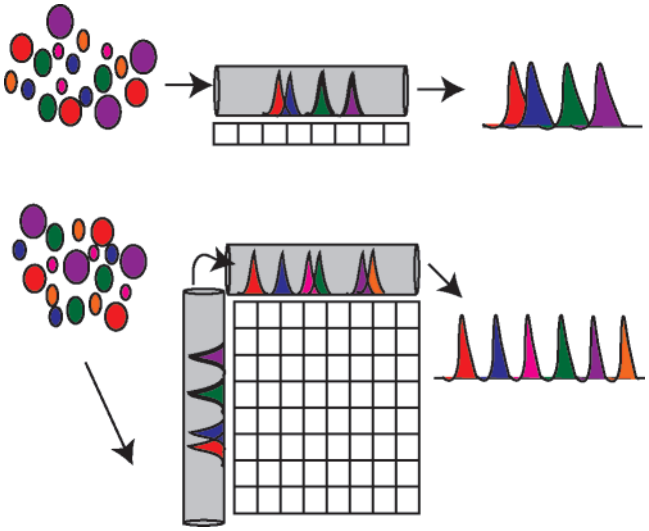

Figure 2. Visualization of peak capacity in both 1D and 2D separations. Identical samples of six species are separated by 1D and 2D techniques. Although the column shown for 1D separation has a theoretical peak capacity of eight (indicated by the boxes below the column), the 1D technique is able to clearly resolve only four distinct peaks. The addition of a second chromatographic dimension greatly improves the theoretical peak capacity ($8 \times 8 = 64$) as shown in the boxes below the columns. The second column is able to improve the separation of overlapped peaks so that clearly resolved peaks from all six species can be clearly identified.

They found that there were approximately 50−100% more components in the actual chromatograms than there were peaks.[21] This indicates that 1D chromatography is just not capable of adequately resolving peaks for complex samples.

2.2. Method Selection for 2D Separations

With the large number of separation techniques available, choosing the right combination for multidimensional separations can be complicated. To determine which separation techniques are best, the different types of separations must be discussed. Giddings proposed that 1D separations fall into one of two categories: selective (S) or nonselective (N).[16] Selective displacements are defined as providing different displacements for different components. In other words, they are what we would define as a separation technique. Nonselective displacements have equal displacement for all components. This type of displacement can be caused by mobile phase flow, column rotation, or other types of bulk sample transport.[16] Selective and nonselective displacements can be combined in three different ways: S × S, S × N, and N × N.

S × S combinations can be further broken down into two specific cases. The first case involves the chromatograph of the first displacement, being independent of the chromatograph in the second displacement. This type of displacement pairing is termed S × S_i. The second case is when both displacements are highly correlated. This case is shown as S × S_c.[16] These displacement pairings can be classified into either discrete separations or continuous separations. Discrete separations involve depositing a small amount of sample at a point and developing in both dimensions to produce discrete zones. These are the separations most frequently used in an analytical capacity. Continuous separations involve sample being continuously fed onto the separation bed. The sample components are separated differentially and can be continuously collected as components streams.[16] Discrete separations work best with S × S_i displacements. This is due to the enhanced resolution from increased separation space available and the fact that the independent displacement mechanisms can use that space. S × S_c and S × N displacements cannot adequately utilize the additional space for separation. N × N displacements do not really separate because there is no differential displacement. A perfect example of a discrete separation is a 2D TLC experiment.[16] Continuous separations operate well with S × S_i, S × S_c, or S × N displacements. This is because elution occurs continuously along one dimension, whereas both separation processes occur simultaneously. N × N displacements again do not yield any separation. Rotational chromatography and deflection electrophoresis are prime examples of continuous displacements[16] (Figure 3). Nonselective displacements have not been used in proteomics analysis to date and are discussed in this section to provide a complete theoretical understanding of multidimensional separations. All separations discussed in the remaining portions of this review are selective displacements.

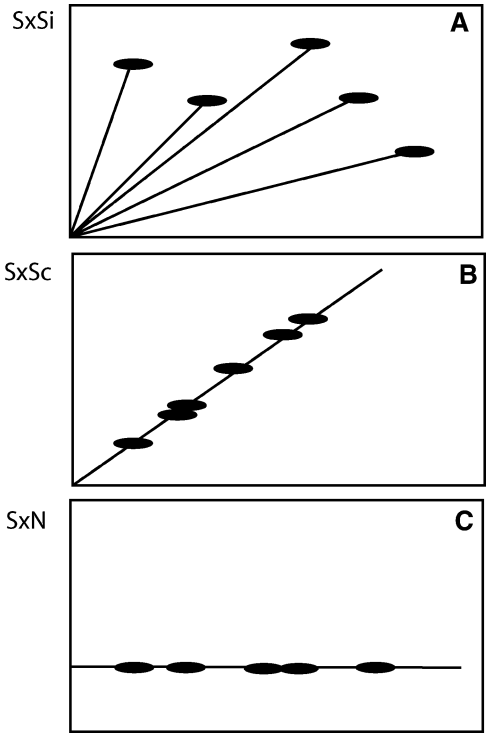

Figure 3. Representation of displacements that yield separation: (A) S × S_i displacements require that the first displacement is independent of the second displacement and require the use of different separation mechanisms. (B) S × S_c displacements involve two highly correlated mechanisms. (C) S × N displacements result when a selective displacement is coupled to a nonselective displacement.

Giddings has identified 15 1D displacements that could potentially serve as 2D separation components (Table 1). This leads to 225 possible combinations for 2D separations.[16] Not all of these combinations will yield valid 2D methods, so the question becomes how to combine techniques for 2D separations. Giddings outlines four criteria for successful technique selection. The first criterion is displacement choice. If the separation is to be a discrete separation then combining techniques that lead to S × S_i displacements is the best choice. However, if continuous separation is desired, the S

Table 1. Giddings' 15 One-Dimensional Displacements[a]

type of displacement	controlling property	example technique
bulk displacement	nonselective	N/A
flow	nonselective	N/A
chromatographic	partition coefficient	HPLC, GC, a number of other related techniques
field flow fractionation	field interaction parameter	field flow fractionation
electrophoretic	electrical mobility	capillary electrophoresis
isoelectric	isoelectric point	chromatofocusing, capillary isoelectric focusing
isotachophoretic	electrical mobility	capillary isotachophoresis
dielectrophorectic	dielectric constant	dielectrophoresis
sedimentation	sedimentation coefficient	gravitational sedimentation, sedimentation equilibrium
isopycnic sedimentation	density	sucrose density gradient, glycerol density gradient
magnetic gradient	magnetic susceptibility	magentic field gradient, pulse field gradient NMR
thermal diffusion	thermal diffusion coefficient	Clusius−Dickel column, temperature gradient
thermogravitational	thermal diffusion factor	thermogravimetry
diffusophoretic	interfacial energy	phase separation
photophoretic	photophoretic mobilty	laser photophoresis

[a] These displacements can serve as the building blocks for multidimensional separations. Each is given with the physical property that controls the separation as well as an example of each technique. All displacements are selective unless otherwise specified.

× N displacements are preferred due to the simplicity of N-type displacements.[16] Another criterion to be considered is the power of the displacement. Techniques for 2D separations should have high peak capacities to maximize the resolution of the separation.[16] Sample imposed restrictions create the third criterion. Some displacements can be ruled out automatically because they are simply not compatible with the sample.[16] The final selection criterion is the compatibility of the displacements. Compatibility depends on whether the displacements are to take place sequentially or simultaneously. Sequential displacements offer more flexibility because displacement conditions can be applied separately. Simultaneous displacements must occur in the same space at the same time.[16]

2.3. Planar 2D Separations versus Coupled Column 2D Separations

So far, the theoretical development has focused on planar 2D separations. However, these methods are not ideal for most separation applications. This work needed to be expanded to columns. Guiochon and co-workers proposed a system that was an intermediate between planar systems and coupled column systems.[22] Giddings offers his thoughts on coupled column 2D separations.[17] It is reasonably simple to meet the independent separations in the definition of multidimensional separations. The question becomes whether the resolution can be maintained as dictated by the second criterion. The answer depends on the amount of time required to feed the sample from the first column to the second column. If that time is short, so that the peaks resolved in the first column enter the second column, then resolution will be maintained or improved and the second criterion met. In this case, resolution can be gained in the first column or in the second column. If the time from the first column to the second column is too long, then any resolution gained will be lost and the system does not meet the second criterion.[17] Coupled column systems offer more flexibility and can be locally more powerful than 2D planar techniques. Coupled column systems offer the power to choose time intervals, sample subsets, etc. It should be noted that any multidimensional separation technique is useful only so long as any interfering components migrate independently during the separation.[17]

3. Early Multidimensional Separations of Amino Acids and Peptides

3.1. Thin Layer Chromatography

Multidimensional separations were reported as early as the 1940s. For example, in 1944, Consden et al. reported the separation of amino acids from mixtures of amino acids and wool hydrolysates.[23] Using cellulose as a support, amino acid separation was tested with a variety of solvent combinations for both 1D and 2D chromatography. Other factors such as temperature, salt content, and the effect of copper already present in the paper were also considered. One-dimensional chromatograms were performed on strips of paper like any thin layer chromatography experiment. Two-dimensional chromatograms required some additional processing. A drop of sample was placed in a corner of a piece of filter paper. The sample was allowed to develop in one solvent system and then dried. The chromatogram was rotated 90° from its starting point and then allowed to develop in a second solvent system. It was determined that six factors, some of which are mentioned above, affect the reproducibility of retention factor (R_F) values. Of those six factors, only extraneous substances present in the paper will alter the order of the bands in a 1D chromatogram. In the case of a 2D chromatogram, the pattern is modified, but recognition of the amino acids is still possible. The authors also noted that in a 2D chromatogram, regardless of solvent, glycine and its straight-chain homologues lie in a smooth curve.[23] Branched-chain homologues show a slight deviation from the curve. This led to the conclusion that a 1D chromatogram could demonstrate the presence of multiple constituents, whereas a 2D chromatogram could conveniently show members of a homologous series.[23]

In 1948, Haugaard and Kroner reported a 2D method that combined chromatography with electrophoresis to separate amino acids with overlapping R_F values.[24] By weaving metal ribbon, primarily nickel, platinum, or aluminum, into paper that had been dipped into pH 6.2 phosphate buffer and then dried, they were able to create leads for electrodes. A mixture of 10 amino acids, including 2 acidic amino acids and 2 basic amino acids, then underwent a typical 1D TLC experiment with phenol as the developing solvent. The paper chromatogram was then connected to electrodes, and a potential of 100V was applied. The acidic amino acids migrated toward

the anode, whereas the basic amino acids migrated toward the cathode. The neutral acids were unaffected by the voltage and did not migrate. The voltage allowed for the separation of amino acids with similar R_F values by their isoelectric points.[24]

Durrum extended the previous reports to examine electrophoresis in two dimensions using buffers with different pH values.[25] The first phase of the experiment was carried out with an ammonium hydroxide solution as the first electrolyte buffer. The paper was then dried, and the second phase of the experiments was carried out in an acetic acid electrolyte buffer. This method gave reproducible results for mixtures of amino acids as well as pancreatic casein hydrolysates. The analysis of blood serum proteins provided more of a challenge.[25] The patterns of the proteins were consistent with amino acid residues and could be reproduced. However, they were also quite sensitive to certain changes in handling, particularly with respect to the amount of time the paper was dried in between phases of the experiment and the degree of dryness. The author hypothesized that protein denaturation at the surface played some kind of role, considering the extreme pH range used in the buffers. To minimize this effect, different electrolytes and pH ranges were selected for electrophoresis. For the first phase, pH 5 acetate buffer was used, and pH 8.6 barbiturate buffer was selected for the second phase. Altough this helped with the protein denaturation, it created a new problem. In the case of the amino acid mixture and the casein hydrolysates, the first phase was volatile and was removed during the drying period between phases. For the blood serum, both phases are nonvolatile salts. The salt from the first phase was not removed before the second phase leading to a background salt content which could obscure separation. Durrum devised two possible resolutions to this issue.[25] The first was to increase the running time of the second phase. The second solution was to run the first phase on a strip of paper rather than the entire sheet. The strip, while still moist, could then be affixed directly to the sheet of paper to run the second phase. This approach reduced the salt concentration so that it remained sufficiently ahead of the amino acid residues[25] (Figure 4).

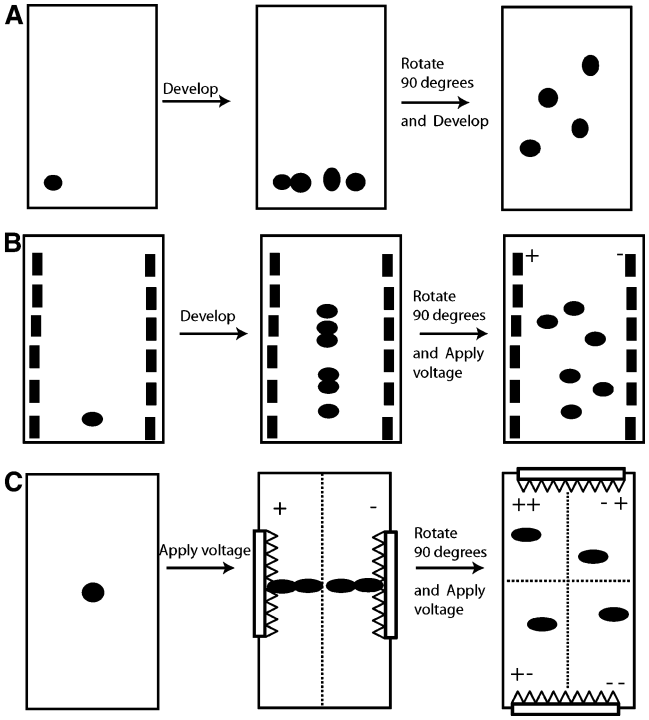

Figure 4. Early methods of multidimensional separations. (A) An example of 2D TLC performed by Consden and co-workers. The sample is spotted in one corner and developed in the first dimension with the first solvent system. After drying, the sample is rotated 90° and developed in the second dimension with a different solvent system. (B) An example of TLC combined with electrophoresis, similar to the work of Haugaard and Kroner. The sample is spotted in the center and allowed to develop as in thin layer chromatography. After drying, a voltage is applied to metal ribbons woven into the paper, and an electrophoretic separation occurs. (C) A 2D electrophoresis system similar to that of Durrum. The sample is spotted in the middle of the paper, and a voltage is applied in one direction with the first buffer system. After drying, the paper is rotated 90°, and a voltage is applied with a second buffer system.

3.2. Coupled Column Chromatography

The 1950s and 1960s marked a period of rapid growth in the field of chromatography. The development of gas chromatography (GC)[26] and later HPLC[27] led to a boom in 1D separation techniques, whereas 2D separations were placed on the backburner. Finding ways to couple instrumentation for 2D techniques was more difficult and expensive than it was to utilize the previously discussed 2D planar techniques. Gas chromatography advanced multidimensional separations in 1968 with Deans' report of an improved system design that made it easier and less expensive to couple two GC columns.[28] Thus, the concept of column switching became more practical to most laboratories.

A similar improvement occurred for HPLC in 1973. Huber and co-workers introduced a high-pressure micro-multiport switching valve.[29] A switching valve for HPLC faces some unique conditions. The valve must be able to operate frequently at high pressure. It must also be able to pass peaks without causing a significant amount of broadening. The valve designed by the authors was a pneumatic piston with one inlet that could be connected to a maximum of three columns. A key characteristic of this valve is a smooth flow path. If the sample was not able to move smoothly through the valve, additional mass transport could occur, which would increase broadening of the peaks. To determine if the valve design had any effect on peaks passing through it, the authors calculated the standard deviation of recorded peaks. The peaks passed through the valve to the detector via capillary tubing. It was determined that the standard deviation from the valve would not have a significant impact on any chromatographic applications.[29]

Multidimensional chromatography was often referred to as column switching or coupled column chromatography (CCC) during this period. However, it is important to keep in mind that not all column switching or CCC is in fact truly multidimensional. A prime example of this is given in a paper by Davis and Kissinger.[30] Although it does involve column switching, both columns contained reversed phase resin. Because the separation mechanisms are not independent, this system does not meet Giddings' criteria[17] and cannot be considered to be multidimensional.

Multidimensional HPLC separations fall into two main categories of either offline or online techniques.[31] Offline techniques involve the collection of sample peaks at the detector exit of the first column. The samples are treated, if necessary, and then re-injected onto the second column. Online techniques require the coupling of two columns through a switching valve. Both offline and online techniques have advantages and disadvantages. Off-line techniques allow

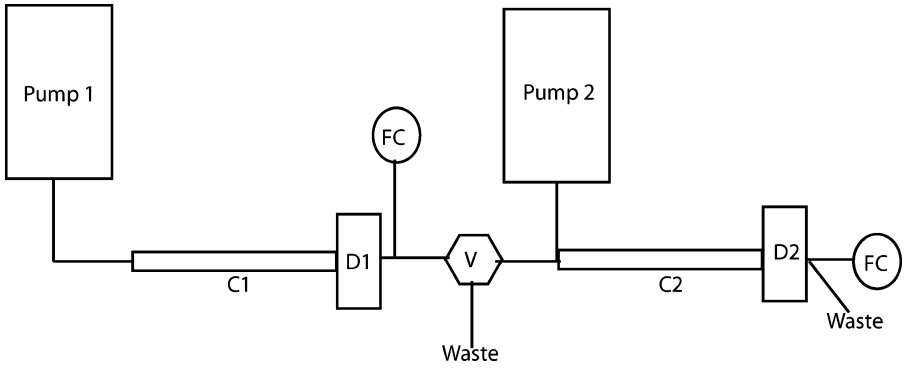

Figure 5. Generalization of coupled column chromatography. The first HPLC (pump 1) sends the sample through the first column (C1), into the detector (D1), onto the fraction collector (FC) to be collected for other analyses, or in the case of off-line separation treatment before injection to the next HPLC (pump2). For online separations, the sample can then be sent to the next HPLC by means of a switching valve (V). The second HPLC sends the sample through the second column (C2) to the second detector (D2) and an additional fraction collector.

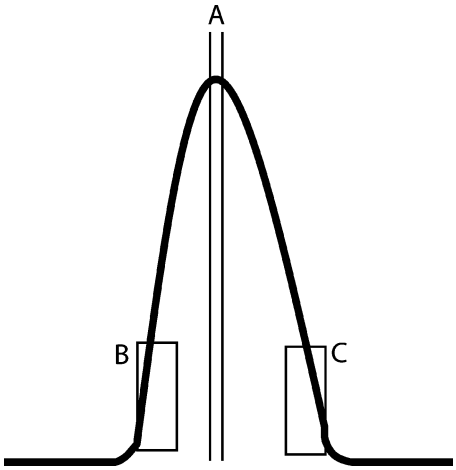

Figure 6. Graphical representation of (A) heart cutting, (B) front cutting, and (C) end cutting.

the separation conditions of both dimensions to be fully and independently optimized. In addition, off-line techniques allow for sample manipulation between dimensions, if necessary. Online techniques are usually preferred because of the reduced sample handling, decreased analysis time, and the ability to be automated[31] (Figure 5). Online techniques can be further broken down into additional subcategories. The most common is heart cutting. Heart cutting involves trapping, usually in a loop, of a defined volume, typically the maximum of the sample peak, after its detection on the first column and its direction to the second column. The next most common subcategory is online concentration. This method diverts the mobile phase with the peak of interest from the first column to be held at the head of the second column for a period of time before injection. During this time, the analyte(s) concentrate at the head of the second column. This technique is frequently used when the mobile phase of the first column is less compatible with the second column. Front cutting and end cutting are similar to heart cutting and occur at the beginning and tail of peaks, respectively[31] (Figure 6).

A number of researchers began to utilize multidimensional chromatography for a variety of applications.[32−35] It even became the subject of literature reviews.[36] There are a few papers from the period that stand out. The first is a report from Erni and Frei.[37] Although this particular paper does not deal with the separations of proteins or peptides, it did have an impact on future work in that area. By combining size exclusion chromatography (SEC) with reversed phase liquid chromatography (RPLC), they were able to separate senna glycosides from a complex plant extract. The authors used an online concentration technique. The sample first passed through the size exclusion column with a mobile phase of pH 7 aqueous buffer. Each fraction was collected in a loop and injected to the top of the reversed phase column. The sample concentrated there and was separated with a step gradient of acetonitrile in 0.1 N sodium bicarbonate/water. The flow rate of the size exclusion column was slow enough that the loop had time to fill, whereas the flow rate of the reversed phase column was fast enough that the loop completely flushed during the injection time. Analysis of seven fractions from SEC by RPLC showed that this method was better able to resolve the components of the extract than either technique individually.[37]

3.3. Multidimensional Chromatographic Separation of Proteins and Peptides

In another important body of work, tryptic peptides of human immunoglobulin D were separated by macroreticular cation exchange chromatography (CX) and RPLC.[38] This was an offline method in which the digested sample was separated by the CX with a gradient of water and a pH 6.2 buffer. The appropriate fractions were pooled and lyophilized before injection on the reversed phase column, which had a linear gradient of 0.1% trifluoroacetic acid to 40% propanol with 0.1% trifluoroacetic acid (TFA). The peaks were collected and later sequenced by Edman degradation. The chromatographic data were placed in profile to create a 3D visualization. This method of data handling provided a convenient summary of the peptides and helped to facilitate the recognition of overlapping peaks. By using 2D HPLC, Takahashi and co-workers were able to determine the sequence of the protein in only a few months.[38]

Later, Takahashi and co-workers developed an online system consisting of anion exchange chromatography (AE) and RPLC for the mapping of very large proteins, in particular human ceroplasmin.[39] The digested protein was loaded onto the AE column and eluted directly onto the RP column with a stepwise gradient of two different pH 8.0 Tris buffers. A linear gradient of acetonitrile with 0.1% TFA was applied to elute the peaks from the reversed phase column. Ceroplasmin was expected to yield 103 peaks, but the chromatographic data revealed approximately 260 peaks. The additional peaks were attributed to duplicated peptides in

adjacent stepwise elutions and incomplete cleavage during digestion.[39] Human ceroplasmin is a glycoprotein known to have four sites of attachment. It is possible to isolate these sites by gel filtration and RPLC, but it is a very time-consuming process. Using their 2D separation technique, the authors were able to isolate all four sites in 16 h and with higher yields than the manual technique.[39] This same separation system was also used by Takahashi et al. for peptide mapping analysis of human serum albumin and four of its known genetic variants.[40] Using the same instrumental setup as the previously discussed work,[39] the authors were able to analyze the genetic variants and identify the variant peptides in each of them. It was further proposed that 2D HPLC could be used to screen for genetic variants and even to study genetic relationships of protein variants found in different populations.[40]

Using an AE column and a reversed phase column in an instrumental configuration similar to Takahashi's,[39,40] Matsuoka et al. were able to separate complex peptide samples.[41] By separating various amounts of digested bovine calmodulin, a detection limit of 10 ng was established, and they succeeded in separating a peptide fraction from a crude brain extract into 150 single peaks with only one 80 min analysis.[41] The work with brain extract was extended to create a protein map with 2D HPLC, and by directly injecting the soluble extract of bovine cerebellum, the authors were able to identify approximately 200 peaks.[42] Many of the peaks were sufficiently pure to be utilized in additional characterization. At the time of publication, 15 proteins from the brain had been identified.[42]

3.4. Comprehensive Multidimensional Separations

A new era of multidimensional chromatography began in 1990. Bushey and Jorgenson published two reports[43,44] of what they termed "comprehensive" 2D chromatography. The difference between previous techniques and Bushey and Jorgenson's method is that comprehensive chromatography analyzes the entire effluent from the first column directly onto the second column. One of the biggest advantages is the time this method saves. The entire chromatograph can be completed in nearly the same amount of time it takes to run just the first column.[43]

The first paper coupled CX chromatography in the first dimension with SEC in the second dimension with the goal of separating a mixture of proteins and blood serum using UV detection.[43] The method is similar to that of Erni and Frei,[37] with the notable difference that the whole effluent is sampled with a much shorter analysis time. The flow rate of the CX column was considerably slower than that of the size exclusion column. This permitted the effluent from the first column to be sampled frequently by the second column, which means that peak profiles could be obtained in both dimensions, not just the second dimension. Peak profiling in both dimensions allows for a 3D visualization of the chromatographic data, similar to that of Takahashi and coworkers.[38] A slower first dimension flow rate also helped to decrease peak broadening in the second dimension.

There are three advantages to placing the ion exchange column before the size exclusion column. First, there is no need to run a gradient in the second dimension. This means that the starting conditions do not need to be regenerated in the column before the next injection can be made. This allows the second column to have a faster flow rate and more frequent sampling. The second advantage is that SEC has a definite beginning and end to its analysis times. This helps prevent overlapping peaks from run to run in the second dimension. Finally, all small ions and buffer components that normally appear as baseline disturbances are eluted at the end of the chromatograph and are easily ignored in data analysis. There is one major issue with the comprehensive method—that of detection. Because the entire effluent from the first column passes through the second column, a larger column is needed to handle the volume. This larger column can increase the dilution and make detection more difficult. The authors hypothesize that some form of peak compression may help to alleviate this.[43] They also note that peak compression measures may be difficult to implement and did not find it necessary in their application.[43]

The second paper by Bushey and Jorgenson coupled reversed phase chromatography with capillary zone electrophoresis (CZE) to separate the peptides from digested ovalbumin using fluorescence detection.[44] Although the dissimilarity of these two techniques lends itself to the orthogonality of the 2D separation, it also increases the difficulty of coupling the techniques. In addition to the difficulty in physically coupling these two types of chromatographic systems, there are some analytical dissimilarities that should be considered. One aspect of dissimilarity to be considered is the organic mobile phase used in RPLC and its compatibility with CZE. The authors, however, did not report any difficulties with the organic mobile phase in the electrophoresis.[44] Another possible incompatibility is the differing system volumes. The microbore reversed phase column used produced a far greater volume than the capillary electrophoresis column could sample. To combat this problem, the authors suggest using an open tubular column or a packed capillary column in the first dimension to better match the volume required by the electrophoresis.[44] CZE has also been coupled to SEC.[45] The coupling of two microcapillary techniques resulted in improved resolution in the separation of human blood serum.[45] Proteins, however, have a tendency to adsorb strongly to silica, which can make the detection of proteins difficult.

In 1995, Holland and Jorgenson reported the use of microcapillary columns in 2D comprehensive chromatography in the separation of tryptic peptides from porcine thyroglobulin and the catecholamine content of a single bovine chromaffin cell by AE chromatography coupled to RPLC.[46] Microcolumns can easily be fabricated to be much longer than standard columns without appreciable band broadening. The increase in length increases the number of theoretical plates, which in turn improves the resolution. This improved resolution was important in the analysis of the tryptic porcine thyroglobulin peptides. From the published peptide sequence, the authors deduced that there should be 211 tryptic peptides. The chromatography revealed 150 peaks, most of which were believed to be single component peaks.[46] Microcolumns are also better suited for the analysis of small volumes of sample. They allow for lower volumes of mobile phase so that smaller volumes of sample can be loaded and run without the significant dilution that can make detection difficult. This is highlighted by the analysis of the single bovine chromaffin cells. The authors were able to distinguish peaks not attributable to background from a few picoliters of sample, the estimated volume of a single chromaffin cell.[46]

That same year Moore and Jorgenson reported a comprehensive 3D system.[47] Their system coupled SEC with the

previously discussed RPLC/CZE system in the separation of ovalbumin peptides. The primary advantage of a 3D system is a further increase in peak capacity. The authors estimate that the peak capacity of the reversed phase column is 23, that of the capillary electrophoresis column is 24, and that of the size exclusion column is 5. According to Giddings,[16] the peak capacity of the reversed phase/capillary electrophoresis system is the product of their individual peak capacities, $23 \times 24 = 552$. By extension, the peak capacity of the 3D systems should be the product of all three peak capacities, $23 \times 24 \times 5 = 550 \times 5 = 2760$. Thus, by adding another dimension, even though it individually has a modest peak capacity, the peak capacity of the whole system was greatly increased.[47]

The chromatographic data were plotted in separation volume plots, where each axis represents a different separation method. Peaks are created by layering 2D slices of data to create elliptical stacks marking elution locations. From these data volumes any 2D combination can be extracted and further analyzed. From this it can be seen that there are no overlapping peaks. The 3D system suffers from some disadvantages. The first issue is sample dilution during analysis. The authors address this issue by using highly concentrated samples. However, this may not be an appropriate method if the sample size is limited. The second problem involves temperature control. To be able to strictly compare retention/migration times, the temperature must be carefully controlled to avoid drift.

3.5. Comprehensive Multidimensional Separations Coupled to Mass Spectrometry

A traditional method for sequencing proteins and peptides often involved stepwise chemical degradation from the N to the C terminus (Edman degradation). Although highly reliable, this method is time-consuming and cannot be used on proteins or peptides with N-terminal modifications. For these reasons, Edman degradation has predominately been replaced by mass spectrometry. Opiteck et al. first reported the combination of multidimensional separation methods with mass spectrometry in 1997.[48] The mass spectrometer essentially adds a third dimension because it has the ability to identify the components of coeluting peaks when they cannot be completely resolved by chromatography. The addition of mass spectrometry makes it possible to quickly analyze uncharacterized samples.

The successful coupling of multidimensional separations and mass spectrometry for protein and peptide analysis could not have been achieved were it not for advances in ionization techniques for mass spectrometry. During the 20th century, several methods were developed for ionizing molecules for MS analysis, including electron ionization, photon ionization, chemical ionization, and Penning ionization. Each of these methods is able to change the proton/electron ratio to create gas-phase ions. However, when these ionization methods are applied to biological molecules such as proteins, severe decomposition of the sample is observed. Only in the past two decades have "soft" ionization methods, such as MALDI and ESI,[3−5,49] been used for the analysis of large biomolecules. Consequently, these methods resulted in the development of powerful instrumentation that is used today for proteomics research.

During a MALDI experiment, the analyte of interest is mixed with an organic solution called a matrix. The analyte mixture is then allowed to dry and crystallize. Laser pulses are applied to the matrix, and the energy absorbed is transferred to the analyte, resulting in desorption and ionization. Separation methods are difficult to achieve during MALDI experiments because of the crystallized matrix that is needed for MALDI analysis. However, recent LC-MALDI platforms have been developed, which will be discussed in more detail later in this review.

During ESI, the analyte of interest is pushed through a metal capillary that contains a high potential energy. As the liquid emerges from the capillary, a strong electric field causes charge separation to occur between the ions. The charge separation is offset by the surface tension, which results in the formation of a Taylor cone at the capillary tip.[50] As the repulsion force of the ions exceeds the surface tension, the liquid is forced from the capillary and forms a jet from which the charged droplets emerge. Typically, nitrogen is used to help generate the liquid aerosol and evaporate the charged droplets. As the solution surrounding the droplets evaporates, electrostatic repulsion of the similarly charged ions causes them to become unstable and separate into smaller droplets. This process is referred to as Coulombic fission and occurs until the individual ions become free of solvent. The ESI source is positioned toward the entrance of the mass spectrometer, where a vacuum pulls the individual gas-phase ions into the instrument for analysis.

Unlike MALDI, in which singly charged ions are generally detected, ESI produces multiply charged ions that allow for detection of high molecular weight molecules when using mass spectrometers with a low mass limit. Additional advantages of ESI include the ability to use lower flow rates (3−6 μL/min) and increased sample-to-volume ratio of the droplet, which allow for faster evaporation time and longer measurement time. These principles have given rise to many modified ESI techniques, but the most widely used is nanospray.[51−53]

Although nanospray is similar to ESI, there are several notable differences between the two techniques. The most essential feature of nanospray is the use of a microcapillary instead of the larger capillary used by ESI. The first microcapillary, described by Wilm and Mann, consists of a pulled glass capillary with an outer diameter (o.d.) of 5−8 μm and an orifice of 1−2 μm.[52] The smaller capillary allows for an even slower flow rate (nanoliters per minute) and increased sample-to-volume ratio compared to ESI. The volume of the droplets produced by nanospray is ~100 times smaller in volume than the droplets created by ESI, which allows more efficient ionization of the sample.[53] An electric potential can be applied by coating the glass microcapillary with a conductive material, such as gold,[53] placing a metal cross between the HPLC and the glass microcapillary to which the voltage is applied,[54] or applying the voltage prior to the nanoLC column using a polyether ether ketone (PEEK) tee or cross.[55]

In the past decade, nanospray ionization in combination with 1D and 2D separation methods has been the forefront of proteomic research. NanoLC columns can be made or purchased and consist of a 50−100 μm pulled glass capillary packed with different stationary phases such as C18 or strong cation exchange (SCX) resins. Samples can be loaded on the nanoLC column in a number of ways including loop injection, autosamplers, bomb loading, or trap column. When such small columns are loaded, they frequently become plugged. Typically, 75 or 100 μm i.d. glass capillaries are

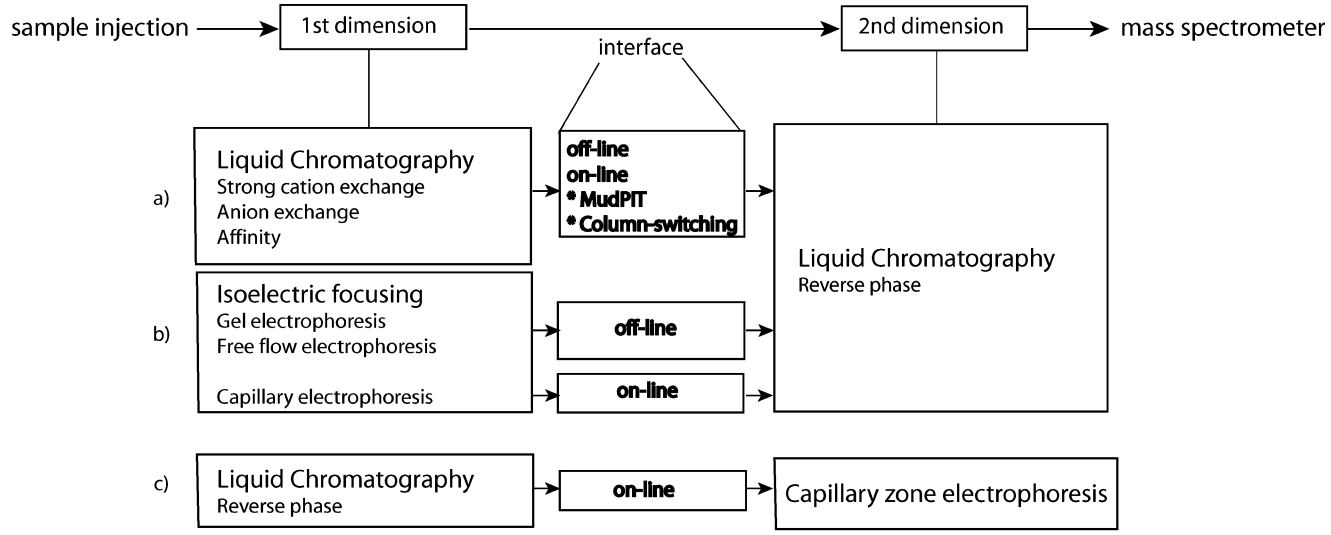

Figure 7. Fournier et al.

Figure 7. Two-dimensional (2D) peptide separation methods for shotgun proteomics analysis. (a) This method couples two liquid chromatography separations. In the first dimension, peptides are separated on the basis of charge or affinity and in the second dimension, on the basis of hydrophobicity. The two liquid chromatography separation methods can be coupled in offline or online modes. The online modes can be performed by MudPIT or a column-switching system. (b) This method couples a first separation based on the isoelectric point and a second separation by liquid chromatography based on hydrophobicity. In the first dimension peptides can be separated by isoelectric focusing through electrophoresis on immobilized pH gradient gel strips (IPG-IEF) or in solution, by capillary electrophoresis (CIEF) or free-flow electrophoresis (FFE-IEF). The IPG-IEF or FFE-IEF and CIEF systems are respectively coupled with the liquid chromatography method in off-line and on-line modes. (c) A third separation method couples liquid chromatography separation based on hydrophobicity with capillary zone electrophoresis. The separation systems are interfaced in an online mode.

used as a compromise between sensitivity (smaller columns) and routine experimental difficulties such as loading capacities and linear flow rates.

In a paper that is an important precursor to shotgun proteomics, Opiteck et al. reported a completely different type of 3D system in 1997.[48] This was the first report of comprehensive multidimensional chromatography coupled to a mass spectrometer for the analysis of *Escherichia coli* cell lysate.[48] The chromatographic system of Bushey and Jorgenson's first paper[43] is modified to deliver sample to the mass spectrometer. After passing through the reversed phase columns, the sample passes a detector and is then split. A fraction of the sample is diverted to the mass spectrometer while the rest of the sample flows to waste or fraction collection. Ion exchange chromatography in the first dimension and reversed phase chromatography in the second dimension were the techniques of choice. By using these modes in this order it was possible to run the second dimension quickly because RPLC is better suited for faster flow rates. It also provided a mobile phase that is more compatible with ESI-MS.[48] This system also has the advantage of being able to desalt proteins online.[48] The data gained from the addition of the mass spectrometer can be a valuable asset in peak identification. The mass spectrum of a peak from RPLC of *E. coli* lysate gave a molecular weight of 40,702.[48] After searching the Swiss-Prot database, the authors determined that there were 38 proteins within 2% of this molecular weight, but there were only 2 proteins within 0.2% of this molecular weight.[48] With the identity of the peak narrowed to two candidates, it would be completely feasible to correctly identify the protein in this peak by further testing with the remaining sample that was diverted to the fraction collector.

4. Shotgun Proteomics–2D Chromatography and MS/MS

The 2D peptide separation methods reported following Giddings' criteria include chromatographic techniques based on hydrophobicity, charge, molecular weight, or functionality of peptides. The separation of peptide mixtures by LC/LC methods has been performed using several orthogonal combinations such as strong cation exchange/reversed phase liquid chromatography (SCX/RPLC), anion exchange chromatography/reversed phase liquid chromatography (AE/RPLC), size exclusion chromatography/reversed phase liquid chromatography (SEC/RPLC), and affinity chromatography/reversed phase liquid chromatography (AFC/RPLC). In most shotgun proteomics analyses, the second dimension is performed by RPLC because the mobile phase is compatible with the mass spectrometer (Figure 7).

Of the many different 2D separation techniques developed,[56,57] the vast majority that are used include SCX coupled to RPLC. This particular 2D approach first separates the peptides on the basis of charge and then by hydrophobicity. There are two main approaches for applying 2D separation methods, offline and online. When using SCX and RPLC for offline approaches, the first dimension (SCX) is not directly coupled to the second dimension (RP) or SCX-RP. Fractions from the SCX column are collected and later subjected to the RP column. The online approach employs coupling the two chromatographic methods together so that the eluent from the first dimension (SCX) is directly eluted onto the second dimension (RP) or SCX/RP, thus avoiding the need for fraction collection. Online approaches are substantially faster than off-line approaches, and sample loss is minimized due to the direct coupling of the two dimensions. There are different variations of the online approach

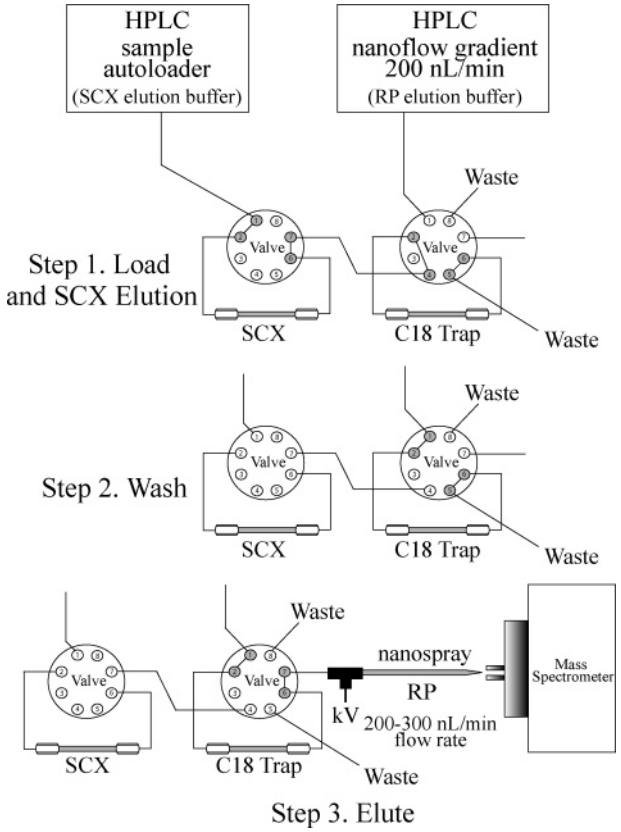

Figure 8. Automated online LC/MS/MS system using SCX and RP chromatography. In step 1, the sample is loaded onto the SCX column and the peptides bind according to their charge. The peptides are then eluted onto a RP trap column with a series of salt (KCl, NaCl, ammonium acetate or formate) bumps. In step 2, after collection of the peptides on the RP trap, the SCX valve is closed and the RP trap is washed free of salts to prevent them from entering the analytical RP column and mass spectrometer. In step 3, the RP valve is configured to elute the peptides from the trap to the analytical RP column using an acetonitrile gradient. The peptides are eluted from the analytical column directly into the mass spectrometer for analysis.

such as using separated columns for the SCX and RP connected by switching valves, or using multidimensional protein identification technology (MudPIT), where the SCX and RP stationary phases are packed together in the same microcapillary column. Once the peptides are separated by 2D chromatography, they are eluted directly into a tandem mass spectrometer. Traditional 2D LC/MS/MS methods use electrospray ionization and nanospray ionization; however, new methods, to be discussed later in this review, have been developed that incorporate 2D LC with MALDI (LC-MALDI).[58−63]

4.1. Strong Cation Exchange and Reversed Phase Systems

4.1.1. Column Switching SCX/RP/MS/MS Systems

Commonly, 2D SCX-RPLC switching systems connect the SCX column with a C18 trap RP column to a capillary C18 RP separation column by two or more switching valves (Figure 8). The peptide mixture is loaded onto the SCX under acidic conditions so that the positively charged peptides bind to the SCX column. By increasing the salt concentration, the peptides are displaced according to their charge and trapped in the C18 trap RP column. The trap C18 RP column inserted between the two analytical columns allows removal of salts from the cation exchange column before MS analysis. The trapped peptides are then eluted to the RP analytical column and separated using an organic phase gradient, usually acetonitrile. The SCX/RP/MS/MS approach was first described in the analysis of 10 protein standards by Opiteck et al.[48] In this arrangement, an HPLC was used to elute proteins from the SCX column with ammonium formate into a two-position, eight-port valve containing two loops. As the SCX eluent entered the first loop, another HPLC pumped the content to a second loop and then onto a RP column. Eluent from the RP column was directed through a UV detector and a flow splitter into the mass spectrometer. After identifying a simple complex protein mixture using this approach, they tested the system with a more complex protein mixture, an *E. coli* lysate.[48] From the *E. coli* lysate analysis, several proteins were identified using the Swiss-Prot database. Although this approach is much faster and far less prone to sample loss, Opiteck et al. concluded that the peak capacity of this approach was inferior in comparison to 2D polyacrylamide gel electrophoresis.[48]

The RP-trap setup has allowed well-characterized SCX elution buffers (i.e., phosphate buffers containing KCl and NaCl) for the analysis of human lung fibroblasts,[64] the *Chlamydia trachomatis* proteome,[65] or with ammonium chloride as shown in virus,[66] *Rickettsia prowazekii*,[67] *Saccharomyces cerevisiae*[68] and human breast,[69] pancreatic,[70] and liver[71] cancer cells. The RP-trap setup has also been used to desalt SCX eluents containing ammonium formate for the analysis of the *E. coli* proteome[72−74] and breast cancer cells[75] or ammonium acetate for the analysis of apoptosis,[76] human pleural effusion,[77] erythrocytes,[78] human Jurkat cells,[79] milk proteins,[80] and retinal development in fish[81] (Table 2). A recent example of this approach can be found in a report by Lim and Kassel.[82] To enrich the sample for phosphopeptides the peptides were loaded on the SCX column, eluted by ammonium formate, and separated on a RP column interfaced with an LCQ.[82] Their strategy was to maximize the binding of nonphosphorylated peptides and minimize the binding of phosphorylated peptides on SCX.[82] With this method, phosphopeptides are eluted from SCX at the earliest steps because their charge should be lower than +1 due to the presence of the phosphate group.[82] The RP-trap system has become a valuable asset for many online approaches and can be purchased from numerous HPLC and mass spectrometer manufacturers.

Link et al. developed a new and improved SCX/RP/MS/MS approach for the direct analysis of large protein complexes (DALPC).[83] This is the first description of analyzing a complex peptide mixture using the SCX/RP/MS/MS approach.[83] DALPC pairs cation exchange chromatography and reversed phase chromatography with tandem mass spectrometry using electrospray ionization. After digestion, acidified peptides are loaded on the cation exchange column. A fraction of the peptides present are moved onto the reversed phase column by a salt step. Peptides are retained on this column for desalting and then eluted to the mass spectrometer by an acetonitrile gradient. The reversed phase column is re-equilibrated, and the process is repeated with the salt concentration increasing at each salt step. Upon completion of the analysis, the obtained mass spectra are then compared with translated genomic sequences. By changing the number of salt steps and the slope of the acetonitrile gradient, DALPC can be tuned to suit the complexity of the sample.[83] Analysis of the *S. cerevisiae*

Table 2. Proteomics Analysis Using an Online SCX/RPLC Column-Switching System for Peptide Separation

application	elution salt	gradient salt conc (mM)	online system	mass spectrometer	no. of proteins identified	database searching
Chlamydia trachomatis proteome[65]	KCl	25−200	one 10-port valve	Q-TOF Global Ultima	117	ProteinLynx
liver cancer cells[71]	NH_4Cl	25−800	two 6-port valves	LCQ DecaXP	644	SEQUEST
Rickettsia prowazekii proteome[67]	NH_4Cl	10−800	one 6-port valve one 10-port valve	LCQ Deca XP Plus	252	SEQUEST
adenovirus type 5 proteome[66]	NH_4Cl	50−500	two 6-port valves	LCQ DecaXP	11	SEQUEST
S. cerevisiae proteome[68]	NH_4Cl	25−800	two 6-port valves	LCQ Deca XP Plus	577	SEQUEST
pancreatic cancer cells[70]	NH_4Cl	50−600	two 6-port valves	LCQ DecaXP	46	SEQUEST
human breast cancer cells[69]	NH_4Cl	100	two 6-port valves	LCQ DecaXP Plus	N/A[a]	SEQUEST
nuclear proteins from human Jurkat cells[79]	$CH_3CO_2NH_4$	50−2000	two 10-port valves	LCQ	174	MASCOT
apoptosis-induced DLD-1cells[76]	$CH_3CO_2NH_4$	5−2000	two 10-port valves	QSTAR XL	480	MASCOT
human pleural effusion proteome[77]	$CH_3CO_2NH_4$	10−2000	two 10-port valves	LCQ DecaXP Plus	1415	SEQUEST
erythrocytes proteome[78]	$CH_3CO_2NH_4$	10−2000	two 10-port valves	LCQ DecaXP Plus	272	SEQUEST
retinal development in salmonid fish[81]	$CH_3CO_2NH_4$	100−2000	two 10-port valves	QSTAR-QqTOF	N/A	ProICAT SP2
E. coli proteome[72]	HCO_2NH_4	20−1000	two 6-port valves	Agilent 1100 series MSD Ion Trap SL	450	MASCOT
human breast cancer cells[75]	HCO_2NH_4	25−250	one 6-port valve	LCQ Deca	798	SEQUEST

[a] N/A, values are not available.

80S ribosome demonstrated the capability of the DALPC approach.[83] The 80S ribosome was purified from yeast, and 120 μg of protein was digested with trypsin under denaturing conditions. Twelve peptide fractions were independently separated on the SCX column using KCl, eluted onto the RP column, and analyzed by tandem mass spectrometry. As a point of reference, the same amount of digested peptides was analyzed by RP/MS/MS and 2D gel electrophoresis for comparison. The results indicated that 75 of the 78 predicted ribosomal proteins were identified by DALPC compared to 56 identified by RP/MS/MS and 64 by 2D gel electrophoresis. These results demonstrated the power of DALPC over traditional separation methods for analysis by mass spectrometry. Interestingly, Link and colleagues followed up the DALPC analysis of ribosomes with a series of studies functionally characterizing novel ribosomal components.[84−86]

One limitation of the use of two columns for the DALPC method is the large amount of sample required to overcome dilution, a problem noted by Jorgenson and co-workers.[46,47] To alleviate this problem, a biphasic microcapillary column was designed with sequential cation exchange and reversed phase resins.[83] This column was used in conjunction with a simplified electrospray interface.[83] With the biphasic column, the flow rate could be greatly reduced, which helped to improve the detection limit by 2 orders of magnitude.[83] However, the continued use of KCl and potassium phosphate proved to be problematic. This work became the foundation for the MudPIT approach.[87,88]

4.1.2. Directly Coupled SCX/RP/MS/MS Systems

MudPIT is a fully automated, online, coupled 2D column SCX/RP/MS/MS approach designed for the analysis of complex peptide mixtures.[83,87,88] MudPIT typically consists of a 100 μm i.d. × 365 μm o.d. fused silica column that is packed with RP and SCX HPLC grade materials using a pressurization vessel (Figure 9A). The complex peptide mixture is then loaded onto the biphasic column and interfaced with a quaternary HPLC. The column then acts as the ion source for the mass spectrometer (Figure 9B). The peptide mixture is ionized and subjected to a series of salt bumps for the elution of peptides, which allows only a fraction of the peptides to be eluted from the SCX resin onto the RP resin. Following each salt bump, an acetonitrile gradient is performed for the direct elution into the tandem mass spectrometer. This was first described by Link et al.

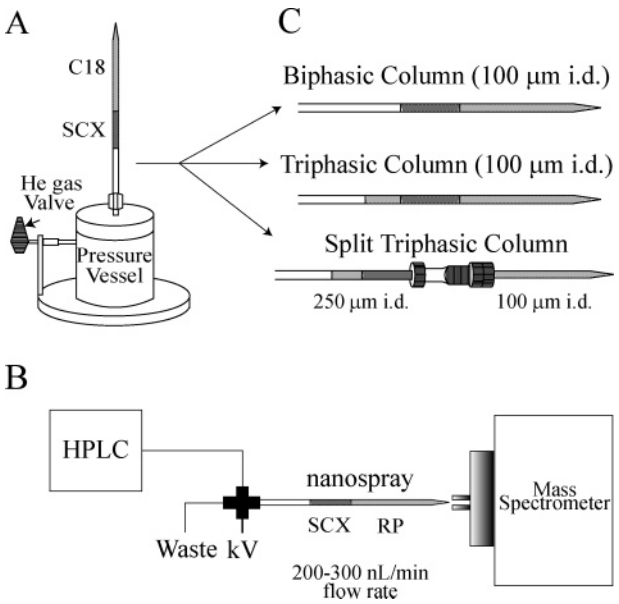

Figure 9. MudPIT system initially consists of packing a bi- or triphasic column with RP and SCX resin. (A) The resin is packed into a 100 μm i.d. column offline using a pressure vessel. The resin is pushed through the MudPIT column using helium at high pressure (500−1000 psi). Once the column is packed, the pressure vessel is used to load the sample. (B) During a MudPIT experiment, a series of salt bumps, followed by acetonitrile gradients, is used to elute the peptides directly into the mass spectrometer. The MudPIT approach requires the use of volatile salts such as ammonium acetate or ammonium formate. (C) Different columns can be used for MudPIT analysis. A biphasic column consists of only RP resin followed by SCX. Initially, samples that were loaded onto a biphasic column had to be desalted offline before MudPIT analysis. The triphasic MudPIT column is similar to the biphasic column with the addition of RP resin added after the SCX resin. This allowed the samples to be desalted on-line before entering the mass spectrometer and results in decreased sample handling. The split triphasic column was designed for holding more RP and SCX resin for larger samples. Two parts of the split column are connected via a microfilter assembly. One side of the column consists of 250 μm i.d. fused silica, and the other side consists of a 100 μm tip.

with the analysis of the 40S and 60S ribosomal subunits from *S. cerevisiae*[83] using KCl. After several methodological improvements and changes to the protocol set forth by Link et al., a large-scale analysis of the *S. cerevisiae* proteome was performed using MudPIT.[87,88] One of the significant differences between DALPC and MudPIT is the substitution of ammonium acetate for KCl and potassium phosphate, which allowed for the exclusion of a desalting step prior to MS/MS acquisition and direct electrospray into the mass spectrometer. Three portions of the yeast proteome were analyzed with three separate MudPIT columns.[87] After analysis of three different samples, Washburn et al. were able to detect and identify 1484 proteins and 5540 peptides. The 1484 proteins were representative of both soluble and insoluble fractions of the yeast lysate. MS/MS analysis of the heavily washed insoluble fraction contributed 2114 peptides of the 5540 peptides identified. Because of the largely unbiased nature of MudPIT, proteins with a variety of physical properties (i.e., p*I*, MW, abundance, and hydrophobicity) were identified. Although ammonium acetate is predominately used for MudPIT analysis, ammonium formate can also be used. The tandem arrangement of the MudPIT column appears to be problematic at first glance. However, MudPIT columns rely on two separate gradients to elute the SCX and RP dimensions, resulting in the ability to treat each portion of the column independently.

Producing complex peptide mixtures from biological samples typically involves the use of urea and other salts for optimum protease digestion. Initially, samples had to be desalted offline prior to loading onto a biphasic column for MudPIT analysis. Like the offline SCX-RP/MS/MS approach, this led to loss of sample and longer analysis times. As a result, a triphasic column was developed, which consisted of a small amount of RP prior to the SCX resin (RP/SCX/RP). The triphasic column was designed for desalting in the first dimension[89] and reduced sample handling. In addition to the triphasic column, the split-three-phase column was developed for larger amounts of RP/SCX resin and samples to be loaded (Figure 9C). The split-three-phase column consists of a 250 μm i.d. $\times$ 365 μm o.d. fused silica on one side of an M-520 inline microfilter assembly and a 100 μm i.d. $\times$ 365 μm o.d. tip on the other side as previously described.[90] Although the MudPIT approach requires the use of salt bumps, in contrast to SCX gradients, and also has the inability to use phosphate buffers containing KCl or NaCl, it has proven to be a successful approach for the proteomics analysis of complex biological samples.

Early work after the development of MudPIT largely focused on the analysis of whole proteomes of organisms or subcellular components of whole organisms. Aside from the initial description of MudPIT carried out on yeast whole cell lysates,[87] this technology was also applied to an analysis of rice leaf, roots, and seed, where a total of 2528 unique proteins were detected and identified.[91] In an early series of studies, MudPIT was used to characterize the life cycle stages of the causative agent of malaria, *Plasmodium falciparum*,[92] and the model rodent parasite *Plasmodium yoelli yoelli*.[93] These studies were followed by the identification of antigens in *P. falciparum*,[94] analyses of rhoptry[95,96]and Mauer's cleft[97] proteins, and a comparison of transcript and protein levels in *P. falciparum* across the life cycle.[98] Finally, a complete integration of genomic, transcriptomic, and proteomic datasets was carried out on *Plasmodium berghei* and *Plamodium chabaudi*.[99] Other examples of MudPIT and other biphasic columns are also present in the literature, including the analysis of the human plasma proteome,[100] the *Arabidopsis* cell wall proteome,[101] and mouse hearts[102] and the identification of proteins from uredospores,[103] human cerebrospinal fluid,[104] human tissue,[105] human urine,[106] human plasma,[107] and breast cancer cells.[108] The split-three-phase column was proven to be useful in the analysis of a microbial biofilm,[109] which is far more complex than any single-organism proteomics and indicates the resolving power of the split-three-phase column for the separation of complex peptide mixtures.

By coupling MudPIT to purification schemes to isolate organelles and portions of organelles, detailed information regarding protein localization is obtainable. In one tour de force, MudPIT was used in conjunction with advanced mathematical analyses and immunofluorescence to determine the subcellular localization of proteins in mouse tissues.[110] Additional examples include an analysis of proteins in the Golgi of rat livers,[111] mammalian cytokinetic midbody proteins,[112] and chromatin-associated proteins in *Caenorhabditis elegans* sperm.[113] In other research, the focus was on the detection, identification and characterization of membrane proteins.[114−118] For example, Wu et al. described an improved approach for analyzing membrane proteins via MudPIT,[114] and Schirmer et al. identified novel proteins localized to the nuclear envelope via MudPIT.[115] In a cancer-related analysis, Durr et al. analyzed the endothelial cell surface of lung,[116] which led to a subtractive proteomics analysis of lung endothelial cell surface to find potential therapeutic targets for cancer.[119]

MudPIT has also been widely used to characterize multiprotein complexes. Affinity purification of a multiprotein complex is particularly feasible in today's world with the availability of generic affinity purification approaches such as the tandem affinity purification (TAP) tag[120]or FLAG tag.[121] By coupling affinity purification of multiprotein complexes with MudPIT, detailed analyses of multiprotein complexes have been described, often resulting in the detection of novel protein−protein interactions. The analysis of mammaliam Mediator is an example of the use of the biphasic MudPIT column for the analysis of multiprotein complexes.[122,123] Finally, SAGA[124] and Swi/Snf[125] have also been analyzed via MudPIT without the use of the triphasic column (Table 3).

To minimize sample loss and maximize sensitivity and efficiency, the triphasic column[89,90] has been proven to be particularly useful for analyzing multiprotein complexes. Examples of multiprotein complex analysis via MudPIT using the triphasic column[89,90] include the analysis of the chromatin remodeling complexes Swi/Snf,[126,127] SAGA,[127,128] COMPASS,[129] Rpd3,[130,131] TRAP/TIP60,[132] SRCAP, and INO89[133] (Table 3). Additional examples of protein complex analysis using the triphasic column include the analysis of the Smc5−Smc6 complex involved in DNA repair[134] and the double-stranded RNA processing complex DCR-1.[135]

MudPIT has proven to have sufficient separation capacity, when coupled with digestion strategies to generate high sequence coverage of proteins, to detect a variety of post-translational modifications. MacCoss et al. described an approach in which a sample is split into three portions and each portion is digested with a unique enzyme and analyzed by MudPIT.[136] After database searching with SEQUEST, a variety of post-translational modifications were found on Cdc2p protein complexes and human lens tissue.[136] The general approach described by MacCoss et al. has since been

Table 3. Biphasic and Triphasic MudPIT Approaches

application	elution salt	mass spectrometer	no. of proteins identified	database search algorithm
human plasma proteome[100]	$CH_3CO_2NH_4$	Q-TOF API US	112	Swiss-Prot
ribosomal proteins of *S. cerevisiae*[83]	KCl	LCQ	30	SEQUEST
proteome analysis of *S. cerevisiae*[87]	$CH_3CO_2NH_4$	LCQ	1484	SEQUEST
P. falciparum proteome analysis[92,94,98]	$CH_3CO_2NH_4$	LCQ DECA	2415[92] 27[94] 2904[98]	SEQUEST
mouse organ composition[110]	$CH_3CO_2NH_4$	LCQ DECA	4768	SEQUEST
Arabidopsis cell wall proteome[101]	$CH_3CO_2NH_4$	Q-TOF2	89	MASCOT
analysis of mouse hearts[102]	$CH_3CO_2NH_4$	LCQ DECA	782	SEQUEST
human tissue profiling[105]	$CH_3CO_2NH_4$	LCQ DECA XP	1713	SEQUEST
analysis of microbial biofilm[109]	$CH_3CO_2NH_4$	LCQ and LTQ	2146	SEQUEST
uredospore proteins from *U. appendiculatus*[103]	$CH_3CO_2NH_4$	LCQ DECA XP	N/A[a]	MASCOT
human cerebrospinal fluid[104]	$CH_3CO_2NH_4$	LCQ DECA	249	SEQUEST
breast cancer cells[108]	$CH_3CO_2NH_4$	LCQ DECA/LTQ	3715	SEQUEST
bovine microtubules[89]	$CH_3CO_2NH_4$	LCQ	62	SEQUEST
proteomics analysis of *P. berghei*[99]	$CH_3CO_2NH_4$	LCQ DECA	1836	SEQUEST
P. yoelli yoelli proteome analysis[93]	$CH_3CO_2NH_4$	LCQ DECA	N/A	SEQUEST
proteomics analyses of rhoptry and Mauer's cleft[95−97]	$CH_3CO_2NH_4$	LCQ DECA	N/A[95] 148[96] 10[97]	SEQUEST
analysis of the Golgi in rat liver[111]	$CH_3CO_2NH_4$	LCQ DECA	421	SEQUEST
mammalian midbody proteome analysis[112]	$CH_3CO_2NH_4$	LCQ DECA	577	SEQUEST
C. elegans sperm proteome[113]	$CH_3CO_2NH_4$	LCQ DECA	132	SEQUEST
membrane protein analysis in rat brains[114]	$CH_3CO_2NH_4$	LCQ DECA	1610	SEQUEST
analysis of rat lung tissue[116]	$CH_3CO_2NH_4$	LCQ DECA XP	450	SEQUEST
nuclear membrane proteins in rat liver[115]	$CH_3CO_2NH_4$	LCQ-DECA	N/A	SEQUEST
identification mammalian HAT and SRCAP protein complexes[132]	$CH_3CO_2NH_4$	LCQ DECA XP	10	SEQUEST
identification of mammalian mediator subunits[122,123]	$CH_3CO_2NH_4$	LCQ DECA XP	N/A[122] 37[123]	SEQUEST
components of the *S. cerevisiae* SAGA complex[124,127,128]	$CH_3CO_2NH_4$	LCQ	18[124] 18[127] 18[128]	SEQUEST
components of the *S. cerevisiae* Swi/Snf complex[126,127]	$CH_3CO_2NH_4$	DECA XP	381[126] 18[127]	SEQUEST
components of the *S. cerevisiae* COMPASS complex[129]	$CH_3CO_2NH_4$	DECA XP	6	SEQUEST
identification of *S. cerevisiae* Rpd3 complex components[130,131]	$CH_3CO_2NH_4$	DECA XP	12[130] 14[131]	SEQUEST
identification of mammalian INO80 complex components[133]	$CH_3CO_2NH_4$	DECA XP	13	SEQUEST
analysis of the mammalian TRAP/TIP60 complex[132]	$CH_3CO_2NH_4$	DECA XP	17	SEQUEST
analysis of the Smc5−Smc6 complex in *S. cerevisiae*[134]	$CH_3CO_2NH_4$	LCQ DECA/LTQ	8	SEQUEST
analysis of the DCR-1 complex in *C. elegans*[135]	$CH_3CO_2NH_4$	LCQ DECA	20	SEQEUST

[a] N/A, values are not available.

used to analyze phosphorylation,[137−141] ubiquitinylation,[142] SUMOylation,[143−145] lipid modifications,[146] and methylation[111] of protein complexes and cellular compartments. With advances in database searching related to phosphorylation,[147,148] SUMOylation,[149] and modifications in general,[150] these types of analyses should only improve in quality.

Finally, MudPIT has proven to be increasingly useful for the quantitative analysis of proteomes and multiprotein complexes. Early work demonstrating MudPIT as a tool for quantitative proteomics analysis used ^{14}N and ^{15}N in the yeast *S. cerevisiae*.[151−153] In these studies, MS scans were used to determine the ratios of ^{14}N and ^{15}N peptides eluting into the mass spectrometer in a manner similar to the one described in detail by MacCoss et al.[154] In an important study, Liu et al. demonstrated the power of the spectral counting method for quantitative analysis via MudPIT,[155] and this approach was demonstrated to correlate strongly with quantitative analysis using peptide ion chromatograms.[156,157] Recently, the ability to carry out statistical analysis of quantitative proteomics datasets using a normalized spectral counting approach has been shown with isotopically labeled *S. cerevisiae*[158]and with label-free human Mediator complexes.[159]

Although MudPIT has proven to be a powerful technology and has been used extensively to foster biological discovery, it is not a comprehensive technique. No proteomics technology has yet to be shown to detect and identify all proteins in a given biological sample. Because this is often an important goal for certain types of experimentation, many researchers have employed multiple approaches to attempt to more completely characterize a proteome. Examples of such studies include a report by Breci et al. using MudPIT, gas-phase fractionation, and gel electrophoresis to characterize the *S. cerevisiae* proteome,[160] and two studies looking at the serum proteome using multiple techniques.[161,162] Another issue with MudPIT is the challenge of using an autosampler and maintaining zero dead volumes. One system to overcome this expanded on the vented column approach,[163] for which a vented fully inline triphasic RP/SCX/RP column was designed,[164] which may help to improve the automation of MudPIT. The vented column approach consists of a PEEK cross filled with packing material (Figure 10). The analytical column is attached to this packed cross. The packed cross permits the use of higher flow rates during the loading because flow can be directed to the waste line and controlled by a two-position valve.

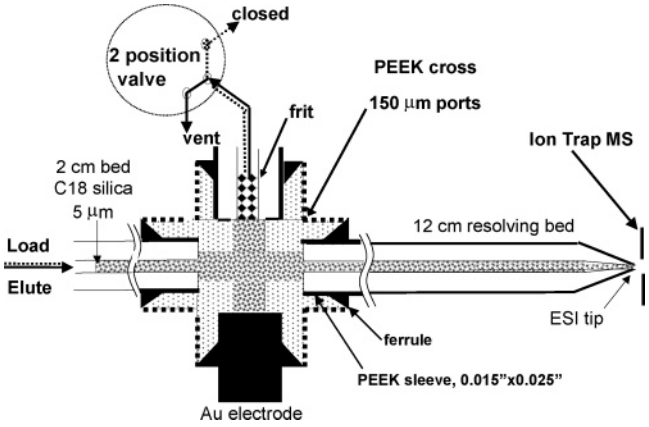

Figure 10. Vented column for desalting SCX fractions prior to mass spectrometry analysis. Reprinted with permission from ref 163. Copyright 2002 American Chemical Society.

4.1.3. Uncoupled SCX-RP/MS/MS Systems

The uncoupled/offline approach involves collecting fractions from the SCX separation and later subjecting them to RP chromatography before MS/MS (Figure 11). The first 2D separation of tryptic peptides using SCX and RP as described by Takahashi et al.[38] has previously been discussed. The authors concluded that the 2D separation method using SCX and RP provided much better resolution of complex peptide mixtures compared to either method alone.[38] Offline fraction collection allows for a variety of buffers and elution conditions to be used for the optimization of each separation. For example, phosphate buffers and salts common to SCX gradient elutions such as NaCl and KCl, which are not compatible with mass spectrometers, may be used and desalted prior to MS/MS.[165−173] Several methods may be used for desalting SCX fractions, including offline methods such as dialysis or desalting columns[165,166] on the RP column prior to MS/MS acquisition,[167,174] use of a vented column,[169] or use of a RP-trap column.[170,175] Numerous examples of 2D separation methods using KCl with the SCX-RP/MS/MS approach are present in the literature. For example, analysis of the *S. cerevisiae* proteome by Peng et al. involved digestion of 1 mg of yeast lysate followed by SCX chromatography for separation in the first dimension, where the lysate was eluted from a SCX column with fractions collected every minute during an 80 min KCl gradient.[169] The fractions were analyzed by RP/MS/MS using a vented column as described by Licklider et al.[163] (Figure 10). Other examples of SCX-RP/MS/MS include identification of the human saliva proteome,[170] the *Bifidobacterium infantis* proteome,[174] phosphoproteins in HeLa cells,[166] phosphoproteins in mouse brains,[165] the *S. cerevisiae* proteome,[169] human heart mitochondria,[171] tomatoes,[172] and human lumbar cerebrospinal fluid[173] (Table 4).

In addition to NaCl and KCl, volatile salts, such as ammonium formate[168,176−182] and ammonium acetate,[183−187] have been used with the SCX-RP/MS/MS approach. Once the SCX fractions are collected, they are lyophilized, resuspended in a low concentration of acetonitrile containing 0.1% formic acid, and directly analyzed using RP/MS/MS. Recent SCX-RP/MS/MS applications using ammonium formate include analysis of the human blood serum proteome,[176] human blood plasma proteome,[168,178] a mouse cortical neuron proteome,[180] and the mouse brain proteome.[181,182] Recent SCX-RP/MS/MS applications using ammonium acetate include analysis of the *Salmonella typhimurium* proteome,[183]

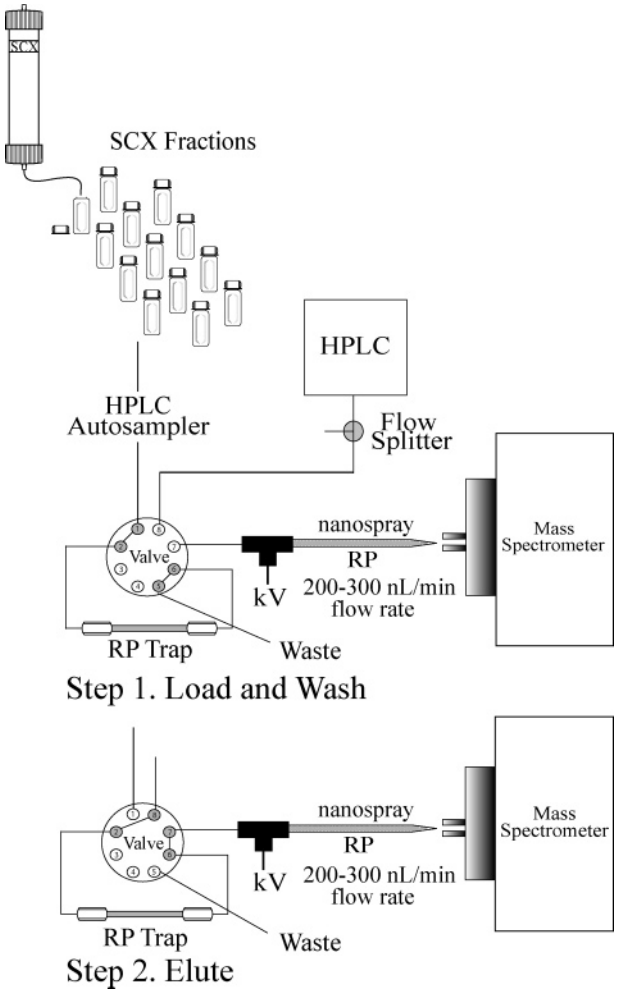

Figure 11. Offline LC/MS/MS system using SCX and RP chromatography. In this approach, the SCX chromatography is performed offline. A peptide mixture is loaded onto the SCX column and eluted using salts such as KCl, NaCl, and ammonium acetate or formate. In step 1, the fractions are placed into an autosampler and loaded onto a C18 RP trap column and washed free of salts. In step 2, the valve is configured to elute the peptides from the C18 RP trap column to the analytical C18 RP column. An acetonitrile gradient is applied, and the peptides are eluted directly into the mass spectrometer.

diseased urine,[184] the *Trypanosoma cruzi* proteome,[185,187] and mouse serum[186] (Table 4).

The flexibility of the offline approach allows incorporation of additional separation techniques such as SEC. For example, in the analysis of human mammary epithelial cells,[188] proteins were separated using SEC followed by fraction collection. After tryptic digestion, the SEC fractions were subjected to SCX chromatography. SCX fractions were then collected and analyzed using RP/MS/MS. Other advantages of using offline separation methods include the ability to carry out salt gradients rather than pulses, or bumps, as well as using a variety of elution conditions. In addition, offline approaches are able to use larger sample volumes and protein amounts, and they have flexible loading capacities.[169] Furthermore, the concentration of SCX peptide fractions can be estimated with UV absorbance, prior to RP chromatography, to maximize sample loading.[169] This is important because the sensitivity of peptide detection is largely dependent on the sample concentration. The main disadvantage of such offline methods is the substantial amount of time required for the optimization of each

Table 4. Offline 2D Approaches Using SCX and RP

application	elution salt	mass spectrometer	no. of identified proteins	database search algorithm
mouse brain phosphoproteins[165]	KCl	LCQ DECA XP	N/A[a]	SEQUEST
HeLa cell phosphoproteins[166]	KCl	LCQ DECA XP	967	SEQUEST
S. cerevisiae proteome analysis[169]	KCl	LCQ DECA XP	1504	SEQUEST
human saliva proteome[170]	KCl	LCA Classic	102	SEQUEST
B. infantis proteome analysis[174]	KCl	Q-TOF	136	SEQUEST
human heart mitochondrial proteome[171]	KCl	QSTAR	314	SONAR
component analysis of tomatoes (*S. lycopersicum*)[172]	KCl	Q-TOF-2	N/A	ProteinLynx
human cerebrospinal fluid proteome[173]	KCl	LCQ Classic	148	SEQUEST, MASCOT
human urinary peptides[184]	$CH_3CO_2NH_4$	Q-TOF	N/A	MASCOT
S. typhimurium proteome analysis[183]	$CH_3CO_2NH_4$	LCQ	816	SEQUEST
mouse serum proteome analysis[186]	$CH_3CO_2NH_4$	LCQ DECA XP	4567	SEQUEST
T. cruzi proteome analysis[185,187]	$CH_3CO_2NH_4$	Q-TOF-2	2784[185] 444[187]	MASCOT
human blood serum[176]	HCO_2NH_4	LCQ DECA XP	490	SEQUEST
mouse cortical neuron proteome[180]	HCO_2NH_4	LCQ DECA XP	4542	SEQUEST
murine natural killer cells[177]	HCO_2NH_4	LCQ DECA XP	2563	SEQUEST
analysis of human mammary epithelial cells[188]	HCO_2NH_4	LCQ DECA XP	1574	SEQUEST
human blood plasma proteome[168,178]	HCO_2NH_4	LCQ DECA XP	174[168] 804[178]	MASCOT
characterization of the mouse brain[181]	HCO_2NH_4	LTQ	7792	SEQUEST
analysis of membrane proteins in mouse brain[182]	HCO_2NH_4	LTQ-FT−ICR	1213	MASCOT

[a] N/A, values are not available.

separation. However, the time required for such offline 3D and 2D separations can be reduced by using automated sample-handling methods between separations. Other disadvantages of offline separation methods include sample loss between separations as well as the extra manual manipulation needed for performing two or three column separations and mass spectrometer analysis. There have been very few studies attempting to systematically compare different approaches. In one such study, Gilar et al. analyzed a five-protein mixture using a variety of LC/LC/MS/MS setups and concluded that a RP/RP system using significantly different pH values provided the highest peak capacity.[189] These types of comparisons of approaches will prove to be useful to shotgun proteomics going forward.

4.2. SCX and RP with MALDI/MS

Even though SCX and RP coupled to ESI have dominated shotgun proteomics, SCX and RP have been successfully coupled to MALDI systems. The first challenge that needed to be overcome was the coupling of liquid chromatography to a MALDI source. Although 1D separations are not within the scope of this review, it must be mentioned that several methods present in the literature use RPLC/MALDI for method development[190−194] and may be useful for implementing 2D separation techniques. The wide range of LC-MALDI methods present in the literature suggests that a well-defined common method has not yet been developed.

Introduction of a liquid solution into a MALDI source was first developed by Li et al.[195] This method was similar to the continuous-flow fast atom bombardment probe and consisted of a fused silica capillary placed orthogonally to the analyzer. The matrix solution was injected with the sample into the fused silica capillary, and the desorption laser energized the analyte solution at the tip of the capillary before entering the flight tube. Analysis of myoglobin was used to compare the flow tube with a solid probe. Li et al. reported that the signal from the flow tube lasted much longer than the signal from the solid tube, but the intensity of myoglobin was 5−10-fold less.[195] Li and co-workers later improved their method by increasing the resolution, but concluded that new and improved matrix solutions would be required for useful MALDI data because of adduct formation.[196]

Murray and co-workers focused their efforts on aerosol sample introduction for MALDI.[197−202] Aerosol sample introduction allows for faster sample evaporation due to the increased surface area and is the most successfully used LC/MS technique for sample introduction to date. During aerosol sample introduction, the analyte is mixed with a matrix solution and passed through a gas nebulizer tube to form the aerosol. Next, the aerosol is passed through a drying tube and into the ion source, where a laser irradiates the particles to form ions that are then accelerated into the mass spectrometer.[200] Murray et al. demonstrated the use of aerosol sample introduction with the coupling of RPLC to MALDI for the analysis of commonly used MALDI analytes.[199]

Murray and co-workers later developed a rotating ball inlet for use with liquid introduction MALDI,[203] which was modified from a rotating quartz wheel developed by Preisler et al.[204] This method consisted of a capillary that contacts a rotating stainless steel ball on which MALDI was carried out. The analyte−matrix solution was delivered through the capillary to the rotating ball, where it passed through a vacuum and was subjected to irradiation with a pulsed laser. The ions where then accelerated into the mass spectrometer. The results from former LC-MALDI/MS methods mentioned above have proven to be beneficial for future methods, but the lack in sensitivity and resolution left a need for new and improved designs.

More recent LC-MALDI methods are focused on offline coupling of LC to MALDI by depositing LC eluants onto MALDI target plates (Figure 12). The majority of 2D separation methods for MALDI are identical to the SCX-RP/MS/MS methods discussed above, by which the SCX chromatography is performed offline and fractions are collected, desalted, and loaded onto a RP column. Eluent from the RP column is then mixed with a matrix online, using a Micro Tee fitting, and spotted out on a MALDI target plate for analysis. As mentioned, offline 2D separations allow

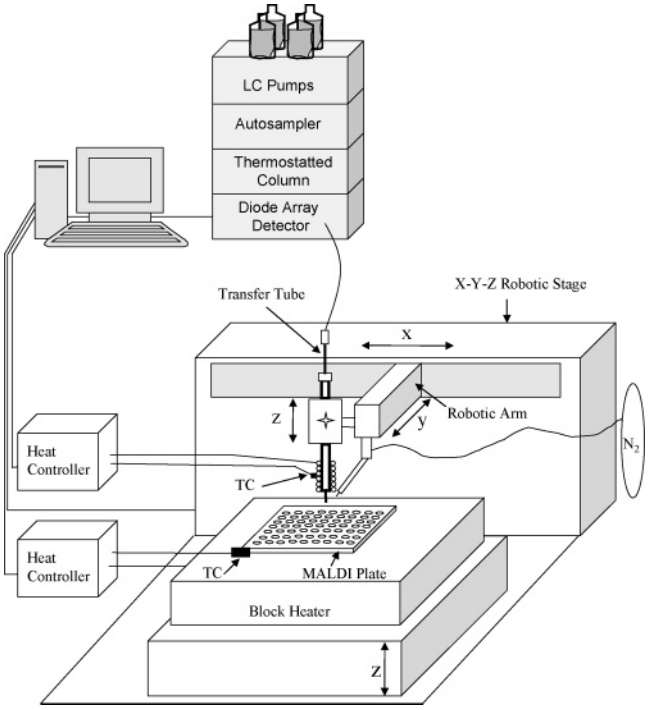

Figure 12. Heated droplet interface used in combination with LC/MALDI. Reprinted with permission from ref 191. Copyright 2004 American Chemical Society.

for flexibility with regard to buffers and salts. Examples of SCX/RP and SCX-RP MALDI/MS/MS using phosphate buffers containing KCl include the analysis of the *S. cerevisiae* proteome[58] and the mouse myelin sheath.[59] The analysis of the yeast proteome was used to compare three different separation methods combined with MALDI/MS/MS, SCX-RP, AE-RP, and SCX/RP.[58] Analysis using SCX-RP consisted of a KCl gradient for SCX elution, followed by fraction collection and RP/MALDI/MS/MS.[58] The AE-RP analysis consisted of a NaCl gradient for AE elution, followed by fraction collection and RP.[58] Finally, the SCX/RP consisted of a potassium phosphate step gradient followed by a RP tap column for desalting and analysis using RP.[58] Results suggested that all three methods are comparable, with 1003 proteins identified using SCX-RP, 1215 proteins identified using AE-RP, and 1187 proteins identified using SCX/RP.[58] Additional examples of SCX-RP include the analysis of the *Bacillus subtilis* proteome using ammonium chloride[60] and of HCT116 cells using SCX-Stage Tips.[61] Analysis of protein complexes such as the growth receptor-bound protein 2 (GRB2) complex using ammonium acetate has also been reported,[62] as well as AFC-RP/MALDI for the purification of DNA binding proteins.[63] The successful analysis of a wide variety of applications using 2D LC coupled to MALDI demonstrates the progress that has been made over the years.

4.3. Alternative 2D Separation Approaches for Shotgun Proteomics

4.3.1. Anion Exchange and Reversed Phase LC

The first system combining anion exchange and reversed phase separation methods was developed by Holland and Jorgenson[46] for the separation of peptide mixtures of biological amines. The peptides were first separated on an AE column, eluted using guanidine thiocyanate and directly transferred to an RP column to be detected by laser-induced fluorescence (LIF). These separation methods were later used by Mawuenyega et al.[205] for a large-scale protein identification of the *C. elegans* proteome. Peptides were eluted from the anion exchange column by a NaCl gradient, trapped on a C18-trap RP column, and separated on an analytical RP column prior to tandem mass spectrometry analysis.[205] The same platform has been used for proteomics analysis of *E. coli*[206] and mouse embryonic stem cells.[207] A large portion of the proteome has been covered including membrane proteins and proteins with extreme p*I* values (between 3.42 and 13.1[206]) and molecular masses (between 6 and 1369 kDa[205]).

4.3.2. Affinity Chromatography and Reversed Phase

Affinity chromatography methods have been applied in shotgun proteomics for the study of post-translational modifications (PTMs). PTMs are covalent modifications of proteins altering their activity, stability, or localization and, consequently, regulating cellular processes. PTM analyses in shotgun proteomics are typically performed by multidimensional peptide separations followed by MS analysis to map the modifications along with the protein sequence. If mass spectrometry-based methods have continuously increased the sensitivity of PTMs detection, they remain limited by the low frequency and low stoichiometry of the modification event. To reduce sample complexity, the modified peptides are typically enriched through affinity chromatography based on properties of the covalent modification. Glycosylation and phosphorylation are the most widely studied PTMs, and enrichment techniques for these modifications have been well described.

4.3.2.1. Phosphopeptide Enrichment Techniques. Phosphorylation, which plays a key role in regulating protein activity, cellular signaling, and metabolism, is the most widely studied PTM. Phosphorylation is a reversible modification of proteins that occurs on serine, threonine, and tyrosine residues by the addition of a phosphate group. Phosphopeptides can be enriched on the basis of the electronegativity of the phosphate group using immobilized metal affinity chromatography (IMAC). Phosphopeptides, negatively charged under acidic conditions, bind to IMAC columns complexed with metal ions such as Fe^{3+} or Ga^{3+}.[208,209] Unbound non-phosphopeptides are removed from the column by an acidic wash, and phosphopeptides are eluted under alkaline conditions. The phosphorylated peptides are then separated by RPLC and analyzed by tandem mass spectrometry (Figure 13). One limit of this method is that peptides containing carboxylic groups could bind to IMAC. This problem has been solved by Ficarro et al.,[208] by converting the free carboxylic acid groups into methyl ester derivatives. The IMAC strategy has been used for several phosphoproteomic analyses in model organisms and cellular organelles, such as yeast,[208] colon carcinoma human cells,[210] rat liver,[211] human cardiac cells,[212] postsynaptic density,[213] and human T cells.[214,215] This approach has also been combined with deuterium labeling of phosphopeptides during the esterification step to analyze changes in phosphorylation in capacitated human sperm cells,[216] starved human lung cancer cells,[217] interferon α-treated human T cells,[218] and pregnant rat uteri treated with 8-bromo-cGMP.[219]

Although IMAC-Fe^{3+} has been a successful method for phosphopeptide enrichment, Pozewitz and Tempst[209] have also reported an efficient binding of phosphopeptides using

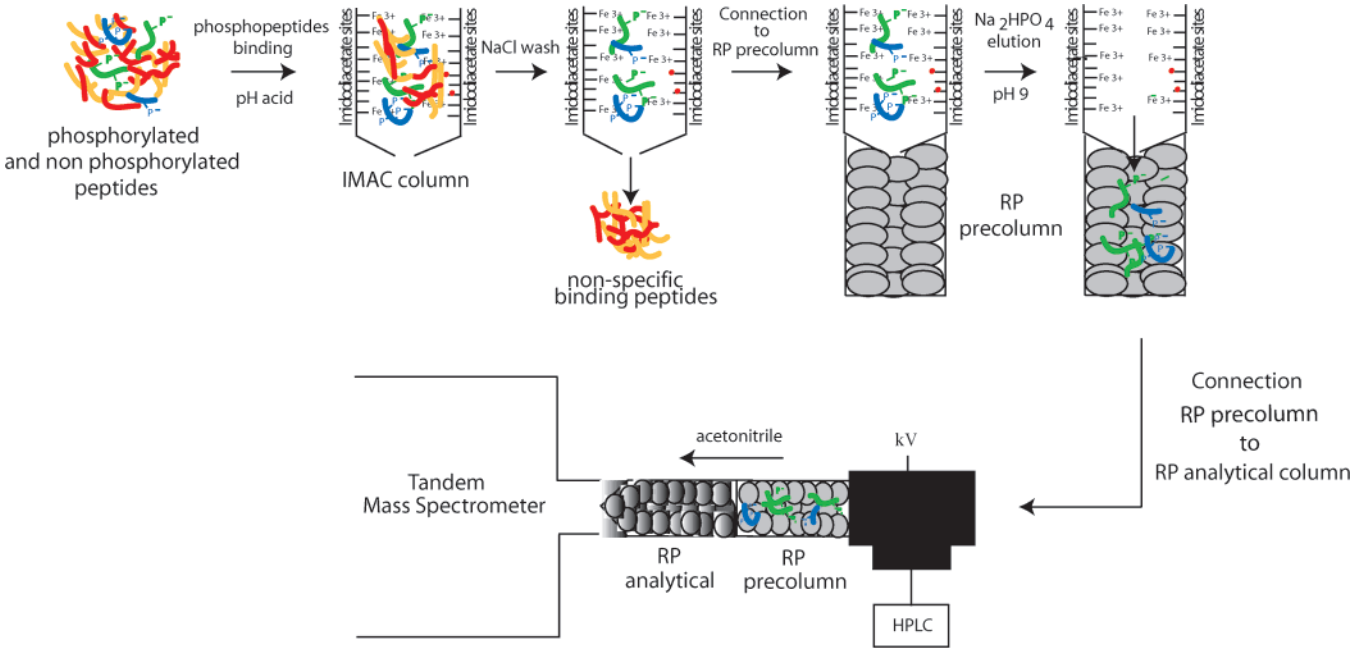

Figure 13. Phosphoproteomic analysis by immobilized metal affinity chromatography (IMAC) combined with reversed phase liquid chromatography (RPLC) and tandem mass spectrometry. This schematic diagram is based on the protocol established by Ficarro et al.[208] The peptide mixture containing phosphorylated and nonphosphorylated peptides is loaded, under acidic conditions, onto an IMAC column complexed with Fe^{3+}. After a salt wash (NaCl), the unbound peptides are removed and the phosphopeptides are directly eluted from the IMAC column to a RP column (RP precolumn) using a phosphate buffer (Na_2HPO_4). The RP precolumn is next connected to a RP analytical column connected online with a tandem mass spectrometer. The elution of peptides from the RP column to the mass spectrometer is performed using acetonitrile.

Ga^{3+} as the metal ion. Without the esterification step, this method has been used for phosphoproteomic analysis of mouse brain cells,[220] human pituitary gland,[221] and B lymphoma cells[222] (Table 5). Recently, by combining both methods, IMAC-Fe^{3+} and -Ga^{3+}, Puente et al.[223] performed the analysis of regulatory kinase pathways of myogenesis. To increase the resolving power of peptide separation, the IMAC-RPLC system has been coupled to ion exchange chromatography. After peptides are separated by ion exchange chromatography, elution fractions are collected before IMAC-RPLC analysis. Combining anion exchange chromatography with IMAC-Fe^{3+}/RPLC/MS/MS, Nuhse et al. performed phosphoproteome analysis of membrane proteins from plants.[224] SCX has been also used as the first separation dimension, taking advantage of the early elution of phosphopeptides.[165,166] The SCX/IMAC-Fe^{3+}/RPLC separation system has been described for the phosphoproteomic analyses of mouse postsynaptic density[225] and synapses[226,] and for the vasopressin signaling pathway in renal cells.[227] Combined with a stable isotope labeling strategy, using stable isotopic labeling of amino acids in cell culture (SILAC), Gruhler et al.[228] performed a quantitative analysis of the yeast pheromone pathway. Most of the phosphoproteome analyses using IMAC have been performed in an offline mode, because the use of salts is incompatible with MS analyses. Recently, Wang et al.[229] developed an automated online IMAC/RPLC separation using switching valves. The peptides, eluted from IMAC with ammonium phosphate, are trapped on a C18-RP column before analysis by RPLC. As described for the SCX-RPLC separation system, the C18-RP trap column allows for desalting.

As an alternative of IMAC, Pinkse et al.[230] developed an automated online 2D HPLC system using titanium dioxide based solid-phase material. TiO_2 has amphoteric ion-exchange properties able to retain organic phosphates. In the TiO_2-RPLC separation system, phosphopeptides are trapped on the TiO_2 column under acidic conditions and eluted to the RP column under alkaline conditions, using ammonium bicarbonate buffer[230] or ammonium hydroxide.[231] Peptides are next analyzed by RPLC/MS/MS. The protocol of Pinkse et al. was later improved by Larsen et al. ,[231] to decrease the binding of non-phosphorylated peptides with acidic function that can be retained on TiO_2. This strategy has been recently applied for the phosphoproteomic analysis of chicken eggshell matrix.[232] Finally, chemical derivatization methods based on β-elimination have been developed to enrich for phosphopeptides, but these methods have several chemical conversion steps, and the recovery is therefore likely to be low.[233,234]

4.3.2.2. Glycopeptide Enrichment Techniques. Glycosylation is one of the most common post-translational modifications playing roles in immune recognition, receptor binding, inflammation, and pathogenicity. Abberant patterns have been associated with diseases such as diabetes and cancer, highlighting the need of methods for glycoprotein identification to provide insights into the identification of new therapeutic targets. Protein glycosylation is characterized by the covalent attachment of glycans on amino acids in asparagine-linked (N-linked) or serine/threonine-linked (O-linked) form. The N-linked glycans are formed with either high-mannose type oligosaccharides or other types of saccharides, such as glucose. The O-linked glycans consist of *N*-acetylgalactosamine (GalNac) followed by other carbohydrates such as galactose or sialic acid.

Two enrichment methods have been recently described for global identification of N-linked glycoproteins. The first method is based on the coupling of N-linked glycoproteins onto a solid support using hydrazine chemistry.[235] Hydrazide chemistry allows capturing of glycoproteins via hydrazone covalent bonds formed between hydrazide and the aldehyde

Table 5. Phosphoproteomics Analysis by Immobilized Metal Affinity Chromatography (IMAC) Combined with Reversed Phase Liquid Chromatography (RPLC) and Tandem Mass Spectrometry

application	phosphopeptide enrichment	elution salt	mass spectrometer	no. of phospho-sites	database searching
S. cerevisiae phosphoproteome[208]	esterification−IMAC-Fe^{3+}	50 mM Na_2HPO_4, pH 9	LCQ	383 S/T/Y[a]	SEQUEST
rat liver phosphoproteome[211]	esterification−IMAC-Fe^{3+}	250 mM Na_2HPO_4, pH 8	QSTAR XL	300 S/T	MASCOT
colon cancer cells phosphoproteome[210]	esterification−IMAC-Fe^{3+}	250 mM Na_2HPO_4, pH 8	QSTAR XL	238 S/T/Y	MASCOT
phosphotyrosine profiling in human T cells[214]	esterification−IMAC-Fe^{3+}	50 mM K_2HPO_4, pH 9	LCQ	65 T	SEQUEST
phosphotyrosine profiling in human T cells[215]	esterification−IMAC-Fe^{3+}	50 mM Na_2HPO_4, pH 9	LCQ DECA	N/A[b]	SEQUEST
human cardiac cells phospho-proteome[212]	esterification−IMAC-Fe^{3+}	50 mM Na_2HPO_4, pH 9	LCQ DECA	75 S/T/Y	SEQUEST MASCOT
mouse postsynaptic density phospho-proteome[213]	esterification−IMAC-Fe^{3+}	100 mM Na_2HPO_4, pH 9	Q-TOF	83 S/T	MASCOT
capacitation in human sperm cells[216]	esterification−IMAC-Fe^{3+}	50 mM Na_2HPO_4, pH 9	LCQ	56 S/T/Y	SEQUEST
serum starvation in human lung cancer cells[217]	esterification−IMAC-Fe^{3+}	100 mM Na_2HPO_4, pH 7.5	LCQ Deca	38 S/T	MASCOT SCOPE
8-bromo-cGMP effect in pregnant rat uteri[219]	esterification−IMAC-Fe^{3+}	0.5% NH_4OH, pH 9.5	Q-TOF	N/A	MASCOT
INFa signaling pathway in human T cells[218]	esterification−IMAC-Fe^{3+}	2% NH_4OH	Q-TOF Ultima mass	57 S/TY	MASCOT
plasma membrane phosphoproteome of *A. thaliana*[224]	IMAC-Fe^{3+}	dilute NH_3, pH 10.5, or 50 mM $(NH_4)_2HPO_4$, pH 9	Q-TOF Ultima mass	N/A	MASCOT
yeast pheromone signaling pathway[228]	IMAC-Fe^{3+}	50 μM KH_2PO_4/NH_3, pH 10	LTQ-FT	729 S/T/Y	MASCOT
mouse postsynaptic density phospho-proteome[225]	IMAC-Fe^{3+}	100 mM Na_2HPO_4, pH 9	Q-TOF	723 S/T/Y	MASCOT
mouse synapse phosphoproteome[220]	esterification−IMAC-Ga^{3+}	200 mM Na_2HPO_4	Q-TOF	331 S/T/Y	MASCOT
B lymphoma cells phosphoproteome[222]	IMAC-Ga^{3+}	200 mM Na_2HPO_4	LCQ DECA	193 S/T/Y	SEQUEST
human pituitary gland phospho-proteome[221]	IMAC-Ga^{3+}	200 mM Na_2HPO_4, pH 8.4	LCQ DECA	9 S	SEQUEST MASCOT
vasopressin signaling pathway in renal cells[227]	IMAC-Ga^{3+}	200 mM Na_2HPO_4, pH 8.4	LTQ-FT	714 S/T/Y	SEQUEST
myoblasts phosphoproteome[223]	IMAC-Fe^{3+} IMAC-Ga^{3+}	250 mM Na_2HPO_4, pH 9	Q-TOF Ultima	47 S/T/Y	MASCOT

[a] S/T/Y are the single letter amino acid codes for serine, threonine, and tyrosine. [b] N/A, values are not available.

groups from oligosaccharides by conversion of the cis-diol groups via periodate oxidation. The nonglycosylated peptides are released by proteolysis, whereas glycopeptides are released by peptide *N*-glycosidase F (PNGase F) treatment. The *N*-glycopeptides are next analyzed by LC/MS/MS or LC-MALDI tandem mass spectrometry. The site of carbohydrate attachment to N-linked glycoproteins is found within a consensus sequence N−X−S(T), where X is any amino acid except proline. This consensus sequence is typically used to increase the confidence in the identity of the modification site after database searching. Morever, the glycan structures are heterogeneous and *N*-glycopeptides masses are unpredictable. The PNGase F treatment overcomes this problem by removing the glycan motif and causing a mass shift of 1 mass unit by deamidation of the asparagine to aspartic acid. The limit of this approach is that any information of the glycan structure is lost. This strategy was first described by Zhang et al. for the identification of glycoproteins from cancer cell lines and human serum.[235] Because the majority of human plasma proteins are believed to be glycosylated, the glycoprotein enrichment strategy has been widely used to reduce sample complexity in human plasma proteome analyses.[236] Sun et al. have recently used this method for *N*-glycopeptide enrichment from a cisplatin-resistant ovarian cell line, where the enrichment step at the peptide level circumvents the problem of membrane protein solubility resulting in reduced sample complexity.[237]

The second method is based on *N*-glycopeptide enrichment by lectin affinity chromatography (LTA). Lectins are protein receptors found in a variety of species from plants to humans. They recognize and bind specifically to monosaccharides and can be classified by the sugar that they recognize. For example, the lectins BS-II from *Bandeiraea simplicifolia* and the agglutinin LTA from *Lotus tetragonolobus* have been used to select glycopeptides containing *N*-acetylglucosamine (GlcNAc)[238] or α-fucose(1−6)-βglcNAc-AsN.[239,240] After binding on the lectin resin, glycopeptides are released by PNGase F treatment prior to fractionation by RPLC and MS/MS analysis. Using this strategy, Geng et al. identified *N*-glycoproteins from known and unknown structures from human blood serum and cancer cell line.[238] Coupled with a stable isotopic labeling strategy this method allowed the identification of aberrant glycosylations of cancer patients.[240] To increase the fidelity of the glycosylation site identification, specific tagging of the modified residues has been performed using $H_2{}^{18}O$ during PNGase F treatment. This leads to the production of ^{18}O-labeled aspartate residues that serve as unique identification markers for the glycosylation sites. This method has been successfully applied for the N-linked glycoproteomics analysis of *T. cruzi*[241] and combined also with a 2D-LC separation system into a systematic approach called IGOT (isotope-coded glycosylation-site-specific tagging) for the large-scale identification of glycoproteins in *C. elegans*.[242] Compared to hydrazine chemistry methods, LTA chromatography is advantageous because glycan motifs are not destroyed and can be predicted depending on the lectin that has been used. However, LTA chromatography is limited with respect to O-glycoproteomics analysis because of the low affinity of lectins for *O*-glycan motifs. The lectin wheat germ agglutinin (WGA), which binds weakly to GlcNAc and sialic acid motifs, has been used for a direct GlcNAc-modified peptide enrichment for the glycoproteomic

analysis of mouse postsynaptic density.[243] To increase the specificity, an enzymatic tagging strategy has been developed by which a radiolabeled galactose was transferred to the terminal GlcNAc residues after galactosytransferase treatment.[244] Next, the labeled glycopeptide could be enriched from complex mixtures by *Ricinus communis* agglutin I (RCA I) lectin affinity chromatography by retardation during the elution process.[244]

O-Glycopeptide enrichment has been more successfully performed using serial lectin affinity chromatography (SLAC). On the basis of the knowledge that lectins interact with a specific structural motif in a glycan,[245] some methods have combined different lectins in multiple affinity steps to select different types of glycan motifs. Durham and Regnier[245] used SLAC for the study of O-glycosylation of the human proteome using a combination of different types of lectins, concanavalin A (ConA) and jacalin, which have respective specificities for N-type and O-type glycans. The ConA affinity chromatography is performed first to remove the high mannose, hybrid, and biantennary *N*-glycopeptides before selection of O-modified glycopeptides on to the jacalin column.[245] Qiu and Regnier used a combination of *Sambucus nigra* agglutinin (SNA) and ConA to identify sialic acid modified glycoproteins in human serum.[246] This strategy has been also used for N-glycoproteomics analysis as described by Hirabayashi et al. using ConA combined with galectin LEC-6 (Gal6) to identify glycoproteins from *C. elegans*.[247] Glycoproteins were first selected using ConA, specific for high mannose type glycans from which the flow through was used for Gal6 lectin binding, specific for complex type glycans. After proteolytic digestion of the enriched glycoproteins, the glycopeptides were selectively recaptured with the same types of columns.[247]

O-Glycans are labile residues and are often lost during collision-induced dissociation, decreasing the efficiency of peptide fragmentation. The identification of the modified residues is therefore difficult. In addition, the consensus site for O-glycosylation is not well-defined. To overcome these limitations, chemical derivatization techniques have been used. In chemical derivatization, the glycan motif is replaced or modified by a covalent residue used as a marker to identify the sites of modification and provides a tool for glycopeptide enrichment by affinity chromatography. Wells et al. described a method for the rat nuclear pore complex analysis, which relies on β-elimination followed by Michael addition with dithiothreitol (BEMAD).[248] *O*-GlcNac is removed from glycopeptides by β-elimination and replaced by dithiothiothiol (DTT) to be enriched by thiol chromatography prior LC/MS/MS analysis. Although BEMAD can be used to map sites, it requires extensive controls to distinguish a peptide that contains a phosphate, *O*-GlcNac, or a complex O-linked carbohydrate group. As an alternative, Khidekel et al. performed a high-throughput analysis of *O*-GlcNac glycosylated proteins from the mammalian brain using a strategy by which GlcNac modified proteins are specifically labeled by a ketone-biotin tag after engineered galactosyltranferase treatment.[249] After proteolytic digestion of the labeled proteins, glycopeptides are selected using avidin affinity chromatography prior to LC/MS/MS analysis.[249] Chemical derivatization has been proven to be a powerful tool but still suffers from lack of specificity.

4.3.3. Isoelectric Focusing and Reversed Phase

Currently, the most common technique used for the first dimension in shotgun proteomics is SCX chromatography. As an alternative approach, isoelectric focusing (IEF) performing peptide separation based on peptide p*I* values has been employed. Peptide separation based on p*I* is orthogonal to RPLC. In addition, p*I* can be used as a constraint for database searching to filter out the false-positive identifications, which increases the confidence of sequence matches. IEF for peptide separation was first described using immobilized pH gradient isoelectric focusing (IPG-IEF) on IPG-IEF gel strips.[250] IPG-IEF gel strips are made with buffering acrylamide derivatives containing either a free carboxylic group or a tertiary amino group to create the pH gradient. IPG-IEF gel strips are typically used for the first dimension of protein separation by 2D gel electrophoresis. Cargile et al.[250] explored the use of IPG-IEF as a first dimension for peptide separation in shotgun proteomics. Peptides loaded onto IPG-IEF strips migrate through the IPG-gel strip upon application of an electric field until they reach the pH at which their net charge is zero. Next, gel strips are cut into sections and peptides are extracted to be further separated by RPLC and analyzed by tandem mass spectrometry.[250−252] This technology has been applied to *E. coli*[250]- and *Rattus norvegicus*[253] proteome analysis. Compared to SCX as the first dimension, the IPG-IEF separation leads to 13% more protein identification from testis samples of *R. norvegicus*, including peptides having a p*I* range that could not be resolved by SCX separation.[253] A main advantage of this technique is that the p*I* can be used as a tool for filtering protein and peptide database search results.[254]

To increase the sample recovery, which can be limited by the need to extract peptides from gel matrices, isoelectric focusing peptide separations have been performed in solution, by capillary electrophoresis (CIEF) and free-flow electrophoresis (FFE-IEF). The basic instrument for a CIEF experiment consists of a high-voltage supply (0−30 kV), a fused silica capillary, two buffer reservoirs, two electrodes, and a detector. To perform CIEF, ampholytes or zwitterions are used to create a pH gradient inside the capillary. The mixture of analytes will migrate under the electric field toward the cathode or anode until the analytes reach their isoelectric point. The first online system combining CIEF and RPLC separation mode was developed by Chen et al. for the 2D separation of *Drosophila* protein extract using UV detection.[255] Additional studies describe improvements in sample loading and analyte concentration.[256] The same authors developed a system coupling CIEF-RPLC for peptide separations connected online with a tandem mass spectrometer for the analysis of yeast cell lysates.[257] In this approach, peptides are first loaded onto a fused silica capillary and, after electrofocusing (pH range from 3 to 10), trapped on a C18-trap RP column prior to RPLC separation. CIEF and RPLC have also been connected by switching valves. The authors identified 1894 unique peptides and 1132 proteins over a pH range between 3.8 and 10.2.[257] Using the same system, analyses of tumor tissue[258] and human salivary proteome[259] have been performed.

Whereas high resolution can be achieved by a CIEF/RPLC system, proteomics applications are limited by the small loading capacity of CIEF. FFE-IEF has emerged as a good alternative. Using FFE-IEF, the separation is performed in a continuously flowing solution without system regeneration by which analytes are continuously injected and collected without any limitation on the total amount analyzed. The FFE-IEF system consists of two plates, between which ampholytes are introduced to create a pH gradient, usually

between 3 and 10 pH units. Analytes in solution are continuously injected and separated upon the application of an electric field. This separation method is mostly combined in an offline mode to a RPLC separation followed by MS/MS analysis. The strategy has been applied to resolve the chromatin-enriched fraction from *S. cerevisiae*[260] and the human saliva proteome.[261] This technology has been recently combined with RPLC performed at different pH values to perform a four-dimensional separation of peptides from human plasma.[262] In several proteomics applications, (hydroxy)prolylmethyl cellulose HPMC is used to minimize the electroosmotic flow during FFE-IEF. However, the large molecular weight of HPMC and its chemical properties interfered with the MS signal. Malmstrom et al. used mannitol and urea as a promising alternative to resolve the *Drosophila melanogaster* proteome.[263] Although FFE-IEF combined with RPLC has proved to be an efficient separation system for complex mixtures, some experimental challenges remain, especially directly interfacing FFE and RPLC-MS online, which currently limits the efficiency of this approach.

4.3.4. Reversed Phase and Capillary Zone Electrophoresis

As previously shown, capillary electrophoresis (CE) and liquid chromatography (LC) have been coupled in the CE-LC mode, but CE is a fast separation method compared to LC. This makes it an appropriate choice as the second dimension because of its ability to sample the first dimension at a relatively high frequency. CZE is the simplest form of capillary electrophoresis by which each component in the sample is separated according to its apparent mobility or mass-to-charge ratio. CZE has been used for peptide separation first separated by LC. Compared to the LC−LC system, fewer proteomics studies have been performed using an online LC-CZE-MS platform. This is because they are limited, first, by the difficulty to interface LC-CZE in an online fashion and, second, by the challenge of interfacing the LC-CZE system online with a mass spectrometer. LC-CZE systems are frequently connected online using a switching valve or fluid tee. Peptides or proteins are often detected by UV or LIF. Bushey and Jorgenson developed the first automated RP-CE system, which coupled a RPLC column with a fused silica capillary connected with sample loops and a switching valve.[44] This system was used to separate peptide mixtures using UV or LIF for detection.

The RP-CZE interface was later improved by using transverse flow-gated interfaces or optical gating. The transverse flow-gated interface was first reported by Lemmo and Jorgenson[264] and improved by Hooker and Jorgenson.[265] Using this interface, the LC column inlet and the CE capillary outlet are positioned in opposite directions separated by a narrow channel (Figure 14). During a run, the CE buffer is continuously flowing to the narrow channel driving the LC effluents to waste. To perform a sample separation, the buffer flow is stopped and the effluents from the LC column migrate to the CE capillary by electromigration. This system eliminates the use of sample loops, which require extra columns and make the interface more complicated.

This system was further improved using optical gating, which performs faster injections.[45,266,267] The optical gating interface couples the LC and CZE by a tee. The effluent from the liquid chromatograph migrates to the CZE capillary by electromigration. A laser beam is focused on the CE capillary just beyond the interface tee, which degrades the fluorescently labeled sample passing through the capillary. When an injection is desired, the beam laser is blocked, allowing the undegraded sample to be introduced in the CE capillary for further separation (Figure 15).

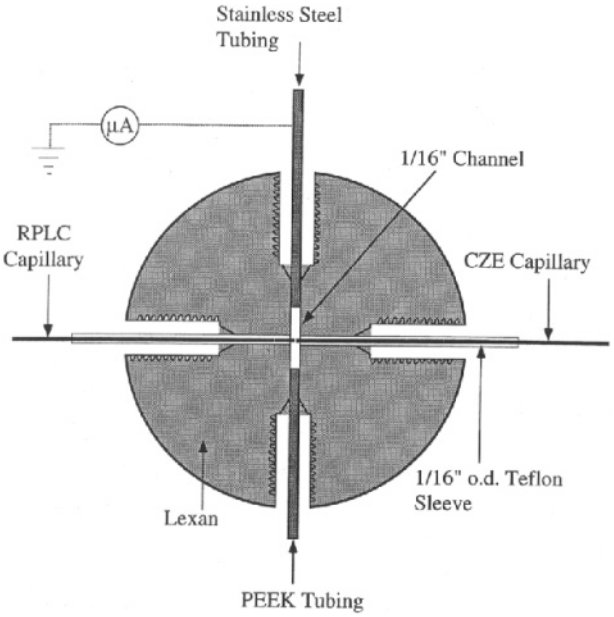

Figure 14. Schematic of the clear flow gating interface for the coupling of reversed phase high-performance liquid chromatography with capillary zone electrophoresis. Reprinted with permission from ref 265. Copyright 1997 American Chemical Society.

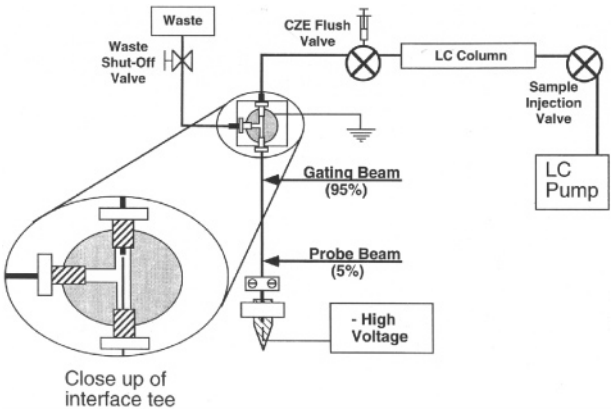

Figure 15. Schematic diagram of the system coupling reversed phase high-performance liquid chromatography system with fast capillary zone electrophoresis. Reprinted with permission from ref 266. Copyright 1995 American Chemical Society.

Recently, Bergstrom et al. reported a 2D LC-CZE system for peptide mixture analysis interfaced with electrospray ionization Fourier transform ion cyclotron resonance (FTI-CR).[268] The challenge of this system relies also on the development of a new interface for the online coupling of LC to CZE. CZE has also been coupled to affinity liquid chromatography by Cao et al., who developed an online IMAC-CZE-ESI-MS system for separating and resolving phosphopeptide mixtures.[269] The peptides were first loaded on the IMAC column charged with Fe^{3+}, and after elution at low pH, peptides were transferred online for CZE separation and directly analyzed by mass spectrometry.[269]

5. Database Searching and Dataset Assembly

5.1. Collision-Induced Dissociation

Several methods exist for peptide fragmentation, but the most commonly used method is collision-induced dissociation (CID).[270] CID is a mechanism whereby the peptide of interest is allowed to collide with an inert gas, often helium, resulting in peptide fragmentation.[271] The amino acid composition of a peptide is identified from a consecutive series of ions having a mass that corresponds to the residue masses of amino acids.[272] The bond cleavage of a peptide can occur at three different bonds, C−C, C−N, and N−C, which yields six different fragments designated a_n, b_n, and c_n when the positive charge is at the N-terminal side and x_n, y_n, and z_n when the positive charge is at the C-terminal side. Differences in peptide fragmentation can occur at low (<100 eV) and high energy (>1000 eV). CID of peptides is typically performed under low energy, which primarily gives rise to y_n and b_n ions.[270] Fragmentation of peptides, whether by low or high energy, is dependent upon the physicochemical properties of the amino acids present in the peptide.[273,274] For example, charged ions from small peptides containing few or no basic amino acids have comparable y_n and b_n ion abundance, whereas larger peptides predominately result in b_n ions.[275] Interpretation of the tandem mass spectra is based upon the mechanisms and fragmentation pathways of the ions.

Proteomic studies of such complex peptide mixtures result in thousands to millions of MS/MS spectra (Figure 16). The analysis of such large datasets is a daunting task, and interpretation of such complex uninterpreted MS/MS spectra requires sophisticated algorithms. Over the past decade, several methods have been developed for analyzing such complex datasets, including database searching algorithms, which will be the main focus in this section. Although there are many database searching algorithms for proteomics analysis[276−286] (Table 6), only the most commonly used algorithms SEQUEST,[6] MASCOT,[287] OMSSA,[288] and X!Tandem[289] will be discussed in detail.

5.2. Database Searching Algorithms

5.2.1. SEQUEST

The SEQUEST algorithm was the first algorithm designed for high-throughput peptide identification from uninterpreted MS/MS data.[6,8,290] SEQUEST uses a cross-correlation function (Xcorr) to evaluate similarities between an experimental mass spectrum and a predicted spectrum from a database. Initially, a preliminary score (S_p) is calculated, which restricts the number of peptide sequences that will be used for calculating the Xcorr. S_p takes into consideration the intensity of the experimental MS peaks when matching them with the predicted MS peaks as well as the length of the peptide. Following the S_p, the Xcorr is calculated between the experimental spectra and the top 500 predicted spectra determined from the S_p. The Xcorr is independent of database size and determines the best match between the experimental and predicted spectra. However, Xcorr is based upon the characteristics of the analyzed peptide, such as mass and charge state, and has been observed to show bias for larger peptides. Therefore, normalization of Xcorr values across peptide features such as peptide mass and charge state has been performed to correct this.[291] The normalized Xcorr value ranges from 1.0 to 0.0, with 1.0 being the best matched experimental and predicted spectra.[291] Another important score that is determined by SEQUEST is the ΔC_n.[8,290] In essence, the ΔC_n value represents how different the best Xcorr is from lower Xcorr values. The ΔC_n value is dependent upon the database size due to the possibility of having similar sequences in larger databases, as well as the possibility of random matches between peaks in the experimental and theoretical spectra. The goal of all database searching algorithms is to interpret experimental tandem mass spectra and assign a confidence score to the interpretation. Peptides can have multiple charge states depending on their length and sequence, and in a MudPIT experiment, for example, any given peptide could be detected and identified in multiple charge states. For example, the SEQUEST interpretation of a peptide that appeared as both a +1 and a +2 peptide is shown in Figure 17, and the SEQUEST interpretation of a peptide that appeared as both a +2 and a +3 peptide is shown in Figure 18.

5.2.2. MASCOT

MASCOT, developed by Perkins et al.,[287] is based on the probability-based scoring algorithm MOWSE (Molecular Weight Search).[292] When using MASCOT, the MS/MS ion search will match the experimental MS/MS spectrum with a theoretical spectrum from the database of protein sequences and group these peptides into protein matches. The MOWSE algorithm will calculate the probability that the observed match between the experimental data and mass values calculated from a candidate peptide or protein sequence is a random event. This probability should be small for the correct peptide assignment because many peaks will match. The score reported by MASCOT is calculated on the basis of this probability (P) as $-10 \times \log(P)$. The score threshold for MASCOT is 45.

5.2.3. X!Tandem

X!Tandem is a more recent database searching algorithm developed for the identification of peptides from MS/MS spectra.[289,293] X!Tandem, like MASCOT, uses a probability-based scoring method instead of a cross-correlation scoring method as seen in SEQUEST. X!Tandem initially matches experimental MS/MS spectra with predicted spectra from a protein database, similar to both MASCOT and SEQUEST. Next, y and b ion peaks that are not present in both the experimental and predicted spectra are removed, thereby allowing the spectra to be simplified. An S_p value is then calculated using only the intensities of the y and b peaks, differing slightly from the S_p calculated by SEQUEST, which is then used to calculate a hyperscore.

X!Tandem assumes that the peptide with the highest hyperscore is the correct peptide match. Histograms are then made of all the hyperscores, and data from the right side of the histogram plots are log-transformed. The log-transformed data are fit to a straight line, which is then extrapolated through the X-axis. The point at which the extrapolated line reaches the largest hyperscore (log of hyperscore) is called the E value. This value is comparable to the ΔC_n in SEQUEST. Although X!Tandem has already assumed that the predicted spectra with the highest hyperscore is the correct match, this match is considered to be significant if the E value is greater than the point at which the extrapolated line crosses the X-axis. The main advantage of X!Tandem over other database searching algorithms is its speed, which is ~200 times faster when performing general searches and

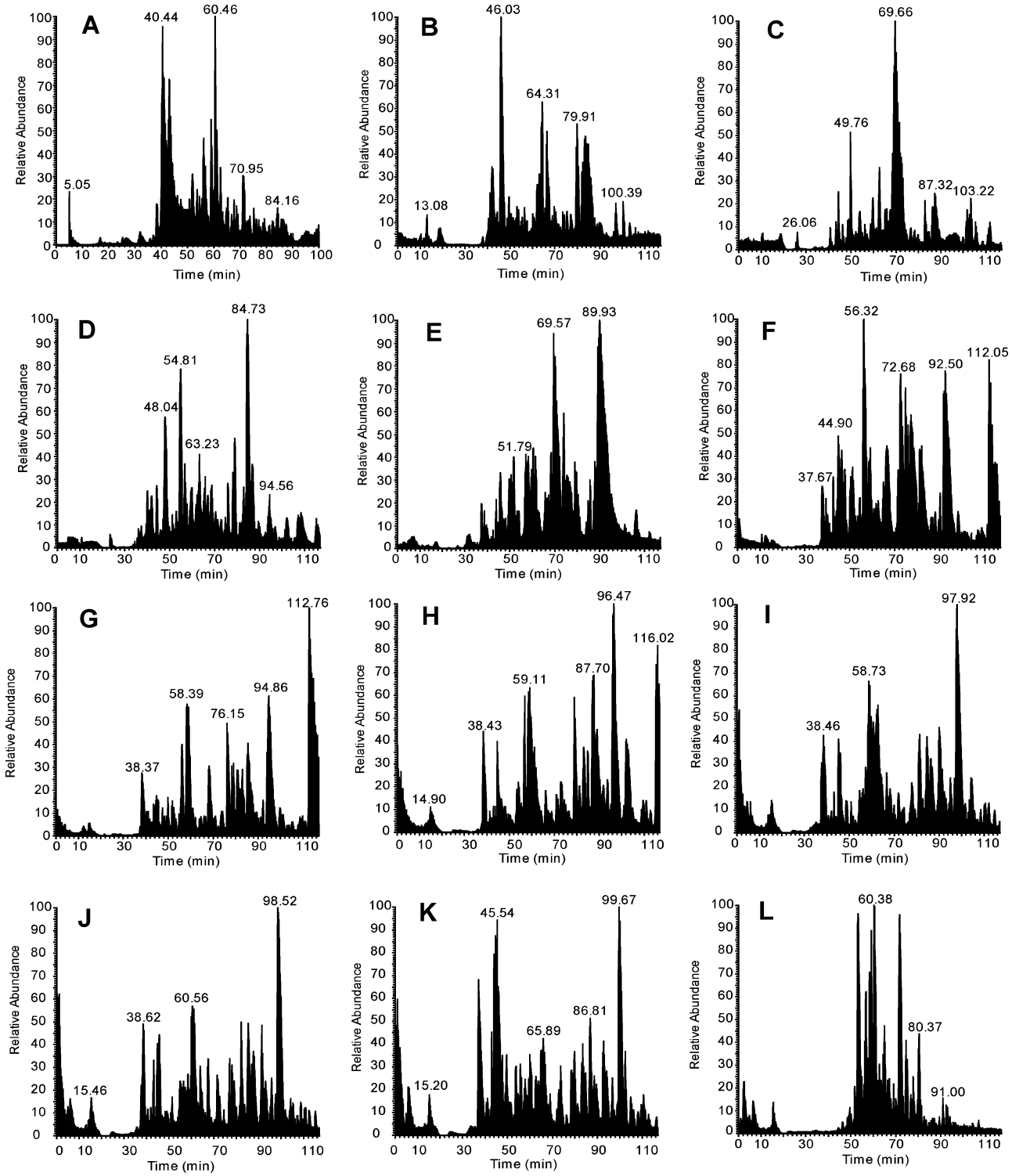

Figure 16. Base peak chromatograms from a MudPIT analysis of enriched *S. cerevisiae* plasma membranes. A complex peptide mixture generated from a *S. cerevisiae* sample enriched for plasma membranes was analyzed on a triphasic MudPIT column using a 12-step chromatography analysis as described previously.[158] Shown in (A) is the first step, which contains a gradient of 100% buffer A [5% acetonitrile (ACN) and 0.1% formic acid] for 10 min to 40% buffer B (80% ACN and 0.1% formic acid) over 60 min up to 100% buffer B by 100 min. Shown in (B) is the second step, which contains 3 min of buffer A, followed by 2 min of 5% buffer C (500 mM ammonium acetate, 5% ACN, and 0.1% formic acid), followed by 5 min of buffer A, followed by a linear gradient to 15% buffer B over 15 min followed by a linear gradient to 45% buffer B over 92 min. Shown in (C−K) are the 3rd through 11th steps, which are the same as the second step except the 2 min buffer C step is 10, 20, 30, 40, 50, 60, 70, 80, and 90% buffer C, respectively. Shown in (L) is the last step, which consists of 2 min of 100% buffer A, followed by 20 min of 100% buffer C, followed by 5 min of 100% buffer A, followed by a linear gradient to 20% buffer B over 10 min, followed by a 68 min linear gradient to 70% buffer B, followed by a 5 min linear gradient to 100% buffer B, followed by 5 min at 100% buffer B, and finally a 2 min wash with 100% buffer A.

∼1000 times faster when searching for specific modifications such as PTMs. Another important improvement for X!Tandem is the ability to use parallel processing for accelerated analyses.[294] X!Tandem is open source and a freely available

Table 6. Commonly Used Database Searching Algorithms for MS/MS Analysis

algorithm	type[a]	URL
Sonar[281]	score	http://bioinformatics.genomicsolutions.com/service/prowl/sonar.html
ProSight PTM[283]	score	https://prosightptm.scs.uiuc.edu
SALSA[282,284]	score	http://www.mc.vanderbilt.edu/lieblerlab/salsa.php
SEQUEST[6]	score	http://fields.scripps.edu/sequest
MS-Tag[277−279]	score	http://prospector.ucsf.edu/
Phenyx[280]	prob	http://www.phenyx-ms.com/about/about_phenyx.html
ProbID[286]	prob	N/A
SCOPE[276]	prob	N/A
MASCOT[287]	prob.	http://www.matrixscience.com
X!Tandem[289,293]	prob	http://www.thegpm.org/TANDEM/index.html
OMSSA[288]	prob	http://pubchem.ncbi.nlm.nih.gov/omssa/
Peaks[285]	prob	http://www.bioinfor.com:8080/peaksonline/

[a] The "type" indicates the scoring method implemented in each algorithm: score, a score is determined by protein or peptide to determine a goodness of match between a theoretical vs observed mass spectrum; prob, probabilisitc algorithms determine directly the confidence in a match between a theoretical vs observed mass spectrum.

database searching approach.

5.2.4. Open Mass Spectrometry Search Algorithm (OMSSA)

The OMSSA[288] scores peptide hits using a probability-based method that compares experimental fragments with those calculated from the library produced by an in silico digestion of a protein database. The statistical model is similar to the one used in the BLAST algorithm.[295] It uses a probabilistic model based on a Poisson distribution. Calculating a distribution of random matches allows the significance of a hit to be expressed as the probability of the hit being random. Each hit is given an *E* value, which represents the expected number of random hits from a search library to a given spectrum such that the random hits have an equal or better score than the hit. Note that a hit with a lower *E* value is more significant than a hit with a high *E* value. For example, an *E* value of 2.0 indicates that two hits can be expected from a random search of a sequence search library that will have an equal or better score than the current hit being scored. A particulary appealing feature of OMSSA is that it is freely accessible to the public.

5.3. Statistical Validation

A large problem with peptide and protein identification via database searching is that some top scoring peptide matches are falsely identified.[169] Therefore, the resulting peptide and protein identifications must be filtered by threshold models or statistical validations. The process by which database searching algorithms identify peptides from complex MS/MS spectra appears to be fairly straightforward. Database searching algorithms use predicted MS/MS spectra within the database and compare them to an experimental MS/MS spectrum. The algorithms then determine the best-matched predicted spectrum to determine the peptide sequence. However, there is no search algorithm to date that is 100% accurate in identifying peptides. Moreover, searching algorithms usually generate a significant number of incorrect peptide assignments.[296] This can be alleviated by the use of statistical models to help validate the peptide assignments generated by the numerous search algorithms.[297]

Peptide Prophet is a statistical tool developed by Keller et al.[297] First, Peptide Prophet converts the scores that are generated from a particular searching database (for example, XCorr, ΔC_n, and S_p scores from SEQUEST) into a single discriminant score.[297] Next, a histogram is made that contains the scores generated by Peptide Prophet, which indicates the distribution of correct and incorrect values, assuming that the distributions are standard statistical distributions. The correct and incorrect distributions are drawn using curve-fitting and Bayesian statistics. These distributions are then used to calculate the probability that a match is correct on the basis of a given discriminant score value. Peptide Prophet and the use of one discriminant value allow data from multiple instruments or different software to be compared.[297]

Although improvements have been made for validating peptide assignments, the objective of proteomics research is to identify proteins from complex mixtures. A statistical model called Protein Prophet has been developed, which uses assigned peptides from MS/MS spectra to compute protein probabilities in a given sample.[298] This open-source software can quickly and accurately determine the probability of a protein's existence in a given sample using MS/MS spectra from high-throughput analyses of complex peptide mixtures.[298] Protein Prophet uses accurately assigned peptides and the probabilities that they are correct to calculate the probability that a particular protein is present in the sample. Accurate peptide probabilities can be calculated from the Peptide Prophet software.[297] After peptides are assigned to individual MS/MS spectra, the peptides are then clustered according to their corresponding proteins. According to Choudhary et al., peptides that have been correctly assigned tend to correspond to proteins that are identified by other correctly assigned peptides, which they call "multihit" proteins.[299] Conversely, incorrectly assigned peptides tend to correspond to proteins that have been identified by other incorrect peptide assignments.[299] From these observations made by Choudhary et al., Protein Prophet clusters the accurately assigned peptides into groups that correspond to the same protein.[299] This allows further validation of the peptide assignment. Protein Prophet determines the estimated number of sibling peptides (NSP) for each protein, which are peptides that correspond to multihit proteins.[298] When using complex databases, one particular peptide can be assigned to multiple proteins. This can be due to protein paralogues, splicing variants, or redundant entries within the database. Such peptides have been described as "degenerate".[300] Protein Prophet deals with degenerate peptides by distributing them across all corresponding proteins to create a smaller protein list sufficient to account for the identified peptides.[298] Redundant database entries are reduced to a single identification, and proteins that cannot be differentiated on the basis of peptide identification are grouped together. Protein Prophet has been evaluated using MS/MS spectra from complex samples, and results suggest that Protein Prophet was able to discriminate between correct and incorrect protein identifications including identifications based on a single peptide.[301] An additional tool for probabilistic determination of protein identifications from interpreted tandem mass spectra results is PROT_PROB,[302] which serves a function similar to that of Protein Prophet.

5.4. False Discovery Rates

With the large datasets generated via shotgun proteomics analyses, false-positive identifications become important

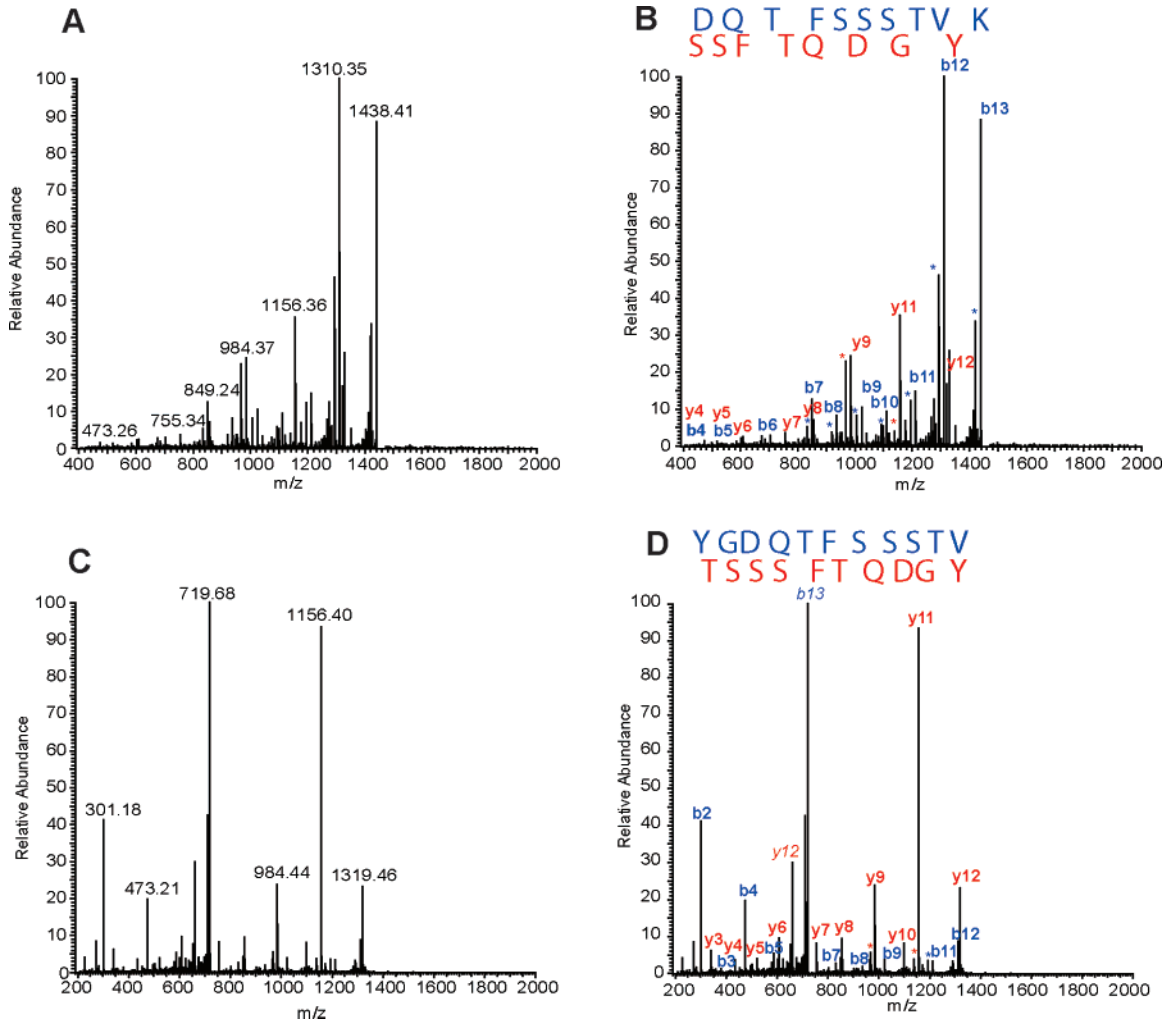

Figure 17. Interpretation of +1 and +2 peptide tandem mass spectra (MS/MS). The fully tryptic peptide K.HYGDQTFSSSTVK.R, where the period represents the trypsin cut, from the *S. cerevisiae* integral membrane protein Pma1 was detected and identified as a +1 peptide and a +2 peptide in an analysis of enriched *S. cerevisiae* plasma membranes.[158] (A) The +1 peptide prior to interpretation is shown with the relative abundance of each ion at a particular mass to charge ratio (*m*/*z*). (B) The SEQUEST interpretation of the MS/MS shown in (A) where b (blue) and y (red) ions are shown along with ammonia ion neutral losses (∗) immediately to the left of corresponding b (blue ∗) or y (red ∗) ions. The text in blue letters represents the +1 b ions series, and the text in red letters represents the +1 y ion series. The SEQUEST interpreted MS/MS had an Xcorr of 4.0185 and a DeltaCn of 0.4779. (C) The same fully tryptic peptide was seen as a +2 ion in the analysis, and the uninterpreted version of this ion is shown. (D) The SEQUEST interpretation of the MS/MS shown in (C), where b (blue) and y (red) ions are shown along with ammonia ion neutral losses (*) immediately to the left of corresponding b (blue ∗)or y (red ∗) ions. The text in blue letters represents the +1 b ions series, and the text in red letters represents the +1 y ion series. The SEQUEST interpreted MS/MS had an Xcorr of 3.6457 and a DeltaCn of 0.4500.

considerations. One approach to estimate false discovery rates (FDRs) is to include randomized (reversed or shuffled) database sequences when uninterpreted MS/MS spectra are searched.[169,303] Two different types of searches have been described: in one instance the MS/MS spectra are searched against the database of interest and a randomized database independently.[179] In a second instance the original database of interest and a randomized database are joined (concatenated) and searched simultaneously.[304] This combined target/decoy database search strategy has been proposed to be the most robust method.[305,306] Furthermore, the database search parameters and the instrumentation used are important aspects to consider when the goal is to minimize false positives.[307] An additional method proposed recently to reduce false-positive rates uses peptide centric databases[308] that may prove to be important in the future.

These methods estimate the number of false-positive identifications by counting the number of matches from a randomized sequence database using the same criteria. Estimated error rates at the protein or peptide level can be calculated using several formulas. Higdon et al.[304] calculate the FPR for protein identification by

$$\mathrm{FDR} = \mathrm{FP}/(\mathrm{FP} + \mathrm{TP}) \quad (4)$$

where FP is false positive and TP, true positive. Blaker et al.[309] divides the number of false-positive protein identifications by the total true-positive protein identifications.

$$\mathrm{FDR} = \mathrm{FP}/\mathrm{TP} \quad (5)$$

Elias et al.[306] estimated the FDR at the peptide level by doubling the number of false-positive hits and dividing by the total number of hits.

$$\mathrm{FDR} = 2\mathrm{FP}/(\mathrm{TP} + \mathrm{FP}) \quad (6)$$

In these studies, the authors use confidence values ranging from 95%[309] to 99%.[306] The major database search engines

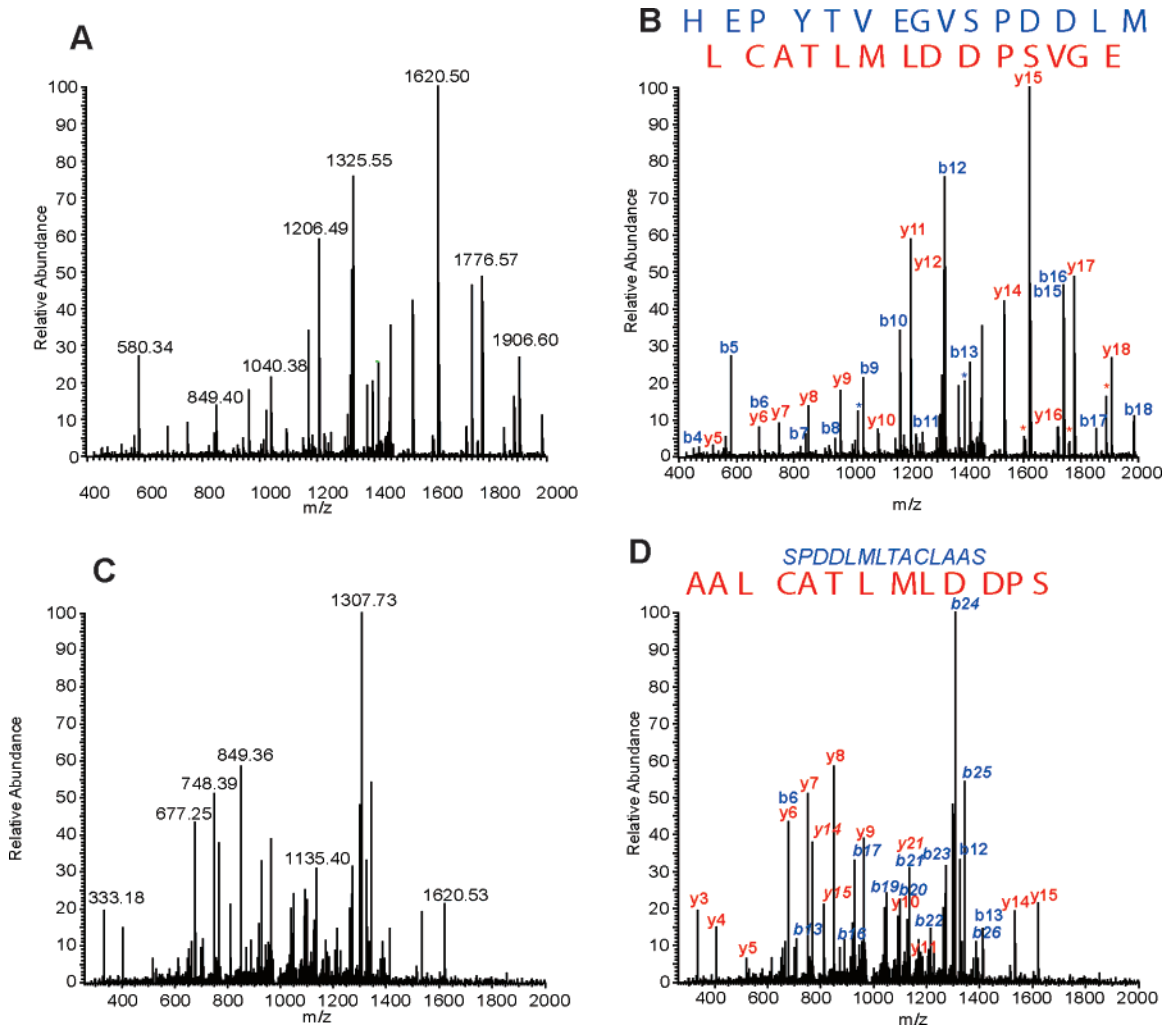

Figure 18. Interpretation of +2 and +3 peptide tandem mass spectra (MS/MS). The fully tryptic peptide K.LSLHEPYTVEGVSPDDLMLTACLAASR.K, where the period represents the trypsin cut, from the *S. cerevisiae* integral membrane protein Pma1 was detected and identified as a +2 peptide and a +3 peptide in an analysis of enriched *S. cerevisiae* plasma membranes.[158] (A) The +2 peptide prior to interpretation is shown with the relative abundance of each ion at a particular mass to charge ratio (*m*/*z*). (B) The SEQUEST interpretation of the MS/MS shown in (A) where b (blue) and y (red) ions are shown along with ammonia ion neutral losses (∗) immediately to the left of corresponding b (blue ∗)or y (red ∗) ions. The text in blue letters represents the +1 b ions series, and the text in red letters represents the +1 y ion series. The SEQUEST interpreted MS/MS had an Xcorr of 6.1168 and a DeltaCn of 0.4444. (C) The same fully tryptic peptide was seen as a +3 ion in the analysis, and the uninterpreted version of this ion is shown. (D) The SEQUEST interpretation of the MS/MS shown in (C), where b (blue) and y (red) ions are shown along with ammonia ion neutral losses (∗) immediately to the left of corresponding b (blue ∗)or y (red ∗) ions. The text in blue letters represents the +2 b ions series, and the text in red letters represents the +1 y ion series. The text in blue letters is in italics and smaller font because these are +2 b ions rather than +1 b ions as shown in the rest of the figure. The SEQUEST-interpreted MS/MS had an Xcorr of 6.1581 and a DeltaCn of 0.5443.

are also beginning to be systematically tested for their respective FDRs, given certain samples and filtering criteria.[310] In general, the estimation of FDRs by these methods is being reported with greater frequency in shotgun proteomics papers and will likely be required in the future.

5.5. Dataset Organization

Database searching algorithms are responsible only for assigning peptide sequences to experimental MS/MS spectra. Deriving information about protein content of a sample requires additional software. The large datasets produced by database searching algorithms contain tens of thousands of peptides with different confidence levels. These peptides must be filtered and assembled to obtain valuable information pertaining to the protein content of the original sample. Although numerous programs for filtering and assembling peptides for protein identification are available, only two will be discussed in detail, DTASelect[311] and INTERACT.[312]

INTERACT is a program written to perform interactive analysis of SEQUEST MS/MS database search results. It allows multiple datasets to be analyzed using a form-based HTML interface interacting with a web-server common gate interface (CGI) program.[312] INTERACT takes in a single or multiple SEQUEST summary HTML files as input and allows the user to quickly filter and sort the search results using different criteria, such as scores, amino acid composition, protein descriptions, and enzyme digest specificity.[312]

DTASelect was developed to analyze and assemble complex proteomics data resulting from SEQUEST analysis.[311] DTASelect performs its analysis in three phases: summarization, evaluation, and reporting.[311] The summarization phase includes collecting the data from SEQUEST, such as the Xcorr and ΔC_n values. During this phase DTASelect refers back to the SEQUEST database to obtain protein sequences as well as protein names. The evaluation phase includes the filtering criteria established by the user for accepting or reject-

ing the assigned peptides generated by SEQUEST. This can include cutoff values for Xcorr or ΔC_n, as well as accepting half or fully tryptic peptides, if trypsin was used. During the reporting phase, DTASelect creates a primary report of the filtered data in an easily viewed HTML format. Within the HTML file, links to the MS/MS spectra are present for each assigned peptide. In addition to the primary report, several other reports can be created in text format, which can later be manipulated using a spreadsheet. DTASelect is a powerful tool for the analysis of assigned MS/MS spectra and for the creation of a final, confidence-derived dataset.

Additional software, called Contrast, has been developed for comparing DTASelect filtered datasets.[311] Similar to DTASelect, Contrast allows the user to select filtering criteria for comparing several DTASelect datasets.[311] When Contrast is run, the filtering criteria are applied to each DTASelect dataset chosen by the user. Contrast then creates a master list of all the proteins and loci from all DTASelect datasets of interest. The proteins in the master list are then sorted and grouped by their appearance in the DTASelect datasets. The output file is in HTML format with links to corresponding DTASelect.html files.

Given the significant challenges in organizing filtered shotgun proteomics datasets, several laboratories have designed computational approaches to make this process more efficient. McAfee and colleagues have recently described the BIGCAT platform used to store, analyze, and disseminate mass spectrometry based proteomics analyses.[313] The Emili laboratory has developed a suite of approaches[314−316] that have played critical roles in their large-scale proteomics analyses of protein localization[110] and protein interaction networks.[317,318] The abundance of software tools available for proteomics analyses makes choosing a particular platform quite challenging. However, progress has been made by the Human Proteome Organization's (HUPO) Protein Standards Initiative (PSI) to define data standards for the proteomics community. This includes standards for comparison, exchange, and verification of proteomics data.[319−324] Also involved in this effort is the PRoteomics IDEntifications database (PRIDE), which is a public database for proteomics data.[325−327]

For public repositories to work, a unified format for storing and disseminating spectral data needs to be developed. For SEQUEST, MS2 and SQT files make data analysis sharing easier.[328] As of 2006 two common file formats for encoding raw mass spectrometer data had been implemented (mzXML and mzDATA) to compare datasets generated on different instrumentation.[319,329,330] Although slightly different, mzXML and mzDATA capture information from mass spectrometry experiments and unify the results so that cross-comparisons can be made. Because of the similarities between mzXML and mzDATA, PSI has developed a new format to replace them called mzML.[320] In addition to mzML, other such repositories, PEDRo[331−333] and Proteios,[334,335] have been developed to capture information from any proteomics experiment including MALDI, 2D-gel, or LC/MS experiments. The hope is that a common and open data representation format will facilitate and improve the searching, comparing, and sharing of proteomics datasets with other open-source tools.[336−338]

6. Conclusion

The use of multidimensional separations in the field of shotgun proteomics has been critical in the initial comprehensive analyses of a number of different organisms. Proteomics analyses are frequently performed to obtain information ranging from simple protein identification to more complex post-translational modification studies and have even extended to dynamic studies in response to various stimuli. The coupling of different chromatographic techniques, primarily SCX and RP separations with the addition of affinity chromatography, has granted a view into a realm previously unknown.

Shotgun proteomics is still unable to detect and identify all of the proteins present in a given sample. The complexity of the sample peptide mixture often exceeds the separation capacity of current multidimensional systems. Given the dynamic range issue in proteomics, when one is trying to detect and identify a low-abundance protein among the background of many high-abundance proteins, multidimensional separations play a key role. However, sample prefractionation and the removal of high-abundance proteins to enrich the lower abundance proteins are necessary for reducing sample complexity. The human serum proteome is likely to span >10 orders of magnitude for the dynamic range.[339] Highly abundant proteins such as albumin and transferrin must be removed from the serum through affinity methods[176] and fractionation techniques.[340]

Other advances in separations will also have a great impact. A key aspect of the work by Giddings and colleagues demonstrated the importance of orthogonality and how this increases the number of theoretical plates in a given analysis.[16−21,341] Another way to increase the number of theoretical plates in a chromatography analysis is to have smaller particle size, which then requires higher pressures for analytical use. Indeed, ultrahigh-pressure reversed phase liquid chromatography (UHPLC) has become an active area of research.[342−345] Smaller reversed phase particles have been synthesized and used in UHPLC.[346,347] As solutions to the challenges posed by the use of UHPLC are developed,[348,349] the implementation of UHPLC in a shotgun proteomics system is now possible with the introduction of commercially available systems. In fact, early efforts to implement a ultrahigh-pressure MudPIT system have been promising.[350] However, to have a fully integrated orthogonal two-dimensional UHPLC shotgun proteomics system, research into small particle and high-pressure resistant strong cation exchange particles is needed. The coupling of such a system to advanced mass spectrometry systems should result in dramatic improvements in shotgun proteomics analyses.

Over the past few years, there have been several significant mass spectrometry advances regarding ion trap instrumentation, which has been the most widely used form of mass spectrometry in shotgun proteomics. A traditional disadvantage of ion trap instrumentation is the lack of resolution when compared to TOF-MS or FT-MS. However, in 2004 Syka et al. described the coupling of the linear ion trap (LTQ) to FT-MS and demonstrated the performance of this system on the analysis of histone modifications.[351] Following the introduction of this instrument, improvements of peptide mass measurement[352] have improved proteomics analyses of prostate cancer[353] and of histone phosphorylation.[354]

One issue with FT-MS and the LTQ/FT-MS is the need for supercooling of the magnet using liquid nitrogen and liquid helium. This makes routine maintenance of this instrument significantly more challenging than traditional ion trap instrumentation. However, for many researchers the dramatic improvements in resolution are worth it. In 2000, Makarov described a novel mass analyzer that traps ions in

an electrostatic field and in which ions orbit around an axial electrode; the analyzer has very high resolution.[355] This mass analyzer became known as the Orbitrap and was subsequently successfully coupled to electrospray ionization.[356] This allowed the development of a new mass spectrometer with many features desirable for proteomics analysis[357] that has been demonstrated to have parts per million mass accuracy.[358,359] Coupling of an LTQ to an Orbitrap promises to provide fast high-resolution analyses of proteomics samples without having to deal with the liquid nitrogen and liquid helium needed for FT-MS.

In a more specific application, the analysis of phosphopeptides is an area of continual research because phosphoserine and phosphothreonine often undergo neutral loss reactions in ion trap instrumentation, making the identification of peptides containing these species challenging. The development of ETD coupled to ion trap instrumentation[360] has the potential to greatly facilitate phosphorylation analysis[361] and also allow for sequencing of whole proteins.[362] Although there is relatively little literature on the use of ETD in proteomics applications, this promises to change upon the introduction of commercial instrumentation coupling ETD to mass spectrometry systems.

Whereas all areas of proteomics benefit from advances in mass spectrometry, shotgun proteomics has the distinct advantage of benefiting from continual research and development in the field of peptide and protein separations. Improvements in the sensitivity of mass spectrometry, as well as advances in separations science, will have a significant impact in the continuing evolution of shotgun proteomics. Finally, the implementation of these technological advances into shotgun proteomics workflows will continue to enable biological discoveries.

7. References

(1) O'Farrell, P. H. *J. Biol. Chem.* **1975**, *250*, 4007.
(2) Klose, J. *Humangenetik* **1975**, *26*, 231.
(3) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299.
(4) Whitehouse, C. M.; Dreyer, R. N.; Yamashita, M.; Fenn, J. B. *Anal. Chem.* **1985**, *57*, 675.
(5) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64.
(6) Eng, J.; McCormack, A. L.; Yates, J. R., 3rd *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976.
(7) Mann, M.; Wilm, M. *Anal. Chem.* **1994**, *66*, 4390.
(8) Yates, J. R., 3rd; Eng, J. K.; McCormack, A. L.; Schieltz, D. *Anal. Chem.* **1995**, *67*, 1426.
(9) Yates, J. R., 3rd. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 297.
(10) Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Rochon, Y.; Aebersold, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9390.
(11) Fountoulakis, M.; Takacs, M. F.; Takacs, B. *J. Chromatogr., A* **1999**, *833*, 157.
(12) Fountoulakis, M.; Takacs, M. F.; Berndt, P.; Langen, H.; Takacs, B. *Electrophoresis* **1999**, *20*, 2181.
(13) Santoni, V.; Molloy, M.; Rabilloud, T. *Electrophoresis* **2000**, *21*, 1054.
(14) Corthals, G. L.; Wasinger, V. C.; Hochstrasser, D. F.; Sanchez, J. C. *Electrophoresis* **2000**, *21*, 1104.
(15) Oh-Ishi, M.; Satoh, M.; Maeda, T. *Electrophoresis* **2000**, *21*, 1653.
(16) Giddings, J. C. *Anal. Chem.* **1984**, *56*, 1258A.
(17) Giddings, J. C. *J. High Resolut. Chromatogr. Chromatogr. Commun.* **1987**, *10*, 319.
(18) Giddings, J. C. *Anal. Chem.* **1967**, *39*, 1027.
(19) Davis, J. M.; Giddings, J. C. *Anal. Chem.* **1983**, *55*, 418.
(20) Davis, J. M.; Giddings, J. C. *Anal. Chem.* **1985**, *57*, 2168.
(21) Davis, J. M.; Giddings, J. C. *Anal. Chem.* **1985**, *57*, 2178.
(22) Guiochon, G.; Beaver, L. A.; Gonnord, M. F.; Siouffi, A. M.; Zakaria, M. *J. Chromatogr. Sci.* **1983**, *255*, 415.
(23) Consden, R.; Gordon, A. H.; Martin, A. J. *Biochem. J.* **1944**, *38*, 224.
(24) Haugaard, G.; Kroner, T. D. *J. Am. Chem. Soc.* **1948**, *70*, 2135.
(25) Durrum, E. L. *J. Colloid Sci.* **1951**, *6*, 274.
(26) Jones, A. J.; Martin, A. J. P. *Analyst* **1952**, *77*, 915.
(27) Brown, P. R. *Anal. Chem.* **1990**, *62*, 995A.
(28) Deans, D. R. *Chromatographia* **1968**, *1*, 18.
(29) Huber, J. F. K.; Van der Linden, R.; Ecker, E.; Oreans, M. *J. Chromatogr.* **1973**, *83*, 267.
(30) Davis, G. C.; Kissinger, P. T. *Anal. Chem.* **1979**, *51*, 1960.
(31) Majors, R. E. *J. Chromatogr. Sci.* **1980**, *18*, 571.
(32) Johnson, E. L.; Gloor, R.; Majors, R. E. *J. Chromatogr.* **1978**, *149*, 571.
(33) Huber, J. F. K.; Eisenbeiss, F. *J. Chromatogr.* **1978**, *149*, 127.
(34) Furst, P.; Zimmerman, L.; Oules, R.; Yahiel, V.; Johnson, C.; Bergstrom, J. *Anal. Biochem.* **1982**, *122*, 394.
(35) Dolphin, R. J.; Willmott, F. W.; Mills, A. D.; Hoogeveen, L. P. *J. Chromatogr.* **1976**, *122*, 259.
(36) Freeman, D. H. *Anal. Chem.* **1981**, *53*, 2.
(37) Erni, F.; Frei, R. W. *J. Chromatogr.* **1978**, *149*, 561.
(38) Takahashi, N.; Takahashi, Y.; Putnam, F. W. *J. Chromatogr.* **1983**, *266*, 511.
(39) Takahashi, N.; Ishioka, N.; Takahashi, Y.; Putnam, F. W. *J. Chromatogr.* **1985**, *326*, 407.
(40) Takahashi, N.; Takahashi, Y.; Ishioka, N.; Blumberg, B. S.; Putnam, F. W. *J. Chromatogr.* **1986**, *359*, 181.
(41) Matsuoka, K.; Taoka, M.; Isobe, T.; Okuyama, T.; Kato, Y. *J. Chromatogr.* **1990**, *515*, 313.
(42) Isobe, T.; Uchida, K.; Taoka, M.; Shinkai, F.; Manabe, T.; Okuyama, T. *J. Chromatogr.* **1991**, *588*, 115.
(43) Bushey, M. M.; Jorgenson, J. W. *Anal. Chem.* **1990**, *62*, 161.
(44) Bushey, M. M.; Jorgenson, J. W. *Anal. Chem.* **1990**, *62*, 978.
(45) Larmann, J. P.; Lemmo, A. V.; Moore, A. W.; Jorgenson, J. W. *Electrophoresis* **1993**, *14*.
(46) Holland, L. A.; Jorgenson, J. W. *Anal. Chem.* **1995**, *67*, 3275.
(47) Moore, A. W., Jr.; Jorgenson, J. W. *Anal. Chem.* **1995**, *67*, 3456.
(48) Opiteck, G. J.; Lewis, K. C.; Jorgenson, J. W.; Anderegg, R. J. *Anal. Chem.* **1997**, *69*, 1518.
(49) Dole, M.; Mach, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. P.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240.
(50) Taylor, G. I. *Proc. R. Soc. London A* **1964**, *280*, 383.
(51) Wilm, M.; Shevchenko, A.; Houthaeve, T.; Breit, S.; Schweigerer, L.; Fotsis, T.; Mann, M. *Nature* **1996**, *379*, 466.
(52) Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1.
(53) Wilm, M. a. M. M. *Int. J. Mass Spectrom. Ion Processes* **1994**, *136*, 167.
(54) Davis, M. T.; Stahl, D. C.; Hefta, S. A.; Lee, T. D. *Anal. Chem.* **1995**, *67*, 4549.
(55) Gatlin, C. L.; Kleemann, G. R.; Hays, L. G.; Link, A. J.; Yates, J. R., 3rd. *Anal. Biochem.* **1998**, *263*, 93.
(56) Evans, C. R.; Jorgenson, J. W. *Anal. Bioanal. Chem.* **2004**, *378*, 1952.
(57) Nagele, E.; Vollmer, M.; Horth, P.; Vad, C. *Expert Rev. Proteomics* **2004**, *1*, 37.
(58) Hattan, S. J.; Marchese, J.; Khainovski, N.; Martin, S.; Juhasz, P. *J. Proteome Res.* **2005**, *4*, 1931.
(59) Vanrobaeys, F.; Van Coster, R.; Dhondt, G.; Devreese, B.; Van Beeumen, J. *J. Proteome Res.* **2005**, *4*, 2283.
(60) Wolff, S.; Otto, A.; Albrecht, D.; Zeng, J. S.; Buttner, K.; Gluckmann, M.; Hecker, M.; Becher, D. *Mol. Cell. Proteomics* **2006**, *5*, 1183.
(61) Saito, H.; Oda, Y.; Sato, T.; Kuromitsu, J.; Ishihama, Y. *J. Proteome Res.* **2006**, *5*, 1803.
(62) Zhen, Y.; Xu, N.; Richardson, B.; Becklin, R.; Savage, J. R.; Blake, K.; Peltier, J. M. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 803.
(63) Stapels, M. D.; Barofsky, D. F. *Anal. Chem.* **2004**, *76*, 5423.
(64) Davis, M. T.; Beierle, J.; Bures, E. T.; McGinley, M. D.; Mort, J.; Robinson, J. H.; Spahr, C. S.; Yu, W.; Luethy, R.; Patterson, S. D. *J. Chromatogr. B: Biomed. Sci. Appl.* **2001**, *752*, 281.
(65) Skipp, P.; Robinson, J.; O'Connor, C. D.; Clarke, I. N. *Proteomics* **2005**, *5*, 1558.
(66) Chelius, D.; Huhmer, A. F.; Shieh, C. H.; Lehmberg, E.; Traina, J. A.; Slattery, T. K.; Pungor, E., Jr. *J. Proteome Res.* **2002**, *1*, 501.
(67) Chao, C. C.; Chelius, D.; Zhang, T.; Daggle, L.; Ching, W. M. *Proteomics* **2004**, *4*, 1280.
(68) Li, R. X.; Zhou, H.; Li, S. J.; Sheng, Q. H.; Xia, Q. C.; Zeng, R. *J. Proteome Res.* **2005**, *4*, 1256.
(69) Ru, Q. C.; Zhu, L. A.; Silberman, J.; Shriver, C. D. *Mol. Cell. Proteomics* **2006**, *5*, 1095.
(70) Mauri, P.; Scarpa, A.; Nascimbeni, A. C.; Benazzi, L.; Parmagnani, E.; Mafficini, A.; Della Peruta, M.; Bassi, C.; Miyazaki, K.; Sorio, C. *FASEB J.* **2005**, *19*, 1125.
(71) Li, C.; Hong, Y.; Tan, Y. X.; Zhou, H.; Ai, J. H.; Li, S. J.; Zhang, L.; Xia, Q. C.; Wu, J. R.; Wang, H. Y.; Zeng, R. *Mol. Cell. Proteomics* **2004**, *3*, 399.
(72) Vollmer, M.; Nagele, E.; Horth, P. *J. Biomol. Tech.* **2003**, *14*, 128.
(73) Nagele, E.; Vollmer, M.; Horth, P. *J. Chromatogr., A* **2003**, *1009*, 197.

(74) Kang, D.; Nam, H.; Kim, Y. S.; Moon, M. H. *J. Chromatogr., A* **2005**, *1070*, 193.
(75) Xiang, R.; Shi, Y.; Dillon, D. A.; Negin, B.; Horvath, C.; Wilkins, J. A. *J. Proteome Res.* **2004**, *3*, 1278.
(76) Gu, S.; Du, Y.; Chen, J.; Liu, Z.; Bradbury, E. M.; Hu, C. A.; Chen, X. *J. Proteome Res.* **2004**, *3*, 1191.
(77) Tyan, Y. C.; Wu, H. Y.; Lai, W. W.; Su, W. C.; Liao, P. C. *J. Proteome Res.* **2005**, *4*, 1274.
(78) Tyan, Y. C.; Jong, S. B.; Liao, J. D.; Liao, P. C.; Yang, M. H.; Liu, C. Y.; Klauser, R.; Himmelhaus, M.; Grunze, M. *J. Proteome Res.* **2005**, *4*, 748.
(79) Mitulovic, G.; Stingl, C.; Smoluch, M.; Swart, R.; Chervet, J. P.; Steinmacher, I.; Gerner, C.; Mechtler, K. *Proteomics* **2004**, *4*, 2545.
(80) Panchaud, A.; Kussmann, M.; Affolter, M. *Proteomics* **2005**, *5*, 3836.
(81) Allison, W. T.; Veldhoen, K. M.; Hawryshyn, C. W. *Mol. Vis.* **2006**, *12*, 655.
(82) Lim, K. B.; Kassel, D. B. *Anal. Biochem.* **2006**, *354*, 213.
(83) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R., 3rd. *Nat. Biotechnol.* **1999**, *17*, 676.
(84) Dong, J.; Lai, R.; Jennings, J. L.; Link, A. J.; Hinnebusch, A. G. *Mol. Cell. Biol.* **2005**, *25*, 9859.
(85) Fleischer, T. C.; Weaver, C. M.; McAfee, K. J.; Jennings, J. L.; Link, A. J. *Genes Dev.* **2006**, *20*, 1294.
(86) Gerbasi, V. R.; Weaver, C. M.; Hill, S.; Friedman, D. B.; Link, A. J. *Mol. Cell. Biol.* **2004**, *24*, 8276.
(87) Washburn, M. P.; Wolters, D.; Yates, J. R., 3rd. *Nat. Biotechnol.* **2001**, *19*, 242.
(88) Wolters, D. A.; Washburn, M. P.; Yates, J. R., 3rd. *Anal. Chem.* **2001**, *73*, 5683.
(89) McDonald, W. H.; Ohi, R.; Miyamoto, D. T.; Mitchison, T. J.; Yates, J. R. *Int. J. Mass Spectrom.* **2002**, *219*, 245.
(90) Florens, L.; Washburn, M. P. *Methods Mol. Biol.* **2006**, *328*, 159.
(91) Koller, A.; Washburn, M. P.; Lange, B. M.; Andon, N. L.; Deciu, C.; Haynes, P. A.; Hays, L.; Schieltz, D.; Ulaszek, R.; Wei, J.; Wolters, D.; Yates, J. R., 3rd. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11969.
(92) Florens, L.; Washburn, M. P.; Raine, J. D.; Anthony, R. M.; Grainger, M.; Haynes, J. D.; Moch, J. K.; Muster, N.; Sacci, J. B.; Tabb, D. L.; Witney, A. A.; Wolters, D.; Wu, Y.; Gardner, M. J.; Holder, A. A.; Sinden, R. E.; Yates, J. R.; Carucci, D. J. *Nature* **2002**, *419*, 520.
(93) Carlton, J. M.; Angiuoli, S. V.; Suh, B. B.; Kooij, T. W.; Pertea, M.; Silva, J. C.; Ermolaeva, M. D.; Allen, J. E.; Selengut, J. D.; Koo, H. L.; Peterson, J. D.; Pop, M.; Kosack, D. S.; Shumway, M. F.; Bidwell, S. L.; Shallom, S. J.; van Aken, S. E.; Riedmuller, S. B.; Feldblyum, T. V.; Cho, J. K.; Quackenbush, J.; Sedegah, M.; Shoaibi, A.; Cummings, L. M.; Florens, L.; Yates, J. R.; Raine, J. D.; Sinden, R. E.; Harris, M. A.; Cunningham, D. A.; Preiser, P. R.; Bergman, L. W.; Vaidya, A. B.; van Lin, L. H.; Janse, C. J.; Waters, A. P.; Smith, H. O.; White, O. R.; Salzberg, S. L.; Venter, J. C.; Fraser, C. M.; Hoffman, S. L.; Gardner, M. J.; Carucci, D. J. *Nature* **2002**, *419*, 512.
(94) Doolan, D. L.; Southwood, S.; Freilich, D. A.; Sidney, J.; Graber, N. L.; Shatney, L.; Bebris, L.; Florens, L.; Dobano, C.; Witney, A. A.; Appella, E.; Hoffman, S. L.; Yates, J. R., 3rd; Carucci, D. J.; Sette, A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9952.
(95) Ling, I. T.; Florens, L.; Dluzewski, A. R.; Kaneko, O.; Grainger, M.; Yim Lim, B. Y.; Tsuboi, T.; Hopkins, J. M.; Johnson, J. R.; Torii, M.; Bannister, L. H.; Yates, J. R., 3rd; Holder, A. A.; Mattei, D. *Mol. Microbiol.* **2004**, *52*, 107.
(96) Sam-Yellowe, T. Y.; Florens, L.; Wang, T.; Raine, J. D.; Carucci, D. J.; Sinden, R.; Yates, J. R., 3rd. *J. Proteome Res.* **2004**, *3*, 995.
(97) Sam-Yellowe, T. Y.; Florens, L.; Johnson, J. R.; Wang, T.; Drazba, J. A.; Le Roch, K. G.; Zhou, Y.; Batalov, S.; Carucci, D. J.; Winzeler, E. A.; Yates, J. R., 3rd. *Genome Res.* **2004**, *14*, 1052.
(98) Le Roch, K. G.; Johnson, J. R.; Florens, L.; Zhou, Y.; Santrosyan, A.; Grainger, M.; Yan, S. F.; Williamson, K. C.; Holder, A. A.; Carucci, D. J.; Yates, J. R., 3rd; Winzeler, E. A. *Genome Res.* **2004**, *14*, 2308.
(99) Hall, N.; Karras, M.; Raine, J. D.; Carlton, J. M.; Kooij, T. W.; Berriman, M.; Florens, L.; Janssen, C. S.; Pain, A.; Christophides, G. K.; James, K.; Rutherford, K.; Harris, B.; Harris, D.; Churcher, C.; Quail, M. A.; Ormond, D.; Doggett, J.; Trueman, H. E.; Mendoza, J.; Bidwell, S. L.; Rajandream, M. A.; Carucci, D. J.; Yates, J. R., 3rd; Kafatos, F. C.; Janse, C. J.; Barrell, B.; Turner, C. M.; Waters, A. P.; Sinden, R. E. *Science* **2005**, *307*, 82.
(100) Choi, K. S.; Song, L.; Park, Y. M.; Marshall, J.; Lund, A. L.; Shion, H.; Park, E. M.; Chae, H. Z.; Park, J. H. *Prep. Biochem. Biotechnol.* **2006**, *36*, 3.
(101) Bayer, E. M.; Bottrill, A. R.; Walshaw, J.; Vigouroux, M.; Naldrett, M. J.; Thomas, C. L.; Maule, A. J. *Proteomics* **2006**, *6*, 301.
(102) Pan, Y.; Kislinger, T.; Gramolini, A. O.; Zvaritch, E.; Kranias, E. G.; MacLennan, D. H.; Emili, A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2241.
(103) Cooper, B.; Garrett, W. M.; Campbell, K. B. *Proteomics* **2006**, *6*, 2477.
(104) Wenner, B. R.; Lovell, M. A.; Lynn, B. C. *J. Proteome Res.* **2004**, *3*, 97.
(105) Cagney, G.; Park, S.; Chung, C.; Tong, B.; O'Dushlaine, C.; Shields, D. C.; Emili, A. *J. Proteome Res.* **2005**, *4*, 1757.
(106) Ru, Q. C.; Katenhusen, R. A.; Zhu, L. A.; Silberman, J.; Yang, S.; Orchard, T. J.; Brzeski, H.; Liebman, M.; Ellsworth, D. L. *J. Chromatogr., A* **2006**, *1111*, 166.
(107) Ru, Q. C.; Zhu, L. A.; Katenhusen, R. A.; Silberman, J.; Brzeski, H.; Liebman, M.; Shriver, C. D. *J. Chromatogr., A* **2006**, *1111*, 175.
(108) Sandhu, C.; Connor, M.; Kislinger, T.; Slingerland, J.; Emili, A. *J. Proteome Res.* **2005**, *4*, 674.
(109) Ram, R. J.; Verberkmoes, N. C.; Thelen, M. P.; Tyson, G. W.; Baker, B. J.; Blake, R. C., 2nd; Shah, M.; Hettich, R. L.; Banfield, J. F. *Science* **2005**, *308*, 1915.
(110) Kislinger, T.; Cox, B.; Kannan, A.; Chung, C.; Hu, P.; Ignatchenko, A.; Scott, M. S.; Gramolini, A. O.; Morris, Q.; Hallett, M. T.; Rossant, J.; Hughes, T. R.; Frey, B.; Emili, A. *Cell* **2006**, *125*, 173.
(111) Wu, C. C.; MacCoss, M. J.; Mardones, G.; Finnigan, C.; Mogelsvang, S.; Yates, J. R., 3rd; Howell, K. E. *Mol. Biol. Cell* **2004**, *15*, 2907.
(112) Skop, A. R.; Liu, H.; Yates, J., 3rd; Meyer, B. J.; Heald, R. *Science* **2004**, *305*, 61.
(113) Chu, D. S.; Liu, H.; Nix, P.; Wu, T. F.; Ralston, E. J.; Yates, J. R., 3rd; Meyer, B. J. *Nature* **2006**, *443*, 101.
(114) Wu, C. C.; MacCoss, M. J.; Howell, K. E.; Yates, J. R., 3rd. *Nat. Biotechnol.* **2003**, *21*, 532.
(115) Schirmer, E. C.; Florens, L.; Guan, T.; Yates, J. R., 3rd; Gerace, L. *Science* **2003**, *301*, 1380.
(116) Durr, E.; Yu, J.; Krasinska, K. M.; Carver, L. A.; Yates, J. R.; Testa, J. E.; Oh, P.; Schnitzer, J. E. *Nat. Biotechnol.* **2004**, *22*, 985.
(117) Huang, J. K.; Phillips, G. R.; Roth, A. D.; Pedraza, L.; Shan, W.; Belkaid, W.; Mi, S.; Fex-Svenningsen, A.; Florens, L.; Yates, J. R., 3rd; Colman, D. R. *Science* **2005**, *310*, 1813.
(118) Fischer, F.; Wolters, D.; Rogner, M.; Poetsch, A. *Mol. Cell Proteomics* **2006**, *5*, 444.
(119) Oh, P.; Li, Y.; Yu, J.; Durr, E.; Krasinska, K. M.; Carver, L. A.; Testa, J. E.; Schnitzer, J. E. *Nature* **2004**, *429*, 629.
(120) Puig, O.; Caspary, F.; Rigaut, G.; Rutz, B.; Bouveret, E.; Bragado-Nilsson, E.; Wilm, M.; Seraphin, B. *Methods* **2001**, *24*, 218.
(121) Einhauer, A.; Jungbauer, A. *J. Biochem. Biophys. Methods* **2001**, *49*, 455.
(122) Tomomori-Sato, C.; Sato, S.; Parmely, T. J.; Banks, C. A.; Sorokina, I.; Florens, L.; Zybailov, B.; Washburn, M. P.; Brower, C. S.; Conaway, R. C.; Conaway, J. W. *J. Biol. Chem.* **2004**, *279*, 5846.
(123) Sato, S.; Tomomori-Sato, C.; Parmely, T. J.; Florens, L.; Zybailov, B.; Swanson, S. K.; Banks, C. A.; Jin, J.; Cai, Y.; Washburn, M. P.; Conaway, J. W.; Conaway, R. C. *Mol. Cell* **2004**, *14*, 685.
(124) Powell, D. W.; Weaver, C. M.; Jennings, J. L.; McAfee, K. J.; He, Y.; Weil, P. A.; Link, A. J. *Mol. Cell. Biol.* **2004**, *24*, 7249.
(125) Kobor, M. S.; Venkatasubrahmanyam, S.; Meneghini, M. D.; Gin, J. W.; Jennings, J. L.; Link, A. J.; Madhani, H. D.; Rine, J. *PLoS Biol.* **2004**, *2*, E131.
(126) Graumann, J.; Dunipace, L. A.; Seol, J. H.; McDonald, W. H.; Yates, J. R., 3rd; Wold, B. J.; Deshaies, R. J. *Mol. Cell. Proteomics* **2004**, *3*, 226.
(127) Lee, K. K.; Prochasson, P.; Florens, L.; Swanson, S. K.; Washburn, M. P.; Workman, J. L. *Biochem. Soc. Trans.* **2004**, *32*, 899.
(128) Lee, K. K.; Florens, L.; Swanson, S. K.; Washburn, M. P.; Workman, J. L. *Mol. Cell. Biol.* **2005**, *25*, 1173.
(129) Schneider, J.; Wood, A.; Lee, J. S.; Schuster, R.; Dueker, J.; Maguire, C.; Swanson, S. K.; Florens, L.; Washburn, M. P.; Shilatifard, A. *Mol. Cell* **2005**, *19*, 849.
(130) Carrozza, M. J.; Florens, L.; Swanson, S. K.; Shia, W. J.; Anderson, S.; Yates, J.; Washburn, M. P.; Workman, J. L. *Biochim. Biophys. Acta* **2005**, *1731*, 77.
(131) Carrozza, M. J.; Li, B.; Florens, L.; Suganuma, T.; Swanson, S. K.; Lee, K. K.; Shia, W. J.; Anderson, S.; Yates, J.; Washburn, M. P.; Workman, J. L. *Cell* **2005**, *123*, 581.
(132) Cai, Y.; Jin, J.; Florens, L.; Swanson, S. K.; Kusch, T.; Li, B.; Workman, J. L.; Washburn, M. P.; Conaway, R. C.; Conaway, J. W. *J. Biol. Chem.* **2005**, *280*, 13665.
(133) Jin, J.; Cai, Y.; Yao, T.; Gottschalk, A. J.; Florens, L.; Swanson, S. K.; Gutierrez, J. L.; Coleman, M. K.; Workman, J. L.; Mushegian, A.; Washburn, M. P.; Conaway, R. C.; Conaway, J. W. *J. Biol. Chem.* **2005**, *280*, 41207.
(134) Pebernard, S.; Wohlschlegel, J.; McDonald, W. H.; Yates, J. R., 3rd; Boddy, M. N. *Mol. Cell. Biol.* **2006**, *26*, 1617.

(135) Duchaine, T. F.; Wohlschlegel, J. A.; Kennedy, S.; Bei, Y.; Conte, D.; Jr.; Pang, K.; Brownell, D. R.; Harding, S.; Mitani, S.; Ruvkun, G.; Yates, J. R., 3rd; Mello, C. C. *Cell* **2006**, *124*, 343.
(136) MacCoss, M. J.; McDonald, W. H.; Saraf, A.; Sadygov, R.; Clark, J. M.; Tasto, J. J.; Gould, K. L.; Wolters, D.; Washburn, M.; Weiss, A.; Clark, J. I.; Yates, J. R., 3rd. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7900.
(137) Sedgwick, C.; Rawluk, M.; Decesare, J.; Raithatha, S.; Wohlschlegel, J.; Semchuk, P.; Ellison, M.; Yates, J., 3rd; Stuart, D. *Biochem. J.* **2006**, *399*, 151.
(138) Wolfe, B. A.; McDonald, W. H.; Yates, J. R., 3rd; Gould, K. L. *Dev. Cell* **2006**, *11*, 423.
(139) Widlund, P. O.; Lyssand, J. S.; Anderson, S.; Niessen, S.; Yates, J. R., 3rd; Davis, T. N. *Mol. Biol. Cell* **2006**, *17*, 1065.
(140) Kesti, T.; McDonald, W. H.; Yates, J. R., 3rd; Wittenberg, C. *J. Biol. Chem.* **2004**, *279*, 14245.
(141) de Bruin, R. A.; McDonald, W. H.; Kalashnikova, T. I.; Yates, J., 3rd; Wittenberg, C. *Cell* **2004**, *117*, 887.
(142) Flick, K.; Ouni, I.; Wohlschlegel, J. A.; Capati, C.; McDonald, W. H.; Yates, J. R.; Kaiser, P. *Nat. Cell Biol.* **2004**, *6*, 634.
(143) Raffa, G. D.; Wohlschlegel, J.; Yates, J. R., 3rd; Boddy, M. N. *J. Biol. Chem.* **2006**, *281*, 27973.
(144) Wohlschlegel, J. A.; Johnson, E. S.; Reed, S. I.; Yates, J. R., 3rd. *J. Proteome Res.* **2006**, *5*, 761.
(145) Wohlschlegel, J. A.; Johnson, E. S.; Reed, S. I.; Yates, J. R., 3rd. *J. Biol. Chem.* **2004**, *279*, 45662.
(146) Willert, K.; Brown, J. D.; Danenberg, E.; Duncan, A. W.; Weissman, I. L.; Reya, T.; Yates, J. R., 3rd; Nusse, R. *Nature* **2003**, *423*, 448.
(147) Beausoleil, S. A.; Villen, J.; Gerber, S. A.; Rush, J.; Gygi, S. P. *Nat. Biotechnol.* **2006**, *24*, 1285.
(148) Schwartz, D.; Gygi, S. P. *Nat. Biotechnol.* **2005**, *23*, 1391.
(149) Pedrioli, P. G.; Raught, B.; Zhang, X. D.; Rogers, R.; Aitchison, J.; Matunis, M.; Aebersold, R. *Nat. Methods* **2006**, *3*, 533.
(150) Tsur, D.; Tanner, S.; Zandi, E.; Bafna, V.; Pevzner, P. A. *Nat. Biotechnol.* **2005**, *23*, 1562.
(151) Washburn, M. P.; Ulaszek, R. R.; Yates, J. R., 3rd. *Anal. Chem.* **2003**, *75*, 5054.
(152) Washburn, M. P.; Ulaszek, R.; Deciu, C.; Schieltz, D. M.; Yates, J. R., 3rd. *Anal. Chem.* **2002**, *74*, 1650.
(153) Washburn, M. P.; Koller, A.; Oshiro, G.; Ulaszek, R. R.; Plouffe, D.; Deciu, C.; Winzeler, E.; Yates, J. R., 3rd. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3107.
(154) MacCoss, M. J.; Wu, C. C.; Liu, H.; Sadygov, R.; Yates, J. R., 3rd. *Anal. Chem.* **2003**, *75*, 6912.
(155) Liu, H.; Sadygov, R. G.; Yates, J. R., 3rd. *Anal. Chem.* **2004**, *76*, 4193.
(156) Old, W. M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K. G.; Mendoza, A.; Sevinsky, J. R.; Resing, K. A.; Ahn, N. G. *Mol. Cell. Proteomics* **2005**, *4*, 1487.
(157) Zybailov, B.; Coleman, M. K.; Florens, L.; Washburn, M. P. *Anal. Chem.* **2005**, *77*, 6218.
(158) Zybailov, B.; Mosley, A. L.; Sardiu, M. E.; Coleman, M. K.; Florens, L.; Washburn, M. P. *J. Proteome Res.* **2006**, *5*, 2339.
(159) Paoletti, A. C.; Parmely, T. J.; Tomomori-Sato, C.; Sato, S.; Zhu, D.; Conaway, R. C.; Weliky, Conaway, J.; Florens, L.; Washburn, M. P. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18928.
(160) Breci, L.; Hattrup, E.; Keeler, M.; Letarte, J.; Johnson, R.; Haynes, P. A. *Proteomics* **2005**, *5*, 2018.
(161) Barnea, E.; Sorkin, R.; Ziv, T.; Beer, I.; Admon, A. *Proteomics* **2005**, *5*, 3367.
(162) Li, X.; Gong, Y.; Wang, Y.; Wu, S.; Cai, Y.; He, P.; Lu, Z.; Ying, W.; Zhang, Y.; Jiao, L.; He, H.; Zhang, Z.; He, F.; Zhao, X.; Qian, X. *Proteomics* **2005**, *5*, 3423.
(163) Licklider, L. J.; Thoreen, C. C.; Peng, J.; Gygi, S. P. *Anal. Chem.* **2002**, *74*, 3076.
(164) Guzzetta, A. W.; Chien, A. S. *J. Proteome Res.* **2005**, *4*, 2412.
(165) Ballif, B. A.; Villen, J.; Beausoleil, S. A.; Schwartz, D.; Gygi, S. P. *Mol. Cell. Proteomics* **2004**, *3*, 1093.
(166) Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villen, J.; Li, J.; Cohn, M. A.; Cantley, L. C.; Gygi, S. P. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12130.
(167) DeSouza, L.; Diehl, G.; Rodrigues, M. J.; Guo, J.; Romaschin, A. D.; Colgan, T. J.; Siu, K. W. *J. Proteome Res.* **2005**, *4*, 377.
(168) Fujii, K.; Nakano, T.; Kawamura, T.; Usui, F.; Bando, Y.; Wang, R.; Nishimura, T. *J. Proteome Res.* **2004**, *3*, 712.
(169) Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. *J. Proteome Res.* **2003**, *2*, 43.
(170) Wilmarth, P. A.; Riviere, M. A.; Rustvold, D. L.; Lauten, J. D.; Madden, T. E.; David, L. L. *J. Proteome Res.* **2004**, *3*, 1017.
(171) Gaucher, S. P.; Taylor, S. W.; Fahy, E.; Zhang, B.; Warnock, D. E.; Ghosh, S. S.; Gibson, B. W. *J. Proteome Res.* **2004**, *3*, 495.
(172) America, A. H.; Cordewener, J. H.; van Geffen, M. H.; Lommen, A.; Vissers, J. P.; Bino, R. J.; Hall, R. D. *Proteomics* **2006**, *6*, 641.
(173) Noben, J. P.; Dumont, D.; Kwasnikowska, N.; Verhaert, P.; Somers, V.; Hupperts, R.; Stinissen, P.; Robben, J. *J. Proteome Res.* **2006**, *5*, 1647.
(174) Vitali, B.; Wasinger, V.; Brigidi, P.; Guilhaus, M. *Proteomics* **2005**, *5*, 1859.
(175) Vollmer, M.; Horth, P.; Nagele, E. *Anal. Chem.* **2004**, *76*, 5180.
(176) Adkins, J. N.; Varnum, S. M.; Auberry, K. J.; Moore, R. J.; Angell, N. H.; Smith, R. D.; Springer, D. L.; Pounds, J. G. *Mol. Cell. Proteomics* **2002**, *1*, 947.
(177) Blonder, J.; Rodriguez-Galan, M. C.; Chan, K. C.; Lucas, D. A.; Yu, L. R.; Conrads, T. P.; Issaq, H. J.; Young, H. A.; Veenstra, T. D. *J. Proteome Res.* **2004**, *3*, 862.
(178) Qian, W. J.; Jacobs, J. M.; Camp, D. G., 2nd; Monroe, M. E.; Moore, R. J.; Gritsenko, M. A.; Calvano, S. E.; Lowry, S. F.; Xiao, W.; Moldawer, L. L.; Davis, R. W.; Tompkins, R. G.; Smith, R. D. *Proteomics* **2005**, *5*, 572.
(179) Qian, W. J.; Liu, T.; Monroe, M. E.; Strittmatter, E. F.; Jacobs, J. M.; Kangas, L. J.; Petritis, K.; Camp, D. G., 2nd; Smith, R. D. *J. Proteome Res.* **2005**, *4*, 53.
(180) Yu, L. R.; Conrads, T. P.; Uo, T.; Kinoshita, Y.; Morrison, R. S.; Lucas, D. A.; Chan, K. C.; Blonder, J.; Issaq, H. J.; Veenstra, T. D. *Mol. Cell. Proteomics* **2004**, *3*, 896.
(181) Wang, H.; Qian, W. J.; Chin, M. H.; Petyuk, V. A.; Barry, R. C.; Liu, T.; Gritsenko, M. A.; Mottaz, H. M.; Moore, R. J.; Camp, Ii, D. G.; Khan, A. H.; Smith, D. J.; Smith, R. D. *J. Proteome Res.* **2006**, *5*, 361.
(182) Le Bihan, T.; Goh, T.; Stewart, I. I.; Salter, A. M.; Bukhman, Y. V.; Dharsee, M.; Ewing, R.; Wisniewski, J. R. *J. Proteome Res.* **2006**, *5*, 2701.
(183) Coldham, N. G.; Woodward, M. J. *J. Proteome Res.* **2004**, *3*, 595.
(184) Cutillas, P. R.; Norden, A. G.; Cramer, R.; Burlingame, A. L.; Unwin, R. J. *Clin. Sci. (London)* **2003**, *104*, 483.
(185) Atwood, J. A., 3rd; Weatherly, D. B.; Minning, T. A.; Bundy, B.; Cavola, C.; Opperdoes, F. R.; Orlando, R.; Tarleton, R. L. *Science* **2005**, *309*, 473.
(186) Hood, B. L.; Zhou, M.; Chan, K. C.; Lucas, D. A.; Kim, G. J.; Issaq, H. J.; Veenstra, T. D.; Conrads, T. P. *J. Proteome Res.* **2005**, *4*, 1561.
(187) Weatherly, D. B.; Astwood, J. A., 3rd; Minning, T. A.; Cavola, C.; Tarleton, R. L.; Orlando, R. *Mol. Cell. Proteomics* **2005**, *4*, 762.
(188) Jacobs, J. M.; Mottaz, H. M.; Yu, L. R.; Anderson, D. J.; Moore, R. J.; Chen, W. N.; Auberry, K. J.; Strittmatter, E. F.; Monroe, M. E.; Thrall, B. D.; Camp, D. G., 2nd; Smith, R. D. *J. Proteome Res.* **2004**, *3*, 68.
(189) Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. *Anal. Chem.* **2005**, *77*, 6426.
(190) Bodnar, W. M.; Blackburn, R. K.; Krise, J. M.; Moseley, M. A. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 971.
(191) Zhang, B.; McDonald, C.; Li, L. *Anal. Chem.* **2004**, *76*, 992.
(192) Chen, H. S.; Rejtar, T.; Andreev, V.; Moskovets, E.; Karger, B. L. *Anal. Chem.* **2005**, *77*, 2323.
(193) Chen, H. S.; Rejtar, T.; Andreev, V.; Moskovets, E.; Karger, B. L. *Anal. Chem.* **2005**, *77*, 7816.
(194) Scholten, A.; Visser, N. F.; van den Heuvel, R. H.; Heck, A. J. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 983.
(195) Li, L.; Wang, A. P. L.; Coulson, L. D. *Anal. Chem.* **1993**, *65*, 493.
(196) Whittal, R. M.; Russon, L. M.; Li, L. *J. Chromatogr., A* **1998**, *794*, 367.
(197) Murray, K. K.; Russel, D. H. *Anal. Chem.* **1993**, *65*, 2534.
(198) Murray, K. K.; Russel, D. H. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 1.
(199) Murray, K. K.; Lewis, T. M.; Beeson, M. D.; Russel, D. H. *Anal. Chem.* **1994**, *66*.
(200) Fei, X.; Wei, G.; Murray, K. K. *Anal. Chem.* **1996**, *68*, 1143.
(201) Fei, X.; Murray, K. K. *Anal. Chem.* **1996**, *68*, 3555.
(202) He, L.; Murray, K. K. *Anal. Chem.* **1997**, *69*, 3613.
(203) Orsnes, H.; Graf, T.; Degn, H.; Murray, K. K. *Anal. Chem.* **2000**, *72*, 251.
(204) Preisler, J.; Foret, F.; Karger, B. L. *Anal. Chem.* **1998**, *70*, 5278.
(205) Mawuenyega, K. G.; Kaji, H.; Yamauchi, Y.; Shinkawa, T.; Saito, H.; Taoka, M.; Takahashi, N.; Isobe, T. *J. Proteome Res.* **2003**, *2*, 23.
(206) Taoka, M.; Yamauchi, Y.; Shinkawa, T.; Kaji, H.; Motohashi, W.; Nakayama, H.; Takahashi, N.; Isobe, T. *Mol. Cell. Proteomics* **2004**, *3*, 780.
(207) Nagano, K.; Taoka, M.; Yamauchi, Y.; Itagaki, C.; Shinkawa, T.; Nunomura, K.; Okamura, N.; Takahashi, N.; Izumi, T.; Isobe, T. *Proteomics* **2005**, *5*, 1346.
(208) Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M. *Nat. Biotechnol.* **2002**, *20*, 301.
(209) Posewitz, M. C.; Tempst, P. *Anal. Chem.* **1999**, *71*, 2883.
(210) Kim, J. E.; Tannenbaum, S. R.; White, F. M. *J. Proteome Res.* **2005**, *4*, 1339.

(211) Moser, K.; White, F. M. *J. Proteome Res.* **2006**, *5*, 98.
(212) Ruse, C. I.; Tan, F. L.; Kinter, M.; Bond, M. *Proteomics* **2004**, *4*, 1505.
(213) Trinidad, J. C.; Thalhammer, A.; Specht, C. G.; Schoepfer, R.; Burlingame, A. L. *J. Neurochem.* **2005**, *92*, 1306.
(214) Salomon, A. R.; Ficarro, S. B.; Brill, L. M.; Brinker, A.; Phung, Q. T.; Ericson, C.; Sauer, K.; Brock, A.; Horn, D. M.; Schultz, P. G.; Peters, E. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 443.
(215) Brill, L. M.; Salomon, A. R.; Ficarro, S. B.; Mukherji, M.; Stettler-Gill, M.; Peters, E. C. *Anal. Chem.* **2004**, *76*, 2763.
(216) Ficarro, S.; Chertihin, O.; Westbrook, V. A.; White, F.; Jayes, F.; Kalab, P.; Marto, J. A.; Shabanowitz, J.; Herr, J. C.; Hunt, D. F.; Visconti, P. E. *J. Biol. Chem.* **2003**, *278*, 11579.
(217) He, T.; Alving, K.; Feild, B.; Norton, J.; Joseloff, E. G.; Patterson, S. D.; Domon, B. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 363.
(218) Zheng, H.; Hu, P.; Quinn, D. F.; Wang, Y. K. *Mol. Cell. Proteomics* **2005**, *4*, 721.
(219) Huang, S. Y.; Tsai, M. L.; Wu, C. J.; Hsu, J. L.; Ho, S. H.; Chen, S. H. *Proteomics* **2006**, *6*, 1722.
(220) Collins, M. O.; Yu, L.; Coba, M. P.; Husi, H.; Campuzano, I.; Blackstock, W. P.; Choudhary, J. S.; Grant, S. G. *J. Biol. Chem.* **2005**, *280*, 5972.
(221) Giorgianni, F.; Beranova-Giorgianni, S.; Desiderio, D. M. *Proteomics* **2004**, *4*, 587.
(222) Shu, H.; Chen, S.; Bi, Q.; Mumby, M.; Brekken, D. L. *Mol. Cell. Proteomics* **2004**, *3*, 279.
(223) Puente, L. G.; Voisin, S.; Lee, R. E.; Megeney, L. A. *Mol. Cell. Proteomics* **2006**, *5*, 2244.
(224) Nuhse, T. S.; Stensballe, A.; Jensen, O. N.; Peck, S. C. *Mol. Cell. Proteomics* **2003**, *2*, 1234.
(225) Trinidad, J. C.; Specht, C. G.; Thalhammer, A.; Schoepfer, R.; Burlingame, A. L. *Mol. Cell. Proteomics* **2006**, *5*, 914.
(226) Munton, R. P.; Tweedie-Cullen, R.; Livingstone-Zatchej, M.; Weinandy, F.; Waidelich, M.; Longo, D.; Gehrig, P.; Potthast, F.; Rutishauser, D.; Gerrits, B.; Panse, C.; Schlapbach, R.; Mansuy, I. M. *Mol. Cell. Proteomics* **2006**.
(227) Hoffert, J. D.; Pisitkun, T.; Wang, G.; Shen, R. F.; Knepper, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 7159.
(228) Gruhler, A.; Olsen, J. V.; Mohammed, S.; Mortensen, P.; Faergeman, N. J.; Mann, M.; Jensen, O. N. *Mol. Cell. Proteomics* **2005**, *4*, 310.
(229) Wang, J.; Zhang, Y.; Jiang, H.; Cai, Y.; Qian, X. *Proteomics* **2006**, *6*, 404.
(230) Pinkse, M. W.; Uitto, P. M.; Hilhorst, M. J.; Ooms, B.; Heck, A. J. *Anal. Chem.* **2004**, *76*, 3935.
(231) Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. *Mol. Cell Proteomics* **2005**, *4*, 873.
(232) Mann, K.; Olsen, J. V.; Macek, B.; Gnad, F.; Mann, M. *Proteomics* **2007**, *7*, 106.
(233) Oda, Y.; Nagasu, T.; Chait, B. T. *Nat. Biotechnol.* **2001**, *19*, 379.
(234) Zhou, H.; Watts, J. D.; Aebersold, R. *Nat. Biotechnol.* **2001**, *19*, 375.
(235) Zhang, H.; Li, X. J.; Martin, D. B.; Aebersold, R. *Nat. Biotechnol.* **2003**, *21*, 660.
(236) Liu, T.; Qian, W. J.; Gritsenko, M. A.; Camp, D. G., 2nd; Monroe, M. E.; Moore, R. J.; Smith, R. D. *J. Proteome Res.* **2005**, *4*, 2070.
(237) Sun, B.; Ranish, J. A.; Utleg, A. G.; White, J. T.; Yan, X.; Lin, B.; Hood, L. *Mol. Cell. Proteomics* **2007**, *6*, 141.
(238) Geng, M.; Zhang, X.; Bina, M.; Regnier, F. *J. Chromatogr., B: Biomed. Sci. Appl.* **2001**, *752*, 293.
(239) Xiong, L.; Regnier, F. E. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2002**, *782*, 405.
(240) Xiong, L.; Andrews, D.; Regnier, F. *J. Proteome Res.* **2003**, *2*, 618.
(241) Atwood, J. A., 3rd; Minning, T.; Ludolf, F.; Nuccio, A.; Weatherly, D. B.; Alvarez-Manilla, G.; Tarleton, R.; Orlando, R. *J. Proteome Res.* **2006**, *5*, 3376.
(242) Kaji, H.; Saito, H.; Yamauchi, Y.; Shinkawa, T.; Taoka, M.; Hirabayashi, J.; Kasai, K.; Takahashi, N.; Isobe, T. *Nat. Biotechnol.* **2003**, *21*, 667.
(243) Vosseller, K.; Trinidad, J. C.; Chalkley, R. J.; Specht, C. G.; Thalhammer, A.; Lynn, A. J.; Snedecor, J. O.; Guan, S.; Medzihradszky, K. F.; Maltby, D. A.; Schoepfer, R.; Burlingame, A. L. *Mol. Cell. Proteomics* **2006**, *5*, 923.
(244) Hayes, B. K.; Greis, K. D.; Hart, G. W. *Anal. Biochem.* **1995**, *228*, 115.
(245) Durham, M.; Regnier, F. E. *J. Chromatogr., A* **2006**.
(246) Qiu, R.; Regnier, F. E. *Anal. Chem.* **2005**, *77*, 7225.
(247) Hirabayashi, J.; Hayama, K.; Kaji, H.; Isobe, T.; Kasai, K. *J. Biochem. (Tokyo)* **2002**, *132*, 103.
(248) Wells, L.; Vosseller, K.; Cole, R. N.; Cronshaw, J. M.; Matunis, M. J.; Hart, G. W. *Mol. Cell. Proteomics* **2002**, *1*, 791.
(249) Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13132.
(250) Cargile, B. J.; Talley, D. L.; Stephenson, J. L., Jr. *Electrophoresis* **2004**, *25*, 936.
(251) Cargile, B. J.; Bundy, J. L.; Freeman, T. W.; Stephenson, J. L., Jr. *J. Proteome Res.* **2004**, *3*, 112.
(252) Cargile, B. J.; Sevinsky, J. R.; Essader, A. S.; Stephenson, J. L., Jr.; Bundy, J. L. *J. Biomol. Tech.* **2005**, *16*, 181.
(253) Essader, A. S.; Cargile, B. J.; Bundy, J. L.; Stephenson, J. L., Jr. *Proteomics* **2005**, *5*, 24.
(254) Cargile, B. J.; Stephenson, J. L., Jr. *Anal. Chem.* **2004**, *76*, 267.
(255) Chen, J.; Lee, C. S.; Shen, Y.; Smith, R. D.; Baehrecke, E. H. *Electrophoresis* **2002**, *23*, 3143.
(256) Chen, J.; Gao, J.; Lee, C. S. *J. Proteome Res.* **2003**, *2*, 249.
(257) Chen, J.; Balgley, B. M.; DeVoe, D. L.; Lee, C. S. *Anal. Chem.* **2003**, *75*, 3145.
(258) Wang, Y.; Rudnick, P. A.; Evans, E. L.; Li, J.; Zhuang, Z.; Devoe, D. L.; Lee, C. S.; Balgley, B. M. *Anal. Chem.* **2005**, *77*, 6549.
(259) Guo, T.; Rudnick, P. A.; Wang, W.; Lee, C. S.; Devoe, D. L.; Balgley, B. M. *J. Proteome Res.* **2006**, *5*, 1469.
(260) Xie, H.; Bandhakavi, S.; Griffin, T. J. *Anal. Chem.* **2005**, *77*, 3198.
(261) Xie, H.; Rhodus, N. L.; Griffin, R. J.; Carlis, J. V.; Griffin, T. J. *Mol. Cell. Proteomics* **2005**, *4*, 1826.
(262) Immler, D.; Greven, S.; Reinemer, P. *Proteomics* **2006**, *6*, 2947.
(263) Malmstrom, J.; Lee, H.; Nesvizhskii, A. I.; Shteynberg, D.; Mohanty, S.; Brunner, E.; Ye, M.; Weber, G.; Eckerskorn, C.; Aebersold, R. *J. Proteome Res.* **2006**, *5*, 2241.
(264) Lemmo, A. V.; Jorgenson, J. W. *Anal. Chem.* **1993**, *65*, 1576.
(265) Hooker, T. F; Jorgenson, J. W. *Anal. Chem.* **1997**, *69*, 4134.
(266) Moore, A. W., Jr.; Jorgenson, J. W. *Anal. Chem.* **1995**, *67*, 3448.
(267) Moore, A. W., Jr.; Larmann, J. P., Jr.; Lemmo, A. V.; Jorgenson, J. W. *Methods Enzymol.* **1996**, *270*, 401.
(268) Bergstrom, S. K.; Dahlin, A. P.; Ramstrom, M.; Andersson, M.; Markides, K. E.; Bergquist, J. *Analyst* **2006**, *131*, 791.
(269) Cao, P.; Stults, J. T. *J. Chromatogr., A* **1999**, *853*, 225.
(270) Papayannopoulos, I. A. *Mass Spectrom. Rev.* **1995**, *14*, 49.
(271) Hunt, D. F.; Yates, J. R., 3rd; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6233.
(272) Hunt, D. F.; Zhu, N. Z.; Shabanowitz, J. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 122.
(273) Tabb, D. L.; Smith, L. L.; Breci, L. A.; Wysocki, V. H.; Lin, D.; Yates, J. R., 3rd. *Anal. Chem.* **2003**, *75*, 1155.
(274) Schutz, F.; Kapp, E. A.; Simpson, R. J.; Speed, T. P. *Biochem. Soc. Trans.* **2003**, *31*, 1479.
(275) van Dongen, W. D.; Ruijters, H. F.; Luinge, H. J.; Heerma, W.; Haverkamp, J. *J. Mass Spectrom.* **1996**, *31*, 1156.
(276) Bafna, V.; Edwards, N. *Bioinformatics* **2001**, *17* (Suppl. 1), S13.
(277) Chalkley, R. J.; Baker, P. R.; Hansen, K. C.; Medzihradszky, K. F.; Allen, N. P.; Rexach, M.; Burlingame, A. L. *Mol. Cell. Proteomics* **2005**, *4*, 1189.
(278) Chalkley, R. J.; Baker, P. R.; Huang, L.; Hansen, K. C.; Allen, N. P.; Rexach, M.; Burlingame, A. L. *Mol. Cell. Proteomics* **2005**, *4*, 1194.
(279) Clauser, K. R.; Baker, P.; Burlingame, A. L. *Anal. Chem.* **1999**, *71*, 2871.
(280) Colinge, J.; Masselot, A.; Giron, M.; Dessingy, T.; Magnin, J. *Proteomics* **2003**, *3*, 1454.
(281) Field, H. I.; Fenyo, D.; Beavis, R. C. *Proteomics* **2002**, *2*, 36.
(282) Hansen, B. T.; Jones, J. A.; Mason, D. E.; Liebler, D. C. *Anal. Chem.* **2001**, *73*, 1676.
(283) LeDuc, R. D.; Taylor, G. K.; Kim, Y. B.; Januszyk, T. E.; Bynum, L. H.; Sola, J. V.; Garavelli, J. S.; Kelleher, N. L. *Nucleic Acids Res.* **2004**, *32*, W340.
(284) Liebler, D. C.; Hansen, B. T.; Davey, S. W.; Tiscareno, L.; Mason, D. E. *Anal. Chem.* **2002**, *74*, 203.
(285) Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2337.
(286) Zhang, N.; Aebersold, R.; Schwikowski, B. *Proteomics* **2002**, *2*, 1406.
(287) Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* **1999**, *20*, 3551.
(288) Geer, L. Y.; Markey, S. P.; Kowalak, J. A.; Wagner, L.; Xu, M.; Maynard, D. M.; Yang, X.; Shi, W.; Bryant, S. H. *J. Proteome Res.* **2004**, *3*, 958.
(289) Craig, R.; Beavis, R. C. *Bioinformatics* **2004**, *20*, 1466.
(290) Yates, J. R., 3rd; Morgan, S. F.; Gatlin, C. L.; Griffin, P. R.; Eng, J. K. *Anal. Chem.* **1998**, *70*, 3557.
(291) MacCoss, M. J.; Wu, C. C.; Yates, J. R., 3rd. *Anal. Chem.* **2002**, *74*, 5593.
(292) Pappin, D. J.; Hojrup, P.; Bleasby, A. J. *Curr. Biol.* **1993**, *3*, 327.
(293) Craig, R.; Beavis, R. C. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2310.
(294) Duncan, D. T.; Craig, R.; Link, A. J. *J. Proteome Res.* **2005**, *4*, 1842.
(295) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res.* **1997**, *25*, 3389.
(296) Keller, A.; Purvine, S.; Nesvizhskii, A. I.; Stolyar, S.; Goodlett, D. R.; Kolker, E. *Omics* **2002**, *6*, 207.

(297) Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R. *Anal. Chem.* **2002**, *74*, 5383.

(298) Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. *Anal. Chem.* **2003**, *75*, 4646.

(299) Choudhary, J. S.; Blackstock, W. P.; Creasy, D. M.; Cottrell, J. S. *Proteomics* **2001**, *1*, 651.

(300) Rappsilber, J.; Mann, M. *Trends Biochem. Sci.* **2002**, *27*, 74.

(301) Kolker, E.; Purvine, S.; Galperin, M. Y.; Stolyar, S.; Goodlett, D. R.; Nesvizhskii, A. I.; Keller, A.; Xie, T.; Eng, J. K.; Yi, E.; Hood, L.; Picone, A. F.; Cherny, T.; Tjaden, B. C.; Siegel, A. F.; Reilly, T. J.; Makarova, K. S.; Palsson, B. O.; Smith, A. L. *J. Bacteriol.* **2003**, *185*, 4593.

(302) Sadygov, R. G.; Liu, H.; Yates, J. R. *Anal. Chem.* **2004**, *76*, 1664.

(303) Moore, R. E.; Young, M. K.; Lee, T. D. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 378.

(304) Higdon, R.; Hogan, J. M.; Van Belle, G.; Kolker, E. *Omics* **2005**, *9*, 364.

(305) Elias, J. E.; Gygi, S. P. *Nat. Methods* **2007**, *4*, 207.

(306) Elias, J. E.; Haas, W.; Faherty, B. K.; Gygi, S. P. *Nat. Methods* **2005**, *2*, 667.

(307) Xie, H.; Griffin, T. J. *J. Proteome Res.* **2006**, *5*, 1003.

(308) Yen, C. Y.; Russell, S.; Mendoza, A. M.; Meyer-Arendt, K.; Sun, S.; Cios, K. J.; Ahn, N. G.; Resing, K. A. *Anal. Chem.* **2006**, *78*, 1071.

(309) Blackler, A. R.; Klammer, A. A.; MacCoss, M. J.; Wu, C. C. *Anal. Chem.* **2006**, *78*, 1337.

(310) Kapp, E. A.; Schutz, F.; Connolly, L. M.; Chakel, J. A.; Meza, J. E.; Miller, C. A.; Fenyo, D.; Eng, J. K.; Adkins, J. N.; Omenn, G. S.; Simpson, R. J. *Proteomics* **2005**, *5*, 3475.

(311) Tabb, D. L.; McDonald, W. H.; Yates, J. R., 3rd. *J. Proteome Res.* **2002**, *1*, 21.

(312) Han, D. K.; Eng, J.; Zhou, H.; Aebersold, R. *Nat. Biotechnol.* **2001**, *19*, 946.

(313) McAfee, K. J.; Duncan, D. T.; Assink, M.; Link, A. J. *Mol. Cell. Proteomics* **2006**, *5*, 1497.

(314) Kislinger, T.; Rahman, K.; Radulovic, D.; Cox, B.; Rossant, J.; Emili, A. *Mol. Cell. Proteomics* **2003**, *2*, 96.

(315) Listgarten, J.; Emili, A. *Mol. Cell. Proteomics* **2005**, *4*, 419.

(316) Radulovic, D.; Jelveh, S.; Ryu, S.; Hamilton, T. G.; Foss, E.; Mao, Y.; Emili, A. *Mol. Cell. Proteomics* **2004**, *3*, 984.

(317) Butland, G.; Peregrin-Alvarez, J. M.; Li, J.; Yang, W.; Yang, X.; Canadien, V.; Starostine, A.; Richards, D.; Beattie, B.; Krogan, N.; Davey, M.; Parkinson, J.; Greenblatt, J.; Emili, A. *Nature* **2005**, *433*, 531.

(318) Krogan, N. J.; Cagney, G.; Yu, H.; Zhong, G.; Guo, X.; Ignatchenko, A.; Li, J.; Pu, S.; Datta, N.; Tikuisis, A. P.; Punna, T.; Peregrin-Alvarez, J. M.; Shales, M.; Zhang, X.; Davey, M.; Robinson, M. D.; Paccanaro, A.; Bray, J. E.; Sheung, A.; Beattie, B.; Richards, D. P.; Canadien, V.; Lalev, A.; Mena, F.; Wong, P.; Starostine, A.; Canete, M. M.; Vlasblom, J.; Wu, S.; Orsi, C.; Collins, S. R.; Chandran, S.; Haw, R.; Rilstone, J. J.; Gandi, K.; Thompson, N. J.; Musso, G.; St Onge, P.; Ghanny, S.; Lam, M. H.; Butland, G.; Altaf-Ul, A. M.; Kanaya, S.; Shilatifard, A.; O'Shea, E.; Weissman, J. S.; Ingles, C. J.; Hughes, T. R.; Parkinson, J.; Gerstein, M.; Wodak, S. J.; Emili, A.; Greenblatt, J. F. *Nature* **2006**, *440*, 637.

(319) Orchard, S.; Hermjakob, H.; Taylor, C. F.; Potthast, F.; Jones, P.; Zhu, W.; Julian, R. K., Jr.; Apweiler, R. *Proteomics* **2005**, *5*, 3552.

(320) Orchard, S.; Hermjakob, H.; Taylor, C.; Aebersold, R.; Apweiler, R. *Proteomics* **2005**, *5*, 4651.

(321) Hermjakob, H.; Montecchi-Palazzi, L.; Bader, G.; Wojcik, J.; Salwinski, L.; Ceol, A.; Moore, S.; Orchard, S.; Sarkans, U.; von Mering, C.; Roechert, B.; Poux, S.; Jung, E.; Mersch, H.; Kersey, P.; Lappe, M.; Li, Y.; Zeng, R.; Rana, D.; Nikolski, M.; Husi, H.; Brun, C.; Shanker, K.; Grant, S. G.; Sander, C.; Bork, P.; Zhu, W.; Pandey, A.; Brazma, A.; Jacq, B.; Vidal, M.; Sherman, D.; Legrain, P.; Cesareni, G.; Xenarios, I.; Eisenberg, D.; Steipe, B.; Hogue, C.; Apweiler, R. *Nat. Biotechnol.* **2004**, *22*, 177.

(322) Omenn, G. S.; States, D. J.; Adamski, M.; Blackwell, T. W.; Menon, R.; Hermjakob, H.; Apweiler, R.; Haab, B. B.; Simpson, R. J.; Eddes, J. S.; Kapp, E. A.; Moritz, R. L.; Chan, D. W.; Rai, A. J.; Admon, A.; Aebersold, R.; Eng, J.; Hancock, W. S.; Hefta, S. A.; Meyer, H.; Paik, Y. K.; Yoo, J. S.; Ping, P.; Pounds, J.; Adkins, J.; Qian, X.; Wang, R.; Wasinger, V.; Wu, C. Y.; Zhao, X.; Zeng, R.; Archakov, A.; Tsugita, A.; Beer, I.; Pandey, A.; Pisano, M.; Andrews, P.; Tammen, H.; Speicher, D. W.; Hanash, S. M. *Proteomics* **2005**, *5*, 3226.

(323) Adamski, M.; Blackwell, T.; Menon, R.; Martens, L.; Hermjakob, H.; Taylor, C.; Omenn, G. S.; States, D. J. *Proteomics* **2005**, *5*, 3246.

(324) Hermjakob, H. *Proteomics* **2006**, *6* (Suppl. 2), 34.

(325) Jones, P.; Cote, R. G.; Martens, L.; Quinn, A. F.; Taylor, C. F.; Derache, W.; Hermjakob, H.; Apweiler, R. *Nucleic Acids Res.* **2006**, *34*, D659.

(326) Martens, L.; Hermjakob, H.; Jones, P.; Adamski, M.; Taylor, C.; States, D.; Gevaert, K.; Vandekerckhove, J.; Apweiler, R. *Proteomics* **2005**, *5*, 3537.

(327) Hermjakob, H.; Apweiler, R. *Expert Rev. Proteomics* **2006**, *3*, 1.

(328) McDonald, W. H.; Tabb, D. L.; Sadygov, R. G.; MacCoss, M. J.; Venable, J.; Graumann, J.; Johnson, J. R.; Cociorva, D.; Yates, J. R., 3rd. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2162.

(329) Pedrioli, P. G.; Eng, J. K.; Hubley, R.; Vogelzang, M.; Deutsch, E. W.; Raught, B.; Pratt, B.; Nilsson, E.; Angeletti, R. H.; Apweiler, R.; Cheung, K.; Costello, C. E.; Hermjakob, H.; Huang, S.; Julian, R. K.; Kapp, E.; McComb, M. E.; Oliver, S. G.; Omenn, G.; Paton, N. W.; Simpson, R.; Smith, R.; Taylor, C. F.; Zhu, W.; Aebersold, R. *Nat. Biotechnol.* **2004**, *22*, 1459.

(330) Keller, A.; Eng, J.; Zhang, N.; Li, X. J.; Aebersold, R. *Mol. Syst. Biol.* **2005**, *1*, 1.

(331) Taylor, C. F.; Paton, N. W.; Garwood, K. L.; Kirby, P. D.; Stead, D. A.; Yin, Z.; Deutsch, E. W.; Selway, L.; Walker, J.; Riba-Garcia, I.; Mohammed, S.; Deery, M. J.; Howard, J. A.; Dunkley, T.; Aebersold, R.; Kell, D. B.; Lilley, K. S.; Roepstorff, P.; Yates, J. R., 3rd; Brass, A.; Brown, A. J.; Cash, P.; Gaskell, S. J.; Hubbard, S. J.; Oliver, S. G. *Nat. Biotechnol.* **2003**, *21*, 247.

(332) Garwood, K.; McLaughlin, T.; Garwood, C.; Joens, S.; Morrison, N.; Taylor, C. F.; Carroll, K.; Evans, C.; Whetton, A. D.; Hart, S.; Stead, D.; Yin, Z.; Brown, A. J.; Hesketh, A.; Chater, K.; Hansson, L.; Mewissen, M.; Ghazal, P.; Howard, J.; Lilley, K. S.; Gaskell, S. J.; Brass, A.; Hubbard, S. J.; Oliver, S. G.; Paton, N. W. *BMC Genomics* **2004**, *5*, 68.

(333) Garwood, K. L.; Taylor, C. F.; Runte, K. J.; Brass, A.; Oliver, S. G.; Paton, N. W. *Bioinformatics* **2004**, *20*, 2463.

(334) Levander, F.; Krogh, M.; Warell, K.; Garden, P.; James, P.; Hakkinen, J. *Proteomics* **2007**, *7*, 668.

(335) Garden, P.; Alm, R.; Hakkinen, J. *Bioinformatics* **2005**, *21*, 2085.

(336) Bellew, M.; Coram, M.; Fitzgibbon, M.; Igra, M.; Randolph, T.; Wang, P.; May, D.; Eng, J.; Fang, R.; Lin, C.; Chen, J.; Goodlett, D.; Whiteaker, J.; Paulovich, A.; McIntosh, M. *Bioinformatics* **2006**, *22*, 1902.

(337) Rauch, A.; Bellew, M.; Eng, J.; Fitzgibbon, M.; Holzman, T.; Hussey, P.; Igra, M.; Maclean, B.; Lin, C. W.; Detter, A.; Fang, R.; Faca, V.; Gafken, P.; Zhang, H.; Whiteaker, J.; States, D.; Hanash, S.; Paulovich, A.; McIntosh, M. W. *J. Proteome Res.* **2006**, *5*, 112.

(338) MacLean, B.; Eng, J. K.; Beavis, R. C.; McIntosh, M. *Bioinformatics* **2006**, *22*, 2830.

(339) Anderson, N. L.; Anderson, N. G. *Mol. Cell. Proteomics* **2002**, *1*, 845.

(340) Tirumalai, R. S.; Chan, K. C.; Prieto, D. A.; Issaq, H. J.; Conrads, T. P.; Veenstra, T. D. *Mol. Cell. Proteomics* **2003**, *2*, 1096.

(341) Davis, J. M.; Giddings, J. C. *J. Chromatogr.* **1984**, *289*, 277.

(342) MacNair, J. E.; Patel, K. D.; Jorgenson, J. W. *Anal. Chem.* **1999**, *71*, 700.

(343) MacNair, J. E.; Lewis, K. C.; Jorgenson, J. W. *Anal. Chem.* **1997**, *69*, 983.

(344) Wu, N.; Lippert, J. A.; Lee, M. L. *J. Chromatogr., A* **2001**, *911*, 1.

(345) Xiang, Y.; Yan, B.; Yue, B.; McNeff, C. V.; Carr, P. W.; Lee, M. L. *J. Chromatogr., A* **2003**, *983*, 83.

(346) Xiang, Y.; Yan, B.; McNeff, C. V.; Carr, P. W.; Lee, M. L. *J. Chromatogr., A* **2003**, *1002*, 71.

(347) Mellors, J. S.; Jorgenson, J. W. *Anal. Chem.* **2004**, *76*, 5441.

(348) Xiang, Y.; Liu, Y.; Stearns, S. D.; Plistil, A.; Brisbin, M. P.; Lee, M. L. *Anal. Chem.* **2006**, *78*, 858.

(349) Eschelbach, J. W.; Jorgenson, J. W. *Anal. Chem.* **2006**, *78*, 1697.

(350) Motoyama, A.; Venable, J. D.; Ruse, C. I.; Yates, J. R., 3rd. *Anal. Chem.* **2006**, *78*, 5109.

(351) Syka, J. E.; Marto, J. A.; Bai, D. L.; Horning, S.; Senko, M. W.; Schwartz, J. C.; Ueberheide, B.; Garcia, B.; Busby, S.; Muratore, T.; Shabanowitz, J.; Hunt, D. F. *J. Proteome Res.* **2004**, *3*, 621.

(352) Haas, W.; Faherty, B. K.; Gerber, S. A.; Elias, J. E.; Beausoleil, S. A.; Bakalarski, C. E.; Li, X.; Villen, J.; Gygi, S. P. *Mol. Cell. Proteomics* **2006**, *5*, 1326.

(353) Everley, P. A.; Bakalarski, C. E.; Elias, J. E.; Waghorne, C. G.; Beausoleil, S. A.; Gerber, S. A.; Faherty, B. K.; Zetter, B. R.; Gygi, S. P. *J. Proteome Res.* **2006**, *5*, 1224.

(354) Garcia, B. A.; Joshi, S.; Thomas, C. E.; Chitta, R. K.; Diaz, R. L.; Busby, S. A.; Andrews, P. C.; Ogorzalek, Loo, R. R.; Shabanowitz, J.; Kelleher, N. L.; Mizzen, C. A.; Allis, C. D.; Hunt, D. F. *Mol. Cell. Proteomics* **2006**, *5*, 1593.

(355) Makarov, A. *Anal. Chem.* **2000**, *72*, 1156.

(356) Hardman, M.; Makarov, A. A. *Anal. Chem.* **2003**, *75*, 1699.

(357) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. *J. Mass Spectrom.* **2005**, *40*, 430.

(358) Olsen, J. V.; de Godoy, L. M.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. *Mol. Cell. Proteomics* **2005**, *4*, 2010.

(359) Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528.
(360) Xia, Y.; Chrisman, P. A.; Erickson, D. E.; Liu, J.; Liang, X.; Londry, F. A.; Yang, M. J.; McLuckey, S. A. *Anal. Chem.* **2006**, *78*, 4146.
(361) Schroeder, M. J.; Webb, D. J.; Shabanowitz, J.; Horwitz, A. F.; Hunt, D. F. *J. Proteome Res.* **2005**, *4*, 1832.
(362) Coon, J. J.; Ueberheide, B.; Syka, J. E.; Dryhurst, D. D.; Ausio, J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9463.

CR068279A