Proteomics of Integral Membrane Proteins—Theory and Application

Anna E. Speers and Christine C. Wu*

Department of Pharmacology, University of Colorado School of Medicine, P.O. Box 6511, MS 8303, Aurora, Colorado 80045

Received June 6, 2007

Contents

1. Introduction

Integral membrane proteins (IMPs) lie at the critical junctions between intracellular compartments and cells and their environment. As such, IMPs are in a unique position to mediate a host of cellular processes, including intercellular communication, vesicle trafficking, ion transport, protein translocation/integration, and propagation of signaling cascades.[1−3] Thus, it is not coincidence that some of the largest classes of drug targets—G-protein coupled receptors (GPCRs), ion channels, transporters, cytochrome P450—are IMPs.[4,5] The hydrophobic core of the phospholipid bilayer

* To whom correspondence should be addressed: telephone, 303-724-3351; fax, 303-724-3663; e-mail, Christine.Wu@uchsc.edu.

10.1021/cr068286z CCC: $65.00 © 2007 American Chemical Society
Published on Web 08/08/2007

Anna Speers studied chemistry (B.A.) at Amherst College. She then attended The Scripps Research Institute (TSRI) as an HHMI fellow, where she worked for Dr. Benjamin F. Cravatt developing methods for click-chemistry activity-based protein profiling. Anna obtained her Ph.D in chemistry from TSRI in 2005. She is currently conducting postdoctoral research at the University of Colorado School of Medicine with Dr. Christine Wu. Her current research focus is method development for shotgun proteomic analysis of membrane proteins and chemical methods for their topological characterization.

Christine Wu received her B.S. in biochemistry from UCLA and her Ph.D. in cell biology from the University of Colorado Health Sciences Center, focusing on the Golgi proteome and molecular mechanisms of secretion using 2D gel-based proteomics platforms in the laboratory of Dr. Kathryn Howell. Her interest in gel-free proteomics technology lead her to The Scripps Research Institute, where she conducted postdoctoral research with Dr. John Yates III, developing shotgun methods targeted toward the analysis of complex membrane-enriched samples. She is currently an Assistant Professor in the Department of Pharmacology at the University of Colorado School of Medicine, where she continues to pursue quantitative, global analysis strategies for integral membrane proteins.

is the driving force behind IMP structure. Because peptide bonds are highly polar, proteins must adopt secondary structures that shield the backbone from the hydrophobic lipid core by allowing extensive hydrogen bonding between backbone amides and carbonyls. As a result, membrane-spanning proteins are characterized by one of two structural features: α-helical bundles or β-barrels.

IMP Structure. *β-Barrel* proteins, or porins, exist in the outer membranes of Gram-negative bacteria, chloroplasts, and mitochondria, where they regulate membrane integrity and allow for the passive influx/efflux of small molecules. For bacteria at least, β-barrel proteins likely account for only a few percent of all open reading frames (ORFs).[6] The β-strands are amphiphilic in nature, having alternating polar and hydrophobic residues in contact with the central pore and lipids, respectively. As a result, the overall hydropathy of β-barrel IMPs is similar to that of soluble proteins,[7−9] and they tend not to present the same analytical challenge as α-helical IMPs.[10]

α-Helix Bundle proteins, the subject of this review, are abundant species found in all membrane types, except for the outer membrane of Gram-negative bacteria. IMPs are predicted to make up 20−25% of all ORFs in most genomes.[11,12] They are divided into the following categories: *Bitopic*, or single-pass, IMPs have one transmembrane α-helical domain (TMD) with exposed globular domains on either side of the membrane. Bitopic proteins often act as cell surface markers, receptors, or adhesion factors, with the cytoplasmic domains operating in cellular signaling pathways or in contact with the cytoskeleton.[13] *Polytopic* IMPs have multiple TMDs arranged in a bundle oriented approximately normal to the membrane plane.[14] Many small molecule transport proteins in eubacteria, archaea, fungi, and plants are polytopic IMPs with 6 or 12 TMDs, while the 7-TMD GPCRs account for almost 5% of mammalian protein coding genes.[11] *Monotopic*, or membrane-anchored proteins, are not transmembrane proteins but are rather tethered to the membrane bilayer by a lipid anchor, such as glycosylphosphatidylinositol (GPI). Such proteins can be dissociated upon cleavage of the anchor with phospholipases, and they have the hydrophilic characteristics of soluble proteins.[15] *Membrane-associated* proteins are bound to the membrane surface via noncovalent interactions with phospholipid head groups or membrane-embedded proteins. Many such proteins can be readily solubilized by treatment with high-pH or high-salt buffers, and they do not generally present the same analytical challenge as IMPs.

Proteomics of Intact Proteins vs Peptides. Nearly all high-throughput proteomic platforms use mass spectrometry (MS) to identify proteins and elucidate the details of their primary sequences. Most studies take the *bottom-up* approach, where intact proteins are digested by enzymes or chemical means into peptides for MS or tandem MS (MS/MS) analysis (see next paragraph). Bottom-up proteomics can yield quantitative information as well, using methods such as spectral counting or stable isotope labeling. The downside to this approach is that much of the sequence information can be lost, as the task of identifying all peptides resulting from a complex digest is currently unfeasible, due to limits in separation, instrumentation cycle times, and variable peptide ionization/fragmentation efficiencies. Sequence coverage can be increased with replicate analyses, but one is still left with the task of reconstructing detailed protein information from an array of peptide data, and the ability to distinguish between protein isoforms, splice variants, or modified states involving multiple, varied post-translational modifications (PTMs) may be lost. As such, the bottom-up approach is most useful for global proteomic surveys and providing leads for further, detailed studies. The complementary *top-down* approach uses MS and MS/MS methods for the detailed study of intact proteins. Here, there is no disconnect in information, allowing the complete classification of any sequence variations and all PTMs. It is much less high-throughput, and is mostly used for single protein analysis due to difficulty in separating complex mixtures and the need for sufficient quantities of material. As such, this review will be limited to the more global bottom-up approach for IMPs. Whitelegge et al.[16] have recently reviewed top-down proteomics for membrane proteins.

Common Proteomic Platforms Defined. Most IMP bottom-up proteomic platforms can be grouped into one of three categories: (1) separation of proteins by two-dimensional polyacrylamide gel electrophoresis (2D PAGE or 2DE), followed by in-gel trypsin digest and MS analysis; (2) separation of proteins by one-dimensional sodium dodecyl sulfate PAGE (1D SDS-PAGE or 1DE), in-gel trypsin digest, peptide separation by reversed-phase (RP) microcapillary liquid chromatography (μLC), and MS analysis; or (3) shotgun[17] analysis of complex protein digests by chromatographic, electrophoretic, isoelectric, and/or affinity separation followed by MS analysis. The most widely used shotgun platform is Multidimensional Protein Identification Technology (MudPIT),[18,19] which involves peptide separation by strong cation exchange (SCX)—either on-line or off-line—followed by RPμLC.

Analysis of peptides by MS can be accomplished in one of two ways: in the case of highly resolved gel spots from 2DE, matrix-assisted laser desorption/ionization (MALDI)[20] peptide mass fingerprinting (PMF)[21] is commonly used. The technique measures the mass of four or five peptide ions using time-of-flight (TOF), from which the parent protein can be assigned. Alternatively, one can use MALDI or electrospray ionization (ESI)[22] MS/MS instruments (e.g., ion trap, TOF/TOF, Q-TOF) capable of providing fragmentation spectra of peptide ions from which their sequences can be determined. The higher information content allows for many protein assignments based only on one peptide, and locations of PTMs can be more precisely determined. MS-based proteomics and instrumentation has recently been reviewed by Aebersold and colleagues.[23]

Review Summary. Problems with α-helical IMPs are twofold: (1) their hydrophobic TMDs, which resist interaction with aqueous buffers, and (2) their general low abundance. Successful strategies for the global analysis of membrane proteins must take into account both of these factors for sample preparation and analysis. Fortunately, much effort in recent years has led to the development of multiple techniques tailored to the analysis of IMPs. This review will discuss (1) enrichment strategies, including subcellular fractionation, removal of peripheral membrane proteins, and delipidation; (2) principles and uses of common denaturing/solubilizing agents; (3) separation strategies, including gel-based and gel-free platforms; (4) digestion strategies specifically targeted toward soluble and/or membrane-embedded domains; (5) quantification methods tailored to IMPs; (6) principles of ionization and fragmentation of hydrophobic peptides; and (7) global topology assignment by prediction algorithms and experimental means. Where instructive, insight into the underlying chemistry is provided.

Method Evaluation. For the different techniques discussed throughout this review, the total proteins identified, IMP enrichment and diversity, and TMD coverage are included whenever that information is available. These parameters can be used as benchmarks for relative comparison. In general, the common topology prediction algorithms used to estimate IMP enrichment (e.g,. TMHMM(v1 or 2),[12] SOSUI,[24] HMMTOP2[25]) are highly capable of distinguishing between soluble proteins and IMPs, allowing for comparison across methods, and any outliers will be noted. Two good numbers to keep in mind for evaluation of enrichment strategies are that 20−25% of ORFs are predicted to encode IMPs and that, in an unfractionated proteome, IMPs typically make up only ∼5−15% of identified proteins. A key indicator of comprehensive IMP analysis is the identification of numerous complex polytopic IMPs, which may be harder to identify than simpler proteins with 1−2 TMDs owing to their relatively higher TMD/soluble domain ratio. As such, IMP-specific protocols should always report not only the numbers but also the kinds of IMPs identified. It should also be noted that, in much of the literature, the term "membrane protein" typically includes integral, associated, and anchored proteins, and thus, it should not be interpreted to imply *integral* membrane protein.

Another parameter to consider is whether or not TMD-containing peptides are identified or if sequence coverage is restricted to soluble domains. Because IMPs may have a significant portion of their sequence embedded in the membrane, analysis of TMD peptides increases the probability of identifying IMPs, in addition to providing a more comprehensive study. TMD coverage can be assessed either directly, by looking at overlap between identified peptides and TMDs predicted by topology algorithms, or indirectly, via hydropathy analysis using the GRAVY (grand average of hydropathy) index, introduced by Kyte and Doolittle.[26] While IMPs may have either positive (hydrophobic) or negative (hydrophilic) GRAVY scores depending on the ratio of soluble to membrane-embedded residues, positive GRAVY scores are highly correlated with TMDs due to their high content of aliphatic residues. As such, the presence of peptides with positive GRAVY scores is indicative of a good TMD-targeted identification strategy.

It must be emphasized that the relative success of different methods is highly dependent upon individual expertise and available instrumentation, and methods should be judged based on the best, most reproducible examples. An additional consideration is the level of automation and computation power available, which can have a very large impact on the time frame for completing specific experiments. Attempts were made to include all recent (∼2003−2007) research that demonstrates a novel technology applied to IMPs and/or an efficient IMP enrichment strategy, or is otherwise instructive. We apologize in advance to any author whose work we have inadvertently overlooked. Impressive work on the proteomics of IMPs has been made for a variety of organisms; however, because plant membrane proteomics has been addressed in recent reviews,[15,27−29] the subject will not be discussed here.

2. Enrichment Strategies

2.1. Organellar Proteomics and Subfractionation

Fractionation for organisms with smaller proteomes, such as bacteria and yeast, often involves cell lysis and sequential centrifugation steps to remove cell debris and isolate the membrane from the soluble fraction.[30−32] Further purification can be accomplished by sucrose density gradient centrifugation[33,34] or sodium carbonate wash to remove membrane associated proteins[35,36] (see section 2.2.1). However, for the more complex mammalian proteomes, which may contain upward of 50 000 proteins,[37] fractionation on multiple levels—tissue, organelle, protein, and peptide—is essential for comprehensive analysis. One of the most widely used techniques for subcellular fractionation is density gradient centrifugation (e.g., using sucrose, sorbitol, Ficoll, or Percoll), which has been used to isolate numerous subcellular structures, for example, plasma membranes,[38,39] mitochondria,[40−42] mitochondrial inner/outer membranes,[43,44] Golgi,[45] clathrin-coated vesicles,[46] synaptic vesicles,[47−49] and platelet membranes,[50] based on the properties of specific membrane

subtypes, prior to proteomic analysis. Free-flow electrophoresis (see section 4.2.1) has been used to isolate plasma membrane vesicles,[47,51,52] detergent-resistant membranes,[53] and mitochondria.[54] Alternatively, subcellular structures can be affinity purified by immunoisolation, for example, synaptic vesicles[47,49] and caveolae.[55,56]

Of all subcellular structures of higher eukaryotes, the plasma membrane (PM) has probably received the most attention by the proteomics community. Composed of the lipid bilayer and membrane-associated and embedded proteins, it is responsible for maintaining a physical barrier between a eukaryotic cell and its environment. The protein components carry out many important biological functions, including intercellular communication, transport of ions/solutes, and signal transduction. As such, characterizing the plasma membrane proteome, particularly IMPs, which bridge the intra- and extracellular environments and directly effect cellular interaction with the outside environment, is of paramount importance. However, it can be difficult to isolate PM because (1) density-separated fractions may be contaminated by other cellular membranes, (2) it is relatively low-abundant compared the entire cellular membrane complement, (3) its physiochemical properties make it similar to other membrane components, and (4) it has a tendency to exist as multiple structures.[57]

2.1.1. Colloidal Silica

One solution to the problem of PM isolation from cultured cells was presented by Jacobson and co-workers,[58,59] in which they applied a thin layer, or pellicle, of cationic, aluminum chlorohydroxide-coated colloidal silica to intact cells. The anionic phospholipid head groups of the membrane, as well as carbohydrate moieties of glycoproteins, form a strong ionic interaction with the positively charged silica—the same principle used in the polylysine coating of cell culture dishes to promote cell adherence. The cationic particles are then cross-linked using the anionic polymer poly(acrylic acid), forming a supportive matrix. Upon cell lysis, the PM remains associated with the support in the form of large, open sheets, which can be readily isolated by centrifugation owing to the increased density afforded by the pellicle. Membranes can then be washed with sodium carbonate and solubilized by SDS. If adherent cultured cells are treated with the colloidal silica, the pellicle-coated apical surfaces can be selectively removed, while the basolateral membrane remains attached to the plate, providing a means to differentially profile the two cell surfaces.[59]

Rahbar and Fenselau[57] applied the colloidal silica technique to MCF-7 cells using 1DE coupled with μLC-MS/MS, identifying 366 proteins, with an estimated 43% localized to the PM. In a more recent study, 540 proteins were identified with ~40% localized to the PM.[60]

The colloidal silica technique is not limited to cultured cells. For example, Schnitzer and colleagues have isolated PM directly from lung tissue.[61,62] In one experiment, isolated PMs were analyzed by MudPIT, giving 450 identified proteins, of which 81% were PM-localized. Of the total PM proteins, 31% were IMPs or GPI-anchored proteins (known in literature, or predicted by TMHMM). Thus, colloidal silica can provide a viable means to enrich for PM, although, like any method, it may require a certain degree of skill and optimization to achieve better enrichment for PM than that afforded by density gradient centrifugation (e.g., 42−51% for rat liver PM reported by Zhang et al.).[38,39]

2.1.2. Aqueous–Polymer Two-Phase Partitioning

As discussed in a recent review,[63] partitioning of membranes in aqueous−polymer two-phase systems is one method for affinity purifying the PM. If two structurally distinct water-soluble polymers, such as polyethylene glycol (PEG) and dextran, are mixed in aqueous solution, above a critical concentration, the polymers will eventually separate into two distinct phases, with membranes preferentially segregating into the top (PEG), more hydrophobic phase according to the specific surface properties, dictated, at least in part, by their lipid composition. In general, partition favors PM > Golgi > lysosomes > ER > mitochondria. To improve separation, a secondary partitioning step, employing a PM-specific affinity ligand such as wheat germ agglutinin (WGA) conjugated to one of the polymers, generally dextran (as its higher MW means that its partitioning behavior is less influenced by the ligand), can be used. Ideally, polymer concentrations are chosen so that any contaminating membranes left over from the first step will partition into the PEG phase, while the affinity ligand causes the target membrane to partition into the dextran phase. Due to its sensitivity to salt, temperature, affinity ligands, and exact polymer concentration, application of the affinity two-phase partition method requires carefully controlled conditions and optimization for specific tissue sources.[63] Two-phase partitioning is also compatible with nonionic detergent-solubilized membranes for partitioning of proteins.[64]

When Schindler et al.[65] applied two-phase affinity partitioning to brain plasma membranes followed by 1DE-μLC-MS/MS, out of ~500 total proteins, 42% were estimated to be PM proteins and 49% were estimated to be IMPs, including a few (at least 50) with >2 TMDs (DAS-TMfilter; may overpredict TMDs with a higher rate than TMHMM[66]).

In another report, Cao et al.[67] analyzed rat liver PM purified by a sucrose density gradient and two-phase partitioning, identifying 428 proteins following 1DE-μLC-MS/MS. Plasma membrane enrichment was estimated at 67%, with a fair number (87, 43%) of proteins integral to the plasma membrane, including a few (15) having more than two TMDs (TMHMM).

From these limited examples, it appears that two-phase partitioning method does provide enrichment of PM, on par or slightly better than what could be expected from density gradient centrifugation. Like the colloidal silica technique, aqueous−polymer two-phase partitioning requires a certain degree of expertise as well as differential optimization for each sample.

2.2. Enriching for Integral Membrane Proteins

2.2.1. Removing Membrane-Associated Proteins

Because IMPs are often in very low abundance, multiple enrichment strategies are beneficial. Membrane-associated proteins interact with polar lipid head groups or IMPs and can be stripped (to some degree) using high-ionic-strength or high-pH buffers. Owing to their interactions with the lipid hydrocarbon chains, IMPs and anchored proteins are generally only solubilized using detergents or organic solvents.[13,68] For ionic dissociation, typical salts include sodium chloride, potassium chloride, sodium bromide, and potassium bromide. Concurrent sonication allows stripping of both sides of the membrane.

Alkaline treatment using sodium carbonate (or sodium hydroxide) is widely used to enrich for IMPs; see, e.g., refs

Table 1. Recommended Strategies for Removing Membrane-Associated Proteins and Lipids (Section 2.2)

	comments
removing associated proteins	
high-pH (sodium carbonate) wash	removes many (not all) membrane-associated proteins
high-ionic-strength (salt) wash	removes many (not all) membrane-associated proteins
enzymatic membrane shaving	enzyme removes membrane-associated proteins missed by high-salt/high-pH, see section 5.1
removing lipids	
methanol/chloroform precipitation	removes lipids that can interfere in digestion, chromatography, MS analysis

35, 39, 43, 50, 69, and 70. Additionally, the high-pH ($\geq$11) buffers cause membrane vesicles to open, allowing release of trapped soluble and membrane-associated proteins, with dissociation facilitated by sonication or mechanical agitation.[71−73] Pasini et al.[74] report that, with increasing number and strength of carbonate washes, the number of identified IMPs (as well as GPI-anchored proteins) remains stable, but membrane-associated proteins experience a significant decrease in abundance.

A combination of high-ionic-strength and high-pH washes can result in further enrichments of membranes; however, Fischer et al.[75] note that the efficacy of such washes can be membrane-dependent. Even in combination with sonication and detergents for inhibiting vesicle formation, washing procedures were ineffective at removing soluble and peripheral membrane proteins from the bacterium *C. glutamicum*, possibly due to its specific cell wall composition. Their solution to removal of soluble and membrane associated proteins was to predigest all loosely associated proteins (including IMP soluble domains) with trypsin, a strategy that has also been applied to mitochondria.[76] (Recommended strategies are summarized in Table 1.)

2.2.2. Delipidation

A second sample preparation concern specific to IMPs is the removal of lipids, which can interfere with enzymatic digestion (a property exploited for membrane shaving experiments, see section 5.1), gel electrophoresis,[74,77,78] and μLC.[73] The most widely used method (Table 1) involves protein precipitation by adding methanol/chloroform to a proteomic preparation,[79] whereby lipids partition into the chloroform layer and proteins precipitate at the chloroform/aqueous methanol interface. Separation can also be achieved via protein precipitation with cold acetone, either alone or in combination with other organics (e.g., refs 80−83), resulting in the selective solubilization of the lipid component. Similarly, acetone washing following trichloroacetic acid (TCA) precipitation will also remove lipids,[50,78] though it should be noted that TCA precipitation itself does not result in delipidation. Alternatively, ethanol treatment can be useful for the precipitation of proteins/solubilization of lipids.[74]

After precipitation, proteins are resolubilized using various buffers, denaturants, and detergents with the aid of heat, sonication, and/or mechanical agitation. In the case of methanol/chloroform, it has been found that sonication of the protein pellet in aqueous methanol produces a fine protein powder that is much more readily solubilized.[84]

3. Protein Solubilization and Denaturation

Protein solubilization and denaturation are of special consideration for IMPs, as their hydrophobic domains resist exposure to aqueous solvents, causing aggregation, adsorption, and precipitation, leading to sample loss and hindering enzymatic access during digestion. Thus, reagents must be specifically chosen so as to maintain IMP solubility and facilitate digestion without (1) interfering with separation, (2) overly attenuating proteolytic activity, or (3) compromising MS analysis. The properties of chaotropes, detergents, and other denaturants are discussed below, and a summary of solubilizing/denaturing agents is given in Table 2.

3.1. Chaotropes

Chaotropes are strong denaturing agents that stabilize unfolded protein states (as compared to native) via hydrogen bonds and electrostatic interactions.[85−87] The most common reagents are urea, thiourea, and guanidinium chloride. Urea and/or thiourea, in combination with various detergents, are often used to solubilize proteins for separation by isoelectric focusing (IEF; see section 4.1.1). Thiourea is better than urea at disrupting hydrophobic interactions, thus its inclusion for membrane protein applications. It is also considerably more soluble in aqueous solution if a high concentration of urea is present.[88] Typical trypsin digestion protocols for soluble proteins call for the use of urea or guanidinium chloride (thiourea can inhibit proteases[89]) to facilitate enzymatic attack of otherwise occluded domains. While urea will not extract most IMPs (particularly polytopic proteins) from the membrane,[73,90] urea does significantly facilitate digestion of exposed soluble domains in on-membrane digestion experiments (see sections 5.1.1 and 5.2.1). It should be noted that, when heated above 37 °C, urea can cause problematic carbamylation of N-termini and lysine residues.[88] Small molecules such as chaotropes (also salts, buffers) do not interfere with peptide analysis by standard μLC-MS/MS methods, as, unlike peptides, they do not bind to ion exchange or reversed-phase resins, and they are thus removed before peptides elute from the column. A separate desalting step is often necessary prior to MALDI analysis to prevent interference by high concentrations of chaotropes.[91]

3.2. Detergents

Detergents are a class of amphipathic molecules containing both hydrophilic and hydrophobic domains. This particular structure allows for self-association (e.g., micelle formation) and binding to hydrophobic surfaces such as those found in TMDs. In a broad sense, detergents are lipid-mimetic, but they are distinguished by the concentrations at which they self-associate and the particular higher-order structures they can adopt.[92] Detergents can be classified into four main groups—linear-chain ionic, nonionic, bile acid, and zwitterionic—with each class differing in its ability to solubilize and denature IMPs. Also included are detergents specifically designed to be compatible with MS analysis.

3.2.1. Ionic Detergents

Ionic detergents have a cationic or anionic head group attached to a hydrocarbon chain. The classic example of an anionic detergent is SDS, which is extremely efficient at solubilizing and denaturing proteins[93] and preventing adsorption onto container walls.[94] As such, SDS is instrumental

Table 2. Summary of Solubilizing/Denaturing Agents (Section 3)

class	common reagents	general characteristics	compatibility	
			enzymatic digest	MS
chaotropes	urea, thiourea, guanidinium chloride	strong denaturing agents, low potential for membrane disruption, used in conjunction with detergents for IMP solubilization and/or digestion	urea and guanidinium chloride generally compatible	may require desalting prior to MALDI, removed by μLC (does not bind to solid phase)
detergents				
ionic	SDS	strongest denaturing agent with high potential for membrane disruption, recommend solubilization at $\geq$1% with heating, generally used prior to SDS-PAGE or when affinity purification allows removal	generally compatible at 0.1%	causes severe ionization suppression, somewhat removed by SCX
nonionic	Triton X-100, NP-40, Brij, OG, digitonin, CYMAL-5	mild denaturants, moderate potential for membrane disruption, generally used for solubilization prior to IEF or to aid in in-gel digest (OG, CYMAL-5)	generally compatible	causes low to moderate ionization suppression, largely removed by SCX
bile acid salts	SDC, sodium cholate	mild denaturants, moderate potential for membrane disruption	2% SDC compatible with trypsin	SDC removed by acidification
zwitterionic	CHAPS, ASB-14	moderate denaturants, more potential for membrane disruption than nonionic/bile acid detergents	generally compatible	causes low to moderate ionization suppression
MS compatible	RapiGest (ionic), PPS (zwitterionic), Invitrosol (proprietary)	moderate denaturants, moderate potential for membrane disruption	generally compatible at ~0.1%	RapiGest/PPS removed by acidification; Invitrosol compatible with MALDI and ESI
organic solvents	60% methanol/ aqueous buffer	strongly denaturing, high potential for membrane disruption	attenuates enzyme activity	directly compatible or can be diluted/removed by evaporation
organic acids	90% FA, 70% TFA	strongly denaturing, high potential for membrane disruption	used with CNBr digest	directly compatible or can be diluted

for solubilization prior to SDS-PAGE, but its incompatibility with IEF has led to significant problems for membrane proteomics by traditional 2DE (see section 4.1.1). At low concentrations (0.1%), SDS is readily compatible with most digestive enzymes.

SDS is an efficient denaturant and solubilizing agent by virtue of two chemical properties: (1) its long, flexible hydrocarbon tail forms hydrophobic interactions with polypeptide chains—irrespective of amino acid identity or sequence—breaking existing intra-protein interactions, and 2) its anionic head group associates with positively charged side chains, disrupting ionic protein−protein interactions and maintaining electrostatic repulsion, preventing protein aggregation. SDS binds proteins with a 1.4/1 w/w ratio, encasing the protein in a virtual shell, which (1) may sterically hinder digestion enzymes and (2) can be difficult to remove completely by dialysis or ion exchange.[95] This can be quite problematic for proteomic MS applications because SDS can interfere with μLC[96−98] and severely suppress ionization by MALDI[94,99] and ESI.[100]

3.2.2. Nonionic Detergents

Nonionic detergents have polyoxyethylene (e.g., Triton X-100, NP-40, Brij) or glycosidic [e.g., *n*-octyl glucoside (OG), 5-cyclohexyl-1-pentyl-β-D-maltoside (CYMAL-5), digitonin] polar head groups combined with a hydrophobic tail. Because nonionic detergents disrupt lipid−lipid and lipid−protein interactions rather than protein−protein interactions, they are considered relatively mild detergents. For some classes, this effect can be strongly influenced by the hydrocarbon chain length, with shorter chain detergents capable of greater solubilization and structural disruption.[101,102] (Indeed, the insolubility of lipid rafts in nonionic detergents is often exploited for their isolation.[103,104]) Despite their mild nature, inclusion of small amounts of nonionic detergents such as OG[105−107] or CYMAL-5[108] facilitates recovery of IMP peptides during in-gel digest. Both reagents appear to be generally compatible with MALDI and μLC-MS/MS analysis, though some interference for the former was reported.[105,108] In general, the nonionic glycosidic detergents are fairly ESI-compatible at low (0.01−0.1%) concentrations (as measured by continuous infusion with myoglobin).[100] Triton X-100 tends to be problematic for MALDI[94] and ESI.[100]

3.2.3. Bile Acid Salts

Bile acid salts, such as sodium deoxycholate (SDC) and sodium cholate, are also ionic detergents, but they have significantly different properties than linear-chain ionic detergents such as SDS. They are steroidal compounds, having a polar and apolar face rather than a distinct "head" and "tail", and they have significantly less solubilizing/denaturing capability.[109] Zhou et al.[110] evaluated the use of SDC in place of SDS for in-solution and in-gel digestion of rat hippocampal plasma membrane as assessed by MALDI-PMF and μLC-MS/MS. Performing in-gel digest in the presence of 0.1% SDC improved membrane proteome analysis, giving 26 vs 19 protein IDs and 10 vs 5 IMPs as compared to the case of no additive detergent. For in-solution digest, SDC (in contrast to SDS) was found to be compatible with tryptic digest at concentrations up to 2%. Use of the bile salt also resulted in more total protein identifications (71 vs 31) and more IMPs (50 vs 22) by μLC-MS/MS. Because SDC precipitates at low pH, acidification prior to sample analysis results in SDC forming a gel-like precipitate, facilitating removal from solution and preventing any potential ion suppression.

3.2.4. Zwitterionic Detergents

Zwitterionic detergents such as CHAPS or the sulfobetaines (e.g., ASB-14), have intermediate properties, being better solubilizing agents than bile salts/nonionic detergents but not nearly as strong as linear-chain ionic species.[109] For example, CHAPS, a synthetic derivative[111] of the naturally occurring bile salts, is generally nondenaturing[112] but is a better solubilizing agent than the bile salts and more disruptive to protein−protein interactions than either the bile salts or Triton X-100.[113] It also does not appear to significantly interfere with trypsin digest[94] or ESI (0.1%, continuous infusion of myoglobin),[100] but it can suppress ionization using MALDI.[94]

3.2.5. "MS-Compatible" Detergents

As mentioned above, many detergents can significantly interfere with MS analysis.[100] To circumvent some of these problems, "MS-compatible" detergents have been introduced. These include the acid-labile surfactants RapiGest (ionic, Waters)[114] and PPS (zwitterionic, Protein Discovery),[115] which have a labile ketal functional group between the hydrophilic head group and the hydrophobic tail and thus break down into innocuous nonsurfactant byproducts upon hydrolysis. Invitrogen has also marketed MALDI- and LC/MS-compatible Invitrosol reagents (proprietary), which either do not interfere with PMF/protein analysis or have orthogonal LC elution (high organic %) to most peptides, respectively. An acetal-based combination detergent/MALDI matrix has also been reported,[116] as well as fluoride-cleavable and photolabile detergents.[117]

To compare the relative efficiencies of some "MS-compatible" detergents for the analysis of a complex proteome, Blackler et al.[118] solubilized an enriched HeLa cell membrane preparation in 1% solutions of RapiGest, PPS, Invitrosol, SDS, CHAPS, or Triton X-100 or in 8 M urea. Solutions were diluted to 0.1% detergent (or 2 M urea) for trypsin digest. The acid-labile surfactants (RapiGest, PPS) were hydrolyzed, and all samples were analyzed by μLC-MS/MS. It was found, not surprisingly, that SDS performed very poorly due to significant ion suppression. In contrast, the commercially marketed MS-compatible detergents RapiGest and PPS performed relatively well, with ∼220 proteins identified for each condition and IMPs making up 20−25% of the total. Invitrosol, CHAPS, Triton X-100, and 8 M urea performed moderately well, with a little over half as many proteins identified and a similar percentage (20−25%) of IMPs. It is somewhat surprising that Trition X-100 gave decent results, given its reputation as an MS-incompatible reagent and the presence of obvious detergent peaks in the chromatogram. In a similar MudPIT study by Ruth et al.,[119] RapiGest was found to be fairly comparable to 8 M urea as a solubilizing agent prior to trypsinization. Because the relative success of each method depends on a variety of factors relating to digestion efficiency, chromatography, and MS detection/identification, the strengths and weaknesses of each reagent are difficult to tease apart based on these few experiments. The outcome for each detergent is also highly sample and protocol dependent. It is clear, however, that the relative effectiveness of a detergent as a solubilizing/denaturing agent is only one of the factors that must be considered for the successful analysis of IMPs using a shotgun approach.

3.2.6. Principles of Detergent Solubilization

For the more mild (nonionic) classes of detergents, solubilization is a process by which detergent molecules first partition into the membrane vesicle bilayer, until the native membrane can accommodate no more detergent without significant structural alteration due to the destabilizing effects of detergent−detergent interactions. This results in bilayer fragmentation and the formation of mixed detergent/phospholipid micelles. Individual proteins/oligomers thus become embedded in their own small (nonsedimenting) mixed micelles or sheet structures, with hydrophobic TMDs largely encased in a detergent coating. At this point, proteins are considered solubilized, and any further addition of detergent simply leads to dilution of phospholipid. It should be noted that encasement of hydrophobic regions may not be complete, potentially leading to protein aggregation.[102,120,121]

In contrast, SDS is postulated to achieve solubilization by a different mechanism. SDS is observed to interact more strongly with proteins than with lipids, effecting protein extraction and unfolding prior to solubilization of the bilayer, as demonstrated by the model protein Ca_2^+ ATPase.[121] Second, due to the low flip-flop rate of charged detergents such as SDS, membrane solubilization is thought to proceed by outer-leaflet phospholipid extraction into detergent micelles, leading to destabilization, fragmentation, and eventual micelle solubilization, rather than supersaturation of the native membrane. Additionally, unlike more mild detergents, SDS is capable of solubilizing protein aggregates formed during the intermediate stages of solubilization.[121]

3.2.7. Detergent Concentrations

The critical micelle concentration (CMC) is the detergent concentration at which self-association gives rise to micelle formation. The CMC decreases with alkyl chain length and increases with double bonds and branch points; for ionic detergents, CMC lessens with greater concentrations of counterion.[109] It is at or near the CMC that enough detergent is present to disrupt membrane vesicles so that solubilization of IMPs can begin to occur.[102,120] In general, detergent concentrations in the range of 1−2% are used to solubilize proteins, which is above the CMC of most common detergents (for list see le Maire et al.[102] and Banerjee et al.[112]). It should be noted that CMCs for all detergents vary according to the particular characteristics and composition of the solution (detergent, protein, lipid, salt, pH, temperature, etc.), and the properties of mixed systems are not readily predicted based on the characteristics of less complex systems.[92]

In the case of SDS, for optimal solubilization/denaturation, membrane preparations can be dissolved in ≥1% SDS and heated briefly (∼100 °C) to facilitate complete denaturation. Because most digestive enzymes will be inactivated at such high SDS concentrations as well, the SDS concentration should be diluted to ∼0.1% for trypsin digest. This reduction in SDS below the CMC (∼0.23%)[112] will allow for some degree of protein refolding, particularly for the more hydrophobic species, but it provides better denaturation than if proteins were initially solubilized in 0.1% SDS.[95]

3.2.8. Phase Separation

The cloud point, or critical point, is the temperature at which a clear, homogeneous nonionic detergent solution becomes turbid upon heating and eventually separates into

Table 3. Summary of Separation Strategies (Section 4)

	IDs[a]	%IMP	comments
		Gel-Based Protein Separation	
IEF/SAS-PAGE	1000's	<15%	not advised for IMPs, good examples (~30% IMP) extremely limited/not easily reproduced
BN/SDS-PAGE	low 100's	~50%	good separation of protein complexes
16-BAC/SDS-PAGE	low 100's	~50%	not highly orthogonal separation
dSDS-PAGE	~100	~30% (?)[b]	not highly orthogonal separation, limited examples
1DE-μLC	high 100's	~50%	well established, very widely used, best gel-based method
		Solution-Phase Protein Separation	
solution IEF/1DE-μLC or 2DE	low 100's	~50%	limited examples
ion exchange/1DE-μLC	low 100's	~50%	limited examples
RPLC/1DE or μLC	low 100's	~20% (?)	limited examples, best at elevated temperature
		Shotgun Methods	
μCIEF/μLC	1000's	~25% (?)	limited examples, not applied to optimized samples, promising technique
MudPIT	1000's	~65%	well established, very widely used
μLC or MudPIT + heat	high 100's	~80−98%	combined with new TMD-targeted sample preparation technique (see section S.2.1), greatly improves recovery of hydrophobic peptides

[a] Numbers are representative of the average protein identifications (IDs) or IMP enrichment. [b] (?) indicates higher enrichment may be possible, but current examples are limited

detergent-rich and detergent-poor phases. This behavior has been exploited for the purification of membranes, which tend to partition into the detergent-rich phase.[92] Triton X-114 has a particularly advantageous cloud point of 22 °C; thus, proteins can be solubilized at or near 0 °C, partitioned into the detergent phase upon heating to 30 °C, and then separated by centrifugation.[122] However, proteins besides IMPs tend to partition into the detergent-rich phase, giving incomplete separation.[9] Aqueous−polymer systems (e.g., PEG/dextran) will also separate according to the same principle (for a review, see ref 63).

3.3. Aqueous–Organic Solvents

Aqueous−organic solvent systems (acetonitrile−water, methanol−water) are an alternative to detergents for facilitating protein digestion. Unlike most detergents, organic−aqueous solutions are directly compatible with μLC-MS/MS or can be easily removed by evaporation, reducing sample handling and potential loss. While such solutions aid in the denaturation and solubilization of substrate proteins,[123−125] trypsin and other proteases retain a useful degree of activity.[124,126,127] The denaturation potential of organic solvents is due to their increased hydrophobicity relative to water, which allows for some stabilizing effect on nonpolar residues exposed in the unfolded state. Studies by Russell et al.[125] compared digestion of model proteins in organic−aqueous vs aqueous buffers, with 60% methanol, 40% acetone, 40% acetonitrile, or 80% acetonitrile affording superior digestion. Trifluoroethanol has also been advanced as an effective organic cosolvent for protein solubilization.[128−131]

In addition to protein denaturation, organic solvents can have disruptive effects on membrane bilayers and, thus, can be used to solubilize IMPs as well. In particular, 60% methanol has proven effective in solubilizing detergent resistant lipid rafts prior to trypsin digestion[69,132−134] (see also section 5.2). Effects of short-chain alcohols (e.g., methanol, ethanol) on lipid membranes have been studied extensively in the literature (see refs 135−137 and references therein). Owing to their different hydrophobicities, ethanol and methanol tend to interact in distinct ways with the membrane. Ethanol localizes to the hydrophilic headgroup region, just below the membrane−water interface, forming extensive hydrogen bonds with resident lipids. In contrast, methanol maintains a shell of water molecules—and possibly substantial contact with bulk water—deforming the membrane by forming deep pockets or wells in the membrane surface, where it accumulates within/below the lipid headgroup region; however, due to its solvation shell, it does not directly interact with lipids.[135,136] In general, as the number of CH_2 groups increases, an alcohol's interactions with lipid tails become more favorable, allowing for faster crossover rates and increased incorporation into the bilayer. This partitioning alters vesicle shape and increases membrane fluidity, disorder, and permeability.[135−137] In excess, short-chain alcohols have the potential to severely compromise membrane integrity, leading to IMP solubilization and denaturation.

3.4. Organic Acids

Organic acids [e.g., formic acid (FA), trifluoroacetic acid (TFA)] are also capable of membrane disruption and IMP solubilization, and they are compatible with downstream μLC-MS/MS analysis. The technique was originally applied to membrane preparations by Washburn et al.,[18] and it has since been used by a number of groups. Martosella et al.[138] used 80% FA to solubilize IMPs prior to RPLC separation, achieving good recovery of very hydrophobic lipid raft proteins. In another example, Da Cruz et al.[43] solubilized a membrane preparation in a minimal volume of 90% FA prior to dilution with an ammonium carbonate/8 M urea buffer for trypsin digestion, identifying a number of IMPs by off-line MudPIT. However, by far the most common application is membrane proteome solubilization in 90% FA followed by in-solution digestion with cyanogen bromide (CNBr) for methionine-directed cleavage[18,19,45,73,75,84,139−141] (see also sections 5.2 and 6.3). In contrast, in-gel digestion with CNBr has been performed using 50−70% TFA to aid in protein denaturation.[36,105,106]

4. Separation Strategies for Proteins and Peptides

A summary of separation strategies is given in Table 3.

4.1. Gel-based Methods for Protein Separation

4.1.1. Isoelectric Focusing (IEF)/SDS-PAGE

IEF followed by SDS-PAGE (traditional 2DE) has long been the standard method for the analysis of complex protein

mixtures. This technique allows for the resolution of >1000 proteins and can be combined with quantitation techniques such as difference gel electrophoresis (DIGE).[142] Resolved protein spots are excised from the gel and subject to in-gel tryptic digestion and analysis by MS (either PMF or MS/MS) to identify proteins of interest. However, standard techniques are, by and large, not compatible with the proteomics of IMPs. Proteins are amphoteric compounds, having both acidic and basic functional groups. As such, there is a specific pH, defined as the isoelectric point (pI), at which a protein carries no net charge and is thus stable in an electric field. This property is exploited for first dimension IEF, which separates proteins based on their pI. Proteins are introduced to a pH gradient (e.g., created by carrier ampholites), and an electrical current is applied. At a pH either above or below their pI, proteins are negatively or positively charged, respectively, and migrate toward the opposite pole until their pI is reached. Under these neutral conditions, solubility in aqueous buffer is at a minimum, and highly hydrophobic proteins tend to precipitate, precluding transfer into the second dimension. In the best case scenario, polytopic proteins will partially transfer into the SDS-PAGE second dimension, compromising visualization, accurate quantification, and MS identification.[34] Strong detergents such as SDS, capable of solubilizing complex multipass IMPs, are incompatible with the low-ionic-strength requirements of IEF. Proteins coated with ionic detergents have drastically altered pIs, resulting in poor resolution (or complete loss of protein due to migration to the electrode). Additionally, ionic compounds in solution will act as mobile ion exchangers, slowing and/or skewing protein migration.[88] As such, solubilizing agents are limited to compounds that are nonionic or zwitterionic over the entire pH range, which are not nearly as effective at solubilizing IMPs. However, some IMPs can be successfully solubilized using urea and thiourea in combination with various detergents (e.g., dodecyl maltoside, Brij56, C13E10, ASB14, and C7BzO).[88,143,144]

However, this does not alleviate the low transfer rates into the second dimension, as demonstrated by protein staining of the IEF strip following transfer.[34] These inherent difficulties are now widely recognized by the proteomics community. And, while there are occasional reports[35,145] of obtaining high percentages (>30%) of IMPs using traditional 2DE, complex multipass and/or very hydrophobic proteins are nearly always severely underrepresented. With regard to work by Aivaliotis et al.,[145] the prediction algorithm they used (TMpred) may overestimate TMDs in soluble proteins by 20%[33] to 50%,[146] so the actual enrichment they achieved is questionable. In response to these limitations, there are now a variety of alternative gel-based techniques that are now widely used for the identification of IMPs.

4.1.2. Blue Native (BN)/SDS-PAGE

BN-PAGE was originally developed by Schagger et al.[147,148] for the determination of the mass and oligomeric state of mitochondrial membrane protein complexes. As the name implies, the method maintains (in many cases) enzyme activity and native protein−protein interactions due to use of mild reagents and a running pH of 7.5. Proteins are initially solubilized in a minimal amount of mild, nonionic detergent (e.g., digitonin, Triton X-100, dodecylmaltoside), and the anionic dye Coomassie brilliant blue G-250 is added. Due to its hydrophobicity and relatively poor solubility in water, Coomassie binds exposed hydrophobic surfaces of proteins, including all IMPs and many soluble proteins. This net negative charge confers electrophoretic mobility, aqueous solubility, and reduced aggregation, even in the absence of detergents. Unlike SDS, which binds with a highly predictable detergent/protein ratio of 1.4 g/g, binding of Coomassie to proteins is much more variable. As such, a different mechanism—molecular sieving, rather than migration rates—allows determination of molecular weight, up to 10 000 kDa. While the lower acrylamide concentrations of SDS-PAGE allow for essentially unrestricted movement of all proteins (i.e., all but the extremely high-molecular-mass proteins would run out of the gel provided sufficient time), in contrast, it is the pore size of BN gradient gels that determines the end-point of migration for individual proteins/complexes.[148,149] BN-PAGE does not have the protein precipitation problems associated with IEF because proteins are never net neutral: the anionic Coomassie dye remains associated with proteins and, additionally, 6-aminocaproic acid or 6-aminohexanoic acid is substituted for NaCl, allowing for electrophoretic separation with constant ionic strength—as opposed to having the salt migrate ahead of the protein—helping to maintain solubility.[150] Following BN-PAGE, a second dimension SDS-PAGE step can be added to further resolve the individual components of each complex.

A related approach is clear native (CN)-PAGE, which follows the same protocol as BN, but with the Coomassie dye omitted. It is typified by lower resolution and is restricted to proteins with an intrinsic net negative charge at the running pH (typically 7.5), since there is no ionic additive to confer a charge shift.[148,149] However, Wittig et al.[151] recently described a "high-resolution" version of CN-PAGE. As in previous applications of CN- and BN-PAGE, the cathode buffer contains a nonionic detergent (dodecylmaltoside or digitonin); however, the anionic detergent SDC is substituted for Coomassie, maintaining membrane protein solubility, imparting proteins with a net negative charge, and increasing resolution to the level of BN-PAGE.

BN/SDS-PAGE has proven generally successful for the identification of IMPs.[152] Stenberg et al.[153] identified 44 proteins—representing 34 protein complexes—from inner membrane vesicles isolated from *E. coli*. According to TMHMM prediction, 55% of proteins contained at least one TMD. In an analysis of murine intestinal brush border membranes, Babusiak et al.[154] reported 27 predicted IMPs (SOSUI) out of 55 identified proteins (49%) using the BN/SDS-PAGE technique. Importantly, complex multispanning IMPs are amenable to separation.[41,153−155] In terms of overall numbers, BN/SDS-PAGE is able to resolve at least several hundred spots per gel.[156] Reifschneider et al.[41] identified ∼140 proteins from crude mitochondria (19% IMP), and Lasserre et al.[30] were able to assign over 160 proteins representing 124 membrane protein complexes from an *E. coli* membrane preparation.

4.1.3. Benzyldimethyl-n-hexadecylammonium Chloride (16-BAC)/SDS-PAGE

Proteins analyzed by 16-BAC/SDS-PAGE are separated according to molecular mass in a discontinuous acidic gradient (pH 4.0−1.5) using the cationic detergent 16-BAC in the first dimension and standard SDS-PAGE in the second dimension. Even though protein migration is dictated by molecular weight in both dimensions, the two detergents have different intrinsic binding properties, which allows for reasonable separation of proteins of similar molecular

weight.[157] As a solubilizing agent for membrane proteins, 16-BAC is fairly potent—superior to the nonionic/zwitterionic detergents used for IEF—but not quite as good as SDS.[158] Because only ~20% of the gel surface is utilized for separation, resolution is compromised, so only a few hundred proteins can be visualized at one time. Accordingly, protein isoforms with variable modifications will tend to migrate together. Due to this comigration, some additional separation of in-gel digested peptides (μLC) prior to identification by MS is generally necessary.[159]

Using 16-BAC/SDS-PAGE, Zahedi et al.[159] identified 146 proteins from whole mitochondria and 57 proteins from enriched mitochondrial membrane preparations. Applying the SOSUI topology prediction algorithm, 47% and 70% of identified proteins, respectively, were characterized as IMPs. Of the latter data set, more than two-thirds of identified proteins were reported to have pIs greater than pH 8 and thus were difficult to resolve by standard 2DE. Several proteins also exhibited high GRAVY scores ($\geq$0.25), outside the typical range of 2DE. The identification of such protein species selectively in 16-BAC/SDS-PAGE is consistent with efficient transfer into the second dimension, which was confirmed by staining the first-dimension lanes following second dimensional electrophoresis.

Moebius et al.[50] applied 16-BAC/SDS-PAGE to the analysis of the human platelet membrane proteome and were able to identify 233 total proteins using μLC-MS/MS analysis. Of those proteins, ~36% had at least 1 TMD; however, the vast majority (~93%) had only 1−2 TMDs (TMHMM). This bias may not be inherent to the 16-BAC/SDS-PAGE system but rather a factor of in-gel trypsin digest in general. Indeed, a 1D SDS-PAGE analysis yielded a sample of similar composition: 40% of 140 proteins with at least one TMD and only 7% with more than two TMDs. Morciano et al.[49] experienced similar problems. They separated immunoisolated synaptic vesicles by 16-BAC/SDS-PAGE, identifying over 100 proteins by in-gel trypsin digest and MALDI PMF. While some complex IMPs were included in their identifications, others could only be detected by immunoblotting, suggesting inefficient in-gel digestion and/or extraction; however, lack of peptide separation prior to MS analysis could also have been a contributing factor.

4.1.4. SDS/SDS-PAGE

A third alternative form of 2DE is SDS/SDS (dSDS) PAGE, where a low percentage acrylamide gel is used for the first dimension and a high percentage one is used for the second. dSDS separation has been reported to result in the migration of complex TMD proteins to a different region of the gel as compared to hydrophilic species,[160] facilitating spot picking for analysis. This variant of 2DE has not been as widely implemented as BN/SDS or 16-BAC/SDS-PAGE.

Hunzinger et al.[40] report the identification of 41 proteins following dSDS separation (as compared to 71 for BN/SDS and 30 for 16-BAC/SDS using MALDI PMF). Burre et al.[47] applied dSDS to the analysis of the synaptic vesicle proteome, resulting in the identification of 96 proteins after MALDI-MS/MS analysis. Proteins were reported to be less well resolved as compared to separation by 16-BAC/SDS-PAGE, but only 59 proteins were identified by the latter technology. 1DE-SDS-PAGE-μLC-MS/MS, the subject of the next section, outperformed both methods, with 143 identified proteins. Integral membrane protein percentages for the three methods were 25% for dSDS, 19% for 16-BAC/SDS-PAGE, and 29% for 1D-SDS (prediction method not reported).

4.1.5. 1D-SDS-PAGE

For complex proteomes, 1D-SDS-PAGE (1DE) alone is insufficient for adequate separation prior to direct MALDI analysis. For example, Moebius et al. report finding 10 or more proteins per band in a typical 1DE separation, vs 1−4 proteins per spot in 16-BAC/SDS-PAGE separation.[50] However, 1DE can be combined with any number of solution-phase chromatographic techniques. The most common method is RPμLC, which separates analytes based on hydrophobicity (see section 4.2.3 for discussion) and can be directly interfaced with ESI or MALDI MS. A 1DE approach has the advantage of being well-established, easy to use, and highly reproducible, capable of separating proteins with a wide range of molecular masses, pI values, and hydrophobicities, largely thanks to the use of SDS. While μLC has some bias against long, hydrophobic peptides, it is expected that not all peptides generated from digestion will have such characteristics, allowing for protein identification, if not full sequence coverage.

Using the 1DE-μLC-MS/MS platform, many researchers report the identification of ~125−200 proteins, with anywhere from ~20 to 55% IMPs.[31−34,47] However, there are several illustrative examples of the power of this technique for identifying fairly large numbers of proteins, in some cases rivaling shogun approaches. Zhao et al.[161] identified 898 cell-surface biotinylated proteins by 1DE-μLC-MS/MS, with an estimated 53% IMPs (SOSUI). Rahbar and Fenselau were able to identify 540 proteins from colloidal silica-enriched plasma membrane,[60] and Cao et al.[67] identified 428 proteins from rat liver plasma membrane, with ~29% estimated enrichment of IMPs (TMHMM). Schindler et al.[65] applied 1DE-μLC-MS/MS to an isolated rat brain plasma membrane fraction, resulting in the identification of 506 proteins, of which 49% were IMPs and 67% had more than one TMD (DAS-TMfilter, may overestimate as compared to TMHMM).

By far the highest identification rate came from Park et al.,[162] who identified over 1300 proteins (60% integral, anchored, or associated) from analysis of an enriched membrane fraction isolated from human brain tissue.

Thus, in terms of overall numbers, 1DE-μLC-MS/MS definitely ranks as a global proteomics platform. The routine identification of several hundreds of proteins puts the resolving power in line or ahead of the alternative 2DE approaches discussed above (BN/SDS, 16-BAC/SDS, dSDS), with a more straightforward implementation. This is not surprising given that separation in the second dimension is based on peptide rather than protein migration, providing a more orthogonal separation. (However, it should be noted that BN/SDS-PAGE is unique in its ability to analyze native protein complexes.) Moderate enrichment of IMPs (~55%) can be attained using 1DE, though hydrophobic domains may be underrepresented in MS analysis (see next section). Like the other gel-based approaches, analysis of gel bands is a time-consuming process, though potentially highly automated with the appropriate robotic equipment.

4.1.6. Analytical Considerations for Gel-Based Methods

4.1.6.1. In-Gel Digest. Several authors report underrepresentation of hydrophobic TMD peptides following gel-based separation and tryptic peptide extraction,[48,50,153,155,160,163]

suggesting an inherent bias against (1) generation of analyzable TM peptides via in-gel trypsinization, due to the paucity of Lys/Arg residues in TMDs and/or structural inaccessibility, as hydrophobic domains may reaggregate after SDS is removed, or (2) the physical extraction of long/hydrophobic peptides from the gel. Application of different digestion strategies, such as CNBr/TFA,[36,105,106] can help to generate shorter, less hydrophobic peptides, achieving higher representation of complex IMPs. Extraction may also be facilitated by inclusion of small amounts of mild detergents (OG, CYMAL-5, SDC); see sections 3.2.2 and 3.2.3 for discussion. However, these techniques have not yet been widely implemented.

4.1.6.2. Peptide Separation before MS Analysis. An important factor limiting overall protein identifications by gel-based methods is whether or not peptides are further separated prior to MS analysis. Given that most gel spots will contain multiple proteins and that IMPs are generally low-abundant proteins, always adding a separation step (e.g., μLC) prior to MS analysis should aid significantly in the identification of species that may otherwise remain obscured. For example, when Rahbar and Fenselau[57] analyzed colloidal silica-isolated PM by 1DE coupled with either nanoESI-MS/MS (direct infusion) or μLC-MS/MS, protein identifications increased from ~45 to 366 upon inclusion of the μLC separation step.

4.2. Gel-free Methods for Protein Separation

4.2.1. Solution-Phase IEF

Solution-phase IEF has yet to be extensively applied to the separation of membrane proteins, so its general utility is somewhat unclear; however, it has the potential to circumvent some of the precipitation and transfer problems associated with gel-based IEF. Solution-phase IEF can be carried out using several platforms. Multichambered devices are composed of a series of sample chambers separated by liquid-permeable membranes.[164,165] The pH gradient can be established by carrier ampholites either in solution or immobilized on the membranes themselves[166] in the presence of an electric field. Protein precipitation on membranes may still be an issue when protein pI coincides with membrane pH.[167] Alternatively, free-flow IEF[164,165,168] uses a single-chamber apparatus and relies on the continuous transport of a proteomic sample in a pH-graded solution established by carrier ampholites flowing as a thin, laminar film. An induced electric field perpendicular to the direction of flow acts as a deflecting force, allowing precise resolution and collection of proteins based on their pI. Precipitation is theoretically minimized as compared to the case of gel-based IEF, as proteins spend very little time at their pI, and detergents can be added to the counterflow buffer. After solution-phase IEF, proteins within each chamber/fraction can be further separated by gel-based or gel-free techniques.

Weber et al.[169] separated enriched (nitrogen cavitation, sodium carbonate treated) peroxisomal membrane proteins by free-flow IEF. After subsequent 1DE separation and MALDI-PMF analysis, 35 proteins were identified, of which at least 15 had at least one predicted TMD (consensus of five different algorithms).

Using a multichambered Rotofor apparatus, Peirce et al.[170] reported the identification of 127 proteins by solution-phase IEF followed by IDE-μLC-MS/MS analysis of avidin−biotin affinity purified lymphocyte plasma membrane proteins. Of those proteins, many were not detected by conventional 2DE, and 74 (58%) were reported to contain at least one TMD (prediction algorithm not reported). Pedersen et al.[35] fractionated yeast membranes (enriched via sodium carbonate treatment) using a multicompartment electrolyzer equipped with isoelectric membranes. The basic proteins (pH 7−10.5) were then further separated by conventional 2DE. In-gel digestion and MALDI-PMF analysis gave 323 protein identifications, of which 105 (33%) contained a predicted TMD (via MIPS yeast database, HMMTOP, and/or literature searches). Moreover, 50% of identified peptides belonging to IMPs overlapped with a TMD. Such good identification of IMPs, complete with TMD coverage, is exceedingly rare in the 2DE literature, highlighting the importance of optimized enrichment and fractionation strategies for identification of low-abundance species.

4.2.2. Anion and Cation Exchange

Anion and strong cation exchange separation techniques can be used for either proteins or peptides, and they are compatible with the inclusion of mild detergents. Anion exchange separates analytes as negatively charged species at high pH via interaction with a positively charged (e.g., ammonium) stationary phase and competing negative counterions (e.g., Cl^-, $HCOO^-$). Conversely, strong cation exchange (SCX) is based on the separation of positively charged analytes at low pH based on their interaction with a negatively charged (e.g., phosphonic/sulfonic acid) stationary phase and competing positively charged ions (e.g., Na^+, K^+, NH_4^+).[171]

From a membrane-enriched fraction of the bacterium *C. glutamicum*, Schluesener et al.[36] separated proteins solubilized in the zwitterionic detergent ASB-14 via anion exchange prior to 1DE. Analysis by MALDI-PMF and μLC-MS/MS gave 170 protein identifications, 29% of which were membrane-integral (TMHMM). Proteins exhibited wide pI (3.7−10.6) and molecular mass ranges (10−120 kDa), and the technique was preferred to 16BAC/SDS-PAGE.[172] In another example of anion exchange prior to 1DE, Schmitt et al.[68] identified 92 proteins by MALDI-PMF and MS/MS from an *N. crassa* mitochondrial outer membrane preparation, including a few TMD-containing IMPs.

Delom et al.[173] applied anion exchange, 1D-lithium dodecyl sulfate-PAGE, and MALDI-PMF to the *Saccharomyces cerevisiae* plasma membrane proteome, identifying 86 proteins with 53% IMPs according to TMAP (which may overpredict as compared to TMHMM[146]). Using cation exchange chromatography, Bagshaw et al.[174] separated a lysosomal membrane fraction using an 8 M urea buffer containing FA and an alkyl glucoside detergent. 1DE-μLC-MS/MS was then carried out on the resulting fractions, giving 215 protein identifications, including a number of IMPs.

Unfortunately, no assessment of protein recovery from ion exchange columns was reported, so the relative success of these analyses is difficult to gauge. However, adding such a fractionation step prior to 1D-SDS significantly increases the number of gel slices and, thus, the analysis time; however, this effort has so far resulted in an overall protein identification rate/IMP enrichment that is similar to that obtained by 1DE-μLC-MS/MS.

4.2.3. Reversed-Phase Liquid Chromatography (RPLC)

Reversed-phase resin consists of silica modified with long straight-chain alkyl silanes, which preferentially bind analytes

based on hydrophobicity, such that the most hydrophobic molecules are the most highly retained. As the mobile phase ramps from a highly polar solvent (water) to a less polar solvent (e.g., acetonitrile or methanol), interaction between the analyte and the mobile phase becomes more favorable, allowing transfer from the stationary phase to the mobile phase and subsequent elution. The major difference between RPLC for peptides and proteins is that resins with larger pore sizes are used for proteins. Solubility and elution can also be more of a problem for IMPs, as they have a potentially much larger hydrophobic surface area than the corresponding peptides. RPLC is not amenable to the inclusion of detergents; however, the organic component of the solution phase can aid in solubility to some degree.

McDonald et al.[175] used a combination of chromatofocusing and RPLC (proteins solubilized in 20% acetonitrile, 2% TFA) at 50 °C to separate inner mitochondrial proteins, analyzing the tryptic digest of each fraction by μLC-MS/MS. They identified over 200 proteins, including some IMPs and proteins with positive GRAVY scores. In comparison, fewer than 100 proteins were identified by standard 2DE.

Martosella et al.[138] found that protein solubilization in 80% FA and separation using a macroporous RP C18 column at 80 °C was optimal for protein separation and, importantly, recovery. A cursory survey of proteins further separated by 1DE and analyzed by μLC-MS/MS gave around 22% IMPs (of 158 proteins; as assessed by GOMiner). The authors note that the elevated temperature is crucial to protein elution from the column. This study represents one of the more promising methods for the LC separation of IMPs.

4.3. Gel-free Peptide Separation: Shotgun Proteomics

Shotgun proteomics[17] is the separation of complex proteomic digests followed by either ESI[22] or MALDI[20] mass spectrometric analysis. A variety of techniques for peptide separation can be used either alone or in combination (e.g., size exclusion, anion exchange, hydrophobic or hydrophilic interaction chromatography, monolithic columns, capillary electrophoresis, capillary isoelectric focusing, and free-flow electrophoresis; for reviews, see refs 171 and 176−179). The most widely used shotgun technique—MudPIT—is a combination of SCX and RP separation.

One advantage of the shotgun approach is the greater solubility of peptides as compared to proteins in aqueous buffers. A gel-free platform also has the added advantage of eliminating problems associated with successful digestion and extraction of hydrophobic peptides from gel slices. However, because proteins are digested en mass prior to separation, the complexity problem becomes significantly magnified. For example, a yeast proteome with 6300 proteins gives 340 000 theoretical tryptic peptides.[140] In addition, validation of protein identifications using biophysical parameters, such as pI and molecular weight, cannot be carried out. However, excellent methods exist to estimate false discovery rates (e.g., as calculated from matches to a randomized protein database.[180,181] Unlike 2DE, shotgun methods can be readily tailored to the analysis if IMPs and, importantly, TMDs (see sections 4.3.4 and 5.2.1).

Most shotgun studies interface separation with ESI rather than MALDI MS due to the extreme ease with which ESI can be combined on-line with μLC separation. However, new techniques have allowed for the off-line deposition of μLC-separated proteins onto MALDI plates, and the LC-MALDI method has been applied in several cases.[182−185]

4.3.1. 1D vs 2D

In general, 2D (or 3D) μLC separations are generally considered superior to a 1D approach for complex samples. Typically, a single RP separation is insufficient for in-depth analysis, maxing out at ~750 proteins on a midrange (fast scanning, low sensitivity) instrument (10 μg tryptic protein digest, 12 cm RP, 3 h gradient, Thermo-Fisher LTQ), which is far below the number expected to make up complex proteomes. Because ions are chosen for fragmentation in a semirandom fashion based largely on abundance, low-abundance proteins can easily be overlooked in a complex sample.[186] Besides the enrichment strategies discussed throughout this review, there are other measures that can be taken to maximize identifications from a single RPLC run. For example, Smith and colleagues utilize long (65 cm) μLC columns run at ultrahigh pressure[187,188] and acquire MS/MS spectra in narrow *m*/*z* ranges (gas-phase fractionation).[133,187−189] In one study of *P. aeruginosa*, 623 proteins (44% IMP by PSORT) were identified using an LCQ Deca ion trap, a fairly impressive number considering the LCQ Deca is a relatively slow scanning instrument of lower sensitivity.[187]

4.3.2. Microcapillary IEF (μCIEF)/μLC-MS/MS

One potentially promising alternative to first dimension SCX separation is solution-phase IEF. In addition to its application for separation of intact proteins, IEF can be applied to peptides by performing the separation in a fused silica microcapillary tube (~60−80 cm, 100 μm i.d., 360 μm o.d.). Separation by μCIEF (also referred to as CIEF) has been used as the first dimension separation in a shotgun technique, whereby fractions of focused peptides are further separated by μLC for analysis by ESI MS/MS. For μCIEF, a digested protein sample is combined with carrier ampholites and introduced to a hydroxypropyl-coated microcapillary. The ends of the capillary are in contact with dilute solutions of ammonium hydroxide (catholyte) and acetic acid (anolyte). Focusing is initiated by the application of an electric field over the entire microcapillary and takes ~30 min. Like the cases of gel and solution-based methods, the pI range can be varied from ~3 to 12. Eluate from the column is loaded onto a series of RP trap columns through an injection/switching valve interface. The trap columns serve a dual purpose: holding peptides awaiting separation and providing a mechanism for ampholyte removal. From there, sequential fractions are eluted onto a μLC column for separation and analysis by tandem MS.[190] The two separation methods are highly orthogonal, with peptides segregated first by pI and then by hydrophobicity. The percentage of peptides found in more than one fraction (10−25%) is reported to be significantly less than that obtained by MudPIT separation.[191]

Wang et al.[192] used μCIEF/μLC to analyze microdissected ovarian tumor tissue. After removal of soluble proteins from the sample, the pellet was solubilized in 1% SDS and then dialyzed against 100 mM Tris buffer to remove excess detergent prior to trypsinization. Following μCIEF, tryptic peptides were separated into 18 fractions, each subject to μLC-MS/MS analysis. Their study resulted in the identification of 3303 proteins, 773 (23%) of which were predicted to contain at least one TMD by TMHMM, making this the largest number of total membrane proteins identified from a

single microdissected tumor sample. The authors also confirm good separation using the μCIEF technique, as 75% of peptides were identified from only one fraction. However, the 23% IMP enrichment is quite low (28% predicted IMPs in the database), and no TMD-containing peptides were identified.

The yeast proteome was subject to a similar analysis.[193] Cells were lysed and soluble proteins removed following centrifugation. The pellet was washed with 8 M urea and solubilized in 1% SDS. After dialysis to remove SDS and digestion with trypsin, peptides were subject to separation and analysis as described above, except μCIEF was conducted in the presence of 4 M urea to help maintain peptide solubility. A total of 2513 proteins were identified, with 634 (25%) predicted to contain at least one TMD and 407 (16%) predicted to contain two or more TMDs by TMHMM. In comparison, the entire yeast proteome (6598 ORFs) is predicted to contain 884 (13%) IMPs with two or more TMDs.

In both of theses studies, and in a recent analysis of formalin vs fresh frozen tissue,[194] IMPs were not significantly enriched beyond that expected for the unfractionated cell/tissue proteome as a whole. These statistics could likely be improved through the application of more stringent enrichment conditions, such as high-salt or high-pH washes. Additionally, no TMD-containing peptides were identified, which is not too surprising given that no detergent was added for trypsin digestion, which is unlikely to generate peptides from hydrophobic domains in sufficient quantity or of amenable size for analysis. There is also the possibility that TMD peptides did not make it through the separation. Given the promise of this technique for the orthogonal separation of large numbers of peptides, easily comparable to that of MudPIT (see below), it would be interesting to see μCIEF/μLC-MS/MS applied to a more stringently prepared membrane protein sample.

4.3.3. MudPIT

The most widely used shotgun proteomic platform is MudPIT,[18,19] which involves peptide separation by SCX—either on-line or off-line—followed by RPμLC. For on-line MudPIT (referred to simply as MudPIT), both the SCX and RP resins are packed in-tandem in a fused silica microcapillary column (360 μm o.d., 25−150 μm i.d.; ∼5−10 μm tip opening) and eluted directly into an ESI tandem mass spectrometer.

Many of the proteomic studies discussed throughout this review use on-line and off-line MudPIT approaches; indeed, in the following sections discussing specific digestion strategies (section 5), affinity purification (section 6), and quantitation strategies (section 7), this shotgun approach is used for nearly all sample analysis. Protein identifications in excess of 1000 are typically reported, with IMP enrichments anywhere from ∼20 to 65%.[69,70,75,195−198] A few more examples, not discussed elsewere, are presented.

Schirmer et al.[70] used MudPIT to analyze the membrane proteome of the rat nuclear envelope, generating two enriched fractions following salt/detergent or high-pH washes. Over 2000 unique proteins were identified, including a large number of IMPs, 67 of which were previously unknown. The group used a subtractive approach to identify proteins unique to the nuclear envelope by excluding identifications (>300) also found in a microsomal membrane preparation.

Wu et al.[45] characterized a stacked Golgi fraction using MudPIT, identifying over 421 proteins, including 110 bona fide Golgi proteins, of which 70 (64%) were predicted IMPs (HMMTOP2). Modification searching identified numerous proteins with Arg dimethylation.

In order to maximize proteome coverage of the yeast mitochondrial proteome, Reinders et al.[42] applied off-line MudPIT, 1DE-μLC-MS/MS, and standard 2DE-MALDI PMF to the yeast mitochondrial proteome, giving 491 proteins (26% IMP) for off-line MudPIT, 630 (23% IMP) for 1DE-μLC-MS/MS, and 169 (5.9% IMP) for classic 2DE (assigned from the *Saccharomyces* Genome Database). The numbers for MudPIT and 1DE were both fairly good, and they included many low-abundant, extreme pI, and high molecular weight proteins, as well as hydrophobic peptides. They were also considerably more comprehensive than 2DE, which resulted in only five unique protein identifications. Notably, MudPIT and 1DE appeared largely orthogonal, overlapping by approximately one-third (three replicate runs per condition). Given the large number of identified proteins, this is suggestive of some true degree of orthogonality between the two methods rather than just undersaturation of sampling.

An important parameter to consider for shogun proteomics, or any type of global analysis, is saturation of peptide sampling. It has been estimated that at least nine MudPIT runs are required to achieve >95% saturation, in agreement with studies by Durr et al., who note that 7−10 runs were needed for significant analytical completeness.[62] While the above study by Reinders et al. only used three replicates, it is encouraging to find *any* implementation of replicate sampling in the literature. One common pitfall resulting from undersampling is the overinterpretation of comparative experiments. For example, McCarthy et al.[199] applied differential detergent fractionation (digitonin, Triton X-100, deoxycholate/Tween-40, SDS) to the separation of the chicken B-cell proteome for off-line MudPIT analysis. The authors reported significant orthogonality (location, class) between the different types of membrane proteins isolated, which may indeed be the case, but it is difficult to reach such a conclusion with any measure of certainty when only ∼20 IMPs were identified in each fraction, representing an extremely small percentage of the total membrane protein content.

Like most separation techniques discussed, MudPIT requires skilled implementation and proper equipment in order to obtain optimal results. For example, Cutillas et al.[51] report that they prefer 1DE-μLC-MS/MS to off-line MudPIT because, in their hands, MudPIT was not as robust a technique and they did not have adequate software for handling large datasets.

Overall, MudPIT experiments are capable of routinely identifying thousands of proteins, representing 2−5-fold more capacity than 1DE-μLC-MS/MS. However, the typical IMP enrichment for most experiments, 50−65% at most, is still only comparable to the 1DE approach. And any outliers, like the 81% estimated enrichment (via TMpred) reported by Da Cruz et al.,[43] are likely overestimated by 20−50%.[33,146] It might seem as though there is a ceiling limiting the degree to which IMPs can be enriched; however, if combined with optimized preparation and analysis protocols, shotgun proteomics is indeed capable of far better results, as discussed below and in section 5.2.1.

4.3.4. μLC at Elevated Temperature

A common finding in proteomic studies of IMPs is that TMD-containing peptides are severely underrepresented in

Table 4. Summary of Solubilization/Digestion (Section 5), Affinity Purification (Section 6), and Quantitation (section 7) Strategies

solubilization/digestion techniques	comments
soluble domains (membrane shaving)	
hppK	nonspecific protease, not reliant on specific cleavage sites in soluble domains, high-pH opens vesicles allowing access to both sides of membrane
Lys-C	requires Lys in soluble domains
trypsin	requires Lys or Arg in soluble domains
embedded and soluble domains	
hppK-CNBr/FA	combines benefits of nonspecific protease shaving with targeted TMD solubilization/analysis, gives ~97% IMPs and ~68% TMD peptides
60%MeOH/trypsin	efficient solubilization strategy for targeting both soluble and TM domains, gives ~45% IMPs
trypsin/hppK-60%MeOH/trypsin	combines benefits of nonspecific protease shaving with good solubilization of TMD regions, gives ~40% IMPs
trypsin-60%MeOH/trypsin/chymotrypsin	some enrichment afforded by shaving, good solubilization and use of orthogonal enzymes for TMD regions, gives 20−50% IMPs, targets hydrophobic peptides

affinity purification	comments
glycosylation	good for identifying sites of glycosylation, not yet effective as global IMP enrichment strategy, IDs in low 100's, ~65% IMPs
biotinylation	widely used, established strategy for cell-surface IMPs, some topological assessment also possible, higher numbers (mid−high 100's) than glycosylation, ~60% IMPs
protein sequence tag	limited examples, a de-enrichment strategy, IDs in low 100's, 20−70% IMPs

quantitation	comments
ICAT	Cys modification, compatible with SDS and 60% methanol
HysTag	Cys modification, limited examples, limited analysis of IMP recovery
ICPL	Lys modification, no affinity tag, limited examples
^{18}O	modification of peptide N-terminus, not sequence limited, compatible with 60% methanol, more IDs as compared to ICAT

shotgun MS/MS analysis (see, e.g., refs 10 and 119). As will be discussed in section 8, hydrophobic peptides are expected to ionize better and fragment at least as well as hydrophilic species, so other factors, namely sample loss and masking by more abundant species, are the likely culprits. With regard to sample loss, a recent study[84] suggests that conducting μLC separation at elevated temperature improves elution of hydrophobic peptides, a phenomenon that has been previously demonstrated for intact lipid raft proteins.[138] To assess the effects of temperature on hydrophobic peptide recovery, a sample enriched in TMD-containing peptides was prepared according to a recently optimized high-pH−proteinase K (hppK)−CNBr/FA protocol (see section 5.2.1). Membrane-embedded peptides, resulting from proteinase K shaving of enriched HeLa cell membranes, were further digested with CNBr in 90% FA. When this TMD-enriched sample was analyzed by μLC-MS/MS, it was found that increasing the separation temperature from room temperature to 60 °C resulted in a 4-fold increase in protein identifications (28 vs 105) and a 5-fold increase in unique peptide identifications (69 vs 327). Moreover, with elevated temperature, the observed peptide population was more hydrophobic, with an average GRAVY score of 0.63 vs 0.41 for room temperature. Overall, the IMP enrichment was estimated at 98%, with TMD-containing peptides making up 63% to 68% of the sample at room temperature and 60 °C, respectively (TMHMM). Peptides with high TMD overlap (≥75% of peptide sequence) in particular showed a significant increase from 11% at 20 °C to 27% at 60 °C.

Importantly, the high-temperature runs seem to be both necessary and sufficient to capture the majority of proteomic information, as the protein and peptide populations observed at elevated temperature were largely inclusive of those found at room temperature. While longer and/or more hydrophobic peptides were observed, as expected, to elute slightly later in the gradient, high organic was not required, as nearly all peptides eluted by ~45% organic, well within the window for optimized chromatography. Thus, the ability of elevated temperature to induce peptide elution appears to be a unique phenomenon, not replicated by simply using a higher percent organic in the mobile phase.[84]

High temperature was also beneficial for the μLC-MS/MS analysis of a high-pH-enriched plasma membrane trypsin digest. Long, TMD-length (>20 residues) peptides with substantial (>75%) TMD overlap were observed almost exclusively at high temperature (3 vs 25 peptides at 20 and 60 °C, respectively). However, as those numbers suggest, TMD-containing peptides made up an extremely low 1−2% of the overall peptide population, owing to the less efficient trypsin digestion and enrichment strategy as compared to the hppK-CNBr/FA protocol.

Due to these dramatic increases in hydrophobic protein and peptide identifications, re-integration of the column heater functionality with the μLC platform should prove highly beneficial for the comprehensive analysis of IMPs and TMDs, especially when combined with an optimal enrichment strategy[73] for targeting TMDs.

5. Solubilization and Digestion Techniques for Shotgun Proteomics

A summary of the solubilization and digestion techniques is given in Table 4 and Figure 1.

5.1. Targeting Soluble Domains: Membrane Shaving

5.1.1. The High-pH—Proteinase K (hppK) Method

The problem analyzing IMPs can be distilled down to an inability to achieve and maintain dissolution in aqueous

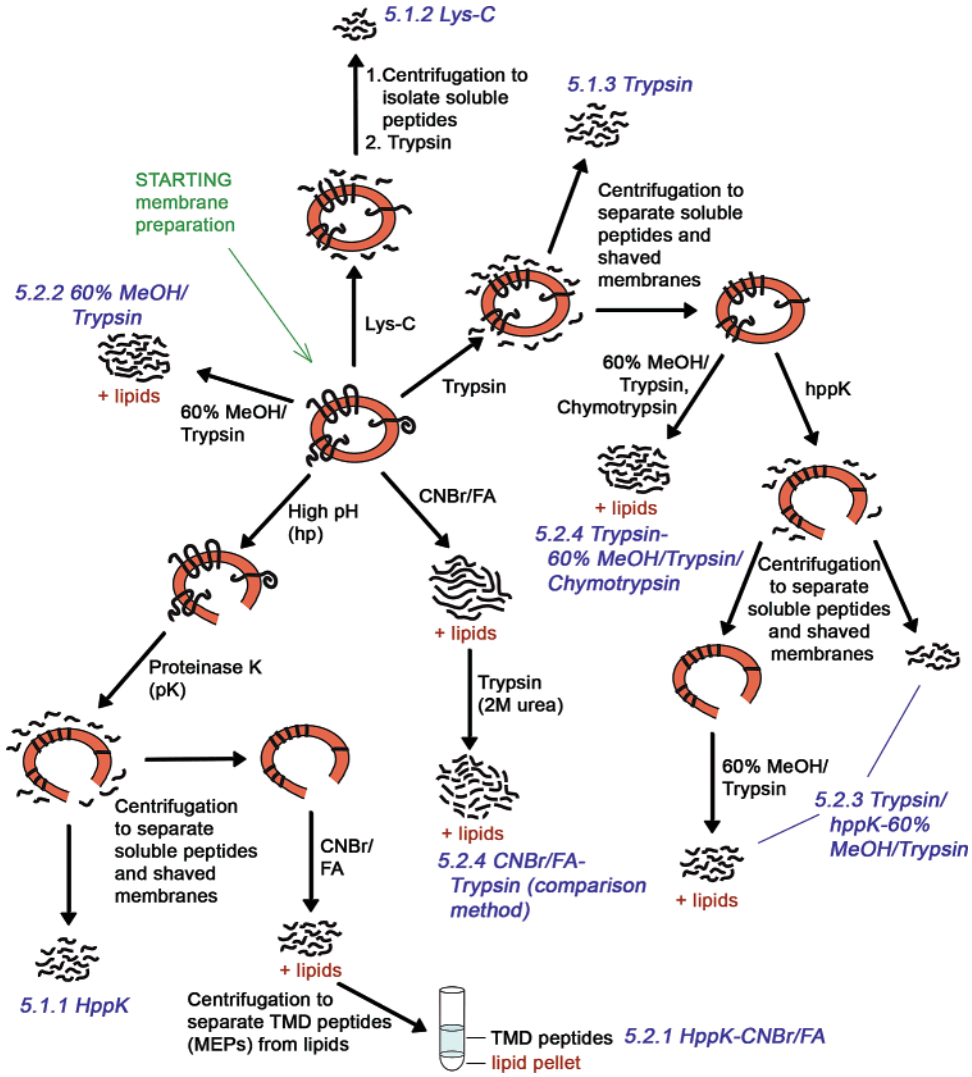

Figure 1. Summary of solubilization/digestion techniques discussed in section 5 for generation of complex protein digests for analysis by MudPIT. The starting material (center, indicated by a green arrow) is an intact cell or membrane vesicle preparation, ideally enriched by high-salt and/or high-pH washing to remove membrane-associated proteins and soluble proteins trapped inside vesicles. For each digest generated for analysis, the section/method name is provided (blue italics). MeOH = methanol.

media without the aid of lipids or lipid-mimetic detergents that interfere with separation, digestion, or MS analysis. One possibility is to simply ignore the problematic hydrophobic regions and restrict analysis to the hydrophilic extramembrane domains. Several years ago, Wu et al.[196] reported a method whereby membranes are first induced to form open vesicles upon mechanical agitation in high-pH buffer. However, the bilayer itself, along with the embedded proteins, remains intact. Upon treatment with the nonspecific protease, proteinase K, all exposed soluble domains and soluble proteins are digested. A nonspecific enzyme is used so that digestion efficiency is not reliant upon the presence of specific cleavage sites in all extramembrane loops. The proteolyzed soluble domains can then be analyzed by MudPIT. The original implementation of the hppK method resulted in the identification of some 1600 proteins from a mouse brain homogenate, of which 28% were IMPs (TMHMM), with over half having at least two TMDs. This percentage is fairly representative of IMPs in the sample composition, as no effort was made to exclude soluble proteins. The hppK method was also applied to analysis of the Golgi proteome and, in combination with a separate CNBr/FA digest, identified over 100 Golgi proteins, the majority of which (68%) were IMPs.[45]

The hppK method, or any of the on-membrane digestion schemes discussed below, has the potential to be used for the assessment of membrane protein topology (see also discussion in ref 200). In the case of hppK, exposed external IMP soluble domains (of PM or other organelles) can be completely digested by proteinase K at neutral pH, where the protease has maximal activity, into very short di/tripeptides. The hppK method is then carried out as described above, except only the soluble domains formerly protected on the interior of the membrane vesicles remain for digestion, giving the relative sidedness of IMPs.[196] One caveat to this approach is that topology analysis can be complicated, to some degree, by vesicles ruptured during cell lysis re-forming with an inside-out conformation and/or contamination between PM and intracellular membranes.

One downside often mentioned regarding proteinase K is that it is a nonspecific protease and, thus, prone to generate much larger numbers of peptides than proteases with specific cleavage sites.[75] However, proteinase K does appear to have certain preferred sites that may be strongly influenced by

local structural motifs, limiting the number of possible peptides (unpublished results). As such, for specialized applications such as membrane shaving, it can be used without significant problems.

5.1.2. Lys-C

Membrane shaving can also be accomplished using proteases with specific cleavage sites. Nielsen et al.[195] digested intact, extensively enriched mouse hippocampal plasma membranes with endoproteinase Lys-C. The resulting peptides were isolated by centrifugation and separated by RPLC and then were further digested with trypsin. MS/MS analysis resulted in 1600 protein identifications. Some 60% of proteins were annotated as membrane or membrane-associated, but the total percent of IMPs was not reported. In a more recent study by Wisniewski and colleagues,[197] salt and sodium carbonate washed mouse fore- and hindbrain membranes were digested with Lys-C, and soluble peptides were isolated and digested with trypsin. Samples were analyzed by off-line MudPIT, identifying 1213 proteins, including 459 (38%) predicted IMPs (TMHMM). One might expect that a more general protease than Lys-C would produce more peptides from IMP soluble domains; however, their choice of Lys-C was likely dictated by its potential compatibility with their HysTag affinity/quantification reagent,[201] which has an internal Arg cleavage site and, thus, would not be compatible with trypsin or any nonspecific protease.

5.1.3. Trypsin

Rodriguez-Ortega et al.[202] analyzed trypsin- or proteinase K-shaved bacterial surface proteins from intact group A *Streptococcus*. Using a combination of MudPIT-MALDI-MS/MS and μLC-MS/MS for analysis of the isolated soluble peptides, a total of 72 proteins were identified, 51% of which contained at least one TMD according to PSORT. The authors note an almost complete lack of cytoplasmic proteins (6%) observed in the protease-susceptible digest, indicating that the method is highly specific for surface-exposed proteins. Moreover, the group correlated the experimentally derived topology with that predicted by PSORT, and they found good correlation for 26/37 (70%) of IMPs. For the remaining 11 proteins, experimental evidence was strong enough to call into question the PSORT assignment for at least six topologies. Thus, data derived from experiments that allow assessment of IMP topology can be used for enhancing structural information and refining topology prediction algorithms (see also section 9.3). Importantly, bacteria surface shaving gives useful information regarding the accessibility of potential antigens for vaccine development, as readily cleaved proteins should also be antibody accessible. Membrane shaving with trypsin has also proven useful for assessing the membrane-embedded nature of potential mitochondrial IMPs.[44,68,203]

5.2. Targeting Membrane Embedded and Soluble Domains

For any proteomic analysis, it is desirable to maximize sequence coverage; however, many digestion strategies, as discussed in the preceding section, are targeted toward hydrophilic domains. Because IMPs often have a significant portion of their sequence embedded in the membrane, even the probability of *identifying* the protein may decrease using methods restricted to soluble domains. Additionally, exposed loops may be short, especially for IMPs with numerous TMDs,[11] or heavily modified (e.g., glycosylated), further complicating identification unless steps are taken to remove (e.g., PNGaseF, β-elimination) and/or account for modifications during analysis and sequencing.[204] As such, targeting TMD domains should be a particularly beneficial strategy. However, owing to the length (17−40+ residues)[205] and extreme hydrophobicity (due in large part to membrane occlusion and a lack of tryptic cleavage sites[206]), they can be difficult to extract from gels (see section 4.1.6) and elute from RP columns,[84] and they may be out of the *m*/*z* range of typical MS/MS instruments (see section 8.2). According to a study by Eichacker et al.,[10] the analysis of TMDs can be facilitated by a reduction of hydrophobicity and/or length, most easily accomplished by alternate solubilization and cleavage strategies.

5.2.1. HppK−CNBr/formic acid

As discussed in section 5.1, one approach for studying IMPs is to shave off the soluble domains from intact membranes using strategies such as hppK digestion.[196] The soluble (protease-accessible) peptide fraction is then isolated by centrifugation for subsequent shotgun analysis; however, analysis of the membrane-embedded peptide (MEP) fraction is less straightforward. In addition to the difficult biophysical properties of TMDs mentioned above, due to their high lipid content, MEP samples tend to be highly viscous, leading to loading and elution problems during μLC separation. A recently reported protocol, employing CNBr digestion in 90% FA and an optimized method for lipid removal, resolved both of these issues.[73] According to the hppK protocol,[196] a sodium-carbonate-washed membrane fraction is digested with proteinase K in high-pH buffer, removing accessible IMP domains and any remaining peripherally associated proteins. The shaved membranes are reisolated by centrifugation and then solubilized with 90% FA for CNBr digestion.[73] Unlike the charged residues Arg and Lys, which occur relatively infrequently in TMDs,[206] Met is predicted to occur in nearly half of all eukaryotic TMDs, making CNBr digest a useful strategy for reducing TMD peptide length.[10] Additionally, from a steric hindrance standpoint, chemical digestion methods do not pose the same accessibility problem as enzymatic digestion in cases where cleavage sites may be occluded by lipids or protein folds. To remove lipids prior to analysis, the CNBr digested sample is suspended in an aqueous−organic buffer, which fortuitously allows for the precipitation of lipids upon centrifugation. Analysis of the MEP samples by μLC-MS/MS at elevated temperature (see section 4.3.4) resulted in identification of 285 proteins, with 97% predicted IMPs and 83% with more than two TMDs (TMHMM). Using a MudPIT approach (also at elevated temperature) significantly increased the number of protein identifications to 639, with enrichment dropping slightly (81% IMP, 56% with more than two TMD), as the improved chromatographic resolution allowed more low-abundant species (in this case the depleted soluble proteins) to be identified.

Importantly, the overall TMD/protein distribution was very similar to that predicted by *in-silico* genome-wide analyses,[207] suggesting that the hppK−CNBr/FA method is not biased against any particular protein class. The hppK-CNBr/FA method did indeed result in good coverage of membrane-spanning regions, with an impressive 68% of all identified peptides overlapping with a predicted TMD for the 1D

analysis and 49% for the 2D, by far the highest report to date. Importantly, the increased coverage resulted in a substantial increase in protein identifications as well; 55% of all proteins identified in the MEP sample were not found during analysis of the soluble domains.[73] Thus, restricting analysis to the soluble domains is a significant handicap regarding comprehensive IMP identification.

5.2.2. 60% Methanol/Trypsin

In a different solubilization approach, Blonder et al. took advantage of the membrane-disruptive potential of methanol (see section 3.3) for the specific solubilization/digestion of IMPs. Their detergent-free sample preparation method involves the solubilization of high-pH-enriched membrane fractions in a 60% methanol buffer, trypsin digestion, and analysis by shotgun proteomics. While trypsin activity is attenuated in the presence of 60% methanol (to ~20% as compared to aqueous conditions), it is sufficient for comprehensive digestion.[208] The method was originally applied to *D. radiodurans*[188] and resulted in the identification of 215 inner membrane IMPs, corresponding to 21% of the predicted population (PSORT). In contrast, a control experiment in which methanol was omitted identified only 2% of predicted IMPs. The method does not appear to be biased against any particular type of IMP; among the IMPs identified using 60% methanol solubilization are numerous complex, highly hydrophobic species.[188] The methanol/trypsin method has been applied to a variety of samples.[69,132−134,187,189,208−210] One study of lipid rafts identified 358 proteins with 44% IMPs, of which one-third contained more than two TMDs (TMHMM).[189] An inventory of the human epidermis membrane proteome identified 1306 proteins, of which ~46% were IMPs, as extrapolated from TMHMM analysis of 223 identified hypothetical proteins.[69]

Zhang et al.[95] compared solubilization/trypsinization of an *E. coli* membrane proteome in 60% methanol vs 1% SDS (10× dilution before digestion) followed by off-line MudPIT (where SCX serves a dual function—peptide fractionation and removal of SDS not tightly bound to peptides). The authors report that more proteins were detected by 60% methanol solubilization as compared to SDS (358 vs 299) and the identified proteins and peptides were overall significantly more hydrophobic, indicating enhanced digestion of IMPs/TMDs in addition to peptide detection.

5.2.3. Trypsin/HppK—60% Methanol/Trypsin

Wei et al.[211] compared two approaches to the on-membrane digestion of *Shigella flexneri* 2a membrane proteins. In an initial experiment, a purified membrane pellet was digested according to the hppK protocol and analyzed by MudPIT, giving 331 protein identifications, which included (according to supplemental material) ~125 (38%) IMPs (32% IMPs with more than two TMDs; TMHMM). In an effort to increase the numbers of IMPs, the authors modified their protocol to remove interfering membrane-associated proteins and added additional digestion steps. As such, the isolated membranes were first washed with high-salt buffer and high-pH buffer containing DTT. Thereafter, membranes were isolated and resuspended at neutral pH, and exposed domains were digested with trypsin. Membranes were reisolated and subject to hppK, and the remaining membrane-embedded fragments were finally solubilized in 60% methanol for trypsin digest. MudPIT analysis of all isolated peptides gave 528 protein identifications with ~223 (42%) IMPs (TMHMM). While the overall percent of IMPs showed modest improvement, the overall number of IMPs increased by almost two-thirds and, importantly, more complex proteins (>2 TMD) were identified with a much higher frequency (~50% vs ~32%). This improvement is likely do to a concerted effort to enrich for IMPs by combined washing and on-membrane digestion and to generate as many usable peptides as possible from orthogonal digestion conditions.

5.2.4. Trypsin—60% Methanol/Trypsin/Chymotrypsin

Fischer et al.[75] compared two enrichment/digestion methods for the MudPIT analysis of the membrane proteome of Gram-positive *C. glutamicum*. In the first protocol, the membrane preparation was predigested with trypsin (in a nondenaturing buffer) to remove soluble proteins/extracellular domains. Shaved membranes were then isolated, solubilized (or, perhaps more appropriately, destabilized) in 60% methanol, and then redigested using trypsin followed by chymotrypsin, giving 274 IMPs (prediction method not reported). In a second protocol, membranes were washed with high-pH buffer, digested with CNBr in 90% FA, and then digested with trypsin in 2 M urea, giving 202 IMPs. Both protocols gave ~22% IMPs out of total protein identifications (authors' supplemental Table 2), with ~70% containing more than two TMDs (authors' supplemental Table 1). While this is not a particularly significant overall enrichment, it is worth considering that (1) previous 2DE experiments (internal citations) were not able to identify any IMPs, (2) in the current study, approximately one-third of all known IMPs (660 annotated) were identified in each experiment, and (3) a separate μLC-MS/MS analysis gave an IMP enrichment of >50%, so possibly the use of SCX, or some other MudPIT-specific variable, is compromising elution of hydrophobic peptides, thus reducing IMP identifications. As a comparison between the two methods, the authors note that trypsin shaving−60% methanol digestion gave more hydrophobic proteins (GRAVY 0.36 vs 0.28) and more TM domain-containing peptides (135 vs 32) than carbonate washing/CNBr/FA digestion. From trials varying each of the steps independently, it appears that both the enrichment-via-shaving and solubilization/digestion strategies contributed to the relative success of the first method.

Of note is the use of chymotrypsin with 60% methanol rather than trypsin alone, which has the potential to generate more peptides from cleavages in TMD domains since chymotrypsin cleaves at hydrophobic residues (Phe, Tyr, Tryp) more likely than Arg and Lys to be present in TMDs.[206] The trypsin/chymotrypsin combination has also been predicted by *in silico* analysis by Fischer and Poetsch[212] to yield good coverage of IMPs.

6. Affinity Purification

A summary of affinity purification techniques is given in Table 4.

6.1. Glycosylated Proteins

While not specific to IMPs, glycosylation can be exploited for affinity purification. Membrane proteins, especially those found in the extracellular membrane and recirculating intercellular vesicles such as endosomes and lysosomes, are often heavily glycosylated.[213−215] Appendage of carbohydrate moieties can occur at Ser or Thr residues (*O*-linked) or at Asn (*N*-linked) in the following sequence motif: N−X−S/

T, where X denotes any residue except Pro.[216] A variety of methods exist for affinity purification of glycoproteins and glycopeptides, including lectin-mediated affinity purification and chemical derivatization methods. To facilitate separation, analysis, and sequencing of modified peptides, the carbohydrate moiety is generally removed prior to MS analysis, by either enzymatic or chemical means.[217] This literature has recently been reviewed.[204,218]

For example, Ghosh et al.[219] applied lectin affinity capture (Con A and WGA) to a crude K562 cell lysate, resulting in the identification of ~150 proteins by RPLC MALDI-MS/MS analysis of tryptic peptides, a third of which were classified as IMPs (ENSEMBL). This enrichment is fairly good given that no prefractionation was attempted, suggesting that the use of membrane enriched samples could increase the IMP enrichment dramatically.

Following up on an earlier study,[220] Fan and colleagues[221] applied lectin affinity chromatography to the proteomic analysis of *N*-glycosylated *C. elegans* membrane proteins. However, instead of enriching glycoproteins directly, the sample was predigested with trypsin, and the resulting glycopeptides were affinity purified on a Con A lectin column. Bound glycopeptides were eluted with excess monosaccharide and deglycosylated using PNGase F, resulting in a signature Asn to Asp conversion (which can be exploited for ^{18}O labeling to enhance the mass shift). Using a combination of MALDI-MS/MS, RPLC-MALDI-MS/MS, and MudPIT analysis, 117 glycoproteins were identified, including 73 (62%) predicted IMPs (SOSUI variant).

In another study, Atwood et al.[222] applied lectin affinity capture to an organellar subfraction of the human parasite *Trypanosoma cruzi*. Trypsin digested glycopeptides were affinity purified on a Con A lectin column. Eluted peptides were then deglycosylated with PNGase F. Analysis by μLC-MS/MS gave 17 glycoproteins, 15 (88%) of which were predicted IMPs (TMHMM). While the numbers are rather low, possibly the result of extensive sample filtration and desalting, the IMP enrichment compares quite favorable to the non-glycan-enriched organellar subfraction, which yielded 42 proteins, only 6 (14%) of which had predicted TMD domains.

In an older study using hydrazide chemistry, Zhang et al.[223] demonstrated the affinity purification of *N*-linked glycopeptides from cell surface IMPs from a prostate cancer epithelial cell line. Glycoproteins from an enriched membrane fraction were first oxidized with periodate to convert carbohydrate *cis*-diols to aldehydes, which can form a covalent hydrazone linkage with immobilized hydrazines. (For quantitation purposes, immobilized peptides can be isotopically labeled at the *N*-terminus.) Subsequent treatment with PNGase F results in cleavage at the peptide−carbohydrate bond, releasing bound species. Following analysis by μLC-MS/MS, 64 glycoproteins were identified, including 45 (70%) IMPs (SWISS-PROT and/or PSORT II).

An *in silico* evaluation of 261 CD molecules, representing a diversity of plasma membrane proteins, suggests that ~97% could theoretically be enriched by glycosylation-based enrichment, with ~50% yielding at least one glycosylated peptide identifiable by trypsin digestion and MS analysis.[224] However, the enrichment potential suggested by this analysis has yet to be borne out experimentally. Overall, affinity purification of glycopeptides from membrane fractions has not yielded particularly high numbers of proteins (~150 or less), so its utility as a general IMP enrichment strategy is limited. It appears, for now, to be more suited for the specific analysis of glycoproteins and their sites of glycosylation.

6.2. Biotinylation

Biotinylation reagents, consisting of a biotin moiety linked to an amine-reactive *N*-hydroxysuccinimide (NHS) or *N*-hydroxysulfosuccinimide (sulfo-NHS; increased water solubility) group, are generally regarded as membrane-impermeable and, thus, are applicable to labeling of IMPs exposed at the cell surface. Sulfo-NHS-SS-biotin, with an internal disulfide bond, is a more hydrophilic reagent and is reported to be superior to the other commonly used reagent, sulfo-NHS-LC-biotin (LC denotes hydrocarbon spacer), in terms of plasma membrane specificity and labeling efficiency.[170] The disulfide allows for cleavage from avidin resin upon treatment with reducing agent, leaving the biotin moiety attached to the avidin/streptavidin affinity support upon elution. Following labeling, enrichment can be performed at either the peptide or protein level; enrichment of proteins allows for broader sequence coverage, whereas enrichment of peptides allows for targeted analysis of labeled residues.

Nunomura et al.[225] labeled the cell surface of mouse embryonic stem cells with sulfo-NHS-LC-biotin. Isolated plasma membranes were digested with trypsin, and labeled peptides were affinity purified via avidin. Eluted peptides were analyzed by MudPIT, giving 608 unique peptide identifications, 551 of which were biotinylated. Specific fragmentation of the biotin tag in addition to sequence information confirmed labeling in the tandem MS spectra. In total, 324 proteins were identified, 62% of which were predicted IMPs (SOSUI). An interesting benefit afforded by the membrane-impermeability of the biotinylating reagent is that it should only label lysines exposed on the cell surface, which allowed for topological assessment of 122 single pass IMPs, localizing the N-terminus to the extracellular space in 80% of cases, and the rest unambiguously assigned by selective C-terminal modification.

Zhao et al.[161] labeled a human carcinoma cell line with sulfo-NHS-SS-biotin and collected labeled proteins on magnetic streptavidin beads. Affinity captured proteins were then subject to high-salt and high-pH washes. Proteins were eluted upon treatment with DTT and analyzed by 1DE-μLC-MS/MS, giving 898 proteins, with an estimated 53% IMPs (SOSUI) and an additional 96 with lipid-anchored consensus sequences. The authors compared this result to a previous study,[226] in which only ~16% of 918 total identifications were IMPs or lipid-anchored. The major difference between the two studies was the high-salt and high-pH washes, demonstrating the importance of removing cytosolic and membrane-associated proteins for detection of the lower abundant IMPs.

In another study, Sostaric et al.[227] undertook the global profiling of the plasma membrane proteome of oviducal epithelial cells using a biotinylation approach. Cultured cells were labeled with sulfo-NHS-LC-biotin and labeled proteins were avidin affinity purified. Proteins were either analyzed by 2DE-μLC-MS/MS (40 proteins) or digested and analyzed by 1DE-μLC-MS/MS (276 proteins), 56% of which were IMPs (SOSUI).

In a different approach, hydrophobic, cell-permeable biotinylating reagents were used by Tang et al.[228] to profile intact *S. oneidensis* bacteria. It was anticipated that the lipophilic reagents would remain closely associated with the membrane, where they would be in a position to preferen-

tially react with membrane and membrane-associated proteins. Of 410 proteins identified, 174 were localized to the cell envelope, with 56 (32%) predicted IMPs (TMHMM).

Cell surface biotinylation can only be applied to cells rather than tissue, limiting its utility. Other concerns with biotinylation include possible contamination from labeling internal components of nonviable cells, or the failure to label cell-surface proteins with few reactive lysine residues (most Lys are cytoplasmically located), due to small extracellular regions or occluding PTMs (e.g., glycosylation).[225,229] Additionally, biotinylation of cell surface proteins is less selective with bacteria containing large porins on the outer membrane (e.g., *E. coli*) that allow the hydrophilic reagent to be internalized into the periplasmic space.[230]

However, despite these potential problems, cell surface biotinylation has proven to be a robust technique, likely due to the near-quantitative labeling afforded by amine modification by NHS derivatives (or thiols by iodoacetamide) and the exploitation of avidin−biotin binding for purification. In terms of overall numbers, significantly more proteins per experiment are identified using this technique as compared to glycoprotein affinity purification. Biotinylation experiments can also provide important topological information by identifying domains exposed on the cell surface or labeled organelles, which can be helpful in defining protein structure as well as refining structure prediction algorithms.

6.3. Protein Sequence Tag

The Protein Sequence Tag (PST) technology is more appropriately classified as a chemical modification strategy for selective *de-enrichment* via affinity purification of unwanted peptide species. The method allows for the specific isolation of N-terminal peptides from each polypeptide generated by CNBr cleavage. Following solubilization in FA and CNBr cleavage, polypeptides are isolated by size exclusion. Free thiols (Cys) and free amines (Lys, N-termini) are then blocked via reaction with iodoacetamide and an NHS derivative of *N*,*N*-dimethylglycine, respectively. Trypsin digestion then breaks each polypeptide down into one peptide with a blocked N-terminus and multiple peptides with free N-termini. (Note: due to Lys modification, cleavage takes place only at Arg.) All peptides with reactive amines are then scavenged from solution using activated ester resin, leaving behind a much smaller population of N-terminal peptides from each original CNBr fragment, which are then analyzed by off-line MudPIT. As applied to yeast mitochondria, 175 proteins were identified, of which 41 (23%) were predicted IMPs and included 20 with more than 2 TMDs (authors' Table 1, TMHMM predictions). This result compared quite favorably to a classic 2DE approach, where only 5 (4%) of 112 proteins identified were membrane-integral and were limited to 1−2 TMD-containing proteins (authors' Table 2, TMHMM predictions).[140] By adding a sodium carbonate wash to their purification protocol, Hamon and colleagues were able to increase their IMP enrichment to 72% (106 of 148 proteins identified) and the number of proteins with more than 2 TMDs to 59 (authors' Table 1, TMHMM predictions). Overall, each protein was identified by ~1.5 peptides, of which >80% corresponded to the N-terminus of a CNBr fragment, demonstrating fairly selective isolation of modified peptides.[139]

7. Quantitation

Like most methods in proteomics, quantification strategies (summarized in Table 4) were not necessarily developed with membrane proteins in mind, and thus, they may require optimization. Because gel-based separation techniques compatible with IMPs do not have the resolving power of classical 2DE, gel-based quantification is possible[231,232] but of limited utility. As a result, most current IMP quantification strategies involve shotgun proteomics analysis (for review and discussion of strategies, advantages, and disadvantages, see refs 233−236).

As compared to the case of soluble proteins, all quantification strategies may be hindered by the reduced number of peptides generated from IMPs using standard solubilization/digestion techniques. This section is mostly limited to discussion of techniques specifically designed or adapted for IMPs. In the absence of IMP-specific protocols, references are given for the application of standard quantification techniques to membrane preparations, though these reports generally lack any discussion specific to their relative ability to quantify IMPs.

Shotgun quantification strategies can be divided into two basic categories: label-free and stable isotope labeling. Label-free quantification strategies have no particular bias against IMPs other than the general considerations of efficient solubilization, digestion, and μLC separation. These methods are based on quantitation of peak areas/intensities (see, for example, refs 197, 237, and 238) or the number of MS/MS spectra for a particular peptide (see ref 198, for example).

Stable isotope labeling involves the incorporation of heavy atoms into proteins/peptides by metabolic means or post-processing chemical modification. During μLC-MS/MS analysis, isotopomeric peptides bearing heavy and light tags elute simultaneously, and the relative ion peak areas can be quantified, giving an indirect estimate of relative protein abundance. Metabolic incorporation of amino acids can be done at the cellular or organismal level using ^{15}N, ^{2}H, or ^{13}C isotopes (see, for example, refs 60 and 239−242) and the protein labeling itself is not inherently biased against IMPs.

Alternatively, chemical modification can be performed at the protein (e.g., ICAT, HysTag, and ICPL) or peptide (e.g., iTRAQ, ^{16}O/^{18}O) level. Labeling at the protein level typically involves modification of nucleophilic amino acids (mostly Cys or Lys) with isotopically labeled tags. However, due to the experimental challenges of targeting the reactive sites (if present) in TMDs, labeling is largely restricted to solvent-exposed soluble domains. Consequently, there are significantly fewer quantifiable peptides/proteins for IMPs as compared to soluble proteins using Lys or Cys labeling.[232] However, several groups have adapted labeling strategies specifically for membrane-embedded proteins, which are discussed below.

For labeling at the peptide level, the general concern of reduced number of quantifiable peptides generated with standard solubilization/digestion techniques applies. However, the actual chemical modification step, such as peptide N-terminal modification with NHS-derivatized isobaric iTRAQ tags,[243] should not be especially biased against membrane proteins, given that peptide solubility properties are much improved as compared to those of their parent IMPs. There are a few examples of iTRAQ applied to membrane proteomes.[244−247] The exception is when digestion and label incorporation are coupled, as in ^{16}O/^{18}O labeling; then, the efficient solubilization of membrane proteins becomes a significant concern, as addressed below.

7.1. ICAT

Quantification by ICAT[248] involves the differential labeling of two proteomic samples with light and heavy ($^{12}C/^{13}C$) forms of a cysteine-reactive iodoacetyl-biotin affinity tag. Samples are then combined, digested, affinity purified, and analyzed by μLC-MS/MS. Because the standard ICAT labeling buffer (~6 M urea/0.1% SDS) may not be sufficient to solubilize/denature all IMPs, alternative strategies may be desired to maximize cysteine labeling. Prokai et al.[249] used the nonionic detergent *n*-octyl glucoside for solubilization of synaptic plasma membranes, quantifying ~75 proteins, including some IMPs.

Reports differ as to the compatibility of ICAT with higher percentages of SDS. Aebersold and colleagues have used 0.5% SDS[250] but also report decreased labeling at those concentrations.[251] Ramus et al. have used concentrations as high as 4% to successfully label select proteins with ICAT.[252] However, to remove SDS, the group performed 1DE, necessitating in-gel digest and affinity purification of extracted peptides. Alternatively, one could precipitate the protein (e.g., methanol−chloroform) and resuspend in a MS-compatible denaturant (e.g., 60% methanol, RapiGest, PPS) for digest, dilute to 0.1% SDS prior to digest, or take advantage of the affinity tag for protein purification.

The use of aqueous−organic solvents for solubilization is also compatible with cysteine labeling.[134,209] Complete solubilization of proteins using 60% methanol was shown to significantly increased the number of cysteine labeled peptides obtained as compared to the standard protocol.[209]

7.2. HysTag

Similar to ICAT reagent, HysTag was also designed to modify cysteines, and it has been applied to quantify membrane proteins.[201,253] The reagent is a decapeptide with the sequence $Hys_6AlaArgAla(D_4)Cys$ that, upon activation with 2,2′-dipyridyl disulfide (DPDS), reacts with cysteine residues via disulfide bond formation. The reagent has an internal trypsin cleavage site (Arg) that allows for removal of the $Hys_6AlaArg$ portion of the tag following affinity purification by metal-affinity or cation-exchange chromatography. Heavy and light forms are distinguished by four D/H atoms in the alanine C-terminal to the trypsin cleavage site. It should be noted that the use of deuterium as an isotopic label is less optimal than ^{13}C, as isotopomeric peptides do not necessarily coelute. Consequently, peptide retention times may need to be corrected, and ion pairs may be subject to differential ion suppression due to different ionization environments, potentially leading to skewed quantification. Prior to labeling with HysTag, proteins (in intact membranes) are reduced with DTT (membrane-permeable) to expose reactive cysteines, and excess reducing reagent is removed by isolation of membranes by centrifugation. (This is not a concern for ICAT, as modification results in an acetyl moiety, which is stable under reducing conditions.) Reaction with the tag is performed in 4 M urea, and labeled proteins are digested with Lys-C (rather than trypsin, to avoid premature cleavage of the tag). The HysTag protocol was used to quantify mouse brain membrane proteins, allowing quantification of several hundred proteins via μLC-MS/MS (281 in ref 201 and 555 in ref 253); however, because no breakdown of IMP percentage was reported, it is hard to gauge how well this strategy worked to identify IMPs. However, to increase targeting of IMPs, it should be possible to combine HysTag labeling and protein digestion with more aggressive solubilization strategies in order to target otherwise inaccessible cysteines. Alternative proteolysis could also improve the numbers of useful peptides generated.

7.3. ICPL

In contrast to ICAT and HysTag that target cysteine residues, isotope coded protein labeling (ICPL) involves the reaction of primary amines (lysines and protein N-termini) with $^{12}C/^{13}C$ derivatives of 6-nicotinoyl-NHS.[254] Bisle et al.[232] applied this quantification strategy to the halophilic archaeon *H. salinarum*, and they were able to quantify 175 proteins by μLC-MALDI-MS/MS, of which 101 (58%) were IMPs (TMHMM). As compared to cysteine labeling strategies, lysine labeling increased the number of theoretically quantifiable peptides by approximately three times (though largely restricted to soluble domains); however, the number of quantifiable peptides for membrane proteins is still low in contrast to that for soluble species, which could be ameliorated if different digestion strategies were employed. Also, the ICPL reagent contains no affinity tag, so labeled peptides are not enriched before analysis; however, for the relatively simple genomes of prokaryotes such as *H. salinarum*, this is not a significant obstacle.

7.4. ^{18}O

Heavy oxygen labeling involves the enzyme-mediated incorporation of ^{18}O from bulk water into peptide C-termini (for a recent review, see ref 255). Specifically, digestion with the endoproteinases trypsin, Lys-C, and Glu-C results in the net incorporation of two oxygen molecules from bulk water into the C-terminus of the peptide product. This phenomenon is due to the fact that, following cleavage, peptides continue to interact with the protease, undergoing further binding/hydrolysis cycles until equilibration is achieved. (This is in contrast to chymotrypsin and Asp-N, which only incorporate one oxygen atom from bulk water.)[256] The 4 amu difference between isotopomeric peptides resulting from digestion in $H_2{}^{16}O$ and $H_2{}^{18}O$ can be successfully detected by MS, allowing relative quantitation of two samples.

Blonder et al. developed their 60% methanol/trypsin protocol[208] specifically for the improved digestion of IMPs; however, the reduced trypsin activity in the high-organic buffer is not sufficient for complete isotope incorporation, due to the required repetitive interaction with the enzyme. To remedy this problem, Blonder et al.[132] modified the ^{18}O labeling protocol by decoupling protein digestion and isotopic labeling, whereby trypsinization of the sample in 60% methanol/$H_2{}^{16}O$ is followed by trypsin-mediated labeling in 20% methanol/$H_2{}^{18}O$, in which enzymatic activity is sufficiently improved to allow for complete label incorporation. The method was applied to the lipid raft proteome of differentially induced vero cells, and an off-line MudPIT analysis resulted in the quantitation of 1417 peptides from 585 unique proteins, including 135 (23%) IMPs (TMHMM). Importantly, TMD peptides were among those quantified.

In a separate report, the IMP-optimized $^{16}O/^{18}O$ labeling strategy was applied to a lipid raft proteome in combination with ICAT labeling of cysteine-containing peptides.[134] A total of 706 and 129 proteins were quantified by each method, respectively (analysis by off-line MudPIT), with over 600 proteins uniquely identified by $^{16}O/^{18}O$, but only 32 uniquely identified by ICAT.

Enzymatically mediated $^{16}O/^{18}O$ labeling is a relatively simple protocol, requiring minimal sample manipulation, and allows for (1) labeling of all tryptic peptides (i.e., as compared to only cysteine-containing), (2) retention of any PTMs, (3) quantitation of differences as low as 1.5 fold, and, as demonstrated, (4) adaptation for membrane proteins. One additional point to consider with quantitation methods is that, while the ability to quantify many peptides per protein may provide seemingly redundant information, the increase sequence coverage is extremely important when post-translational modifications are considered.

8. MS Analysis of Hydrophobic Peptides

8.1. Ionization

8.1.1. ESI

ESI is the one area where studying hydrophobic peptides can actually be advantageous. ESI generates aerosolized droplets with a charged surface, containing the excess protons produced during ESI, and a neutral interior, consisting of solvent, electrolytes, and charged analytes paired by counterions. As the solvent evaporates, ions are released into the gas phase. Analytes with the largest proportion of nonpolar functional groups are most likely to be found at the droplet surface and have the highest ionization efficiencies. According to the equilibrium partitioning model of Cech and Enke[257] (see also ref 258), each analyte has a partitioning coefficient (*K*) that is the ratio of its concentration on the droplet surface to that in the neutral interior. Analytes with higher *K* values thus exist predominantly on the surface of the droplet and are in a more favorable position to carry the excess charge. This factor is critical for successful ESI analysis, as only ions that are (1) part of the charged droplet surface phase and (2) capable of carrying a charge will be present in the MS spectrum. Because hydrophobic peptides will be more likely to exist on the droplet surface where their hydrophobic regions can be desolvated, they are better positioned to compete for excess charge and thus experience greater ionization efficiency. It should be noted that even extremely hydrophobic peptides still have a highly polar peptide backbone, allowing them to carry a charge and thus satisfying criterion 2, above. From analysis of a series of tripeptides, hydrophobic peptides were found to suppress ionization of more polar peptides. In turn, hydrophobic peptide ionization can be suppressed by species with even higher *K* values, such as surfactants. This study was further extended to relate RP chromatographic retention time with ionization efficiency. In accordance with their previous results, it was found that, in general, compounds with higher retention times (typically longer and/or more hydrophobic) also had greater ESI response.[259]

8.1.2. MALDI

In contrast, MALDI ionization proceeds by a different mechanism, which is highly dependent upon peptide interaction with the matrix. Different functional groups—basic (especially Arg), polar, or aromatic residues—tend to ionize with highest efficiency.[260−262] In a comparison between MALDI-TOF/TOF and ESI-QTOF analysis of a trypsin digest of *E. coli* DNA-binding proteins, Stapels and Barofsky[182] reported that peptides identified by ESI were more hydrophobic in character.

8.2. Fragmentation

In addition to successful ionization, peptide identification is also highly dependent upon fragmentation efficiency. Most peptide sequencing experiments for routine proteomic analysis rely on low-energy fragmentation techniques such as collision induced dissociation (CID), where internal energy is imparted to peptide ions by collision with noble gas atoms. As explained by the Mobile Proton Model of Wysocki and colleagues,[263] energy imparted by collision causes the ionizing proton(s) to transfer intramolecularly until backbone amide protonation (believed to occur at the carbonyl oxygen)[264,265] causes amide bond destabilization and subsequent fragmentation. Such cleavage events are referred to as "charge-directed". The initiating charge is then retained on the most basic fragment, giving rise to a *b* or *y* ion; the complementary fragment (*y* or *b*) may or may not be observed depending on whether it is charged or neutral.[266] Other types of ions, such as *a* (*b* − CO), * ($-NH_3$), and ° ($-H_2O$) are also common in MS/MS spectra.[267] The peptide bond cleavages observed and their relative intensities are dependent upon numerous variables, including peptide ion charge state, residue content, sequence, size, and gas-phase secondary structure.

A variety of studies[266,268−272] have allowed some generalization of fragmentation pathways as follows: the most important variables are the mobility of proton(s), the position and identity of basic residues, and the presence of Pro. Basic residues are of primary importance because they have the ability to localize or "sequester" protons, so more energy is required to intramolecularly transfer the proton to a backbone carbonyl. The most basic residue is Arg, followed by His, Lys, N-terminal amines, and the amide carbonyl. This ranking is in accordance with resonance structures capable of delocalizing the extra charge. When a proton is localized to Arg, the activation energy required to mobilize the proton may be more than is needed to initiate so-called "charge-remote" fragmentation, which does not involve the proton. In such cases, cleavage C-terminal to acidic residues dominates. In the absence of Arg or when the number of protons exceeds the number of Arg residues, then the proton is mobile, and fragmentation occurs through the more common charge-directed pathways. Frequently observed cleavage sites are N-terminal to Pro and/or are near other basic residues (e.g., His)[264] or basic secondary structure regions (C-terminus of helical peptides).[268] With a mobile proton in the absence of Pro, energy differences between different backbone protonation sites are considerably reduced, with the proton having no strong preference for a particular amide bond. Resulting cleavage is thus largely nonselective, providing rich fragmentation patterns for sequencing.

Concerning peptide size, longer peptides (not surprisingly) tend to ionize in higher charge states, increasing the chance of having a mobile proton and good fragmentation. Indeed, singly charged peptides are reported to often generate poor tandem MS spectra due to dominance of selective cleavage events.[270] However, there are disadvantages to sequencing very long peptides. Most sequencing algorithms (e.g., Mascot, Sequest) are optimized to handle lower charge states (+1 to +3 ions), from which high-confidence assignments are relatively more straightforward to make: longer/higher charge state peptides have more possible fragment ions, increasing the probability of multiple ion assignments to a single peak and randomly matching peaks. Internal ions, resulting from two intramolecular cleavages, are also more

Table 5. Recommended Topology Prediction Algorithms for Estimating Percent IMP Enrichment (Section 9.2)

program	website	comments
TMHMM2	http://www.cbs.dtu.dk/services/TMHMM/	very good at distinguishing soluble from membrane-integral proteins; conservative IMP prediction
SOSUI	http://bp.nuap.nagoya-u.ac.jp/sosui/	very good at distinguishing soluble from membrane-integral proteins
HMMTOP2	http://www.enzim.hu/hmmtop/index.html	good at distinguishing soluble from membrane-integral proteins
PHOBIUS	http://phobius.cgb.ki.se/	same as TMHMM2, but can distinguish N-terminal signal peptides from TMDs

abundant. Instrumentation can also be a limiting factor, as low-charge parent ions of long peptides may lie outside the typical full scan range, and significant numbers of high-mass-fragment ions may lie outside the MS/MS scan range (up to ~2000 *m*/*z*).[270] That being said, it has been shown that analysis of higher charge states (+4 to +7) is possible by deconvoluting spectra generated on high-mass-accuracy instruments prior to a database search, allowing identification of peptides up to 3400 Da.[273]

It follows from these observations that hydrophobic peptides, being generally longer and lacking in basic residues, should provide good fragmentation spectra, provided that digestion strategies are used that avoid excessively long peptides. Thus, in general, failure to identify TMDs tends to result from problems in sample preparation/separation steps leading up to MS analysis, but not from MS itself.

9. Global Topology Assignment

9.1. Integral Membrane Protein Structure

In order to maintain energetically favorable interactions with the bilayer core, the primary sequences of TMDs are composed largely of hydrophobic (mostly aliphatic) residues, and they must be at least 15 residues in length to span the entire membrane.[274] The aromatic amino acids Tyr and Trp are relatively abundant around the lipid−water interface, while charged and polar residues are largely restricted to the soluble domains.[206] Additionally, because cytoplasmically oriented loop regions do not have to be translocated across the bilayer upon folding, they tend to be longer and contain a larger number of positively charged residues (which require more energy expenditure to translocate),[274] providing a "positive-inside rule"[275,276] for the prediction of IMP orientation. Owing to their strong constitutional bias and the propensity to adopt a semiregular secondary structure, predicting the general topology of α-helical IMPs—the number and location of TMDs and their orientation with respect to the cytoplasm—should be relatively straightforward, certainly easier than a priori prediction of soluble protein structure.[274,277]

However, much of what is known about IMP structure is based on the high-resolution structures of only a handful of proteins, and as more structures are solved, the topological landscape becomes increasingly diverse. As of publication, there are only ~125 unique high-resolution IMP structures in the Protein Data Bank,[278] which are listed at http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html. The number of entries for IMPs doubles every ~3 years, but it still lags behind structural assignment of soluble proteins by ~15 years.[279] This discrepancy is largely accounted for by the challenging biochemical properties IMPs, which can be difficult to recombinantly express, solubilize in active form, and purify[1−3] and which are notoriously hard to crystallize despite some recent technical advances. Similarly, structural NMR studies can be complicated by the presence of a lipid bilayer and require significant amounts of purified material.[2]

With regard to their structural diversity, it is now clear that TM helical structures are not nearly as limited as first thought. Membrane spanning regions may be very long, containing 40 residues or more,[205] highly tilted, flexed, kinked, or interrupted by short intramembrane breaks in helicity.[280−282] Helices may not necessarily span the entire membrane—turning back to form a re-entrant loop rather than an actual *trans*membrane domain.[283,284]

There are also rare cases of dual topology proteins, which coexist in opposite orientations in equal abundance,[285] and "frustrated" proteins that can adopt multiple orientations due to conflicting structural requirements[286] (for reviews, see refs 6 and 287). Thus, topological characterization is a challenging task, and advances will be facilitated as more high-resolution structures become available for training and testing and more structural constraints are derived by experimental means.[288]

9.2. Topology Prediction Algorithms

A variety of algorithms have been developed for IMP topology prediction (for reviews, see refs 13 and 274).

Kyte and Doolittle[26] first introduced the hydropathy index in 1982 for identifying probable TMDs using a "sliding window" approach, whereby the average hydropathy of the 19 contiguous residues in each window was calculated, with higher scores being indicative of greater preference for the lipid environment. From there, the first significant advance was TopPred,[289] which combined the sliding window approach for TMD prediction with the "positive-inside rule" for predicting orientation with respect to the cytoplasm. In a departure from a fixed length window, Jones et al.[290] introduced MEMSAT, which was able to predict the most probable TMD length, in addition to location and orientation. The algorithm took into account the experimentally derived propensity of each amino acid to be in one of five locations—an inside or outside loop, inside end, middle, or outside end of a helix—for its predictions. The first method to use a neural network, trained on experimentally characterized proteins, was PHDhtm.[291] It was soon joined by other algorithms using circular Hidden Markov models (HMM)—namely TMHMM and HMMTOP—to choose between different possible states (seven for TMHMM, five for HMMTOP) and pick the most likely topology by comparison with experimentally derived models. With over 30 topology prediction algorithms available to chose from, consensus programs have also been developed that take into account the predictions of multiple algorithms, which tend to cancel out individual errors and give higher accuracy than individual methods.[292]

While several recent studies[66,283,293] have attempted to compare prediction programs, interpretation of results is difficult due to (1) the paucity of high-resolution structures to use as a benchmark and (2) the fact that most of those structures were likely part of the programs' initial training

set. Some programs perform better or worse than others with respect to specific parameters, but no one method is superior in all regards. Common pitfalls include assigning TMDs in soluble proteins, classifying signal peptides as TMDs, and misassigning TMDs due to minimum and maximum length constraints.[274] As such, it is recommended that multiple prediction algorithms be used to obtain the most reliable topology assignments. According to a survey by Punta et al.,[274] programs that tend to rank highly overall are TMHMM2,[12] HMMTOP2,[25] PHDhtm,[291] SPLIT4,[294] and TMAP.[295] Their review contains an excellent list of recommended prediction methods and useful links.

For the analysis of the IMP content of a particular dataset, TMHMM(v1 or 2)[296] and SOSUI[24] stand out with respect to their ability to distinguish soluble proteins from IMPs, misclassifying fewer than ~1% of globular proteins. Another commonly used algorithm, HMMTOP2,[25] gives slightly higher false positives (6%) but misses fewer true IMPs (0−1% vs 4−8% for TMHMM and SOSUI).[66] As mentioned above, one of the most significant shortcomings of these prediction algorithms is their tendency to misclassify signal sequences as TMDs.[66] However, this problem has been addressed with the development of Phobius,[297,298] which allows for the simultaneous prediction of both TMDs (using TMHMM2) and signal peptides (using SignalP-HMM).[299]

It should be noted that many non-HMM topology prediction methods (e.g., TMpred,[300] MEMSAT1.5,[290] Toppred2,[301] and DAS[302]) *may* predict TMDs in a considerable number of soluble proteins,[66,146] potentially giving highly inflated estimates of IMP enrichment. For a summary of recommended algorithms for estimating IMP content, see Table 5.

9.3. Experimental Constraints Improve Topology Prediction

Providing prediction algorithms with constraints, such as the relative location of the C-terminus, has been shown to improve the prediction models.[293] von Heijne and colleagues used global strategies to define the location of the C-terminus for both *E. coli*[303] and *S. cerevisiae*[304] IMPs, and by homology, they extended the topology information to over 51 000 bacterial inner membrane proteins[305] and ~15 000 eukaryotic IMPs.[304]

For *E. coli*, the reporter proteins alkaline phosphatase (PhoA) and green fluorescent protein (GFP) were fused to the C-terminus of 573 candidate IMPs. Because PhoA is active only in the periplasm[306] and GFP fluoresces only in the cytoplasm,[307] simple, high-throughput assays can determine the orientation of protein C-termini.[308,309] Without the experimental constraints, TMHMM predicted the correct orientation for only 78% of the IMPs in the dataset, and results were generally lower scoring than after the C-terminal location was defined. In addition, they identified a number of dual topology proteins and identified IMPs with N_{in}-C_{in} topologies as a predominant structural class.[303] Similarly, for the 51 000+ proteins found by homology searching, unrestrained TMHMM prediction provided the correct orientation in only 69% of cases. Thus, from an initial ~600 protein dataset, significantly improved topology maps were obtained for more than 30% of all predicted bacterial inner membrane proteins.[308] However, for the majority of eukaryotic IMP families, no bacterial homologues exist.[288] From the initial *E. coli* data set, only ~4000 eukaryotic homologues could be assigned.[304]

As such, Kim et al.[304] used the known topology reporter HA/Suc2/His4C[310,311] to create C-terminal fusion proteins for 617 *S. cerevisiae* IMPs. Suc2 and His4C are only glycosylated if translocated to the ER lumen, and His4C, the catalytic domain of His4p histidinol dehydrogenase, can act on its substrate only if located in the cytosol. Endoglycosidase H digestion and growth assays (of a *his4* mutant strain expressing the fusion protein in minimal media containing histidinol) can thus be used to assess the location of the C-terminus for each fusion IMP: if lumenally oriented, the protein would be glycosylated and not grow in minimal media supplemented with histidinol; if cytosolically located, protein would not be glycosylated but would grow under restricted conditions. Topology information was obtained for ~550 IMPs, for which prior, unconstrained TMHMM searches resulted in correct orientation for only 69%. The overall topology distribution was similar to that of *E. coli*, except N_{out}-7TMD-C_{in} proteins (e.g., GPCRs) were of higher abundance. Through homology searching, the C-terminal constraints could be extended to over 13 000 eukaryotic IMPs. Importantly, the C-terminal predictions for eukaryotic proteins found by homology searching of both the *E. coli* and yeast data sets were in nearly 100% agreement, validating topology extrapolation for homologous proteins.[304] This is an important finding, as similar C-terminal tagging studies for higher eukaryotes are not yet feasible on a global scale.[312]

10. Summary and Conclusion

The hydrophobic nature of the phospholipid bilayer core is an extremely inhospitable environment for otherwise polar species. However, the necessity of communication, signaling, and transport across bilayers forced the evolution of integral membrane proteins. They adapted by tailoring their primary sequence and secondary structures to maximize energetically favorable hydrophobic interactions between lipids and residue side chains, while at the same time shielding the polar backbone by extensive intra- (α-helix) or interstrand (β-sheet) hydrogen bonding. Unfortunately, this adaptation away from an aqueous environment has proven quite problematic for proteomic applications. Coupled with their low abundance, IMPs are difficult species to study both individually and on a global scale. Fortunately, there now seems to be a critical mass of research labs devoted to the study of these biomolecules. Indeed, there has been a virtual explosion of studies devoted to the proteomic characterization of IMPs in the past few years, advancing different methods for enrichment, solubilization, affinity purification, quantitation, and analysis. In terms of mature strategies for IMP proteomics, 1DE-μLC and MudPIT are clearly the forerunners. Alternate 2DE approaches, such as BN/SDS, 16-BAC/SDS, and dSDS, while superior to classic 2DE in terms of compatibility with very hydrophobic species, do not have the resolving power to give them much advantage over a 1DE separation. Additionally, like 1DE, they require an orthogonal peptide separation step to maximize identifications. While several techniques have been suggested for improving the gel extraction of hydrophobic peptides, they have not been widely implemented, limiting TMD coverage and hindering identification of IMPs. Until recently, shotgun methods suffered from the same affliction, with IMP enrichment maxing out ~65% and TMD-containing peptides routinely underrepresented. However, that shortcoming can largely be alleviated with protocols that specifically target TMD domains. Given that membrane-embedded regions may make

up a large percentage of IMP structure, the success of this strategy is not surprising. The targeted enrichment of TMDs is achieved by removing *both* soluble proteins and soluble IMP domains (via proteinase K shaving at high pH). Moreover, to deal with TMD hydrophobicity, enrichment is combined with appropriate digestion strategies (CNBr) to reduce TMD length/hydrophobicity and shotgun separation at elevated temperature to improve hydrophobic peptide recovery. Fortunately, there are now a number of groups interested in similar solubilization/digestion strategies. With continued refinements of sample preparation and shotgun separation/analysis, hopefully it will soon be possible to routinely identify thousands of IMPs from highly enriched samples, allowing IMP characterization, at least on a global profiling level, to advance to that of soluble proteins.

11. Abbreviations

1DE	one-dimensional gel electrophoresis, specifically SDS-PAGE
2DE	two-dimensional gel electrophoresis
traditional 2DE	IEF/SDS-PAGE
16-BAC	benzyldimethyl-*n*-hexadecylammonium chloride
BN	blue native
CID	collision-induced dissociation
μCIEF	microcapillary isoelectric focusing, same as CIEF
CN	clear native
CNBr	cyanogen bromide
CMC	critical micelle concentration
dSDS	2D SDS/SDS-PAGE
DTT	dithiothreitol
ESI	electrospray ionization
FA	formic acid
GPCR	G-protein coupled receptor
GPI	glycosylphosphatidylinositol
GRAVY	grand average of hydropathy (see introduction)
hppK	high-pH-proteinase K
ICAT	isotope-coded affinity tag
ICPL	isotope-coded protein labeling
ID	identification
IEF	isoelectric focusing
IMP	integral membrane protein, specifically α-helical
μLC	microcapillary liquid chromatography, RP unless otherwise specified
LC	liquid chromatography
MALDI	matrix-assisted laser desorption/ionization
MeOH	methanol
MEP	membrane-embedded peptide
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MudPIT	multidimensional protein identification technology
NHS	*N*-hydroxysuccinimide
OG	*n*-octyl glucoside
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PEG	polyethylene glycol
pI	isoelectric point
PM	plasma membrane
PMF	peptide mass fingerprinting
PTM	post-translational modification
QTOF	quadropole time-of-flight
RP	reversed-phase
SCX	strong cation exchange
SDC	sodium deoxycholate
SDS	sodium dodecyl sulfate
TCA	trichloroacetic acid
TMD	transmembrane domain, specifically a transmembrane α-helix
TOF	time-of-flight
WGA	wheat germ agglutinin

12. Acknowledgments

Financial support was provided by NIH Grants DA021744 and AA016171.

13. References

(1) Ott, C. M.; Lingappa, V. R. *J. Cell Sci.* **2002**, *115*, 2003.
(2) Torres, J.; Stevens, T. J.; Samso, M. *Trends Biochem. Sci.* **2003**, *28*, 137.
(3) Zhou, C.; Zheng, Y.; Zhou, Y. *Genomics, Proteomics Bioinf.* **2004**, *2*, 1.
(4) Hopkins, A. L.; Groom, C. R. *Nat. Rev. Drug Discovery* **2002**, *1*, 727.
(5) Russ, A. P.; Lampel, S. *Drug Discovery Today* **2005**, *10*, 1607.
(6) Elofsson, A.; von Heijne, G. *Annu. Rev. Biochem.*, in press.
(7) von Heijne, G. *Prog. Biophys. Mol. Biol.* **1996**, *66*, 113.
(8) Chen, C. P.; Rost, B. *Appl. Bioinf.* **2002**, *1*, 21.
(9) Santoni, V.; Molloy, M.; Rabilloud, T. *Electrophoresis* **2000**, *21*, 1054.
(10) Eichacker, L. A.; Granvogl, B.; Mirus, O.; Muller, B. C.; Miess, C.; Schleiff, E. *J. Biol. Chem.* **2004**, *279*, 50915.
(11) Wallin, E.; von Heijne, G. *Protein Sci.* **1998**, *7*, 1029.
(12) Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. *J. Mol. Biol.* **2001**, *305*, 567.
(13) Hurwitz, N.; Pellegrini-Calace, M.; Jones, D. T. *Philos. Trans. R. Soc. London, B: Biol. Sci.* **2006**, *361*, 465.
(14) Bowie, J. U. *J. Mol. Biol.* **1997**, *272*, 780.
(15) Elortza, F.; Mohammed, S.; Bunkenborg, J.; Foster, L. J.; Nuhse, T. S.; Brodbeck, U.; Peck, S. C.; Jensen, O. N. *J. Proteome Res.* **2006**, *5*, 935.
(16) Whitelegge, J.; Halgand, F.; Souda, P.; Zabrouskov, V. *Expert Rev. Proteomics* **2006**, *3*, 585.
(17) Yates, J. R., III. *J. Mass Spectrom.* **1998**, *33*, 1.
(18) Washburn, M. P.; Wolters, D.; Yates, J. R., III. *Nat. Biotechnol.* **2001**, *19*, 242.
(19) Wolters, D. A.; Washburn, M. P.; Yates, J. R., III. *Anal. Chem.* **2001**, *73*, 5683.
(20) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299.
(21) Yates, J. R., III; Speicher, S.; Griffin, P. R.; Hunkapiller, T. *Anal. Biochem.* **1993**, *214*, 397.
(22) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64.
(23) Domon, B.; Aebersold, R. *Science* **2006**, *312*, 212.
(24) Hirokawa, T.; Boon-Chieng, S.; Mitaku, S. *Bioinformatics* **1998**, *14*, 378.
(25) Tusnady, G. E.; Simon, I. *Bioinformatics* **2001**, *17*, 849.
(26) Kyte, J.; Doolittle, R. F. *J. Mol. Biol.* **1982**, *157*, 105.
(27) Rolland, N.; Ferro, M.; Ephritikhine, G.; Marmagne, A.; Ramus, C.; Brugiere, S.; Salvi, D.; Seigneurin-Berny, D.; Bourguignon, J.; Barbier-Brygoo, H.; Joyard, J.; Garin, J. *J. Exp. Bot.* **2006**, *57*, 1579.
(28) Komatsu, S.; Konishi, H.; Hashimoto, M. *J. Exp. Bot.* **2007**, *58*, 103.
(29) Ephritikhine, G.; Ferro, M.; Rolland, N. *Plant Physiol. Biochem.* **2004**, *42*, 943.
(30) Lasserre, J. P.; Beyne, E.; Pyndiah, S.; Lapaillerie, D.; Claverol, S.; Bonneu, M. *Electrophoresis* **2006**, *27*, 3306.
(31) Graham, R. L.; O'Loughlin, S. N.; Pollock, C. E.; Ternan, N. G.; Weatherly, D. B.; Jackson, P. J.; Tarleton, R. L.; McMullan, G. *J. Proteome Res.* **2006**, *5*, 2465.
(32) Graham, R. L.; Pollock, C. E.; O'Loughlin, S. N.; Ternan, N. G.; Weatherly, D. B.; Tarleton, R. L.; McMullan, G. *Proteomics* **2007**, *7*, 82.
(33) Goo, Y. A.; Yi, E. C.; Baliga, N. S.; Tao, W. A.; Pan, M.; Aebersold, R.; Goodlett, D. R.; Hood, L.; Ng, W. V. *Mol. Cell. Proteomics* **2003**, *2*, 506.
(34) Klein, C.; Garcia-Rizo, C.; Bisle, B.; Scheffer, B.; Zischka, H.; Pfeiffer, F.; Siedler, F.; Oesterhelt, D. *Proteomics* **2005**, *5*, 180.
(35) Pedersen, S. K.; Harry, J. L.; Sebastian, L.; Baker, J.; Traini, M. D.; McCarthy, J. T.; Manoharan, A.; Wilkins, M. R.; Gooley, A. A.; Righetti, P. G.; Packer, N. H.; Williams, K. L.; Herbert, B. R. *J. Proteome Res.* **2003**, *2*, 303.
(36) Schluesener, D.; Fischer, F.; Kruip, J.; Rogner, M.; Poetsch, A. *Proteomics* **2005**, *5*, 1317.
(37) Rabilloud, T. *Proteomics* **2002**, *2*, 3.
(38) Zhang, L.; Xie, J.; Wang, X.; Liu, X.; Tang, X.; Cao, R.; Hu, W.; Nie, S.; Fan, C.; Liang, S. *Proteomics* **2005**, *5*, 4510.

(39) Zhang, L. J.; Wang, X. E.; Peng, X.; Wei, Y. J.; Cao, R.; Liu, Z.; Xiong, J. X.; Yin, X. F.; Ping, C.; Liang, S. *Cell. Mol. Life Sci.* **2006**, *63*, 1790.
(40) Hunzinger, C.; Wozny, W.; Schwall, G. P.; Poznanovic, S.; Stegmann, W.; Zengerling, H.; Schoepf, R.; Groebe, K.; Cahill, M. A.; Osiewacz, H. D.; Jagemann, N.; Bloch, M.; Dencher, N. A.; Krause, F.; Schrattenholz, A. *J. Proteome Res.* **2006**, *5*, 625.
(41) Reifschneider, N. H.; Goto, S.; Nakamoto, H.; Takahashi, R.; Sugawa, M.; Dencher, N. A.; Krause, F. *J. Proteome Res.* **2006**, *5*, 1117.
(42) Reinders, J.; Zahedi, R. P.; Pfanner, N.; Meisinger, C.; Sickmann, A. *J. Proteome Res.* **2006**, *5*, 1543.
(43) Da Cruz, S.; Xenarios, I.; Langridge, J.; Vilbois, F.; Parone, P. A.; Martinou, J. C. *J. Biol. Chem.* **2003**, *278*, 41566.
(44) Zahedi, R. P.; Sickmann, A.; Boehm, A. M.; Winkler, C.; Zufall, N.; Schonfisch, B.; Guiard, B.; Pfanner, N.; Meisinger, C. *Mol. Biol. Cell* **2006**, *17*, 1436.
(45) Wu, C. C.; MacCoss, M. J.; Mardones, G.; Finnigan, C.; Mogelsvang, S.; Yates, J. R., III; Howell, K. E. *Mol. Biol. Cell* **2004**, *15*, 2907.
(46) Blondeau, F.; Ritter, B.; Allaire, P. D.; Wasiak, S.; Girard, M.; Hussain, N. K.; Angers, A.; Legendre-Guillemin, V.; Roy, L.; Boismenu, D.; Kearney, R. E.; Bell, A. W.; Bergeron, J. J.; McPherson, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3833.
(47) Burre, J.; Beckhaus, T.; Schagger, H.; Corvey, C.; Hofmann, S.; Karas, M.; Zimmermann, H.; Volknandt, W. *Proteomics* **2006**, *6*, 6250.
(48) Coughenour, H. D.; Spaulding, R. S.; Thompson, C. M. *Proteomics* **2004**, *4*, 3141.
(49) Morciano, M.; Burre, J.; Corvey, C.; Karas, M.; Zimmermann, H.; Volknandt, W. *J. Neurochem.* **2005**, *95*, 1732.
(50) Moebius, J.; Zahedi, R. P.; Lewandrowski, U.; Berger, C.; Walter, U.; Sickmann, A. *Mol. Cell. Proteomics* **2005**, *4*, 1754.
(51) Cutillas, P. R.; Biber, J.; Marks, J.; Jacob, R.; Stieger, B.; Cramer, R.; Waterfield, M.; Burlingame, A. L.; Unwin, R. J. *Proteomics* **2005**, *5*, 101.
(52) Senis, Y. A.; Tomlinson, M. G.; Garcia, A.; Dumon, S.; Heath, V. L.; Herbert, J.; Cobbold, S. P.; Spalton, J. C.; Ayman, S.; Antrobus, R.; Zitzmann, N.; Bicknell, R.; Frampton, J.; Authi, K. S.; Martin, A.; Wakelam, M. J.; Watson, S. P. *Mol. Cell. Proteomics* **2007**, *6*, 548.
(53) Katsumata, O.; Kimura, T.; Nagatsuka, Y.; Hirabayashi, Y.; Sugiya, H.; Furuyama, S.; Yanagishita, M.; Hara-Yokoyama, M. *Biochem. Biophys. Res. Commun.* **2004**, *319*, 826.
(54) Zischka, H.; Braun, R. J.; Marantidis, E. P.; Buringer, D.; Bornhovd, C.; Hauck, S. M.; Demmer, O.; Gloeckner, C. J.; Reichert, A. S.; Madeo, F.; Ueffing, M. *Mol. Cell. Proteomics* **2006**, *5*, 2185.
(55) Sprenger, R. R.; Fontijn, R. D.; van Marle, J.; Pannekoek, H.; Horrevoets, A. J. *Biochem. J.* **2006**, *400*, 401.
(56) Ostrom, R. S.; Insel, P. A. *Methods Mol. Biol.* **2006**, *332*, 181.
(57) Rahbar, A. M.; Fenselau, C. *J. Proteome Res.* **2004**, *3*, 1267.
(58) Chaney, L. K.; Jacobson, B. S. *J. Biol. Chem.* **1983**, *258*, 10062.
(59) Stolz, D. B.; Jacobson, B. S. *J. Cell Sci.* **1992**, *103 (Pt 1*), 39.
(60) Rahbar, A. M.; Fenselau, C. *J. Proteome Res.* **2005**, *4*, 2148.
(61) Oh, P.; Li, Y.; Yu, J.; Durr, E.; Krasinska, K. M.; Carver, L. A.; Testa, J. E.; Schnitzer, J. E. *Nature* **2004**, *429*, 629.
(62) Durr, E.; Yu, J.; Krasinska, K. M.; Carver, L. A.; Yates, J. R., III; Testa, J. E.; Oh, P.; Schnitzer, J. E. *Nat. Biotechnol.* **2004**, *22*, 985.
(63) Schindler, J.; Nothwang, H. G. *Proteomics* **2006**, *6*, 5409.
(64) Everberg, H.; Leiding, T.; Schioth, A.; Tjerneld, F.; Gustavsson, N. *J. Chromatogr., A* **2006**, *1122*, 35.
(65) Schindler, J.; Lewandrowski, U.; Sickmann, A.; Friauf, E.; Nothwang, H. G. *Mol. Cell. Proteomics* **2006**, *5*, 390.
(66) Chen, C. P.; Kernytsky, A.; Rost, B. *Protein Sci.* **2002**, *11*, 2774.
(67) Cao, R.; Li, X.; Liu, Z.; Peng, X.; Hu, W.; Wang, X.; Chen, P.; Xie, J.; Liang, S. *J. Proteome Res.* **2006**, *5*, 634.
(68) Schmitt, S.; Prokisch, H.; Schlunck, T.; Camp, D. G., 2nd; Ahting, U.; Waizenegger, T.; Scharfe, C.; Meitinger, T.; Imhof, A.; Neupert, W.; Oefner, P. J.; Rapaport, D. *Proteomics* **2006**, *6*, 72.
(69) Blonder, J.; Terunuma, A.; Conrads, T. P.; Chan, K. C.; Yee, C.; Lucas, D. A.; Schaefer, C. F.; Yu, L. R.; Issaq, H. J.; Veenstra, T. D.; Vogel, J. C. *J. Invest. Dermatol.* **2004**, *123*, 691.
(70) Schirmer, E. C.; Florens, L.; Guan, T.; Yates, J. R., III; Gerace, L. *Science* **2003**, *301*, 1380.
(71) Fujiki, Y.; Hubbard, A. L.; Fowler, S.; Lazarow, P. B. *J. Cell Biol.* **1982**, *93*, 97.
(72) Howell, K. E.; Palade, G. E. *J. Cell Biol.* **1982**, *92*, 822.
(73) Blackler, A. R.; Speers, A. E.; Wu, C. C. *Nat. Methods*, submitted.
(74) Pasini, E. M.; Kirkegaard, M.; Mortensen, P.; Lutz, H. U.; Thomas, A. W.; Mann, M. *Blood* **2006**, *108*, 791.
(75) Fischer, F.; Wolters, D.; Rogner, M.; Poetsch, A. *Mol. Cell. Proteomics* **2006**, *5*, 444.
(76) Forner, F.; Arriaga, E. A.; Mann, M. *J. Proteome Res.* **2006**, *5*, 3277.
(77) Schroder, B.; Hasilik, A. *Anal. Biochem.* **2006**, *357*, 144.
(78) Simoes-Barbosa, A.; Santana, J. M.; Teixeira, A. R. *Electrophoresis* **2000**, *21*, 641.
(79) Wessel, D.; Flugge, U. I. *Anal. Biochem.* **1984**, *138*, 141.
(80) Mastro, R.; Hall, M. *Anal. Biochem.* **1999**, *273*, 313.
(81) le Maire, M.; Deschamps, S.; Moller, J. V.; Le Caer, J. P.; Rossier, J. *Anal. Biochem.* **1993**, *214*, 50.
(82) Borner, G. H.; Sherrier, D. J.; Weimar, T.; Michaelson, L. V.; Hawkins, N. D.; Macaskill, A.; Napier, J. A.; Beale, M. H.; Lilley, K. S.; Dupree, P. *Plant Physiol.* **2005**, *137*, 104.
(83) Carboni, L.; Piubelli, C.; Righetti, P. G.; Jansson, B.; Domenici, E. *Electrophoresis* **2002**, *23*, 4132.
(84) Speers, A. E.; Blackler, A. R.; Wu, C. C. *Anal. Chem.* **2007**, *79*, 4613.
(85) Moglich, A.; Krieger, F.; Kiefhaber, T. *J. Mol. Biol.* **2005**, *345*, 153.
(86) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1992**, *226*, 491.
(87) Tobi, D.; Elber, R.; Thirumalai, D. *Biopolymers* **2003**, *68*, 359.
(88) Gorg, A.; Weiss, W.; Dunn, M. J. *Proteomics* **2004**, *4*, 3665.
(89) Castellanos-Serra, L.; Paz-Lago, D. *Electrophoresis* **2002**, *23*, 1745.
(90) Wei, J.; Sun, J.; Yu, W.; Jones, A.; Oeller, P.; Keller, M.; Woodnutt, G.; Short, J. M. *J. Proteome Res.* **2005**, *4*, 801.
(91) Naldrett, M. J.; Zeidler, R.; Wilson, K. E.; Kocourek, A. *J. Biomol. Tech.* **2005**, *16*, 423.
(92) Garavito, R. M.; Ferguson-Miller, S. *J. Biol. Chem.* **2001**, *276*, 32403.
(93) Reynolds, J. A.; Tanford, C. *J. Biol. Chem.* **1970**, *245*, 5161.
(94) Zhang, N.; Li, L. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 889.
(95) Zhang, N.; Chen, R.; Young, N.; Wishart, D.; Winter, P.; Weiner, J. H.; Li, L. *Proteomics* **2007**, *7*, 484.
(96) Bosserhoff, A.; Wallach, J.; Frank, R. W. *J. Chromatogr.* **1989**, *473*, 71.
(97) Kawasaki, H.; Suzuki, K. *Anal. Biochem.* **1990**, *186*, 264.
(98) Vissers, J. P.; Chervet, J. P.; Salzmann, J. P. *J. Mass Spectrom.* **1996**, *31*, 1021.
(99) Zischka, H.; Gloeckner, C. J.; Klein, C.; Willmann, S.; Swiatek-de Lange, M.; Ueffing, M. *Proteomics* **2004**, *4*, 3776.
(100) Loo, R. R.; Dales, N.; Andrews, P. C. *Protein Sci.* **1994**, *3*, 1975.
(101) Lund, S.; Orlowski, S.; de Foresta, B.; Champeil, P.; le Maire, M.; Moller, J. V. *J. Biol. Chem.* **1989**, *264*, 4907.
(102) le Maire, M.; Champeil, P.; Moller, J. V. *Biochem. Biophys. Acta* **2000**, *1508*, 86.
(103) Chamberlain, L. H. *FEBS Lett.* **2004**, *559*, 1.
(104) Raimondo, F.; Ceppi, P.; Guidi, K.; Masserini, M.; Foletti, C.; Pitto, M. *Expert Rev. Proteomics* **2005**, *2*, 793.
(105) Quach, T. T.; Li, N.; Richards, D. P.; Zheng, J.; Keller, B. O.; Li, L. *J. Proteome Res.* **2003**, *2*, 543.
(106) van Montfort, B. A.; Doeven, M. K.; Canas, B.; Veenhoff, L. M.; Poolman, B.; Robillard, G. T. *Biochem. Biophys. Acta* **2002**, *1555*, 111.
(107) van Montfort, B. A.; Canas, B.; Duurkens, R.; Godovac-Zimmermann, J.; Robillard, G. T. *J. Mass Spectrom.* **2002**, *37*, 322.
(108) Katayama, H.; Tabata, T.; Ishihama, Y.; Sato, T.; Oda, Y.; Nagasu, T. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2388.
(109) Seddon, A. M.; Curnow, P.; Booth, P. J. *Biochem. Biophys. Acta* **2004**, *1666*, 105.
(110) Zhou, J.; Zhou, T.; Cao, R.; Liu, Z.; Shen, J.; Chen, P.; Wang, X.; Liang, S. *J. Proteome Res.* **2006**, *5*, 2547.
(111) Hjelmeland, L. M. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 6368.
(112) Banerjee, P.; Joo, J. B.; Buse, J. T.; Dawson, G. *Chem. Phys. Lipids* **1995**, *77*, 65.
(113) Kalipatnapu, S.; Chattopadhyay, A. *IUBMB Life* **2005**, *57*, 505.
(114) Yu, Y. Q.; Gilar, M.; Gebler, J. C. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 711.
(115) Norris, J. L.; Porter, N. A.; Caprioli, R. M. *Anal. Chem.* **2003**, *75*, 6642.
(116) Norris, J. L.; Porter, N. A.; Caprioli, R. M. *Anal. Chem.* **2005**, *77*, 5036.
(117) Norris, J. L.; Hangauer, M. J.; Porter, N. A.; Caprioli, R. M. *J. Mass Spectrom.* **2005**, *40*, 1319.
(118) Blackler, A. R.; MacCoss, M. J.; Wu, C. C. Comparison of Acid-Labile and Traditional Detergents for Membrane Solubilization and Digestion of Membrane Proteins. In 55th ASMS Conference on Mass Spectrometry, Indianapolis, IN, June 3−7, 2007.
(119) Ruth, M. C.; Old, W. M.; Emrick, M. A.; Meyer-Arendt, K.; Aveline-Wolf, L. D.; Pierce, K. G.; Mendoza, A. M.; Sevinsky, J. R.; Hamady, M.; Knight, R. D.; Resing, K. A.; Ahn, N. G. *J. Proteome Res.* **2006**, *5*, 709.
(120) Kragh-Hansen, U.; le Maire, M.; Noel, J. P.; Gulik-Krzywicki, T.; Moller, J. V. *Biochemistry* **1993**, *32*, 1648.
(121) Kragh-Hansen, U.; le Maire, M.; Moller, J. V. *Biophys. J.* **1998**, *75*, 2932.
(122) Bordier, C. *J. Biol. Chem.* **1981**, *256*, 1604.
(123) Fink, A. L.; Painter, B. *Biochemistry* **1987**, *26*, 1665.
(124) Welinder, K. G. *Anal. Biochem.* **1988**, *174*, 54.

(125) Russell, W. K.; Park, Z. Y.; Russell, D. H. *Anal. Chem.* **2001**, *73*, 2682.
(126) Simon, L. M.; Laszlo, K.; Vertesi, A.; Bagi, K.; Szajani, B. *J. Mol. Catal. B: Enzym.* **1998**, *4*, 41.
(127) Simon, L. M.; Kotorman, M.; Garab, G.; Laczko, I. *Biochem. Biophys. Res. Commun.* **2001**, *280*, 1367.
(128) Wang, H.; Qian, W. J.; Mottaz, H. M.; Clauss, T. R.; Anderson, D. J.; Moore, R. J.; Camp, D. G., 2nd; Khan, A. H.; Sforza, D. M.; Pallavicini, M.; Smith, D. J.; Smith, R. D. *J. Proteome Res.* **2005**, *4*, 2397.
(129) Zhang, H.; Lin, Q.; Ponnusamy, S.; Kothandaraman, N.; Lim, T. K.; Zhao, C.; Kit, H. S.; Arijit, B.; Rauff, M.; Hew, C. L.; Chung, M. C.; Joshi, S. B.; Choolani, M. *Proteomics* **2007**, *7*, 1654.
(130) Deshusses, J. M.; Burgess, J. A.; Scherl, A.; Wenger, Y.; Walter, N.; Converset, V.; Paesano, S.; Corthals, G. L.; Hochstrasser, D. F.; Sanchez, J. C. *Proteomics* **2003**, *3*, 1418.
(131) Ru, Q. C.; Zhu, L. A.; Katenhusen, R. A.; Silberman, J.; Brzeski, H.; Liebman, M.; Shriver, C. D. *J. Chromatogr., A* **2006**, *1111*, 175.
(132) Blonder, J.; Hale, M. L.; Chan, K. C.; Yu, L. R.; Lucas, D. A.; Conrads, T. P.; Zhou, M.; Popoff, M. R.; Issaq, H. J.; Stiles, B. G.; Veenstra, T. D. *J. Proteome Res.* **2005**, *4*, 523.
(133) Blonder, J.; Hale, M. L.; Lucas, D. A.; Schaefer, C. F.; Yu, L. R.; Conrads, T. P.; Issaq, H. J.; Stiles, B. G.; Veenstra, T. D. *Electrophoresis* **2004**, *25*, 1307.
(134) Blonder, J.; Yu, L. R.; Radeva, G.; Chan, K. C.; Lucas, D. A.; Waybright, T. J.; Issaq, H. J.; Sharom, F. J.; Veenstra, T. D. *J. Proteome Res.* **2006**, *5*, 349.
(135) Patra, M.; Salonen, E.; Terama, E.; Vattulainen, I.; Faller, R.; Lee, B. W.; Holopainen, J.; Karttunen, M. *Biophys. J.* **2006**, *90*, 1121.
(136) Pinisetty, D.; Moldovan, D.; Devireddy, R. *Ann. Biomed. Eng.* **2006**, *34*, 1442.
(137) Ly, H. V.; Longo, M. L. *Biophys. J.* **2004**, *87*, 1013.
(138) Martosella, J.; Zolotarjova, N.; Liu, H.; Moyer, S. C.; Perkins, P. D.; Boyes, B. E. *J. Proteome Res.* **2006**, *5*, 1301.
(139) Prinz, T.; Muller, J.; Kuhn, K.; Schafer, J.; Thompson, A.; Schwarz, J.; Hamon, C. *J. Proteome Res.* **2004**, *3*, 1073.
(140) Kuhn, K.; Thompson, A.; Prinz, T.; Muller, J.; Baumann, C.; Schmidt, G.; Neumann, T.; Hamon, C. *J. Proteome Res.* **2003**, *2*, 598.
(141) 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.
(142) Unlu, M.; Morgan, M. E.; Minden, J. S. *Electrophoresis* **1997**, *18*, 2071.
(143) Ahram, M.; Springer, D. L. *Expert Rev. Proteomics* **2004**, *1*, 293.
(144) Rabilloud, T.; Luche, S.; Santoni, V.; Chevallet, M. *Methods Mol. Biol.* **2007**, *355*, 111.
(145) Aivaliotis, M.; Haase, W.; Karas, M.; Tsiotis, G. *Proteomics* **2006**, *6*, 217.
(146) Moller, S.; Croning, M. D.; Apweiler, R. *Bioinformatics* **2001**, *17*, 646.
(147) Schagger, H.; von Jagow, G. *Anal. Biochem.* **1991**, *199*, 223.
(148) Schagger, H.; Cramer, W. A.; Vonjagow, G. *Anal. Biochem.* **1994**, *217*, 220.
(149) Wittig, I.; Braun, H. P.; Schagger, H. *Nat. Protoc.* **2006**, *1*, 418.
(150) Lin, S. X.; Gangloff, A.; Huang, Y. W.; Xie, B. X. *Anal. Chim. Acta* **1999**, *383*, 101.
(151) Wittig, I.; Karas, M.; Schagger, H. *Mol. Cell Proteomics*, in press.
(152) Krause, F. *Electrophoresis* **2006**, *27*, 2759.
(153) Stenberg, F.; Chovanec, P.; Maslen, S. L.; Robinson, C. V.; Ilag, L. L.; von Heijne, G.; Daley, D. O. *J. Biol. Chem.* **2005**, *280*, 34409.
(154) Babusiak, M.; Man, P.; Petrak, J.; Vyoral, D. *Proteomics* **2007**, *7*, 121.
(155) Granvogl, B.; Reisinger, V.; Eichacker, L. A. *Proteomics* **2006**, *6*, 3681.
(156) Bunai, K.; Yamane, K. *J. Chromatogr., B* **2005**, *815*, 227.
(157) Macfarlane, D. E. *Anal. Biochem.* **1989**, *176*, 457.
(158) Hartinger, J.; Stenius, K.; Hogemann, D.; Jahn, R. *Anal. Biochem.* **1996**, *240*, 126.
(159) Zahedi, R. P.; Meisinger, C.; Sickmann, A. *Proteomics* **2005**, *5*, 3581.
(160) Rais, I.; Karas, M.; Schagger, H. *Proteomics* **2004**, *4*, 2567.
(161) Zhao, Y.; Zhang, W.; Kho, Y.; Zhao, Y. *Anal. Chem.* **2004**, *76*, 1817.
(162) Park, Y. M.; Kim, J. Y.; Kwon, K. H.; Lee, S. K.; Kim, Y. H.; Kim, S. Y.; Park, G. W.; Lee, J. H.; Lee, B.; Yoo, J. S. *Proteomics* **2006**, *6*, 4978.
(163) Brouillard, F.; Bensalem, N.; Hinzpeter, A.; Tondelier, D.; Trudel, S.; Gruber, A. D.; Ollero, M.; Edelman, A. *Mol. Cell. Proteomics* **2005**, *4*, 1762.
(164) Righetti, P. G.; Castagna, A.; Antonioli, P.; Boschetti, E. *Electrophoresis* **2005**, *26*, 297.
(165) Righetti, P. G.; Castagna, A.; Herbert, B.; Candiano, G. *Biosci. Rep.* **2005**, *25*, 3.
(166) Herbert, B.; Righetti, P. G. *Electrophoresis* **2000**, *21*, 3639.
(167) Ho, E.; Hayen, A.; Wilkins, M. R. *Proteomics* **2006**, *6*, 5746.
(168) Moritz, R. L.; Simpson, R. J. *Nat. Methods* **2005**, *2*, 863.
(169) Weber, G.; Islinger, M.; Weber, P.; Eckerskorn, C.; Volkl, A. *Electrophoresis* **2004**, *25*, 1735.
(170) Peirce, M. J.; Wait, R.; Begum, S.; Saklatvala, J.; Cope, A. P. *Mol. Cell. Proteomics* **2004**, *3*, 56.
(171) Swanson, S. K.; Washburn, M. P. *Drug Discovery Today* **2005**, *10*, 719.
(172) Schluesener, D.; Rogner, M.; Poetsch, A. *Anal. Bioanal. Chem.*, in press.
(173) Delom, F.; Szponarski, W.; Sommerer, N.; Boyer, J. C.; Bruneau, J. M.; Rossignol, M.; Gibrat, R. *Proteomics* **2006**, *6*, 3029.
(174) Bagshaw, R. D.; Mahuran, D. J.; Callahan, J. W. *Mol. Cell. Proteomics* **2005**, *4*, 133.
(175) McDonald, T.; Sheng, S.; Stanley, B.; Chen, D.; Ko, Y.; Cole, R. N.; Pedersen, P.; Van Eyk, J. E. *Mol. Cell. Proteomics* **2006**, *5*, 2392.
(176) Ye, M. L.; Jiang, X. G.; Feng, S.; Tian, R. J.; Zou, H. F. *Trends Anal. Chem.* **2007**, *26*, 80.
(177) Evans, C. R.; Jorgenson, J. W. *Anal. Bioanal. Chem.* **2004**, *378*, 1952.
(178) Josic, D.; Clifton, J. G. *J. Chromatogr., A* **2007**, *1144*, 2.
(179) Issaq, H. J.; Chan, K. C.; Janini, G. M.; Conrads, T. P.; Veenstra, T. D. *J. Chromatogr., B* **2005**, *817*, 35.
(180) Blackler, A. R.; Klammer, A. A.; MacCoss, M. J.; Wu, C. C. *Anal. Chem.* **2006**, *78*, 1337.
(181) Elias, J. E.; Gygi, S. P. *Nat. Methods* **2007**, *4*, 207.
(182) Stapels, M. D.; Barofsky, D. F. *Anal. Chem.* **2004**, *76*, 5423.
(183) Li, N.; Shaw, A. R.; Zhang, N.; Mak, A.; Li, L. *Proteomics* **2004**, *4*, 3156.
(184) Ji, C.; Li, L. *J. Proteome Res.* **2005**, *4*, 734.
(185) Zhang, N.; Li, N.; Li, L. *J. Proteome Res.* **2004**, *3*, 719.
(186) Liu, H.; Sadygov, R. G.; Yates, J. R., III. *Anal. Chem.* **2004**, *76*, 4193.
(187) Blonder, J.; Goshe, M. B.; Xiao, W.; Camp, D. G., 2nd; Wingerd, M.; Davis, R. W.; Smith, R. D. *J. Proteome Res.* **2004**, *3*, 434.
(188) Blonder, J.; Goshe, M. B.; Moore, R. J.; Pasa-Tolic, L.; Masselon, C. D.; Lipton, M. S.; Smith, R. D. *J. Proteome Res.* **2002**, *1*, 351.
(189) Karsan, A.; Blonder, J.; Law, J.; Yaquian, E.; Lucas, D. A.; Conrads, T. P.; Veenstra, T. *J. Proteome Res.* **2005**, *4*, 349.
(190) Chen, J.; Balgley, B. M.; DeVoe, D. L.; Lee, C. S. *Anal. Chem.* **2003**, *75*, 3145.
(191) Guo, T.; Lee, C. S.; Wang, W.; DeVoe, D. L.; Balgley, B. M. *Electrophoresis* **2006**, *27*, 3523.
(192) Wang, W.; Guo, T.; Rudnick, P. A.; Song, T.; Li, J.; Zhuang, Z.; Zheng, W.; Devoe, D. L.; Lee, C. S.; Balgley, B. M. *Anal. Chem.* **2007**, *79*, 1002.
(193) Wang, W.; Guo, T.; Song, T.; Lee, C. S.; Balgley, B. M. *Proteomics* **2007**, *7*, 1178.
(194) Guo, T.; Wang, W.; Rudnick, P. A.; Song, T.; Li, J.; Zhuang, Z.; Weil, R. J.; Devoe, D. L.; Lee, C. S.; Balgley, B. M. *J. Histochem. Cytochem.* **2007**, *55*, 763.
(195) Nielsen, P. A.; Olsen, J. V.; Podtelejnikov, A. V.; Andersen, J. R.; Mann, M.; Wisniewski, J. R. *Mol. Cell. Proteomics* **2005**, *4*, 402.
(196) Wu, C. C.; MacCoss, M. J.; Howell, K. E.; Yates, J. R., III. *Nat. Biotechnol.* **2003**, *21*, 532.
(197) 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.
(198) Zybailov, B.; Mosley, A. L.; Sardiu, M. E.; Coleman, M. K.; Florens, L.; Washburn, M. P. *J. Proteome Res.* **2006**, *5*, 2339.
(199) McCarthy, F. M.; Burgess, S. C.; van den Berg, B. H.; Koter, M. D.; Pharr, G. T. *J. Proteome Res.* **2005**, *4*, 316.
(200) Bledi, Y.; Inberg, A.; Linial, M. *Briefings Funct. Genomics Proteomics* **2003**, *2*, 254.
(201) Olsen, J. V.; Andersen, J. R.; Nielsen, P. A.; Nielsen, M. L.; Figeys, D.; Mann, M.; Wisniewski, J. R. *Mol. Cell. Proteomics* **2004**, *3*, 82.
(202) Rodriguez-Ortega, M. J.; Norais, N.; Bensi, G.; Liberatori, S.; Capo, S.; Mora, M.; Scarselli, M.; Doro, F.; Ferrari, G.; Garaguso, I.; Maggi, T.; Neumann, A.; Covre, A.; Telford, J. L.; Grandi, G. *Nat. Biotechnol.* **2006**, *24*, 191.
(203) Sickmann, A.; Reinders, J.; Wagner, Y.; Joppich, C.; Zahedi, R.; Meyer, H. E.; Schonfisch, B.; Perschil, I.; Chacinska, A.; Guiard, B.; Rehling, P.; Pfanner, N.; Meisinger, C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13207.
(204) Budnik, B. A.; Lee, R. S.; Steen, J. A. *Biochim. Biophys. Acta* **2006**, *1764*, 1870.
(205) Chen, C. P.; Rost, B. *Protein Sci.* **2002**, *11*, 2766.
(206) Ulmschneider, M. B.; Sansom, M. S.; Di Nola, A. *Proteins* **2005**, *59*, 252.
(207) Sadka, T.; Linial, M. *Bioinformatics* **2005**, *21 Suppl 1*, i378.
(208) Blonder, J.; Conrads, T. P.; Yu, L. R.; Terunuma, A.; Janini, G. M.; Issaq, H. J.; Vogel, J. C.; Veenstra, T. D. *Proteomics* **2004**, *4*, 31.
(209) Goshe, M. B.; Blonder, J.; Smith, R. D. *J. Proteome Res.* **2003**, *2*, 153.
(210) Blonder, J.; Conrads, T. P.; Veenstra, T. D. *Expert Rev. Proteomics* **2004**, *1*, 153.

(211) Wei, C.; Yang, J.; Zhu, J.; Zhang, X.; Leng, W.; Wang, J.; Xue, Y.; Sun, L.; Li, W.; Wang, J.; Jin, Q. *J. Proteome Res.* **2006**, *5*, 1860.
(212) Fischer, F.; Poetsch, A. *Proteome Sci.* **2006**, *4*, 2.
(213) Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364.
(214) Schartz, N. E.; Chaput, N.; Andre, F.; Zitvogel, L. *Curr. Opin. Mol. Ther.* **2002**, *4*, 372.
(215) Roth, J. *Chem. Rev.* **2002**, *102*, 285.
(216) Bause, E. *Biochem. J.* **1983**, *209*, 331.
(217) Springer, D. L.; Auberry, D. L.; Ahram, M.; Adkins, J. N.; Feldhaus, J. M.; Wahl, J. H.; Wunschel, D. S.; Rodland, K. D. *Dis. Markers* **2003**−**2004**, *19*, 219.
(218) Bond, M. R.; Kohler, J. J. *Curr. Opin. Chem. Biol.* **2007**, *11*, 52.
(219) Ghosh, D.; Krokhin, O.; Antonovici, M.; Ens, W.; Standing, K. G.; Beavis, R. C.; Wilkins, J. A. *J. Proteome Res.* **2004**, *3*, 841.
(220) Fan, X.; She, Y. M.; Bagshaw, R. D.; Callahan, J. W.; Schachter, H.; Mahuran, D. J. *Anal. Biochem.* **2004**, *332*, 178.
(221) Fan, X.; She, Y. M.; Bagshaw, R. D.; Callahan, J. W.; Schachter, H.; Mahuran, D. J. *Glycobiology* **2005**, *15*, 952.
(222) Atwood, J. A., III; Minning, T.; Ludolf, F.; Nuccio, A.; Weatherly, D. B.; Alvarez-Manilla, G.; Tarleton, R.; Orlando, R. *J. Proteome Res.* **2006**, *5*, 3376.
(223) Zhang, H.; Li, X. J.; Martin, D. B.; Aebersold, R. *Nat. Biotechnol.* **2003**, *21*, 660.
(224) Nicholson, I. C.; Ayhan, M.; Hoogenraad, N. J.; Zola, H. *J. Leukocyte Biol.* **2005**, *77*, 190.
(225) Nunomura, K.; Nagano, K.; Itagaki, C.; Taoka, M.; Okamura, N.; Yamauchi, Y.; Sugano, S.; Takahashi, N.; Izumi, T.; Isobe, T. *Mol. Cell. Proteomics* **2005**, *4*, 1968.
(226) Zhao, Y.; Zhang, W.; White, M. A.; Zhao, Y. *Anal. Chem.* **2003**, *75*, 3751.
(227) Sostaric, E.; Georgiou, A. S.; Wong, C. H.; Watson, P. F.; Holt, W. V.; Fazeli, A. *J. Proteome Res.* **2006**, *5*, 3029.
(228) Tang, X.; Yi, W.; Munske, G. R.; Adhikari, D. P.; Zakharova, N. L.; Bruce, J. E. *J. Proteome Res.* **2007**, *6*, 724.
(229) Pasquali, C.; Fialka, I.; Huber, L. A. *J. Chromatogr., B* **1999**, *722*, 89.
(230) Sabarth, N.; Lamer, S.; Zimny-Arndt, U.; Jungblut, P. R.; Meyer, T. F.; Bumann, D. *J. Biol. Chem.* **2002**, *277*, 27896.
(231) Burre, J.; Beckhaus, T.; Corvey, C.; Karas, M.; Zimmermann, H.; Volknandt, W. *Electrophoresis* **2006**, *27*, 3488.
(232) Bisle, B.; Schmidt, A.; Scheibe, B.; Klein, C.; Tebbe, A.; Kellermann, J.; Siedler, F.; Pfeiffer, F.; Lottspeich, F.; Oesterhelt, D. *Mol. Cell. Proteomics* **2006**, *5*, 1543.
(233) Chen, X.; Sun, L.; Yu, Y.; Xue, Y.; Yang, P. *Expert Rev. Proteomics* **2007**, *4*, 25.
(234) Ong, S. E.; Mann, M. *Nat. Chem. Biol.* **2005**, *1*, 252.
(235) Yan, W.; Chen, S. S. *Briefings Funct. Genomics Proteomics* **2005**, *4*, 27.
(236) 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.
(237) Scheurer, S. B.; Rybak, J. N.; Roesli, C.; Brunisholz, R. A.; Potthast, F.; Schlapbach, R.; Neri, D.; Elia, G. *Proteomics* **2005**, *5*, 2718.
(238) Foster, L. J.; Zeemann, P. A.; Li, C.; Mann, M.; Jensen, O. N.; Kassem, M. *Stem Cells* **2005**, *23*, 1367.
(239) Liang, X.; Zhao, J.; Hajivandi, M.; Wu, R.; Tao, J.; Amshey, J. W.; Pope, R. M. *J. Proteome Res.* **2006**, *5*, 2632.
(240) Foster, L. J.; De Hoog, C. L.; Mann, M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5813.
(241) McClatchy, D. B.; Dong, M. Q.; Wu, C. C.; Venable, J. D.; Yates, J. R., III. *J. Proteome Res.* **2007**, *6*, 2005.
(242) Wu, C. C.; MacCoss, M. J.; Howell, K. E.; Matthews, D. E.; Yates, J. R., III. *Anal. Chem.* **2004**, *76*, 4951.
(243) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. *Mol. Cell. Proteomics* **2004**, *3*, 1154.
(244) Guo, Y.; Singleton, P. A.; Rowshan, A.; Gucek, M.; Cole, R. N.; Graham, D. R.; Van Eyk, J. E.; Garcia, J. G. *Mol. Cell. Proteomics* **2007**, *6*, 689.
(245) Radosevich, T. J.; Reinhardt, T. A.; Lippolis, J. D.; Bannantine, J. P.; Stabel, J. R. *J. Bacteriol.* **2007**, *189*, 1109.
(246) Lund, T. C.; Anderson, L. B.; McCullar, V.; Higgins, L.; Yun, G. H.; Grzywacz, B.; Verneris, M. R.; Miller, J. S. *J. Proteome Res.* **2007**, *6*, 644.
(247) Chen, X.; Walker, A. K.; Strahler, J. R.; Simon, E. S.; Tomanicek-Volk, S. L.; Nelson, B. B.; Hurley, M. C.; Ernst, S. A.; Williams, J. A.; Andrews, P. C. *Mol. Cell. Proteomics* **2006**, *5*, 306.
(248) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994.
(249) Prokai, L.; Zharikova, A. D.; Stevens, S. M., Jr. *J. Mass Spectrom.* **2005**, *40*, 169.
(250) Han, D. K.; Eng, J.; Zhou, H.; Aebersold, R. *Nat. Biotechnol.* **2001**, *19*, 946.
(251) Smolka, M. B.; Zhou, H.; Purkayastha, S.; Aebersold, R. *Anal. Biochem.* **2001**, *297*, 25.
(252) Ramus, C.; Gonzalez de Peredo, A.; Dahout, C.; Gallagher, M.; Garin, J. *Mol. Cell. Proteomics* **2006**, *5*, 68.
(253) Olsen, J. V.; Nielsen, P. A.; Andersen, J. R.; Mann, M.; Wisniewski, J. R. *Brain Res.* **2007**, *1134*, 95.
(254) Schmidt, A.; Kellermann, J.; Lottspeich, F. *Proteomics* **2005**, *5*, 4.
(255) Miyagi, M.; Rao, K. C. *Mass Spectrom. Rev.* **2007**, *26*, 121.
(256) Schnolzer, M.; Jedrzejewski, P.; Lehmann, W. D. *Electrophoresis* **1996**, *17*, 945.
(257) Cech, N. B.; Enke, C. G. *Anal. Chem.* **2000**, *72*, 2717.
(258) Cech, N. B.; Enke, C. G. *Mass Spectrom. Rev.* **2001**, *20*, 362.
(259) Cech, N. B.; Krone, J. R.; Enke, C. G. *Anal. Chem.* **2001**, *73*, 208.
(260) Zhu, Y. F.; Lee, K. L.; Tang, K.; Allman, S. L.; Taranencko, N. I.; Chen, C. H. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1315.
(261) Krause, E.; Wenschuh, H.; Jungblut, P. R. *Anal. Chem.* **1999**, *71*, 4160.
(262) Olumee, Z.; Sadeghi, M.; Tang, X. D.; Vertes, A. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 744.
(263) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365.
(264) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35*, 1399.
(265) Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Garbuzynskiy, S. O.; Galzitskaya, O. V.; Surin, A. K.; Zubarev, R. A. *Angew. Chem., Int. Ed.* **2007**, *46*, 1481.
(266) Tabb, D. L.; Huang, Y.; Wysocki, V. H.; Yates, J. R., III. *Anal. Chem.* **2004**, *76*, 1243.
(267) Tabb, D. L.; Smith, L. L.; Breci, L. A.; Wysocki, V. H.; Lin, D.; Yates, J. R., III. *Anal. Chem.* **2003**, *75*, 1155.
(268) Sudha, R.; Panda, M.; Chandrasekhar, J.; Balaram, P. *Chemistry* **2002**, *8*, 4980.
(269) Breci, L. A.; Tabb, D. L.; Yates, J. R., III; Wysocki, V. H. *Anal. Chem.* **2003**, *75*, 1963.
(270) Huang, Y.; Triscari, J. M.; Tseng, G. C.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *Anal. Chem.* **2005**, *77*, 5800.
(271) Nair, H.; Wysocki, V. H. *Int. J. Mass Spectrom.* **1998**, *174*, 95.
(272) Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W. Q.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 5142.
(273) Shaffer, S. A.; Scherl, A.; Gallis, B.; Aker, J.; Taylor, G.; Brittnacher, M.; Goodlett, D. R. Identifying Peptides with Higher Order Charge-States: An Optimized Approach for Characterizing Membrane Proteins. In 55th ASMS Conference on Mass Spectrometry, Indianapolis, IN, June 3−7, 2007.
(274) Punta, M.; Forrest, L. R.; Bigelow, H.; Kernytsky, A.; Liu, J.; Rost, B. *Methods* **2007**, *41*, 460.
(275) von Heijne, G.; Gavel, Y. *Eur. J. Biochem.* **1988**, *174*, 671.
(276) Nilsson, J.; Persson, B.; von Heijne, G. *Proteins* **2005**, *60*, 606.
(277) Bowie, J. U. *Protein Sci.* **1999**, *8*, 2711.
(278) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
(279) White, S. H. *Protein Sci.* **2004**, *13*, 1948.
(280) Riek, R. P.; Rigoutsos, I.; Novotny, J.; Graham, R. M. *J. Mol. Biol.* **2001**, *306*, 349.
(281) Yeagle, P. L.; Bennett, M.; Lemaitre, V.; Watts, A. *Biochem. Biophys. Acta* **2007**, *1768*, 530.
(282) Screpanti, E.; Hunte, C. *J. Struct. Biol.*, in press.
(283) Cuthbertson, J. M.; Doyle, D. A.; Sansom, M. S. *Protein Eng., Des. Sel.* **2005**, *18*, 295.
(284) Viklund, H.; Granseth, E.; Elofsson, A. *J. Mol. Biol.* **2006**, *361*, 591.
(285) Rapp, M.; Granseth, E.; Seppala, S.; von Heijne, G. *Nat. Struct. Mol. Biol.* **2006**, *13*, 112.
(286) Gafvelin, G.; von Heijne, G. *Cell* **1994**, *77*, 401.
(287) von Heijne, G. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 909.
(288) Fleishman, S. J.; Unger, V. M.; Ben-Tal, N. *Trends Biochem. Sci.* **2006**, *31*, 106.
(289) von Heijne, G. *J. Mol. Biol.* **1992**, *225*, 487.
(290) Jones, D. T.; Taylor, W. R.; Thornton, J. M. *Biochemistry* **1994**, *33*, 3038.
(291) Rost, B.; Fariselli, P.; Casadio, R. *Protein Sci.* **1996**, *5*, 1704.
(292) Nilsson, J.; Persson, B.; Von Heijne, G. *Protein Sci.* **2002**, *11*, 2974.
(293) Melen, K.; Krogh, A.; von Heijne, G. *J. Mol. Biol.* **2003**, *327*, 735.
(294) Juretic, D.; Zoranic, L.; Zucic, D. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 620.
(295) Persson, B.; Argos, P. *J. Protein Chem.* **1997**, *16*, 453.
(296) Sonnhammer, E. L.; von Heijne, G.; Krogh, A. *Proc.*—*Int. Conf. Intell. Syst. Mol. Biol.* **1998**, *6*, 175.
(297) Kall, L.; Krogh, A.; Sonnhammer, E. L. *J. Mol. Biol.* **2004**, *338*, 1027.

(298) Kall, L.; Krogh, A.; Sonnhammer, E. L. *Nucleic Acids Res.*, in press.
(299) Nielsen, H.; Krogh, A. *Proc.—Int. Conf. Intell. Syst. Mol. Biol.* **1998**, *6*, 122.
(300) Hofmann, K.; Stoffel, W. *Biol. Chem. Hoppe-Seyler* **1993**, *374*, 166.
(301) Claros, M. G.; von Heijne, G. *Comput. Appl. Biosci.* **1994**, *10*, 685.
(302) Cserzo, M.; Wallin, E.; Simon, I.; von Heijne, G.; Elofsson, A. *Protein Eng.* **1997**, *10*, 673.
(303) Daley, D. O.; Rapp, M.; Granseth, E.; Melen, K.; Drew, D.; von Heijne, G. *Science* **2005**, *308*, 1321.
(304) Kim, H.; Melen, K.; Osterberg, M.; von Heijne, G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11142.
(305) Granseth, E.; Daley, D. O.; Rapp, M.; Melen, K.; von Heijne, G. *J. Mol. Biol.* **2005**, *352*, 489.
(306) Manoil, C.; Beckwith, J. *Science* **1986**, *233*, 1403.
(307) Feilmeier, B. J.; Iseminger, G.; Schroeder, D.; Webber, H.; Phillips, G. J. *J. Bacteriol.* **2000**, *182*, 4068.
(308) Drew, D.; Sjostrand, D.; Nilsson, J.; Urbig, T.; Chin, C. N.; de Gier, J. W.; von Heijne, G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2690.
(309) Rapp, M.; Drew, D.; Daley, D. O.; Nilsson, J.; Carvalho, T.; Melen, K.; De Gier, J. W.; Von Heijne, G. *Protein Sci.* **2004**, *13*, 937.
(310) Deak, P. M.; Wolf, D. H. *J. Biol. Chem.* **2001**, *276*, 10663.
(311) Kim, H.; Melen, K.; von Heijne, G. *J. Biol. Chem.* **2003**, *278*, 10208.
(312) Davis, T. N. *Curr. Opin. Chem. Biol.* **2004**, *8*, 49.

CR068286Z