Directions in Automated Sample Preparation of Proteins

David A. Wells and David A. Weil

TECAN US

The overall process for proteomics comprises manual, disjointed steps, and there is a real need for improved strategies for the high-throughput analysis of protein expression and function. The authors examine current procedures for the sample preparation of proteins; subsequent protein separation, identification and analysis; and the emerging use of automated procedures in the overall process.

David A. Wells provides education, training and consulting services through Sample Prep Solutions Co., 672 Huntington Court, St. Paul, Minnesota 55119 USA; e-mail dawells@sampleprep.com. **David A. Weil** is a senior research specialist at 3M's Corporate Analytical Technology Center, 201-BW-09, St. Paul, Minnesota 55144 USA; e-mail daweil@mmm.com.

The study of genomics and proteomics is resulting in a greater understanding of the relationships and interactions among genes and proteins in living organisms. The initial excitement from these disciplines was that scientists were discovering more-selective drug targets and biomarkers. Individual protein identification was found to be of little value toward this goal. Instead, systems biology, in which protein–protein and protein–ligand interactions occur, is beginning to lead to a better understanding of disease progression and to targeted drug therapies. This research is driving demand for fully automated protocols for industrial-scale proteomics. These systems encompass sample preparation, separation, characterization and bioinformatics.

The objectives of this article are to review current protein sample preparation procedures with existing deficiencies and examine future trends toward automating and miniaturizing this overall process.

Proteomics: An Overview

A rapidly evolving approach for drug discovery focuses on the fundamental biological processes occurring within an organism, specifically the interaction between a genome and its proteome. The term proteome refers to the full complement of proteins expressed by the genome (1). The human genome sequence recently has been elucidated (2, 3), as has that of many other organisms. The interactions among the genes, RNA molecules and proteins in each cell of a functioning organism are highly intricate. Unlike the Human Genome Project, in which the genome proved to be static, the proteome of a cell or tissue is highly dynamic and constantly changes with respect to its surrounding environment, physiological state, stress, drug administration, health or disease.

The Human Proteome Project faces the challenge of describing all the proteins expressed by the genome and their levels throughout body tissues under various environmental conditions of normal, stressed and diseased states. Scientists estimate that although the human genome contains 20,000–30,000 genes, the number of proteins could be 50,000–500,000 (4). These greater numbers of proteins are due to the many post-translational modifications — such as phosphorylation, acetylation, sulfation and glycosylation — which result in different protein reactivities. However, simply cataloging proteins is insufficient because it does not define protein–protein interac-
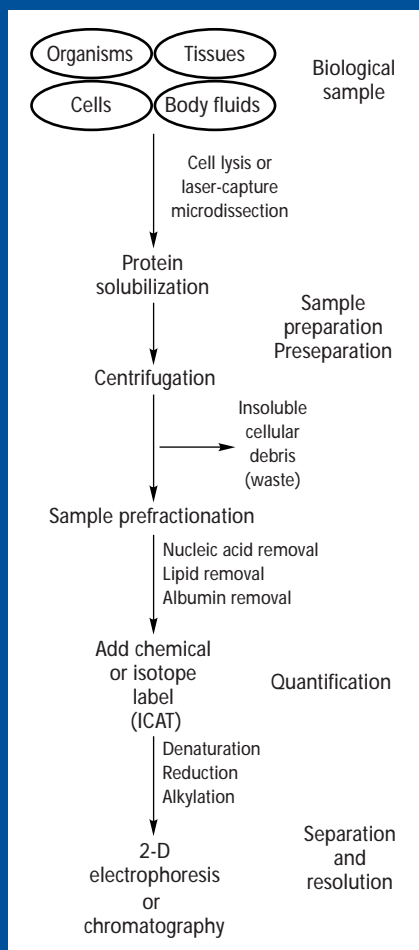
Figure 1. Diagram of the typical series of sample preparation steps required for protein purification before separation.

tions or structure–function relationships in different cell types.

Proteomics has many more technical challenges than does genomics. Although polymerase chain reaction (PCR) technology works for DNA, no similar method is available to amplify proteins. Therefore, analytical issues — such as sample handling, separations of pico- to femtomole amounts of proteins in complex biological matrices and isolation of the ultra low-level significant proteins — must be solved. The high sensitivity of the mass spectrometer has enabled the detection of ultra low levels of proteins, but sample preparation remains vital to remove proteins of high abundance (for example, IgG, human serum albumin and protein A). Additional tools required for proteomics include protein separation using gels and multidimensional liquid chromatography (LC) with detection by mass spectrometry (MS), and the use of protein databases

and bioinformatics. The industrialization of proteomics requires alliances among manufacturers of systems and technologies that link fluid transport, consumables, digestions, robotics, detection and data analysis through a common software platform for the combined processes to be automatable.

Today, the overall process for proteomics generally comprises manual and disjointed steps for sample preparation, protein separation and characterization. Scientists have a real need to develop and implement improved strategies for the high-throughput analysis of protein expression and function to quickly realize the tremendous medical advances that are envisioned through successful use of the information gained from proteomics. Here, we will examine current procedures for the sample preparation of proteins; the subsequent steps of protein separation, identification and analysis; and the use of and need for automated procedures as part of the overall process.

Classic Protein Sample Preparation Techniques

Sample preparation can be used to isolate proteins of interest from biological cells or organisms before analysis. It is an important step toward achieving accurate, reproducible and meaningful results. Proteins isolated from these sources contain contaminants such as keratin, albumin, serum proteins, nucleic acids, lipids, carbohydrates and polysaccharides that are present naturally. In addition, various inorganic salts, buffers, reducing agents, surfactants, detergents and preservatives can be added to a sample to retain enzymatic or biological activity. The presence of these extraneous materials can cause problems such as smearing, masking and poor reproducibility in electrophoresis separations. These materials also affect performance when using LC techniques with subsequent detection or quantitation by MS.

No universal sample preparation procedure exists to isolate all proteins in a mixture because proteins are present in multiple forms, are found within different cell locations (for example, membrane or cytoplasm) and have varying solubilities. Common protein sample preparation processes include desalting, concentration, centrifugation, dialysis, filtration and ul-

trafiltration, precipitation and lyophilization. Distinctive characteristics of the protein — such as isoelectric point, molecular weight, shape, solubility and hydrophobicity — guide the design of this purification. Because proteins are very fragile, analysts must take care during sample preparation to avoid the introduction of unwanted modifications that can change conformation and biological activity.

Several sample preparation processes can be combined to develop a purification strategy (Figure 1). Typically, a cell lysis procedure is performed by sonication, enzyme treatment or mechanical means to release the proteins. Subsequently, analysts can use a solubility scheme to yield a representative protein sample. Nucleic acids commonly are removed by precipitation techniques or by sonication, which breaks them into smaller fragments. Lipids are removed with excess detergent or with precipitation. Each cell or tissue type requires a specific methodology. Link (6) has compiled the protocols used to isolate protein from mouse and human tissue, body fluids and microorganisms. After the initial purification into a soluble crude sample, a procedure known as prefractionation can enrich and purify specific protein components before separation and analysis. This technique uses various chromatographic and electrophoretic methods (7).

An important part of the sample preparation process is to remove extraneous proteins present in high abundance, such as albumin. Albumin is the major acidic protein component in serum, and its removal can enhance the sensitivity of assay techniques and improve the effectiveness and binding capacity of affinity purification media. Various commercially available kits contain reagents, supplies and protocols for performing sample preparation processes such as the removal of albumin, DNA, salts and buffers, as well as the isolation and solubilization of specific protein groups (for example, membrane proteins, acidic and basic proteins and nuclear proteins). Table I lists typical examples of these kits. Earlier albumin removal kits used membranes and columns that contained resins with various immobilized textile dyes, and their efficiency was rated at only approximately 70%. Newer generation albumin removal kits are based
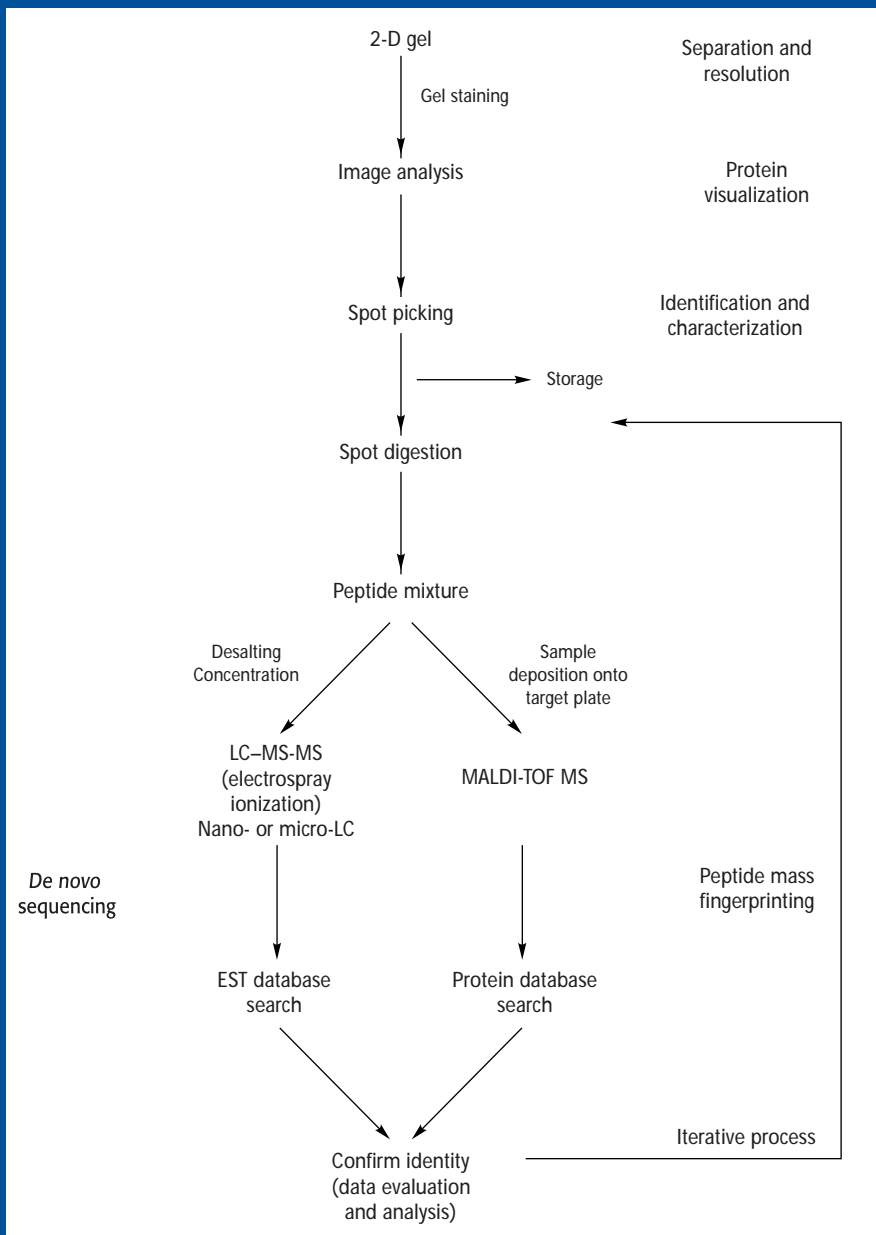
Sample Preparation

2-D gel	Separation and resolution
↓ Gel staining	
Image analysis	Protein visualization
↓	
Spot picking	Identification and characterization
↓ → Storage	
Spot digestion	
↓	
Peptide mixture	

Desalting Concentration / Sample deposition onto target plate

LC–MS-MS (electrospray ionization) Nano- or micro-LC   **MALDI-TOF MS**

De novo sequencing   Peptide mass fingerprinting

EST database search   **Protein database search**

Confirm identity (data evaluation and analysis)   Iterative process

Figure 2. Diagram of the typical procedures for the analysis and identification of proteins by MS.

tial sequence analysis (after digestion). This identification would be impossible without databases such as expressed sequence tags, protein and peptide sequences, and powerful data mining algorithms (4).

Sample Preparation for Proteomics

Proteomics research actually is separated into three areas: functional, structural and expression proteomics, each of which has its own set of sample preparation protocols. Expression proteomics currently is the most active and recognized area of proteomics in drug discovery and pharmaceutical research. It is defined as the identification and quantitation of proteins present in healthy and diseased tissues. Functional proteomics analyzes nondenatured proteins under conditions that keep protein complexes together (11,12), and structural proteomics aims to elucidate the structures of the functional active sites of each human protein. The ultimate goal is to develop drugs that are highly specific for these active sites (4).

The tremendous analytical challenge of proteomics arises from the fact that the proteome is a collection of 30–80% of gene products expressed at both low levels (10–100 copies per cell) and high levels (10,000–100,000 copies per cell). These numbers represent a dynamic range of at least six orders of magnitude. As a point of reference, most eukaryotic cells contain approximately 20,000 proteins that have an average molecular weight of 50 kDa. Enzymatic digestion yields approximately 30 peptides per protein, or roughly 6,000,000 unique peptides. Certainly, these numbers present colossal technical challenges in terms of analytical sample throughput, detection and data analysis.

The sample preparation techniques of greatest interest in expression proteomics focus on prefractionating and enriching proteins before their separation by preparative electrophoresis or chromatography. Detection is either by matrix-assisted laser desorption ionization (MALDI) or atmospheric-pressure ionization (API) MS. Figure 2 shows a typical workflow diagram of the total process for identifying differential protein expression levels.

The classic methods for extracting and isolating proteins as used in protein chem-

upon polyclonal antibodies, and these kits approach 95% removal. Analysts must take care to avoid removing low-abundance proteins, and the newer selective removal kits seem to do a good job. After proteins are purified, some of the common separation methods are single- and two-dimensional (2-D) electrophoresis, LC, size-exclusion chromatography, gel filtration and affinity chromatography.

Classic protein chemistry typically involves the study of one protein at a time. It is beyond the scope of this article to describe the protein purification procedures used for preparing an extract for enzyme purification or in studies of enzyme ac-

tivity, protein sequencing or structure analysis. These classic laboratory procedures are discussed in the literature (6, 8–10). In contrast to classic protein chemistry, which studies a single component, the focus of proteomics is on the interaction of multiple, distinct proteins and their roles within the larger system. The major difference between classic protein chemistry and proteomics is that the interest in classic protein chemistry is focused on isolating an individual protein and identifying its total sequence, whereas the goal in proteomics is to characterize a complex mixture of proteins present in various expressed levels and identify them using par-

Sample Preparation

istry also are used in proteomics. In addition, newer methods such as laser-capture microdissection increase the differentiation between diseased and healthy tissues (13). Labeling with stable isotopes (for example, ^{13}C, ^{2}H and ^{15}N) (14) or with isotope -coded affinity tags (ICATs) commonly is used in affinity-capture procedures for differentially expressed proteins at low abundance (15). A kit based on the ICATs technology is available from Applied Biosystems (Foster City, California, USA).

Two-dimensional electrophoresis. Two-dimensional electrophoresis is used to separate proteins in the first dimension by iso-electric focusing based upon isoelectric point and in the second dimension by molecular weight using sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS–PAGE). Visualization of the proteins separated in the gels is performed using inorganic and organic dyes. High-resolution 2-D imaging and analysis software programs summarize the results and help to identify target proteins for excision and additional study.

Several precautions go along with using 2-D gels. One important point is that most biologically significant proteins are present at low levels and, therefore, generally are not detected (16). Efficient sample preparation techniques can remove most highly abundant interfering proteins and improve reproducibility in 2-D electrophoresis separations. Prefractionation according to a physicochemical parameter such as iso-electric point can be performed to enable a higher loading capacity onto gels and minimize protein–protein interactions (17). It also is important to remove insoluble materials from the sample before running

Table I. Typical examples of sample preparation and automation products

Company	Pre-2-D gel kits
Agilent Technologies	Immunodepletion column kits
Amersham Biosciences	Ettan 2-D cleanup kit, Ettan SDS–PAGE cleanup kit, Ettan 2-D Quant, DeStreak rehydration solution kit, PlusOne kits
Applied Biosystems	Affinity Depletion kit, Poros anti-HSA
Beckman Coulter*	
Biacore*	
BioMolecular Technologies*	
Bio-Rad Laboratories	ReadyPrep sequential extraction kit, ReadyPrep reagents
Bruker Daltonics	
Ciphergen Biosystems*	
Dionex*	
Eprogen	
Genomics Solutions	
Glygen	
Gyros	
Hamilton Co.	
Harvard Bioscience	
LC Packings	
Leap Technologies*	
MDS Proteomics*	
MDS Sciex*	ProteoSpin Depletion
Michrom BioResources	
Micro-Tech Scientific*	
Millipore	ZipPlate Micro-SPE plates, Montage Albumin Deplete kit
New Objectives*	
Pall*	Ultrafiltration spin columns and membranes
PerkinElmer Life Sciences	
Pierce Biotechnology	2-D Sample Prep kit for nuclear proteins, 2-D Sample Prep kit for membrane proteins, SwellGel Blue Albumin removing kit, PrepTide peptide isolation kits
ProMetic Life Sciences*	Mimetic Ligand albumin removal kit
Proteome Systems	ProteoPrep Membrane Extraction, ProteoPrep, Universal Extraction, ProteoPrep Sample Extraction, ProteoPrep QuantiPro
Qiagen*	PhosphoProtein purification system
Sequenom*	
Shimadzu Biotech*	
Sigma-Aldrich*	QuantiPro protein concentration, ProteoPrep membrane extraction kit, ProteoPrep universal extraction kit, ProteoPrep protein precipitation kit
Tecan U.S.	
Thermo Finnigan	
Vivascience*	Albumin removal kit, DNA removal kit, acidic protein purification kits, basic protein purification kit
Waters	
Zymark*	

INFORMATION CURRENT AS OF JUNE 2003.

In-Gel Digestion Kits	2-D Gel to MALDI Products	Other Products
		Capillary chromatography, including columns; nanochromatography, including columns; multidimensional chromatography
Ettan 2-D kits	Ettan Spot Picker, Ettan Digester, Ettan 2-D system	
	Proteomics Solution 1 system, SymBiot I system (spot picker)	
	ProteomeLab PA 800 and DU 800 systems	Biomek FX and Biomek2000 liquid handler
		Biacore 3000 Surface Plasmon Resonance MS system
Pi(3) reagent kit ReadyPrep digestion kit	ProteomeWorks system	MassPrep, Clonis Cellular LCM
	MAP II, MAP II/8, Proteineer sp, Proteineer dp	AnchorChip sample targets
		ProteinChip
		Multidimensional chromatography, nanochromatography
		ProteoSep nonporous chromatography, ProteoView Virtual 2D Gel software, protein separations
	Investigator ProPrep, Investigator ProMS, ProPick spot picker, ProGest gel digester	
		NuTip micropipette tip, TopTip micropipette tip, SyringeTip, LC Fibersorbent-impregnated capillary tubing
		Gyrolab MALDI SP1, Gyrolab workstation, Gyrolab cutter
	Microlab Star	Microlab 4000
		Ultra-Micro PrepTip, Micro PrepTip, Ultra-Micro Tip columns, Micro Tip columns, Micro ProTip, Macro ProTip, Ultra-Micro ProTip
		Probot Micro fraction collector, UltiMate micropump/detection, Famos Micro autosampler, Monolyths capillary columns, PreMap Micro columns
	2D iD gel processor	Multidimensional chromatography
	2D iD gel imaging/spot cutting system	
		Magic microfraction collector, Magic autosampler
		Ultra-Plus II MD system, Endurance 96- and 384-well autosampler
Montage in-gel digest kit		ZipTip pipette tips
		PicoFrit columns, IntegraFrit columns, IntegraFrit sample trap columns, Silica tips, Taper tips, Pico tips
	ProXCision gel cutter, MultiProbe II HT EX, MultiProbe II EX	Evolution P(3) system
ProteoProfile in-gel digest kit	Xcise gel digestion system, ProteomIQ system	
	MassArray 20000	
	Xcise gel digestion system	
ProteoProfile in-gel digestion kit		
	ProTeam Digest workstation, Genesis RSP 200	ProTeam Free Flow Electrophoresis system; ProTeam Digest Advanced Digest, ProTeam IEF, ProTeam 2-D, Protein crystallography, Cell maintenance, and Protein purification workstations
	ProteomeX workstation	
		LC–MALDIprep
		Liquid-handling systems

Figure 3. The ProTeam Digest workstation automates sample processing (digestion, desalting and concentration) and spotting onto the MALDI target plate. (Photograph courtesy of Tecan U.S.)

Figure 4. The Ettan Spot Handling workstation represents an integrated 2-D MS platform for proteomics analysis consisting of instruments, chemicals and software. (Photograph courtesy of Amersham Biosciences.)

the gel because hydrophobic membrane proteins or tissue proteins act differently than do other proteins. It is important that the protein be soluble before 2-D electrophoresis separation. Various inorganic and organic buffers must be added to disrupt disulfide bonds and noncovalent interactions (18).

Several companies now produce 2-D gels, sample cleanup kits, reagents, buffers and organic–inorganic dyes to optimize the separation and isolation of targeted proteins. Manufacturers include Agilent Technologies Inc. (Wilmington, Delaware, USA), Amersham Biosciences (Piscataway, New Jersey, USA), Bio-Rad Laboratories (Hercules, California, USA), Millipore Corp. (Bedford, Massachusetts, USA), PerkinElmer Life Sciences Inc. (Boston, Massachusetts, USA), Pierce Biotechnology Inc. (Rockford, Illinois, USA) and Proteome Systems Ltd. (North Ryde, New South Wales, Australia).

Extraction and digestion from gels. In 2-D electrophoresis, most protein is trapped inside the gel and the challenge is to extract it before the analysis step. Enzymatic digestion (for example, using trypsin) of excised protein spots involves placing the small gel pieces into a centrifuge tube and sequentially performing several steps: reduction, alkylation, washing and dehydration. Buffer and enzyme then can be introduced, and the protein can be digested by incubation at 25 °C or 37 °C overnight. Several manufacturers are exploring methods to dramatically reduce the time required for enzymatic digestion by using very small sample volumes with an on-line enzymatic digestion technique (19, 20). After digestion, several time-consuming manual steps to remove the enzyme, buffers and salts must be completed before a sample is ready for introduction into a mass spectrometer. Several manufacturers provide in-gel digestion kits designed for preparing as many as 96 protein samples in 96-well filtration plates. These kits include enzymes, reagents and sample cleanup products that desalt and concentrate protein-digested samples for MS analysis (Table I) (21).

The procedure of excising protein spots from gels followed by enzymatic digestion presents several opportunities to reduce the loss of low-level proteins and increase throughput. Several manufacturers have developed automated robotic and liquid-handling workstations that start with a 2-D gel, excise the targeted protein spots, digest and concentrate them and, in some cases, spot the final samples onto plates for MS analysis. The following sections discuss how automation, online enzymatic digestions and microfluidics are being applied to significantly reduce analysis time.

Detection by MS. The detection method of choice in proteomics is MS using MALDI, nanoelectrospray ionization tandem MS (MS-MS) or reversed-phase microcapillary LC electrospray ionization. Several reviews describe details of the use of MS in proteomics (22–24). The initial applications of MALDI and electrospray ionization for protein detection used purified single proteins and peptides for peptide mass fingerprinting and de novo peptide sequence analysis. However, experiments using biologically significant samples from 2-D gels and cell extracts exposed a weak-

ness in the procedure — an intolerance to detergents, dyes and organic and inorganic buffers. The ultra high sensitivity of these techniques also makes them more sensitive to the presence of abundant concentrations of protein contaminants such as keratin, protein A and albumin. Aside from quantitation, the detection sensitivity and accuracy of mass measurement of proteins and peptides are affected by the presence of these contaminants. Early users of MALDI and electrospray ionization faced the challenge of developing sample preparation products that reproducibly removed salt, detergents, lipids and highly abundant proteins without eliminating biologically significant proteins that were present in low-copy numbers.

Microcolumns and micropipette tips packed with stationary-phase sorbent materials have emerged as cost-effective, easy-to-use products for concentrating and desalting protein and peptide samples before MALDI analysis. These products have very small dead volumes and can be used with ultra small sample quantities. Manufacturers of these micropipette tip products include Glygen Corp. (Columbia, Maryland, USA), Harvard Bioscience/AmiKa (Holliston, Massachusetts, USA) and Millipore.

ZipTip micropipette tips (Millipore) contain small beds of C18, C4, strong cation-exchange or metal chelate chromatography media fixed at the end of the 10-µL pipette tip. Concentrated, purified samples are eluted in a 1–4 µL volume.

Harvard Bioscience/AmiKa and Glygen also developed the PrepTip and NuTip pipette tips, respectively, in which the interior walls are coated with the sample binding material, rather than resin being placed at the base of the pipette tip; thus, sample flows freely through the opening. The range of chemistries in these products has expanded to include several varieties of reversed-phase (C4, C8, C18 and hydrophobic), normal-phase (silica, amino, cyano and hydrophilic), ion-exchange and affinity purification materials that accommodate sample volumes in the ranges of 1–10 µL, 10–100 µL and 100–500 µL. ProTip (Harvard Bioscience/AmiKa) differs from PrepTip in that it contains a small amount of hydrophilic polymer material. This hydrophilic polymer binds proteins present within organic solvent but

does not bind solutions of salts, buffers or detergents. The protein is eluted from the pipette tip using a very small volume of aqueous solution.

PureTip (Harvard Bioscience/AmiKa) and TopTip pipette tips (Glygen) perform a rapid cleanup of biological samples using size-exclusion chromatography. The Zip-Plate plate (Millipore) provides the ZipTip pipette tips in a 96-well format; a protocol combines in-gel digestion with spotting of the purified and concentrated peptides onto a sample target for analysis by MALDI-MS.

Multidimensional LC. Chromatographers seek alternatives to electrophoresis and gel technology because the sheer numbers of proteins to be separated and analyzed demand a high-throughput system with greater preparative capacity and reproducibility. Capillary LC is a natural choice because it is a proven separation technique that can be fully automated and interfaced with tandem MS for on-line detection. This combination also is attractive because the mass spectrometer offers additional resolution by mass-to-charge ratio beyond the separation achieved on-column. Further discussion of protein separations and interfacing of LC with MS is outside the scope of this article, but reviews of the current state of protein separations for proteomics are available (24, 25–27). Multidimensional protein-identification technology (MudPIT) uses two chromatography steps interfaced back-to-back and various configurations; for example, separation on the first-dimension column with fraction collection followed by injection of the collected fractions onto the second-dimension column; directly coupled LC columns and multidimensional LC using column switching (28).

A 2-D LC separation typically involves an ion-exchange, affinity, size-exclusion or chromatofocusing column coupled with a reversed-phase column in either the first or the second dimension. A mixed-bed microcapillary column that contains ion-exchange and reversed-phase moieties on a resin bed also can be useful. These chromatographic approaches have been used in place of 2-D gels for various analysis protocols, such as shotgun analysis of all proteins (250–5000 proteins), analysis of a single class of proteins (10–250 proteins) and analysis of specific protein(s) (1–10 proteins).

The development of on-the-fly acquisition of tandem mass spectra with data-dependent instrument control and chromatographic control has greatly increased throughput of mass spectrometers. Intelligent algorithms provide automated analysis of peptide mass spectra for protein identification (22). In addition, researchers can achieve a comprehensive separation of complex peptide mixtures as well as the resolution of intact proteins. This LC approach has the ability to detect low-copy proteins using nano LC–MS-MS and certain classes of proteins that are not easily observed (for example, large proteins, hydrophobic membrane proteins and very acidic or very basic proteins) on 2-D gels (29). However, the complete resolution of all proteins in a proteome will require the use of more than two dimensions and a prefractionation step. Toward this goal, Michrom BioResources Inc. (Auburn, California, USA) manufactures a multidimensional separations module to provide high-throughput proteomics assays.

Automation of Sample Preparation Processes

The purification of proteins includes manual and disjointed steps as part of a multiple-step process. Automation of these individual procedures is an important goal for proteomics laboratories to meet high-throughput demands. The choices for automation differ in complexity according to the task required.

Figure 2 outlines common options for automation within the overall process, such as the use of laser capture microdissection to isolate healthy cells from diseased ones, excision of targeted proteins from 2-D gels, digestion of these excised gel spots or plugs, desalting and concentration following digestion and spotting of the peptide samples onto MALDI target plates for analysis. The automation of these individual steps exists as task-specific modules. The stained spots of interest can be extracted from a gel by instruments such as the Investigator ProPic workstation (Genomic Solutions Inc., Ann Arbor, Michigan, USA) and the gelPix high-throughput protein spot excision system (Genetix Ltd., New Milton, UK); an on-board high-resolution camera acquires an image of the gel, analyzes it using soft-

ware and generates a protein target excision map; selected proteins are excised through a multiple excision head unit in parallel. Other typical instruments capable of spot picking from gels are made by Amersham Biosciences, Bio-Rad Laboratories, Bruker Daltonics Inc. (Billerica, Massachusetts, USA), PerkinElmer Life Sciences and Proteomics Solutions. Extracted protein plugs can be placed into 96-well (and sometimes 384-well) plates and cataloged. Proteolytic enzymes (for example, trypsin) can be added to the removed pieces of gel and digested with heat in an instrument such as the Investigator ProGest gel digester (Genomic Solutions) or the MassPrep multifunctional robotic protein-handling system (Micromass/Waters Corp., Milford, Massachusetts, USA). This type of module performs all necessary wash steps and digests proteins into peptides. These systems generally use an in-gel digestion kit such as the Montage kits from Millipore. The work-station module uses a diluter with dual-syringe probes to reduce pipetting time. Multiple

units can be networked for higher throughput and controlled by a remote computer.

The polypeptide fragments can be desalted and concentrated using an automated MALDI preparation station. This workstation essentially is a modified liquid handler for one-, four- or eight-channel pipetting, and frequently it has a robotic arm that shuttles microplates in, out and around the deck. Porter (30, 31) covers the basics of liquid-handling workstations.

ZipTips and other disposable particle-loaded tips containing solid-phase extraction media are used commonly for this sample preparation step, and the tips are compatible with most liquid-handling workstations. The act of spotting places the peptide samples onto specific media, called MALDI targets, and immobilizes them in a suitable matrix for analysis. Typical examples of MALDI sample preparation and spotting workstations are the Investigator ProMS MALDI preparation station (Genomic Solutions), MAP II/8

AutoPrep robotic system (Bruker Daltonics), MassPrep, ProTeam Advanced Digest workstation (Tecan U.S. Inc., Durham, North Carolina, USA [see Figure 3]) and SymBiot I sample workstation (Applied Biosystems). Workstations such as the Hamilton Microlab Star also can be used simply to spot target plates with high positional accuracy. Modules are available that combine two distinct functions, such as protein digestion with MS sample spotting; the Investigator ProPrep Proteomic workstation (Genomic Solutions) represents an example of this approach to automation.

Total automated systems have been developed by assembling individual modules in series in an environmentally controlled atmosphere under computer control. Examples of total systems include the Ettan Spot Handling 2-D workstation (Amersham Biosciences [see Figure 4]), the Investigator Proteomics system (Genomic Solutions), the Pioneer suite (Bruker Daltonics), the ProTeam product suite (Tecan U.S.), the ProteomeWorks system (Micromass/Waters with Bio-Rad), the Proteomics Solution 1 system (Applied Biosystems) and the Xcise system (Proteome Systems). Some key abilities of these systems are high-resolution imaging, accurate spot excision, peptide purification, spotting onto MALDI target plates and integrated software. Table I lists typical manufacturers of products for automated 2-D spot picking, digestion, MALDI plate preparation, plate spotting and total integrated systems.

Multidimensional LC with MS detection also is amenable to automation by adding column-switching capabilities with plumbing to accommodate multiple columns in parallel. Chromatography companies are responding to this need with many products, the discussion of which is outside the scope of this article. Typical examples of such systems include the NanoLC Proteomics system (Eksigent Technologies, Livermore, California, USA) and Proteome X workstation (Thermo Finnigan, San Jose, California, USA); Eprogen (Darien, Illinois, USA) has a 2-D protein separation technology called ProteoSep that is useful for this application.

The demand for automated protein processing methods continues as companies are challenged to meet many needs: greater sample throughput, reproducibility, sensitivity for low-abundance proteins, enhanced low-volume liquid handling and

high-density sample spotting onto target plates (32). The important role for automation in proteomics is discussed further in texts (33–35) and in product reviews (36, 37).

Advancements in Sample Preparation

The industrialization of proteomics cannot occur simply from automating 2-D electrophoresis and MS. Advancements in protein separations, labeling chemistry, nanoscale sample handling, on-probe membrane digestions and protein chips are being investigated as methods to reduce the time, steps, sample loss and cost per analysis. The development of next-generation products will occur from leveraging technologies, advancements in engineering and liquid handling and low-cost manufacturing from the semiconductor industry in collaboration with tool manufacturers.

The presence of multiple proteins at varying expression levels per gel is a major limitation of 2-D electrophoresis-based proteomics. On-target concentration by products such as Bruker Daltonics' AnchorChip sample targets has lowered the limits of detection but has not eliminated the ion-suppression effects that can prevent detection of low-level proteins. Waters has shown a reduction of ion suppression using the LC–MALDIprep system to continuously deposit protein digest from reversed-phase LC onto tracks of a precoated target (38). Mussellman and Johnscher (39) discussed other commercially available systems such as the Probot Micro fraction collector (LC Packings, San Francisco, California, USA) in a recent review of LC–MALDI products.

Sample preparation for pre-MS protein analysis from 2-D electrophoresis requires several labor-intensive and time-consuming sample-handling steps. Gyros AB (Uppsala, Sweden) has developed the Gyrolab MALDI SP1 workstation using microfluidics on a CD to simultaneously digest 96 protein samples. The product also can concentrate, desalt and mix peptides with a MALDI matrix in a 200-nL final volume. Pretreatment of the surfaces reduced nonspecific binding and improved peptide recovery. Andersson (40) obtained sensitivity in the 400-amol range from a bovine serum albumin digest.

Extensive research also is taking place to develop commercial products that reduce digestion time from hours to minutes by using a variety of methods such as on-line microcolumn capture and digestions (19), micromembrane digestions (20), nonporous polyurethane membranes (41) and on-probe nanovials (42). Clearly, proteomic sample preparation is undergoing tremendous advances in response to high-throughput needs, and the development of improved automated systems that meet these goals will continue to intensify.

References

1. V.C. Wasinger, S.J. Cordwell, A. Cerpapoljak et al., *Electrophoresis* **16,** 1090–1094 (1995).
2. The International Human Genome Mapping Consortium, *Nature* **409,** 934–941 (2001).
3. The Celera Genomics Sequencing Team, *Science* **291,** 1304–1351 (2001).
4. D.C. Liebler, *Introduction to Proteomics: Tools for the New Biology* (Humana Press, Totowa, New Jersey, 2002).
5. D.A. Wells and L.L. Herron, *LCGC* **20(5),** 416–428 (2002).
6. A.J. Link, Ed., *2-D Proteome Analysis Protocols* (Humana Press, Totowa, New Jersey, 1999).
7. P.G. Righetti, A. Castagna and B. Herbert, *Anal. Chem.* **73,** 320A–326A (2001).
8. M.P. Deutscher, Ed., *Methods in Enzymology, vol. 182: Guide to Protein Purification* (Academic Press, New York, 1990).
9. R. Westermeier and T. Naven, *Proteomics in Practice: A Laboratory Manual of Proteome Analysis* (Wiley-VCH, Weinheim, Germany, 2002).
10. J.E. Coligan, B.M. Dunn, H.L. Ploegh et al., Eds., *Current Protocols in Protein Science* (John Wiley & Sons, New York, 2002).
11. A.J.S. Rayl, *Scientist* **15(7),** 10 (2001).
12. M.J. Hubbard, *Proteomics* **2,** 1069–1078 (2002).
13. R.E. Banks, M.J. Dunn, M.A. Forbes et al., *Electrophoresis* **20,** 689–700 (2000).
14. Y. Oda, K. Huang, F.R. Cross et al., *Proc. Nat. Acad. Sci. USA* **96,** 6591–6596 (1999).
15. S.P. Gygi, B. Rist, S.A. Gerber et al., *Nat. Biotechnol.* **17,** 994–999 (1999).
16. T.J. Griffin and R. Aebersold, *J. Biol. Chem.* **276,** 45497–45500 (2001).
17. T. Rabilloud, *Proteomics* **2,** 3–10 (2002).
18. V.M. Govorun and A.I. Archakov, *Biochemistry (Moscow)* **67,** 1109–1123 (2002).
19. D. Craft, A. Doucette and L. Li, *J. Proteome Res.* **1,** 537–547 (2002).
20. Y. Jiang and C.S. Lee, *J. Chromatogr. A.* **924,** 315–322 (2001).
21. Pluskal, A. Bogdanova, M. Lopez et al., *Proteomics* **2,** 145–150 (2002).
22. J.R. Yates III, *J. Mass Spectrom.* **33,** 1–19 (1998).
23. R. Aebersold and D.R. Goodlett, *Chem. Rev.* **101,** 269–295 (2001).
24. K.B. Tomer, *Chem. Rev.* **101,** 297–328 (2001).
25. J. Peng and S.P. Gygi, *J. Mass Spectrom.* **36,** 1083–1091 (2001).
26. S. Kazmi and I.S. Krull, *PharmaGenomics* **2,** 14–29 (2001).
27. S. Kazmi and I.S. Krull, *PharmaGenomics* **2,** 16–27 (2001).
28. T. Wehr, *LCGC* **20(10),** 954–962 (2002).
29. B.F. Cravatt and E.J. Sorensen, *Curr. Opin. Chem. Biol.* **4,** 663–668 (2000).
30. G. Porter, *Am. Biotech. Lab.* **20(7),** 40 (2002).
31. G. Porter, *PharmaGenomics* **3,** 38–44 (2002).
32. G. Porter, *Am. Biotech. Lab.* **21(3),** 70–71 (2003).
33. S. Pennington and M. Dunn, Eds., *Proteomics: From Protein Sequence to Function* (Springer-Verlag, New York, 2000).
34. M. Quadroni and P. James, *Electrophoresis* **20,** 664–677 (1999).
35. M.F. Lopez, *Electrophoresis* **21,** 1082–1093 (2000).
36. J. Handley, *Anal. Chem.* **73,** 379A–383A (2001).
37. C.M. Harris, *Anal. Chem.* **73,** 447A–451A (2001).
38. D.B. Wall, S.J. Berger, J.W. Finch et al., *Electrophoresis* **23,** 3193–3204 (2002).
39. B. Musellman and K. Johnscher, *Am. Genomic/Proteomic Technol.* **22,** 24–27 (2002).
40. P.E. Andersson, *PharmaGenomics* **4,** 38–44 (2003).
41. C.H. Lee, M.E. McComb, M. Bromirski et al., *Rapid Commun. Mass Spectrom.* **15,** 191–202 (2001).
42. D. Ericsson, S. Ekstrom, J. Nilsson et al., *Proteomics* **1,** 1072–1081 (2001).**PG**

Help Us Serve You Better.
Please take a moment to complete our issue feedback survey, online at **www.pharmagenomicsonline.com.** *Thank You. Your input is important.*