

Chapter 2

Proteomics: State of the Art and Its Relevance for Gene Therapy

Sergiu Scobioala, Rainer Klocke, Günter Michel,
and Sigrid Nikol

The molecular mechanisms underlying most diseases, including those of the cardiovascular system, are widely unknown. Basically, pathological changes in the organism arise from protein alterations. Proteomics comprises a set of tools allowing the identification of protein alterations, i.e. changes of protein abundance and posttranslational modifications, associated with diseases. The linkage of information about such protein changes with functional alterations as revealed by physiological studies constitutes functional proteomics that enables the disclosure of disease mechanisms.

Disease-linked protein alterations include those of suitable candidates for drug targets and disease biomarkers as well as therapeutic proteins/peptides. Since gene therapy depends on the function of a therapeutic protein encoded by a “therapeutic” gene, proteomic analyses provide the basis for the design and application of gene therapies.

The storage and administration of experimental data obtained by the application of proteomic analyses is supported by species- and tissue-specific protein databases and specific software. Publications in this field are reviewed in this chapter.

Sergiu Scobioala Department for Cardiology and Angiology, Westfälische Wilhelms University, Münster, Germany; and Interdisciplinary Center for Clinical Research (IZKF), Münster, Germany.
Rainer Klocke, Günter Michel, and Sigrid Nikol Department for Cardiology and Angiology, Westfälische Wilhelms University, Münster, Germany.

Cardiovascular Research: New Technologies, Methods, and Applications, edited by Gerard Pasterkamp and Dominique P. V. de Kleijn. Springer, New York, 2005.

1. INTRODUCTION

1.1. Definitions

The term *proteome* was created by Wasinger *et al.* (1995) and Wilkins *et al.* (1996a) and defined as the cell- and tissue-specific patterns of proteins expressed by the genome of an organism. Proteomics provides the technologies to investigate the proteome. Traditionally, proteomic analyses or “proteomics” means the separation of a large number of proteins from a cell line or organism at a given time point in 2-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and the cataloging of the spots to create databases. Today, proteomics also comprises the identification of proteins and their posttranslational modifications (PTMs) by mass spectrometry (MS), their functional analysis, e.g., by interaction studies using techniques such as the yeast two-hybrid system or the phage display method, and their subcellular localization.

1.2. Importance of Proteomics

Because of different mechanisms, e.g. differential splicing and alternative codon usage, one gene can encode various proteins. Preliminary studies suggest an average number of protein variants per gene of three to more than six in human beings (Wilkins *et al.*, 1996b). Thus, the human organism may contain more than half a million proteins, which is not evident from the genome sequence. This complexity is further enhanced by different types of modifications contributing to the regulation of protein activity at various levels of protein expression (Fig. 1). Examples are PTMs, e.g. phosphorylation, glycosylation, peptide cleavage, and complex formation of the proteins. Consequently, the functional complexity of an organism significantly exceeds that indicated by its genome sequence.

Since proteins are the main functional entities of the cell they should be expected to provide the most relevant information regarding cellular functions. Together with an often poor correlation between mRNA abundance and the quantity of the corresponding protein present within a cell, a recent study analyzing human liver samples determined the correlation coefficient between the amount of mRNA

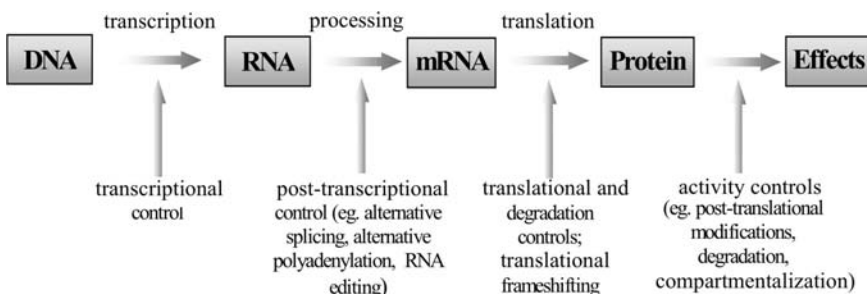


FIGURE 1. The ways in which gene expression can be regulated or modified from transcription to posttranslation. (See Color Plate 3.)

present and the corresponding protein abundance to be 0.48 only (Seilhamer and Anderson, 1999); this indicates the need for large-scale proteomics studies that complement the widely performed transcriptome analyses by DNA microarray hybridization technologies.

Because of the central importance of proteins for the functions of cells, tissues, organs, and whole organisms, disease processes are predominantly caused by changes in protein expression, subcellular localization, and PTM. Accordingly, the identification of such changes by proteomics is essential to define candidates for drug targets, therapeutic proteins, and disease biomarkers. Since gene therapy, although primarily based on the delivery of a “therapeutic” gene to the diseased organ, is an approach whose effect is finally due to the function of the therapeutic protein encoded by the “therapeutic” gene, proteomics provides the basis required for the design and application of gene therapies. In addition, as it is the case for conventional types of therapies too, proteomics enables the molecular monitoring of therapeutic effects and undesirable side effects by providing and exploiting disease biomarkers.

2. TECHNOLOGIES

Proteomics comprises a vast number of techniques, which basically enables sample preparation, separation, imaging, and identification of proteins, as well as the determination of the nature and position of protein modifications (Fig. 2)

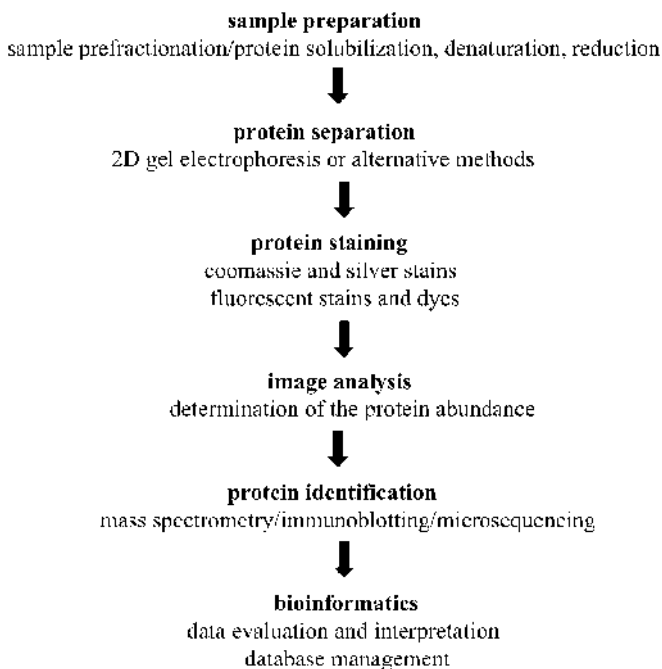


FIGURE 2. Elements of proteomics-based studies.

(Westbrook *et al.*, 2001) and the identification of protein–protein interactions. 2D-PAGE and MS are currently the central proteome analysis techniques.

2.1. Protein Separation by 2D-PAGE

2.1.1. Sample Preparation

Sample preparation involves the isolation and homogenization of cells or tissues, protein dissociation into its polypeptide subunits, solubilization, denaturation (unfolding of secondary and tertiary structures), and reduction of internal disulfide bonds to break secondary structures.

Tissue samples are usually isolated by microsurgery. However, since tissue samples often contain many different contaminating cell types, e.g. cells of connective tissue (fibroblasts), only a subset of the cells present will actually contribute proteins that are relevant for the study. To circumvent this problem laser capture microdissection was developed (Fig. 3) (Simone *et al.*, 2000a). Specific cell subpopulations isolated by laser capture microdissection from frozen or fixed tissue sections under direct microscopic visualization have been used for highly

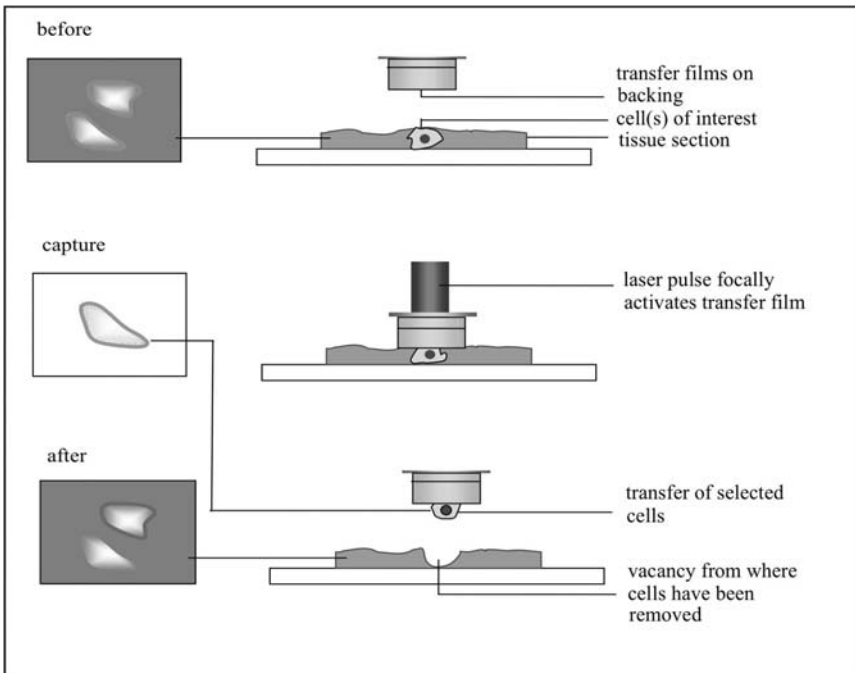


FIGURE 3. Laser capture microdissection. A transparent polymer film is placed in direct contact with the surface of a heterogeneous tissue section. Laser energy is used to activate the polymer directly over the selected cells. The activated region captures the selected cells, which can be lifted away from unwanted tissue. (See Color Plate 3.)

Table 1
Methods for Protein Prefractionation

Prefractionation Method	Principle	Reference
High-speed centrifugation	Differential sedimentation under centrifugal force	Celis, 1998
Affinity-based purification	Sterical interaction of proteins with immobilized receptors on specific surface	Ianello and Jeffrey, 1990
Liquid-phase separation	Separation of proteins in the liquid phase according to their physical characteristics—charge, size, etc.	Bier, 1998
Precipitation	Differential solubility of proteins	Görg <i>et al.</i> , 1997
Selective “clearing”	Selective filtration of proteins of interest	Walsh <i>et al.</i> , 1984
Sequential extraction	Differential solubility of proteins	Weiss <i>et al.</i> , 1992

sensitive and reproducible proteomic analysis using 2D-PAGE and other analytical methods (Simone *et al.*, 2000a, 2000b).

Deep-frozen cells or tissues are homogenized by different techniques such as grinding in a liquid nitrogen-cooled mortar, use of microdismembrators, sonication, or shearing-based methods. Because of the high diversity of proteins in eukaryotic tissues, it is sometimes advisable to carry out a prefractionation step to enrich the desired proteins (e.g. low abundant proteins or alkaline proteins) on the basis of their physical characteristics—e.g. isoelectric point (pI), molecular weight, hydrophobicity. Prefractionation of proteins can be achieved by various methods (Table 1).

The most frequently used buffer for sample solubilization is based on O’Farrel’s lysis buffer and contains the chaotropic agent urea, the non-ionic detergent NP-40, and the reducing agent dithiothreitol (O’Farrel, 1975). For the solubilization of certain classes of proteins, in particular membrane, membrane-associated, or other hydrophobic proteins, modified variants of O’Farrell’s lysis buffer containing various combinations of reducing agents, chaotropic agents, and detergents (Herbert, 1999; Herbert and Molloy, 1998; Herbert and Sanchez, 1997) are available. Although the presence of the relatively high concentrations of chaotropic agents and detergents in the lysis buffer usually minimizes protease activity for the duration of the sample-processing period (Rabilloud, 1996), it is sometimes advisable to include protease inhibitors to protect sensitive proteins from proteolytic degradation.

2.1.2. 2D-PAGE

The first step of proteome analysis consists in the separation of complex mixtures of proteins obtained from cells, tissues, and organs.

Currently, the best method for separating complex protein mixtures is 2D-PAGE. One can routinely resolve 2000–3000 proteins from a single sample of a protein extract. Using large-scale gels (Jungblut *et al.*, 1994) even up to 10,000 proteins can be resolved. 2D-PAGE involves the separation of proteins on the basis of their isoelectric point by isoelectric focusing (IEF) in the first dimension and based on molecular mass in the second dimension. The result is a gel

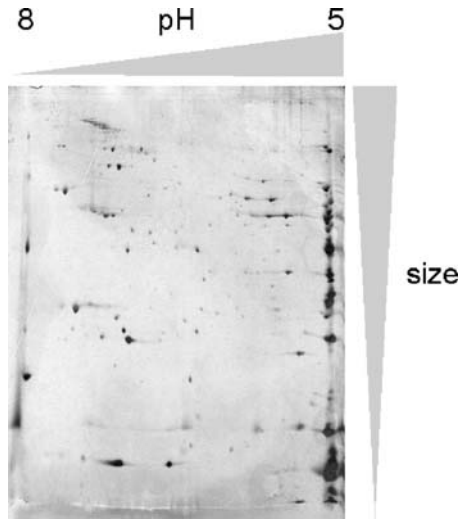


FIGURE 4. Silver-stained 2D-PAGE gels of a mouse heart (left ventricle). (*See Color Plate 4.*)

which, when stained, reveals a large number of spots (Fig. 4). Any given spot may contain more than one protein and any given protein may be contained in multiple spots depending on, for example, its posttranslational modification states.

2.1.2a. Separation in the First Dimension Proteins are amphoteric molecules and hence will migrate in the presence of an electric field based on the total sum of the charges present on the molecule at a given pH. Following application of current, the charged polypeptides will migrate in a polyacrylamide gel strip or a tube gel that contains a pH gradient until they reach the pH at which their overall charge is neutral (pI), hence producing a gel strip containing discrete protein bands along its length.

The currently preferred approach for protein separation in the first dimension consists in the use of Immobiline reagent-based (Amersham Pharmacia Biotech) immobilized pH gradients (IPGs). IPG strips are composed of acidic and basic buffering groups covalently linked to the polyacrylamide matrix. They are available in various lengths and pH ranges. Wide pH (3–10) gradients can be used to provide an overview of the protein diversity in a sample. Narrow-range IPG strips covering a pH range of three pH steps are used to achieve an optimum resolution in order to avoid the presence of multiple proteins in a single spot for unambiguous protein identification (Dunn, 2000).

Before separation in the second dimension, iodoacetamide is added for *S*-carboxymethylation in order to prevent S–S bridges from re-forming after they have been reduced using DDT or mercaptoethanol.

2.1.2b. Separation in the Second Dimension Since the IPG strips containing the separated, carboxymethylated proteins are placed on an SDS polyacrylamide gel, the proteins will migrate in a direction rectangular to their

Table 2
Comparison of Staining Techniques

Parameter	Silver Stain	Coomassie Blue Dye	Fluorescent Dye
Sensitivity	+++ ^a	++ ^b	+++ ^a
Reproducibility	+++ ^a	++	+++ ^a
Compatibility with mass spectrometry	++ ^b	+++ ^d	+++ ^a
Linearity of staining intensity	++ ^c	+++	+++ ^a
Simplicity and rapidity of staining procedure	+	++	++ ^a

^aGörg *et al.*, 2000.

^bHames, 1990.

^cGiometti *et al.*, 1991.

^dMacri *et al.*, 2000.

former movement when current is applied. Specific gel characteristics such as size, thickness, and percentage of acrylamide, the inclusion of a stacking gel as well as the buffering conditions may be varied according to the specific requirements (Macri and Rapundalo, 2001). Second dimension 2D-PAGE systems are available in both horizontal and vertical configurations. (Görg *et al.*, 1995).

2.1.3. Protein Staining and Labeling

Protein spots in 2D gels can be detected by Coomassie staining, zinc or silver staining, ³²P or ³⁵S radiolabeling, fluorescent staining, and/or immunodetection.

Although radiolabeling is the most sensitive detection method, silver chemistry and Coomassie staining are the most commonly applied (Table 2). Introduction of new fluorescent dyes with a variety of detection wavelengths and sensitivities has great application potential.

2.1.3a. Coomassie Stains Stains with the Coomassie blue dyes G-250 or R-250 cannot be used to quantify proteins and the sensitivity is clearly below silver and fluorescent stains (Hames, 1990). However, in contrast to most silver stain methods, Coomassie blue staining does not greatly interfere with subsequent protein identification by mass spectrometry (Macri and Rapundalo, 2001).

2.1.3b. Silver Stain With a sensitivity 10–50 times that of Coomassie blue, silver stain allows the detection of as little as 0.1 ng of protein per spot. Either silver diammine (alkaline method) or silver nitrate (acidic method) is mainly used as the silvering agent (Rabilloud, 1992).

2.1.3c. Fluorescent Stains and Dyes For fluorescent staining three approaches are possible (Görg *et al.*, 2000): (1) pre-electrophoretic derivatization of proteins, where proteins are coupled with a fluorescent dye, e.g. monobromobimane and 2-methoxy-2, 4-diphenyl-3(2*H*)-furanone, prior to the IEF step of 2D-PAGE, (2) labeling of proteins after the first dimension, and (3) postelectrophoretic labeling after 2D-PAGE with dyes like SYPRO Orange, SYPRO Red, or

SYPRO Ruby. The fluorescently labeled proteins can be imaged with a UV transilluminator, a blue-light transilluminator, or laser-scanning instruments (Macri and Rapundalo, 2001). The best results were obtained with SYPRO Ruby staining. The patterns obtained with silver staining and SYPRO Ruby staining are similar, but not identical (Görg *et al.*, 2000; Görg and Weiss, 2000).

2.1.4. Technical Problems and Limitations of Protein Separation by 2D-PAGE and Their Detection

Streaks are one of the most common technical problems associated with 2D-PAGE (Görg and Weiss, 2000). Horizontal streaks are localized in different areas of the 2D gels and are caused by (1) reduced solubility of certain proteins during IEF due to protein overload or protein–contaminant interaction, (2) considerable over focusing times, or (3) depletion of DDT during IEF (Görg *et al.*, 1995).

Vertical “empty” lines can be caused by urea crystals formed during IEF on the surface of the IPG strip (cf. Görg *et al.*, 2000).

Background smear can be due to poor quality of chemicals used for sample preparation or electrophoresis (cf. Görg *et al.*, 2000).

There are still technical problems to resolve all the proteins within a proteome by 2D-PAGE. Two circumstances hampering total proteome resolution by 2D-PAGE are insufficient sensitivity to detect low-abundance proteins and the underrepresentation of certain classes of proteins, e.g. basic and membrane proteins. To overcome these problems sample prefractionation methods (see above) can be applied to achieve subproteomic enrichment (Table 1).

2.1.5. Image Analysis

Image analysis follows 2D-PAGE and protein staining. Image acquisition is accomplished with an imaging densitometer, document scanner, charged-coupled device camera, or storage phosphoimager (Macri and Rapundalo, 2001). The resulting image either is acquired digitally or has to be converted to a digital format. Several computer software packages are available for the analysis of 2D-PAGE gel images (Table 3). These software packages accomplish the detection and quantification of spots and they enable both multiple image alignments and image comparisons. The data obtained by quantitative analysis of matched protein spots are generally exported to a spreadsheet program for statistical analysis. Such export files usually contain quantitative measurements as well as protein spot characteristics such as x/y coordinates, area, and volume of specific protein spots (Macri and Rapundalo, 2001).

2.2. Alternative Analytical Protein Separation Technologies

2D-PAGE is the most popular analytical protein separation technique for proteome analysis. However, other separation methods have been developed and may be used for proteome analysis, provided that they achieve a sufficient degree of

Table 3
Software for the Analysis of 2D-PAGE

Software	Company	Web site
Proteomweaver	Definiens	www.definiens.com
Delta 2D	Decodon	www.decodon.de
Gellab	Scanalytics	www.scanalytics.com
ImageMaster 2-D Elite	Amersham Biosciences	www.amershambiosciences.com
Image1Q	Proteome Systems	www.proteomesystems.com
Investigator 2-D	BST Scientific	www.bst-asia.com
Kepler	Large Scale Biology	www.lsb.com
Melanie	Geneva Bioinformatics	www.genebio.com
PDQuest	Bio-Rad	www.discover.bio-rad.com
Phoretix 2D/Progenesis	Nonlinear Dynamics	www.nonlinear.com
ProteinMime	Scimagix	www.scimagix.com
Z3	Compugen	www.2dgels.com

resolution or may be suitable to isolate and display a subproteome not amenable to separation by 2D-PAGE. Examples for such technologies are capillary zone electrophoresis (Jensen *et al.*, 2000; Shen *et al.*, 2000), affinity chromatography (Damer *et al.*, 1998; Ping *et al.*, 2001), or the recently developed protein chip method (Fung *et al.*, 2000, 2001; Nelson *et al.*, 2000; von Eggeling *et al.*, 2000; Weinberger *et al.*, 2001; Williams and Addona, 2000).

2.3. Protein Identification and Characterization by MS

2D-PAGE does not provide direct clues to the identities of the separated proteins. The current method of choice for the identification of proteins is MS since it is faster, cheaper, and more accurate than earlier techniques like microsequencing by automated Edman degradation. In addition, MS is much more sensitive, can deal with protein mixtures, offers much higher throughput, and is available to automation. MS-based mass determination is based on the conversion of proteins or peptides into gas-phase ions with an ionization source. Mass spectrometers are classified according to the ionization source and mass analyzer employed. Matrix-assisted laser desorption ionization (MALDI) and electrospray ionization are the most widely used ionization techniques. Basically, ionized peptides are separated according to their mass-to-charge ratio and detected by mass analyzers. MALDI-derived peptides are usually analyzed in so-called time-of-flight analyzers.

Protein identification by MS is usually accomplished by two main approaches. The first technique is called mass fingerprinting and was initially suggested by Henzel *et al.* (1993). It is generally employed to identify proteins purified by 2D-PAGE and relies on in-gel digestion of gel-separated proteins into peptides by sequence-specific proteases (i.e. trypsin) or chemical reagents (i.e. CNBr). These sequence-specific protein digests provide characteristic mass profiles or “fingerprints” since most amino acid residues have a unique mass (Fig. 5).

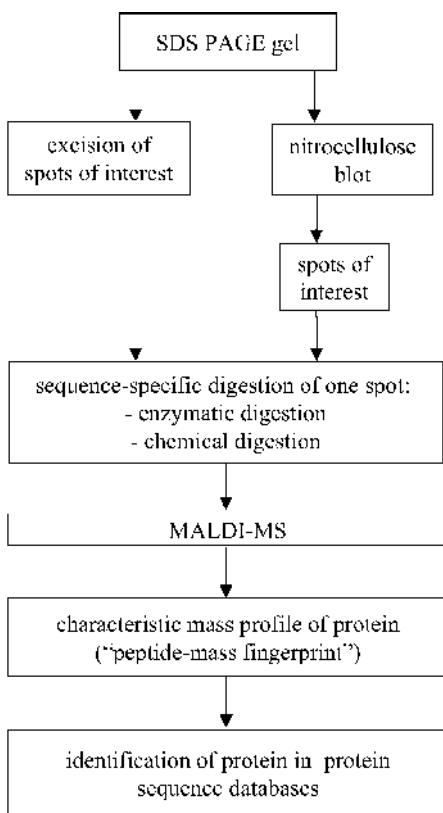


FIGURE 5. Mass spectrometric protein identification separated by 2D-PAGE.

The success of peptide mass fingerprinting (PMF) also depends on the existence of comprehensive, searchable databases for the species under investigation. Since the presence of full length cDNA sequences in the database is at least instrumental, PMF works effectively for species whose genomes are completely sequenced.

The main reason for analyzing peptides rather than proteins is that the molecular weight of proteins is usually not sufficient for database identification. However, even PMF databases are often insufficient for unambiguous protein identification. Therefore, PMF frequently has to be combined with additional MS techniques to obtain further information. Tandem mass spectrometry (MS/MS) can be instrumental to achieve unambiguous protein identification. In MS/MS, an ionized peptide is selected by the primary MS and fragmented by collision with inert gas. The resulting fragments are then analyzed by the supplementary MS to identify and locate the amino acids in the peptide. The identity and location within the peptide of only two amino acids can be sufficient to identify the peptide by comparison with the databases.

In addition to protein identification MS can be used to identify PTMs, e.g. phosphorylation, glycosylation, and sulfation. Since PTMs can determine activity, stability, localization, and turnover of proteins, their identification is very important for the determination of protein functions. In comparison with standard biochemical methods (see below), the use of MS for PTM identification is more powerful since it provides information about the type of a modification and its localization within the protein.

Efforts have been made to miniaturize and automate protein preparation using microfabricated “chips.” Such surface-enhanced laser desorption/ionization chips combine protein chip technology with MS (Fung *et al.*, 2000). Samples are first prepared and fractionated on chips with varying chromatographic properties, e.g. anion exchange, cation exchange, metal affinity, or reverse phase. After the chips are washed to remove unbound proteins, the bound proteins are identified in a time-of-flight mass spectrometer.

2.4. Technologies to Analyze PTMs

Besides MS (see above) further technologies including a number of classic biochemical techniques may be used to investigate PTMs. The use of biochemical methods as lectin binding, acid/Schiff staining, or enzymatic deglycosylation allows the identification of glycosylated proteins (Baker *et al.*, 1992; Ianello and Jeffrey, 1990). Phosphorylated proteins have been identified by ^{32}P radiolabeling in living cells (Hansen and Moller, 1993; Mason *et al.*, 1998), enzymatic dephosphorylation (Arrell *et al.*, 2000), or the use of antiphosphospecific antibodies (Soskic *et al.*, 1999).

Proteolysis, another type of PTM, may play a role in the etiology or pathology of cardiovascular diseases. For example, using direct Edman sequencing, ischemia/reperfusion-associated myosin light chain 1 proteolysis at amino acid residue 19 has been detected (Arrell *et al.*, 2001a).

2.5. Protein Bioinformatics

Bioinformatics is of central importance for proteomics. Special software packages enable the quantitative analysis and database collection of the results of 2D-PAGE proteome analysis (Dunn, 2000). In addition, special bioinformatics tools have been developed for the detailed characterization of proteins, e.g. for the calculation of their physicochemical properties (e.g. pI, M_r) and the prediction of their potential posttranslational modifications (i.e. MOD, FIND, BOLD, NetPhas) (Banks *et al.*, 2000). Most of these tools are available via proteomics servers like NIH (<http://www.ncbi.nlm.nih.gov>), EBI (<http://www.ebi.ac.uk>), and ExPASy (<http://www.expasy.ch/tools/>).

SWISS-PROT and SWISS-2DPAGE are typical examples of annotated protein and two-dimensional electrophoresis databases (Fig. 6). Such databases ideally should provide a visual image of a 2D-PAGE gel from which a protein spot of interest

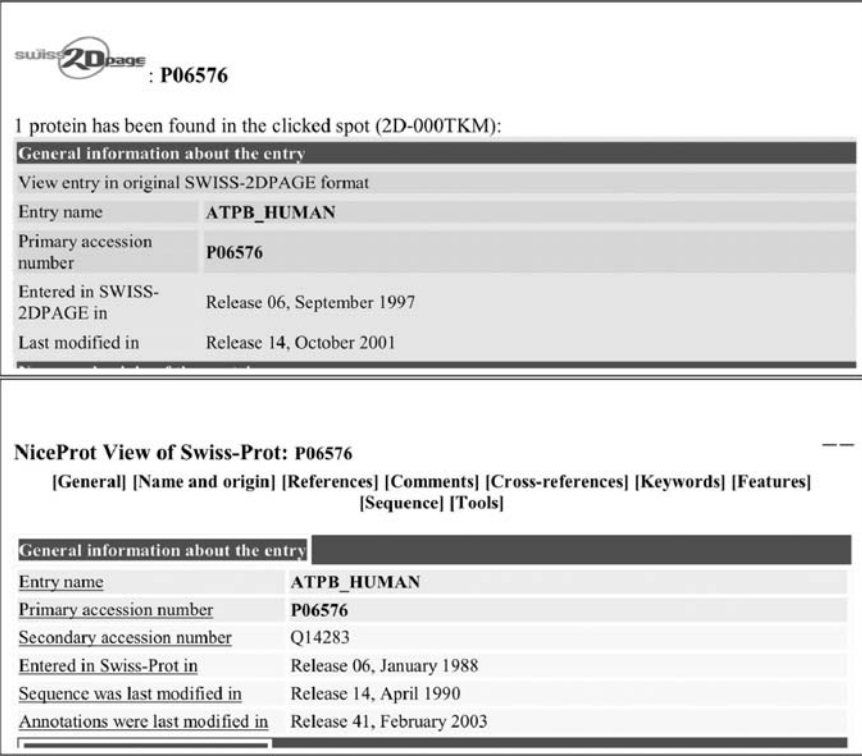


FIGURE 6. Annotated protein in SWISS-2DPAGE and SWISS-PROTS databases. (See Color Plate 4.)

can be selected to obtain specific information about that protein (e.g. M_r , pI , amino acid sequence if available, function, any PTMs) (Arrell *et al.*, 2001b). In addition, complex experimental details, relevant biochemical pathways, and disease implications should be highlighted (Banks *et al.*, 2000).

Special databases of human cardiac proteins are HSC-2DPAGE (Evans *et al.*, 1997), HEART-2DPAGE, HP-2DPAGE, and RAT HEART-2DPAGE. These databases are accessible through www.expasy.ch/ch2d/2d-index.html (Table 4). They contain information on several hundreds of cardiac proteins.

2.6. Technologies to Analyze Protein–Protein Interactions

Protein–protein interactions have a high impact on the determination of biological functions and their knowledge can potentially be exploited for therapeutic purposes. Currently, the most important approaches to study protein–protein interactions are the purification of protein complexes, the yeast two-hybrid system, and the phage display method.

Table 4
2D-PAGE Protein Databases for the Heart Accessible via the World Wide Web

Database	Web Address	Source	Protein Spots Separated	Protein Species Separated	Organization
HSC-2DPAGE	http://www.harefield.nthames.nhs.uk/nhli/protein	Human heart (LV)	~1500	~153	Heart Science Centre, Harefield Hospital
		Dog heart (LV)	1212	80	
		Rat heart (LV)	1188	30	
HEART-2DPAGE	http://www.chemie.fu-berlin.de/user/pleiss/dhzb.html	Human heart (RA, RV)	3300	150	German Heart Institute, Berlin
HP-2DPAGE	http://www.mdc-berlin.de/~emu/heart/	Human heart (RV)	~ 3300	70	MDC, Berlin
RAT HEART-2D-PAGE	http://gelmatching.inf.fu-berlin.de/~pleiss/2d/	Rat heart	>3000	64	German Heart Institute, Berlin

2.6.1. Purification of Protein Complexes

Protein–protein interactions can be studied by the purification of entire multiprotein complexes by affinity-based methods and their subsequent analysis by 2D-PAGE in combination with MS (Pandey and Mann, 2000). One way to identify the interaction partners of a protein consists in the coimmunoprecipitation of the protein and its interaction partners with an antibody. Usually, the protein is genetically tagged with an epitope for which an antibody is commercially available. The recombinant full length cDNA encoding the tagged protein is then transfected into a suitable cell. Subsequently, the overexpressed protein is immunoprecipitated from the cell extract together with its interaction partners.

Alternatively, a multiprotein complex can be purified based on its interaction with a “bait” fusion protein which is immobilized on a solid support. After washing off the proteins that interact nonspecifically, the protein complex is eluted and analyzed by 2D-PAGE in combination with subsequent MS for protein identification (Fig. 7). Such studies provide insight into mechanisms. Since no assumptions are made about the complex, unsuspected connections between cellular processes may be recognized.

2.6.2. Yeast Two-Hybrid System

The yeast two-hybrid system is a powerful tool to study protein–protein interactions (Fields and Song, 1989). It is based on the modular structure of the yeast transcription factors that contain a DNA-binding domain (DBD) and an activation domain (AD) (Auerbach *et al.*, 2001). Using DNA cloning techniques, a protein

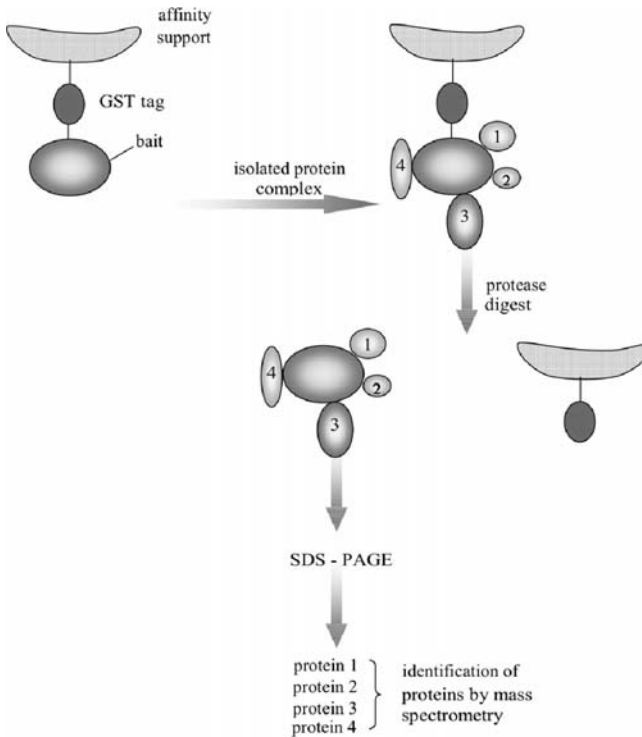


FIGURE 7. Analysis of protein–protein interactions. The protein of interest is expressed as a fusion protein with a cleavable affinity tag to identify interacting proteins. It is immobilized onto agarose beads using a glutathione S-transferase tag. Nuclear cell extracts are incubated with the beads and the beads washed extensively. Thrombin is used to cleave between the glutathione S-transferase and the “bait” protein, which results in the elution of all proteins that are specifically bound to “bait” (Pandey and Mann, 2000). The eluted proteins are resolved by 2D-PAGE and analyzed by MS. The success of the above-mentioned strategies relies on sufficient affinity of the protein complex to the bait and on optimized conditions for purifications steps. (See Color Plate 5.)

of interest X (“bait”) is fused to the DBD of a transcription factor whereas a potentially interacting protein Y (“prey”) is fused to the AD. The DBD-X and AD-Y fusion proteins are then expressed in a yeast strain that contains a plasmid with the binding site for the DBD upstream of a reporter gene. In the case of physical interaction of DBD-X and AD-Y, the DNA-binding and activation domains will come into close proximity, resulting in the reconstitution of a functional transcription factor that will activate transcription of the reporter gene (Fig. 8a).

In principle, no previous knowledge about the interacting proteins is necessary for a screen to be performed. Thus, the yeast two-hybrid system is suitable to identify protein–protein interactions in a genome-wide scale. Basically, two libraries of cDNAs comprising open reading frames (ORFs) fused to the DBD or AD of a transcription factor are used. Two different amino acid-deficient yeast strains containing the different libraries are generated. After mating the haploid strains

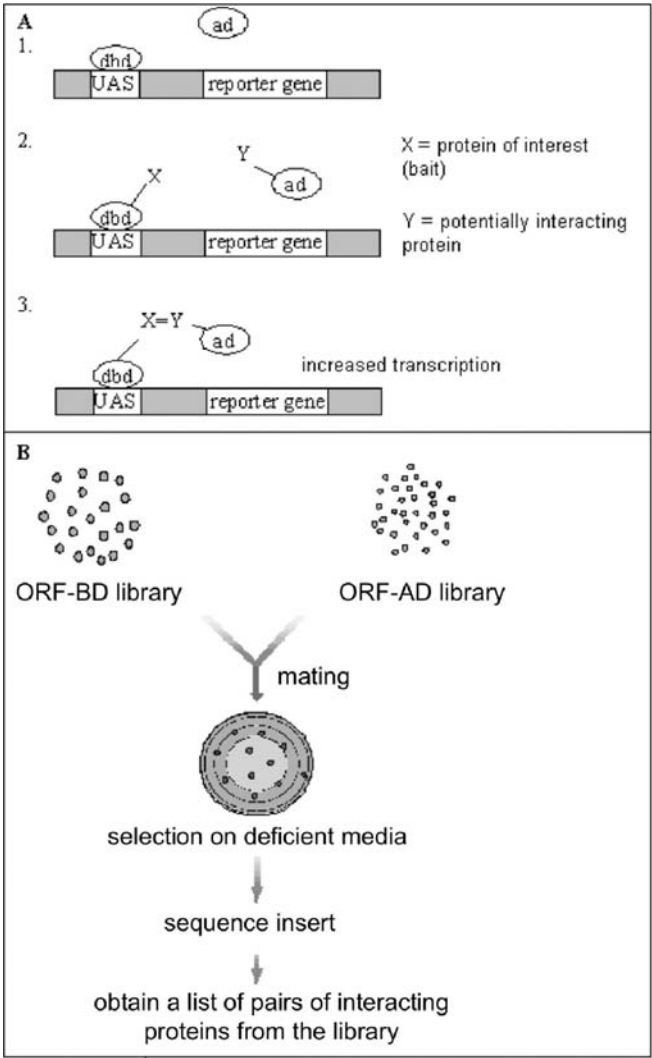


FIGURE 8. The yeast two-hybrid system. (a) *Schematic representation of the yeast two-hybrid system.* (1) The two separated domains of a transcription factor are not functional and therefore do not induce transcription of the reporter gene. (2) The DBD and AD are fused to two proteins of interest and co-expressed in a yeast reporter strain. (3) If DBD-X and AD-Y interact, the fusion proteins are assembled at the binding site of the reporter gene, which leads to activation of transcription. (b) *Library-based yeast two-hybrid screening method.* In this strategy, two different yeast strains containing two different cDNA libraries are prepared. In one case, the open reading frames (ORFs) are expressed as GAL 4-BD fusions and in other case, they are expressed as GAL 4-AD fusions. The two yeast strains are then mated and diploids selected on deficient media. Thus, only the yeast cells expressing interacting proteins survive. The inserts from both the plasmids are then sequenced to obtain a pair of interacting genes. (Modified from Pandey and Mann, 2000). (See Color Plate 6.)

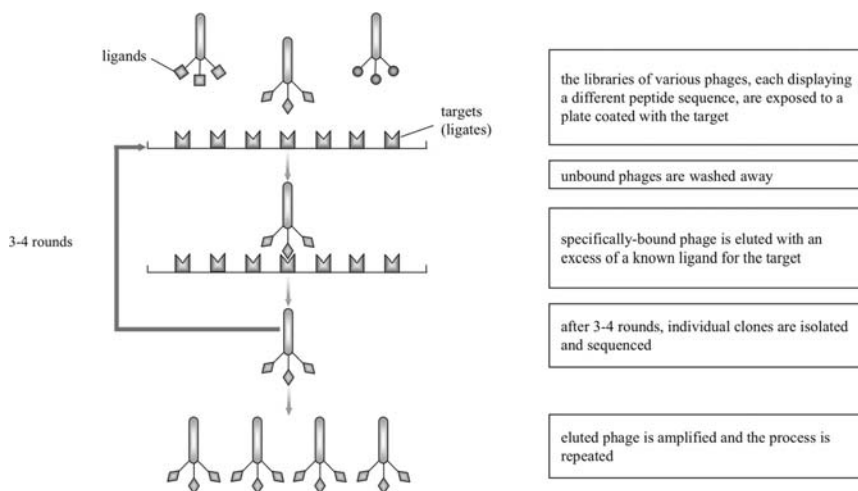


FIGURE 9. Affinity-selection of ligands from phage display libraries. Different peptide sequences can be screened for binding to a specific target molecule. After several rounds only the clones with the peptides that bind to the target are isolated. (See Color Plate 7.)

diploids are selected on deficient media. Thus, only the yeast cells expressing interacting proteins survive (Fig. 8b). The ORF cDNAs encoding the interacting proteins are identified by sequencing the relevant clones (Pandey and Mann, 2000).

The range of the conventional yeast two-hybrid system is limited by the fact that hydrophobic transmembrane proteins and membrane-associated proteins may not enter into the nucleus. Therefore, alternative membrane-associated two-hybrid systems have been developed (reviewed in Auerbach *et al.*, 2002).

2.6.3. Phage Display

The phage display method is based on bacteriophage particles expressing a peptide or protein of interest fused to a capsid or coat protein. The DNA encoding the virus proteins and also the fusion protein is localized within the protein coat of the virus particle. Consequently, the virus provides a physical link between the protein/peptide and the genetic information that encodes it. Libraries of peptides can be screened for binding to a variety of target molecules (antibodies, enzymes, cell-surface receptors) by incubating the phage library with the target molecules. Whereas the nonbinding phage is washed away, the remaining phages are amplified. After several rounds of binding under increasing selection pressure, individual clones are characterized by DNA sequencing (Fig. 9).

Phage display can be used to screen for peptide epitopes, peptide ligands, enzyme substrates, or single-chain antibody fragments in solution. Like the two-hybrid system it is simple and can be performed in an automated high-throughput setting (Pandey and Mann, 2000). Whether the former or the latter is preferable

depends on the particular class of proteins being studied, since the interactions take place in solution as opposed to the nucleus of the yeast cell. In addition, phage display is applicable to transcription factors, in contrast to the yeast two-hybrid system.

3. CONCLUSIONS AND PERSPECTIVES

Pathological changes within an organism arise from protein alterations. Proteomics provides a powerful set of tools for the large-scale study of protein expression and PTM. Accordingly, the application of proteomics is the strategy of choice to elucidate the mechanisms of diseases and to identify protein-based drug targets, disease biomarkers and therapeutic peptides and proteins, thus providing the basis of pharmacological and gene-based therapies.

2D-PAGE and MS are still the basic methods of protein identification, enabling the establishment of protein databases. Many complementary technologies have been developed recently or are currently being developed. These include protein arrays, the yeast-two hybrid system, phage-display libraries, surface-enhanced laser desorption/ionization (Auerbach *et al.*, 2001; Pandey and Mann, 2000). Either alone or in combination they will gain important roles in proteomics. The further development of high-quality bioinformatic tools and disease- and species-specific databases will be necessary to complement proteomics technologies (Jager *et al.*, 2002).

ACKNOWLEDGMENT. We thank the Interdisciplinary Center for Clinical Research (IZKF, project B15) for supporting Dr Sergiu Scobioala.

REFERENCES

- Arrell, D. K., Au, B. P. Y., Fraser, H., Sasaki, N., and Van Eyk, J. E., 2000, Pharmacological preconditioning mediates altered phosphorylation of the "unphosphorylatable" myosin light chain 1, *Circulation* **102**(Suppl. II):II-202.
- Arrell, D. K., Neverova, I., Fraser, H., Marban, E., and Van Eyk, J. E., 2001a, Proteomic analysis of pharmacologically preconditioned cardiomyocytes reveals novel phosphorylation of myosin light chain 1, *Circulation Res.* **14**:480–487.
- Arrell, D. K., Neverova, I., and Van Eyk, J. E., 2001b, Cardiovascular proteomics. Evolution and potential, *Circulation Res.* **88**:763–773.
- Auerbach, D., Galeuchet-Schenk, B., Hottiger, M. O., and Stagljar, I., 2002, Genetic approaches to the identification of interactions between membrane proteins in yeast, *J. Recept. Signal Transduct. Res.* **22**(1–4):471–481.
- Auerbach, D., Schenk, B., Hübner, B., Meier, M., Hottiger, M. O., Hübscher, U., and Stagljar, I., 2001, Proteomics in vivo: using yeast as a genetic tool to study protein–protein interactions, *BIOforum Int.* **4**:171–173.
- Baker, C. S., Corbett, J. M., May, A. J., Yacoub, M. H., and Dunn, M. J., 1992, A human myocardial two-dimensional electrophoresis database: protein characterisation by microsequencing and immunoblotting, *Electrophoresis* **13**(9–10):723–726.

- Banks, R. E., Dunn, M. J., Hochstrasser, D. E., Sanchez, J. C., Blackstock, W., Pappin, D. J., and Selby, P. J., 2000, Proteomics: new perspectives, new biomedical opportunities, *Lancet* **356**(9243):1749–1756.
- Bier, M., 1998, Recycling isoelectric focusing and isotachopheresis, *Electrophoresis* **19**(7):1057–1063.
- Celis, J., 1998, *Cell Biology: A Laboratory Handbook*, 2nd edn. San Diego: Academic Press.
- Damer, C. K., Partridge, J., Pearson, W. R., and Haystead, T. A., 1998, Rapid identification of protein phosphatase 1-binding proteins by mixed peptide sequencing and data base searching. Characterization of a novel holoenzymic form of protein phosphatase 1, *J. Biol. Chem.* **273**(38):24396–24405.
- Dunn, M. J., 2000, Studying heart disease using the proteomic research, *Drug Dis. Today* **2**:76–84.
- Evans, G., Wheeler, C. H., Corbett, J. M., and Dunn, M. J., 1997, Construction of HSC-2DPAGE: a two-dimensional gel electrophoresis database of heart proteins, *Electrophoresis* **18**(3–4):471–479.
- Fields, S., and Song, O. K., 1989, A novel genetic system to detect protein–protein interactions, *Nature* **340**(6230):245–246.
- Fung, E. T., Thulasiraman, V., Weinberger, S. R., and Dalmasso, E. A., 2001, Protein biochips for differential profiling, *Curr. Opin. Biotechnol.* **12**(1):65–69.
- Fung, E. T., Wright, G. L., and Dalmasso, E. A., 2000, Proteomic strategies for biomarker identification: progress and challenges, *Curr. Opin. Mol. Therap.* **2**, 643–649.
- Giometti, C. S., Gemmell, M. A., Tollaksen, S. L., and Taylor, J., 1991, Quantitation of human leukocyte proteins after silver staining: a study with two-dimensional electrophoresis, *Electrophoresis* **12**:536–543.
- Görg, A., Obermaier, C., Boguth, G., Csordas, A., Diaz, J. J., and Madjar, J. J., 1995, Very alkaline immobilized pH gradients for two-dimensional electrophoresis of ribosomal and nuclear proteins, *Electrophoresis* **18**(3–4):328–337.
- Görg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., and Weiss, W., 2000, The current state of two-dimensional electrophoresis with immobilized pH gradients, *Electrophoresis* **21**(6):1037–1053.
- Görg, A., Obermaier, C., Boguth, G., and Weiss W., 1999, Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: wide pH gradients up to pH 12, longer separation distances and simplified procedures, *Electrophoresis* **20**(4–5):712–717.
- Görg, A., and Weiss, W., 2000, *Proteome Research: Two-Dimensional Electrophoresis and Identification Methods*, pp. 57–106. Berlin: Springer- Verlag.
- Hames, B. D., 1990, *A Practical Approach*, pp. 51–85. New York: Oxford University Press.
- Hansen, K., and Moller, J. V., 1993, Application of two-dimensional gel analysis to identification and characterization of tyrosine phosphorylated substrates for growth factor receptors, *Electrophoresis* **14**:112–126.
- Henzel, W. J., Billeci, T. M., Stults, J. T., Wong, S. C., Grimley, C., and Watanabe, C., 1993, Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases, *Proc. Natl. Acad. Sci. U.S.A.* **90**(11):5011–5015.
- Herbert, B. R., 1999, Advances in protein solubilisation for two-dimensional electrophoresis, *Electrophoresis* **20**:660–663.
- Herbert, B. R., Molloy, M. P., Gooley, A. A., Walsh, B. J., Bryson, W. G., and Williams, K. L., 1998, Improved protein solubility in two-dimensional electrophoresis using tributyl phosphine as reducing agent, *Electrophoresis* **19**(5):845–851.
- Herbert, B. R., and Sanchez, J. C., 1997, *Proteome Research: New Frontiers in Functional Genomics*. Heidelberg: Springer-Verlag.
- Ianello, R. C., and Jeffrey, P. L., 1990, Glycoproteins of rat skeletal muscle sarcolemma: characterization by two-dimensional gel electrophoresis and effect of denervation, *Exp. Neurol.* **108**:156–161.
- Jager, D., Jungblut, P. R., and Muller-Werdan, U., 2002, Separation and identification of human heart proteins, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **771**:131–153.
- Jensen, P. K., Pasa-Tolic, L., Peden, K. K., Martinovic, S., Lipton, M. S., Anderson, G. A., Tolic, N., Wong, K. K., and Smith, R. D., 2000, Mass spectrometric detection for capillary isoelectric focusing separations of complex protein mixtures, *Electrophoresis* **21**(7):1372–1380.

- Jungblut, P., Otto, A., Zeindl-Eberhart, E., Plessner, K. P., Knecht, M., Regitz-Zagrosek, V., Fleck, E., and Wittmann-Liebold, B., 1994, Protein composition of the human heart: the construction of a myocardial two-dimensional electrophoresis database, *Electrophoresis* **15**(5):685–707.
- Macri, J., McGee, B., Thomas, J. N., Du, P., Stevenson, T. I., Kilby, G. W., and Rapundalo, S. T., 2000, Cardiac sarcoplasmic reticulum and sarcolemmal proteins separated by two-dimensional electrophoresis: surfactant effects on membrane solubilization, *Electrophoresis* **21**(9):1685–1693.
- Macri, J., and Rapundalo, S. T., 2001, Application of proteomics to the study of cardiovascular research, *Trends Cardiovasc. Med.* **2**:66–75.
- Mason, G. G., Murray, R. Z., Pappin, D., and Rivett, A. J., 1998, Phosphorylation of ATPase subunits of the 26S proteasome, *FEBS Lett.* **430**(3):269–274.
- Nelson, R. W., Nedelkov, D., and Tubbs, K. A., 2000, Biosensor chip mass spectrometry: a chip-based proteomics approach, *Electrophoresis* **21**(6):1155–1163.
- O'Farrel, P. H., 1975, High resolution two-dimensional electrophoresis of protein, *J. Biol. Chem.* **250**:4007–4021.
- Pandey, A., and Mann, M., 2000, Proteomics to study genes and genomes, *Nature* **405**:321–329.
- Ping, P., Zhang, J., Pierce, W. M., Jr., and Bolli, R., 2001, Functional proteomic analysis of protein kinase C epsilon signaling complexes in the normal heart and during cardioprotection, *Circulation Res.* **88**(1):59–62.
- Rabilloud, T., 1992, A comparison between low background silver diammine and silver nitrate protein stains, *Electrophoresis* **13**:429–439.
- Rabilloud, T., 1996, Solubilization of proteins for electrophoretic analyses, *Electrophoresis* **17**:813–29.
- Seilhamer, J., and Anderson, L., 1999, A comparison of selected m-RNA and protein abundances in human liver, *Electrophoresis* **18**:533–537.
- Shen, Y., Berger, S. J., Anderson, G. A., and Smith, R. D., 2000, High-efficiency capillary isoelectric focusing of peptides, *Anal. Chem.* **72**(9):2154–2159.
- Simone, N. L., Paweletz, C. P., Charboneau, L., Petricoin, E. F., III, and Liotta, L. A., 2000a, Laser capture microdissection: beyond functional genomics to proteomics, *Mol. Diagn.* **5**(4): 301–307.
- Simone, N. L., Remaley, A. T., Charboneau, L., Petricoin, E. F., III, Glickman, J. W., Emmert-Buck, M. R., Fleisher, T. A., and Liotta, L. A., 2000b, Sensitive immunoassay of tissue cell proteins procured by laser capture microdissection, *Am. J. Pathol.* **156**(2):445–452.
- Soskic, V., Gorlach, M., Poznanovic, S., Boehmer, F.-D., and Godovac-Zimmermann, J., 1999, Functional proteomics analysis of signal transduction pathways of the PDGF beta receptor, *Biochemistry* **38**:1757–1764.
- von Eggeling, F., Davies, H., Lomas, L., Fiedler, W., Junker, K., Claussen, U., and Ernst, G., 2000, Tissue-specific microdissection coupled with ProteinChip array technologies: applications in cancer research, *Biotechnologies* **29**:1066–1070.
- Walsh, M. J., Limos, L., and Tourtellotte, W. W., 1984, Two-dimensional electrophoresis of cerebrospinal fluid and ventricular fluid proteins, identification of enriched and unique proteins, and comparison with serum, *J. Neurochem.* **43** 1277–1285.
- Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., Gooley, A. A., Wilkins, M. R., Duncan, M. W., Harris, R., Williams, K. L., and Humphrey-Smith, I., 1995, Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*, *Electrophoresis* **16**(7):1090–1094.
- Weinberger, S. R., Dalmaso, E. A., and Fung, E. T., 2001, Current achievements using protein chip array technology, *Curr. Opin. Chem. Biol.* **6**:86–91.
- Weiss, W., Postel, W., and Görg, A., 1992, Application of sequential extraction procedures and glycoprotein blotting for the characterization of the 2-D polypeptide patterns of barley seed proteins, *Electrophoresis* **13**:770–773.
- Westbrook, J. A., Jun, X. Y., Wait, R., and Dunn, M. J., 2001, A combined radiolabelling and silver staining technique for improved visualisation, localization and identification of proteins separated by two-dimensional gel electrophoresis, *Proteomics* **1**:370–376.
- Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphrey-Smith, I., Hochstrasser, D. F., and Williams, K. L., 1996a, Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it, *Biotechnol. Genet. Eng. Rev.* **13**:19–50.

- Wilkins, M. R., Sanchez, J. C., Williams, K. L., and Hochstrasser, D. F., 1996b, Current challenges and future applications for protein maps and posttranslational vector maps in proteome projects, *Electrophoresis* **17**:830–838.
- Williams, C., and Addona, T. A., 2000, The integration of SPR biosensors with mass spectrometry: possible application for proteome analysis, *Trends Biotechnol.* **18**:45–48.

Cardiovascular Research
New Technologies, Methods, and Applications
Pasterkamp, G.; de Kleijn, D. (Eds.)
2006, XV, 239 p., Hardcover
ISBN: 978-0-387-23328-4