Proteomics is the study of the “proteome” or the entire protein complement of a genome, a term first coined merely 6 years ago. It is now known, however, that the proteome is far more complex than previously suggested by the one-gene, one-protein adage. The proteome consists of all proteins present in a cell or organism at a given time, including not only those translated directly from genetic material but also the variety of modified proteins arising from alternative splicing of transcripts, from extensive posttranslational processing, or from some combination of the two, resulting in modifications that have the potential to alter protein structure and/or function (Figure 1).

Since proteins are involved in virtually every cellular function, control every regulatory mechanism, and are modified in disease (as the cause or effect), the proteome dictates the phenotype of the cell and, collectively, the tissue or organ that the cells comprise. This phenotype varies under normal conditions, such as cell cycle stage, differentiation, function, and age, or as a result of the onset of or interventions in response to acute insults or chronic diseases (for general reviews, see Haynes and Yates, Dutt and Lee, Hoving et al, and Blackstock and Weir; cardiovascular/medical-based reviews, Jungblut et al, Dunn, and Banks et al). Acute insults lead to rapid posttranslational modification (PTM) of proteins, whereas in chronic disease states, cotranslational and posttranslational protein modifications occur in concert with altered gene expression, leading to varied protein levels. For specific proteins, disease-induced modification will substan-

Original received January 5, 2001; revision received March 9, 2001; accepted March 13, 2001.
From the Departments of Physiology (D.K.A., I.N., J.E.V.E.) and Biochemistry (J.E.V.E.), Queen’s University, Kingston, Ontario, Canada.
Correspondence to J.E. Van Eyk, 429 Botterell Hall, Queen’s University, Kingston, Ontario, Canada K7L 3N6. E-mail JVE1@post.queensu.ca
© 2001 American Heart Association, Inc.

Circulation Research is available at http://www.circresaha.org
tially affect function, which in turn has the potential to affect other proteins. The result is a dynamic, ongoing process of protein expression and modification. Proteomics is aimed at identifying and characterizing these protein changes.

Conceptually, proteomics is simple. In practice, it is technically challenging. This is due to a combination of the dynamic phenotypic modulation described above, which must be addressed for proper experimental design and interpretation of the seemingly limitless data that proteomic studies can provide. As outlined in Figure 2, numerous steps are involved in elucidation of a proteome. Although a vast array of techniques may be used, proteomics basically involves protein separation, protein identification, and characterization of the nature and position of protein modifications (for books exclusively on proteomics, see Link, Wilkins et al, and Rabilloud). Two-dimensional gel electrophoresis (2-DE) and mass spectrometry (MS) are currently the mainstay of proteomic analysis techniques, but additional methods are exploited for reproducible and complete separation of complex protein mixtures, as well as for characterization of PTMs.

Here, we overview the application of proteomics to cardiovascular research. We briefly describe considerations for appropriate experimental design, techniques used for proteomic analyses, various proteomic approaches, and examples of these approaches, including specific cardiovascular examples, where available. Although the recent advent of proteomics translates to only a sporadic selection of current examples in cardiovascular research, they will increase dramatically over the next few years. Integration of proteomics with functional data from established biochemical and physiological methods should lead, in the future, to development of functional proteomics, clarifying proteome dynamics in cardiovascular disease.

**Techniques**

**Considerations for Experimental Design**

Any proteomic analysis is very costly, laborious, and time-consuming, regardless of whether or not the resources for full-scale automation are available. In other words, do not embark on a proteomic analysis on a whim. As in any experimental procedure, the approach is critical, so be sure that a proteomic analysis is necessary before jumping headlong into the fray. A proteomic study provides you with mountains of data, which may or may not be interpretable if the question you are attempting to address is too complex. This complexity may also be further exacerbated by a number of factors. Because proteomic analyses look for protein modifications, and these modifications may vary with time, it may be necessary in many cases to make observations of the proteome at a variety of time points. Once modified proteins have been found, subsequent analysis of each and every one of them may require completely unique experimental conditions. Thus, the more straightforward the design of the experiment, the more likely it is that you will be able to gather information from which reliable conclusions may be drawn. In short, proteomic tools may provide us with information never before imaginable, but proper design and implementation of the study are paramount.

**Sample Preparation**

For accurate proteomic analyses, sample preparation is of the utmost importance. Analysis of a proteome at a specific moment of interest requires that it be arrested precisely at that moment. Otherwise, further cellular processes may alter it through addition or elimination of modifications. Key to the success of resolving any proteome is maximization of protein solubility. Although the proteome may be analyzed as a whole in a single solubilization step, it is often preferable to prepare a set of subproteomes. Fractionation methods exploiting specific protein characteristics, such as their inherent chemical properties (biospecificity; hydrophobicity; charge) or differential cellular compartmentalization, may thus be used to selectively solubilize specific subproteomes. This simplifies the daunting task of attempting to resolve thousands of proteins simultaneously, with the added benefit that a greater proportion of protein will most likely be solubilized. Of course, with each additional fractionation step, a subproteomic approach also increases the potential for undesired protein modification. Whether analyzing an entire...
proteome or a group of subproteomes, which is often guided by the specific research objectives, sample preparation must be reproducible as well as compatible with subsequent methods of protein separation and identification.

Two-Dimensional Gel Electrophoresis
The workhorse of proteomic protein separation is 2-DE. Separation in the first dimension is carried out by isoelectric focusing (IEF), which focuses proteins by relative charge according to their inherent charge, or isoelectric point (pI). Proteins are then resolved orthogonally in the second dimension by their relative molecular mass ($M_r$), typically by SDS-PAGE. Technical aspects of 2-DE have been widely reviewed.22–24

After completion of 2-DE, a number of options for detection of protein spots exist, the more common including Coomassie, zinc, or silver staining; $^{32}$P or $^{35}$S radiolabeling; and/or immunodetection. Although radiolabeling is the most sensitive detection method, silver is the most commonly applied. Introduction of new fluorescent dyes with a variety of detection wavelengths and sensitivities approaching that of silver has great application potential. Experimental requirements, however, generally dictate the choice of detection method.

Image analysis follows 2-DE. Adequate image reproduction is essential to properly map detected spots. Specialized image analysis software is now available from a number of sources to deal with the tremendous complexity of protein spot patterns. Such software must fulfill a number of requirements, including spot detection and quantification, and the abilities to perform both multiple image alignments and image comparisons.25,26 Although not absolutely essential, statistical functions are often incorporated into these software packages for complete quantitative image analysis.

Alternative Protein Separation Methods
Although 2-DE is the most popular protein separation technique for proteomic analyses, other separation methods may be warranted, provided that they have the ability either to achieve the resolution of 2-DE or to isolate a subproteome not amenable to separation by 2-DE. Well-established separation methods based on the physical attributes of proteins, alone or in combination, include ion exchange, size exclusion, reversed-phase high-performance liquid chromatography (HPLC) (for review, see Opitcek et al.27 and Patterson; noncardiovascular, see Link et al.29), capillary IEF, and capillary zone electrophoresis (for review, see Manabe).20,31,32 Noncardiovascular, Jensen et al.32 and Shen et al.33. Another separation method, affinity chromatography, is extremely selective, exploiting protein affinity for antibodies, specific target proteins, or chemical moieties (cardiovascular, see Damer et al.34 and Ping et al.35). Affinity-based separation may be achieved by conventional column chromatography, immunoprecipitation, or by the newly developed "protein chip" method (for review, see Williams and Addona).36 Noncardiovascular, eg, Nelson et al.37 and von Eggeling et al.38.

Some of these methods are advantageous in that they may be carried out under native or denaturing conditions. This is possible because proteins fractionated by many of these alternative methods remain in a soluble state, unlike protein subjected to 2-DE. Isolation of proteins in their native state thus provides the opportunity for in vitro biochemical assays after separation. A further benefit, in terms of proteomics, is that maintenance of solubility allows such methods to be linked directly to MS.

Protein Identification by Mass Spectrometry
After separation and detection, proteins of interest must be identified. The most significant breakthrough in the evolution of proteomics is the development of MS for protein identification (for review, see References 39 to 43). Over the past few years, improvements in MS have made it an unrivaled technique, by virtue of its accuracy of mass detection, its detection sensitivity, its ability to deal simultaneously with mixtures of multiple proteins, and its amenity to automation and therefore, high throughput.

MS instruments range, in relative terms, from simple (MALDI-TOF) to highly complex tandem (MS/MS). In all mass spectrometers, peptides are ionized from the sample. This is achieved either by matrix-assisted laser desorption/ionization (MALDI) of a solid-state sample or by electrospray ionization (ESI) directly from the liquid phase. Ionized peptides are separated on the basis of their mass-to-charge ratio and detected according to their time-of-flight (TOF) distribution or analyzed by quadrupole mass filters. In tandem MS/MS, an ionized peptide of interest is selected by the first MS and fragmented by collision with inert gas, and the resulting fragments are then analyzed in the second MS. Modern ESI-based MS/MS may use LC systems such as capillary zone electrophoresis or very low flow-rate reversed-phase HPLC before ionization to fractionate complex peptide mixtures. At the very least, all MS provides precise peptide masses, whereas more sophisticated instruments (particularly tandem MS) also allow peptide sequence determination.

A common approach to rapid MS protein identification is peptide mass mapping. Peptide mapping relies on in-gel digestion of proteins by sequence-specific proteases (ie, trypsin, Asp-N, Lys-C) or chemical reagents (ie, CNBr). Since most amino acid residues have a unique mass, protein digestion will yield a set of distinct peptides specific to each protein. A mass spectrum of eluted peptides results, therefore, in a unique peptide mass fingerprint (PMF). The set of peptide masses obtained by MS is then used to search against protein databases created by "in silico" cleavage of all known, predicted, or partial protein sequences (for review, see Fenyo).34 The efficiency of this technique is such that it has become commonplace for rapid protein identification (noncardiovascular, see Conrads et al.46 Karaoglu et al.46 and Berndt et al.47). For unambiguous protein identification, additional protein sequence information is often required, which can be achieved by tandem MS. A sequence of only five amino acid residues is often sufficient to identify a protein, unless it is obtained from a highly conserved structural or binding motif. In such cases, additional sequence information will be necessary to narrow down the possibilities and unequivocally identify an unknown protein. As MS instruments evolve, it is hoped that they will all eventually be
capable of yielding sequence data, increasing the rigor of protein identification over that of mass fingerprinting.

Application of Proteomics to Cardiovascular Research

How can proteomic analysis benefit the study of cardiovascular disease? Changes to the cardiovascular system arise from or have the potential to alter, collectively or individually, proteomes of cardiac muscle and components of the vascular system, including smooth muscle and endothelial cells. Such changes may be documented through an integrated series of proteomic approaches. Traditional proteomic methods of mapping and identifying proteins give researchers the ability to develop protein databases. Global changes may be identified by protein profiling by comparison to either experimental controls or protein databases or both. Investigation of the molecular mechanisms of proteome change may be carried out by broad-based screening or more focused approaches. In these cases, modified proteins are identified, as is the nature of each modification, and in some cases, the actual site of these modifications may be determined. The next logical step would be to link information about proteome changes directly to functional consequences of these changes. This incorporation of proteome information with functional biochemical or physiological assays may be called “functional” proteomics. Proteomics as a field is just now approaching the level at which full-scale functional proteomic analyses are possible, and unfortunately, there are as yet no published cardiovascular examples of proteomic studies of this magnitude. The following is a description of each of these proteomic approaches.

Traditional Proteomics: Creation of Protein Inventories

Traditional proteomic analyses seek to resolve the entire array of proteins present in a cell during a particular condition or at a given moment in time. Consequently, this provides a protein inventory for that particular set of circumstances. Initial proteomic papers, which focused on 2-DE separation of myocardial tissue proteins, were at best able only to tentatively identify a handful of proteins. Importantly, however, they demonstrated the potential of proteomics—the feasibility of simultaneously separating and detecting hundreds of proteins, with the potential to identify and characterize differences between experimental samples.

Compilation of information obtained from such inventories facilitates protein database construction. For proteomics, these databases ideally should provide a visual image of a 2-D gel from which one may select a protein spot of interest and obtain information about that protein (i.e., M, pl, amino acid sequence if available, function, any PTMs, method of identification, etc) either directly from within the database itself or by cross-referencing to other protein databases (i.e., SWISS-PROT, TrEMBL). Although current 2-D gel analysis software is capable of producing such databases, technical issues with 2-DE made database reliability questionable until only very recently. These issues included difficulties with IEF reproducibility and the inability to resolve all the proteins within a proteome. Two particular aspects influencing total proteome resolution by 2-DE are ineffectiveness at distinguishing lower-abundance proteins, especially in the presence of highly abundant proteins, and underrepresentation of specific classes of proteins (particularly basic and membrane proteins).

The most crucial improvement in IEF reproducibility was the introduction of immobilized pH gradient (IPG) gels. In the past, poor IEF gel-to-gel reproducibility existed both within and between laboratories, for a variety of reasons. IPG gels provide consistently reliable pH gradients, a lack of which was the main drawback to carrier ampholyte–based 2-DE. IPG gels are now commercially available in a wide variety of gradients, from pH 3 to 12, in both broad (pH 3 to 10) and narrow (a variety of single pH unit) “zoom” ranges. Coupling of zoom gels with a subproteomic approach facilitates increased protein resolution and therefore detection of both lower-abundance proteins and protein modifications that might otherwise go undetected.

Detection of both lower-abundance proteins and subtle changes in PTMs is facilitated by reduction of proteome complexity through a subproteomic approach. For example, our recently developed selective extraction method, called “IN Sequence,” enriches for high-abundance myofilament proteins in a single extract. This is particularly useful because fractionation facilitates not only the investigation of myofilament proteins but also that of numerous lower-abundance proteins present in other extracts. Furthermore, this method allowed detection and quantification of a very subtle change in protein phosphorylation (of myosin light chain 1, MLC1), which was otherwise obscured in whole-cell homogenates.

Other recent improvements are now addressing particular classes of proteins that have been problematic for 2-DE. Both basic and membrane proteins have historically been underrepresented on 2-D gels. Basic proteins are difficult to focus during IEF because of electroendosmotic effects at high pH and the resulting cathodic drift within the gel. Recent efforts have been applied to overcome this, with strategies ranging from addition of organic solvents to reduce cathodic drift to subproteomic enrichment of basic proteins by selective precipitation. Another problem with basic proteins, as well as membrane and membrane-associated proteins, is their tendency to aggregate during protein separation. This arises from their extreme hydrophobicity. Progress has recently been made in 2-D analysis of such proteins by differential combinations of sample treatment, fractionation, and detergent application (for review, see Pasquale et al; noncardiovascular, see Santoni et al). Although not yet applied to cardiovascular research, these advances may prove useful in the study of proteins associated with, for example, Ca handling, mitochondria, second messenger signaling transduction cascades, etc. An interesting study by Macri and colleagues demonstrated that altering solubilization conditions improved the resolution and detection of membrane proteins from cardiac sarcoplasmic reticulum and sarcosomal fractions, while having little effect on whole-tissue homogenates. These differences in ability to detect these proteins were apparent only when membrane proteins were already enriched (by use of subproteomics). Despite this advance, not
all membrane proteins were detected by 2-DE (ie, sarcoplasmic/endoplasmic reticulum Ca\textsuperscript{2+}-ATPase, SERCA\textsuperscript{17}), indicating that further refinement of such techniques is still required and further exemplifying the highly complex considerations that one must deal with in the design and implementation of a proteomic study.

Many of these advances, however, are contributing toward improved reliability of 2-D gel protein databases. Pioneering proteomic work by the laboratories of Dunn and Jungblut led to the creation of numerous online 2-D databases of human, dog, mouse, and rat myocardium.\textsuperscript{65–75} They are only partially complete (with roughly 200 proteins identified), but they provide a foundation for the inventory of these particular tissues. These freely accessible works in progress are tremendously important, because they provide researchers with a basis for visualization of changes in protein patterns resulting from the conditions of their particular study. For proteomics to fulfill its potential, comprehensive protein inventories must be prepared for the variety of species and tissues studied in cardiovascular research, and they must remain freely accessible.

**Protein Profiling or Protein Signatures**

For certain applications, protein identification is not always a necessity. Simply monitoring a protein profile after resolution by one or more separation methods is often sufficient to address whether two or more experimental conditions induce the same protein changes. This is advantageous when determining, for example, molecular pathways of action by multiple drugs in pharmaceutical drug discovery programs. Direct comparison of their protein profiles or “protein signatures” facilitates rapid screening of differences between various treatments without an absolute necessity for protein identification (for review, see Steiner and Witzmann\textsuperscript{76}). The assumption underlying this approach is that there is a consistent and predictable pattern of protein change or modification associated with every particular cellular event or phenotype. Testing of the validity of phenotype consistency will, with time, be achieved through proteomics. Such an approach may, of course, also be applied to comparison of drug-treated versus untreated, mutant versus wild-type, or healthy versus diseased samples.

One recent technical advance that will help in this approach is the introduction of fluorescent dyes, which have sensitivity similar to that of silver stain, but with a greater dynamic linear staining range (for review, see Patton\textsuperscript{77}). Unlike silver, however, the use of two or more dyes that fluoresce at different wavelengths provides the opportunity to coelectrophorese differentially stained samples. Direct comparison of samples may thus be achieved within the same gel, reducing both the subjectivity of gel matching and the time required for analysis. Such a method is applicable regardless of experimental question.

**Molecular Mechanisms of Cardiovascular Disease**

Although traditional proteomics and protein profiling provide important research information, the ultimate goal of developing proteomic techniques for cardiovascular research lies in its ability to characterize molecular mechanisms of disease. In this case, identification of the protein and the nature of its modification are essential. Proteome variations that one might expect depend on characteristics of the particular study being conducted. Altered levels of more than forty proteins have been identified to date in a variety of cardiovascular proteomic experiments (Table).

For chronic conditions or disease states, modification of the proteome often manifests itself as altered protein levels due to specific gene upregulation or downregulation, isof orm switching, or de novo protein synthesis. For example, the abundance of heat-shock proteins (HSP) and mitochondrial and other proteins involved in energy production was documented to vary significantly in a number of heart failure models.\textsuperscript{78–84} Isoform switching has also been reported in the case of ventricular expression of atrial MLC1 and MLC2 in neonatal rat myocytes from a phenylephrine-stimulated hypertrophy model.\textsuperscript{85} De novo synthesis of HSP72 and two unknown proteins related to the HSP70 family was shown to occur in heat-stressed endothelial cells.\textsuperscript{81} In addition, a number of putative PTMs have been documented in dilated cardiomyopathy,\textsuperscript{85,66,68,78,79,82–84} For example, detailed 2-DE analysis of dilated cardiomyopathy--diseased human myocardial tissue revealed more than fifty HSP27 protein species by immunoblotting.\textsuperscript{56,82} Although only nine were finally confirmed to belong to HSP27 and none were analyzed for the presence of PTMs, this illustrates the potentially large number of PTMs possible for a single protein.

In acute conditions, in which there is often insufficient time to recruit de novo transcription and translation, PTMs would be the primary mechanism of protein change, resulting from modification of specific amino acids. To date, cardiovascular proteomic examples of PTM identification are lacking. A prime example of such a study is the identification of phosphorylation and palmitoylation in a membrane-receptor signal transduction cascade.\textsuperscript{85} Of course, one should not overlook the possibility of variations in protein levels evinced by protein degradation as a response to acute injury, such as with troponin I in myocardial stunning.\textsuperscript{86,87}

Molecular mechanisms are addressed in one of two ways: by broad-based screening or by a more focused proteomic approach. Broad-based screening is applicable in situations in which little is known of the molecular mechanisms or for an unbiased look at the entire proteome to identify previously unknown protein changes involved in the disease or condition. This extends traditional or protein profiling approaches to the next logical step, identification of the modifications. Cardiovascular examples of broad-based screening include extensive studies on human and bovine cardiomyopathies, as well as many of the studies mentioned above that document changes in protein levels (see Table). Focused proteomics, conversely, analyzes only a discrete subproteome. This may be applied to situations in which a molecular mechanism is understood but some components of this mechanism are not. One such example, in cardiovascular research, was carried out by use of protein kinase C (PKC) monoclonal antibodies to immunoprecipitate proteins involved in PKC signaling cascades.\textsuperscript{35} This recent study identified a large number of previously unsuspected proteins that may be downstream targets of PKC signaling during myocardial preconditioning.
## Protein Changes Documented by Proteomic Studies in Cardiovascular Disease

<table>
<thead>
<tr>
<th>Protein Identified</th>
<th>Protein Change</th>
<th>Model</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proteins associated with energy/metabolism</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. ATPase synthase</td>
<td>Increase</td>
<td>Human DCM 65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decrease</td>
<td>Dog DCM 78</td>
<td></td>
</tr>
<tr>
<td>b. Cytochrome c oxidase VA</td>
<td>Decrease</td>
<td>Dog DCM 79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decrease</td>
<td>Bovine DCM 84</td>
<td></td>
</tr>
<tr>
<td>c. Cytochrome b₅</td>
<td>Decrease</td>
<td>Dog DCM 79</td>
<td></td>
</tr>
<tr>
<td>d. Fatty acid–binding protein</td>
<td>Decrease</td>
<td>Dog DCM 78</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bovine DCM 84</td>
<td></td>
</tr>
<tr>
<td>e. 3,2-trans-Enoyl-CoA-isomerase</td>
<td>Variable decrease</td>
<td>Dog DCM 79</td>
<td></td>
</tr>
<tr>
<td>f. Dihydrolipoamide dehydrogenase precursor</td>
<td>Decreased</td>
<td>Dog DCM 78</td>
<td></td>
</tr>
<tr>
<td>g. Isocitrate dehydrogenase</td>
<td>Decreased</td>
<td>Bovine DCM 84</td>
<td></td>
</tr>
<tr>
<td>h. Creatine kinase M-chain</td>
<td>Minor decrease</td>
<td>Dog DCM 78, 79, 84</td>
<td></td>
</tr>
<tr>
<td>i. Hydroxymethyl glutaryl CoA synthase</td>
<td>Variable decrease</td>
<td>Dog DCM 79</td>
<td></td>
</tr>
<tr>
<td>j. NADH ubiquinone oxidoreductase (75-kDa subunit)</td>
<td>Decrease</td>
<td>Rat phen. hypertrophy 80</td>
<td></td>
</tr>
<tr>
<td>k. Ubiquinol cytochrome c reductase core protein 1</td>
<td>Decrease</td>
<td>Bovine DCM 84</td>
<td></td>
</tr>
<tr>
<td>l. Pyruvate dehydrogenase (E1 component)</td>
<td>Increase</td>
<td>Dog DCM 79</td>
<td></td>
</tr>
<tr>
<td>m. Triosephosphate isomerase</td>
<td>Increase/variable: 3 possible PTMs</td>
<td>Dog DCM 79, 79</td>
<td></td>
</tr>
<tr>
<td>n. Phosphoglycerate mutase (muscle)</td>
<td>Variable increase</td>
<td>Dog DCM 79</td>
<td></td>
</tr>
<tr>
<td>o. Isocitrate dehydrogenase (mitochondrial subunit a)</td>
<td>Decrease</td>
<td>Dog DCM 78</td>
<td></td>
</tr>
<tr>
<td>p. Elongation factor Tu (mitochondrial preP43)</td>
<td>Not present</td>
<td>Dog DCM 78</td>
<td></td>
</tr>
<tr>
<td>q. Mitochondrial thioredoxin-dependent peroxide reductase</td>
<td>Decrease</td>
<td>Bovine DCM 84</td>
<td></td>
</tr>
<tr>
<td><strong>Stress-induced/heat-shock proteins (possible protective proteins)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. HSP 72</td>
<td>Novel expression</td>
<td>Heat-shock cells 81</td>
<td></td>
</tr>
<tr>
<td>b. HSP 70 (inducible)</td>
<td>Decrease</td>
<td>Dog DCM 78, 79</td>
<td></td>
</tr>
<tr>
<td>c. Mitochondrial HSP70 precursor</td>
<td>Decrease</td>
<td>Bovine DCM 84</td>
<td></td>
</tr>
<tr>
<td>d. Mitochondrial stress protein (HSP70-related)</td>
<td>Increase</td>
<td>Dog DCM 79</td>
<td></td>
</tr>
<tr>
<td>d. HSP 60</td>
<td>Decrease</td>
<td>Dog DCM 78</td>
<td></td>
</tr>
<tr>
<td>e. Mitochondrial matrix protein p1 (membrane-bound HSP60)</td>
<td>Decrease: 2 possible PTMs</td>
<td>Rat phen. hypertrophy 80</td>
<td></td>
</tr>
<tr>
<td>f. Chaperon cofactor a</td>
<td>Increase</td>
<td>Rat phen. hypertrophy 80</td>
<td></td>
</tr>
<tr>
<td>g. αB-Crystallin</td>
<td>Increase: possible PTM</td>
<td>Human and dog DCM 65, 68, 78</td>
<td></td>
</tr>
<tr>
<td>h. HSP 27</td>
<td>Increase: possible PTM</td>
<td>Human DCM 65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Increase</td>
<td>Rat phen. hypertrophy 80</td>
<td></td>
</tr>
<tr>
<td>i. Unknown 34- and 40-kDa (related to HSP70 family)</td>
<td>Novel expression</td>
<td>Heat-shock cells 81</td>
<td></td>
</tr>
<tr>
<td><strong>Proteins involved in contraction/cytoskeleton</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Desmin</td>
<td>Decrease in intact: increase in low Mr</td>
<td>Dog DCM 78, 79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Increase</td>
<td>Bovine DCM 84</td>
<td></td>
</tr>
<tr>
<td>b. Actin</td>
<td>Increase: possible PTM</td>
<td>Dog DCM 79</td>
<td></td>
</tr>
<tr>
<td>c. MLC2 (ventricular isoform)</td>
<td>Variable: possible PTM</td>
<td>Human DCM 65, 83</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Increase: atrial isoform expression: possible PTM</td>
<td>Rat phen. hypertrophy 80</td>
<td></td>
</tr>
<tr>
<td>d. MLC1 (atrial isoform)</td>
<td>Increase: possible PTM</td>
<td>Rat phen. hypertrophy 80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Variable</td>
<td>Human DCM 83</td>
<td></td>
</tr>
<tr>
<td>e. SERCA 2A</td>
<td>Decrease</td>
<td>Dog DCM 78</td>
<td></td>
</tr>
</tbody>
</table>
Table continued:

<table>
<thead>
<tr>
<th>Protein Identified</th>
<th>Protein Change</th>
<th>Model</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Farnesyl diphosphate farnesyl transferase</td>
<td>Variable decrease</td>
<td>Dog DCM</td>
<td>79</td>
</tr>
<tr>
<td>b. α-1-Antiproteinase precursor</td>
<td>1 form increase/1 form decrease: possible PTM</td>
<td>Bovine DCM</td>
<td>84</td>
</tr>
<tr>
<td>c. Galactoside 3(4)-l-fucosyltransferase</td>
<td>Decrease</td>
<td>Dog DCM</td>
<td>79</td>
</tr>
<tr>
<td>d. IgG Fc receptor II</td>
<td>Increase</td>
<td>Dog DCM</td>
<td>79</td>
</tr>
<tr>
<td>e.Isovaleryl CoA dehydrogenase</td>
<td>Decrease</td>
<td>Bovine DCM</td>
<td>84</td>
</tr>
<tr>
<td>f. ras-Related protein</td>
<td>Increase</td>
<td>Dog DCM</td>
<td>79</td>
</tr>
<tr>
<td>g. Syndecan-2</td>
<td>Decrease</td>
<td>Dog DCM</td>
<td>79</td>
</tr>
<tr>
<td>h. Major allergen can F1</td>
<td>Decrease</td>
<td>Dog DCM</td>
<td>79</td>
</tr>
<tr>
<td>i. Cystatin c</td>
<td>Decrease</td>
<td>Dog DCM</td>
<td>79</td>
</tr>
<tr>
<td>j. Ubiquitin carboxyl terminal hydrolase</td>
<td>Increase</td>
<td>Bovine DCM</td>
<td>84</td>
</tr>
<tr>
<td>k. Nucleoside diphosphate kinase A</td>
<td>Increase</td>
<td>Rat phen. hypertrophy</td>
<td>80</td>
</tr>
<tr>
<td>l. Inosine 5'-monophosphate dehydrogenase</td>
<td>Decrease</td>
<td>Dog DCM</td>
<td>79</td>
</tr>
</tbody>
</table>

DCM indicates models of dilated cardiomyopathy including human myocardium from an explanted/transplanted or donor heart\cite{65,83}; fast pacing–induced failure dog model\cite{69,70}; and genetic bovine model\cite{84}; rat phen. hypertrophy, phenylephrine-stimulated hypertrophy of cultured neonatal rat myocytes were compared with control untreated cells\cite{80}; heat-shock cells, cultured endothelial cells that underwent heat shock were compared with control untreated cells.\cite{81}

Documented protein changes observed in “proteomic broad-based screening” studies involving (1) simultaneous analysis of ten or more proteins, (2) used 2-D gel electrophoresis, and (3) identification of proteins by immunological, amino acid composition, N-terminal amino acid sequencing, and/or MS.

Protein Identification

PMF protein identification follows an involved series of steps after protein separation. After complete in-gel protease digestion or chemical cleavage of a protein spot and full recovery of peptides from the gel (when 2-DE is the separation method of choice), peptide mass determination is carried out by MS. Accuracy of MS measurement increases the likelihood of proper matching of peptide fragments during subsequent bioinformatic protein database searches. Because protein matching depends on a minimum number of matching peptide mass fragments and therefore a certain proportion of total amino acid sequence coverage, the ability to optimize all these conditions dictates the likelihood of PMF protein identification.\cite{82}

As described in the Techniques section, there are a number of commonly used exoproteases and endoproteases for in-gel digestion, as well a number of technical tips to improve peptide digestion and recovery.\cite{88,90} For optimization of PMF, it is important to note that the choice of protease or combination of proteases can have a dramatic effect on coverage of protein sequence. For example, human myocardial MLC1 sequence coverage was 56% with Asp-N digestion, compared with 44% for trypsin, whereas a combination of both Asp-N and trypsin increased coverage to 80%, thereby increasing the probability of correct protein identification.\cite{89} Multiple protease digestion, however, may not always be beneficial. For example, the specificity of trypsin and/or Lys-C would lead to far too small and far too numerous peptide fragments if used to digest troponin I, whereas a partial digest or a single cleavage by Asp-N does allow correct identification. Thus, optimal conditions for fragmentation of a particular protein must be determined empirically and are not always inherently obvious.

PMF protein identification is also complicated in the study of species underrepresented in protein databases. This is often the case with cardiovascular proteomic studies, because some species traditionally used as cardiovascular models (rabbit, dog, and swine) currently face this problem. This harkens back to the importance of protein inventories. For underrepresented species, one needs to maximize amino acid sequence (protein) coverage, which may necessitate multiple PMFs. Low amino acid sequence homology between proteins present in databases and the protein under examination may further complicate cross-species identification. In this case, greater sequence coverage is also beneficial because it increases the likelihood of matching a highly conserved region of the protein.\cite{92} Database searches based on newly developed
algorithms for protein matches based on weak amino acid sequence homologies\textsuperscript{93} should improve the ability to positively identify such proteins.

Even the best strategy may lead to false-positive protein identification, for a number of reasons. Partial amino acid sequences obtained by MS may be too short or not specific enough (ie, in nonconserved regions of the protein). Database size may be insufficient, with either few representatives of the protein or an absence of that protein sequence for a particular species of interest. A small number of detected peptide fragments would result in poor sequence coverage, a problem that may be exacerbated by the presence of PTMs, which might also contribute to decreased numbers of matching peptides. Also, if peptide masses were not obtained within a reasonable peptide mass tolerance, results might be misleading, suggesting incorrect protein identity, or inconclusive, with no likely matches obtained. Finally, previously unidentified proteins would be absent from any protein database. Therefore, even with the advantages of MS, confidence in protein identification is supported by means of additional identification methods.

Protein quantification is important in assessment of changes in gene expression resulting from disease or a particular cellular intervention. One new technology, which allows accurate quantification by MS, is stable isotope labeling. This ability to quantify samples represents an exciting new dimension for MS. The method involves growing cultured cells in two differently labeled media,\textsuperscript{94} or more recently, by labeling after sample preparation, through the use of isotope-code affinity tag peptide labeling.\textsuperscript{95} Proteins from two different experimental conditions are reacted with the isotope-code affinity tag reagent (consisting of biotin and a specific cysteine [thiol] reactive group), one labeled with hydrogen ions, the other with deuterium ions, to produce a mass difference. With or without further protein separation, a particular protein (or group of proteins) is digested, and the biotin-containing cysteine-derivatized fragments are isolated via avidin affinity. The ratio of the light to heavy peptide fragments determined by MS indicates differences in the quantity of that particular protein between the experimental conditions. Although this is the first flexible method for quantitative MS, the technique is limited in terms of PTM identification. Because protein derivatization occurs after sample harvesting, it may lead to further PTMs that were not the result of original experimental conditions.

**Identification and Characterization of PTMs**

PTM identification and characterization have been lacking within the field of proteomics. Different strategies may be used to investigate PTMs, including MS and a number of classic biochemical techniques. Over the years, biochemical methods that have been exploited include identification of glycosylated proteins using periodic acid/Schiff staining, lectin binding, or enzymatic deglycosylation (eg, see Packer et al\textsuperscript{96} and Ianello and Jeffrey\textsuperscript{97}), whereas phosphorylated proteins have been identified by radiolabeling of cells with \textsuperscript{32}P\textsuperscript{98–101} enzymatic dephosphorylation,\textsuperscript{59} or the use of antiphosphospecific antibodies.\textsuperscript{102} Proteolysis is another PTM that may play a role in the phenotype of numerous cardio-vascular diseases. For example, determination of N-terminal cleavage sites is possible by direct Edman sequencing, as used in the identification of MLC1 proteolysis at amino acid residue 19 resulting from severe ischemia/reperfusion.\textsuperscript{103}

Relative to standard biochemical methods, the use of MS for PTM identification is powerful because it can provide, in a single step, information about the type of modification as well as mapping the modified amino acid residue. In addition, whereas various biochemical methods are individually geared toward identifying one form of PTM, bioinformatic analysis of MS PMF data may be carried out for concurrent identification of up to twenty-two known PTMs or may be used to predict possible modifications on the basis of amino acid sequence homology. Several web-based programs are available to assist in identifying PTMs (ie, MODFIND,\textsuperscript{104} BOLD,\textsuperscript{105} Net-Phos,\textsuperscript{106}) and new ones are continually being introduced (eg, GlycoSuite\textsuperscript{107}).

Since pI shifts of protein spots on 2-D gels may also be indicative of PTM, determination of the modifications produced by these shifts in the myocardium (or sample of interest) are essential for revealing their roles in physiological function. Ultimately, the type of modification and the modified amino acid residue(s) must be determined, through either biochemical techniques, amino acid sequencing, or MS analysis. In the case of MS, the sample must be analyzed by an instrument that further fragments peptides, allowing their analysis at the amino acid level. This may now be done by a MALDI-TOF-PSD (postsource decay), but many researchers rely on the more complex LC-nESI triple quadrupole or the hybrid MALDI quadrupole TOF (ie, Q-TOF). The more sophisticated the MS instrument, the greater the likelihood of accurately identifying a specific PTM on a particular protein. Even though these experiments are very time-consuming, they are absolutely critical for determination of underlying molecular mechanisms of protein change.

**Functional Proteomics**

Today we are only on the threshold of proteomics. Continuing technological developments and introduction of new approaches are leading to functional proteomic studies: a conjunction of functional data from established biochemical and physiological methods with proteome information of the cellular or organ phenotype. Because any phenotype is the sum of the contribution of all proteins present in the cell, understanding of the precise role of each individual protein modification requires detailed physiological analysis. These analyses may range from in vitro biochemical studies to production of transgenic mice. Full-scale functional proteomics is thus a highly complex process.

To completely understand how every protein modification either contributes to or plays a role in an observed phenotype, it is important to delineate the exact sequence of events that occur. A time course of protein change with measured alterations in physiological and biochemical parameters allows a reconstruction of the events leading up to and/or resulting from an altered phenotype. As a result, it may be possible to determine exact functional changes arising from specific protein modifications or of specific protein modifi-
cations arising from a particular functional change. In this way, functional proteomics may reveal whether a protein modification is the cause or the result of a particular disease process. To add to the complexity of the issues involved in a functional proteomic study, it must also be mentioned that a single observed phenotype may arise from multiple pathways. For example, an observed systolic contractile dysfunction might arise as a result of modification of calcium-handling proteins, of myofilament proteins, or some combination of the two. We hope that this provides some idea of the daunting logistics of incorporating functional data into the already complex process of conducting a proteomic analysis, which we hinted at earlier. It will not be easily amenable for a single laboratory to tackle all aspects of a functional proteomic study on its own. Therefore, the feasibility of such studies will most likely require collaboration between laboratories involved in the biochemistry, physiology, and pathophysiology of cardiovascular disease that are properly equipped to provide specific individual components of an integrated functional proteomic effort.

Incorporation of proteomics within this established cardiovascular research framework provides a means of identifying and characterizing complex protein changes associated not only with cardiovascular dysfunction but also with pharmacological interventions taken in response to dysfunction. Proteomics has clearly revealed dynamic and remarkable changes that occur at the cellular level, many of which have never previously been observed. Successful incorporation of new technologies and the coordinated application of the variety of proteomic approaches now available will help to unravel the intricacies of proteome change in cardiovascular disease.

Acknowledgments

Dr Van Eyk receives funding from the Canadian Institutes of Health Research (grant-in-aid 49843) and the Ontario Heart and Stroke Foundation (grant-in-aid T-3759). Dr Van Eyk is a Heart and Stroke Foundation of Canada Scholar, and D. Kent Arrell is an Ontario Graduate Scholar in Science and Technology.

References


Cardiovascular Proteomics: Evolution and Potential
D. Kent Arrell, Irina Neverova and Jennifer E. Van Eyk

Circ Res. 2001;88:763-773
doi: 10.1161/01.RES.88.8.763

Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2001 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://circres.ahajournals.org/content/88/8/763

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published
in Circulation Research can be obtained via RightsLink, a service of the Copyright Clearance Center, not the
Editorial Office. Once the online version of the published article for which permission is being requested is
located, click Request Permissions in the middle column of the Web page under Services. Further information
about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation Research is online at:
http://circres.ahajournals.org/subscriptions/