Proteomics in Cardiovascular Biology and Medicine

Joseph Loscalzo, MD, PhD

The sequencing of the human genome is undoubtedly one of the major accomplishments of biomedical science.1,2 Knowledge of the precise sequence of all human genes provides unparalleled access to a complete understanding of human biology in all of its complexity. Or does it? The completion of the human genome sequence showed far fewer genes than expected, 30,000, which is not that different from that of the lowly earthworm, with at least 17,300.3 This number of genes was viewed with surprise by some investigators as too few to account for human biological diversity in form and function. The genome-centric view of the biological universe, however, is rather myopic and fails to take into consideration what protein chemists have known for almost a century. Proteins, the ultimate products of the human genome, define biology. Proteins are directly responsible for all biological form and function. It is estimated that there are ~6 to 7 times as many distinct proteins (ie, ~200,000) as genes in humans, in part owing to splicing and exchange of various structural cassettes among genes during transcription. Proteins are not only more abundant than the genes that encode them, but they are also much more structurally complex with primary, secondary, tertiary, and quaternary structural elements; addition-
ally, they have greatly varied biochemical functions that critically depend on structure. Furthermore, mature proteins are also subject to a host of post-translational modifications, including proteolysis, sulfhydryl oxidation and disulfide bond formation, phosphorylation, glycosylation, S-nitrosation, fatty acylation, and oxidation. These biochemical modifications often yield products with functions different from those of the unmodified parent protein, and many of these modifications, such as oxidation, reflect the consequences of environmental modulation of genetic determinants. Taken together, the many post-translational changes in protein structure and function add incredible complexity to that of the basic genome and constitute what has become the protein equivalent of the genome, namely, the proteome. In parallel to genomics, proteomics is, thus, defined as the sequence, modification, and function of all proteins in a biological system.

Cardiovascular Medicine and the Proteome

Why does the cardiologist need to know about this technically complex and conceptually arcane field of biology? The main reason for reviewing this topic is that the field of proteomics, although relatively new and rapidly changing, truly has the potential to revolutionize how we diagnose disease, assess risk, determine prognosis, and target therapeutic strategies among individuals with cardiovascular disease. In addition and importantly, it defines a rational biological basis for linking the consequences of environmental factors directly to gene products. Simple examples will illustrate this point. The reason that some individuals who smoke have extensive atherothrombotic disease whereas others do not may lie in the proteome, not the genome, and the susceptibility of specific proteins to oxidative modification by specific components of cigarette smoke. Oxidation of low-density lipoprotein, a key to understanding atherogenesis, is reflected in the proteome, not the genome, because oxidative modifications of the amino acid side chains of apoprotein B (as well as of lipid components of the particle) define this atherogenic species. In hyperhomocysteinemia, mixed disulfide bond formation between cysteinyl side chains of proteins and homocysteine may modify protein function and contribute to the risk of atherothrombosis. In diabetics, chemical modifications of proteins by glycation (eg, hemoglobin A1c), oxidation, and glycoxidation (eg, advanced glycation end products) are important determinants of the adverse clinical consequences of hyperglyce-

From the Whitaker Cardiovascular Institute and Evans Department of Medicine, Boston University School of Medicine, Boston, Mass.

Correspondence to Joseph Loscalzo, MD, PhD, Boston University School of Medicine, Whitaker Cardiovascular Institute, 700 Albany St, W507, Boston, MA 02118. E-mail jloscalz@bu.edu

(Circulation. 2003;108:380-383.)

© 2003 American Heart Association, Inc.

Circulation is available at http://www.circulationaha.org

DOI: 10.1161/01.CIR.0000079867.56212.17

380
mic states, including diabetic vascular disease.

Human proteomics is in its infancy, yet it will undoubtedly prove to be the true key to understanding human biology and disease. To expect knowledge of the human genome alone to offer this detailed insight is as simplistic as expecting that the floor plan of a house will adequately describe the actual house in all of its complexity—for example, the appearance of the kitchen cabinets changing over time with oxidation of the wood grain, the nuance of shadow altering the sense of space in the family room as the sun sets, etc. If the genome is the floor plan, the proteome is the house. This comment is not hyperbole. Rather, to one trained in protein chemistry, the complexity of biological systems is a consequence of proteins interacting with proteins, proteins changing conformation and function by interacting with cofactors, proteins changing structure and function by undergoing covalent post-translational modifications, and proteins interacting with networks of proteins in stoichiometrically and geometrically defined patterns. Networks of proteins working in concert, simultaneously or ad seriatum, comprise metabolic or other functional pathways essential for cell and organism behavior. Taken together, variations in these complex pathways define cell and organism phenotype, underlie biodiversity, and serve as the basis for disease susceptibility and pathotype. Neither DNA sequence nor mRNA expression analysis can substitute for an analysis of the proteome. One might argue that analysis of all mRNA species should offer similar information; however, the correlation coefficient for the abundance of mRNA and the corresponding abundance of protein into which it is translated is not greater than 0.48.4 This, in part, can be explained by differences in the half-lives of an mRNA and its corresponding protein. Furthermore, mRNA analysis cannot provide any information on the abundance of proteolyzed or post-translationally modified polypeptides, let alone the function of these molecular species. The relationships among the elements of the genome and the proteome are illustrated in Figure 1.

**Analysis of the Proteome**

To illustrate the true power of proteomics in cardiovascular disease, we need to review briefly the currently available methods for analyzing the proteome. Before proceeding, it is also important to point out that this technology is changing very rapidly and will increase dramatically in sensitivity, specificity, throughput, and analytical impact over the next several years.

There are 5 basic elements to any proteomic analysis: sample acquisition, protein extraction, protein separation, protein sequence determination, and sequence comparison to referential databases for protein identification (Figure 2). Sample acquisition can be as straightforward as obtaining a plasma sample or tissue biopsy from an individual, or it may involve the rather precise removal of a cell or cluster of cells from a biopsy specimen by using laser-capture microdissection methods. Once the plasma, tissue, or cell sample has been obtained, the specimen is subjected to conventional methods of protein analysis with the intention of removing all DNA, RNA, carbohydrates, and lipids. Typically, this step is accomplished by chemical extraction, generally with methanol.

Extracted proteins must next be separated for further identification. Conventionally, this step has been accomplished by 2-dimensional gel electrophoresis. In the first dimension, proteins are separated according to molecular mass, whereas in the second dimension, they are separated according to isoelectric point (or net charge). When first developed, 2-dimensional protein gel electrophoresis was believed to provide unique and unequivocal protein separation with each “spot” on the gel corresponding to a single protein. Subsequent analysis us-
ing highly sensitive mass spectrometry techniques, however, has shown that this view is incorrect, and many if not most spots on a 2-dimensional gel contain multiple protein constituents. Furthermore, sensitivity is also limited by the staining method used to detect protein spots on the gel. Thus, owing to the lack of sensitivity and specificity, alternative methods for separating and identifying separated proteins have evolved. Many of these involve liquid chromatographic methods, which utilize solid- and liquid-phase media to separate proteins according to specific biochemical properties, such as molecular mass, isoelectric point, or hydrophobicity. These liquid chromatographic separations can be performed in series to improve resolving power. Furthermore, if one is interested in a specific class of proteins (such as those bearing a sulfhydryl group) or in a specific post-translational modification (such as phosphorylation), unique columns that contain antibodies specific for these functionalities can be used to separate these groups of proteins from all others by so-called affinity chromatography. For many of the liquid chromatography approaches, proteins are subjected to proteolytic digestion (typically with trypsin) to afford a multitude of peptides derived from each protein.

Once separated, the resulting peptides require identification. Currently available methods all use some form of mass spectrometry. Mass spectrometry is a rapidly evolving methodology that converts proteins or peptides to charged species that can be separated on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry is a rapidly evolving methodology that converts proteins or peptides to charged species that can be separated on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002. Mass spectrometry requires that proteins or peptides are first converted to gas-phase ions within a specific region of the instrument, the ionization source. The ions are then separated with a mass analyzer on the basis of their mass-to-charge ratio (m/z). These methods have been considered a major advance in the identification of polypeptides from the proteome, and for their contributions to this analytical field, John Fenn and Koichi Tanaka shared the Nobel Prize in Chemistry in 2002.
trypsin to yield a tractable number of smaller peptides. These peptides are then subjected to either electrospray ionization or MALDI mass spectrometry. The resulting charged peptides that are detected in this phase of the analysis can next be subjected to high-speed collision with an inert gas, such as argon, yielding smaller-sized charged fragments that can be pieced together, akin to a jigsaw puzzle, to reconstruct peptide sequence. Peptide sequences identified with these methods must next be analyzed by comparison with known database sequences to determine the unequivocal identity of the protein. An example of the application of these methods is presented in Figure 3.

Proteomic analysis is currently limited by sensitivity, specificity, and throughput. Sensitivity is rapidly improving, with detection at the attomole (10^-18 mol) level achieved by current mass spectrometry methods, although this benchmark is not yet routine. Specificity continues to improve, especially with application of serial liquid chromatographic methods in place of 2-dimensional protein gel electrophoresis. Throughput remains a problem in that conventional mass spectrometers must still sequence peptide ions one at a time; however, some newer devices are designed to accommodate multiplexing of samples.

Once proteins in a given proteome have been identified, their relative abundance levels need to be determined, especially if a purpose of the experiment is to determine the comparative abundance of a protein or proteins in normal and diseased states. To determine relative abundance, novel methods have been developed using stable isotope tags (isotope-coded affinity tag, or ICAT method) that react with specific functional groups. Using 2 tags of different mass, 1 of natural abundance and the other isotopically labeled, quantitative differences in abundance of proteins between 2 different samples can be readily determined.

Using these methods, increasing numbers of proteins whose functions are not yet known have been identified. Thus, a thorough analysis of the proteome should include some measure of function using cultured cells capable of synthesizing the protein of interest. This approach is, of course, akin to functional genomic analysis and its importance in gauging the significance of a gene of unknown function or a mutation or polymorphism of uncertain effect.

Conclusions

Applications of proteomics to medicine are limited at the current time, but are rapidly evolving. Cancer biologists have made the first attempts to utilize proteomics for diagnostic and prognostic purposes. Studies of proteomic patterns in sera of patients with breast cancer provide an example of this strategy. One need not know the function of a protein using this approach; rather, proteomic patterns in sera offer the possibility of identifying simple associations for diagnosis, prognosis, and response to therapies.

The application of proteomics to cardiovascular disease holds great promise. It will very likely be the case in a few short years that analysis of a simple plasma sample will provide unique prognostic information about an individual’s risk of atherothrombotic disease or heart failure or about the prognosis of these disorders once established. Similarly, proteomic analysis of a myocardial biopsy specimen may provide useful prognostic information in patients with unexplained heart failure or in cardiac transplant recipients. In addition, therapeutic strategies in these diseases can be tailored to the overabundance, deficiency, or altered function of a specific protein or proteins. Most importantly, understanding the proteome will offer the real possibility of understanding the function of the cardiovascular system in all of its complexity, leading to possible approaches for systems modification to correct dysfunction or enhance basal function.

Acknowledgments

This work is supported in part by grants HL61828 and HL58976, the Specialized Center of Research in Ischemic Heart Disease grant P01HL55993, and the Cardiovascular Proteomics Center grant N01HV28178, all from the National Institutes of Health. The author thanks Drs Catherine Costello, Jane Freedman, Diane Handy, and Steven Naylor for helpful suggestions and assistance with figures, and Stephanie Tribuna for expert secretarial assistance.

References

Proteomics in Cardiovascular Biology and Medicine
Joseph Loscalzo

doi: 10.1161/01.CIR.0000079867.56212.17
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2003 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/108/4/380

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation 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 is online at:
http://circ.ahajournals.org//subscriptions/