Chapter 1   The origin and scope of proteomics

1.1 INTRODUCTION 1

1.2 THE BIRTH OF LARGE-SCALE BIOLOGY AND THE "OMICS" ERA 1

1.3 THE GENOME, TRANSCRIPTOME, PROTEOME, AND METABOLOME 6

1.4 FUNCTIONAL GENOMICS 8

Transcriptomics is the systematic, global analysis of mRNA 8
Large-scale mutagenesis and interference can also determine the functions of genes on a global scale 11

1.5 THE NEED FOR PROTEOMICS 15

1.6 THE SCOPE OF PROTEOMICS 17

Protein identification and quantitation are the most fundamental aspects of proteomic analysis 17
Important functional data can be gained from sequence and structural analysis 18
Interaction proteomics and activity-based proteomics can help to link proteins into functional networks 19

1.7 CURRENT CHALLENGES IN PROTEOMICS 20

Chapter 2   Strategies for protein separation

2.1 INTRODUCTION 23

2.2 GENERAL PRINCIPLES OF PROTEIN SEPARATION IN PROTEOMICS 23

2.3 PRINCIPLES OF TWO-DIMENSIONAL GEL ELECTROPHORESIS 25

Electrophoresis separates proteins by mass and charge 25
Isoelectric focusing separates proteins by charge irrespective of mass 26
SDS-PAGE separates proteins by mass irrespective of charge 28

2.4 THE APPLICATION OF 2DGE IN PROTEOMICS 29

The four major advantages of 2DGE are robustness, reproducibility, visualization, and compatibility with downstream microanalysis 29
The four major limitations of 2DGE are resolution, sensitivity, representation, and compatibility with automated protein analysis 30
The resolution of 2DGE can be improved with giant gels, zoom gels, and modified gradients, or by pre-fractionating the sample 30
The sensitivity of 2DGE depends on the visualization of minor protein spots, which can be masked by abundant proteins 31
The representation of hydrophobic proteins is an intractable problem reflecting the buffers required for isoelectric focusing 32
Downstream mass spectrometry requires spot analysis and picking 34

2.5 PRINCIPLES OF MULTIDIMENSIONAL LIQUID CHROMATOGRAPHY 34

Protein and peptide separation by chromatography relies on differing affinity for stationary and mobile phases 34
Affinity chromatography exploits the specific binding characteristics of proteins and/or peptides 36
Size exclusion chromatography sieves molecules on the basis of their size 36
Ion exchange chromatography exploits differences in net charge 37
Reversed-phase chromatography and hydrophobic interaction chromatography exploit the affinity between peptides and hydrophobic resins 38

2.6 MULTIDIMENSIONAL LIQUID CHROMATOGRAPHY STRATEGIES IN PROTEOMICS 39

Multidimensional liquid chromatography is more versatile and more easily automated than 2DGE but lacks a visual dimension 39
The most useful MDLC systems achieve optimal peak capacity by exploiting orthogonal separations that have internally compatible buffers 40
MudPIT shows how MDLC has evolved from a laborious technique to virtually hands-free operation 41
RP-RPLC and HILIC-RP systems offer advantages for the separation of certain types of peptide mixtures 44
Affinity chromatography is combined with MDLC to achieve the simplification of peptide mixtures 44
Chapter 3  Strategies for protein identification  

3.1 INTRODUCTION  

3.2 PROTEIN IDENTIFICATION WITH ANTIBODIES  

3.3 DETERMINATION OF PROTEIN SEQUENCES BY CHEMICAL DEGRADATION  
Complete hydrolysis allows protein sequences to be inferred from the content of the resulting amino acid pool  
Edman degradation was the first general method for the de novo sequencing of proteins  
Edman degradation was the first protein identification method to be applied in proteomics, but it is difficult to apply on a large scale  

3.4 MASS SPECTROMETRY—BASIC PRINCIPLES AND INSTRUMENTATION  
Mass spectrometry is based on the separation of molecules according to their mass/charge ratio  
The integration of mass spectrometry into proteomics required the development of soft ionization methods to prevent random fragmentation  
Controlled fragmentation is used to break peptide bonds and generate fragment ions  
Five principal types of mass analyzer are commonly used in proteomics  

3.5 PROTEIN IDENTIFICATION USING DATA FROM MASS SPECTRA  
Peptide mass fingerprinting correlates experimental and theoretical intact peptide masses  
Shotgun proteomics can be combined with database searches based on uninterpreted spectra  
MS/MS spectra can be used to derive protein sequences de novo  

Chapter 4  Strategies for protein quantitation  

4.1 INTRODUCTION  

4.2 QUANTITATIVE PROTEOMICS BASED ON 2DGE  
The quantitation of proteins in two-dimensional gels involves the creation of digital data from analog images  
Spot detection, quantitation, and comparison can be challenging without human intervention  

4.3 MULTIPLEXED IN-GEL PROTEOMICS  
Difference in-gel electrophoresis involves the simultaneous separation of comparative protein samples labeled with different fluorophores  
Parallel analysis with multiple dyes can also be used to identify particular structural or functional groups of proteins  

4.4 QUANTITATIVE MASS SPECTROMETRY  
Label-free quantitation may be based on spectral counting or the comparison of signal intensities across samples in a narrow m/z range  
Label-based quantitation involves the incorporation of labels that allow corresponding peptides in different samples to be identified by a specific change in mass  
ICAT reagents are used for the selective labeling of proteins or peptides  
Proteins and peptides can also be labeled nonselectively  
Isobaric tagging allows protein quantitation by the detection of reporter ions  
Metabolic labeling introduces the label before sample preparation but is limited to simple organisms and cultured cells  

Chapter 5  The analysis of protein sequences  

5.1 INTRODUCTION  

5.2 PROTEIN FAMILIES AND EVOLUTIONARY RELATIONSHIPS  
Evolutionary relationships between proteins are based on homology  
The function of a protein can often be predicted from its sequence  

5.3 PRINCIPLES OF PROTEIN SEQUENCE COMPARISON  
Protein sequences can be compared in terms of identity and similarity  
Homologous sequences are found by pairwise similarity searching  
Substitution score matrices rank the importance of different substitutions  
Sequence alignment scores depend on sequence length  
Multiple alignments provide more information about key sequence elements  

5.4 STRATEGIES TO FIND MORE DISTANT RELATIONSHIPS  
PSI-BLAST uses sequence profiles to carry out iterative searches  

74
Pattern recognition methods incorporate conserved sequence signatures

5.5 THE RISK OF FALSE-POSITIVE ANNOTATIONS

Chapter 6  The analysis of protein structures

6.1 INTRODUCTION

6.2 STRUCTURAL GENOMICS AND STRUCTURE SPACE
Coverage of structure space is currently uneven
Structure and function are not always related

6.3 TECHNIQUES FOR SOLVING PROTEIN STRUCTURES
X-ray diffraction requires well-ordered protein crystals
NMR spectroscopy exploits the magnetic properties of certain atomic nuclei
Additional methods for structural analysis mainly provide supporting data

6.4 PROTEIN STRUCTURE PREDICTION
Structural predictions can bridge the gap between sequence and structure
Protein secondary structures can be predicted from sequence data
Tertiary structures can be predicted by comparative modeling if a template structure is available
Ab initio prediction methods attempt to construct structures from first principles
Fold recognition (threading) is based on similarities between nonhomologous folds

6.5 COMPARISON OF PROTEIN STRUCTURES

6.6 STRUCTURAL CLASSIFICATION OF PROTEINS

6.7 GLOBAL STRUCTURAL GENOMICS INITIATIVES

Chapter 7  Interaction proteomics

7.1 INTRODUCTION

7.2 METHODS TO STUDY PROTEIN–PROTEIN INTERACTIONS
Genetic methods suggest interactions from the combined effects of two mutations in the same cell or organism
Protein interactions can be suggested by comparative genomics and homology transfer

Affinity-based biochemical methods provide direct evidence that proteins can interact
Interactions between proteins in vitro and in vivo can be established by resonance energy transfer
Surface plasmon resonance can indicate the mass of interacting proteins

7.3 LIBRARY-BASED METHODS FOR THE GLOBAL ANALYSIS OF BINARY INTERACTIONS

7.4 TWO-HYBRID/PROTEIN COMPLEMENTATION ASSAYS
The yeast two-hybrid system works by assembling a transcription factor from two inactive fusion proteins
Several large-scale interaction screens have been carried out using different yeast two-hybrid screening strategies
Conventional yeast two-hybrid screens have a significant error rate

7.5 MODIFIED TWO-HYBRID SYSTEMS FOR MEMBRANE, CYTOSOLIC, AND EXTRACELLULAR PROTEINS

7.6 BACTERIAL AND MAMMALIAN TWO-HYBRID SYSTEMS

7.7 LUMIER AND MAPPIT HIGH-THROUGHPUT TWO-HYBRID PLATFORMS

7.8 ADAPTED HYBRID ASSAYS FOR DIFFERENT TYPES OF INTERACTIONS

7.9 SYSTEMATIC COMPLEX ANALYSIS BY TANDEMD AFFINITY PURIFICATION–MASS SPECTROMETRY

7.10 ANALYSIS OF PROTEIN INTERACTION DATA

7.11 PROTEIN INTERACTION MAPS

7.12 PROTEIN INTERACTIONS WITH SMALL MOLECULES

Chapter 8  Protein modification in proteomics

8.1 INTRODUCTION

8.2 METHODS FOR THE DETECTION OF POST-TRANSLATIONAL MODIFICATIONS

8.3 ENRICHMENT STRATEGIES FOR MODIFIED PROTEINS AND PEPTIDES

8.4 PHOSPHOPROTEOMICS
Protein phosphorylation is a key regulatory mechanism
Separated phosphoproteins can be detected with specific staining reagents
Sample preparation for phosphoprotein analysis typically involves enrichment using antibodies or strongly cationic chromatography resins 173

8.5 ANALYSIS OF PHOSPHOPROTEINS BY MASS SPECTROMETRY
A combination of Edman degradation and mass spectrometry can be used to map phosphorylation sites 176
Intact phosphopeptide ions can be identified by MALDI-TOF mass spectrometry 176
Phosphopeptides yield diagnostic marker ions and neutral loss products 177

8.6 QUANTITATIVE ANALYSIS OF PHOSPHOPROTEINS

8.7 GLYCOPROTEOMICS
Glycoproteins represent more than half of the eukaryotic proteome 181
Glycans play important roles in protein stability, activity, and localization, and are important indicators of disease 183
Conventional glycoanalysis involves the use of enzymes that remove specific glycan groups and the separation of glycoproteins by electrophoresis 184
Glycoprotein-specific staining allows the glycoprotein to be studied by 2DGE 187
There are two principal methods for glycoprotein enrichment that have complementary uses 188
Mass spectrometry is used for the high-throughput identification and characterization of glycoproteins 189

Chapter 9 Protein microarrays 191
9.1 INTRODUCTION 191
9.2 THE EVOLUTION OF PROTEIN MICROARRAYS 191
9.3 DIFFERENT TYPES OF PROTEIN MICROARRAYS 193
Analytical, functional, and reverse microarrays are distinguished by their purpose and the nature of the interacting components 193
Analytical microarrays contain antibodies or other capture reagents 194
Functional protein microarrays can be used to study a wide range of biochemical functions 196
9.4 THE MANUFACTURE OF FUNCTIONAL PROTEIN MICROARRAYS—PROTEIN SYNTHESIS 197
Proteins can be synthesized by the parallel construction of many expression vectors 197

9.5 THE MANUFACTURE OF FUNCTIONAL PROTEIN MICROARRAYS—PROTEIN IMMOBILIZATION 201
9.6 THE DETECTION OF PROTEINS ON MICROARRAYS 203
Methods that require labels can involve either direct or indirect detection 203
Label-free methods do not affect the intrinsic properties of interacting proteins 204
9.7 EMERGING PROTEIN CHIP TECHNOLOGIES 207
Bead and particle arrays in solution represent the next generation of protein microarrays 207
Cell and tissue arrays allow the direct analysis of proteins in vivo 207

Chapter 10 Applications of proteomics 211
10.1 INTRODUCTION 211
10.2 DIAGNOSTIC APPLICATIONS OF PROTEOMICS 212
Proteomics is used to identify biomarkers of disease states 212
Biomarkers can be discovered by finding plus/minus or quantitative differences between samples 215
More sensitive techniques can be used to identify biomarker profiles 218
10.3 APPLICATIONS OF PROTEOMICS IN DRUG DEVELOPMENT 219
Proteomics can help to select drug targets and develop lead compounds 219
Proteomics is also useful for target validation 222
Chemical proteomics can be used to select and develop lead compounds 222
Proteomics can be used to assess drug toxicity during clinical development 224
10.4 PROTEOMICS IN AGRICULTURE 225
Proteomics provides novel markers in plant breeding and genetics 225
Proteomics can be used for the analysis of genetically modified plants 227
10.5 PROTEOMICS IN INDUSTRY—IMPROVING THE YIELD OF SECONDARY METABOLISM 228

Glossary 231
Index 248