This book was carefully produced. Nevertheless, authors and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.
Preface to the Second Edition

Protein chemistry is facing a dramatic change these days. And it is not a step to higher sensitivity in protein analysis which pushes the field of proteins – but what might have been expected to be the bottleneck for further developments: It is a change in our attitude to the proteins. They are not any longer treated like individual components, in most cases with long lasting investigations, time and labor intensive. Instead, they are considered as members of pathways or networks, and protein studies may comprise hundreds or even thousands of components at once. Technology-driven developments permit the handling of large numbers of proteins in parallel or with high throughput. Namely, gel electrophoresis, mass spectrometry, and bioinformatics are the key players which open a new period in protein chemistry: 2D-gel electrophoresis can separate highly complex protein mixtures from cells or tissues; automated sample handling followed by mass analysis quickly determines peptide fragments; and finally bioinformatics enables us to identify a protein in the database within seconds. The improvements of various technologies on the one hand, but also their combined effort on the other hand opened the new research strategy called “proteomics”.

The buzz word “proteomics” has gained enormous attention, but nevertheless protein chemistry is more than that. There are still other basic techniques, e.g., chromatography and Edman sequencing, which are indispensable to study proteins, their functions and their modifications. All this is the subject of this book in its second edition: the microcharacterization of proteins. We have rearranged the chapters to render the modern working procedures and we have added relevant information with respect to proteome research.

We are really fascinated by the great diversity of proteins and the challenging attempt to explore the protein world.

Darmstadt
Martinsried
Bochum

November 1998

Roland Kellner
Friedrich Lottspeich
Helmut E. Meyer
Contributors

Appel, Dr. Ron D.; Medical Informatic Division, Geneva University Hospital, CH-1211 Geneva 14

Bahr, Dr. Ute; Institute of Analytical Chemistry, University of Frankfurt, D-60590 Frankfurt

Bairoch, Dr. Amos; Swiss Institute of Bioinformatics, CH-1211 Geneva 4

Binz, Dr. Pierre-Alain; Central Clinical Chemistry Laboratory, Geneva University Hospital, CH-1211 Geneva 14

Behnke, Dr. Beate; Institute of Sanitary Engineering, University of Stuttgart, D-70569 Stuttgart

Blüggel, Dr. Martin; Institute for Physiological Chemistry, University of Bochum, D-44780 Bochum

Bold, Dr. Peter; Institute for Laser Medicine, University of Düsseldorf, D-40225 Düsseldorf

Chaurand, Dr. Pierre; Institute for Laser Medicine, University of Düsseldorf, D-40225 Düsseldorf

Eckerskorn, PD Dr. Christoph; Toplab GmbH, D-82152 Martinsried

Frishman, Dr. Dimitrij; MPI for Biochemistry, D-82152 Martinsried

Gasteiger, Dr. Elisabeth; Medical Informatic Division, Geneva University Hospital, CH-1211 Geneva 14

Görg, Prof. Angelika; Technical University of Munich, D-85350 Weihenstephan

Hillenkamp, Prof. Franz; Institute of Medical Physics and Biophysics, University of Münster, D-48149 Münster

Hochstrasser, Prof. Denis F.; Central Clinical Chemistry Laboratory, Geneva University Hospital, CH-1211 Geneva 14

Houthaeve, Dr. Tony; University of Gent, B-9000 Gent
Contributors

Immler, Dr. Dorian; Institute for Physiological Chemistry, University of Bochum, D-44780 Bochum

Karas, Prof. Michael; Institute of Analytical Chemistry, University of Frankfurt, D-60590 Frankfurt

Kempter, Dr. Christoph; Institute of Sanitary Engineering, University of Stuttgart, D-70569 Stuttgart

Lohaus, Dr. Christiane; Institute for Physiological Chemistry, University of Bochum, D-44780 Bochum

Maierl, Dr. Andreas; MPI for Biochemistry, D-82152 Martinsried

Mehl, Dr. Ehrenfried; MPI for Biochemistry, D-82152 Martinsried

Metzger, Prof. Jörg W.; Institute of Sanitary Engineering, University of Stuttgart, D-70569 Stuttgart

Mewes, Dr. Hans-Werner; MPI for Biochemistry, D-82152 Martinsried

Schägger, Prof. Hermann; Biochemistry Department, University of Frankfurt, D-60590 Frankfurt

Schwer, Dr. Christine; Schering AG, D-13342 Berlin

Serwe, Dr. Maria; Institute for Physiological Chemistry, University of Bochum, D-44780 Bochum

Siethoff, Dr. Christoph; Institute for Physiological Chemistry, University of Bochum, D-44780 Bochum

Spengler, Prof. Bernhard; Institute for Physical Chemistry, University of Würzburg, D-97074 Würzburg

Westermeier, Dr. Reiner; Amersham Pharmacia Biotech, D-79111 Freiburg

Wilkins, Dr. Marc R.; Australian Proteome Analysis Facility, Macquarie University, Sydney
Contents

Section I: Overview

I.1 Microcharacterization of Proteins 3
   Friedrich Lottspeich and Roland Kellner

   1 General Aspects 4
   2 From a Cell to a Protein Sequence 4
   3 Genome and Proteome 9

Section II:
Microseparation Techniques and Sample Preparation

II.1 Microseparation Techniques I:
High Resolution Gel-Electrophoretic Techniques:
   Qualitative, Quantitative and
   Micropreparative Applications 13
   Angelika Görg and Reiner Westermeier

   1 Introduction 13
   2 Theory 13
   3 Media and Equipment 15
   3.1 Media 15
   3.2 Equipment 15
   4 Gel Electrophoretic Methods 16
   4.1 Disc Electrophoresis 16
   4.2 Gradient Gel Electrophoresis 16
   4.3 Additives in Electrophoresis 16
II.2 Microseparation Techniques II:  
*Gel Electrophoresis for Sample Preparation in Protein Chemistry*  
_Hermann Schägger_

1 Introduction 25  
2 Denaturing Techniques 26  
2.1 Commonly Used SDS-Polyacrylamide Gel Electrophoresis Techniques for Protein Separation 26  
2.2 Blue-SDS-PAGE for Recovery of Membrane Proteins from Gels 28  
2.3 Electroelution of Proteins after Blue-SDS-PAGE 29  
2.4 Electroblotting of Blue and Colorless SDS Gels 29  
2.5 Isoelectric Focusing in the Presence of Urea 30  
3 Native Techniques 31  
3.1 Colorless-Native-PAGE 31  
3.2 Blue-Native-PAGE 31  
3.3 Native Isoelectric Focusing 32  
4 References 33

II.3 Microseparation Techniques III:  
*Electroblotting*  
_Chris Eckerskorn_

1 Introduction 35  
2 Electroblotting 35  
2.1 Polyacrylamide Gel Electrophoresis 36  
2.2 Blot Systems 37  
2.2.1 Tank Blotting 37  
2.2.2 Semidy Blotting 38  
2.3 Blotting Parameters 40  
2.3.1 The Blotting Process 40  
2.3.2 Transfer Buffers 41  
2.3.3 Addition of SDS 41  
2.3.4 Addition of Methanol 42  
2.3.5 Influence of Protein Concentration 42  
3 Blotting Membranes 43  
4 References 46
Microseparation Techniques IV: Analysis of Peptides and Proteins by Capillary Electrophoresis

Christine Schwer

1 Introduction 49
2 Theory 49
2.1 Capillary Isotachophoresis 50
2.2 Capillary Zone Electrophoresis 51
2.3 Electroosmotic Flow 52
3 Instrumentation 54
3.1 Injection 54
3.2 Detection 57
4 Applications 60
4.1 Peptide Separations 60
4.2 Protein Separations 60
5 References 64

Microseparation Techniques V: High Performance Liquid Chromatography

Maria Serwe, Martin Blüggel and Helmut E. Meyer

1 Introduction 67
2 Principle of HPLC 67
3 Getting Started 69
3.1 Solvents 69
3.2 Pump 70
3.3 Gradient 71
3.4 Pre-Column Split 72
3.5 Sample Preparation 72
3.6 Injector 73
3.7 Tubings 73
3.8 In-Line Filter, Guard Column 73
3.9 Column 74
3.10 Detection 76
3.11 Fractionation 77
4 Applications 78
5 References 84
Contents

II.6 Sample Preparation I: 
**Removal of Salts and Detergents** 87  
*Dorian Immler, Ehrenfried Mehl, Friedrich Lottspeich and Helmut E. Meyer*

1 Introduction 87  
2 Removal of Salts or Polar Components 88  
2.1 Protein-Binding Membranes 88  
2.2 High Performance Liquid Chromatography 89  
2.3 Desalting on Microcolumns 90  
3 Removal of Detergents and Apolar Contaminants 90  
3.1 Detergents 90  
3.2 Protein-binding Membranes 91  
3.3 Precipitation 91  
3.4 SDS-PAGE 93  
3.5 Concentrating Gels 94  
4 “Golden Rules” for Protein and Peptide Handling 94  
5 References 95

II.7 Sample Preparation II: 
**Chemical and Enzymatic Fragmentation of Proteins** 97  
*Roland Kellner and Tony Houthaeve*

1 Strategy 97  
2 Denaturation, Reduction and Alkylation 99  
3 Enzymatic Fragmentation 101  
3.1 Enzymes 102  
3.2 Practical Considerations 104  
3.2.1 Practical Considerations for On-Membrane Digestions 104  
3.2.2 Practical Considerations for In-Gel Digestions 105  
3.2.3 Automation of Digest Procedures 107  
4 Chemical Fragmentation 108  
4.1 Cyanogen Bromide Cleavage 108  
4.2 Partial Acid Hydrolysis 110  
4.3 Hydroxylamine Cleavage of Asn-Gly Bonds 111  
4.4 Cleavage at Tryptophan 113  
4.5 Cleavage at Cysteine 113  
5 References 113
Section III: Bioanalytical Characterization

III.1 Amino Acid Analysis 119
Roland Kellner, Helmut E. Meyer and Friedrich Lottspeich

1 Introduction 119
2 Sample Preparation 120
2.1 Peptides and Proteins 120
2.1.1 Enzymatic Hydrolysis 121
2.1.2 Acid Hydrolysis 121
2.1.3 Alkaline Hydrolysis 122
2.2 Free Amino Acids 123
3 Derivatization 123
3.1 Post-Column Derivatization 125
3.1.1 Ninhydrin 125
3.1.2 Orthophthalaldehyde 125
3.1.3 Fluorescamine 127
3.2 Pre-Column Derivatization 127
3.2.1 Phenylisothiocyanate 128
3.2.2 Orthophthalaldehyde 128
3.2.3 Fluorenylmethyl Chloroformate 128
3.2.4 Dabsyl Chloride 128
3.2.5 Dansyl Chloride 129
3.2.6 Chiral Reagents 129
4 Data Evaluation 131
5 Instrumentation 132
6 Discussion 134
7 References 136

III.2 Chemical Methods for Protein Sequence Analysis 141
Friedrich Lottspeich, Tony Houthaeve and Roland Kellner

1 The Edman Degradation 141
1.1 Coupling, Cleavage and Conversion 141
1.2 Identification of the PTH Amino Acids 143
2 Instrumentation 145
2.1 The Liquid Phase Sequencer 146
2.2 The Solid Phase Sequencer 147
2.3 The Gas Phase Sequencer 147
2.4 The Pulsed Liquid Phase Sequencer 148
2.5 The Biphasic Column Sequencer 148
3 Difficulties of Amino Acid Sequence Analysis 149
3.1 The Sample and Sample Matrices 149
3.2 Difficulties with the Edman Chemistry 151
Contents

4 Applications 182
  4.1 Sample Preparation 182
  4.2 Ion Fragmentation 183
  4.3 Molecular Weight Determination of Proteins and Glycoproteins 183
    4.3.1 Accuracy of Mass Determination 187
    4.3.2 Sensitivity and Mass Range 187
  4.4 Analysis of Oligonucleotides 188
  4.5 Analysis of Glycans and Glycoconjugates 189
  5 Combination of MALDI with Biochemical Methods 190
    5.1 Peptide Mapping of Digested Proteins by MALDI 190
    5.2 Combination of MALDI and Gel Electrophoresis 191
    5.3 Combination of MALDI with Capillary Zone Electrophoresis 192
  6 References 192

III.5 MALDI Postsource Decay Mass Analysis 197
Bernhard Spengler, Pierre Chaurand and Peter Bold

1 Introduction 197
  2 Methodology and Principal Mechanisms of Postsource Decay 198
    2.1 Internal Energy Uptake and Ion Stability 198
    2.2 Instrumentation for MALDI-PSD Analysis 199
    2.3 Precursor Ion Selection 201
  3 Applications and Spectra Interpretation 201
    3.1 Sequence Analysis of Peptides 204
    3.2 Primary Structure Analysis of Unknown Peptides 209
    3.3 Primary Structure Analysis of Modified Peptides and Other Biomolecules 209
  4 Potential and Perspectives 209
  5 References 210

III.6 Electrospray Mass Spectrometry 213
Jörg W. Metzger, Christoph Eckerskorn, Christoph Kempter and Beate Behnke

1 Introduction 213
  2 Instrumentation 213
    2.1 The Electrospray Source 214
    2.2 The Mass Analyzer 217
    2.3 The Detector 218
  3 Mass Spectra of Proteins 219
  4 Coupling of Chromatographic Methods to the Mass Spectrometer 221
    4.1 On-Line HPLC-MS 221
    4.2 Coupling of a Protein Sequencer to an ESI Mass Spectrometer 224
    4.3 Microcapillary LC Coupled to Mass Spectrometry 225
Contents

4.4 Capillary Electrophoresis Coupled to Mass Spectrometry 227
5 Purity Control of Synthetic Peptides 228
5.1 Sample Introduction with an Autosampler 228
5.2 Off-Line HPLC-MS 229
5.3 Characterization of Combinatorial Compound Collections 229
6 References 232

III.7 Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) 235
Jörg W. Metzger
1 Introduction 235
2 Principle 236
2.1 Resolution, Mass Accuracy, Mass Range and Sensitivity 240
2.2 Tandem Mass Spectrometry 241
3 Electrospray and FT-ICR-MS 242
4 References 243

III.8 Sequence Analysis of Proteins and Peptides by Mass Spectrometry 245
Christoph Siethoff, Christiane Lohaus and Helmut E. Meyer
1 Introduction 245
2 Basics in Peptide Fragmentation 247
3 Instrumentation and Generation of MS/MS Data Sets 249
3.1 MALDI-PSD Time-of-Flight Mass Spectrometry 249
3.2 ESI-Triple Stage Quadrupole Mass Spectrometry 251
3.3 ESI-Ion Trap Mass Spectrometry 252
Section IV: Computer Sequence Analysis

IV.1 Internet Resources for Protein Identification and Characterization 277
Pierre-Alain Binz, Marc R. Wilkins, Elisabeth Gasteiger, Amos Bairoch, Ron D. Appel and Denis F. Hochstrasser

1 Introduction 277
2 General Approach 278
2.1 Experimental Attributes and Choice of Programs 278
2.2 Information in Databases 279
2.2.1 The Annotations in SWISS-PROT 279
2.2.2 SWISS-PROT Supplement TrEMBL 279
3 Identification and Characterization Tools at ExPASy 281
3.1 Identification Tools 281
3.1.1 Identification with Sequence Tags: TagIdent 282
Identification with Amino Acid Composition: AACompIdent 283
Identification with Combined Attributes: MultiIdent 284
Identification with Peptide Mass Fingerprints: PeptIdent 286
Characterization Tools 287
Compute pI/MW 287
PeptideMass 288
ProtParam 289
ProtScale 289
FindMod 290
Other Identification and Characterization Tools on the Internet 294
The ProteinProspector Server (University of California San Francisco Mass Spectrometry Facility) 294
The PROWL Site from the MS Groups at Rockefeller and NY Universities 295
Peptidesearch at the EMBL Protein & Peptide Group 296
The MOWSE Database Search Tool from the Daresbury Laboratory 296
The Masssearch Page of the Computational Biochemistry Research Group Server at the ETH Zurich 296
Conclusion 297
References 297
Abbreviations

API  atmospheric pressure ionization
ATZ  anilinothiazolinone
AUFS  absorption units full scale

BLAST  basic local alignment search tool
BNPS  bromine adduct of 2-(2-nitrophenylsulfonyl)-3-indolenine

CAPS  3-cyclohexylamino-1-propanesulfonic acid
CE  capillary electrophoresis
CEC  capillary electrophoracography
CGE  capillary gel electrophoresis

CHAPS  3-(3-cholamidopropyl)dimethylammonio-1-propane sulfonate
CID  collision-induced dissociation

CIEF  capillary isoelectric focusing
CORBA  common object request broker architecture
CTAB  cetyltrimethylammonium bromide
CZE  capillary zone electrophoresis

Da  Dalton
DABS-Cl  4-dimethylaminoazobenzene sulfonyl chloride
DC  direct current

DHB  dihydroxy benzoic acid
DNA  desoxyribonucleic acid

DNP  2,4-dinitrophenol
DMPTU  dimethylphenylthiourea
DPTU  diphenylthiourea
DPU  diphenylurea
DTPA  diethylene triamine-penta-acetic acid

DTT  dithiothreitol
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDTA</td>
<td>ethylenediamine tetraacetic acid</td>
</tr>
<tr>
<td>EOF</td>
<td>electroosmotic flow</td>
</tr>
<tr>
<td>ESI</td>
<td>electrospray ionisation</td>
</tr>
<tr>
<td>EST</td>
<td>expressed sequence tag</td>
</tr>
<tr>
<td>FAB</td>
<td>fast atom bombardment</td>
</tr>
<tr>
<td>FID</td>
<td>flame ionization detector</td>
</tr>
<tr>
<td>FLASH</td>
<td>fast look-up algorithm for string homology</td>
</tr>
<tr>
<td>FLEC</td>
<td>1-(9-fluorenlyl)ethyl chloroformate</td>
</tr>
<tr>
<td>FMOC</td>
<td>9-fluorenlymethyloxycarbonyl</td>
</tr>
<tr>
<td>FT</td>
<td>fourier transform</td>
</tr>
<tr>
<td>HFA</td>
<td>hexafluoroacetone</td>
</tr>
<tr>
<td>HIC</td>
<td>hydrophobic interaction chromatography</td>
</tr>
<tr>
<td>HILIC</td>
<td>hydrophilic interaction chromatography</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>ICR</td>
<td>ion cyclotron resonance</td>
</tr>
<tr>
<td>IEC</td>
<td>ion exchange chromatography</td>
</tr>
<tr>
<td>IEF</td>
<td>isoelectric focusing</td>
</tr>
<tr>
<td>IPG</td>
<td>immobilized pH gradient</td>
</tr>
<tr>
<td>IR</td>
<td>infrared</td>
</tr>
<tr>
<td>ITP</td>
<td>isotachophoresis</td>
</tr>
<tr>
<td>LC</td>
<td>liquid chromatography</td>
</tr>
<tr>
<td>MALDI</td>
<td>matrix assisted laser desorption ionization</td>
</tr>
<tr>
<td>MECC</td>
<td>micellar electrokinetic capillary chromatography</td>
</tr>
<tr>
<td>MES</td>
<td>2-((N-morpholino)ethanesulfonic acid</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>MoAb</td>
<td>monoclonal antibody</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectrometry</td>
</tr>
<tr>
<td>NEPHGE</td>
<td>non equilibrium pH gradient gel electrophoresis</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>NTCB</td>
<td>2-nitro-5-thiocyanobenzoic acid</td>
</tr>
<tr>
<td>OPA</td>
<td>orthophthalaldehyde</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
</tbody>
</table>
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>polymeric chain reaction</td>
</tr>
<tr>
<td>PEDANT</td>
<td>protein extraction, description, and analysis tool</td>
</tr>
<tr>
<td>PIR</td>
<td>protein identification resource</td>
</tr>
<tr>
<td>PITC</td>
<td>phenylisothiocyanate</td>
</tr>
<tr>
<td>PSD</td>
<td>post source decay</td>
</tr>
<tr>
<td>PTC</td>
<td>phenylthiocarbamyl</td>
</tr>
<tr>
<td>PTH</td>
<td>phenylthiohydantoin</td>
</tr>
<tr>
<td>PTM</td>
<td>post-translational modification</td>
</tr>
<tr>
<td>PVDF</td>
<td>polyvinylidene difluoride</td>
</tr>
<tr>
<td>RF</td>
<td>radio frequency</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RPC</td>
<td>reverse phase chromatography</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SEC</td>
<td>size exclusion chromatography</td>
</tr>
<tr>
<td>SWISS-PROT</td>
<td>protein database from the Swiss Institute of Bioinformatics</td>
</tr>
<tr>
<td>TCA</td>
<td>trichloroacetic acid</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N,N',N'-tetramethylethyldiamin</td>
</tr>
<tr>
<td>TFA</td>
<td>trifluoroacetic acid</td>
</tr>
<tr>
<td>TH</td>
<td>thiohydantoin</td>
</tr>
<tr>
<td>TIC</td>
<td>total ion current</td>
</tr>
<tr>
<td>TOF</td>
<td>time-of-flight</td>
</tr>
<tr>
<td>TrEMBL</td>
<td>protein translations from EMBL database</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolett</td>
</tr>
</tbody>
</table>
Section I: Overview
Microcharacterization of Proteins

*Friedrich Lottspeich and Roland Kellner*

At the beginning of the 20th century it became apparent to biochemists that proteins are macromolecules which consist of amino acid components. Emil Fischer deduced from hydrolyzation experiments and the very first peptide syntheses [1903] that proteins are built up of numerous linked α-amino acids. The nature of different amino acids and the peptide bond linkage became clear. However, small synthetic peptides did not exhibit the properties of proteins. It was not realized until 1950 that all molecules of a given protein have one unique amino acid sequence. Several key developments helped to reveal the molecular architecture of proteins, namely the development of chromatography by Martin [1941] and amino acid analysis by Stein and Moore [1948]. One milestone was the finding of Frederick Sanger, who identified phenylalanine as an end-group of insulin by its reaction with 2,4-dinitrofluorobenzene. The Sanger method forms yellow 2,4-dinitrophenyl (DNP) derivatives with amino groups. At the same time Pehr Edman established the stepwise degradation of proteins using phenylisothiocyanate as a reagent. Then in 1953 Sanger and Tuppy were able to determine the amino acid sequence of insulin, which was the first protein characterized on the molecular level.

And furthermore human insulin became the first biotechnology product in 1982. A synthetic gene was constructed corresponding to the amino acid sequence of insulin and the foreign DNA was inserted into the bacterium Escherichia coli. The bacteria served as a host and expressed the human gene, synthesizing the recombinant protein insulin. Genetic engineering and recombinant proteins revolutionized the classical field of biochemistry. Very successful techniques were established to synthesize oligonucleotides and to analyze cDNA sequences rapidly and with high throughput. The PCR technology allowed to multiply single DNA chains and therefore to overcome sensitivity problems. The opportunity for automatization of genetic techniques initiated giant research activities like the human genome project. Numerous genome projects for smaller species could already be finished and there are only doubts if the complete human genome will be identified in the year 2001 or in 2005 or even faster. Genomics was the key word for life science activities in the last years. And only recently the twin name came up – proteomics. The technology-driven development of protein chemistry permits to investigate new biologically relevant questions with enormous business opportunities, especially in pharmaceutical areas. The characterization of proteins is a very exciting field these days.
1 General Aspects

In the classical view, the reason for a protein purification is the observation of a biological phenomenon and search for its molecular basis. Thus, a single or a few proteins which are correlated with observed activity are isolated and characterized at the molecular level. Knowledge of the primary structure of these proteins as well as their modifications and processing sites is of fundamental importance. To acquire this knowledge, particularly when only a small amount of protein is available, may be difficult and laborious. All techniques available at the time have to work in a concerted action to elucidate the structure of the protein in a reasonable time. The entire amino acid sequence of a protein is now more easily obtained by deduction from its DNA sequence rather than analyzing the total protein by protein chemical amino acid sequence determination. However, some protein chemical information is always required to isolate the DNA or to check the accuracy of the DNA sequence. Important information, like the actual N-terminus or C-terminus of a protein, can only be determined by protein chemical methods. Probably the most important task of protein chemistry determining the post-translational modifications, about which little can be learned from the analysis of DNA sequences, but which often regulate the activity of the proteins. Supplementation with immunologic, crystallographic or NMR data provides additional information which is not available by analysis of the DNA or protein sequences alone. Modern protein chemistry is one of the important members in the orchestra of all the different disciplines in the biosciences, providing the basis for a detailed understanding of structure-function relationships.

In recent years additional areas of responsibility have been allocated to protein chemistry, mainly due to the developments in recombinant DNA techniques. Projects to analyze the human genome or the genomes of other organisms at the DNA or RNA level are the focus of strong international research activities. However, even today it is apparent that the meaningfulness of the genomic analyses will be limited, unless at the same time more attention is paid to the function of the huge number of sequenced or characterized pieces of DNA. The connective link between the genome and the multiple cellular functions of an organism are the proteins. The quantitative appearance and certain posttranslational modifications have major influence on the function and activity of a protein. Thus, the analysis of proteins, which are the real players and tools in the cell, is a supplementary and inevitable approach to understand the biological events in a cell. Figure 1 outlines how protein analysis links genomic and protein information at the molecular level.

2 From a Cell to a Protein Sequence

Nowadays it is possible to separate most of the cellular proteins by two dimensional (2D) gel electrophoresis and to check their amount quantitatively. In this way, a picture of the protein expression for a certain cell stage can be obtained. In a kind of subtractive approach, different metabolic stages of cells can be compared with the help of computerized image analysis and data handling. The subtle protein changes that arise
Microcharacterization of Proteins

Figure 1. Protein analysis as an interface between genomic and protein information.
in all complex biological systems against a background of constitutive proteins have to be recognized and further analyzed with modern protein chemical micro methods.

The amounts of the most abundant proteins in a 2D gel electrophoretic separation are in the very low picomole or femtomole level, corresponding to few micrograms or even submicrogram amounts of material. It is because the amounts of interesting proteins are so small that it has been so imperative to improve the protein analytical methods towards higher sensitivity and more speed. Sequence analysis, for example, has come to a level where a few picomoles of a protein can be sequenced: that is roughly the amount of the 100–300 most abundant cellular proteins separated in 2D gel electrophoresis. All the other most commonly used methods in protein chemistry, like separation techniques, amino acid analysis or mass spectrometry, today work in a similar sensitivity range. An urgent need also to attack the proteins present in even lower quantities is a strong driving force for further developments in methodology and instrumentation. Improvements in sensitivity to the femtomole level for all the protein chemical methods mentioned is within the reach of modern micro protein chemistry. The consequence will be that thousands of proteins of a cell can be analyzed.

Manipulating these small quantities requires several peculiarities to be taken into consideration. Proteins adsorb strongly to any kind of surface and microgram amounts may be lost in a few minutes to the wall of a vessel. Consequently, several techniques commonly used on the macro scale, like ultrafiltration, dialysis or lyophilization, cannot be recommended when working with micro amounts. Sometimes adsorbed protein can be recovered by treatment with concentrated formic acid or by incubating with detergent solutions. In general a good recovery is achieved for peptide material dissolving it in aqueous trifluoroacetic acid with a few percent of acetonitrile.

At the micro level contamination becomes a major problem. Laboratory dust and impurities in solvents, reagents and equipment are almost inevitable sources of contamination. It is obvious that automated separation and reaction devices which use minimum volumes of solvents and reagents and which keep the sample in a closed environment should be used for preference. However, so far, common laboratory equipment seldom is in accordance with these requirements, and thus instrument development is called for.

As a consequence of the special situation with micro amounts, careful planning of the purification strategy is imperative to minimize all the handling and transfer steps and to keep contamination as low as possible. Often, early fractionation steps by multistep extraction or precipitation techniques, commonly used in large scale protein purification, cannot be adopted at the micro scale. Even at the early stages the most efficient separation techniques available have to be adopted in an optimal sequence. Detection methods like mass spectrometry, or diode array spectroscopy yielding multiple information are strongly recommended. In Figure 2 several strategies are summarized which have been developed and successfully applied in the last years.

Starting materials like a cell or a complex protein mixture already enriched in the protein of interest by conventional purification steps is applied to a high-resolving separation method. In Figure 3 the different fractionation methods are compared according to their capability to separate molecules of different sizes. Immediately it becomes clear why gel electrophoresis is usually the method of choice to separate complex protein mixtures. For small molecules (e.g., amino acids or peptides) chro-