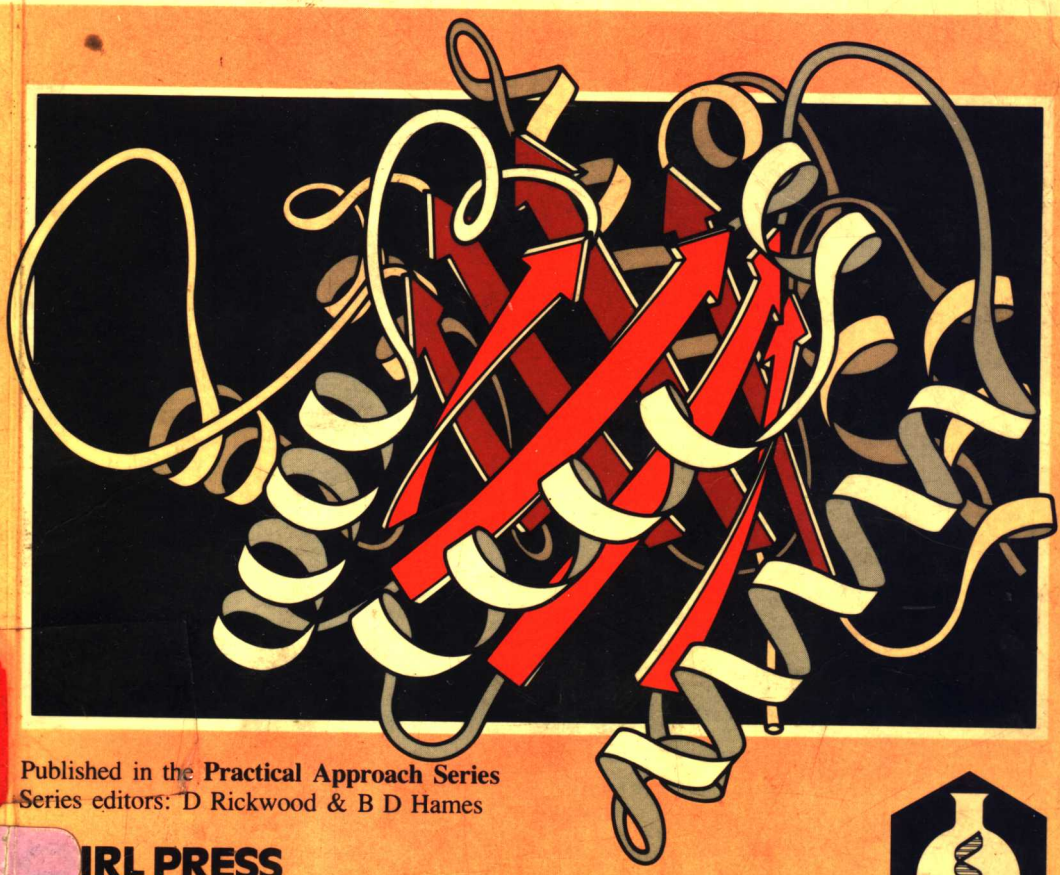
Protein structure

a practical approach

Edited by
T E Creighton



Published in the **Practical Approach Series**
Series editors: D Rickwood & B D Hames

IRL PRESS
at
OXFORD UNIVERSITY PRESS
Oxford New York Tokyo

Protein structure

a practical approach

Edited by

T E Creighton

MRC Laboratory of Molecular Biology, Hills Road,
Cambridge CB2 2QH, UK

IRL PRESS
at
OXFORD UNIVERSITY PRESS
Oxford New York Tokyo

IRL Press
Eynsham
Oxford
England

First Published 1989

British Library Cataloguing in Publication Data

Protein structure
1. Proteins. Structure
I. Creighton, Thomas E., *1940–* II. Series
547.7'5044

Library of Congress Cataloging in Publication Data

Protein structure.
(Practical approach series)
Includes bibliographies and index.
1. Proteins—Analysis. 2. Proteins—Structure.
3. Biochemistry—Technique. I. Creighton, Thomas E.,
1940– . II. Series. [DNLM: 1. Proteins—analysis.
QH 55 P9684]
QP551.P69764 1988 574.19'245 88-31965

ISBN 0 19 963000 3 (hardbound)
0 19 963001 1 (softbound)

Previously announced as:

ISBN 1 85221 137 7 (hardbound)
ISBN 1 85221 138 5 (softbound)

Typeset by Paston Press, Loddon, Norfolk, and printed by
Information Printing Ltd, Oxford, England

Preface

Techniques for the experimental study of protein structure need not be complex, and this volume presents a collection of those most likely to be useful in the average biochemical laboratory. Most of the procedures use relatively simple electrophoretic, chromatographic and spectrophotometric techniques, so all that is required is some simple electrophoresis equipment, a column and a fraction collector, and a UV spectrophotometer. No expensive, complicated items of equipment are required, so the procedures should be generally useful, and the dangers of becoming wedded or addicted to a single instrument or technique are minimal.

With the procedures described here, it is possible to estimate the molecular weight of a specific protein (Chapter 1), to resolve it from all other proteins in virtually any mixture (Chapters 2 and 3), and to identify it given just a portion of its amino acid sequence or a clone of its gene (Chapter 4). The covalent structure can be characterized by the peptide maps it produces (Chapter 5), the numbers of certain amino acid residues present in the protein can be counted (Chapter 6), and the number and identities of any disulphide bonds present can be established (Chapter 7).

The three-dimensional folded conformations of proteins are crucial for their biological properties. Chapter 8 provides guidelines as to what can and cannot be inferred about the structure and properties of a protein from just its gene or amino acid sequence, while the procedures for getting proteins to adopt their appropriate folded conformations are described in Chapter 9. Simple procedures for monitoring the overall conformational properties of protein are described in Chapters 10–12, using electrophoretic, spectral, and immunochemical techniques. Measuring the conformational stability of a protein is described in Chapter 13, and guidelines on how to optimize that stability are given in the final chapter.

Techniques for purifying proteins and determining their primary structures, either directly or by gene cloning, are not given, for they are the subjects of entire volumes in this series (*Protein Purification: A Practical Approach*; *Protein Sequencing: A Practical Approach*; *DNA Cloning: A Practical Approach*). Likewise, procedures for determining the detailed three-dimensional structures of proteins by X-ray diffraction or by nuclear magnetic resonance are not provided here. Neither are techniques that an average molecular biologist would be expected to use directly. Both require items of equipment that are not available in most laboratories and a virtual full-time commitment to just one of the techniques. Therefore, if you want to characterize your protein further, either by determining its complete primary structure or its atomic-resolution three-dimensional structure, you are advised to prepare large quantities of protein, plus perfect single crystals, if possible, and to take them to your nearest, or most friendly, laboratory specializing in these techniques.

In the meantime, use the relatively simple procedures of this volume to obtain a low-resolution 'picture' of your protein that will be useful for understanding its biological function. For more detailed studies of the functional properties of the

protein, you are advised to acquire and use a copy of the companion volume, *Protein Function: A Practical Approach*. For a comprehensive description of the properties of proteins, my volume *Proteins: Structures and Molecular Properties* (W. H. Freeman, New York, 1983) should be useful.

T. E. Creighton

Contributors

T. Arakawa
Graduate Department of Biochemistry, Brandeis University, Waltham, MA 02254, USA

P. Argos
European Molecular Biology Laboratory, Postfach 10,2209, D-6900 Heidelberg, FRG

E. A. Carrey
Department of Biochemistry, University of Dundee, Dundee DD1 4HN, UK

T. E. Creighton
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

L. Djavadi-Ohaniance
Departement de Biochimie et Génétique Moléculaire, Institut Pasteur, F-75724 Paris Cedex 15, France

B. Friguet
Departement de Biochimie et Génétique Moléculaire, Institut Pasteur, F-75724 Paris Cedex 15, France

M. E. Goldberg
Departement de Biochimie et Génétique Moléculaire, Institut Pasteur, F-75724 Paris Cedex 15, France

D. P. Goldenberg
Department of Biology, University of Utah, Salt Lake City, UT 84112, USA

S. M. Hanash
University of Michigan Medical School, C. S. Mott Children's Hospital, Department of Pediatric Hematology, R4451 Kresge I, Box 0510, Ann Arbor, MI 48109, USA

M. Hollecker
European Molecular Biology Laboratory, Postfach 10,2209, D-6900 Heidelberg, FRG. Present address: C.N.R.S., Centre de Biophysique Moléculaire, F-45071 Orléans Cedex 2, France

G. Jackowski
Department of Clinical Biochemistry, University of Toronto, Research Institute, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, Canada M5G 1X8. Present address: Division of Cardiovascular Research, Vioclone Biologicals Inc., Mississauga, Ontario, Canada

R. Jaenicke
Institut für Biophysik und physikalische Biochemie, Universität Regensburg, Universitätsstrasse 31, D-8400 Regensburg, FRG

R. Kuick
University of Michigan Medical School, C. S. Mott Children's Hospital, Department of Pediatric Hematology, R4451 Kresge I, Box 0510, Ann Arbor, MI 48109, USA

C. N. Pace
Biochemistry Department, Texas A & M University, College Station, TX 77843, USA

P. G. Righetti
Chair of Biochemistry, Faculty of Pharmacy and Department of Biomedical Sciences and Technologies, University of Milano, Via Celoria 2, Milano 20133, Italy

R. Rudolph
Institut für Biophysik und physikalische Biochemie, Universität Regensburg, Universitätsstrasse 31, D-8400 Regensburg, FRG

K. H. Scheidtmann
Institut für Immunbiologie, Albert-Ludwigs-Universität Freiburg, Stefan Meier Str. 8, 78 Freiburg, FRG

F. X. Schmid
Institut für Biophysik und physikalische Biochemie, Universität Regensburg, Universitätsstrasse 31, D-8400 Regensburg, FRG

Y. P. See
Department of Clinical Biochemistry, University of Toronto, Research Institute, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, Canada M5G 1X8

B. A. Shirley
Biochemistry Department, Texas A & M University, College Station, TX 77843, USA

J. R. Strahler
University of Michigan Medical School, C. S. Mott Children's Hospital, Department of Pediatric Hematology, R4451 Kresge I, Box 0510, Ann Arbor, MI 48109, USA

J. A. Thomson
Biochemistry Department, Texas A & M University, College Station, TX 77843, USA

S. N. Timasheff
Graduate Department of Biochemistry, Brandeis University, Waltham, MA 02254, USA

Contents

ABBREVIATIONS xvii

1. ESTIMATING MOLECULAR WEIGHTS OF POLYPEPTIDES BY SDS GEL ELECTROPHORESIS 1
Y. P. See and G. Jackowski

Introduction 1
Electrophoresis Systems 2
Slab gel apparatus 3
General reagents 3
Preparation of general stock solutions 3
Continuous buffer system 3
Discontinuous buffer system 5
Linear acrylamide gradient gels 7
Staining and De-staining 8
Coomassie blue staining 8
Silver staining 9
Molecular Weight Determination 10
Simple gel systems 10
Multiple acrylamide concentrations 10
Gradient gels 10
Molecular weight limits 12
Protein standards 12
Problems Related to Sample Preparation 13
Amount of protein in sample 13
Preparation and treatment of sample 13
Problems Encountered in Molecular Weight Determinations 15
General problems 15
Analysis of multiple acrylamide concentrations 17
Linear acrylamide gradient gels 18
Conclusion 19
References 19

2. ISOELECTRIC FOCUSING OF PROTEINS IN CONVENTIONAL AND IMMOBILIZED pH GRADIENTS 23
P. G. Righetti

Introduction 23
Sample Handling 23
Standardization 23
Time 24
Temperature 24
Sterility 24

Air 24
Freezing and thawing 24
Buffers 24
On Protein Heterogeneity 25
Genetic factors 25
Post-synthetic modification 27
Conventional Isoelectric Focusing 27
Applications and limitations 28
Carrier Ampholytes 29
Equipment 31
Electrophoretic equipment 31
Polymerization cassette 32
Gel polymerization 36
Running the gel 39
Staining protocols and densitometry 41
Trouble-shooting 41
Artefacts 42
Titration Curves 42
Immobilized pH Gradients 44
The Immobiline matrix 45
Narrow and ultra-narrow pH gradients 46
Extended pH gradients 46
IPG Methodology 52
Casting an Immobiline gel 52
Re-swelling of dry Immobiline gels 56
Storage of the Immobiline chemicals 56
Mixed-bed, CA–IPG gels 57
Effect of salts on proteins in IPG 58
Illustrating the Resolving Power 61
Acknowledgements 62
References 62

3. TWO-DIMENSIONAL POLYACRYLAMIDE GEL ELECTROPHORESIS OF PROTEINS **65**
J. R. Strahler, R. Kuick and S. M. Hanash

Introduction 65
Sample Preparation and Electrophoretic Methods 68
Sample preparation 68
Materials 72
Equipment 72
Sample solubilization 74
First dimension IEF with carrier ampholytes 76

First dimension IEF with immobilized pH gradients 77
Second dimension SDS–acrylamide gel electrophoresis 79
Silver staining 81
Fluorography 82
Electrophoretic transfer of proteins to membranes 84
Gel Analysis 85
Sources of extraneous variability 85
Measuring spot sizes 88
Comparing spot sizes 89
Acknowledgements 91
References 91

4. IMMUNOLOGICAL DETECTION OF PROTEINS OF KNOWN SEQUENCE **93**
K. H. Scheidtmann

Introduction 93
Generation of Specific Antibodies 93
Antibodies Against Synthetic Peptides 94
Criteria for selection of peptides 94
Coupling of peptide to carrier protein 96
Immunization 99
Testing antisera 99
Antibodies Against Fusion Proteins Produced by Gene Expression 100
Preparation of fusion protein for immunization 100
Detection of Proteins by Immunological Methods 101
Detection of Proteins in Cell Extracts 101
Preparation of whole cell extracts 101
Preparation of 'soluble' cell extracts 101
Extraction of nuclei 105
Opening cells by 'Douncing' 105
Cell fractionation 105
Radiolabelling of cells 106
Evaluation of Detection Methods 106
Western blotting 107
Immunoprecipitation 109
Detection of Proteins *In Situ* 112
Immunofluorescence 112
Other staining procedures 112
Further Applications 113
Immunoaffinity chromatography 114
Assay for enzymatic or DNA-binding activities 114
Acknowledgements 114
References 114

5. PEPTIDE MAPPING **117**
E. A. Carrey

Introduction: Strategies 117
Hydrolysis 117
Separation 118
Identification and analysis 118
Cleavage of Peptide Bonds 118
Chemical methods 119
Proteolytic enzymes 120
Controlled proteolysis and total proteolysis 124
Proteolysis in SDS solutions 128
Separation of Polypeptide Fragments 129
Polyacrylamide gel electrophoresis 130
Isoelectric focusing 131
Chromatography 131
Electroblots as the Second Dimension 132
Electroblotting procedure 132
Identification of polypeptide fragments 133
Mapping radiolabelled fragments 136
Use of the electroblot in sequencing strategies 138
Further Analytical Dimensions 139
Phospho-peptides 139
Isoelectric focusing/SDS electrophoresis maps 141
Off-diagonal maps 142
Acknowledgements 143
References 143

6. COUNTING INTEGRAL NUMBERS OF RESIDUES BY CHEMICAL MODIFICATION **145**
M. Hollecker

Introduction 145
Advantages of the method 145
Limitations of the method 146
Cysteine Residues 146
Choice of reagents for modifying thiol groups specifically 146
Experimental procedure 146
Lysine Residues 148
Choice of reagent 148
Experimental procedure 148
Electrophoretic Resolution of the Modified Species 149
Gel preparation 150
Counting cysteine residues 152
Counting amino groups 153
References 153

7. DISULPHIDE BONDS BETWEEN CYSTEINE RESIDUES **155**
T. E. Creighton

Introduction 155
Determining the Number of Disulphide Bonds Present 156
Reducing protein disulphides 156
Ellman assay for thiols 157
Preserving Protein Disulphides 158
Blocking thiols 158
Unmasked buried thiols 158
Fragmentation of the Protein 159
Cleavage by thermolysin 159
Identifying Disulphide-linked Peptides 160
Peptide separation methods 160
Diagonal electrophoresis 160
References 166

8. PREDICTIONS OF PROTEIN STRUCTURE FROM GENE AND AMINO ACID SEQUENCES **169**
P. Argos

Introduction 169
Finding Protein Coding Regions in Genetic Sequences 169
With amino acid composition 169
With oligopeptide preferences 171
Aligning Amino Acid Sequences 172
Using simple amino acid comparisons 172
Using 'windows' of sequences 173
Significance of the alignment 176
Multiple sequence alignment 178
Sequence patterns as an alignment aid 179
Finding repeats within a sequence 180
Recommendations 181
Secondary Structure Prediction 181
Methods for prediction 181
Secondary structure prediction: accuracy and utility 183
Prediction of trans-membrane spans 183
Antigenic sites, signal and target sequences 184
Help for Site-Directed Mutagenesis 184
Model Building from a Known Tertiary Structure 186
Computer Program Packages 187
Acknowledgements 188
References 188

9. FOLDING PROTEINS **191**
R. Jaenicke and R. Rudolph

Introduction 191
Basic Techniques for Monitoring Protein Folding 191
Catalytic function 192
Other functional properties 194
Conformational characteristics 194
Limited proteolysis 194
Unfolding and Refolding of Proteins 198
Denaturants 198
Choice of unfolding and refolding conditions 199
Optimization of reconstitution 200
Equilibrium transitions 203
Kinetics 204
Folding Proteins with Concomitant Disulphide Bond Formation 206
Denaturation/reduction 208
Renaturation/re-oxidation 208
Association of Oligomeric Proteins 211
Dissociation/association 211
Reactivation 212
Refolding 214
Re-association 214
Effects of Ligands and Other Proteins 218
Co-factors 218
Co-enzymes and substrates 221
Effect of enzymes catalysing protein folding 222
References 222

10. ANALYSIS OF PROTEIN CONFORMATION BY GEL ELECTROPHORESIS **225**
D. P. Goldenberg

Introduction and Basic Principles 225
Non-denaturing Gel Electrophoresis of Proteins 228
Apparatus 228
Buffer systems 229
Gel composition 233
Polymerization catalysts 233
Preparing slab gels 234
Sample preparation, application and electrophoresis 235
Staining gels 236
Drying gels 237
Quantification of band intensities 238
Transverse Gradient Gel Electrophoresis 239
Apparatus for transverse gradient gel electrophoresis 241

Preparing the solutions for urea gradient gel electrophoresis 242
Casting transverse gradient gels 244
Sample preparation, application and electrophoresis 245
Interpreting urea gradient gel patterns 247
Acknowledgements 250
References 250

11. SPECTRAL METHODS OF CHARACTERIZING PROTEIN CONFORMATION AND CONFORMATIONAL CHANGES 251
F. X. Schmid

Introduction 251
Absorbance 252
Spectrophotometers 257
Samples and cuvettes 258
Instrumental settings 260
Experimental protocol for determining the difference spectrum between the native and unfolded states of a protein 261
Additional techniques 264
Fluorescence 264
Fluorescence of the aromatic amino acids 265
Fluorescence of proteins 265
Environmental effects on tyrosine and tryptophan residue emission 269
Measurement of protein fluorescence 270
Circular Dichroism 276
CD spectra of native and unfolded proteins 278
Measurement of circular dichroism 278
Experimental guide to measuring the CD spectra of proteins 281
Evaluation of the data 282
Determination of protein secondary structure by CD 283
References 284

12. IMMUNOCHEMICAL ANALYSIS OF PROTEIN CONFORMATION 287
B. Friguet, L. Djavadi-Ohaniance and M. E. Goldberg

Introduction 287
Antibody Selection: How to Obtain the Appropriate Antibodies 288
Equipment 289
Buffers and reagents 289
Indirect ELISA procedure 290
Important remarks: what should be remembered? 291
Antibody Characterization 292
Antibody purification 292
Isotype determination 293

Discrimination between anti-native and anti-denatured antibodies 294
Do the antibodies recognize different epitopes? 295
Antibodies as Conformational Probes 299
Preparation of 'fragments antigen binding' (Fab) 299
Stoichiometry determination 300
Affinity measurements in solution 301
Conformational changes detected by affinity measurement 308
Conclusion 308
References 309

13. MEASURING THE CONFORMATIONAL STABILITY OF A PROTEIN **311**
C. N. Pace, B. A. Shirley and J. A. Thomson

Introduction 311
Selecting a Technique to Follow Unfolding 311
Determining Unfolding Curves 313
Equilibrium and reversibility 315
Urea and GdmCl unfolding 316
Thermal unfolding 318
Analysing Unfolding Curves 319
Urea and GdmCl unfolding 321
Thermal unfolding 323
Determining Differences in Stability 325
Concluding Remarks 328
Acknowledgements 329
References 329

14. STABILIZATION OF PROTEIN STRUCTURE BY SOLVENTS **331**
S. N. Timasheff and T. Arakawa

Introduction 331
Principles of Structure Stabilization by Solvent Components 331
Sources of Exclusion 335
Balance Between Co-solvent Exclusion and Binding 337
Co-solvent Interactions in the Denaturation Reaction 339
Practical Considerations 341
Acknowledgements 344
References 344

APPENDIX **347**

Suppliers of Specialist Items

INDEX **349**

CHAPTER 1

Estimating molecular weights of polypeptides by SDS gel electrophoresis

YEW PHEW SEE and GEORGE JACKOWSKI

1. INTRODUCTION

Estimation of the molecular weight (M_r) of a polypeptide is of central importance to the characterization of proteins and is finding increasing use in molecular biology. It is our aim in this chapter to discuss very briefly the theoretical aspects of estimating the molecular weights of polypeptides using sodium dodecyl sulphate (SDS)–polyacrylamide gel electrophoresis and to provide a collection of proven procedures and advice on their use.

Molecular weight determination of polypeptides using SDS–polyacrylamide gel electrophoresis (PAGE) was first introduced, empirically, by Shapiro *et al.* (1). It was confirmed and extended by Weber and Osborn (2) and Dunker and Rueckert (3). These authors showed that the relative mobilities of the polypeptides were related to their molecular weights. Explanation for this behaviour of polypeptides in SDS–PAGE was soon provided by Pitt-Rivers and Impiombato (4) and Reynolds and Tanford (5, 6). They found that, under appropriate conditions, all reduced polypeptides bind the same amount of SDS on a weight basis (1.4 g SDS/g polypeptide). Viscosity analysis suggested that the reduced polypeptide–SDS complexes formed rod-like particles, with lengths proportional to the molecular weight of the polypeptides. Unreduced polypeptides containing intact disulphide bonds do not bind the optimum amount of SDS and have different hydrodynamic volumes. Gel filtration analysis showed that these unreduced polypeptide–SDS complexes were smaller than the corresponding reduced polypeptide–SDS complexes (6, 7).

Although viscosity analysis suggested that the reduced polypeptide–SDS complex had a rod-like structure (6), the exact structure of this complex is still unknown. Some protein chemists mistakenly believe that SDS dissociates proteins in the same manner as guanidinium chloride to give a completely random coil. Optical rotatory dispersion analysis showed that the SDS–polypeptide complexes contain significant proportions of α-helical structures (6, 8, 9). Mattice *et al.* (9) had suggested, however, that the 'overall conformation more closely approximates that of a random coil than a rod'. From viscosity analysis, Reynolds and Tanford (6) proposed the prolate ellipsoid model for the SDS–polypeptide structure. Other models such as the 'necklace' model (10), the deformable prolate ellipsoid model (11) and the amino acid-enhanced structural change model (9)

have been proposed. The validity of these models was examined by Mattice *et al.* (9) who concluded that 'none of the models provides a molecular explanation for the ability of diverse proteins to form complexes that contain 1.4 g SDS/g of protein'.

The amount of SDS bound by a protein is dependent on the structure of the protein (reduced or unreduced), the temperature and the ionic strength of the solution (4, 5). Unreduced proteins bind much less SDS than reduced proteins, and increasing ionic strength decreases the amount of SDS bound (4, 5, 12). SDS binds mainly to the hydrophobic regions of polypeptides, whereas the hydrophilic regions bind much less SDS (13). Why proteins with different hydrophobicities bind the same amount of SDS per gram is not understood. Glycoproteins with more than 10% carbohydrate bind much less SDS than other proteins of the same molecular weight (14, 15).

Separation of SDS–polypeptide complexes by SDS–PAGE is analogous to gel filtration. The separations are dependent on the sizes of the SDS–polypeptide complexes and not necessarily on the sizes or molecular weights of the original polypeptides or proteins. The electrophoretic mobility is proportional to the molecular weight of a polypeptide or protein only when the charge/mass ratios and the shapes of all the SDS–polypeptide complexes are the same. Therefore valid molecular weight standards are required for molecular weight determination. (For more detailed discussion, see references 16–18.) Consequently, molecular weight determination of polypeptides in SDS–polyacrylamide gels is still an empirical procedure.

For proper molecular weight determination, the following criteria should be met.

(i) All polypeptides (standard and unknown) must bind the optimum amount of SDS and thus have a constant charge/mass ratio.
(ii) All SDS–polypeptide complexes (standard and unknown) must assume the same conformation or shape.
(iii) Electrophoresis of standard and unknown proteins must be run under the same conditions and at the same time (preferably on the same gel).

Three procedures are available for molecular weight determination using SDS–PAGE. They are:

(i) single gel system—plotting R_f versus log M_r;
(ii) multiple gel system—plotting K_r versus M_r;
(iii) gradient gel system—plotting log (%T) versus log M_r.

Each of these systems will be described and the various problems encountered will be discussed.

2. ELECTROPHORESIS SYSTEMS

SDS–PAGE may be run on either disc or slab gel systems. Presently, the slab gel system is most widely used for molecular weight determination. The advantages of this system are listed below.