Protein Function

A PRACTICAL APPROACH

SECOND EDITION

Edited by
T. E. CREIGHTON

The Practical Approach Series
Series Editor: B. D. Hames

Protein Function

A Practical Approach

Edited by

T. E. CREIGHTON

European Molecular Biology Laboratory
Heidelberg, Germany

IRL PRESS
at
OXFORD UNIVERSITY PRESS
Oxford New York Tokyo

Oxford University Press, Great Clarendon Street, Oxford OX2 6DP

Oxford New York
Athens Auckland Bangkok Bogota Bombay Buenos Aires
Calcutta Cape Town Dar es Salaam Delhi Florence Hong Kong
Istanbul Karachi Kuala Lumpur Madras Madrid Melbourne
Mexico City Nairobi Paris Singapore Taipei Tokyo Toronto

and associated companies in
Berlin Ibadan

Oxford is a trade mark of Oxford University Press

Published in the United States
by Oxford University Press Inc., New York

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
(Data available)

ISBN 0 19 963616 8 (Hbk)
ISBN 0 19 963615 X (Pbk)

Two volume set
ISBN 0 19 963617 6 (Hbk)
ISBN 0 19 963620 6 (Pbk)

Typeset by Footnote Graphics, Warminster, Wilts
Printed in Great Britain by Information Press, Ltd, Eynsham, Oxon.

Protein Function

The Practical Approach Series

SERIES EDITOR

B. D. HAMES
Department of Biochemistry and Molecular Biology
University of Leeds, Leeds LS2 9JT, UK

★ indicates new and forthcoming titles

Affinity Chromatography
★ Affinity Separations
Anaerobic Microbiology
Animal Cell Culture (2nd edition)
Animal Virus Pathogenesis
Antibodies I and II
★ Antibody Engineering
Basic Cell Culture
Behavioural Neuroscience
Biochemical Toxicology
Bioenergetics
Biological Data Analysis
Biological Membranes
Biomechanics—Materials
Biomechanics—Structures and Systems
Biosensors
★ Calcium-PI signalling
Carbohydrate Analysis (2nd edition)
Cell–Cell Interactions
The Cell Cycle
Cell Growth and Apoptosis
Cellular Calcium
Cellular Interactions in Development
Cellular Neurobiology
Clinical Immunology
★ Complement
Crystallization of Nucleic Acids and Proteins
Cytokines (2nd edition)
The Cytoskeleton
Diagnostic Molecular Pathology I and II
Directed Mutagenesis
★ DNA and Protein Sequence Analysis
DNA Cloning 1: Core Techniques (2nd edition)
DNA Cloning 2: Expression Systems (2nd edition)
★ DNA Cloning 3: Complex Genomes (2nd edition)
★ DNA Cloning 4: Mammalian Systems (2nd edition)
Electron Microscopy in Biology
Electrophysiology
Enzyme Assays

★ Epithelial Cell Culture
Essential Developmental Biology
Essential Molecular Biology I and II
Experimental Neuroanatomy
★ Extracellular Matrix
Flow Cytometry (2nd edition)
★ Free Radicals
Gas Chromatography
Gel Electrophoresis of Nucleic Acids (2nd edition)
Gel Electrophoresis of Proteins (2nd edition)
Gene Probes 1 and 2
Gene Targeting
Gene Transcription
Glycobiology
Growth Factors
Haemopoiesis
Histocompatibility Testing
HIV Volumes 1 and 2
Human Cytogenetics I and II (2nd edition)
Human Genetic Disease Analysis
★ Immunochemistry 1
★ Immunochemistry 2
Immunocytochemistry
In Situ Hybridization
Iodinated Density Gradient Media
★ Ion Channels
Lipid Analysis
Lipid Modification of Proteins
Lipoprotein Analysis
Liposomes
Mammalian Cell Biotechnology
Medical Bacteriology
Medical Mycology
Medical Parasitology
Medical Virology
★ MHC 1
★ MHC 2
Microcomputers in Biology
Molecular Genetic Analysis of Populations
Molecular Genetics of Yeast
Molecular Imaging in Neuroscience
Molecular Neurobiology
Molecular Plant Pathology I and II
Molecular Virology
Monitoring Neuronal Activity
Mutagenicity Testing
★ Neural Cell Culture
Neural Transplantation
★ Neurochemistry (2nd edition)
Neuronal Cell Lines
NMR of Biological Macromolecules
Non-isotopic Methods in Molecular Biology
Nucleic Acid Hybridization
Nucleic Acid and Protein Sequence Analysis
Oligonucleotides and Analogues
Oligonucleotide Synthesis
PCR 1
PCR 2
Peptide Antigens

Photosynthesis: Energy Transduction
Plant Cell Biology
Plant Cell Culture (2nd edition)
Plant Molecular Biology
Plasmids (2nd edition)
★ Platelets
Pollination Ecology
Postimplantation Mammalian Embryos
Preparative Centrifugation
Prostaglandins and Related Substances
Protein Blotting
Protein Engineering
Protein Function (2nd edition)
Protein Phosphorylation
Protein Purification Applications
Protein Purification Methods
Protein Sequencing
Protein Structure (2nd edition)
★ Protein Structure Prediction
Protein Targeting
Proteolytic Enzymes
Pulsed Field Gel Electrophoresis
Radioisotopes in Biology
Receptor Biochemistry
Receptor–Ligand Interactions
RNA Processing I and II
★ Subcellular Fractionation
Signal Transduction
Solid Phase Peptide Synthesis
Transcription Factors
Transcription and Translation
Tumour Immunobiology
Virology
Yeast

Preface

This second edition continues the tradition of the first, to present experimental procedures that will be applicable to many proteins and that can be performed in an average laboratory, without the need for specialized equipment.

A first priority in characterizing any new protein is to identify it by its primary structure. Very often, a new protein is found to be homologous in sequence to a protein of known function and to share at least some of those functional properties. The latest techniques for identifying proteins are described in Chapter 1.

It is very often the case that a new protein is first identified genetically, so its gene and primary structure are known before the protein itself has been noticed. It is not unusual for the protein in its natural state to be present in only minute quantities, so that its purification is impractical. This is most often the case with proteins of the greatest functional specific activity and of the greatest pharmaceutical use. This low level of natural occurrence is no longer a major obstacle to characterizing and making practical use of the protein, for it is almost routine to produce a protein in large quantities by expressing its cloned gene. The most useful host for this procedure is *Escherichia coli*, and the procedures for expressing any cloned gene in this way are described in Chapter 2.

Genes expressed in *E. coli* or other hosts often produce their polypeptide chains in an unfolded, insoluble state. This is frequently the greatest obstacle to producing large quantities of a protein in a useful form. The expressed protein must be induced to adopt its native, functional state before it will be useful. The general procedures for doing so are described in Chapter 3.

A common theme of protein function is that it invariably involves the protein interacting physically with other molecules; a protein never acts in isolation, but always acts upon something. So a major concern is to characterize the interaction of the protein with these other molecules. Chapter 4 describes the methods for doing so, concentrating on membrane-bound proteins, especially receptors, which play such a central role. The thorny question of the significance and molecular basis of co-operativity in ligand binding, especially the more frequent negative co-operativity, is addressed, along with suggested procedures for clarifying the situation.

Electrophoresis is a major technique in studying protein structure (see *Protein structure: a practical approach* and *Gel electrophoresis of proteins: a practical approach*), and it is becoming a very useful technique for studying protein function also. Central to this is the ability to blot a protein band or spot from a polyacrylamide gel onto a membrane to which it sticks tightly. Somewhat miraculously, a significant fraction of these bound molecules can be induced to refold and to regain their ligand-binding properties. As described in Chapter 5, these functional properties can be characterized very easily.

Biologically relevant ligand binding occurs almost invariably at specific sites on proteins, and it is important to identify and characterize all such binding sites. One of the most direct methods for doing so is by affinity labelling, described in Chapter 6. A reactive group is incorporated into a ligand and thereby reacts with the protein much more rapidly when bound than when free in solution, due to the very high 'effective concentrations' that can occur in ligand–protein complexes. With larger ligands, such as other protein molecules, the interacting macromolecules can be identified by cross-linking them covalently in the complex, using reagents with two reactive groups, one at each end of a suitable linker moiety. Techniques for doing so are described in Chapter 7. The procedures described in these two chapters are illustrated for specific classes of ligands, but they should be readily adapted to other complexes and ligands.

One of the most biologically important areas of protein function is in the control of gene expression, which invariably involves proteins binding to DNA and to RNA. Most of the regulatory proteins occur in very small quantities within the cell and have consquently been very difficult to study. Many of the techniques used have relied upon the properties of the nucleic acids, rather than the protein itself. These involve the identification of DNA–protein complexes by the change in electrophoretic mobility of a small fragment of DNA produced by its binding to a protein; such complexes are usually extremely tight and long-lived, so that they can survive an electrophoretic separation, as in 'bandshift gels'. The specific sites on the DNA occupied by the protein can be identified by the aptly-named 'footprinting' technique. Chapter 8 describes these techniques using purified sequence-specific DNA-binding proteins, while Chapter 9 describes how to use them to identify such proteins in crude extracts, and then how to purify them with the use of DNA affinity chromatography. (The general techniques of protein purification and affinity chromatography with other ligands are described in the volumes *Protein purification: a practical approach* and *Affinity chromatography: a practical approach*.)

The quickest way to identify the functional groups of a protein, even with the advent of protein engineering, remains the classical approach of chemical modification, described in Chapter 10. Once they have been identified, their functional roles can be delineated using the more specific approach of site-directed mutagenesis and protein engineering.

Comparable simple techniques for characterizing protein structure are to be found in the companion volume *Protein structure: a practical approach*. For a comprehensive description of the properties of proteins, see the second edition of my volume *Proteins: structures and molecular properties* (W. H. Freeman, New York, 1993).

Heidelberg, Germany T. E. C.
March 1997

Contributors

RON D. APPEL
Computer Center, Geneva University Hospital, 24 Rue Micheli-du-Crest, CH-1211 Geneva 14, Switzerland.

STEVE BASS
Department of Molecular Biology, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080–4990, USA.

GERALD BÖHM
Institut für Biotechnologie, Martin-Luther-Universität Halle, Kurt-Mothes-Strasse 3, D-06120 Halle (Saale), Germany.

ROBERTA F. COLMAN
Department of Chemistry and Biochemistry, University of Delaware, Newark, DE 19716, USA.

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

DENIS F. HOCHSTRASSER
Clinical Chemistry Laboratory, Geneva University Hospital, 24 Rue Micheli-du-Crest, CH-1211 Geneva 14, Switzerland.

TAIJI IMOTO
Graduate School of Pharmaceutical Sciences, Kyushu University 62, Maidashi, Higashi-ku, Fukuoka 812–82, Japan.

RAINER JAENICKE
Institut für Biophysik und physikalische Biochemie, Universität Regensburg, Universitätsstrasse 31, D-93053 Regensburg, Germany.

KEN B. KAPLAN
Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139–4307, USA.

RONALD KLUGER
Department of Chemistry, University of Toronto, 80 St. George Street, Toronto, Ontario, Canada M5S 3H6.

ALEXANDER LEVITZKI
Department of Biological Chemistry, Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel.

HAUKE LILIE
Institut für Biotechnologie, Martin-Luther-Universität Halle, Kurt-Mothes-Strasse 3, D-06120 Halle (Saale), Germany.

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

RAINER RUDOLPH
Institut für Biotechnologie, Martin-Luther-Universität Halle, Kurt-Mothes-Strasse 3, D-06120 Halle (Saale), Germany.

JEAN-CHARLES SANCHEZ
Clinical Chemistry Laboratory, Geneva University Hospital, 24 Rue Michelidu-Crest, CH-1211 Geneva 14, Switzerland.

PETER K. SORGER
Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139–4307, USA.

ANNE K. SOUTAR
MRC Lipoprotein Team, Clinical Sciences Centre, Royal Postgraduate Medical School, Hammersmith Hospital, London, UK.

DAVID P. WADE
MRC Lipoprotein Team, Clinical Sciences Centre, Royal Postgraduate Medical School, Hammersmith Hospital, London, UK.

MARC R. WILKINS
Clinical Chemistry Laboratory, Geneva University Hospital, 24 Rue Michelidu-Crest, CH-1211 Geneva 14, Switzerland.

KEITH L. WILLIAMS
Macquarie University Centre for Analytical Biotechnology, Sydney N.S.W. 2109, Australia.

HIDENORI YAMADA
Department of Bioengineering Science, Faculty of Engineering, Okayama University, Tsushima-Naka, Okayama 700, Japan.

MARIA YANG
Department of Quality Control; Clinical Development, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080–4990, USA.

Abbreviations

AAA	amino acid analysis
AMPS-BDB	adenosine 5′-*O*-[*S*-(4-bromo-2,3- ioxobutyl)] thiophosphate
AMPS-BOP	adenosine 5′-*O*-[*S*-(3-bromo-2-oxopropyl)] thiophosphate
A_x	absorbance at wavelength *x* nm
2-BDB-TεA 2′,5′-DP	2-[4-bromo-2,3-dioxobutylthio]-1,N^6-ethenoadenosine 2′,5′-diphosphate
BPG	2,3-bis(phospho)glycerate
BPTI	bovine pancreatic trypsin inhibitor (aprotinin, Trasylol™)
BSA	bovine serum albumin
BSP	bromosulfophthalein
Caps	3-(cyclohexylamino)-1-propanesulfonic acid
CBS-	4-carboxybenzenesulfonyl-
CD	circular dichroism
cDNA	complementary DNA
CDNB	1-chloro-2,4-dinitrobenzene
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propane sulfonate
2D	two-dimensional
DBBF	fumaroyl *bis*(3,5-dibromosalicylate)
DBST	trimesyl dibromosalicyl
DEAE	diethylaminoethyl
DMEM	Dulbecco's modified Eagle medium
DMS	dimethylsulfate
DMSO	dimethyl sulfoxide
DNase	deoxyribonuclease
DOC	deoxycholate
DON	6-diazo-5-oxonorleucine
DTE	dithioerythritol
DTNB	5,5′-dithionitrobenzoic acid
DTT	dithiothreitol
ECL	enhanced chemiluminescence
EDC	1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide hydrochloride
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EGTA	ethylene glycol-bis(β-aminoethyl ether)-*N,N,N′,N′*-tetraacetic acid

ELISA	enzyme-linked immunosorbent assay
ε_x	molar absorbance at wavelength x, when specified
Fab	immunoglobulin fragment, antigen binding
FCS	fetal calf serum
Fmoc	9-fluorenylmethyl chloroformate
4-FSB	4-(fluorosulfonyl) benzoic acid
5′-FSBA	5′-*p*-fluorosulfonylbenzoyl adenosine
5′-FSBAzA	5′-*p*-fluorosulfonylbenzoyl-8-azidoadenosine
GdmCl	guanidinium chloride
$(GlcNAc)_3$	trimer of *N*-acetylglucosamine
GMPS-BDB	guanosine 5′-*O*-[*S*-(4-bromo-2,3-dioxobutyl)] thiophosphate
GMPS-BOP	guanosine 5′-*O*-[*S*-(3-bromo-2-oxopropyl)] thiophosphate
GSH	glutathione, thiol form
GSSG	glutathione, disulfide form
Hepes	*N*-2-hydroxyethyl piperazine-*N*′-2-ethanesulfonic acid
HSE	heat shock element
HSF	heat shock factor
IB	inclusion body
Ig	immunoglobulin
IPG	immobilized pH gradient
IU	International Unit; catalysis of 1 μmol substrate per min
kb	kilobase
kDa	kilodalton
LDH	lactate dehydrogenase
LDL	low-density lipoprotein
MAP	methyl acetyl phosphate
mBBr	monobromobimane
M_2C_2H	4-(*N*-maleimidomethyl) cyclohexane-1-carboxyl hydrazide
mol. wt	molecular weight (dimensionless)
Mops	3-(*N*-morpholino)propanesulfonic acid
M_r	relative molecular mass (dimensionless)
mRNA	messenger RNA
MS	mass spectrometry
NAD^+	nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
NBS	*N*-bromosuccinimide
NMNS-BOP	nicotinamide ribose 5′-*O*-[*S*-(3-bromo-2-oxopropyl)] thiophosphate
NMR	nuclear magnetic resonance
NTB	2-nitro-5-thiobenzoic acid

NTSB	2-nitro-5-thiosulfobenzoate
oligo	oligonucleotide
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDA	piperazine diacrylamide
PDI	protein disulfide isomerase
PEG	polyethylene glycol
pI	isoelectric point
Pipes	piperazine-*N,N'*-bis(2-ethanesulfonic acid)
PMSF	phenylmethylsulfonyl fluoride
PPI	peptidyl-prolyl-*cis/trans*-isomerase
PTH	phenylthiohydantoin
PVDF	polyvinylidene difluoride
rGH	rat growth hormone
RNase	ribonuclease
RP	reversed-phase
RU	resonance units
S-BDB-G	*S*-(4-bromo-2,3-dioxobutyl)-glutathione
SD	Shine–Dalgarno
SDS	sodium dodecyl sulfate
SPDP	*N*-succinimidyl 3-(2-pyridyldithio)propionate
SPR	surface plasmon resonance
STI	soybean trypsin inhibitor
STII	stable enterotoxin II
TB	Tris–borate
TBE	Tris–borate–EDTA
TCA	trichloroacetic acid
TEMED	*N,N,N',N'* tetramethylethylenediamine
TFA	trifluoroacetic acid
TIR	translation initiation region
TLC	thin-layer chromatography
TNBS	2,4,6-trinitrobenzenesulfonic acid
TNM	tetranitromethane
TPCK	*N*-tosyl-L-phenylalanine chloromethyl ketone
tRNA	transfer RNA
TSH	thyroid stimulating hormone
TTDS	trimesyl tris(3,5-dibromosalicylate)
TTK	transcription factor Tramtrack
U	unit
UV	ultraviolet
WSB	washing sample buffer

Contents