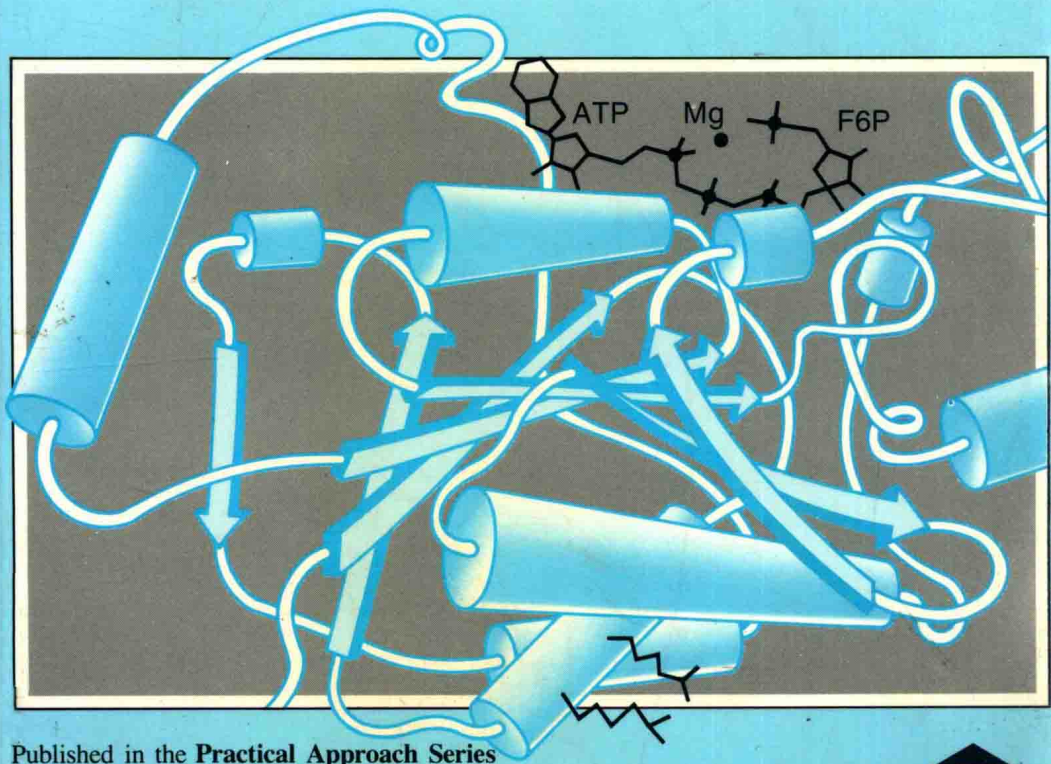
Protein function

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 function

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 function.
1. Proteins
I. Creighton, Thomas E. *1940–*
II. Series
547.7′5

Library of Congress Cataloging in Publication Data

Protein function.
(Practical approach series)
Includes bibliographies and index.
1. Proteins—Analysis. 2. Proteins—Structure-activity relationships.
3. Proteins—Affinity labeling. 4. Biochemistry—Technique.
I. Creighton, Thomas E., 1940– . II. Series. [DNLM 1. Proteins—physiology. QU 55 P96645]
QP551.P69583 1988 574.19′245 88-31963
ISBN 0 19 963006 2 (hardback)
ISBN 0 19 963007 0 (softback)

Previously announced as:
ISBN 1 85221 139 3 (hardback)
ISBN 1 85221 140 7 (softback)

Typeset by Paston Press, Loddon, Norfolk.
Printed by Information Printing Ltd, Oxford, England.

Protein function

a practical approach

TITLES PUBLISHED IN THE PRACTICAL APPROACH SERIES

Series editors:

Dr D Rickwood

Department of Biology, University of Essex
Wivenhoe Park, Colchester, Essex CO4 3SQ, UK

Dr B D Hames

Department of Biochemistry, University of Leeds
Leeds LS2 9JT, UK

Affinity chromatography
Animal cell culture
Antibodies
Biochemical toxicology
Biological membranes
Carbohydrate analysis
Centrifugation (2nd Edition)
DNA cloning
Drosophila
Electron microscopy in molecular biology
Gel electrophoresis of nucleic acids
Gel electrophoresis of proteins
Genome analysis
HPLC of small molecules
HPLC of macromolecules
Human cytogenetics
Human genetic diseases
Immobilised cells and enzymes
Iodinated density gradient media
Light microscopy in biology
Lymphocytes
Lymphokines and interferons
Mammalian development
Microcomputers in biology
Microcomputers in physiology
Mitochondria
Mutagenicity testing
Neurochemistry
Nucleic acid and protein sequence analysis
Nucleic acid hybridisation
Oligonucleotide synthesis
Photosynthesis: energy transduction
Plant cell culture
Plant molecular biology
Plasmids
Prostaglandins and related substances
Protein function
Protein structure
Spectrophotometry and spectrofluorimetry
Steroid hormones
Teratocarcinomas and embryonic stem cells
Transcription and translation
Virology
Yeast

Preface

The functional properties of proteins vary enormously. Some are simply structural, some catalyse chemical reactions; others transmit information, or interconvert the various forms of energy: chemical, light, movement. Consequently, a comprehensive guide to studying protein function would encompass all of biochemistry.

This volume does not attempt that, but concentrates on those aspects of protein function that are common to most, if not all, proteins. The first of these is the dependence of the functional properties upon the covalent structure and the conformation of each protein. The first priority in any study of protein function should be to preserve the structural and conformational integrity of the protein; to study a 'dying' protein is a hopeless task. Therefore, the very first chapter provides guidelines as to how to maintain protein function, both by minimizing covalent alterations of its structure and by maintaining its folded conformation. Further procedures for studying protein structure and conformation are given in the companion volume *Protein Structure: A Practical Approach*.

The other 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. Therefore, a primary concern is to characterize the interaction of the protein with these other molecules. General procedures for measuring the most basic parameters, the number of ligand molecules bound by a protein and their relative affinities, are described in Chapter 2. Such studies have a very long history, but there are a surprising number of misconceptions about the intepretation of binding data when multiple ligands are bound to the same protein. Consequently, a major part of such a chapter is pointing out how *not* to proceed, and this chapter also provides numerous examples of the errors commonly made in such studies.

Electrophoresis is a major technique in studying protein structure (see *Protein Structure: A Practical Approach* and *Gel Electrophoresis of Proteins: A Practical Approach*), and is also becoming a very useful technique in studying protein function, with the recent advances in blotting techniques described in Chapter 3. The proteins in an electrophoretic gel can be transferred to a membrane, to which they stick tightly. Although I personally find it difficult to understand how it happens, it is an undisputed fact that a sufficient fraction of the protein molecules refold to a sufficient extent to exhibit ligand-binding ability, in spite of having been denatured and stuck to the membrane. With this simple technique, the abilities of numerous proteins to bind virtually any ligand can be tested very simply.

Biologically relevant ligand binding invariably occurs at specific sites on proteins, so 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 4. A reactive group is incorporated into a ligand and 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 molecules can be identified by cross-linking them covalently in the complex. Techniques for doing so are described in Chapter 5. The procedures described in these two chapters are illustrated for a specific class of ligands, but should be readily adapted to other complexes.

Very many proteins consist of multiple polypeptide chains, usually as relatively autonomous structural subunits. This oligomeric structure often has profound implications for the function of the protein, but in many cases the functional implications are not at all obvious. Very many simple, but ingenious, techniques have been devised to examine the roles of subunits in protein function, and these are described in Chapter 6.

One of the most biologically important areas of protein function is in the control of gene expression, which invariably involves protein binding to DNA and RNA. Most of the regulatory proteins occur in very small quantities within cells and have consequently been very difficult to study. Most of the techniques used have relied upon the properties of the nucleic acids, rather than the protein. These involve the identification of DNA–protein complexes by the change in electrophoretic mobility of a small fragment of DNA produced by a protein binding to it; such complexes are extremely tight and consequently dissociate so slowly that the complex 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 7 describes these techniques using purified proteins, while Chapter 8 describes how to use them to identify sequence-specific DNA-binding proteins in crude mixtures, then 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 Approqch*.)

The functional groups involved in protein function often have somewhat unusual physical properties, and these may be characterized by the relatively simple technique of competitive labelling described in Chapter 9. This technique can also be used to identify binding sites on proteins for ligands by comparing the protein with the protein–ligand complex, since interaction with another molecules usually causes changes in the reactivities of the functional groups involved. The classical technique for identifying functional groups involved in protein function is to examine the functional effects of chemically modifying the various classes of reactive groups. A number of new approaches and reagents have been developed in recent years, and some of these are described in Chapter 10.

Finally, the most specific modifications of protein structure are those produced by the recently-developed techniques of site-directed mutagenesis, and this technique has become so widely used that no volume on protein function would be complete without it, even though the procedures described do not actually involve the protein. A gene for the protein is required, but this can now be obtained almost routinely by the procedures described in *DNA Cloning: A Practical Approach*, volumes 1–3) or by gene synthesis (*Oligonucleotide Synthesis: A Practical*

Approach). The procedures described in Chapter 11 are some of the most recent and most efficient yet devised.

Use of the techniques described in this volume should provide much information about the functional properties of any protein, but the procedures must be used appropriately, taking into account all the relevant properties of proteins. For a comprehensive description, the reader is referred to my volume (*Proteins: Structures and Molecular Properties*. W. H. Freeman, New York, 1983). For relatively simple techniques to characterize a protein's structure, the companion volume *Protein Structure: A Practical Approach* is highly recommended.

T. E. Creighton

Contributors

G. Ammerer
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK. Present address: Institute of Molecular Pathology, Dr. Bohr Gasse, A1030 Vienna and Institut für Allgemeine Biochemie, Universität Wien, A1090 Vienna, Austria

C. Casiano
Department of Biological Chemistry, University of California School of Medicine, Davis, CA 95616, USA

R. F. Colman
Department of Chemistry and Biochemistry, University of Delaware, Newark, DE 19716, USA

F. Eckstein
Max-Planck-Institut für Experimentelle Medizin, Abteilung Chemie, Herman-Rein-Strasse 3, D-3400 Göttingen, FRG

E. Eisenstein
Molecular Biology-Virus Laboratory, 229 W. M. Stanley Hall, University of California, Berkeley, CA 94720, USA

T. Imoto
Faculty of Pharmaceutical Sciences, Kyushu University 67, Maidashi, Higashi-ku, Fukuoka 812, Japan

H. Kaplan
Department of Biochemistry, University of Ottawa, Ottawa, Canada K1N 9B4

A. M. Klibanov
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

I. M. Klotz
Department of Chemistry, Northwestern University, Evanston, IL 60201, USA

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

J. R. Sayers
Max-Planck-Institut für Experimentelle Medizin, Abteilung Chemie, Herman-Rein-Strasse 3, D-3400 Göttingen, FRG

H. K. Schachman
Molecular Biology-Virus Laboratory, 229 W. M. Stanley Hall, University of California, Berkeley, CA 94720, USA

D. Shore
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK. Present address: Department of Microbiology, Columbia University, New York, NY 10032, USA

P. K. Sorger
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

A. K. Soutar
MRC Lipid Research Unit, Hammersmith Hospital, DuCane Road, London W12 0HS, UK

R. R. Traut
Department of Biological Chemistry, University of California School of Medicine, Davis, CA 95616, USA

D. B. Volkin
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

D. P. Wade
MRC Lipid Research Unit, Hammersmith Hospital, DuCane Road, London W12 0HS, UK

H. Yamada
Faculty of Pharmaceutical Sciences, Kyushu University 67, Maidashi, Higashi-ku, Fukuoka 812, Japan

N. M. Young
Division of Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Canada K1A 0R6

N. Zecherle
Department of Biological Chemistry, University of California School of Medicine, Davis, CA 95616, USA

Abbreviations

AP1	yeast homologue of mammalian activator protein 1
APOP	*N*-[4-(*p*-azidosalicylamido)butyl]-3′-(2′-pyridyldithio)propionamide
APTP	*N*(4-azidophenyl)phthalimide
ATCase	aspartate transcarbamoylase
X-BDB-TA 5′-DP	X-[4-bromo-2,3-dioxobutylthio]-adenosine 5′-diphosphate
2-BDB-TAMP	2-[4-bromo-2,3-dioxobutylthio]adenosine 5′-monophosphate
BSA	bovine serum albumin
CBS-Lys	N^{ε}-(4-carboxybenzenesulphonyl) lysine
CBS-Tyr	*O*-(4-carboxybenzenesulphonyl) tyrosine
Chaps	3-[(3-cholamidopropyl)dimethyl ammonio]-1-propane sulphonate
CP1	centromere-binding protein
DMF	dimethylformamide
DMS	dimethylsulphate
DMSO	dimethyl sulphoxide
DNP	dinitrophenyl
DNP-F	1-fluoro-2,4-dinitrobenzene
dNTPαS	deoxynucleoside 5′-*O*-(1-thio) triphosphate
DTBP	dithiobispropionimidate
DTT	dithiothreitol
EDC	1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide
EDTA	ethylenediamine tetraacetic acid
5′-FSBA	5′-*p*-fluorosulphonylbenzoyl adenosine
5′-FSBεA	5′-*p*-fluorosulphonylbenzoyl-1,N^6-enthenoadenosine
5′-FSBG	5′-*p*-fluorosulphonylbenzoyl guanosine
GdmCl	guanidinium hydrochloride
HSE	heat shock element segment of DNA
HSF	heat shock transcription factor protein from yeast
Hepes	*N*-hydroxyethyl-piperazine-*N*′-2-ethanesulphonic acid
HPLC	high-performance liquid chromatography
IPTG	isopropyl-β-D-thiolgalactopyranoside
2-IT	2-iminothiolane
LDL	low-density lipoprotein
Mes	2-(N-morpholino) ethane sulphonic acid
NBS	*N*-bromosuccinimide
NMR	nuclear magnetic resonance
PADR	2′-phosphoadenosine 5′-diphosphoribose
PAGE	polyacrylamide gel electrophoresis
PALA	*N*-phosphonacetyl-L-aspartate
PEG	polyethylene glycol
Pipes	piperazine-*N*,*N*′-bis-(2-ethanesulphonic acid)

PMSF	phenylmethylsulphonyl fluoride
PRTF	pheromone receptor transcription factor protein
PTH	phenylthiohydantoin
RAP1	repressor activator protein 1
RF IV	double-stranded closed-circular M13 DNA
SBF-B	silencer binding factor-B/ARS binding factor 1 protein
SDS	sodium dodecyl sulphate
SPDP	*N*-succinimidyl 3-(2-pyridyldithio)propionate
2-TA 2′,5′-DP	2-thioadenosine 2′,5′-bisphosphate
TB	90 mM Tris–borate buffer, pH 8.3
TBE	90 mM Tris–borate buffer, pH 8.3, 2 mM EDTA
TCA	trichloroacetic acid
TEMED	*N,N,N′,N′*-tetramethylethylene diamine
TFIIIA	transcription factor IIIA protein from *Xenopus*
THPA	3,4,5,6-tetrahydrophthalic anhydride
TLC	thin-layer chromatography
TLCK	N_a-*p*-tosyl-L-lysine chloromethyl ketone
TNBS	trinitrobenzene sulphonic acid
TNM	tetranitromethane
TPCK	*N*-tosyl-L-phenylalanine chloromethyl ketone
Tris	tris (hydroxymethyl) amino methane
VLDL	very low-density lipoprotein
X-gal	5-bromo-4-chloro-3-indolyl-β-galactoside

Contents