RNA:Protein Interactions

A PRACTICAL APPROACH

Edited by
CHRISTOPHER W. J. SMITH

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

http://www.oup.co.uk/PAS

RNA:Protein Interactions

A Practical Approach

Edited by

CHRISTOPHER W. J. SMITH

Department of Biochemistry
University of Cambridge

Oxford New York Tokyo
OXFORD UNIVERSITY PRESS
1998

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

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

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 963651 6 (Hbk)
ISBN 0 19 963650 8 (Pbk)

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

The Practical Approach Series

SERIES EDITOR

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

See also the Practical Approach web site at **http://www.oup.co.uk/PAS**

★ 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
★ Antisense Technology
★ Applied Microbial Physiology
Basic Cell Culture
Behavioural Neuroscience
Bioenergetics
Biological Data Analysis
Biomechanics—Materials
Biomechanics—Structures and Systems
Biosensors
Carbohydrate Analysis (2nd edition)
Cell–Cell Interactions
The Cell Cycle
Cell Growth and Apoptosis
Cellular Calcium
Cellular Interactions in Development
Cellular Neurobiology
★ Chromatin
Clinical Immunology
★ Complement
Crystallization of Nucleic Acids and Proteins
Cytokines (2nd edition)
The Cytoskeleton
Diagnostic Molecular Pathology I and II
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)
★ Drosophila (2nd edition)
Electron Microscopy in Biology

Electron Microscopy in Molecular 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 (3rd edition)

Gene Probes 1 and 2

Gene Targeting

Gene Transcription

★ Genome Mapping

Glycobiology

★ Growth Factors and Receptors

Haemopoiesis

Histocompatibility Testing

HIV Volumes 1 and 2

★ HPLC of Macromolecules (2nd edition)

Human Cytogenetics I and II (2nd edition)

Human Genetic Disease Analysis

★ Immunochemistry 1

★ Immunochemistry 2

Immunocytochemistry

★ *In Situ* Hybridization (2nd edition

Iodinated Density Gradient Media

Ion Channels

★ Light Microscopy (2nd edition)

Lipid Modification of Proteins

Lipoprotein Analysis

Liposomes

Mammalian Cell Biotechnology

Medical Parasitology

Medical Virology

★ MHC Volumes 1 and 2

★ Molecular Genetic Analysis of Populations (2nd edition)

Molecular Genetics of Yeast

Molecular Imaging in Neuroscience

Molecular Neurobiology

Molecular Plant Pathology I and II

Molecular Virology

Monitoring Neuronal Activity

Mutagenicity Testing

★ Mutation Detection

Neural Cell Culture

Neural Transplantation

★ Neurochemistry (2nd edition)

Neuronal Cell Lines

NMR of Biological Macromolecules

Non-isotopic Methods in Molecular Biology

Nucleic Acid Hybridization

Oligonucleotides and Analogues

Oligonucleotide Synthesis

PCR 1

PCR 2

★ PCR 3: PCR *In Situ* Hybridization

Peptide Antigens

Photosynthesis: Energy Transduction

Plant Cell Biology

Plant Cell Culture (2nd edition)

Plant Molecular Biology

Plasmids (2nd edition)

★ Platelets

Postimplantation Mammalian Embryos

Preparative Centrifugation

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

RNA Processing I and II

★ RNA:Protein Interactions

★ Signalling by Inositides

Subcellular Fractionation

Signal Transduction

Transcription Factors

Tumour Immunobiology

Preface

Recent years have seen a resurgence of interest in the structure and biology of RNA, witnessed by the formation in 1993 of the 'RNA Society' and the launch of its journal *RNA*. In part, this renewal has arisen from the realization that RNA, once thought of as a passive carrier of information (mRNA) or as a scaffold for proteins (rRNA), in fact harbours great diversity in structure and versatility in function. Similar to DNA both chemically and in its capacity as an information carrier, in its ability to adopt complex globular folds, and to mediate chemical reactions, RNA is more reminiscent of proteins. The appreciation of this point has led to a more 'ribocentric' view of processes such as pre-mRNA splicing in the spliceosome and peptide-bond formation in the ribosome, in which the snRNAs and rRNAs are thought to have the primary responsibility for catalysis. Nevertheless, RNA in the cell is always found in complex with proteins and even bona fide ribozymes such as RNase P and group II introns require protein cofactors for their function *in vivo*. Interest in RNA also stems from the fact that gene expression entails multiple steps that are mediated and regulated by RNA:protein interactions. These are mostly post-transcriptional—pre-mRNA splicing, polyadenylation, editing, transport, cytoplasmic targeting, translation, mRNA turnover—but some are even transcriptional, as illustrated by HIV Tat protein and λ N protein, which control different aspects of transcription. Indeed most stages of the life cycle of retroviruses such as HIV involve critical RNA:protein recognition events. The detection and characterization of RNA:protein interactions is therefore an important challenge facing many investigators. This book attempts to bring together a comprehensive assembly of the techniques currently used in studying RNA:protein interactions. The approaches covered range from those that would initially be used in detecting a protein that binds to a particular RNA site to those that would be used to analyse the interactions between pure RNA and protein in precise molecular detail. In choosing topics and authors for this volume I was particularly mindful of the previous volumes in this series: *The ribosome and protein synthesis* and *RNA processing*. Given that the ribosome and spliceosome are two of the major theatres of RNA:protein interactions, it is not surprising that these previous volumes contain many techniques applicable to general RNA:protein problems and both are still highly recommended. However, the expanding interest in various aspects of RNA:protein interactions has been accompanied by some major technical innovations. Happily, a number of the pioneers of these advances have agreed to describe them in this book. While the authors have drawn working examples from their own specialized systems, they have attempted to highlight the general principles and applicability of their techniques. In keeping with the traditions of the *Practical Approach* series the chapters

are written in a clear explanatory style and are liberally peppered with protocols.

The first two chapters deal with large scale synthesis of RNA. Mike Gait and colleagues describe the current technology for automated chemical synthesis of oligoribonucleotides including incorporation of base-, sugar-, and phosphate-modified residues. They also describe a number of applications of chemically synthesized RNA, drawing upon their own experience with the HIV-1 Tat:TAR interaction. Kyoshi Nagai and Gabriele Varani's groups have been at the forefront of structural investigations of RNA:protein complexes using both X-ray crystallography and NMR. They detail methods for large scale enzymatic synthesis of RNA for these investigations including the use of internal ribozymes to guarantee homogeneity of product and also the production of isotopically labelled nucleotides for heteronuclear NMR. The chemical and enzymatic approaches are combined in the third chapter where Melissa Moore describes the oligonucleotide-mediated RNA ligation technique that she and Phil Sharp developed for synthesis of long site-specifically modified RNAs. She and Charles Query detail applications of this technology for site-specific RNA:protein cross-linking. Doug Black and Leslie Bell have collaborated to produce a chapter on one of the 'work-horse' techniques for studying nucleic acid:protein interactions, the gel shift assay. They detail methods for using this assay with both crude cell extracts, including identification of components of gel shifted complexes, as well as more quantitatively with pure proteins. The following four chapters deal with a variety of strategies for isolating novel RNA binding proteins or their cDNAs. Richard Jackson, Nancy Standart, and colleagues detail the various affinity chromatographic methods that often provide the crucial final step in a protein purification strategy. Michael Rosbash and colleagues outline synthetic lethal genetic approaches that they have used to uncover interacting partners in the yeast spliceosome. Jeffrey Willusz writes about the 'Northwestern' method for directly screening expression libraries with a labelled RNA ligand as probe. Next, Marvin Wickens, Stan Fields, and colleagues describe their recent elaboration of the yeast 'two-hybrid' system into a 'three-hybrid' system capable of detecting RNA:protein interactions *in vivo*. This method already has a track record in the identification of novel RNA binding proteins and, as outlined in the chapter, it has a number of other potential applications. Harada and Frankel describe an *in vivo* selection system in *E. coli* that has been used to select optimal RNA binding polypeptides for a particular RNA ligand. The emphasis in this method has been to test combinatorial variants of known RNA binding proteins for improved or novel specificities. Merryman and Noller give details of various chemical probing methods that are used to map RNA:protein and RNA:RNA interactions, drawing upon their extensive experience of footprinting and modification interference assays in the ribosome. Jutta Rinke-Appel and Richard Brimacombe then detail site-specific cross-linking methods that have been used to study RNA:RNA and

RNA:protein contacts in the ribosome, complementing the examples from the spliceosome given by Moore and Query. Finally, Andrew Ellington and colleagues describe the *in vitro* selection of nucleic acid ligands from large degenerate sequence pools (often referred to as SELEX). Many users of this book would be content to use this approach to define the optimal RNA binding site for confirmed RNA binding proteins. However, this chapter also outlines the more ambitious approach (100 ml PCR reactions!) of finding specific RNA partners, 'aptamers', for proteins that never knew that they could bind RNA. This is an area that is currently of intense interest as it promises the potential to develop RNA as a therapeutic tool.

I thank all of the authors for agreeing to contribute to this book, for their diligence, and for at times accepting suggestions for changes. I think their efforts have combined to produce a clear and up-to-date text that contains all the basic approaches as well as the cutting-edge techniques.

Cambridge C. W. J. S.
July 1997

Contributors

LESLIE BELL
Molecular Biology Program, Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089–1340, USA.

SABINE BELL
Department of Chemistry, University of Indiana, Bloomington, IN 47405, USA.

DOUGLAS L. BLACK
Molecular Biology Institute, University of California, Los Angeles, CA 90095, USA.

RICHARD BRIMACOMBE
Max-Planck-Institut für Molekulare Genetik, Ihnestrasse 73, 14195-Berlin, Germany.

F. MAIKE BRÜCK
Department of Chemistry, University of Indiana, Bloomington, IN 47405, USA.

RAYMOND CHAN
Molecular Biology Institute, University of California, Los Angeles, CA 90095, USA.

RICHARD C. CONRAD
The Center for Aptamer Research, Indiana Molecular Biology Institute, and Department of Chemistry, University of Indiana, Bloomington, IN 47405, USA.

DAVID J. EARNSHAW
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

ANDREW D. ELLINGTON
The Center for Aptamer Research, Indiana Molecular Biology Institute, and Department of Chemistry, University of Indiana, Bloomington, IN 47405, USA.

MARK A. FARROW
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

STANLEY FIELDS
Departments of Genetics and Medicine, Markey Molecular Medicine Center, University of Washington, Box 357360, Seattle, WA 98195, USA.

JAN H. FOGG
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

ALAN D. FRANKEL
Department of Biochemistry and Biophysics, 513 Parnassus Avenue, University of California, San Francisco, CA 94143–0448, USA.

MICHAEL J. GAIT
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

RICHARD L. GRENFELL
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

KAZUO HARADA
Department of Biochemistry and Biophysics, 513 Parnassus Avenue, University of California, San Francisco, CA 94143–0448, USA.

RICHARD J. JACKSON
Department of Biochemistry, University of Cambridge, Old Addenbrookes Site, 80 Tennis Court Road, Cambridge CB2 1GA, UK.

ANN KAMINSKI
Department of Biochemistry, University of Cambridge, Old Addenbrookes Site, 80 Tennis Court Road, Cambridge CB2 1GA, UK.

BRIAN KRAEMER
Department of Biochemistry, 420 Henry Mall, University of Wisconsin, Madison, WI 53706, USA.

CHUCK MERRYMAN
Center for Molecular Biology of RNA, Sinsheimer Laboratories, University of California, Santa Cruz, CA 95064, USA.

HOSUNG MIN
Molecular Biology Institute, University of California, Los Angeles, CA 90095, USA.

MELISSA J. MOORE
Howard Hughes Medical Institute, Department of Biochemistry, Brandeis University, Waltham, MA 02254, USA.

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

NIKOLAI A. NARYSHKIN
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

HARRY F. NOLLER
Center for Molecular Biology of RNA, Sinsheimer Laboratories, University of California, Santa Cruz, CA 95064, USA.

DIRK H. OSTARECK
Gene Expression Program, EMBL, 1 Meyerhofstrasse, Heidelberg 69117, Germany.

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

STEPHEN R. PRICE
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

CHARLES C. QUERY
Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY 10461, USA.

JUTTA RINKE-APPEL
Max-Planck-Institut für Molekulare Genetik, Ihnestrasse 73, 14195-Berlin, Germany.

MICHAEL ROSBASH
Howard Hughes Medical Institute, Biology Department, Brandeis University, Waltham, MA 02254, USA.

DHRUBA SENGUPTA
Departments of Genetics and Medicine, Markey Molecular Medicine Center, University of Washington, Box 357360, Seattle, WA 98195, USA.

TERENCE V. SMITH
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

NANCY M. STANDART
Department of Biochemistry, University of Cambridge, Old Addenbrookes Site, 80 Tennis Court Road, Cambridge CB2 1GA, UK.

FRANCOISE STUTZ
Howard Hughes Medical Institute, Biology Department, Brandeis University, Waltham, MA 02254, USA.

JIE TANG
Howard Hughes Medical Institute, Biology Department, Brandeis University, Waltham, MA 02254, USA.

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

JIWU WANG
Molecular Biology Program, Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089–1340, USA.

MARVIN WICKENS
Department of Biochemistry, 420 Henry Mall, University of Wisconsin, Madison, WI 53706, USA.

JEFFREY WILUSZ
Department of Microbiology and Molecular Genetics, UMDNJ—New Jersey Medical School, Newark, NJ 07103, USA.

BEILIN ZHANG
Department of Biochemistry, 420 Henry Mall, University of Wisconsin, Madison, WI 53706, USA.

Abbreviations

A_{260}	absorbance at 260 nm (or other specified wavelength)
ABG-NHS	*N*-(*p*-azidobenzoyl)-glycine *N*-hydroxysuccinimide ester
acp^3U	3-(3-amino-3-carboxy-propyl)-uridine
AMA	methylamine:concentrated ammonium hydroxide (1:1)
AMV RTase	avian myblastosis virus reverse transcriptase
APAB	*p*-azido phenacyl bromide
APS	ammonium persulfate
BAPTD	3-[3-(bromoacetylamino)phenyl]-3-(trifluoromethyl) diazirine
BIV	bovine immunodeficiency virus
bp	base pair
BSA	bovine serum albumin
CAB-NHS	ε-caproylamidobiotin-*N*-hydroxysuccinimide
CDI	carbodiimide
cDNA	complementary DNA
CMCT	1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-*p*-toluene sulfonate
CPG	controlled pore glass support
c.p.m.	counts per minute
CTAB	cetyltrimethylammonium bromide
Da	Dalton
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCS	downstream control sequence
ddH_2O	double distilled water
ddNTP	dideoxynucleotide triphosphate
DEAE	diethylaminoethyl
DEPC	diethyl pyrocarbonate
DMF	*N,N*-dimethylformamide
DMFa	dimethylformamidine
DMS	dimethyl sulfate
DMSO	dimethyl sulfoxide
DMTr	dimethoxytrityl
DNase	deoxyribonuclease
dNTP	deoxynucleotide triphosphate
dsRNA:DNA	double-stranded RNA or DNA
DTNB	5,5′-dithio-*bis*(nitrobenzoic acid); Ellman's reagent
DTT	dithiothreitol
EDITH	3-ethoxy-1,2,4-dithiazoline-5-one
EDTA	ethylenediaminetetraacetic acid
EF-G	elongation factor-G (bacterial translation factor)

EF-Tu	elongation factor-Tu (bacterial translation factor)
EGTA	ethylene*bis*(oxyethylenenitrilo)tetraacetic acid
EMS	ethylmethanesulfonate
EMSA	electrophoretic mobility shift assay
FES	filter elution solution
FOA	fluoro-orotic acid
Fpmp	1-(2-fluorophenyl)-4-methoxypiperidin-4-yl
GST	glutathione *S*-transferase
h	hour(s)
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	human immunodeficiency virus
hnRNP	heterogeneous nuclear ribonucleoprotein particle
HPLC	high-performance liquid chromatography
HRV	human rhinovirus
IF-1,2,3	initiation factors 1, 2, and 3 (bacterial translation factors)
IgG	immunoglobulin class G
IPTG	isopropyl-β-D-thiogalactopyranoside
IRE	iron-response element
IRP1	iron regulatory protein, iron-response protein
MALDI-TOF	matrix-assisted laser desorption time-of-flight (mass spectrometry)
Mes	4-morpholineethanesulfonic acid
min	minute(s)
MUD	mutant U1 die
NC	nitrocellulose
NMP	nucleotide monophosphate
nt	nucleotide
NTP	nucleotide triphosphate
ONPG	*o*-nitrophenyl-β-D-galactopyranoside
ORF	open reading frame
PAB(P)	poly(A) binding protein
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
pcv	packed cell volume
PEG	polyethylene glycol
PMSF	phenylmethylsulfonyl fluoride
PNK	polynucleotide kinase
PTB	polypyrimidine tract binding protein
PVP	polyvinyl pyrrolidone
RBD	RNA binding domain
RNase	ribonuclease
rNMP	nucleotide monophosphate
RNP	ribonucleoprotein

rNTP	nucleotide triphosphate
RRE	rev-response element
RRM	RNA recognition motif
RT-PCR	reverse transcription followed by polymerase chain reaction
S	Svedberg
SDS	sodium dodecyl sulfate
sec	second(s)
SELEX	systematic evolution of ligands by exponential enrichment
snRNP	small nuclear ribonucleoprotein particle
ssRNA:DNA	single-stranded RNA or DNA
TAR	*trans*-activation response element
TB	Tris, borate buffer
TBDMS	*t*-butyldimethylsilyl group
TBE buffer	Tris, borate, EDTA buffer
TCA	trichloroacetic acid
TDB	4-trifluoromethyl diazirino-benzoyl *N*-hydroxysuccinimide ester
TE	Tris–HCl, EDTA buffer
TEA·3HF	triethylamine trihydrofluoride
TEAA	triethylammonium acetate
TEMED	tetraethylmethylenediamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
U2AF	U2 snRNP auxiliary factor
UV	ultraviolet
YPD/YEPD	yeast extract, peptone, glucose (dextrose)
YPG/YEPG	yeast extract, peptone, glycerol
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
YFG	'your favourite gene'

Contents