

laboratory techniques

in biochemistry and molecular biology

general editors: T. S. WORK and E. WORK

**determination of
sequences in RNA**

G. G. BROWNLEE

DETERMINATION OF SEQUENCES IN RNA

G. G. Brownlee

*Medical Research Council,
Laboratory of Molecular Biology,
Hills Road, Cambridge, England*



1972

NORTH-HOLLAND PUBLISHING COMPANY—AMSTERDAM · LONDON
AMERICAN ELSEVIER PUBLISHING CO., INC.—NEW YORK

© 1972 North-Holland Publishing Company

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

Library of Congress Catalog Card Number: 76-157037

ISBN North-Holland – series 0 7204 4200 1

– part 3.I: 0 7204 4209 5

ISBN American Elsevier: 0 444 10102 0

63 illustrations & graphs, 30 tables

Published by:

NORTH-HOLLAND PUBLISHING COMPANY – AMSTERDAM

Sole distributors for the U.S.A. and Canada:

AMERICAN ELSEVIER PUBLISHING COMPANY, INC.

52 VANDERBILT AVENUE, NEW YORK, N.Y. 10017

This book is the pocket-edition of Volume 3, Part I, of the series 'Laboratory Techniques in Biochemistry and Molecular Biology'.

Volume 3 of the series contains the following parts:

Part I Determination of sequences in RNA, by G.G. Brownlee

Part II Techniques of lipidology: isolation, analysis and identification of lipids,
by M. Kates

Printed in The Netherlands

Contents

<i>Preface</i>	8
<i>List of abbreviations</i>	9
<i>Chapter 1. Introduction</i>	11
1.1. Historical aspects	11
1.2. General approach to sequence determination	14
1.3. Nomenclature.	16
<i>Chapter 2. Sequence of non-radioactive RNA</i>	18
2.1. Introduction	18
2.2. Enzymic digestion	21
2.2.1. Complete digestion by T_1 -ribonuclease	21
2.2.2. Complete digestion by T_1 -ribonuclease and alkaline phosphatase	22
2.2.3. Complete digestion by pancreatic ribonuclease	22
2.2.4. Partial digestion	23
2.2.4.1. Partial T_1 -ribonuclease	23
2.2.4.2. Partial pancreatic ribonuclease	24
2.3. Fractionation of oligonucleotides	25
2.3.1. Complete T_1 -ribonuclease, and combined T_1 -ribonuclease and alkaline phosphatase digests.	25
2.3.2. Complete pancreatic ribonuclease digests	28
2.3.3. Partial enzyme digests	28
2.3.3.1. 'Half' molecules produced by partial T_1 -RNase	28
2.3.3.2. Extensive partial T_1 -RNase digests	30

2.3.3.3. Partial P-RNase digests	34
2.4. Desalting of nucleotide peaks from columns	35
2.4.1. Gel filtration	35
2.4.2. Adsorption onto DEAE-cellulose	36
2.4.3. Dialysis	37
2.5. Yields of isolated oligonucleotides	37
2.6. Sequence in oligonucleotides	38
2.6.1. Sequence methods	42
2.6.1.1. Alkaline hydrolysis	43
2.6.1.2. T_2 -RNase	46
2.6.1.3. Snake venom phosphodiesterase	48
2.6.1.4. Pancreatic RNase digestion of T_1 -oligonucleotides and T_1 -RNase digestion of P-RNase oligonucleotides	48
2.6.1.5. U_2 -RNase digestion of T_1 -end products	49
2.6.1.6. Partial digestion with snake venom phosphodiesterase	49
2.6.1.7. Micrococcal nuclease	51
2.6.1.8. Polynucleotide phosphorylase	51
2.6.1.9. Spectral identification	52

Chapter 3. High-voltage paper electrophoresis 54

3.1. Introduction	54
3.2. High-voltage electrophoresis	54
3.3. The 'hanging' electrophoresis tank	55
3.4. 'Up-and-over' tanks	56
3.5. Power supply and safety precautions	58
3.6. Loading, buffers and operating	64

Chapter 4. A two-dimensional ionophoretic fractionation method for labelled oligonucleotides 67

4.1. Introduction	67
4.2. Enzymatic digestion	69
4.3. Two-dimensional ionophoretic fractionation procedure	70
4.4. Structure of nucleotides	75
4.4.1. Composition	76
4.4.2. Digestion with enzymes	77
4.4.2.1. Pancreatic ribonuclease	77
4.4.2.2. T_1 -RNase	78
4.4.2.3. Micrococcal nuclease	78
4.4.2.4. Venom phosphodiesterase	80
4.4.2.5. Partial digestion with spleen phosphodiesterase	81

4.5. Application of the two-dimensional method for fingerprinting 16S and 23S rRNA	82
4.5.1. T ₁ -RNase digests	83
4.5.2. Pancreatic RNase digests	89
4.6. Results of partial spleen phosphodiesterase digestion of oligonucleotides	89
4.6.1. Partial spleen phosphodiesterase digestion of oligonucleotides derived from complete T ₁ -RNase digestion	90
4.6.2. Nucleotides from pancreatic ribonuclease digests	96
4.7. Conclusion	99

Chapter 5. Sequence of 5S RNA 100

5.1. Introduction	100
5.2. Complete T ₁ -RNase digestion	103
5.2.1. Secondary splits	107
5.2.2. Partial digestion of large oligonucleotides with snake venom phosphodiesterase	109
5.2.3. End groups	117
5.3. Complete pancreatic RNase digestion	118
5.3.1. Partial digestion with spleen phosphodiesterase	120
5.4. Relative molar yields and discussion	122
5.5. 'Homochromatography' and fractionation of partial enzymatic digests	130
5.5.1. A two-dimensional procedure using DEAE-paper	132
5.5.2. A two-dimensional procedure using DEAE-cellulose thin layers	136
5.5.2.1. Preparation of thin layers	136
5.5.2.2. Fractionation procedure	137
5.5.2.3. Homomixtures	140
5.5.2.4. Elution and analysis	141
5.6. Partial enzymatic digestion	142
5.6.1. Extensive digestion with T ₁ -RNase	143
5.6.2. 'Half-molecules' with T ₁ -RNase	148
5.6.3. Pancreatic RNase	150
5.6.4. Spleen acid RNase	155
5.7. Chemical blocking with water-soluble carbodiimide	155
5.8. Chemical blocking by partial methylation	165
5.9. Derivation of the sequence by overlapping	173
5.9.1. 5'-terminal sequence (residues 1-16)	175
5.9.2. 3'-terminal sequence (residues 103-120)	176
5.9.3. Residues 76-105	176
5.9.4. Residues 14-67	177
5.9.5. Residues 65-79	179
5.10. Discussion	180

5.10.1. Sequence variations	180
5.10.2. Secondary structure	181
5.10.3. Sequence homologies	185
<i>Chapter 6. Methods for sequencing large oligonucleotides isolated as end-products of T₁-RNase digestion</i>	<i>187</i>
6.1. Introduction	187
6.2. Isolation of large oligonucleotides	189
6.3. Deduction of the sequence of the oligonucleotide AAUUAACUAUCCAAUUUUCG	192
6.4. Deduction of the sequence of the oligonucleotide AUAUUUCAUACCACAAG	194
6.5. Cleavage next to C's	196
6.6. Cleavage next to A's with RNase U ₂	199
6.7. Partial spleen acid RNase digestion.	200
6.8. Partial pancreatic RNase digestion	201
<i>Chapter 7. Minor bases and tRNA</i>	<i>202</i>
7.1. Introduction	202
7.2. Purification of ³² P-tRNA.	203
7.3. Minor bases	204
7.4. Identification of minor bases in tyrosine suppressor tRNA of <i>E. coli</i>	208
7.5. Complete hydrolysis with T ₂ -RNase	210
7.6. A cross-linkage in <i>E. coli</i> tRNA ₁ ^{Val}	211
<i>Chapter 8. Limited sequence objectives: end-groups of RNA and sequences adjacent to minor bases in ribosomal RNA</i>	<i>213</i>
8.1. Introduction	213
8.2. End groups in RNA	214
8.2.1. 5'-end nucleotide.	214
8.2.2. 5'-end oligonucleotide	216
8.3. 3'-terminal oligonucleotide - a 'diagonal' method	218
8.4. Sequence around methylated bases in 16S and 23S RNA of <i>E. coli</i>	220
<i>Chapter 9. In vitro labelled ³²P-RNA.</i>	<i>225</i>
9.1. Introduction	225
9.2. Polynucleotide phosphokinase labelling.	227
9.3. Synchronised synthesis of RNA using purified replicase	231

<i>Acknowledgements</i>	235
<i>Appendices</i>	237
1 Tentative rules for representation of nucleosides and polynucleotides	237
2 Addresses of suppliers of enzymes, electrophoresis supports and fine chemicals, etc.	239
3 Addresses of suppliers of equipment	240
4 ϵ_{nm} for some nucleotides and nucleosides at pH 7.0 and 260 m μ	241
5 Preparation of radioactively labelled RNA	241
6 Sequence of some oligonucleotides from rRNA	250
7 Summary of conditions of enzyme digestion and of fractionation systems for radioactive sequencing	250
<i>References</i>	256
<i>Subject index</i>	261

Dedicated to my mother

Determination of sequences in RNA

LABORATORY TECHNIQUES IN BIOCHEMISTRY AND MOLECULAR BIOLOGY

Edited by

T. S. WORK - *N.I.M.R., Mill Hill, London*

E. WORK - *Imperial College, London*

Advisory board

G. POPIAK - *U.C.L.A.*

S. BERGSTROM - *Stockholm*

K. BLOCH - *Harvard University*

P. SIEKEVITZ - *Rockefeller University*

E. SMITH - *U.C.L.A.*

E. C. SLATER - *Amsterdam*



NORTH-HOLLAND PUBLISHING COMPANY - AMSTERDAM · LONDON
AMERICAN ELSEVIER PUBLISHING CO., INC. - NEW YORK

Contents

<i>Preface</i>	8
<i>List of abbreviations</i>	9
<i>Chapter 1. Introduction</i>	11
1.1. Historical aspects	11
1.2. General approach to sequence determination	14
1.3. Nomenclature.	16
<i>Chapter 2. Sequence of non-radioactive RNA</i>	18
2.1. Introduction	18
2.2. Enzymic digestion	21
2.2.1. Complete digestion by T_1 -ribonuclease	21
2.2.2. Complete digestion by T_1 -ribonuclease and alkaline phosphatase	22
2.2.3. Complete digestion by pancreatic ribonuclease	22
2.2.4. Partial digestion	23
2.2.4.1. Partial T_1 -ribonuclease	23
2.2.4.2. Partial pancreatic ribonuclease	24
2.3. Fractionation of oligonucleotides	25
2.3.1. Complete T_1 -ribonuclease, and combined T_1 -ribonuclease and alkaline phosphatase digests.	25
2.3.2. Complete pancreatic ribonuclease digests	28
2.3.3. Partial enzyme digests	28
2.3.3.1. 'Half' molecules produced by partial T_1 -RNase	28
2.3.3.2. Extensive partial T_1 -RNase digests	30

2.3.3.3. Partial P-RNase digests	34
2.4. Desalting of nucleotide peaks from columns	35
2.4.1. Gel filtration	35
2.4.2. Adsorption onto DEAE-cellulose	36
2.4.3. Dialysis	37
2.5. Yields of isolated oligonucleotides	37
2.6. Sequence in oligonucleotides	38
2.6.1. Sequence methods	42
2.6.1.1. Alkaline hydrolysis	43
2.6.1.2. T_2 -RNase	46
2.6.1.3. Snake venom phosphodiesterase	48
2.6.1.4. Pancreatic RNase digestion of T_1 -oligonucleotides and T_1 -RNase digestion of P-RNase oligonucleotides	48
2.6.1.5. U_2 -RNase digestion of T_1 -end products	49
2.6.1.6. Partial digestion with snake venom phosphodiesterase	49
2.6.1.7. Micrococcal nuclease	51
2.6.1.8. Polynucleotide phosphorylase	51
2.6.1.9. Spectral identification	52

Chapter 3. High-voltage paper electrophoresis 54

3.1. Introduction	54
3.2. High-voltage electrophoresis	54
3.3. The 'hanging' electrophoresis tank	55
3.4. 'Up-and-over' tanks	56
3.5. Power supply and safety precautions	58
3.6. Loading, buffers and operating	64

Chapter 4. A two-dimensional ionophoretic fractionation method for labelled oligonucleotides 67

4.1. Introduction	67
4.2. Enzymatic digestion	69
4.3. Two-dimensional ionophoretic fractionation procedure	70
4.4. Structure of nucleotides	75
4.4.1. Composition	76
4.4.2. Digestion with enzymes	77
4.4.2.1. Pancreatic ribonuclease	77
4.4.2.2. T_1 -RNase	78
4.4.2.3. Micrococcal nuclease	78
4.4.2.4. Venom phosphodiesterase	80
4.4.2.5. Partial digestion with spleen phosphodiesterase	81

4.5. Application of the two-dimensional method for fingerprinting 16S and 23S rRNA	82
4.5.1. T ₁ -RNase digests	83
4.5.2. Pancreatic RNase digests	89
4.6. Results of partial spleen phosphodiesterase digestion of oligonucleotides	89
4.6.1. Partial spleen phosphodiesterase digestion of oligonucleotides derived from complete T ₁ -RNase digestion	90
4.6.2. Nucleotides from pancreatic ribonuclease digests	96
4.7. Conclusion	99

Chapter 5. Sequence of 5S RNA 100

5.1. Introduction	100
5.2. Complete T ₁ -RNase digestion	103
5.2.1. Secondary splits	107
5.2.2. Partial digestion of large oligonucleotides with snake venom phosphodiesterase	109
5.2.3. End groups	117
5.3. Complete pancreatic RNase digestion	118
5.3.1. Partial digestion with spleen phosphodiesterase	120
5.4. Relative molar yields and discussion	122
5.5. 'Homochromatography' and fractionation of partial enzymatic digests	130
5.5.1. A two-dimensional procedure using DEAE-paper	132
5.5.2. A two-dimensional procedure using DEAE-cellulose thin layers	136
5.5.2.1. Preparation of thin layers	136
5.5.2.2. Fractionation procedure	137
5.5.2.3. Homomixtures	140
5.5.2.4. Elution and analysis	141
5.6. Partial enzymatic digestion	142
5.6.1. Extensive digestion with T ₁ -RNase	143
5.6.2. 'Half-molecules' with T ₁ -RNase	148
5.6.3. Pancreatic RNase	150
5.6.4. Spleen acid RNase	155
5.7. Chemical blocking with water-soluble carbodiimide	155
5.8. Chemical blocking by partial methylation	165
5.9. Derivation of the sequence by overlapping	173
5.9.1. 5'-terminal sequence (residues 1-16)	175
5.9.2. 3'-terminal sequence (residues 103-120)	176
5.9.3. Residues 76-105	176
5.9.4. Residues 14-67	177
5.9.5. Residues 65-79	179
5.10. Discussion	180

5.10.1. Sequence variations	180
5.10.2. Secondary structure	181
5.10.3. Sequence homologies	185
<i>Chapter 6. Methods for sequencing large oligonucleotides isolated as end-products of T₁-RNase digestion</i>	<i>187</i>
6.1. Introduction	187
6.2. Isolation of large oligonucleotides	189
6.3. Deduction of the sequence of the oligonucleotide AAUUAACUAUCCAAUUUUCG	192
6.4. Deduction of the sequence of the oligonucleotide AUAUUUCAUACCACAAG	194
6.5. Cleavage next to C's	196
6.6. Cleavage next to A's with RNase U ₂	199
6.7. Partial spleen acid RNase digestion.	200
6.8. Partial pancreatic RNase digestion	201
<i>Chapter 7. Minor bases and tRNA</i>	<i>202</i>
7.1. Introduction	202
7.2. Purification of ³² P-tRNA.	203
7.3. Minor bases	204
7.4. Identification of minor bases in tyrosine suppressor tRNA of <i>E. coli</i>	208
7.5. Complete hydrolysis with T ₂ -RNase	210
7.6. A cross-linkage in <i>E. coli</i> tRNA ₁ ^{Val}	211
<i>Chapter 8. Limited sequence objectives: end-groups of RNA and sequences adjacent to minor bases in ribosomal RNA</i>	<i>213</i>
8.1. Introduction	213
8.2. End groups in RNA	214
8.2.1. 5'-end nucleotide.	214
8.2.2. 5'-end oligonucleotide	216
8.3. 3'-terminal oligonucleotide - a 'diagonal' method	218
8.4. Sequence around methylated bases in 16S and 23S RNA of <i>E. coli</i>	220
<i>Chapter 9. In vitro labelled ³²P-RNA.</i>	<i>225</i>
9.1. Introduction	225
9.2. Polynucleotide phosphokinase labelling.	227
9.3. Synchronised synthesis of RNA using purified replicase	231

<i>Acknowledgements</i>	235
<i>Appendices</i>	237
1 Tentative rules for representation of nucleosides and polynucleotides	237
2 Addresses of suppliers of enzymes, electrophoresis supports and fine chemicals, etc.	239
3 Addresses of suppliers of equipment	240
4 ϵ_{nm} for some nucleotides and nucleosides at pH 7.0 and 260 m μ	241
5 Preparation of radioactively labelled RNA	241
6 Sequence of some oligonucleotides from rRNA	250
7 Summary of conditions of enzyme digestion and of fractionation systems for radioactive sequencing	250
<i>References</i>	256
<i>Subject index</i>	261

Dedicated to my mother

Preface

The aim of this monograph is to survey the methods that have recently been developed for the determination of the sequence of small RNA molecules. I have attempted to give the necessary experimental details for a reader to reproduce these methods in the laboratory and apply them to his own particular sequence problem. The chapter on electrophoresis equipment and technique will, I hope, be especially useful to those envisaging setting up the equipment for sequence determination. In this book, I have covered both the classical methods and the newer radioactive methods. However, my emphasis is naturally towards the latter as it is with these methods that I am most familiar. The recent spate of reports in scientific journals of the sequence of various transfer RNA molecules by radioactive methods have shown that this is a powerful and rapid approach. However, here I have not concerned myself with discussing the results of sequence work and they are only discussed in so far as is necessary to illustrate the methodology of sequence determination. Nor is any reference made to the methods of sequence determination of DNA, in which there has been less progress than with RNA. At the time of writing the longest known RNA sequence is that of the 6S RNA determined by the radioactive approach. It is certain, however, that larger sequences will be established by using these methods. Perhaps we may conjecture that the time will not be too distant when we may know the entire RNA sequence of the ribosome.

List of abbreviations

(excluding those defined in appendix 1 and table 7.2, p. 205)

tRNA	Transfer RNA.
rRNA	Ribosomal RNA.
mRNA	Messenger RNA.
G*	Unknown derivative of guanosine.
Py	Any pyrimidine nucleoside.
Pu	Any purine nucleoside.
R _u	Chromatographic or electrophoretic mobility of an unknown compound relative to uridine 2'(3')-phosphate = 1.00.
P _i	Orthophosphate.
P-RNase (RNase A)	Pancreatic ribonuclease.
T ₁ -RNase (RNase T ₁)	Takadiastase T ₁ ribonuclease.
T ₂ -RNase (RNase T ₂)	Takadiastase T ₂ ribonuclease.
U ₂ -RNase (RNase U ₂)	Ustilago U ₂ ribonuclease.
DNase	Pancreatic deoxyribonuclease.
mA	Milli-amperes.
C	Curie.
dpm	Radioactive disintegrations per minute.
p.f.u.	Plaque-forming units.
u	Units.
EDTA	Ethylenediaminetetraacetic acid.
Tris	Tris-(hydroxymethyl)-aminomethane.