

# DNA ISOLATION AND SEQUENCING

**Bruce A. Roe, Judy S. Crabtree and Akbar S. Khan**

*Department of Chemistry and Biochemistry, The University of Oklahoma, Norman, OK 73019, USA*

**JOHN WILEY & SONS**

Chichester • New York • Brisbane • Toronto • Singapore

Published in association with BIOS Scientific Publishers Limited

© 1996 John Wiley & Sons Ltd, Baffins Lane, Chichester, West Sussex PO19 1UD, UK, tel (01243) 779777. Published in association with BIOS Scientific Publishers Ltd, 9 Newtec Place, Magdalen Road, Oxford OX4 1RE, UK.

All rights reserved. No part of this book may be reproduced by any means, or transmitted, or translated into a machine language without the written permission of the publisher.

***British Library Cataloguing in Publication Data***

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

0 471 96324 0

***Library of Congress Cataloging in Publication Data***

Roe, Bruce A.

DNA isolation and sequencing / Bruce A. Roe.  
p. cm.—(Essential techniques)

Includes bibliographical references and index.

ISBN 0-471-96324-0 (alk. paper)

I. Nucleotide sequence ———Methodology. 2. DNA— Separation  
tion

—Methodology. I. Title. II. Series: Essential techniques series.  
QP625.N89R64 1996

574.873282— dc20

95-48844

CIP

Typeset by Footnote Graphics Ltd, Warminster, UK

Printed and bound in UK by Biddles Ltd, Guildford, UK

The information contained within this book was obtained by BIOS Scientific Publishers Limited from sources believed to be reliable. However, while every effort has been made to ensure its accuracy, no responsibility for loss or injury occasioned to any person acting or refraining from action as a result of the information contained herein can be accepted by the publishers, authors or editors.

# ABBREVIATIONS

Amp	ampicillin	MTBE	modified Tris-borate EDTA
APS	ammonium persulfate	PCR	polymerase chain reaction
BSA	bovine serum albumin	PEG	polyethylene glycol
c <sup>7</sup> dGTP	7-deaza-dGTP	RF	replicative form
CIAP	calf intestinal alkaline phosphatase	SDS	sodium dodecylsulfate
CPG	controlled-pore glass	SSC	standard saline-citrate
DMF	dimethylformamide	STS	sequence-tagged sites
DMSO	dimethylsulfoxide	TB	Terrific Broth
DMT	dimethyloxytrityl	TEMED	<i>N,N,N',N'</i> -tetramethylenediamine
DTT	dithiothreitol	Tet	tetracycline
EDTA	ethylenediamine tetraacetate	Tris	Tris(hydroxymethyl)aminomethane
EtBr	ethidium bromide	UV	ultraviolet
FSB	frozen storage buffer	X-gal	5-bromo-4-chloro $\beta$ -D-galactopyranoside
IP <sub>2</sub> G	isopropyl $\beta$ -D-thiogalactopyranoside		
KDB	Klenow dilution buffer		

## PREFACE

This manual is a compilation of many of the everyday methods used in the average molecular biology laboratory, with emphasis on the techniques for large-scale DNA sequencing. Various forms of the manual have been in use in our laboratory for several years, and this latest version has been updated to include more detailed DNA sequencing protocols and DNA sequencing automation techniques.

This manual has been written in a protocol format, with little theoretical discussion. For theory and additional information, users of this manual are referred back to the

original literature, or to other textual manuals such as that published by Sambrook, Fritsch and Maniatis (1989) see reference 1, Chapter 1.

The following persons are acknowledged for contributing methods and suggestions during the assembly of this manual: Stephanie Chissoe, Sandy Clifton, Dennis Burian, Rick Wilson, Din-Pow Ma, James Wong, Leslie Johnston-Dow, Elaine Mardis, Zhili Wang, Kala Iyer, Steve Toth, Guozhang Zhang, Hua Qin Pan and other members of the Roe laboratory, both past and present.

*Bruce A. Roe*

# SAFETY

Attention to safety aspects is an integral part of all laboratory procedures and national legislations impose legal requirements on those persons planning or carrying out such procedures. While the authors, editor and publisher believe that the recipes and practical procedures, as set forth in this book, are in accord with current recommendations and practice at the time of publication, they accept no legal responsibility for any errors or omissions, and make no warranty, expressed or implied, with respect to material contained herein. It remains the responsibility of the reader to ensure that the procedures which are followed are carried out in a safe manner and that all necessary safety instructions and national regulations are implemented.

In view of ongoing research, equipment modifications and changes in governmental regulations, the reader is urged to review and evaluate the information provided by the manufacturer, for each reagent, piece of equipment or device, for any changes in the instructions or usage and for added warnings and precautions.

These are freely accessible using:

<http://joule.pcl.ox.ac.uk/MSDS/>.

Other safety information on the Internet can be accessed on:

<gopher://atlas.chem.utah.edu/11/MSDS>

<gopher://ginfo.cs.fit.edu:70/lm/safety/msds>

<http://physchem.ox.ac.uk/MSDS>

<http://www.fisher1.com/Fischer/Alphabetical Index.html>

<http://www.pp.orst.edu>

You are actively encouraged to check these data sheets to confirm our assignments and for more detailed information on individual hazards; however the author, editor and publisher can accept no responsibility for any material contained in these data sheets. Furthermore, you must always follow the precautions outlined on labels and data sheets provided by individual manufacturers.

**Radiation.** The use of radioisotopes is subject to legislation and requires permission in most countries. Furthermore, national guidelines for their use and disposal must be rigorously adhered to. The procedures in protocols that use radioisotopes must only be carried out by individuals

All procedures mentioned within this book must be carried out under conditions of good laboratory practice in accordance with local and national guidelines. Some procedures involve specific hazards, including but not limited to hazards in the following categories:

**Chemical.** A number of the reagents are known to be carcinogenic, mutagenic, toxic, inflammable, highly reactive or otherwise hazardous. Substances known to be hazardous have been marked with the symbol  $\Delta$  in the list of reagents (but not subsequently) for each protocol, or if they appear as alternatives to the main protocol, the *first time* they appear in the notes. The reader should consult the safety notes on these pages before embarking on any of the procedures covered. This is in no way meant to imply that undesignated chemicals are nonhazardous, and all laboratory chemicals should be handled with extreme caution. Information is not available on the possible hazards of many compounds. The criteria we have generally used for denoting a substance with  $\Delta$  is based upon a hazard level of 2 or more (on a scale 0–4) in any of the categories in the Baker Saf-T-Data™ system used in the material safety data sheets (MSDS) held at the University of Oxford, UK.

who have received training in the use of such material using the appropriate facilities, protection and personal monitoring procedures.

**Biological.** Antibodies, sera and cells (particularly, but not exclusively, those of human and nonhuman primate origin) pose a significant biological hazard. All such materials, whatever their origin, may harbor human pathogens and should be handled as potentially infectious material in accordance with local guidelines. Any recombinant DNA work associated with protocols is likely to require permission from the relevant regulatory body and you must consult your local safety officer before embarking upon this work.

**Electrical.** Many of the procedures in this book use electrical equipment. Electrophoresis techniques may present particular hazards of this nature.

**Lasers.** Flow cytometers and certain other types of laboratory equipment contain lasers. Users should ensure they are fully aware of the potential hazards of using such equipment.

# CONTENTS

Abbreviations	viii	5 Elution of DNA fragments from agarose	16
Preface	ix	6 Kinase end-labeling of DNA	17
Safety	x	7 Bacterial cell maintenance and storage	18
		8 Fragment purification on Sephacryl S-500 spin columns	20
<b>I General methods</b>	<b>1</b>		
Methods available	1		
Phenol extraction of DNA samples	1		
Concentration of DNA by ethanol precipitation	1		
Restriction digestion	2		
Agarose gel electrophoresis	2		
Elution of DNA fragments from agarose	4		
Kinase end-labeling of DNA	4		
Bacterial cell maintenance and storage	5		
Fragment purification on Sephacryl S-500 spin columns	5		
Notes on precipitation of nucleic acids	6		
References	1		
<i>Protocols</i>			
1 Phenol extraction of DNA samples	8		
2 Concentration of DNA by ethanol precipitation	10		
3 Restriction digestion	12		
4 Agarose gel electrophoresis	14		
		<b>II Random subclone generation</b>	<b>22</b>
		Methods available	22
		Sonication	22
		Nebulization	23
		Random fragment end-repair, size selection and phosphorylation	24
		DNA ligation	25
		Competent cell preparation	26
		Bacterial cell transformation	27
		References	22
		<i>Protocols</i>	
		9 Sonication	30
		10 Nebulization	32
		11 Random fragment end-repair, size selection and phosphorylation	34

12	DNA ligation	37	23	Klenow catalyzed radiolabeled DNA sequencing	92
13	Competent cell preparation	38	24	Modified T7 catalyzed radiolabeled DNA sequencing	94
14	Bacterial cell transformation	40	25	Radiolabeled sequencing gel preparation, loading and electrophoresis	96
<b>III</b>	<b>Methods for DNA isolation</b>	<b>42</b>	26	<i>Taq</i> polymerase-catalyzed cycle sequencing using fluorescent labeled primers	100
	Methods available	42	27	Fluorescent terminator reactions	102
	Large-scale double-stranded DNA isolation	42	28	Sequenase <sup>TM</sup> catalyzed sequencing with dye-labeled terminators	106
	Miniprep double-stranded DNA isolation	44	29	Fluorescent DNA sequencing	110
	Large-scale M13 RF isolation	46	30	Double-stranded sequencing of cDNA clones containing long poly(A) tails using anchored poly(dT) primers	114
	Single-stranded M13 DNA isolation using phenol	48	31	cDNA sequencing based on PCR and random shotgun cloning	116
	Biomek-automated modified Eperon isolation procedure	48			
	96-Well double-stranded template isolation	50			
	Genomic DNA isolation from blood	51			
	References	42			
	<i>Protocols</i>				
15	Large-scale double-stranded DNA isolation	52	<b>V</b>	<b>Additional methods</b>	<b>118</b>
16	Miniprep double-stranded DNA isolation	57		Methods available	118
17	Large-scale M13 RF isolation	60		Polymerase chain reaction (PCR)	118
18	Single-stranded M13 DNA isolation using phenol	64		Purification of PCR fragments for cloning	119
19	Biomek automated modified Eperon isolation procedure	66		Preparation of <i>Sma</i> I-linearized, dephosphorylated double stranded M13 RF cloning vector	120
20	96-Well double-stranded template isolation	68		Synthesis and purification of oligonucleotides	120
21	Genomic DNA isolation from blood	72			





# I GENERAL METHODS

## Methods available

### *Phenol extraction of DNA samples (see Protocol 1)*

Phenol extraction is a common procedure used to purify a DNA sample [1]. Typically, an equal volume of TE-saturated phenol is added to an aqueous DNA sample in a microcentrifuge tube. The mixture is vortexed vigorously, and then centrifuged to achieve phase separation. The upper, aqueous layer is removed carefully to a new tube, avoiding the phenol interface, and is then subjected to two ether extractions to remove residual phenol. An equal volume of water-saturated ether is added to the tube, the mixture is vortexed and the tube is centrifuged to allow phase separation. The upper, ether layer is removed and discarded, including phenol droplets at the interface. After this extraction is repeated, the DNA is concentrated by ethanol precipitation.

### *Concentration of DNA by ethanol precipitation (see Protocol 2)*

Typically, 2.5–3 volumes of an ethanol/acetate solution is added to the DNA sample in a microcentrifuge tube, which is placed in an ice-water bath for at least 10 min. Frequently, this precipitation is performed by incubation at  $-20^{\circ}\text{C}$  overnight [1]. To recover the precipitated DNA, the tube is centrifuged, the supernatant discarded, and the DNA pellet is rinsed with a more dilute ethanol

## References

1. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
2. Studier, F.W. (1973) *J. Mol. Biol.* **79**:237.
3. Richardson, C.C. (1971) in *Procedures in Nucleic Acids Research*, Vol 2 (G.L. Cantoni and D.R. Davies, eds), p. 815. Harper and Row, New York.
4. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene*, **33**:103.
5. Jernpeth, J., Greener, A., Short, J.M., Viola, J. and Kretz, P.L. (1992) *Strategies*, **5**:81.
6. Woodcock, D.M., Crowther, D.M., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z. and Graham, M.W. (1989) *Nucleic Acids Res.* **17**:3469.
7. Murray, N.E., Brammar, W.J. and Murray, K. (1977) *Mol. Gen. Genet.* **150**:53.
8. Messing, J. (1983) *Methods Enzymol.* **101**:20.
9. Bachman, B.J. (1990) *Microbiol. Rev.* **54**:130.

solution. After a second centrifugation, the supernatant again is discarded, and the DNA pellet is dried in a Savant Speed-Vac.

### ***Restriction digestion*** (see *Protocol 3*)

Restriction enzyme digestions are performed by incubating double-stranded DNA molecules with an appropriate amount of restriction enzyme, in its respective buffer as recommended by the supplier, and at the optimal temperature for the specific enzyme. The optimal sodium chloride concentration in the reaction varies for different enzymes, and a set of three standard buffers containing three concentrations of NaCl are prepared and used when necessary. Typical digestions include a unit of enzyme per microgram of starting DNA, and one enzyme unit usually (depending on the supplier) is defined as the amount of enzyme needed to digest one microgram of double-stranded DNA completely in 1 h at the appropriate temperature. These reactions usually are incubated for 1–3 h, to ensure complete digestion, at the optimal temperature for enzyme activity, typically 37°C.

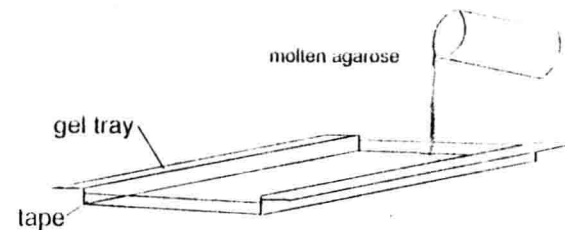
### ***Agarose gel electrophoresis*** (see *Protocol 4*)

Agarose gel electrophoresis [2] is employed to check the progression of a restriction enzyme digestion, to determine quickly the yield and purity of a DNA isolation or polymerase chain reaction (PCR) and to size-fractionate DNA molecules, which then could be eluted from

## **Protocols provided**

1. *Phenol extraction of DNA samples*
2. *Concentration of DNA by ethanol precipitation*
3. *Restriction digestion*
4. *Agarose gel electrophoresis*
5. *Elution of DNA fragments from agarose*
6. *Kinase end labeling of DNA*
7. *Bacterial cell maintenance and storage*
8. *Fragment purification on Sephacryl S-500 spin columns*

the gel. Prior to gel casting, dried agarose is dissolved in buffer by heating, and the warm gel solution then is poured into a mold (made by wrapping clear tape around and extending above the edges of an 18 cm × 18 cm glass plate), which is fitted with a well-forming comb. The percentage of agarose in the gel varies. Although 0.7% agarose gels typically are used, in cases where the accurate size-fractionation of DNA molecules smaller than 1 kb is required, a 1%, 1.5% or 2% agarose gel is prepared, depending on the expected size(s) of the fragment(s). Ethidium bromide is included in the gel matrix to enable fluorescent visualization of the DNA fragments under ultraviolet (UV) light. Agarose gels are submerged in electrophoresis buffer in a horizontal electrophoresis apparatus. The DNA samples are mixed with gel tracking dye and loaded into the sample wells. Electrophoresis is usually at 150–200 mA for 0.5–1 h at room temperature, depending on the desired separation. When low-melting agarose is used for preparative agarose gels, electrophoresis is at 100–120 mA for 0.5–1 h, again depending on the desired separation, and a fan is positioned such that the heat generated is dissipated rapidly. Size markers are co-electrophoresed with DNA samples, when appropriate for fragment size determination. Two size markers are used,  $\pi$ X174 cleaved with restriction endonuclease *Hae*III to identify fragments between 0.3 and 2 kb and  $\lambda$  phage cleaved with restriction endonuclease *Hind*III to identify fragments between 2 and 23 kb. After electrophoresis, the gel is placed on a UV light box



**Figure 1.** Preparation of a mold tray for agarose gels using plastic tape. Figure reproduced from Jones, P., Qui, J. and Rickwood, D. (1994) *RNA Isolation and Analysis* published by BIOS Scientific Publishers, Oxford.

and a picture of the fluorescent ethidium bromide stained DNA separation pattern is taken with a Polaroid camera.

***Elution of DNA fragments from agarose (see Protocol 5)***

DNA fragments are eluted from low-melting temperature agarose gels using an unpublished procedure first developed in our laboratory. Here, the band of interest is excised with a sterile razor blade, placed in a microcentrifuge tube, frozen at  $-70^{\circ}\text{C}$  and then melted. Then, TE-saturated phenol is added to the melted gel slice, and the mixture is again frozen and then thawed. After this second thawing, the tube is centrifuged and the aqueous layer removed to a new tube. Residual phenol is removed with two ether extractions, and the DNA is concentrated by ethanol precipitation.

***Kinase end-labeling of DNA (see Protocol 6)***

Typical 5'-kinase labeling reactions include the DNA to be labeled,  $[\gamma\text{-}^{32}\text{P}]\text{dATP}$ , T4 polynucleotide kinase and buffer [3]. After incubation at  $37^{\circ}\text{C}$ , reactions are heat inactivated by incubation at  $80^{\circ}\text{C}$ . Portions of the reactions are mixed with gel loading dye and loaded into a well of a polyacrylamide gel and electrophoresed. The gel percentage and electrophoresis conditions vary depending on the sizes of the DNA molecules of interest. After electrophoresis, the gel is dried and exposed to X-ray film for radiolabeled DNA sequencing.

### ***Bacterial cell maintenance and storage*** (see Protocol 7)

Three strains of *Escherichia coli* are used in these studies: JM101 for M13 infection and isolation [4], XL1BMR1<sup>+</sup> (Stratagene) for M13 or pUC-based DNA transformation [5] and ED8767 for cosmid DNA transformation [6, 7 and H. Revel, personal communication]. To maintain their respective F' episomes necessary for M13 viral infection [8], JM101 is streaked on to a M9 minimal media (modified from that given in ref. 1) plate and XL1BMR1<sup>+</sup> is streaked on to an LB plate [1] containing tetracycline. ED8767 is streaked on to an LB plate containing ampicillin. These plates are incubated at 37°C overnight. For each strain, 3 ml of appropriate liquid media are inoculated with a smear of several colonies and incubated at 37°C with shaking at 250 r.p.m. for 8 h, and those cultures are then transferred into 50 ml of respective liquid media and incubated for an additional 12–16 h. Glycerol is added to a final concentration of 20%, and the glycerol stock cultures are distributed in 1.3 ml aliquots and frozen at –70 °C until use [1].

### ***Fragment purification on Sephacryl S-500 spin columns***

(see Protocol 8)

DNA fragments larger than a few hundred base pairs can be separated from smaller fragments by chromatography on a size exclusion column such as Sephacryl S-500. To simplify this procedure, a mini-spin column method has been developed.

## Notes on precipitation of nucleic acids

### General rules

Most nucleic acids may be precipitated by addition of monovalent cations and 2–3 vol. of cold 95% ethanol, followed by incubation at 0 to  $-70^{\circ}\text{C}$ . The DNA or RNA may then be pelleted by centrifugation at 10000–13000 r.p.m. for 15 min at  $4^{\circ}\text{C}$ . A subsequent wash with 70% ethanol, followed by brief centrifugation, removes residual salt.

The general procedure for precipitating DNA and RNA is:

- 1 Add 0.1 vol. of 3 M sodium acetate, pH 4.5 to the nucleic acid solution to be precipitated. 5 M ammonium acetate, pH 7.4, NaCl and LiCl may be used as alternatives to sodium acetate. DNA also may be precipitated by addition of 0.6 vol. of isopropanol.
- 2 Add 2 vol. of cold 95% ethanol.
- 3 Place at  $-70^{\circ}\text{C}$  for at least 30 min, or at  $-20^{\circ}\text{C}$  overnight.

or alternatively:

- 1 Combine 95 ml of 100% ethanol with 4 ml of 3 M sodium acetate, pH 4.5 and 1 ml of sterile water. Mix by inversion and store at  $-20^{\circ}\text{C}$ .
- 2 Add 2.5 vol. of cold ethanol/acetate solution to the nucleic acid solution to be precipitated.
- 3 Place at  $-70^{\circ}\text{C}$  for at least 30 min or  $-20^{\circ}\text{C}$  for 2 h to overnight.

### Oligonucleotides

Add 0.1 vol. of 3 M sodium acetate, pH 4.5, and 3 vol. of cold 95% ethanol. Place at  $-70^{\circ}\text{C}$  for at least 1 h.

## **RNA**

Add 0.1 vol. of 1 M sodium acetate, pH 4.5, and 2.5 vol. of cold 95% ethanol. Precipitate large volumes at  $-20^{\circ}\text{C}$  overnight. Small volume samples may be precipitated by placing in powdered dry ice or a dry ice-ethanol bath for 5–10 min.

## **Isobutanol concentration of DNA**

DNA samples may be concentrated by extraction with isobutanol. Add slightly more than one volume of isobutanol, vortex vigorously and centrifuge to separate the phases. Discard the isobutanol (upper) phase, and extract once with water-saturated diethyl ether to remove residual isobutanol. The nucleic acid then may be ethanol precipitated as described previously.

## **Notes on phenol extraction of nucleic acids**

The standard and preferred way to remove proteins from nucleic acid solutions is by extraction with neutralized phenol or phenol:chloroform. Generally, samples are extracted by addition of 1.0 vol. of neutralized (with TE buffer, pH 7.5) phenol to the sample, followed by vigorous mixing for a few seconds to form an emulsion. Following centrifugation for a few minutes, the aqueous (top) phase containing the nucleic acid is recovered and transferred to a clean tube. Residual phenol is then removed by extraction with an equal volume of water-saturated diethyl ether. Following centrifugation to separate the phases, the ether (upper) phase is discarded and the nucleic acid is ethanol precipitated as described previously.

A 1:1 mixture of phenol and chloroform is also useful for the removal of protein from nucleic acid samples. Following extraction with phenol:chloroform, the sample should be extracted once with an equal volume of chloroform, and ethanol precipitated as described above.



# Protocol 1. Phenol extraction of DNA samples

## Reagents

95% Ethanol/0.12 M sodium acetate (ethanol/acetate $\Delta$ )  
TE-saturated phenol $\Delta$ : add an equal amount of 10 mM Tris-HCl, pH 7.5-8.0, 1 mM Na<sub>2</sub>EDTA, pH 8.0 to ultrapure phenol $\Delta$ , mix well, allow phases to separate, remove and discard upper (aqueous) phase. Repeat until the pH of the aqueous phase is between 7.5 and 8.0 (store at 4°C)  
1.0 M Tris-HCl, pH 7.6

1.0 M Tris-HCl, pH 8.0  
Water-saturated diethyl ether

## Equipment

Microcentrifuge  
1.5 ml Microcentrifuge tubes  
Vortex mixer

## Procedure

- 1 Add an equal volume of TE-saturated phenol to the DNA sample contained in a 1.5 ml microcentrifuge tube and vortex for 15–30 sec.
- 2 Centrifuge the emulsion for 5 min at 12 000 r.p.m. at room temperature to separate the phases.
- 3 Remove about 90% of the upper, aqueous layer to a clean tube, carefully avoiding proteins. Add an equal volume of 1:1 TE-saturated phenol:chloroform, centrifuge and remove to a clean tube as above. This additional extraction is not usually necessary if care is taken during the first phenol extraction.
- 4 Add an equal volume of water-saturated ether to the phenol extracted

## Notes

- This procedure will take about 15 min.
- (i) Pure phenol is colorless, a pinky color indicates oxidation and in this case the phenol needs to be distilled before use.