Molecular Genetics

Part III

Chromosome Structure

Edited by

J. HERBERT TAYLOR

Molecular Genetics

Edited by

J. HERBERT TAYLOR
Institute of Molecular Biophysics
Florida State University
Tallahassee, Florida

Part III

Chromosome Structure

1979

ACADEMIC PRESS • NEW YORK SAN FRANCISCO LONDON
A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1 7DX

Library of Congress Cataloging in Publication Data

Taylor, James Herbert, Date ed.
Molecular genetics.

(Molecular biology; an international series of monographs and textbooks, v. 4, pt. 1–

CONTENTS: ––v. 3.
Chromosome structure.
1. Genetics––Collected works. I. Title.
QH426.5.T38 574.8'732 62–22114
ISBN 0–12–684403–8

PRINTED IN THE UNITED STATES OF AMERICA

79 80 81 82 9 8 7 6 5 4 3 2 1

Preface

Twelve years have elapsed since "Molecular Genetics," Part II, was published. Although there was no intention of including another volume at that time, new information has accumulated at such an alarming rate that concepts which were widely held must now be highly modified or even abandoned. The first part was written soon after the structure of DNA proposed by Watson and Crick had been accepted and numerous advances were being made based on the proposed structure and relating to replication, chromosome reproduction, mechanisms of mutation, coding, and the molecular control of transcription and the protein synthetic system. Many of the presentations have withstood the test of time and require limited modification, while others are no longer relevant.

The most recent and rapid advances impinge on the problems of chromosome structure and the functional role of the structural organization of the genetic apparatus. Therefore, in preparing to expand this treatise, the decision was made to consider a number of topics related to the structure and modification of DNA, chromatin, and the higher order organization insofar as treatment at the molecular level was possible.

We begin with the restriction enzymes (site-specific endonucleases), which have been so useful in making many of the analyses possible and reveal interesting and still poorly understood roles that these enzymes may serve in the cells which produce them. Another group of nucleases, which have been less useful to the molecular biologist but seem to be so essential to the replication, organization, and function of DNA, are the topoisomerases (nicking and closing enzymes). Many of the earlier speculations on the unwinding of DNA and the problems posed have been solved in evolution by this interesting group of enzymes, which appear to conserve the bond energy for reversible reactions that were not envisioned a few years ago. The role of methylation of DNA has puzzled and intrigued some of us since it was discovered many years ago. Its role in the modification–restriction systems has revived interest, and it is likely that surprises await us in this area. A major role for methylation in eukaryotes has yet to be discovered even though it is of almost universal occurrence in the higher forms. Insects may present an important exception, but some other type of DNA modification may be substituted in these animals.

Transcription was beginning to be understood, we thought, but now cloning and sequence analysis has changed concepts of posttran-

scriptional modifications of the premessenger RNA that reveal a new dimension in the organization of the whole genome. These developments are so recent that we have not assessed their full impact in this volume. The significance and role in the evolution of the genome will have to wait for later treatment. However, in Chapter IV, Oscar Miller and his associates give us a molecular view of the organization and operation of the genetic apparatus which has both astonished and pleased the chemically oriented molecular biologist as well as the electron microscopist.

The cloning and sequencing of DNA are beyond the scope of our treatment and are not far enough advanced to be covered fully, but Biro and Weissman present a basis for following the new developments and also discuss certain regulatory features of the genetic systems so far examined on the basis of sequence information.

The major changes in concepts of chromatin structure and packaging of DNA that have evolved from studies of nuclease digests and electron micrographs have been traced in the last two chapters. The first, written by Rill, considers the nucleosome and its substructure, with emphasis on histone–DNA interactions and arrangements. The second treats the higher orders of organization and possible subunits of chromosomes based on the knowledge gained from the analysis of the nucleosomes and their components.

Plans for this volume were made with the hope that all chapters could be written simultaneously and that the information in all would be equally up-to-date. Such plans seem never to work in reality and, as in earlier volumes, there was considerable variation in the time the different chapters were finally completed. All authors were given a chance to update their material, but there are limits to how much one can revise manuscript after its initial conception. Thanks are due to those who finished early and had to bear with delays of others, but we trust most of the work is durable enough to withstand the test of time. If the latest references are missing from some chapters it is probably related to this variation in completion time.

We wish to thank all of the contributors and especially those scientists and publishers who generously contributed illustrations, graphs, and other illustrative materials to the volume at the request of the various authors. For expediting the final stages of publication and for managing many of the technical details, the publisher is due much of the credit. We hope you will find the collection timely, informative, and interesting reading.

J. HERBERT TAYLOR

Contents

Chapter I

The Role of Restriction Endonucleases in Molecular Genetics

MARC ZABEAU AND RICHARD J. ROBERTS

Chapter II

DNA Topoisomerases: Enzymes That Catalyze the Concerted Breaking and Rejoining of DNA Backbone Bonds

JAMES C. WANG AND LEROY F. LIU

Chapter III

Enzymatic Methylation of DNA: Patterns and Possible Regulatory Roles

J. HERBERT TAYLOR

Chapter IV

Transcriptional Units in Eukaryotic Chromosomes

ANN L. BEYER, STEVEN L. MCKNIGHT, AND OSCAR L. MILLER, JR.

Chapter V

Recognition and Control Sequences in Nucleic Acids

P. ANDREW BIRO AND SHERMAN M. WEISSMAN

Chapter VI

Nucleosomes: Composition and Substructure

RANDOLPH L. RILL

Chapter VII

Nucleosomes and Higher Levels of Chromosomal Organization

JOHN C. HOZIER

Chapter I

The Role of Restriction Endonucleases in Molecular Genetics

MARC ZABEAU AND
RICHARD J. ROBERTS

I. Introduction

The acronyms echoing through the halls of our academic institutions usually reflect the haute couture of scientific research. Although old friends such as DPN, FAD, PEP still visit occasionally, a new circle has developed, and *Eco*RI, *Sal*I, and *Pst*I have moved into the vocabulary of the molecular biologist. With them have come terms such as agarose gels, ligase, maps, vectors, blotting, molecular cloning, and *restriction endonucleases*. The new acronyms serve to identify this

MOLECULAR GENETICS, PART III

ISBN 0-12-684403-8

latter class of important enzymes, which have revitalized molecular biology and have finally allowed direct access to the mysteries of the eukaryotic chromosome.

The information contained within the genome of every organism provides a precise program upon which its biological processes depend. To comprehend these processes in molecular terms, a detailed analysis of gene organization and structure is essential. Molecular approaches have been hampered by the extreme complexity of the DNA molecules encoding this program. Even a simple bacterial chromosome, such as that of *Escherichia coli*, consists of a single polynucleotide chain containing several million nucleotides. The analysis of such a complex molecule requires that it be dissected into discrete segments, amenable to biochemical analysis. The restriction endonucleases have made this feasible. A new technology is now available to investigate the organization of chromosomes, and to analyze genes at both the functional and structural levels. In addition, restriction enzymes have played a key role in the development of recombinant DNA procedures which permit genes to be isolated and manipulated in a fashion hitherto unimaginable.

The study of restriction enzymes can be traced back to the early 1950s when Luria and his collaborators reported the phenomenon of host-controlled variation (Luria and Human, 1952; Bertani and Weigle, 1953; Luria, 1953). They showed that the ability of bacteriophages to grow on particular strains was dependent upon the specific "modifications" induced by the host in which they had been propagated previously. It was concluded that these bacteria must contain some "specificity systems" able to restrict the host range of phages. Extensive genetic and biochemical studies of these systems in *E. coli* strains led to the identification of the two components involved (Linn and Arber, 1968; Meselson and Yuan, 1968). The first was an endodeoxyribonuclease (restriction endonuclease) which could distinguish between host DNA and foreign DNA. This was made possible by strain-specific modification, accomplished by the second component of the system—a modification enzyme. Usually this is mediated by methylation of specific DNA sequences, within either the host DNA or phage DNA grown on that host, which then prevents their cleavage by the restriction enzyme.

Despite continued interest in the biological role of these enzymes from *E. coli*, work on them has been overshadowed by the events which followed the characterization of a similar enzyme, *Hin*dII, from *Haemophilus influenzae* Rd (Kelly and Smith, 1970; Smith and Wilcox, 1970). For this enzyme, unlike the *E. coli* enzymes, cleaved DNA

at specific sites.[1] Only 6 years have elapsed since the first report appeared exploiting the specificity of *Hin*dII (Danna and Nathans, 1971), and yet the present proliferation of papers is almost overwhelming. More than 140 similar enzymes are now known, and a highly sophisticated technology is being applied to the study of gene structure and function. This chapter will attempt to provide a summary of the general properties of restriction enzymes and to describe their various applications in molecular genetics. Several earlier reviews have appeared (Arber, 1965, 1971, 1974; Arber and Linn, 1969; Boyer, 1971; Meselson *et al.*, 1972; Nathans and Smith, 1975; Roberts, 1976).

II. General Properties of Restriction Enzymes

A. Introduction

The term "restriction endonuclease" was originally used to designate an endodeoxyribonuclease that was involved in a genetically defined process of host-controlled restriction. The first enzymes (*Eco*B and *Eco*K) were isolated from *E. coli* strains B and K (Linn and Arber, 1968; Meselson and Yuan, 1968) and were detected by their ability to selectively degrade, *in vitro*, DNA isolated from bacteriophages susceptible to *in vivo* restriction. Only phage DNAs which lacked the proper strain-specific modification were cleaved, whereas modified phage DNA was resistant to *in vitro* degradation. This property of restriction enzymes to degrade unmodified DNA selectively, but not modified DNA, was subsequently exploited to detect site-specific endonucleases in other bacterial strains (Smith and Wilcox, 1970; Yoshimori, 1971; Gromkova and Goodgal, 1972; Middleton *et al.*, 1972). These new enzymes differed significantly from the *Eco*B and *Eco*K

[1] The nomenclature used throughout this review is detailed in Smith and Nathans (1973). Restriction endonucleases bear a three-letter system name that abbreviates the genus and species of the organism from which they were isolated. Where necessary, a fourth letter is added to designate the strain. Roman numbers following the system name are assigned to differentiate multiple enzymes from the same source. Where only one enzyme has been isolated, the Roman number I is used to avoid later confusion if a second enzyme should be discovered. The prefix endo R (endonuclease R) is omitted to conserve space, and also because in most cases, the endonucleases have not been shown to form part of a genetic restriction-modification system. Examples are *Hin*dII, one of multiple restriction enzymes from *Haemophilus influenzae* serotype d; *Hin*fI, an enzyme from *Haemophilus influenzae* serotype *f*; *Alu*I, an enzyme from *Arthrobacter luteus*; and *Hph*I, an enzyme from *Haemophilus parahaemolyticus* (in this case, *Hpa* was already used for enzymes from *Haemophilus parainfluenzae*).

endonucleases, and it soon became apparent that at least two different types of restriction endonucleases must exist (Boyer, 1971). The principal difference, which was to have far-reaching consequences, lay in the nature of the degradation products, for while the type I enzymes (*Eco*B and *Eco*K) gave a heterogeneous array of products, the type II enzymes gave a specific set of discrete fragments. The development of an agarose gel electrophoresis system to fractionate DNA fragments of different sizes (Aaij and Borst, 1972), gave a simple and rapid assay for the type II enzymes (Sharp *et al.*, 1973). This assay has been used to screen many bacterial strains, and a large number of type II enzymes have been isolated (Roberts, 1976). Activities identified using this gel assay procedure are most properly designated site-specific endonucleases; nevertheless they are often referred to as "restriction enzymes," even though most have not been shown to participate in a restriction–modification system. Throughout this chapter, we will use the terms specific endonuclease, restriction endonuclease, and restriction enzyme interchangeably. After a brief description of the type I enzymes, the rest of this chapter will focus on the type II enzymes, which are rapidly becoming indispensable as the molecular scalpels of the contemporary biologist.

B. TYPE I RESTRICTION ENZYMES

Interest in the type I enzymes, *Eco*B and *Eco*K, has centered around their role in the biological process of host-controlled restriction and modification (Arber and Linn, 1969; Arber, 1974). Both restriction enzymes and their companion modification enzymes have been purified to near homogeneity, and their subunit structure, catalytic properties, and cleavage mechanism have been examined (Eskin and Linn, 1972a; Lautenberger and Linn, 1972; Yuan *et al.*, 1975). Although both *Eco*B and *Eco*K bind to specific sites on the DNA (Arber and Kuhnlein, 1967; Murray *et al.*, 1973b; Horiuchi *et al.*, 1975; Brack *et al.*, 1976b), they show no cleavage specificity (Horiuchi and Zinder, 1972; Murray *et al.*, 1973a). In addition, their endonucleolytic activity requires the cofactors Mg^{2+}, ATP, and *S*-adenosylmethionine (Linn and Arber, 1968; Meselson and Yuan, 1968; Roulland-Dussoix and Boyer, 1969). The purified enzymes exist as complexes composed of three nonidentical subunits (Eskin and Linn, 1972a; Meselson *et al.*, 1972), and this complex has been shown to catalyze both endonucleolytic cleavage (restriction) and methylation (modification) (Haberman *et al.*, 1972; Vovis *et al.*, 1974; Vovis and Zinder, 1975). Following cleavage, the restriction endonuclease is converted into a potent

ATPase (Eskin and Linn, 1972b; Yuan *et al.*, 1972). Enzymes exhibiting properties similar to those described for *Eco*K and *Eco*B have been isolated from *Haemophilus influenzae* strains R_d and R_f (Gromkova *et al.*, 1973; Piekarowicz *et al.*, 1976).

The main difference between the type I and the type II enzymes is that the latter recognize a specific sequence and *cleave* at a *specific* site. They require only Mg^{2+} as a cofactor and have a much simpler subunit structure. In particular, the restriction and modification enzymes exist as separate entities. Two enzymes, *Eco*P1 and *Eco*P15, specified by bacteriophage P1 (Meselson and Yuan, 1968; Haberman, 1974) and plasmid P15 (Reiser and Yuan, 1977) share properties with both type I and type II enzymes, but can be distinguished from both. They may be the progenitors of yet another type of restriction enzyme. Although they possess a subunit structure similar to that of the type I enzymes (Arber, 1974) and can catalyze both cleavage and methylation, they show no absolute requirement for *S*-adenosylmethionine, yet are stimulated by it, and do not catalyze a massive ATP hydrolysis (Haberman, 1974; Reiser and Yuan, 1977). Both enzymes display a cleavage specificity comparable with that of the type II enzymes, although complete digest patterns have not yet been observed (Risser *et al.*, 1974; Reiser and Yuan, 1977).

C. TYPE II RESTRICTION ENZYMES

1. *Detection and Purification Procedures*

The first procedures used to assay restriction enzyme activity were based upon the selective degradation of foreign DNA as opposed to host DNA. Degradation was measured as the loss of biological activity (Takano *et al.*, 1966; Linn and Arber, 1968; Meselson and Yuan, 1968; Gromkova and Goodgal, 1972; Takanami, 1973; Takanami and Kojo, 1973; Bron *et al.*, 1975) or as change in either sedimentation velocity (Meselson and Yuan, 1968; Roulland-Dussoix and Boyer, 1969) or viscosity (Smith and Wilcox, 1970; Middleton *et al.*, 1972). A filter-binding assay has also been described (Reiser and Yuan, 1977). These rather laborious assays, which are still the only ones available to monitor the type I enzymes, have now been superceded by the agarose gel assay (Sharp *et al.*, 1973). This assay takes advantage of the fact that a type II restriction enzyme generates a specific set of fragments upon digestion of a small substrate DNA. When fractionated by agarose gel electrophoresis in the presence of ethidium bromide, these fragments can be visualized directly by their fluorescence upon uv irradiation of the gel. When slab gels with multiple slots are used (Sugden *et al.*,

1975), numerous assays can be performed simultaneously, thus permitting the direct visualization of the results of a chromatographic fractionation. The discrete banding pattern not only shows the elution profile of the enzyme but also reveals the presence of two different restriction enzymes within the same bacterial strain. Moreover, the presence of contaminating nonspecific nucleases in fractions containing a restriction enzyme can be inferred from the sharpness of the bands obtained under various digestion conditions. The simultaneous detection of both the desired specific endonuclease and the undesired nonspecific nucleases illustrates the power of this assay.

Most purification procedures are aimed at quickly obtaining an enzyme preparation which is relatively free of contaminating nonspecific nucleases and the required degree of purity strongly depends upon its particular usage. In comparative restriction enzyme analysis and genome mapping experiments, less pure enzyme preparations suffice, whereas DNA sequencing requires highly purified enzymes devoid of nonspecific contaminants. Only for studies of catalytic properties is homogeneous protein needed or its acquisition attempted. Because restriction enzymes are isolated from widely different bacterial sources each containing a different set of contaminants, it would be naive to suppose that a general purification procedure exists for all restriction enzymes. Nevertheless, many of the schemes reported in the literature (see the reference list in Table I) often represent only minor variations of that used for purifying *Hin*dII (Smith and Wilcox, 1970). The first steps in the isolation usually involve the preparation of a high-speed supernatant of the cell lysate and the removal of nucleic acids, by gel filtration or precipitation with either streptomycin sulfate or polyethylene imine. Further purification is achieved by column chromatography and, as with most enzymes involved in nucleic acid metabolism, phosphocellulose has proved immensely useful. Other ion exchangers, such as DEAE-cellulose, QAE-Sephadex, etc., have been used extensively, and, recently, several more exotic adsorbents have become fashionable. Columns of single-stranded DNA agarose (Schaller *et al.*, 1972) have sometimes given dramatic purification (Sack, 1974; P. A. Myers and R. J. Roberts, unpublished results) as have the hydrophobic matrices provided by the ω-aminoalkyl-Sepharose derivatives (Shaltiel and Er-El, 1973; Gelinas *et al.*, 1977b; Mann *et al.*, 1978). Heparin-agarose has been introduced recently and may have general utility (Bickle *et al.*, 1977). Finally, rapid and specialized procedures have been devised for purifying *Eco*RI (Bingham *et al.*, 1977; Sumegi *et al.*, 1977) and *Bgl*II (Bickle *et al.*, 1977).

One of the key factors responsible for the success in purifying re-

striction enzymes has been their quite remarkable stability. Indeed, many enzymes will continue to digest DNA in a linear fashion for periods in excess of 12 hours. This must reflect both their inherent stability and also the absence of significant amounts of proteases in partially purified enzyme preparations. Since assays of crude cell extracts rarely give distinct fragment patterns, due to the high concentration of nonspecific nucleases, it is difficult to quantitate the amounts of enzyme originally present. For this reason, enzyme yield is usually described in terms of the amount of enzyme finally obtained. Phage λ DNA is a commonly used substrate for monitoring cleavage, and the yield is conveniently expressed in arbitrary units, where one unit is defined as the amount of enzyme necessary to completely digest 1 μg of λ DNA in 1 hour at 37°C. This unit definition must be viewed with caution, because it does not necessarily give an accurate reflection of the total amount of DNA which one might expect to cleave with a given amount of enzyme. Both the degree of purity of the enzyme and the DNA concentration can markedly influence the cleavage efficiency. Although a more rigorous definition of a unit would be desirable, the kinetic parameters necessary to establish an absolute rate are difficult to obtain when homogeneous enzyme preparations are not available. A satisfactory unit has been measured only in the case of *Eco*RI, where one unit is defined as the amount of enzyme that cleaves 1 pmole of phosphodiester bonds per minute (Greene *et al.*, 1975; Modrich and Zabel, 1976).

2. *Characterization*

The key feature which distinguishes one type II enzyme from another lies in the specificity of the double-strand break, so that the most useful characteristic is the nature of the recognition sequence and cleavage site. This contrasts with the usual situation, where enzyme characterization involves detailed studies of kinetic parameters, catalytic properties, and protein structure. Consequently, great efforts have been made to elucidate the nucleotide sequence which they recognize and the positions of cleavage relative to that sequence, while only limited data are available concerning kinetic parameters, etc. A summary of the most recent data for the well-characterized enzymes is presented in Table I, and the partially characterized enzymes are listed in Table II.

An important first step in the characterization of a new restriction enzyme involves a description of the fragment patterns that are obtained upon digestion of various substrate DNAs. Comparison of these fragment patterns with those obtained using enzymes of known speci-

TABLE I

TYPE II RESTRICTION ENDONUCLEASES AND THEIR RECOGNITION SEQUENCES

Enzyme[a]	Recognition sequence[b]	Ends[c]	Microorganism	References[d,e]
*Acc*I	GT$^{\downarrow}$ ($^{C}_{A}$)($^{G}_{T}$)AC	2b 5′-ext.	*Acinetobacter calcoaceticus*	50
*Alu*I	AG$^{\downarrow}$ CT	Flush	*Arthrobacter luteus*	35
*Asu*I	G$^{\downarrow}$ GNCC	3b 5′-ext.	*Anabaena subcylindrica*	21
*Ava*I	C$^{\downarrow}$ PyCGPuG	4b 5′-ext.	*Anabaena variabilis*	22, 29
*Bal*I	TGG$^{\downarrow}$ CCA	Flush	*Brevibacterium albidum*	15
*Bam*HI	G$^{\downarrow}$ GATCC	4b 5′-ext.	*Bacillus amyloliquefaciens*	37, 46
*Bbv*I	GC($^{A}_{T}$)GC	?	*Bacillus brevis*	16
*Bcl*I	T$^{\downarrow}$ GATCA	4b 5′-ext.	*Bacillus caldolyticus*	5, 39
*Bgl*II	A$^{\downarrow}$ GATCT	4b 5′-ext.	*Bacillus globigii*	33, 47, 52
*Eco*RI	G$^{\downarrow}$ AATTC	4b 5′-ext.	*Escherichia coli* RY13	18, 20, 49
*Eco*RII	$^{\downarrow}$CC($^{A}_{T}$)GG	5b 5′-ext.	*Escherichia coli* R245	4, 6, 49
*Fnu*DII	CG$^{\downarrow}$ CG	Flush	*Fusobacterium nucleatum* D	26
*Hae*I	($^{A}_{T}$)GG$^{\downarrow}$ CC($^{T}_{A}$)	Flush	*Haemophilus aegyptius*	31
*Hae*II	PuGCGC$^{\downarrow}$ Py	4b 3′-ext.	*Haemophilus aegyptius*	3, 34, 45
*Hae*III	GG$^{\downarrow}$ CC	Flush	*Haemophilus aegyptius*	7, 27
*Hga*I	5′-GACGC(N)$_5^{\downarrow}$ 3′-CTGCG(N)$_{10\uparrow}$	5b 5′-ext.	*Haemophilus gallinarum*	9, 43, 44
*Hha*I	GCG$^{\downarrow}$ C	2b 3′-ext.	*Haemophilus haemolyticus*	36
*Hind*II	GTPy$^{\downarrow}$ PuAC	Flush	*Haemophilus influenzae* R_d	24, 42
*Hind*III	A$^{\downarrow}$ AGCTT	4b 5′-ext.	*Haemophilus influenzae* R_d	32
*Hin*fI	G$^{\downarrow}$ ANTC	3b 5′-ext.	*Haemophilus influenzae* R_f	23, 28, 30
*Hpa*I	GTT$^{\downarrow}$ AAC	Flush	*Haemophilus parainfluenzae*	13, 19, 40
*Hpa*II	C$^{\downarrow}$ CGG	2b 5′-ext.	*Haemophilus parainfluenzae*	13, 19, 40
*Hph*I	5′-GGTGA(N)$_8^{\downarrow}$ 3′-CCACT(N)$_{7\uparrow}$	1b 3′-ext.	*Haemophilus parahaemolyticus*	25, 28
*Kpn*I	GGTAC$^{\downarrow}$ C	4b 3′-ext.	*Klebsiella pneumoniae*	41, 48
*Mbo*I	$^{\downarrow}$GATC	4b 5′-ext.	*Moraxella bovis*	14
*Mbo*II	5′-GAAGA(N)$_8^{\downarrow}$ 3′-CTTCT(N)$_{7\uparrow}$	1b 3′-ext.	*Moraxella bovis*	10, 11, 14
*Mnl*I	5′-CCTC(N)$_{5-10}^{\downarrow}$ 3′-GGAG(N)	?	*Moraxella nonliquefaciens* ATCC 17953	51

*Pst*I	CTGCA↓G	4b 3′-ext.	*Providencia stuartii* 164	8, 41
*Pvu*II	CAG↓CTG	Flush	*Proteus vulgaris*	16
*Sac*I	GAGCT↓C	4b 3′-ext.	*Streptomyces achromogenes*	2
*Sac*II	CCGC↓GG	2b 3′-ext.	*Streptomyces achromogenes*	2
*Sal*I	G↓TCGAC	4b 5′-ext.	*Streptomyces albus* G	1
*Taq*I	T↓CGA	2b 5′-ext.	*Thermus aquaticus* YTI	38
*Xba*I	T↓CTAGA	4b 5′-ext.	*Xanthomonas badrii*	53
*Xho*I	C↓TCGAG	4b 5′-ext.	*Xanthomonas holcicola*	17
*Xma*I	C↓CCGGG	4b 5′-ext.	*Xanthomonas malvacearum*	12

[a] Restriction enzymes are named in accordance with the proposal of Smith and Nathans (1973).

[b] Recognition sequences are written from 5′ → 3′; one strand only is presented and the cleavage site is indicated by an arrow. For example, G↓GATCC is the abbreviation for $\begin{matrix} 5'\text{-}G\downarrow G\text{-}A\text{-}T\text{-}C\text{-}C\text{-}3' \\ 3'\text{-}C\text{-}C\text{-}T\text{-}A\text{-}G\uparrow G\text{-}5' \end{matrix}$

[c] b, Base; ext., extension.

[d] Where more than one reference is given, the one underlined gives the purification procedure.

[e] Key to references: 1. Arrand *et al.* (1978); 2. J. R. Arrand, P. A. Myers, and R. J. Roberts, unpublished results; 3. B. G. Barrell and P. Slocombe, unpublished results; 4. Bigger *et al.* (1973); 5. A. H. A. Bingham, R. J. Sharp, and A. Atkinson, unpublished results; 6. Boyer *et al.* (1973); 7. Bron and Murray (1975); 8. Brown and Smith (1976); 9. Brown and Smith (1977); 10. Brown *et al.* (1979); 11. Endow (1977); 12. Endow and Roberts (1977); 13. Garfin and Goodman (1974); 14. Gelinas *et al.* (1977a); 15. Gelinas *et al.* (1977b), 16. T. R. Gingeras and R. J. Roberts, unpublished results; 17. Gingeras *et al.* (1978); 18. Greene *et al.* (1974); 19. Gromkova and Goodgal (1972); 20. Hedgpeth *et al.* (1972); 21. S. G. Hughes, T. Bruce, and K. Murray, unpublished results; 22. S. G. Hughes and K. Murray, unpublished results; 23. C. A. Hutchison, III, and B. G. Barrell, unpublished results; 24. Kelly and Smith (1970); 25. Kleid *et al.* (1976); 26. A. Lui, B. C. McBride, and M. Smith, unpublished results; 27. Middleton *et al.* (1972); 28. J. H. Middleton, P. V. Stankus, M. H. Edgell, and C. A. Hutchison, III, unpublished results; 29. Murray *et al.* (1976); 30. K. Murray, A. Morrison, and R. J. Roberts, unpublished results; 31. K. Murray, A. Morrison, H. W. Cooke, and R. J. Roberts, unpublished results; 32. Old *et al.* (1975); 33. Pirrotta (1976); 34. Roberts *et al.* (1975); 35. Roberts *et al.* (1976a); 36. Roberts *et al.* (1976b); 37. Roberts *et al.* (1977); 38. Sato *et al.* (1977); 39. D. Sciaky and R. J. Roberts, unpublished results; 40. Sharp *et al.* (1973); 41. Smith *et al.* (1976); 42. Smith and Wilcox (1970); 43. Takanami (1973); 44. Takanami (1974); 45. Tu *et al.* (1976); 46. Wilson and Young (1975); 47. Wilson and Young (1976); 48. R. Wu and R. J. Roberts, unpublished results; 49. Yoshimori (1971); 50. M. Zabeau and R. J. Roberts, unpublished results; 51. M. Zabeau, R. Greene, P. A. Myers, and R. J. Roberts, unpublished results; 52. B. S. Zain and R. J. Roberts, unpublished results; 53. Zain and Roberts (1977).