Methods in Mol

ENZ
MO
B

Enzymes of Molecular Biology

Methods in Molecular Biology

John M. Walker, SERIES EDITOR

16. **Enzymes of Molecular Biology,** edited by *Michael M. Burrell, 1993*
15. **PCR Protocols:** *Current Methods and Applications,* edited by *Bruce A. White, 1993*
14. **Glycoprotein Analysis in Biomedicine,** edited by *Elizabeth F. Hounsell, 1993*
13. **Protocols in Molecular Neurobiology,** edited by *Alan Longstaff and Patricia Revest, 1992*
12. **Pulsed-Field Gel Electrophoresis:** *Protocols, Methods, and Theories,* edited by *Margit Burmeister and Levy Ulanovsky, 1992*
11. **Practical Protein Chromatography,** edited by *Andrew Kenney and Susan Fowell, 1992*
10. **Immunochemical Protocols,** edited by *Margaret M. Manson, 1992*
9. **Protocols in Human Molecular Genetics,** edited by *Christopher G. Mathew, 1991*
8. **Practical Molecular Virology,** edited by *Mary K. L. Collins, 1991*
7. **Gene Transfer and Expression Protocols,** edited by *Edward J. Murray, 1991*
6. **Plant Cell and Tissue Culture,** edited by *Jeffrey W. Pollard and John M. Walker, 1990*
5. **Animal Cell Culture,** edited by *Jeffrey W. Pollard and John M. Walker, 1990*
4. **New Nucleic Acid Techniques,** edited by *John M. Walker, 1988*
3. **New Protein Techniques,** edited by *John M. Walker, 1988*
2. **Nucleic Acids,** edited by *John M. Walker, 1984*
1. **Proteins,** edited by *John M. Walker, 1984*

Methods in Molecular Biology • 16

Enzymes of Molecular Biology

Edited by

Michael M. Burrell

Advanced Technologies (Cambridge) Limited, Cambridge, UK

Humana Press  **Totowa, New Jersey**

999 Riverview Drive, Suite 208
Totowa, New Jersey 07512

Printed in the United States of America.

Library of Congress Cataloging in Publication Data

Main entry under title:

Methods in molecular biology.

Enzymes of molecular biology / edited by Michael M. Burrell.
p. cm. — (Methods in molecular biology ; 16)
Includes index.
ISBN 0-89603-234-5
1. Enzymology. 2. Molecular biology. I. Burrell, Michael M.
II. Series: Methods in molecular biology (Totowa, NJ) ; 16.
QP601.E567 1993
574.19'25—dc20 92-45040
CIP

Preface

The scientist's understanding of the cell at the molecular level has advanced rapidly over the last twenty years. This improved understanding has led to the development of many new laboratory methods that increasingly allow old problems to be tackled in new ways. Thus the modern scientist cannot specialize in just one field of knowledge, but must be aware of many disciplines.

To aid the process of investigation, the *Methods in Molecular Biology* series has brought together many protocols and has highlighted the useful variations and the pitfalls of the different methods. However, protocols frequently cannot be simply taken from the shelf. Thus the starting sample for a chosen protocol may be unavailable in the correct state or form, or the products of the procedure require a different sort of processing. Therefore the scientist needs more detailed information on the nature and requirements of the enzymes being used. This information, though usually available in the literature, is often widely dispersed and frequently occurs in older volumes of journals; not everyone has comprehensive library facilities available. Also many scientists searching out such information are not trained enzymologists and may be unaware of some of the parameters that are important in a specific enzyme reaction.

The present book, *Enzymes of Molecular Biology*, provides a companion volume to the *Methods in Molecular Biology* series—a reference text designed to minimize the time scientists must spend searching the literature to discover how best to make their reactions work efficiently. The intention is to provide sufficient information for even a nonenzymologist to design an experiment, and we have therefore brought together information about a broad range of enzymes commonly used as tools in molecular biology. Within the constraints of producing a sensibly sized volume, enzymes have been chosen that modify both nucleic acids and proteins. The chapters have been arranged to provide some background information on each selected enzyme and those parameters and properties important in its use. Each chapter starts with a description of both the source and use of the enzyme under discussion, and then

provides details on the size and structure of the protein. This is followed by a discussion of those specific parameters—such as pH, ionic strength, activators, inhibitors, K_m, and substrate concentration—that must be satisfied to achieve an optimized reaction. The chapters then each conclude with exemplary practical procedures and protocols that put the previous discussion in context.

Enzymes of Molecular Biology will be useful to graduates and undergraduates coming to a topic for the first time and to every investigator working in a new system or area. The intention is to provide a good starting point for further inquiry

Michael M. Burrell

Contents

Contributors

JÜRGEN ALVES • *Medizinische Hochschule Hannover, Zentrum Biochemie, Abteilung Biophysikalische Chemie, Hannover, Germany*

MICHAEL M. BURRELL • *Advanced Technologies (Cambridge) Ltd., Cambridge, UK*

DEBORAH A. COOK • *School of Natural Sciences, The University of Hertfordshire, Hatfield, UK* (Present address: *Du Pont, UK, Stevenage, UK*)

JAMES M. D'ALESSIO • *Molecular Biology Research and Development, Bethesda Research Laboratories, Life Technologies, Inc., Gaithersburg, MD*

ROBERT GEIGER • *Medizinische Hochschule Hannover, Zentrum Biochemie, Abteilung Biophysikalische Chemie, Hannover, Germany*

GARY F. GERARD • *Molecular Biology Research and Development, Bethesda Research Laboratories, Life Technologies, Inc., Gaithersburg, MD*

HORACE B. GRAY, JR. • *Department of Biochemical and Biophysical Sciences, University of Houston, TX*

FRANK GROSSE • *Max-Planck-Institut fur Experimentelle Medizin, Abteilung Chemie, Gottingen, Germany* (Present address: *Heinrich Pette Institut fur Experimentelle Virologie, Abteilung Tumorvirologie, Hamburg, Germany*)

DAVID P. HORNBY • *Krebs Institute, Sheffield University, Sheffield, UK*

AXEL LANDGRAF • *Medizinische Hochschule Hannover, Zentrum Biochemie, Abteilung Biophysikalische Chemie, Hannover, Germany*

TAO LU • *Department of Biochemical and Biophysical Sciences, University of Houston, TX*

ANDREAS MANNS • *Max-Planck-Institut fur Experimentelle Medizin, Abteilung Chemie, Gottingen, Germany*

MARTIN J. MAUNDERS • *Advanced Technologies (Cambridge) Ltd., Cambridge, UK*

ALFRED PINGOUD • *Medizinische Hochschule Hannover, Zentrum Biochemie, Abteilung Biophysikalische Chemie, Hannover, Germany*

ANDREW J. SHARP • *Plant Molecular Biology Group, The University of Hertfordshire, Hatfield, UK*

ROBERT J. SLATER • *School of Natural Sciences, The University of Hertfordshire, Hatfield, UK*

PATRICIA J. SWEENEY • *Division of Biological Sciences, The University of Hertfordshire, Hatfield, UK*

JOHN M. WALKER • *Division of Biological Sciences, The University of Hertfordshire, Hatfield, UK*

A. FRED WEIR • *Advanced Technologies (Cambridge) Ltd., Cambridge, UK*

JULIA S. WINDER • *Division of Biological Sciences, The University of Hertfordshire, Hatfield, UK* (Present address: *Cantab Pharmaceuticals, Cambridge, UK*)

HEINER WOLFES • *Medizinische Hochschule Hannover, Zentrum Biochemie, Abteilung Biophysikalische Chemie, Hannover, Germany*

Chapter 1

Nucleases

An Overview

A. Fred Weir

Enzymes able to digest nucleic acids are of course essential to molecular biology, indeed the whole technology was founded on the discovery of bacterial enzymes that cleave DNA molecules in a base-specific manner. These enzymes, the type II restriction endonucleases, are perhaps the best studied of the nucleases as to both their in vivo role and their use as tools in the techniques of molecular biology. However, the nucleases are ubiquitous in living organisms and function in all situations where partial or complete digestion of nucleic acid is required. These situations not only include degradation and senescence but also replication and recombination, although it must be noted that, to date, evidence for the involvement of nucleases in the latter two processes in eukaryotes is largely circumstantial. The significance of nucleases in the functioning of nucleic acids as the genetic material can be gaged however by considering that several enzymes implicated in DNA replication, recombination, and repair have integral exo- or endodeoxyribonuclease activity. For example, the 5'–3' and 3'–5' exonuclease activity of DNA polymerases and the endo-DNase activity of topoisomerases (e.g., *see* ref. *1*).

As well as the restriction endonucleases, various other nuclease enzymes have been used as tools in molecular biology, the purpose of this chapter is to give some background on the main deoxyribonu-

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

cleases (DNases) and then to focus on the techniques in which they are used. The enzymes that molecular biologists use as tools are dealt with in separate chapters in this volume.

1.1. Nomenclature

Anyone who has tried the isolation of a DNase enzyme will know that the presence of multiple types of nuclease activity makes this process fraught with difficulty. In this section, consideration will be given to the properties of the DNase enzymes with a view to understanding their nomenclature, which for the most part is somewhat confusing (Table 1).

Nucleases, although a large group in themselves, are part of a larger group of enzymes, the phosphodiesterases, which are able to catalyze the cleavage of phosphate-ester bonds. Schmidt and Laskowsi *(2)* identified three types of nuclease enzymatic activity: DNases, ribonucleases (RNases), and exonucleases. On this definition, it is apparent that so-called DNases and RNases cleave their substrates endonucleolytically, i.e., at internal sites, and that this activity is distinct and separable from any exonuclease activity. In practical terms, this definition holds true in that an endo-DNase will not digest DNA molecules to completion, i.e., to nucleotide monomers; only when exonuclease activity is present will the digestion of DNA go to completion.

A second confusing element in the nomenclature of nucleases, and DNases in particular, is the presence of single-stranded DNases, e.g., mung bean nuclease and nuclease S1 from *Aspergillus*. These enzymes, although having high specificity for single-stranded DNA molecules, will, at high concentrations and in preparations not purified to homogeneity, digest native (double-stranded) DNA molecules albeit at reduced rates. For an example of this, Weir and Bryant *(3)* have isolated a nuclear-located DNase from the embryo axes of pea that has a low, but measurable activity on native DNA but rapidly catalyzes the hydrolysis of heat-denatured DNA. It is not known so far whether these two activities are separable, but evidence from similar enzymes suggests that these activities are part of the same protein molecule. DNases then, tend to be classified as to "what they do best"; e.g., the DNase of Weir and Bryant would be called a single-strand specific endo-DNase. In the following discussion, the examples are from the DNase class of nucleases, however all the points considered can be equally applied to the RNases.

Table 1
The Nomenclature of Nuclease

EC number	Reaction	Example
3.1.11	Exodeoxyribonucleases producing 5'-phosphomonoesters.	Exonuclease III EC 3.1.11.2
3.1.13	Exoribonucleases producing 5'-phosphomonoesters.	Exo-RNase H EC 3.1.13.2
3.1.14	Exoribonucleases producing other than 5'-phosphomonoesters.	Yeast RNase EC 3.1.14.1
3.1.15	Exonucleases active with either ribo- or deoxyribonucleic acids and producing 5'-phosphomonoesters.	Venom exonuclease EC 3.1.15.1
3.1.16	Exonucleases active with either ribo- deoxyribonucleic acids and producing other than 5'-phosphomonoesters.	Spleen exonuclease EC 3.1.16.1
3.1.21	Endodeoxyribonucleases producing 5'-phosphomonoesters.	DNase I EC 3.1.21.1 Type II restriction DNases EC 3.1.21.4
3.1.22	Endodeoxyribonucleases producing other than 5'-phosphomonoesters.	DNase II EC 3.1.22.1
3.1.25	Site-specific endodeoxyribonucleases: specific for altered bases.	
3.1.26	Endoribonucleases producing 5'-phosphomonoesters.	RNase H EC 3.1.26.4
3.1.27	Endoribonucleases producing other than 5'-phosphomonoesters.	RNase T1 EC 3.1.27.3
3.1.30	Endonucleases active with either ribo- or deoxyribonucleic acids and producing 5'-phosphomonoesters.	*Aspergillus* nuclease S1 and Mung bean nuclease EC 3.1.30.1
3.1.31	Endonucleases active with either ribo- or deoxyribonucleic acids and producing other than 5'-phosphomonoesters.	Micrococcal nuclease EC 3.1.31.1

1.1.1. Criteria Used for Classification

1.1.1.1. Exo- vs Endonucleolytic Cleavage

Exo-DNases cleave from the ends of DNA molecules releasing phosphomononucleotides. Cleavage can be either in the 3' to 5' direction releasing 5' phosphomononucleotides or in the 5' to 3' direction to yield 3' phosphomononucleotides. An example of a widely used exonuclease is exonuclease III from *Escherichia coli* (EC 3.1.11.2), which will digest one strand of a double-stranded DNA molecule from a 3' overhang or blunt end. This property has been used to produce bidirectional or unidirectional nested deletion of templates for sequencing.

Endo-DNases cleave at internal phosphate bonds. Cleavage of double-stranded DNA substrates can be by a "single-hit" or a "double-hit" mechanism *(4)* or by a combination of both (*see* Chapter 2, Section 2.4.). Essentially this means that the enzymes can either cleave the two strands of the DNA molecule at points opposite or at sites on the two strands that are well away from each other. The scission of the molecule will take place at a relatively faster rate in the former case as compared to the latter. The prime example of an endo-DNase is pancreatic DNase (DNase I, EC 3.1.21.1). Under optimal conditions this enzyme uses a double-hit mechanism for cleavage of substrates.

1.1.1.2. Base Specificity at or near the Site of Cleavage

None of the eukaryotic enzymes so far isolated appear to have such specificity, however there is evidence that enzymes with optimal activity on single-stranded DNAs will preferentially cleave at A-T rich sites in native DNA molecules *(3,5,6)*. As already mentioned, the Type II restriction endonucleases have absolute specificity for a group of bases at or near the cleavage site.

1.1.1.3. Site of Cleavage

The site of cleavage can be on either side of the phosphate bond leading to a 5' or a 3' monoesterified product. No enzymes have been isolated that can split the internucleotide bond on either side. This property is particularly important if the DNA molecule is to be subsequently made blunt-ended for a ligation experiment. DNA molecules left with a 5' overhang from a staggered cut are usually filled in with the Klenow fragment of *E. coli* DNA polymerase, whereas those with a 3' overhang have the overhang cleaved back with the exonuclease activity of T4 DNA polymerase to create blunt ends.

1.1.1.4. Cleavage of Native or Single-Stranded DNA

Nucleases tend to have a "preference" for cleavage of single-stranded DNA or double-stranded DNA substrates. In general single-strand specific DNases, such as nuclease SI from *Aspergillus* (EC 3.1.30.1), will digest native DNA if the enzyme concentration is high. Nuclease Sl is used to analyze the structure of DNA-RNA hybrids and in cDNA synthesis where it opens the hairpin loop generated during the synthesis.

1.1.1.5. General

DNases have also in the past been classified according to their pH optima and requirements for the presence of metal ions; there are two major mammalian DNases—one working at neutral pH the other in acidic conditions—and both require Mg ions. Other nuclease enzymes exist that require Ca^{2+} ions (pea nuclear DNase *[3]*) or Zn^{2+} ions (nuclease S1). In addition to the aforementioned properties, it is important to note that many crude preparations of DNase exhibit nonspecificity for the sugar moiety of nucleic acids, i.e., they will cleave both RNA and DNA. In these cases it is obviously essential to remove contaminating RNase activity before the enzyme is used to remove DNA from a preparation of RNA.

From the foregoing discussion, it can be seen that the nucleases are a complex group of enzymes. However, when it comes to their use as tools in molecular biology, the situation is very much simplified as only a handful of enzymes are used routinely in experimental protocols. The enzymes DNases I and II, exonucleases, nuclease S1, *Bal*31, and RNase will be discussed in much more detail in the following sections with special emphasis on the experimental protocols for their use in molecular biology techniques.

References

1. Roca, J. and Mezquita, C. (1989) DNA topoisomerase II activity in nonreplicating, transcriptionally inactive, chicken late spermatids. *EMBO J.* **8,** 1855–1860.
2. Schmidt, G. and Laskowski, Sr., M. (1969) Phosphate ester cleavage, in *Enzymes*, vol. 5b (Walker, P. D., Lardy, H., and Myrback, A., eds.), Academic, New York, pp. 3–35.
3. Weir, A. F. and Bryant, J. A. (1989) Partial purification and properties of a chromatin-bound deoxyribonuclease from the embryo axes of germinating pea. *Phytochemistry* **28,** 1339–1343.
4. Bernardi, G. and Sadron, C. (1964) Studies on acid deoxyribonuclease. I. Kinetics of the initial degradation of DNA by acid DNase. *Biochemistry* **3,** 1411–1418.

5. Felsenfeld, G. (1978) Chromatin. *Nature* **271,** 115–121.
6. Kroeker, W. D. and Fairley, J. L. (1975) Specific limited cleavage of bihelical deoxyribonucleic acid by wheat seedling nuclease. *J. Biol. Chem.* **250,** 3773–3778.

CHAPTER 2

Deoxyribonuclease I (EC 3.1.21.1) and II (EC 3.1.22.1)

A. Fred Weir

1. Introduction

From Chapter 1 on nucleases, we know that the term DNase refers to an enzyme that endonucleolytically cleaves DNA molecules. This chapter deals with those DNases that preferentially catalyze the hydrolysis of double-stranded DNA (ds DNA) and that have found a use in the various techniques employed in molecular biology. The Type II restriction endonucleases obviously fall into this category, however they are a large subject on their own and therefore are dealt with in a separate section of this book. DNases acting preferentially on single-stranded DNA (ss DNA) substrates are also dealt with in another section of this chapter.

Despite the ubiquitous nature of nucleases in living organisms, it is perhaps surprising to find that, apart from the Type II restriction endonucleases, only one ds DNA DNase enzyme is used routinely in molecular biology. The reasons for this are twofold. First, as outlined in Chapter 1, it is very difficult to isolate a DNase in a form that is completely free of accompanying RNase and exonuclease contamination and indeed these activities may be part of the same enzyme molecule. Obviously, it would not be desirable to remove DNA from a precious RNA preparation with a DNase that has an integral RNase activity. The second reason is that the first DNase isolated from tissues, DNase I (or pancreatic DNase), performs all the tasks very well and can be isolated readily in a pure form; why use or look for another enzyme? DNase I

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

has been commercially available for at least 25 years, most of the work on the properties of the enzyme being done in the 1960s and, to date, no DNase has been found to replace it in the dual tasks of complete degradation of "nuisance material" DNA and the partial hydrolysis of DNA molecule in such techniques as nick translation.

In the following sections, the properties of DNase I are discussed in detail with special emphasis directed toward comparing the activity of DNase I with that of other types of DNase, most notably DNase II.

2. DNase I (EC 3.1.21.1)

2.1. Reaction

Bovine pancreatic DNase, or more usually, DNase I, catalyzes the hydrolysis of ds DNA molecules but, at high concentrations of enzyme, ss DNA will also be digested. "Complete" hydrolysis results in the formation of small oligonucleotide products that are resistant to further cleavage, but are acid soluble (*see* section on assay of DNase activity); cleavage results in the formation of 5' monoesterified products. Cleavage of DNA substrate with DNase II (spleen DNase or acid DNase, EC 3.1.22.1), in contrast, results in the formation of 3' monoesterified products.

2.2. pH Optimum

Possibly the major reason that DNase I is preferred to DNase II is that DNase I has optimum activity in the region of pH 7–8, whereas DNase II, as its alternative name describes, has a pH optimum in acidic condition: pH 4.2–5.5.

2.3. Activators and Inhibitors

DNase I has an absolute requirement for divalent metal cations. The most commonly used is Mg^{2+}, however Mn^{2+}, Ca^{2+}, Co^{2+}, and Zn^{2+} will also activate DNase I. Concentrations of Mg^{2+} above approx 50 m*M* become inhibitory, which is not the case for Co^{2+} and Mn^{2+}. Monovalent metal ions are also inhibitors of the enzyme activity. In the presence of Ca^{2+}, Mg^{2+} has a synergistic effect, i.e., the rate of hydrolysis of DNA in the presence of both ions is more than the sum of the rates of hydrolysis of DNA in the presence of each ion separately. A total of 0.1 m*M* Ca^{2+} is sufficient to give this enhanced reactivity in the presence of 10 m*M* Mg^{2+}, however the rate of hydrolysis of DNA is still greatest in the presence of Mn^{2+} ions (*see* Section 2.4.).

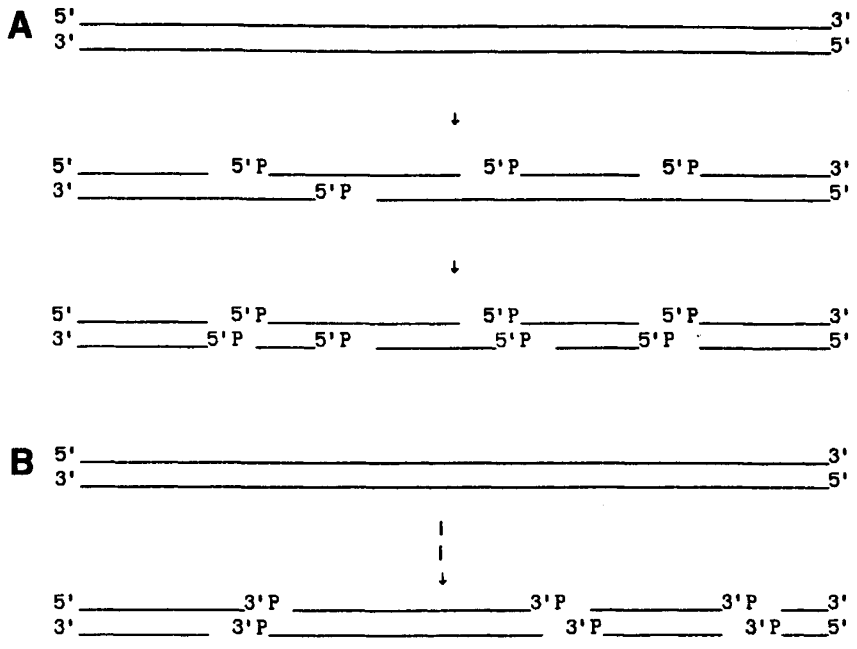


Fig. 1. **A.** Double-hit mechanism of DNase I in presence of Mg^{2+} ions. **B.** Single-hit mechanism of DNase II.

Apart from the monovalent cations, there is no general inhibitor of DNase I such as those available for the inhibition of RNases. The only real way to combat DNase activity, which may be a worry during DNA extractions, is to do the extraction as quickly as possible and at low temperatures. Inclusion of EDTA in the extraction buffer is a good idea, but note that Jones and Boffey *(1)* recently discovered a DNase in the leaves of wheat seedlings that appears to be stimulated in the presence of EDTA.

2.4. Kinetics

Using a variety of methods, including light-scattering, viscometry, and sedimentation analysis, it can be shown that there are two different types of mechanism for the cleavage of ds DNA substrates by DNase (Fig. 1). DNase I, at low concentrations and under the usual assay conditions, inserts nicks at random points in each strand of the DNA at points away from each other. This is termed a "double-hit" mechanism (Fig. 1A); complete scission of the molecule will not occur until two

nicks are opposite. Monitoring of the reaction therefore will not indicate the presence of DNA molecules of intermediate size until after a lag phase *(2a)*. In contrast, DNase II was shown to cleave high-mol-wt DNA substrates on both strands at points opposite to each other resulting in the complete scission of the molecule *(3)*. This process was termed a single-hit mechanism (Fig. 1B), as scission of the DNA occurred from a single encounter with the enzyme molecule; during a digestion, intermediate size molecules will appear immediately. An interesting feature of DNase I is that the cleavage mechanism can be altered from double-hit to a DNase II-like single-hit mechanism by using high concentrations of the enzyme or by altering the divalent cation from Mg^{2+} to Mn^{2+} or Co^{2+}. Using ss DNA as substrate, Melgar and Goldthwaite *(2b)* showed that in the presence of Mn^{2+} ions, DNase I had vastly increased V_{max} as compared to the activity in the presence of Mg^{2+}, whereas there was little change in the K_m. The increased rate of hydrolysis in the presence of Mn^{2+} could of itself lead to the formation of intermediate size fragments in the short periods observed without there being a change in the actual mechanism of cleavage of a double-stranded substrate; the greater the number of random nicks, the greater the likelihood of there being scission of the molecules. If viscosity measurements are used to follow the progress of a DNase digestion, a relationship can be obtained when the log of a function of the change in viscosity of the DNA solution is plotted against log time. The slope of the resulting line, n, gives an indication of the mechanism of the reaction: A value of approx 1.0 indicates single-hit kinetics, whereas a value between 1 and 2 indicates a predominantly double-hit mechanism. When the hydrolysis of ds DNA by DNase I in the presence of Mn^{2+} is monitored by viscometry at low temperature, i.e., at low rate of hydrolysis, a value of $n = 1.16$ is obtained, indicating that the reaction is predominantly of the single-hit kind. Monovalent cations also lower the rate of hydrolysis of ds DNA by DNase I in the presence of Mn^{2+}, and under these conditions the n value changes from approx 1.0 (single-hit mechanism) to values approaching 2.0 (double-hit mechanism). The inhibition of DNase I by monovalent cations is probably a result of competition for effector sites that directly or indirectly affect the active site of the enzyme.

The process by which Mn^{2+} promotes the hydrolysis of ds DNA by DNase I to switch from a double-hit mechanism to a single-hit mechanism is not known but may involve either (1) the ability of Mn^{2+} to

promote the binding of two DNase I molecules at opposite sites on the two DNA strands (it is known that active DNase II is dimeric); or (2) the ability of Mn^{2+} to enable the DNase I enzyme molecule to flip from the hydrolyzed strand to the opposite strand, which is then also cleaved *(2b)*.

2.5. Specificity

There is evidence that certain deoxynucleotide sequences can be hydrolyzed in the presence of Mn^{2+} but not in the presence of Mg^{2+}. For example, poly (dG:dC) can be hydrolyzed by DNase I with Mn^{2+} but not with Mg^{2+}*(4)*. This resistance is not attributable to secondary structure of the double-helix, as DNase I in the presence of Mg^{2+} is not able to hydrolyze the dC strand of a polymer consisting of dI:dC. Therefore, in addition to altering the kinetics of the reaction, different metal ions can also place qualitative constraints on hydrolysis of certain residues.

3. Experimental Procedures

3.1. Introduction

DNase I is the enzyme of choice for all molecular biology techniques (apart from restriction digestions) that require a double-stranded DNase enzyme. It is readily available in pure form, i.e., exonuclease and RNase free, from a number of commercial sources, has high activity on ds DNA, and at high concentrations, will also cleave ss DNA. The reaction does not go to completion, to produce deoxynucleotide monomers, but a certain limit size is reached, after which no further digestion takes place. DNase I cleavage leads to the formation of 5' monoesterified products and the enzyme has optimum activity in the pH 7.0 region. DNase I is inhibited by monovalent cation but has an absolute requirement for divalent metal ions. The nature of the cation affects the mode of action of the enzyme and its ability to cleave at certain deoxynucleotide residues. The enzyme has optimal activity in the presence of Mn^{2+} cations, but it is noteworthy that molecular biologists use Mg^{2+} as the cofactor. This is probably attributable to the fact that many of the enzymes used in conjunction with DNase also require Mg^{2+} as cofactor, e.g., DNA polymerase I in the nick-translation protocol, and that many of the techniques require only limited digestion of the DNA.

3.2. Assay of DNase Activity

Many methods have been used in the past to assay DNase activity: Melgar and Goldthwaite *(2a)* used a somewhat complicated method whereby T4 DNA was labeled with ^{32}P and embedded in a polyacrylamide gel. After exposure of the gel to DNase, the gel was pelleted and the amount of radioactive DNA fragments released into the supernatant was measured. More recent methods (e.g., ref. *5*) have again used radiolabeled substrates, but this time the rate of hydrolysis is monitored by precipitating the DNA fragments produced out of aqueous solution with trichloroacetic acid (TCA) in the presence of a carrier and then measuring acid-soluble or acid-insoluble radioactivity. For most molecular biologists however, all that is required is an assay method to show whether DNase is responsible for the loss of DNA from a preparation, and if so, to show where the contamination is coming from. This is best accomplished by using a covalently closed circular (ccc) plasmid as substrate in a component analysis. In the following example, the components of a restriction enzyme digestion of a plasmid are to be analyzed.

1. Set up several reactions depending on the number of components to be tested, remember that the controls are all important.
   a. ccc plasmid in sterile distilled water (500 ng in 10 µL)
   b. ccc plasmid in buffer
   c. ccc plasmid in buffer + bovine serum albumin (BSA)
   d. ccc plasmid in buffer + enzyme
   e. ccc plasmid in buffer + BSA + enzyme
2. Incubate at 37°C for 30–60 min.
3. Run digests on an agarose gel.
4. Stain with ethidium bromide and view under UV illumination.

Endo-DNase activity is indicated by excessive conversion, as compared to the controls, of ccc plasmid to open circles or to linears. Usually endo-DNase contamination will be accompanied by exonuclease activity, so there will be complete degradation of the plasmid to a smear of material down the gel.

3.3. Removal of DNase Activity

DNase enzymes are heat labile and can be removed from solutions by autoclaving and from glassware and the like by baking or autoclaving. To make sure that solutions of RNase I are DNase free, boil the

RNase solution for 15 min (RNase I is a remarkably stable enzyme) and cool to room temperature. Dispense into aliquots and store frozen.

As already mentioned (Section 2.2.), during the preparation of DNA, endogenous DNases can be inhibited by the inclusion of EDTA in the homogenization buffer. It was also pointed out that a DNase activity had recently been discovered that was apparently stimulated by EDTA *(1)*. The prevalence of this type of nuclease activity is not known, but if it is suspected, damage can be limited by performing extractions at low temperatures and proceeding to the protein removal step as rapidly as possible.

3.4. Removal of RNase from DNase Solutions

RNase-free preparations of DNase I are now commercially available, however some of these still may not be entirely satisfactory. Various column chromatography methods have been used to remove RNase activity from DNase solutions (e.g., ref. *6*) and an effective method for disabling the RNase A enzyme has been described by Zimmerman and Sadeen *(7)* and modified by Gurney and Gurney *(8)*. This method depends on the alkylation of a histidine in the active site of RNase A (a likely contaminant of DNase I as they are derived from the same tissue) using sodium iodoacetate. A brief protocol is outlined in the following:

1. Use autoclaved micropipet tips and glassware.
2. Dissolve 10 mg commercial DNase in 2 mL of 2.5 m*M* HCl.
3. Dialyze the solution for a few hours at 40°C against 1 L of 2.5 m*M* HCl with stirring, then overnight against 1 L of fresh 2.5 m*M* HCl again at 4°C with stirring.
4. Freshly prepare 1*M* sodium iodoacetate. Mix 2.5 mL of 0.2*M* sodium acetate, pH 5.3, 2.0 mL of dialyzed enzyme, and 7.5 mL of 1*M* sodium iodoacetate. Incubate the mixture at 55°C for 60 min. A precipitate will form.
5. Dialyze the solution overnight against 1 L of 2.5 m*M* HCl at 4°C.
6. Centrifuge solution at 10,000*g* for 30 min at 0°C.
7. Carefully pipet off the supernatant into a sterilized tube or bottle with a tight fitting cap.
8. The protein concentration should be 3–4 mg/mL. The DNase should be stored at 4°C, the activity should be stable (unfrozen) for at least 2 yr.

RNases can also be inhibited by a number of compounds including vanadyl ribonuclease complex, heparin, and their own specific ribo-

nuclease inhibitors. RNases are dealt with in much more detail elsewhere in this volume (Chapter 13).

3.5. "Degradative" Uses of DNase I

Stock solutions of DNase I should be made up at 2 mg/mL in 10 m*M* Tris-HCl, pH 7.5, 50 m*M* NaCl, 10 m*M* $MgCl_2$, 1 m*M* dithiothreitol, and 50% glycerol, and stored frozen in small aliquots. Removal of DNA from preparations can be accomplished by incubation with 20–50 μg/mL DNase I in the presence of 50 m*M* Tris-HCl, pH 7.5, 10 m*M* $MgCl_2$, at 37°C for 60 min.

A method for preparing random oligonucleotide fragments from calf thymus DNA for use in the oligolabeling protocol can be found in *(9)*.

3.6. Nick Translation

The alternative method for producing radiolabeled probes other than oligolabeling is nick translation. In this method DNase I is used to nick the DNA to be labeled creating free 3'OH groups. *Escherichia coli* DNA polymerase I catalyzes the addition of nucleotides to the 3'OH termini and the 5'–3' exonuclease activity of the same enzyme removes nucleotides from the 5' end thus translating the nick. Linear, supercoiled, nicked, or gapped circular ds DNA molecules can be labeled to specific activities of $>10^8$ cpm/μg with ^{32}P labeled deoxynucleotide triphosphates. Good protocols for the method are given in refs. *10* and *11* and Chapter 3. Meinkoth and Wahl *(10)* pointed out that, at saturating levels of nucleotide triphosphates, the size of the labeled fragments is determined by the DNase concentration: Fragments of 500–1500 nucleotides long produce optimal signal:noise ratios in the subsequent hybridization. The necessary concentration of DNase required can be determined by doing a titration using from 0–100 pg/mL DNase in the reactions and running the labeled DNA on acrylamide gels *(9)*. DNase concentration, which gives approx 30% incorporation of label, is optimal.

3.7. DNA Footprinting by DNase I Protection

DNase I has been used in the past to dissect the structure of chromatin and to isolate nucleosomes (e.g., *see 12,13*), but in recent years molecular biologists have become interested in the proteins that bind to DNA and possibly regulate the expression of genes. The isolation and characterization of such transactivating factors can be accomplished by using the powerful technique of DNA footprinting. In this

technique the DNA of interest is labeled at either the 3' or 5' end and protein is bound to the DNA; the nucleoprotein complex is then subject to limited digestion with DNase I. Under the conditions used, DNase will exhibit double-hit kinetics and, therefore, will insert random single-strand nicks, however, cleavage will be inhibited if protein is already bound to the DNA. The DNA is then purified, denatured, and run on a sequencing polyacrylamide gel. Autoradiography of the labeled DNA in the gel should show a ladder of fragments representing cleavage at approximately every base for the control with no protein present. However, if protein has been bound to a particular sequence, there will be no cleavage by DNase I and there will be a gap in the ladder. This gap is known as a footprint; the sequence of the footprint can be determined if a sequencing reaction is run alongside the DNase I protection assay. Detailed protocols for the technique are given in refs. *14* and *15*, also note that Stratagene now markets a DNA footprinting kit using the DNase I protection method. Exonuclease III protection has also been used for DNA footprinting, but in this technique the DNA must be labeled at the 5' end (*see* ref. *15*).

3.8. Concluding Remarks

The DNases are a ubiquitous and diverse group of enzymes that play an integral part in the functioning of DNA as the genetic material. Much of the work on DNases has been directed toward finding useful tools for molecular biology rather than toward elucidating their physiological roles; despite this, only one of the DNases, DNase I, has found regular use in techniques where limited digestion or complete digestion of DNA is required. The prospects are that other eukaryotic DNase will be found with more specific activities than DNase I, e.g., the enzymes involved in DNA repair and DNases that may function in recombination events. Analysis and use of these enzymes will possibly give us a greater insight into how DNA does function as the genetic material.

References

1. Jones, M. and Boffey, S. A. (1984) Deoxyribonuclease activities of wheat seedlings. *FEBS Lett.* **174,** 215–218.

2a. Melgar, E. and Goldthwaite, D. A. (1968) Deoxyribonucleic acid nucleases. 1. The use of a new method to observe the kinetics of deoxyribonucleic acid degradation by deoxyribonuclease I, deoxyribonuclease II and *Escherichia coli* endonuclease I. *J. Biol. Chem.* **243,** 4401–4408.

2b. Melgar, E. and Goldthwaite, D. A. (1968) Deoxyribonucleic acid nucleases. II. The effects of metals on the mechanism of action of deoxyribonuclease I. *J. Biol. Chem.* **2,** 4409–4415.
3. Bernardi, G. (1968) Acid deoxyribonuclease. *Adv. Enzymol.* **31,** 1–49.
4. Bollum, F. J. (1968) Degradation of the homopolymer complexes polydeoxyadenylate-polydeoxythymidylate, polydeoxyinosinate-polydeoxycytidylate and polydeoxyguanylate-polydeoxycytidilate by deoxyribonuclease I. *J. Biol. Chem.* **240,** 2599–2601.
5. Weir, A. F. and Bryant, J. A. (1989) Partial purification and properties of a chromatin-bound deoxyribonuclease fron the embryo axes of germinating pea. *Phytochemistry* **2,** 1339–1343.
6. Maxwell, I. H., Maxwell, F., and Hahn, W. F. (1977) Removal of RNase activity from DNase by affinity chromatography on agarose-coupled aminophenylphosphoryluridine-2'(3')-phosphate. *Nucleic Acids Res.* **4,** 241–246.
7. Zimmerman, S. B. and Sandeen, G. (1966) The ribonuclease activity of crystallized pancreatic deoxyribonuclease. *Anal. Biol. Chem.* **1,** 269–277.
8. Gurney, Jr.,T. and Gurney, E. G. (1984) Preparation of RNase free DNase by alkylation, in *Methods in Molecular Biology, vol. 2: Nucleic Acids* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 13–20.
9. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
10. Meinkoth, J. and Wahl, M. (1987) Nick translation. *Methods Enzymol.* **152,** 91–94.
11. Mathew, C. G. P. (1984) Radiolabeling of DNA by nick translation, in *Methods in Molecular Biology, vol. 2: Nucleic Acids* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 257–261.
12. Cheah, K. S. E. and Osborne, D. J. (1977) Analysis of nucleosomal deoxyribonucleic acid in a higher plant. *Biochem. J.* **1,** 141–144.
13. Grellet, F., Penon, P., and Cooke, R. (1980) Analysis of DNA associated with nucleosomes in pea chromatin. *Planta* **148,** 346–353.
14. Plumb, M. A. and Goodwin, G. H. (1988) Detection of sequence-specific protein-DNA interactions by the DNA-footprinting technique, in *Methods in Molecular Biology, vol. 4: New Nucleic Acid Techniques* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 139–164.
15. Hennighauen, L. and Lubon, H. (1987) Interaction of protein with DNA *in vitro. Methods Enzmol.* **152,** 721–735.

CHAPTER 3

DNA Polymerases (EC 2.7.7.7)

Martin J. Maunders

1. Introduction

DNA polymerases catalyze the synthesis of DNA molecules from monomeric deoxynucleotide triphosphate units. This definition encompasses those enzymes classed as DNA-dependent DNA polymerases (EC 2.7.7.7, which require both a DNA template strand and a DNA primer to which the monomeric units can be added), reverse transcriptases (EC 2.7.7.49, RNA-dependent DNA polymerases that utilize an RNA template), and terminal deoxynucleotidyl transferases (EC 2.7.7.31, which require no template). This chapter will consider only DNA-dependent DNA polymerases (EC 2.7.7.7, DNA nucleotidyltransferases, DNA-directed).

DNA-dependent DNA polymerases are a class of enzyme that is essential for the replication and maintenance of all organisms. All perform essentially the same reaction, the addition of monomeric units to synthesize a complementary copy of an existing DNA template. This activity can be used in molecular biology for a wide range of techniques. These include the labeling of DNA molecules, terminally or throughout the sequence, for use as probes or for sequence determination, the modification of DNA termini for facilitating subsequent manipulations, the production of double-stranded cDNA, the extension of mismatched sequences for site-directed mutagenesis, the extension of synthetic DNA sequences for gene synthesis, and the specific amplification of limiting quantities of DNA for analysis.

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

The choice of enzyme depends on the reaction conditions employed and the desirability of other inherent enzymatic activities. The range of enzymes available is rapidly expanding, ensuring a concomitant increase in the applications of DNA polymerases in molecular biology.

1.1. Prokaryotic DNA Polymerases

The classical *E. coli* DNA polymerase family consists of DNA polymerases I, II, and III (Pol I, Pol II, and Pol III). Pol III is largely responsible for chromosomal replication, and it initiates the growth of both DNA chains on RNA primers. Pol I "fills in" the gaps left by the intermittent nature of Pol III activity on the 3'–5' strand. Apart from this activity, the major role of Pol I is in DNA repair rather than replication. The role of Pol II is still largely unclear.

The three enzymes have widely differing properties. Pol I, which is the most widely used in molecular biology, is described in the next section. Pol II has a polymerase activity of only 5% that of Pol I and has no 5'–3' exonuclease activity. It can only act on duplexes with gaps of <100 bases and cannot replicate long single-stranded stretches from short primers. Pol III has a similar template requirement to Pol II, but is 300 times more active. Pol III holoenzyme consists of at least 13 components *(1)*, although various smaller complexes are active in vitro.

In addition to the *E. coli* enzymes, DNA polymerases from other bacteria are now attaining more prominence in molecular biology, particularly those from thermophilic organisms. Other useful DNA polymerases have been isolated from bacteriophage-infected *E. coli* cells, especially the enzymes encoded by bacteriophages T4 and T7.

1.2. Eukaryotic DNA Polymerases

Eukaryotic organisms possess three main DNA polymerases, Pol α, Pol β, and Pol γ. Pol α is considered to be the major DNA polymerase *(2)* and is responsible for chromosomal replication. The function of Pol β is unclear, whereas Pol γ, being located in both the nucleus and the mitochondria, is believed to be involved in mitochondrial DNA replication. The three enzymes all lack 3'–5' exonuclease activity, which has since been detected on a fourth DNA polymerase, Pol δ *(3)*. This finding supports the idea that eukaryotic DNA polymerases work in concert during DNA replication.

2. DNA Polymerase I (EC 2.7.7.7)

DNA polymerase I, Pol I, or Kornberg polymerase is the product of the *pol A* gene of *E. coli*, which has now been cloned to yield an overproducing strain.

2.1. Enzyme Data

The enzyme comprises a single polypeptide chain of mol mass 109,000 Da, with various enzymic activities. These are a 5'–3' polymerase, a 3'–5' exonuclease (proofreading activity), a 5'–3' exonuclease, and a phosphate-exchange activity. The polymerase activity catalyzes the formation of a complementary DNA strand to an existing single-stranded template, by the extension of a DNA primer possessing a 3'-hydroxyl terminus. The latter may be provided merely by a short gap or even a nick in the DNA duplex. The presence of these multiple activities allows the use of DNA polymerase I to perform several complex manipulations.

DNA polymerase I requires the presence of magnesium ions, as do all DNA polymerases. (In some cases, manganese ions may substitute, but reaction rates are lower, and there is a loss of fidelity.) The pH optimum for the enzyme is 7.4, and sulfhydryl reagents are required.

Enzymic activity is measured by the polymerization of labeled nucleotides into double-stranded DNA using poly-dAdT as both template and primer. One unit of activity catalyzes the incorporation of 10 nmol into an acid precipitable form in 30 min at 37°C *(4)*. DNA Polymerase I can be obtained at a purity of 5000 U/mg.

2.2. Uses of DNA Polymerase I

There are several uses of DNA polymerase I in molecular biology. First, it is used for the incorporation of labeled nucleotides into DNA probes by nick translation *(5)*. This requires the action of deoxyribonuclease I (*see* Chapter 2) to introduce nicks into the duplex with exposed 3'-hydroxyl termini to act as a substrate for the DNA polymerase I. 5'–3' Polymerization occurs concomitantly with 5'–3' exonuclease activity, which has the effect of physically translating the "nick" along the DNA molecule and allowing the introduction of labeled nucleotides into the newly replaced DNA strand. Reaction conditions are such that one complete exchange of the DNA duplex is achieved

in this manner. DNA polymerase I is the only common DNA polymerase that can carry out this coupled reaction by virtue of its 5'–3' exonuclease function. A typical nick translation protocol is given in Section 2.3.1.

Disadvantages of nick translation with regard to other methods of probe labeling include the facts that the specific activity that may be achieved is not high, and that at temperatures above 20°C artifactual sequences are produced, such as "snapback DNA," where the newly synthesized strand disassociates from the template and is then itself copied to form inverted DNA sequences.

Another use of DNA polymerase I is in repair synthesis, where the enzyme is used to "fill in" protruding 5' termini. End-labeling can also be achieved by removal of 3' protruding termini by the 3'–5' exonuclease activity and then subsequently filling in the 3' recess so formed.

DNA polymerase I can also be used in cDNA synthesis, where it is employed in the synthesis of the second DNA strand. The enzyme has an inherent RNase H activity that is of value here. However, cDNA synthesis is largely performed by DNA polymerase I large (Klenow) fragment or reverse transcriptase.

2.3. Experimental Procedures

DNA polymerase I is best stored in 50 m*M* potassium phosphate, pH 7.0, containing 250 µ*M* DTT and 50% glycerol. At –20°C, it is stable for over 6 mo.

2.3.1. Protocol for Nick Translation with DNA Polymerase I

2.3.1.1. Materials Required

1. Probe DNA.
2. Nick translation buffer:
   - 500 m*M* Tris-HCl, pH 7.8
   - 50 m*M* $MgCl_2$
   - 100 m*M* β-Mercaptoethanol
   - 100 µg/mL Bovine serum albumin
3. 200 µ*M* dGTP.
4. 200 µ*M* dCTP.
5. 200 µ*M* dTTP.
6. 0.1 µg/mL DNase I (prepare by serial dilution of 1 mg/mL stock).
7. DNA polymerase I (diluted to 2 U/µL in storage buffer).

8. 0.5*M* EDTA, pH 8.0.
9. TNE buffer:
   - 10 m*M* Tris-HCl, pH 8.0
   - 100 m*M* NaCl
   - 1 m*M* EDTA
10. TE buffer:
   - 10 m*M* Tris-HCl, pH 8.0
   - 1 m*M* EDTA
11. Sephadex G-50 equilibrated in TNE buffer.
12. α-^{32}P dATP (10 mCi/mL, 3000 Ci/mmol).
13. Sterile distilled water.

2.3.1.2. Equipment Required

1. Water bath (15°C).
2. Water bath (70°C) or dry heating block.
3. Bench-top centrifuge.
4. Sterile glass wool.
5. Sterile 1.5-mL microfuge tubes.
6. Pipets and sterile pipet tips.
7. Ice bucket.

2.3.1.3. Protocol

1. In a sterile tube on ice mix the following:
   - Probe DNA (up to 1 µg)
   - 5 µL Nick translation buffer
   - 200 µ*M* dGTP, dCTP, dTTP (5 µL each)
   - 5 µL α-^{32}P dATP (50 µCi, 20 µ*M*)
   - 2 µL DNase I (200 pg)
   - 1 µL DNA polymerase I (2 U)
   - Sterile distilled water to 50 µL
2. Incubate at 15°C for 60 min.
3. Stop the reaction by addition of 2 µL of 0.5*M* EDTA, pH 8.0, or by heating to 70°C for 10 min.
4. Purify the labeled probe from unincorporated nucleotides by a spun column method. Prepare a column of Sephadex G-50 in a 200-µL pipet tip plugged with glass wool. Wash the column with 500 µL TNE buffer, by supporting it in a microfuge tube and spinning at 2000*g* for a few seconds. Load the nick translation reaction onto the column, and spin for 30 s. Elute the labeled probe with a futher 100 µL of TNE (or TE) buffer and a further spin. Pool the run-through and eluate, and then discard the column containing the unincorporated nucleotides.

3. Klenow Fragment

3.1. Enzyme Data

By treatment of DNA polymerase I with the protease subtilisin, a polypeptide of mol mass 35,000 Da is removed that carries the 5'–3' exonuclease activity *(6)*. The remaining moiety, of mol mass 76,000 Da, retains all the other enzymic activities and is commonly termed the DNA polymerase I large fragment or Klenow fragment.

3.2. Uses of Klenow Fragment

It is impossible to use Klenow fragment for nick translation because of the lack of a 5'–3' exonuclease activity. However, this limitation is turned to advantage in other applications of the enzyme by rendering the polymerization activity more controllable.

The primer extension activity of Klenow fragment has made it widely used in dideoxynucleotide sequencing methods *(7)*, whereby DNA templates up to 350 bases long can be sequenced at a rate of 45 nucleotides/s. Several different DNA polymerases are commonly used for sequencing reactions. Each has distinct abilities when encountering particularly "difficult" regions of the template sequence.

The enzyme can also be used to produce labeled double-stranded DNA probes (actually only labeled in one or the other of the two strands) by the extension of oligomeric primers of random sequence *(8)*. The oligolabeling technique produces probes of a much higher specific activity than nick translation, with over 55% of the labeled nucleotide incorporated in 30 min at 37°C. Although the entire probe sequence is represented in the labeling mix, individual molecules are commonly only a few hundred bases in length, which may be a disadvantage for some applications. A typical reaction protocol is given in Section 3.3.1.

Single-stranded (strand-specific) probes can also be prepared by end-labeling of protruding, flush, or recessed 3' termini. This is achieved by removal of the 3' nucleotides by means of the 3'–5' exonuclease activity (which is more active on single-stranded than double-stranded DNA), to form a recessed 3' end. Following the addition of high concentrations of deoxynucleotide triphosphates (one or more of which may be labeled), the polymerase activity resynthesizes the excised region *(9,10)*. This process has advantages over nick translation in that

the labeled probe produced is strand-specific, and there is no production of hairpin-like artifacts.

This reaction can of course be performed with unlabeled deoxynucleotides to produce flush or modified termini for subsequent cloning reactions. The fill-in reaction (*see* Section 3.3.2.) can be controlled by the withholding of one particular deoxynucleotide to generate novel cohesive ends. Excess Klenow fragment will also yield only partially filled termini because of the 3'–5' exonuclease activity.

Klenow fragment is also used for elongating oligomeric primers in site-directed mutagenesis to produce double-stranded DNA. However, the enzyme tends to displace mismatched oligomers and so remove the intended mutation. T4 DNA polymerase is perhaps more useful in this context.

The enzyme can also be used for the synthesis of the second DNA strand during cDNA synthesis *(11,12)*. Finally, Klenow fragment was originally used in the polymerase chain reaction, before the introduction of thermostable DNA polymerases permitted the process to be automated.

3.3. Experimental Procedures

Klenow fragment has similar physical properties to Pol I and should be stored under similar conditions. The following sections describe experimental protocols for two techniques utilizing Klenow fragment: radiolabeling DNA by extension of random primers and in-filling of cohesive termini.

3.3.1. Protocol for Labeling with Random Hexameric Primers

3.3.1.1. Materials Required

1. Probe DNA.
2. Reaction buffer:
   - 450 m*M* HEPES, pH 6.6
   - 50 m*M* $MgCl_2$
   - 10 m*M* Dithiothreitol
   - 2 mg/mL Bovine serum albumin
   - 2.5 m*M* dGTP
   - 2.5 m*M* dCTP
   - 2.5 m*M* dTTP
3. Klenow fragment (diluted to 2 U/μL in storage buffer).
4. α-^{32}P dATP (10 mCi/mL, 3000 Ci/mmol).
5. Sterile distilled water.

3.3.1.2. Equipment Required

1. 1.5-mL Microfuge tubes.
2. Microfuge.
3. Boiling water bath or dry heating block.
4. Water bath (37°C).
5. Pipets and sterile pipet tips.
6. Ice bucket.

3.3.1.3. Protocol

1. Dilute 50 ng probe DNA to 34 µL with sterile water.
2. Denature by boiling for 3 min. Spin down condensation and stand on ice.
3. Add 10 µL of reaction buffer.
4. Add 5 µL (50 µCi) α-^{32}P dATP.
5. Add 1 µL (2 U) Klenow fragment.
6. Incubate at 37°C for 30 min or at room temperature for 1–2 h.
7. Add 200 µL sterile water, and terminate the reaction by boiling for 3 min.
8. Stand the reaction on ice and use (unpurified) within 10 min.

3.3.2. Protocol for In-Filling Cohesive Termini

3.3.2.1. Materials Required

1. Target DNA.
2. Reaction buffer:
   - 70 m*M* Tris-HCl, pH 7.5
   - 70 m*M* $MgCl_2$
   - 10 m*M* Dithiothreitol
   - 250 µ*M* dATP
   - 250 µ*M* dGTP
   - 250 µ*M* dCTP
   - 250 µ*M* dTTP

   N.B. Alternatively, the in-filling reaction can be performed in most restriction endonuclease buffers.
3. 500 m*M* NaCl.
4. Klenow fragment (diluted to 0.5 U/µL in storage buffer).
5. Sterile distilled water.

3.3.2.2. Equipment Required

1. 1.5-mL Microfuge tubes.
2. Water bath (25°C).
3. Water bath or dry heating block (70°C).
4. Pipets and sterile pipet tips.

3.3.2.3. PROTOCOL

1. Dilute 1 µg of target DNA to 44 µL with distilled water.
2. Add 5 µL of reaction buffer.
3. Optionally, add NaCl to a concentration of 0–50 m*M*, reducing the volume of water use in step 1 *pro rata*.
4. Add 1 µL (0.5 U) of Klenow fragment.
5. Incubate at 25°C for 15 min.
6. Terminate the reaction by incubating at 70°C for 10 min.

4. T4 DNA Polymerase

T4 DNA polymerase is produced by the *E. coli* bacterium when infected with bacteriophage T4.

4.1. Enzyme Data

It is a monomeric enzyme of mol mass 114,000 Da, that will sequentially add mononucleotide units to a DNA primer by means of its 5'–3' polymerase activity. The enzyme cannot utilize nicked DNA as a substrate, since its lack of a 5'–3' exonuclease activity *(13)* prevents it from displacing the existing downstream DNA strand. The enzyme has a highly active 3'–5' proofreading exonuclease activity, 200-fold more active than that of Klenow fragment, which again is more active against single- than double-stranded DNA.

4.2. Uses of T4 DNA Polymerase

T4 DNA polymerase is used to label 3' termini in the same manner as Klenow fragment. The reaction can be performed at 37°C or at room temperature, but incubation at 12°C maximizes the polymerase activity over that of the exonuclease.

The exonucleolysis/in-filling process can of course be performed in the absence of labeled deoxynucleotide triphosphate to convert cohesive or "ragged" termini to "polished" flush ends suitable for cloning. A further use of T4 DNA polymerase is in the process of "gap-filling" during site-directed mutagenesis. This entails the conversion of a partially single-stranded molecule to a fully double-stranded entity. For this application, it is usual to include the T4 gene 32 protein, which stimulates DNA synthesis *(14)*.

4.3. Experimental Procedures

T4 DNA polymerase will function in most restriction endonuclease buffers. An alternative reaction buffer is 50 m*M* Tris-HCl, pH 8.5, 15 m*M* ammonium sulfate, 7 m*M* $MgCl_2$, 0.1 m*M* EDTA, 10 m*M* β-mercaptoethanol, 33 μ*M* dNTPs, and 200 μg/mL BSA *(15)*. The enzyme requires Mg^{2+} (at least 6 m*M*) for activity. Its pH optimum is in the range 8.0–9.0, exhibiting 50% activity at pH 7.5 and 9.7. The reaction temperature should be 37°C.

The enzyme is best stored in 10 m*M* Tris-HCl, pH 8.0, 100 m*M* NaCl, 2.5 m*M* $MgCl_2$, 0.5 m*M* EDTA, 2 m*M* DTT, and 50% glycerol *(15)*. It is stable if stored undiluted in this buffer at –20°C for several months.

5. T7 DNA Polymerase

This DNA polymerase is also produced by bacteriophage-infected *E. coli*, although it is now obtained from a cloned gene.

5.1. Enzyme Data

It is the only common DNA polymerase that consists of more than one subunit. The functional enzyme is composed of the T7 gene 5 protein (mol mass 80,000–85,000 Da), which carries all the enzymic activities, plus the *E. coli* thioredoxin protein (mol mass 12,000 Da) required to associate with the template.

The enzyme has a very active 5'–3' polymerase activity that can incorporate 300 nucleotides/s. It is also the most highly processive of all the DNA polymerases in that it can extend a primer for several kilobases without the enzyme dissociating from the template. The enzyme is also very tolerant of nucleotide analogs, which renders it very useful for DNA sequencing and oligolabeling techniques.

T7 DNA polymerase lacks a 5'–3' exonuclease activity *(16)*, but like T4 DNA polymerase, it has a potent 3'–5' exonuclease. In this case, the activity is 1000-fold that exhibited by Klenow fragment. The 3'–5' exonuclease activity can be inactivated by incubation with a reducing agent, oxygen, and low concentrations of Fe^{2+}. This modification does not affect the polymerization ability of the enzyme *(17)*. Chemical modification removes 99% of the exonuclease activity, and total removal has been achieved by genetic engineering. These modified versions of T7 DNA polymerase are readily available commercially.

5.2. Uses of T7 DNA Polymerase

A major use of native or modified T7 DNA polymerase is in DNA sequencing when fragments of over 1,000 bases can be analyzed from one reaction. The enzyme is also useful for other applications when extension of a primer over long stretches of template are required. The other uses of this enzyme include the rapid end-labeling of DNA probes and the extension of primers during site-directed mutagenesis.

5.3. Experimental Procedures

A suitable reaction buffer for T7 DNA polymerase is 40 m*M* Tris-HCl, pH 7.5, 10 m*M* $MgCl_2$, 1 m*M* DTT, 100 µg/mL BSA, 300 µ*M* dNTPs *(18)*, and reactions should be performed at 37°C. The enzyme is stable for over 6 mo when stored at –20°C in 25 m*M* Tris-HCl, pH 7.5, 250 m*M* NaCl, 5 m*M* DTT, and 50% glycerol.

6. *Taq* Polymerase

Taq polymerase is a DNA polymerase isolated from the bacterium *Thermus aquaticus (19)*. Similar enzymes have been isolated from other thermophilic organisms, including *Thermus thermophilus*, *Th. flavus*, *Th. litoralis*, *Pyrococcus furiosus*, and *Bacillus stearothermophilus*. *Taq* polymerase preparations have been isolated from several different strains of *Th. aquaticus* itself, and each has slightly different characteristics. The most common source is still strain YT-1, although the enzyme is also available from recombinant clones.

6.1. Enzyme Data

Taq polymerase is a monomeric enzyme with a mol mass in the range 60,000–97,000 Da depending on the source. The enzyme exhibits a 5'–3' polymerase activity with a high processivity and lacks a 3'–5' exonuclease (proofreading) activity, but does possess a 5'–3' exonuclease (strand replacement) activity. A modified form of the enzyme is now available from which this latter activity has been removed. DNA polymerases from *Thermus litoralis* and *Pyrococcus furiosus* do exhibit the 3'–5' exonuclease activity *(20,21)*, which renders DNA replication less error-prone.

Taq polymerase has another activity in that it can incorporate an additional nucleotide (usually deoxyadenosine) at the 3' terminus of any DNA duplex in a nontemplate-dependent reaction *(22)*. This addition often hinders subsequent manipulations, but can be turned to advantage as a means of increasing cloning efficiency *(23)*.

The major distinguishing feature of *Taq* polymerase is its extreme thermal stability. The enzyme can withstand temperatures in excess of 95°C for prolonged periods, and in fact, its optimum for reaction is 75°C. The rate of reaction is reduced to 50% at 60°C, and to 10% at 37°C. A range of DNA polymerases from other species of thermostable bacteria (including those mentioned earlier) are now becoming available, which will broaden the spectrum of possible reaction conditions.

6.2. Uses of Taq *Polymerase*

Taq polymerase is rapidly replacing other DNA polymerases in many laboratory techniques, primarily because of its thermal stability and tolerance of temperature changes. This feature allows reactions to be performed at elevated temperatures, whereby DNA duplex melting and annealing can be rigorously controlled. This ensures the presence of a suitable single-stranded template for high-fidelity replication, and also the stringency of primer association can be precisely defined. Additional benefits are that contaminating enzyme activities are largely eradicated, and complex but routine procedures can be automated.

The major use of *Taq* polymerase at present is in the polymerase chain reaction (PCR) *(24)*. This technique is a simple method of amplifying minute quantities of DNA for a variety of subsequent procedures, including cloning, sequencing, hybridization, and genome mapping *(25)*. The enzyme from *Thermus thermophilus* (when used in a manganese-containing buffer) has an additional reverse transcriptase activity, which extends the use of PCR directly to cDNA synthesis. The PCR has applications beyond the reseach laboratory, including uses in forensic science, disease diagnosis or prognosis, paternity testing, and in animal and plant breeding programs.

Another use of *Taq* polymerase is directly in DNA sequencing, where the high temperatures employed help reduce problems caused by secondary structure in the template and allow an increase in the stringency of primers used. *Taq* polymerase has, for the reasons mentioned earlier, become a very important tool for the molecular biologist and is described in detail in Chapter 4.

References

1. McHenry, C. S. (1988) DNA Polymerase III holoenzyme of *Escherichia coli*. *Ann. Rev. Biochem.* **57**, 519–550.

2. Loeb, L. A., Liu P. K., and Fry M. (1986) DNA polymerase α: Enzymology, function, fidelity and mutagenesis, in *Progress in Nucleic Acid Research and Molecular Biology*, vol. 33 (Cohn, W. E., ed.), Academic, New York, pp. 57–110.
3. Tsang-Lee, M. Y. W., Tan, C.-K., Downey, K. M., and So, A. G. (1981) Structure and functional properties of calf thymus DNA polymerase δ, in *Progress in Nucleic Acid Research and Molecular Biology*, vol. 26 (Cohn, W. E., ed.), Academic, New York, pp. 83–96.
4. Richardson, C. C., Schildkraut, C. L., Aposhian, H. V., and Kornberg, A. (1964) Enzymatic synthesis of deoxyribonucleic acid. XIV. Further purification and properties of deoxyribonucleic acid polymerase of *Escherichia coli*. *J. Biol. Chem.* **239,** 222–232
5. Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P. (1977) Labeling DNA to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* **113,** 237–251.
6. Klenow, H., Overgaard-Hansen, K., and Patkar, S. A. (1971) Proteolytic cleavage of DNA polymerase into two different catalytic fragments. *Eur. J. Biochem.* **22,** 371–381.
7. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
8. Feinberg, A. P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132,** 6–13.
9. Telford, J. L., Kressman, A., Koski, R. A., Grosschedl, R., Muller, F., Clarkson, S. G., and Birnstiel, M. L. (1979) Delimitation of a promoter for RNA polymerase III by means of a functional test. *Proc. Natl. Acad. Sci. USA* **76,** 2590–2594.
10. Cobianchi, F. and Wilson, S. H. (1987) Enzymes for modifying and labeling DNA and RNA. *Methods Enzymol.* **152,** 94–110.
11. Rougeon, F., Kourilsky, P., and Mach, B. (1975) Insertion of a rabbit β-globin gene sequence into an *E.coli* plasmid. *Nucleic Acids Res.* **2,** 2365–2378
12. Gubler, U. (1987) Second strand cDNA synthesis: classical method. *Methods Enzymol.* **152,** 325–329.
13. Englund, P. (1971) Analysis of nucleotide sequences at 3' termini of duplex deoxyribonucleic acid with the use of the T4 deoxyribonucleic acid polymerase. *J. Biol. Chem.* **246,** 3269–3276.
14. Craik, C. S., Largman, C., Fletchet, T., Roczniak, S., Barr, P. J., Fletterick, R., and Rutter, W. J. (1985) Redesigning trypsin: alteration of substrate specificity. *Science* **228,** 291–297.
15. Biochemicals for Molecular Biology (1990/1991) T4 DNA polymerase. Boehringer Mannheim GmbH, Mannheim, Germany, p. 67.
16. Tabor, S., Huber, H. E., and Richardson, C. C. (1987) *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J. Biol. Chem.* **262,** 16,212–16,223.
17. Tabor, S. and Richardson, C. C. (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* **84,** 4767–4771.

18. Molecular and Cell Biology Catalogue (1992/1993) FPLC pure T7 DNA Polymerase, cloned. Pharmacia P-L Biochemicals Inc., Milwaukee, WI. p. 172.
19. Chien, A., Edgar, D. B., and Trela, J. M. (1976) Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus. J. Bacteriol.* **127,** 1550–1557.
20. New England Biolabs Catalogue (1990/1991) Vent™ DNA polymerase. New England Biolabs Inc., Beverly, MA. p. 47.
21. Bergseid, M., Scott, B., Mathur, S., Nielson, K., Shoemaker, D., and Mathur, E. (1991) A high fidelity thermostable DNA polymerase isolated from *Pyrococcus furiosus. Strategies in Molecular Biology 4.* Stratagene Cloning Systems Inc., La Jolla, CA, pp. 34–35
22. Clark, J. M. (1988) Novel non-templated nucleotide addition reactions catalysed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **16,** 9677–9686.
23. Mead, D. A., Pey, N. K., Herrnstadt, C., Marcil, R. A., and Smith, L. M. (1991) A universal method for the direct cloning of PCR amplified nucleic acid. *Biotechnology* **9,** 657–663.
24. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985) Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anaemia. *Science* **230,** 1350–1354.
25. Ehrlich, H. A., Gelfand, D., and Sninsky, J. (1991) Recent advances in the polymerase chain reaction. *Science* **252,** 1643–1651.

CHAPTER 4

Taq Polymerase (EC 2.7.7.7)

With Particular Emphasis on Its Use in PCR Protocols

Axel Landgraf and Heiner Wolfes

1. Introduction

Taq polymerase (EC 2.7.7.7) is a thermostable DNA-dependent DNA polymerase that was first isolated in 1976 from *Thermus aquaticus* strain YT-1 (ATTC # 25 104)*(1)*. It catalyzes the template-directed polymerization of dNTPs at high temperatures.

5'-dGdTdGdC	*Taq* pol.	5'-dGdTdGdCdAdCdGdAdGdAdG
3'-dCdAdCdGdTdGdCdTdCdTdC	→	3'-dCdAdCdGdTdGdCdTdCdTdC
	+dNTPs	

When *Taq* polymerase was isolated in 1976, nobody could have known the impact the application of *Taq* polymerase to the polymerase chain reaction (PCR) would have on molecular biology *(2,3)*. In an issue of *Science* in 1989, the enzyme was dubbed molecule of the year *(4)*.

2. Enzymology of the *Taq* Polymerase

Taq polymerase constitutes only 0.01–0.02% of the total protein of *Thermus aquaticus*. The increased demand for the enzyme led to the cloning of the gene in *E. coli*. An advantage of cloning an enzyme from a thermophile organism in a mesophile bacterium is that, in general, contaminants in the enzyme preparation should not be active at a temperature optimal for the *Taq* polymerase *(5)*.

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

The sequence for the enzyme allowed a better characterization of the *Taq* polymerase *(5)*. Given the biochemical, biophysical, and sequence data, it can be concluded that the two enzyme preparations published in the original paper *(1,5)* represented either two different enzymes or that the smaller protein was a fragment of the constitutive *Taq* polymerase. The characteristic features of the commercially available cloned *Taq* polymerase are analogous to those of the latter protein.

The *Taq* polymerase gene is 2499 nucleotides long and codes for a protein of 832 amino acids. The theoretical mass is 93,900 Da, which is in fairly good agreement with the apparent mass of 97,300 Da obtained from gel electrophoresis. A sequence comparison with polymerase I from *E. coli* shows significant similarities in the 5'–3' exonuclease domain (amino acids 1–410) as well as in the polymerase domain (amino acids 410–832). Regions to which definite functions could be assigned in polymerase I are highly conserved in *Taq* polymerase.

No proofreading activity (3'–5' exonuclease) could be identified in *Taq* polymerase *(6)*. There is no homology in *Taq* polymerase to a domain ranging from amino acid 300 to 410 in polymerase I, to which proofreading activity is attributed in the *E. coli* enzyme.

Taq polymerase catalyzes the DNA-dependent polymerization of dNTPs. One unit of the enzyme is defined as the amount of enzyme that will incorporate 10 nmol of radioactively labeled dTTP into acid-insoluble material at 80°C in 30 min *(1)*. The enzyme could be purified to a specific activity of 200,000 U/mg protein. As an optimal reaction temperature, 75–80° C was established, at which the incorporation rate is approx 150 nucleotides/s/enzyme molecule *(7)*. In the M13 system, incorporation rates of <60, 24, 1.25, and 0.25 nucleotides/s/enzyme molecule were measured at 70, 55, 37, and 22°C respectively *(8)*. Incorporation rates decrease at temperatures >90°C; the stability of the enzyme also decreases drastically. The half-life time of the enzyme declines from 130 to 40 to between 5 and 6 min at 92.5, 95, and 97.5°C respectively *(7)*.

Enzyme activity is dependent on bivalent cations. Concentrations of 2 m*M* $MgCl_2$ are optimal. Maximum polymerization rates are obtained with 0.7–0.8 m*M* dNTPs. Substrate inhibition is observed at dNTP concentrations of 4–6 m*M*. Monovalent cations also have an effect on the activity of the enzyme. Optimum conditions are 50 m*M* KCl, whereas inhibition is reported at concentrations >75 m*M* KCl *(8)*. NaCl, NH_4Cl, and NH_4 acetate cannot substitute for KCl without a decrease in specific

Table 1
Inhibitor Effects on *Taq* Polymerase I Activity[a]

Inhibitor	Concentration	% Activity
Urea	<0.5*M*	100
	1.0*M*	118
	1.5*M*	107
	2.0*M*	82
SDS	0.001%	105
	0.01%	10
	0.1%	<0.1
Ethanol	<3%	100
	10%	110
DMSO	<1%	100
	10%	53
	20%	110
DMF	<5%	100
	10%	82
	20%	17

[a]Based on results from ref. 7.

activity. Denaturating agents, detergents, and organic solvents in low concentration are tolerated by *Taq* polymerase; at higher concentrations, inhibition is observed.

A compilation of the effects of several substances is shown in Table 1. These concentrations are not obligatorily applicable to the PCR; e.g., 0.5*M* Urea are inhibitory to the PCR, whereas up to 1.5*M* Urea do not alter the incorporation yields of *Taq* polymerase. The inhibition by 0.01% SDS can be reversed by adding 0.1% Tween 20 or NP40 *(7)*.

Taq polymerase is an enzyme with high processivity *(8)*. An absolute requirement is a DNA template with a primer to initiate the reaction. The primer must be dephosphorylated on the 3' end *(7)*. *Taq* polymerase has a strand displacement 5'–3' exonuclease activity. 5'-phosphorylated oligodeoxynucleotides are not degraded by the 5'–3' exonuclease activity of the enzyme. Studies of the fidelity of DNA synthesis by *Taq* polymerase revealed error rates of 1.1×10^{-4} for misincorporation and 2.5×10^{-5} for frameshift mutations *(6)*. An error rate of 2×10^{-4} was determined by PCR experiments followed by sequencing *(3)* and by denaturing gradient gel electrophoresis *(9)*. T-C transitions are occurring more frequently than other misincorporation errors *(6,10)*.

3. Polymerase Chain Reaction (PCR)

The first publication concerning the polymerase chain reaction (PCR) appeared in 1985 *(2)*. In this protocol, Klenow polymerase was used as the replicating enzyme, thus necessitating the addition of new polymerase after each denaturation step. In 1988, the first use of *Taq* polymerase for the PCR was reported *(3)*, leading the way to an automation of the protocol. The application of *Taq* polymerase to the PCR was the basis for the success of the technique. The enzyme is extremely suitable for the following reasons:

- *Taq* polymerase is stable up to 95°C; thus it is not necessary to replenish the enzyme after each PCR cycle;
- The maximal enzyme activity is between the temperature range of 70–75°C, which minimizes secondary structures of the template, resulting in high polymerization yield; and
- The annealing temperature can be chosen from 30–70°C, allowing an optimal adaptation of cycle parameters to appropriate annealing temperatures of the primers; therefore, byproducts are hardly generated.

3.1. Fundamentals of the PCR

The PCR is an enzymatic chain reaction that leads to the amplification of specific DNA sequences from a given template *(11)*. In addition to the template, two primers (typically: oligodeoxynucleotides of 20 nucleotides in length) are needed, whose sequence must be complementary to the template DNA. The amplification is performed in the presence of dNTPs as substrate and catalyzed by *Taq* polymerase. The sequence between the two regions complementary to the primers is doubled by repeating the cycle, which consists of a series of temperature steps in which the DNA is denatured, the primers are annealed, and the polymerization of the specific sequence occurs. The newly synthesized sequence is available as template in the next cycle; therefore, theoretically, a twofold increase of template is achieved in each step. In theory, 30 cycles should yield an amplification of a specific sequence by a factor of $2^{30} \approx 1 \times 10^9$, as long as substrates or enzyme is not limiting (Fig. 1).

The PCR technique reaches outstanding results because of the fact that the chain reaction requires the specific annealing of two primers to a complementary sequence. Unspecific annealing of primers or nicks in the template does not lead to exponential, but only linear

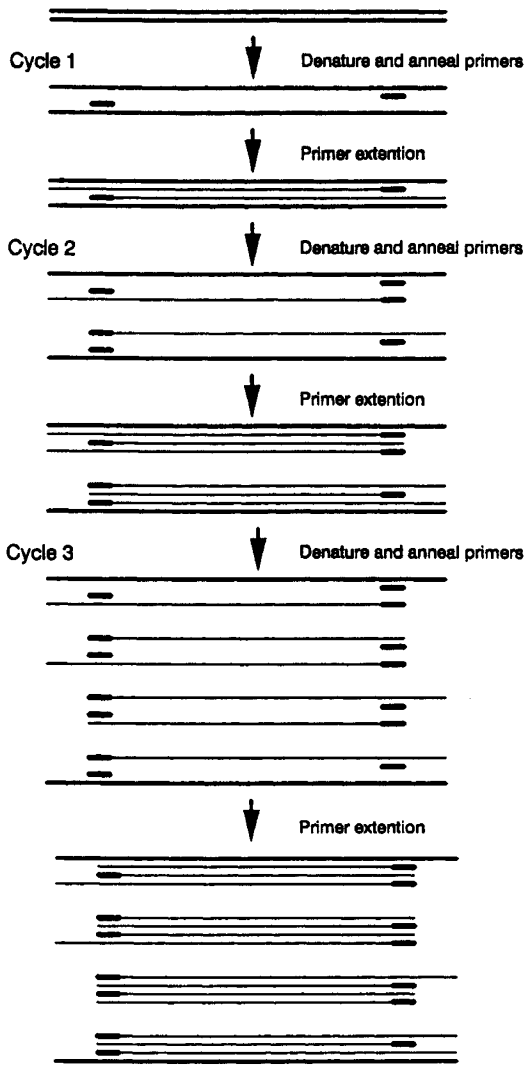


Fig. 1. Principle of the polymerase chain reaction.

amplification. Therefore, this system offers an optimal ratio between specific and unspecific signal.

Characteristically, <1 μg of template in a vol of 100 μL is used in the PCR. In this experimental outline, a single copy gene of the human genome would be in a concentration of 2.5 f*M*. Primers and enzyme are used in large excess (1 μ*M* and 1 n*M*, respectively). The efficiency of

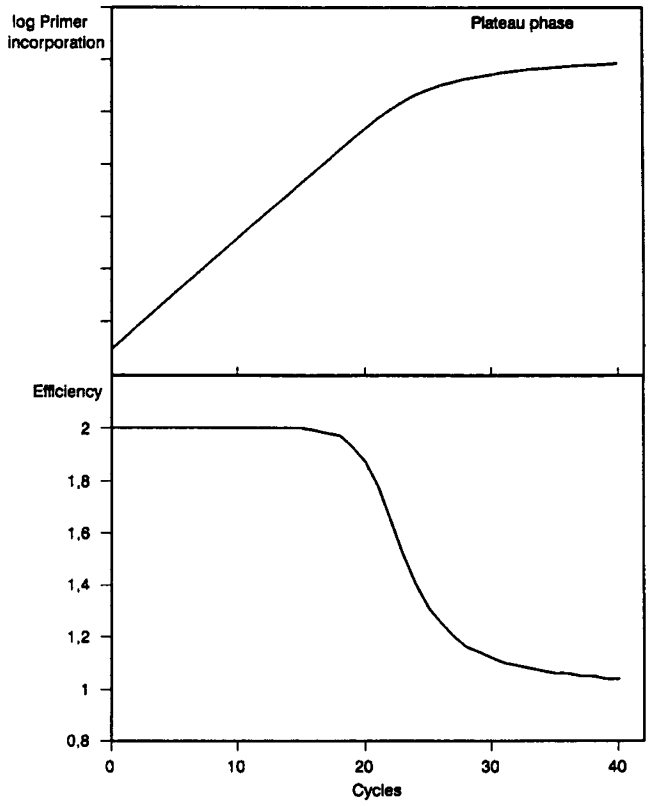


Fig. 2. Typical PCR. Changes in efficiency and product yield. Graphs were calculated on the basis of: 1 μ*M* of each primer spanning a stretch of 300 bp, 200 μ*M* each dNTP, 2 U enzyme 1 n*M* (processing 10 templates/cycle, estimated enzyme degradation 5%/cycle); and 200-ng template (0.05 aMol single copy gene). Volume of 100 μL.

this reaction will be close to the theoretical value of 2, which means that the sequence is doubled in every cycle. After 20–30 cycles are completed, the product (and therefore the template) is amplified by a factor of 10^6. Now the template concentration is in the same order as the enzyme concentration. The efficiency of the PCR decreases steadily, and exponential growth is no longer observed (Fig. 2). This phase of the reaction is called plateau phase *(12)*. Hence, the product yield is dependent on the enzyme concentration *(13)* (Fig. 3).

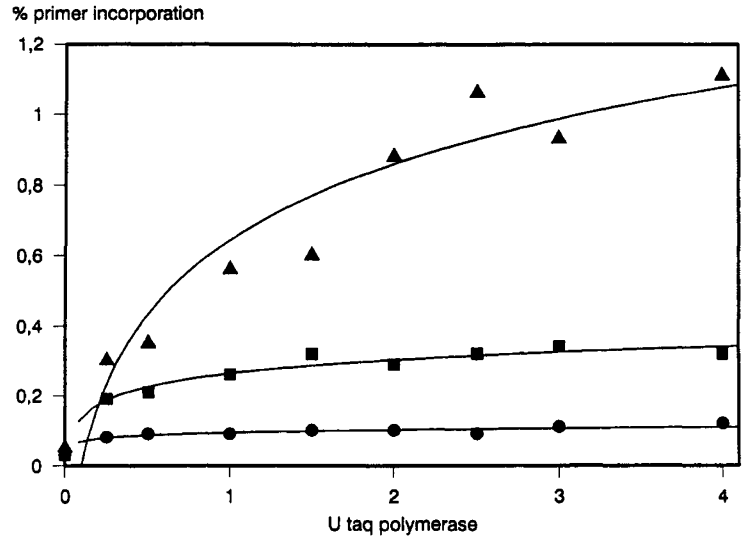


Fig. 3. Dependency of primer incorporation on enzyme amounts. ●, 22 cycles; ■, 24 cycles; ▲, 26 cycles.

The initial template concentration hardly influences the product yield in the plateau phase. Even 1–10 copies of target sequence will be amplified to give the plateau concentration after 40–50 cycles. Under the experimental conditions mentioned earlier, incorporation of 10–100 n*M* primer can be expected, which corresponds to roughly 1 µg DNA. Therefore, the amplification has to reach the plateau phase, if the product is to be analyzed by gel electrophoresis and ethidium bromide staining.

The PCR is applicable for quantitative determinations only prior to the onset of the plateau phase. Otherwise, quantitative analysis has to be performed using competing primer pairs in an experiment (Section 4.7.).

The efficiency, defined as the amplification factor per cycle, has a theoretical value of 2, in accordance with the doubling of the sequence in every step. However, in practice, this is not achieved. The efficiency is dependent on the copy number of the template: The lower the copy number, down to a certain limit, the better the PCR works. An efficiency close to the theoretical value (>1.9) can only be expected with copy numbers of the template in the range of 10^7–10^8 *(14)*. Starting

from this limit, the yield of the reaction is restricted by too low enzyme concentration or by competitive reannealing of the fragments *(12)*. For the amplification of long fragments, the reaction efficiency is very dependent on the fragment length: The efficiency of 1.9 for a fragment of 1000 nucleotides declines to 1.4 for a fragment of 7 kb *(15)*.

3.2. Application of Taq *Polymerase in the PCR*

The application of *Taq* polymerase in the PCR requires certain prerequisites to be fulfilled.

3.2.1. Buffers

The buffer used most commonly in the PCR consists of 10 m*M* Tris-HCl, pH 8.4, 50 m*M* KCl, 1.5–2.5 m*M* $MgCl_2$, and 0.01% gelatin as stabilizer *(2)*. The Mg^{2+} ion concentration influences predominantly the specificity of the PCR. With increasing Mg^{2+} concentrations, the amplification of unspecific sequences increases. At high dNTP concentrations, it should also be taken into account that the concentration of free ions is reduced as a result of Mg^{2+} binding to dNTPs and to DNA. It was reported that, in an incorporation experiment, high dNTP concentrations (4–6 m*M*) were inhibitory *(8)*.

A different widely used buffer system is composed of 67 m*M* Tris-HCl, pH 8.8, 6.7 m*M* $MgCl_2$, 16.6 m*M* $[NH_4]_2SO_4$, 10 m*M* β-mercaptoethanol, and 0.02% bovine serum albumin (BSA) or gelatine. In principle, gelatine should be given preference over BSA, because BSA is likely to denature. On the other hand, in some protocols the use of stabilizing proteins is totally abandoned *(8)*. Sometimes 10% dimethylsulfoxide (DMSO) is added to the buffers mentioned above *(10)*.

The Michaelis Menten constant K_m of *Taq* polymerase for each dNTP has a value of 10–15 μ*M* *(16)*. A normal PCR experiment contains 50–200 μ*M* of each dNTP. This concentration is sufficient to amplify even large fragments to microgram yields without any limitation of dNTP substrate. Even after 60 cycles, a pool of about 50% dNTPs is still available *(8)*. Sometimes the dNTP concentration is lowered to 10 μ*M* in order to achieve an optimum incorporation of radioactive labeled deoxynucleotides *(17)*. Modified nucleotides, e.g., deoxynucleoside-α-thiotriphosphates (dNTPαS) *(18)*, 7-deaza-2'-deoxyguanosine-5'-triphosphate d(c^7GTP) *(8)*, 5-Biotin *(11)*-2' deoxyuridine-5'- triphosphate (Bio1 1dUTP) *(16)*, are incorporated at high rates by *Taq* polymerase.

3.2.2. Primer

In theory, at least a 16 bp of primer are necessary for an unambiguous hybridization to a given stretch of DNA in the human genome (3×10^9 bp). In practice, primers 20–25 nucleotides long are used in PCR experiments, with the G-C content varying from 40–60% in order to ensure specific amplification. Primers shorter than 20 nucleotides are used successfully for allele-specific hybridization *(19)*. Except for this purpose, utilization of shorter primers should be avoided, since they might detach from the template at the high temperature where polymerization is effective.

To determine the approximate melting temperatures of a primer–DNA duplex, the equation:

$$T_m\ [°C] = 2[T + A] + 4[G + C]$$

can be used *(20)*. More exact numbers can be calculated by applying algorithms *(21)* that take the occurrence of intramolecular interactions, e.g., hairpin structures, into account.

In established oligodeoxynucleotide synthesis, a product yield of more than 80% for short primers can be expected *(20)*. Therefore, a purification of the crude synthesis products is most often not required. Depending on circumstances, which demand purification of the crude products, e.g., an oligodeoxynucleotide has to be labeled prior to the PCR, two methods are suitable for purification: (1) When little amounts of primers are needed, denaturing PAGE with 7*M* urea followed by detection by shadow casting *(22)* is a convenient method for purification. Several oligodeoxynucleotides can be purified in parallel. Since urea is an inhibitor of *Taq* polymerase, it must be removed by chromatographic techniques after elution from the gel slice. (2) Chromatographic procedures are a better choice for the purifications of large quantities of oligodeoxynucleotides. Reversed-phase HPLC techniques and ion-exchange chromatography are commonly used *(23)*.

Primers should always be stored lyophilized in aliquots at –20°C. It is convenient to prepare 10X stock solutions with 1–10 μ*M* concentrations of the primer. A primer concentration of 0.1–1 μ*M* is optimal for the PCR. An increase in the oligodeoxynucleotide concentration will produce a rise in nonspecific bands. One example of the production of an unspecific band is owing to partial primer duplex formation and polymerization and leads to a product of less than the length of the two primers.

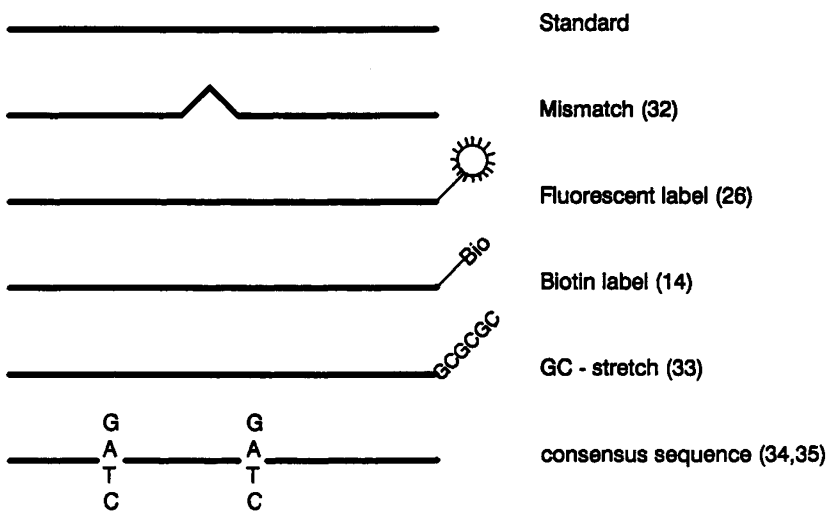


Fig. 4. Versatility of PCR primers.

Apparently, primer pairs are formed and elongated by the enzyme *(24)*. Biased primer concentration results in asymmetric amplification of one strand *(25)* (Section 4.3.).

PCR is capable of dealing with many different problems by utilization of primers specifically designed for a particular purpose (Fig. 4). Primers carrying a chemical label at the 5' end, e.g., oligodeoxynucleotides substituted with a biotin or fluorescent dye, are well accepted by *Taq* polymerase. Oligodeoxynucleotides can be synthesized with a primary amino group (aminolink) at the 5' end, to which biotin *(14)* or fluorescent markers *(26)* may be coupled. As a consequence, 5' biotinylated primers offer the option to exploit avidin-coupled dye reactions as well as the immobilization of PCR products via an avidin *(14)* or streptavidin bridge *(27)* to a matrix. The use of fluorescent primers offers a variety of applications in qualitative and quantitative analysis, e.g., nonisotopic sequencing reactions *(28,29)* or color coupling tests *(30)*.

A perfect homology to the target sequence is not a *conditio sine qua non* for primers, although a high degree of sequence homology is necessary. Taking into account that the enzyme is working downstream from the 3' end of the primer, only the last three bases adjacent to this region need to be correctly base paired, in order to initiate polymerization *(31)*. The 5' region of the primer is less sensitive to

base mismatches. Therefore, new restriction sites can easily be introduced into an amplification product *(32)*. It is also possible to use primers elongated by a stretch of 40 GC bp *(33)*. Successful amplification of sequences derived from known amino acid sequences utilizing primers with mixed bases at one position has been reported *(34,35)*.

3.2.3. Template

Normally, very much less than 1 μg DNA is used for an experiment. The sensitivity of the method allows a dilution of the sample down to the theoretical limit of 1 molecule/test, even for routine diagnostics *(36,37)*. The requirements for DNA purity are rather modest. Intact cells may be used; they are lysed during the first denaturation step of the PCR procedure. Some authors recommend a proteinase K digest in the presence of nonionic detergents in order to render the DNA more accessible *(38)*. However, there is no advantage to employing higher concentrations of template; this will make a contamination with inhibitors more likely. Without any DNA isolation, amplification of bacterial sequences is feasible directly by incubating 1×10^3 cells at 95°C for 10 min *(39)*. At higher cell concentrations, the PCR is inhibited. Hematin (hemin) has also been found to inhibit the PCR at levels as low as 0.8 μ*M* *(38)*. Therefore, only nuclear fractions from whole blood samples should be analyzed by the PCR. The sensitivity of the technique allows the amplification of degraded DNA from very old, but well-preserved samples *(40)*. A preceding reverse transcription in the same buffer system permits the direct amplification of RNA sequences *(41,42)* (Section 4.2.).

3.2.4. Taq *Polymerase*

Taq polymerase 1 is commercially available from several manufacturers at a concentration of 5 U/μL. One unit corresponds to 50 fmol of enzyme *(8)*. *Taq* polymerase lots from different suppliers may lead to considerably different quantitative results. One to four units per 100-μL reaction aliquot are sufficient to amplify a sequence to yield amounts of DNA detectable by ethidium bromide staining. Larger enzyme quantities will provoke the amplification of unspecific products *(3)*. If the denaturation temperature for double-stranded DNA is optimal, enzyme damage can be minimized, and enzyme activity will be present in the reaction mixture even after 50 or more cycles. The enzyme dilution should always be made up freshly by pipeting from stock solution into incubation buffer (Section 3.2.6.).

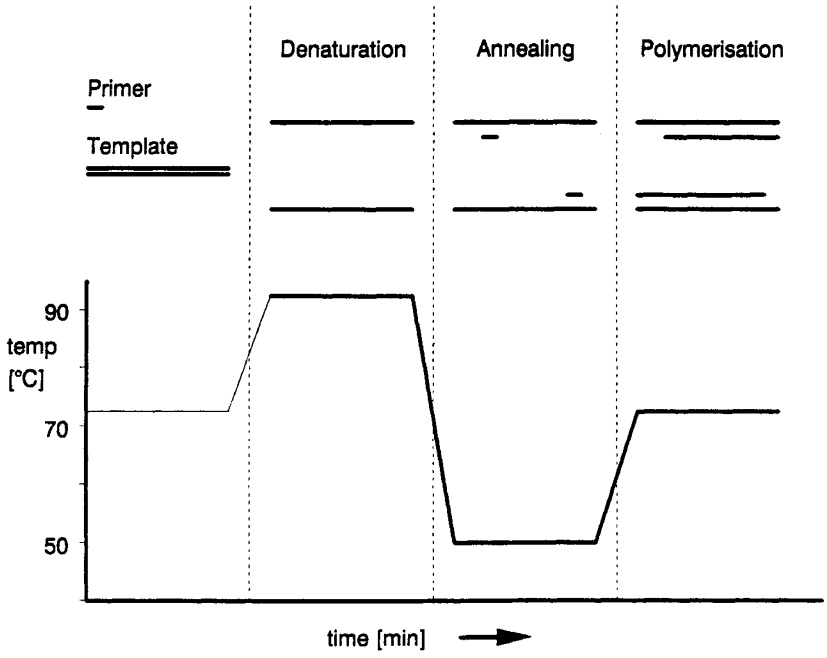


Fig. 5. Typical PCR temperature protocol.

3.2.5. Thermocycler

The PCR reaction requires a repeated cycling through temperature profiles (Fig. 5). Sample volumes of 20–100 µL are used in small reaction tubes, which will facilitate the heat transfer from the incubation device. Temperature control is usually achieved by automated thermocyclers:

- Water bath devices;
- Heating blocks with conventional electrical heating and liquid cooling systems; or
- Heating based on other technologies, e.g., Peltier elements or fan heating.

It is the main feature of the devices to generate fast temperature shifts in the range of 30–95°C. Deviations concerning block uniformity or precision in repeated cycles of more than 1°C are not tolerable. A prerequisite for allele-specific amplification experiments is a highly reproducible annealing temperature.

With regard to the adaptation of published protocols to a specific thermocycler, it is very important to consider the way the actual temperature is determined and regulated. Three principles can be distinguished:

- Temperature assessment in the heating block: (a) the time reading starts, when the desired value in the block is reached or (b) the time reading starts independently of the effective temperature of the block;
- Temperature assessment in a reference vial: the time reading starts, when the rated temperature in the reference is reached; and
- Temperature assessment in water baths: the time reading is independent of the actual sample temperature.

Different from the programmed temperature profile of the heating block, the actual profile in the reaction vial is distinctly more shallow, since the heat transfer through the material of the vessel is fairly bad.

For a successful PCR experiment, a set of defined temperatures must be adjusted:

- To denature the template, 90–95°C must be reached for a short interval. Prolonged incubation times at this temperature will lead to severe enzyme losses by degradation. It was demonstrated *(43)* that complete strand separation is not a necessary condition for annealing, rather, a partial melting of the strands is sufficient to permit hybridization of the primers.
- The annealing temperature must be attained for a few seconds only. At a primer concentration of 1 μ*M*, the half-life time of the template–primer hybrid is not more than 0.4 s at a temperature 20°C below the T_m value *(44)*.
- As mentioned earlier, the incorporation rate of the enzyme is 60–150 nucleotides/s at 70–75°C; thus, a few seconds are adequate for the amplification of sequences shorter than 1000 nucleotides. An incubation time of 1 min/1000 nucleotides of sequence length is a good rule of thumb, when sequences exceeding a length of 1000 nucleotides or more are to be amplified *(24)*.

3.2.6. Standard PCR Protocol

Following this protocol, a product band should be detectable from a single copy gene. Prepare the following solution in a final volume of 30–100 μL:

- 1 μg Genomic DNA as template;
- 1 μ*M* Primers, 20–25 bp in length, flanking a sequence of 100–1000 bp;
- 10 m*M* Tris-HCl, pH 8.4;
- 50 m*M* KCl;
- 1.5 m*M* $MgCl_2$;
- 0.01% Gelatine; and
- 200 μ*M* of each dNTP.

Top this mixture carefully with paraffin oil, incubate at 95°C for 3 min, add 1–4 U of *Taq* polymerase/100 µL reaction vol, and repeat 30 cycles (temperature measurement in the heating block) of:

- 1 min at 74°C;
- 1 min at 92°C; and
- 1 min at 50°C.

End the procedure with an incubation at 74°C for 3 min. Analyze a 10-µL aliquot on an agarose or PAGE gel; stain with ethidium bromide.

3.2.7. Reliable PCR Experiments

In clinical diagnosis, reliability is the most important criterion for the application of a test. The extreme sensitivity of the PCR requires high standards for the conditions of operation. Even minor contaminations with foreign DNA will render results uncertain. With this caveat in mind, some precautions are necessary to avert false positive results *(30)*.

- Separate all PCR materials and equipment from processes involved in DNA isolation.
- Aliquot all buffer, dNTP, and primer stock solutions.
- Pipet the DNA last.
- Always process controls in parallel: (a) a negative control with all components, but lacking the template, and (b) a positive control with a control template in a concentration similar to the template under investigation.
- Perform only the minimum number of amplification cycles necessary to detect the product. This is crucial for experiments in which the loss of a band (e.g., demonstration of the deletion of particular a sequence) has to be detected *(45)*.

Recently, it was reported *(46)* that irradiation at 254 nm for 5 min of the PCR mix lacking template and enzyme will reduce DNA contamination by a factor of 10^5.

4. PCR Applications

As an overview, the next paragraphs will deal with the application of the PCR. For more detailed information, *see* the literature cited *(47–49)*.

4.1. PCR as a Standard Method in Clinical Diagnosis

The scope of PCR application in clinical diagnosis ranges from prenatal diagnosis for sickle cell anemia *(2)* to postmortem analysis of

tissues of cancer metastases. We would like to restrict this chapter to some examples with characteristic features of PCR application in clinical diagnosis.

The reasons for the success of the PCR as a standard method in clinical diagnosis are its outstanding sensitivity, reliability, and versatility. Diagnostic screening for mutations by the PCR should always be complemented with a statistical analysis; the probability of false diagnosis is on the order of 1% for a single-locus analysis in single cells *(50)*.

The extreme sensitivity of the technique can trace one target sequence in a background of 10^5–10^6 cells *(3)*. The PCR makes affirmative statements on the genetic information of single somatic or sperm cells possible *(36)*. In a serial examination, 700 different sperms were typified, demonstrating a frequency of recombination of 0.16 in one individual *(51)*. An extension of this test series could allow the detection of recombination frequencies <0.001; hence, data would be available that are more exact than information from linkage analysis of pedigree studies.

Multiple gene loci can be amplified in parallel. Six domains of the Duchenne Muscle Dystrophy (DMD) gene were amplified in a single experiment *(45)*. Thus, it is possible to monitor by PCR 70% of the deletions in the 700-kb DMD gene. A simple test for β-thalassemia was developed by combining the PCR amplification with subsequent dot blot hybridization to allele specific oligodeoxynucleotides *(52)*.

The PCR was established to amplify fragments up to 3000 bp to amounts detectable by staining *(24,53)*. The amplification of larger fragments is feasible, although with little efficiency: Hypervariable minisatellite alleles up to 5–10 kb were faithfully amplified *(15)*. Up to several micrograms of product can be achieved by the PCR, amounts sufficient as substrate in binding experiments, e.g., footprint analysis *(54)*.

4.2. PCR Applications in Combination with Nonradioactive Detection Techniques

The use of radioactive probes should be avoided in routine screening tests. Even DNA of low abundance can be analyzed by the PCR followed by restriction analysis of the products *(55)*. The definite sequence is enhanced to high copy numbers; therefore, radioactive detection systems are dispensable. Product identification can be easily done, for example, by hybridization to biotinylated oligodeoxynucleotides and subsequent detection via avidin-linked enzymes *(56)*.

Allele-specific oligodeoxynucleotide (ASO) PCR *(57)* is based on the fact that only fully complementary primers will initiate the reaction. Therefore, amplification will be prohibited, if the complementary sequence is deleted or mutated. A primer specific for the mutated sequence is used as positive control. Point mutations, deletions, and insertions in the α-1-antitrypsin gene were monitored *(58)* with primers harboring a 3' base mismatch with respect to the unmutated gene. One 3' mismatch is necessary and sufficient to ensure reliable detection. Primers bearing a mismatch located in the center of the primer sequence are not suitable for allele-specific hybridization.

A different technique to find point mutations is competitive oligodeoxynucleotide priming (COP)*(19)*. When the annealing of a primer mixture is performed at low stringency, the perfectly matched primer will be favored with a level of discrimination greater than 100:1. An extension of COP and ASO-PCR involves the use of oligodeoxynucleotides conjugated with a fluorescent dye marker; the products are analyzed in a fluorimeter, hence rendering the electrophoretic analysis obsolete *(30,59)*.

4.3. Application of Specifically Designed Primers

For the reason that a primer needs to be complementary only at its 3' end to the sequence to be analyzed, mismatched primers can be of advantage in introducing new restriction sites in the amplified product. The product can be traced by restriction enzyme analysis even at high background *(55)*. Also, the fragment produced can be cloned readily, one more benefit of this method *(32)*.

Sometimes it is difficult to detect a single mutation in longer fragments using denaturation gradient gel electrophoresis (DGGE). This drawback can be overcome by attachment of a 40-bp GC-rich stretch to the primers *(60)*. The analysis of the PCR product, melting at a higher temperature, will enhance the sensitivity of this method.

Mixed oligodeoxynucleotide primers designed on the basis of the amino acid sequence of urate oxidase were used successfully to clone the gene for this enzyme *(34)*. If the primers crosshybridize to repetitive sequences, even an optimal adaptation of the annealing temperature will not eliminate unspecific bands. In this case, the amplification products can be diluted and reamplified with a second set of primers *(61,62)*. The use of dc^7GTP in a ratio of 3:1 over GTP is recommended to improve the PCR of sequences with distinct secondary structures *(8,63)*.

Normally, it is a prerequisite for the PCR to know the sequences flanking the region to be amplified. Different methods were developed to circumvent this drawback: Circularization of DNA fragments and subsequent cleavage in one known site result in a constellation, where an unknown region is flanked by sequenced regions *(64)*.

Normally, the PCR is applied to generate well-defined sequences. Whole genome PCR involves the ligation of oligodeoxynucleotides prior to the amplification, and constitutes a pool of material for further manipulation *(65)*. With this technique, possible pitfalls in microcloning of dissected chromosome fragments could also be bypassed. After the ligation of oligodeoxynucleotides to restriction fragments, the DNA was amplified and cloned in high yields *(66,67)*.

Ligation of known sequences to chemically cleaved DNA templates and subsequent PCR amplification allows a specific in vivo footprint analysis *(68)* as well as detection of methylation patterns *(69)*. The principles of this ligation-mediated PCR application are depicted in Fig. 6. Random cleavage sites are generated in genomic DNA using Maxam-Gilbert chemistry. The strands are denatured, and a gene-specific primer is annealed and extended with DNA polymerase. Oligodeoxynucleotide linkers are blunt-end ligated to the double-stranded products. The linker-ligated fragments are amplified with the PCR and analyzed on sequencing gels.

4.4. Amplification of mRNA

The amplification of RNA is performed on a cDNA template produced by reverse transcriptase. A small and not very efficient reverse transcriptase activity was detected in *Taq* polymerase *(70)*. Molony Murine Leukemia Virus (MMLV) reverse transcriptase requires buffer conditions similar to *Taq* polymerase. Therefore, after addition of poly dT or sequence-specific primers and reverse transcriptase, the PCR can be executed directly in the same vial *(71)*. Avian Myeloblastosis Virus (AMV) reverse transcriptase and *Taq* polymerase may also be added together to the RNA; the reactions will be started subsequently by choosing the proper reaction temperatures *(42)*.

Starting with cDNA, fragments as long as 1.6 kb could be cloned *(72)*. The extreme sensitivity of the PCR allows semiquantitative recordings on specific mRNA content of single cells *(73)*. Eleven out of 12 patients suffering from chronic myelogenous leukemia (CML) who underwent bone marrow transplantation still produced transcripts specific for CML.

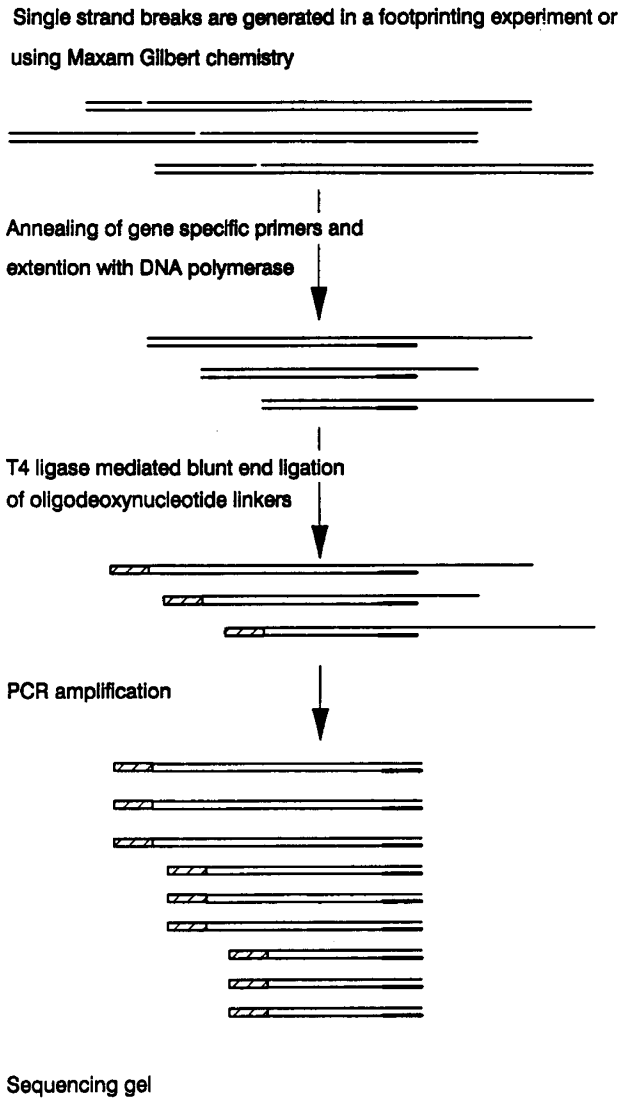


Fig. 6. Application of PCR to in vivo footprinting and detection of methylation pattern.

Although the patients were cytogenetically inconspicious even 3–6 yr after surgery, CML-specific transcripts could be detected by PCR of the reverse transcribed RNA *(74)*.

If the boundaries of genes are not exactly known, ligation of specific sequences to the cDNA will render an amplification feasible. Primers complementary to the ligated oligodeoxynucleotides permit the PCR

of unknown sequences. This method is termed "anchor PCR" *(75–77)*. An alternative to cDNA synthesis is a combination of consecutive RNA–DNA hybridization, S1 nuclease digest, and PCR. This technique was applied successfully to trace CML-specific transcripts *(78)*.

4.5. Sequencing of PCR Products

Sequencing PCR-amplified products is a time-saving procedure in the analysis of genes. No time-consuming cloning of DNA is necessary. On the contrary, errors made by *Taq* polymerase will lead to false sequence information, if cloned PCR products are sequenced *(3)*. Sequencing revealed that, in 22 clones created from amplified DNA, none had a faultless sequence *(10)*. In a direct sequencing reaction of amplified DNA, on the average, these shortcomings will not become visible. Double-strand sequencing of a single copy gene was utilized to typify mutations in the globin gene *(79)*.

In order to increase the efficiency ot the sequencing reaction, diverse strategies were developed aimed at yielding single-stranded DNA *(80)*. When biased primer concentrations are used, a phase of normal PCR will precede a reaction producing dominantly single-stranded DNA *(8,25)*. A well-balanced ratio of the primer concentration is crucial. If the primer concentration is too low, too little DNA will be available for single-strand synthesis. On the other hand, the ratio of single-stranded to double-stranded product will be inferior at primer concentrations that are too high. In experiments working with asymmetric primer concentrations, the PCR efficiency decreases to a factor of 1.7 as compared to amplifications with stoichiometrlc concentrations.

Single-stranded DNA can be recovered by affinity chromatographic purification after amplification with biotinylated primers *(27)*. In an elegant experiment *(81)*, a phage promoter was attached 5' onto at least one of the PCR primers. The amplified segments were transcribed into RNA and sequenced with reverse transcriptase. Another way to generate single-stranded DNA involves a 5'-phosphorylated primer and phage λ exonuclease, which specifically attacks double-stranded DNA only if there is a 5' terminal phosphate *(82)*. With the dideoxy sequencing method, single copy genes can easily be sequenced with fluorescent dye-labeled primers *(29)*. As an alternative, Maxam-Gilbert sequencing of PCR products *(83)* or sequencing after incorporation of dNTPαS during the PCR is possible *(18)*. When fluorescent dye-labeled primers are utilized, the cleavage products can be sequenced directly *(28)*.

In the plateau phase of the reaction, the processivity of *Taq* polymerase decreases at excess template concentrations. This phenomenon is exploited in a very simple DNA sequencing method, where phosphothioates are incorporated into the newly synthesized DNA. The product is digested with snake venom phosphodiesterase and analyzed on a sequencing gel. No product purification is necessary applying this technique *(84)*.

4.6. PCR and Site-Directed Mutagenesis

High yields of the desired mutations can be produced, when mismatched primers are employed in the PCR *(85)*. Deletions, insertions, and mutations were created in plasmids with the support of mismatched primers and PCR *(86)*. Furthermore, these primers may generate a new restriction site, serving as an option to introduce amplified fragments with the cassette method.

Amplification of the whole plasmid offers the opportunity to introduce mutations in the absence of restriction sites *(53)*; partial deletions in the plasmid are also feasible *(87)*. In 70% of the amplified plasmids, 3' additions of one base not present in the template were observed. Prior to cloning, these protruding ends had to be removed by the Klenow fragment of DNA polymerase to restore the gene sequence *(53)*. The PCR has proved useful in the precise gene fusion at any chosen location *(88)*. Sequencing of mutant clones after mutagenesis is compulsory, given the relatively high error rate of *Taq* polymerase (1.1×10^{-4} *[6]*, 2×10^{-4} *[3,9]*). At least 20% of the clones derived from a 250-bp fragment amplified 10^6-fold will contain one error in the sequence (cf Fig. 7).

4.7. Quantitative Interpretation of PCR Results

Standard PCR in itself is not a quantitative method, when the plateau is reached. For a quantitative interpretation of PCR experiments, low cycle numbers and template concentrations below enzyme concentrations are optimal, i.e., conditions where amplification is not restricted by limiting parameters prevailing in the plateau phase. Hybridization of ^{32}P-labeled oligodeoxynucleotides revealed a linear correlation between template DNA and detected signal, provided that 500–4000 copies of DNA as template and 30 amplification cycles were used *(17)*. In order to increase the incorporation rate, the dNTP concentration was lowered to 10 μ*M* (Section 3.2.1.).

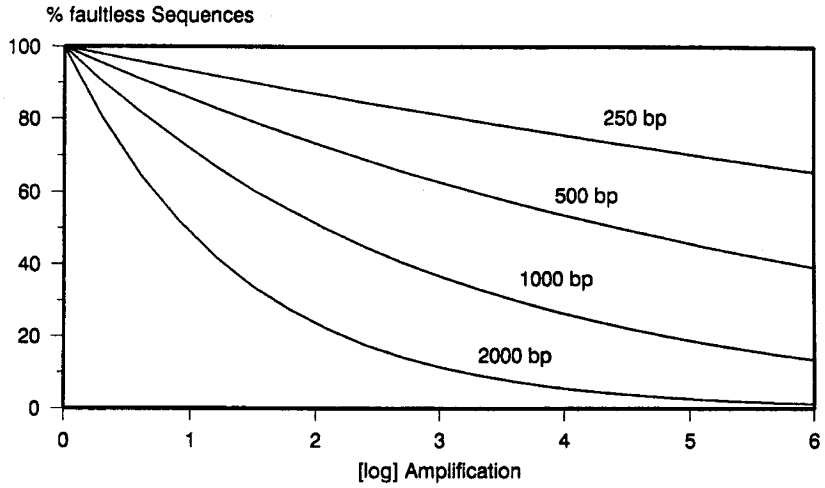


Fig. 7. Dependence of error rate and fragment length.

In a technique referred to as differential PCR, alterations of genes can be monitored when two genes are coamplified in one reaction vessel. In this method, the primers are competing for the limiting enzyme; therefore, gene copy numbers can be estimated semiquantitatively even in the plateau phase of the PCR *(89)*.

Quantitation of mRNA transcription in different cell types was successfully performed using Southern hybridization to an internal standard and to the amplified product *(90)*. Monitoring antiviral therapy benefits from the exceptional sensitivity of the PCR as well *(91)*.

A test system where standards identical to the target sequence, except for one restriction site, are coamplified provides an elegant way for the absolute determination of mRNA amounts. Consequently, both signals are enhanced to the same extent; variations in the reaction efficiency will not influence the result. After completion of the PCR, the standard is cleaved and compared with the target sequence *(43)*. An alternative to blotting techniques is the utilization of biotinylated primers; the PCR products are immobilized on avidin agarose and hybridized to specific probes *(14)*.

5. Sequencing with *Taq* Polymerase

Although *Taq* polymerase is considered often to be only useful for the PCR, last but not least, it should be pointed out that the enzyme is ideal for dideoxy sequencing, the reasons being:

- The enzyme is highly processive;
- Secondary structures are suppressed at the optimal temperature range of 70–80°C;
- dc^{7}GTP is accepted as building block; this nucleotide analog suppresses secondary structures as well; and
- *Taq* polymerase is less expensive than T7 polymerase and its modified form, sequenase.

In asymmetric PCR *(8)* and in single-strand sequencing in the M13 system *(92)*, ddNTPs are accepted with different efficiencies compared to dNTPs. Optimal are dNTP/ddNTP ratios of 1:32 (dATP/ddATP), 1:16 (dCTP/ddCTP), 1:6 (dGTP/ddGTP), and 1:48 (dTTP/ddTTP) *(8)*.

References

1. Chien, A., Edgar, D. B., and Trela, J. M. (1976) Deoxyribonucleic acid polymerase from the extreme thermophile thermus aquaticus. *J. Bacteriol.* **127,** 1550–1557.
2. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. (1985) Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230,** 1350–1354.
3. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239,** 487–491.
4. Koshland, D. E. (1989) The molecule of the year. *Science* **246,** 1541.
5. Lawyer, F. C., Stoffel, S., Saiki, R. K., Myambo, K., Drummond, R., and Gelfand, D. H. (1989) Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J. Biol. Chem.* **264,** 6427–6437.
6. Tindall, K. R. and Kunkel, T. A. (1988) Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* **27,** 6008–6013.
7. Gelfand, D. H. (1989) *Thermus aquaticus* DNA polymerase, in *Current Communications in Molecular Biology: Polymerase Chain Reaction* (Erlich, H. A., Gibbs, R., and Kazazian, H. H., eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 11–17.
8. Innis, M. A., Myambo, K. B., Geifand, D. H., and Brow, M. A. D. (1988) DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Natl. Acad. Sci. USA* **85,** 9436–9440.
9. Keohavong, P. and Thilly, W. G. (1989) Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci. USA* **86,** 9253–9257.
10. Dunning, A. M., Talmud, P., and Humphries, S. E. (1988) Errors in the polymerase chain reaction. *Nucleic Acids Res.* **16,** 10,393.

11. Mullis, K. B. (1989) The polymerase chain reaction: why it works, in *Current Communications in Molecular Biology: Polymerase Chain Reaction* (Erlich, H. A., Gibbs, R., and Kazazian, H. H., eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 237–243.
12. Erlich, H. A., Gelfand, D. H., and Saiki, R. K. (1988) Specific DNA amplification. *Nature* **331,** 461,462.
13. Landgraf, A. (1991) Qualitative und quantitative DNA Analytik mit der Polymerase Ketten Reaktion. PhD thesis, University of Hannover, Germany.
14. Syvänen, A. C., Bengtström, M., Tenhunen, J., and Söderlund, H. (1988) Quantification of polymerase chain reaction products by affinity based hybrid collection. *Nucleic Acids Res.* **16,** 11,327–11,338.
15. Jeffreys, A. J., Wilson, V., Neumann, R., and Keyte, J. (1988) Amplification of human minisatellites by the polymerase chain reaction: Towards DNA fingerprinting of single cells. *Nucleic Acids Res.* **16,** 10,953–10,972.
16. Williams, J. F. (1989) Optimization strategies for the polymerase chain reaction. *BioTechniques* **7,** 762–769.
17. Abbott, M. A., Poiesz, B. J., Byrne, B. C., Kwok, S., Sninsky, J. J., and Ehrlich, G. D. (1988) Enzymatic gene amplification: Qualitative and quantitative methods for detecting proviral DNA amplified in vitro. *J. Infect. Dis.* **158,** 1158–1169.
18. Nakamaye, K. L., Gish, G., Eckstein, F., and Vosberg, H. P. (1988) Direct sequencing of polymerase chain reaction amplified DNA fragments through the incorporation of deoxynucleoside alpha-thiotriphosphates. *Nucleic Acids Res.* **16,** 9947–9960.
19. Gibbs, R. A., Nguyen, P. N., and Caskey, C. T. (1989) Detection of single DNA base differences by competitive oligonucleotide priming. *Nucleic Acids Res.* **17,** 2437–2448.
20. Winnacker, E. L. (1987) *From Genes to Clones.* VCH Verlagsgesellschaft, Weinheim, Germany.
21. Rychlik, W. and Rhoads, R. E. (1989) A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and *in vitro* amplification of DNA. *Nucleic Acids Res.* **17,** 8543–8551.
22. Atkinson, T. and Smith, M. (1984) Solid phase synthesis of oligonucleotides by the phosphite-triester method, in *Oligonucleotide Synthesis—a Practical Course* (Gait, M. J., ed.), IRL, Oxford, pp. 56–84.
23. Pingoud, A., Fliess, A., and Pingoud, V. (1989) HPLC of oligonucleotides, in *HPLC of macromolecules—a Practical Approach* (Oliver, R. W. A., ed.), IRL, Oxford, pp. 183–208.
24. Saiki, R. K. (1998) Optimization of the polymerase chain reaction, in *Current Communications in Molecular Biology: Polymerase Chain Reaction* (Erlich, H. A., Gibbs, R., and Kazazian, H. H., eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 25–29.
25. Gyllensten, U. B., and Erlich, H. A. (1988) Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci. USA* **85,** 7652–7656.

26. Smith, L. M., Sanders, J. Z., Kaiser, R. J., Huges, P., Chris, D., Heiner, C., Kent, S. B. H., and Hood, L. E. (1986) Fluorescence detection in automated DNA sequence analysis. *Nature* **321,** 674–679.
27. Mitchell, L. G. and Merril, C. R. (1989) Affinity generation of single-stranded DNA for dideoxy sequencing following the polymerase chain reaction. *Anal. Biochem.* **178,** 239–242.
28. Silver, J. and Keerikatte, V. (1989) Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus. *J. Virol.* **63,** 1924–1928.
29. McBride, L. J., Koepf, S. M., Gibbs, R. A., Salser, W., Mayrand, P. E., Hunkapiller, M. W., and Kronick, M. N. (1989) Automated DNA sequencing methods involving polymerase chain reaction. *Clin. Chem.* **35,** 2196–2201.
30. Chehab, F. F. and Kan, Y. W. (1989) Detection of specific DNA sequences by fluorescence amplification: A color complementation assay. *Proc. Natl. Acad. Sci. USA* **86,** 9178–9182.
31. Sommer, R. and Tautz, D. (1989) Minimal homology requirements for PCR primers. *Nucleic Acids Res.* **17,** 6749.
32. Scharf, S. J., Horn, G. T., and Erlich, H. A. (1986 Direct cloning and sequence analysis of enzymatically amplified sequences. *Science* **233,** 676–678.
33. Sarkar, G. and Sommer, S. S. (1988) RNA amplification with transcript sequencing (RAWTS). *Nucleic Acids Res.* **16,** 5197.
34. Lee, C. C., Wu, X., Gibbs, R. A., Cook, R. G., Muzny, D. M., and Caskey, C. T. (1988) Generation of cDNA probes directed by amino acid sequence: cloning of urate oxidase. *Science* **239,** 1288–1291.
35. Nunberg, J. H., Wright, D. K., Cole, G. E., Petrovskis, E. A., Post, L. E., Compton, T., and Gilbert, J. H. (1989) Identification of the thymidine kinase gene of feline herpesvirus: Use of degenerate oligonucleotides in the polymerase chain reaction to isolate herpes virus gene homologs. *J. Virol.* **63,** 3240–3249.
36. Li, H., Gyllensten, U. B., Cui, X., Saiki, R. K., Erlich, H. A., and Arnheim, N. (1988) Amplification and analysis of DNA sequences in single human sperm and diploid cells. *Nature* **335,** 414–417.
37. Cui, X., Li, H., Goradia, T. M., Lange, K., Kazaziar, H. H., Jr., Galas, D., and Arnheim, N. (1989) Single-sperm typing: Determination of genetic distance between the ^{G}gamma-globin and parathyroid hormone loci by using the polymerase chain reaction and allele specific oligomerization. *Proc. Natl. Acad. Sci. USA* **86,** 9389–9393.
38. Higuchi, R. (1989) Simple and rapid preparation of samples for PCR, in *PCR Technology—Principles and Applications for DNA Amplification* (Erlich, H. A., ed.), Stockten, New York, pp. 31–38.
39. Olive, D. M. (1989) Detection of enterotoxigenic *Escherichia Coli* after polymerase chain reaction amplification with a thermostable DNA polymerase. *J. Clin. Microbiol.* **27,** 261–265.
40. Pääbo, S., Higuchi, R. G., and Wilson, A. C. (1989) Ancient DNA and the polymerase chain reaction. The emerging field of molecular archaeology. *J. Biol. Chem.* **264,** 9709–9712.

41. Sherman, D. R., Geliebter, J., and Cross, G. A. M. (1989) Rapid and simple amplification of a specific RNA sequence by the polymerase chain reaction. *Trends Genet.* **5,** 137.
42. Goblet, C., Prost, E., and Whalen, R. G. (1989) One-step amplification of transcripts in total RNA using the polymerase chain reaction. *Nucleic Acids Res.* **17,** 2144.
43. Becker-Andre, M. and Hahlbrock, K. (1989) Absolute mRNA quantification using the polymerase chain reaction (PCR). A novel approach by a PCR aided transcript mration assay (PATTY). *Nucleic Acids Res.* **17,** 9437–9446.
44. Hearst, J. E. (1988) A photochemical investigation of the dynamics of oligonucleotide hybridization. *Ann. Rev. Phys. Chem.* 291–315.
45. Chamberlain, J. S., Gibbs, R. A., Ranier, J. E., Nguyen, P. N., and Caskey, C. T. (1988) Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res.* **16,** 11,141–11,156.
46. Sarkar, G. and Sommer, S. S. (1990) Shedding light on PCR contamination. *Nature* **343,** 27.
47. Innis, M. A., Gelfand, D. H., and White,T. J., (eds.) (1989) *PCR Protocols and Applications: A Laboratory Manual.* Academic, New York.
48. Erlich, H. A., Gibbs, R. A., and Kazazian, H. H., Jr. (eds.) (1989) *Current Communications in Molecular Biology—Polymerase Chain Reaction.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
49. Erlich, H. A. (ed.) (1989) *PCR Technology—Principles and Applications of DNA Technologies.* Stockten, New York.
50. Krawczak, M., Reiss, J., Schmidtke, J., and Rösler, U. (1989) Polymerase chain reaction: Replication errors and reliability ot gene diagnosis. *Nucleic Acids Res.* **17,** 2197–2202.
51. Sommerfelt, H., Grewal, H. M. S., and Bhan, M. K. (1990) Simplified and accurate nonradioactive polynucleotide gene probe assay for identification of enterotoxigenic *Escherichia coli. J. Clin. Microbiol.* **28,** 49–54.
52. Saiki, R. K., Chang, C. A., Levenson, C. H., Warren, T. C., Boehm, C. D., Kazazian, J., and Erlich, H. A. (1988) Diagnosis of sickle cell anemia and beta-thalassemia with enzymatically amplified DNA and nonradioactive allele-specific oligonucleotide probes. *N. Engl. J. Med.* **319,** 537–541.
53. Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G., and Galas, D. J. (1989) A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Res.* **17,** 6545–6551.
54. Higuchi, R., Krummel, B., and Saiki, R. K (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interaction. *Nucleic Acids Res.* **16,** 7351–7367.
55. Haliassos, A., Chomel, J. C., Grandjouan, S., Kruh, J., Kaplan, J. C., and Kitzis, A. (1989) Detection of minority point mutations by modified PCR technique: A new approach for a sensitive diagnosis of tumor-progression markers. *Nucleic Acids Res.* **17,** 8093–8095.
56. Saiki, R. K., Bugawan, L., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1986) Analysis of enzymatically amplified beta-globin and HLA-DQ-alpha DNA with allele—specific oligonucleotide probes. *Nature* **324,** 163–166.

57. Bos, J. L., Fearon, E. R., Hamilton, S. R., Verlan-de Vries, M., van Boom, J. H., van der Eb, A. H., and Vogelstein, B. (1987) Prevalence of ras gene mutations in human colorectal cancers. *Nature* **327,** 293–297.
58. Okayama, H., Curiel, D. T., Brantly, M. L., Holmes, M. D., and Crystal, R. G. (1989) Rapid, nonradioactive detection of mutations in the human genome by allele-specific amplification. *J. Lab. Clin. Med.* **114,** 105–113.
59. Chehab, F. F. and Kan, Y. W. (1990) Detection of sickle cell anaemia mutation by colour DNA amplification. *Lancet* **335,** 15–17.
60. Sheffield, V. C., Cox, D. R., Lerman, L. S., and Myers, R. M. (1989) Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proc. Natl. Acad. Sci. USA* **86,** 232–236.
61. Haqqi, T. M., Sarkar, G., David, C. S., and Sommer, S. S. (1988) Specific amplification with PCR of a refractory segment of genomic DNA. *Nucleic Acids Res.* **16,** 11,844.
62. Wrischnik, L. A., Higuchi, R. G., Stoneking, M., Erlich, H. A., Arnheim, N., and Wilson, A. C. (1987) Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. *Nucleic Acids Res.* **15,** 529–542.
63. McConlogue, L., Brow, M. A. D., and Innis, M. A. (1988) Structure-independent DNA amplification by PCR using 7-deaza-2'-deoxyguanosine. *Nucleic Acids Res.* **16,** 9869.
64. Triglia, T., Peterson, M. G., and Kemp, D. J. (1988) A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. *Nucleic Acids Res.* **16,** 8186.
65. Kinzler, K. W. and Vogelstein, B. (1989) Whole genome PCR: Application to the identification of sequences bound by gene regulatory proteins. *Nucleic Acids Res.* **17,** 3645–3653.
66. Lüdecke, H. J., Senger, G., Claussen, U., and Horsthemke, B. (1989) Cloning defined region of the human genome by microdissection of banded chromosomes and enzymatic amplification. *Nature* **338,** 348–350.
67. Saunders, R. D. C., Glover, D. M., Ashburner, M., Siden-Kiamos, I., Louis, C., Monastirioti, M., Savakis, C., and Kafatos, F. (1989) PCR amplification of DNA microdissected from a single polytene chromosome band: A comparison with conventional microcloning. *Nucleic Acids Res.* **17,** 9027–9037.
68. Mueller, P. R. and Wold, B. (1989) In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **246,** 780–786.
69. Pfeifer, G. P., Steigerwald, S. D., Mueller, P. R., Wold, B., and Riggs, A. D. (1989) Genomic sequencing and methylation analysis by ligation mediated PCR. *Science* **246,** 810–813.
70. Jones, M. D. and Foulkes, N. S. (1989) Reverse transcription of mRNA by *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res.* **17,** 8387,8388.
71. Doherty, P. J., Huesca-Contreras, M., Dosch, H. M., and Pan, S. (1989) Rapid amplification of complementary DNA from small amounts of unfractionated RNA. *Anal. Biochem.* **177,** 7–10.

72. Lomeli, H., Tyagi, S., Pritchard, C. G., Lizardi, P. M., and Kramer, F. R. (1989) Quantitative assays based on the use of replicatable hybridization probes. *Clin. Chem.* **35,** 1826–1831.
73. Rappolee, D. A., Brenner, C. A., Schuitz, D. M., Mark, D., and Werb, Z. (1988) Developmental expression of PDGF, TGF-α, and TGF-β genes in preinplantation mouse embryos. *Science* **241,** 1823–1825.
74. Gabert, J., Thuret, I., Lafage, M., Carcassonne, Y., Maraninchi, D., and Mannoni, P. (1989) Detection of residual *bcr/abl* translocation by polymerase chain reaction in chronic myeloid leukaemia patients after bone-marrow transplantation. *Lancet* **2,** 1125–1128.
75. Loh, E. Y., Elliott, J. F., Cwirila, S., Lanier, L. L., and Davis, M. M. (1989) Polymerase chain reaction with single-sided specifity: analysis of t-cell receptor chain. *Science* **243,** 217–220.
76. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85,** 8998–9002.
77. Belyavsky, A., Vinogradova, T., and Rajewsky, K. (1989) PCR-based cDNA library construction: General cDNA libraries at the level of a few cells. *Nucleic Acids Res.* **17,** 2919–2932.
78. Lee, M. S., Chang, K. S., Freireich, E. J., Kantarian, H. M., Talpaz, M., Trujillo, J. M., and Stass, S. A. (1988) Detection of minimal residual bcr/abl transcripts by a modified polymerase chain reaction. *Blood* **72,** 893–897.
79. Wong, C., Dowling, C. E., Saiki, R. K., Higuchi, R. G., Erlich, H. A., and Kazazian, H. H., Jr. (1987) Characterization of beta-thalassaemia mutations using direct genomic sequencing of amplified single copy DNA. *Nature* **330,** 384–386.
80. Gyllensten, U. B. (1989) PCR and DNA sequencing. *BioTechniques* **7,** 700–709.
81. Stoflet, E. S., Koeberl, D. D., Sarkar, G., and Sommer, S. S. (1988) Genomic amplification with transcript sequencing. *Science* **239,** 491–494.
82. Higuchi, R. G. and Ochman, H. (1989) Production of single-stranded DNA templates by exonuclease digestion following the polymerase chain reaction. *Nucleic Acids Res.* **17,** 5865.
83. Cariello, N. F., Scott, J. K., Kat, A. G., Thilly, W. G., and Keohavong, P. (1988) Resolution of a missense mutant in human genomic DNA by denaturating gradient gel electrophoresis and direct sequencing using in vitro DNA amplification: HPRT. *Am. J. Hum. Genet.* **42,** 726–734
84. Olsen, D. B. and Eckstein, F. (1989) Incomplete primer extension during *in vitro* DNA amplification catalyzed by Taq polymerase; exploitation for DNA sequencing. *Nucleic Acids Res.* **17,** 9613–9620.
85. Rochlitz, C. F., Scott, G. K., Dodson, J. M., and Benz, C. C. (1988) Use of the polymerase chain reaction technique to create base-specific ras oncogene mutations. *DNA* **12,** 515–519.
86. Vallette, F., Mege, E., Reiss, A., and Adesnik, M. (1989) Construction of mutant and chimeric genes using the polymerase chain reaction. *Nucleic Acids Res.* **17,** 723–733.

87. Mole, S. E., Iggo, R. D., and Lane, D. P. (1989) Using the polymerase chain reaction to modify expression plasmids for epitope mapping. *Nucleic Acids Res.* **17,** 3319.
88. Yon, J. and Fried, M. (1989) Precise gene fusion by PCR. *Nucleic Acids Res.* **17,** 4895.
89. Frye, R. A., Benz, C. C., and Liu, E. (1989) Detection of amplified oncogenes by differential polymerase chain reaction. *Oncogene* **4,** 1153–1157.
90. Chelly, J., Kaplan, J. C., Maire, P., Gautron, S., and Kahn, A. (1988) Transcription of the dystrophin gene in human muscle and non-muscle tissues. *Nature* **333,** 858–860.
91. Larzul, D., Brechot, C., and Guesdon, J.-L. (1989) A highly sensitive detection method for hepatitis b viral DNA sequences: monitoring of antiviral therapy, in *Molecular Probes Technology and Medical Applications* (Albertini, A., Paoletti, R., and Reisfeld, R. A., eds.), Raven, New York, pp. 97–104.
92. Peterson, M. G. (1988) DNA sequencing using *Taq* polymerase. *Nucleic Acids Res.* **16,** 10,915.

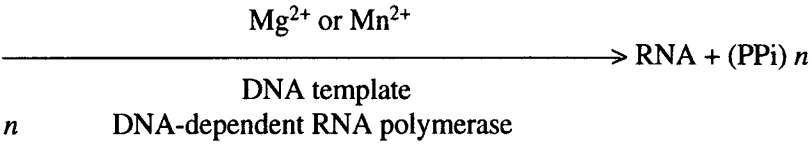
Chapter 5

Eukaryotic Nuclear RNA Polymerases (EC 2.7.7.6)

Deborah A. Cook and Robert J. Slater

1. Introduction

DNA-dependent RNA polymerases (ribonucleoside triphosphate: RNA nucleotidyltransferase, EC 2.7.7.6) catalyze the synthesis of RNA from ribonucleoside triphosphates in the presence of a DNA template and divalent cation *(1)*.



The RNA molecules are synthesized complementary and antiparallel to one of the DNA strands in a 5' to 3' direction. The ribonucleotides are covalently joined together by internucleoside 3' to 5' phosphodiester bonds with concomitant release of inorganic pyrophosphate.

In order for the polymerase to do this, it must "recognize" the beginning of the DNA sequence to be transcribed within the double-stranded template; insert the correct nucleotide residue into each position, as determined by the template sequence; carry out synthesis, so that RNA is synthesized from beginning to end as a consequence of one polymerase binding event; and recognize termination signals in order to end RNA synthesis at the appropriate point. To achieve all this, the polymerase must be able to recognize regulatory sites on the DNA

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

template and interact with protein factors that modulate the activity of the enzyme. Several reviews have appeared on eukaryotic RNA polymerases *(2–4)*.

2. Biological Function

Bacterial RNA polymerase containing a single apo enzyme is responsible for all the synthesis of RNA in the cell. There are approx 7000 RNA polymerase molecules present in an *Escherichia coli* cell, and many of these are involved in transcription, although the actual number depends on the growth conditions *(5)*. The RNA polymerase best characterized is that of *E. coli*, but its structure is similar in all other bacteria studied. The holoenzyme has a mol mass of about 480,000 Da and is made up of five subunits, α, β, β', σ, and ω, there being two copies of α and one each of the others: $\alpha_2\beta\beta'\sigma\omega$. The holoenzyme can be separated into two components, the apo enzyme ($\alpha_2\beta\beta'\omega$) and the sigma factor (σ). The sigma subunit of the *E. coli* polymerase has a specific role as an initiation factor for transcription: It enables the enzyme to find consensus promoter sequences. Only the holoenzyme can initiate transcription; the sigma "factor" is then released, leaving the core enzyme to undertake elongation. Thus, the core enzyme has the ability to synthesize RNA on a DNA template; its affinity for DNA is the result of electrostatic attractions, but is nonspecific. It is the sigma factor that ensures correct transcription initiation involving the stable binding of the RNA polymerase to the promoter *(5)*. Eukaryotes, however, contain multiple DNA-dependent RNA polymerases. Those that transcribe nuclear genomes are generally referred to as RNA polymerases I, II, and III transcribing rRNA, mRNA and tRNA, and 5SrRNA, respectively (Table 1).

2.1. Properties

Eukaryotic RNA polymerases transcribe both double-stranded and single-stranded DNA templates in vitro, and single-stranded templates generally promote higher rates of transcription *(6)*. RNA polymerases of plants and animals are capable of initiating transcription at nicks in the DNA template in vitro *(3,8)*. Purified eukaryotic RNA polymerases are capable of some degree of selective binding and initiation of RNA chains on deproteinized DNA templates *(9,10)*, although this is nonspecific, and in no case has the selectivity been correlated with a promoter or initiation site. From a number of studies with animal cell-free

Table 1
Properties of Eukaryotic RNA Polymerases[a]

	Eukaryotic RNA polymerases		
	I	II	III
Cellular location	Nucleolus	Nucleoplasm	Nucleoplasm
Transcription product	rRNA	hnRNA snRNA	tRNA 5SrRNA
α-Amanitin sensitivity (μg/mL required for 50% inhibition)	Nonsensitive	Very sensitive (0.01–0.05)	Moderately sensitive (5–1000)
Subunit structure	Two large subunits 100,000–240,000 and several small subunits 10,000–95,000		
Ionic strength optima $(NH_4)_2SO_4$ M			
Soluble enzyme	0.03–0.05	0.025–0.1	0.04–0.01, 0.18–0.3
Chromatin-bound enzyme	0.05–0.1	0.25–0.5	0.05–0.1
Mn^{2+}/Mg^{2+} activity ratio	~5:1	~7:1	~3:1
Poly (dA/dT)/DNA activity ratio	2:3	1:2	1:3
Chromatographic elution $(NH_4)_2SO_4$ M			
DEAE-Sephadex	~0.1	~0.2	0.2–0.3
DEAE-cellulose	~0.1	~0.2	~0.1

[a]Taken from refs. *1,4,6,7.*

extracts, it is apparent that additional factors are required to obtain faithful transcription in vitro, so the enzymes are not often used in their pure state because they lack specificity. Eukaryotic RNA polymerases differ in subunit structure, type of genome transcribed, chromatographic properties, divalent cation, ionic strength optima, and α-amanitin sensitivity. The poisonous mushroom *Amanita phalloides* produces α-amanitin, a bicyclic octopeptide that blocks the elongation step of transcription by direct interaction with RNA polymerases that differ in their sensitivity (Table 1). RNA polymerase I is insensitive to the toxin, but is responsive to actinomycin D, which binds to DNA preventing elongation. RNA polymerase II is very sensitive and RNA polymerase III moderately so. Thus, the various RNA polymerases I, II, and III are readily distinguished (Table 1). Intact cells are not freely permeable to α-amanitin; therefore, it is necessary to use a whole-cell or nuclear extract.

2.2. Ion Requirement

The requirement for a divalent cation by RNA polymerases indicates binding of nucleotides to the enzyme as metal chelates, for example: MgATP *(11)*. The optimum divalent cation concentration is generally in the range of 5–10 m*M* Mg^{2+} or 1–2 m*M* Mn^{2+} *(2)*. In addition to the above requirements, RNA polymerase activity is stimulated by the inclusion of monovalent cations, such as $(NH_4)_2SO_4$, NH_4Cl, KCl, or NaCl. This is especially evident when engaged RNA polymerases are assayed in isolated nuclei or chromatin. With isolated plant or animal nuclei, RNA polymerase I and III activity is most active at 0.05–0.10*M* $(NH_4)_2SO_4$, whereas RNA polymerase II activity is optimized at 0.25–0.50*M* *(12–15)*.

2.3. Structure

The nuclear RNA polymerases have a complex subunit structure. Each has a mol mass of approx 500,000 Da and is made up of two large subunits, each having a mol mass >100,000 Da, which seem to be characteristic for each class of enzyme, and a number of smaller subunits, with mol wt from 10,000 Da to just less than 100,000 Da *(2,4)*. The situation is further complicated by the phenomenon of microheterogeneity within RNA polymerase classes *(7)*.

3. Transcription Factors

The basic components required for efficient and accurately regulated eukaryotic transcription initiation include two types of DNA elements known as promoters and upstream regulatory sequences; two sets of proteins known as general transcription factors and specific or regulatory proteins, and the RNA polymerase enzyme. Transcription can occur when the eukaryotic RNA polymerase recognizes a preexisting DNA–protein complex. RNA polymerase II has four general transcription factors associated with it; these have been partially purified and designated TFIIA, B, D, and E *(16–19)*. TFIIA is required for the efficient interaction of TFIID with the promoter element known as the TATA box *(20,21)*, which is found in nearly all eukaryotic promoters, with a consensus sequence "TATAAA" and located approx 25–30 bp upstream of the transcription start site in mammalian promoters *(22–25)*. It has been shown that TFIIE specifi-

cally interacts with RNA polymerase II and TFIIB *(26,27)*. These transcription factors range in mol mass from 25,000 Da for yeast *(28)* to 76,000 Da for TFIIE *(24)*.

Although promoters themselves are sufficient to direct basal level gene expression, specific transcription requires the assistance of additional transcription factors *(29)*. These sequence-specific DNA binding proteins are directly involved in the recognition of mRNA transcription initiation in higher organisms; for example, CAT transcription factor (CTF) is responsible for selective recognition of eukaryotic promoters that contain the sequence CCAAT *(30,31)*, whereas the Sp1 protein binding regions contain one or more GC boxes (GGGCGG), which can be present in either orientation with respect to transcription. Not all GC boxes bind Sp1 with equal affinity, and sequences outside the core hexanucleotide seem to modulate the efficiency of binding (for review, *see* refs. *25,32*).

In HeLa cell factor, termed the major late transcription factor or upstream stimulatory factor, MLTF/USF binds to an upstream consensus sequence GGCCACGTGACC, which functions in a bidirectional manner. MLTF/USF may act synergistically with TFIID, since binding of MLTF/USF enhances binding of TFIID *(33,34)*. The binding sequence occurs in a number of cellular promoters, for example, mouse and rat γ-fibrinogen *(35)*. Another promoter-specific transcription factor is the heat shock transcription factor (HSTF). All major heat shock genes contain multiple HSE (heat shock regulatory elements) with the consensus sequence C—GAA—TCC—G *(36)*. In HeLa cells, heat shock causes more HSTF to bind to the HSE, whereas in yeast, HSTF binds constitutively and heat shock causes HSTF to become phosphorylated *(37)*.

3.1. In Vitro Transcription Systems

The development of methods for preparing soluble whole-cell and nuclear extracts from mammalian cells *(38–41)* has led to a dramatic increase in our understanding of the mechanistic details of complex biochemical reactions, such as transcription, replication, and RNA processing.

In vitro transcription systems were partially developed to study the biochemical mechanisms of DNA–protein interaction, gene expression, and regulation. To date, several cell-free transcription systems have been developed, the first by Wu in 1978 *(42)*; which demon-

strated accurate transcription of exogenous DNA by polymerase III, and others such as those developed by Weil et al. *(38)*; Manley et al. *(39)*; Dignam et al. *(40)*; and Zerivitz and Akusjarvi *(41)*.

The major advantage of soluble cell-free transcription systems is that they can be fractionated by standard techniques in order to identify those components involved in the transcription process *(43)*. The purified enzymes alone do not direct selective and accurate transcription from purified DNA templates, but they appear to initiate and terminate at a large number of sites, which may not be relevant to transcription in vivo *(3,4)*. The most extensive use of in vitro transcription systems to date has been to study the structure of promoter sites and the transcription factors needed to bring about selective transcription *(44)*.

4. Experimental Procedures

4.1. TCA Precipitation Assay

This procedure can be used as an estimate of nucleoside triphosphate incorporation into trichloroacetic acid (TCA) precipitable products following an in vitro transcription reaction. ^{3}H-UTP is usually the radiolabel of choice. The reaction (as detailed in Section 4.2., but for the substitution of ^{32}P-labeled GTP with 1.0 µCi of ^{3}H-UTP) is stopped by the addition of 10% TCA. Acid-insoluble material is then collected on glass fiber disks, washed twice with 5% TCA (with sodium pyrophosphate to reduce nonspecific binding of the radiolabel), and finally with alcohol to remove unincorporated radiolabel. The filters are dried and the radioactivity estimated by using a scintillation counter; hence, an estimation of nucleotide incorporation can be made *(1)*.

4.2. Run-Off Transcription

RNA polymerase does not accurately terminate transcription in vitro. One procedure therefore is to use as template a cloned gene fragment that includes the promoter, but is truncated downstream. Using the "run-off" assay system, the whole-cell or nuclear lysate, DNA template, and radiolabeled (^{32}P) nucleotide are incubated at 30°C for 60 min. RNA is synthesized from the truncated DNA template and analyzed electrophoretically to detect the species of RNA with the predicted length.

A typical 50-µL in vitro transcription reaction would contain the following:

- 20–30 µL Whole-cell or nuclear extract;
- 1–2.5 µg Template DNA;
- 500 µ*M* each of ATP, CTP, and UTP;
- 4 m*M* Creatine phosphate;
- 50 µ*M* Unlabeled nucleoside triphosphate (GTP)
- 10 µCi Labeled nucleoside triphosphate (^{32}P-GTP specific activity).

The volume is made up to 50 µL with sterile distilled water and incubated at 30°C for 1 h. Suitable controls include omission of template DNA or unlabeled nucleoside triphosphates, and addition of RNase or α-amanitin.

This system uses template DNA that has been cleaved by a restriction enzyme cutting downstream from its putative start site. RNA polymerases that transcribe this DNA will stop or "fall off" when they reach the end of the DNA. If a substantial number of enzymes initiate transcription at the same site, then a population of molecules will migrate as a single band during gel electrophoresis. If DNA segments cleaved by different restriction enzymes are used as templates in separate reaction mixtures, the transcription start site can be deduced by comparison of the sizes of RNAs produced; this can then be used for promoter mapping.

4.3. Primer Extension

This method is commonly used to measure the precise 5' end of transcripts, based on the fact that reverse transcriptase can transcribe RNA into DNA. Therefore, a single-strand oligonucleotide of DNA primer hybridized to RNA can be extended until the end of the RNA is reached. Primer extension utilizes a primer of specific sequence to determine the amount and length of a specific RNA. This sort of analysis has been widely used to characterize mRNA made both in vivo and in vitro *(45)*.

Primer extension is used to map the 5' termini of an RNA transcript; hence, the cap site of transcription can be determined. In addition, precursors and processing intermediates of mRNA can be detected. A short known fragment (30–80 bp) of DNA corresponding to a sequence near to the anticipated cap site of the gene of interest is labeled at the 5' terminus; this is termed the primer (a synthetic oligonucleotide or restriction fragment). The test RNA, which has been digested with RNase-free DNase to remove template DNA, can be synthesized in an in vitro transcription

system (with the omission of the radiolabeled triphosphate, as described in Section 4.2.) and hybridized with an excess of the single-stranded DNA primer. The primer and the RNA are then incubated at 85°C for 10 min to denature the nucleic acids, and annealed by incubating at a suitable temperature, for example, 55°C for 8–12 h. The exact temperature needs to be optimized by preliminary experiments. The length and GC content of the primer are used to determine the temperature of annealing; the standard temperature of 55°C is suitable for a 20 base oligonucleotide sequence with a GC content of approx 50% *(46,47)*. The primer:RNA hybrid is then dissolved in reverse transcriptase buffer, approx 50 U of reverse transcriptase added to extend the primer to produce a cDNA copy complementary to the RNA template and incubated at 42°C for 90 min. If analyzing more than 10 µg of RNA in order to prevent distortion of the sequencing gel, a RNase digestion is needed at this stage.When the reverse transcriptase reaches the end of the sequence (the cap site), it "falls off." The length of the resulting end-labeled cDNA is measured by denaturing gel electrophoresis usually on a DNA sequencing gel, hence, the sequence prior to the cap site can also be determined.

Single-stranded DNA primers are preferred for experiments, since they eliminate the formation of DNA:DNA hybrids and can be designed to hybridize to specific sequences of the mRNA.

4.4. Nuclease S1 RNA Mapping

The enzyme S1 nuclease degrades single-stranded DNA to yield 5' phosphoryl mono or oligonucleotide *(48)*. Primer extension cannot be used to identify the 3' end of a transcript. Nuclease S1 mapping is therefore the method of choice for this purpose. The procedure was originally described by Berk and Sharp in 1977 *(49)*; it has been used to map the location of 5' and 3' termini of mRNA or DNA templates, and locate 5' and 3' splice points *(50)*.

End-labeled RNA or DNA probes derived from various segments of the genomic DNA are hybridized to mRNA favoring DNA:RNA hybrid formation *(51)*. Products of the hybridization are digested with nuclease S1 under conditions favoring single-stranded nucleic acid degradation only. Then the products are analyzed by gel electrophoresis to yield important quantitative and qualitative information about the mRNA structure. The basic method is as follows: Probe and RNA are

mixed together, resuspended in hybridization buffer (40 m*M* PIPES [pH 6.4], 1 m*M* EDTA [pH 8.0], 0.4*M* NaCl and 50% formamide), and incubated at 85°C for 10 min to denature the nucleic acids. Hybridization follows for 12–16 h, but again the temperature is dependent on the GC content of the DNA. Nuclease S1 at 100–1000 U/mL and single-stranded carrier DNA at 20 μg/mL are then added for single-strand digestion, but the temperature depends on the degree of digestion required. Loops of DNA will be degraded at 20°C, and 37–45°C will digest single-stranded regions of DNA:RNA hybrids. The reactions are cooled to 0°C and nuclease stop mixture (containing carrier RNA at 50 μg/cm^3) added. The products can then be phenol:chloroform extracted, and the nuclease S1-resistant hybrids analyzed by alkaline or neutral agarose gel electrophoresis. In the case where a radiolabeled DNA probe has been chosen, this can be analyzed by electrophoresis through a polyacrylamide/urea gel and an autoradiogram established. A detailed description of nuclease-S1 mapping was presented in a review by Favalaro et al. in 1980 *(52)*.

5. Cell and Nuclear Extracts

5.1. Soluble Cell Extract Transcription System for RNA Polymerase II

A soluble transcription system is a cell extract that will faithfully transcribe an exogenous DNA as template. For this, the extract should contain all the necessary factors needed for the correct initiation and elongation of the RNA chain, and be low in nuclease activity to prevent degradation of the exogenous DNA template or transcription products during the incubation. Establishing soluble systems provides a starting point for the separation, isolation, and analysis of those factors required (in addition to RNA polymerase) for accurate transcription in vitro *(45)*.

To date, many eukaryotic and viral genes have been sequenced, and putative controlling sequences have been deduced. These observations can be verified by in vitro manipulation of the DNA template and then using the DNA in an in vitro transcription system to test the effect.

Weil et al. in 1979 *(38)* prepared a high-speed supernatant (S100) cellular extract from cultured human cells. This S100 extract can be used for transcription experiments once supplemented with purified

RNA polymerase II. RNA polymerase II from wheat germ does not function in this system, indicating that it is a combined effect of RNA polymerase II and specific factors that mediate accurate and selective transcription. Fractionated S100 does contain multiple factors required for accurate transcription by purified RNA polymerase II *(16,22)*.

5.2. HeLa Whole Cell Lysate (39)

The extract is prepared by gentle lysis of cells with a Dounce homogenizer. Ammonium sulfate is added to approx 20% saturation, and the nuclei lysed. Cell debris and broken nuclei are removed by centrifugation at 175,000*g* for 3 h. Proteins in the supernatant are resuspended and dialyzed for 12 h. The dialysate is centrifuged to remove insoluble material, and supernatant removed to form the extract. This can be stored in small aliquots with repeated thawing and refreezing for up to 1 yr at –70°C. The extract normally contains 6–30 mg/mL of protein.

The rate of RNA polymerase II elongation is approx 300 nucleotides/min approx 1/10th that in vivo *(53)*. The in vitro temperature of 30°C should not be increased to 37°C, otherwise RNase degradation may become a problem *(39)*. The reaction mixture is detailed in Section 4.2.

Each whole cell extract must be characterized by DNA and extract titration, since there is a threshold DNA concentration below which no accurate transcription occurs and an inhibitory effect at high DNA concentrations. Specific transcription can be obtained when the final concentration of protein in the extract is in the range 6–18 mg/mL. The extract contains high levels of 18S and 28S rRNA. Therefore, following extraction, it is possible that only 50–75% of the sample may be loaded onto the gel to avoid overloading. Since the extract is made from whole cells, weak promoters may not be as easily detected, although this does depend on the level of background nonspecific transcription occurring.

5.3. Nuclear Extracts from Human Cells (40)

The extract is prepared by isolating nuclei in hypotonic buffer, lysis of cells with a Dounce homogenizer, extraction of the nuclei at 0.42*M* NaCl, high-speed centrifugation to clear the supernatant, and concentration of the extracted nucleic acids and various transcription factors by ammonium sulfate precipitation. Since the nuclei are isolated at very low ionic strength (0.01*M* KCl), retention of nuclear components is facilitated and contaminating cytoplasmic proteins are discarded in the supernatant following pelleting of the nuclei. Extraction at 0.42*M*

NaCl appears to release all of the nuclear components required for specific transcription, including RNA polymerase II, without releasing tightly bound nuclear proteins that can inhibit specific transcription *(40)*.

5.4. Nuclear Extract (41)

This method uses lysolecithin (a natural membrane lipid—lysophosphatidylcholine) to disrupt plasma membranes and requires no detergents or lysis with a Dounce homogenizer. Lysolecithin nuclear extracts are said to be competent for RNA polymerase II and III transcription, DNA replication, pre-mRNA splicing, and sequence-specific DNA–protein binding. Nuclear extracts can be prepared on a small scale (10^7 cells), as well as for preparative purposes by this method.

Lysolecithin is highly toxic and cells lyse within 90 s. An extract made from HeLa cells at a density of $4–6 \times 10^5$/mL provides a suitable nuclear extract for in vitro transcription analysis. Using HeLa suspensions at 10^8 cells/mL, 100% lysis can be obtained with a concentration of lysolecithin at 300–800 µg/mL. In the protocol *(41)*, a concentration of 400 µg/mL was chosen.

References

1. Slater, R. J. (1987) *In vitro* transcription, in *Techniques in Molecular Biology*, vol. 2 (Walker, J. M. and Gaastra, W., eds.), Groom Helm, London and Sidney, pp. 203–227.
2. Jacob, S. T. (1973) Mammalian RNA polymerases. *Prog. Nucleic Acids Res. Mol. Biol.* **13,** 93–126.
3. Chambon, P. (1975) Eukaryotic nuclear RNA polymerases. *Ann. Rev. Biochem.* **44,** 613–638.
4. Roeder, R. G. (1976) Eukaryotic nuclear RNA polymerases, in *RNA Polymerase* (Losick, R. and Chamberlin, M., eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 285–329.
5. Chamberlin, M. (1982) Bacterial DNA-dependent RNA polymerases, in *The Enzymes*, 3rd ed., vol. 15B (Boyer, P., ed.), Academic, New York, pp. 61–108.
6. Guilfoyle, T. J. (1983) DNA-dependent RNA polymerases of plants and lower eukaryotes, in *Enzymes of Nucleic Acid Synthesis and Modification*, vol. 2 (Jacob, S. T., ed.), CRC, Cleveland, OH, pp. 1–41.
7. Beebee, T. J. C. and Butterworth, P. H. W. (1981) Eukaryotic DNA-dependent RNA polymerases: An evaluation of their role in the regulation of gene expression, in *Eukaryotic Gene Regulation*, vol. 2 (Kolodny, G. M., ed.), CRC, Cleveland, OH, pp. 1–56.
8. Chandler, D. W. and Gralla, J. (1980) Specific binding and protection of form II SV40 deoxyribonucleic acid by ribonucleic acid polymerase II from wheat germ. *Biochemistry* **19,** 1604–1612.

9. Seidman, S., Surzycki, S. J., Delorbe, W., and Gussin, G. N. (1979) Interaction between wheat germ RNA polymerase II and adenovirus 2 DNA. Evidence for two types of stable binary complexes. *Biochemistry* **18,** 3363–3371.
10. Grellet, F., Cooke, R., Teissere, M., Delseny, M., Xech, J., and Penon, P. (1981) Electron microscopic mapping of wheat germ RNA polymerase II binding sites on cloned CaMV DNA. *Nucleic Acids Res.* **9(16),** 3927–3939.
11. Mildvan, A. S. and Loeb, L. A. (1979). The role of metal ions in the mechanism of DNA and RNA polymerases. *CRC. Crit. Rev. Biochem.* **6,** 219–244.
12. Chen, Y. M., Lin, C. Y., Chang, H., Guilfoyle, T. J., and Key, J. L. (1975) Isolation and properties of nuclei from control and auxin treated soybean hypocotyl. *Plant Physiol.* **56,** 78–82.
13. Lin, C. Y., Guilfoyle, T. J., Chen, Y. M., Nagao, R. T., and Key, J. L. (1975) Isolation of nucleoli and localisation of ribonucleic acid polymerase I from soybean hypocotyl. *Plant Physiol.* **56,** 850–852.
14. Guilfoyle, T. J. and Key, J. L. (1977) Purification and characterization of soybean DNA-dependent RNA polymerases and the modulation of their activities during development, in *Nucleic Acids and Protein Synthesis in Plants* (Bogorad, L. and Weil, J., eds.), Plenum, New York, pp. 37–63.
15. Guilfoyle, T. J. (1980) Transcription of the cauliflower mosaic virus genome in isolated nuclei from turnip leaves. *Virology* **107,** 71–80.
16. Matsui, T., Segall, J., Weil, A., and Roeder, R. G. (1980) Multiple factors for accurate initiation of transcription by purified RNA polymerase II. *J. Biol. Chem.* **255(24),** 11,992–11,996.
17. Samuels, M., Fire, A., and Sharp, P. A. (1982) Separation and characterization of factors mediating accurate transcription by RNA polymerase II. *J. Biol. Chem.* **257,** 14,419–14,427.
18. Dignam, J. D., Martin, P. L., Shastry, B. S., and Roeder, R. G. (1983) Eukaryotic gene transcription with purified components. *Methods Enzymol.* **104,** 582–598.
19. Sawadogo, M. and Roeder, R. G. (1985) Factors involved in specific transcription by human RNA polymerase II: Analysis by a rapid and quantitative in vitro assay. *Proc. Natl. Acad. Sci. USA* **82,** 4394–4398.
20. Sawadogo, M. and Roeder, R. G. (1985) Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* **43,** 165–175.
21. Nakajima, N., Horikoshi, M., and Roeder, R. G. (1988) Factors involved in specific transcription by mammalian RNA polymerase II. Purification, genetic specificity and TATA box-promoter interaction of TFIID. *Cell Biol.* **8,** 4028–4040.
22. Davison, B. L., Egly, J., Mulvihill, E. R., and Chambon, P. (1983) Formation of stable preinitiation complexes between class B transcription factors and promoter sequences. *Nature* **301,** 680–686.
23. Fire, A., Samuels, M., and Sharp, P. A. (1984) Interactions between RNA polymerase II factors and template leading to accurate transcription. *J. Biol. Chem.* **259,** 2509–2516.
24. Grosschedl, R. and Birnstiel, M. L. (1980) Spacer DNA sequences upstream of the T-A-T-A-A-A-T-A sequence are essential for promotion of H2A histone gene transcription in vivo. *Proc. Natl. Acad. Sci. USA* **77,** 7102–7106.

25. Dynan, W. S. and Tjian, R. (1985) Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. *Nature* **316,** 774–778.
26. Zheng, X-M., Moncollin, V., Egly, J. M., and Chambon, P. (1987) A general transcription factor forms a stable complex with RNA polymerase II. *Cell* **50,** 361–368.
27. Reinberg, D. and Roeder, R. G. (1987) Factors involved in specific transcription by mammalian RNA polymerase II: Purification and functional analysis of factors IIB and IIE. *J. Biol. Chem.* **262,** 3310–3321.
28. Buratowski, S., Hahn, S., Sharp, P. A., and Guarente, L. (1988) Function of a yeast TATA element binding protein a mammalian transcription system. *Nature* **334,** 37–42.
29. Ptashne, M. (1986) Gene regulation by proteins acting nearby and at a distance. *Nature* **225,** 683–689.
30. Jones, K. A., Yamamoto, K. R., and Tjian, R. (1985) Two distinct transcription factors bind to the HSV thymidine kinase promoter *in vitro*. *Cell* **42,** 559–572.
31. Jones, K. A., Kadonaga, J. T., Rosenfield, P. J., Kelly, T. J., and Tjian, R. (1987) A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell* **48,** 79–89.
32. Kadonaga, J. T., Jones, K. A., and Tjian, R. (1986) Promoter-specific activation of RNA polymerase II transcription by Sp1. *TIBS* **11,** 20–23.
33. Sawadogo, M. and Roeder, R. G. (1985) Interaction of a gene specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* **43,** 165–169.
34. Carthew, R. W., Chodosh, L. A., and Sharp, P. A. (1985) An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. *Cell* **43,** 439–444.
35. Chodosh, L. A., Carthew, R. W., Morgan, J. G., Crabtree, G. R., and Sharp, P. A. (1987) The adenovirus major late transcription factor activates the rat γ-fibrinogen promoter. *Science* **238,** 684–691.
36. Pelham, H. R. B. (1985) Activation of heat shock genes in eukaryotes. *Trends Genet.* **1,** 31–40.
37. Sorger, P. K., Lewis, M. J., and Pelham, H. R. B. (1987) Heat shock factor is regulated differently in yeast and HeLa cells. *Nature* **329,** 81–85.
38. Weil, P. A., Luse, D. S., Segall, J., and Roeder, R. G. (1979) Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* **18,** 469–484.
39. Manley, J. L., Fire, A., Cano, A., Sharp, P. A., and Gefter, M. L. (1980) DNA dependent transcription of adenovirus genes in a soluble whole cell extract. *Proc. Natl. Acad. Sci. USA* **77,** 3855–3859.
40. Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acid. Res.* **11,** 1475–1489.
41. Zerivitz, K. and Akusjarvi, G. (1989) An improved nuclear extract preparation method. *Gene Anal. Techn.* **6,** 101–109.
42. Wu, G-J. (1978) Adenovirus DNA-directed transcription of 5.5S rRNA *in vitro*. *Proc. Natl. Acad. Sci. USA* **75,** 2175–2179.

43. Heintz, N. and Roeder, R. (1982) Transcription of eukaryotic genes in soluble cell-free systems., in *Genetic Engineering Principals and Methods*, vol. 4 (Setlow, J. and Hollaener, A., eds.), Plenum, New York and London.
44. Manley, J. L. (1982) Transcription of mammalian genes in vitro, in *Genetic Engineering*, vol. 4 (Setlow, J. K. and Hollaender, A., eds.), Plenum, New York, pp. 37–56.
45. McKnight, S. L. and Kisbury, R. (1982) Transcription control signals of a eukaryotic protein coding gene. *Science* **217,** 316–324.
46. Saiki, R. K. (1989) Basic methodology, in *PCR Technology: Principles and Applications for DNA Amplification* (Erlich, H. A., ed.), Macmillan, Basingstoke, UK, pp. 7–38.
47. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
48. Vogt, V. M. (1973) Purification and further properties of single-stranded specific nucleases from *Aspergillus Oryzae. Eur. J. Biochem.* **33,** 192–198.
49. Berk, A. J. and Sharp, P. A. (1977) Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. *Cell* **12,** 721–728.
50. Weaver, R. F. and W[illegible]smann, C. (1979) Mapping by a modification of the Berk-Sharp procedure. 5' termini of 15s β-globin mRNA precursor and mature 10s β-globin mRNA have identical map coordinates. *Nucleic Acids Res.* **7,** 1175–1180.
51. Casey, J. and Davidson, N. (1977) Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucleic Acids Res.* **4,** 1539–1545.
52. Favalaro, J., Treisman, R., and Kamen, R. (1980) Transcription maps of polyoma virus specific RNA: Analysis by two-dimensional nuclease S1 gel mapping. *Methods Enzymol.* **65,** 718–720.
53. Manley, J. L., Fire, A., Samules, M., and Sharp, P. A. (1983) In vitro transcription: whole-cell extract. *Methods Enzymol.* **101,** 568–582.

CHAPTER 6

Reverse Transcriptase (EC 2.7.7.49)

The Use of Cloned Moloney Murine Leukemia Virus Reverse Transcriptase to Synthesize DNA from RNA

Gary F. Gerard and James M. D'Alessio

1. Introduction

The conversion of mRNA into cDNA is the essential first step in the study of eukaryotic cell products expressed from cloned genes. The key enzyme used first in this process, retroviral RNA-directed DNA polymerase (reverse transcriptase), catalyzes the synthesis of a DNA copy of an RNA template in the presence of a suitable primer. Reverse transcriptase (RT) was discovered in 1970 *(1,2)* and was first used to copy a eukaryotic cell mRNA in 1971 *(3–5)*; the first cDNA clones prepared using RT were reported in 1976 *(6)*. Until recently, the only enzyme available was purified from avian myeloblastosis virus (AMV). The overall quality and consistency of commercially available AMV RT preparations have improved dramatically in the last six years, although considerable differences in performance characteristics still exist among enzyme preparations from different commercial suppliers *(7)*. In addition, parameters for carrying out first-strand cDNA synthesis from poly(A)$^+$ mRNA populations with AMV RT have been optimized thoroughly *(8)*; standard optima exist for pH, nucleotide concentration, and monovalent and divalent cation concentration; also essential stimulatory additives (such as sodium pyrophosphate and spermidine-HCl) have been identified. Furthermore, conditions have

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

been established for proper handling and storage of the enzyme. Conditions for second-strand cDNA synthesis with first-strand product synthesized by AMV RT have also been optimized *(9,10)*.

In 1985, Moloney murine leukemia virus (M-MLV) RT was cloned, overexpressed, and purified from *E. coli (11–14)*. A series of reports have been published delineating optimal conditions for use of this alternative enzyme in first- and second-strand cDNA synthesis *(15–18)*. M-MLV RT has been used widely to synthesize cDNA and is frequently used instead of AMV RT for carrying out cDNA synthesis.

In 1989, a modified version of M-MLV RT lacking RNase H activity became available *(7,19)*. RT without interfering RNase H activity catalyzes more efficient synthesis of full-length cDNA than either native AMV or cloned M-MLV RT possessing RNase H activity *(7)*. The enzyme lacking RNase H activity is designated M-MLV H^- RT. We describe here some of the properties of both forms of M-MLV RT and protocols for using M-MLV H^- RT in first- and second-strand cDNA synthesis.

2. The Enzyme

2.1. Sources

Commercial preparations of purified M-MLV RT are derived from two sources *(12–14)*. Both proteins differ from the native enzyme in the amino acids at their amino and carboxy termini because of changes in sequence information introduced into the gene during the cloning procedure. The protein encoded by plasmid pB6B15.23 *(14)* has 25 additional amino acids fused at the amino terminus, is missing seven amino acids present at the carboxy terminus of the native enzyme, and has nine additional carboxy-terminal amino acids not found in the M-MLV *pol* gene. The RT encoded by pRT601 *(12,13)* has one additional amino acid (methionine) at its amino terminus, the 671 amino acids present in the native enzyme, and 12 additional amino acids at its carboxy terminus. The same number of molecules of each enzyme is needed to catalyze cDNA synthesis, despite their structural and apparent catalytic distinctions (*see* Table 1). Ten times as many units (*see* Section 2.2.) of pRT601 RT as pB6B15.23 RT are required to copy efficiently an mRNA population *(11)*. Not surprisingly, the specific activity of pRT601 RT (350,000 U/mg of protein) is 10 times greater than that of

Table 1
Comparison of Some Properties of the Commercially Available RNase H-containing M-MLV RTs

Plasmid encoding enzyme	Enzyme specific activity, U[a]/mg[b]	Units used typically in cDNA synthesis	Molar ratio of enzyme to mRNA[c]
pRT601	350,000	200	5:1
pB6B15.23	35,000	20	5:1

[a]Units are determined and defined as described in ref. *(20)*.
[b]Commercial preparations of both enzymes are nearly homogeneous.
[c]Assumes an average mRNA is 2 kb in length, and a typical reaction contains 1 µg.

pB6B15.23 RT (35,000–40,000 U/mg of protein) in the same unit assay *(14,19)*. Since commercial preparations of both proteins are nearly homogeneous, an equal number of molecules of each enzyme is used during cDNA synthesis.

M-MLV H⁻ RT is encoded by a modified p601 plasmid and includes the same first 497 amino acids in cloned M-MLV RT *(19)*, plus five additional amino acids at the carboxy terminus encoded by the *pol* gene read out of frame. The specific activity of M-MLV H⁻ RT is 100,000 U/mg of protein *(19)*.

2.2. Unit Assay

The same unit assay is used by all manufacturers to define the DNA polymerase activity of RT. Poly(A) • oligo$(dT)_{12-18}$ is used as the template-primer, and the unit is defined as nmol of dTMP incorporated in 10 min at 37°C *(20)*. The unit assay reaction conditions are not optimal for copying mRNA with either AMV, M-MLV RT, or M-MLV H⁻ RT.

2.3. Storage

Cloned M-MLV RT and M-MLV H⁻ RT are stable for long periods of time (>6 mo) when stored at –20°C in a buffer containing 0.1*M* NaCl, nonionic detergent, and 50% glycerol. Freezing the enzyme at –70°C does not affect the unit activity, but does decrease the functional activity (ability to copy long mRNAs). When stored in 50% glycerol-containing solutions, the enzymes should never be placed at temperatures at which 50% glycerol freezes.

2.4. Catalytic Properties

There are two major difficulties with RT-catalyzed cDNA synthesis: (1) mass yields of cDNA transcripts generally do not exceed 50%, and (2) many cDNA transcripts are less than full length. When AMV RT was first used to synthesize cDNA, these difficulties were caused in part by RNase contaminants in AMV RT preparations. However, commercial preparations of AMV RT are now much cleaner. The high level of expression of cloned M-MLV RT and M-MLV H^- RT in *E. coli* has made it possible to eliminate detectable RNase contamination from enzyme preparations. However, results still fall short of the ultimate objective of synthesizing one full-length cDNA copy of every mRNA in a population, and this is owing in part to the catalytic properties of all RTs.

2.4.1. DNA Polymerase

RT resembles other template-directed DNA polymerases. It catalyzes the synthesis of DNA in the presence of a preformed primer with a 3'-OH hydrogen bonded to a template polynucleotide, a divalent metal ion, and a mixture of all four dNTPs *(21)*. In contrast to most other DNA polymerases, RT can use either RNA or DNA to prime DNA synthesis from an RNA or DNA template. As a DNA polymerase, RT displays low processivity *(22,23)* and elongates DNA chains slowly *(23)*. RT tends to pause at certain sequence and secondary structure-specific domains in RNA *(7)*. As discussed below (*see* Section 2.4.2.), such pausing can prevent an mRNA molecule from being copied all the way to its 5' end.

There is no endogenous 3'–5' exonuclease editing function associated with RT *(21)*. In spite of this, the error rate of cloned M-MLV RT is low (1 mismatched residue in 30,000 incorporated; ref. *24*). When the RNase H activity of cloned M-MLV RT was deleted to produce M-MLV H^- RT, the error rate did not change *(19)*.

2.4.2. RNase H

The RNase H activity of native RT specifically degrades the RNA strand of RNA • DNA *(21)* and can act independently of the RT DNA polymerase activity *(25)*. The initiation of cDNA synthesis depends on the hybridization of a DNA primer to mRNA, usually at the poly(A) tail. When this occurs, the hybrid is not only a priming site for the polymerase activity of RT, but is also a substrate for the RT RNase H

activity as well *(7,26)*. This sets up a competition between the two activities, and the extent to which the RNase H activity destroys the hybrid prior to the initiation of polymerization determines the maximum number of priming events that can actually occur *(26)*. This reduces the yield of cDNA by effectively removing a portion of the mRNA from the reaction.

When RT is synthesizing cDNA, RT RNase H activity degrades the template that has already been copied because it is in hybrid form. If the scissions in the mRNA occur near the point of chain growth, the uncopied portion of the mRNA can dissociate from the transcriptional complex, resulting in termination of cDNA synthesis for that template. This problem is exacerbated by the tendency of RT to pause during reverse transcription *(7)* and results in truncated cDNA molecules.

When M-MLV H^- RT is used, RNA cleavage and its deleterious effects are eliminated, resulting in more efficient synthesis of full-length cDNA *(7)*. However, M-MLV H^- RT still tends to pause during polymerization. Thus, some RNA molecules are still not copied to their 5' ends by M-MLV H^- RT, even though the RNA remains intact during DNA synthesis. Other modifications to the RT polypeptide or addition of accessory proteins (e.g., a nucleic acid binding protein) will be required to reduce pausing and termination of transcription.

2.5. RNA Template

The quality of the mRNA template in a cDNA synthesis reaction dictates the maximum amount of sequence information that can be converted into cDNA. Thus, it is important to optimize the isolation of mRNA from a given biological source and to prevent adventitious introduction of RNases into a preparation that has been carefully rendered RNase-free. We recommend procedures that rely on chaotropic salts to inactivate RNases for extraction of RNA from any source *(27,28)*. The presence of the poly(A) tail at the 3' end of mRNA has been exploited to fractionate mRNA from a total RNA population by affinity chromatography on oligo(dT) cellulose *(29)*. Most cDNA libraries are made from poly(A)-selected mRNA. When the quantity of the RNA source is limiting and affinity chromatography is not feasible, it is possible to synthesize effectively cDNA from unfractionated mRNA. This is done routinely in preparation for amplifying cDNA by the polymerase chain reaction technique *(30–33)*. Both M-MLV RT

and M-MLV H⁻ RT are much less sensitive than AMV RT to inhibition by tRNA and rRNA (*16*; Gerard, G., unpublished), and well suited to copying unfractionated mRNA.

2.6. Primer

The primer for cDNA synthesis is usually oligo$(dT)_{12-18}$ present in molar excess relative to the 3' poly(A) tails of mRNA. This method of priming usually results in a cDNA library in which the 3' ends of RNAs are overrepresented because of the catalytic properties of RT described in Section 2.4.2. This bias can be partially overcome by using random hexamer primers instead of oligo(dT) or as a supplement to an oligo(dT)-primed reaction *(10,34)*. When random hexamers are used, the priming events are more evenly distributed over entire mRNA molecules. However, some sequence information corresponding to the 3' end of mRNA is lost as the result of randomizing primer location. In addition, rRNA in an RNA population is also copied when random hexamers are used.

The ratio of random hexamer primers to mRNA is critical. Figure 1 illustrates that there must be a balance between the desired average length and the mass yield of cDNA product. With M-MLV H⁻ RT, approx 10 primers/mRNA gives a reasonable yield without sacrificing product length. The ratio is higher with M-MLV RT, approx 40:1. When oligo(dT) or random primers are used to prime first-strand cDNA synthesis, the termini of the ensuing double-stranded cDNA are functionally identical, precluding directional cloning of the cDNA. All members of a directional library contain cDNA inserts cloned in a specific orientation relative to the transcriptional polarity of the original mRNA. Directional cloning is desirable when cDNA libraries will be screened by expression *(35,36)* and is essential in some strategies for generating subtracted, i.e., partial, cDNA libraries *(37–39)*.

cDNA can be cloned directionally by introducing two different restriction endonuclease recognition sites at its termini. This is most simply achieved by using a primer-adapter to initiate first-strand synthesis. A primer-adapter has the structure 5'-p$(dX)_n$-$(dT)_{15}$-OH-3', where $(dX)_n$ encodes sequence information for one or more restriction endonuclease sites. The primer-adapter introduces desired restriction endonuclease site(s) at the end of double-stranded cDNA corresponding to the 3' end of mRNA. Subsequent ligation of linkers or adapters to the DNA can be used to introduce a different restriction endonuclease site(s) at the opposite end of the DNA.

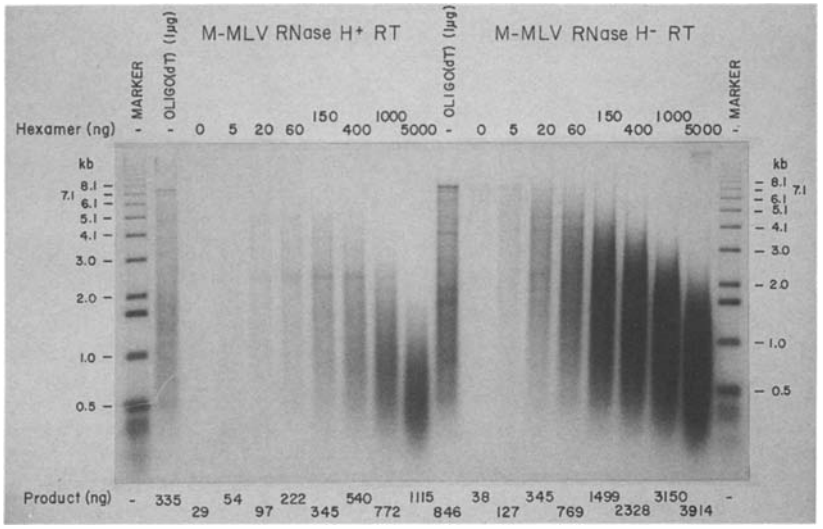


Fig. 1. Effect of random hexamer primer concentration on the mass and length of cDNA synthesized from mRNA by M-MLV RT and M-MLV H⁻ RT. An autoradiograph is shown of [^{32}P]cDNA synthesized from a mixture of 0.67 μg each of Ad-2 mRNA, Hela mRNA, and AMV 35S RNA with the indicated amounts of oligo(dT)$_{12-18}$ or random hexamers. DNA was fractionated on an alkaline 1.4% agarose gel.

2.7. Reaction Conditions

2.7.1. pH and Metal Ion Requirements

Table 2 gives the monovalent and divalent metal ion requirements of M-MLV RT and M-MLV H⁻ RT. The Mg^{2+} optima are quite sharp at 3 m*M*. KCl concentrations between 0 and 100 m*M* give product yields that are 75% or greater of those achieved at the optimum, 75 m*M*. Although the pH activity curve is broad, the yield of full-length product from a long RNA template (7.5 kb) for both enzymes is maximal at pH 8.3 at 22°C (at 37°C, pH 8.0).

2.7.2. dNTPs

The K_ms of M-MLV RT for dCTP and dATP are 31 and 24 μ*M*, respectively *(17)*. The K_ms of M-MLV H⁻ RT for dNTPs have not been determined, but are assumed to be similar. In spite of these relatively low values, high concentrations (≥500 m*M*) of dNTPs in a reaction mixture result in the most efficient synthesis of full-length cDNAs.

Table 2
Optima of M-MLV RT and M-MLV H⁻ RT[a]

$MgCl_2$, m*M*	% Activity		KCl, m*M*	% Activity		pH, 22°C	% Activity	
	H^+	H^-		H^+	H^-		H^+	H^-
1	68	49	0	75	93	7.2	90	ND[b]
2	88	93	50	95	100	7.5	95	ND[b]
3	100	100	75	100	98	7.9	97	ND[b]
4	88	ND[b]	100	86	83	8.3	100	ND[b]
5	83	77	125	72	66	8.5	98	ND[b]
7.5	62	ND[b]	150	54	44			
10	57	58	200	23	6			
15	ND	33						
20	20	15						

[a]Optima were determined from yields (incorporation after 60 min). Optima based on initial rates of incorporation are identical.

[b]ND is not determined.

2.7.3. Sodium Pyrophosphate

Sodium pyrophosphate at a concentration of 4 m*M* stimulates the synthesis of full-length cDNA from long mRNAs by AMV RT *(40,41)*, although there is disagreement on the mechanism by which sodium pyrophosphate produces this effect *(42,43)*. Sodium pyrophosphate also inhibits hairpin primed double-stranded cDNA synthesis by AMV. Sodium pyrophosphate inhibits the polymerase activity of M-MLV RTs and should not be added to reactions catalyzed by these enzymes.

2.7.4. Polyanions

Spermidine-HCl at 0.5 m*M* has been reported to stimulate the activity of AMV RT *(8)* and Rauscher murine leukemia virus RT *(44)*. However, spermine and spermidine inhibit both M-MLV RTs (*8*; Gerard, G., unpublished).

2.7.5. Bovine Serum Albumin and RNase Inhibitors

In many cDNA synthesis protocols, it is recommended that bovine serum albumin be added to reaction mixtures to stabilize RT. Bovine serum albumin has no effect on the activity of the M-MLV RTs. This may be because of the high level of RT protein added to a typical cDNA synthesis reaction.

The cloned M-MLV RTs are purified free of detectable RNase contamination. If precautions are taken to avoid adventitious introduction of RNase contaminants into reaction mixtures, RNase inhibitors are not necessary, and addition of such inhibitors can increase the chances of introducing contaminants into a reaction mixture.

2.7.6. Actinomycin D

Actinomycin D at 50–100 μg/mL inhibits the synthesis of double-stranded from single-stranded DNA *(45)*. It is added during first-strand cDNA synthesis to inhibit hairpin primed double-stranded DNA synthesis. It probably interacts with 3' single-stranded cDNA ends in such a way that DNA • DNA hybridization needed for hairpin formation is inhibited *(46)*. In first-strand cDNA reactions catalyzed by M-MLV RT, the proportion of cDNA molecules generated with exposed 3' single-stranded ends can be controlled to a large extent in the absence of actinomycin D by adjusting the ratio of enzyme to mRNA in the reaction. No actinomycin D is required in reactions catalyzed by M-MLV H^- RT, since first-strand cDNA product remains hybridized to intact mRNA template, and hairpin primed DNA synthesis cannot occur.

2.7.7. Inhibitors

Glycerol concentrations as high as 19% can be tolerated in RT reaction mixtures (Gerard, G., unpublished). Both M-MLV RTs are inhibited by phosphate, pyrophosphate, and polyamines.

2.7.8. Reducing Agent

Both M-MLV RTs require dithiothreitol (DTT) for optimal activity. A minimum concentration of 5 m*M* is necessary, although 10 m*M* are optimal. Aqueous, unbuffered solutions of 0.1*M* DTT are stable at 20°C, even after repeated freeze-thaw cycles. However, DTT is oxidized quite rapidly in buffered solutions above pH 7.5, so that reaction buffers that include DTT should be prepared freshly and discarded after use. No data are available on the effect of β-mercaptoethanol on M-MLV RTs.

2.7.9. Temperature

When either M-MLV RT or M-MLV H^- RT is used to synthesize first-strand cDNA, the cDNA libraries with the largest average insert size result when first-strand cDNA synthesis is performed at 37°C.

Table 3
Comparison of the Reaction Optima of M-MLV RT, M-MLV H⁻ RT, and AMV RT

Parameter	Enzyme		
	M-MLV RT	M-MLV H⁻ RT	AMV RT
pH		8.3	8.3 8.4
[KCl]	75 m*M*	75 m*M*	50–100 m*M*
[$MgCl_2$]	3 m*M*	3 m*M*	10 m*M*
Temperature	37°C	37°C (42°C)[a]	42°C
[DTT]	10 m*M*	10 m*M*	1 m*M*
[dNTPs]	0.5 m*M*	0.5 m*M*	1 m*M*
Other components	None	None	4 m*M* NaPPi 0.5 m*M* Spermidine-HCl

[a]M-MLV H⁻ RT functions equally well at 37 and 42°C.

M-MLV RT and M-MLV H⁻ RT can be used at incubation temperatures as high as 45 and 50°C, respectively.

2.7.10. Summary

Table 3 summarizes the reaction optima of M-MLV RT, M-MLV H⁻ RT, and AMV RT.

3. The Second-Strand Reaction

We have chosen to describe exclusively a protocol (*see* Section 4.) for a one-tube format double-stranded cDNA synthesis procedure because it is fast, maximizes DNA product recovery, and yields product that can be modified and cloned into most vectors. The second-strand cDNA is synthesized by nick-translation replacement of mRNA. First described by Okayama and Berg *(47)*, and later popularized by Gubler and Hoffman *(48)*, second-strand synthesis is catalyzed by *E. coli* DNA polymerase I in combination with *E. coli* RNase H and *E. coli* DNA ligase. Inclusion of *E. coli* DNA ligase in the reaction has been shown to improve the cloning efficiency of double-stranded cDNA synthesized from longer (>2 kb) mRNAs *(18)*.

We must make several additional comments concerning the second-strand reaction. The reaction is incubated at 16°C to reduce the tendency of DNA polymerase I at higher temperatures to strand displace rather than nick translate. The pH of the first-strand reaction is reduced in the

second-strand reaction from 8.3 to 7.5 to suppress the 5'–3' exonuclease activity of DNA polymerase I. At higher pH, the 5'–3' exonuclease will attack exposed 5' single-stranded DNA ends more readily; in particular, the 5' end of a primer-adapter used to prime first-strand cDNA synthesis can be degraded extensively enough so that the restriction endonuclease recognition sites are lost. This can lower the output of clones if the cloning strategy relies on cleavage of one of these sites. Furthermore, the amount of sequence information corresponding to the length of the 5'-most RNA primer in the reaction normally lost during synthesis will presumably be minimized by the decrease to pH 7.5 *(18)*. The minimum number of bases lost because of the combined action of RNase H and DNA polymerase I 5'–3' and 3'–5' exonuclease is approx 8 *(18)*. Finally, T4 DNA polymerase is added at the end of second-strand synthesis to ensure that the termini of the double-stranded cDNA are blunt.

4. Experimental Procedures

The following sections refer to conditions optimized for M-MLV H^- RT.

4.1. Laboratory Practices

It cannot be overemphasized that successful cDNA synthesis demands an RNase-free environment at all times. In general, this will require the same level of care used to maintain aseptic conditions when working with microorganisms. In addition, there are several guidelines that should be followed:

1. Never assume that anything is RNase-free except for sterile pipets, centrifuge tubes, culture tubes, and similar labware that is explicitly stated to be sterile.
2. Avoid using any recycled bottles and other glassware unless they have been specifically rendered RNase-free by rinsing with 0.5*N* NaOH followed by copious amounts of sterile, distilled water. Alternatively, glassware can be baked at 150°C for 4 h.
3. Dedicate laboratory glassware for use with RNA and clearly mark it. Do not send anything outside of the laboratory to be washed once it has been rendered RNase free.
4. Microcentrifuge tubes can generally be taken from an unopened box, autoclaved, and used for all cDNA work. If necessary, they can be soaked overnight in a 0.01% aqueous solution of diethylpyrocarbonate (DEPC), rinsed with autoclaved, distilled water, and then autoclaved.

5. If made with RNase-free labware, most solutions can be made from reagent-grade materials and distilled water, and autoclaved. Solutions that are heat-sensitive should be made with autoclaved, distilled water and sterile-filtered to 0.2 μm in disposable plasticware.
6. If all else fails, most aqueous buffer solutions can be treated with 0.01% (v/v) DEPC and autoclaved. Buffers containing primary amines (such as Tris) cannot be effectively treated by this method.
7. Dedicate a separate set of automatic pipets for manipulating RNA and the buffers and enzymes used to synthesize cDNA.
8. Always wear disposable gloves to prevent contamination.

4.2. Materials

4.2.1. Reagents, Supplies, and Equipment

1. M-MLV H^- RT(BRL) (SuperScript™ RT).
2. *E. coli* DNA polymerase I.
3. *E. coli* DNA ligase.
4. *E. coli* RNase H.
5. T4 DNA polymerase.
6. $pd(N)_6$ (Random Hexamers) (Pharmacia LKB Biotech., Inc., Piscataway, NJ).
7. $pd(T)_{12-18}$ [$oligo(dT)_{12-18}$].
8. [α-^{32}P]dCTP (>3000 Ci/mmol).
9. GF/C glass fiber filters (1 × 2 cm).
10. dATP, dCTP, dGTP, dTTP.
11. Autoclaved 0.5- and 1.5-mL microcentrifuge tubes.
12. Automatic pipets capable of dispensing 1–20 μL and 20–200 μL.
13. Autoclaved, disposable tips for automatic pipets.
14. Disposable gloves.
15. 37 and 16°C water baths.
16. Phenol:$CHCl_3$:isoamyl alcohol (25:24:1 [v/v/v]): Saturate a bottle of redistilled phenol with distilled water. When the phases have separated, remove some of the bottom layer (phenol) with a glass pipet and transfer it to a clean glass bottle. Add an equal vol of $CHCl_3$:isoamyl alcohol (24:1 [v/v]), and store at 4°C until use. This solution should not be kept for more than 1 wk. The water-saturated phenol should have the excess water removed from it before it is stored at –20°C until needed again.
17. 10% (w/v) TCA containing 1% (w/v) sodium pyrophosphate (store at 4°C).
18. Yeast tRNA, 1 mg/mL: Place 25 mg of yeast tRNA (BRL) in 2.5 mL of 0.4*M* NaCl–0.05*M* Tris-HCl (pH 7.5); add 0.1 mL of 10% sodium dodecyl sulfate and 1 mL of 1 mg/mL proteinase K (BRL). Incubate at 37°C for 2 h. Extract at room temperature three times with an equal vol of phenol:$CHCl_3$:isoamyl alcohol, and precipitate the tRNA from the

aqueous phase with 2.5 vol of ethanol. Dissolve the tRNA pellet in 0.4*M* NaCl–0.05*M* Tris-HCl, and repeat the precipitation. Dissolve the pellet in sufficient autoclaved, distilled, deionized water to bring the concentration to 1 mg/mL (17.7 A_{260} U/mg). Store the solution at –20°C.

19. Autoclaved, distilled water.
20. Ethanol: 95% ethanol.
21. Bovine serum albumin (BSA), 50 mg/mL.

4.2.2. Solutions

1. dNTPs (10 m*M*): Prepare 50-m*M* stocks of each dNTP by dissolving 13–15 mg of each sodium salt in 100 µL of H_2O. Adjust to pH 6–7 by adding the appropriate amount of 1*N* NaOH (20–40 µL). Use pH paper. Add 4 µL of 1*M* Tris-HCl (pH 7.5) to buffer the solution, and bring the volume of the solution to 500 µL with water. Check the concentration of each solution at a 1/1000 dilution in water by determining the absorbance at 260 nm (E_{260} [1 cm/pH 7] for dCTP, dGTP, dATP, and dTTP is 7.4, 12.8, 15.3, and 8.4 cm^{-1} mM^{-1}, respectively). Prepare a mixture of all four dNTPs in water at a final concentration of 10 m*M* of each dNTP. Store all solutions at –20°C.
2. 7.5 *M* Ammonium acetate.
3. 0.1*M* Dithiothreitol (DTT) (store at –20°C).
4. 0.5*M* EDTA: Adjust pH to 7.5 with NaOH.
5. 5X First-Strand Buffer: 250 m*M* Tris-HCl (pH 8.3 at room temperature), 375 m*M* KCl, and 15 m*M* $MgCl_2$ (store at –20°C).
6. 5X Second-Strand Buffer: 94 m*M* Tris-HCl (pH 6.9 at room temperature), 453 m*M* KCl, 23 m*M* $MgCl_2$, 750 µ*M* β-NAD, and 50 m*M* $(NH_4)_2SO_4$ (store at –20°C).
7. 10X DNA Synthesis Buffer: 200 m*M* Tris-HCl (pH 8.3 at room temperature), 500 m*M* KCl, 25 m*M* $MgCl_2$, and 1 mg/mL BSA (store at –20°C).

4.3. Model Protocols

4.3.1. One-Tube Double-Stranded cDNA Synthesis: First-Strand Reaction

1. In a 1.5-mL tube, mix on ice 1–5 µg poly(A)$^+$ mRNA and one of the following:
   - 0.5–2.5 µg $pd(T)_{12-18}$
   - 40–200 ng $pd(N)_6$ or
   - A 10-fold molar excess of a primer-adapter (0.2–1.0 µg $pd[N]_{30}$)

   Add water to a vol of 7 µL; heat at 70°C for 10 min and chill on ice.
2. In the tube on ice, mix 4 µL 5X First-Strand Buffer, 2 µL 0.1*M* DTT, 1 µL 10 m*M* dNTPs, 1 µL [α-^{32}P]dCTP (1 µCi/µL), and sufficient water to

bring the final vol to 20 µL after RT is added; if second-strand cDNA is to be labeled, omit the [α-^{32}P]dCTP.

3. If pd(T)$_{12-18}$ or primer-adapter is present, equilibrate the tube to 37°C before adding 200–1000 U (1–5 µL) of M-MLV H$^-$ RT (SuperScript™ RT, use 200 U/µg mRNA); if pd(N)$_6$ is used, immediately add the appropriate amount of RT, and incubate at room temperature for 10 min; finally incubate all reaction types at 37°C for 1 h.
4. The final reaction composition is 50 m*M* Tris-HCl (pH 8.3),75 m*M* KCl, 3 m*M* $MgCl_2$, 10 m*M* DTT, 500 µ*M* each of dATP, dCTP, dGTP, and dTTP, 50–250 µg/mL mRNA, either 25–125 µg/mL pd(T)$_{12-18}$, 2–10 µg/mL pd(N)$_6$, or 10–50 µg/mL 30 base primer-adapter, and 10,000–50,000 U/mL SuperScript™ RT.
5. Place the tube on ice. Remove 2 µL from the reaction mixture, and add them to 48 µL of 20 m*M* EDTA (pH 7.5) containing 5 µg of tRNA. This will be used to calculate first-strand yield and to analyze the product by gel electrophoresis (*see* Section 4.4.).
6. Use the remaining 18 µL of first-strand reaction mixture (or the entire 20-µL reaction mixture if ^{32}P was not used) to carry out second-strand synthesis.

4.3.2. One-Tube Double-Stranded cDNA Synthesis: Second-Strand Reaction

1. This protocol is suitable for 1–5 µg mRNA originally in the 20-µL first-strand reaction mixture.
2. On ice, add the following reagents in the order shown to the first-strand reaction mixture tube: water to bring the final reaction mixture vol to 150 µL, 30 µL 5X Second-Strand Buffer, 3 µL 10 m*M* dNTPs, 10 U *E. coli* DNA ligase, 40 U *E. coli* DNA polymerase I, and 2 U *E. coli* RNase H. If the second-strand product is to be labeled, add 10 µCi of [α-^{32}P]dCTP after the unlabeled dNTPs. Vortex the tube gently, and incubate at 16°C for 2 h.
3. The final composition is 25 m*M* Tris-HCl (pH 7.5),100 m*M* KCl, 5 m*M* $MgCl_2$, 10 m*M* $(NH_4)_2SO_4$, 0.15 m*M* β-NAD, 250 µ*M* each of dATP, dCTP, dGTP, dTTP, 1.2 m*M* DTT, 67 U/mL DNA ligase, 267 U/mL DNA polymerase I, and 13 U/mL RNase H.
4. Add 20 U of T4 DNA polymerase, and continue incubating at 16°C for 5 min.
5. Place the tube on ice. If [α-^{32}P]dCTP was added to the second-strand reaction, remove 10 µL from the tube and add it to 40 µL of 20 m*M* EDTA containing 5 µg of tRNA. This will be used to calculate yield and to analyze the products by gel electrophoresis (Section 4.4.).

6. Add 10 µL of 0.5*M* EDTA to the remaining reaction mixture on ice. The product can now be phenol-extracted and ethanol-precipitated in preparation for addition of linkers or adapters *(35,37,49)*.

4.3.3. One-Tube Double-Stranded cDNA Synthesis: First-Strand Reaction in Preparation for Amplification of Specific cDNA Sequences by PCR

1. The following protocol is designed for synthesis of first-strand cDNA suitable for PCR amplification with specific primers. For strategies designed for PCR amplification of total cDNA libraries, *see* ref. *50* and Chapter 4.
2. In a 0.5-mL tube, mix on ice 1 µg total RNA (smaller amounts of poly [A]+ mRNA can be used) and one of the following: 0.5 µg $pd(T)_{12-18}$, or 100 ng $pd(N)_6$. Add water to a vol of 14 µL; heat at 70°C for 10 min and chill on ice.
3. In the tube on ice, mix 2 µL 10X DNA Synthesis Buffer, 2 µL 0.1*M* DTT, and 1 µL 10 m*M* dNTPs.
4. If $pd(T)_{12-18}$ is used, equilibrate the tube to 37°C, and add 200 U (1 µL) of SuperScript™ RT; if $pd(N)_6$ is used, immediately add 1 µL of SuperScript™ RT, and incubate at room temperature for 10 min; finally, incubate all reaction types at 37°C for 1 h.
5. The final reaction composition is 20 m*M* Tris-HCl (pH 8.3), 50 m*M* KCl, 2.5 m*M* $MgCl_2$, 10 m*M* DTT, 500 µ*M* each of dATP, dCTP, dGTP, and dTTP, 50 µg/mL RNA, 25 µg/mL $pd(T)_{12-18}$ or 5 µg/mL $pd(N)_6$, 100 µg/mL BSA, and 10,000 U/mL SuperScript™ RT.
6. Heat the tube at 90°C for 5 min, place the tube on ice for 5 min, and collect the tube contents by brief centrifugation.
7. Add 2 µL of RNase H (2 U/µL) to the tube, and incubate for 20 min at 37°C before proceeding with PCR amplification. Digestion of the first-strand cDNA–RNA hybrid to remove RNA is essential to subsequent amplification of the cDNA.
8. Place the tube on ice. Add directly to the tube, and mix 8 µL 10X DNA Synthesis Buffer, appropriate amounts of PCR primers *(51,52)*, sufficient water to bring the final vol to 100 µL after *Taq* DNA polymerase is added, and *Taq* DNA polymerase (1–4 U).
9. Layer 100 µL of mineral oil over the reaction mixture.
10. Carry out PCR amplification *(51,52)*.

4.3.4. [^{32}P]cDNA Synthesis

1. The following protocol is designed to synthesize high specific activity (4×10^8 cpm/µg) [^{32}P]cDNA that represents all sequences in an RNA population. The ratio of random primer to RNA template has been se-

lected to maximize product yield and length. Such labeled cDNA can be used as a hybridization probe or as starting material for generating a subtracted cDNA library *(38,39)*.

2. In a 1.5-mL tube, dry 10 µL [α-^{32}P]dCTP (10 µCi/µL, >3000 Ci/mmol). Add on ice 4 µL 5X First-Strand Buffer, 2 µL 0.1*M* DTT, 1 µL 10 m*M* each of dATP, dGTP, and dTTP, 4 µL 0.1 m*M* dCTP, 0.5 µg mRNA, 0.3 µg pd(N)$_6$, and sufficient water to bring the final vol to 20 µL after RT is added. Gently vortex the tube, and add 0.5 µL (100 U) of M-MLV H$^-$ RT. Incubate at room temperature for 10 min and at 37°C for 1 h.
3. Add 5 µL 0.5*M* EDTA and an equal vol (25 µL) of 0.6*N* NaOH, and incubate at 68°C for 30 min. Remove unincorporated dNTPs by chromatography in 0.1*M* NaCl, 10 m*M* Tris-HCl (pH 7.5), and 0.1 m*M* EDTA over a Sephadex G-50 column. Recover the product by ethanol precipitation.
4. Approximately 70% of the total dCTP in the reaction should be incorporated, resulting in the synthesis of 1.5×10^8 cpm of labeled product from 0.5 µg of mRNA. Up to 5 µg of RNA can be copied in a 20-µL reaction. If more than 0.5 µg of RNA is used, the amount of dCTP, [α-^{32}P]dCTP, pd(N)$_6$, and RT used should be increased proportionately.

4.4. Product Analysis

4.4.1. First-Strand Reaction

1. Spot duplicate aliquots (10 µL) from the diluted sample (Section 4.3.1., step 5) on separate glass fiber filters. Dry one filter and count in scintillant to determine the specific activity of the dCTP in the reaction mixture. Wash the other filter in TCA-sodium pyrophosphate (Section 4.2.1.), and dry and count the filter to determine yield. The yield of the first-strand reaction is calculated from the amount of acid-precipitable radioactivity determined. In order to perform the yield calculation, the specific activity of the [α-^{32}P]dCTP in the reaction must be determined. The specific activity (SA) is defined as the counts per minute (cpm) of an aliquot of the reaction mixture divided by the quantity (pmol) of the same nucleotide in the aliquot.

$$\text{SA (cpm/pmol dCTP)} = (\text{cpm}/10\ \mu\text{L})/(200\ \text{pmol dCTP}/10\ \mu\text{L}) \quad (1)$$

The SA should be approx 200 cpm/pmol. Once the SA is known, the amount of cDNA in the first-strand reaction can be calculated from the amount of acid-precipitable radioactivity determined from the washed filter:

$$\text{Amount of cDNA (µg)} = [(\text{cpm}) \times (50\ \mu\text{L}/10\ \mu\text{L}) \times (20\ \mu\text{L}/2\ \mu\text{L}) \times (4\ \text{pmol dNTP/pmol dCTP})]/[(\text{cpm/pmol dCTP}) \times (3030\ \text{pmol dNTP/µg cDNA})] \quad (2)$$

The correction in the numerator takes into account that, on the average, four nucleotides will be incorporated into the cDNA for every dCTP. The factor in the denominator is the amount of nucleotide that corresponds to 1 µg of single-stranded DNA.

2. Precipitate the product in the remaining diluted sample by adding 0.5 vol of 7.5*M* ammonium acetate and 2 vol of cold ethanol, and then centrifuging. Wash the pellet with 70% ethanol, centrifuge again, and dry the pellet after removing the ethanol. The size of the cDNA product can be analyzed by alkaline agarose gel electrophoresis *(53)*. Dissolve the sample pellet in 10 µL sample buffer (30 m*M* NaOH, 1 m*M* EDTA, 10% glycerol, 0.01% bromophenol blue). Appropriate [^{32}P]DNA size markers should be placed in sample buffer, e.g., BRL 1-kb DNA Ladder. The gel (1.4% [w/v]) should be cast in 30 m*M* NaCl, 2 m*M* EDTA, and then equilibrated for at least 3 h in alkaline electrophoresis buffer (30 m*M* NaOH, 2 m*M* EDTA) before loading the samples. For an 11 × 14 cm horizontal gel, electrophoresis should be for 5–6 h at 50 V or for 16–18 h at 15 V. Dry the gel under vacuum, heat for 1–2 h, and expose the gel to X-ray film overnight at room temperature.

4.4.2. Second-Strand Reaction

1. Spot duplicate aliquots (10 µL) from the diluted sample (Section 4.3.2., step 5) on separate glass fiber filters, and proceed with determining the [α-^{32}P]dCTP specific activity and the amount of acid-precipitable radioactivity as in Section 4.4.1., step 1. The SA of [α-^{32}P]dCTP is given by:

$$\text{SA (cpm/pmol dCTP)} = (\text{cpm}/10\ \mu\text{L})/(500\ \text{pmol dCTP}/10\ \mu\text{L}) \quad (3)$$

The specific activity should be approx 500 cpm/pmol. The amount of cDNA synthesized in the second-strand reaction is given by:

$$\text{Amount of cDNA } (\mu\text{g}) = [(\text{cpm}) \times (50\ \mu\text{L}/10\ \mu\text{L}) \times (150\ \mu\text{L}/10\ \mu\text{L}) \times (4\ \text{pmol dNTP/pmol dCTP})]/[(\text{cpm/pmol dCTP}) \times (3030\ \text{pmol dNTP}/\mu\text{g cDNA})] \quad (3)$$

2. Ethanol precipitate the DNA from the remaining diluted sample and proceed with alkaline agarose gel analysis as described in Section 4.4.1., step 2.

References

1. Baltimore, D. (1970) Viral RNA-dependent DNA polymerase. *Nature* **226,** 1209–1211.
2. Termin, H. M. and Mizutani, S. (1970) RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **226,** 1211–1213.

3. Ross, J., Aviv, H., Scolnick, E., and Leder, P. (1972) *In vitro* synthesis of DNA complementary to purified rabbit globin mRNA. *Proc. Natl. Acad. Sci. USA* **69,** 264–268.
4. Verma, I. M., Temple, G. F., Fan, H., and Baltimore, D. (1972) *In vitro* synthesis of DNA complementary to rabbit reticulocyte 10S RNA. *Nature New Biol.* **235,** 163–167.
5. Kacian, D. L., Spiegelman, S., Bank, A., Terada, M., Metafora, S., Dow, L., and Marks, P. A. (1972) *In vitro* synthesis of DNA components of human genes for globins. *Nature New Biol.* **235,** 167–169.
6. Maniatis, T., Kee, S. G., Efstratiadis, A., and Kafatos, F. C. (1976) Amplification and characterization of a β-globin gene synthesized *in vitro*. *Cell* **8,** 163–182.
7. Gerard, G. F., D'Alessio, J. M., and Kotewicz, M. L. (1989) cDNA synthesis by cloned Moloney murine leukemia virus reverse transcriptase lacking RNase H activity. *Focus* **11,** 66–69.
8. Krug, M. S. and Berger, S. L. (1987) First strand cDNA synthesis primed with oligo(dT). *Methods Enzymol.* **152,** 316–325.
9. Gubler, U. (1987) Second-strand cDNA synthesis: mRNA fragments as primers. *Methods Enzymol.* **152,** 330–335.
10. Gubler, U. and Chua, A. O. (1991) The establishment of cDNA libraries in lambda gt10, in *Essential Molecular Biology. A Practical Approach*, vol. 2 (Brown, T. A., ed.), IRL, Washington, DC, pp. 39–56.
11. Gerard, G. F. (1985) Comparison of cDNA synthesis by avian and cloned murine reverse transcriptase. *Focus* **7(1),** 1–3.
12. Kotewicz, M. L., D'Alessio, J. M., Driftmier, K. M., Blodgett, K. P., and Gerard, G. F. (1985) Cloning and overexpression of Moloney murine leukemia virus reverse transcriptase in *E. coli*. *Gene* **35,** 249–258.
13. Gerard, G. F., D'Alessio, J. M., Kotewicz, M. L., and Noon, M. C. (1986) Influence on stability in *E. coli* of the carboxy-terminal structure of cloned Moloney murine leukemia virus reverse transcriptase. *DNA* **5,** 271–279.
14. Roth, M. J., Tanese, N., and Goff, S. P. (1985) Purification and characterization of murine retroviral reverse transcriptase expressed in *E. coli*. *J. Biol. Chem.* **260,** 9326–9335.
15. D'Alessio, J. M., Noon, M. C., Ley, H. L., and Gerard, G. F. (1986) One-tube double-stranded cDNA synthesis using cloned M-MLV reverse transcriptase. *Focus* **9(1),** 1–4.
16. Gerard, G. F. (1986) Making effective use of cloned M-MLV reverse transcriptase. *Focus* **9(2),** 5,6.
17. Gerard, G. F. (1988) Synthesis of high specific activity cDNA. *Focus* **10,** 12,13.
18. D'Alessio, J. M. and Gerard, G. F. (1988) Second-strand cDNA synthesis with *E. coli* DNA polymerase I and RNase H: fate of information at the mRNA 5' terminus and the effect of *E. coli* DNA ligase. *Nucleic Acids Res.* **16,** 1999–2014.
19. Kotewicz, M. L., Sampson, C. M., D'Alessio, J. M., and Gerard, G. F. (1988) Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res.* **16,** 265–277.

20. Houts, G. E., Miyagi, M., Ellis, C., Beard, D., and Beard, J. M. (1979) Reverse transcriptase from avian myeloblastosis virus. *J. Virol.* **29,** 517–522.
21. Gerard, G. F. (1983) Reverse transcriptase, in *Enzymes of Nucleic Acid Synthesis and Modification, vol 1: DNA Enzymes* (Jacob, S. T., ed.), CRC, Boca Raton, FL, pp. 1–38.
22. Matson, S. W., Fay, P. J., and Bambara, R. A. (1980) Mechanism of inhibition of the avian myeloblastosis virus deoxyribonucleic acid polymerase by adriamycin. *Biochemistry* **19,** 2089–2096.
23. Huber, H. E., McCoy, J. M., Seehra, J. S., and Richardson, C. C. (1989) Human immunodeficiency virus 1 reverse transcriptase. Template binding, processivity, strand displacement synthesis, and template switching. *J. Biol. Chem.* **264,** 4669–4678.
24. Roberts, J. D., Beberek, K., and Kunkel, T. A. (1988) The accuracy of reverse transcriptase from HIV-1. *Science* **242,** 1171–1173.
25. Verma, I. M. (1977) The reverse transcriptase. *Biochem. Biophys. Acta* **473,** 1–38.
26. Berger, S. L., Wallace, D. M., Puskas, R. S., and Eschenfeldt, W. H. (1983) Reverse transcriptase and its associated ribonuclease H: interplay of two enzyme activities controls the yield of single-stranded complementary deoxyribonucleic acid. *Biochemistry* **22,** 2365–2372.
27. Han, J. H., Stratowa, C., and Rutter, W. J. (1987) Isolation of full-length putative rat lysophospholipase cDNA using improved methods for mRNA isolation and cDNA cloning. *Biochemistry* **26,** 1617–1625.
28. Han, J. H. and Rutter, W. J. (1988) Isolation of intact mRNA and construction of full length cDNA libraries: use of a new vector, λgt22, and primer-adapters for directional cDNA cloning, in *Genetic Engineering, Principles and Methods*, vol. 10 (Setlow, J. K., ed.), Plenum, New York, pp. 195–219.
29. Aviv, H. and Leder, P. (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid cellulose. *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
30. Rappolee, D. A., Mark, D., Banda, M. J., and Werb, Z. (1988) Wound macrophages express TGF-α and other growth factors in vivo: analysis by mRNA phenotyping. *Science* **241,** 708–712.
31. Doherty, P. J., Huesca-Contreras, M., Dosch, H. M., and Pan, S. (1989) Rapid amplification of complementary DNA from small amounts of unfractionated RNA. *Anal. Biochem.* **177,** 7–10.
32. Grady, L. J. and Campbell, W. P. (1989) Amplification of large RNAs (>1.5 Kb) by polymerase chain reaction. *BioTechniques* **7,** 798–800.
33. Carothers, A. M., Urlaub, G., Mucha, J., Grunberger, D., and Chasin, L. A. (1989) Point mutation analysis in a mammalian gene: rapid preparation of total RNA, PCR amplification of cDNA, and *Taq* sequencing by a novel method. *BioTechniques* **7,** 494–499.
34. Rutledge, R. G., Seligy, V. L., Cote, M. J., Dimock, K., Lewin, L. L., and Tenniswood, M. P. (1988) Rapid synthesis and cloning of complementary DNA from any RNA molecule into plasmid and phage M13 vectors. *Gene* **68,** 151–158.

35. Helfman, D. M., Fiddes, J. C., and Hanahan, D. (1987) Directional cDNA cloning in plasmid vectors by sequential addition of oligonucleotide linkers. *Methods Enzymol.* **152,** 349–359.
36. Huynh, T. V., Young, R. A., and Davis, R. W. (1985) Constructing and screening cDNA libraries in λgt10 and λgt11, in *DNA Cloning, A Practical Approach*, vol. 1 (Glover, D. M., ed.), IRL, Washington, D.C., pp. 49–78.
37. Polazzolo, M. J. and Meyerowitz, E. M. (1987) A family of lambda phage cDNA cloning vectors, λSWAJ, allowing the amplification of RNA sequences. *Gene* **52,** 197–206.
38. Duguid, J. R., Rohwer, R. G., and Seed, B. (1988) Isolation of cDNAs of scrapie-modulated RNAs by subtractive hybridization of a cDNA library. *Proc. Natl. Acad. Sci. USA* **85,** 5738–5742.
39. Travis, G. H. and Sutcliffe, J. G. (1988) Phenol emulsion-enhanced DNA-driven subtractive cDNA cloning: isolation of low abundance monkey cortex-specific mRNAs. *Proc. Natl. Acad. Sci. USA* **85,** 1696–1700.
40. Kacian, D. L. and Myers, J. C. (1976) Synthesis of extensive, possibly complete, DNA copies of poliovirus RNA in high yields and at high specific activities. *Proc. Natl. Acad. Sci. USA* **73,** 2191–2195.
41. Myers, J. C., Spiegelman, S., and Kacian, D. L. (1977) Synthesis of full-length DNA copies of avian myeloblastosis virus RNA in high yields. *Proc. Natl. Acad. Sci. USA* **74,** 2840–2843.
42. Myers, J. C. and Spiegelman, S. (1978) Sodium pyrophosphate inhibition of RNA-DNA hybrid degradation by reverse transcriptase. *Proc. Natl. Acad. Sci. USA* **75,** 5329–5333.
43. Srivastava, A. and Modak, M. J. (1979) Reverse transcriptase-associated RNase H. Part IV. Pyrophosphate does not inhibit RNase H activity of AMV DNA polymerase. *Biochem. Biophys. Res. Commun.* **91,** 892–899.
44. Marcus, S. L., Smith, S. W., and Bacchi, C. J. (1981) Polyamines stimulate natural RNA-directed DNA synthesis by Rauscher murine leukemia virus DNA polymerase. *Biochem. Biophys. Res. Commun.* **99,** 1361–1368.
45. McDonnell, J. P., Garapin, A. C., Levinson, W. E., Qunitrell, N., Fanshier, L., and Bishop, J. M. (1970) DNA polymerases of Rous sarcoma virus: delineation of two reactions with actinomycin. *Nature* **228,** 433–435.
46. Bunte, T., Novak, U., Friedrich, R., and Moelling, K. (1980) Effect of actinomycin D on nucleic acid hybridization. The cause of erroneous DNA elongation during DNA synthesis of RNA tumor viruses *in vito. Biochem. Biophys. Acta.* **610,** 241–247.
47. Okayama, H. and Berg, P. (1982) High-efficiency cloning of full-length cDNA. *Mol. Cell Biol.* **2,** 161–170.
48. Gubler, U. and Hoffman, B. J. (1983) A simple and very efficient method for generating cDNA libraries. *Gene* **25,** 263–269.
49. Wu, R., Wu, T., and Ray, A. (1987) Adaptors, linkers, and methylation. *Methods Enzymol.* **152,** 343–349.
50. Belyavsky, A., Vinogradova, T., and Rajewsky, K. (1989) PCR-based cDNA library construction: general cDNA libraries at the level of a few cells. *Nucleic Acids Res.* **17,** 2919–2932.

51. Saiki, R. K (1989) The design and optimization of the PCR, in *PCR Technology. Principles and Applications for DNA Amplification* (Erlich, H. A., ed.), Stockton, New York, pp. 7-16.
52. Innis, M. A. and Gelfand, D. H. (1990) Optimization of PCRs, in *PCR Protocols. A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Shinsky, J. J., and White, T. J., eds.), Academic, New York, pp. 3–12.
53. McDonnell, M. W., Simon, M. N., and Studier, F. W. (1977) Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. *J. Mol. Biol.* **10,** 119–146.

CHAPTER 7

Terminal Deoxyribonucleotidyl Transferase (EC 2.7.7.31)

Frank Grosse and Andreas Manns

1. Introduction

1.1. The Reaction

Terminal deoxyribonucleotidyl transferase (TdT) (EC 2.7.7.31) catalyzes the addition of deoxynucleoside triphosphates to the 3' ends of oligo- and polynucleotide primers. Different from all the other known DNA polymerases, TdT-catalyzed DNA synthesis is not template-directed. The reaction can be described as:

$$n\ \mathrm{dNTP} + \mathrm{d(pX)}_m \xrightarrow[\mathrm{TdT}]{\mathrm{Me}^{2+}} \mathrm{d(pX)}_m\mathrm{(pN)}_n + n\mathrm{PP}_i \qquad (1)$$

Only one dNTP is necessary for polymerization, although any dNTP (and derivatives thereof) or any combination of dNTPs will serve as substrate.

1.2. Biological Function of TdT

TdT activity is only found in the nuclei of pre-T and pre-B lymphocytes during immunopoiesis. Therefore, it has been proposed that this enzyme might play a role in creating somatic diversity of immunoglobulin and T-cell receptor genes *(1)*. This hypothesis has gained indirect support by the demonstration that the existence of N regions, i.e., small numbers of nucleotides that are not encoded by the chromosomes, correlates with the expression of TdT. Besides this circumstantial evidence for a role as an active mutator, after 25 years of intensive

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

research, still nothing is known about the biological function of TdT. Nevertheless, TdT is an important marker for the diagnosis and classification of pre-B-cell and pre-T-cell leukemias and differential diagnosis of myeloid leukemias. The use of TdT as a marker in leukemia diagnosis is the subject of two excellent reviews *(2,3)*.

1.3. Applications

Because of its terminal addition properties, TdT has been widely employed for the production of synthetic homo- and heteropolymers. A wide variety of polymers can be synthesized from derivatives of dNTPs, which are *N*-acetylated or *N*-alkylated. A detailed discussion of the mechanism of these reactions, the statistics of polymerization, inhibitors, and practical synthetic applications has been provided by Bollum *(4)*.

Homopolymeric tailing of linear duplex DNA for in vitro genetic recombination is the most common application today. In a key experiment of genetic engineering, linear bacteriophage P22 DNA was tailed with oligo(dA), and another set of linear P22 was tailed with the complementary oligo(dT). After annealing, the dA/dT-tailed recombinant molecule was ligated to give covalently closed dimeric circles *(5)*. Details of the tailing reaction were recently reviewed *(6,7)*.

Another application utilizes the limited addition of rNTPs *(8)*, ddNTPs *(9)*, and cordycepin triphosphate *(10,11)* for terminal additions. Since these substrates lead to chain terminations, 3' ends can be extended in a controlled manner. This has been exploited for the radioactive labeling of the 3'-hydroxyl ends of single- and double-stranded DNA. A modification of this method is the 3' labeling of DNA primers by biotin-11-dUTP *(12,13)* or fluorescent succinylfluorescein-labeled dideoxnucleoside triphosphates *(14)*. TdT has also been used to elongate primers by only one phosphorothioate deoxynucleotide. The resulting extension products were purified by means of a mercury beaded column. This approach has been used for the introduction of mispaired primer ends for site-directed mutagenesis protocols *(6)*.

2. The Enzyme

2.1. Structure

TdT is a monomeric enzyme with a mol mass of about 60,000 Da, whose cDNA has been cloned and expressed in various organisms (*see*, e.g., *[15]*). Dependent on the biological source, the enzyme con-

sists of 508–529 amino acids. A high degree of sequence homology (>80%) has been observed between TdTs from different species. There is also a remarkable similarity between TdT and the cellular repair polymerase, DNA polymerase β. Earlier preparations represented a proteolytically degraded form of the enzyme that consists of two subunits. The smaller 8000 Da α-subunit corresponds to amino acids 403–508 of human TdT, and the larger 24,000 Da β-subunit corresponds to amino acids 159–402 of the undegraded 60,000 Da form *(3)*. The undegraded form of the enzyme has been reported to exhibit a severalfold lower turnover number than the (commercially available) degraded enzyme *(16)*. It is not yet known whether proteolytic degradation changes enzymic properties other than k_{cat}.

The two-subunit form of the enzyme is commercially available from many biochemical companies. Although many firms deliver an excellent enzyme, every now and then we have observed nuclease contamination in different batches. Therefore, it is recommended that the specifications of different lots be checked, particularly for nuclease contamination, and moreover, it is wise to include experimental controls that would detect exonuclease and endonuclease activity (*see*, e.g., *[17]*).

2.2. Reaction Conditions

2.2.1. Reaction Buffer

TdT is peculiar with regard to the reaction conditions in several respects. The activity is strongly inhibited by the ammonium cation as well as chloride, iodide, and phosphate anions stimulating a quest for the optimal buffer. In general, potassium or sodium cacodylate buffers are preferred, since they were shown to be optimal for polypurine and polypyrimidine synthesis *(18)*. However, cacodylate buffers suffer from important drawbacks. First, cacodylate (dimethyl arsenic acid) is toxic; second, cacodylate might be contaminated by heavy metal ions, which must be removed prior to use, e.g., by treatment with a complexing chelate resin, and third, SH-protective compounds, such as DTT, which are mandatory for an enzyme's activity, react themselves with cacodylate to yield garlic-like smelling and highly toxic sulfur-cacodylic compounds. Thus, an alternative for cacodylate buffers is highly desirable. The only systematic study for replacing cacodylate by other buffer substances revealed that TdT activity in

2(*N*-morpholino)ethane sulfonic acid (MES) has only one-quarter the activity of TdT in cacodylate. With Tris-HCl, only 2% of maximal activity was observed *(19)*. We found that 100 m*M* Tris-acetate, pH 7.2, is an attractive alternative to cacodylate buffers, giving nearly the same activity as Chelex-treated potassium cacodylate.

2.3. Divalent Cations

The polymerization reaction requires the presence of a divalent cation, with an order of efficiency of $Mg^{2+} > Zn^{2+} > Co^{2+} > Mn^{2+}$ *(20)* for the elongation of oligonucleotide primers with dAMP. dGTP is also optimally added in the presence of Mg^{2+}; on the other hand, the polymerization of pyrimidines is best in Co^{2+}-containing buffers *(18)*. In genetic engineering, most tailing reactions are performed in the presence of either Co^{2+} for pyrimidine additions or Mn^{2+} for purine additions. The frequent use of Mn^{2+} for purine additions is based on the finding that Mn^{2+} permits TdT to extend duplex termini *(6)*, and that contaminating nucleases are less active or even inactive in the presence of Mn^{2+}.

2.4. Parameters Influencing the Reaction

TdT binds its substrates with rather high K_m values of 100 μ*M* for dATP and dGTP, 500 μ*M* for dTTP and dCTP, and 1 μ*M* for oligonucleotide primers and up to 1 m*M* for homopolymer primer ends *(18)*. Thus, for optimal reaction rates, all the substrates should be present in highly concentrated solutions. This is most readily achieved by working in small volumes. The number of nucleotides added to a distinct primer molecule will in principle be determined by the ratio of Mol dNTPs:Mol 3'-OH-termini in the reaction mixture. Thus, if there is a 100-fold molar excess of dNTP over primer ends, nearly 100 nucleotides will be incorporated per primer with a distribution of products that ideally obeys the Poisson distribution (for details, *see* ref. *4*). Such a behavior is based on the assumption that all primer ends are equally utilized by the enzyme, which is only true for the homopolymeric extension of an isohomopolymeric primer, e.g., (dC)-tailing of $(dC)_{10}$. On the other hand, tailing of restriction fragments is far from being ideal, mainly because differences in primer binding exist that are dependent on the availability of the 3'-OH group and the nucleotide composition of the primer termini. Assume a blunt-ended DNA that should become extended by an oligo(dC) tail. The very first nucleotide incorporation step at the blunt end will be rather slow, because of the poor availability of a free,

i.e., temporarily melted 3' terminus. The melting (or breathing) process of the primer end, in turn, is dependent on the relative base composition: Ends that are rich in G:C base pairs are more stable and thus less accessible to the elongation reaction than A:T-rich ends. As soon as one nucleotide is inserted, the primer has a 3'-OH overhang, which will be elongated more easily. As a consequence, a few blunt-end primers will be elongated to a larger extent than initially expected. Furthermore, TdT binds about 10-fold stronger to a terminal dG or dA than to a dC or dT *(18)*. This again might lead to a bias in the expected distribution of elongated termini, particularly if different primer ends are involved. All these effects lead to more or less unpredictable results in both the fraction of primers elongated and the number of nucleotides added.

For the extension of restriction fragments with subsequent insertion into vectors, very long extensions are as harmful as very short ones. Long stretches of homopolymers within a plasmid or phage vector are not very stable, and tend to delete or otherwise hinder the vector propagation. Very short stretches or a large fraction of primer ends without any addition drastically reduce the cloning efficiency. Therefore, a careful control of the lengths of tails by direct visualization from a gel is required before the annealing reaction with a complementary tailed counterpart is performed. This is not necessary for tailing with dGMP, because this reaction ceases after the incorporation of about 20 nucleotides owing to aggregate formation of the newly generated oligo(dG) stretches *(4)*. Thus, by using dG/dC-tailing, only the lengths of the dC-tail must be controlled and optimized. For the extension of restriction fragments, the amount of enzyme relative to DNA ends is a critical factor. It is conceivable that at a low ratio of enzyme to 3'-OH ends the enzyme tends to elongate an already elongated primer in preference to a new one, yielding few primers containing many nucleotides and many primers without any addition. To ensure that most of the primers will be elongated, particularly when the reaction is carried out with 3'-recessive or blunt ends, a 10- to 100-fold molar excess of TdT over primer ends is necessary *(21)*. Terminal additions to blunt-ended DNA and 3'-recessive-ended DNA also require a partial melting of such ends. Therefore, conditions that destabilize DNA duplexes, such as low salt buffers and/or the replacement of Mg^{2+} by Co^{2+} or Mn^{2+}, are favorable. Likewise, λ exonuclease can be used to trim the 5' ends and thereby converting a blunt or 3'-recessive end into a 3' overhang *(5)*.

3. Experimental Procedures

In the following, we will give some examples for the practical use of terminal transferase. We readdress the issue of homopolymer synthesis (cf ref. *4*) by providing data for the use of phosphorothioate derivatives of the four dNTP substrates (dNTPαS). Homopolymers containing phosphorothioate diester linkages are interesting mainly because they are expected to be more stable against nuclease degradation than phosphatediester homopolymers *(22,23)*. Their thermodynamic behavior in an annealed state is not yet known. Furthermore, $(dCS)_{28}$ has been shown to display anti-HIV-1 activity, both in vivo and in vitro *(24)*.

We further describe reaction conditions for the addition of dNTPαS to 3'-hydroxyl ends of linear double-stranded DNA. The insert DNA will be tailed uniformly with $(dG)_{20}$. The less valuable linearized vector will be tailed with oligo(dC). After annealing, the dG/dC-tailed recombinant molecule will be used to transform competent *E. coli* cells. Within the bacteria, remaining single-stranded DNA will be filled up by DNA polymerase I and ligated to covalently closed circular DNA. As a sidereaction, intracellular nucleases degrade the recombinant molecule, preferably at its vulnerable single-stranded regions. This reduces transformation efficiencies. Replacing the dNTPs by their α-phosphorothioate analogs yields single-stranded DNA intermediates that are less susceptible to degradation.

3.1. Materials

1. Terminal deoxynucleotidyl transferase (Stratagene, Heidelberg, Germany).
2. Cacodylic acid (Sigma, Deisenhofen, Germany).
3. Solid KOH.
4. Chelex 100; 200–400 mesh, sodium form (BioRad, Richmond, CA).
5. $p(dC)_{10}$ (Pharmacia, Freiburg, Germany).
6. 0.1 *M* $CoCl_2$.
7. 0.1 *M* $MnCl_2$.
8. 0.1 *M* Dithiothreitol (Boehringer, Mannheim, Germany).
9. Deoxynucleoside triphosphates (Boehringer).
10. [^{3}H]Deoxynucleoside triphosphates, 600 Ci/mmol (Amersham, Braunschweig, Germany).
11. Deoxynucleoside-5'-*O*-(1-thio)-triphosphates, dNTPαS (Amersham).
12. [^{35}S]Deoxynucleoside-5'-*O*-(1-thio)-triphosphates, 1300 Ci/mmol (NEN, Bad Homburg, Germany, Amersham).
13. DE81 filter disks (Whatman, Maidstone, UK).

14. 0.5 *M* Na_2HPO_4.
15. Competent cells of *E. coli* DH5α (Gibco-BRL, Eggenstein, Germany).
16. 10 m*M* Tris-HCl, pH 7.2, 1 m*M* EDTA, 100 m*M* NaCl.

3.2. Methods

3.2.1. Preparation of the Tailing Buffer

1. Equilibrate 5 g Chelex 100 with 3*M* potassium acetate, pH 7. After 5 min at room temperature, remove excess liquid by passing the slurry through a glass sintered funnel by applying water aspirator vacuum. Wash the Chelex in the funnel with 10 mL distilled water by applying water aspirator vacuum.
2. Prepare a 1.2*M* solution of cacodylic acid. Add KOH pellets, until a pH of approx 7 is obtained. Adjust to pH 7.2 by the dropwise addition of 1*M* KOH. Dilute with distilled water to obtain a 1*M* stock solution.
3. Add the equilibrated Chelex to the potassium cacodylate solution, and stir for about 5 min. Remove the ionic exchanger by filtration.
4. Prepare a 5X C-Tailing Buffer by pipeting in the following order:
   a. 10 mL 1*M* potassium cacodylate, pH 7.2
   b. 8.8 mL distilled water
   c. 0.2 mL 0.1*M* DTT
   d. 1 mL 0.1*M* $CoCl_2$

   Store in 500-µL aliquots at –20°C.
5. Prepare a 5X G-Tailing Buffer by pipeting in the following order:
   a. 10 mL 1*M* potassium cacodylate, pH 7.2
   b. 8.8 mL distilled water
   c. 0.2 mL 0.1*M* DTT
   d. 1 mL 0.1*M* $MnCl_2$

   Store in 500-µL aliquots at –20°C.

3.2.2. Preparation of Phosphorothioate-Containing Homopolymers

1. Mix:

$p(dC)_{10}$ (10 A_{260} U/mL)	1.5 µL (10 µ*M* 3'-OH)
5X C-tailing buffer	5 µL
[α-^{35}S]dCTP (1300 Ci/mmol, 13 µCi/µL)	2 µL
5 m*M* dCTPαS	2.5 µL
Distilled water	13 µL
Terminal transferase (17 U/µL)	1 µL

2. Incubate for 60 min at 37°C.
3. Spot a 2-µL aliquot onto a DE81-filter disk, dry under a fan, and count the total radioactivity in a toluene-based scintillant.

4. After measuring the total radioactivity, place the filter disk on a suction device. Rinse several times with ethanol (to remove toluene), and then dry under a fan.
5. Keep the dried filter for 5 min in a 0.5*M* solution of Na_2HPO_4. Then place it on a suction device, and rinse five times with 1 mL (each) of 0.5*M* Na_2HPO_4, two times with 1 mL water, and two times with 1 mL of ethanol. Then dry under a fan, and count the dCMPS incorporation in a toluene-based scintillant.
6. Calculate the amount of dCS-tails added to $(dC)_{10}$ from the amount of filter-retained radioactivity:

$$\begin{aligned}&(\text{cpm binding to DE81/Total cpm}) \times \\ &(500\ \mu M\ \text{dCMPS}/10\ \mu M\ 3'\text{-OH}) = \text{dCMPS additions/primer}\end{aligned} \quad (2)$$

7. Ethanol-precipitate the remaining DNA in the presence of potassium acetate.
8. Control the tail lengths by denaturing polyacrylamide gel electrophoresis.

3.2.3. Tailing of Restriction Fragments with dNTPαS

3.2.3.1. Tailing Double-Stranded DNA with Oligo(dGS)

1. Prepare a reaction mixture containing:

Double-stranded insert DNA (dried pellet)	About 2.5 pmol 3' ends
5X G-tailing buffer	5 µL
[α-^{35}S]dGTP (1300 Ci/mmol, 13 µCi/µL)	2 µL
100 µ*M* dGTPαS	6 µL
Distilled water	10 µL
Terminal transferase (17 U/µL)	2 µL

2. Incubate at 37°C for 30 min.
3. Spot a 2-µL aliquot onto a DE81-filter disk, and proceed as described in Section 3.2.2.
4. Calculate the amount of dG-tails added to the double-stranded DNA from the amount of radioactivity bound to the DE81 filter:

$$\begin{aligned}&(\text{cpm binding to DE81/Total cpm}) \times (24\ \mu M\ \text{dGMPS}/20) \\ &= 3'\text{-OH termini of insert DNA}\ (\mu M)\end{aligned} \quad (3)$$

If the expected number of 20 for dGMPS added to each 3' terminus is not achieved or (likewise) the calculated number of 3'-OH termini of insert DNA is too low, the tailing procedure should be repeated after ethanol precipitation of the DNA. Because dG- and dGS-containing tails cannot exceed 20 residues, there is no danger of generating tails that are too long.

5. Ethanol-precipitate the remaining DNA in the presence of potassium acetate.

3.2.3.2. Tailing Linearized Vector DNA with Oligo(dCS)

1. Prepare a reaction mixture with:

Linearized vector DNA (dried pellet)	10 pmol 3' ends
5X C-tailing buffer	20 µL
0.1 m*M* dCTPαS	30 µL
[α-^{35}S]dCTP (1300 Ci/mmol, 13 µCi/µL)	2 µL
Distilled water	38 µL
Terminal transferase (17 U/µL)	10 µL

2. Spot a 2-µL aliquot onto a DE81 filter disk, dry under a fan, and count the total radioactivity in a toluene-based scintillant.
3. Incubate at 20°C, take 2-µL aliquots every 15 min, and determine as soon as possible the incorporated amount of dCS per 3'-OH terminus by using the DE81 filter technique described in Section 3.2.2. When about 20 dCS residues are incorporated per primer, stop the reaction by heating to 65°C for 10 min. Chill on ice.
4. Ethanol-precipitate DNA in the presence of potassium acetate.
5. Take an aliquot, cleave with a restriction endonuclease that cuts close to the extension, and control the tail lengths by polyacrylamide gel electrophoresis.
6. If necessary, repeat the procedure by incubating for a shorter or longer time.

3.2.3.3. Annealing of the Tailed DNA Components and Transformation of Competent *E. coli* Cells

1. Dissolve the (dGS)-tailed insert DNA and the (dCS)-tailed vector DNA in 10 µL (each) of 10 m*M* Tris-HCl, pH 7.2, 1 m*M* EDTA, and 100 m*M* NaCl.
2. Mix aliquots of the two solutions in an approximate molar ratio of 1:1 (3' ends) for each component.
3. Incubate at 65°C for 10 min in an incubator block.
4. Turn off the incubator block, and let slowly cool down to room temperature.
5. Take an aliquot with about 1 fmol annealed construct for the transformation of 100 µL competent *E. coli* cells.

The same protocol is applicable for tailing with unmodified deoxynucleoside triphosphates. Transformation efficiencies are about threefold higher with phosphorothioate-tailed constructs than those obtained with normal dG/dC-tailed constructs.

References

1. Baltimore, D. (1974) Is terminal deoxynucleotidyl transferase a somatic mutagen in lymphocytes? *Nature (London)* **248,** 409–411.
2. Bollum, F. J. (1979) Terminal deoxynucleotidyl transferase as a hematopoietic cell marker. *Blood* **54,** 1203–1215.
3. Bollum, F. J. and Chang, L. M. S. (1986) Terminal transferase in normal and leukemic cells. *Adv. Cancer Res.* **47,** 37–61.
4. Bollum, F. J. (1974) Terminal deoxynucleotidyl transferase, in *The Enzymes* (Boyer, P. D., ed.), Academic, New York, pp. 145–171.
5. Lobban, P. E. and Kaiser, A. D. (1973) Enzymatic end-to-end joining of DNA molecules. *J. Mol. Biol.* **78,** 453–471.
6. Deng, G. and Wu, R. (1983) Terminal transferase: use in the tailing of DNA and for *in vitro* mutagenesis. *Methods Enzymol.* **100,** 96–116.
7. Eschenfeldt, W. H., Puskas, R. S., and Berger, S. L. (1987) Homopolymeric tailing. *Methods Enzymol.* **152,** 337–342.
8. Kössel, H. and Roychoudhury, R. (1971) Synthetic polynucleotides. The terminal addition of riboadenylate acid to deoxynucleotides by terminal deoxynucleotidyl transferase as a tool for the specific labeling of deoxynucleotides at the 3'-ends. *Eur. J. Biochem.* **22,** 271–276.
9. Cozzarelli, N. R., Kelly, R. B., and Kornberg, A. (1969) Enzymatic synthesis of DNA. Hydrolysis of a 5'-triphosphate-terminated polynucleotide in the active center of DNA polymerase. *J. Mol. Biol.* **45,** 513–531.
10. Gaastra, W. and Klemm, P. (1984) Radiolabeling of DNA with 3' terminal transferase, in *Methods in Molecular Biology, vol 2: Nucleic Acids* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 269–271.
11. Tu, C.-D. and Cohen, S. N. (1980) 3'-End labeling of DNA with [α-^{32}P]cordycepin-5'-triphosphate. *Gene* **10,** 177–183.
12. Vincent, C., Tchen, P., Cohen-Solal, M., and Kourilsky, P. (1982) Synthesis of 8-(2,4-dinitrophenyl-2,6-aminohexyl)aminoadenosine-5'-triphosphate: Biological properties and potential uses. *Nucleic Acids Res.* **10,** 6787–6796.
13. Kumar, A., Tchen, P., Roullet, F., and Cohen, J. (1988) Nonradioactive labeling of synthetic oligonucleotide probes with terminal deoxynucleotidyl transferase. *Anal. Biochem.* **169,** 376–382.
14. Trainor, G. L. and Jensen, M. A. (1988) A procedure for the preparation of fluorescence-labeled DNA with terminal deoxynucleotidyl transferase. *Nucleic Acids Res.* **16,** 11,846.
15. Chang, L. M. S., Rafter, E., Rusquet-Valerius, R., Peterson, R. C., White, S. T., and Bollum, F. J. (1988) Expression and processing of recombinant human terminal transferase in the baculovirus system. *J. Biol. Chem.* **263,** 12,509–12,513.
16. Chang, L. M. S. and Bollum, F. J. (1986) Molecular biology of terminal transferase. *Crit. Rev. Biochem.* **21,** 27–52.
17. Fisher, P. A., Wang, T. S.-F., and Korn, D. (1979) Enzymological characterization of DNA polymerase α. *J. Biol. Chem.* **254,** 6128–6137.
18. Kato, K.-I., Gonçalves, J. M., Houts, G. E., and Bollum, F. J. (1967) Deoxynu-

cleotide-polymerizing enzymes of calf thymus gland. II. Properties of the terminal deoxynucleotidyltransferase. *J. Biol. Chem.* **22,** 2780–2789.
19. Chiprich, T. P. (1978) The effect of different buffers on terminal deoxynucleotidyl transferase activity. *Biochim. Biophys. Acta* **518,** 535–538.
20. Deibel, M. and Coleman, M. S. (1980) Biochemical properties of purified human terminal deoxynucleotidyltransferase. *J. Biol. Chem.* **255,** 4206–4212.
21. Brutlag, D., Fry, K., Nelson, T., and Hung, P. (1977) Synthesis of hybrid bacterial plasmids containing highly repeated satellite DNA. *Cell* **10,** 509–519.
22. Eckstein, F. (1985) Nucleoside phosphorothioates. *Ann. Rev. Biochem.* **54,** 367–402.
23. Eckstein, F. and Gish, G. (1989) Phosphorothioates in molecular biology. *Trends Biochem. Sci.* **14,** 97–100.
24. Majumdar, C., Stein, C. A., Cohen, J. S., Broder, S., and Wilson, S. H. (1989) Stepwise mechanism of HIV reverse transcriptase: primer function of phosphorothioate oligodeoxynucleotide. *Biochemistry* **28,** 1340–1346.

CHAPTER 8

Restriction Enzymes

Alfred Pingoud, Jürgen Alves, and Robert Geiger

1. Introduction

Restriction enzymes are endonucleases that recognize specific double-stranded DNA sequences and cleave the DNA in both strands, e.g.:

$$\begin{array}{lcl} \text{-GpApApTpTpC-} & & \text{-G} \qquad\qquad\quad \text{pApApTpTpC-} \\ & \rightarrow & \qquad\qquad + \\ \text{-CpTpTpApApG-} & & \text{-CpTpTpApAp} \qquad\qquad\quad \text{G-} \end{array} \qquad (1)$$

They have been identified in many prokaryotic organisms and are considered to be part of a defense system directed against foreign DNA. Site-specific endodeoxyribonucleases are given EC numbers, e.g., EC 3.1.21.4 for a Type II enzyme *(see below)*.

The phenomenon of restriction and modification was discovered by Bertani and Weigle *(1)*, who noticed that bacteriophage λ grown on *E. coli* C could not be efficiently propagated on *E. coli* K. Once established there, however, it would grow poorly on *E. coli* C. Similar responses of bacteriophages to their hosts were observed in other systems. The molecular basis for this phenomenon was identified by Arber and Dussoix *(2)*, who demonstrated that the bacterial host contains a restriction enzyme that destroys the bacteriophage DNA and a modification enzyme that protects the host DNA against the nucleolytic activity of its own restriction enzyme. It was suggested that these restriction enzymes recognized a specific DNA sequence and subsequently cleaved the DNA, unless it was modified by the companion

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

Table 1
Classification of Restriction Enzymes

	Type I	Type II	Type III
Example	*Eco*K	*Eco*RI	*Eco*P
Subunits	Three different	Two identical	Two different
Activity	Restriction, modification, topoisomerase, ATPase	Only restriction	Restriction, modification, ATPase
Cofactor requirements	Mg^{2+}, ATP, S-AdoMet	Mg^{2+}	Mg^{2+}, ATP, (S-AdoMet)
Recognition sequence	AACNNNNGTGC	GAATTC	AGACC
Position of cleavage	Variable and great distance from the recognition site	Within the recognition site	25 bp away from the recognition site

modification enzyme, which was supposed to recognize the same specific DNA sequence on the host DNA, methylate it, and thereby prevent cleavage. This was verified in vitro by Meselson and Yuan *(3)*, who demonstrated that bacteriophage λ DNA isolated from an *E. coli* C culture was degraded by the *E. coli* K restriction enzyme, whereas bacteriophage λ DNA isolated from an *E. coli* K culture was not. A similar experiment was carried out by Linn and Arber *(4)* with bacteriophage fd DNA. Later Smith and Wilcox *(5)* showed that the *Haemophilus influenzae* restriction enzyme cleaved DNA from the bacteriophage P22, but had no effect on the chromosomal DNA.

On detailed biochemical characterization of purified restriction enzymes *(3–6)*, it became apparent that they differed in their basic enzymology, in particular their subunit composition, cofactor requirement, and mode of cleavage. Three classes are recognized: Type I (EC 3.1.21.3), Type II (EC 3.1.21.4), and Type III (EC 3.1.21.5) (Table 1). Type I enzymes, such as *Eco*K *(3)*, typically are composed of three nonidentical subunits, need Mg^{2+} ions, ATP, and *S*-adenosylmethionine to be active, and cleave the DNA at apparently random sites far away from the recognition site. Type II enzymes, such as *Hin*dII *(5)*, typically exist as dimers of two identical subunits, require only Mg^{2+} ions for their activity, and cleave DNA within or very close to the recogni-

tion sequence. Type III enzymes, such as *Hin*fIII *(7)*, are ATP and Mg^{2+} dependent, but do not show a stringent requirement for *S*-adenosylmethionine (Table 1). The classification of restriction enzymes as Type I or Type II enzymes was proposed by Boyer *(8)*. The currently used and accepted nomenclature for individual restriction enzymes was introduced by Smith and Nathans *(9)*: It uses a three-letter abbreviation (in italics) for the organism from which the enzyme has been isolated. Often a fourth letter is included to designate a particular strain. A roman numeral is added to differentiate enzymes of different specificity isolated from the same source. For example, *Hin*dII is the name for the second restriction enzyme isolated from *H. influenzae* strain d. The official nomenclature includes a capital letter R followed by a raised dot in front of the enzyme symbol, e.g., R•*Hpa*II, to distinguish the restriction enzyme from the corresponding modification enzyme, e.g., M•*Hpa*II. Most often, however, the letter R• is not used, such that *Hpa*II, for example, denotes the restriction endonuclease and not the modification methyltransferase.

The Type II restriction enzymes that recognize defined DNA sequences 4 to 8 bp in length and cleave the DNA within or close to the recognition sequence have turned out to be of paramount importance for the analysis and engineering of the genetic material. They will be the subject of this chapter. Excellent reviews have been published that deal with the genetics *(10)* and enzymology of restriction enzymes *(11–14)*. Two volumes of the series *Gene Amplification and Analysis* deal with restriction enzymes *(15,16)*. Experimental details for the use of restriction enzymes can be found in various laboratory manuals (e.g., *17–21*), as well as in vol. 65 of *Methods in Enzymology (22)*. Comprehensive lists of the properties and sources of known restriction enzymes appear regularly *(23,24)*.

The present chapter will focus on the structure and mechanism of action of restriction enzymes. Information will be emphasized that is important for the use of restriction enzymes for various purposes.

2. Enzymology of Type II Restriction Enzymes (EC 3.1.21.4)

Well over 1200 different restriction enzymes have been discovered in a wide variety of prokaryotes, (*see* compilation by Roberts *[24]*); among them, many are isoschizomers, i.e., enzymes that recognize the

same DNA sequence. They all cleave the DNA within or next to the recognition site producing "sticky" (with a 5' or 3' overhang) or "blunt" ends, viz.

	-GpTpCpGpApC-		-G	pTpCpGpApC-
*Sal*I		→	+	
	-CpApGpCpTpG-		-CpApGpCpTp	G-
			(sticky end with 5' overhang)	
	-CpGpApTpCpG-		-CpGpApT	pCpG-
*Pvu*I		→	+	
	-GpCpTpApGpC-		-GpCp	TpApGpC-
			(sticky end with 3' overhang)	
	-GpApTpApTpC-		-GpApT	pApTpC-
*Eco*RV		→	+	
	-CpTpApTpApG-		-CpTpAp	TpApG-
			(blunt end with no overhang)	(2)

Altogether, about 130 different specificities have been described. Different strains of individual species often contain different restriction enzymes: in *E. coli,* e.g., 141 restriction enzymes with 40 different specificities are known *(24)*. Few restriction enzymes have been shown to be part of a restriction/modification system. Even fewer have been purified to homogeneity and characterized biochemically. Most of our knowledge on the mechanism of action of restriction endonucleases rests on studies of these enzymes.

2.1. Purification

Bacteria that produce Type II restriction enzymes usually synthesize very small amounts of these proteins. Typically, between 1 and 10 mg of the homogeneous enzyme can be isolated/kg wet cell paste in a multistep purification scheme. The isolation procedure in general involves breaking up the cells by enzymatic or mechanical procedures, high-speed centrifugation, removal of nucleic acids by precipitation with polyethyleneimine or binding to DEAE-cellulose at an ionic strength that prevents binding of the restriction enzyme to the DNA, and a series of chromatographic steps. Phosphocellulose, hydroxyapatite, and heparin-agarose chromatography seem to be particularly suited, and have been widely employed, for the purification of restriction

enzymes. In general, two or three chromatographic steps yield restriction enzyme preparations that are devoid of nonspecific nucleases *(25,26)*. In combination, they might yield a homogeneous enzyme preparation. Quite often, however, it is necessary to add an affinity chromatography step on blue sepharose, which contains an immobilized Cibacron dye, as well as DNA-cellulose or variants thereof, as final purification steps in the isolation of restriction enzymes. Recently, an eicosomer (oligonucleotide) affinity matrix containing the recognition sequence of *Eco*RI was produced that allows the enzyme to be purified to near homogeneity in a single step starting from a crude cell extract *(27)*. It remains to be shown whether this procedure is of general utility. The possibility of cloning the genes for restriction enzyme into expression vectors has led to a considerable simplification of the isolation procedures for these enzymes. Typically, 100–1000 mg of homogeneous protein can be isolated from 1 kg of wet cell paste using two chromatographic steps only (e.g., *28,29*). Care has to be taken that overproduction does not lead to aggregation and inactivation *(30)* (*see* note added in proof at end of chapter).

2.2. Molecular Properties

To date about 30 Type II restriction enzymes have been sequenced (Table 2). Some of them show a high degree of homology, such as *Eco*RI and *Rsr*I *(47)*,*Taq*I and *Tth*HB81 (Zebala, Mayer, and Barany, personal communication), and *Bsu*BI and *Pst*I (Xu, Kapfer, Walter, and Trautner, personal communication). Others have stretches of homology, such as *Bcn*I and *Mva*I (Butkus, Padegimiene, Menkevicius, Butkiene, Timinskas, and Janulaitis, personal communication), as well as *Bsu*FI and *Msp*I *(31)*. It is noteworthy that these pairs of restriction enzymes are isoschizomers. This indicates that their common function at least in part is reflected in a common structure. A similar significance can be attributed to our finding that several restriction enzymes (*Dpn*II, *Eco*RV, *Hin*cII), which cleave the DNA within the sequence G-/-C or G— /—C where the dashes stand for As or Ts, have a common protein sequence motif—Ser Gly X X X Asn X Ile-. This sequence in *Eco*RV is involved in DNA recognition as shown by site-directed mutagenesis studies. In addition, *Eco*RV shows a stretch of sequence homology to *Sma*I. In *Eco*RV, this stretch corresponds to a region implicated in backbone contacts outside the recognition sequence. Otherwise,

Table 2
Structural Properties of Selected Restriction Enzymes

Enzyme	Subunit		Reference
	Amino acid res.	Mol wt, kDa	
*Bcn*I	211	24	Butkus, Padegimiene, Menkevicius, Butkiene, Timinskas, and Janulaitis, personal communication
*Bsu*BI	311	43	Xu, Kapfer, Walter, and Trautner, personal communication
*Bsu*FI	395	46	Kapfer et al., 1991 *(31)*
*Bsu*RI	576	66	Kiss et al., 1985 *(32)*
*Dde*I	240	28	Sznyter et al., 1987 *(33)*
*Dpn*I	254	30	Lacks et al., 1986 *(34)*
*Dpn*II	288	34	Lacks et al., 1986 *(34)*
*Eco*RI	276	31	Greene et al., 1981 *(35)*; Newman et al., 1981 *(36)*
*Eco*RII	402	45	Kossykh et al., 1989 *(37)*; Bhagwat et al., 1990 *(38)*
*Eco*RV	245	29	Bougueleret et al., 1984 *(39)*
*Fok*I	583	67	Kita et al., 1989 *(40)*; Looney et al., 1989 *(41)*
*Hha*II	227	26	Schoner et al., 1983 *(42)*
*Hinc*II	257	30	Waite-Rees, Moran, Slatko, Honstra-Coe, and Benner, personal communication
*Hinf*I	272	31	Chandrasegaran et al., 1988 *(43)*
*Hpa*I	254	30	Waite-Rees, Moran, Slatko, Honstra-Coe, and Benner, personal communication
*Msp*I	262	30	Lin et al., 1989 *(44)*
*Mva*I	259	30	Xu, Kapfer, Walter, and Trautner, personal communication
*Pae*R7	246	27	Theriault et al., 1985 *(45)*
*Pst*I	326	37	Walder et al., 1984 *(46)*
*Rsr*I	276	31	Stephenson et al., 1989 *(47)*
*Sau*96I	268	31	Szilak et al., 1990 *(48)*
*Sin*I	230	27	Karreman and de Waard, 1988 *(49)*
*Sma*I	247	29	Heidmann et al., 1989 *(50)*
*Taq*I	234	28	Slatko et al., 1987 *(51)*
*Tth*HB81	263	31	Zebala, Mayer, and Barany, personal communication

there is no sequence similarity over an extended stretch of amino acids. Lauster *(52)*, however, has detected intra- and intermolecular homologies comprising 10–20 amino acid residues: *Bsu*RI, *Pae*R7, and *Pst*I each contain a twofold repeat. *Bsu*RI in addition contains a fourfold repeat that is homologous to regions found also in *Eco*RI, *Eco*RV, and *Pae*R7. Most restriction enzymes are active as dimers of identical subunits with subunit mol mass ranging from approx 25,000 to 35,000 Da, exceptions being *Bsu*BI, *Bsu*FI, *Bsu*RI, *Eco*RII, and *Fok*I with mol mass of 43,000, 46,000, 66,000, 45,000, and 67,000 Da, respectively. The *Bsu*RI enzyme, which has been reported to be active as a monomer, might be—as suggested by sequence repetitions—a quasi dimer. The *Fok*I enzyme differs from most type II restriction enzymes, in as much as it recognizes an asymmetric sequence. A few restriction enzymes have been reported to be small and monomeric, e.g., *Bgl*I, *Bsp*RI, and *Pal*I *(53–55)*. These differences in quarternary structure may reflect major differences in reaction mechanisms, or that the state of aggregation depends critically on the absence or presence of the substrate.

The three-dimensional structure of two restriction enzymes, *Eco*RI and *Eco*RV, has been solved *(29,56;* Winkler unpublished). Work on crystallographic analysis of *Hha*II is in progress *(57)*. Although there is no similarity in the topological arrangement of the secondary structure elements of *Eco*RI and *Eco*RV, the overall shape is similar. In both cases, the two identical subunits are arranged such that a deep cleft is formed that—as is evident from the X-ray structure analysis of the protein DNA cocrystals—constitutes the DNA binding site. It is noteworthy that both *Eco*RI and *Eco*RV form complexes with twofold symmetry, which means that the two identical subunits of these restriction enzymes are engaged in the same set of interactions with the two halves of their palindromic recognition sequences, as originally suggested by Kelly and Smith *(58)*.

2.3. Reaction Mechanisms

The presence of Mg^{2+} ion is essential for the enzymatic activity of all restriction enzymes. For *Eco*RI, it was shown that other divalent cations can be used: Mn^{2+}, Co^{2+}, and Zn^{2+} *(59)*. Presumably, Mg^{2+} participates in the catalytic event by polarizing the phosphodiester bond to be cleaved and/or activating water to form the required nucleophile.

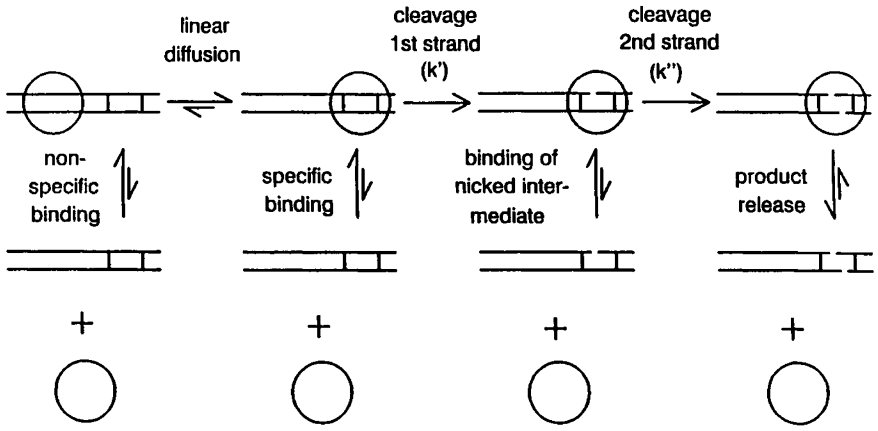


Fig. 1. Generalized scheme to illustrate the events following the association of a macromolecular DNA with a restriction enzyme.

Figure 1 depicts the reaction catalyzed by a restriction enzyme. It shows (exemplified for *Eco*RI) that this enzyme interacts not only with its recognition sequence, but also with lower affinity with nonspecific sequences *(60,61)*. This results in competitive inhibition of the enzyme. Assuming that Michaelis-Menten conditions are valid, the steady-state rate of reaction, v, is given by:

$$v = [V_{max} \cdot F \cdot C^{total}]/[K_m \cdot F + C^{total}] \qquad F = 1/[1 + L \cdot K_m/K_i] \qquad (3)$$

where V_{max} and K_m are the intrinsic maximum velocity and Michaelis constant, K_i, is the inhibitory constant of the nonspecific DNA, L is the relative concentration of nonspecific over specific DNA and C^{total} the sum of the total concentration of specific and nonspecific DNA *(62)*. This implies that the addition of DNA not containing a recognition sequence (as a matter of fact, single-stranded DNA and RNA have a similar effect!) to a DNA substrate leads in a concentration-dependent manner to an inhibition of the cleavage reaction *(62,63)*. On the other hand, nonspecific binding might under certain conditions help restriction enzymes to localize their specific recognition site by facilitated diffusion, provided nonspecific and specific sites are on the same DNA molecule. It has indeed been demonstrated that nonspecifically bound restriction enzymes slide along the DNA in an apparently random process until the specific sequence is reached, a process that

accelerates the rate of the cleavage reaction *(64–67)* by lowering the apparent K_m. Linear diffusion along the DNA may also be responsible for the processive cleavage of DNA in two (or more) sites by a restriction enzyme: Depending on reaction conditions, in particular the ionic strength, two sites close to each other may be cleaved in a processive manner, when the enzyme remains associated with the DNA after the first cleavage, rather than in a distributive manner, when the enzyme dissociates after the first cleavage *(68–70)*.

2.3.1. Affinity

Although Mg^{2+} ions are necessary for the phosphodiester bond cleavage, they are not needed for the strong binding of restriction enzymes to DNA (*see* note added in proof at end of chapter). The affinity of restriction enzymes for their substrate in the absence of DNA is very high: measurements carried out with macromolecular DNA indicate that, e.g., *Eco*RI binds to its site with an equilibrium constant of the order of 10^{10}–$10^{11} M^{-1}$, to nonspecific DNA with an equilibrium constant of the order of $10^6 M^{-1}$ *(71,72)*. Similar results have been obtained with dodekadeoxynucleotides or larger oligodeoxynucleotides *(72,73)*, but not, however, with smaller ones, which are not bound as firmly *(74)*. Obviously, oligodeoxynucleotides that are too small to fill the complete DNA binding site of the restriction enzyme are not bound with the maximum affinity. It is noteworthy that, for all restriction enzymes analyzed in this context, residues external to the recognition site are important for enzyme–DNA interaction. The strength of binding between *Eco*RI and DNA decreases with increasing ionic strength *(60,72,73)* demonstrating that electrostatic interactions are involved in complex formation. Although so far no precise data have been obtained for the affinity of restriction enzymes for their substrate in the presence of Mg^{2+} ions, presteady-state *(75)* as well as steady-state experiments indicate that restriction enzymes display a high affinity for their substrate. Apparent K_m values determined for eight restriction enzymes and high-mol-wt substrates range from 0.5–17 n*M* (Table 3). Although these values have been determined with different substrates, i.e., substrates with variable proportions of nonspecific and specific DNA sequences, as well as under considerably different buffer conditions and temperatures, they demonstrate that restriction enzymes in general have a very low K_m for high-mol-wt DNA. In contrast,

Table 3
Steady-State Parameters for DNA Cleavage by Restriction Enzymes

Enzyme	Substrate	K_M, n*M*	k_{cat}, min^{-1}	Reference
*Bam*HI	SV40	3.6	1.5	Nardone and Chirikjian, 1987 *(66)*
*Bgl*II	p2H1	–	1.75	Imber and Bickle, 1981 *(76)*
*Cau*I	ϕX174 RF	<1	1.7	Bennett, 1987 *(77)*
*Cau*II	ϕX174 RF	2.5	305	Bennett, 1987 *(77)*
*Eco*RI	SV40	30	3	Greene et al., 1975 *(78)*
	ColE1	8	4	Modrich and Zabel, 1976 *(79)*
	phage λ	10	1.3	Berkner and Folk, 1977 *(80)*
	pBR322	5	1.8	Langowski et al., 1980 *(68)*
	pMB9	2.5	7.7	Halford and Johnson, 1981 *(81)*
	pUC8	50	19	Wolfes et al., 1986 *(82)*
	ds d(TGAATTCA)	7000	4	Greene et al., 1975 *(78)*
	ds d(GGAATTCC)	200	18	Brennan et al., 1986 *(83)*
	ds d(CTGAATTCAG)	120	4	McLaughlin et al., 1987 *(84)*
	ds d(CGCGAATTCGCG)	130	0.68	McLaughlin et al., 1987 *(84)*
	ds d(TATAGAATTCTAT)	16	18	Alves et al., 1989 *(85)*
*Eco*RV	pAT153	<0.5	0.9	Halford and Goodal, 1988 *(86)*
	ds d(AAAGATATCTT)	4000	7	Fliess et al., 1988 *(87)*
*Hha*II	pSKII	17	11	Kaddurah-Daouk et al., 1985 *(88)*
*Mva*I	ds d(ACCTACCTGGTGGT)	210	0.45	Kubareva et al., 1988 *(89)*
*Rsr*I	pBR322	14	6.5	Aiken and Gumport, 1988 *(90)*
*Sal*GI	pMB9	0.5	0.1	Maxwell and Halford, 1982 *(91)*

oligodeoxynucleotide substrates have much higher K_m values *(78,83,84,87,92,93)*, presumably because they do not fill the complete DNA binding site of the enzyme, do not adopt the same conformation as when part of a high-mol-wt DNA molecule, and because of the absence of facilitated diffusion processes. Pohl et al. *(94)* have noticed that there is a correlation between K_m values and the length of the recognition sequence: Restriction enzymes recognizing a hexanucleotide sequence have on average a 10 times lower K_m than enzymes that recognize a tetranucleotide sequence.

2.3.2. Rate of Reaction

Restriction enzymes are remarkably slow enzymes, with turnover numbers (k_{cat}) around 1–10 min^{-1}. It has been argued that this is because of the physical constraint that k_{cat}/K_m cannot exceed the diffusion controlled limit for a bimolecular association rate constant *(14,95)*, which means that with a K_m of 1 n*M* k_{cat} cannot be significantly larger than 1 s^{-1}. As a matter of fact, single turnover experiments with *Eco*RI have shown that the intrinsic rate constants of phosphodiester bond cleavage (k' and k" in Fig. 1) come close to this limit *(83,96–98)*. The disparity between k_{cat} and k' or k" as measured in steady-state and presteady-state experiments with *Eco*RI and pBR322 indicates that it is not the phosphodiester bond cleavage that is rate limiting for the enzymatic turnover. Depending on experimental conditions, site localization, a conformational change of the enzyme, or product release may be the rate-determining step for the *Eco*RI catalyzed cleavage of DNA.

The temperature dependence of the activity of various restriction enzymes has been analyzed in terms of the apparent activation energy, E_a, for the enzymatic reaction and the apparent inactivation enthalpy, E_i *(94)*. An average value of 50 ± 23 kJ/mol for E_a and 190 ± 75 kJ/mol for E_i was determined. Most restriction enzymes show a temperature optimum for DNA cleavage between 25 and 45°C.

The rate of cleavage has been observed to be largely independent of the superhelicity of the substrate *(99)*, but to depend very much on the sequences adjacent to the recognition site. For example, the 5 *Eco*RI sites in λ DNA are cleaved with considerably different rates *(100–102)*. Similar observations were made with *Eco*RI sites in other DNA substrates *(103,104)*, as well as for other restriction enzymes, e.g., *Pst*I *(105)*. Studies with synthetic oligodeoxynucleotides that differ in their

sequences adjacent to the recognition site have shown for *Eco*RI that AT base pairs enhance and GC base pairs slow down the hydrolytic activity of *Eco*RI, presumably because the structure of the recognition sequence or its propensity to adopt a certain conformation is modulated by the flanking sequences such that it becomes a better or worse substrate for *Eco*RI *(106)*. Similar mechanisms are likely to operate with all restriction enzymes *(107)*.

2.3.3. Order of Reaction

In Fig. 1, the DNA cleavage reaction catalyzed by a restriction enzyme is formulated as two partial reactions that lead to the hydrolysis of a phosphodiester bond in each of the two individual strands of the DNA duplex. Depending on the DNA substrate and the reaction conditions, these two partial reactions can be separated kinetically, when the rate of cleavage of the first phosphodiester bond is smaller than the rate of dissociation of the enzyme from the nicked substrate, or the two partial reactions take place in a concerted fashion, when the rate of cleavage of the second phosphodiester bond is higher than the rate of dissociation *(75,108–111)*. In general, restriction enzymes follow a concerted reaction mechanism under optimum reaction conditions, unless the base composition of the sequences flanking the recognition site is very different. In these cases, the rates of cleavage of the first and second strand can become very different, such that the enzyme has an increased probability to dissociate from the nicked intermediate. DNA substrates in which the symmetrical recognition site for a restriction enzyme has been made unsymmetrical by introducing phosphorothioate groups in one strand *(112,113)* or a single mismatch base pair *(114)* are cleaved much better in the unmodified than in the modified strand, which allows the isolation of a nicked intermediate. The preferential attack of the unmodified strand in double-stranded DNA containing phosphorothioates by restriction enzymes has been taken advantage of to remove the template single strand by exonuclease digestion for in vitro mutagenesis experiments.

Phosphorothioate substituted oligodeoxynucleotides have been used to elucidate the stereochemical course of the *Eco*RI catalyzed cleavage reaction: Hydrolysis of the R_p diastereomer of d(pGGsAATTCC) proceeds with inversion of configuration at phosphorous, suggesting that hydrolysis occurs by a direct nucleophilic attack of water at the

scissile phosphodiester bond, rather than with an involvement of a covalent enzyme intermediate *(115)*. It remains to be established whether this is true for other restriction enzymes (*see* note added in proof at end of chapter). Since some nucleases show retention of configuration at phophorous, and others show inversion of the configuration, it might well be that not all restriction endonucleases utilize the same mechanism.

2.4. Specificity

The specificity of restriction endonucleases for their recognition sites is very high. Under optimum reaction conditions, even sites that differ in only one base pair from the canonical site are not cleaved, unless large enzyme concentrations are used or the reaction is allowed to proceed for sufficient time. For example, a standard pUC8 DNA digestion carried out with micromolar concentrations of *Eco*RI leads to cleavage within seconds not only at the canonical sequence -GAATTC-, but after several hours of incubation also at two other sites, -TAATTC- and -GAGTTC-; two other sequences in pUC8 that also differ in only one base pair from the canonical site, -TAATTC- and -GAATAC-, however, are refractory to cleavage (Pingoud and Alves, unpublished). Similarly, *Eco*RV cleaves not only its recognition sequence, -GATATC-, but also, at very small rates, several alternative DNA sequences, e.g., in pAT153 -GAAATC- *(116)* and -GTTATC- *(117)*. A detailed study using a series of nine tetradekadeoxynucleotides that contain all possible single base pair substitutions of the *Eco*RI recognition site has shown that, depending on the position within the sequence and the kind of substitution, the oligonucleotides are cleaved more or less readily, with rates varying from 1/1000 to $<1/10^6$ of the rate with which the canonical sequence is cleaved *(114)* (*see* note added in proof at end of chapter).

Considerable progress has been made in recent years in understanding the molecular basis of the specificity of restriction enzymes, mainly because of chemical modification studies, in which the structural elements of the DNA substrates needed for efficient cleavage were identified (ref. 5–64 in *[87]*). In spite of the intrinsic problem of differentiating between direct and indirect effects caused by the chemical modification of the oligodeoxynucleotide on the acceptance as a substrate by the enzyme, it is clear from these studies that a lot but not all of the structural information present in the sequence of bases that constitute

the recognition site is used by the enzyme to differentiate between specific and nonspecific DNA *(118)*. For example, *Eco*RV needs the presence of the thymin methyl groups in its recognition sequence -GATATC- for a productive interaction, whereas *Eco*RI accepts oligodeoxynucleotides as substrates that contain deoxyuridine in place of thymidine in the recognition sequence -GAATTC- *(83,87,119–121)*. Studies with modified oligonucleotides have also shown that isoschizomers do not necessarily have the same structural requirements for DNA cleavage *(89,107,122)*. Furthermore, these studies have demonstrated that some restriction enzymes fail to cleave DNA modified by methylation at a position in the recognition sequence other than the one modified by the corresponding modification methyltransferase. For example, *Eco*RV does not cleave DNA containing the sequences -G^{m6}ATATC- or -GAT^{m6}ATC- *(87)*, although only the former one is formed in vivo by the action of the *Eco*RV methylase *(123)*.

The most detailed insight into the understanding of sequence specificity has been provided by the X-ray structure analysis of an *Eco*RI-oligodeoxynucleotide complex *(56)*. This complex was crystallized in the absence of Mg^{2+}, the essential cofactor for catalysis and, therefore, does not represent a *bona fide* enzyme substrate complex. Nevertheless, the structure analysis provides sufficiently detailed information to propose models for the mechanism of recognition. According to the suggestion of McClarin et al. *(56)*, the specificity of *Eco*RI is based on 12 precisely directed hydrogen bonds, six from each of the two identical subunits to the recognition sequence -GAATTC-: Arg200 is supposed to form two hydrogen bonds to guanine; Glu144 and Arg145 supposedly interact with the two adenine residues via two hydrogen bonds each (*see* note added in proof at end of chapter). Site-directed mutagenesis experiments, however, in which these amino acid residues of *Eco*RI have been replaced suggest that the cleavage specificity of *Eco*RI is determined by factors in addition to direct hydrogen bonding to the bases of the recognition sequence *(82,85,124–126)*. Since the precise geometry of the phosphodiester backbone is determined by the sequence, also backbone contacts may contribute to specificity *(127)*. It seems that sequence discrimination by *Eco*RI is redundant. Presumably this is true for restriction enzymes in general, since only an overdetermined discrimination is fail safe.

2.4.1. Methylation

Methylation of critical residues of the recognition sequence in general prevents a specific interaction between restriction enzymes and DNA. In vivo, modification enzymes protect the host DNA through methylation against cleavage by the corresponding restriction enzyme. Methylation usually occurs at the amino group at position 6 of an adenine residue (^{m6}A) or at position 5 of a cytosine residue (^{m5}C), and rarely at position 4 of a cytosine residue (^{m4}C) (for a review, cf *[128]*). Methylation renders the DNA resistant to cleavage at the modified site. Methylation at other positions within the recognition sequence by other methyltransferases, including *dam*, *dcm*, and eukaryotic methyltransferases, may or may not affect the rate of DNA cleavage by a restriction enzyme. For example, DNA containing the site -AGGCCTGG- is not cleaved by *Aat*I or *Stu*I when the DNA is extracted from a *dcm*$^+$ strain of *E. coli*, since in this case, the recognition site for *Aat*I or *Stu*I overlaps with a *dcm* methylation site *(129)*. Although in this case isoschizomers exhibit the same sensitivity, others do not, e.g., *Hpa*II is sensitive to methylation within its recognition sequence -CCGG-, whereas *Msp*I is not *(130)*, or only to a very minor degree *(131)*. This differential sensitivity of *Hpa*II and *Msp*I can be used to analyze the methylation state of genes of higher eukaryotes, which depending on the transcriptional activity are more or less methylated at -CG- sequences.

Likewise, the isoschizomers *Mbo*I, *Sau*3AI, *Dpn*I, and *Dpn*II cleave DNA within the *dam* methylase recognition site -GATC-, but differ with respect to their sensitivity toward methylation: *Mbo*I and *Dpn*II cleave only -GATC-, and *Sau*3AI recognizes both -GATC- and -G^{m6}ATC-, whereas *Dpn*I only cleaves -G^{m6}ATC- *(132)*. The effects of site-specific methylation on restriction as well as modification enzymes have been reviewed *(128)* and will, therefore, not be discussed in more detail. It should be emphasized, however, and this will be mentioned later (Section 3.7.), that the combined use of modification and restriction enzymes can be utilized to narrow down the specificity of a restriction enzyme.

2.4.2. "Star" Activity

The specificity of restriction enzymes can be relaxed by suboptimum buffer conditions. This phenomenon, defined as "star" activity, was first observed by Polisky et al. *(133)*, who noticed that, under

conditions of elevated pH and low ionic strength, *Eco*RI cleaves DNA not only within the canonical recognition sequence -GAATTC-, but also at other sites. A variety of other conditions reduce the sequence specificity of *Eco*RI, including the presence of Mn^{2+} instead of Mg^{2+} *(59,134)* and the presence of organic solvents, such as glycerol or dimethyl sulfoxide *(135–138)*. It was suggested that this "*Eco*RI* activity" was directed against sites containing the sequence -AATT-. Later, it was deduced by Gardner et al. *(139)* that under "star" conditions *Eco*RI recognizes sites that differ in a single position from the canonical *Eco*RI sequence; any substitution can occur, with the exception of A → T or T → A changes within the central tetranucleotide.

Several other restriction enzymes have been shown to exhibit "star" activities, e.g., *Ava*I *(140)*, *Bam*HI *(138,141)*, *Bst*I *(142)*, *Bsu*I *(143)*, *Eco*RV *(144)*, *Hae*III *(140)*, *Hha*I *(138)*, *Hin*dIII *(134,145)*, *Hpa*I *(140)*, *Pst*I *(138)*, *Pvu*II *(146)*, *Sal*I *(138,140)*, *Sst*I *(138,140)*, *Sst*II *(140)*, *Taq*I *(147)*, and *Xba*I *(138,140)*. For *Bam*HI, *Eco*RI, *Eco*RV, *Hin*dIII, *Pvu*II, and *Taq*I the "star" site cleavage sequences have been determined to differ from the normal recognition sequence only by a single base pair.

"Star" patterns have been observed in restriction digests also under optimum buffer conditions when high concentrations of the restriction enzymes were used; the kinetics of cleavage are different, however, since suboptimal buffer conditions decrease the high accuracy of restriction enzymes by lowering the rate of cleavage at canonical sites and increasing the rate of cleavage at degenerate sites, whereas at high concentrations of enzyme under optimum buffer conditions, cleavage at canonical sites occurs with a very much higher rate than at degenerate sites. It has been shown that the presence of spermine and spermidine in the reaction buffer effectively suppresses "star" activity of *Bam*HI, *Bsu*RI, *Eco*RI, *Eco*RV, *Hin*dIII, *Pst*I, and *Sal*I *(148)*.

2.4.3. Single-Strand Cleavage

Several restriction enzymes were reported to cleave single-stranded DNA, for example, *Hae*III *(149)*, *Hha*I and *Sfa*I *(150)*, and *Hae*III, *Hha*I, *Hin*fI, *Hpa*II, and *Mbo*II *(151)*. Detailed mechanistic studies, however, demonstrated that most likely "single-strand" cleavage occurs at transiently formed double-stranded sites *(152)*. This is presumably also true for oligodeoxynucleotide substrates; the cleavage of d(GAACCGGAGA)

or d(TCTCCGGTT) by *Msp*I *(153)* may be owing to duplex formation via the palindromic -CCGG- core sequence, similarly described for the cleavage of "single-stranded" oligodeoxynucleotides by *Eco*RI *(106)*. In a more recent report, Bischofsberger et al. *(154)* concluded from cleavage experiments carried out with immobilized oligodeoxynudeotides that *Eco*RI is capable of cleaving single strands. Since it cannot be completely ruled out that even oligodeoxynucleotides covalently linked to oligo-dT-cellulose can in part form duplex molecules via the palindromic recognition site, this conclusion should be met with reservation. Similarly, the reported cleavage of DNA in DNA × RNA hybrid double strands by *Eco*RI, *Hin*dII, *Sal*I, *Msp*I, *Hha*I, *Alu*I, *Taq*I, and *Hae*II *(155)* presumably can be explained by normal DNA cleavage, since the substrate was produced by the reverse transcription of a viral RNA by AMV polymerase, was not characterized with respect to a DNA × RNA hybrid structure and was not analyzed as to what extent it was contaminated with normal DNA double strands, which are also produced by AMV polymerase.

2.5. Inhibition

The cleavage of DNA by restriction enzymes can be inhibited by covalent modification of the substrate or the enzyme, as well as by complex formation with low-mol-wt ligands. Chemical modification of restriction enzymes has been used to characterize their active sites: *Bgl*I catalyzed cleavage of DNA is inhibited by 2,3 butane-dione modification of Arg residues, DNA binding remaining unaffected *(53)*. For *Eco* RI, it was shown that the modification of Lys residues with methyl acetimidate inactivates the enzyme; the relevant Lys residues are protected against chemical modification by complex formation with the substrate *(67)*. It was also shown that Glu residues are essential for catalysis of *Eco*RI, since modification with 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-*p*-sulfonate inactivates the enzyme *(156)*. Rosebengal-sensitized photooxidation of His residues leads to the inactivation of *Eco*RI, but not *Eco*RV *(157)*. In contrast, although modification of a single sulfhydryl group by a fluorescent maleimide derivative leads to the total loss of activity in *Eco*RV, it does not affect the activity of *Eco*RI *(158)*. In a detailed study, Nath *(159)* has analyzed the effect of several sulfhydryl reagents on restriction enzyme activities: *Sal*I, *Bgl*II, *Mpa*I, *Eco*RI, and *Sst*II proved to be insensitive

to *p*-mercuribenzoate, *N*-ethyl maleimide, 5,5'-dithiobis (2-nitrobenzoic acid), and iodoacetate, whereas *Hin*dIII, *Bam*HI, *Pvu*I, *Ava*I, *Sma*I, and *Pst*I could be inhibited by some, but not all of the tested sulfhydryl reagents.

The restriction enzyme catalyzed cleavage of DNA can be inhibited by ligands that specifically bind to DNA such as certain antibiotics, (distamycinA, netropsin, actinomycinD, anthramycin, olivomycin), carcinogens (6,4'-diamidino-2-phenylindole), cytostatic drugs (*cis*-dichlorodiamineplatinumII), and dyes (proflavine, ethidium bromide, Hoechst 33258) *(74,160–171)*. This inhibition is dependent on the recognition sequence and its flanking region, which reflects the preferential binding of these drugs to AT- or GC-rich regions. It is, therefore, possible in some cases to suppress DNA cleavage by a restriction enzyme at one site without affecting the cleavage at another site. The effective concentration for inhibition may be different for single-strand ("nicking") and double-strand cleavage *(171)*, an effect that has been taken advantage of to study the kinetics of DNA cleavage *(81,172)* and may be used to prepare open circular DNA with a specific phosphodiester bond cleaved. Sequence-specific binding of oligopyrimidine nucleotides to double-stranded DNA, which results in triple helix formation, can be used, as was shown for *Ksp*632I *(173)* and *Eco*RI *(174)*, to inhibit restriction endonuclease catalyzed cleavage of DNA at particular sites.

3. Practical Considerations for the Use of Restriction Enzymes

3.1. Storage of Restriction Enzymes

The stability of preparations of purified restriction enzymes critically depends on adequate storage conditions, in particular, on buffer composition, enzyme concentration, presence of additives, and an appropriate storage temperature.

Different buffer compositions are recommended for the storage of different enzymes: For most restriction enzymes 10–50 m*M* Tris-HCl, pH 7.5, a minimum of 50 m*M* NaCl or KCl, 1 m*M* EDTA, and 60% (v/v) glycerol have proven to be optimal. Many restriction enzymes require additional additives, most commonly 1 m*M* 1,4-dithiothreitol, 1,4-dithioerythritol or β-mercaptoethanol (to prevent oxidation of

cysteine residues), 0.01–0.1% (w/v) Triton X-100, Tween, Lubrol, or other detergents, as well as 0.1 mg/mL nuclease-free bovine serum albumin (to prevent aggregation and precipitation).

Restriction enzymes should be stored in unfrozen solution at temperatures below 0°C, preferably at –20°C in a 60% glycerol solution. If stored frozen at –70°C, repeated thawing and freezing should be avoided (the heat generated during the defrost cycles of frost-free freezers may adversely affect the stability of a restriction enzyme preparation!)*(175,176)*.

3.2. Definition of Specific Activity

The standard activity measurement for restriction enzyme preparations is based on the determination of the minimum amount of enzyme required for the complete digestion of a standard DNA substrate, such as bacteriophage λ DNA. Some restriction enzymes require λ DNA isolated from dam⁻ or dcm⁻ *E. coli* strains, λ DNA fragments, or other DNA substrates. It must be emphasized that the activity of a given restriction enzyme is dependent on the substrate, mainly because of the influence of sequences flanking the recognition site. This is particularly evident with many restriction enzymes that recognize sites composed of GC base pairs, such as *Nae*I, *Nar*I, *Sac*II, and *Xma*II. *Nae*I, for example, is known to cleave pBR322 DNA at positions 401, 769, and 946 more than 10 times faster than at position 1283, which is cleaved as slowly as the single site in λ DNA. With a given DNA substrate, the rate of cleavage may depend on the topological state: *Eco*RI, for example, attacks supercoiled pBR322 DNA more slowly than linear or open circular BR322 DNA *(99)*.

One unit of a given restriction enzyme is defined as the activity that cleaves 1 µg of λ DNA in 1 h under optimum buffer conditions at (normally) 37°C. A considerable number of enzymes have abnormal temperature optima, e.g., 25°C: *Sma*I, *Bcl*I (at pH 8.5); 30°C: *Apa*I; 55°C: *Bsp*MII, *Bcl*I (at pH 7.5); 65°C: *Bsp*MI, *Taq*I. Volume activity refers to the concentration of the enzyme, i.e., U/µL.

Many restriction enzymes show a concentration-dependent change in specificity: Loss of activity at low enzyme concentrations may be owing to adsorption of protein at the wall of the test tube. At high concentration, aggregation may be responsible for an apparent lower activity. For one restriction enzyme, *Eco*RII, a dependence of the

Table 4
Resolution of Different Gel Systems for Electrophoresis of DNA Fragments

Separation range of linear DNA in bp	Appropriate gel system
5–10	20%PAA
10–50	15%PAA
50–100	10% PAA
100–200	5% PAA
200–1000	3% PAA
100–500	3% Agarose
200–3000	2% Agarose
500–6000	1% Agarose
1000–20,000	0.5% Agarose
5000–60,000	0.3% Agarose
50,000–several million	1% Agarose (FIGE)

specific activity on substrate concentration was observed: It was demonstrated that cleavage of refractory DNA recognition sites could be achieved by addition of DNA that contains an abundance of *Eco*RII-sensitive sites. It was argued that *Eco*RII may be the prototype restriction endonuclease that requires at least two simultaneously bound substrates for its activation *(177)*. A similar observation was recently made for *Nae*I *(178)*. It was shown that some *Nae*I sites are cleaved rapidly and that some are almost totally resistant to cleavage. *Nae*I could be activated to cleave these resistant sites in the presence of excess cleavable *Nae*I sites. Interestingly, with spermidine present, resistant sites were cleaved rapidly and cleavable DNA inhibited their cleavage.

3.3. Assay for Restriction Endonuclease Activity

The earliest attempts to determine the activity of a restriction enzyme made use of the drop in viscosity of DNA solutions on nucleolytic degradation of the DNA *(5)*. This rather insensitive and imprecise method was subsequently replaced by the analysis of the reaction products obtained from restriction enzyme catalyzed digests of DNA by gel electrophoresis in which fragments are separated according to size *(179)* (for the separation properties of different electrophoresis gels, *see* Table 4). DNA bands are usually visualized by staining the gel, preferentially *after* electrophoresis, in a solution of ethidium bromide, which binds to DNA by intercalation and thereby increases its fluorescence quantum yield about 20-fold. As little as nanogram

quantities of DNA can be seen on inspection of the gel under UV light. Silver staining allows the detection of subnanogram quantities of DNA. Alternatively, radioactively labeled DNA substrates may be used: After digestion and electrophoresis, DNA bands are then visualized by autoradiography. The detection limit of this procedure is in the picogram range with ^{32}P-labeled DNA.

There have been efforts to devise spectrophotometric assays to measure the activity of restriction enzymes. One assay makes use of DNA immobilized on cellulose, which on incubation with restriction enzymes leads to the release of DNA fragments into the solution and a concomitant increase in UV absorbance *(180)*. Although this assay is less sensitive than the electrophoretic procedure, it allows a continuous recording of the progress of the reaction. This is also possible with a fluorimetric assay, which takes advantage of the fact that superhelical DNA binds less ethidium bromide than nicked circular or linear DNA *(81,172)*. Therefore, cleavage of superhelical DNA in the presence of ethidium bromide leads to increased ethidium binding and fluorescence. Both the spectrophotometric and the fluorimetric assays suffer from the disadvantage that they do not discriminate between specific and nonspecific endonucleolytic activities (*see* note added in proof at end of chapter).

The following procedure may be used as a standard assay for most restriction enzymes:

- Incubate a solution of 250 ng λ DNA in 10 µL of the appropriate reaction buffer at the desired reaction temperature;
- Add 5 µL of a defined dilution of the restriction enzyme, freshly prepared with reaction buffer;
- Incubate 15 min at the desired temperature;
- Stop the reaction by the addition of 5 µL of a solution, containing 250 m*M* EDTA, pH 8.0, 1.2% (w/v) SDS, 25% (w/v) sucrose, and 0.2% (w/v) bromophenol blue;
- Incubate at 70°C for 5 min;
- Chill on ice;
- Load 15 µL of the incubation mixture onto 1–2% agarose gels; and
- Carry out electrophoresis.

The agarose concentration for the electrophoresis gels depends mainly on fragment length produced by each enzyme: e.g., *Eco*RI-λ-fragments (ranging from 3530–21,230 bp) are resolved in 1% agarose

gels, whereas complete *Eco*RV-λ-fragments (ranging from 270–5770 bp) require separation on at least 1.5% agarose gels (*see* Table 4). Volume activity (*see* Section 3.2.) toward λ DNA is calculated by determining the lowest enzyme dilution that allows complete cleavage of all recognition sites to yield the characteristic cleavage pattern for the particular enzyme.

It should be kept in mind that the volume activity measured with λ DNA does not directly indicate how much enzyme is required for the complete cleavage of another DNA substrate. Apart from influences concerning DNA conformation and flanking sequences discussed above, it is necessary to consider the concentration of sites for each substrate. For example: λ DNA with a total length of 48,500 bp has five *Eco*RI sites, i.e., 1 site/9700 bp, pBR322 with a total length of 4363 bp harbors only one *Eco*RI site, i.e., 1 site/4363 bp. Therefore, in pBR322, there are 2.2 times more *Eco*RI sites/μg DNA than in λ DNA and more enzyme will be required. The exact amount of additional enzyme will depend on the kinetic properties of the enzyme.

3.4. Optimum Reaction Conditions

The specific activity and accuracy of restriction enzymes mainly depend on appropriate reaction conditions with respect to pH, ionic strength, cofactors, stabilizing additives, reaction time, and temperature. Optimum reaction conditions are normally indicated by the enzyme suppliers. Table 5 summarizes reaction conditions regarded as optimum for the commercially available enzymes.

3.4.1. Buffer pH

Whereas most restriction enzymes work well in the pH range from 7.2–8.0, some enzymes have more narrow pH optima. For example, the activity of *Aat*II significantly decreases outside the pH range 7.4–7.8. In several cases, it has been observed that high pH (>8.5) may induce star activity (e.g., *Eco*RI, *Dde*I). The recommended reaction pH refers to the optimum reaction temperature. The pH of the reaction buffers, therefore, should be adjusted at this temparature. Often the use of a 20–50 m*M* Tris-HCl solution is recommended as buffer. The Tris-HCl system, however, is markedly temperature-dependent because of a ΔpK_a/°C of –0.031, which results in a pH shift of approx 0.25 pH U/10°C. Also, not all pH electrodes are suitable for use with Tris buffers.

3.4.2. Ionic Strength

The accuracy and activity of restriction enzymes are strongly affected by the ionic strength of the reaction milieu. The required ionic strength is achieved by the addition of NaCl or KCl to the Tris-HCl buffer. Whereas most restriction enzymes accept both KCl or NaCl some enzymes strongly prefer potassium salts (e.g., *Sma*I). It must be emphasized that low ionic strength may induce star activity (e.g., *Bam*HI, *Eco*RI, *Eco*RV, *Xba*I), and that high salt concentrations may activate contaminant nonspecific endonuclease activities or inhibit the restriction enzyme.

Some suppliers (i.e., AGS, Amersham, BRL, Boehringer, Stratagene) have elaborated simplified buffer sets to be used with their enzymes. Using only three different types of buffers, however, as recommended for example by Davis et al. *(18)* (i.e., a low, a medium, and a high salt buffer), does not meet the requirements of all enzymes and results, therefore, in many cases in a lower specific activity. McClelland et al. *(181)* and Pharmacia-LKB have developed one buffer system to be used with all enzymes. The potassium glutamate buffer "KGB" and the "One-Phor-All PLUS" buffer must be diluted before use according to the specific requirements of several enzymes. Although the simplified buffer systems in general *cannot be recommended* for the economic use of restriction enzymes, they allow the easy codigestion of DNA with several enzymes. However, if different ionic conditions are essential for site-specific cleavage with several restriction endonucleases, the DNA should be digested consecutively, beginning with the enzyme that has its optimum activity at the lower ionic strength. When consecutive digestion is not possible because of incompatibility of the reaction conditions, the fragments generated by the first cleavage reaction must be purified by phenol/chloroform extraction and ethanol precipitation prior to incubation with the second enzyme. The double digestion of multiple cloning sites in plasmid vectors with two different restriction enzymes may cause severe problems; it is often not possible to digest overlapping or directly neighboring restriction sites in polylinkers.

3.4.3. Essential Cofactors

Mg^{2+} ions are an absolute requirement for the catalytic activity of restriction enzymes. Chelating agents, such as EDTA, therefore inhibit the nucleolytic activity, whereas specific DNA binding is not disturbed (*see*

Table 5
Commercially Available Restriction Endonucleases

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Aat*I	*Eco*147I, *Stu*I	AGG/CCT	. + . +	7.5	60(KCl)	37	.	.	–+	6 0 7	>100 m*M* salt and >10% glycerol decrease activity, activity varies depending on substrate
*Aat*II		GACGT/C	....	7.5	50(KCl)	37	.	.	+	10 1 0	Activity decreases sharply outside the optimal pH range: 7.4–7.8
*Acc*I		GT/MKAC	. . . +	7.5	10	37	5	.	–	9 2 1	>50 m*M* salt decreases activity, at 55°C fivefold increased activity, nicking occurs slowly in the unmethylated strand of a hemi-methylated sequence GTMKA^{m5}C
*Acc*II	*Bce*RI, *Bsp*50I, *Bst*UI, *Mvn*I, *Tha*I	CG/CG	. . + +	7.5	25	37	.	.	–	157 23 0	

*Acc*III	*Bsp*MII, *Kpn*2I, *Mro*I	T/CCGGA	+.−−	7.7	150	60	.	.	.	24 1 0	
*Acy*I	*Aha*II, *Aos*II, *Bbi*II, *Hgi*DI, *Hin*1I	GR/CGYC	....	8.0	100	37	.	.	.	40 6 0	At 50°C increased activity
*Afl*I	*Ava*II, *Eco*47I, *Sin*I	G/GWCC	....	7.7	50	37	.	.	.	35 8 6	
*Afl*II	*Bfr*I	C/TTAAG	..−−	8.0	50	37	.	.	.	3 0 1	
*Afl*III		A/CRYGT	....	7.5	150	37	.	.	.	20 1 0	
*Aha*I	*Apy*I, *Bsp*NI, *Bst*GII, *Bst*NI, *Mva*I	CC/WGG	....	7.5	25	37	.	.	.	114 10 0	
*Aha*II	*Acy*I, *Aos*II, *Bbi*II, *Hgi*DI, *Hin*1I	GR/CGYC	..++	8.0	100	37	.	.	−	40 6 0	<100 m*M* salt decreases activity, GRGC^{m5}CC is cut faster than GRGCCC
*Aha*III	*Dra*I	TTT/AAA	..−−	7.7	0	37	.	.	.	13 3 12	<50 m*M* salt decreases activity
*Alu*I		AG/CT	...+	7.6	50	37	5	.	+	143 16 34	<50 m*M* salt decreases activity, activity varies depending on substrate
*Alw*21I	*Asp*HI, *Hgi*AI	GWGCW/C	....	7.5	120	37	.	.	.	28 8 0	
*Alw*26I		5'GTCTCN/3' 3'CAGAGN$_5$/5'	....	9.0	100	37	.	+	.	37 3 2	
*Alw*44I	*Apa*LI, *Sno*I	G/TGCAC	....	7.5	20	37	.	.	.	4 3 0	
*Alw*I		5'GGATCN$_4$/3' 3'CCTAGN$_5$/5'	+...	7.4	0	37	.	.	+	58 12 6	Fragments are poorly ligated
*Alw*NI		CAGNNN/CTG	.+..	7.4	50	37	.	.	+	41 1 2	

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Aoc*I	*Axy*I, *Bsu*36I, *Cvn*I, *Eco*81I, *Eco*O109, *Mst*II, *Sau*I	CC/TNAGG	....	7.7	150	37	.	.	.	2 0 0	
*Aoc*II	*Bsp*1286I, *Nsp*II, *Sdu*I	GDGCH/C	....	7.5	50	37	.	.	.	38 10 4	
*Aos*I	*Avi*II, *Fdi*II, *Fsp*I, *Mst*I	TGC/GCA	....	7.5	80(KCl)	37	.	.	.	15 4 0	
*Aos*II	*Acy*I, *Aha*II, *Bbi*II, *Hgi*DI, *Hin*1I	GR/$^{+}$CGYC	..+.				.	.	.	40 6 0	
*Apa*I		GGGCC/*C^{+}	..++	7.4	0	30	5	.	+	1 0 1	The buffer should contain 0.01% Triton, 50 m*M* KCl can be used, >50m*M* salt decreases activity, at 37°C decreased activity
*Apa*LI	*Alw*44I, *Sno*I	G/TGCA^{0}C	....	7.5	0	37	.	.	+	4 3 0	>25 m*M* salt decreases activity
*Apy*I	*Aha*I, *Bsp*NI, *Bst*GII, *Bst*NI, *Mva*I	CC/$^{+0}$WGG	.−..	7.5	0	30	.	+	.	71 6 16	The buffer should contain 0.01% Triton
*Aqu*I	*Ava*I, *Avr*I, *Nsp*7524III, *Nsp*III	*C/YCGRG	....	7.5	0	37	.	.	.	8 1 0	

*Ase*I	*Asn*I	AT/TAAT	. . – –	7.5	150(KCl)	37	.	.	.	17 1 3	
*Asn*I	*Ase*I	AT/TAAT	. . – –	8.0	100	37	.	.	.	17 1 3	
*Asp*700	*Xmn*I	GAANN/NNTTC	. . . .	7.5	50	37	.	.	.	24 2 0	
*Asp*718		$\overset{0}{\text{G}}/\text{GTAC}\overset{+}{\text{C}}$	. + + +	8.5	75	37	.	.	.	2 0 1	
*Asp*HI	*Alw*21I, *Hgi*AI	GWGCW/C	. . . .	8.0	100	37	.	.	.	28 8 0	
*Asp*I	*Tth*111I	GACN/NNGTC	. . . .	8.0	100	37	.	.	.	2 1 0	
*Asu*I	*Cfr*13I, *Nsp*IV, *Sau*96AI, *Sdy*I	G/GNCC	. . . .	7.6	50	37	.	.	.	74 15 11	
*Asu*II	*Bsp*119I, *Bst*BI, *Csp*45I, *Lsp*I, *Nsp*7524V, *Nsp*V	$\overset{0}{\text{TT}}/\text{CGAA}$	. . – –	7.4	20	37	.	.	+	7 0 0	>50 m*M* salt decreases activity
*Ava*I	*Aqu*I, *Avr*I, *Nsp*7524III, *Nsp*III	$\overset{+}{\text{C}}/\text{Y}\overset{0}{\text{C}}\overset{+}{\text{G}}\text{R}\overset{+}{\text{G}}$	. . + .	8.0	50	37	5	+	–+	8 1 0	The buffer should contain 0.01% Triton, largely insensitive to salt conc., at 45°C increased activity, >10% glycerol decreases activity, nicking occurs slowly in the unmethylated strand of the hemimethylated sequence CYCG^{m6}AG
*Ava*II	*Afl*I, *Eco*47I, *Sin*I	$\text{G}/\text{GWC}\overset{++}{\text{C}}$	. + + +	8.0	50	37	3	.	+	35 8 6	The buffer should contain 0.01% Triton, needs clean DNA for good activity, >50 m*M* salt decreases activity

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Ava*III	*Eco*T22I, *Nsi*I	ATGCA/T	....	8.0	100	37	.	.	.	14 0 3	
*Avi*II	*Aos*I, *Fdi*II, *Fsp*I, *Mst*I	TGC/GCA	....	7.5	150	37	.	.	.	15 4 0	
*Avr*I	*Aqu*I, *Ava*I, *Nsp*7524III, *Nsp*III	C/YCGRG	....				.	.	.	8 1 0	>100 mM salt decreases activity
*Avr*II		C/CTAGG	....	7.4	50	37	.	.	–	2 0 2	
*Axy*I	*Aoc*I, *Bsu*36I, *Cvn*I, *Eco*81I, *Eco*O109, *Mst*II, *Sau*I	CC/TNAGG	....	7.5	100	37	.	.	.	2 0 0	
*Bal*I	*Msc*I	TGG/CCA	.+..	8.0	0	37	5	.	+	18 1 0	Needs clean DNA for good activity, at 25°C increased stability, >25 m*M* salt decreases activity, cleavage in *dcm* methylated DNA is 50-fold slower than in unmethylated DNA
*Bam*HI	*Bst*I	G/GATCC[i]	−−..	7.9	100	37	.	+	–	5 1 1	Needs clean DNA for good activity, >100 m*M* salt decreases activity

*Ban*I	*Eco*64I	$G\overset{0}{}/GYRCC$	....	7.4	20	50	.	.	+	25 9 1	>50 m*M* salt decreases activity, *Ban*I gives various rate effects when the recognition sequence is ^{m5}C-methylated at different sites
*Ban*II	*Eco*24I	$GRG\overset{+}{}CY/C$	....	7.4	50	37	.	.	+	7 2 2	Largely insensitive to salt concentration, >5% glycerol decreases activity
*Ban*III	*Bsc*I, *Bsp*XI, *Bsu*15I, *Cla*I	$AT/CG\overset{+}{}AT$	+...	7.5	80(KCl)	37	.	.	+	15 1 0	>3%glycerol and <50 m*M* salt decreases activity
*Bbe*I		GGCGC/C	....	7.1	0	37	.	.	.	1 4 0	
*Bbi*II	*Acy*I, *Aha*II, *Aos*II, *Hgi*DI, *Hin*1I	GR/CGYC	....	7.5	0	37	.	.	.	40 6 0	
*Bbr*PI	*Eco*72I, *Pma*CI, *Pml*I	CAC/GTG	....	8.0	100	37	.	.	.	3 0 0	
*Bbu*I	*Sph*I	GCATG/C	....	7.5	0	37	.	.	+	6 1 2	
*Bbv*I		$5'GCAGCN_8/3'$ $3'CGTCGN_{12}/5'$ (* on both)	....	8.0	25	37	.	.	+	199 21 22	Largely insensitive to salt concentration, activity varies depending on substrate
*Bce*RI	*Acc*II, *Bsp*50I, *Bst*UI, *Mvn*I, *Tha*I	CG/CG	....				.	.	.	157 23 0	

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl..	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Bcl*I		T/GATCA	+ . – –	7.4	50	50	5	.	–	8 0 1	Needs clean DNA for good activity, 100 m*M* KCl can be used instead of 50 m*M* NaCl, <50 m*M* and>100 m*M* salt decreases activity
*Bcn*I	*Nci*I	CC/SGG[i]	. – + +	7.5	50	37	.	.	.	114 10 0	
*Bfr*I	*Afl*II	C/TTAAG	....	7.5	100	37	.	.	.	3 0 1	
*Bgl*I		GCCNNNN/NGGC	. – + +	7.4	75	37	1	.	+	29 3 1	Needs clean DNA for good activity, at pH 9.5 sixfold activity, <50 m*M* salt decreases activity, rate of cleavage at certain hemimethylated^{m5}C sites varies, complete methylation in overlapping M'*Hae*III sites inhibits cleavage *by Bgl*I.

*Bgl*II		A/GATCT	−...	7.4	100	37	1	.	−	6 0 0	At pH 9.5 sixfold activity, <50 m*M* salt decreases activity
*Bsa*AI		YAC/GTR	....	7.5	100	37	.	.	.	14 1 0	
*Bsc*I	*Ban*III, *Bsp*XI, *Bsu*15I, *Cla*I	AT/CGAT	....	7.5	50	37	.	.	.	15 1 0	
*Bse*PI	*Bss*HII	G/CGCGC	....				.	.	.	6 0 0	
*Bsm*AI		5'GTCTCN/3' 3'CAGAGN_5/5'	....	7.4	100	50	.	.	.	37 3 2	
*Bsm*I		5'GAATGCN/3' 3'CTTAC/G 5'	....	7.4	50	60	.	.	.	36 1 3	
*Bsp*50I	*Acc*II, *Bce*RI, *Bst*UI, *Mvn*I, *Tha*I	CG/CG	....	7.5	50	37	.	.	.	157 23 0	
Bsp 119I	*Asu*II, *Bst*BI, *Csp*45I, *Lsp*I, *Nsp*7524V, *Nsp*V	TT/CGAA	....	9.0	100	37	.	+	.	7 0 0	
*Bsp*120I		G/GGCCC	....	7.8	20	37	.	.	.	1 0 1	
Bsp 1286I	*Aoc*II, *Nsp*II, *Sdu*I	GDGCH/C	....	7.5	0	37	.	.	.	38 10 4	>50 m*M* salt decreases activity
*Bsp*AI	*Mbo*I, *Nde*II, *Sau*3AI	/GATC	....	7.4	80(KCl)	50	.	.	.	116 22 8	
*Bsp*HI	*Rsp*XI	T/CATGA	+.−−	7.4	100(KCl)	37	.	.	+	8 4 2	
*Bsp*MI		5'ACCTGCN$_4$/3' 3'TGGACGN$_8$/5'	..++	7.5	150	37	.	.	+	41 1 0	<100 m*M* salt decreases activity, the single site in pBR322 is nearly resistant to cleavage

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Bsp*MII	*Acc*III, *Kpn*2I, *Mro*I	T/CCGGA	– . + +	7.5	150	60	.	.	.	24 1 0	<100 m*M* salt decreases activity
*Bsp*NI	*Aha*I, *Apy*I, *Bst*GII, *Bst*NI, *Mva*I	CC/WGG	. . – –	8.0	100	37	.	.	.	114 10 0	
*Bsp*RI	*Bsu*RI, *Hae*III, *Pal*I	G/GCC	. . . .	8.2	25	37	.	+	.	149 22 18	
*Bsp*XI	*Ban*III, *Bsc*I, *Bsu*15I, *Cla*I	AT/CGAT	+ . . .	8.0	100	37	.	.	.	15 1 0	
*Bsr*I		5'ACTGGN/3' 3'TGAC/C5'	. . . .	7.8	150(KCl)	65	.	.	.	110 19 11	
*Bss*HII	*Bse*PI	G/CGCGC	. . . .	7.4	25	50	.	.	–	6 0 0	Largely insensitive to salt concentration
*Bst*I	*Bam*HI	G/GATCC	– . . +	7.0	100	37	.	+	–	5 1 1	At 55°C increased activity
*Bst*BI	*Asu*II, *Bsp*119I, *Csp*45I, *Lsp*I, *Nsp*7524V, *Nsp*V	TT/CGAA	. . . .	7.5	50	65	.	.	.	7 0 0	
*Bst*EII	*Eco*91I, *Eco*O65I	G/GTNACC	. – – –	8.0	150	60	1	.	–	13 0 0	50 m*M* KCl can be used instead of 150 m*M* NaCl, <100 m*M*NaCl decreases activity, at 37°C only 1/4 of the activity at 60°C

*Bst*GII	*Aha*I, *Apy*I, *Bsp*NI, *Bst*NI, *Mva*I	CC/WGG	....				.	.	.	71	6	17	
*Bst*NI	*Aha*I, *Apy*I, *Bsp*NI, *Bst*GII, *Mva*I	0 0 CC/WGG	.−−−	8.0	150	60	.	.	.	71	6	17	Fragments are poorly ligated, at 37°C only 1/3 of the activity at 60°C, <100 m*M* salt decreases activity
*Bst*UI	*Acc*II, *Bce*RI, *Bsp*50I, *Mvn*I, *Tha*I	+ + CG/CG	..++	8.0	0	60	.	.	.	157	23	0	
*Bst*XI		+ + CCANNNNN/ NTGG	...+	7.6	150	55	.	.	−	13	0	1	<50 m*M* salt decreases activity
*Bst*YI	*Mfl*I, *Xho*II	0 R/GATCY[i]	+...	8.0	0	60	.	.	.	21	8	3	
*Bsu*15I	*Ban*III, *Bsc*I, *Bsp*XI, *Cla*I	AT/CGAT	....	7.5	50	37	.	.	.	15	1	0	
*Bsu*36I	*Aoc*I, *Axy*I, *Cvn*I, *Eco*81I, *Eco*O109, *Mst*II, *Sau*I	CC/TNAGG	....	7.4	100	37	.	.	.	2	0	0	
*Bsu*RI	*Bsp*RI, *Hae*III, *Pal*I	* G/GCC	...+	7.5	50	37	.	.	.	149	22	18	Nicking occurs in the unmethylated strand of the hemimethylated sequence GG^{m5}CC
*Ccr*I	*Pae*R7I, *Sex*I, *Xho*I	C/TCGAG	....	7.7	100	37	.	.	.	1	1	0	Was previously called *Ccr*II
*Cel*II	*Esp*I	GC/TNAGC	....	7.5	150	37	.	.	.	6	0	1	

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Cfo*I	*Hha*I	GCG/C	. . + .	7.6	25	37	1	.	–+	215 31 2	>50 m*M* salt strongly reduces activity, activity varies depending on substrate
*Cfr*I	*Eae*I	Y/GGCCR	. . . +	8.5	25	37	.	.	.	39 6 0	
*Cfr*91	*Xcy*I, *Xma*I	C/CCGGG[i]	. . – –	7.5	200(Glut.)	37	.	+	.	3 0 0	For complete digestions Na-glutamate should be used and the DNA concentration >50 µg/mL
*Cfr*10I		R/CCGGY	. . . +	8.5	100(KCl)	37	.	.	–	61 7 1	The buffer should contain SO_4^{2-} and Triton
*Cfr*13I	*Asu*I, *Nsp*IV, *Sau*96AI, *Sdy*I	G/GNCC	. . . +	8.5	50	37	.	.	–	74 15 11	
*Cfr*42I	*Ksp*I, *Mra*I, *Sac*II, *Sst*II	CCGC/GG	. . . .	7.5	0	37	.	.	.	4 0 0	Shows dramatic differences in the rates of cleavage at different recognition sites
*Cla*I	*Ban*III, *Bsc*I, *Bsp*XI, *Bsu*15I	AT/CGAT	+ . + .	7.9	50	37	5	.	–+	15 1 0	
*Csp*6I		G/TAC	. . . .	7.8	20	37	.	.	.	113 3 12	

*Csp*45I	*Asu*II, *Bsp*119I, *Bst*BI, *Lsp*I, *Nsp*7524V, *Nsp*V	TT/CGAA	....	7.5	60	37	.	.	+	7 0 0	
*Cvn*I	*Aoc*I, *Axy*I, *Bsu*36I, *Eco*81I, *Eco*O109, *Mst*II, *Sau*I	CC/TNAGG	....	7.4	50(KCl)	37	1	.	+	2 0 0	Fragments are poorly ligated
*Dde*I		C/TNAG	...+	7.5	150	37	1	+	–+	104 8 20	Various DNAs require different amounts of *Dde*I for complete digestion, <50 m*M* salt and >10%glycerol decrease cleavage
*Dpn*I		GA/TC	..––	7.4	150	37	1	.	–+	116 22 –	Cleaves only DNA, when both internal adenine residues are *(dam-)*methylated, <100 m*M* salt decreases activity
*Dra*I	*Aha*III	TTT/AAA	..––	8.0	25	37	1	.	–+	13 3 12	Looses activity upon dilution, is sensitive to higher and lower salt concentration
*Dra*II	*Eco*O109I, *Pss*I	RG/GNCCY	.+++	8.5	50	37	.	.	.	3 4 3	
*Dra*III		CACNNN/GTG	....	8.2	100	37	.	.	+	10 0 1	50 m*M* NaCl/50 m*M* KCl can be used instead of 100 m*M* KCl, <50 m*M* and >100 m*M* salt decrease activity

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Drd*I		GACNNNN/NNGTC	....	7.4	50(KCl)	37	.	.	.	3 2 1	
*Dsa*I		C/CRYGG	....	7.5	100	55	.	.	.	25 0 6	
*Dsa*V	*Scr*FI	CC/NGG	....	7.9	70(KAc)	60	.	.	.	185 16 17	
*Eae*I	*Cfr*I	Y/GGCCR	.+++	7.4	50(KCl)	37	.	.	+	39 6 0	Is very sensitive to higher and lower salt concentration
*Eag*I	*Ecl*XI, *Eco*52I, *Xma*III	C/GGCCG	..++	8.2	150	37	.	.	+	2 1 0	<50 m*M* salt decreases activity
*Ear*I	*Ksp*632I	5'CTCTTCN/3' 3'GAGAAGN$_4$/5'	....	7.4	0	37	.	.	.	34 2 1	
*Ecl*136II		GAG/CTC	....	7.8	20	37	.	+	.	2 0 0	
*Ecl*XI	*Eag*I, *Eco*52I, *Xma*III	C/GGCCG	....	8.0	100	37	.	.	.	2 1 0	
*Eco*24I	*Ban*II	GRGCY/C	....	8.0	0	37	.	.	.	7 2 2	
*Eco*31I		5'GGTCTCN/3' 3'CCAGAGN$_5$/5'	.+..	7.5	25	37	.	.	.	2 1 0	
*Eco*47I	*Afl*I, *Ava*II, *Sin*I	G/GWCC	.+++	7.5	100	37	.	.	–	35 8 6	>10% glycerol decreases activity
*Eco*47III		AGC/GCT	..+.	7.5	100	37	.	.	–	2 4 1	
*Eco*52I	*Eag*I, *Ecl*XI, *Xma*III	C/GGCCG	....	8.0	100	37	.	.	–	2 1 0	

*Eco*57I		5'GTGAAGN$_{16}$/3' .... 3'CACTTCN$_{14}$/ 5'		7.8	20	37	.	.	.	1 4 1	Is stimulated by 10 µM S-adenosyl-methionine, cleavage is never complete
*Eco*64I	*Ban*I	G/GYRCC	....	8.5	50	37	.	.	.	25 9 1	
*Eco*72I	*Bbr*PI, *Pma*CI, *Pml*I	CAC/GTG	....	7.8	150	37	.	+	.	3 0 0	
*Eco*81I	*Aoc*I, *Axy*I, *Bsu*36I, *Cvn*I, *Eco*O109, *Mst*II, *Sau*I	CC/TNAGG	....	8.0	40(KCl)	37	.	.	–	2 0 0	
*Eco*91I	*Bst*EII, *Eco*O65I	G/GTNACC	....	8.5	50	37	.	.	.	13 0 0	
*Eco*105I	*Sna*BI	TAC/GTA	....	7.8	20	37	.	+	.	1 0 0	
*Eco*130I	*Eco*T14I, *Sty*I	C/CWWGG	....	8.5	100	37	.	.	.	10 1 8	
*Eco*147I	*Aat*I, *Stu*I	AGG/CCT	....	8.0	25	37	.	+	.	6 0 7	
*Eco*NI		CCTNN/ NNNAGG	....	7.5	50	37	.	.	+	9 1 2	Fragments are poorly ligated
*Eco*O109I	*Dra*II, *Pss*I	RG/GNCCY	.+.+	8.0	25	37	.	+	.	3 4 3	40 m*M* KCl can be used instead of 25 m*M* NaCl, largely insensitive to salt concentration
*Eco*O109	*Aoc*I, *Axy*I, *Bsu*36I, *Cvn*I, *Eco*81I, *Mst*II, *Sau*I	CC/TNAGG	....				.	.	.	2 0 0	
*Eco*O65I	*Bst*EII, *Eco*91I	G/GTNACC	....	7.5	150	37	.	.	.	13 0 0	

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Eco*RI		G/AATTC	. . + +	7.5	50	37	5	+	+	5 1 1	<50 m*M* salt results in reduced specificity, shows a reduced rate of cleavage at hemimethylated GAATT^{m5}C sequences and does not cut an oligonucleotide that contains ^{m5}C in both strands
*Eco*RII		/CCWGG	. + . +	7.5	50	37	1	.	+	71 6 17	<50 m*M* salt decreases activity
*Eco*RV		GAT/ATC	. . – –	7.9	150	37	1	+	+	21 1 1	<150 m*M* salt decreases activity, cuts the fully ^{m5}C-substituted bacteriophage XP12 DNA
*Eco*T14I	*Eco*130I, *Sty*I	C/CWWGG	. . . .	7.5	100	37	.	.	.	10 1 8	
*Eco*T22I	*Ava*III, *Nsi*I	ATGCA/T	. . . .	7.5	125	37	.	.	+	14 0 3	
*Ehe*I		GGC/GCC	. . + .	7.7	100	37	.	.	.	1 4 0	
*Esp*I	*Cel*II	GC/TNAGC	. . . +				.	.	.	6 0 1	
*Fdi*II	*Aos*I, *Avi*II, *Fsp*I, *Mst*I	TGC/GCA	. . . .	7.4	0	37	.	.	.	15 4 0	

*Fnu*4HI		$\overset{+}{G}C/NG\overset{+}{C}$	. . + +	7.5	50	37	.	.	−+	379	42	24	Fragments are poorly ligated, >50 m*M* salt decrease activity
*Fok*I		$\overset{0}{5'GGATGN_9/3'}$ $3'CCTACN_{13}/5'$ $\quad 0\,0 \quad 0$	− − − −	7.5	50	37	.	.	+	150	12	11	20 m*M* KCl can be used instead of 50 m*M* NaCl, largely insensitive to salt concentration, cuts about two- to four-fold more slowly at $CATC^{m5}C$ than at CATCC
*Fsp*I	*Aos*I, *Avi*II, *Fdi*II, *Mst*I	$TG\overset{+}{C}/GCA$	. . + .	7.4	75	37	.	.	−	15	4	0	Sensitive to higher and lower salt conc.
*Gsu*I		$5'CTGGAGN_{16}/3'$ $3'GACCTCN_{14}/5'$	. . . .	8.0	25	37	.	.	.	25	4	6	Is stimulated by S-adenosylmethionine, the buffer should contain 0.02% Triton
*Hae*II		$RG\overset{+}{C}GC/Y$	. . + .	7.4	50	37	1	.	−+	48	11	1	The buffer should contain 0.01% Triton, needs clean DNA for good cleavage, >100 m*M* salt decreases activity, gives various rate effects when the recognition sequence is ^{m5}C-methylated at different sites, is rather unstable during storage

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ	P	S[h]	Comments
*Hae*III	*Bsp*RI, *Bsu*RI, *Pal*I	G/GCC	. – – +	7.4	50	37	2	+	–	149	22	18	The buffer should contain 0.01% Triton, active up to 70°C, >50 m*M* salt decreases activity, nicking occurs in the unmethylated strand of the hemi-methylated sequence $GG^{m5}CC$
*Hap*II	*Hpa*II, *Msp*I	C/CGG	. . + +	7.5	0	37	.	.	–	328	26	1	
*Hga*I		5'GACGCN$_5$/3' 3'CTGCGN$_{10}$/5'	. . + +	7.4	50	37	.	.	+	102	11	0	>50 m*M* salt decreases activity
*Hgi*AI	*Alw*21I, *Asp*HI	GWGCW/C	. . . .	8.0	150	37	.	.	+	28	8	0	<150 m*M* salt decreases activity
*Hgi*DI	*Acy*I, *Aha*II, *Aos*II, *Bbi*II, *Hin*1I	GR/CGYC	. . . .				.	.	.	40	6	0	
*Hha*I	*Cfo*I	GCG/C	. . + +	8.0	25	37	1	+	+	215	31	2	15 m*M* KCl can be used instead of 25 m*M* NaCl
*Hin*1I	*Acy*I, *Aha*II, *Aos*II, *Bbi*II, *Hgi*DI	GR/CGYC	. . . .	8.5	25	37	.	.	.	40	6	0	

*Hin*6I	*Hin*PI, *Sci*NI	G/CGC	....	8.5	50	37	.	.	.	215 31 2	
*Hinc*II	*Hin*dII	GTY/RAC	..−−	7.4	50	37	1	+	+	35 2 7	<50 m*M* salt decreases activity, needs clean DNA for good activity
*Hin*dII	*Hinc*II	GTY/RAC	....	7.6	50	37	.	.	.	35 2 7	
*Hin*dIII		A/AGCTT	....	8.0	50	37	2	+	−	6 1 6	<50 m*M* and >100 m*M* salt decrease activity, activity varies depending on substrate
*Hin*fI		G/ANTC	..−−	7.4	100	37	1	+	−+	148 10 10	Unstable in solutions containing <20 μg/mL protein, <50 m*M* salt decreases activity, cuts unmethylated GANTC faster than hemimethylated GANT^{m5}C, which is cut faster than fully ^{m5}C-methylated sites, the rate difference between unmethylated and fully methylated DNA is only about 10-fold
*Hin*PI	*Hin*6I, *Sci*NI	G/CGC	..+.	8.0	50	37	.	.	+	215 31 2	Largely insensitive to salt concentration

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Hpa*I		GTT/AAC	. . – –	7.4	20(KCl)	37	1	+	–	14 0 4	Needs clean DNA for good activity, is rather unstable during storage
*Hpa*II	*Hap*II, *Msp*I	C/CGG	. . + +	7.4	10(KCl)	37	5	.	–	328 26 1	>50 m*M* KCl decreases activity, nicking occurs in the unmethylated strand of the hemi-methylated sequence ^{m5}CCGG
*Hph*I		5'GGTGAN$_8$/3' 3'CCACTN$_7$/5'	+ . . .	7.4	10(KCl)	37	.	.	+	168 12 4	>100 m*M* salt decreases activity
*Kpn*2I	*Acc*III, *Bsp*MII, *Mro*I	T/CCGGA	– . . .	7.5	20	37	.	+	.	24 1 0	
*Kpn*I		GGTAC/C	. – – –	7.5	0	37	1	+	–+	2 0 1	The buffer should contain 0.01% Triton, >50 m*M* salt decreases activity, gives rise to large fragments with eukaryotic DNA, needs clean DNA for good activity,

											activity varies depending on substrate
*Ksp*632I	*Ear*I	5'CTCTTCN/3' 3'GAGAAGN$_4$/5'	....	7.9	70(KAc)	37	.	.	.	34 2 1	
*Ksp*I	*Cfr*72I, *Mra*I, *Sac*II, *Sst*II	CCGC/GG	....	7.5	0	37	.	.	.	4 0 0	
*Lsp*I	*Asu*II, *Bsp*119I, *Bst*BI, *Csp*45I, *Nsp*7524V, *Nsp*V	TT/CGAA	....	7.5	0	60	.	.	.	7 0 0	
*Mae*I		C/TAG	....	8.4	250	45	.	.	.	13 5 12	
*Mae*II		A/CGT	..+−	8.2	250	50	.	.	.	143 10 0	Nicking occurs slowly in the unmethylated strand of the hemimethylated sequence A^{m5}CGT
*Mae*III		/GTNAC	....	8.2	275	55	.	.	.	156 17 14	
*Mam*I		GATNN/NNATC	....	7.5	100	37	.	.	.	21 1 3	
*Mbo*I	*Bsp*AI, *Nde*II, *Sau*3AI	/GATC	+.−−	7.4	100	37	1	.	+	116 22 8	Needs clean DNA for good activity, <50 m*M* salt decreases activity
*Mbo*II		5'GAAGAN$_8$/3' 3'CTTCTN$_7$/5'	+.−−	7.4	10(KCl)	37	3	.	−+	130 11 16	Needs clean DNA for good activity, largely insensitive to salt concentration, cuts the fully ^{m5}C-methylated bacteriophage XP12

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
											DNA, although certain hemimethylated ^{m5}C-containing substrates are reported not to be cut
*Mfl*I	*Bst*YI, *Xho*II	R/GATCY	+. . +	7.5	0	37	.	.	.	21 8 3	Cuts more slowly at ^{m6}AGATCY sites
*Mlu*I		A/CGCGT	. . + .	7.5	100	37	5	.	–	7 0 0	<50 m*M*, >100 m*M* salt or >5% glycerol decrease activity
*Mnl*I		5'CCTCN$_7$/3' 3'GGAGN$_7$/5'	. . . .	7.6	50	37	.	.	+	262 26 51	>100 m*M* salt decreases activity
*Mra*I	*Cfr*72I, *Ksp*I, *Sac*II, *Sst*II	CCGC/GG	. . . .	7.0	20	45	.	.	.	4 0 0	
*Mro*I	*Acc*III, *Bsp*MII, *Kpn*2I	T/CCGGA	– . . .	7.9	70(KAc)	37	.	.	.	24 1 0	
*Msc*I	*Bal*I	TGG/CCA	. . . .	7.6	50	37	.	.	.	18 1 0	
*Mse*I		T/TAA	. . – –	7.4	50	37	10	.	+	195 15 47	>50 m*M* NaCl decreases activity
*Msp*I	*Hap*II, *Hpa*II	C/CGG	. . – +	7.7	50	37	.	.	+	328 26 1	The buffer should contain 0.02% Triton, largely insensitive to salt concentration, cuts very slowly at GGC^{m5}CGG sites

*Mst*I	*Aos*I, *Avi*II, *Fdi*II, *Fsp*I	TGC/GCA	....	7.7	150	37	.	.	.	15 4 0	<50 m*M* salt decreases activity
*Mst*II	*Aoc*I, *Axy*I, *Bsu*361, *Cvn*I, *Eco*81I, *Eco*O109, *Sau*I	CC/TNAGG	....	7.7	150	37	.	.	+	2 0 0	<100 m*M* salt decreases activity
*Mva*I	*Aha*I, *Apy*I, *Bsp*NI, *Bst*GII, *Bst*NI,	CC/WGG[i]	.−..	8.0	150(KCl)	37	.	.	-	71 6 17	NaCl can be used instead of KCl
*Mvn*I	*Acc*II, *Bce*RI, *Bsp*50I, *Bst*UI, *Tha*I	CG/CG	....	7.5	50	37	.	.	.	157 23 0	
*Nae*I		GCC/GGC	..++	8.0	25	37	.	.	+	1 4 1	20 m*M* KCl can be used instead of 25 m*M* NaCl, some sites are cleaved much faster than others that are nearly resistant, >100 m*M* salt decreases activity
*Nar*I	*Nun*II	GG/CGCC	.−+.	7.4	0	37	24	.	+	1 4 0	Some sites are cleaved much faster than others that are nearly resistant, >50 m*M* salt decreases activity

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Nci*I	*Bcn*I	$^{0\ +}$CC/SGG	. . + +	7.4	25	37	5	.	−+	114 10 0	Cleavage results were supposed to produce 3'P and 5'OH ends, a finding that was questioned recently *(77)*, >50 m*M* salt decreases activity
*Nco*I		$^{+}$C/CATGG	. . . .	7.9	150	37	3	.	+	4 0 3	<100 m*M* salt decreases activity
*Nde*I		$^{0\ +\ \ +}$CA/TATG	. . − −	7.8	150	37	<1	.	−	7 1 2	Needs clean DNA for good activity, <150 m*M* salt decreases activity
*Nde*II	*Bsp*AI, *Mbo*I, *Sau*3AI	$^{+\ \ 0}$/GATC	+ . − −	7.6	150	37	5	.	+	116 22 8	
*Nhe*I		G/CTAGC^{+}	. . + +	7.4	50	37	.	.	−+	1 1 0	>5% glycerol or >100 m*M* salt decreases activity
*Nla*III		CATG/	. . . .	7.6	0	37	.	.	+	181 26 17	The buffer should contain 50 m*M* $(NH_4)_2SO_4$
*Nla*IV		GGN/NCC	. . . .	7.4	0	37	.	.	+	82 24 16	The buffer should contain 50 m*M* $(NH_4)_2SO_4$

*Not*I		GC/GGCCGC (++ 0)	. . + +	7.9	150	37	.	.	+	0 0 0	The buffer should contain 0.01% Triton
*Nru*I	*Spo*I	TCG/CGA (+)	+ . . .	8	50+50(KCl)	37	1	.	–+	5 1 0	<50 m*M* salt decreases activity
*Nsi*I	*Ava*III, *Eco*T22I	ATGCA/T (+)	. . . .	8.4	150	37	1	.	–	14 0 3	
*Nsp*7524I	*Nsp*I, *Nsp*HI	RCATG/Y	. . . .	8.0	80(KCl)	37	.	.	–+	32 4 2	
*Nsp*7524III	*Aqu*I, *Ava*I, *Avr*I, *Nsp*III	C/YCGRG	. . . .	7.5	50	37	.	.	– +	8 1 0	
*Nsp*7524V	*Asu*II, *Bsp*119I, *Bst*BI, *Csp*45I, *Lsp*I, *Nsp*V	TT/CGAA	. . . .	8.0	60	37	.	.	.	7 0 0	
*Nsp*I	*Nsp*7524I, *Nsp*HI	RCATG/Y	. . . .	7.5	50	50	.	.	.	32 4 2	
*Nsp*II	*Aoc*II, *Bsp*1286I, SduI	GDGCH/C	. . . .	8.5	0	37	.	.	.	38 10 4	
*Nsp*III	*Aqu*I, *Ava*I, *Avr*I, *Nsp*7524III	C/YCGRG	. . . .	8.0	25(KCl)	37	.	.	.	8 1 0	
*Nsp*IV	*Asu*I, *Cfr*13I, *Sau*96AI, *Sdy*I	G/GNCC	. . . .	8.0	0	37	.	.	.	74 15 11	
*Nsp*V	*Asu*II, *Bsp*119I, *Bst*BI, *Csp*45I, *Lsp*I, *Nsp*7524V	TT/CGAA	. . . .	8.0	25	50	.	.	.	7 0 0	
*Nsp*BII		CMG/CKG	. . . .	8.0	20	37	.	.	.	75 6 4	
*Nsp*HI	*Nsp*7524I, *Nsp*I	RCATG/Y	. . . .	8.0	20	37	.	.	.	32 4 2	
*Nun*II	*Nar*I	GG/CGCC	. . . .	8.0	20	37	.	.	.	1 4 0	
*Pae*R7I	*Ccr*I, *Sex*I, *Xho*I	C/TCGAG (*)	. . . .	7.4	0	37	.	+	–	1 0 0	>100 m*M* salt decreases activity
*Pal*I	*Bsp*RI, *Bsu*RI, *Hae*III	G/GCC	. . . .	7.5	0	37	.	.	–+	149 22 18	

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Pfl*MI		CCANNNN/NTGG	. + . .	7.4	50	37	.	.	+	14 2 2	Particular sites in λ-DNA are cleaved at significantly lower rates than those found with other substrates, >100 m*M* salt decreases activity
*Ple*I		5'GAGTCN$_4$/3' 3'CTCAGN$_5$/5'	. . . .	7.8	0	37	.	.	+	61 4 5	
*Pma*CI	*Bbr*PI, *Eco*72I, *Pml*I	CAC/GTG	. . . .	7.5	20	37	.	.	.	3 0 0	
*Pml*I	*Bbr*PI, *Eco*72I, *Pma*CI	CAC/GTG	. . . .	7.4	0	37	.	.	+	3 0 0	
*Ppu*MI		RG/GWCCY	. . . .	7.4	50	25	.	.	–	3 2 1	Activity decreases twofold when digestions are performed at 37°C, >50 m*M* salt decreases activity
*Pss*I	*Dra*II, *Eco*O109I	RG/GNCCY	. . . .				.	.	.	3 4 3	
*Pst*I		$\overset{+}{C}TGCA\overset{*}{}/G$	. . . .	7.5	100	37	1	+	–+	28 1 2	Largely insensitive to salt concentration
*Pvu*I	*Xor*II	$CGA\overset{0}{T}/\overset{*}{}CG$	– . + +	7.4	150	37	3	.	–	3 1 0	The buffer should contain 0.01%Triton, *Pvu*I cleaved

*Pvu*II		CAG/CTG[i]	. . . +	7.5	60	37	1	+	–+	15 1 3	DNA isolated from dam+ *E. coli* cells is ligated with reduced efficiency, <100 m*M* salt decreases activity The buffer should contain 0.01%Triton, largely insensitive to salt concentration
*Rsa*I		GT/AC	. . – –	8.0	50	37	1	.	+	113 3 12	Largely insensitive to salt concentration, >10% glycerol decreases activity
*Rsp*XI	*Bsp*HI	T/CATGA	+ . . .	7.5	50+50(KCl)	37	.	.	.	8 4 2	
*Rsr*II		CG/GWCCG	. . + +	8.0	10	37	.	.	+	5 0 0	>25 m*M* salt decreases activity
*Sac*I	*Sst*I	GAGCT/C	. . . .	7.4	0	37	.	.	+	2 0 0	Needs clean DNA for good activity, >50 m*M* salt decreases activity
*Sac*II	*Cfr*721, *Ksp*I, *Mra*I, *Sst*II	CCGC/GG	. . . +	7.5	0	37	.	.	–+	4 0 0	Particular sites in λ-DNA and φX-DNA are cleaved at significantly lower rates than those found in other substrates, >50 m*M* salt

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Sal*I		G/TCGAC	. . +.	8.0	150	37	2	+	−+	2 1 0	decreases activity, needs clean DNA for good activity
*Sau*3AI	*Bsp*AI, *Mbo*I, *Nde*II	/GATC	− . + +	7.5	100	37	1	+	+	116 22 8	<150 m*M* salt decreases activity, needs clean DNA for good activity, activity varies depending on substrate β-mercaptoethanol and >10% glycerol decreases activity, largely insensitive to salt concentration, nicking occurs in the unmethylated strand of the hemi-methylated sequence GAT^{m5}C, cuts at reduced rate at ^{m6}AGATC
*Sau*961	*Asu*I, *Cfr*131, *Nsp*IV, *Sdy*I	G/GNCC	. + + +	8.0	60	37	5	.	−+	74 15 11	Largely insensitive to salt concentration
*Sau*I	*Aoc*I, *Axy*I, *Bsu*36I, *Cvn*I, *Eco*81I, *Eco*O109, *Mst*II	CC/TNAGG	. . . .	7.2	75	37	.	.	.	2 0 0	

*Sca*I		AGT/ACT	....	7.4	150	37	1	+	–+	5 1 0	>10% glycerol decreases activity
*Sci*NI	*Hin*6I, *Hin*PI	G/CGC	....				.	.	.	215 31 2	
*Scr*FI	*Dsa*V	$^{0\ +}$CC/NGG	.+++	7.6	50	37	.	.	+	185 16 17	Fragments are difficult to ligate, <50 m*M* salt decreases activity
*Sdu*I	*Aoc*II, *Bsp*1286I, *Nsp*II	GDGCH/C	....	7.7	0	37	.	.	.	38 10 4	
*Sdy*I	*Asu*I, *Cfr*13I, *Nsp*IV, *Sau*96AI	G/GNCC	....				.	.	.	74 15 11	
*Sec*I		CC/NNGG	....				.	.	.	105 8 16	
*Sex*I	*Ccr*I, *Pae*R7I, *Xho*I	C/TCGAG	....				.	.	.	1 0 0	
*Sfa*NI		5'GCATCN$_4$/3' 3'CGTAGN$_9$/5'	....	7.5	150	37	.	.	+	169 22 6	<100 m*M* salt decreases activity
*Sfi*I		$^{0\ \ 0\ +}$GGCCNNNN $^{0+}$/NGGCC	..++	7.8	50	50	.	.	–	0 0 1	
*Sin*I	*Afl*II, *Ava*II, *Eco*471	$^{+}$G/GWCC	...+	7.4	20	37	.	.	+	35 8 6	Very sensitive to increased salt concentrations
*Sma*I		$^{+\ 0\ +}$CCC/GGG[i]	..++	8.0	20(KCl)	25	1	.	–+	3 0 0	Absolute requirement for potassium salt, needs clean DNA for cleavage, at 37°C the enzyme has a half life of only 15 min, nick-

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
											ing occurs in the unmethylated strand of the hemi-methylated sequence $CC^{m5}CGGG$
*Sna*BI	*Eco*105I	TAC/GTA	....	7.7	50	37	.	.	–	1 0 0	>50 m*M* salt decreases activity
*Sno*I	*Alw*44I, *Apa*LI	G/TGCAC	....	7.5	50	45	.	.	.	4 3 0	
*Spe*I		$\overset{+}{A}$/CTAGT	..––	7.5	100	37	.	.	+	0 0 0	Three cleavage sites in the DNA of adenovirus 2, <50 m*M* or >100 m*M* salt decrease activity
*Sph*I	*Bbu*I	GCATG/$\overset{0}{C}$	..––	7.4	100	37	1	.	+	6 1 2	May aggregate with DNA, phenol extraction or proteinase cleavage is recommended before gel analysis, <100 m*M* salt decreases activity
*Spl*I		C/GT$\overset{+}{A}$CG	....	7.5	100	37	.	.	.	1 0 0	

*Spo*I	*Nru*I	TCG/CGA	+...	7.5	50(KCl)	37	.	.	.	5 1 0	
*Ssp*I		AAT/ATT	..−−	7.4	100	37	.	.	+	20 1 6	<50 m*M*, >100 m*M* salt or > 5% glycerol decrease activity
*Sst*I	*Sac*I	GAGCT/C	....	8.0	50	37	5	+	+	2 0 0	>50 m*M* salt decreases activity
*Sst*II	*Cfr*72I, *Ksp*I, *Mra*I, *Sac*II	CCGC/GG	....	8.0	50	37	5	+	−	4 0 0	
*Stu*I	*Aat*I, *Eco*1471	AGG/CCT	.+.+	8.0	100	37	5	.	−+	6 7 0	The buffer should contain 0.01% Triton, largely insensitive to salt concentration
*Sty*I	*Eco*130I, *Eco*T14I	C/CWWGG	....	8.5	100	37	1	.	+	10 1 8	<100 m*M* salt or >5% glycerol decrease activity
*Taq*I	*Tth*Hb81	T/CGA	+.−−	8.4	100	65	3	.	−	121 7 1	At 37°C only half of the activity at 65°C, >100 m*M* salt decreases activity
*Tha*I	*Acc*II, *Bce*RI, *Bsp*50I, *Bst*UI, *Mvn*I	CG/CG	..+.	8.0	0	60	1	.	−	157 23 0	At 37°C only half of the activity at 65°C, >50 m*M* salt decreases activity
*Tth*111I	*Asp*I	GACN/NNGTC	....	7.4	50	65	.	+	−	2 1 0	Fragments are difficult to ligate, >100 m*M* salt decreases activity
*Tth*Hb81	*Taq*I	T/CGA	+.−−	8.0	100	60	.	.	−	121 7 1	

Table 5 *(Continued)*

Name	Isoschizomers[a]	Sequence[b]	Inhib. by methyl.[c]	Buffer cond.[d] pH	mM NaCl	°C	Stable Activ., h[e]	Star activ.[f]	Complete denat. at 65°C[g]	Cleavage sites in λ P S[h]	Comments
*Xba*I		T/CTAGA	+...	7.9	100	37	2	+	−+	1 0 0	<50 m*M* salt decreases activity, activity varies depending on substrate
*Xcm*I		$CCAN_5/N_4TGG$	....	8.0	0	37	.	.	.	12 0 1	
*Xcy*I	*Cfr*9I, *Xma*I	C/CCGGG	....				.	.	+	3 0 0	
*Xho*I	*Ccr*I, *Pae*R7I, *Sex*I	C/TCGAG	..+.	8.0	150	37	5	.	+	1 0 0	The buffer should contain 0.01% Triton, needs clean DNA for good cleavage, the enzyme is stable to phenol, therefore proteinase cleavage is recommended before phenol extraction, is rather unstable during storage
*Xho*II	*Bst*YI, *Mfl*I	R/GATCY	−..+	7.7	0	37	.	.	−	21 8 3	The buffer should contain 0.01% Triton, has only poor activity on DNA isolated from *dam*⁻ *E. coli* cells,

											>50 m*M* salt decreases activity, nicking occurs slowly in the unmethylated strand of the hemi-methylated sequence RGAT^{m5}CY
*Xma*I	*Cfr*9I, *Xcy*I	$\overset{+\ 0\ 0}{\text{C/CCGGG}}$	. . – .	7.5	25	37	16	.	+	3 0 0	>50 m*M* salt decreases activity
*Xma*III	*Eag*I, *Ecl*XI, *Eco*52I	$\overset{*}{\text{C/GGCCG}}$	. . . +	8.2	0	25	5	.	– +	2 1 0	Marked differences in cleavage rates at different sites, >100 m*M* salt decreases activity
*Xmn*I	*Asp*700	$\overset{+\ 0}{\text{GAANN/NNTTC}}\overset{+}{\text{.}}$	. . + +	8.0	0	37	.	.	+	24 2 0	>50 m*M* salt decreases activity
*Xor*II	*Pvu*I	$\overset{0\ +}{\text{CGAT/CG}}$	– . + +	7.4	50(KCl)	37	1	.	–	3 1 0	Very sensitive to increased salt concentration

[a]Only enzymes that cleave the DNA at the same phosphodiester bond are listed here.

[b]Methylation at ^{5m}C or ^{6m}A that do (+ ,*) or do not (o) interfere with cleavage are indicated. The modification by the canonical methylase is marked by *. B = G, T, or C, D = A, G, or T, H = A, C, or T, K = G or T, M = A or C, N = A, T, G, or C, R = A or G, S = G or C, V = A, G, or C, W = A or T, Y = C or T.

[c]The inhibitory effect of methylation by the *E. coli dam* (+ . . .) and *dcm* (. + . .)methylases or the eukaryotic CG (. . + .) and CNG (. . . +) methylation is indicated by +. If methylation does not inhibit the cleavage this is indicated by –.

[d]In most cases the buffer conditions are recommendations given by the manufacturers. Conditions that decrease the activity are specified under comments.

Table 5 *(Continued)*

[e]The time for which the enzyme stays fully active.

[f]Enzymes for which a different cleavage pattern has been detected at special buffer conditions or after a prolonged incubation time using high enzyme concentrations are indicated by +.

[g]Complete denaturation of the enzyme by an incubation for 15 min at 65°C is indicated by +, –+ marks a partial inactivation. Enzymes that stay active under these conditions are indicated by –.

[h]The number of cleavage sites in λ-DNA (λ), pBR322 DNA (P), and SV40 DNA (S) are indicated.

[i]*Bam*HI: the canonical methylation is GGAT^{m4}CC but GGAT^{m5}CC does also inhibit the cleavage,
*Bcn*I: the canonical methylation is C^{m4}CSGG,
*Bst*YI: RGAT^{m4}CY is not cleaved,
*Cfr*9I: the canonical methylation is C^{m4}CCGGG, CC^{m4}CGGG does also inhibit the cleavage,
*Mva*I: the canonical methylation is C^{m4}CWGG, ^{m4}CCWGG does also inhibit the cleavage,
*Pvu*II: the canonical methylation is CAG^{m4}CTG.

note added in proof at end of chapter). A concentration of 5–10 mM $MgCl_2$ seems to be optimal. Some enzymes (e.g., *Nla*III and *Nla*IV) additionally require 50 m*M* $(NH_4)_2SO_4$ to be activated.

3.4.4. Stabilizing Additives

Stabilizing additives, already mentioned as components of the storage buffer (Section 3.1.), are used to maintain the activity and accuracy of restriction enzymes during the reaction.

3.4.5. Reaction Time and Temperature

The reaction time needed to digest a given amount of DNA is determined by the activity of the restriction enzyme. However, it is not useful to exceed the expected lifetime of a restriction enzyme under the reaction conditions indicated in Table 5. The repeated addition of the required activity may overcome problems concerning the extreme instability of some enzymes (e.g., *Sca*I, *Sph*I, *Sty*I, *Cfo*I). Care has to be taken that accumulation of glycerol caused by the repeated addition of enzymes does not induce "star" activity.

Temperature is a critical parameter for the optimum use of restriction enzymes. Whereas most restriction enzymes have temperature optima around 37°C, some enzymes, in particular those isolated from cryo- or thermophilic bacteria, need relatively low or high temperatures, respectively, for catalytic activity. The optimum incubation temperatures are summarized in Table 5. For incubation at elevated temperatures, one should seal the reaction vessels or overlay the reaction mixture with parafine oil.

3.5. Assays for Purity of Restriction Endonuclease Preparations

Commercially available restriction enzyme preparations are normally tested by the suppliers for the absence of the following contaminants: other specific restriction endonucleases, nonspecific endonucleases, 5'-exonucleases, 3'-exonucleases, and phosphatases. The absence of these enzyme activities can be assayed by the fragment-pattern-stability assay, plasmid-nicking assay, 5'-exonuclease-phosphatase assay, 3'-exonuclease assay, and ligation-recut assay (*see* the following sections). Depending on the specific application, the absence of contaminant activities is more or less critical: Whereas

very pure enzyme preparations are needed for DNA cloning, digestion of genomic DNA for Southern analysis is not impaired by a contaminant phosphatase activity.

3.5.1. Fragment-Pattern-Stability Assay

Contamination of enzyme preparations with other specific endonucleases is assayed in λ DNA cleavage experiments as described above, however, with incubation periods of up to several hours. The absence of a contaminating site-specific endonuclease can be verified by examination of the generated λ-restriction-fragment pattern. It should be kept in mind, however, that the prolonged incubation of a DNA substrate with a large amount of a restriction enzyme preparation may produce an altered banding pattern because of the limited accuracy of all restriction enzymes. The sites cleaved under these conditions are normally the same as those that are attacked under "star" conditions after shorter digestion times. Nonspecific nuclease contaminations cause degradation of the individual fragments to yield a "smear" on the agarose gel.

3.5.2. Plasmid-Nicking Assay

The absence of contaminating nonspecific endonucleases may be assayed by incubating any supercoiled plasmid or phage DNA that does not harbor a site for the specific endonuclease for several hours. The cleavage of a single phosphodiester bond in the superhelical DNA results in a conversion to open circular DNA, which can be easily detected by electrophoresis on agarose gels containing 0.01 mg/mL ethidium bromide.

3.5.3. 5'-Exonuclease / Phosphatase Assay

To detect any contamination with 5'-exonuclease or phosphatase activities, the following standard assay can be used. λ DNA is completely cleaved with a restriction enzyme producing a large number of fragments with 5' protruding ends (e.g., *Hpa*II, *Hap*II, *Msp*I). The DNA fragments are labeled using T4 polynucleotide kinase (*see* Chapter 20) and $\gamma[^{32}P]$-ATP and ADP in an exchange reaction. The resulting labeled DNA fragments are purified by chromatography over DE52 spun columns. The radioactively labeled DNA is then incubated with the enzyme preparation to be tested. During this incubation, any contamination that liberates phosphate or mononucleotides from the 5' end can

be detected *(182)*. The radioactive products of these side reactions can be quantitatively determined by homochromatography on DEAE- or PEI-cellulose thin-layer plates or by HPLC on a reversed-phase column.

3.5.4. 3'-Exonuclease Assay

λ-DNA restriction fragments with 5' protruding ends (*see* Section 3.5.3.) are labeled at their 3' ends by incorporation of α[^{32}P]- or α[^{35}S]-dNTPs using Klenow polymerase (*see* Chapters 3 and 4). The 3'-labeled DNA is incubated after separation of the unincorporated activity with the enzyme preparation to be tested. Any 3'-exonuclease activity attacking the 3' terminally labeled fragments will liberate [^{32}P]- or [^{35}S]-labeled mononucleotides.

3.5.5. Ligation-Recut Assay

The functional integrity of the fragment termini produced after cleavage of any plasmid DNA with restriction enzymes is assayed in the ligation-recut assay. Initially, the substrate DNA is completely digested by incubating 10 µg of DNA with 50 U of the restriction enzyme in 50 µL of the appropriate reaction buffer over a period of 1–2 h. Restriction fragments are extracted twice with phenol/chloroform (1:1) and precipitated after addition of 1/10 vol 3*M* sodium acetate, pH 4.8, by 2 vol ethanol. Sticky ends are ligated with 0.1 U, blunt ends with 1–10 U T_4-DNA ligase/µg DNA in a total volume of 10 µL by incubation for 16 h at 12–16°C in 20 m*M* Tris-HCl, pH 7.5, 10 m*M* $MgCl_2$, 10 m*M* DTE or DTT, and 0.6 m*M* ATP. Ligation is terminated by heating the reaction mixture for 10 min at 70°C. An aliquot (approx 500 ng) is analyzed by agarose gel electrophoresis to determine the extent to which the fragments could be rejoined. The products of the ligation are recut with the same restriction enzyme under the reaction conditions mentioned above. The fragments produced are analyzed by agarose gel electrophoresis. A normal banding pattern indicates that 5' and 3' termini remained intact, demonstrating the absence of contaminating endo- and exonuclease activities.

3.6. Nonspecific Inhibition of Restriction Enzymes

In general, restriction enzymes are inhibited by chelation of the essential cofactor Mg^{2+}. Some enzymes are very sensitive to variations in ionic strength or can be specifically inhibited by thiol reagents. Impurities introduced together with the substrate may cause a significant reduction of activity or, in extreme cases, completely inhibit restriction

enzymes. Substrate impurities vary with the method used to isolate and purify the DNA. In particular, "miniprep" methods developed for rapid plasmid isolation tend to give DNA material containing impurities that may interfere with the restriction digests: salt contained in the lysis buffer, organic solvents remaining from the extraction with phenol/chloroform or from the precipitation with alcohols, as well as cellular proteins, carbohydrates, lipids, and other nucleic acids that were not completely removed during the purification.

Salt and organic solvents can be removed by repeated washing of the DNA after an ethanol precipitation with 70% ethanol at room temperature prior to drying the pellet in a desiccator, or alternatively by dialysis against 0.3*M* sodium acetate, pH 4.8, and precipitation by ethanol. Contaminating proteins, especially nonspecific nucleases, must be separated by phenol/chloroform extraction after incubation with proteinase K. Contaminating cellular RNA can be destroyed by the addition of DNase-free RNase to the lysis buffer (0.1 mg/mL). Chromosomal DNA fragments must be separated by chromatographic procedures or density gradient centrifugation.

3.7. Specific Inhibition of Restriction Enzymes by Methylation of the Substrate

As discussed in Section 2.4.1., the activity and specificity of restriction enzymes are influenced by the methylation state of the DNA substrate (for a review, cf *[128]*). Site-specific methylation of DNA is owing to the following methyltransferase activities: modification methyltransferases as part of bacterial restriction/modification systems, modification methyltransferases as part of bacterial DNA-repair systems, and eukaryotic cytosine-specific methyltransferases. Of interest, in particular during cloning work, is DNA methylation by the *dam* and *dcm* methyltransferases, which are present in the majority of the commonly used *E. coli* strains. The *dam* methyltransferase transfers a methyl group from *S*-adenosylmethionine to the N^6 position of adenine within the sequence GATC. The *dcm* methyltransferase specifically methylates the internal cytosine at C^5 within the sequence CCWGG. Methylation by *dam* or *dcm* methyltransferases inhibits DNA cleavage by certain restriction enzymes whose recognition sites overlap or are identical with the *dam* or *dcm* modification sites (e.g., *Taq*I, *Mbo*I). Other restriction enzymes are not influenced by *dam* or *dcm* methyla-

tion (e.g., *Bam*HI, *Bst*NI). Another class of restriction enzymes (e.g., *Cfu*I, *Dpn*I) only cleaves the sequence G^{m6}ATC when modified by the *dam* methylase or exclusively cuts methylated *dcm* sites, C^{m5}CWGG (e.g., *Apy*I). The sensitivity toward site-specific methylation also may prevent cleavage of eukaryotic DNA whose cytosine residues may be methylated in CG or CNG sequences. Whereas *Hpa*II and *Sma*I are inhibited by ^{m5}CG, *Msp*I and *Xma*I are not influenced by this type of methylation. The inability of some restriction enzymes to cleave methylated DNA can be utilized to obtain large fragments even with enzymes recognizing hexanucleotide palindromes. The sensitivity of the commercially available restriction enzymes toward different types of methylation is summarized in Tables 5, 6, and 7.

To avoid site-specific *dam* and/or *dcm* methylation of recognition sites in *E. coli* vectors, it is necessary to propagate the vectors in strains lacking one or both of the methyltransferase systems (e.g., *E. coli* GM1674, which has the genotype *dam*$^-$ *dcm*$^-$). It must be taken into account, however, that *E. coli* strains, deficient in *dam* or *dcm* methylation, have elevated rates of mutation.

Site-specific methylation in vitro, using Type II modification enzymes or the *dam* or *dcm* methyltransferases, can be used to block the subset of restriction sites that overlap with the methylated sequence (Table 6). Therefore, the *Msp*I methyltransferase (CCGG) can block the cleavage of DNA by the restriction endonuclease *Bam*HI (GGATCC) in the sequences GGATCCGG and CCGGATCC. For restriction endonucleases recognizing degenerate sequences, a new specificity can result. For example, *Hinc*II recognizes the sequence GTYRAC. The *Taq*I methyltransferase inhibits the cleavage of GTCGAC by modifying TCGA. Thus, the specificity of *Hinc*II is reduced to GTYAAC, when the DNA has been methylated by M•*Taq*I.

DNA methylation by Type II modification enzymes can be blocked by prior methylation with other methyltransferases. If the first methylation does not inhibit the cleavage with the corresponding endonuclease, one can selectively cleave those recognition sites where the second methylation was blocked. For example, methylation with M•*Msp*I prevents methylation with M•*Bam*HI at GGATCCCGG sequences. These sequences are cut by *Bam*HI, whereas all other recognition sites for this enzyme are methylated by its corresponding methylase (*see* footnote *d* in Table 6).

Table 6
Commercially Available Modification Methylases, Their Effect of DNA Cleavage by Certain Restriction Endonucleases, and New Cleavage Specificities Created by Their Use Together with Restriction Endonucleases

		Buffer conditions						
Name	Sequence[a]	pH	Mg^{2+} m*M*	NaCl m*M*	SAM[b] μ*M*	°C	Enzyme	Resulting specificity in methylated DNA[c]
M•*Alu*I	AG^{m5}CT	7.5	0	0	80	37	*Aqu*I	C/YCGRG excluding AG*CTCGRG*
							*Ava*I	C/YCGRG excludingAG*CTCGRG*
	Complete inhibition of DNA cleavage catalyzed by *Hin*dIII, *Pvu*I, *Sac*I, and *Sst*I						*Ban*II	GGGCY/C
							*Bsp*1286I	GKGCH/C
							*Dde*I	C/TNAG excluding AGC*TNAG*
							*Esp*I	BGC/TNAGC
							*Fnu*4HI	GC/NGC excluding A*GCTGC*
							*Hgi*AI	GTGCW/C
							*Nhe*I	BG/CTAGC
							*Pst*I	CTGCA/G excluding AG*CTGCAG*
							*Xho*I	C/TCGAG excluding AG*CTCGAG*
M•*Bam*HI	GGAT^{m4}CC^{d}	7.5	0	50	80	37	*Cfr9*I	C/CCGGG excluding GGAT*CCCGGG*
							*Eco*RII	/CCWGG excluding GGAT*CCWGG*
							*Hpa*II	C/CGG excluding GGAT*CCGG*
							*Msp*I	C/CGG excluding GGAT*CCGG*
							*Mva*I	CC/WGG excluding GGAT*CCWGG*
							*Nco*I	C/CATGG excluding GGAT*CCATGG*
							*Sma*I	CCC/GGG excluding GGAT*CCCGGG*
							*Xma*I	C/CCGGG excluding GGAT*CCCGGG*
M•*Cla*I	ATCG^{m6}AT	7.5	0	0	80	37	*Eco*RV	GAT/ATC excluding ATC*GATATC*
							*Hin*fI	G/ANTC excluding ATC*GATTC*
	Complete inhibition of DNA cleavage catalyzed by *Ban*III						*Mbo*I	/GATC excluding ATC*GATC*
							*Nde*II	/GATC excluding ATC*GATC*
							*Taq*I	T/CGA excluding A*TCGA*T

*Dpn*I only cleaves M•*Cla*I methylated DNA in the sequence ATCGA/TCGAT

M•*Eco* dam	$G^{m6}\overset{0}{A}TC$	7.5	0	0	80	37	*Acc*III	T/CCGGA excluding *TCCGGA*TC
							*Ban*III	AT/CGATD
	Complete inhibition of DNA cleavage catalyzed by *Alw*I, *Bcl*I, *Bst*YI, *Mfl*I, and *Nde*II						*Bsp*HI	T/CATGA excluding *TCATGA*TC
							*Cla*I	AT/CGATD
							*Hph*I	$GGTGAN_8$/ excluding *GGTGA*TC
							*Mbo*II	$GAAGAN_8$ excluding *GAAGA*TC
							*Nru*I	TCG/CGA excluding *TCGCGA*TC
							*Rsp*XI	T/CATGA exduding *TCATGA*TC
							*Spo*I	TCG/CGA excluding *TCGCGA*TC
							*Taq*I	T/CGA excluding *TCGA*TC
							*Tth*HB81	T/CGA excluding *TCGA*TC
							*Xba*I	T/CTAGA excluding *TCTAGA*TC
M•*Eco*RI	$G\overset{+}{A}^{m6}A\overset{0}{T}TC$	8.0	0	100	80	37		
M•*Hae*III	$GG^{m5}CC$	8.5	0	50	80	37	*Apy*I	CC/WGG excluding GG*CCWGG*
							*Aqu*I	C/YCGRG excluding GG*CCCGRG*
	Complete inhibition of DNA cleavage catalyzed by *Aat*T, *Bcl*I, *Eae*I, *Eag*I, *Not*I, and *Stu*I						*Ava*I	C/YCGRG excluding GG*CCCGRG*
							*Ban*II	GAGCY/C
							*Bgl*I	$HGCCN_4$/NGGC
							*Bsp*1286I	GMGCH/C
							*Bst*XI	$CCAN_5$/NTGG excluding GG*CCAN$_6$TGG*
							*Cfr*9I	C/CCGGG excluding GG*CCCGGG*
							*Cfr*10I	R/CCGGY excluding G*GCCGGY*
							*Fnu*4HI	GC/NGC excluding GG*CCGC*
							*Hpa*II	C/CGG excluding GGCCGG
							*Msp*I	C/CGG excluding GGCCGG
							*Nae*I	HGCC/GGC
							*Nco*I	C/CATGG excluding GG*CCATGG*
							*Sac*II	CCGC/GG excluding GG*CCGCGG*
							*Sau*96I	G/GWCC
							*Sma*I	CCC/GGG excluding GG*CCCGGG*

Table 6 *(Continued)*

Name	Sequence[a]	Buffer conditions: pH	Mg^{2+} m*M*	NaCl m*M*	SAM[b] μ*M*	°C	Enzyme	Resulting specificity in methylated DNA[c]
							*Xma*I	C/CCGGG excluding GG*CCCGGG*
M•*Hha*I	$G^{m5}CGC$ (m5C: 0)	7.5	0	0	80	37	*Acc*II	HCGCG
							*Aha*II	GA/CGYC
	Complete inhibition of DNA cleavage catalyzed by *Cfo*I, *Eco*47III, *Fsp*I, *Hae*II, *Hin*PI, and *Nar*I						*Aos*II	GA/CGYC
							*Bsp*MI	ACCTGCN_4/ excluding *ACCTGC*GC
							*Bst*UI	HCG/CG
							*Fnu*4HI	GC/NGC excluding *GCNGC*GC
							*Hga*I	GACGCN_5 excluding *GACGC*GC
							*Tha*I	HCG/CG
M•*Hpa*II	$C^{m5}CGG$ (m5C: +)	7.5	0	0	80	37	*Aha*II	GR/CGYC excluding *GRCGCC*GG
							*Apa*I	GGGCC/C excluding *GGGCCC*GG
	Complete inhibition of DNA cleavage catalyzed by *Bsp*MII, *Hap*II, *Nae*I, *Nci*I, and *Sma*I						*Asp*718	G/GTACC excluding *GGTACC*GG
							*Ava*I	C/TCGRG
							*Ava*II	G/GWCC excluding *GGWCC*GG
							*Bgl*I	GCCN_4/NGGC excluding *GCCGGN$_3$GG* C
							*Dra*II	RG/GNCCY excluding *RGGNCCC*GG
							*Eae*I	Y/GGCCR excluding *YGGCCG*G
							*Eag*I	C/GGCCGH
							*Eco*47I	G/GWCC excluding *GGWCC*GG
							*Rsr*II	CG/GWCCGH
							*Scr*FI	CC/WGG
M•*Hph*I	$T^{m5}CACC$[e] (m5C: 0)	7.5	0	0	80	37	*Aha*II	GR/CGYC excluding *GRCGTC*ACC
							*Eco*RI	G/AATTC excluding *GAATTC*ACC
							*Hin*fI	G/ANTC excluding *GANTC*ACC
							*Sau*3AI	/GATC excluding *GATC*ACC
							*Xmn*I	GAAN_2/N_2TTC excluding *GAAN$_4$TTC*ACC

M•*Msp*I	$^{m5}\overset{+}{C}CGG$	7.5	0	100	80	37	*Apa*I	GGGCC/C excluding *GGGCCC*GG
							*Ava*II	G/GWCC excluding *GGWCC*GG and excluding *GGWCC*CGG
	Complete inhibition of DNA cleavage catalyzed by *Bsp*MII, *Cfr*10I, *Hpa*II, and *Nae*I						*Bam*HI	G/GATCC excluding *GGATCC*GG
							*Bgl*I	GCCN$_4$/NGGC excluding *GCCGGN$_3$GGC* and excluding *GCCCGGN$_2$GCC*
							*Bsp*MI	ACCTGCN$_4$/ excluding *ACCTGC*CGG
							*Bst*I	G/GATCC excluding *GGATCC*GG
							*Bsu*RI	G/GCC excluding *GGCC*GG
							*Cfr*I	Y/GGCCR excluding *YGGCCG*G
							*Cfr*13I	G/GNCC excluding *GGNCC*GG
							*Eae*I	Y/GGCCR excluding *YGGCCG*G
							*Eag*I	C/GGCCGH
							*Eco*47I	G/GWCC excluding *GGWCC*CGG
							*Eco*RI	G/AATTC excluding *GAATTC*CGG
							*Fnu*4HI	GC/NGC excluding *GCNGC*CGG
							*Hae*III	G/GCC excluding *GGCC*GG
							*Hga*I	GACGCN$_5$ excluding *GACGC*CGG
							*Hha*I	GCG/C excluding *GCGC*CGG
							*Mfl*I	R/GATCY excluding *RGATCC*GG
							*Nhe*I	G/CTAGC excluding *GCTAGC*CGG
							*Rsr*II	CG/GWCCGH
							*Sau*3AI	/GATC excluding *GATC*CGG
							*Sau*96I	G/GNCC excluding *GGNCC*GG and excluding *GGNCC*CGG
							*Scr*FI	CC/WGG
							*Sin*I	G/GWCC excluding *GGWCC*GG
							*Xho*II	R/GATCY excluding *RGATCC*GG
							*Xma*III	C/GGCCGH
							*Xmn*I	GAAN$_2$/N$_2$TTC excluding *GAAN$_4$TTC*CGG

Table 6 *(Continued)*

Name	Sequence[a]	Buffer conditions: pH	Mg^{2+} m*M*	NaCl m*M*	SAM[b] μ*M*	°C	Enzyme	Resulting specificity in methylated DNA[c]
M•PstI	CTGC^{m6}AG	7.5	0	0	80	37	*Alu*I	AG/CT excluding *AGCT*GCAG
M•*Taq*I	$\overset{o}{T}CG^{m6}A$	7.4	6	100	80	65	*Acc*I	GT/AKAC
							*Alu*I	AG/CT excluding TCG*AGCT*
	Complete inhibition of DNA cleavage catalyzed by *Ban*III, *Cla*I, *Csp*45I, *Pae*R7I, *Sal*I, *Tth*HB81, and *Xho*I						*Ava*I	C/CCGRG
							*Eco*RI	G/AATTC excluding TC*GAATTC*
							*Eco*RV	GAT/ATC excluding TC*GATATC*
							*Hinc*II	GTT/RAC
							HindII	*GTT/RAC*
							HinfI	*G/ANTC excluding TCGANTC*
							MboI	*/GATC excluding TCGATC*
							NdeII	*/GATC excluding TCGATC*
							SpeI	*A/CTAGT excluding TCGACTAGT*
							XmnI	*GAAN$_2$/N$_2$TTC excluding TCGAAN$_4$TTC*
*Dpn*I only cleaves M•*Taq*I methylated DNA in the sequence TCGA/TCGA								

[a]Methylation at ^{5m}C or ^{6m}A is indicated, which does (+) or does not (o) interfere with methylation at indicated position. B = G, T, or C, D = A, G, or T, H = A, C, or T, K = G or T, M = A or C, N = A, T, G, or C, R = A or G, S = G or C, V = A, G, or C, W = A or T, Y = C or T.

[b]SAM = S-adenosylmethionine.

[c]The recognition sequence of the endonuclease is italicized. DNA cleavage of the following restriction endonucleases is blocked by more than one modification methylase (number of methylases), which can further reduce the subset of cleavage sites: *Aqu*I(2), *Aha*II(3), *Apa*I(2), *Ava*I(4), *Ava*II(2), *Ban*II(2), *Bgl*I(3), *Bsp*MI(2), *Bsp*1286I(2), *Cfr*I(2), *Eae*I(2), *Eag*I(2), *Eco*47I(2), *Eco*RI(3), *Eco*RV(2), *Fnu*4HI(4), *Hin*fI(3), *Hpa*II(2), *Mbo*I(2), *Msp*I(2), *Nco*I(2), *Nde*II(2), *Nhe*I(2), *Rsr*II(2), *Sau*3AI(2), *Sau*96I(2), *Sma*I(2), *Taq*I(2), *Xma*I(2), *Xmn*I(3). Methylation by two different enzymes can lead to new nonpalindromic cleavage specificities:

M•*Alu*I + M•*Hae*III methylation followed by *Ban*II cleavage GAGCC/C
M•*Alu*I + M•*Hae*III methylation followed by *Bsp*1286I cleavage GWGCK/C
M•*Hae*III + M•*Taq*I methylation followed by *Ava*I cleavage C/TCGGG
[d]M•*Bam*HI can be inhibited by prior methylation with M•*Hpa*II or M•*Msp*I, which does not inhibit DNA cleavage by the *Bam*HI endonuclease. This leads to the new specificities: G/GATCCGG and G/GATCCCGG.
[e]The modification of DNA by methylation of the adenine in the complementary strand inhibits the methylation with M•*Hph*I.

3.8. "Star" Activity

"Star" activity, discussed in Section 2.4.2., is defined as a relaxed specificity of restriction enzymes under suboptimum reaction conditions and is related to the naturally observed limited accuracy of these enzymes when used in high concentration over a prolonged time. The relatively low specific activities of most commercial enzyme preparations, normally 10 U/μL, do not lead to a "star" pattern under normal buffer conditions.

3.9. Terminating the Cleavage Reaction

Depending on the specific purpose for which a restriction digest has to be carried out, the digestion reaction is terminated by inhibiting, destroying, or extracting the restriction enzyme. This can be achieved by chelation of the essential cofactor Mg^{2+}, thermal denaturation of the enzyme, chemical denaturation of the enzyme, or extraction of the enzyme by phenol/chloroform. It should be kept in mind that chelation only blocks catalytic activity, but does not prevent DNA binding of the restriction enzymes. DNA binding may interfere with the electrophoretic analysis of the digest. It is recommended, therefore, to denature the restriction enzymes prior to separation of the cleavage products by gel electrophoresis by the addition of a loading buffer containing 1% SDS and 100 m*M* EDTA. Thermal denaturation after restriction digestion and prior to further reactions, e.g., ligation, is achieved by heating the reaction mixture to 65°C for 10 min, which is sufficient to inactivate most enzymes or to 95°C for 10 min for some of them. This is, however, not effective with thermostable enzymes (cf Table 5). In this case, it is necessary to denature and remove the restriction enzyme by a phenol/chloroform extraction normally followed by ethanol precipitation of the DNA fragments.

3.10. Analysis of the Cleavage Products

The analytical separation of DNA restriction fragments commonly is carried out by electrophoretic techniques that allow a precise determination of the size of DNA molecules. For oligodeoxynucleotides, homochromatography *(183,184)* or HPLC methods (for a review, cf *[185]*) can be used as an alternative.

Electrophoretic separation of DNA cleavage products can be achieved in agarose, polyacrylamide, or agarose-polyacrylamide composite gels under native or denaturating conditions. Whereas condi-

Table 7
Isoschizomer Pairs that Can Be Used in the Analysis of the Methylation State of DNA

	Methylated sequence	Cut by	Not cut by
Methylation at CG	$C^{m5}CGG$	*Msp*I	*Hpa*II or *Hap*II
	$CC^{m5}CGGG$	*Xma*I or *Cfr*9I	*Sma*I
	$TC^{m5}CGGA$	*Acc*III	*Bsp*MII
	$GAT^{m5}C$	*Mbo*I	*Sau*3AI
	$GGTAC^{m5}C$	*Kpn*I	*Asp*718
	$GGWC^{m5}C$	*Afl*I	*Ava*II or *Eco*47I
Methylation at GATC	$G^{m6}ATC$	*Sau*3AI	*Mbo*I or *Nde*II
	$RG^{m6}ATCY$	*Xho*II or *Bst*YI	*Mfl*I
	$TCCGG^{m6}A$	*Bsp*MII or *Mro*I	*Acc*III
Methylation at CCWGG	$C^{m5}CWGG$	*Bst*NI or *Mva*I	*Eco*RII
Methylation at CNG	$C^{m5}CWGG$	*Bst*NI or *Mva*I	*Eco*RII
	$GGTA^{m5}CC$	*Kpn*I	*Asp*718
	$T^{m5}CCGGA$	*Acc*III	*Bsp*MII

tions that preserve the native structure of DNA are used to separate double-stranded DNA, denaturing conditions (e.g., electrophoresis at alkaline pH or in the presence of chaotropic agents) are used to analyze single-stranded DNA fragments. Depending on the size of the fragments to be separated, different gel systems are used (Table 4).

4. Analytical Applications

Tables 8–11 summarize the recognition sequences and sites of cleavage of commercially available restriction enzymes. Main suppliers of restriction and modification enzymes are listed in Table 12.

4.1. Restriction Mapping of DNA Molecules

A set of different restriction endonucleases can be used to produce a map of their recognition sequences in a given DNA. Such a map allows the positioning of functional sites and the location of hybridization probes, as well as the unambiguous identification of this DNA. In rare cases, for example with some hemoglobinopathias, point mutations can be identified by restriction mapping if they impair the cleavage of one of these enzymes. To obtain an informative map, one should choose restriction enzymes that cut the DNA at least two and up to about ten times within the stretch of interest. An unequivocal assignment of the

Table 8
Restriction Enzymes Recognizing Palindromic
Tetra- and Hexanucleotide Sequences[a]

	AATT	ACGT	AGCT	ATAT	CATG	CCGG	CGCG	CTAG	GATC	GCGC	GGCC	GTAC	TATA	TCGA	TGCA	TTAA
▼□□□□									MboI+							
□□□□▼		MaeII				HpaII		MaeI		HinPI+	HaeIII	Csp6I		TaqI+		MseI
□□▼□□			AluI				AccII+		DpnI			RsaI				
□□□▼□										HhaI+						
□□□□▼					NlaIII											
A□□□□T			HindIII		AflII	CfrI0I	MluI AflII	SpeI	BglII XhoII+							
A□□□□T														ClaI+		AseI+
A□□□□T				SspI						Eco47II	StuI+	ScaI				
A□□□□T																
A□□□□T					NspHI					HaeII					NsiI+	
C□□□□G					NcoI StyI+ DsaI	XmaI+ AvaI+	DsaI	AvrII StyI+			Eco52I+ EagI+	SplI		XhoI+ AvaI+		AflII+
C□□□□G				NdeI	SecI	SecI	SecI	SecI								
C□□□□G		BbrPI+ BsaAI+	PvuII NspBII			SmaI	NspBII									
C□□□□G							SacII+		PvuI							
C□□□□G															PstI	
G□□□□C	EcoRI						BssHII+	NheI	BamHI+ XhoII+	BanI+	Bsp120I	Asp718 BanI+		SalI	ApaLI	
G□□□□C		AcyI+								NarI AcyI+			AccI	AccI		
G□□□□C			Eci136II	EcoRV		NaeI			NlaIV	EheI NlaIV	NlaIV	NlaIV		HincII+		HpaI HincII+
G□□□□C			SacI+ BanII+ SduI+ HgiAI+													
G□□□□C		AatII			SphI+ NspHI+					BbeI HaeII	ApaI BanII+ SduI+	KpnI			SduI+ HgiAI+	
T□□□□A					BspHI+	AccIII+		XbaI	BclI		EcoI+					
T□□□□A														NspV+		
T□□□□A		SnaBI+ BsaAI+					NruI+			FspI+	BalI+					DraI+
T□□□□A																
T□□□□A																

[a]The column on the left defines the length of the recognition sequence and the phosphodiester bond cleaved and gives the first and last nucleotides for hexanucleotide sequences. The row on top gives the central tetranucleotide of all recognition sequences. Listed in the table are commercially available restriction enzymes. For sequences that are recognized by different isoschizomers (*see* Table 5) a + marks the name of the enzyme sold by most manu-

Table 9
Restriction Enzymes Recognizing Palindromic Penta-, Hepta-, and Octanucleotide Sequences

Recognition sequence comprising 5 bp		Recognition sequence comprising 7 bp		Recognition sequence comprising 8 bp	
*Afl*I[+]	G/GWCC[a]	*Ppu*MI	RG/GWCCY[b]	*Not*I	GC/GGCCGC
*Aha*I[+]	CC/WGG	*Rsr*II	CG/GWCCG		
*Bcn*I[+]	CC/SGG				
*Eco*RII	/CCWGG				

[a]Listed in the table are commercially available restriction enzymes. For sequences that are recognized by different isoschizomers (*see* Table 5), a + marks the name of the enzyme sold by most manufacturers.

[b]R = A or G, S = G or C, W = A or T, Y = T or C.

cleavage sites is only possible if each restriction fragment for a given enzyme is cut by one of the other enzymes. This information leads to the linkage of neighboring sequences. Therefore, a series of single and double digests with all chosen enzymes and enzyme combinations followed by the determination of the length of the fragments produced has to be carried out to construct a restriction map. Partial digests of end-labeled DNA produced by one enzyme can be employed for the alignment of individual sites. Alternatively, two-dimensional gel electrophoresis can be used for the alignment of restriction fragments: the cleavage products of the first digest are analyzed in the first dimension; separated fragments are then digested with another enzyme (or digested to completion with the same enzyme when a partial digest was performed) and analyzed in the second dimension. The second cleavage reaction can easily be performed by diffusing the appropriate buffer and the enzyme into the gel, and after the reaction is completed

facturers. In addition, in each box enzymes that recognize degenerate sequences are listed. The table allows one to identify a restriction enzyme for any given palindromic tetra- or hexanucleotide recognition sequence and to select restriction enzymes that produce a desired 5' or 3' projecting single-strand region after cleavage. For example, if one needs DNA to be cleaved within the sequence AGATCT between A and G, one should look for A/ ----T in the column on the left and for GATC on the row on top. This defines *Bgl*II as the restriciton enzyme needed. An inspection of the column designated by GATC shows that *Bgl*II produces a 5' projecting single-stranded region after cleavage of the DNA, which is compatible with products of an *Mbo*I, *Bam*HI, or *Bcl*I as well as some of the products of an *Xho*II cleavage, because *Xho*II has a degenerate recognition sequence.

Table 10
Restriction Enzymes Recognizing Palindromic Sequences
with One to Nine Unspecified Nucleotides in the Center

Recognition sequence comprising 6 defined bp[a]		Recognition sequence comprising 4 defined bp		Recognition sequence comprising 8 defined bp	
*Alw*NI	CAGNNN/CTG	*Dde*I	C/TNAG	*Sfi*I	GGCCNNNN/NGGCC
*Axy*I+	CC/TNAGG	*Fnu*4HI	GC/NGC		
*Bgl*I	GCCNNNN/NGGC	*Hin*fI	G/ANTC		
*Bst*EII+	G/GTNACC	*Mae*III	/GTNAC		
*Bst*XI	CCANNNNN/NTGG	*Sau*96I+	G/GNCC		
*Dra*II+	RG/GNCCY[b]				
*Dra*III	CACNNN/GTG				
*Drd*I	GACNNNN/NNGTC				
*Eco*NI	CCTNN/NNNAGG				
*Esp*I+	GC/TNAGC				
*Mam*I	GATNN/NNATC				
*Pfl*MI	CCANNNN/NTGG				
*Tth*111I+	GACN/NNGTC				
*Xcm*I	CCANNNNN/NNNNTGG				
*Xmn*I+	GAANN/NNTTC				

[a]Listed in the table are commercially available restriction enzymes. For sequences that are recognized by different isoschizomers (*see* Table 5), a + marks the name of the enzyme sold by most manufacturers.

[b]R = A or G, Y = T or C

Table 11
Restriction Enzymes Recognizing Nonpalindromic Sequences

Recognition sequence comprising 5 bp		Recognition sequence comprising 6 bp		Recognition sequence comprising 4 bp	
*Alw*26I	5' GTCTCN/3'	*Bsm*I	5' GAATGCN/3'	*Mnl*I	5' CCTCN$_7$/3'
	3' CAGAGN$_5$/5'		3' CTTAC/G 5'		3' GGAGN$_7$/5'
*Alw*I	5' GGATCN$_4$/3'	*Bsp*MI	5' ACCTGCN$_4$/3'		
	3' CCTAGN$_5$/5'		3' TGGACGN$_8$/5'		
*Bbv*I	5' GCAGCN$_8$/3'	*Ear*I	5' CTCTTCN/3'		
	3' CGTCGN$_{12}$/ 5'		3' GAGAAGN$_4$/5'		
*Bsm*AI	5' GTCTCN/3'	*Eco*31I	5' GGTCTCN/3'		
	3' CAGAGN$_5$/5'		3' CCAGAGN$_5$/5'		
*Bsr*I	5' ACTGGN/3'	*Eco*57I	5' GTGAAGN$_{16}$/3'		
	3' TGAC/C 5'		3' GACTTCN$_{14}$/5'		
*Fok*I	5' GGATGN$_9$/3'	*Gsu*I	5' CTGGAGN$_{16}$/3'		
	3' CCTACN$_{13}$/5'		3' GACCTCN$_{14}$/5'		
*Hga*I	5' GACGCN$_5$/3'	*Ksp*632I	5' CTCTTCN/3'		
	3' CTGCGN$_{10}$/5'		3' GAGAAGN$_4$/5'		
*Hph*I	5' GGTGAN$_8$/3'				
	3' CCACTN$_7$/5'				
*Mbo*II	5' GAAGAN$_8$/3'				
	3' CTTCTN$_7$/5'				
*Ple*I	5' GAGTCN$_4$/3'				
	3' CTCAGN$_5$/5'				
*Sfa*NI	5' GCATCN$_4$/3'				
	3' CGTAGN$_9$/5'				

Table 12
Main Suppliers of Restriction Enzymes

Amersham Buchler (Braunschweig, Germany)
Applied Genetechnology Systems (AGS, Heidelberg, Germany)
Appligene (Illkirch, France)
Biozym Diagnostik (Hameln, Germany)
Boehringer (Mannheim, Germany)
GIBCO-BRL (Gaithersburg, MD)
ICN Biomedicals (Cleveland, OH)
International Biotechnologies, Inc. (IBI, New Haven, CT)
Janssen Biochimica (Beerse, Belgium)
New England Biolabs (Beverly, MA)
New Brunswick Scientific Company (Hatfield, UK)
Paesel & Lorei (Frankfurt/Main, Germany)
Pharmacia-LKB (Uppsala, Sweden)
Promega (Madison, WI)
Sigma (St. Louis, MO)
Stehelin (Basel, Switzerland)
Stratagene (La Jolla, CA)
United States Biochemical Corporation (USB, Cleveland, OH)

by loading the gel slice on top of another gel for the analysis in the second dimension. For small DNAs, such as plasmid DNAs, there is no difficulty in finding a suitable set of enzymes for restriction mapping. For longer DNAs, one has to find rare cutting enzymes to obtain well-resolved restriction fragments. This puts the upper limit for restriction mapping in the range of the length of small prokaryotic genomes *(186,187)*.

Restriction mapping may be used to obtain an estimate of the number of copies of genes present in a genome. Digestion of the total genomic DNA by a restriction enzyme that has no site within the gene (more precisely the DNA region complementary to the probe) to be analyzed followed by electrophoresis, transfer to nitrocellulose, and hybridization to a specific nucleic acid probe identifies a number of bands that correspond to the minimum number of genes within this genome.

4.2. Restriction Fragment Length Polymorphisms

The pattern of DNA restriction fragments is characteristic for a given DNA. Point mutations or rearrangements of the DNA lead to a change in the length of one or more restriction fragments. With a specific

probe that hybridizes only to one sequence in a Southern blot experiment, one will find in general only one restriction fragment for a given allele. During evolution, different alleles can acquire mutations in the sequences next to the hybridization position that may result in a different length of a particular restriction fragment. This phenomenon is called a restriction fragment length polymorphism (RFLP) for this pair of enzyme and probe. Such RFLPs are useful for the identification of genetic diseases, because the gene defect linked to the RFLP can be identified in the absence of a phenotypic abnormality or before its manifestation *(188,189)*. RFLPs can also be used for forensic purposes or to clarify kinship relationships between individuals.

4.3. DNA Fingerprints

The pattern of different DNA fragments produced by digesting genomic DNA with a restriction enzyme can be used to identify not only the species, but also different strains of a species or even different individuals, provided an enzyme is used that has many cleavage sites, such that a pattern rich in information is obtained *(190)*. This approach has been successfully used for the typing of microorganisms and for an estimation of the evolutionary relatedness of various species. If the genomic DNA is too complex, such that individual restriction fragments cannot be resolved as with eukaryotic genomes, one has to label a subset of all these fragments by hybridization with a DNA probe that has a sequence complementary to a middle repetitive DNA sequence in this genome. This procedure results in a pattern of approx 20–50 bands that can be considered as a genetic fingerprint. Genetic fingerprinting begins to form an important part of casework procedure in forensic science laboratories *(191)*.

4.4. Analysis of the Methylation State of DNA from Natural Sources

Eukaryotic DNA is methylated to a variable extent on the 5 position of cytosine, and the subject is discussed in detail in Sections 2.4.1. and 3.1. There is evidence that the methylation state is correlated with the level of gene expression. Isoschizomer pairs of restriction endonucleases, which allow the analysis of the methylation state of DNA, are given in Table 7.

4.5. Cleavage of Single-Stranded DNA

The substrate of restriction endonucleases is double-stranded DNA, but some virus DNAs are single stranded and some protocols in molecular biology, such as DNA sequencing and site-directed mutagenesis, use single-stranded DNA. Because of internal sequence homologies, these DNAs show a lot of secondary structures that can reconstitute recognition sequences of restriction enzymes (*see* Section 2.4.3.). Therefore, enzymes with 4- or 5-bp long recognition sites will often cleave these single-stranded DNAs.

Type II restriction enzymes with nonpalindromic recognition sequences cleaving more than 5 bp outside of this sequence (e.g., *Fok*I, *Gsu*I, *Eco*57I, cf Table 11) can be used to cleave any specific DNA sequence in single-stranded DNA. This reaction is based on the use of synthetic oligodeoxynucleotides that form an adaptor comprising the double-stranded recognition sequence linked to an overhanging single-stranded sequence that is complementary to the desired cleavage site in the target DNA. After hybridization of this adaptor to the target sequence, the restriction enzyme binds to the recognition sequence within the oligodeoxynucleotide and cleaves the target DNA as shown in Fig. 2 *(192,193)*. In principle, this procedure should be applicable also to double-stranded target DNA.

4.6. Labeling Restriction Fragments at Their Ends

For a variety of analytical purposes, radioactive labeling of restriction fragments is necessary (*see* Sections 3.5.3. and 3.5.4.). Normally, ethidium bromide is used to stain double-stranded restriction fragments for analytical applications. If the amount of a produced fragment is less than about 10 ng, the fluorescence of this dye may not be strong enough for the detection of the fragment. By radioactive labeling, the detection limit can be reduced up to 1000-fold. There are several methods for labeling the 5' or 3' end of each strand in the DNA fragment (for detailed protocols, cf *[21,194,195]*).

All restriction enzymes (according to Bennett *[77] Nci*I is not an exception to this rule as previously believed) generate 5'-phosphate and 3'-OH groups at the termini of the DNA fragments (*see* note added in proof at end of chapter). In order to obtain 5'-labeled termini, restriction fragments must be dephosphorylated with alkaline phosphatase,

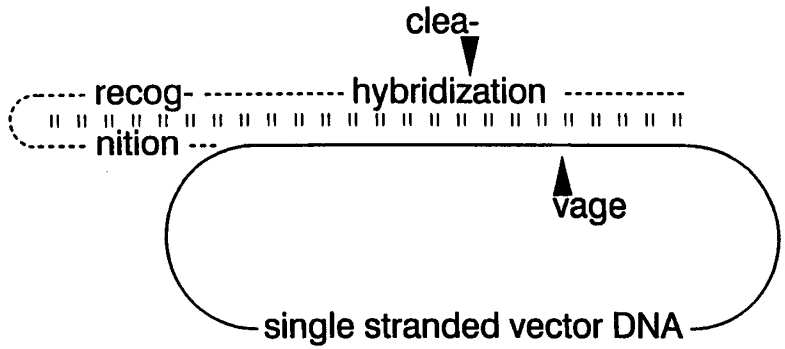


Fig. 2. Site-directed cleavage of single-stranded DNA using an adapter oligodeoxynucleotide containing a recognition site for a type IIs restriction enzyme (e.g., *Fok*I) and a sequence complementary to the DNA sequence to be cleaved.

which has to be removed by phenol extraction, or with HK™-phosphatase, which can be heat inactivated, before this end can be radioactively labeled using T4 polynucleotide kinase and $\gamma[^{32}P]$- or $\gamma[^{35}S]$-ATP *(196)*. The yield of the phosphorylation reaction depends on the accessibility of the 5'-OH group: It is higher for 5' overhanging ends as compared to blunt or 3' overhanging ends. The DNA can be dephosphorylated also by the kinase itself by enhancing the reverse reaction using ADP in excess over the radioactive ATP *(197)*.

Labeling at the 3'-OH group is carried out using calf thymus terminal deoxynucleotidyl transferase and $\alpha[^{32}P]$-cordycepine triphosphate, which functions as a chain terminator. Also for this reaction, the yield depends on the accessibility of the 3'-OH terminus; the yield can be improved by thermal denaturation of the DNA ends.

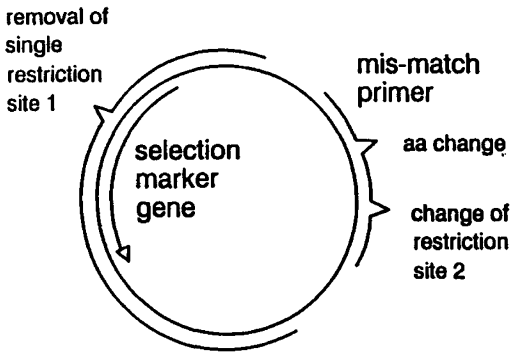
Both these methods result in the labeling of both 5' or 3' ends of the two strands forming the double-stranded restriction fragment. For labeling only one end of a fragment, one can carry out a DNA polymerization reaction (fill-in reaction) with a suitable mixture of dNTPs. DNA polymerases can only work on a 5'-overhanging end. With two restriction enzymes, two different ends are generated. This offers the possibility of labeling only one of these ends by including only those radioactive $\alpha[^{32}P]$- or $\alpha[^{35}S]$-dNTPs in the polymerization mixture that are required by the fill-in reaction for one end.

5. Preparative Applications

For a variety of different purposes, in particular for the cloning of DNA and sequencing using the Maxam-Gilbert method *(182)*, it is necessary to produce fragments in preparative amounts. This is done by a scale-up of analytical digests and/or an extension of the incubation time. The derived restriction fragment is usually isolated from the product mixture either by preparative electrophoresis or when larger amounts are needed by HPLC (for detailed protocols *see* refs. *21,185*).

An application for which restriction digests have to be carried out on a preparative scale that is becoming increasingly more important is the preparation of DNA for site-directed mutagenesis. The gapped-duplex protocol *(198)* requires, in addition to the single-stranded circular template DNA, a complementary piece of DNA that covers most of the template and leaves uncovered the region in which the mutation is to be introduced. The gap is prepared by a suitable restriction digest of the double-stranded mutagenesis vector, e.g., M13RF, pEMBL, and so on, isolation of the appropriate fragment, and annealing to the single-stranded mutagenesis vector.

Preparative restriction digests can also be used to increase the molar yield of a mutagenesis experiment. Suppose the circular template strand of a gapped-duplex carries a particular singular restriction site within the selection marker gene (e.g., the β-lactamase gene), whereas the gap-forming complementary strand owing to a silent point mutation does not contain this site. Annealing of the mismatch oligodeoxynucleotide, polymerization, and ligation will yield a double-stranded circular DNA that carries in its template strand the wild-type sequence of the gene as well as the recognition sequence for the restriction enzyme and, in its newly formed strand, the mutated sequence of the gene, but no recognition sequence for the restriction enzyme. Transformation into repair-deficient *E. coli* cells will yield a more or less equal distribution of two plasmids in the clones, only one of which carries the mutated gene. The yield of this one can be improved by isolating plasmid DNA from the *E. coli* culture and cleavage of the DNA by the restriction enzyme, which attacks only the wild-type DNA. Retransformation of the DNA resulting from this reaction will lead to preferential transformation of the supercoiled DNA, i.e., the mutant DNA. This procedure leads to much increased yields of the



1. Hybridization:
   ss vector DNA +
   ds gap fragment +
   mis-match primer
2. Fill-in with DNA polymerase
   and ligase
3. Transformation into a repair
   deficient strain
4. Cleavage of the plasmid DNA
   with restriction enzyme 1
   (enrichment of mutant plasmid)
5. Transformation into a normal
   strain
6. Screening with restriction
   enzyme 2 (identification of
   mutant plasmid)
7. DNA sequencing

Fig. 3. Use of restriction enzyme catalyzed DNA cleavage to increase the marker yield of a site-directed mutagenesis experiment and to screen for positive clones.

desired mutant. Identification of positive clones can be done by screening for a restriction site introduced together with the desired mutation on the mismatch primer or direct sequencing (Fig. 3).

Note Added in Proof

After completion of this review, many new discoveries regarding restriction enzymes were made, only a few of which and only those that are of general importance are considered in this note.

The structure of the *Eco*RI-DNA recognition complex has been revised *(199)*: According to the new structure (reviewed in ref. *200)*, the recognition interactions comprise 16 protein-base hydrogen bonds to the purines and pyrimidines, and also van der Waal's contacts to all of the pyrimidines of the recognition sequence. Furthermore, hydrogen bonds and electrostatic contacts to the phosphates within and outside of the recognition sequence most likely contribute to the specificity of DNA binding by *Eco*RI. Direct interactions are buttressed by indirect ones between amino acid residues involved directly in base-specific contacts and amino acids nearby. In contrast to the earlier

recognition model, the new one is highly redundant, a result that had emerged already from previous site-directed mutagenesis experiments on *Eco*RI (reviewed in ref. *201*). The revised structure of the *Eco*RI-DNA complex is in agreement with previous *(201)* and more recent *(202)* mutational analyses.

Although the structure of the *Eco*RV-DNA complex has not yet been published, its general features have been reviewed *(203,204)*. The DNA is highly deformed in the complex. Recognition interactions comprise 12 protein-base hydrogen bonds and several van der Waal's contacts, as well as electrostatic contacts. Results from site-directed-mutagenesis experiments support the recognition model derived from the X-ray structure analysis *(205–207)*. Different from *Eco*RI, *Eco*RV does not bind to DNA in a specific manner when Mg^{2+} is absent *(208)*; it does so, however, in the presence of Mg^{2+} *(209)*. *Eco*RV does not seem to be an exception, because, as with *Taq*I, specificity is observed only late in the enzymatic cycle *(210)*. The structural reason for the different behavior of *Eco*RI and *Eco*RV—as the paradigmatic examples of restriction enzymes that exhibit DNA binding specificity already in the absence or only in the presence of Mg^{2+}—could be the degree of overlap of the DNA binding site and the catalytic center. In *Eco*RI, which produces a four-base extension, most of the specificity determining contacts can presumably be formed in the absence of Mg^{2+}, because electrostatic repulsion between DNA and protein owing to several negatively charged amino acid residues of the Mg^{2+} binding site occurs only at the end of the recognition sequence. In *Eco*RV, which is a blunt-end cutter, negatively charged amino acid residues of the Mg^{2+} binding site are located vis-à-vis the center of the recognition sequence in the protein-DNA complex; specificity determining contacts, therefore, can hardly be formed in the absence of Mg^{2+}.

Structural and mechanistic studies on *Eco*RI and *Eco*RV have not only made it possible to identify amino acids that are involved in specific DNA binding, but also those that are likely to have a primarily catalytic function *(205,206)*. Based on this information, and knowing that *Eco*RI and *Eco*RV cleave the phosphodiester bond with inversion of configuration *(115,211)*, a catalytic mechanism, which is sufficiently general to apply for many restriction enzymes *(212)*, was proposed for these enzymes. Important features of this mechanism include water activation by a phosphate residue next to the scissile phosphodiester

bond, polarization of the P-O bond by Mg^{2+} and protonation of the leaving group by a Mg^{2+} bound water molecule. Site-directed-mutagenesis experiments have made it possible to define in other restriction enzymes (e.g., *Bam*HI *[213]*) amino acid residues involved in catalysis.

The observation that many restriction enzymes, depending on sequence context and reaction conditions, cleave the two strands of the DNA in separate binding events has given rise to the assumption that the accuracy of restriction enzymes can be increased by proofreading using DNA ligase. This was shown to be the case in vitro and in vivo *(214)*. The requirement for proofreading could explain why all restriction enzymes produce 5' phosphorylated ends and free 3'-ends, because these are the substrates for DNA-ligase.

The purification of restriction enzymes is comparatively easy when their genes are cloned in expression vectors. A further simplification can be achieved when their genes are fused with DNA sequences coding for affinity "tags." For example, C-terminal His_6 tags allow one-step-purification of the *Eco*RI:His_6 fusion protein. The resulting preparation is homogeneous, fully active, and devoid of contaminating nonspecific nuclease activities *(215)*.

Assaying the enzymatic activity of restriction enzymes has been simplified considerably by a recently developed continuous spectrophotometric assay, which takes advantage of the hyperchromic effect produced by dissociation of the product double strands after cleavage of short oligodeoxynucleotides *(216)*.

References

1. Bertani, G. and Weigle, J. J. (1953) Host controlled variation in bacterial viruses. *J. Bacteriol.* **65,** 113–121.
2. Arber, W. and Dussoix, D. (1962) Host specificity of DNA produced by *Escherichia coli.* I: Host controlled modification of bacteriophage λ. *J. Mol. Biol.* **5,** 18–29.
3. Meselson, M. and Yuan, R. (1968) DNA restriction enzyme from *Escherichia coli. Nature* **217,** 1110–1114.
4. Linn, S. and Arber, W. (1968) In vitro restriction of phage fd replicative form. *Proc. Natl. Acad. Sci. USA* **59,** 1300–1305.
5. Smith, H. O. and Wilcox, K. W. (1970) A restriction enzyme from Haemophilus influenzae. I: Purification and general properties. *J. Mol. Biol.* **51,** 379–391.
6. Roulland-Dussoix, D. and Boyer, H. W. (1969) The *Escherichia coli* B restriction endonuclease. *Biochim. Biophys.* **195,** 219–232.

7. Kauc, L. and Piekarowicz, A. (1978) Purification and properties of a new restriction endonuclease from Haemophilus influenzae Rf. *Eur. J. Biochem.* **92,** 417–426.
8. Boyer, H. W. (1971) DNA restriction and modification mechanisms in bacteria. *Ann. Rev. Microbiol.* **25,** 153–176.
9. Smith, H. O. and Nathans, D. (1973) A suggested nomenclature for bacterial host modification and restriction systems and their enzymes. *J. Mol. Biol.* **81,** 419–433.
10. Wilson, G. G. (1988) Type II restriction-modification systems. *Trends Genetics* **4,** 314–318.
11. Modrich, P. and Roberts, R. J. (1982) Type II restriction and modification enzymes, in *Nucleases* (Linn, S. M. and Roberts, R. J., eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 109–154.
12. Bickle, T. A. (1987) DNA restriction and modification systems, in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology,* vol. 1 (Neidhardt, F. C., ed.), American Society for Microbiology, Washington, DC, pp. 692–696.
13. Malcolm, A. D. B. and Snousnou, G. (1987) Restriction enzymes and DNA, in *Topics in Molecular and Structural Biology*, vol. 9 (Neidle, S., ed.), Macmillan, London, pp. 193–222.
14. Bennett, S. P. and Halford, S. E. (1989) Recognition of DNA by type II restriction enzymes, in *Current Topics in Cellular Regulation*, vol. 30, Academic, New York, pp. 57–109.
15. Chirikjian, J. G. (1981) *Restriction Endonucleases.* Elsevier, New York.
16. Chirikjian, J. G. (1987) *Restriction Endonucleases and Methylases.* Elsevier, New York.
17. Fuchs, R. and Blakesley, R. (1983) Guide to the use of type II restriction endonucleases. *Methods Enzymol.* **100,** 3–37.
18. Davis, L. G., Dibner, M. D., and Battey, J. F. (1986) *Basic Methods in Molecular Biology.* Elsevier, New York.
19. Ansubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith J. A., and Struhl, K. (1987) *Current Protocols in Molecular Biology.* Wiley, New York.
20. Perbal, B. (1988) *A Practical Guide to Molecular Cloning.* Wiley Interscience, New York.
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
22. Grossman, L. and Moldave, K (1980) *Methods Enzymol.* **65,** 1–968.
23. Kessler, C. and Höltke, H.-J. (1986) Specificity of restriction endonucleases and methylases—a review (ed. 2). *Gene* **47,** 1–153.
24. Roberts, R. J. (1989) Restriction enzymes and their isoschizomeres. *Nucleic Acids Res.* **17,** r347–r388.
25. Greene, P. J., Heyneker, H. L., Bolivar, F., Rodriguez, R. L., Betlach, M. C., Covarrubias, A. A., Backman, K., Russel, D. J., Tait, R., and Boyer, H. W. (1978) A general method for the purification of restriction enzymes. *Nucleic Acids Res.* **5,** 2373–2380.

26. Pirrotta, V. and Bickle, T. A. (1980) General purification schemes for restriction endonucleases. *Methods Enzymol.* **65,** 89–95.
27. Blanks, R. and McLaughlin, L. W. (1988) An oligodeoxynucleotide affinity column for the isolation of sequence specific DNA binding proteins. *Nucleic Acids Res.* **16,** 10,283–10,299.
28. Cheng, S.-C., Kim, R., King, K., Kim, S.-H., and Modrich, P. (1984) Isolation of gram quantities of *Eco*RI restriction and modification enzymes from an overproducing strain. *J. Biol. Chem.* **259,** 11,571–11,575.
29. d'Arcy, A., Brown, R. S., Zabeau, M., van Resandt, R. W., and Winkler, F. K. (1985) Purification and crystallization of the *Eco*RV restriction endonuclease. *J. Biol. Chem.* **260,** 1987–1990.
30. Luke, P. A. and Halford, S. E. (1985) Solubility of *Eco*RI restriction endonuclease and its purification from an over-producing strain. *Gene* **37,** 241–246.
31. Kapfer, W., Walter, J., and Trautner, T. A. (1991) Cloning, characterization and evolution of the *Bsu*FI restriction endonuclease gene of *Bacillus subtilis* and purification of the enzyme. *Nucleic Acids Res.* **19,** 6457–6463.
32. Kiss, A., Posfai, G., Keller, C. C., Venetianer, P., and Roberts, R. J. (1985) Nucleotide sequence of the *Bsu*RI restriction-modification system. *Nucleic Acids Res.* **13,** 6403–6421.
33. Sznyter, L. A., Slatho, B., Moran, L., O'Donnell, K. H., and Brooks, J. E. (1987) Nucleotide sequence of the *Dde*I restriction-modification system and characterization of the methylase protein. *Nucleic Acids Res.* **15,** 8249–8266.
34. Lacks, S. A., Mannarelli, B. M., Springhorn, S. S., and Greenberg, B. (1986) Genetic basis of the complementary *Dpn*I and *Dpn*II restriction systems of *S. pneumoniae*: an intercellular cassette mechanism. *Cell* **46,** 993–1000.
35. Greene, P. J., Gupta, M., and Boyer, H. W. (1981) Sequence analysis of the DNA encoding the *Eco*RI endonuclease and methylase. *J. Biol. Chem.* **256,** 2143–2151.
36. Newman, A. K., Rubin, R. A., Kim, S.-H., and Modrich, P. (1981) DNA sequences of structural genes for *Eco*RI DNA restriction and modification enzymes. *J. Biol. Chem.* **256,** 2131–2139.
37. Kossykh, V., Repyk, A., Kaliman, A., and Buryanov, Y. (1989) Nucleotide sequence of the *Eco*RII restriction endonuclease gene. *Biochim. Biophys. Acta* **1009,** 290–292.
38. Bhagwat, A. S., Johnson, B., Weule, K., and Roberts, R. J. (1990) Primary sequence of the *Eco*RII endonuclease and properties of its fusions with β-galactosidase. *J. Biol. Chem.* **265,** 767–773.
39. Bougueleret, L., Schwanstein, M., Tsugita, A., and Zabeau, M. (1984) Characterization of the genes coding for the *Eco*RV restriction and modification system of *Escherichia coli*. *Nucleic Acids Res.* **12,** 3659–3676.
40. Kita, K, Kotani, H., Sugisaki, H., and Takanami, M. (1989) The *Fok*I restriction-modification system. *J. Biol. Chem.* **264,** 5751–5756.
41. Looney, M. C., Moran, L. S., Jack, W. E., Feehery, G. R., Benner, J. S., Slatko, B. E., and Wilson, G. G. (1989) Nucleotide sequence of the *Fok*I restriction-modification system: separate strand-specificity domains in the methyltransferase. *Gene* **80,** 193–203.

42. Schoner, B., Kelly, S., and Smith, H. O. (1983) The nucleotide sequence of the *Hha*II restriction and modification genes from *Haemophilus haemolyticus*. *Gene* **24,** 227–236.
43. Chandrasegaran, S., Lunnen, K. O., Smith, H. O., and Wilson, G. C. (1988) Cloning and sequencing the *Hin*fI restriction and modification genes. *Gene* **70,** 387–392.
44. Lin, P. M., Lee, C. H., and Roberts, R. J. (1989) Cloning and characterization of the genes encoding the *Msp*I restriction modification system. *Nucleic Acids Res.* **17,** 3001–3011.
45. Theriault, G., Roy, P. H., Howard, K., Benner, J. S., Brooks, J. E., Waters, A. F., and Gingeras, T. R. (1985) Nucleotide sequence of the *Pae*R7 restriction modification system and partial characterization of its protein products. *Nucleic Acids Res.* **13,** 8441–8461.
46. Walder, R. Y., Walder J. A., and Donelson, J. E. (1984) The organization and complete nucleotide sequence of the *Pst*I restriction-modification system. *J. Biol. Chem.* **259,** 8015–8026.
47. Stephenson, F. H., Ballard, B. T., Boyer, H. W., Rosenberg, J. M., and Greene, P. J. (1989) Comparison of the nucleotide and amino acid sequences of the *Rsr*I and *Eco*RI restriction endonucleases. *Gene* **85,** 1–13.
48. Szilak, L.,Venetianer, P., and Kiss, A. (1990) Cloning and nucleotide sequence of the genes coding for the *Sau*961 restriction and modification enzymes. *Nucleic Acids Res.* **18,** 4659–4664.
49. Karreman, C. and de Waard, A. (1988) Cloning and complete nucleotide sequences of the typeII restriction-modification genes of *Salmonella infantis*. *J. Bact.* **170,** 2527–2532.
50. Heidmann, S., Seifert, W., Kessler, C., and Domdey, H. (1989) Cloning, characterization and heterologous expression of the *Sma*I restriction-modification system. *Nucleic Acids Res.* **17,** 9783–9796.
51. Slatko, B. E., Benner, J. S., Jager-Quinton, T., Moran, L. S., Simcox, T. G., van Cott, E. M., and Wilson, G. G. (1987) Cloning sequencing and expression of the *Taq*I restriction modification system. *Nucleic Acids Res.* **15,** 9781–9796.
52. Lauster, R. (1989) Evoution of type II DNA methyltransferases. A gene duplication model. *J. Mol. Biol.* **206,** 313–321.
53. Lee, Y. H. and Chirikjian, J. B. (1979) Sequence-specific endonuclease BglI: Modification of lysine and arginine residues of the homogeneous enzyme. *J. Biol. Chem.* **254,** 6838–6841.
54. Koncz, C., Kiss, A., and Venetianer, P. (1978) Biochemical characterisation of the restriction-modification system of *Bacillus sphaericus*. *Eur. J. Biochem.* **89,** 523–529.
55. Bakzi, K. and Rushizky, G. (1979) Purification of the restriction enzyme *Pal*I. *Anal. Biochem.* **99,** 207–212.
56. McClarin, J. A., Frederick, C. A., Wang, B.-C. Greene, P., Boyer, H. W., Grable, J., and Rosenberg, J. M. (1986) Structure of the DNA-*Eco*RI endonuclease recognition complex at 3 Å resolution. *Science* **234,** 1526–1541.

57. Chandrasegaran, S., Smith, H. O., Amzel, M. L., and Ysern, X. (1986) Preliminary X-ray diffraction analysis of HhaII endonuclease-DNA cocrystals. *Proteins Struct. Funct. Genet.* **1,** 263–266.
58. Kelly, T. J., Jr. and Smith, H. O. (1970) A restriction enzyme from Haemophilus influenzae II. Base sequence of the restriction site. *J. Mol. Biol.* **51,** 393–409.
59. Woodhead, J. L., Bhave, N., and Malcolm, A. D. B. (1981) Cation dependence of restriction endonuclease *Eco*RI activity. *Eur. J. Biochem.* **115,** 293–296.
60. Goppelt, M., Pingoud, A., Maass, G., Mayer, H. Köster, H., and Frank, R. (1980) The interaction of the *Eco*RI restriction endonuclease with its substrate. *Eur. J. Biochem.* **104,** 101–107.
61. Woodhead, J. L. and Malcolm, A. D. B. (1980) Non-specific binding of restriction endonuclease *Eco*RI to DNA. *Nucleic Acids Res.* **8,** 389–402.
62. Langowski, J., Pingoud, A., Goppelt, M., and Maass, G. (1980) Inhibition of *Eco*RI action by polynucleotides. A characterization of the non-specific binding of the enzyme to DNA. *Nucleic Acids Res.* **8,** 4727–4736.
63. Hinsch, B., Mayer, H., and Kula, M.-R. (1980) Binding of nonsubstrate nucleotides to a restriction endonuclease: a model for the interaction of *Bam*HI with its recognition sequence. *Nucleic Acids Res.* **8,** 2547–2559.
64. Jack, W. E., Terry, J. B., and Modrich, P. (1982) Involvement of outside DNA sequences in the major kinetic path by which *Eco*RI endonuclease locates and leaves its recognition sequence. *Proc. Natl. Acad. Sci. USA* **79,** 4010–4014.
65. Ehbrecht, H.-J., Pingoud, A., Urbanke, C., Maass, G., and Gualerzi, C. (1985) Linear diffusion of restriction endonucleases on DNA. *J. Biol. Chem.* **260,** 6160–6166.
66. Nardone, G., George, J., and Chirikjian, J. G. (1986) Differences in the kinetic properties of *Bam*HI endonuclease and methylase with linear DNA substrates. *J. Biol. Chem.* **26,** 12,128–12,133.
67. Woodhead, J. L. and Malcolm, A. D. B. (1980) Non-specific binding of restriction endonuclease *Eco*RI to DNA. *Nucleic Acids Res.* **8,** 389–402.
68. Langowski, J., Alves, A., Pingoud, A., and Maass, G. (1983) Does the specific recognition of DNA by the restriction endonuclease *Eco*RI involve a linear diffusion step? Investigation of the processivity of the EcoRI endonuclease. *Nucleic Acids Res.* **11,** 501–513.
69. Terry, B. J., Jack, W. E., and Modrich, P. (1985) Facilitated diffusion during catalysis by *Eco*RI endonuclease. *J. Biol. Chem.* **260,** 13,130–13,137.
70. Yolov, A. A., Gromova, E. S., and Shabarova, Z. A. (1985) Processive cleavage of concatamer DNA duplexes by *Eco*RII restriction endonuclease. *Mol. Biol. Rep.* **10,** 173–176.
71. Jack, W. E., Rubin, R. A., Newman, A., and Modrich, P. (1980) Restriction endonucleases, in *Gene Amplification and Analysis,* vol. 1 (Chirikjian, J. G., ed.), Elsevier, New York, pp. 165–179.
72. Terry, B. J., Jack, W. E., Rubin, R. A., and Modrich, P. (1983) Thermodynamic parameters governing interaction of *Eco*RI endonuclease with specific and nonspecific DNA sequences. *J. Biol. Chem.* **258,** 9820–9825.

73. Jen-Jacobson, L., Kurpiewski, M., Lesser, D., Grable, J., Boyer, H. W., Rosenberg, J. M. and Greene, P. (1983) Coordinate ion pair formation between *Eco*RI endonuclease and DNA. *J. Biol. Chem.* **256,** 14,638–14,646.
74. Goppelt, M., Langowski, J., Pingoud, A., Haupt, W., Urbanke, C., Mayer, H., and Maass, G. (1981) The effect of several nucleic acid binding drugs on the cleavage of dGGAATTCC and pBR322 by the *Eco*RI restriction endonuclease. *Nucleic Acids Res.* **9,** 6115–6127.
75. Alves, J., Urbanke, C., Fliess, A., Maass, G., and Pingoud, A. (1989) Fluorescence stopped-flow kinetics of the cleavage of synthetic oligodeoxynucleotides by the *Eco*RI restriction endonuclease. *Biochemistry* **28,** 7879–7888.
76. Imber, R. and Bickle, T. A. (1981) Purification and properties of the restriction endonuclease *Bgl*II from *Bacillus globigii. Eur. J. Biochem.* **117,** 395–399.
77. Bennett, S. P. (1987) Restriction enzymes with asymmetric sites. Ph. D. Thesis, University of Bristol.
78. Greene, P. J., Poonian, M. S., Nussbaum, A. L., Tobias, L., Garfin, D. E., Boyer, H. W. and Goodman, H. M. (1975) Restriction and modification of a self-complementary octanucleotide containing the *Eco*RI substrate. *J. Mol. Biol.* **99,** 237–261.
79. Modrich, P. and Zabel, D. (1976) *Eco*RI endonuclease; physical and catalytic properties of the homogeneous enzyme. *J. Biol. Chem.* **251,** 5866–5874.
80. Berkner, K. L. and Folk, W. R. (1977) *Eco*RI cleavage and methylation of DNAs containing modified pyrimidines in the recognition sequence. *J. Biol. Chem.* **252,** 3185–3193.
81. Halford, S. E. and Johnson, N. P. (1981) The *Eco*RI restriction endonuclease, covalently closed DNA and ethidium bromide. *Biochem. J.* **199,** 767–777.
82. Wolfes, H., Alves, J., Fliess, A., Geiger, R., and Pingoud, A. (1986) Site directed mutagenesis experiments suggest that Glu111, Glu144, and Arg145 are essential for endonucleolytic activity of *Eco*RI. *Nucleic Acids Res.* **14,** 9063–9080.
83. Brennan, C. A., van Cleve, M. D., and Gumport, R. I. (1986) The effects of base analogue substitutions on the cleavage by the *Eco*RI restriction endonuclease of octadeoxyribonucleotides containing modified *Eco*RI recognition sequences. *J. Biol. Chem.* **261,** 7270–7278.
84. McLaughlin, L. W., Bensler, F., Graeser, E., Piel, N., and Scholtissek S. (1987) Effects of functional group changes in the *Eco*RI recognition site on the cleavage reaction catalyzed by the endonuclease. *Biochemistry* **26,** 7238–7245.
85. Alves, J., Rüter, T., Geiger, R., Fiess, A., Maass, G., and Pingoud, A. (1989) Changing the hydrogen-bonding potential in the DNA binding site of *Eco*RI by site-directed mutagenesis drastically reduces the enzymatic activity, not, however, the preference of this restriction endonuclease for cleavage within the site -GAATTC-. *Biochemistry* **28,** 2678–2684.
86. Halford, S. E. and Goodall, A. J. (1988) Modes of DNA cleavage by the *Eco*RV restriction endonuclease. *Biochemistry* **27,** 1771.
87. Fliess, A., Wolfes, H., Seela, F., and Pingoud, A. (1988) Analysis of the recognition mechanism involved in the *Eco*RV catalyzed cleavage of DNA using modified oligodeoxynucleotides. *Nucleic Acids Res.* **16,** 11,781–11,787.

88. Kaddurah-Daouk R., Che, P., and Smith, H. O. (1985) Catalytic properties of the *Hha*II restriction endonuclease. *J. Biol. Chem.* **260,** 15,345–15,351.
89. Kubareva, E. A., Pein, G.-D., Gromova, E. S., Kuznezova, S. A., Tashlitzki, V. N., Cech, D., and Shaborova, Z. A. (1988) The role of modification in oligonucleotides in sequence recognition by *Mva*I restriction endonuclease. *Eur. J. Biochem.* **175,** 615–618.
90. Aiken, C. and Gumport, R. I. (1988) Restriction endonuclease *Rsr*I from *Rhodobacter sphaeroides*, an isoschizomer of *Eco*RI: purification and properties. *Nucleic Acids Res.* **16,** 7901–7916.
91. Maxwell, A. and Halford, S. E. (1982) The *Sal*GI restriction endonuclease. *Biochem. J.* **203,** 93–98.
92. Dwyer-Hallquist, P., Kezdy, F. J., and Agarwal, K. L. (1982) Interaction of the *Hpa*I endonuclease with synthetic oligonucleotides. *Biochemistry* **21,** 4693–4700.
93. Kita, K., Hiraoka, N., Kimizuka, F., Obayashi, A., Kojima, H., Takahashi, H., and Saito, H. (1985) Interaction of the restriction endonuclease *Sca*I with its substrates. *Nucleic Acids Res.* **13,** 7015–7024.
94. Pohl, F. M., Thomae, R., and Karst, A. (1982) Temperature dependence of the activity of DNA-modifying enzymes: endonucleases and DNA ligase. *Eur. J. Biochem.* **123,** 141–152.
95. Fersht, A. (1985) *Enzyme Structure and Mechanism* (2nd ed.), W. H. Freeman, New York.
96. Langowski, J., Urbanke, C., Pingoud, A., and Maass, G. (1981) Transient cleavage kinetics of the *Eco*RI restriction endonuclease measured in a pulsed quench-flow apparatus: enzyme concentration dependent activity change. *Nucleic Acids Res.* **9,** 3483–3490.
97. Halford, S. E. and Johnson, N. P. (1983) Single turnovers of the *Eco*RI restriction endonuclease. *Biochem. J.* **211,** 405–415.
98. King, K., Benkovic, S. J., and Modrich, P. (1989) Glu-111 is required for activation of the DNA cleavage center of *Eco*RI endonuclease. *J. Biol. Chem.* **264,** 11,807–11,815.
99. Snousnou, G. and Malcolm, A. D. B. (1984) Supercoiling and the mechanism of restriction endonucleases. *Eur. J. Biochem.* **138,** 275–280.
100. Thomas, M. and Davis, R. W. (1975) Studies on the cleavage of bacteriophage Lambda DNA with *Eco*RI restriction endonuclease. *J. Mol. Biol.* **91,** 315–328.
101. Halford, S. E., Johnson, N. P., and Grinsted, J. (1980) The *Eco*RI restriction endonuclease with bacteriophage λ DNA. *Biochem. J.* **191,** 581–592.
102. Berkner, K. L. and Folk, W. R. (1983) An assay for the rates of cleavage of specific sites in DNA by restriction endonucleases: its use to study the cleavage of phage λ DNA by *Eco*RI and phage P22 DNA containing thymine or 5-bromouracil by *Hin*dIII. *Biochemistry* **129,** 446–456.
103. Goldstein, L., Thomas, M., and Davis, R. W. (1975) *Eco*RI endonuclease cleavage map of bacteriophage P4-DNA. *Virology* **66,** 420–427.
104. Forsblom, S., Rigler, R., Ehrenberg, M., Pettersson, U., and Philipson, L. (1976) Kinetic studies of the cleavage of adenovirus DNA by restriction endonuclease *Eco*RI. *Nucleic Acids Res.* **3,** 3255–3269.

105. Armstrong, K. and Bauer, W. K (1982) Preferential site-dependent cleavage by restriction endonuclease *Pst*I. *Nucleic Acids Res.* **70,** 993–1007.
106. Alves, J., Pingoud, A., Haupt, W., Langowski, J., Peters, F., Maass, G., and Wolff, C. (1984) The influence of sequence adjacent to the recognition site on the cleavage of oligodeoynucleotides by the *Eco*RI endonuclease. *Eur. J. Biochem.* **140,** 83–92.
107. Wolfes, H., Fliess, A., and Pingoud, A. (1985) A comparison of the structural requirements for DNA cleavage by the isoschizomers *Hae*III, *Bsp*RI, and *Bsu*RI. *Eur. J. Biochem.* **150,** 105–110.
108. Ruben, G., Spielman, P., Tu, C. D., Jay, E., Siegel, B., and Wu, R. (1977) Relaxed circular SV40 DNA as cleavage intermediate of two restriction endonucleases. *Nucleic Acids Res.* **4,** 1803–1813.
109. Rubin, R. A. and Modrich, P. (1978) Substrate dependence of mechanism of *Eco*RI endonuclease. *Nucleic Acids Res.* **5,** 2991–2997.
110. Halford, S. E., Johnson, N. P., and Grinsted, J. (1979) The reactions of the *Eco*RI and other restriction endonucleases. *Biochem. J.* **179,** 353–365.
111. Hinsch, B. and Kula, M.-R. (1981) Reaction kinetics of some important site-specific endonucleases. *Nucleic Acids Res.* **9,** 3159–3174.
112. Potter, B. V. L. and Eckstein, F. (1984) Cleavage of phosphothioate-substituted DNA by restriction endonucleases. *J. Biol. Chem.* **259,** 14,243–14,248.
113. Taylor, J. W., Ott, J., and Eckstein, F. (1985) The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA. *Nucleic Acids Res.* **13,** 8765–8785.
114. Thielking, V., Alves, J., Fliess, A., Maass, G., and Pingoud, A. (1990) The accuracy of the *Eco*RI restriction endonuclease: binding and cleavage studies with oligodeoxynucleotide substrates containing degenerate recognition sequences. *Biochemistry,* **29,** 4682–4691.
115. Connolly, B., Eckstein, F., and Pingoud, A. (1984) The stereochemical course of the restriction endonuclease *Eco*RI catalyzed reaction. *J. Biol. Chem.* **259,** 10,760–10,763.
116. Landgraf, A. (1987) Genauigkeit der Restriktionsendonuklease *Eco*RV. Diploma thesis, University of Hannover.
117. Taylor, J. W. and Halford, S. E. (1989) Discrimination between DNA sequences by the *Eco*RV restriction endonuclease. *Biochemistry* **28,** 6198–6207.
118. Bodnar, J. W., Zempsky, W., Warder, D., Bergson, C., and Ward, D. C. (1983) Effect of nucleotide analogs on the cleavage of DNA by the restriction enzymes *Alu*I, *Dde*I, *Hin*fI, *Rsa*I, and *Taq*I. *J. Biol. Chem.* **258,** 15,206–15,213.
119. Fliess, A., Wolfes, H., Rosenthal, A., Schwellnus, K, Blöcker, H., Frank, R., and Pingoud, A. (1986) Role of thymidine residues in DNA recognition by the *Eco*RI and *Eco*RV restriction endonucleases. *Nucleic Acids Res.* **14,** 3463–3474.
120. McLaughlin, L. W., Leong, T., Benseler, F., and Piel, N. (1988) A new approach to the synthesis of a protected 2-aminopurine derivative and its incorporation into oligodeoxynucleotides containing the *Eco*RI and *Bam*HI recognition sites. *Nucleic Acids Res.* **16,** 5631–5644.

121. Mazzarelli, J., Scholtissek, S., and McLaughlin, L.W. (1989) Effects of functional group changes in the *Eco*RV recognition site on the cleavage reaction catalyzed by the endonuclease. *Biochemistry* **28,** 4616–4622.
122. Baumstark, B. R., Roberts, R. J., and RajBhandary, U. L. (1979) Use of short synthetic DNA duplexes as substrates for the restriction endonucleases *Hpa*II and *Mno*I. *J. Biol. Chem.* **254,** 8943–8950.
123. Nwosu, V. U., Connolly, B. A., Halford, S. E., and Garnett, J. (1988) The cloning, purification and characterization of the *Eco*RV modification methylase. *Nucleic Acids Res.* **16,** 3705–3720.
124. Yanofsky, S. D., Love, R., McClarin, J. A. Rosenberg, J. M., Boyer, H. W., and Greene, P. J. (1987) Clustering of null mutations in the *Eco*RI endonuclease. *Proteins: Struct. Funct. Genet.* **2,** 273–282.
125. Geiger, R., Rüter, T., Alves, J., Fliess, A., Wolfes, H., Pingoud, V., Urbanke, C., Maass, G., Pingoud, A., Düsterhöft, A., and Kröger, M. (1989) Genetic engineering of *Eco*RI mutants with altered amino acid residues in the DNA binding site: physicochemical investigations give evidence for an altered monomer/dimer equilibrium for the Gln144Lys145 and Gln144Lys145Lys200 mutants. *Biochemistry* **28,** 2667–2677.
126. Needles, M. C., Fried, S. R., Love, R., Rosenberg, J. M., Boyer, H. W., and Greene P. J. (1989) Determinants of *Eco*RI endonuclease sequence discrimination. *Proc. Natl. Acad. Sci. USA* **86,** 3579–3583.
127. Oelgeschläger, T., Geiger, R., Rüter, T., Alves, J., Fliess, A., and Pingoud, A. (1990) Probing the function of individual amino acid residues in the DNA binding site of *Eco*RI restriction endonuclease by analysing the toxicity of genetically engineered mutants. *Gene,* **89,** 19–27.
128. Nelson, M. and McClelland, M. (1989) Effect of site-specific methylation on DNA. Modification methyltransferases and restriction endonucleases. *Nucleic Acids Res.* **17,** r389–r415.
129. Song, Y.-H., Rüter, T., and Geiger, R. (1988) DNA cleavage by *Aat*I and *Stu*I is sensitive to *Escherichia coli* dcm methylation. *Nucleic Acids Res.* **16,** 2718.
130. Waalwijk, C. and Flavell R. A. (1978) *Msp*I, an isoschizomer of *Hpa*II which cleaves both unmethylated and methylated *Hpa*II sites. *Nucleic Acids Res.* **5,** 3231–3236.
131. Sneider, T. W. (1980) The 5'-cytosine in CCGG is methylated in two eucaryotic DNAs and *Msp*I is sensitive to methylation at this site. *Nucleic Acids Res.* **8,** 3829–3840.
132. Lacks, S. A. and Greenberg, B. (1977) Complementary specificity of restriction endonucleases of *Diplococcus pneumoniae* with respect to DNA methylation. *J. Mol. Biol.* **114,** 153–168.
133. Polisky, B., Greene, P., Garfin, D. E., McCarthy, B. J., Goodman, H. M., and Boyer, H. W. (1975) Specificity of substrate recognition by the *Eco*RI restriction endonuclease. *Proc. Nat. Acad. Sci. USA* **72,** 3310–3314.
134. Hsu, M. and Berg, P. (1978) Altering the specificity of restriction endonuclease: effect of replacing Mg^{2+} with Mn^{2+}. *Biochemistry* **17,** 131–138.

135. Goodman, H. M., Greene, P. J., Garfin, D. E., and Boyer, H. W. (1977) DNA site recognition by the *Eco*RI restriction endonuclease and modification methylase, in *Nucleic Acids-Protein Recognition* (Vogel, H. J., ed.), Academic, New York, pp. 239–259.
136. Tikchonenko, T. L, Karamov, E. V., Zavizion, B. A., and Naroditsky, B. S. (1978) *Eco*RI* activity: enzyme modification or activation of accompanying endonucleases? *Gene* **4,** 195–212.
137. Mayer, H. (1978) Optimization of the *Eco*RI*-activity of *Eco*RI endonuclease. *FEBS Lett.* **90,** 341–344.
138. Malyguine, E., Vannier, P., and Yot, P. (1980) Alteration of the specificity of restriction endonucleases in the presence of organic solvents. *Gene* **8,** 163–177.
139. Gardner, R. C., Howarth, A. J., Messing, J., and Shepherd, R. J. (1982) Cloning and sequencing of restriction fagments generated by *Eco*RI. *DNA* **1,** 109–115.
140. George, J. and Chirikjian, J. G. (1982) Sequence-specific endonuclese *Bam*HI: relaxation of sequence recognition. *Proc. Natl. Acad. Sci. USA* **79,** 2432–2438.
141. George, J., Blakesley, R. W., and Chirilian, J. G. (1980) Sequence-specific endonuclease *Bam*HI: effect of hydrophobic reagents on sequence recognition and catalysis. *J. Biol. Chem.* **255,** 6521–6524.
142. Clarke, C. M. and Hartley, B. S. (1979) Purification, properties and specificity of the restriction endonuclease from *Bacillus stearothermophilus. Biochem. J.* **177,** 49–62.
143. Heininger, K., Hörz, W., and Zachau, H. G. (1977) Specificity of cleavage by a restriction nuclease from *Bacillus subtilis. Gene* **1,** 291–303.
144. Halford, S. E., Lovelady, B. M., and McCallum, S. A. (1986) Relaxed specificity of the *Eco*RV restriction endonuclease. *Gene* **41,** 173–181.
145. Nasri, M. and Thomas, D. (1986) Relaxation of recognition sequence of specific endonuclease *Hin*dIII. *Nucleic Acids Res.* **14,** 811–821.
146. Nasri, M., Sayadi, S., and Thomas, D. (1985) Relaxation of *Pvu*II recognition sequence. *FEBS Lett.* **185,** 101–104.
147. Barany, F. (1988) The *Taq*I "star" reaction: strand preferences reveal hydrogen-bond donor and acceptor sites in canonical sequence recognition. *Gene* **65,** 149–165.
148. Pingoud, A. (1985) Spermidine increases the accuracy of type II restriction endonucleases. *Eur. J. Biochem.* **147,** 105–109.
149. Horiuchi, K. and Zinder, N. D. (1975) Site-specific cleavage of single-stranded DNA by a *Haemophilus* restriction endonuclease. *Proc. Natl. Acad. Sci. USA* **72,** 2555–2558.
150. Blakesley, R. W. and Wells, R. D. (1975) "Single-stranded" DNA from fX174 and M13 is cleaved by certain restriction endonucleases. *Nature* **257,** 421,422.
151. Godson, G. N. and Roberts, R. J. (1976) A catalogue of cleavages of ϕX174, S13, G4 and ST-1 DNA by 26 different restriction endonucleases. *Virology* **73,** 561–567.
152. Blakesley, R. W., Dodgson, J. B., Nes, I. F., and Wells, R. D. (1977) Duplex regions in "single-stranded" ϕX174 DNA are cleaved by a restriction endonuclease from *Haemophilus aegypticus. J. Biol. Chem.* **252,** 7300–7306.

153. Yoo, O. J. and Argarwal, K. L. (1980) Isolation and characterization of two proteins possessing *Hpa*II methylase activity. *J. Biol. Chem.* **255,** 6445–6449.
154. Bischofsberger, N., Ng, P. G., Webb, T. R., and Matteucci, M. D. (1987) Cleavage of single stranded oligonucleotides by *Eco*RI restriction endonuclease. *Nucleic Acids Res.* **15,** 709–716.
155. Molloy, P. L. and Symons, R. H. (1980) Cleavage of DNA•RNA hybrids by type II restriction enzymes. *Nucleic Acids Res.* **8,** 2939–2946.
156. Woodhead, J. L. and Malcolm, A. D. B. (1981) The essential carboxyl group in restriction endonuclease *Eco*RI. *Eur. J. Biochem.* **120,** 125–128.
157. Wolfes, H., Fliess, A., Winkler, F., and Pingoud, A. (1986) Cross-linking of bromodeoxyuridine-substituted oligonucleotides to the *Eco*RI and *Eco*RV restriction endonucleases. *Eur. J. Biochem.* **159,** 267–273.
158. Luke, P. A. and Halford, S. E. (1986) Fluorescent labeling of *Eco*RI and *Eco*RV. *Biochem. Soc. Trans.* **14,** 259–260.
159. Nath, K. (1981) Effect of sulfhydryl group inhibitors on restriction endonuclease activities. *Arch. Biochem. Biophys.* **212,** 611–617.
160. Kania, J. and Fanning, T. G. (1976) Use of a sequence-specific DNA-binding ligand to probe the environments of *Eco*RI restriction endonuclease cleavage sites. *Eur. J. Biochem.* **67,** 367–371.
161. Nosikov, V. V., Braga, E. A., Karlishev, A. V., Zhuze, A. L., and Polyanovsky, O. L. (1976) Protection of particular cleavage sites of restriction endonucleases by distamycin A and actinomycin D. *Nucleic Acids Res.* **3,** 2293–2301.
162. Kuroyedov, A. A., Grokhovsky, S. L., Zhuze, A. L., Nosikov, V. V., and Polyanovsky, O. L. (1977) Distamycin A and its analogs as agents for blocking of endo-R-*Eco*RI activity. *Gene* **1,** 389–395.
163. Nosikov, V. V. and Sain, B. (1977) Protection of particular R•*Hin*dIII cleavage sites by distamycin A, propyldistamycin and netropsin. *Nucleic Acids Res.* **4,** 2263–2273.
164. Polyanovsky, O. L., Nosikov, V. V., Zhuze, A. L., Braga, E., and Karlyshev, A. V. (1979) Regulation of restriction endonuclease activity with antibiotics. *Adv. Enzyme Regul.* **17,** 307–321.
165. Cohen, G. L., Ledner, J. A., Bauer, W. R., Ushay, H. M., Caravana, C., and Lippard, S. J. (1980) Sequence dependent binding of *cis*-dichlordiammineplatinium(II) to DNA. *J. Am. Chem. Soc.* **102,** 2487–2488.
166. Sumner, W. and Bennett, G. N. (1981) Anthramycin inhibition of restriction endonuclease cleavage and its use as a reversible blocking agent in DNA constructions. *Nucleic Acids Res.* **9,** 2105–2119.
167. Ushay, M. H., Tullius, T. D., and Lippard, S. J. (1981) Inhibition of the *Bam*HI cleavage and unwinding of pBR322 deoxyribonucleic acid by the antitumor drug *cis*-dichlorodiammineplatinium(I). *Biochemistry* **20,** 3744–3748.
168. Malcolm, A. D. B. and Moffatt, J. R. (1981) Differential reactivities at restriction enzyme sites. *Biochim. Biophys. Acta* **655,** 128–135.
169. Malcolm, A. D. B., Moffatt, J. R., Fox, K. R., and Waring, M. J. (1983) Differential inhibition of a restriction enzyme by quinoxaline. *Biochim. Biophys. Acta* **699,** 211–216.

170. Nilsson, M.-G., Skarped, C., and Magnusson, G. (1982) Structure at restriction endonuclease *Mbo*I cleavage sites protected by actinomycin D or distamycin A. *FEBS Lett.* **145,** 360–364.
171. Osterlund, M., Luthman, H., Nisson, S. V., and Magnusson, G. (1982) Ethidium bromide-inhibited restriction endonucleases cleave one strand of circular DNA. *Gene* **20,** 121–125.
172. Langowski, J. and Urbanke, C. (1982) A continuous fluorimetric assay for the *Eco*RI restriction endonuclease. *Fresen. Z. Anal. Chem.* **311**, 367.
173. Francois, J.-C., Saison-Behmoaras, T., Tuong, N. T., and Helene, C. (1989) Inhibition of restriction endonuclease cleavage via triple helix formation by homopyrimidine oligonucleotides. *Biochemistry* **28,** 9617–9619.
174. Hanvey, J. C., Shimizu, M., and Wells, R. D. (1990) Site specific inhibition of *Eco*RI restriction/modification enzymes by a DNA triple helix. *Nucleic Acids Res.* **18,** 157–161.
175. Notes on enzymes dispatched in dry ice. (1988) *Boehringer Colloquium* **4,** 8.
176. Enzyme storage in frost-free freezers. (1985) *BRL Focus* **7,** 5.
177. Krüger, D. H., Barcak G. J., Renter, M., and Smith, H. O. (1988) *Eco*RII can be activated to cleave refractory DNA recognition sites. *Nucleic Acids Res.* **16,** 3997–4008.
178. Conrad, M. and Topal, M. O. (1989) DNA and spermidine provide a switch mechanism to regulate the activity of restriction enzyme *Nae*I. *Proc. Natl. Acad. Sci. USA* **86,** 9707–9711.
179. Sharp, P. A., Sugden, B., and Sambrock, J. (1973) Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. *Biochemistry* **12,** 3055–3063.
180. Janulaitis, A. A. and Vaitkevitchius, D.P. (1981) A spectrophotometric procedure for the determination of activity of restriction endonucleases. *Anal. Biochem.* **116,** 116–122.
181. McClelland, M., Hanish, J., Nelson, M., and Patel, Y. (1988) KGB: a single buffer for all restriction endonucleases. *Nucleic Acids Res.* **16,** 364.
182. Maxam, A. M. and Gilbert, W. (1980) Sequencing end labeled DNA with base specific chemical cleavages. *Methods Enzymol.* **65,** 499–560.
183. Jay, E., Banbara, R., Padmanabham, R., and Wu, R. (1974) DNA-sequence analysis: a general, simple and rapid method for sequencing large oligonucleotide fragments by mapping. *Nucleic Acids Res.* **1,** 331–353.
184. Tu, C. P. D., Jay, E., Bahl, C. P., and Wu, R. (1976) A reliable mapping method for sequence determination of oligodeoxyribonucleotides by mobility shift analysis. *Anal. Biochem.* **74,** 73–79.
185. Pingoud, A., Fliess, A., and Pingoud, V. (1989) HPLC of oligonucleotides, in *HPLC of Macromolecules, a Practical Approach* (Oliver, R. W. A., ed.), IRL Press, Oxford, UK, pp. 183–208.
186. Smith, C. L., Econome, J. G., Schutt, A., Klco, S., and Cantor, C. R. (1987) Physical map of the *Escherichia-coli*-K12 genome. *Science* **236,** 1448–1453.
187. Bautsch, W. (1989) Rapid physical mapping of the *Mycoplasma mobile* genome by two-dimensional field inversion gel electrophoresis techniques. *Nucleic Acids Res.* **16,** 11,461–11,467.

188. Botstein, D., White, R. L., Skolnik, M., and Davis, R. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32,** 314–333.
189. Orkin, S. H. (1986) Reverse genetics and human disease. *Cell* **47,** 845–850.
190. Jeffreys, A. J. (1987) Highly variable minisatellites and DNA fingerprints. *Biochem. Soc. Trans.* **15,** 309–317.
191. Gill, P., Lygo, J. E., Fowler, S. J., and Werrett, D. J. (1987) An evaluation of DNA fingerprinting for forensic purposes. *Electrophoresis* **8,** 38–44.
192. Szybalski, W. (1985) Universial restriction endonuclease: Designing novel cleavage specificities by combining adapter oligodeoxynucleotide and enzyme moieties. *Gene* **40,** 169–173.
193. Podhajska, A. J. and Szybalski, W. (1985) Conversion of the *Fok*I endonudease to a universial restriction enzyme: Cleavage of phage M13mp7 DNA at predetermined sites. *Gene* **40,** 175–182.
194. Gaastra, W. and Josephson, J. (1985) Radiolabeling of DNA using polynucleotide kinase, in *Methods in Molecular Biology, vol. 2: Nucleic Acids* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 263–268.
195. Gaastra, W. and Klemm, P. (1985) Radiolabeling of DNA with 3'-terminal transferase, in *Methods in Molecular Biology, vol. 2: Nucleic Acids* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 269–272.
196. Kleppe, K. and Lillehaug, J. R. (1978) Polynucleotide kinase. *Adv. Enzymol.* **48,** 245–275.
197. Berkner, K. L. and Folk W. R. (1977) Polynucleotide kinase exchange reaction, quantitative assay for restriction endonudease-generated 5'-phosphoryl termini in DNAs. *J. Biol. Chem.* **252,** 3176–3184.
198. Kramer, W., Drutsa, V., Jansen, H.-W., Kramer, B., Pilugfelder, M., and Fritz, H. J. (1984) The gapped duplex DNA approach to oligonudeotide-directed mutation construction. *Nucleic Acids Res.* **12,** 9441–9456.
199. Kim, Y., Grable, J. C., Love, R., Greene, P. J., and Rosenberg, J. M. (1990) Refinement of *Eco*RI endonuclease crystal structure: A revised protein chain tracing. *Science* **249,** 1307–1309.
200. Rosenberg, J. M. (1991) Structure and function of restriction endonucleases. *Curr. Op. Struct. Biol.* **1,** 104–113.
201. Heitman, J. (1992) How the *Eco*RI endonuclease recognizes and cleaves DNA. *BioEssays* **14,** 445–454.
202. Jeltsch, A., Alves, J., Oelgeschläger, T., Wolfes, H., Maass, G., and Pingoud, A. (1992) A mutational analysis of the function of Gln115 in the *Eco*RI restriction endonuclease, a critical amino acid for recognition of the inner thymidine in the sequence -GAATTC- and for coupling specific DNA binding to catalysis. *J. Mol. Biol.,* in press.
203. Winkler, F. K. (1992) Structure and function of restriction endonucleases. *Curr. Op. Struct. Biol.* **2,** 93–99.
204. Halford, S. E., Taylor, J. D., Vermote, C. L. M., and Vipond, I. B. (1993) Mechanism of action of restriction endonuclease *Eco*RV, in *Nucleic Acids and Molecular Biology*, vol. 7 (Eckstein, F. and Lilley, D. M. J., eds.), Springer, Berlin, in press.

205. Thielking, V., Selent, U., Koehler, E., Wolfes, H., Pieper, U., Geiger, R., Urbanke, C., Winkler, F. K., and Pingoud, A. (1991) Site-directed mutagenesis studies with *Eco*RV restriction endonuclease to identify regions involved in recognition and catalysis. *Biochemistry* **30,** 6416–6422.
206. Selent, U., Rüter, T., Köhler, E., Liedtke, M., Thielking, V., Alves, J., Oelgeschläger, T., Wolfes, H., Peters, F., and Pingoud, A. (1992) Site directed mutagenesis study to identify amino acid residues involved in the catalytic function of the restriction endonuclease *Eco*RV. *Biochemistry* **31,** 4808–4815.
207. Vermote, C. L. M., Vipond, B., and Halford, S. E. (1992) *Eco*RV restriction endonuclease: communication between DNA recognition and catalysis. *Biochemistry* **31,** 6089–6097.
208. Taylor, J. D., Badcoe, I. G., Clarke, A. R., and Halford, S. E. (1991) *Eco*RV restriction endonuclease binds all DNA sequences with equal affinity. *Biochemistry* **30,** 8743–8753.
209. Thielking, V., Selent, U., Köhler, E., Landgraf, A., Wolfes, H., Alves, J., and Pingoud, A. (1992) Mg^{2+} confers DNA binding specificity to the *Eco*RV restriction endonuclease. *Biochemistry* **31,** 3727–3732.
210. Zebala, J. A., Choi, J., Trainor, G. L., and Barany, F. (1992) DNA recognition of base analogue and chemically modified substrates by the *Taq*I restriction endonuclease. *J. Biol. Chem.* **267,** 8106–8116.
211. Grasby, J. A. and Connolly, B. A. (1992) The stereochemical outcome of the hydrolysis reaction catalyzed by the *Eco*RV restriction endonuclease. *Biochemistry*, in press.
212. Jeltsch, A., Alves, J., Maass, G., and Pingoud, A. (1992) On the catalytic mechanism of *Eco*RI and *Eco*RV: A detailed proposal based on biochemical results, structural data and molecular modelling. *FEBS Lett.* **304,** 4–8.
213. Xu, S. and Schildkraut, I. (1991) Isolation of *Bam*HI variants with reduced cleavage activities. *J. Biol. Chem.* **266,** 4425–4429.
214. Taylor, J. D., Goodall, A. J., Vermote, C. L., and Halford, S. E. (1990) Fidelity of DNA recognition by the *Eco*RV restriction/modification system *in vivo*. *Biochemistry* **29,** 10,727–10,733.
215. Rotzal, L. (1992) Molekulargenetische Konstruktion, Aufreinigung und Charakterisierung des *Eco*RI:His_6 Fusionsproteins. Diploma thesis. University of Hannover.
216. Waters, T. R. and Connolly, B. A. (1992) Continuous spectrophotometric assay for restriction endonucleases using synthetic oligodeoxynucleotides and based on the hyperchromic effect. *Anal. Biochem.* **204,** 204–209.

CHAPTER 9

DNA Methyltransferases (EC 2.1.1.72 and EC 2.1.1.73)

David P. Hornby

1. Introduction

DNA methyltransferases (Mtases) catalyze the transfer of the *S*-methyl group of *S*-adenosylmethionine (SAM) to deoxycytosine (dC) or deoxyadenine (dA) bases within defined DNA sequences *(1)*. Individual enzymes are specific for one or the other base, and modify at the 6-NH_2 of dA (EC 2.1.1.72) or at the N^4 or 5-C position of dC (EC 2.1.1.73) depending on the particular enzyme *(2)*. The reaction is predominantly irreversible. Enzymes, such as O^6-methylguanine DNA Mtase, that participate in DNA repair processes will not be discussed here.

Both mammalian and bacterial DNA Mtases have a sequence-specificity component that is an integral part of the reaction. However, sequence recognition is much more relaxed in the case of mammalian Mtases. Moreover, all known mammalian DNA Mtases are dC-specific, whereas both dA and dC specificities are found in bacteria. A minimum of four and a maximum of eight specific bases are found in the DNA recognition sites of bacterial enzymes, whereas mammalian Mtases require only a CpG doublet. (There is some evidence that CpN sequences *[3]* are also modified in other higher organisms and that plant DNA Mtases can also methylate CpNpG motifs *[4]*.) The biological consequences of DNA methylation are diverse, but are usually the result of a modulation of protein:nucleic acid interactions induced by methylation *(1)*. Thus, for example, in prokaryotes, site-specific DNA methylation is suffi-

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

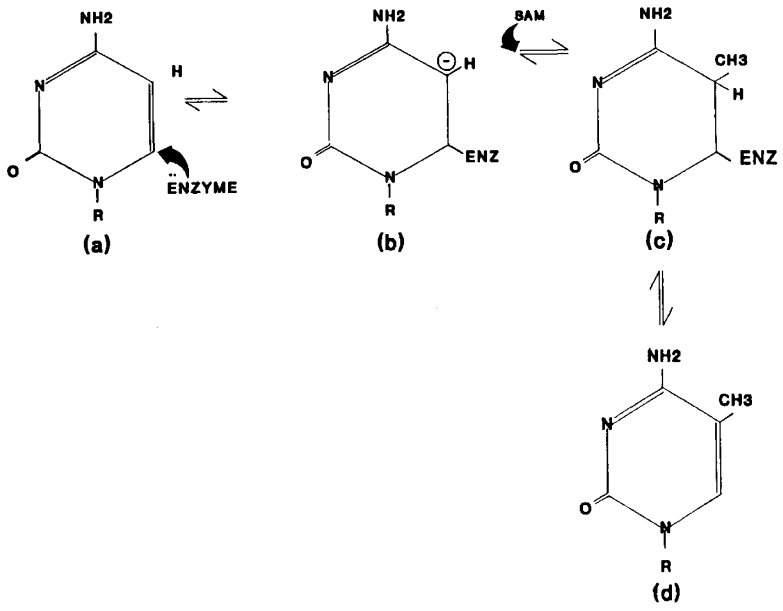


Fig. 1. Proposed mechanism of action of 5' cytosine-specific DNA methyltransferases based on the studies of thymidilate synthase and *Hha* I DNA methyltransferase.

cient to render a sequence refractory to endonucleolytic cleavage by a restriction enzyme of the same specificity (*see* Chapter 8, Sections 2.4.1. and 3.7.). This is the molecular basis of restriction and modification. In mammalian cells, cytosine methylation has been implicated in the regulation of gene transcription and in the inactivation of the X chromosome, although the molecular details of these processes remain unresolved *(5)*.

The reaction catalyzed by all DNA Mtases is shown below. SAM invariably acts as the methyl donor, and the natural acceptor is chromosomal DNA, but in vitro, plasmid DNA, phage DNA, or an appropriate oligonucleotide duplex can also act as methyl acceptors.

$$\text{DNA} + \text{SAM} \rightarrow \text{MeDNA} + \text{SAH}$$

The products of the reaction are methylated DNA and *S*-adenosylhomocysteine (SAH). The site of modification is enzyme-specific. Little is known about the mechanism of dA methylation, whereas studies on the mechanistically related enzyme thymidilate synthase have led to the proposed scheme for dC methylation shown in Fig.1 *(6)*. Interestingly, a catalytic cysteine flanked on the N-terminal side by a proline

residue is found in all of the pyrimidine-specific Mtases, which modify at a carbon atom so far characterized. Indeed, amino acid sequences of the dC DNA Mtases from mouse *(7)*, human (Andrews, P., Burton D. R., and Hornby, D. P., unpublished results), and a variety of bacteria *(8)*, although otherwise quite dissimilar, share a pro-cys dipeptide along with a putative specificity determining sequence toward the C-terminus of the polypeptide.

It has been suggested *(6)* that the first stage of C-5, dC methylation involves the nucleophilic attack by the cysteine-sulfhydryl group at the C-6 position of the heterocycle resulting in the transient formation of a covalent enzyme:nucleic acid adduct. Subsequent abstraction of a proton at the C-5 position following the addition of the activated methyl group of SAM leads to the formation of 5-methyldC with the concomitant release of the enzyme (*see* Fig. 1).

2. Enzyme Requirements

2.1. S-*Adenosylmethionine (SAM)*

Storage of this substrate is particularly important, since the half-life of SAM at pH 7.5 and 37°C is around 20 h. Therefore, SAM should be prepared immediately prior to assay, or if prolonged storage is desired, the solution should be made 5 m*M* with respect to sulfuric acid, 10% with respect to ethanol, and at a SAM concentration sufficient to facilitate subsequent dilution into a well-buffered assay mixture (e.g., between 1 and 10 m*M*). In the latter context, [^{3}H-methyl] SAM is supplied by the manufacturers in dilute sulfuric acid. In general, the apparent K_m for SAM for those enzymes that have been purified and kinetically characterized lies between 1 and 100 μ*M*, and therefore, a final concentration of 100 μ*M* of the substrate in the reaction incubation should be saturating. Of course, there may be instances where the latter concentration is either subsaturating ($K_m > 50$ μ*M*) or inhibitory (substrate inhibition), but this will only emerge following a detailed investigation of individual Mtase enzymes.

2.2. DNA

In order to assay the activity of a DNA Mtase during a purification protocol, [^{3}H-methyl] SAM is often employed as the methyl donor. In this way, a sensitive estimation of methyltransfer can be obtained by liquid scintillation counting. Alternatively, methylation of a DNA

substrate can be monitored by subsequent restriction enzyme analysis. In each case, the DNA substrate should be carefully chosen in order to maximize the incorporation of methyl groups during the incubation. For example, the rate of methylation of calf thymus DNA by a vertebrate DNA Mtase will be slower than that of a bacterial DNA preparation, since the frequency of methylated CpG dinucleotides will be much greater in the former substrate. Furthermore, it has been shown that in the case of mammalian and some (but not all) prokaryotic DNA Mtases, hemimethylated DNA is the preferred substrate. Therefore, a judicious choice of DNA substrate can improve the sensitivity of the assay considerably. In the case of dA methylation, a chromosomal DNA preparation from a *dam*$^-$ strain of *E. coli* will be free of adenine methylation and is therefore a convenient substrate for monitoring activity during purification.

More defined substrates may be required in those cases where the sequence specificity of the enzyme is known. For example, in order to determine the specific activity of the enzyme *Eco*RI Mtase, a plasmid or oligonucleotide duplex containing one or more GAATTC recognition sequences would be a suitable substrate. Incubation of the latter enzyme preparation with a plasmid containing no *Eco*RI sites would serve as a control against nonspecific methylation. This is especially important during the early stages of a protein purification.

2.3. Cofactor Requirements

Restriction and modification enzymes are classically divided into three groups that reflect their structural and functional complexity. The Type I modification enzymes are the most complex and require Mg^{2+} ions (~5 m*M*), and ATP (100 μ*M*) for optimal activity. In general, Type II and III enzymes catalyze methyltransfer in the absence of ATP and Mg^{2+}, although there are some exceptions, such as *Taq*I Mtase, which do require divalent cations. Mammalian DNA Mtases require the presence of neither of the above cofactors, although it has been reported that Co^{2+} (100 μ*M*) ions stimulate some mammalian enzymes.

2.4. Ionic Strength

All enzymes have an intrinsic optimal ionic strength for activity, and this can only be determined on an empirical basis. However, in general, the optimum salt (usually NaCl or KCl) concentration is that of the cor-

responding restriction endonuclease in the case of prokaryote DNA Mtases. Mammalian dC Mtases have been shown to be extremely sensitive to KCl: a concentration of 25 m*M* giving optimal activity with the majority of mammalian enzymes.

2.5. Buffer Components

Again for any given enzyme, the buffer that balances optimum rates with complementary stabilization of the enzyme can only be determined by a program of empirical experiments. Tris buffers have been traditionally employed for restriction and modification enzymes, although there is no *a priori* reason why this should be the case. Indeed, phosphate buffers are generally easier to titrate to a given pH and are excellent buffers at pH 7.4, the pH optimum of many if not all of the DNA Mtases so far isolated. For convenience, however, the three types of restriction enzyme buffers described by Maniatis et al. *(9)* are suitable for most purposes. The buffering agent is usually employed at a concentration of between 50 and 100 m*M* and is often supplemented with dithiothreitol (1 m*M*) or β-mercaptoethanol (5 m*M*) and bovine serum albumin (100 μg/mL). The latter reagents are not always necessary and are occasionally inhibitory. If a new Mtase is being isolated, it is best to determine the enzyme activity in the presence and absence of these reagents.

The reaction is usually carried out at 37°C for between 1 and 2 h and quenched by the addition of an equal vol of phenol. The reaction vol should be around 20 μL, although preparative incubations of 200 μL may be required if the DNA is to be analyzed further.

2.6. Experimental Procedures

2.6.1. Materials

1. 50 m*M* Tris-HCl, pH 8.0, 100 m*M* NaCl, 1 m*M* EDTA.
2. Bio-Rad A 0.5*M* gel.
3. 0.3 *M* Potassium acetate (or ammonium acetate).
4. Ethanol.
5. [^{3}H-methyl]-*S*-adenosylmethionine (SAM).
6. 100 m*M* NaCl.

2.6.2. Assay

The quantitative estimation of DNA Mtase activity can be acheived in several ways. However, most methods are based on one of two principles. The first and perhaps the simplest involves the enzyme

catalyzed transfer of tritiated methyl groups to DNA from [^{3}H-methyl] SAM followed by separation of the reaction components either by selective precipitation or by chromatography. The tritiated DNA produced in the incubation is subsequently estimated by liquid scintillation counting.

In the tritium transfer method, the DNA Mtase enzyme is incubated with the appropriate substrates, and the reaction mixture quenched with phenol. The tritiated DNA can then be isolated by chromatography through a Pasteur pipet-sized column of Bio-Rad A 0.5*M* gel-filtration medium (or similar). Typically, the reaction mixture is loaded onto the column and eluted with a buffer containing 50 m*M* Tris-HCl, pH 8, 100 m*M* NaCl, and 1 m*M* EDTA. The DNA elutes well in advance of the unreacted SAM. Alternatively, the DNA can be precipitated with 3 vol of cold ethanol in the presence of 0.3*M* potassium (or ammonium) acetate. If precipitation is used, care should be taken not to "carry over" any unreacted SAM. Moreover, stringent controls in the absence of individual reaction components should be included in order to verify the assay results. This is particularly important in assays of crude extracts, which probably contain protein and RNA Mtases.

A different strategy for estimating DNA Mtase activity involves the use of DNA fragments containing specific enzyme recognition sites. This approach is particularly convenient in assaying the DNA Mtases of bacteria. In one method, a DNA fragment containing the methylation site of, for example, *Eco*RI Mtase is incubated with the enzyme in the presence of unlabeled SAM, and the modified DNA so produced is extracted with phenol and precipitated with ethanol. The DNA is subsequently incubated with *Eco*RI restriction enzyme. The extent of methylation can be determined by the extent of resistance to endonuclease digestion. This approach can be made highly sensitive by incorporating radioactively labeled DNA into the reaction mixture. The DNA fragments can either be resolved by gel electrophoresis or by HPLC.

3. Applications of DNA Methyltransferases

DNA Mtases have been incorporated into a number of molecular biology protocols and generally bring a greater degree of flexibility to molecular cloning strategies. Some of the uses of DNA MTases are outlined in the following sections.

3.1. Construction of Gene Libraries

A partial restriction digest of genomic DNA from any organism can form the starting material for the construction of a genomic library. In order to take advantage of the versatile bacteriophage λ vectors, which contain single unique restriction sites (e.g., λgt11, λgt10, and so on), DNA must be either digested with *Eco*R1 (compatible with the unique cloning site of λgt10 and λgt11) or with any enzyme followed by the addition of *Eco*RI linkers to the ends of the DNA fragments. The "activation" of the linkers requires an incubation of the fragments with an excess of *Eco*RI. As a consequence, DNA that has been digested with, say, *Sau*IIIA and therefore probably contains one or more internal *Eco*RI sites will be further reduced in size. In order to prevent this secondary fragmentation of the DNA, the genomic digest is incubated with *Eco*RI Mtase in the presence of SAM prior to the addition of the linkers. In this way, only the tandem linker sites are cleaved, and the DNA fragment remains intact. This procedure can of course be applied to any restriction enzyme, providing a compatible Mtase is available. Moreover, the principle of protection of specific restriction sites can be extended to other related DNA manipulations as required. This strategy is particularly suitable for the construction of cDNA libraries (*see* Chapters 17 and 19 in vol. 4 of this series).

3.2. Modifying Restriction Sites

Methylation of a DNA sequence that is the target for a restriction enzyme with partially degenerate specificity by a DNA Mtase of overlapping specificity can reduce the number of such sites that are cleaved by the cognate restriction enzyme. For example, as shown in Fig. 2, the enzyme *Hin*cII recognizes and cleaves the sequence GTPyPuAC. The two sequences GTCAAC and GTCGAC, which are found in the plasmid pBR322, are both sites for *Hin*cII. However, the sequence G*TCGA*C, but not the sequence GTCAAC, is a target for the TCGA-specific *Taq*I Mtase. Therefore, methylation of pBR322 by *Taq*I Mtases produces a plasmid with effectively a single *Hin*cII site. This approach has been promoted by New England Biolabs *(10)*, who offer an extensive range of DNA Mtases for such purposes.

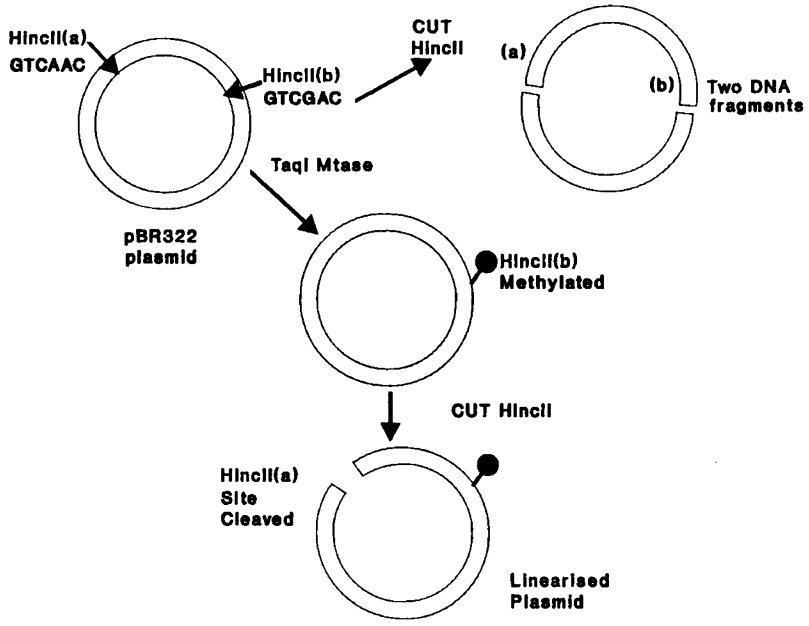


Fig. 2. Selective inactivation of partially degenerate restriction sites.

In a second approach, the use of a DNA Mtase to protect a restriction site involves the methylation of bases that are part of overlapping Mtase sites. Thus, in the example shown in Fig. 3, *Bam*HI and *Msp*I sites overlap by two bases in the sequence GGAT*CCGG*. Methylation of the 5' dC of the *Msp*I site renders the *Bam*HI site refractory to *Bam*HI endonuclease. Those *Bam*HI sites that are not followed by a GpG dinucleotide are not methylated by *Msp*I Mtase and are therefore not protected.

3.3. In Vitro Methylation of Eukaryotic Genes

Methylation of dC bases in mammalian DNA has been implicated in the control of gene transcription. The availability of methylated and unmethylated DNA for the experimental investigation of this phenomenon is highly desirable. In order to reproduce eukaryotic methylation patterns in vitro, the DNA must either be synthesized chemically from a 5-methyldCTP precursor or prepared enzymically using a preparation of a CpG-specific Mtase. Alternative but somewhat less satisfactory Mtases include *Hpa*II (CCGG-specific) and *Hha*I (CGCG), which

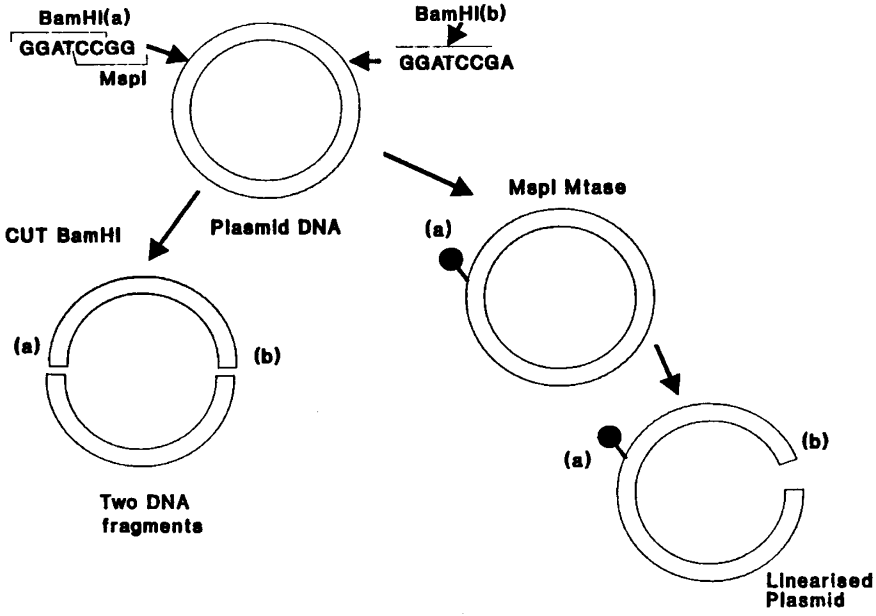


Fig. 3. Inactivation of restriction sites using DNA Mtase of overlapping specificity.

clearly would not methylate the full complement of CpG dinucleotides in a given sequence of DNA. Unfortunately, no commercial preparations of a CpG Mtase are currently available. However, a number of purification schedules have been published for such enzymes from a range of mammalian sources. A preparation *(11)* obtained from mouse erythroleukemia cells has been well characterized and is probably the enzyme of choice for this type of work.

As before, the DNA fragment or oligonucleotide duplex to be modified is incubated with the Mtase preparation as described earlier. In order to determine the extent and sequence-specific nature of the methylation, the DNA can be subjected to strategic restriction analysis (if the DNA sequence is known). A more thorough analysis involves the comparative determination of the sequences of modified and unmodified DNA by the Maxam Gilbert method *(12)*. Methylated cytosines do not react with hydrazine and, therefore, do not give bands in the C-track of the sequencing gel. Following successful methylation, the DNA can be used in comparative binding studies with transcription factors or can be used to transfect cells, and so forth.

4. Notes

1. Methylated DNA does not transform as efficiently as unmodified DNA into a number of common strains of *E. coli*. In general, strain C600 is suitable for Me-CG-modified DNA, RR1 for Me-CC modification, and K803 can tolerate both. This problem arises owing to the action of the *mcrA mcrB* genes of *E. coli*, which lead to restriction of cytosine-modified DNA. A detailed account of this problem has been given by Raleigh *(2)*.
2. 5-*d*-Azacytidine, when incorporated into DNA, is a potent inhibitor of the Mtase reaction. It is thought that the analog traps the covalent enzyme-SH:pyridine intermediate, since methyltransfer to the 5-N position is mechanistically improbable *(6)*.The peptide methinin is another potent inhibitor of mammalian DNA Mtase activity. However, its mode of action is obscure.

Acknowledgments

I would like to thank all of the members of my laboratory for their helpful comments on the manuscript.

Abbreviations

Mtase, methyltransferase; SAM, *S*-adenosylmethionine; dA, deoxyadenine; dC, deoxycytidine; N, any deoxynucleotide; p, phosphate; Me, methyl; SAH, *S*-adenosylhomocysteine; *dam*$^-$, DNA adenine methylation; ATP, adenosine triphosphate; HPLC, High-performance liquid chromatography; Py, pyrimidine; Pu, purine.

References

1. Razin, A. and Riggs, A.D. (1980) DNA methylation and gene function. *Science* **210,** 604–610.
2. Raleigh, E. A. (1987) Restriction and modification *in vivo* by *Escherichia coli* K12. *Methods Enzymol.* **152,** 130–141.
3. Woodcock, D. M., Crowther, P. J., and Diver, W. P. (1987) The majority of deoxycytidines in human DNA are not in the CpG dinucleotide. *Biochem. Biophys. Res. Comm.* **145,** 888–894.
4. Gruenbaum,Y., Naveh-many, T., Cedar, H., and Razin, A. (1982) Sequence specificity of methylations in higher plants. *Nature* **292,** 860–862.
5. Cedar, H. (1988) DNA methylation and gene activity. *Cell* **53,** 3,4.
6. Santi, D. V., Garrett, C. E., and Barr, P. J. (1983) On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs. *Cell* **33,** 9,10.
7. Bestor, T. H., Laudano, A., Mattaliano, R., and Ingram, V. (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. *J. Mol. Biol.* **202,** 971–984.

8. Lauster, R.,Trautner, T. A., and Noyer-Weidner, M. (1989) Cytosine-specific Type II DNA methyltransferases. *J. Mol. Biol.* **206,** 305–312.
9. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
10. New England Biolabs Catalog (1988/1989) New England Biolabs, Beverly, MA.
11. Bestor, T. H. and Ingram, V. (1983) Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity and mode of interaction with DNA. *Proc. Natl. Acad. Sci. USA* **80,** 5559–5563.
12. Maxam, A. M. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65,** 499–560.

CHAPTER 10

DNA and RNA Ligases (EC 6.5.1.1, EC 6.5.1.2, and EC 6.5.1.3)

Martin J. Maunders

1. Introduction

Ligases are a class of enzymes that catalyze the joining of nucleic acid molecules by the formation of phosphodiester bonds between their termini *(1)*. The nucleic acid substrate to be linked may be DNA or RNA depending on the type of ligase involved.

The enzymes are widespread and have been identified in a range of organisms, including bacteria *(2,3)*, phage-infected bacteria *(4)*, yeasts *(5)*, amphibians *(6)*, and mammals *(5)*, including *Homo sapiens (7)*. Many ligase-deficient and ligase-overproducing mutants have been isolated, and levels of enzyme in the *E. coli* bacterium can vary 1000-fold (from 0.01 to 10 × normal) without deleterious effect.

2. DNA Ligases (EC 6.5.1.1–2)

The role of DNA ligases in vivo is believed to include the joining of short DNA fragments formed during DNA replication *(8)*, and so enabling DNA synthesis to progress in an overall 3'–5' direction on the antiparallel strand of the double helix, while continual 5'–3' synthesis proceeds on the other strand.

Other possible functions of DNA ligases include roles during genetic recombination and in the repair of UV-damaged DNA *(9)*, though ligase$^-$ bacterial mutants can in general perform these functions sufficiently well. In yeast, however, there is evidence to support the idea of a role for DNA ligases in UV damage repair *(10)*.

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

In vitro a variety of activities have been detected. These include the joining of hydrogen-bonded cohesive DNA termini (with either 5' or 3' projecting single-stranded regions), the joining of blunt-ended double-stranded DNA molecules, the resealing of single-stranded "nicks" in DNA duplexes, and the interconversion of ATP and AMP by the exchange of pyrophosphate groups.

The two most intensively studied and widely used DNA ligases are that from the *E. coli* bacterium (EC 6.5.1.2) *(11,12)* and that occurring in *E. coli* that has been infected with bacteriophage T4 (EC 6.5.1.1) *(13)*. This section of the chapter will concentrate largely on these two enzymes.

E. coli DNA ligase has been purified from several strains that overproduce this enzyme *(11,12,14)*. The T4 enzyme is prepared from *E. coli* that has been infected with either wild-type or replication-deficient T4 phage *(15)*. Both the *E. coli* and T4 DNA ligase genes have been cloned into recombinant λ vectors.

2.1. The Enzyme

E. coli DNA ligase is a monomeric enzyme of asymmetric shape ($S_{20,w}$ = 3.9S) with a mol mass of 77,000 Da as determined by sedimentation equilibrium *(14)*, or 74,000 ± 3000 Da by polyacrylamide gel electrophoresis (PAGE) *(16)*. T4 DNA ligase is also an elongated monomer ($S_{20,w}$ = 3.5) of mol mass 68,000 ± 6800 Da determined by gel filtration *(18)* or 63,000 ± 3200 Da by PAGE *(17)*. T4 DNA ligase exhibits five species on isoelectric focusing. The adenylated form of the enzyme displays a major band with a *pI* of 6.0, and the non-adenylated form has a value of 6.2 for this band *(15)*.

2.2. Enzyme Reaction

2.2.1. Nucleic Acid Substrate

DNA ligase requires as a substrate two DNA termini, the 5' terminus, carrying a phosphate group, and the 3' terminus, a hydroxyl group. These termini must reside on a double-stranded molecule (DNA:DNA or DNA:RNA). Both strands of the duplex may terminate, in the form of a staggered end or a blunt end, and the ligase then requires a second similar double-stranded terminus and proceeds to join the two in an intermolecular reaction. Alternatively, the two termini may be provided by a nick in just one strand of a duplex, which the enzyme will then seal.

The enzyme will not ligate a 3' phosphate group to a 5' hydroxyl, a 3' hydroxyl to a 5' hydroxyl, a 3' dideoxy nucleoside to a 5' phosphate, or a 3' hydroxyl to a 5' triphosphate terminus. If only one of the 5' termini of a double-stranded break is phosphorylated, DNA ligase can only rejoin that one strand and the resultant product is a nicked molecule, since the second strand cannot be ligated. However, this double-stranded DNA molecule is intact and can then be used in transformation reactions enabling the ligation to be completed in vivo.

Only T4 DNA ligase will ligate blunt-ended DNA molecules, the reaction proceeding by way of nicked intermediates *(14)*. Unlike joining cohesive ends, the rate is not linearly dependent on enzyme concentration. The enzyme can also ligate at a mispaired 3' base *(18,19)* and will ligate DNA/RNA hybrids, but with much reduced activity.

The reverse reaction is also catalyzed *(20)*, the enzyme behaving as an AMP-dependent endonuclease, yielding nicked DNA. The two activities of DNA ligase result in the slow (but eventually complete) relaxation of supercoiled DNA.

The only action on RNA performed by *E. coli* DNA ligase is the ligation of the 3' hydroxyl of an RNA strand to a phosphorylated 5' DNA terminus *(21)*. T4 DNA ligase has some activity in joining RNA molecules annealed to DNA and even RNA:RNA *(22)*.

E. coli DNA ligase has a K_m for the 5' phosphate group of 2.5–5.6 $\times 10^{-8} M$ *(23)*. The value for T4 DNA ligase is $6 \times 10^{-7} M$ for cohesive ends *(24)*, $5 \times 10^{-5} M$ for blunt ends, and $1.5 \times 10^{-9} M$ for repairing nicks *(17)*.

2.2.2. Cofactors

DNA ligase requires a nucleotide cofactor for reaction, forming a covalent AMP-enzyme intermediate. *E. coli* DNA ligase utilizes NAD (as do ligases from *B. subtilis, T. thermophilus,* and *S. typhimurium*). The enzyme has a high specificity for NAD with a K_m of 3×10^{-8} to $7 \times 10^{-6} M$ *(25)*.

T4 DNA ligase (and T7 and mammalian ligases) requires ATP. It will also utilize dATP (at 0.5% of the rate), which acts as a competitive inhibitor with ATP *(17)*. The value of $K_{m\ \mathrm{ATP}}$ for the joining reaction is $1.4–10.0 \times 10^{-5} M$ and the $K_{i\ \mathrm{dATP}}$ is $3.5 \times 10^{-5} M$. For the pyrophosphate exchange reaction used in the enzyme assay *(17)*, $K_{m\ \mathrm{ATP}}$ is $2 \times 10^{-6} M$ and $K_{i\ \mathrm{dATP}}$ is $1 \times 10^{-5} M$ *(1)*. T7 DNA ligase can in fact utilize either ATP or dATP *(26)*.

2.2.3. Temperature

Rates of reaction are very temperature-sensitive, the optimum depending on the lengths of the oligonucleotides being joined, and being greater than the T_m of the substrate. *E. coli* DNA ligase has an optimum temperature of 10–15°C for ligating cohesive ends *(27)*.

Another bacterial ligase from the thermophilic bacterium *Thermus thermophilus (28,29)* is becoming widely used as further applications of the polymerase chain reaction are developed *(30)*. The enzyme is similar to *E. coli* DNA ligase in size, pH optimum, cofactor and cation requirement, and response to activators. However, its temperature optimum is much higher, being 24–37°C for ligating cohesive termini, and 65–72°C for closing nicks (within a range of 15–85°C), which may result in more widespread applications of the technique of ligation.

T4 DNA ligase has an optimum temperature for ligating cohesive ends of 4°C *(31)* and for sealing nicks of 37°C *(32)*. The optimal temperature for blunt-end ligation is 25°C for 16 mers or longer, smaller molecules requiring lower temperatures consistent with their decreased melting temperatures.

2.2.4. pH

Optimal pH for ligation varies to some extent with the reaction catalyzed and the buffer employed. *E. coli* DNA ligase has a pH optimum for joining DNA strands of 7.5–8.0 in Tris-HCl buffers and of 8.0 in sodium phosphate *(33)*. For the phosphate exchange reaction, the optimal pH is 6.5 in potassium phosphate, the rate of reaction falling to 50% at pH 5.6 and 7.5. This reaction rate falls to 20% in Tris-HCl buffer at pH 8.0.

T4 DNA ligase has a similar pH optimum range for joining DNA, the reaction rates falling to 40% at pH 6.9 and 65% at pH 8.3. The pH optimum for the exchange reaction is also similar to that for *E. coli* DNA ligase. For the sealing of nicks, the optimal pH range is 7.2–7.8 in Tris-HCl, with the rates dropping to 46% at pH 6.9 and to 65% at pH 8.5 *(17,33)*.

2.2.5. Cations

DNA ligases require a divalent cation, with Mg^{2+} being the most commonly utilized. The *E. coli* enzyme requires Mg^{2+} at an optimum concentration of 1–3 m*M*. Low concentrations of Mn^{2+} can substitute, and in the

range 0.2–1.0 m*M* cause a slight increase in activity, though higher concentrations are inhibitory *(33)*. Co^{2+} and Ni^{2+} are inactive, Zn^{2+} has slight activity, and Ca^{2+} has been reported to yield 60% activity in some cases and none in others *(27,33)*.

T4 DNA ligase has a magnesium optimum of 10 m*M*, whereas Mn^{2+} is only 25% as efficient *(17)*. However, in joining DNA:RNA hybrids, Mn^{2+} is twice as effective as Mg^{2+} *(34)*.

2.2.6. Activators and Inhibitors

Ammonium ions in low concentrations stimulate the *E. coli* enzyme, in joining DNA termini, and V_{max} can be increased by up to 20-fold *(23)*. K^+ and Rb^+ show similar stimulation, Cs^+ and Li^+ exert a small effect, and Na^+ is ineffective.

In contrast, the T4 DNA ligase is unaffected by low concentrations of ammonium ions, whereas higher levels (0.2*M*) of NH_4^+, Na^+, K^+, Cs^+, and Li^+ inhibit almost completely. Blunt-end ligation is inhibited by 25 m*M* phosphate and 50 m*M* Na^+ *(35)*.

Polyamines, such as spermine and spermidine, also inhibit, but this can be overcome by increasing the DNA concentration. Spermine is the most effective, inhibiting the joining reaction by 90%. Some workers, however, recommend the use of spermidine *(35,36)*.

2.2.7. Sulfhydryl Reagents

E. coli DNA ligase does not need the presence of sulfhydryl reagents *(37)*. T4, T7, and mammalian DNA ligases do require either β-mercaptoethanol or dithiothreitol (DTT) *(38)*. T4 DNA ligase performs much more efficiently if DTT is the reagent of choice.

2.2.8. Enzyme Assay

There are several types of assays carried out to define the activity of DNA ligase. Functional assays include the ligating of radiolabeled DNA to unlabeled DNA in solution *(33)* or bound to a matrix *(39)*, and the conversion of poly A:T to a form resistant to exonuclease III *(35)*. Direct measurements can be made of the conversion of radiolabeled phosphate monoesters to diesters resistant to alkaline phosphatase activity *(40)*. Also, the restoration of biological activity to DNA previously nicked by pancreatic DNAse or restriction endonucleases can be used as an assay *(3)*.

Three types of units are defined commonly. First, a Weiss unit *(17)* catalyzes the exchange of 1 nmol of ^{32}P from inorganic pyrophosphate to ATP in 20 min at 37°C, 0.015 Weiss units ligate 50 % of the *Hin*dIII fragments from 5 μg λ DNA in 30 min at 16°C. Second, the Modrich-Lehman unit, based on the exonuclease resistance assay *(37)*, is equivalent to 5 Weiss units. Finally, cohesive end units are functionally defined by various commercial suppliers, and are generally much smaller than Weiss units and difficult to relate quantitatively.

3. Experimental Procedures

3.1. Uses of DNA Ligases

The most widespread use of DNA ligase is in the construction of recombinant DNA molecules. This may be necessary for the cloning of cDNA or genomic fragments for construction of libraries, or for mapping, sequencing, or use as probes. The use of ligase in the inverse polymerase chain reaction facilitates the cloning of segments of genomic DNA some distance from known sequences *(41)*. DNA ligase may also be used in the assembly of genes from DNA fragments and synthetic oligomers.

Other uses of DNA ligase include the detection of nicked DNA by the release of AMP *(38)*, nearest neighbor analysis following kinasing *(42,43)*, use in mutagenesis *(44)*, and the making of affinity columns by attaching DNA to solid matrices. In most cases, the enzyme of choice is T4 DNA ligase. *E. coli* ligase is not widely used, because it is inefficient with blunt ends. However, *E. coli* DNA ligase is used for cDNA cloning by replacement synthesis *(45)* where a virtue is made of its inability to ligate RNA to DNA to form spurious products.

Examples of specific ligation protocols may be found in Gaastra and Hansen *(46)*. The following details apply to the use of T4 DNA ligase, except where specific reference is made to the *E. coli* enzyme.

3.2. Storage and Stability

T4 DNA ligase is usually supplied at a concentration of 1–5 U/μL. It can be diluted in reaction buffer immediately prior to ligation, but for longer term storage, a specific storage buffer should be used.

Storage buffers usually contain:

- 10–20 m*M* Tris-HCl (or potassium phosphate), pH 7.4–7.6
- 50–60 m*M* KCl

- 1–5 m*M* DTT
- 200 µg/mL BSA
- 50% Glycerol

High concentrations of enzyme are very stable at –20°C. If diluted to 500 U/mL, the enzyme is 90% stable at this temperature for 6 mo. Lower concentrations are increasingly less stable. For example, a 10 U/mL dilution will lose 40% of its activity in 3 mo. Storage at temperatures lower than –20°C may also be detrimental.

3.3. Reaction Conditions—Cohesive Termini

Reaction conditions can vary with the purity of each batch of enzyme, the purity of the DNA substrate, and the presence of buffer components from any previous or subsequent in vitro reaction step. DNA ligase acts in virtually all restriction enzyme buffers and also in the same buffer used to kinase linkers prior to their ligation *(47)*.

A commonly used buffer for ligation reactions with T4 DNA ligase is given below with a range of acceptable parameters:

20 m*M* Tris-HCl pH 7.6	(20–66 m*M*, pH 7.5–7.8)
10 m*M* $MgCl_2$	(5–50 m*M*)
10 m*M* DTT	(5–20 m*M*)
1 m*M* ATP	(66 µ*M*– 1 m*M*)
0.1 Weiss units enzyme	(0.01–1.0 U/µg DNA)

The inclusion of BSA (Fraction V) at 50 µg/mL is optional.

DNA concentration (*see* Section 3.5.):

Reaction vol 20 µL	(10–100 µL)
Temperature 16°C	(4–25°C)
Time 4 h	(20 min–24 h)

If DNA substrates with cohesive ends are being ligated, the DNA should be heated to 70°C for 10 min to melt any hydrogen bonding between the termini of similar molecules before the various DNA substrates are mixed. After heating, the mixture should be allowed to cool slowly to allow intermolecular reannealing, prior to the addition of the reaction buffer and finally the DNA ligase.

Reaction times and temperatures have an inverse relationship. For example, similar results might be achieved by incubation at 25°C for 1 h, at 15°C for 4–6 h, and at 4°C for 16 h *(48)*.

After reaction, the enzyme can be inactivated at 65°C for 10 min or by addition of dimethyl dicarbonate to 0.1% and heating to 37°C for 10 min. The DNA can then be concentrated by ethanol precipitation, which also removes magnesium ions that may reduce the efficiency of a subsequent transformation step. If ethanol precipitation is not performed, the magnesium concentration should be reduced by fivefold (or greater) dilution of the ligation mixture.

Ligation can be achieved directly on DNA in low melting agarose excised from a gel, though more enzyme is required and the efficiency is 10-fold lower. The gel should be run in TAE buffer (40 m*M* Tris-acetate, 1 m*M* EDTA, pH 7.5–8.0) rather than the more usual Tris-borate-EDTA (TBE). The presence of ethidium bromide in the gel at 0.5 μg/mL causes no ill effects *(47)*.

E. coli DNA ligase will function in a similar buffer to T4 DNA ligase, although the Tris-HCl concentration is usually reduced (20–30 m*M*) and ATP replaced by 25–50 μ*M* NAD. DTT (20 m*M*) is often included in the buffer, although this may be superfluous *(37)*. In some cases, 10 m*M* ammonium sulfate and 10 m*M* KCl are also added.

3.4. Reaction Conditions—Blunt Termini

The ligation of blunt- or flush-ended DNAs is only possible using T4 DNA ligase and not *E. coli* DNA ligase *(19,49)*. The reaction is slower than the ligation of cohesive ends, but is extremely useful in that it can be used to join virtually any two pieces of DNA, be they naturally occurring or synthetic.

In practice, this ability can be made universal by the infilling or "polishing" of cohesive or incompatible termini. This technique can be adapted to infill partially "ragged" termini leaving short cohesive termini that are not self-ligatable, but that will accept inserted DNA *(50,51)*.

Blunt-end ligation can be performed in the standard buffer with the ATP concentration adjusted to 0.5 m*M* *(52)*. More ligase should be used, of the order of 50 U/mL and 1–2 U/μg DNA. Optional additions for blunt-end ligation include 100 μg/mL BSA, 1 m*M* hexamine cobalt chloride, and 0.1–1.0 m*M* spermidine. The latter two reagents can increase the efficiency of linker ligation by fivefold.

The DNA concentration should also be raised. In the case of small oligomers, such as linkers, it is easy to achieve a high concentration of DNA termini of the order of 4–20 μ*M* or up to 100-fold in excess

of the DNA fragment to be linked. The effective concentration of the DNA can be increased by the use of condensing agents or volume excluders, such as polyethylene glycol (PEG), Ficoll, and hexamine cobalt chloride *(53)*, which accelerate the rate of ligation by up to 1000-fold, and permit ligation at lower absolute DNA and enzyme concentrations. Condensing agents alter the distribution of the products and suppress intramolecular reaction (recircularization), most products being linear multimers.

PEG8000 gives maximum stimulation of ligation when present at a concentration of 15%, this effect being achieved at a concentration of 0.5 m*M* ATP and 5 m*M* $MgCl_2$. Even slight increases in the ATP or decreases in the magnesium ion concentrations greatly diminish the effect of PEG *(54)*. PEG enhances the ligation of blunt-ended oligomers as short as 8 bp and also simulates ligation of cohesive ends by 10–100-fold.

Hexamine cobalt chloride stimulation depends on the concentration used, 1.0–1.5 μ*M* having maximal effect. Blunt-end ligation is stimulated 50-fold, but cohesive-end ligation only fivefold. Ligation is possible in the presence of monovalent cations (e.g., 30 m*M* KCl), but the products are then largely recircularized. This reagent does not significantly increase the ligation of short oligomers.

T4 RNA ligase will stimulate T4 DNA ligase activity *(24)*, but is not widely used since PEG is more efficient and inexpensive. RNA ligase cannot perform blunt-end DNA ligations itself, nor will it activate the *E. coli* enzyme.

3.5. Ligation Products

The products of ligation depend largely on the nature of the DNA substrates employed. Removal of the 5'-phosphate groups from one of the substrate molecules (e.g., a vector) by alkaline phosphatase treatment will prevent self-ligation and favor the formation of recombinant products. The use of DNA molecules carrying different cohesive termini at each end (asymmetric cloning) also prevents self-ligation and offers control over the relative orientation of the DNA fragments in the resultant products.

The products obtained also depend on the absolute concentrations and ratios of the DNA substrates present in the mixture. Low DNA concentrations favor intramolecular reaction (recircularization), and

higher concentrations favor intermolecular reaction (oligomerization and formation of recombinant molecules). Condensing agents exert a similar effect. For example, 10% polyethylene glycol enhances intermolecular ligation *(54,55)*. Longer DNA molecules are more prone to recircularization than shorter ones.

Optimal ratios usually have to be determined empirically, though they are fairly broad, and a 30% variation in starting concentrations makes little difference. Some theoretical guidelines have been suggested by Revie et al. *(56)* for vectors of 2.5–7.5 kb and inserts of 0.2–10 kb, which can be briefly summarized as follows;

1. Always use forced directional (asymmetric) cloning and phosphatased vectors where possible.
2. Keep the vector DNA concentration below 1 μg/mL, unless "scavenging" for minute quantities of insert DNA.
3. Under normal conditions, use 3:1 molar ratio of insert:vector, but not more than 5 μg/mL insert DNA.
4. Increase the insert concentration as the vector size decreases.

In practice, a suitable preliminary experiment would use a concentration of 20–60 μg/mL vector (2–3 kb in size) with an equimolar or slightly greater concentration of insert DNA *(47)*.

The above rules apply only to plasmid ligations. For the formation of λ or cosmid concatamers, insert and vector should be ligated at a 1:1 molar ratio in as high a concentration as possible.

In all cases, the ability to select for the recombinant DNA molecule of choice by, for example, transformation and expression of an antibiotic resistance or by enzymic formation of a colored product from a chromogenic substrate would offer much more flexibility in the choice of DNA concentrations and reaction conditions of the ligation.

3.6. Reaction Protocol

A specific example of a ligation reaction protocol for DNA fragments with cohesive termini is as follows:

3.6.1. Materials Required

1. 10 m*M* Tris-HCl / 1 m*M* EDTA, pH 7.6.
2. Phenol.
3. Chloroform.
4. Ethanol.

5. 10X Ligation buffer:
   - 200 m*M* Tris-HCl, pH 7.6
   - 50 m*M* $MgCl_2$
   - 50 m*M* DTT
   - 500 μg/mL BSA.
6. 10 m*M* ATP

3.6.2. Protocol

1. Prepare the fragments to be ligated by digestion of the source DNA species (e.g., vector and foreign genomic DNA) with restriction endonucleases. Purify by phenol/chloroform extraction and ethanol precipitation. Resuspend each DNA in 10 m*M* Tris/1 m*M* EDTA, pH 7.6.
2. Mix aliquots containing 0.1 μg of each DNA species. Add 15 μL water. Warm to 70°C for 10 min, to melt any annealing of cohesive ends between molecules of the same species. Allow to cool slowly to room temperature, and then place on ice.
3. Add 2 μL of 10X ligation buffer, plus 1 μL 10 m*M* ATP.
4. Add 0.1 Weiss unit T4 DNA ligase.
5. Incubate at 16°C for 4 h.
6. Terminate the reaction by heating to 70°C for 10 min. Aliquots of the ligation mixture can then be used for bacterial transformation either directly or after ethanol precipitation.

4. T4 RNA Ligase (EC 6.5.1.3)

Like T4 DNA ligase, this enzyme is found in *E. coli* after infection with T-even phage *(57)*. It has also been cloned and is now available from several overproducing strains of *E. coli*. The enzyme forms a phosphodiester bond between a phosphorylated 5' terminus and a 3' hydroxyl similarly to the DNA ligases *(58–60)*, except that the substrates in this case are in general RNA moieties.

The in vivo role of the enzyme is unclear, although it is probably involved in host RNA modification and transcript splicing *(61)*. The protein does have a morphological role in the noncovalent attachment of tail fibers to the base plate of the phage *(62)*.

4.1. The Enzyme

The mol mass of the enzyme has been determined as 48,000 Da by sedimentation equilibrium, 47,000 Da by gel filtration, and 41,000–45,000 Da by SDS-PAGE *(63,64)*. It is monomeric in solution *(63)* and has an isoelectric point of 6.1 *(65)*.

4.2. Enzymic Reaction

4.2.1. Nucleic Acid Substrate

The normal substrate of RNA ligase is single-stranded RNA, although it will act on a variety of single- or double-stranded RNA or DNA molecules *(60–68)*. RNA ligase will act on very small pieces of ribonucleic acid, with 40 mers being the probable upper size limit *(69)*. The minimum size of the 5' moiety is a ribonucleoside 3'5'-*bis* phosphate (i.e., P-base-P), a 5'-monophosphate being ineffective. The nature of the base has some bearing on the reaction rate, pyrimidines reacting 2–10 times faster than purines, and modified purines slower yet. With these reservations, a wide variety of bases will be accepted, including deoxynucleosides, *O*-methylated-, halogenated-, dihydro-, thio-, and deaza-derivatives, and deaminated purines *(58)*. The minimum size of the 3' moiety is a trinucleoside *bis*-phosphate, base-P-base-P-base-OH.

4.2.2. Cofactors

The enzyme requires ATP (although dATP will substitute) and forms an enzyme–AMP complex during reaction. Low levels of ATP are preferred, often provided by an ATP-generating system *(58)*.

4.2.3. Cations

Magnesium ions are required for reaction, although Mn^{2+} will substitute and often improve the rate of reaction; 1*M* $(NH_4)_2SO_4$ completely inhibits activity.

4.2.4. Enzyme Assay

The enzyme is usually assayed by measuring the conversion of 5'-^{32}P-labeled poly-A to a circular form insensitive to phosphomonoesterase at 37°C. One unit converts 1 nmol of 5'-^{32}P-labeled poly-A at a concentration of 1 μ*M (24)* or 10 μ*M (59)* termini to a phosphatase-resistant form in 30 min at 37°C.

Another definition of the unit of activity is that it ligates 1 pmol of ^{32}P-labeled mononucleoside biphosphate to a larger TCA-precipitable RNA species in 30 min at 4°C. This unit is approximately equivalent to one-sixth of that defined by the phosphatase resistance assay.

A third unit can be defined on the basis of the pyrophosphate exchange reaction. One unit of enzyme catalyzes the exchange of 1 nmol of ^{32}P-labeled inorganic pyrophosphate into ATP in 30 min.

5. Experimental Procedures

5.1. Uses of RNA Ligase

The first obvious use of RNA ligase is in synthesizing oligomers. Modification of preexisting RNA molecules is perhaps a more interesting application. This may involve merely the end-labeling of an RNA moiety *(69,70)*, or alternatively, bases can be removed by periodate oxidation and β-elimination, and then modified bases can be added by RNA ligase to examine the restored function *(71)*. Internal modification can also be achieved after nicking or treatment with RNAse H *(72)*. The enzyme can also be used to join DNAs *(69)*, although this requires a large concentration of enzyme and a longer time.

5.2. Storage and Stability

T4 RNA ligase is available at specific activities of 1000–4000 U/mg protein corresponding to 1–4 U/μL. The enzyme should be diluted to the required concentration in storage buffer, which is similar to that used for T4 DNA ligase. In this case, the Tris-HCl concentration may be raised to 50 m*M*, and KCl may be omitted. The enzyme should be held at –20°C.

5.3. Reaction Conditions

As with all manipulations involving RNA, great care should be taken to eliminate the presence of ribonucleases. In practice, it is usually sufficient to ensure that all solutions and vessels have been autoclaved prior to use.

T4 RNA ligase reaction buffers commonly consists of:

- 50 m*M* HEPES/NaOH pH 7.5 (or 50 m*M* Tris-HCl, pH 8.0)
- 10 m*M* $MgCl_2$ (10–18 m*M*)
- 20 m*M* DTT (3–33 m*M*)
- 1 m*M* ATP (0.1–1.0 m*M*)
- 10% (v/v) DMSO
- 100 μg/mL BSA (10–100 μg/mL)

Common alternative concentrations are shown in parentheses.

The substrate termini should be present in a concentration of 1–100 μ*M* (corresponding to 0.5–20 μg RNA), in a reaction vol of 20–50/μL. The amount of RNA ligase required will vary with each application within the range 0.1–16.0 U.

The temperature of reaction must be optimized for the specific application, conditions ranging from 37°C for 30 min to 4°C overnight. After this time, the reaction should be terminated by boiling for 2 min.

6. Summary

DNA and RNA ligases are widely occurring, readily available, and relatively inexpensive enzymes that are an essential part of the molecular biologist's armory, both for the analysis and the manipulation of nucleic acids. The major drawback of these enzymes lies in their relative lack of substrate specificity as far as nucleotide sequence is involved. This factor necessitates careful experimental design in order to obviate the production of spurious or misinterpreted results.

References

1. Engler, M. J. and Richardson C. C. (1982) DNA ligases, in *The Enzymes XV* Part B (Boyer, P. D., ed.), Academic, New York, pp. 3–29.
2. Gellert, M. (1967) Formation of covalent circles of lambda DNA by *E. coli* extracts. *Proc. Natl. Acad. Sci. USA* **57,** 148–155.
3. Laipis, P. J., Olivera, B. M., and Ganesan, A.T. (1969) Enzymatic cleavage and repair of transforming DNA. *Proc. Natl. Acad. Sci. USA* **62,** 289–296.
4. Weiss, B. and Richardson, C. C. (1967) Enzymatic breakage and joining of deoxyribonucleic acid. I. Repair of single strand breaks in DNA by an enzyme system from *Escherichia coli* infected with T4 bacteriophage. *Proc. Natl. Acad. Sci. USA* **57,** 1021–1028.
5. Soderhall, S. and Lindhahl, T. (1976) DNA ligases of eukaryotes. *FEBS Lett.* **67,** 1–7.
6. Pfeiffer, P. and Vielmetter, W. (1988) Joining of nonhomologous DNA double strand breaks in vitro. *Nucleic Acids Res.* **16,** 907–924.
7. North, P., Ganesh, A., and Thacker, J. (1990) The joining of double strand breaks in DNA by human cell extracts. *Nucleic Acids Res.* **18,** 6205–6210.
8. Okasaki, K., Okasaki, T., Sakabe, K., Sugimoto, K., and Sugino, A. (1968) Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proc. Natl. Acad. Sci. USA* **59,** 598–605.
9. Boyce, R. P. and Howard-Flanders, P. (1964) Release of ultraviolet light induced thymine dimers from DNA of *E. coli* K-12. *Proc. Natl. Acad. Sci. USA* **51,** 293–300.
10. Johnston, L. H. and Nasmyth, K. A. (1978) *Saccharomyces cerevisiae* cell cycle mutant *cdc*9 is defective in DNA ligase. *Nature (London)* **274,** 891–893.
11. Panasenko, S. M., Cameron, J. R., Davis, R. W., and Lehman, I. R. (1977) Five hundred-fold overproduction of DNA ligase after induction of a hybrid lambda lysogen constructed in vitro. *Science* **196,** 188–189.
12. Panasenko, S. M., Alazard, R. J., and Lehman, I. R. (1978) A simple three step procedure for the large scale purification of DNA ligase from a hybrid lambda lysogen constructed in vitro. *J. Biol. Chem.* **253,** 4590–4592.

13. Murray, N. E., Bruce, S. A., and Murray, K. (1979) Molecular cloning of the DNA ligase gene from bacteriophage T4. II. Amplification and preparation of the gene product. *J. Mol. Biol.* **132,** 493–505.
14. Modrich, P., Anraku, Y., and Lehman, I. R. (1973) Deoxyribonucleic acid ligase. Isolation and physical characterization of the homogeneous enzyme from *Escherichia coli*. *J. Biol. Chem.* **248,** 7495–7501.
15. Knopf, K. W. (1977) Simple isolation method and assay for T4 DNA ligase and characterization of the purified enzyme. *Eur. J. Biochem.* **73,** 33–38.
16. Higgins, N. P. and Cozzarelli, N. R. (1980) DNA joining enzymes: a review. *Methods Enzymol.* **68,** 50–71.
17. Weiss, B., Jacquemin-Sablon, A., Live, T. R., Fareed, G. C., and Richardson, C. C. (1968) Enzymatic breakage and joining of deoxyribonucleic acid. VI. Further purification and properties of polynucleotide liase from *Escherichia coli* infected with bacteriophage T4. *J. Biol. Chem.* **243,** 4543–4555.
18. Sgaramella, V. and Khorana, H. G. (1972) Studies on polynucleotides. CXII. Total synthesis of the structural gene for an alanine transfer RNA from yeast. Enzymic joining of chemically synthesized polynucleotides to form the DNA duplex representing nucleotide sequence 1 to 20. *J. Mol. Biol.* **72,** 427–444.
19. Shortle, D., Grisafi, P., Benkovic, S. J., and Botstein, D. (1982) Gap misrepair mutagenesis: efficient site-directed induction of transition, transversion, and frameshift mutations in vitro. *Proc. Natl. Acad. Sci. USA* **79,** 1588–1592.
20. Modrich, P., Lehman, I. R., and Wang, J. C. (1972) Enzymatic joining of polynucleotides. XI. Reversal of *Escherichia coli* deoxyribonucleic acid ligase reaction. *J. Biol. Chem.* **247,** 6370–6372.
21. Nath, K. and Hurwitz, J. (1974) Covalent attachment of polyribonucleotides to polydeoxyribonucleotides catalyzed by deoxyribonucleic acid ligase. *J. Biol. Chem.* **249,** 3680–3688.
22. Kleppe, K., Van de Sande, J. H., and Khorana, H. G. (1970) Polynucleotide ligase catalysed joining of deoxyribo-oligonucleotides on ribopolynucleotide templates and of ribo-oligonucleotides on deoxyribopolynucleotide templates. *Proc. Natl. Acad. Sci. USA* **67,** 68–73.
23. Modrich, P. and Lehman, I. R. (1973) Deoxyribonucleic acid ligase. A steady state kinetic analysis of the reaction catalysed by the enzyme from *Escherichia coli*. *J. Biol. Chem.* **248,** 7502–7511.
24. Sugino, A., Goodman, H. M., Heynecker, H. L., Shine, J., Boyer, H. W., and Cozzarelli, N. R. (1977) Interaction of bacteriophage T4 RNA and DNA ligases in joining duplex DNA at base paired ends. *J. Biol. Chem.* **252,** 3987–3994.
25. Olivera, B. M. and Lehman, I. R. (1967) Diphosphopyridine nucleotide: A cofactor for the polynucleotide joining enzyme from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **57,** 1700–1704.
26. Hinkle, D. C. and Richardson, C. C. (1975) Bacteriophage T7 deoxyribonucleic acid replication in vitro. Purification and properties of the gene 4 protein of bacteriophage T7. *J. Biol. Chem.* **250,** 5523–5529
27. Dugaiczyk, A., Boyer, H. W., and Goodman, H. M. (1975) Ligation of EcoRI endonuclease generated DNA fragments into linear and circular structures. *J. Mol. Biol.* **96,** 171–184.

28. Takahashi, M., Yamaguchi, E., and Uchida, T. (1984) Thermophilic DNA ligase. Purification and properties of the enzyme from *Thermus thermophilus* HB8. *J. Biol Chem.* **259,** 10,041–10,047.
29. Barany, F. (1991) Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc. Natl. Acad. Sci. USA* **88,** 189–193.
30. Saiki, R. K., Scharf, S. J., Faloona, F., Mullis, K. B., Horn, G. T., Ehrlich, H. A., and Arnheim, N. (1985) Enzymatic amplification of β-globin genomic sequences and restriction-site analysis for diagnosis of sickle cell anaemia. *Science* **230,** 1350–1354.
31. Ferretti, I. and Sgaramella, V. (1981) Temperature dependence of the joining by DNA ligase of the termini produced by type II restriction endonuclease. *Nucleic Acids Res.* **9,** 85–93.
32. Pohl, F. M., Thomae, R., and Karst, A. (1982) Temperature dependence of the activity of DNA modifying enzymes: endonucleases and DNA ligase. *Eur. J. Biochem.* **123,** 141–152.
33. Zimmerman, S. B., Little, J. W., Oshinsky, C. K., and Gellert, M. (1967) Enzymatic joining of DNA strands: a novel reaction of diphosphopyridine nucleotide. *Proc. Natl. Acad. Sci. USA* **57,** 1841–1848.
34. Fareed, G. C., Wilt, E. M., and Richardson, C. C. (1971) Enzymatic breakage and joining of deoxyribonucleic acid. VIII. Hybrids of ribo- and deoxyribonucleotide homopolymers as substrates for polynucleotide ligase of bacteriophage T4. *J. Biol. Chem.* **246,** 925–932.
35. Raae, A. J., Kleppe, R. K., and Kleppe, K. (1975) Kinetics and effect of salts and polyamines on T4 polynucleotide ligase. *Eur. J. Biochem.* **60,** 437–443.
36. Modrich, P. and Lehman, I. R. (1970) Enzymatic joining of polynucleotides. IX. A simple and rapid assay of polynucleotide joining (ligase) activity by measurement of circle formation from linear deoxyadenylate-deoxythymidylate copolymer. *J. Biol. Chem.* **245,** 3626–3631.
37. Weiss B. and Richards C. C. (1967) The 5' terminal dinucleotides of the separated strands of T7 bacteriophage deoxyribonucleic acid. *J. Mol. Biol.* **23,** 405–415.
38. Weiss, B., Thompson, A., and Richardson, C. C. (1968) Enzymatic breakage and joining of deoxyribonucleic acid. VII. Properties of the enzyme-adenylate intermediate in the polynucleotide ligase reaction. *J. Biol. Chem.* **243,** 4556–4563.
39. Cozzarelli, N. R., Melechen, N. E., Jovin, T. M., and Kornberg, A. (1967) Polynucleotide cellulose as a substrate for a polynucleotide ligase induced by phage T4. *Biochem. Biophys. Res. Commun.* **28,** 578–586.
40. Weis, B., Live, T. R., and Richardson, C. C. (1968) Enzymatic breakage and joining of deoxyribonucleic acid. V. End group labeling and analysis of deoxyribonucleic acid containing single strand breaks. *J. Biol. Chem.* **243,** 4530–4542.
41. Triglia, T., Peterson, M. G., and Kemp, D. J. (1988) A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. *Nucleic Acids Res.* **16,** 8186.
42. Wu, R. and Kaiser, A. D. (1967) Mapping the 5'-terminal nucleotides of the DNA of bacteriophage lambda and related phages. *Proc. Natl. Acad. Sci. USA* **57,** 170–177.

43. Wu, R. and Kaiser, A. D. (1968) Structure and base sequence in the cohesive ends of bacteriophage lambda DNA. *J. Mol. Biol.* **35,** 523–537.
44. Wallace, R. B., Johnson, P. F., Tanaka, S., Schold, M., Itakura, K., and Abelson, J. (1980) Directed deletion of a yeast transfer RNA intervening sequence. *Science* **209,** 1396–1400.
45. Okayama, H. and Berg, P. (1982) High efficiency cloning of full length cDNA. *Mol. Cell Biol.* **2,** 161–170.
46. Gaastra, W. and Hansen, K. (1984) Ligation of DNA with T4 DNA ligase, in *Methods in Molecular Biology, vol 2. Nucleic Acids* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 225–230.
47. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Strategies for cloning in plasmid vectors. Ligation of termini created by restriction enzymes. Ligases, kinases and phosphatases, in *Molecular Cloning*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 1.53–1.73, 5.10–5.14, 5.61–5.67.
48. Promega Protocols and Applications Guide (1989) Cloning of DNA Inserts. Promega Corp., Madison, WI, p. 37.
49. Sgaramella, V. and Ehrlich, S. D. (1978) Use of the T4 polynucleotide ligase in the joining of flush-ended DNA segments generated by restriction endonucleases. *Eur. J. Biochem.* **86,** 531–537.
50. Hung, M.-C. and Wensink, P. C. (1984) Different restriction enzyme generated sticky ends can be joined in vitro. *Nucleic Acids Res.* **12,** 1863–1874.
51. Korch, C. (1987) Cross index for improving cloning selectivity by partially filling in 5' extensions of DNA produced by type II restriction endonucleases. *Nucleic Acids Res.* **15,** 3199–3220.
52. Ferretti, L. and Sgaramella, V. (1981) Specific and reversible inhibiion of the blunt end joining activity of the T4 DNA ligase. *Nucleic Acids Res.* **9,** 3695–3705.
53. Rusche, J. R. and Howard-Flanders, P. (1985) Hexamine cobalt chloride promotes intermoleclar ligation of blunt end DNA fragments by T4 DNA ligase. *Nucleic Aids Res.* **13,** 1997–2008.
54. Pheiffer, B. H. and Zimmerman, S. B. (1983) Polymer stimulated ligation: enhanced blunt- or cohesive-end ligation of DNA or deoxyribooligonucleotides by T4 DNA ligase in polymer solutions. *Nucleic Acids Res.* **11,** 7853–7871.
55. Hayashi, K., Nakazawa, M., Ishizaki, Y., Hiraoka, N., and Obayashi, A. (1986) Regulation of inter- and intra-molecular ligation in the presence of polyethylene glycol. *Nucleic Acids Res.* **14,** 7617–7631.
56. Revie, D., Smith, D. W., and Yee, T. W. (1988) Kinetic analysis for optimisation of DNA ligation reactions. *Nucleic Acids Res.* **16,** 10,301–10,321.
57. Leis, J., Silbe, R., Malathi, V. G., and Hurwitz, J. (1972) Studies on transcription and ligation of RNA, in *Advances in the Biosciences*, vol. 8 (Raspé, G., ed.), Pergamon, New York, pp. 117–144.
58. Uhlenbeck, O. C. and Gumport, R. I. (1982) T4 RNA ligase, in *The Enzymes XV* Part B (Boyer, P. D., ed.), Academic, New York, pp. 31–58.
59. Silber, R., Malathi, V. G., and Hurwitz, J. (1972) Purification and properties of bacteriophage T4 induced RNA ligase. *Proc. Natl. Acad. Sci. USA* **69,** 3009–3013.

60. Sugino, A., Snopek, T. J., and Cozzarelli, N. R. (1977) Bacteriophage T4 RNA ligase. Reaction intermediates and interaction of substrates. *J. Biol. Chem.* **252,** 1732–1738.
61. Abelson, J. (1979) RNA processing and the intervening sequence problem. *Annu. Rev. Biochem.* **48,** 1035–1069.
62. Snopek, T. J., Wood, W. B., Conley, M. P., Chen, P., and Cozzarelli, N. R. (1977) Bacteriophage T4 RN ligase is gene 63 product, the protein that promotes tail fibre attachment to the baseplate. *Proc. Natl. Acad. Sci. USA* **74,** 3355–3359.
63. Last, J. A. and Anderson,W. F. (1976) Purification and properties of bacteriophage T4-induced RNA ligase. *Arch. Biochem. Biophys.* **174,** 167–176.
64. Moseman-McCoy, M. I., Lubbens, T. H., and Gumport, R. I. (1979) The purification of nuclease-free T4 RNA ligase *Biochim. Biophys. Acta* **562,** 149–161.
65. Vasilenko, S. K., Veniyaminova, A. G., Yamkovoy, V. I., and Maiyorov, V. I. (1979) Bacteriophage T4 RNA ligase. 1. Purification, enzyme assay, dimer form. *Bioorg. Khim.* **5,** 621–627.
66. Walker, G. C., Uhlenbeck, O. C., Bedows, E., and Gumport, R. I. (1975) T4 induced RNA ligase joins single stranded oligoribonucleotides. *Proc. Natl. Acad. Sci. USA* **72,** 122–126.
67. England, T. E., Gumport, R. I., and Uhlenbeck, O. C. (1977) Dinucleoside pyrophosphates are substrates for T4 induced RNA ligase. *Proc. Natl. Acad. Sci. USA* **74,** 4839–4842.
68. Hinton, D. M. and Gumport, R. I. (1979) The synthesis of oligodeoxyribonucleotides using RNA ligase. *Nucleic Acids Res.* **7,** 453–464.
69. Peattie, D. A. (1979) Direct chemical method for sequencing RNA. *Proc. Natl. Acad. Sci. USA* **76,** 1760–1764.
70. Lockard, R. E., Alzner-Deweerd, B., Heckman, J. E., MacGee, J., Tabor, M. W., and RajBhandary, U. L. (1978) Sequence analysis of 5'[^{32}P]labeled mRNA and tRNA using polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **5,** 37–56.
71. Hecht, S. M., Alford, B. L., Kuroda,Y., and Kitano, S. (1978) Chemical aminoacylation of tRNAs. *J. Biol. Chem.* **253,** 4517–4520.
72. Donis-Keller, H. (1979) Site specific enzymatic cleavage of RNA. *Nucleic Acids Res.* **7,** 179–192.

CHAPTER 11

The BAL 31 Nucleases (EC 3.1.11)

Horace B. Gray, Jr. and Tao Lu

1. Introduction

1.1. Source

The extracellular nucleases commonly called the BAL 31 nuclease take their name from the designation given the marine bacterium producing them, which was originally classified as *Pseudomonas* BAL 31 *(1)* and reclassified as belonging to the small genus *Alteromonas (2)* with the species named *espejiana* after its discoverer, a Chilean microbiologist. The nuclease activities were originally found as contaminants in preparations of bacteriophage PM2 grown on this organism, but were shown *(3,4)* to be bacterial products. Only 10–20% of the nuclease activity is found in the periplasm *(5)*. The American Type Culture Collection strain of *Alteromonas espejiana* (ATCC 29659) produces BAL 31 nuclease as proficiently as the strain originally obtained from its discoverer.

1.2. Some Physical Features

Two kinetically and molecularly distinct forms of the nuclease constitute the bulk of the activity in culture supernatants and are the only ones that have been partially characterized *(6)*. The smaller of these single-subunit enzymes, the "slow" (S) form (mol mass 85,000 Da), is derived by proteolysis, mediated by an *Alteromonas*-produced protease, of the "fast" (F) form (mol mass 109,000 Da), which in turn derives from an even larger precursor *(7)*. Conversion of F nuclease to a species indistinguishable in molecular size and catalytic properties from the S enzyme can be done by proteolysis in vitro *(7)*.

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

The BAL 31 nucleases are remarkable for their resistance to inactivation and or denaturation in the presence of detergents, urea, or high concentrations of electrolyte, and are very stable upon extended storage in the cold *(3,4,8)*. The nuclease activities are not highly resistant to inactivation at elevated temperatures, but the overall secondary and tertiary structures of the S enzyme are extremely resistant to disruption: The internally proteolytically cleaved enzyme fails to dissociate at 100°C in the presence of 1% sodium dodecyl sulfate *(7)*.

1.3. Reactions Catalyzed

The nucleases have three general activities against DNA: a 3'–5' exonuclease activity that apparently removes one residue at a time from duplex structures *(9)*, a 5'–3' exonuclease activity that carries out the bulk of the degradation of single-stranded DNA *(10)*, and a much slower (2–3% of the bond cleavage rate for the 5'–3' exonuclease *[10]*) endonuclease activity against single-stranded DNA that is also elicited by a variety of covalent and noncovalent lesions or distortions in duplex DNA *(3,4,6,8,11–14)*. The two forms differ greatly in the rates of catalysis, at given molar concentrations of duplex ends and enzyme, for the exonuclease reaction if the DNA is double-stranded (hence the aforementioned "fast" and "slow" designations); but the two forms are much more comparable in kinetic behavior if the substrate is single-stranded DNA *(6,10)*. The nucleases are not sugar-specific and catalyze the terminally directed hydrolysis of duplex RNA and readily degrade RNA containing nonduplex structure *(15)*. Endonucleolytic cleavage of duplex RNA in response to lesions or distortions has not been examined except that it is likely that the nucleases can cleave in response to a strand break in duplex RNA *(15)*.

The combination of a 3'–5' exonuclease activity on duplex DNA, a 5'–3' mode of attack on single-stranded DNA, and relatively infrequent endonucleolytic attack in single-stranded DNA serves to reduce the length of linear duplex DNA and is referred to as the duplex exonuclease activity. Partially degraded molecules possess the expected 5'-terminated single-stranded "tails," and there was no evidence for 3'-terminated tails *(9)*, implying that the 5'–3' attack is limited to single-stranded DNA. An unexpected aspect of the mechanism is that the average length of the tails, for a constant number of nucleotides removed by the exonuclease action, decreases markedly to a limiting value with

increasing enzyme concentration *(9)*. In the absence of DNA polymerase-mediated repair, ligation of partially shortened duplexes under conditions favoring the joining of fully base-paired ends is undetectable when low concentrations of nuclease are used, but becomes quite substantial (up to 50%) at nuclease concentrations that give rise to tails of minimum average length *(9)*. This strongly suggests that the 5'–3' exonuclease activity, acting on 5'-terminated single-stranded tails that are generated through attack by the 3'–5' exonuclease on duplex ends, terminates at the junction between single-stranded and duplex structures to leave fully base-paired ends on a significant fraction of the molecules in a partially digested population. The length reduction of duplex RNA *(15)* presumably proceeds by a similar mechanism, but has not been characterized. The BAL 31 nucleases are the most efficacious enzymes known for the controlled length reduction of duplex DNA, and are apparently the only enzymes that can mediate this reaction for duplex RNA.

The covalent lesions in duplex DNA that have been shown to elicit cleavage by BAL 31 nuclease, some of which have been examined only in the case of the S form, include strand breaks, nitrous acid-induced inter-strand crosslinks, UV irradiation-induced photoproducts, adducts from reaction with arylating and alkylating agents or with Ag^{+} and Hg^{++} ions, and apurinic sites *(4,10,13,14)*. Quasiduplex DNA, in which there are extra nucleotides in one strand separated by 3 bp, was attacked by the BAL 31 nuclease, but not by the S1 and *Neurospora crassa* nucleases *(16)*. Noncovalent alterations in duplex structure that can cause endonucleolytic attack include those associated with a very moderate degree of negative supercoiling and very high positive supercoiling *(12)*, as well as the junctions between regions of B- and Z-DNA helical structures in the same molecule *(8)*.

The only known nonsubstrate nucleic acids are nonmodified, nonsupercoiled (or slightly positively supercoiled) circular duplex DNAs, which are cleaved at such slow rates that they serve as excellent controls for the detection of lesions or distorted structures in such DNAs *(14)*. The strand break (nick) introduced as the initial cleavage event itself provides a substrate site, so that the other strand is cleaved, usually after the removal of several nucleotides from the originally nicked strand at the site of the nick *(17)*, to produce termini that are then attacked so as to shorten the resulting duplex as previously outlined.

1.4. Uses in Manipulation of Nucleic Acids

The duplex exonuclease activity of the nucleases has been very widely exploited, because the removal of sequences from DNA termini in a controlled manner is desirable in numerous applications. Unidirectional deletions can be done for DNAs cloned into a circular vector if single unique sites for two restriction enzymes are present, one on each side of the cloned insert. Linearization by cleavage with one of the restriction enzymes allows for unidirectional deletions in the insert, whereas subsequent cleavage with the other restriction enzyme releases the partially degraded vector DNA so that the shortened insert can be religated to an intact vector DNA for subcloning. Conversely, manipulations of the vector DNAs themselves in this manner have proven useful—an early application of this method led to a vector DNA modified so that cloned inserts abutted the DNA coding for a signal peptide of the vector, allowing secretion of the protein products of cloned genes into the periplasmic space of *E. coli (18,19)*. Such unidirectional deletions in cloned inserts have been used as a means to sequence, using a single primer for the DNA polymerase-mediated extension, long segments of DNA cloned into M13 phage-derived vectors through the production of a set of progressively shorter unidirectionally deleted derivatives of the original insert *(20)*. The progressive removal of sequences from duplex termini allows the determination of the restriction map of a given DNA by observing the order in which the fragments from digestion with the restriction enzyme in question disappear from gel electrophoretic patterns of progressively shortened samples *(21)*. This technique is greatly enhanced if the fragment to be mapped is in a circular cloning vector as ambiguity resulting from simultaneous degradation from both ends of the fragment is eliminated *(22)*. Cloning vectors containing infrequently cleaved restriction sites have been constructed in this laboratory to take advantage of BAL 31 nuclease-mediated deletions in both sequencing and restriction fragment mapping *(22)*.

The blockage of the exonuclease action by nucleosomes and by interstrand crosslinks, such as can be induced by psoralen derivatives, has been used to map the locations of such structures on duplex DNAs *(23,24)*. The detection of lesions or distortions in duplex DNA through the endonuclease activity has been alluded to earlier.

2. Enzyme Requirements

2.1. pH

The near-neutral pH optima for the nuclease-catalyzed reactions *(4,9,25)* is an advantage for most work. There is a difference in the pH optima for single-stranded and linear duplex substrates for both the S and F nucleases, which are in the ranges 8.5–8.8 and 7.0–8.0 for single-stranded and linear duplex DNA, respectively *(4,9,25)*. Attack on supercoiled DNA, examined only for the S enzyme, was optimal at the same pH as for linear duplex DNA *(4)*.

2.2. Metal Ion Cofactors

Ca^{2+} and Mg^{2+} are both required cofactors *(3)*. Ca^{2+} is essential for activity, with nuclease activities on both single- and double-stranded DNA irreversibly (with respect to the readdition of excess Ca^{2+}) lost in solution if the molar concentration of EDTA exceeds that of this cation *(3,4)*. However, activity can be recovered after electrophoresis under denaturing conditions (sodium dodecyl sulfate [SDS]-polyacrylamide gels) by incubation in a Ca^{2+}-containing buffer *(7)*.

Maximum velocity against both single-stranded and linear duplex DNAs at constant [Ca^{2+}] is achieved between 10 and 15 m*M* Mg^{2+} *(9,25)*. At nominally zero (actually 0.01–0.02 m*M*) Mg^{2+}, there was residual duplex exonuclease activity (8 and 45% for the S and F enzymes, respectively), but virtually none for single-stranded DNA. In corresponding profiles where [Ca^{2+}] was varied, concentrations near 10 m*M* are needed to achieve full velocity for the length reduction of duplex DNA, but activity is maximal on single-stranded DNA at or below 1 m*M* *(9,25)*.

In light of the above, a buffer containing 12.5 m*M* each of Ca^{2+} and Mg^{2+} is recommended, since this confers full activity with respect to both classes of substrate. The 5-m*M* concentrations of these ions in the BAL 31 nuclease buffers recommended by some suppliers would yield only 60–65% of the duplex exonuclease activity.

2.3. Effects of Temperature

The optimal temperature for the activity on single-stranded DNA is near 60°C *(4)*, but the use of such elevated temperatures is impractical. Internal breaks would almost certainly be introduced into duplex DNA because of at least transient thermally mediated unstacking of base

pairs. Moreover, the half-life of the activity on single-stranded DNA and the duplex exonuclease activity are 3–5 min at 50°C *(10)*. The half-life of the activity is at least 20 h at 30°C, which is the recommended temperature for manipulations with these enzymes.

In a storage buffer containing 5 m*M* each of Ca^{2+} and Mg^{2+} near 4°C, the nuclease should retain most of its activity on a time scale of years as inferred from studies in which no activity was lost upon storage for several months *(4)*. Most suppliers provide the enzyme in 50% (v/v) glycerol for storage at –20°C, in which it should maintain full activity indefinitely.

2.4. Effects of Ionic Strength, Protein Denaturing Agents, and Other Potential Inhibitors

A concentration of NaCl of 0.6*M* is present in the usual assay buffer. The activity on single-stranded DNA is optimal in this range *(4)*, but the duplex exonuclease activity appears to increase if the ionic strength is reduced from that of the assay buffer *(26)*. However, the NaCl concentration of the usual buffer will repress the activity of contaminating salt-sensitive nucleases or other factors that can cause adventitious breaks. The nuclease was used at 0.1*M* NaCl with no evidence of internal breaks in nonsupercoiled closed duplex DNA provided that the buffers were autoclaved; omission of this step did lead to the appearance of unexpected strand breaks *(12)*. Such internal breaks must be avoided because each such lesion, after cleavage of the intact strand, gives rise to two new termini, which then serve as substrates for the duplex exonuclease activity.

Very high concentrations of 1:1 electrolytes are tolerated extremely well by the nucleases. With single-stranded DNA as substrate, the S enzyme displays 40 and 27% of its maximal activity in 4.5*M* NaCl and 7*M* CsCl, respectively *(4)*; in a comparison at only two concentrations, the F nuclease was 70% as active in 4.5*M* NaCl as in 0.6*M* NaCl *(8)*. The duplex exonuclease activity of the F nuclease at 4.5*M* NaCl was 26% of that at 0.6*M (8)*.

Protein denaturing agents at concentrations that normally denature proteins are unable to abolish the BAL 31 nuclease activities. Single-stranded DNA was hydrolyzed at 40% of the maximal rate in 6.5*M* urea *(4)*, whereas crude preparations maintained activity against both

single- and double-stranded substrates in the presence of 5% SDS if both Mg^{2+} and Ca^{2+} were present prior to exposure to the detergent *(3)*. Preliminary experiments on the S enzyme showed that substantial activity on single-stranded DNA was observed in up to 6*M* guanidinium hydrochloride *(5)*.

Cleavage of adducts of nonsupercoiled closed duplex DNA with Hg^{2+} or Ag^{+} by the F nuclease *(13,27)* suggested that the enzyme is not highly sensitive to these metal ions, but it must be pointed out that, in the case of Ag^{+}, virtually all the metal ion was bound to DNA nucleotide. The only metal ion tested that appeared to be strongly inhibitory was Zn^{2+} *(28)*. Inhibitors of unknown nature that can very strongly repress the duplex exonuclease activity of the BAL 31 nucleases have been observed in preparations of plasmid DNA prepared by rapid procedures, such as the alkaline lysis protocol of Birnboim and Doly *(29)*, that do not involve purification by centrifugation to equilibrium in CsCl density gradients containing intercalating compounds such as ethidium bromide *(30,31)*. An additional purification step, using small columns of Bio-Gel A-0.5m agarose-based size exclusion resin (Bio-Rad Laboratories, Richmond, CA), has been shown to render the DNA digestible by commercially provided (Gibco BRL, Grand Island, NY) BAL 31 nuclease, thus avoiding the time-consuming equilibrium ultracentrifugation procedure *(30)*.

Another unexplained strong source of inhibition was found as a component of a T4 DNA polymerase repair reaction: DNA that had been thus treated, and the polymerase inactivated, without subsequent precipitation and resuspension of the DNA was virtually unaffected by the duplex exonuclease activity *(5)*. The inhibitor was shown not to be the 5'-deoxynucleoside triphosphate substrates for the polymerase and was not a large macromole since the nuclease activity on the repaired DNA was restored when it was passed through a Sephadex® G25 "spun column."

The 5'-deoxynucleoside monophosphates (5'-dNMPs) resulting from nuclease action *(4)* are apparently inhibitory because the velocity of the duplex exonuclease reaction decreases significantly with time, even though the substrate concentration (concentration of DNA ends) remains constant until molecules begin to be completely degraded *(6)*. The nuclease is known to bind to 5'-dNMPs, because an affinity column consisting of 5'-dNMPs covalently attached to agarose is used in the purification procedure *(4)*.

Systematic studies have been done on the effects of a "chaotropic" agent ($NaClO_4$) and compounds that protect cells from damage due to increases or decreases in extracellular osmolarity *(10)*. However, these are not compounds that are likely to be part of a reaction mixture in general applications of these enzymes, and it is noted here only that $NaClO_4$ rapidly inactivates the nucleases at concentrations in the range of 1–2*M*.

2.5. Effects of Proteolysis

The proteolytic conversion of F to S nuclease that takes place in *Alteromonas* culture supernatants can be mimicked by proteolysis in vitro *(7)*. However, it has also been found that extended exposure to protease of the S nuclease can preferentially remove the duplex exonuclease activity, but the bulk of the activity against single-stranded DNA survives. The ability to linearize a nicked circular DNA correlates with the remaining activity on single-stranded DNA, so that the endonuclease activity in response to lesions is not lost upon proteolysis. Since the bulk of the activity on single-stranded DNA is of the 5'–3' exonuclease variety, whereas attack on duplex DNA is through a 3'–5' mode, the data suggest that the protease-treated nuclease may lose its 3–5' exonuclease capability *(7)*. The observation is interesting, because if conditions for the removal of the duplex exonuclease activity can be optimized, the remarkable power of the enzyme to cleave endonucleolytically in response to alterations in duplex DNA structure could be much better exploited to reveal imperfections, such as small mismatches in nominally duplex DNAs, since the information as to the site of cleavage would not be obliterated by the exonuclease activity.

The internal breaks that were thought to be introduced into the S nuclease by proteolysis to produce the preferential loss of duplex exonuclease activity were not revealed in denaturing polyacrylamide gels, where the proteins are heated in boiling water in the presence of 1% SDS and 20 m*M* β-mercaptoethanol prior to electrophoresis. Rather, the mol mass appeared to be unaffected for samples that had clearly undergone loss of over 90% of the starting duplex exonuclease activity, but retained over half of the activity against single-stranded DNA. When urea was added to the denaturation buffer to 6*M*, the internal breaks were revealed by the progressively more extensive fragmentation of

the nuclease with increasing extent of exposure to protease *(7)*. This indicates a resistance of the breakup of the secondary and tertiary structure of the S nuclease to extreme denaturing conditions that is apparently without parallel. It is interesting that the catalytic activity is maintained near room temperature in the presence of strong denaturants, but is rather thermolabile. A relatively thermally sensitive binding/catalytic site(s) in an otherwise extremely stable protein structure is suggested.

3. Experimental Procedures for Unit Assay and Characterization of the Duplex Exonuclease Activity

3.1. General Considerations

Most commercial preparations of BAL 31 nuclease are mixtures of the F and S forms. A problem with their use for the most common application, the controlled length reduction of duplex DNA, is the possible variance of the relative amounts of F and S nucleases from batch to batch. This is because the F enzyme is derived from a larger precursor by proteolysis, mostly in the culture supernatant, and itself is the source of the smaller S enzyme *(7)*. Hence, the relative amounts of the two species can vary depending on the extent of exposure to the supernatant protease(s) before the nucleases are separated from these. Only one supplier (International Biotechnologies, Inc., New Haven, CT) offers the purified S and F nucleases, which could be expected to be reproducible in duplex exonuclease activity from batch to batch. Since the nuclease is very stable on extended storage in the cold *(4)*, it seems advisable to obtain as large a sample as possible and to characterize it for duplex exonuclease activity unless the pure S or F species is obtained.

The recommended unit assay for general characterization of a sample (e.g., in order to check its activity after long periods of storage) is using single-stranded DNA *(4)* because maximum velocity conditions, desirable in any enzyme assay, are readily achieved. Examination of the velocity vs substrate concentration profile with viral ϕX174 DNA as substrate shows that 90% of the apparent maximum velocity V_{max}^{app} is achieved at a concentration of this single-stranded DNA near 10 µg/mL for the S nuclease and at an even lower concentration for the F nuclease *(6)*. Calculations using recent carefully determined values for the apparent kinetic parameters, which take into account the fact that the major mode of degradation of single-stranded DNA is

exonucleolytic attack from the 5' ends, indicate that 90% of V_{max}^{app} for the S and F nucleases *(10)* can be obtained with approx 34 and 9 μg/mL, respectively, of a DNA the size of ϕX174. Since the denatured calf thymus DNA normally used in this assay is not expected to be larger in molecular weight that ϕX174 DNA and is routinely used at concentrations of at least several hundred μg/mL, the concentration of ends should be much more than sufficient to provide maximum velocity conditions. Data with calf thymus DNA as substrate in the velocity vs substrate concentration curve mentioned above (6) showed that V_{max}^{app} was reached at concentrations near 30 μg/mL, which supports the notion that the standard assay provides maximum velocity conditions. However, this assay cannot be used to predict the kinetics of the duplex exonuclease reaction for mixtures of the two forms.

The use of duplex DNA to assay the nuclease through the release of nucleotides owing to the duplex exonuclease activity, as recommended by some suppliers, will not take place under close to V_{max} conditions unless short duplexes are used at high concentrations. For a DNA of 1000 bp, its concentration would have to be over 150 μg/mL for the F nuclease to achieve 90% of maximum velocity according to the most recent estimates for the apparent K_m *(10)*; for the S nuclease, the corresponding concentration would be in the 2 mg/mL range. DNA of defined molecular size is required so that the substrate concentration (molar concentration of termini) is known. If enzyme is assayed using duplex DNA for the purpose of determining the activity of a particular batch on linear duplexes, care should be taken to keep the substrate concentration in subsequent digests close to those in the pilot experiments. Also, the G + C content of the DNA in the pilot experiments should not differ significantly from that of DNA in experiments based on the pilot studies, unless the F nuclease is used (see Section 3.3.1.). Finally, there is a DNA length dependence of the kinetic parameters. (*see* Section 3.3.3.)

3.2. Assay Using Single-Stranded DNA

This assay has been described in detail *(4)*, but is reproduced here for the convenience of the reader and because of recent modifications *(7)* in the procedure. Calf thymus DNA (Type I, Sigma Chemical Co., St. Louis, MO) is dissolved at a nominal concentration of 1–2 mg/mL in BE buffer (100 m*M* NaCl, 20 m*M* Tris-HCl, 1 m*M* EDTA [pH 8])

by overnight stirring. NaOH (2*M* or higher) is added to a nominal concentration of 0.1*M* and, after several minutes of alkaline denaturation, the solution is neutralized with a severalfold excess of 3*M* sodium phosphate (pH 7). Dialysis into BE buffer is followed by adjustment to the proper concentrations of NaCl, $MgCl_2$, and $CaCl_2$ in 20 m*M* Tris-HCl/1 m*M* EDTA (pH 8) to yield, after addition of enzyme solution, the desired composition of assay buffer: 0.6*M* NaCl, 12.5 m*M* $CaCl_2$, 12.5 m*M* $MgCl_2$ in the aforementioned Tris/EDTA buffer. If 0.1 vol of enzyme solution in CAM buffer (100 m*M* NaCl, 5 m*M* $MgCl_2$, 5 m*M* $CaCl_2$, 20 m*M* Tris-HCl, 1 m*M* EDTA [pH 8]) is mixed with the DNA solution in the assay, the DNA solution should contain 0.656*M* NaCl, 13.3 m*M* $MgCl_2$, and 13.3 m*M* $CaCl_2$, which is conveniently achieved with 5, 0.5, and 0.5*M* stock solutions of the respective chemicals, each containing the Tris/EDTA buffer. CAM buffer (usually containing 50% [v/v] glycerol) is used as the nuclease storage buffer by several suppliers.

After mixing 0.9 vol of the above DNA solution with 0.1 vol of enzyme solution at 30°C, aliquots of 0.4 mL are withdrawn at such times as 5, 10, and 15 min and mixed with 40 µL of 0.5*M* sodium EDTA (pH 8–8.5) followed by 0.8 mL of 10% (w/v) $HClO_4$. This is conveniently done in a 1.5-mL Eppendorf centrifuge tube for centrifugation of the precipitated nondigested DNA after the acidified mixtures have stood on ice for several minutes (but not for extended periods). Five-tenths to 1 mL of supernatant is carefully removed from each tube after centrifugation for approx 10 min, and the value of A_{260} is measured in a 1-cm path length quartz cuvet with a 4 mm wide cavity. The blank (zero time of incubation) is made up separately by mixing 40 µL of 0.5*M* EDTA with 0.4 mL of the denatured DNA solution before addition of the enzyme and $HClO_4$ solution. This ensures that any acid-soluble UV-absorbing material in nuclease preparations will not contribute to the A_{260} values when the blank absorbance is subtracted.

The resulting plot of A_{260} vs time (t) (in minutes), including the point (0,0), should be a good straight line (correlation coefficient from linear least squares analysis 0.995 or greater) from which the value of the slope, $\Delta A_{260}/t$, is obtained. According to Vogt *(32)*, the number of enzyme units in the total mixture is given by:

$$U = (\Delta A_{260}/t)(30.3)(V_{tot}) \quad (1)$$

where V_{tot} is the total volume of the acidified reaction mixture aliquot in mL (V_{tot} = 1.24 mL if the above volumes are used). Division by the volume of enzyme solution per aliquot (0.04 mL in the above) yields the U/mL of nuclease in the solution before dilution into the assay mixture.

For the removal of aliquots at 5-min intervals, $\Delta A_{260}/t$ should be at least 0.02 min^{-1}, which implies that the enzyme solution being assayed (before dilution into the aforementioned reaction mixture) should be in the range of 20 U/mL. Correspondingly, longer assays are needed for accurate results with more dilute nuclease solutions. Only when the absorbance of the acid-soluble released DNA nucleotides significantly exceeds 1 has nonlinearity in the plots of A_{260} vs t been observed.

3.3. Characterization of the Duplex Exonuclease Activity

3.3.1. Kinetics of Molecular Weight Reduction

Testing of each batch of commercial BAL 31 nuclease for its ability to catalyze the duplex exonuclease reaction, where this use is to be made of the enzyme, is necessary as discussed in Section 3.1. This is conveniently done by subjecting restriction enzyme-generated fragments of a relevant unique sequence (e.g., a plasmid DNA or the cloned, isolated DNA fragment that is itself to be treated) to progressive digestion and analyzing the partially degraded fragments by agarose gel electrophoresis using (usually) the nondegraded restriction fragments as molecular size markers. It is advisable to assume that the kinetics will be those of the S nuclease, since this species usually predominates in culture supernatants, and to adjust the amount of enzyme accordingly if significantly faster degradation is observed.

For estimation of the rate of degradation, the equation:

$$v_0/[S] = V_{max}^{app} / [S] + K_m^{app} \quad (2)$$

has been used *(6)*, where v_0 is the initial reaction velocity, [S] is the substrate concentration (mol/L of duplex termini), V_{max}^{app} is the apparant maximum velocity corresponding to the concentration of nuclease used, and K_m^{app} is the apparent Michaelis constant. It is important to note that the kinetic parameters are apparent values and do not necessarily have the interpretations associated with Michaelis-Menten kinetic analysis of reactions in which only a single enzyme-substrate need be considered. At least two such intermediates, one

corresponding to enzyme nonspecifically bound to duplex DNA away from the ends and not catalytically active, and another productive complex with enzyme bound at the termini, are needed to model this reaction *(10)*.

The significance of $v_0/[S]$ is that it is the number of nucleotide residues released per DNA terminus per unit time *(6)*, which is usually the desirable parameter. This can be taken as equal to one-half the number of base pairs removed per terminus for digests of more than approx 50–100 nucleotides removed per terminus at enzyme concentrations that minimize single-stranded tail length (*see* Section 3.3.2.). At very low nuclease concentrations, substantial single-stranded tails may be present *(9)* so that the number of residues removed is not equal to one-half the number of base pairs removed.

Data obtained in recent kinetic studies of the duplex exonuclease activity to determine the effects of the G + C content of the substrate for the S nuclease gave very large standard deviations for the values of K_m^{app} and V_m^{app} so that substitution of these values in Eq. (2) would not give reliable estimates for $v_0/[S]$ *(10)*. The reason proved to be that the values of K_m^{app} tend to be much larger than the largest values of [S] compatible with the photometric technique, even after cleavage of the DNA with a restriction enzyme to increase the number of termini per unit weight concentration of DNA. Hence, the substrate concentration range falls far short of the range of approx 0.33–2 K_m^{app} that is considered to yield optimally accurate results for such determinations *(33)*. In agreement with this, the plots of v_0 vs [S] were all very good straight lines, as expected from Eq. (2) where $[S] << K_m^{app}$; the right-hand side of Eq. (2) then reduces to V_m^{app}/K_m^{app}, which will be constant at a fixed enzyme concentration and is the slope of a plot of v_0 vs [S].

Hence, it is possible, in the range of [S] that will ordinarily be used in such experiments, to calculate $v_0/[S]$ at any given enzyme concentration and data are moreover available to include the dependence on G + C content as described in Section 3.3.3. As an example, consider a test DNA of G + C content near 50 mol percent containing 5000 bp that has been cut into six fragments by a restriction endonuclease, and 2 µg are treated with BAL 31 nuclease in a vol of 100 µL. This would be convenient for removal of several aliquots containing approx 0.5 µg of DNA each for analysis (after quenching the reaction with an excess of EDTA) at different times of exposure (i.e., 5-min intervals) to nuclease

in an ordinary agarose gel electrophoresis experiment. For duplex DNA of this G + C content, the value of v_0/[S] is close to 0.80 residues released/terminus/min at an enzyme concentration of 1.0 U/mL *(10)*. Assuming a desired rate of 10 residues removed/terminus/min, an enzyme concentration of 10/0.8 = 12.5 U/mL should be used independently of the substrate concentration.

The older data for the S nuclease *(6)* give very different maximum velocities/U/L of nuclease and an estimate of K_m^{app} that is at least an order of magnitude lower than those evidenced (although with considerable error) from the more recent experiments *(10)*. Because it turns out that these data were measured under conditions where the total enzyme concentration $[E]_{tot}$ actually was in excess of that of the substrate, the usual assumption of Michaelis-Menten kinetics that $[S] >> [E]_{tot}$ was not valid. The more recent data do not suffer from that limitation, allow the dependence of v_0 on G + C content to be assessed (*see* Section 3.3.3.), and appear to be quite accurate as long as $[S] << K_m^{app}$ a condition that will be met unless concentrations of hundreds of micrograms per milliliter of fragments averaging several thousand base pairs in length are used. Hence, use of the more recent data is recommended.

Recent studies on the duplex exonuclease activity of the F nuclease *(10)* give results much more in agreement with the older data *(6)* and strongly suggest that the values for K_m^{app} are generally more than an order of magnitude lower than those for the S nuclease, which necessitates the use of Eq. (2) without the assumption that [S] can be dropped from the right-hand side. Here, values of K_m^{app} and V_{max}^{app} were determined with good accuracy, because it was possible to access substrate concentrations that were in the range of K_m^{app}.

These studies revealed that both K_m^{app} and V_m^{app} decrease with increasing length of the linear duplex substrate for the F nuclease. Parallel studies could not be done for the S enzyme because of the error in determination of K_m^{app} noted above. These data are consistent with an interpretation of the mechanism in which nonspecific binding to the duplex away from the ends is followed by a "search" process to form a productive complex with nuclease bound at a terminus *(10)*. The values of V_{max}^{app}, but not those of K_m^{app}, seem to level off above a length in the range of 2000–2500 bp. It appears that reasonable kinetic parameters to use for the F nuclease, for molecules in this size range or above, are 2.14 ± 0.04 nmol/U/min for the maximum velocity/U of

nuclease/L and 58 ± 2 n*M* for K_m^{app}. V_{max}^{app} in Eq. (2) is obtained simply by multiplying the normalized maximum velocity above by the nuclease concentration in U/L (the velocity was shown to be linear with enzyme concentration for both the F and S enzymes over a wide range of substrate concentration) *(9,25)*. The dependence of v_0/[S] on %(G + C) is much weaker than for the S enzyme (*see* Section 3.3.3.).

Several comments are in order here. The DNA should be precipitated and resuspended in a small volume of Tris/EDTA buffer, and diluted into a reaction mixture that will give the desired buffer composition (same as that for the assay using single-stranded DNA), since it will be difficult to resuspend directly in the reaction buffer. The expected extent of degradation, at least in the first one or two aliquots, should not be more than approx 20% of the average fragment length so that smearing of the bands of asynchronously digested DNA will not preclude reasonable molecular size estimation from at least some of the fragments. To avoid changes in substrate concentration owing to complete digestion of fragments, a restriction enzyme that produces any very short fragments from the DNA should be avoided. Because the velocity does decrease with time, apparently because of inhibition by released 5'-dNMPs *(6)*, it should not be assumed that extensive digestion of a DNA (more than 25%, for example) will proceed at the same rate as that observed early in the course of a digest.

3.3.2. Effect of Nuclease Concentration

The presence of substantial single-stranded tails on partially digested duplexes can be avoided if the concentration of nuclease is near 2 and 10 U/mL for the F and S enzymes, respectively, or greater *(9)*. The value corresponding to S nuclease should be used for all commercial preparations not sold as separated S and F enzymes. These (or higher) concentrations are compatible with most uses of the enzymes, corresponding to 33 and 8 residues removed/terminus/min from a DNA of near 50% G + C content for the F and S nucleases, respectively, at the substrate concentration (7.3×10^{-8} *M*, needed only for the calculation for F nuclease) of the above example. Significantly, digestion at or above the aforementioned concentrations not only leaves short average tail lengths (approx 7 residues), but allows ligation under conditions favoring blunt-end joining of up to 50% of that observed with DNA known to possess fully base-paired ends *(9)*. Hence, treatment

with S1 nuclease or DNA polymerase to remove single-stranded ends and provide fully based-paired termini is unnecessary, although it is apparent *(9)* that polymerase-mediated repair will enhance the fraction of ligatable termini.

3.3.3. Effects of Base Composition

The duplex exonuclease activity of the S form decreases markedly with increasing G + C content *(8,21)*, and the magnitude of this effect has been examined over the range of 37–66 mol% G + C residues *(10)*. For the S nuclease, the effect is significant and leads to an approx 4.2-fold change in $v_0/[S]$ over this composition range. The dependence of $v_0/[S]$ on G + C content over the above range can be approximated, independently of the substrate concentration as previously noted, by the equation:

$$v_0/[S] = 16.85 - 0.546P + 0.00456P^2 \quad (3)$$

where P is the mol% G + C, and the value of $v_0/[S]$ is that corresponding to an enzyme concentration of 1 U/mL. The desired value of $v_0/[S]$ is thus calculated by multiplication of that from Eq. (3) by the nuclease concentration (U/mL) to be used in the actual experiment. The equation should not be used to extrapolate outside of the above range of base composition ($37 \leq P \leq 66$).

The duplex exonuclease activity of the F enzyme is much less dependent on composition than for the S enzyme, with $v_0/[S]$ varying only by a factor of 1.7 over the above range at the substrate concentration of the example used earlier. Moreover, nearly all the change occurs between approx 52 and 37% G + C. Given the paucity of data, the values for the maximum velocity/U/L of nuclease and K_m^{app} above should be used in the range of 66–52% G + C, and the maximum velocity parameter should be increased by interpolation from the value above to a maximum of 1.7 times that value in the range from 52–37% G + C.

The less pronounced effect of variances of G + C content for the F nuclease suggests its use where it is desired to minimize the effects of local sequence features, such as local regions of high G + C content, on the termination points of digestion of partially degraded samples. With samples that corresponded to S nuclease, sequence analysis of partially degraded samples showed a strong tendency for the nuclease to stop in regions of several consecutive G + C pairs with a strong preference for dG at the 5' end *(18,19)*. When the location of "stop"

sites was analyzed in this laboratory for both the S and F nucleases, this tendency was confirmed for the S enzyme and shown also to be true for the F species *(10)*. This suggests that the factors accounting for the dependence of the rates of 3'-terminal nucleotide removal by the respective nuclease species on G + C content are not a direct result of the relative thermodynamic stability of the two types of base pairs. This is because the F nuclease, which removes 3'-terminal nucleotides more rapidly and shows a lesser dependence of this rate on G + C content, would otherwise be expected to show a lesser tendency to stop in "runs" of several G-C pairs. Further support for the lack of a direct role of relative thermodynamic stability of base pairs in this connection are the facts that stops tend to occur at 5' dG residues, whereas stops at both dC and dG residues should be expected if stability alone were the dominant factor and the kinetic parameters show little dependence on G + C content above about 50%, whereas the melting temperature T_m, an indicator of stability, is linearly dependent on the G + C content. Use of the F nuclease will make average extents of degradation more predictable, however, because localized changes in velocity associated with local differences in G + C content are decreased.

4. The Nuclease as a Probe for Lesions in Nonsupercoiled Closed DNA

It was noted earlier that the nucleases can introduce breaks in double-stranded DNA in response to a wide variety of both covalent and noncovalent distortions of duplex structure, whereas covalently closed circular duplexes that are essentially nonsupercoiled or slightly positively supercoiled (form I° DNA) are extremely refractory to nuclease-mediated attack *(11,13,14)*. This provides for very stringent and general tests of altered duplex structure caused by any perturbation that does not introduce strand breaks (such breaks themselves provide substrate sites) or result in significant negative supercoiling of form I° DNA. Information as to whether duplex DNA structure is perturbed by a given reagent or treatment is valuable as a guide to further studies.

Form I° DNA is readily produced from any supercoiled duplex species by incubation with DNA topoisomerase I. This enzyme is commercially available, but crude preparations *(34)* used as described *(11)* are quite adequate for this purpose. The nicking-closing reaction is routinely monitored by electrophoresis in 1% agarose gels.

Whether the modifying treatment itself introduces strand breaks is readily ascertained by electrophoresis of the treated form I° DNA and nontreated controls in agarose gels containing a low concentration of ethidium bromide (e.g., *11*). Ethidium bromide at 1 μg/mL should also be present in nuclease reaction mixtures to ensure that changes in the ionic environment and/or temperature between the conditions of the nuclease reaction and the topoisomerase treatment *(35)* do not result in even slight negative supercoiling. High enzyme concentrations, on the order of 100 U/mL, should initially be used to provide a stringent test, because the rate of attack in response to various types of lesions varies widely.

The rate of disappearance of form I° DNA, as by agarose gel electrophoresis of aliquots containing equal volumes of the reaction mixture, is the parameter to be assayed, since nuclease at such concentrations will rapidly degrade, via the duplex exonuclease activity, the linear duplexes produced after the introduction of a strand break, and cleavage in the second strand. In fact, this reaction has been so rapid that it led to a novel coupled assay in one study in which the loss of fluorescence of EtdBr owing to the complete digestion of a DNA of 10,000 bp was used to monitor the rate of cleavage in response to the presence of apurinic sites; this digestion was fast compared to the rate of introduction of the initial scission *(14)*. In the only study where direct comparisons were made, the F nuclease was more efficient in cleaving in response to a covalent lesion than the S species *(14)*.

The exonuclease activity unfortunately destroys information as to the site of the initial cleavage in reactions converting an appreciable percentage of the DNA. For a type of distortion that readily elicits cleavage, that resulting from negative supercoiling, it has proven possible to localize sites of BAL nuclease cleavage in experiments in which the overall extent of cleavage was very limited and the DNA was labeled with ^{32}P after cleavage with restriction nucleases so that the small fraction undergoing endonucleolytic attack could be examined in autoradiograms of agarose gels *(36)*. The interesting possibility of using nuclease that has been treated with protease so that the bulk of the duplex exonuclease activity is eliminated in order to localize sites of attack was noted in an earlier section. This will require further experimentation to optimize the conditions for such protease treatment.

Acknowledgments

Work reported here for which support has not been acknowledged elsewhere was supported in part by grant GM-21839 from the National Institutes of Health and in part by Biomedical Research Support Grant S07RR-07147 (University of Houston).

References

1. Espejo, R. T. and Canelo, E. S. (1968) Properties and characterization of the host bacterium of Bacteriophage PM2. *J. Bacteriol.* **95,** 1887–1891.
2. Chan, K. Y., Baumann, L., Garza, M. M., and Baumann, P. (1978) Two new species of *Alteromonas: Alteromonas espejiana* and *Alteromonas undina. Int. J. Syst. Bacteriol.* **28,** 217–222.
3. Gray, H. B., Jr., Ostrander, D. A., Hodnett, J. L., Legerski, R. J., and Robberson, D. L. (1975) Extracellular nucleases of *Pseudomonas* BAL 31. I. Characterization of single strand-specific deoxyriboendonuclease and double-strand deoxyriboexonuclease activities. *Nucleic Acids Res.* **2,** 1459–1492.
4. Gray, H. B., Jr., Winston, T. P., Hodnett, J. L., Legerski, R. J., Nees, D. W., Wei, C.F., and Robberson, D. L. (1981) The extracellular nuclease from *Alteromonas espejiana*: an enzyme highly specific for nonduplex structure in nominally duplex DNAs, in *Gene Amplification and Analysis,* vol. 2, (Chirikjian, J. G. and Papas, T. S., eds.), Elsevier North-Holland, New York, pp. 169–203.
5. Hauser, C. R. and Gray, H. B., Jr., unpublished.
6. Wei, C.-F., Alianell, G. A., Bencen, G. H., and Gray, H. B., Jr. (1983) Isolation and comparison of two molecular species of the BAL 31 nuclease from *Alteromonas espejiana* with distinct kinetic properties. *J. Biol. Chem.* **258,** 13,506–13,512.
7. Hauser, C. R. and Gray, H. B., Jr. (1990) Precursor-product relationship of larger to smaller molecular forms of the BAL 31 nuclease from *Alteromonas espejiana*: Preferential removal of duplex exonuclease activity relative to endonuclease activity by proteolysis. *Arch. Biochem. Biophys.* **276,** 451–459.
8. Kilpatrick, M. W., Wei, C.-F., Gray, H. B., Jr., and Wells, R. D. (1983) BAL 31 nuclease as a probe in concentrated salt for the B-Z DNA junction. *Nucleic Acids Res.* **11,** 3811–3822.
9. Zhou, X.-G. and Gray, H. B., Jr. (1990) Mechanism of exonuclease action of BAL 31 nuclease. *Biochim. Biophys. Acta* **1049,** 83–91.
10. Lu, T. (1992) Some physical and catalytic properties of BAL 31 nuclease. Ph.D. dissertation, University of Houston.
11. Legerski, R. J., Gray, H. B., Jr., and Robberson, D. L. (1977) A sensitive endonuclease probe for lesions in deoxyribonucleic acid helix structure produced by carcinogenic or mutagenic agents. *J. Biol. Chem.* **252,** 8740–8746.
12. Lau, P. P. and Gray, H. B., Jr. (1979) Extracellular nucleases of *Alteromonas espejiana* BAL 31. IV. The single strand-specific deoxyriboendonuclease activ-

ity as a probe for regions of altered secondary structure in negatively and positively supercoiled closed circular DNA. *Nucleic Acids Res.* **6,** 331–357.

13. Wei, C.-F., Alianell, G. A., Gray, H. B., Jr., Legerski, R. J., and Robberson, D. L. (1983) Use of *Alteromonas* BAL 31 nuclease as a probe for covalent alterations in duplex DNA, in *DNA Repair: A Laboratory Manual of Research Procedures,* vol. 2 (Friedberg, E. C. and Hanawalt, P. C., eds.), Marcel Dekker, New York, pp. 13–40.
14. Wei, C.-F., Legerski, R. J., Alianell, G. A., Robberson, D. L., and Gray, H. B., Jr. (1984) A single apurinic site can elicit BAL 31 nuclease-catalyzed cleavage in duplex DNA. *Biochim. Biophys. Acta* **782,** 408–414.
15. Bencen, G. H., Wei, C.-F., Robberson, D. L., and Gray, H. B., Jr. (1984) Terminally directed hydrolysis of duplex ribonucleic acid catalyzed by a species of the BAL 31 nuclease from *Alteromonas espejiana. J. Biol. Chem.* **259,** 13,584–13,589.
16. Evans, D. H. and Morgan, A. R. (1982) Extrahelical bases in duplex DNA. *J. Mol. Biol.* **160,** 117–122.
17. Przykorska, A., Hauser, C. R., and Gray, H. B., Jr. (1988) Circular intermediates with missing nucleotides in the conversion of supercoiled or nicked circular to linear duplex DNA catalyzed by two species of BAL 31 nuclease. *Biochim. Biophys. Acta* **949,** 16–26.
18. Talmadge, K., Stahl, S., and Gilbert, W. (1980) Eukaryotic signal sequence transports insulin antigen in *Escherichia coli. Proc. Natl. Acad. Sci. USA* **77,** 3369–3372.
19. Talmadge, K. and Gilbert, W. (1980) Construction of plasmid vectors with unique *Pst*I cloning sites in a signal sequence coding region. *Gene* **12,** 235–241.
20. Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E., and Surrey, S. (1982) "Nonrandom" DNA sequence analysis in bacteriophage M13 by the dideoxy chain-termination method. *Proc. Natl. Acad. Sci. USA* **79,** 4298–4302.
21. Legerski, R. J., Hodnett, J. L., and Gray, H. B., Jr. (1978) Extracellular nucleases of *Pseudomonas* BAL 31. III. Use of the double-strand deoxyriboexonuclease activity as the basis of a convenient method for the mapping of fragments of DNA produced by cleavage with restriction enzymes. *Nucleic Acids Res.* **5,** 1445–1464.
22. Hauser, C. R. and Gray, H. B., Jr. (1991) Vectors containing infrequently cleaved restriction sites for use in BAL 31 nuclease-assisted and end-label-mediated analysis of cloned DNA fragments. *Gene Anal. Tech. Appl.* **8,** 139–147.
23. Scott, W. A., Walter, C. F., and Cryer, B. L. (1984) Barriers to nuclease Bal31 digestion across specific sites in simian virus 40 chromatin. *Mol. Cell. Biol.* **4,** 604–610.
24. Zhen, W.-P., Buchardt, O., Nielsen, H., and Nielsen, P. E. (1986) Site specificity of psoralen-DNA interstrand cross-linking determined by nuclease *Bal*31 Digestion. *Biochemistry* **25,** 6598–6603.
25. Zhou, X.-G. (1989) Ethidium bromide-mediated renaturation of denatured closed circular DNAs in alkaline solution: mechanistic aspects and fractionation of closed circular DNA on a molecular weight basis. Some catalytic prop-

erties and mechanism of exonuclease action of BAL 31 nuclease. Ph.D. dissertation, University of Houston.
26. Winston, T. P. and Gray, H. B., Jr., unpublished.
27. Wei, C.-F. (1981) I. Heterogeneity in small genome DNas. II. *Alteromonas* BAL 31 nuclease as a probe for covalent alterations in duplex DNA produced by carcinogenic or mutagenic agents. III. Type I DNA topoisomerases from mammalian cell nuclei interlock strands and promote renaturation of denatured closed circular PM2 DNA. Ph.D. dissertation, University of Houston.
28. Wei, C.-F. and Gray, H. B., Jr., unpublished.
29. Birnboim, H. C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7,** 1513–1523.
30. Sharrocks, A. and Hornby, D. P. (1987) A rapid method for BAL 31 deletion analysis. *Nucleic Acids Res.* **15,** 8564.
31. Padilla, A. and Gray, H. B., Jr., unpublished.
32. Vogt, V. M. (1973) Purification and further properties of single-strand-specific nuclease from *Aspergillus oryzae. Eur. J. Biochem.* **33,** 192–200.
33. Segel, I. H. (1975) *Enzyme Kinetics.* Wiley-Interscience, New York, pp. 46,47.
34. Germond, J. E., Hirt, B., Oudet, P., Gross-Belland, M., and Chambon, P. (1975) Folding of the DNA double helix in chromatin-like structures fom simian virus 40. *Proc. Natl. Acad. Sci. USA* **72,** 1843–1847.
35. Bauer, W. R. (1978) Structure and reactions of closed duplex DNA. *Ann. Rev. Biophys. Bioeng.* **7,** 287–313.
36. Müller, U. R. and Wilson, C. L. (1987) The effect of supercoil and temperature on the recognition of palindromic and non-palindromic regions in ϕX174 replicative form DNA by S1 and *Bal*31. *J. Biol. Chem.* **262,** 3730–3738.

CHAPTER 12

Mung-Bean Nuclease 1 (EC 3.1.30.1)

Andrew J. Sharp and Robert J. Slater

1. Introduction

Mung-bean nuclease 1 was first purified by Sung and Laskowski *(1)* in 1962 from mung-bean sprouts *(Phaseolus aureus)*. It belongs to the class of enzymes EC 3.1.30.1., which has a preference for single-stranded nucleic acid substrates, lacks sugar specificity, and hydrolyzes single-stranded substrates to produce products with 5'-phosphoryl and 3'-hydroxyl termini, ranging from mono- to, at least, heptanucleotides. Although it shows a preference for single-stranded nucleic acids over double-stranded of 30,000-fold *(2)*, used in high concentrations with extended incubation times, mung-bean nuclease 1 will completely degrade double-stranded DNA *(3–5)*. Mung-bean nuclease 1 is also reported to show a separate 3'-ω-monophosphatase activity *(6)* (*see* Section 2.7.). Mung-bean nuclease 1 is a zinc metalloenzyme that requires Zn^{2+} and a reducing agent, such as cysteine, for both activity and stability.

Mung-bean nuclease 1 has been used in the removal of protruding tails in double-stranded DNA, in the excising of cloned DNA fragments inserted into vectors, and in other techniques we describe later in this chapter. We also discuss some of the background enzyme data for those wishing to understand the enzyme in more detail and for those wishing to modify standard protocols to suit their own purposes. Although it is not within the scope of this chapter to try to cover all of the available data regarding specificity and mechanism of cleavage, the information provided here should be adequate for most workers seeking a general understanding.

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

2. Enzyme Data

2.1. Purification

Purification of the nuclease activity from mung-bean sprouts has been reported by several authors. Sung and Laskowski *(1)* published their original paper in 1962, where they described the enzyme as the first peak of a series of five peaks eluting from a DEAE-Sephadex column. The column itself was the final step in the purification process, which gave a 2000-fold purification with a 0.5% yield. This was followed by an improved technique in 1968 by Johnson and Laskowski *(7)*. Ardelt and Laskowski *(2)* used an eight-step process that gave a 25,000-fold purification with a yield of 20% and claimed to have purified the enzyme to homogeneity.

2.2. Molecular Mass

The mol mass of the nuclease is between 35,000 *(8)* and 39,000 Da *(9)* as determined by SDS gel electrophoresis. The enzyme is known to be a glycoprotein that is 29% carbohydrate by weight and is also reported to be composed of two subunits (25,000 and 15,000 Da) linked by a disulfide bond(s) *(4)*.

2.3. pH Optimum

Mung-bean nuclease has a pH optimum of 5.0, although this can be varied (4.8–5.5) depending on reaction conditions or in order to enhance certain characteristics of the enzyme over others. Although predominantly single-strand specific, mung-bean nuclease 1 will degrade duplex DNA to mononucleotides at high concentrations *(4)*. Both the hydrolysis of single-stranded and the hydrolysis of double-stranded DNA have the same pH optima. The preference for single-stranded structures over double-stranded increases as the pH increases. At pH values close to 7.0, the preference for super-coiled DNA over the relaxed form is 30,000-fold. Increasing salt concentrations cause the pH optimum to become more acidic.

2.4. Temperature

Mung-bean nuclease 1 is usually assayed and used at 37°C. Ghangas and Wu *(10)* found that "terminally directed" nuclease activity decreased at 5°C, whereas endonucleolytic "nicking" still occurred. Johnson and

Laskowski *(3)* found that a temperature of 37°C improved the hydrolysis of native biosynthetic $d(A\text{-}T)_n$ by auto-acceleration (interpreted as the production of a more favorable substrate, i.e., single strands, during the course of the experiment).

2.5. Salt Optimum

Mung-bean nuclease 1 is sensitive to ionic strength; 0.025–0.050*M* ammonium acetate or Tris has been found to be optimal, whereas concentrations of 0.2–0.4*M* are reported to give 80–90% inhibition of enzyme activity *(1)*. Salt concentration can affect stability (*see* Section 2.8.) and pH optimum (*see* Section 2.3.).

2.6. Assay

The most common assay for mung-bean nuclease 1 (and that which is used by those commercial companies who supply the enzyme) is based on the ability of the enzyme to form acid-soluble products, i.e., the ability to convert denatured DNA to mono- and oligonucleotides that are not precipitated on addition of TCA, but that do contribute to the hyperchromicity of the supernatant. Another assay, which is based on the absolute requirement of the enzyme for Zn^{2+}, involves the use of electrophoresis. Although it is only qualitative, it does have the advantage of being very sensitive, thus only requiring minimal amounts. Heat-denatured DNA is incubated with the enzyme in the presence of Zn^{2+}, in the absence of Zn^{2+}, and in the absence of Zn^{2+}, but in the presence of 0.001*M* EDTA. These DNA samples are then run on a 0.8% agarose gel (100 ng/well is sufficient) and stained with ethidium bromide. The three lanes on the gel should show: complete breakdown of DNA in lane 1 (optimum conditions for the enzyme); no breakdown in lane two (although some breakdown may occur, since the enzyme may lose only 70–80% of its original activity on removal of exogenous Zn^{2+} *[9]*); and no breakdown in lane three (since 0.001*M* EDTA will remove the zinc metalloportion of the enzyme, thereby inhibiting it irreversibly). This last gel lane is important to determine the complete absence of other DNases, since it is known that simple addition of EDTA to an assay mixture will not necessarily protect it from all DNase attack *(11)*. This assay may be used with serial dilutions of the enzyme to give an indication of specific activity.

2.7. Specificity

Mung-bean nuclease 1 has more than one activity. Its predominant mode of action, and that for which it is most commonly used, is as a single-strand-specific, but sugar nonspecific nuclease that will cleave the 3'-phosphodiester bonds of a single-stranded substrate, giving rise to mono-, di-, and trinucleotides.

Mung-bean nuclease 1 will degrade double-stranded DNA under certain conditions, i.e., under conditions that do not favor a tight helical DNA structure, large duplex polymers of DNA are completely degraded from their termini. The terminally directed single-strand- and double-strand-specific activities of the enzyme are both intrinsic properties of the molecule, the evidence being that (1) they are inactivated and reactivated in parallel and (2) the two activities migrate together on analytical gels *(4)*.

Under conditions that favor the tight helical structure of DNA, the enzyme has been shown to catalyze a number of double-stranded cleavages at a limited number of points. This occurs by a two-step mechanism: first, the creation of a single-stranded nick, then a double-stranded scission *(4)*. Other workers have shown that the sensitivity of duplex DNA toward endonucleolytic scission of one strand increases with increasing superhelical density. This is thought to result from tight bends in the DNA. Such distorted regions become favored sites for single-strand-specific nucleases. The degree of super-coiling is important with respect to nuclease cleavage *(12)*. In addition to terminally directed single-stranded and double-stranded nuclease activity and the limited endonucleolytic activity on duplex DNA under the correct conditions, there is another separate activity that appears to reside on the same molecule and copurifies with it. Mikulski and Laskowski *(6)* reported a 3'-ω-monophosphatase activity that cleaved the 3'-terminal phosphate group of 3'-mononucleotides and that was distinct from the 3'-nuclease activity (i.e., that which cleaves the phosphodiester bond between nucleotides of a polynucleotide chain). Both of these activities will result in an —OH group at the 3' position. This monophosphatase activity hydrolyzes ribose mononucleotides 50–100 times faster than it hydrolyzes deoxyribose mononucleotides. This example of activity toward both phosphomonoester and phosphodiester bonds residing in the same enzyme is not unique. Richardson et al. *(13)* reported similar activity with exonuclease III.

The relative base preference of mung-bean nuclease is A > T(U) > C > G as reported by Mikulski and Laskowski *(6)*, who noted that the ribohomopolymer poly-U was preferentially degraded with respect to poly-A. This is thought to be due to the fact that poly-U lacks a relatively ordered secondary structure. The structure of the DNA at A,T-rich regions has been implicated in the apparent preference of the nuclease for these regions. A good example of this is phage λ DNA, which is preferentially cleaved in the A,T-rich central portion of the molecule *(3)*. Other sites are cleaved more slowly. Johnson and Laskowski *(3)* suggest that regions of DNA rich in A,T residues possess a certain degree of single-stranded character (dependent upon region length, temperature, and ionic strength). This suggestion is based on the work of von Hippel and Felsenfeld *(14)* and Wingert and von Hippel *(15)*, who postulate that, at regions below T_m A,T-rich regions undergo local strand separation to a greater extent than G,C-rich regions, a phenomenon known as "structural breathing." A more recent explanation is that "stable DNA unwinding," not "breathing" accounts for mung-bean nuclease 1 hypersensitivity of specific A,T-rich regions *(16)*. These regions of stable DNA unwinding are of potential biological importance, since they have been detected in replication origins and transcriptional regulatory regions. Kowalski et al. *(16)* found that nuclease hypersensitivity of the A,T-rich regions is hierarchical, and either deletion of the primary site or a sufficient increase in the free energy of super-coiling leads to an enhanced nuclease cleavage at an alternative A,T-rich site.

Vernick et al. *(17)* have suggested the involvement of an altered DNA structure near gene boundaries in determining the recognition sites for this enzyme, since they found no direct relationship between dA-dT-richness and site of cleavage. As yet, no specific recognition sequence has been determined for mung-bean nuclease 1, although there is some evidence to suggest that the enzyme may be "region-specific" *(3)*.

2.8. Stability

Pure mung-bean nuclease is stable at pH 7–8 in buffer, although this is not its optimum pH for enzyme activity. However, Kowalski et al. *(9)* reported that the enzyme can be stabilized at pH 5.0 in the presence of 0.1 m*M* Zn acetate, 1.0 m*M* cysteine, and 0.001% Triton X-100. It is stable to heat treatments up to 65°C (60–70°C). The purified enzyme,

stored in 50% (w/v) glycerol, has been shown to be stable for at least 6 mo at 20°C. Johnson and Laskowski *(3)* found that the enzyme, lyophilized and kept in a deep-freeze (temperature not specified), lost 50% of its specific activity in 1 yr.

Stability is affected by surface adsorbtion of the enzyme to the container and by the composition of the container. Stability was found to improve by storage at higher enzyme concentrations *(9)*.

2.9. Inhibitors

Dialysis of mung-bean nuclease 1 against 0.05*M* sodium acetate (pH 5.0) containing 0.001% Triton X-100 results in 70–80% inhibition of the enzyme. This inhibition is reversible on addition of 0.1 m*M* Zn^{2+} and 1.0 m*M* cysteine with incubation at 23°C for 20–30 min. Co^{2+}, Mg^{2+}, Mn^{2+}, Ca^{2+}, Fe^{2+}, and Cu^{2+} are incapable of reactivating the enzyme under the same conditions. Dialysis against 0.001 *M* EDTA will remove the Zn portion of the metalloenzyme resulting in irreversible inactivation. Adjustment of the pH to 8.0 results in 99.99% inhibition of the activity as compared with pH 5.0. The addition of 0.01% SDS at pH 5.0 will inactivate the enzyme completely *(9)*.

2.10. Uses of Mung-Bean Nuclease 1

Mung-bean nuclease 1 shows differing activities under different reaction conditions, and, as such, it has the potential to be used for a variety of purposes and techniques. One of the main uses of mung-bean nuclease 1 is the conversion of protruding termini to blunt ends *(10,18, 19)* (*see* Fig. 1). Mung-bean nuclease 1 may be the enzyme of choice for this technique over nuclease S1, since it may be less vigorous in its actions with respect to cleaving double-stranded DNA.

Ghangas and Wu *(10)* found the terminally directed or exonucleolytic "trimming" activity to be specifically 5'–3'. Proper trimming was obtained when the final blunt end contained a GC base pair at its terminus. This was not always the case if an AT base pair occupied this position. The nucleolytic composition of the overhang did not seem to affect the efficiency or quality of the digestion.

Mung-bean nuclease 1 has also been used: as a structural probe in the detection of complexes of DNA hairpins *(19)*; in high-resolution mapping of termini in RNA transcripts *(20)*; as a probe for DNA secondary structure *(8)*; as a nuclease for gene isolation *(5,21)*; and for the linearization of super-coiled DNA *(12)*.

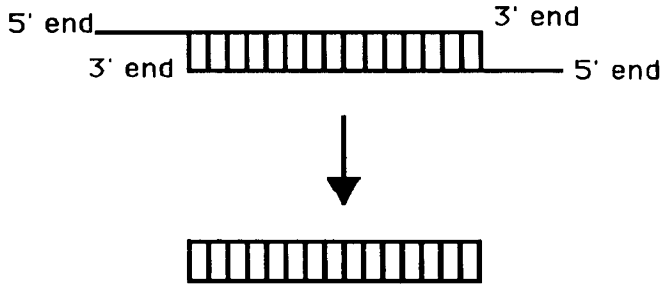


Fig. 1. Conversion of protruding termini to blunt ends.

3. Experimental Procedures for Removal of Protruding Tails from Double-Stranded DNA

This technique is also known as "blunt-ending" and is performed to remove single-stranded "tails" left by digestion with a restriction enzyme, usually prior to cleavage of the DNA with another restriction enzyme. The following protocol is one modified from Maniatis et al. (1989) as part of a larger protocol for engineering a restriction site.

3.1. Materials

1. 10X Mung-bean nuclease buffer:
   - 300 m*M* Sodium acetate (pH 4.5)
   - 500 m*M* Sodium chloride (NaCl)
   - 10 m*M* Zinc chloride ($ZnCl_2$)
   - 50% Glycerol
2. 10 m*M* Tris-HCl (pH 7.5).
3. Mung-bean nuclease 1.5U/μL.

3.2. Methods

For 5 μg of the DNA to be cleaved in 10 μL of 10 m*M* Tris-HCl (pH 7.5), digest with mung-bean nuclease to generate a blunt end as follows:

1. Mix the 10 μL of DNA with:

10X mung-bean nuclease buffer	2 μL
Mung-bean nuclease (1.5 U/μL)	2.5 μL
H_2O	5.5 μL

2. Incubate the reaction for 1 h at 37°C.
3. Extract the reaction once with phenol:chloroform, and transfer the aqueous phase to a fresh microfuge tube.

4. Precipitate the DNA with 2 vol of ethanol at 0°C. Recover the DNA by centrifugation at 12,000*g* for 10 min at 0°C in a microfuge. Remove the supernatant, and wash the pellet with 70% ethanol. Redissolve the DNA in 20 µL of the appropriate buffer (i.e., 1X restriction enzyme buffer if further restrictions are planned).

References

1. Sung, S. C. and Laskowski, M., Sr. (1962) A nuclease from mung bean sprouts. *J. Biol. Chem.* **237,** 506–511.
2. Ardelt, W. and Laskowski, M., Sr. (1971) Mung bean nuclease I. IV. An improved method of preparation. *Biochem. Biophys. Res. Commun.* **44,** 1205–1211.
3. Johnson, P. H. and Laskowski, M., Sr. (1970) Mung bean nuclease I. I. Resistance of double-stranded deoxyribonucleic acid and susceptibility of regions rich in adenosine and thymidine to enzymatic hydrolysis. *J. Biol. Chem.* **245,** 891–898.
4. Kroeker, W. D., Kowalski, D., and Laskowski, M., Sr. (1976) Mung bean nuclease I. Terminally directed hydrolysis of native DNA. *Biochemistry* **15,** 4463–4467.
5. Kroeker, W. D. and Kowalski, D. (1978) Gene-sized pieces produced by digestion of linear duplex DNA with mung bean nuclease. *Biochemistry* **17,** 3236–3243.
6. Mikulski, A. J. and Laskowski, M., Sr. (1970) Mung bean nuclease I. III. Purification procedure and (3')-ω-monophosphatase activity. *J. Biol. Chem.* **245,** 5026–5031.
7. Johnson, P. H. and Laskowski, M., Sr. (1968) Sugar unspecific mung bean nuclease I. *J. Biol. Chem.* **243,** 3421–3424.
8. Martin, S. A., Ullrich, R. C., and Meyer, W. L. (1986) A comparative study of nucleases exhibiting preference for single-stranded nucleic acid. *Biochem. Biophys. Acta.* **867,** 76–80.
9. Kowalski, D., Kroeker, W. D., and Laskowski, M., Sr. (1976) Mung bean nuclease I. Physical, chemical and catalytic properties. *Biochemistry* **15,** 4457–4463.
10. Ghangas, G. S. and Wu, R. (1975) Specific hydrolysis of the cohesive ends of bacteriophage λDNA by three single strand-specific nucleases. *J. Biol. Chem.* **250,** 4601–4606.
11. Jones, M. C. and Boffey, S. A. (1984) Deoxyribonuclease activities of wheat seedlings. *FEBS Lett.* **174,** 215–218.
12. Wang, J. C. (1974) Interactions between twisted DNAs and enzymes: The effects of superhelical turns. *J. Mol. Biol.* **87,** 797–816.
13. Richardson, C. C., Lehmann, I. R., and Kornberg, A. J. (1964) A deoxyribonucleic acid phosphatase–exonuclease from Escherichia coli. *J. Biol. Chem.* **239,** 251–258.
14. von Hippel, P. H. and Felsenfeld, G. (1964) Micrococcal nuclease as a probe of DNA conformation. *Biochemistry* **3,** 27–39.

15. Wingert, L. and Von Hippel, P. H. (1968) The conformation dependent hydroxylysis of DNA by micrococcal nuclease. *Biochem. Biophys. Acta.* **157,** 114–126.
16. Kowalski, D., Natale, D. A., and Eddy, M. J. (1988) Stable DNA unwinding, not "breathing," accounts for single-strand-specific nuclease hypersensitivity of specific A + T-rich sequences. *Proc. Natl. Acad. Sci. USA* **85,** 9464–9468.
17. Vernick, K. D., Imberski, R. B., and McCutchen, T. F. (1988) Mung bean nuclease exhibits a generated gene-excision activity upon *Plasmodium falciporum* genomic DNA. *Nucleic Acids Res.* **16,** 6883–6896.
18. Upcroft, P. and Healy, A. (1987) Rapid and efficient method for cloning of blunt-ended DNA fragments. *Gene* **51,** 69–75.
19. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1989) *Molecular Cloning*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
20. Baumann, U., Lehmann, U., Schwellnus, K., van Boom, J. H., and Kuhn, H. (1987) Complexes of DNA hairpins and a single-stranded oligonucleotide detected by affinity chromatography and mung bean nuclease cleavage. *Eur. J. Biochem.* **170,** 267–272.
21. Murray, M. G. (1986) Use of sodium trichloroacetate and mung bean nuclease to increase sensitivity and precision during transcript mapping. *Anal. Biochem.* **158,** 165–170.

Chapter 13

RNase A (EC 3.1.27.5)

Michael M. Burrell

1. Introduction

The term ribonuclease (RNase) is an imprecise term and is used to cover both enzymes that cause exonucleolytic cleavage and endonucleolytic cleavage of RNA. Exonucleases may cleave the RNA in 3'–5' direction or vice versa, whereas some endoribonucleases have a specific requirement for certain bases. For example the RNase from *Bacillus cereus* cleaves at pyrimidine residues *(1)*. Some enzymes produce 5' phosphomononucleotides, whereas others give rise to 3' phosphomononucleotides. This chapter focuses on the endoribonuclease RNase A (otherwise described as RNase, RNase I, or pancreatic ribonuclease), which shows some base specificity in where it cleaves RNA. The enzyme has been particularly well characterized at the molecular level.

2. The Enzyme

The enzyme was obtained in crystalline form by Kunitz in 1940 *(2)* and the entire amino acid sequence of bovine pancreatic RNase A is now known. It has 124 amino acids and a mol mass of 12,600 Da *(3)*. It is a fairly stable enzyme and contains four disulfide bridges, which occur in all mammalian pancreatic ribonucleases. When the bridges are reductively broken the protein is denatured and becomes inactive. On reoxidation the protein refolds and complete activity is restored *(4)*. It is possible, however, to reduce the bridges only partially and retain enzyme activity. Removal of four peptides at the carboxyl termi-

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

nus destroys enzyme activity. There is a considerable literature on the folding of the RNase A polypeptide and the importance of different residues in this folding *(5)*.

Pancreatic RNases from many species are glycosylated *(6)*. There are four attachment sites for the glycosidic side chains and all involve asparagine residues. It has been suggested that the role of these side chains is to protect the RNase molecule from being absorbed in the gut so that it is transported to the large intestine where it would then be available to hydrolyze the RNA from the cecal microflora. In vivo RNase A protein is first synthesized as a 16,100 Da protein that contains a 25 amino acid extension at the N-terminus, which is removed on secretion. It appears that the presecretory form is also enzymatically active *(7)*.

2.1. The Reaction

RNase A is an endoribonuclease and therefore will cleave RNA to both mononucleotides and oligonucleotides. The enzyme attacks at the 3' phosphate of a pyrimidine nucleotide. Therefore under the appropriate reaction conditions the sequence pG-pG-pC-pA-pG will be cleaved to give pG-pG-pCp and A-pG. A cyclic intermediate is formed (*see* Fig. 1). This specificity for pyrimidine residues should be considered when using RNase A in protocols, such as the one in this chapter, for detecting mismatches. Thus, mismatches involving pyrimidines are likely to be cleaved preferentially to those involving purines unless conditions are adjusted appropriately (*see* Sections 2.3. and 2.4.).

2.2. Assay and Unit Definition

There are several different assays and units of activity used for RNase A. As with all enzyme assays, the purity of the substrate will affect the result. Kunitz *(2)* was the first to crystallize the protein and used an assay that depended on a fixed starting amount of yeast nucleic acid (0.5 mg total phosphate in 1 mL of 0.1*M* acetate buffer at pH 5.0). After 10 min incubation at 25°C the RNA is precipitated with 2 mL of 0.25% uranium acetate and the solution filtered. The filtrate is then assayed for total phosphate. A unit of activity for this assay is defined as the activity that produces 1×10^{-3} mg of soluble phosphorus per mL of digestion mixture.

An alternative assay, which depends on the shift in the absorption spectrum of RNA on digestion with RNase, was introduced by Kunitz in 1946 *(8)*. RNA digestion causes a decrease in the absorption at 300 nm. Kunitz

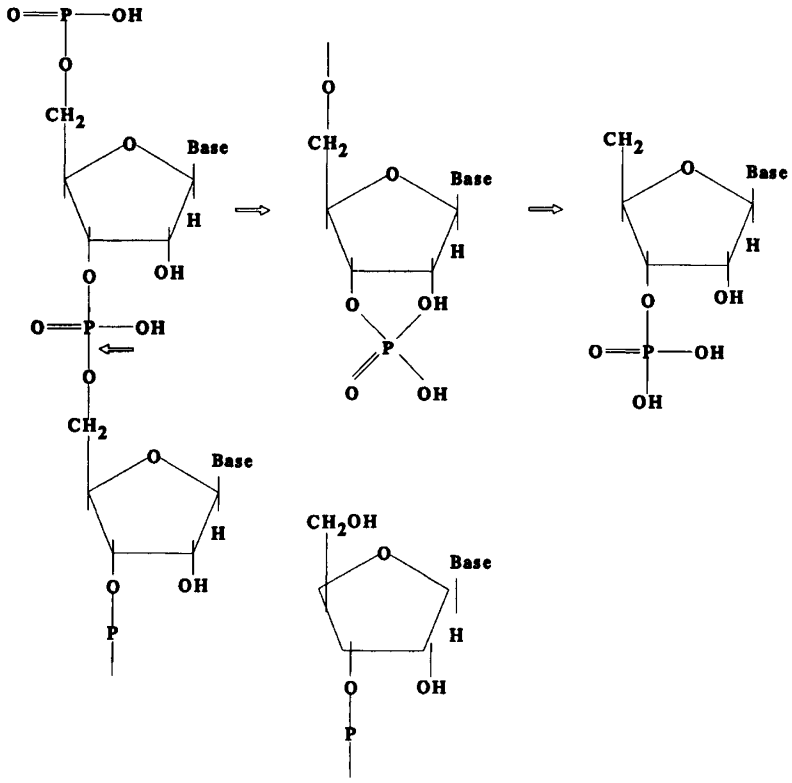


Fig. 1. Cleavage of RNA.

defined a unit of ribonuclease activity as the amount of enzyme capable of causing a decrease of 100%/min in the absorbance at 300 nm of a solution of 0.05% yeast nucleic acid in 0.05*M* acetate buffer at pH 5.0.

Afinsen et al. *(9)* introduced a stopped assay where the activity is measured at a temperature of 25°C and a wavelength of 260 nm with a 0.8% solution of yeast nucleic acid and the undigested RNA is precipitated with 0.75% uranium acetate in 25% perchloric acid. Several other methods of precipitating the undigested RNA have been used such as glacial acetic acid *(2)* or perchlorate alone *(10)*. In addition synthetic polymers are often used in place of extracted RNA as substrate *(11)*.

2.3. Substrate

RNase A has its highest activity with single-stranded RNA. However, if the concentration of RNase exceeds 2.2 Kunitz U/mL and the RNA concentration is of the order of 30 µg/mL or greater, then double-stranded RNA or poly A RNA will be cleaved *(12)*. Thus, at low

concentrations of enzyme, RNA with secondary structures will not be cleaved. If it is important to achieve total RNA hydrolysis either the concentration of enzyme must be high or the ionic conditions adjusted to ensure the RNA is single stranded.

2.4. pH and Ionic Strength

The optimum pH is in the region of 7.6 *(2)*, although other workers have reported different values with different sources of RNA. As indicated earlier, assays are frequently done at pH 5.0 and this is usually because higher pH values can lead to higher background and nonenzymic degradation. The importance of ionic strength is reviewed by Afinsen *(13)* and must be carefully controlled for reproducible results. For natural substrates an ionic strength of 0.1 gives optimal activity. However, 0.35 is optimal for synthetic substrates *(14)*.

At physiological pH and ionic strength, RNase A has very little activity against double-stranded RNA *(12)* or poly A *(15)*. However, at lower ionic strength or as indicated in Section 2.3. at high substrate and enzyme concentrations, these forms of RNA may be degraded *(12,15)*.

2.5. Inhibitors

Many ions, both cations and anions, have been reported to inhibit RNase, but the literature contains many conflicting results. It is clear that the particular conditions of assay can alter the result. Thus Ca^{2+} ions have been reported to stimulate activity *(14)* and to have no effect *(16)*, but the conditions of assay were not the same. In general it would appear that divalent cations can interfere with the assay and may inhibit the reaction. Thus in critical experiments it is probably wise to avoid high concentrations of cation, test for interference under the conditions of use, and include EDTA in the reaction mixture.

The products of digestion, especially pyrimidine nucleotides, may inhibit the reaction *(14)*. Anionic polymers will inhibit at high concentrations probably by combination with the enzyme *(13)*.

Many tissues contain proteins that inhibit RNase. Mammalian placental tissue has commonly been used as a source of this protein. It is an acidic protein with an approximate mol mass of 50,000 Da and forms a 1:1 complex with RNase A. The definition of a unit of inhibitor is the amount that inhibits 5 ng of RNase A by 50%. Reagents that react with free thiol groups will inactivate the inhibitor.

Other mechanisms of inhibiting RNase activity during RNA extraction include the use of sodium dodecyl sulfate, diethyl pyrocarbonate, and 4*M* guanidinium thiocyanate plus 0.1*M* mercaptoethanol *(5)*. Oxovanadium (VO^{2+}) forms a stable complex with nucleotide monophosphates, which competitively inhibits RNase A, but the complex also prevents RNA translation *(17)*. Spermine at 0.13 m*M* will inhibit tRNA hydrolysis by 50% *(18)*. It inhibits the rate of the reaction and does not affect the K_m.

RNase A is a common contaminant of DNase 1 preparations and can be removed by alkylation of a histidine residue at the active site. The procedure was developed by Zimmerman *(19)*. Practical protocols have been discussed by Gurney and Gurney *(20)*.

2.6. Stability

RNase is probably one of the most stable enzymes the molecular biologist will use. It is quite stable over a wide range of pH values below 25°C. At 100°C the enzyme is most stable between pH 2.0 and pH 4.5. For example heating for 30 min at 100°C and pH 3.5 will only destroy 21% of the activity whereas 5 min at pH 9.0 will destroy more than 90% *(2)*.

UV light (254 nm) at pH 5.0 will inactivate the enzyme probably as a result of destruction of the disulfide bridges in the protein *(21)*. Chemical destruction of these disulfide bridges will also destroy activity. If the disulfide bridges are left intact by the method of denaturation then the protein will refold quite rapidly. Enzyme activity can still be restored if the disulfide bridges are reoxidized, but air oxidation requires several hours *(5)*.

3. Experimental Procedures

RNase A can be used in combination with other RNases to characterize RNA directly and provide sequence information. A common use of the enzyme is to test for complimentarity between RNA:DNA hybrids either to remove large unhybridized regions or, under appropriate reaction conditions, to identify single base differences *(22)*. In the latter case, an RNA fragment labeled with ^{32}P is prepared from a DNA template of interest. This is then hybridized to test DNA in solution and the resulting hybrid treated with RNase A. Where single base pair mismatches occur the RNA is cleaved. Therefore, upon electrophoresis, more than one labeled RNA molecule is detected.

From the preceding discussion in this chapter, it is clearly important that, to obtain reproducible results, the purity and concentration of the nucleic acid used must be consistent, the supply and concentration of RNase A standardized, the pH and the ionic strength standardized.

3.1. Materials for Mismatch Analysis

1. ^{32}P-labeled RNA prepared for example with SP6 RNA polymerase *(23)*. Approximately 3 ng of probe RNA containing 1×10^{-5} cpm is required per hybridization.
2. RNase A at a concentration of 2 mg/mL in sterile distilled water and boiled for 10 min (so that it is DNase free).
3. Deionized formamide.
4. Hybridization buffer:
   - 80% Formamide
   - 40 m*M* PIPES (pH 6.4)
   - 0.4 m*M* NaCl
   - 1 m*M* EDTA
5. RNase buffer:
   - 10 m*M* Tris-HCl (pH 7.5)
   - 1 m*M* EDTA
   - 200 m*M* NaCl
   - 100 m*M* LiCl
6. 20% SDS.
7. Proteinase K 10 mg/mL.
8. Phenol:Chloroform (1:1) containing 4% isoamyl alcohol, 0.01% hydroxyquinoline.
9. Carrier transfer RNA (e.g., yeast, wheat).
10. 4*M* Ammonium acetate.

3.2. Methods for Mismatch Analysis

1. Suspend 20–50 ng of cloned test DNA or 3–6 µg of total genomic test DNA in 30 µL of hybridization buffer.
2. Add 0.5 µL of probe RNA, mix the solutions, and centrifuge the droplets to the bottom of the tube.
3. Heat to 90°C for 10 min.
4. Anneal at 45°C for 30 min.
5. Add 350 µL of RNase buffer containing RNase A at 40 µg/mL.
6. Mix by vortexing and centrifuge droplets to bottom of tube.
7. Incubate at 25°C for 30 min.
8. Add 10 µL of 20% SDS and 10 µL of proteinase K (10 mg/mL).
9. Mix and centrifuge droplets to bottom of the tube.

10. Incubate at 37°C for 15–30 min.
11. Extract protein with an equal volume phenol:chloroform solution.
12. Retain aqueous supernatant and add 10–20 µg of carrier tRNA.
13. Precipitate the nucleic acids 2.5 vol with ethanol and 1/40 vol 4*M* ammonium acetate.
14. Analyze labeled fragments by electrophoresis with a suitable polyacrylamide gel to detect the expected size range of fragments.

References

1. Lockard, R. E., Alzner-DeWeerd, B., Heckman, J. E., MacGee, J., Tabor, M. W., and Raj-Bhandary U. L. (1978) Sequence analysis of 5' [^{32}P] labelled mRNA and tRNA using polyacrylamide gel electrophoresis. *Nucleic Acid Res.* **5,** 37–56.
2. Kunitz, M. (1940) Crystalline ribonuclease. *J. Gen. Physiol.* **24,** 15–32.
3. Smyth, D. G., Stein, W. H. and Moore, S. (1963) The sequence of amino acid residues in bovine pancreatic ribonuclease; revisions and confirmations. *J. Biol. Chem.* **238,** 227–234.
4. Afinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* **181,** 223–230
5. Blackburn, P. and Moore, S. (1982) Pancreatic ribonuclease, in *The Enzymes,* vol XV (Boyer, P. D., ed.), Academic, New York, pp. 317–433.
6. Bentiema, J. J., Gaastra, W., Scheffer, A. J., and Welling, G. W. (1976) Carbohydrate in pancreatic ribonucleases. *Eur. J. Biochem.* **63,** 441–448.
7. Haugen, T. H. and Heath, E. C. (1979) De novo biosynthesis of an enzymatically active precursor form of bovine pancreatic RNase. *Proc. Natl. Acad. Sci. USA* **76,** 2689–2693.
8. Kunitz, M. (1946) A spectrophotometric method for the measurement of ribonuclease activity. *J. Biol. Chem.* **164,** 563–568.
9. Afinsen, C. B., Redfield, R. R., Choate, W. L., Page, J., and Carrol, W. R. (1953) Studies on the gross structure, cross linkages and terminal sequences in ribonuclease. *J. Biol. Chem.* **207,** 201–210.
10. Schneider, W. C. and Hogeboom, G. H., (1952) Intracellular distributions of enzymes. *J. Biol. Chem.* **198,** 155–163.
11. Markham, R. and Strominger, J. L. (1956) The action of leaf ribonuclease. *Biochem. J.* **64,** 46p.
12. Libonati, M. (1971) Degradation of poly A and double stranded RNA by aggregates of pancreatic ribonuclease. *Biochem. Biophys. Acta* **228,** 440–445.
13. Anfinsen, C. B. and White, F. H., Jr. (1961) The ribonucleases: Occurrence, structure and properties, in *The Enzymes,* vol.VII, 2nd ed. (Boyer, P., Lardy, H., and Myrback, K., eds.), Academic, New York, pp. 95–123.
14. Davis, F. F. and Allen F. W. (1955) The action of ribonuclease on synthetic substrates. *J. Biol. Chem.* **217,** 13–21.
15. Beers, R. F. (1960) Hydrolysis of polyadenylic acid by pancreatic ribonuclease. *J. Biol. Chem.* **235,** 2393–2398.

16. Bain, J. A. and Rusch, H. P. (1944) A method for the assay of in biological material. *J. Biol. Chem.* **153,** 659–667.
17. Berger, S. L. and Birkenmeier, C. S. (1979) Inhibition of intractable nucleases with ribonucleoside vanadyl complexes: isolation of messenger ribonucleic acid from resting lymphocytes. *Biochemistry* **18,** 5143–5149.
18. Holbrook, D. J., Whichard, L. P., and Washington, M. E. (1975) Effect of nucleic acid binding compounds on the hydrolytic activity of various ribonucleases. *Eur. J. Biochem.* **60,** 317–324.
19. Zimmerman, S. B. and Sandeen, G. (1966) The ribonuclease activity of crystallized pancreatic deoxyribonuclease. *Anal. Biochem.* **14,** 269–277.
20. Gurney, T. and Gurney, E. G. (1984) Preparation of "RNase"-free DNase by alkylation in *Methods in Molecular Biology, vol. 2: Nucleic Acids* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 13–20.
21. Rathinasamy, T. K. and Augenstein, L. G. (1968) Photochemical yields in ribonuclease and oxidized glutathione irradiated at different wavelengths in the ultraviolet. *Biophys. J.* **8,** 1275–1287.
22. Myers, R. M., Larin, Z., and Maniatis, T. (1985) Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science* **230,** 1242–1246.
23. Melton, D. A., Kreig, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. (1984) Efficient in-vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acid Res.* **12,** 7035–7056.

CHAPTER 14

Pronase (EC 3.4.24.4)

Patricia J. Sweeney and John M. Walker

1. Introduction

Pronase is the name given to a group of proteolytic enzymes that are produced in the culture supernatant of *Streptomyces griseus* K-1 *(1–3)*. Pronase is known to contain at least ten proteolytic components: five serine-type proteases, two Zn^{2+} endopeptidases, two Zn^{2+}-leucine aminopeptidases, and one Zn^{2+} carboxypeptidase *(4,5)*. Pronase therefore has very broad specificity, and is used in cases where extensive or complete degradation of protein is required. It has been used, for example, to reveal the protein components of cell organelles by the hydrolysis of tissue slices *(6)*, and as an alternative to proteinase K to remove protein during plasmid DNA *(7)*, chromosomal DNA *(8)*, and RNA isolation *(9–11)*. Another use of pronase is the production of a protein hydrolysate suitable for amino acid analysis *(12,13)*.

Traditionally, protein hydrolysates for amino acid analysis are produced by hydrolysis in 6*N* HCl. However, this method has the disadvantage that tryptophan is totally destroyed, serine and threonine partially (5–10%) destroyed, and most importantly, asparagine and glutamine are hydrolyzed to the corresponding acids. Digestion of the protein/peptide with pronase overcomes these problems, and is particularly useful when the concentration of asparagine and glutamine is required. The use of pronase for this application is detailed in Section 3. The use of pronase in the preparation of optically active amino acids has been described. The hydrolysis of amino amides by pronase results in high-rate significant enantioselectivity and a high degree of conver-

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

sion of substrates *(5)*. Pronase has also been used to study membrane topology and protein membrane translocation in the same manner as described for proteinase K in Chapter 14 *(14,15)*, and has been used to disperse mast cells prior to purification and characterization of these cells *(16)*.

2. Enzyme Data

2.1. Specificity

Because it is a mixture of exo- and endoproteases, pronase has a broad specificity, cleaving nearly all peptide bonds *(17)*.

2.2. Purification

The purification of individual proteolytic components of pronase has been reported by a number of authors *(18–21)*. In 1988, the purification and characterization of an additional amino acid-specific endopeptidase from Pronase were described by Yoshida et al. *(22)*.

2.3. Molecular Mass

Molecular masses of 15,000–27,000 Da have been reported for components of pronase. These weights have usually been determined by gel filtration *(18,20–22)*.

2.4. pH Optimum

Pronase has optimal activity at pH 7–8. However, individual components are reported to retain activity over a much wider pH range *(20–23)*. The neutral components are stable in the pH range 5.0–9.0, in the presence of calcium, and have optimal activity at pH 7.0–8.0. The alkaline components are stable in the pH range 3.0–9.0, in the presence of calcium, and have optimal activity at pH 9.0–10.0 *(4)*. The aminopeptidase and carboxypeptidase components are stable at pH 5.0–8.0 in the presence of calcium *(23)*.

2.5. Assay

The assay is normally based on the hydrolysis of casein, which is followed spectrophotometrically by reaction with Folin and Ciocalteau's reagent. One milliliter of enzyme solution is suitably diluted in 0.067*M* phosphate buffer, pH 7.4. A 2% casein solution, adjusted to pH 7.4 with NaOH, is also prepared. One milliliter of the 2% casein solution is added to the enzyme solution, and the mixture

incubated at 40°C. After 10 min, 2 mL of protein precipitation agent are added (1.63 g trichloroacetic acid, 0.82 g sodium acetate, 0.57 mL glacial acetic acid made up to 100 mL with water). The mixture is incubated for a further 10 min at 40°C, and then centrifuged for 5 min at 3000 rpm. Five milliliters of 0.4*M* sodium carbonate solution and 1 mL Folin and Ciocalteau's reagent (diluted 1:5 just prior to use) are added to 1 mL of supernatant. The mixture is then incubated for a further 10 min at 40°C, and the absorbance at 660 nm then measured. One PUK (Proteolytic Units of Kaken; "Kaken" is a Japanese company) unit is defined as the amount of enzyme that produces an absorbance of 1.0 at 660 nm. Using this assay, the enzyme is usually supplied with a specific activity of about 70,000 PUK/g.

Alternatively, the enzyme is supplied with a specific activity of about 7000 PU/g where one PU (proteolytic unit) is the enzyme activity that liberates Folin-positive amino acids and peptides corresponding to 1 μmol of tyrosine within 1 min under the aforementioned assay conditions. One PUK corresponds to 50 PU. Because of its broad specificity, the enzyme is capable of hydrolyzing many peptide, amide, and ester bonds, including the majority of the specific substrates of most proteolytic enzymes. Pronase also hydrolyzes peptides involving D amino acid residues *(24)*.

2.6. Stability

The calcium ion dependence for the stability of some of the components (mainly exopeptidases) was one of the earliest observations made of Pronase *(2)*. Pronase is therefore normally used in the presence of 5–20 m*M* calcium. The addition of excess EDTA results in the irreversible loss of 70% of proteolytic activity *(18)*. Two peptidase components are inactivated by EDTA, but activity is restored by the addition of Co^{2+} or Ca^{2+}. One of these components, the leucine aminopeptidase, is heat stable up to 70°C. All other components of Pronase lose 90% of their activity at this temperature *(5)*. The leucine aminopeptidase is not inactivated by 9*M* urea, but is labile on dialysis against distilled water *(2)*. Some of the other components of Pronase are also reported to be stable in 8*M* urea *(2)*, and one of the serine proteases retains activity in 6*M* guanidinium chloride *(24)*. Pronase retains activity in 1% SDS (w/v) and 1% Triton (w/v) *(15)*. The enzyme is stable at 4°C for at least 6 mo and is usually stored as a stock solution of 5–20 mg/mL in water at –20°C.

2.7. Inhibitors

Among the alkaline proteases, there are at least three that are inhibited by diisopropyl phosphofluoridate (DFP) *(18)*. In general, the neutral proteinases are inhibited by EDTA, and the alkaline proteinases are inhibited by DFP *(4)*. No single enzyme inhibitor will inhibit all the proteolytic activity in a Pronase sample.

3. Experimental Procedures

3.1. DNA Isolation

When used in DNA isolation, Pronase is generally prepared as a stock solution at about 5–20 mg/mL in water. Prior to storage at –20°C, the solution is heated to 56°C for 15 min followed by a 1-h incubation at 37°C. The purpose of this step is to encourage self-digestion. This eliminates contamination with DNases and RNases. For use, the enzyme is added to the DNA sample (in the presence of 0.5–1% SDS to disrupt DNA–protein interactions) typically at 250–500 µg protein/mL, 37°C, for 1–4 h.

3.2. Protein Hydrolysis

To prepare a protein hydrolysate, dissolve 0.2 µmol of protein in 0.2 mL of 0.05*M* ammonium bicarbonate buffer, pH 8.0 (or 0.2*M* sodium phosphate pH 7.0 if ammonia interferes with the amino acid analysis). Add Pronase to 1% (w/w), and incubate at 37°C for 24 h. To achieve complete hydrolysis, it is usually necessary to make a further addition of aminopeptidase M (*see* Chapter 18) at 4% (w/w), and incubate at 37°C for a further 18 h. The sample is then lyophilized and subjected to amino acid analysis. When using two enzymes in this way, there is often an increase in the background amino acids owing to hydrolysis of each enzyme. It is therefore important to carry out a digestion blank to correct for these background amino acids *(25)*.

References

1. Hiramatsu, A. and Ouchi, T. (1963) On the proteolytic enzymes from the commercial protease preparation of *Streptomyces griseus* (Pronase P). *J. Biochem. (Tokyo)* **54(4),** 462–464.
2. Narahashi, Y. and Yanagita, M. (1967) Studies on proteolytic enzymes (Pronase) of *Streptomyces griseus* K-1. Nature and properties of the proteolytic enzyme system. *J. Biochem. (Tokyo)* **62(6),** 633–641.

3. Wählby, S. and Engström, L. (1968) Studies on *Streptomyces griseus* protease. Amino acid sequence around the reactive serine residue of DFP-sensitive components with esterase activity. *Biochim. Biophys. Acta* **151,** 402–408.
4. Narahashi, Y., Shibuya, K., and Yanagita, M. (1968) Studies on proteolytic enzymes (Pronase) of *Streptomyces griseus* K-1. Separation of exo- and endopeptidases of Pronase. *J. Biochem.* **64(4),** 427–437.
5. Yamskov, I. A., Tichonova, T. V., and Davankov, V. A. (1986) Pronase-catalyzed hydrolysis of amino acid amides. *Enzyme Microb.Technol.* **8,** 241–244.
6. Bouteille, M., Dupuy-Coin, A. M., and Moyne, G. (1975) Techniques of localization of proteins and nucleoproteins in the cell nucleus by high resolution autoradiography and cytochemistry. *Methods Enzymol.* **40,** 3–41.
7. Gartland, K. M. A. (1988) Large-scale isolation of Ti plasmid DNA, in *Methods in Molecular Biology, vol. 4: New Nucleic Acid Techniques* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 465–479.
8. Pollard, J. W., Luqmani, Y., Bateson, A., and Chotai, K. (1984) DNA transformation of mammalian cells, in *Methods in Molecular Biology, vol. 2: Nucleic Acids* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 321–332.
9. Doerfler, W. (1969) Non-productive infection of baby hamster kidney cells (BHK 21) with adenovirus type 12. *Virology* **38(4),** 587–606.
10. Joklik, W. K. (1981) Procedures for studying transcription and translation of viral and host nucleic acids in interferon-treated cells. *Methods Enzymol.* **79,** 307–330.
11. De Maeyer-Guignard, J., De Maeyer, E., and Montagnier, L. (1972) Interferon messenger RNA: Translation in heterologous cells. *Proc. Natl. Acad. Sci. USA* **69,** 1203–1207.
12. Hediger, H., Stevens, R. L., Bradenberger, H., and Schmid, K. (1973) Determination of asparagine, glutamine and pyrrolidonecarboxylic acid in total enzymic hydrolysates of peptides and glycopeptides by gas-liquid chromatography. *Biochem. J.* **133,** 551–561.
13. Nomoto, M., Narahashi, Y. and Murakami, M. (1960) A proteolytic enzyme of *Streptomyces griseus*. Hydrolysis of protein by *Streptomyces griseus* protease. *J. Biochem.* **48(4),** 593–602.
14. Cole, S. T., Chen-Schmeisser, U., Hindennach, I., and Henning, U. (1983) Apparent bacterophage-binding region of an *Escherichia coli* K-12 outer membrane protein. *J. Bacteriol.* **153(2),** 581–587.
15. Chang, C. N., Model, P., and Blobel, G. (1979) Membrane biogenesis: cotranslational integration of the bacteriophage F1 coat protein into an *Escherichia coli* membrane fraction. *Proc. Natl. Acad. Sci. USA* **76(3),** 1251–1255.
16. Fox, C. C., Duorak, A. M., Peters, S. P., Kagey-Sobotka, A., and Lichtenstein, L. M. (1985) Isolation and characterization of human intestinal mucosal mast cells. *J. Immunol.* **135(1),** 483–491.
17. Nomoto, M., Narahashi, Y., and Murakami, M. (1960) A proteolytic enzyme of *Streptomyces griseus*. Substrate specificity of *Streptomyces griseus* protease. *J. Biochem. (Tokyo)* **48(6),** 906–915.

18. Awad, W. M., Soto, A. R., Siegei, S., Skiba, W. E., Bernstrom, G. G., and Ochoa, M. S. (1972) The proteolytic enzymes ofthe K-1 strain of *Streptomyces griseus* obtained from a commercial preparation (Pronase). Purification of four serine endopeptidases. *J. Biol. Chem.* **247,** 4144–4154.
19. Wählby, S. (1968) Studies on *Streptomyces griseus* protease. Separation of DFP-reacting enzymes and purification of one of the enzymes. *Biochim. Biophys. Acta* **151,** 394–401.
20. Gertler, A. and Trop, M. (1971) The elastase-like enzymes from *Streptomyces griseus* (Pronase). Isolation and partial characterization. *Eur. J. Biochem.* **19,** 90–96.
21. Wählby, S. (1969) Studies on *Streptomyces griseus* protease. Purification of two DFP-reactin enzymes. *Biochim. Biophys. Acta* **185,** 178–185.
22. Yoshida, N., Tsuruyama, S., Nagata, K., Hirayama, K., Noda, K., and Makisumi, S. (1988) Purification and characterization of an acidic amino acid specific endopeptidase of *Streptomyces griseus* obtained from a commercial preparation (Pronase). *J. Biochem.* **104,** 451–456.
23. Narahashi, Y. (1970) Pronase. *Methods Enzymol.* **19,** 651–664.
24. Siegel, S., Brady, A. H., and Awad, W. M. (1972) Proteolytic enzymes of the K-1 strain of *Streptomyces griseus* obtained from a commercial preparation (Pronase). Activity of a serine enzyme in 6M guanidinium chloride. *J. Biol. Chem.* **247,** 4155–4159.
25. Garner, M. H., Garner, W. H., and Gurd, F. R. N. (1974) Recognition of primary sequence variations among sperm whale myolobin components with successive proteolysis procedures. *J. Biol. Chem.* **249,** 1513–1518.

CHAPTER 15

Proteolytic Enzymes for Peptide Production

Patricia J. Sweeney and John M. Walker

1. Introduction

There are three main reasons why a protein chemist might wish to cleave a protein of interest into peptide fragments. The first reason is to generate, by *extensive* proteolysis, a large number of relatively small (5–20 residues) peptides either for peptide mapping (*see* vol. 1, Chapter 5) or for purification and subsequent manual sequence determination by the dansyl-Edman method (*see* vol. 1, Chapter 24). The second reason is to generate relatively large peptides (50–150 residues) by *limited* proteolysis for automated sequence analysis, such as with the gas-phase sequencer. The third reason is to prepare, again by limited proteolysis, specific fragments for studies relating structure to function. In each case, the specificity of the enzyme used to generate the peptides is a prime consideration, since the aim is to provide high yields of discrete fragments. It can be appreciated that significantly <100% cleavage at some or all of the cleavage sites on the protein being digested will generate a far more complex mixture of a larger number of polypeptides, each in relatively low yield. It is for this reason that enzymes of high specificity, such as trypsin, which cleaves at the C-terminal side of arginine and lysine residues, are mainly used for peptide production. However, other proteases with considerably less specificity have also found use in peptide production, particularly when limited proteolysis is being used, or where native protein is used as the substrate when only a limited number of susceptible peptide

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

Table 1
Major Cleavage Sites of Some Proteolytic Enzymes

Enzymes	Cleavage site
Chymotrypsin	C-terminal to hydrophobic residues, e.g., Phe, Tyr, Trp, Leu
Clostripain	C-terminal to Arg residues
Elastase	C-terminal to amino acids with small hydrophobic side chains
Endoproteinase Arg-C	C-terminal to Arg residues
Endoproteinase Asp-N	N-terminal to Asp and Cys
Endoproteinase Glu-C	C-terminal to Asp and Glu
Endoproteinase Lys-C	C-terminal to Lys
Pepsin	Broad specificity; preference for cleavage C-terminal to Phe, Leu, and Glu
Thermolysin	N-terminal to amino acids with bulky hydrophobic side-chains, e.g., Ileu, Leu, Val, and Phe
Trypsin	C-terminal to Lys and Arg

bonds are available to the proteases (i.e., those on the surface of the protein). This chapter describes ten proteolytic enzymes that have found extensive use in peptide production in recent years. These are chymotrypsin, clostripain, elastase, endoproteinase Arg-C (submaxillary protease), endoproteinase Asp-N, endoproteinase Glu-C (V8 protease), endoproteinase Lys-C, pepsin, thermolysin, and trypsin. The specificities of these enzymes are summarized in Table 1.

2. Enzyme Data

2.1. Chymotrypsin (EC 3.4.21.1)

2.1.1. General Information

The enzyme is initially synthesized in the pancreas as the inactive zymogen chymotrypsinogen. Chymotrypsinogen is converted to the active enzyme by cleavage by other proteolytic enzymes when pancreatic juice passes into the small intestine. Tryptic cleavage of the bond between Arg 15 and Ileu 16 results in a fully active enzyme (π-chymotrypsin). Further chymotryptic cleavage between Leu 13 and Ser 14 liberates the dipeptide Ser 14-Arg 15, and results in the formation of δ-chymotrypsin *(1)*. Further chymotryptic cleavage liberates the dipeptide Thr 147-Asn 148 resulting in the commonly used and commercially available form of the enzyme, α-chymotrypsin *(2)*. Chymotrypsin is known

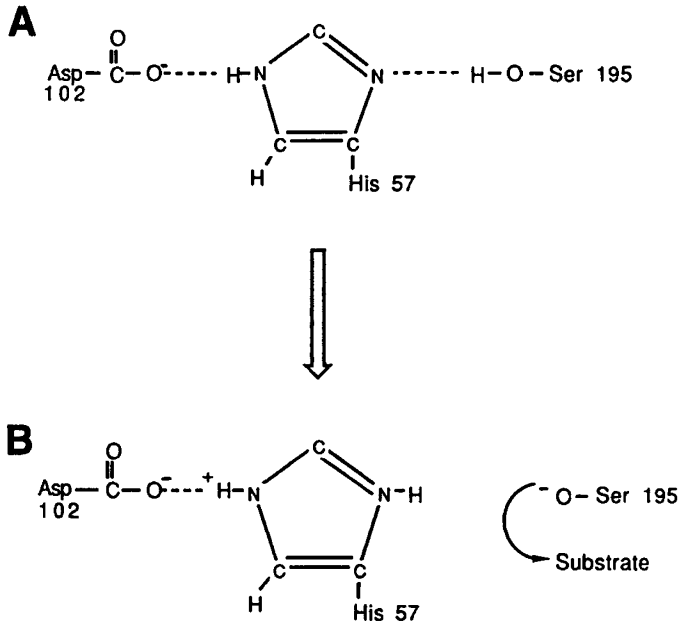


Fig. 1. The catalytic triad at the active-site of chymotrypsin. (**A**) Enzyme alone. (**B**) On addition of substrate, a proton is transferred from serine 195 to histidine 57, resulting in the formation of a strongly nucleophilic oxyanion on Ser 195. The positively charged imidazole ring generated in this process is stabilized by electrostatic interaction with the negatively charged Asp 102. The oxyanion carries out nucleophilic attack on the peptide bond.

as a serine protease, since its catalytic activity results from a strongly ionized serine residue that is generated at the active site as a consequence of the microenvironment of serine residue 195 (*see* Fig. 1). This strongly nucleophilic oxyanion catalyzes the cleavage of peptide bonds. Many inhibitors of serine proteases (such as diisopropylphosphofluoridate [DFP] and phenylmethane sulfonyl fluoride [PMSF]) function by reacting with the active site serine residue. The commercially available source of the enzyme is from bovine pancreas *(3–5)*.

2.1.2. Specificity

Chymotrypsin has relatively broad specificity, cleaving on the C-terminal side of hydrophobic residues, especially phenylalanine, tyrosine, tryptophan, and leucine. However, some significant cleavages have

also been reported C terminal to Met, Ileu, Ser, Thr, Val, His, Gly, and Ala residues. The nature of neighboring residues can affect the rate of cleavage of a particular bond. Irrespective of the nature of *X*, *X*—Pro bonds are not cleaved *(6)*.

2.1.3. Molecular Mass

α-Chymotrypsin (chymotrypsin A) has a mol mass of about 25,000 Da and contains 241 amino acid residues. The molecule has three polypeptide chains (A chain—13 residues; B chain—131 residues; and C chain—97 residues) linked by disulfide bridges. The amino acid sequence is known *(7)*.

2.1.4. pH Optimum

The enzyme has a pH optimum between 7.5 and 8.5 *(2)*.

2.1.5. Assay

The assay is based on measuring the esterolytic activity of the enzyme. The decrease in absorbance caused by the hydrolysis of the ester linkage of *N*-acetyl-L-tyrosine ethyl ester (ATEE) is measured spectrophotometrically at 237 nm.

ATEE (2.8 mL of 2 m*M* in 0.1*M* phosphate buffer, pH 7.0) is incubated at 25°C. At time zero, 20 µL of enzyme (diluted in 1 m*M* HCl) are added. The change in absorbance is monitored with time at 237 nm. The enzyme is usually supplied with a specific activity of 9000–11,000 ATEE U/mg when this assay is used. One ATEE unit is defined as the amount of enzyme that causes a decrease in absorbance of 0.001/min *(8,9)*.

Alternatively, an assay based on hydrolysis of Suc-$(Ala)_2$-Pro-Phe-4-nitroanilide can be used. The reaction is monitored at 410 nm. Using this assay, the enzyme is usually supplied with a specific activity of 60–70 U/mg, where 1 U is defined as the amount of enzyme that hydrolyzes 1 µmol of substrate/min at pH 7.0, and 25°C.

2.1.6. Stability

The enzyme functions in the presence of 2*M* guanidine hydrochloride *(10)* and 0.1% SDS. Stock solutions (10 mg/mL) in 1 m*M* HCl can be stored frozen or for a number of days at 4°C. When diluted for use to 1 mg/mL in 1 m*M* HCl (since autolysis occurs at high pH), the presence of 2 m*M* Ca^{2+} helps stabilize (or possibly activate) the enzyme. The lyophilized enzyme is stable for years at 4°C.

Chymotrypsin is a basic protein, and hence, it sticks easily to glass. Therefore, dilute solutions of the enzyme should be avoided *(2)*.

2.1.7. Inhibitors

The most commonly used serine protease inhibitors are DFP and PMSF (Section 2.1.1.). PMSF is preferred, since it is much less toxic than DFP. PMSF is prepared as a stock solution (1*M*) in a *dry* solvent, such as propan-2-*ol* or methanol (stable for months at 4°C). It is added to the enzyme solution with vigorous stirring to a final concentration of 1 m*M*. It has a half-life of about 1 h in aqueous solution. DFP (highly toxic!) is prepared as a 500 m*M* stock solution in propan-2-*ol* and added to a final concentration of 0.1 m*M*. It also has a half-life of 1 h in aqueous solution. Other low-mol-mass inhibitors include chymostatin (mol mass 605 Da), a microbial amino acid aldehyde, active at 10–100 μ*M* *(11)*, and tosyl phenylalanyl chloromethyl ketone (TPCK) *(12)*, active at 10–100 μ*M*. TPCK reacts with histidine 57 involved in the catalytic triad shown in Fig. 1. High-mol-mass inhibitors include aprotinin, α_1 antitrypsin, and α_2 macroglobulin, and they are all effective at equimolar concentrations with chymotrypsin. The enzyme is also completely inhibited in the presence of 10 m*M* Cu^{2+} and Hg^{2+}, and partially inhibited by 1 m*M* Zn^{2+}. Chymotrypsin that has been completely inhibited by Cu^{2+} can be reactivated by the addition of Ca^{2+} or Versene *(13)*.

2.2. Clostripain (EC 3.4.22.8)

2.2.1. General Information

Clostripain (clostridiopeptidase B) is a sulfhydryl protease isolated from culture filtrates of *Clostridium histolyticum (14,15)*. Sulfhydryl proteases are all characterized by a thiol group at the active site that is essential for proteolytic activity.

2.2.2. Specificity

The enzyme has amidase and esterase activity as well as protease activity. The enzyme primarily cleaves proteins at the C-terminal side of arginine residues, i.e., —Arg—X—bonds, although —Lys—X— bonds are also cleaved, but at a much slower rate *(14)*. Thus, under conditions of controlled hydrolysis, the enzyme may be used to cause cleavage effectively at arginine residues only *(16,17)*.

2.2.3. Molecular Mass

The enzyme has a mol mass of approx 50,000 Da. The amino acid composition has been published *(14)*.

2.2.4. pH Optimum

The enzyme has a pH optimum in the range pH 7.4–7.8 in phosphate or Tris buffer *(14)*.

2.2.5. Assay

The assay is based on the hydrolysis of benzoyl arginine ethyl ester (BAEE). The substrate solution is 0.05*M* sodium phosphate buffer, pH 7.8, 0.25 m*M* BAEE, and 2.5 m*M* dithiothreitol, at 25°C. The reaction is followed by monitoring the change in absorbance at 253 nm with time after addition of enzyme. A molar extinction coefficient of $1150 M^{-1} \cdot cm^{-1}$ is used to express the results as μmol BAEE hydrolyzed/min at 25°C *(18)*. Using this assay, the enzyme is usually supplied with a specific activity of 50–250 U/mg protein, where 1 U is defined as the amount of enzyme that hydrolyzes 1 μmol of BAEE/min at pH 7. 6 at 25°C.

2.2.6. Stability

Enzymic activity is rapidly lost on incubation in the presence of oxidizing agents because of oxidation of the reactive thiol group, but can be regenerated by treatment with dithiothreitol *(14)*. The enzyme is active in 6*M* urea *(19)*.

2.2.7. Inhibitors

The enzyme is inhibited by any thiol reactive agent (e.g., iodoacetic acid), by PMSF (*see* Section 2.1.7. for practical details), and by parachloromercuribenzoate *(20,21)*. The PMSF inhibition can be reversed by the addition of thiol reagents, such as DTT or mercaptoethanol. Since the reactive thiol group is easily oxidized, the enzyme is usually used in the presence of a reducing agent, such as dithiothreitol (1–2 m*M*). Some authors report that binding of calcium is an essential prerequisite of activity *(20)*, and certain suppliers indicate that the enzyme requires activation prior to use, by preincubation in 1 m*M* calcium acetate, 2.5 m*M* DTT for 2–4 h at 25°C. The esterase activity is inhibited in the presence of Co^{2+}, Cu^{2+}, Cd^{2+}, Na^{+}, and K^{+} *(14)*. EDTA completely inhibits activity at a concentration of 10 μ*M*. This is probably the result of the calcium requirement of the enzyme. When per-

forming assays, it should be noted that citrate, borate, Veronal, and Tris anions partially inhibit esterase activity *(18)*.

2.3. Elastase (EC 3.4.21.36)

2.3.1. General Information

The commercial enzyme is normally prepared from porcine pancreas. Elastase I (pancreatopeptidase E) is a serine protease and is produced in the pancreas as the inactive zymogen proelastase. Proelastase is converted to elastase by tryptic cleavage of a single peptide bond *(22–24)*. The isolation of porcine elastase II has also been described, but elastase I is the commercially available enzyme *(25)*.

2.3.2. Specificity

The enzyme has rather a broad specificity cleaving C-terminal to small hydrophobic side chains, e.g., Gly, Ala, Ser, Val, Leu, and Ileu *(26–28)*.

2.3.3. Molecular Mass

Elastase is a single polypeptide chain of 240 residues crosslinked by four disulfide bridges, and has a mol mass of 25,900 Da. The complete amino acid sequence is known *(29)*; the gene has been cloned and mol mass confirmed *(30)*.

2.3.4. pH Optimum

The enzyme has a pH optimum of pH 8.0–8.5 *(31)*.

2.3.5. Assay

An assay based on the hydrolysis of *N*-benzoyl-L-alanine methyl ester (BAME) is used. The reaction conditions consist of 10 m*M* substrate in 0.1*M* KCl and 0.01*M* Tris-HCl, pH 8.0, at 25°C. The enzyme solution is added and the change in absorbance with time is monitored at 253 nm *(24)*.

Alternatively, an assay based on the liberation of congo red from elastin-congo red can be used *(24)*. The enzyme is supplied with a specific activity of 15–20 U/mg using this assay, where 1 U is defined as the amount of enzyme that solubilizes 1 mg of elastin from elastin-congo red in 60 min at 25°C in carbonate buffer, pH 8.8.

Also, hydrolysis of succinyl-(L-ala)$_3$-4-nitroanilide by elastase can be monitored at 410 nm. Using this assay, the enzyme is supplied with a specific activity of 10–15 U/mg where 1 U corresponds to the amount of enzyme that liberates 1 μmol of 4-nitroaniline/min at 25°C, pH 7.8.

2.3.6. Stability

The enzyme is active in 0.1% SDS. The lyophilized enzyme is stable indefinitely at –20°C and for at least 6 mo at 4°C. The enzyme rapidly autolyzes if stored at room temperature near its pH optimum of 8.8. However, in solution at pH 5.0, proteolytic activity is slight, and solutions at this pH can be used at room temperature with minimal autolysis. The addition of 0.1 m*M* Ca^{2+} (or 0.1 m*M* Mg^{2+} to a lesser extent) stabilizes the native form of the enzyme against thermal denaturation, even if a destabilizing ion (e.g., Cu^{2+}) is present *(32)*. Elastase retains activity in 2*M* urea. Note that the enzyme sticks to glass, and therefore, plasticware should be used where possible. The addition of Triton X-100 to incubation solutions at 0.05% helps minimize this adhesion *(31)*.

2.3.7. Inhibitors

Since it is a serine protease, elastase is inhibited by DFP and PMSF (*see* Section 2.1.7. for practical details). However, it is not inhibited by TPCK (*see* Section 2.1.7.). It is inhibited by elastinal, a low-mol-mass (mol mass 513 Da) amino acid aldehyde at 10–100 μ*M*, and also by α_2 macroglobulin (at equimolar concentrations with elastase) and by α_1 antitrypsin. Various sulfonyl fluorides and *p*-dinitrophenyl diethyl phosphate also inhibit the enzyme *(33)*. The activity of elastase has been shown to be appreciably affected by salts. Sodium chloride (50–70 m*M*) gave 50% inhibition *(31)*, and similar effects were observed with potassium chloride, ammonium sulfate, and sodium cyanide *(31)*. Copper sulfate (10 m*M*) gave 50% inhibition, but millimolar concentrations of zinc, manganese, cobalt, magnesium, or calcium had no effect *(34,35)*. The enzyme is unaffected by EDTA *(35–37)*.

2.4. Endoproteinase Arg-C (EC 3.4.21.40)

2.4.1. General Information

The enzyme is prepared from mouse submaxillary glands *(38,39)*. It is also known as submaxillary protease.

2.4.2. Specificity

This enzyme cleaves specifically C terminal to arginine residues, although cleavage at some residues does not always go to completion *(40)*. Other workers have found occasional bonds that are resistant to cleavage. For example, when the immunoglobulin κ chain was used as

substrate, cleavage occurred at most arginine residues, but one Arg—Val bond and one Arg—Arg bond were resistant to cleavage *(41)*. The specificity of the enzyme appears to broaden during long incubations, so it is best to use the enzyme at high ratios for short periods, e.g., 1:50 dilution for 2 h at 37°C *(40)*.

2.4.3. Molecular Mass

The enzyme is a serine protease and has a mol mass of about 28,000 Da *(42)*. The amino acid composition is also known *(43)*.

2.4.4. pH Optimum

The enzyme has a pH optimum between pH 7.5 and 8.5 *(42)*.

2.4.5. Assay

The assay is based on the hydrolysis of N_{α}-*p*-tosyl-L-arginine methyl ester (TAME). The substrate solution is 0.8*M* TAME in 0.1*M* phosphate, pH 8.0. The reaction is started by the addition of 0.1 mL of enzyme (suitably diluted in redistilled water) to 2.9 mL of substrate solution. The change in absorbance with time at 247 nm is monitored, and enzyme activity calculated. Using this assay, the enzyme is supplied with a specific activity of 100–250 U/mg where 1 U is defined as the amount of enzyme that hydrolyzes 1 μmol of TAME/min.

2.4.6. Stability

The enzyme retains 90% activity after 1 h at 25°C in 0.1% SDS. It is unstable at low pH, and the enzyme is denatured in 4*M* urea. Solutions of the enzyme can be stored at –20°C without loss of activity. The enzyme loses specificity during long incubations; digestions proceed more rapidly if the substrate is first denatured. The purified form of the enzyme (enzyme A) undergoes autolysis on storage at pH 8.0 overnight, resulting in the formation of enzyme D, formed by the cleavage of a fragment of mol mass 3000–5000 Da from enzyme A *(44,45)*. Commercially available enzyme is a mixture of enzymes A and D.

2.4.7. Inhibitors

The enzyme is inhibited by low pH (<2), DFP, TLCK *(43)* (*see* Section 2.1.7. for practical details), and α_2 macroglobulin when used in equimolar ratios with the enzyme. The enzyme has also been shown to be inhibited at 1 m*M* concentrations by Hg^{2+} (100% inhibition), Cu^{2+} (80%), Zn^{2+} (65%), Cd^{2+} (50%), and Fe^{2+} (50%) *(42)*.

2.5. Endoproteinase Asp-N

2.5.1. General Information

This enzyme has been isolated from the culture filtrates of *Pseudomonas fragic (46,47)*. The enzyme is a neutral protease and is a metalloenzyme having an essential zinc atom at the active site *(46,47)*.

2.5.2. Specificity

The enzyme cleaves protein substrate on the N-terminal side of aspartic acid residues or cysteic acid residues *(48)*. Cleavage occurred for all such residues in oxidized ribonuclease, but when myoglobin was used, a similar specificity was observed, except that only four out of six aspartyl bonds present were hydrolyzed *(48)*. However, occasional additional cleavage of Glu residues has been reported *(49)*.

2.5.3. Molecular Mass

The mol mass of the enzyme is about 50,000 Da *(46,47)*. The amino acid analysis of the enzyme has been published *(47)*.

2.5.4. pH Optimum

The enzyme has a pH optimum of about pH 7.0, with activity dropping off rapidly above pH 9.0 and below pH 6.0 *(46)*.

2.5.5. Assay

The assay is based on the hydrolysis of azocasein. A 2% azocasein substrate in 50 m*M* phosphate buffer, pH 7.0, is used. Substrate (250 µL) is incubated at 25°C and 150 µL of enzyme then added. After incubation for 20 min, the reaction is terminated by addition of 1.2 mL of 10% (w/v) TCA, and the mixture allowed to stand for a further 15 min to allow complete precipitation of the remaining azocasein. The mixture is centrifuged (8000*g*) or filtered, and 1–2 mL of the supernatant are then combined with 1.4 mL of 1.0*M* NaOH. The absorbance is read at 440 nm, and the concentration of enzyme calculated. One unit of enzyme activity is defined as the amount of enzyme that produces an absorbance change of 1.0 in a 1-cm path length cuvet under the above conditions.

2.5.6. Stability

The enzyme is fully active in 2*M* urea *(48)*. It is stable between pH 7.0 and 9.5, but is rapidly inactivated outside these values *(46)*. The enzyme is stable at temperatures up to 40°C in the absence of Ca^{2+}. The

presence of 5 m*M* $CaCl_2$ appears to stabilize the enzyme somewhat with residual activity still remaining at 60°C *(46)*. The enzyme is stable at 4°C, stored dry. A solution in redistilled water may be used for 1 wk at maximum if stored at 4°C *(44)*. Aqueous solutions of the enzyme are stable up to 1 mo at –20°C. Freeze/thawing of enzyme solutions is not recommended. The enzyme is stable in 0.01% SDS, 0.1*M* guanidine hydrochloride, and 10% (v/v) acetonitrile *(44)*.

2.5.7. Inhibitors

Being a metalloenzyme, the enzyme is inhibited by chelating agents, such as EDTA, and 1,10-phenanthrolene *(46,47)*. Activity can be restored by the addition of zinc or cobalt ions, but not by adding calcium or magnesium ions *(46,47)*. The enzyme is not inhibited by iodoacetic acid, PCMB, dithiothreitol, PMSF, or soyabean trypsin inhibitor *(46,47)*.

2.6. Endoproteinase Glu-C (EC 3.4.21.9)

2.6.1. General Information

The enzyme is purified from the culture filtrates of *Staphylococcus aureus* strain V8 *(50–52)*. The enzyme has also been named staphylococcal protease and V8 protease.

2.6.2. Specificity

The enzyme cleaves C terminal to Asp and Glu residues in protein substrates when used in phosphate buffer, pH 7.8. However, when used in ammonium buffers (ammonium bicarbonate, pH 7.8, or ammonium acetate, pH 4.0), cleavage is restricted to Glu residues only. —Glu—*X*— or Asx—*X* bonds are not cleaved if *X* is Pro or *S*-carboxymethylcysteine *(53)*. If *X* is a bulky hydrophobic residue, cleavage is slow. In addition to the expected cleavages at Asp and Glu, a number of other workers have noted a number of nonspecific cleavages, often of Ser—*X* bonds *(54–56)*.

2.6.3. Molecular Mass

The enzyme is a serine protease (*see* Section 2.1.1.) with a mol mass of 12,000 Da *(50)*. The amino acid composition has been published *(50)*.

2.6.4. pH Optimum

The protease is active in the pH range 3.5–9, and exhibits maximum proteolytic activity at pH 4.0 and 7.8 with hemoglobin as substrate *(55)*.

2.6.5. Assay

The assay is based on the hydrolysis of casein. The reaction conditions consist of 1% casein in 0.1*M* Tris-HCl, pH 7.8. Enzyme is added, and the mixture is incubated at 25°C for 10 min. The reaction is stopped by adding an equal volume of 10% TCA, and the mixture is allowed to stand for a further 15 min. The mixture is then filtered or centrifuged; the absorbance of the supernatant is read at 280 nm, and the enzyme activity calculated. Using this assay, the enzyme is usually supplied with a specific activity of 500 U/mg, where 1 U is the amount of enzyme that caused 0.001 A_{280} unit of change/min at 25°C, pH 7.8.

Alternatively, an assay based on the hydrolysis of *N*-CBZ-L-glutamyl α-phenyl ester can be used. Using this assay, the enzyme is usually supplied with a specific activity of 500–1000 U/mg, where 1 U is defined as the amount of enzyme that will hydrolyze 1 μmol of substrate/min at pH 7.8, at 37°C. (One unit is equivalent to approx 0.004 casein digestion units.)

Also, an assay based on the hydrolysis of carbobenzoxy-Phe-Leu-Glu-4-nitroanilide, at pH 7.8, 410 nm, and 25°C has been described. The enzyme is supplied with a specific activity of 10–20 U/mg when this assay is used, where 1 U hydrolyzes 1 μmol of substrate/min at pH 7.8 and 25°C.

2.6.6. Stability

The enzyme is active in the presence of 0.2% SDS, and retains 50% activity in 4*M* urea *(55)*. The enzyme has been shown to be equally resistant to heat denaturation over a wide range of temperatures. Seventy percent of activity remained after heating at 100°C for 3 min *(50)*. Aqueous solutions are stable to freeze/thawing. The enzyme is stable at pH 4–10, but precipitates below pH 4. It retains activity in 1*M* guanidine hydrochloride and 10% acetonitrile. The enzyme is stable at 4°C, when stored dry. A solution in redistilled water may be used for 1–2 d at maximum, if stored at 4°C.

2.6.7. Inhibitors

Being a serine protease, the enzyme is inhibited by DFP *(50)* (*see* Chymotrypsin, Section 2.1.7., for practical details). The enzyme is not inhibited by EDTA, nor does the addition of a range of divalent metal ions have any effect on enzyme activity *(50)*. The enzyme is also inhibited by α_2 macroglobulin at concentrations equimolar with the enzyme.

2.7. Endoproteinase Lys-C (EC 3.4.99.30)

2.7.1. General Information

This enzyme is a serine protease isolated from culture filtrates of *Lysobacter enzymogenes (57).*

2.7.2. Specificity

The enzyme cleaves specifically C-terminal to lysine residues *(57).* Some minor nonspecific cleavages, notably at Asn—*X* bonds, have also been reported.

2.7.3. Molecular Mass

It has a mol mass of about 33,000 Da.

2.7.4. pH Optimum

The enzyme has a pH optimum in the range 8.5–8.8.

2.7.5. Assay

The assay is based on hydrolysis of Tosyl-Gly-Pro-Lys-4-nitroanilide (commercially sold as Chromozym® PL). Chromozym® PL (9 mg) is dissolved in 1 mL of redistilled water. This solution (0.05 mL) is combined with 1.0 mL of buffer (25 m*M* Tris-HCl, pH 7.7, 1 m*M* EDTA) and incubated at 25°C. The reaction is started by addition of 0.05 mL of enzyme (suitably diluted with the above buffer). The increase in absorbance is monitored at 405 nm, and using an extinction coefficient of 10.4 $mM^{-1} \cdot cm^{-1}$, the enzyme activity of the sample is calculated. Using this assay, the enzyme is supplied with a specific activity of 25–35 U/mg, where 1 U is defined as the amount of enzyme that liberates 1 μmol 4-nitroaniline/min at 25°C from Tosyl-Gly-Pro-Lys-4-nitroanilide.

2.7.6. Stability

The enzyme is active in 0.5% SDS, 5*M* urea, and in 10% acetonitrile. Aqueous solutions are stable at pH 5–12 for up to 1 d at 4°C, or can be stored at –20°C for months. The enzyme is stable at 4°C, stored dry. To avoid autolysis, the incubation temperature should not exceed 37°C *(44).*

2.7.7. Inhibitors

The enzyme is inhibited by DFP (1 m*M*), TLCK (1 m*M*), leupeptin (0.4 m*M*), and aprotinin (100 μg/mL). The enzyme is not inhibited by EDTA (2 m*M*), PMSF (1 m*M*), or α_1-antitrypsin.

2.8. Pepsin (EC 3.4.23.1)

2.8.1. General Information

The term "pepsin" refers to a group of gastric proteinases that are active at acid pH values (pH 1–5). These enzymes are formed by partial proteolysis of the inactive zymogens, the pepsinogens *(58)*. Pepsin is very suitable for investigating arrangement of disulfide bonds in proteins owing to minimization of disulfide interchange reactions because of low pH. The commercially available form of the enzyme is Pepsin A from porcine gastric mucosa *(59)*.

2.8.2. Specificity

Pepsin has broad specificity with a preference for peptides containing linkages with aromatic or dicarboxylic L-amino acids. It preferentially cleaves C terminal to Phe and Leu and to a lesser extent Glu linkages. The enzyme does not cleave at Val, Ala, or Gly.

2.8.3. Molecular Mass

The enzyme has a mol mass of 36,000 Da. The amino acid sequence and composition have been determined *(58,59)*.

2.8.4. pH Optimum

Pepsin is optimally active at pH 2–4 *(58)*.

2.8.5. Assay

The assay is based on hydrolysis of denatured hemoglobin. Five milliliters of hemoglobin solution (20 mg/mL in 60 m*M* HCl) and 0.1 mL enzyme in HCl (10 m*M*) are combined and incubated for 10 min at 37°C. Ten milliliters of TCA (5% [w/v]) are then added, and the mixture centrifuged at 8000*g* or filtered. Five milliliters of filtrate are combined with 10 µL of 0.5*M* NaOH and 3 mL of Folin and Ciocalteau's reagent; after 10 min, the absorbance is read at 578 nm. The quantity of tyrosine released can be determined from a standard curve. Using this assay, the enzyme is supplied with a specific activity of 1 Anson unit/gram, where 1 Anson unit is the amount of enzyme that hydrolyzes hemoglobin, such that hydrolysis products formed/min give the same absorbance with Folin's reagent as 1 mmol of tyrosine.

Alternatively, an assay based on the hydrolysis of acetyl-L-phenylalanyl-3,5-diiodo-L-tyrosine can be used. Using this assay, the enzyme is supplied with a specific activity of approx 150 U/g, where

1 U is defined as the amount of enzyme that hydrolyzes 1 μmol of substrate/min at pH 2.0 and 37°C.

2.8.6. Stability

The enzyme is irreversibly inactivated at pH values >6.0 *(60)*. Below pH 6, it retains activity in 4*M* urea and in 3*M* guanidine hydrochloride *(61)*, and is unaffected by heating to 60°C. The lyophilized enzyme is stable for months at 4°C.

2.8.7. Inhibitors

Pepsin is inhibited by Pepstatin A (mol mass 685.9 Da), diazoketones, phenylacyl bromide, and aliphatic alcohols.

2.9. Thermolysin (EC 3.4.24.4)

2.9.1. General Information

The enzyme is produced in the culture supernatants of the microorganism *Bacillus thermoproteolyticus*. It is a metalloenzyme having an absolute requirement of zinc ions for activity. The enzyme is thermostable *(62)*.

2.9.2. Specificity

In general, thermolysin hydrolyzes peptide bonds N terminal to hydrophobic amino acids with bulky sidechains *(63)*, e.g., Ileu, Leu, Val, and Phe. However, the enzyme is not absolutely specific for these residues, with other workers having identified minor cleavage sites N terminal to most other amino acids. However, the *preferred* sites of cleavage are those N terminal to the bulky hydrophobic residues, and it is this characteristic that has made thermolysin useful in protein studies. Also, thermolysin appears to be able to cleave native proteins to a greater extent than most other proteases. Bonds involving proline of the type —Pro—Ileu— are cleaved, but in sequences of the type *X*—Phe—Pro—, the *X*—Phe— bond is resistant to cleavage. Cleavage is also inhibited by an adjacent α amino or carboxyl group, so thermolysin has no exopeptidase activity *(63,64)*.

2.9.3. Molecular Mass

It is a single polypeptide chain of 316 residues and has a mol mass of 34,600 Da *(65)*. The complete amino acid sequence has been determined *(65)*.

2.9.4. pH Optimum

Thermolysin has optimal activity in the pH range 7.0–8.0, but can be used up to pH 9.0 if required.

2.9.5. Assay

The assay is based on the hydrolysis of casein. One milliliter of borate buffer (0.1*M*, pH 7.2), 1 mL of substrate solution (2 g casein dissolved in water and 10 mL of 1 *M* NaOH, pH adjusted to 7.2, and solution made up to 100 mL), and 0.02 mL of enzyme solution are incubated for 10 min at 37°C; 2.0 mL of TCA (0.1*M*) and 0.2 mL $CaCl_2$ (2 m*M*) are then added; the mixture is incubated for a further 10 min at 37°C, and then centrifuged or filtered. Filtrate (1.5 mL) is combined with 5 mL of Na_2CO_3 (0.4*M*) and 1 mL of Folin and Ciocalteau's reagent (commercial reagent diluted 1:3). This mixture is then incubated for 20 min at 37°C. The absorbance is read at 578 nm, and using standard solutions, the concentration of the enzyme is calculated. The enzyme is usually supplied with a specific activity of 4 U/mg, where 1 U is defined as the amount of enzyme that liberates 1 μmol Folin-positive amino acids and peptides (calculated as tyrosine)/min at pH 7.2 and 37°C.

2.9.6. Stability

The enzyme is thermostable, and at 80°C has a half-life of 1 h *(62)*. There is no loss of activity at 60°C over 1 h. A zinc atom at the active site is essential for activity, but the enzyme also has an absolute requirement for calcium ions for stability, and this should therefore be included in digestion buffers, although most preparations contain sufficient Ca^{2+} to be active without the addition of extra Ca^{2+}. The enzyme is stable in 8*M* urea, and also in 20% methanol or ethanol. However, these conditions enhance thermal denaturation and digestion should then be carried out at 20°C *(62)*. Aqueous solutions of the enzyme can be stored frozen for weeks without loss of activity *(62)*. The enzyme is also active in 1% SDS, and 0.1*M* NaCl. The lyophilized enzyme is stable at 4°C for many months.

2.9.7. Inhibitors

Because of the requirement for a zinc atom at the active site, chelating agents, such as EDTA and 1,10 phenanthroline, are strong inhibitors *(62)*. Complete inhibition occurs in 5 m*M* EDTA at 40°C in 3 min.

Oxalate, citrate, and phosphate also inhibit the enzyme, so phosphate buffers should be avoided. Mercuric chloride and silver nitrate at 5 m*M* also cause complete inactivation. The enzyme is also inhibited by α_2 macroglobulin and diethyl pyrocarbonate. The enzyme is not inhibited by DFP, TPCK, soyabean trypsin inhibitors, cysteine, or sodium cyanide *(62,66)*.

2.10. Trypsin (EC 3.4.21.4)

2.10.1. General Information

Trypsin is a serine protease that is formed by cleavage of the inactive precursor trypsinogen. Activation is achieved by limited proteolytic cleavage at a single peptide bond, Lys_6-Ile_7, near the N terminus of the zymogen. The activation process is catalyzed by a variety of enzymes, including enterokinase, mold proteases, and trypsin itself. The commercial form of the enzyme is from bovine pancreas *(67)*.

2.10.2. Specificity

Trypsin is a highly specific endopeptidase whose protease activity is restricted to the positively charged side chains lysine and arginine *(67,68)*. Cleavage occurs C terminal to these residues. The enzyme also hydrolyzes ester and amide linkages of synthetic derivatives of these amino acids. No hydrolysis at *X*—Pro bonds occurs, and reduced hydrolysis occurs if *X* is preceded or followed by an acidic residue. Some nonspecific cleavage may occur because of the presence of ψ-trypsin in some commercial preparations. Therefore, results from a tryptic digest should be interpreted carefully, since the number of peptides may not add up to the total number of arginine and lysine residues.

2.10.3. Molecular Mass

The bovine enzyme has a mol mass of approx 24,000 Da. The enzyme has been sequenced, and the amino acid composition has been reported *(67)*.

2.10.4. pH Optimum

The enzyme has a pH optimum of 7–9 *(67)*.

2.10.5. Assay

The assay is based on hydrolysis of *N*-benzoyl-L-arginine ethyl ester. The substrate solution is 46.7 m*M* Tris-HCl, pH 8.0, 0.9 m*M* BAEE, and 19 m*M* $CaCl_2$. Following the addition of enzyme, the change in

absorbance at 255 nm at 25°C is measured, and the enzyme concentration calculated using an extinction coefficient of 0.081 $mM^{-1} \cdot cm^{-1}$. The enzyme is supplied with a specific activity of 6000–10,000 U/mg when this assay is used, where 1 U is defined as the amount of enzyme that produces an increase in absorbance at 255 nm of 0.001/min at pH 7.6 and at 25°C.

2.10.6. Stability

Trypsin is stable at pH 3 at low temperatures and can be stored for weeks without loss of activity. The lyophilized enzyme is stable for months at 4°C. Trypsin is reversibly denatured above pH 11 *(69)*. An increase in temperature above pH 8 results in irreversible denaturation. At pH 2–3, it is stable up to 40°C.

Trypsin is fully active in 6.5*M* urea, but is reversibly denatured by 8*M* urea *(70)*. Urea-denatured trypsin regains activity by dialysis against $10^{-3}M$ HCl *(69)*.

The enzyme retains activity in 30% ethanol *(71)*, 0.1% SDS, 1*M* guanidine hydrochloride, and 10% acetonitrile. Trypsin has a tendency to autolyze at pH values >5. The addition of calcium ions (20 m*M*) to trypsin solutions retards autolysis *(67)*.

2.10.7. Activators and Inhibitors

Trypsin is activated by calcium ions and lanthanides *(72)*. The enzyme is inhibited by DFP, TLCK, soybean trypsin inhibitors *(73)*, aprotinin (mol mass 6500 Da), leupeptin (an amino acid aldehyde, mol mass 426.6 Da), and α-macroglobulin *(74)*. Inhibition by benzamidine is reversible *(44)*.

3. Experimental Procedures—Practical Details

3.1. Choice of Buffer

It is generally preferable to use a volatile buffer, particularly if the peptides are to be separated by ion-exchange chromatography or paper chromatography/electrophoresis, where the buffer salts may interfere, particularly if the digested sample is concentrated or dried down at the end of the digestion. Salt concentration is, however, not a problem if reverse-phase HPLC is to be used to separate the peptides. The majority of the enzymes described above have pH optima in the range 7.5–8.5, and the most commonly used volatile buffers in this range are ammonium bicarbonate (100 m*M* gives a pH of 8.0) or *N*-ethyl

morpholine acetate (pH 8.5). For pepsin, 5% acetic acid or 10 m*M* HCl is used. Ammonium acetate (0.1*M*) can be used in the pH range 4.5–5.5. If samples are to be taken for SDS gel electrophoresis, potassium ions should be avoided, since these precipitate with dodecyl sulfate and can aggregate peptides in the electrophoresis wells.

3.2. Condition of the Substrate

The physical state of the substrate is of great importance in determining the degree of digestion. Cleavage of the native protein will often result in large proteolytic fragments, since only these cleavage sites on the surface of the protein will be accessible to the enzyme. However, if it is required that the enzymic digestion go to completion, then the substrate protein should be denatured and disulfide bridges reduced, thus making all potential cleavage sites accessible to the enzyme. Denaturation can be achieved by boiling the protein solution, precipitation in 5% trichloroacetic acid, or treatment in 8*M* urea or 6*M* guanidium hydrochloride. The reduction and alkylation of sulfhydryl groups have been described in vol. 3 of this series, Chapter 7. Denatured proteins are often insoluble, and although proteolytic enzymes will cleave insoluble substrates, it is desirable that the precipitate be as finely divided as possible to allow maximum access of the enzyme. This can be achieved by dissolving the precipitate in 8*M* urea and then removing the urea by dialysis against buffer, or dissolving the precipitate at low pH and then slowly titrating back to the required pH. Many of the enzymes described above remain active in the presence of SDS or urea, and therefore, these agents may be used to denature the protein substrate (which remains soluble) and then enzyme may be added directly. For glycoproteins, the sugar backbone can sterically hinder enzyme access to the polypeptide backbone and severely reduce digestion. Deglycosylation prior to enzymic cleavage can be achieved using appropriate enzymes.

3.3. Reaction Conditions

When attempting to achieve complete digestion, an enzyme to substrate ratio of 1–2% (w/w) is commonly used, usually at 37°C for 2–4 h. However, longer times (24–48 h) with repeated addition of enzyme may have to be used with particularly resistant substrates. In these cases, one should be aware of the increased possibility of generating

proteolytic fragments from the enzyme being used. Bacterial growth and protein cleavage by released enzymes can also be a problem in extended hydrolysis. In all cases, therefore, it is wise to carry out a preliminary small-scale trial digestion to determine the minimum digestion time necessary to generate a stable peptide pattern. For limited proteolysis experiments, much lower enzyme-to-substrate ratios (0.1% or less), lower temperatures (10–20°C), and shorter times (15–30 min) may all be used in conjunction to reduce the amount of proteolysis that occurs. Again, it is necessary to carry out preliminary small-scale trial digests to determine the optimum conditions. Conditions for trial digests are described in Section 3.6.

3.4. Monitoring of Digestion

3.4.1. Limited Digestion

Conditions for controlled proteolysis are far more difficult to predict than the rather obvious extreme conditions used for total proteolysis and are highly dependent on the experimental conditions chosen by the user. Since relatively large peptides are to be generated, the production of peptides in trial digests is best monitored by SDS gel electrophoresis (*see* vol. 1 of this series, Chapter 6). Following addition of the enzyme (1:100 to 1:1000 [w/w]) to either the native or denatured protein at 30–37°C, remove samples containing about 5 μg of protein at regular intervals (say every 15 min for 4 h) and add to SDS sample buffer. Since many proteolytic enzymes (e.g., clostripain, thermolysin, endoproteinase Glu-C, elastase, papain, and chymotrypsin) are still active in 0.1% SDS, it is essential to boil the sample immediately to inactivate the protease. When all samples have been taken, they are run on the gel and the time of digestion that gives the required pattern of peptides used for the large-scale digest. If digestion was too rapid, a repeat trial should be carried out at a lower enzyme-to-substrate ratio and/or a lower temperature. However, if cleavage of the native protein is being attempted, considerably *higher* enzyme-to-substrate ratios may be needed, possibly as high as 1:1 (w/w).

3.4.2. Total Digestion

Following addition of the enzyme (1:10 to 1:100 [w/w]) at 37°C, remove samples containing about 5 μg of protein at 0, 1, 2, 4, and 8 h. A further addition of protease is made (to replace enzyme lost by autolysis), and incubation continued overnight. A further sample is

then taken, a final addition of enzyme made, and incubation continued for a further 8 h, when a final sample is taken. Samples are either immediately frozen for later analysis by reverse-phase HPLC (*see* vol. 1 of this series, Chapter 5) or dried directly onto a cellulose TLC plate (which effectively stops the reaction), and when fully loaded, the plate is run in an appropriate solvent, such as butanol/pyridine/acetic acid/water (15:12:3:10) and peptides identified by spraying with ninhydrin or fluorescamine (*see* vol. 1). Whichever method is used, the time of digestion that gives an unchanging pattern of peptides is used for the large-scale digestion.

3.5. Termination of the Digestion

The simplest method of termination is freezing followed by lyophilization. Such a procedure will retain biological activity of limited proteolysis products where necessary. Where total digestion has taken place, samples may be easily (and more quickly) rotary evaporated to dryness, since one is not attempting to retain any tertiary structures in the proteolysis products. Alternatively, the pH digest may be altered (normally by the addition of acid) to a pH value at which the enzyme is no longer active, or the sample boiled rapidly to denature the enzyme. The reaction can also be stopped by the addition of a specific inhibitor for the enzyme. In general, serine proteases are inhibited by the addition of PMSF to 1 m*M*. (PMSF is prepared as a stable 1*M* stock solution in propan-2-*ol* and added to the reaction with vigorous mixing. PMSF *cannot* be prepared as an *aqueous* stock solution, since it is rapidly hydrolyzed in water with a half-life of about 1 h. Metalloenzymes can be inhibited by the addition of disodium EDTA to a concentration of 10 m*M* (from a 1*M* stock solution), and sulfhydryl proteases inhibited by the addition of iodoacetic acid to a concentration just in excess of that of the sulfhydryl reagent added to activate the enzyme.

3.6. Typical Digestion Conditions

The conditions given below are suitable for producing total digestion, e.g., suitable for peptide mapping.

3.6.1. Chymotrypsin

Dissolve the protein substrate at a concentration up to 10 mg/mL in 0.1*M* ammonium bicarbonate (pH 8.0). Add enzyme solution to give a ratio of 2% (w/w), and incubate for 4 h at 37°C *(10)*.

3.6.2. Clostripain

Note that the enzyme may need to be activated prior to use (*see* Section 2.2.7.). Dissolve the substrate protein at 1–10 mg/mL in 50 m*M* ammonium bicarbonate, 0.2 m*M* calcium acetate, and 2.5 m*M* DTT. Add clostripain at 1–2% (w/w), and incubate for 2–3 h at 37°C. More extensive incubations can result in some cleavage at lysine residues *(75)*.

3.6.3. Elastase

Dissolve the substrate at 1–10 mg/mL in 100 m*M* ammonium bicarbonate at 37°C. Add elastase at 1–2% (w/w), and incubate for 4 h *(76)*.

3.6.4. Endoproteinase Arg-C

Dissolve the substrate in 1% ammonium bicarbonate, pH 8.0. Add enzyme at 2% (w/w), and incubate at 37°C for 8 h *(77)*.

3.6.5. Endoproteinase Asp-N

Dissolve the substrate in 50 m*M* phosphate buffer, pH 7.0, at 37°C. Add enzyme at 5% (w/w) in the same buffer, and incubate for 4–18 h *(78)*.

3.6.6. Endoproteinase Glu-C

For cleavage at glutamate residues, dissolve the proteins substrate in 0.1*M* ammonium bicarbonate, pH 8.0, at a concentration of about 10 mg/mL, and add enzyme at 3% (w/w). Incubate at 37°C for 2–18 h. To extend the cleavage to aspartate residues, perform the same incubation, but using 0.1*M* sodium phosphate buffer, pH 7.8 *(79)*.

3.6.7. Endoproteinase Lys-C

Dissolve the protein substrate in 0.1*M* ammonium bicarbonate, pH 8, at 10 mg/mL. Add enzyme at 2% (w/w), and incubate at 37°C for 2 h. The same amount of enzyme is again added, and after 22 h, the digestion is stopped *(80,81)*.

3.6.8. Pepsin

Dissolve the substrate in 0.01*M* HCl, and adjust the pH to 2.0. Add pepsin to 1% (w/w), and incubate for 1 h at 25°C *(82)*.

3.6.9. Thermolysin

Dissolve the substrate protein in 0.1*M* ammonium bicarbonate containing 5 m*M* $CaCl_2$. Add enzyme to 2% (w/w), and incubate at 37°C for 1 h *(83)*.

3.6.10. Trypsin

Dissolve the substrate in 0.1*M* ammonium bicarbonate. Add enzyme to 1–2% (w/w), and incubate at 37°C for 1–4 h *(84)*.

References

1. Miller, D. D., Horbett, T. A., and Teller, D. C. (1971) Reevaluation of the activation of bovine chymoytrypsinogen A. *Biochemistry* **10,** 4641–4648.
2. Wilcox, P. E. (1970) Chymotrypsinogens—Chymotrypsins. *Methods Enzymol.* **19,** 64–80.
3. Desnuelle, P. (1960) Chymotrypsin. *Enzymes* **4,** 93–120.
4. Hess, G. P. (1971) Chymotrypsin—Chemical properties and catalysis. *Enzymes* **4,** 213–250.
5. Tomlinson, G., Shaw, M. C., and Viswanatha, T. (1976) Chymotrypsin (S). *Methods Enzymol.* **34,** 415–420.
6. Baumann, W. K., Bizzozero, S. A., and Dutler, H. (1970) Specificity of α-chymotrypsin. Dipeptide substrates. *FEBS Lett.* **8,** 257–260.
7. Hartley, B. S. and Kaffman, B. J. (1966) Corrections to the amino acid sequence of bovine chymotrypsinogen A. *Biochem.J.* **10,** 229–231.
8. Schwert, G. W. and Takenaka, Y. (1955) A spectrophotometric determination of trypsin and chymotrypsin. *Biochim. Biophys. Acta.* **16,** 570–575.
9. Kumar, S. and Hein, G. E. (1969) A rapid kinetic assay method for α-chymotrypsin and its application. *Anal. Biochem.* **30,** 203–211.
10. Spackman, D. H., Stein, W. H., and Moore, S. (1960) The disulfide bonds of ribonuclease. *J. Biol. Chem.* **235,** 648–659.
11. Umezawa, H. (1976) Structures and activities of protease inhibitors of microbial origin. *Methods Enzymol.* **45,** 678–700.
12. Jansen, E. F. and Balls, A. K. (1952) The inhibition of β- and γ-chymotrypsin and trypsin by diisopropyl fluorophosphate. *J. Biol. Chem.* **194,** 721–727.
13. Green, M. M., Gladner, J. A., Cunningham, L. W., and Neurath, H. (1952) The effects of divalent cations on the enzymic activities of trypsin and of α-chymotrypsin. *J. Am. Chem. Soc.* **74,** 2122,2123.
14. Mitchell,W. M. and Harrington, W. F. (1968) Purification and properties of clostridiopeptidase B (Clostripain). *J. Biol. Chem.* **243,** 4683–4692.
15. Mitchell, W. M. and Harrington, W. F. (1971) Clostripain. *Enzymes* **3,** 699–719.
16. Cole, P. W., Murakami, K., and Inagami, T. (1971) Specificity and mechanism of Clostripain catalysis. *Biochemistry* **10,** 4246–4252.
17. Mitchell, W. M. (1977) Cleavage at arginine residues by Clostripain. *Methods Enzymol.* **47,** 165–176.
18. Mitchell, W. M. and Harrington, W. F. (1970) Clostripain. *Methods Enzymol.* **19,** 635–642.
19. Shih, J. W. and Hash, J. H. (1971) The N,O-Diacetylmuramidase of *Chalaropsis* species. Amino acid composition and partial structure formula. *J. Biol. Chem.* **246,** 994–1006.

20. Labouesse, B. and Gros, P. (1960) Laclostripaine, protéase de *Clostridium histolyticum* Purification et activation par les thiols. *Bull. Soc. Chim. Biol.* **42,** 543–558.
21. Porter, W. H., Cunningham, L. W., and Mitchell, W. M. (1971) Studies on the active site of Clostripain. The specific inactivation by the chloromethyl ketone derived from α-N-Tosyl-L-Lysine. *J. Biol. Chem.* **246,** 7675–7682.
22. Mandl, I. (1961) Collagenases and elastases. *Adv. Enzymol.* **23,** 163–264.
23. Hartley, B. S. and Shotton, D. M. (1971) Pancreatic elastase. *Enzymes* **3,** 323–373.
24. Shotton, D. M. (1970) Elastase. *Methods Enzymol.* **19,** 113–140.
25. Vered, M., Gertler, A., and Burnstein, Y. (1986) Partial amino acid sequence of porcine elastase II. Active site and the activation peptide regions. *Int. J. Peptide Prot. Res.* **27,** 183–190.
26. Atlas, D. (1975) The mapping of the active site of porcine elastase: size specificity, stereospecificity and the CPK model. "PEPTIDES 1974. Proceedings of the 13th European Peptide Symposium Kiryat Anauim, Israel" (Wolman, Y., ed.), pp. 361–376.
27. Naughton, M. A. and Sanger, F. (1961) Purification and specificity of pancreatic elastase. *Biochem. J.* **78,** 156–163.
28. Narayanon, A. S. and Anwar, R. A. (1969) The specificity of purified porcine elastase. *Biochem. J.* **114,** 11–17.
29. Shotton, D. M. and Hartley, B. S. (1970) Amino acid sequence of porcine pancreatic elastase and its homologies with other serine proteinases. *Nature* **225,** 802–806.
30. Shirasu, Y. Yoshida, H., Mikayama, T., Matsuki, S., Tanaka, J. and Ikenaga, H. (1986) Isolation and expression in *E. coli* of a cDNA clone encoding porcine pancreatic elastase. *J. Biochem.* **99,** 1707–1712.
31. Mandl, I. (1962) Pancreatic elastase. *Methods Enzymol.* **5,** 665–673.
32. Favre-Bonvin, G., Bostancioglu, K., and Wallach, J. M. (1986) Ca^{2+} and Mg^{2+} protection against thermal denaturation of pancreatic elastase. *Biochem. Internatl.* **13,** 983–989.
33. Shotton, D. M. and Watson, H. C. (1970) Three-dimensional structure of tosylelastase. *Nature* **225,** 811–816.
34. Lewis, U. J., Williams, D. E., and Brine, N. G. (1956) Pancreatic elastase: purification, properties and function. *J. Biol. Chem.* **222,** 705–720.
35. Lamy, F., Craig, C. P., and Tauber, S. (1961) Studies on elastase and elastin. Assay and properties of elastase. *J. Biol. Chem.* **236,** 86–91.
36. Gertler, A. and Birk, Y. (1970) Isolation and characterization of porcine proelastase. *Eur. J. Biochem.* **12,** 170–176.
37. Walford, R. L. and Kickhöffen, B. (1962) Selective inhibition of elastolytic and proteolytic properties of elastae. *Arch. Biochem. Biophys.* **98,** 191–196.
38. Schenkein, I ., Boesman, M., Takarsky, E., Fishman, L., and Levy, M. (1969) Proteases from mouse submaxillary gland. *Biochem. Biophys. Res. Commun.* **36,** 156–165.

39. Schenkein, I., Franklin, E. C., and Frangione, B. (1981) Proteolytic enzymes from the mouse submaxillary gland. A partial sequence and demonstration of spontaneous cleavages. *Arch. Biochem. Biophys.* **209,** 57–62.
40. Schenkein, I., Levy, M., Franklin, E. C., and Frangione, B. (1977) Proteolytic enzymes from the mouse submaxillary gland. Specificity restricted to arginine residues. *Arch. Biochem. Biophys.* **182,** 64–70.
41. Wasserman, R. L. and Capra, J. D. (1978) The amino acid sequence of the light chain variable region of a canine myeloma immunoglobulin; Evidence that the V_k subgroups predated mammalian speciation. *Immunochemistry* **15,** 303–305.
42. Boesman, M., Levy, M., and Schenkein, I. (1976) Esteroproteolytic enzymes from the submaxillary gland. Kinetics and other physicochemical properties. *Arch. Biochem. Biophys.* **175,** 463–476.
43. Levy, M., Fishman, L., and Scheinkein, I. (1970) Mouse submaxillary gland proteases. *Methods Enzymol.* **19,** 672–681.
44. Keesey, J. (ed.) (1989) *Biochemica Information: A Revised Biochemical Reference Source.* Boehringer Mannheim Biochemicals, Indianapolis, IA.
45. Schenkein, I., Gabor, M., Franklin, E. C., and Frangione, B. (1979) Structural studies of a male mouse submaxillary gland protease. *Fed. Proc. Fed. Amer. Chem. Soc. Exp. Biol.* **38,** 326.
46. Porzio, M. A. and Pearson, A. M. (1975) Isolation of an extracellular neutral proteinase from *Pseudomonas fragi. Biochim. Biophys. Acta.* **384,** 235–241.
47. Noreaau, J. and Drapeau G. R. (1979) Isolation and properties of the protease from the wild-type and mutant tralns of *Pseudomonas fragi. J. Bacteriol.* **140,** 911–916.
48. Drapeau, G. R. (1980) Substrate specficity of a proteolytic enzyme isolated from a mutant of *Pseudomonas fragi. J. Biol. Chem.* **255,** 839–840.
49. Ingrosso, D., Fowler, A. V., Bleibaum, J., and Clarke, S. (1989) Specificity of Endoproteinase Asp-N *(Pseudomonas fragi)*: Cleavage at glutamyl residues in two proteins. *Biochem. Biophys. Res. Commun.* **162,** 1528–1534.
50. Drapeau, G. R., Boily, Y., and Houmard, J. (1972) Purification and properties of an extracellular protease of *Staphlococcus aureus. J. Biol. Chem.* **247,** 6720–6726.
51. Houmard, D. and Drapeau, G. R. (1972) Staphlococcal protease: A proteolytic enzyme specific for glutamoyl bonds. *Proc. Natl. Acad. Sci.* **69,** 3506–3509.
52. Austen, B. M. and Smith, E. L. (1976) Action of staphlococcal proteinase on peptides of varying chain length and composition. *Biochem. Biophys. Res. Commun.* **72,** 411–417.
53. Wooton, J. C., Baron, A. J., and Fincham, R. S. (1975) The amino acid sequence of *Nemospora* NADP-specific glutamate dehydrogenase. *Biochem. J.* **149,** 749–755.
54. Walker, J. M., Hastings, J. R., and Johns, E. W. (1977) The primary structure of a non-histone chromosomal protein. *Eur. J. Biochem.* **76,** 461–468.
55. Drapeau, G. R. (1976) Protease from *Staphlococcus aureus. Methods Enzymol.* **45,** 469–475.

56. Drapeau, G. R. (1977) Cleavage at glutamic acid with staphlococcal protease. *Methods Enzymol.* **47,** 189–191.
57. Jekel, P. A., Weijer, W. J., and Beintema, J. J. (1983) Use of endoproteinase Lys-C from *Lysobacter enzymogenes* in protein sequence analysis. *Anal. Biochem.* **134,** 347–354.
58. Fruton, J. S. (1971) Pepsin. *Enzymes* **4,** 120–164.
59. Tang, J. (1970) Gastricsin and pepsin. *Methods Enzymol.* **19,** 406–421.
60. Rajagopalan, T. G., Moore, S., and Stein, W. H. (1966) Pepsin from pepsinogen. Purification and properties. *J. Biol. Chem.* **241,** 4940–4950.
61. Perimann, G. E. (1959) Effect of solvents and of temperature on the optical rotary properties of pepsin. *Proc. Natl. Acad. Sci. USA* **45,** 915–922.
62. Matsubara, H. (1970) Purification and assay of thermolysin. *Methods Enzymol.* **19,** 642–651.
63. Heinrikson, R. L. (1977) Applications of thermolysin in protein structural analysis. *Methods Enzymol.* **47,** 75–189.
64. Matsubara, H., Sasaki, R., Singer, A., and Jukes, T. H. (1966) Specific nature of hydrolysis of insulin and tobacco mosaic virus protein by thermolysin. *Arch. Biochem. Biophys.* **115,** 324–331.
65. Titani, K., Hermodson, M. A., Ericson, L. H., Walsh, K. A., and Neurath, H. (1972) Amino acid sequence of thermolysin. *Nature* **238,** 35–37.
66. Latt, S. A., Holmquist, B., and Vallee, B. L. (1969) Thermolysin: A zinc metalloenzyme. *Biochem. Biophys. Res. Commun.* **37,** 333–339.
67. Walsh, K. A. (1970) Trypsinogens and trypsins of various species. *Methods Enzymol.* **19,** 41–63.
68. Mares-Guia, M. and Shaw, E. (1965) Studies on the active center of trypsin. *J. Biol. Chem.* **240,** 1579–1585.
69. Harris, J. I. (1956) Effect of urea on trypsin and α-chymotrypsin. *Nature* **177,** 471–473.
70. Anson, M. L. (1938) Estimation of pepsin, trypsin, papain and cathepsin with haemoglobin. *J. Gen. Physiol.* **22,** 79–89.
71. Schwert, G. W. and Eisenberg, M. A. (1949) Kinetics of the amidase and esterase activities of trypsin. *J. Biol. Chem.* **179,** 665–672.
72. Nord, F. F. and Bier, M. (1953) Mechanism of enzyme action. LV. Interaction between calcium and trypsin. *Biochem.Biophys. Acta* **12,** 56–60.
73. Birk, Y. (1961) Purification and some properties of a highly active inhibitor of trypsin and α-chymotrypsin from soybeans. *Biochim. Biophys. Acta* **54,** 378–381.
74. Shaw, E. and Springhorn, S. (1967) Identification of the histidine residue at the active center of trypsin labeled by TLCK. *Biochem. Biophys. Res. Commun.* **27,** 391–397.
75. Carrey, E. A. and Epand, R. M. (1983) Conformational and biological properties of glucagon fragments containing residues 1-17 and 19-29. *J. Peptide Protein Res.* **22,** 362–370.
76. Grunnet, I. and Knudsen, J. (1983) Medium-chain fatty acid synthesis by goat mammary-gland fatty acid synthetase. The effect of limited proteolysis. *Biochem. J.* **209,** 215–222.

77. Poncz, L. and Dearborn, D. G. (1983) The resistance to tryptic hydrolysis of peptide bonds adjacent to N,N-dimethyllysyl residues. *J. Biol. Chem.* **258,** 1844–1850.
78. Bentz, H., Chang, R-J., Thompson, A. Y., Glaser, C. B., and Rosen, D. M. (1990) Amino acid sequence of bovine osteoinductive factor. *J. Biol. Chem.* **265,** 5024–5029.
79. Tomasselli, A. G., Frank, R., and Schiltz, E. (1986) The complete primary structure of GTP:AMP phosphotransferase from beef heart mitochondria. *FEBS Lett.* **202,** 303–307.
80. Perides, G., Kuhn, S., Scherbarth, A., and Traub, P. (1987) Probing of the structural stability of vimentin and desmin-type intermediate filaments with Ca^{2+}-activated proteinase, thrombin and lysine-specific endoproteinase Lys-C. *Eur. J. Cell Biol.* **43,** 450–458.
81. Steffens, G. J, Gunzler, W. A., Ötting, F., Frankus, E., and Flohé, L. (1982) The complete amino acid sequence of low molecular weight urokinase from human urine. *Hoppe-Seyler's Z. Physiol. Chem.* **363,** 1043–1058.
82. Konigsberg, W., Goldstein, J., and Hill, R. J. (1963) The structure of human haemoglobin VII. The digestion of the β chain of human haemoglobin with pepsin. *J. Biol. Chem.* **238,** 2028–2033.
83. Price, N. C., Duncan, D., and McAlister, J. W. (1985) Inactivation of rabbit muscle phosphoglycerate mutase by limited proteolysis with thermolysin. *Biochem. J.* **229,** 167–171.
84. Eggerer, H. (1984) Hysteretic behaviour of citrate synthase. Site-directed limited proteolysis. *Eur. J. Biochem.* **143,** 205–212.

CHAPTER 16

Proteinase K (EC 3.4.21.14)

Patricia J. Sweeney and John M. Walker

1. Introduction

Proteinase K is a serine protease and the main proteolytic enzyme produced by the fungus *Tritirachium album* Limber *(1)*. The enzyme has a broad specificity, cleaving peptide bonds C-terminal to a number of amino acids. The enzyme is produced, together with other proteases and an aminopeptidase, during stationary phase when the fungus is grown by submerged culture. The enzyme is so named because the organism can grow on native keratin as sole carbohydrate and nitrogen source owing to the enzyme's ability to digest keratin. Because of its broad substrate specificity, high activity, and its ability to digest native proteins, proteinase K has found considerable use in procedures where the inactivation and degradation of proteins is required, particularly during the purification of nucleic acids. The enzyme retains activity in the presence of 0.5% sodium dodecyl sulfate, which is used in mammalian cell lysis. This allows the use of proteinase K in conjunction with cell lysis resulting in the rapid degradation of released intracellular nucleases and the subsequent isolation of intact nucleic acids. Following digestion, degraded protein is routinely removed by phenol extraction. For example, proteinase K has been used to degrade protein during the isolation of high-mol-wt eukaryotic DNA for cloning in phage or cosmid vectors *(2–4)*, to remove protein during plasmid *(5)* and lambda phage DNA *(6)* isolation, and to remove protein from protein DNA complexes produced during DNA footprinting analysis *(7)*. The proteinase K method for extracting RNA is a well established

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

method *(8–10)*. Because of its broad specificity and its ability to function in the presence of detergent, proteinase K has also been used extensively to study membrane protein topology and protein translocation across membranes *(11–19)*. These various uses of proteinase K are further discussed in Section 3. The use of proteinase K to degrade polypeptides to produce a mixture of oligopeptides 2–6 residues in length that is suitable for sequence analysis by GC-MS has also been reported *(20)*.

The chromosomal gene for proteinase K has been characterized and expression of the cDNA in *E. coli* has been achieved *(21)*. The crystal and molecular structure of the enzyme have also been determined by X-ray diffraction studies to 0.15 nm resolution *(22)*.

2. Enzyme Data

2.1. Specificity

The enzyme is an endoprotease having a fairly broad specificity but with a preference for cleavage of peptide bonds C-terminal to aliphatic and aromatic amino acids, especially alanine *(1)*. Although commonly used for its broad specificity, the enzyme has also been used to generate a single cleavage in a native protein. Lebherz et al. used proteinase K to cleave an isoenzyme of creatine kinase at a single position between two alanine residues *(11)*. X-ray and model building studies on the specificity of the active site of proteinase K have been described *(23)*.

2.2. Molecular Mass

Early work based on gel filtration studies suggested a mol mass of 18,500 ± 500 Da *(1)*. Later studies using SDS gel electrophoresis gave values of 27,000 Da *(24)*. The gene sequence was later determined and gave a true value for the mol mass of proteinase K of 28,930 Da *(25)*.

2.3. Isoelectric Point

The isoelectric point of the protein occurs at pH 8.9 *(1)*.

2.4. pH Optimum

The pH activity curve of proteinase K, determined for the hydrolysis of urea-denatured hemoglobin, showed optimal activity in the pH range 7.5–12.0 *(1)*. However, the enzyme is normally used in pH range 7.5–9.0 *(1,26,27)*.

2.5. Assay

The assay is based on the hydrolysis of *N*-acetyl-L-tyrosine ethylester (ATEE). One unit will hydrolyze 1 μmol of ATEE per minute at pH 9.0 and 30°C in Tris-HCl buffer, pH 8.0. Hydrolysis is monitored at 237 nm. Using this assay, the enzyme is usually supplied with a specific activity of approx 300 U/mg. Other suppliers use an assay whereby 1 Anson unit liberates 1 mmol of Folin-positive amino acids per minute at pH 7.5 and 35°C using hemoglobin as substrate. In this case, the enzyme is supplied with a specific activity of approx 20–30 U/mg of protein. Alternatively, an assay based on the hydrolysis of Suc-$(Ala)_3$-NH-Np can be used. The assay is performed in 50 m*M* Tris-HCl, 5 m*M* $CaCl_2$ (pH 8.0) using an enzyme concentration of 5–8 μg/mL and 1 m*M* substrate in a final volume of 1 mL. The reaction mixture is incubated for 60 min at 20°C and then stopped by the addition of 0.2 mL of glacial acetic acid. After a further 15 min, the absorbance at 410 nm of free nitroaniline is measured. Using this assay, the enzyme shows a specific activity of approx 13 U/mg. One unit of activity is the amount of enzyme that liberates 1 mmol of *p*-nitroaniline per minute of reaction *(26,27)*.

2.6. Stability

Proteinase K is an unusually stable enzyme. Studies have shown that the enzymic activity of proteinase K is controlled by calcium. Hence, proteinase K is normally used in the presence of approx 2.5–5 m*M* $CaCl_2$. If calcium is removed by EDTA, followed by gel filtration, enzymic activity drops to 20% of its original value within 6 h, without autolysis. Addition of excess calcium results in an immediate rise of residual activity to 28% of the original value but full activity is not restored *(27)*. The activity of the calcium bound enzyme is at a maximum at 37°C. However, the enzyme demonstrates a broad temperature profile with >80% of its maximum activity being retained between 20 and 60°C. Autolysis of the enzyme occurs during sample preparation for SDS gel electrophoresis and at low concentrations (0.01 mg/mL) in aqueous solution. It does not occur at concentrations >1 mg/mL *(26)*. Proteinase K is fully active in 0.5% (w/v) SDS and is frequently used in the presence of SDS *(2–8)*. It is also active in 1% (w/v) Triton X-100.

Stock solutions of the enzyme are normally prepared in water at 10–25 mg/mL and stored at –20°C. However, proteinase K may be stored in 50 m*M* Tris-HCl, pH 8.0, containing 1 m*M* $CaCl_2$ and is stable for at least 12 mo at 4°C *(28)*.

2.7. Inhibitors

Proteinase K appears to be a serine protease, being inhibited by diisopropyl phosphofluoridate (DFP or Dip F) and phenylmethanesulfonyl fluoride (PMSF) *(1)*. PMSF is normally added to a final concentration of 5 m*M* from a 100 m*M* dry stock solution in DMSO or isopropyl alcohol. It is not, however, inhibited by tosyl lysyl chloromethyl ketone (TLCK or Tos-Lys-CH_2Cl), an inhibitor of trypsin-like serine proteases, or tosyl phenylalanyl chloromethyl ketone (TPCK or Tos-Phe-CH_2Cl), an inhibitor of chymotrypsin-like serine proteases. It is not inhibited by sulfhydryl reagents.

3. Experimental Procedures

3.1. Nucleic Acid Isolation

A stock solution of the enzyme is diluted typically to 50–200 µg/mL in the solution to be digested, normally in the pH range 7.5–8.0 and at 37°C. Incubation times vary depending on the nature of the experiment, but can range from 30 min to 18 h. Although inhibitors of proteinase K are known (*see* Section 2.7.) they are generally not used as the proteinase K is usually denatured by subsequent phenol extractions. Nucleic acid purification protocols involving the use of proteinase K are described in detail in earlier volumes of this series *(2–8)*.

3.2. Protein Studies

When studying protein translocation, membrane topology, and enzyme compartmentation, proteinase K is added to membrane preparations (or in vitro translation systems supplemented with membrane vesicles) to degrade those portions of the proteins that are accessible. Thus, cytoplasmic proteins are completely degraded, completely translocated proteins should be fully protected, whereas transmembrane proteins and those that cross the membrane more than once will be degraded to yield discrete fragments *(11–19)*. Following proteolytic digestion, proteins are analyzed by SDS gel electrophoresis and patterns compared before and after proteolysis to identify protected polypeptides.

However, misleading results can be obtained if exposed regions of the polypeptide are inherently protease resistant. To differentiate genuine protection from protease activity and inherent protease activity, a further digestion is carried out in the presence of a nonionic detergent (e.g., 0.5–2% Triton X-100), which disrupts membranes allowing access of all proteins to proteinase K. Any polypeptide or fragment remaining after this treatment must be inherently protease resistant. The ability of proteinase K to function in the presence of detergents, and its broad specificity, makes the enzyme a very appropriate one for this type of study. Digestion of membrane is carried out at a range of concentrations (generally 20–1000 μg/mL) to find optimum conditions. Proteinase K is normally prepared fresh as a 10X stock solution in 100 m*M* $CaCl_2$. Following digestion, it is necessary to inhibit the enzyme to prevent further degradation of proteins following protein extraction for SDS gel electrophoresis.

References

1. Ebeling, W., Hennrich, N., Klockow, M., Metz, H., Orth, H. D., and Lang, H. (1974) Proteinase K from *Tritirachium album* Limber. *Eur. J. Biochem.* **47,** 91–97.
2. Mathew, C. G. P. (1984) Isolation of high molecular weight eukaryotic DNA, in *Methods in Molecular Biology, vol. 2: Nucleic Acids* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 31–34.
3. Steven, J., McKechnie, D., and Graham, A. (1988) Isolation of high molecular weight DNA suitable for the construction of genomic libraries, in *Methods in Molecular Biology, vol. 4: New Nucleic Acid Techniques* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 221–234.
4. Haley, J. D. (1988) Cosmid library construction, in *Methods in Molecular Biology, vol. 4: New Nucleic Acid Techniques* (Walker, J. M. ed.), Humana, Clifton, NJ, pp. 257–283.
5. Van Helden, P. D. and Hoal, E. G. (1988) Plasmid preparation on Sephacryl S1000, in *Methods in Molecular Biology, vol. 4: New Nucleic Acid Techniques* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 69–74.
6. Bateson, A. N. and Pollard, J. W. (1988) Construction of mammalian genomic libraries using λ replacement vectors, in *Methods in Molecular Biology*, vol. 4: *New Nucleic Acid Techniques* (Walker, J. M. ed.), Humana, Clifton, NJ, pp. 235–355.
7. Plumb, M. A. and Goodwin, G. H. (1988) Detectlon of sequence-specific protein-DNA interaction by the DNA-footprinting techniques, in *Methods in Molecular Biology*, vol. 4: *New Nucleic Acid Techniques* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 139–164.

8. McGookin, R. (1984) RNA extraction by the proteinase K method, in *Methods in Molecular Biology, vol. 2: Nucleic Acids* (Walker, J. M. ed.), Humana, Clifton, NJ, pp. 109–112.
9. Wiegers, U. and Hilz, H. (1971) A new method using 'Proteinase K' to prevent mRNA degradation during isolation from HeLa cells. *Biochem. Biophys. Res. Comm.* **44,** 513–519.
10. Fourcroy, P. (1980) Isolation of undegraded polysomes from radish cotyledons: Use of Proteinase K and cycloheximide. *Phytochemistry* **19,** 7–10.
11. Lebherz, H. G., Burke, T., Shackelford, J. E., Strickler, J. E., and Wilson, K. J. (1986) Specific proteolytic modification of creatine kinase isoenzymes. *Biochem. J.* **233,** 51–56.
12. Randall, L. L. (1983) Translocation of domains of nascent periplasmic proteins across the cytoplasmic membrane is independent of elongation. *Cell* **33,** 231–240.
13. Lipp, J. and Dobberstein, B. (1986) Signal recognition particle-dependent membrane insertion of mouse invariant chain: A membrane spanning protein with a cytoplasmically exposed amino terminus. *J. Cell Biol.* **102,** 2169–2175.
14. Zimmermann, R. and Mollay, C. (1986) Import of honeybee prepromelittin into the endoplasmic reticulum. *J. Biol. Chem.* **261,** 12,889–12,895.
15. Imanaka, T., Small, G. M., and Lazarow, P. B. (1987) Translocation of acyl-CoA oxidase into peroxisomes requires ATP hydrolysis but not membrane potential. *J. Cell Biol.* **105,** 2915–2922.
16. Schleyer, M. and Neupert, W. (1985) Transport of proteins into mitochondria: Translocational intermediates spanning contact sites between outer and inner membranes. *Cell* **43,** 339–350.
17. Müller, M. and Blobel, G. (1984) *In vitro* translocation of bacterial proteins across the plasmamembrane of *Escherichia coli. Proc. Natl. Acad. Sci. USA* **81,** 7421–7425.
18. Wolfe, P. B. and Wickner, W. (1984) Bacterial leader peptidase, a membrane protein without a leader peptide, uses the same export pathway a presecretory proteins. *Cell* **36,** 1067–1072.
19. Connolly, T., Collins, P., and Gilmore, R. (1989) Access of Proteinase K to partially translocated nascent polypeptides in intact and detergent solubilized membranes. *J. Cell Biol.* **108,** 299–307.
20. Krutzsch, H. C. (1986) Polypeptide sequence analysis using gas chromatography-mass spectrometry, in *Methods of Protein Microcharacterization: A Practical Handbook* (Shively, J. E., ed.), Humana, Clifton, NJ, pp. 381–401.
21. Gunkel, F. A. and Gassen, H. G. (1989) Proteinase K from *Tritirachium album* Limber. Characterisation of the chromosomal gene and expression of the cDNA in *Escherichia coli. Eur. J. Biochem.* **179,** 185–194.
22. Betzel, C., Pal, G. P., and Saenger, W. (1988) Three-dimensional structure of Proteinase K at 0.15nm resolution. *Eur. J. Biochem.* **178,** 155–171.
23. Betzel, C., Bellemann, M., Pal, G. P., Bajorath, J., Saenger, W., and Wilson, K. S. (1988) X-ray and model building studies on the specificity of the active site of Proteinase K. *Proteins* **4,** 157–164.

24. Pähler, A., Banerjee, A., Dattagupta, J. K., Fujiwana, T., Lindner, K., Pal, G. P., Suck, D., Weber, G., and Saenger, W. (1984) Three-dimensional structure of fungal Proteinase K reveals similarity to bacterial subtilisin. *EMBO J.* **3,** 1311–1313.
25. Jany, K. D., Lederer, G., and Mayer, B. (1986) Amino acid sequence of Proteinase K from mold *Tritirachium album* Limber. Proteinase K—a subtilisin related enzyme with disulfide bonds. *FEBS Lett.* **199,** 139–144.
26. Bajorath, J., Saenger, W., and Pal, G. P. (1988) Autolysis and inhibition of Proteinase K, a subtilisin-related serine proteinase isolated from the fungus *Tritirachium album* Limber. *Biochem. Biophys. Acta.* **954,** 176–182.
27. Bajorath, J., Hinrichs, W., and Saenger, W. (1988) The enzymatic activity of proteinase K is controlled by calcium. *Eur. J. Biochem.* **176,** 441–447.
28. Keesey, J. (ed.) (1989) *Biochemical Information —A Revised Biochemical References Source.* Pub. Boehringer Mannheim Biochemicals, Indianapolis, IN.

CHAPTER 17

Carboxypeptidase Y (EC 3.4.16.1)

Julia S. Winder and John M. Walker

1. Introduction

Carboxypeptidases are proteolytic enzymes that remove L-amino acids, one residue at a time, from the carboxyl terminus of polypeptide chains, i.e., they are exoproteases. A number of such enzymes have been isolated from plant and animal sources, each differing in their chemical and physical properties and the rate at which they release particular amino acids. The major use of carboxypeptidases in molecular biology is in the determination of the C-terminal amino acid sequence of peptides and proteins (no suitable chemical method exists for the sequential removal of C-terminal amino acids from a polypeptide). The protein or peptide being analyzed is digested with carboxypeptidase and aliquots removed at timed intervals, and analyzed for the presence of free amino acids. The amount of each amino acid released is plotted against time, and the C-terminal sequence deduced from the relative rate of release of each amino acid. Four carboxypeptidases have been used extensively to provide peptide and protein sequence data. These are: carboxypeptidase A (EC 3.4.17.1) from bovine pancreas *(1)*, carboxypeptidase B (EC 3.4.17.2) from porcine pancreas *(1)*, carboxypeptidase C (EC 3.4.12.1) from orange leaves *(2)*, and carboxypeptidase Y (EC 3.4.16.1) from yeast *(3)*. Historically, carboxypeptidases A and B were the first to be discovered and used for sequence determination. Carboxypeptidase A releases most C-terminal amino acids, but will not cleave at arginine, proline, and

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

hydroxyproline, or lysine residues *(1)*. The specificity of carboxypeptidase B is far more restricted, cleaving only C-terminal arginine and lysine residues. Carboxypeptidases A and B, therefore, tended to be used together, but even so, exopeptidase activity was effectively blocked when a proline residue was reached. The isolation of carboxypeptidase C provided an enzyme that combined the specificity of carboxypeptidases A and B, but also cleaved at Pro residues *(2)*, i.e., it cleaves all C-terminal amino acids. Carboxypeptidase Y (CPY, isolated from baker's yeast) has the same broad specificity as carboxypeptidase C, but because of its strong action on protein substrates and its ability to work in the presence of urea and detergents, it is nowadays the enzyme of choice for C-terminal sequence analysis and will be described in this chapter.

2. The Enzyme

2.1. Purification

Large-scale purification of the enzyme from baker's yeast has been described by a number of workers *(3,4)*.

2.2. Specificity

The enzyme cleaves all L-amino acids one residue at a time from the C-terminal of polypeptide chains. However, the rate of release of individual amino acids varies. Catalysis is maximum when the penultimate and/or terminal residues have aromatic or aliphatic side chains *(5)*. When glycine or aspartic acid is in the terminal position, or lysine and arginine in the penultimate position, the release of the amino acid is slow *(3,6)*. The cleavage of tripeptides is difficult, and dipeptides are completely resistant to hydrolysis. C-terminal proline is a good substrate, but a proline residue on the carboxy terminal side of glycine is not likely to be released *(5)*.

The enzyme is a serine carboxypeptidase having a strongly nucleophilic serine residue at the active center *(7,8)*, which is generated by a charge relay system involving a histidine residue *(9)*. It has strong esterase activity toward the substrates of chymotrypsin and also anilidase activity *(10)*. It therefore seems to be quite similar to chymotrypsin in both mechanism and active site, although carboxypeptidase Y is an exopeptidase and chymotrypsin an endopeptidase.

2.3. Physical and Chemical Properties

CPY has a mol mass of 61,000–63,000 Da as determined by SDS gel electrophoresis *(3,11)*. It is a glycoprotein, having a single polypeptide chain of about 430 residues with 16 residues of glucosamine in the carbohydrate moiety. About 96%, i.e., 416 of approx 430 amino acid residues, have been sequenced with tentative assignment of the carbohydrate attachment sites, and the molecule shown to contain one sulfhydryl group, either four or five disulfide bridges, and five methionine residues *(12,13)*.

2.4. pH Optimum

The optimum pH for hydrolysis of acidic amino acids is pH 5.5, while that for neutral and basic amino acids is pH 7 *(5,14)*.

2.5. Assay

The substrate N-carbobenzoxy-Phe-Ala (N-CBZ-Phe-Ala) is prepared as a 100-m*M* stock solution in methanol and diluted when required in 50 m*M* MES buffer, pH 6.75, to give a substrate solution that is 2 m*M* in N-CBZ-Phe-Ala and contains 2% methanol. Ten micrograms of enzyme are added to 3 mL of substrate solution, and the decrease in absorbance at 230 nm measured with time. One unit is defined as the amount of enzyme that will hydrolyze 1.0 µmol of N-CBZ-Phe-Ala to N-CBZ-L-phenylalanine and L-alanine/min at pH 6.75 at 25°C (E^{M}_{230} = 191.5). Using this assay, the enzyme is normally provided with a specific activity of 100–130 U/mg of protein. Esterase and anilidase assays have also been well documented *(10)*.

2.6. Stability

The lyophilized enzyme is stable for 6–12 mo if stored at 4°C. A suspension of the enzyme in saturated ammonium sulfate can be stored at –20°C indefinitely. When dissolved in water and dialyzed against water to give an approx 1% solution of the enzyme, this solution can be aliquoted and stored at –20°C for at least 2 yr. Diluted solutions (<0.1 mg/mL) lose activity fairly quickly and should therefore be prepared just before use *(5)*. The enzyme is fully stable in the presence of 10% methanol at pH 5.5–8.0 for 8 h at 25°C, and in 20% methanol, 80% activity remains after incubation at pH 7 for 24 h *(15)*. Enzyme activity is rapidly lost below pH 3 *(3)* or above 60°C *(11)*.

2.7. Denaturing Agents

Eighty percent of activity remains after incubation with 6*M* urea at 25°C for 1 h *(3)*. The enzyme also retains its activity for extended periods in 1% SDS *(16)*. Its stability to denaturing agents makes CPY very suitable for studying proteins that have inaccessible or poorly accessible C-termini under normal (nondenaturing) conditions *(16)*.

2.8. Inhibitors

The enzyme is a serine protease, and is therefore inhibited by DFP *(17)* and PMSF *(7)*, but is not inhibited by soy bean and lima bean trypsin inhibitors. It is inhibited by *p*-hydroxymercuribenzoate, probably by reaction with a single thiol group thought to be located near or at the substrate binding site *(3,8)*. Enzyme activity is affected by metal ions: Cu^{2+}, Ag^{+}, and Hg^{2+} result in complete loss of activity at $10^{-4}M$, 1 m*M* Cu^{+}, Mg^{2+}, Ca^{2+}, Ba^{2+}, Cr^{2+}, Mn^{2+}, Fe^{2+}, Fe^{3+}, Co^{2+}, or Ni^{2+} results in loss of more than 50% activity *(7)*. Certain organic solvents, such as DMF and ethanol, apparently show competitive inhibition *(5)*. EDTA and *o*-phenanthroline have no effect on enzyme activity *(10)*.

2.9. Additional Comments

Commercial preparations may contain free amino acids owing to autolysis, which should be removed before use. Repeated freeze-thawing of solutions of the enzyme or prolonged storage at room temperature can also lead to autolysis and the liberation of free amino acids. For ammonium sulfate suspensions, centrifuge and wash the pellet in saturated ammonium sulfate before dissolving in buffer. Alternatively, the dissolved enzyme can be dialyzed against the pyridine acetate buffer used for digestion (*see* Section 3.).

3. Experimental Procedure

3.1. Determination of C-Terminal Sequences in Peptides and Proteins (16)

Dissolve 20 nmol of the protein to be studied in 200 µL of digestion buffer (0.1*M* pyridine acetate, 0.1 m*M* norleucine, pH 5.6, containing 1% SDS). The norleucine is used as an internal standard for amino acid analysis to allow for compensation of any handling losses or sampling errors. Heat the solution to 60°C for 20 min to denature the protein. After cooling, remove a 25-µL aliquot as the zero time sample. Add 2

nmol of carboxypeptidase Y in 5–10 µL of 0.1*M* pyridine acetate buffer, pH 5.6, thoroughly mix, incubate at room temperature, and remove 25-µL aliquots at $T = 1, 2, 5, 10, 20, 30$, and 60 min. Add 5 µL of glacial acetic acid to each sample to stop the reaction. Samples are then frozen, lyophilized (pyridine acetate is volatile), and then subjected to amino acid analysis. The small amount of SDS in each sample applied to the amino acid analysis does not interfere with the elution profile or affect the integrity of the machine. The sampling times indicated here should be appropriate for most proteins. However, should the C-terminal sequence be such that a number of slowly released amino acids are present, then the experiment may have to be repeated using longer incubation times or a higher enzyme-to-substrate ratio.

If analyzing peptides, SDS can be omitted from the incubation buffer, since the C-terminal should be readily accessible. In addition, the time intervals for sampling can be reduced, since the rate of appearance of free amino acids will be faster than that for proteins. Use $T = 0, {}^1/_2, 1, 2, 5, 10, 15$, and 20 min. A graph is plotted showing the amount of each amino acid released with time and the C-terminal sequence deduced from the relative rate of release of each amino acid (*see*, e.g., ref. *16*).

References

1. Neurath, H. (1960) Carboxypeptidases A and B, in *The Enzymes,* vol. 4 (Boyer, P. D., Lardy, H., and Myrbäck, K., eds.), Academic, New York, pp. 11–36.
2. Sprossler, B., Heilmann, H. D., Grampp, E., and Uhlig, H. (1971) Carboxypeptidase C from orange leaves. *Hoppe-Seyler's Z. Physiol. Chem.* **352,** 1524–1530.
3. Hayashi, R., Moore, S., and Stein, W. H. (1973) Carboxypeptidase from yeast. *J. Biol. Chem.* **248,** 2296–2302.
4. Johansen, J. T., Breddam, K., and Ottesen, M. (1976) Isolation of carboxypeptidase Y by affinity chromatography. *Carlsberg Res. Commun.* **41,** 1–14.
5. Hayashi, R., Bai, Y., and Hata, T. (1975) Kinetic studies of carboxypeptidase Y. *J. Biochem.* **77,** 69–79.
6. Breddam, K. and Ottesen, M. (1987) Determination of C-terminal sequences by digestion with serine carboxypeptidases. The influence of enzyme specificity. *Carslberg Res. Commun.* **52,** 55–63.
7. Hayashi, R., Bai, Y., and Hata, T. (1975) Further confirmation of carboxypeptidase Y as a metal free enzyme having a reactive serine residue. *J. Biochem.* **77,** 1313–1318.
8. Hayashi, R., Moore, S., and Stein, W. H. (1973) Serine at the active centre of yeast carboxypeptidase. *J. Biol. Chem.* **248,** 8366–8369.
9. Hayashi, R., Bai, Y., and Hata, T. (1975) Evidence for an essential histidine in carboxypeptidase Y. *J. Biol. Chem.* **250,** 5221–5226.

10. Hayashi, R. (1976) Carboxypeptidase Y, in *Methods in Enzymology,* vol. 45 (Lorand, L., ed.), Academic, New York, pp. 568–587.
11. Kuhn, R. W., Walsh, K. A., and Neurath, H. (1974) Isolation and partial characterisation of an acid carboxypeptidase from yeast. *Biochemistry* **13,** 3871–3877.
12. Martin, B. M., Svendsen, I., Viswanatha, T., and Johansen, J. T. (1982) Amino acid sequence of carboxypeptidase Y. 1. Peptides from cleavage with cyanogen bromide. *Carlsberg Res. Commun.* **47,** 1–13.
13. Svendsen, I., Martin, B. M., Viswanatha, T., and Johansen, J. T. (1982) Amino acid sequence of carboxypeptidase Y. II. Peptides from enzymatic cleavages. *Carlsberg Res. Commun.* **47,** 15–27.
14. Bai, Y., Hayashi, R., and Hata, T. (1975) Kinetic studies of carboxypeptidase Y. *J. Biochem.* **78,** 617–626.
15. Hayashi, R., Bai, Y., and Hata, T. (1974) Inhibition of carboxypeptidase Y by chloromethyl ketone derivatives of benzyloxycarbonyl-L-phenylalanine. *J. Biochem.* **76,** 1355–1357.
16. Klemm, P. (1984) Carboxyl-terminal sequence determination of proteins and peptides with carboxypeptidase Y, in *Methods in Molecular Biology, vol. 1: Proteins* (Walker, J. M., ed.), Humana, Clifton, NJ, pp. 255–259.
17. Doi, E., Hayashi, R., and Hata, T. (1967) Purification of yeast proteinases. *Agr. Biol. Chem.* **31,** 160–169.

CHAPTER 18

Aminopeptidases

Aminopeptidase M (EC 3.4.11.2), Pyroglutamate Aminopeptidase (EC 3.4.19.3), and Prolidase (EC 3.4.13.9)

Patricia J. Sweeney and John M. Walker

1. Introduction

Aminopeptidases are proteolytic enzymes that remove L-amino acids sequentially from the amino termini of polypeptide chains. A number of aminopeptidases have been isolated, including leucine aminopeptidase from serine kidney cytosol *(1)*, aminopeptidase P from *E. coli (2)*, proline iminopeptidase from *E. coli*, and swine kidney *(3)*, aminopeptidase B from rat liver *(4)*, and aminopeptidase A from rat kidney *(5)*. However, three aminopeptidases in particular have found routine use in protein chemistry. The first is pyroglutamate aminopeptidase, a thiol exoprotease that cleaves N-terminal pyroglutamyl residues (pyrrolidone carboxylic acid) from peptides and proteins *(6–10)*. N-terminal glutamine residues can readily cyclize to the pyroglutamyl derivative (Fig. 1). This can occur during peptide and protein purification (it is uncertain whether the N-terminal pyroglutamyl residues of a number of naturally occurring peptides and proteins are genuine posttranslational modifications, or were introduced by cyclization of N-terminal glutamine during purification) or during sequence determination when glutamine was the newly liberated N-terminal amino acid. This cyclized derivative does not have a free amino group, and

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

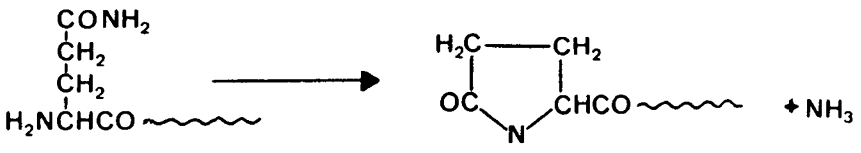


Fig. 1. The cyclization of N-terminal glutamine to pyroglutamic (pyrrolidone carboxylic) acid.

therefore, the peptide or protein is not amenable to sequence determination, unless the pyroglutamyl derivative is removed by pyroglutamate aminopeptidase *(11,12)*. The enzyme was first purified from *Pseudomonas fluoresens (12)*, but nowadays the calf liver enzyme is used, and it is this enzyme that we describe.

Aminopeptidase M, a zinc-containing metalloprotease, from swine kidney microsomes *(13–16)* removes amino acids sequentially from the N-terminals of peptides and proteins. It, therefore, has some use in the determination of N-terminal sequence data, although the Edman degradation would probably be the method of choice for most workers. Aminopeptidase M is more frequently used in the preparation of peptide and protein hydrolysates for amino acid analysis. Traditionally, peptides and proteins are hydrolyzed in 6*N* HCl, but this approach results in the total loss of tryptophan, partial loss (5–10%) of serine and threonine, and hydrolysis of asparagine and glutamine to the corresponding acids. The use of enzymes to produce a peptide/protein hydrolysate overcomes these problems. Although aminopeptidase M is capable of cleaving all possible peptide bonds, in practice, the X—Pro bond is not completely cleaved and the dipeptide X—Pro is released *(17)*.

Aminopeptidase M therefore tends to be used in conjunction with our third enzyme, prolidase. Prolidase, a manganese-containing metalloprotease, has been purified from a number of sources, but the porcine kidney enzyme is generally used. It is more correctly a highly specific *imido*peptidase, since it cleaves the dipeptide X—Pro or X—Hypro *(18)*. The amino acid sequence and gene location of human prolidase have been elucidated *(19)*, and active site modeling studies of the enzyme have been performed *(20)*.

2. Enzyme Data

2.1. Pyroglutamate Aminopeptidase (Calf Liver)

2.1.1. Alternative Names

These are L-pyroglutamyl peptide hydrolase, 5-oxoprolyl-peptidase, pyrrolidonyl peptidase, pyroglutamate aminopeptidase, pyroglutamyl peptidase, and pyroglutamase. *Note:* This enzyme is now classified as EC 3.4.19.3. In earlier literature, the enzyme was classified as EC 3.4.11.8.

2.1.2. Specificity

The enzyme cleaves N-terminal pyroglutamyl residues from peptides and proteins, but not if the following residue is proline. The enzyme has highest specificity when the pyroglutamate residue is linked to alanine *(13)*.

2.1.3. Molecular Mass

A mol mass of 79,000–80,000 Da has been reported for the enzyme, based on gel filtration studies *(21)*.

2.1.4. pH Optimum

The enzyme is active at pH 7–9, but is normally used at pH 8.0 *(22)*.

2.1.5. Assay

The assay is based on the hydrolysis of L-pyroglutamic acid β-naphthylamide. The enzyme buffer used is prepared as follows: 1 L of 0.1*M* Na_2HPO_4 is adjusted to pH 8 with 0.1*M* NaH_2PO_4 to give 1605 mL, and this solution is made 5% (v/v) in glycerol. A deblocking buffer is prepared by making 112 mL of the enzyme buffer, 5 m*M* in DTT, and 10 m*M* in Na_2 EDTA, adjusting the solution to pH 8 and purging with N_2. Enzyme solution (0.4 mL) and 0.5 mL of deblocking buffer are combined, purged with argon, and sealed with a cap. The mixture is incubated for 3 min at 37°C, the substrate added (1.78 mmol pyroglutamyl-β-naphthylamide in 0.1 mL of methanol), and the mixture again incubated for 5 min at 37°C. One milliliter of 25% TCA solution is then added to stop the reaction. To quantify the β-naphthylamine released, 1 mL of this solution is mixed with 1 mL of 0.1% $NaNO_3$. After 3 min, 1 mL of a 0.5% ammonium sulfamate solution is added to destroy excess nitrite. After 5 min, 2 mL of 0.05% *N*-1-naphthylethylenediamine dihydrochloride solution is added, and the solution

incubated at 37°C for 1 h. The solution is then cooled, and the absorbance recorded at 570 nm. The amount of β-naphthylene can be determined using a standard curve *(23)*. Using this assay, the enzyme is normally supplied with a specific activity of 100–350 U/mg protein, where 1 U hydrolyzes 1 nmol of L-pyroglutamic acid β-naphthylamide to L-pyroglutamic acid and β-naphthylamine/min. The enzyme has also been assayed using (Pro-^{3}H)-thyroliberin as substrate *(24)*, and by HPLC methods that use peptides (containing pyroglutamate residues) as substrates *(25)*.

2.1.6. Stability

The lyophilized enzyme is stable at 4°C for months. The enzyme is generally unstable. Its stability is enhanced by sucrose and EDTA in some commercial preparations. It may be reconstituted in solutions containing 5 m*M* DTT and 10 m*M* EDTA, and stored at –20°C, or it can be used for a maximum of 1 wk, if stored at 4°C. The enzyme is stable in 1 *M* urea and 0.1 *M* guanidine hydrochloride *(21)*. Podell and Abraham *(26)* have found the enzyme to be extremely unstable above room temperature, and it was found that deblocking did not occur at 37°C. It was found that an initial incubation at 4°C followed by a second incubation at room temperature (*see* Section 3.1.) was necessary to ensure maximum enzyme stabilizing conditions. Air oxidation causes severe reduction in enzyme activity, and therefore, incubations involving the enzyme should be carried out under nitrogen.

2.1.7. Activation / Inhibition

Podell and Abraham *(26)* have studied factors affecting the stability of the enzyme and have developed a buffer (deblocking buffer; *see* Section 3.1.) that is compatible with the use of the enzyme. The enzyme requires a thiol compound for activation and is inactivated by thiol-blocking compounds, such as iodoacetamide, and divalent metal ions, such as Hg^{2+}. Activity can be restored by short incubations with mercaptoethanol. Dithiothreitol and EDTA are included in the deblocking buffer to overcome inactivation *(6)*.

2.2. Aminopeptidase M (Porcine Kidney Microsomes)

2.2.1. Alternative Names

These are amino acid arylamidase, microsomal alanyl aminopeptidase, and α-aminoacyl peptide hydrolase.

2.2.2. Specificity

The enzyme cleaves N-terminal residues from all peptides having a free α-amino or α-imino group. However, in peptides containing an *X*—Pro sequence, where *X* is a bulky hydrophobic residue (Leu, Tyr, Trp, Met sulfone), or in the case of an N-blocked amino acid, cleavage does not occur. It is for this reason that prolidase (Section 2.3.) is used in conjunction with aminopeptidase M to produce total hydrolysis of peptides.

2.2.3. Molecular Mass

The enzyme has a mol mass of about 280,000 Da based on gel filtration and is composed of 10 subunits (28,000 ± 3000 Da) of two different types *(27)*. The enzyme molecule contains five disulfide bridges, each of which connects two subunits *(27)*. The gene has been cloned, the amino acid sequence determined, and mol mass confirmed *(27,28)*.

2.2.4. pH Optimum

At substrate concentrations used for sequence work, the pH optimum is between 7.0 and 7.5, but at higher substrate concentrations, approaches 9.0 *(15)*.

2.2.5. Assay

The assay is based on hydrolysis of *p*-nitroanilides of amino acids, especially alanine or leucine. The enzyme is added to 0.06*M* phosphate buffer, pH 7.0, containing 1.66 m*M* L-leucine *p*-nitroanilide or L-alanine-*p*-nitroanilide to give a final vol of 2.0 mL. The increase in absorbance at 405 nm is recorded at 37°C. One enzyme unit is defined as the amount of enzyme that produces 1 μmol *p*-nitroanilide/min at 37°C *(16)*. The enzyme is normally supplied with a specific activity of 4 U/mg protein *(16)*. Alternatively, an assay based on the hydrolysis of L-leucinamide at pH 8.5 and 25°C is used. Using this assay, the enzyme is usually supplied with a specific activity of 25 U/mg protein. The hydrolysis of phenylalanyl-3-thia-phenylalanine at pH 8.2 and 25°C has also been used as the basis of an assay that distinguishes between leucine aminopeptidase and aminopeptidase M *(29)*.

2.2.6. Stability

The lyophilized enzyme is stable for several years at –20°C. A working solution can be prepared by dissolving about 0.25 mg of protein in 1 mL of deionized water to give a solution of approx 6 U of activity/mL. This solution can be aliquoted and stored frozen for sev-

eral months at –20°C. The enzyme is reported to be stable at pH 7.0 at temperatures up to 65°C, and is stable between pH 3.5 and 11.0 at room temperature for at least 3 h *(15)*.

2.2.7. Activation / Inhibition

The enzyme is not affected by sulfhydryl reagents, has no requirements for divalent metal ions (unlike the cytosolic leucine aminopeptidase), is stable in the presence of trypsin, and is active in 6*M* urea. It is not inhibited by PMSF, DFP, or PCMB. It is, however, irreversibly denatured by alcohols and acetone, and 0.5*M* guanidinium chloride, but cannot be precipitated by trichloroacetic acid *(15)*. It is inhibited by 1,10-phenanthroline (10*M*) *(16)*.

2.3. Prolidase (Porcine Kidney)

2.3.1. Alternative Names

These are imidodipeptidase, proline dipeptidase, aminoacyl L-proline hydrolyase, and peptidase D.

2.3.2. Specificity

The enzyme is highly specific, and cleaves dipeptides with a prolyl or hydroxyprolyl residue in the carboxyl terminal position *(18,30)*. It has no activity with tripeptides *(19)*. The rate of release is inversely proportional to the size of the amino terminal residue *(19)*. The enzyme's activity depends on the nature of the amino acids bound to the imino acid. For optimal activity, amino acid side chains must be as small as possible and apolar to avoid steric competition with the enzyme receptor site. The enzyme has the best affinity for alanyl proline and glycyl proline. Experimental data have also suggested that prolidase may only cleave the *trans*-form of the peptide bond *(31)*. Active site modeling studies have also been performed *(20)*.

2.3.3. Molecular Mass

The mol mass, determined by SDS-PAGE, is 110,000 Da for the native protein and 53,000 Da for the reduced form. The enzyme is therefore thought to consist of two chains linked by disulfide bonds *(30)*. Prolidase is a glycoprotein containing about 0.5% carbohydrate *(32)*.

2.3.4. pH Optimum

The enzyme has optimal activity at pH 6–8, but it is normally used at pH 7.8–8.0 *(33)*.

2.3.5. Assay

An assay based on the hydrolysis of glycyl-L-proline at 37°C has been described. An incubation mixture (0.5 mL) containing 0.1 mL enzyme preparation in 50 m*M* glycyl-L-proline, 50 m*M* Tris-HCl, pH 7.8, and 1 m*M* $MnCl_2$ is prepared and incubated for 1 h at 37°C. The reaction is then stopped with 1.0 mL of 0.45*M* TCA, and the quantity of proline produced measured using the Chinard colorimetric method *(34)*, with L-proline standards. Chinard's reagent consists of 600 mL glacial acetic acid, 400 mL orthophosphoric acid (6*M*), and 25 g ninhydrin, dissolved at 70°C. One milliliter of glacial acetic acid and 1 mL of Chinard's reagent are added to 0.5 mL of reaction mixture. After 10 min at 90°C, absorbance is read at 515 nm. Dilutions should be made in 0.45*M* TCA. Using this assay, the enzyme is usually supplied with a specific activity of 200–300 U/mg of protein *(33,35)*. Measurement of prolidase activity using an Ala-Pro substrate, isotachophoresis, or chromatographic methods has also been described *(36,37)*.

2.3.6. Stability

The enzyme is most stable at pH 6–8 in the presence of 0.01*M* $MnCl_2$. The lyophilized enzyme is stable for many months when stored at –20°C *(38)*, and is stable for several weeks at 4°C if stored in the presence of 2 m*M* $MnCl_2$ and 2 m*M* β-mercaptoethanol *(30)*.

2.3.7. Activation / Inhibition

Manganous ions are essential for optimal catalytic activity. The enzyme is inhibited by 4-chloromercuribenzoic acid, iodoacetamide, EDTA, fluoride, and citrate. However, if Mn^{2+} is added before iodoacetamide, no inhibition is observed *(31)*.

3. Experimental Procedures

3.1. Removal of Pyroglutamic Acid from the N-Terminus of a Protein or Peptide

This method is based on the procedure described by Podell and Abraham *(26)*. The sample (10 mg) is dissolved in 10 mL of deblocking buffer (0.1*M* sodium phosphate buffer, pH 8.0, 5 m*M* DTT, 10 m*M* EDTA, 5% glycerol) and placed in a screw-top vial. Approximately 25 μg of enzyme are then added, the vial flushed with nitrogen, and sealed. Protein and enzyme are allowed to incubate at 4°C for 9 h with

occasional mixing. Another 25 µg of enzyme are then added, and incubation continued under nitrogen at room temperature for a further 14 h. Since the purpose of deblocking is invariably to render the protein amenable to sequence analysis, the protein can be desalted either by dialysis against 0.05*M* acetic acid and then lyophilized, or by passage through an HPLC gel-filtration column in 0.1% aqueous TFA.

3.2. N-Terminal Sequence Determination by Time-Course Hydrolysis with Aminopeptidase M (39)

Dissolve the polypeptide (1 nmol) in 49 µL of 0.2*M* sodium phosphate buffer, pH 7.0. Add 1 µL (0.005 U) of aminopeptidase M solution, and incubate at room temperature. Timed aliquots (5 µL) are then removed (0, 15, 30, and 45 min, and so on) and analyzed by amino acid analysis for the presence of free amino acids. The amount of each amino acid released is plotted against time, and the order of amino acids deduced from the graph.

3.3. The Use of Aminopeptidase M and Prolidase to Produce a Peptide/Protein Hydrolysate for Amino Acid Analysis (40)

Dissolve the polypeptide (1 nmol) in 24 µL of 0.2*M* sodium phosphate buffer, pH 7.0. Add 1 µg of aminopeptidase M (1 µL), and incubate at 37°C. For peptides containing 2–10 residues, 8 h are sufficient for complete digestion. For larger peptides (11–35 residues), a further addition of enzyme after 8 h is needed, followed by a further 16-h incubation. For polypeptides containing more than 35 residues, digestion with aminopeptidase M alone is insufficient. The polypeptide is first digested with the nonspecific protease pronase (1%, v/v), followed by aminopeptidase M *(41,42)*.

Since in many cases the X—Pro— bond is not completely cleaved by these enzymes, to ensure complete cleavage of proline containing polypeptides, the aminopeptidase M digest should be treated with 1 µg of prolidase for 2 h at 37°C before analysis. The use of two or more proteolytic enzymes to produce hydrolysis will often lead to an increase in the level of background amino acids. A digestion blank should therefore also be analyzed in order to correct for the background amino acids.

References

1. Himmelhoch, S. R. (1970) Leucine aminopeptidase from swine kidney. *Methods Enzymol.* **19,** 508–513.
2. Yaron, A. and Mlynr, D. (1968) Aminopeptidase P. *Biochem. Biophys. Res. Commun.* **32(4),** 658–663.
3. Sarid, S., Berger, A., and Katchalski, E. (1962) Proline iminopeptidase: purification and comparison with iminodipeptidase (Prolinase). *J. Biol. Chem.* **237(7),** 2207–2212.
4. Hopsu, V. K., Makinen, K. K., and Glenner, G. G. (1966) Characterization of aminopeptidase B. Substrate specificity and affector studies. *Arch. Biochem. Biophys.* **114,** 567–575.
5. Glenner, G. G., McMillan, P. J., and Folk, J. E. (1962) A mammalian peptidase specific for the hydrolysis of N-terminal α-L-glutamyl and aspartyl residues. *Nature* **194,** 867.
6. Doolittle, R. F. (1970) Pyrrolidone carboxyl peptidase. *Methods Enzymol.* **19,** 555–569.
7. Szewczuk, A. and Kwiatkowska, J. (1970) Pyrrolidonyl peptidase in animal, plant and human tissues. Occurrence and some properties of the enzyme. *Eur. J. Biochem.* **15,** 92–96.
8. Mudge, A. W. and Fellows, R. E. (1973) Bovine pituitary pyrrolidone carboxyl peptidase. *Endocrinology* **93,** 1428–1434.
9. Browne, P. and O'Cuinn, G. (1983) An evaluation of the role of a pyroglutamyl peptidase, a post-proline cleaving enzyme and a post-proline dipeptidyl aminopeptidase, each purified from the soluble fraction of a guinea-pig brain, in the degradation of thyroliberin in vitro. *Eur. J. Biochem.* **137,** 75–87.
10. O'Connor, B. and O'Cuinn, G. (1985) Purification and kinetic studies on narrow specificity synaptosomal membrane pyroglutamate aminopeptidase from guinea pig brain. *Eur. J. Biochem.* **150,** 47–52.
11. Brandt, A., Glanville, R. W., Hörlein, D., Bruckner, P., Timpl, R., Fietzek, P. P., and Kühn, K. (1984) Complete amino acid sequence of the N-terminal extension of calf-skin type III procollagen. *Biochem. J.* **219,** 625–634.
12. Doolittle, R. F. and Armentrout, R. W. (1968) Pyrrlidonyl peptidase. An enzyme for selective removal of pyrrolidonecarboxylic acid residues from polypeptides. *Biochemistry* **7,** 516–521.
13. Pfleiderer, G., Celliers, P. G., Stanulovic, M., Wachsmuth, E. D., Determann, H., and Braunitzer, G. (964) Eizenschaften und analytische anwendung der aminopeptidase aus nierenpartikceln. *Biochem. Z.* **340,** 552–564.
14. Pfleiderer, G. and Celliers, P. G. (1963) Isolation of an aminopeptidase in kidney tissue. *Biochem. Z.* **339,** 186–189.
15. Wachsmuth, E. D., Fritze, I., and Pleiderer, G. (1966) An aminopeptidase occurring in pig kidney. An improved method of preparation. Physical and enzymic properties. *Biochemistry* **5(1),** 169–174.
16. Pfleiderer, G. (1970) Particle bound aminopeptidase from pig kidney. *Methods Enzymol.* **19,** 514–521.

17. Light, A. (1972) Leucine aminopeptidase in sequence determination of peptides. *Methods Enzymol.* **25B,** 253–262.
18. Sjöstrom, H., Noren, O., and Jossefsson, L. (1973) Purification and specificity of pig intestinal prolidase. *Biochim. Biophys. Acta* **327,** 457–470.
19. Endo, F., Tanoue, A., Nakai, H., Hata, A., Indo, Y., Titani, K., and Matsuela, I. (1989) Primary structure and gene localization of human prolidase. *J. Biol. Chem.* **264(8),** 4476–4481.
20. King, G. F., Crossley, M. J., and Kuchel, P. W. (1989) Inhibition and active-site modelling of prolidase. *Eur. J. Biochem.* **180,** 377–384.
21. Keesey, J. (ed.) (1989) *Biochemica Information: A Revised Biochemical Reference Source.* Boehringer Mannheim Biochemicals, Indianapolis, IN.
22. Personal communication. Boehringer Mannheim.
23. Capecchi, J. T. and Loudon, G. M. (1985) Substrate specificity of pyroglutamylaminopeptidase. *J. Med. Chem.* **28,** 40–143.
24. O'Connor, B. and O'Cuinn, G. (1984) Localization of a narrow-specificity thyroliberin hydrolyzing pyroglutamate aminopeptidase in synaptosomal membranes of guinea-pig brain. *Eur. J. Biochem.* **144,** 271–278.
25. Martini, F., Bossa, F., and Barra, D. (1985) Assay of pyroglutamyl aminopeptidase by high-performance liquid chromatoraphy and its use in peptide sequencing. *Peptides* **6,** 103–105.
26. Podell, D. N. and Abraham, G. N. (1978) A technique for the removal of pyroglutamic acid from the amino terminus of proteins using calf liver pyroglutamate amino peptidase. *Biochem. Biophys. Res. Commun.* **81,** 176–185.
27. Wachsmuth, E. D. (1967) Untersuchungen zur struktur der aminopeptidase aus partikeln von Schweinenieren. *Biochem. Z.* **346,** 467–473.
28. Malfroy, B., Kado-Fong, H., Gros, C., Giros, B., Schwartz, J. C., and Hellmiss, R. (1989) Molecular cloning and amino acid sequence of rat kidney aminopeptidase M: a member of a super family of zinc-metallohydrolases. *Biochem. Biophys. Res. Commun.* **161(1),** 236–241.
29. Hwang, S. Y., Kingsbury, W. D., Hall, N. M., Jakas, D. R., Dunn, G. L., and Gilvarg, C. (1986) Determination of leucine aminopeptidase using phenylalanyl-3-thia-phenylalanine as substrate. *Anal. Biochem.* **154,** 552–558.
30. Manao, G., Nassi, P., Cappugi, G., Camici, G., and Ramponi, G. (1972) Swine kidney prolidase: assay, isolation procedure, and molecular properties. *Physiol. Chem. Phys.* **4,** 75–87.
31. Myara, I., Charpentier, C., and Lemonnier, A. (1984) Minireview: Prolidase and prolidase deficiency. *Life Sci.* **34,** 1985–1998.
32. Sjöström, H. and Norén, O. (1974) Structural properties of pig intestinal proline dipeptidase. *Biochim. Biophys. Acta* **359,** 177–185.
33. Myara, I., Charpentier, C., and Lemonnier, A. (1982) Optimal conditions for prolidase assay by proline colourimetric determination: application to iminodipeptiduria. *Clin. Chim. Acta* **125,** 193–205.
34. Chinard, F. P. (1952) Photometric estimation of proline and ornithine. *J. Biol. Chem.* **199,** 91–95.

35. Richter, A. M., Lancaster, G. L., Choy, F. Y. M., and Hechtman, P. (1989) Purification and characterization of activated human erythrocyte prolidase. *Biochem. Cell Biol.* **67,** 34–41.
36. Mikasa, H. (1984) Measurement of prolidase activity in erythrocytes using isotachophoresis. *J. Chromat.* **310,** 401–406.
37. Yoshimoto, T., Matsubara, F., Kawano, E., and Tsuru, D. (1983) Prolidase from bovine intestine: purification and characterization. *J. Biochem.* **94,** 1889–1896.
38. Hill, R. L. and Schmidt, W. R. (1962) The complete enzymic hydrolysis of proteins. *J. Biol. Chem.* **237,** 389–396.
39. Jones, B. N. (1986) Microsequence analysis by enzymatic methods, in *Methods of Protein Microcharacterization* (Shively, J. E., ed.), Humana, Clifton, NJ, p. 347.
40. Jones, B. N. (1986) Amino acld analysis by *o*-phthaldialdehyde precolumn derivitization and reverse-phase HPLC, in *Methods of Protein Microcharacterization* (Shively, J. E., ed.), Humana, Clifton, NJ, pp. 127,145.
41. Garner, M. H., Garner, W. H., and Gurd, F. R. N. (1974) Recognition of primary sequence variations among sperm whale myoglobin components with successive proteolysis procedures. *J. Biol. Chem.* **249,** 1513–1518.
42. Rothgeb, T. M., Jones, B. N., Hayes, D. F., and Gurd, R. S. (1977) Methylation of glucagon, characterization of the sulfonium derivative, and regeneration of the native covalent structure. *Biochemistry* **16,** 5813–5818.

CHAPTER 19

Alkaline Phosphatase (EC 3.1.3.1)

Martin J. Maunders

1. Introduction

Alkaline phosphatases (or alkaline phosphomonoesterases) catalyze the hydrolysis of phosphate monoesters of a variety of alcohol moieties, being most active at an alkaline pH. The enzymes have been isolated from a variety of sources, including bacteria, fungi, invertebrates, fish, and mammals (being located in many organs, including bone marrow, kidney, placenta, and intestinal mucosa), but have not been isolated from higher plants. The most commonly used alkaline phosphatases (AP) are those from calf intestinal mucosa (called CIAP, CIP, or CAP) and from the bacterium *Escherichia coli* (BAP).

The in vivo function of the enzyme is unclear, since its phosphatase activity is nonspecific. In bone tissue, it may be involved in ossification by formation of calcium phosphate *(1)*, whereas in other mammalian tissues, a role in phosphate transport has been suggested.

The bacterial enzyme is expressed constitutively in most species, an exception being *E. coli*, where synthesis is repressed by orthophosphate *(2)*. On phosphate starvation, BAP is secreted into the periplasmic lumen *(3)*. Again a role has been suggested in a phosphate transport system *(4)*.

The enzyme has two common uses in molecular biology. First, it is used as a reporter in detection systems for particular protein or nucleic acid molecules, and second, to dephosphorylate the termini of nucleic acids enabling subsequent in vitro modification.

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

Detection systems involve the use of an alkaline phosphatase molecule conjugated to a second protein that specifically recognizes the target molecule. This protein may be a ligand, such as streptavidin, which would bind to a biotin-labeled target, or an antibody that offers more versatility. The presence of the target–detector–enzyme complex is revealed by the action of alkaline phosphatase on a chromogenic substrate (such as 5-bromo-4-chloro-3-indolylphosphate/Nitroblue tetrazolium chloride, or BCIP/NBT), or a substrate that will produce luminescence when dephosphorylated (e.g., 3-[2'-spiroadamantane]-4-methoxy-4-[3-phosphoryloxy]-phenyl-1,2-dioxetane, or AMPPD).

Such systems enable the detection of specific proteins by ELISA or western blotting, or of DNA sequences either directly by detection of haptens that have been linked to DNA probes (e.g., biotin, digoxygenin) or indirectly by detection of the expression products of cloned genes. Specific DNA molecules can also be detected by use of oligonucleotides that have been conjugated directly to alkaline phosphatase *(5)*. The numerous variations and intricacies of these AP-linked systems are beyond the scope of this chapter, which will now concentrate on the direct action of alkaline phosphatases on phosphorylated nucleic acids.

The major application of alkaline phosphatase in molecular biology is in dephosphorylating 5' termini of DNA or RNA to prevent self-ligation. This is normally performed on vector DNA to reduce the number of nonrecombinant molecules produced during cloning, but may also be applied to the DNA to be cloned to prevent joining of small noncontiguous fragments that would then give spurious products *(6)*.

Dephosphorylation also enables subsequent tagging with radiolabeled phosphate using the enzyme T4 polynucleotide kinase (*see* Chapter 20). This approach can be used for DNA *(7)* or RNA *(8)* sequencing, and for fragment mapping *(9)*.

2. The Enzyme

2.1. Bacterial Alkaline Phosphatase (BAP)

BAP is a dimer of identical or very similar subunits *(10)*. The mol mass has been determined as 67,000–110,000 Da, varying with pH and ionic strength *(11)* owing in some measure to tetramer formation and also transition to a random coil at low pH. The enzyme is often regarded as a compact sphere of mol mass 80,000 Da *(12)*. The sedimentation

coefficient of BAP, $S^{\circ}_{20,w}$, is 6.0 at pH 8.0 *(11)* and its isoelectric point pH 4.5 *(12)*.

Isozymes have been detected *(13)*—normally three, but more have been reported. Active dimers can be formed from monomers derived from different sources *(13)*.

The dimeric form of the enzyme contains two atoms of zinc, which are required for activity. The dimer dissociates to monomers at pH <3.0 with release of zinc ions, and chelating agents can also remove the zinc atoms, but without the formation of monomers in this case. Some workers report four atoms of zinc per dimer, only two being required for activity *(14)*, and inactivation of the enzyme by metal chelators is biphasic. There is one active site per dimer at low substrate concentrations (<0.1 m*M*) *(15)* and two at higher concentrations (>1 m*M*) *(16)*.

2.2. Calf Intestinal Alkaline Phosphatase (CIAP)

CIAP is a glycoprotein with a mol mass of 100,000–140,000 Da *(17,18)*, comprising two identical or similar subunits and containing four atoms of zinc per dimer *(18,19)*. Isozymes have been found of placental alkaline phosphatase, but not CIAP *(1)*. The enzyme has an isoelectric point of 5.7 independent of temperature (in the range 15–25°C) and ionic strength (0.02–0.5) *(20)*.

3. Enzymic Reaction

3.1. Substrate

Substrates for alkaline phosphatase are varied ranging from the phosphate esters of primary and secondary alcohols, sugar alcohols, and phenols to nucleotides and nucleic acids. This nonspecificity is reflected in the similar rates of hydrolysis of a wide range of substrates *(21)*. *Trans*-phosphorylations also occur, often to alcohol moieties in the buffer other than water, such as Tris or ethanolamine.

Of interest to the molecular biologist is the action on the 5'-phosphate group of single- or double-stranded DNA or RNA that yields a 5'-hydroxyl group *(22,23)*. The enzyme exhibits further activity on RNA and can hydrolyze 5'-di- and triphosphate groups *(24)*, and 3'-phosphate groups *(22)*. It can also cleave 2-', 3'-, and 5'-phosphates of mononucleosides *(21)*. Only phosphate monoesters are susceptible and diesters are not reactive, although some alkaline phosphatase preparations will also act as pyrophosphatases. This activity can be pre-

vented by the inclusion of Mg^{2+} in the reaction buffer, which stimulates the monoesterase activity and almost completely inhibits pyrophosphatase activity *(23)*.

3.2. Temperature

The enzymic reaction occurs over a wide range of temperature, and is usually employed within a range dependent on the substrate. The rate of reaction at 37°C is twice that at 25°C *(25)*.

3.3. pH

As its name implies, the enzyme is active and stable in mildly alkaline solution (pH 7.5–9.5)*(26)*. It is rapidly inactivated at acid pH (e.g., pH 5), because of the loss of Zn^{2+}. Lost activity cannot be restored by addition of zinc ions, but this inactivation is prevented by presence of inorganic phosphate *(27)*. The pH optimum of CIAP increases with substrate concentration and decreases with increased ionic strength *(1)*.

3.4. Cations

Zinc ions are essential for the activity of alkaline phosphatase. Co^{2+}, Mg^{2+}, Mn^{2+}, or Hg^{2+} have been reported to substitute for Zn^{2+} in some cases *(1,21)*. An unusual bacterial alkaline phosphatase (not *E. coli* BAP) has been described that is inhibited by Zn^{2+} ions (0.1 m*M*) and instead requires 5 m*M* $CaCl_2$ for activity *(28)*.

In the case of CIAP, Ca^{2+}, Ni^{2+}, and Cd^{2+} have minimal effect, and Be^{2+} is inhibitory. Zinc itself has been reported to be inhibitory at high concentrations *(1)*, the kinetics of which are complex, and the inhibition is not observed in the presence of glycine *(29)*. The inhibition can also be overcome by the presence of magnesium (30). It has been suggested that magnesium can stimulate the action of additional active sites on the enzyme molecule *(31)*.

3.5. Inhibitors

Inorganic orthophosphate, thiophosphate, and arsenate are strong competitive inhibitors of alkaline phosphatase with low K_i values *(32)* ranging from 0.6–20 μ*M* depending on the buffer employed. Other inhibitors of CIAP include borate, carbonate, pyrophosphate, iodosobenzoate, and iodoacetamide anions, and urea *(1,33)*.

Metal-binding agents that complex with the zinc atoms can also inhibit, including EDTA, cyanide, α,α'-dipyridyl, and 1,10-phenanthroline.

Incubation with Zn^{2+} can reverse some of the inactivation by these agents, but Mg^{2+} and Co^{2+} are much less effective *(29)*.

Many amino acids are weak inhibitors of CIAP, including L-phenylalanine and L-tryptophan *(1,34)*, whereas D-phenylalanine has no effect *(31)*. Inhibition by glycine, cysteine, and histidine is probably owing to chelation of the zinc ions *(35)*. Diisopropylfluorophosphate, an inhibitor of other serine hydrolases, has only a slight effect on alkaline phosphatase (at 1–10 m*M*) *(36)*. Finally, the enzyme is also inhibited by high substrate concentrations (millimolar amounts), the precise mechanism of which is unclear *(1)*.

3.6. Sulfhydryl Reagents

The presence of sulfhydryl reagents is not required for reaction. Alkaline phosphatase is reversibly dissociated by thiol reagents in the presence of urea *(13)*.

3.7. Enzyme Assay and Unit Definition

One unit of alkaline phosphatase is defined as that amount of enzyme that will hydrolyze 1 µmol of 4-nitrophenyl phosphate/min *(37)*. However, other conditions can vary with suppliers.

The assay is normally performed in 1*M* diethanolamine buffer, including 10 m*M* 4-nitrophenyl phosphate and 0.25 m*M* $MgCl_2$. However, a variety of conditions have been employed resulting in assays at pH 8.0 *(12)*, pH 9.6 *(38)*, pH 9.8 *(26,39)*, or at pH 10.5 *(25,40)*. The assay temperature has also been defined as 37°C *(25,26,39,40)* or as 25°C *(12,38)*. The variation in unit definitions is illustrated by the fact that 5 U measured in Diethanolamine buffer at 37°C are equivalent to 1 U in glycine/NaOH at 25°C *(38)*, and the presence of 1*M* diethanolamine in a buffer can double the observed reaction rate *(40)*. This plethora of unit definitions illustrates the wisdom of titrating the quantity of alkaline phosphatase required for any particular operation. A functional unit of activity is sometimes defined as being that quantity of enzyme that will dephosphorylate 1 µg (or 1 pmol) of a particular DNA species in 1 h.

4. Experimental Procedures

4.1. Storage and Stability

The purity of alkaline phosphatase preparations may vary widely, and this has effects on the stability of the enzyme, as well as contaminating catalytic activities. Commercial preparations are usually assayed to be

free of such activities as DNA endo- and exonucleases, RNase, Protease, and adenosine deaminase.

BAP is commonly available at 30–40 U/mg protein (with the proviso mentioned above), and CIAP at up to 900 U/mg protein. The enzyme is supplied at concentrations of 1–50 U/μL, often in a buffer containing 3.2*M* ammonium sulfate (65% saturated), 1 m*M* $MgCl_2$, and 0.1 m*M* $ZnCl_2$. Desalting to remove the ammonium sulfate if present is necessary prior to reaction.

It is possible to dilute the enzyme for a short time in reaction buffer (*see* Section 4.2.), but for longer term storage, the following specific buffer is recommended:

- 10 m*M* Tris-HCl (pH 8.0–8.3)
- 1–5 m*M* $MgCl_2$
- 0.1–0.2 m*M* $ZnCl_2$
- 50% Glycerol

Triethanolamine (30 m*M*, pH 7.6) may be used as an alternative to Tris-HCl, and optional additions include 50 m*M* KCl and 3 m*M* NaCl.

The diluted enzyme is best stored at 4°C rather than at –20°C. Under these conditions, it is stable for 6 mo. The enzyme is in fact stable for several days at room temperature in neutral or mildly alkaline solution, but is inactivated by acid.

4.2. Reaction Conditions

The following reaction conditions are largely optimized for CIAP, although most are applicable to BAP also. The major difference is that reactions with CIAP are best performed at 37°C, whereas BAP is used at 60–65°C. CIAP has largely superseded BAP as the enzyme of choice, since it can be easily inactivated after reaction because of its greater heat lability.

Alkaline phosphatase is active in a wide range of buffers, and it is often possible to add CIAP directly to restriction endonuclease buffers after restriction of the substrate DNA, since the enzyme appears to work equally well as in its own specific buffer. One can add $ZnCl_2$ to 1 m*M* and then proceed *(6)* or even include the CIAP in the digestion mixture to act concurrently with the restriction endonuclease *(39)*.

Where a prepurified DNA substrate is available, the phosphatase reaction can be performed in the following buffer: 10–50 m*M* Tris-HCl,

pH 8.0–9.0 (or 100 m*M* Glycine/NaOH, pH 10.5) and 0.1–1.0 m*M* $ZnCl_2$. In many cases, however, the presence of $ZnCl_2$ in the enzyme storage buffer renders it unnecessary to add further zinc ions to the reaction mixture. Further optional additions include 1 m*M* $MgCl_2$, 1 m*M* Spermidine, and 0.1 m*M* EDTA *(26)*.

The quantity of substrate and enzyme used ideally needs to be titrated for each application *(25)*, but is generally within the range of 1 U/1–100 pmol termini, with the larger enzyme:substrate ratios being employed when the substrate is in the form of flush or recessed termini *(26,38)*. The substrate is usually present at a concentration of 1–50 pmol termini in a reaction vol of 20–200 µL.

Optimum reaction temperatures and times also vary with substrate:

1. DNA carrying protruding phosphorylated termini can be dephosphorylated by reaction at 37°C for 30–60 min, with the optional addition of a further aliquot of enzyme after 30 min *(6,38,41)*.
2. DNA with flush ends will react under similar conditions *(25)*, although reaction at higher temperatures (e.g., 50°C) may be more efficient *(26)*. Two-step incubations of 15 min at 37°C followed by addition of more enzyme and further incubation at 55°C for 45 min, or 15 min each at 37°C and 56°C followed by addition of more enzyme and a repetition of the cycle give effective results *(6)*.
3. DNA with recessed termini are best treated as for flush-ended DNA *(26)*, or by using BAP, higher temperatures can be employed, such as incubations at 60°C for 60 min *(25)*.
4. Dephosphorylation of RNA is also performed at higher temperatures, such as 55°C for 30–60 min *(41)*, possibly with an initial 15 min at 37°C *(6)*.

After reaction, it is necessary to remove or inactivate the alkaline phosphatase, since this will interfere with the efficiency of subsequent ligation and transformation procedures. The reaction can be terminated by adding EDTA to 10 m*M*, and then extracting with phenol. The EDTA step may be improved by heating to 65°C for 60 min or 75°C for 10 min prior to phenol extraction. Other methods of terminating the reaction involve the digestion of CIAP by proteinase K (100 µg/mL in 5 m*M* EDTA, pH 8.0, 0.5% SDS) at 56°C for 30 min, or the chelation of the zinc ions by heating to 65°C for 45–60 min in 50 m*M* EGTA, pH 8.0, or 10 m*M* nitrilo-triacetic acid *(18)*. In all cases, a final extraction with phenol or phenol/chloroform is recommended.

The above procedures refer to CIAP, since it is very difficult to remove BAP after reaction. A newly available bacterial alkaline phosphatase isolated from an antarctic organism *(28)* can be heat killed. After 30 min at 65°C, its activity is reduced to 0.01%.

4.3. Reaction Protocol

An example of a specimen reaction protocol is given in the following:

1. Following restriction enzyme digestion, phenol/chloroform extraction, and ethanol precipitation of the required DNA, dissolve the linearized molecule in 10 m*M* Tris-HCl, pH 8.3.
2. Take an aliquot containing 1 µg DNA.
3. Add 10 µL of the following 10X reaction buffer:
   - 100 m*M* Tris-HCl, pH 8.3
   - 10 m*M* $MgCl_2$
   - 10 m*M* $ZnCl_2$
4. Add water to 100 µL.
5. Add 0.01 U of CIAP (diluted from stock in the specified buffer).
6. Incubate at 37°C for 30 min.
7. Add 5 µL of 10% SDS and 5 µL of 200 m*M* EDTA, pH 8.0. Mix well. Heat to 65°C for 60 min.
8. Extract with phenol/chloroform.
9. Add 10 µL of 3*M* sodium acetate, pH 7.0, and 300 µL ethanol. Precipitate the DNA at –20°C, and resuspend in the requisite buffer for subsequent manipulations (usually ligation).

5. Summary

Alkaline phosphatase offers the molecular biologist another tool for the manipulation of nucleic acids. Pretreatment with the enzyme allows either the specific labeling of particular molecular species or the correct synthesis of recombinant molecules. The wide specificity of AP enables most substrates to be dephosphorylated, and by good experimental design, this can be used to direct subsequent nonspecific reactions, such as DNA ligation, to yield only the desired products. The nature of the enzyme preparation and the variations in substrate require particular reactions to be optimized, and the subsequent removal of such a potent agent requires attention, but the simplicity and effectiveness of the dephosphorylation reaction render it a very widely used technique.

References

1. Fernley, H. N. (1971) Mammalian alkaline phosphatases, in *The Enzymes,* vol. 4, 3rd ed. (Boyer, P. D., ed.), Academic, New York, pp. 417–447.
2. Torrianni, A. (1960) Influence of inorganic phosphate on the formation of phosphatases by *Escherichia coli. Biochim. Biophys. Acta* **38,** 460–469.
3. Malamy, M. and Horecker, B. L. (1964) Release of alkaline phosphatase from cells of *Escherichia coli* upon lysozyme spheroplast formation. *Biochemistry* **3,** 1889–1893.
4. Reid, T. W. and Wilson, I. B. (1971) *E. coli* alkaline phosphatase, in *The Enzymes,* vol. 4, 3rd ed. (Boyer, P. D., ed.) Academic, New York, pp. 373–415.
5. Cambridge Research Biochemical Ltd. Catalog (1991) E-link and E-link plus, oligonucleotide alkaline phosphatase cojugation kits. Cambridge Research Biochemicals Ltd., Northwich, Cheshire, UK, p. 132.
6. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Dephosphorylation of linearised plasmid DNA, in *Molecular Cloning,* 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 1.60,1.61.
7. Maxam, A. M. and Gilbert W. (1980) Sequencing of end-labeled DNA with base specific chemical cleavages. *Methods Enzymol.* **65,** 499–560.
8. Randerath, K., Gupta, R. C., and Randerath, E. (1980) ^{3}H and ^{32}P derivative methods for base composition and sequence analysis of RNA. *Methods Enzymol.* **65,** 638–680.
9. Tu, C.-P. D. and Wu, R. (1980) Sequence analysis of short DNA fragments. *Methods Enzymol.* **65,** 620–638.
10. Rothman, F. and Byrne, R. (1963) Fingerprint analysis of alkaline phosphatase of *Escherichia coli* K12. *J. Mol. Biol.* **6,** 330–340.
11. Applebury, M. L. and Coleman, J. E. (1969) *Escherichia coli* alkaline phosphatase. Metal binding, protein conformation and quaternary structure. *J. Biol. Chem.* **244,** 308–318.
12. Garen, A. and Levinthal, C. (1960) A fine structure genetic and chemical study of the enzyme alkaline phosphatase of *E. coli.* 1. Purification and characterisation of alkaline phosphatase. *Biochim. Biophys. Acta* **38,** 470–483.
13. Levinthal, C., Signer, E., and Fetherolf, K. (1962) Reactivation and hybridisation of reduced alkaline phosphatase. *Proc. Natl. Acad. Sci. USA* **48,** 1230–1237.
14. Lazdunski, C., Petitclerc, C., and Lazdunski, M. (1969) Structure-function relationships for some metalloalkaline phosphatases of *E. coli. Eur. J. Biochem.* **8,** 510–517.
15. Heppel, L. A., Harkness, D., and Hilmoe, R. (1962) A study of the substrate specificity and other properties of the alkaline phosphatase of *Escherichia coli. J. Biol. Chem.* **237,** 841–846.
16. Fife, W. K. (1967) Phosphorylation of alkaline phosphatase *(E. coli)* with *o*- and *p*-nitrophenyl phosphate at pH <6. *Biochem. Biophys. Res. Commun.* **28,** 309–317.

17. Engstrom, L. (1961) Studies on calf intestinal alkaline phosphatase. 1. Chromatographic purification, microheterogeneity and some other properties of the purified enzyme. *Biochim. Biophys. Acta* **52,** 36–48.
18. Efstratiadis, A., Vournakis, J. N., Donis-Keller, H., and Chaconas, G. (1977) End-labeling of enzymically decapped mRNA. *Nucleic Acids Res.* **4,** 4165–4174.
19. Simsek, M., Zeigenmeyer, J., Heckman, J., and RajBhandary, U. L. (1973) Absence of the sequence G-T-ψ-C-G(A) in several eukaryotic cytoplasmic initiator transfer RNAs. *Proc. Natl. Acad. Sci. USA* **70,** 1041–1045.
20. Lazdunski, M., Brouillard, J., and Ouellet, L. (1965) Etude des effets électrostatiques sur le méchanisme d'action catalytique de la phosphatase alcaline intestinale de veau. *Can. J. Chem.* **43,** 2222–2235.
21. Reid, T. W. and Wilson, I. B. (1971) *E. coli* alkaline phosphatase, in *The Enzymes,* vol. 4, 3rd ed. (Boyer, P. D., ed.), Academic, New York, pp. 373–415.
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Treatment of eukaryotic DNA with alkaline phosphatase, in *Molecular Cloning*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 3.38–3.43.
23. Fernley, H. N. and Walker, T. G. (1967) Studies on alkaline phosphatase. Inhibition by phosphate derivatives and the substrate specificity. *Biochem. J.* **104,** 1011–1018.
24. Johnston, D. E. and McClure, W. R. (1976) Abortive initiation of in vitro RNA synthesis on bacteriophage lambda DNA, in *RNA Polymerase* (Losick, L. and Chamberlin, M. J., eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 413–428.
25. Biotechnology Catalog (1990/1991) Alkaline phosphatase. Pharmacia LKB Biotechnology, Uppsala, Sweden, pp. 309–310.
26. Keesey, J. (1987) Alkaline phosphatase, in *Biochemica Information*, Boehringer Mannheim Biochemicals, Indianapolis, IN, pp. 8–11.
27. Fosset, M., Chappelet-Tordo, D., and Lazdunski, M. (1974) Intestinal alkaline phosphatase. Physical properties and quaternary structure. *Biochemistry* **13,** 1783–1788.
28. Epicentre Technologies Product Information (1990) HK phosphatase. EpicentreTechnologies, Madison, WI.
29. Harkness, D. R. (1968) Studies on human placental alkaline phosphatase. 1. Purification and crystallization. *Archiv. Biochem. Biophys.* **126,** 503–512.
30. Morton, R. K. (1955) Phosphomonoesterase of milk. *Methods Enzymol.* **2,** 533–539.
31. Fernley, H. N. and Walker, P. G. (1970) Inhibition of alkaline phosphatase by L-phenylalanine. *Biochem. J.* **116,** 543–544.
32. Fernley, H. N. and Walker, P. G. (1969) Studies on alkaline phosphatase. Transient state and steady state kinetics of *Escherichia coli* alkaline phosphatase. *Biochem. J.* **111,** 187–194.
33. Zittle, C. A. and Della Monica, E. S. (1950) Effects of borate and other ions on the alkaline phosphatase of bovine milk and intastinal mucosa. *Archiv. Biochem. Biophys.* **26,** 112–122.

34. Barman, T. E. (1969) Alkaline phosphatase, in *The Enzyme Handbook*, vol. 2, Springer-Verlag, Berlin, pp. 521,522.
35. Agus, S. G., Cox, R. P., and Griffin, M. J. (1966) Inhibition of alkaline phosphatase by cysteine and its analogues. *Biochim. Biophys. Acta* **118,** 363–370.
36. Dabich, D. and Neuhaus, O. W. (1966) Purification and properties of bovine synovial fluid alkaline phosphatase. *J. Biol. Chem.* **241,** 415–420.
37. Bessey, O. A., Lowry, O. H., and Brock, M. J. (1946) A method for the rapid determination of alkaline phosphatase with five cubic millilitres of serum. *J. Biol. Chem.* **164,** 321–329.
38. Mössner, E., Boll, M., and Pfleiderer, G. (1980) Purification of human and bovine alkaline phosphatases by affinity chromatography. *Z. Physiol. Chem.* **361,** 543–549.
39. Promega Protocols & Applications Guide (1989) Cloning of DNA inserts. Genetic cloning and mapping. Promega Corp., Madison, WI, pp. 37, 158.
40. NEN Research Products Catalog E11, 2nd ed. (1986) Enzymes for nucleic acid synthesis and modification. Du Pont (UK) Ltd., Stevenage, Herts, UK, p. 145.
41. Chaconas, G. and Van de Sande, J. H. (1980) 5'-^{32}P labeling of RNA and DNA restriction fragments. *Methods Enzymol.* **65,** 75–85.

CHAPTER 20

Polynucleotide Kinase (EC 2.7.1.78)

Martin J. Maunders

1. Introduction

The enzyme polynucleotide kinase (PNK, or ATP:5'-dephosphopolynucleotide 5'-phosphatase) catalyzes the transfer of a γ-phosphate group from a 5'-nucleoside triphosphate moiety to a free 5'-hydroxyl of a polynucleotide such as DNA or RNA, to form a 5'-phosphorylated DNA or RNA molecule and a nucleoside diphosphate.

This activity has been identified in *E. coli* infected with bacteriophages T2, T4, or T6 *(1,2)*. No enzyme can be detected in the uninfected bacterium, but a similar kinase activity has been observed in mammalian tissues, including rat liver *(3,4)*, calf thymus *(5)*, and various cell lines including Chinese hamster lung cells *(6)* and HeLa cells *(7)*. The bacteriophage enzymes act on both DNA or RNA, whereas the mammalian enzymes are generally active only on DNA. Exceptions to this are the calf thymus enzyme, which has a slight action on RNA, and the HeLa cell enzyme, which is solely RNA-specific.

The in vivo role of the enzyme is possibly in maintaining DNA or RNA in the 5'-phosphorylated, 3'-hydroxylated state, which is the substrate for many reactions such as ligation and packaging.

Polynucleotide kinase has many uses in molecular biology, however they can be grouped into two classes. The kinasing activity can be used purely to modify DNA, RNA, or synthetic oligonucleotides for subsequent manipulations, or it can enable the radiolabeling of these molecules for subsequent detection in probing, mapping, or sequencing experiments.

From: *Methods in Molecular Biology, Vol. 16: Enzymes of Molecular Biology*
Edited by: M. M. Burrell

The bacteriophage polynucleotide kinases have similar properties, that from T4-infected *E. coli* having been the most studied. In this chapter, I concentrate on T4 polynucleotide kinase as an example of these enzymes.

2. The Enzyme

T4 polynucleotide kinase is encoded by the structural gene *pse*T. This gene also codes for a T4 3'-phosphatase *(8–10)*, whose activity has been identified as residing on the same enzyme molecule. The mutant phage *pse*T1 lacks the phosphatase activity *(11–13)*, but has unaffected kinase activity, and is therefore often used as a source for preparing the enzyme. Like many commercially available enzymes, polynucleotide kinase has also been prepared from a recombinant overproducing strain.

T4 polynucleotide kinase is a tetramer composed of identical subunits, each consisting of 45–55% α-helix and possessing an N-terminal phenylalanine residue *(14)*. The mol mass of the native enzyme has been determined as 140,000 Da by gel filtration *(15)*, and 147,300 Da by centrifugation *(14)*. The size of the denatured and reduced monomers has been measured as 33,000 Da by polyacrylamide gel electrophoresis *(14,15)*, and 33,200 Da by centrifugation *(14)*. The sedimentation coefficients, $S°_{20,w}$, are 2.95S and 6.55S for the monomer and tetramer respectively.

By comparison, the rat liver enzyme has been determined to have a mol mass of 80,000 Da by gel filtration *(16,17)*, and a sedimentation coefficient of 4.4S *(16)*.

3. Enzymic Reaction

3.1. Reaction Catalyzed

T4 polynucleotide kinase catalyzes the transfer of the γ-phosphate group from a 5'-nucleoside triphosphate to the 5'-OH of an acceptor molecule. This may be a nucleoside-3'-phosphate, an oligonucleotide, or a polynucleotide. The reaction is reversible, and the enzyme will catalyze polynucleotide dephosphorylation in the presence of a nucleotide diphosphate such as ADP *(18)*. Excess ADP will cause the reverse reaction to be favored.

The reverse reaction can be utilized for the exchange of labeled phosphate groups between the two substrates. This exchange reaction allows 5'-labeling of polynucleotides without prior removal of the existing 5'-phosphate group.

The associated 3'-phosphatase activity causes the hydrolysis of the 3'-phosphate group of a variety of substrate molecules, including deoxynucleoside 3'-monophosphates, deoxynucleoside 3',5'-diphosphates, and 3'-phosphorylated polynucleotides, to form a 3'-hydroxyl group and release inorganic phosphate.

3.2. Substrate

3.2.1. Acceptor

The 5'-hydroxylated nucleoside moiety in the reaction may comprise single- or double-stranded DNA, RNA, a synthetic oligonucleotide, a nucleoside-3'-monophosphate, or a deoxynucleoside-3'-monophosphate. The enzyme can act on any molecule terminating in a naturally occurring nucleoside. By contrast, rat liver polynucleotide kinase cannot act on RNA or oligonucleotides <10 bases in length.

When the molecule is double-stranded DNA, a protruding 5' terminus is a better substrate than a blunt or recessed end. However, by increasing the concentration of the phosphate donor (usually ATP), all 5' termini can be completely phosphorylated. Increased enzyme concentrations also cause more efficient kinasing of recessed termini.

Nicks in duplex DNA will act as substrates for reaction, but the rate of kinasing is 10- to 30-fold slower than for single-stranded DNA, or protruding 5' termini *(19,20)*, and phosphorylation is incomplete, with only 70% being achievable even after long reaction times *(8)*. Raising the ATP concentration will not promote complete phosphorylation in this case.

The size of the acceptor molecule has little effect on the rate of reaction, within the range 150–50,000 nucleotides. The K_m value for large DNA fragments released by nuclease treatment is 7.6 μM *(21)*. K_m for nucleoside-3'-phosphates and oligonucleotides is 22.2–143.0 μM depending on the 5' base and the length of the oligonucleotide *(22)*.

3.2.2. Donor

The phosphate group donor for the kinase reaction may be any nucleoside triphosphate. Although ATP is used routinely for experimental or assay purposes, CTP, UTP, GTP, dATP, and TTP perform equally well *(21)*. The ATP concentration should be at least 1 μM for the reaction to proceed, and excess ATP is required for optimal kinasing *(23)*.

The following K_m values have been determined for T2 polynucleotide kinase: 14 μ*M* for ATP, 15 μ*M* for UTP, 33 μ*M* for GTP, and 25 μ*M* for CTP *(21)*. With the T4 enzyme, the K_m value for ATP is 13–140 μ*M* depending on the DNA acceptor *(3,24)*.

Under the conditions for phosphate exchange, the K_{mATP} (forward reaction) is 4 μ*M*, and the K_{mADP} (reverse reaction) is 200 μ*M*. The optimal concentrations for exchange are 10 μ*M* ATP and 300 μ*M* ADP *(20)*, but even so reaction is usually incomplete.

For the dephosphorylation of single-stranded oligonucleotides, K_{mADP} is 0.22 μ*M* *(18)*, but again only partial dephosphorylation can be achieved.

3.3. Temperature

As with most *E. coli* bacteriophage-derived enzymes, the optimal reaction temperature is 37°C.

T4 polynucleotide kinase will also perform at lower temperatures, which is a useful attribute. The rate of the kinase reaction at 0°C is reduced to 7% of that at 37°C. However, the exchange reaction rate is reduced much further, to 1.2% of that at 37°C *(25)*, so under these conditions, kinasing is greatly favored.

3.4. pH

The optimum pH range for T4 polynucleotide kinase is 7.4–8.0, with maximum activity in Tris buffers being observed at pH 7.6 *(2)*. The reverse reaction has a pH optimum of 6.2, in imidazole buffer *(20)*.

The 3'-phosphatase activity is greatest at pH 5.9 *(9)*.

3.5. Cations

Polynucleotide kinase has a requirement for magnesium ions for both the forward and exchange reactions *(18)*, with no activity being detectable in their absence. At the optimal pH of 7.6, the optimum magnesium concentration is 10 m*M* *(2)*.

Manganese can partially replace magnesium in some cases. Note that 3.3 m*M* Mn^{2+} will permit reaction at 50% of the maximum rate obtainable with 10 m*M* Mg^{2+}.

The 3'-phosphatase activity similarly requires Mg^{2+}, with the reaction rate falling to 2% in its absence. In this case Co^{2+} can replace Mg^{2+} to some extent *(21)*.

3.6. Activators

The activity of polynucleotide kinase is stimulated by sodium chloride, potassium chloride, and polyamines *(22)*. Spermine promotes tetramer formation *(14)*, and potassium chloride maintains the enzyme in this oligomeric form. A total of 1.7 m*M* of spermine can increase reaction rate by 30-fold *(26)*. Spermidine, in addition to enhancing the reaction rate, also has the ability to inhibit nucleases present in some kinase preparations *(27)*.

Polyethylene glycol (PEG 8000) also improves the efficiency of the reaction *(26)*. Stimulation depends on the PEG concentration, which should ideally be titrated within the range 4–10%. The stimulating action of PEG is owing to it causing the DNA to undergo a "psi" transition and collapse into a highly condensed state *(28)*. This state is only achievable for DNA molecules >300 bp in length, so PEG has little effect on shorter molecules.

3.7. Inhibitors

T4 polynucleotide kinase is inhibited by inorganic phosphate and pyrophosphate *(2,3,26)*. If 70 m*M* sodium or potassium phosphate buffers (pH 7.6) are employed, enzyme activity is reduced to 5% of that observed in Tris buffer. The reaction rate in Tris buffer can be halved by the addition of phosphate to 7–20 m*M* or pyrophosphate to 5 m*M* *(21,29)*. In addition, 50 m*M* of potassium phosphate will cause 60% inhibition of the exchange reaction *(18)*.

Inorganic phosphate is, however, relatively more inhibitory to *E. coli* alkaline phosphatase than to T4 polynucleotide kinase. It can therefore be used to enhance phosphorylation when phosphatase is present *(30)*, e.g., when end-labeling molecules after a prior phosphatasing step.

Another strong inhibitor of polynucleotide kinase is the ammonium ion. For example, 75% inhibition can be caused by the presence of 7 m*M* ammonium sulfate *(29,31)*.

Potassium chloride has an activating effect (as described in Section 3.6.), but high concentrations can be inhibitory *(19)*. This inhibition is not observed in substrates with protruding 5'-OH termini. Sodium chloride, although stimulating kinase activity toward single-stranded substrates *(26)*, can also be inhibitory with some duplexes *(21)*.

3.8. Sulfhydryl Reagents

Sulfhydryl reagents are essential for the action of polynucleotide kinase *(18)*, whose activity falls to 2% in their absence. Maximum activity can be obtained in the presence of 5 m*M* dithiothreitol (DTT) *(32)*. However, 80% of this maximum activity can be achieved using 10 m*M* β-mercaptoethanol in place of DTT, and 70% with 10 m*M* glutathione *(32)*.

3.9. Enzyme Assay and Unit Definition

Assaying the activity of T4 polynucleotide kinase entails measurement of the transfer of a radiolabeled phosphate group from γ-^{32}P ATP to an acid insoluble product *(2)*. The phosphate acceptor is normally a duplex DNA molecule, enzymatically treated to create 5'-hydroxyl groups. This may be achieved by partial digestion using micrococcal nuclease, which specifically produces 5'-hydroxyl and 3'-phosphate termini, or using pancreatic DNAse followed by alkaline phosphatase, which removes the 5'-phosphate groups originally created.

The standard assay conditions are as follows: 70 m*M* Tris-HCl, pH 7.6; 10 m*M* $MgCl_2$; 5 m*M* DTT; 66 μ*M* γ-^{32}P-ATP; and 0.26 m*M* 5'-OH salmon sperm DNA. Incubate at 37°C.

Commercial suppliers of the enzyme may use slightly different assays, including variations in the nature of the DNA acceptor, the concentrations of 5' termini and ATP, and the inclusion of other buffer components such as spermidine. One unit of activity is defined as the amount of enzyme that catalyzes the incorporation of 1 nmol of ^{32}P into an acid insoluble form in 30 min at 37°C.

The 3'-phosphatase activity can be determined by incubation with AMP *(9)*. The enzyme is incubated with 16 m*M* 3'-AMP for 60 min at 37°C, and the released inorganic phosphate measured. Preparations from mutant or recombinant sources that are nominally phosphatase-free hydrolyze <0.1% of the AMP *(23)*.

4. Experimental Procedures

4.1. Uses of Polynucleotide Kinase

The major use of T4 polynucleotide kinase in the molecular biology laboratory is for the specific phosphorylation of the 5' termini of DNA and RNA, either by direct kinasing or by the exchange reaction. This may be for the purposes of labeling the molecule, or merely to enable the molecule to be further manipulated.

End-labeling allows the quantification of the termini, the enzyme often being used in conjunction with alkaline phosphatase to assess the number of 5'-phosphate groups present by measuring the available 5'-hydroxyl groups before and after phosphatasing. It can also be used to characterize cleavage points in nucleic acids.

End-labeled oligonucleotides can be used as primers for sequencing *(33)*. Both oligonucleotides and polynucleotides can be used as hybridization probes *(34–36)* for clone or genomic characterization, restriction mapping using partial digestion techniques, DNA or RNA fingerprinting *(37,38)*, DNA footprinting *(39,40)*, nuclease S1 analysis, physical mapping, and sequence analysis *(41)*.

Phosphorylation reactions where the phosphate moiety is unlabeled (or labeled purely for the purposes of monitoring the reaction) are used in the synthesis of substrates for DNA or RNA ligation *(42)*. These may be vector molecules, genomic fragments, or synthetic oligonucleotides such as linkers. Such manipulations allow the assembly of long nucleic acid molecules from short synthetic precursors.

The mutant 3'-phosphatase-free enzyme is especially useful for RNA analysis since the continued presence of the 3'-phosphate group prevents cyclization of concatenation *(23)*.

Finally the wild-type enzyme may be used as a specific 3'-phosphatase under the right conditions *(43)*.

4.2. Storage and Stability

Enzyme preparations are commercially available at concentrations of 1000–12,000 U/mL, and with specific activities in the range 30,000–40,000 U/mg.

The enzyme is normally stable for up to 18 mo at –20°C in a suitable storage buffer. The composition of this buffer can vary, but a general purpose buffer used for the storage of a wide variety of enzymes can be used, with the optional addition of 0.1–1.0 μ*M* ATP. An example of a typical storage buffer would be: 50 m*M* Tris-HCl, pH 7.5; 25 m*M* KCl; 1 m*M* DTT; 0.1 m*M* EDTA; 0.1 μ*M* ATP; and 50% glycerol.

4.3. Reaction Conditions

T4 polynucleotide kinase will act in restriction enzyme buffers allowing simultaneous restriction digestions *(31)*, but specific conditions have also been defined as described in Section 3.9. Reaction condi-

tions can be varied to give the optimal reaction with a particular substrate. For example, when kinasing blunt or 5'-recessed termini, higher concentrations of ATP and enzyme are necessary *(27)*, and additional buffer components such as PEG may be beneficial *(44)*. Several reaction protocols are given in Section 4.4.

A number of points are worthy of note. First, because of the extreme inhibition of T4 polynucleotide kinase by ammonium salts, the preparation of the polynucleotide substrate should not involve precipitation with ammonium acetate prior to the kinasing step. Second, all reaction buffers used should be Tris- or imidazole-based, and not contain inorganic phosphate; otherwise inhibition will be observed again. Finally, the target molecule should be rigorously purified from small mol-wt fragments, since these would contribute a large number of 5' termini, which could compete in the reaction.

After the kinasing reaction, subsequent purification depends on the use to which the phosphorylated polynucleotide will be put. If it is to be used as a hybdridization probe, further purification is not necessary unless problems with background noise have been encountered. Purification by repeated ethanol precipitation may be sufficient. When working with oligonucleotides of <18 bases in length, precipitation with cetylpyridinium bromide *(27,45)* must be employed. If the molecule is to be used in further manipulations, it may need more rigorous purification by, e.g., column chromatography, spin dialysis, or polyacrylamide gel electrophoresis.

4.4. Reaction Protocols

4.4.1. Materials Required

1. Suitably prepared 5'-OH DNA, RNA, or oligonucleotides.
2. Reaction buffers (use one or the other):
   a. For kinasing, 10X Tris kinase buffer: 500 m*M* Tris-HCl, pH 7.6, 100 m*M* $MgCl_2$, 50 m*M* DTT, 1 m*M* Spermidine, and 1 m*M* EDTA.
   b. For kinasing or phosphate exchange, 10X imidazole buffer: 500 m*M* imidazole-Cl, pH 6.4; 180 m*M* $MgC1_2$; 50 m*M* DTT; 1 m*M* Spermidine; and 1 m*M* EDTA.
3. γ-^{32}P-ATP (3000 or 5000 Ci/mmol).
4. T4 polynucleotide kinase (10 U/μL).
5. 500 m*M* EDTA, pH 8.0.
6. Sterile distilled water.
7. Phenol/chloroform and chloroform for extraction.

8. 7.5*M* Ammonium acetate.
9. Ethanol.
10. TE buffer: 10 m*M* Tris-HCl, pH 8.0, 1 m*M* EDTA.
11. 24% PEG 8000.
12. 1 m*M* ADP solution (for exchange reaction only).
13. 50 n*M* ATP solution (for exchange reaction only).
14. Sephadex G-50 for column or spun column purification after reaction.
15. For measurement of incorporation:
    a. 10% trichloroacetic acid, 5% trichloroacetic acid, and 70% ethanol;
    b. DE81 filters and 500 m*M* sodium phosphate, pH 7.0; or
    c. Polyethyleneimine cellulose TLC strips and 500 m*M* ammonium bicarbonate.

4.4.2. Kinasing Single-Stranded DNA Fragments or Duplexes with Protruding 5' Termini (27,31)

1. Prepare and purify dephosphorylated DNA fragments.
2. To an aliquot containing 1–50 pmol ends add:
   - 5 µL of 10X Tris kinase buffer
   - 15 µL of γ-^{32}P ATP (50 pmol, 3000 Ci/mmol, 10 Ci/µL, 1 µ*M* final concentration)
   - 1 µL of T4 polynucleotide kinase (10 U)
   - Water to 50 µL.
3. Incubate at 37°C for 30 min.
4. Terminate the reaction by adding 2 µL of 500 m*M* EDTA.
5. Purify by phenol/chloroform extraction.
6. Remove unincorporated ATP by column or spun-column chromatography using Sephadex G-50.
7. Add 1/2 vol 7.5*M* ammonium acetate and 2 vol ethanol. Precipitate the DNA at –70°C for 30 min. Centrifuge, drain pellet, and redissolve in 50 µL TE.
8. Determine incorporation of label by TCA precipitation. (*See* Section 4.4.5. for measurement of incorporation in oligonucleotides.)

4.4.3. Kinasing DNA Duplexes with Blunt or Recessed 5' Ends (27)

1. To an aliquot of DNA containing 1–50 pmol ends in 9 µL or less, add:
   - 4 µL of 10X imidazole buffer
   - Water to 13 µL
   - 10 µL of 24% PEG 8000
   - 15 µL of γ-^{32}P ATP (50 pmol, 3000 Ci/mmol, 10 µCi/µL)
   - 2 µL (20 U) of T4 polynucleotide kinase

2. Incubate at 37°C for 30 min.
3. Terminate the reaction by adding 2 µL of 500 m*M* EDTA, pH 8.0.
4. Extract, purify, and measure incorporation as in Section 4.4.2.

4.4.4. Exchange Reaction (20)

1. To 1–50 pmol DNA in a small volume add:
   - 5 µL of 10X imidazole buffer
   - 5 µL of 1m*M* ADP
   - 1 µL of 50 n*M* ATP
   - 15 µL of γ-^{32}P ATP (50 pmol, 3000 Ci/mmol, 10 µCi/µL)
   - Water to 38 µL
   - 10 µL of 24% PEG 8000
   - 2 µL (20 U) of T4 polynucleotide kinase
2. Incubate at 37°C for 30 min.
3. Terminate the reaction by adding 2 µL of 500 m*M* EDTA, pH 8.0.
4. Extract, purify, and measure incorporation as in Section 4.4.2.

4.4.5. Labeling Oligonucleotides (27)

1. To 10 pmol of oligonucleotide in 1 µL add the following:
   - 2 µL of 10X Tris kinase buffer,
   - 5 µL of γ-^{32}P ATP (10 pmol, 5000 Ci/mmol, 10 µCi/µL), and
   - 11.5 µL of water.
2. Mix and take 0.5-µL zero-time aliquot and add this to 10 µL TE.
3. To the remainder of the reaction mix add 1 µL (10 U) T4 polynucleotide kinase.
4. Incubate at 37°C for 45 min.
5. Take another 0.5-µL aliquot and add to 10 µL of TE as before.
6. Terminate the reaction in the remainder of the mix by heating to 68°C for 10 min.
7. Measure incorporation efficiency as follows *(46,47)*: Spot 0.5 µL of each diluted aliquot onto 15-cm long polyethyleneimine cellulose strips. Perform thin layer chromatography using 500 m*M* ammonium bicarbonate as the developing solution. Allow the solvent front to run 10–13 cm. Saran Wrap and autoradiograph, or slice up the strip and measure the radioactivity along the strip by scintillation counting. Compare time-zero and 45-min samples. In this TLC system, oligonucleotides remain at the origin, inorganic phosphate migrates near the solvent front, and ATP occupies an intermediate position.

   An alternative method is to measure incorporation by adsorption to DE81 filters. Oligonucleotides bind tightly, whereas ATP can be washed off with 0.5*M* sodium phosphate, pH 7.0.

8. If the oligonucleotide is not labeled highly enough, add another 10 U of enzyme and incubate for a further 30 min.
9. Purify as required.

Note: Equal concentrations of ATP and 5' ends gives 50% labeling. To obtain high specific activity, increase the ATP:oligonucleotide ratio to 10:1. Only 10% of the label will be transferred, but virtually every oligonucleotide molecule will be labeled and to a specific activity approaching that of the ATP *(27)*.

5. Summary

Polynucleotide kinase provides a simple means of modifying the termini of nucleic acids. It facilitates the labeling of molecules, or the preparation of those molecules for further manipulation. It can be used in concert with other enzymes such as alkaline phosphatase and DNA ligase to enable a series of manipulations to be carried out, or it can perform useful functions alone. The enzyme is especially important for the modification of short synthetic oligonucleotides, but its ability to act on a variety of substrates makes it a tool with a wide range of applications.

References

1. Spencer, E., Loring, D., Hurwitz, J., and Monroy, G. (1978) Enzymatic conversion of 5'-phosphate terminated RNA to 5'-di- and triphosphate terminated RNA. *Proc. Natl. Acad. Sci. USA* **75,** 4793–4797
2. Richardson, C. C. (1965) Phosphorylation of nucleic acid by an enzyme from T4 bacteriophage infected *Escherichia coli. Proc. Natl. Acad. Sci. USA* **54,** 158–163
3. Novogrodsky, A., Tal, M., Traub, A., and Hurwitz, J. (1966) The enzymatic phosphorylation of ribonucleic acid and dexoxyribonucleic acid. II. Further properties of the 5'-hydroxyl polynucleotide kinase. *J. Biol. Chem.* **241,** 2933–2943.
4. Ichimura, M. and Tsukada, K. (1971) Polynucleotide kinase from rat liver nuclear extracts. *J. Biochem.* **69,** 823–828.
5. Austin, G. E., Sirakoff, D., Roop, B., and Moyer, G. H. (1978) Purification and properties of polynucleotide kinase of calf thymus. *Biochim. Biophys. Acta* **522,** 412–422.
6. Zimmerman, S. B. and Pheiffer, B. H. (1981) Eukaryotic Kinases, in *The Enzymes,* 3rd ed., vol. 14 (Boyer, P. D., ed.), Academic, New York, pp. 315–329.
7. Shuman, S. and Hurwitz, J. (1979) 5'-hydroxyl polyribonucleotide kinase from HeLa cell nuclei. Purification and properties. *J. Biol. Chem.* **254,** 10,396–10,404.

8. Weiss, B., Live, T. R., and Richardson, C. C. (1968) Enzymatic breakage and joining of deoxyribonucleic acid. V. End group labelling and analysis of deoxyribonucleic acid containing single strand breaks. *J. Biol. Chem.* **243,** 4530–4542.
9. Cameron, V. and Uhlenbeck, O. C. (1977) 3'-phosphatase activity in T4 polynucleotide kinase. *Biochemistry* **16,** 5120–5126.
10. Sirotkin, K., Cooley, W., Runnels, J., and Snyder, L. R. (1978) A role in true-late gene expression for the T4 bacteriophage 5'-polynucleotide kinase 3'-phosphatase. *J. Mol. Biol.* **123,** 221–233
11. Cameron, V., Soltis, D., and Uhlenbeck, O. C. (1978) Polynucleotide kinase from a T4 mutant which lacks the 3'-phosphatase activity. *Nucleic Acids Res.* **5,** 825–833.
12. DePew, R. and Cozzarelli, N. (1974) Genetics and physiology of bacteriophage T4 3'-phosphatase. Evidence for involvement of the enzyme in T4 DNA metabolism. *J. Virol.* **13,** 888–897.
13. Soltis, D. A. and Uhlenbeck, O. C. (1982) Independent locations of kinase and 3'-phosphatase activities on T4 polynucleotide kinase. *J. Biol. Chem.* **257,** 11,340–11,345.
14. Lillehaug, J. R. (1977) Physicochemical properties of T4 polynucleotide kinase. *Eur. J. Biochem.* **73,** 499–506.
15. Panet, A., van de Sande, J. H., Loewen, P. C., Khorana, H. G., Raae, A. J., Lillehaug, J. R., and Kleppe, K. (1973) Physical characterisation and simultaneous purification of bacteriophage T4 induced polynucleotide kinase, polynucleotide ligase and deoxyribonucleic acid polymerase. *Biochemistry* **12,** 5045–5050.
16. Teraoka, H., Mizuta, K., Satl, F., Shimoyachi, M., and Tsukada, K. (1975) Polynucleotide kinase from rat liver nuclei. Purification and properties. *Eur. J. Biochem.* **58,** 297–302.
17. Levin, C. J. and Zimmerman, S. B. (1976) A deoxyribonucleic acid kinase from nuclei of rat liver. Purification and properties. *J. Biol. Chem.* **251,** 1767–1774.
18. van de Sande, J. H., Kleppe, K., and Khorana, H. G. (1973) Reversal of bacteriophage T4 induced polynucleotide kinase action. *Biochemistry* **12,** 5050–5055.
19. Lillehaug, J. R., Kleppe, R. K., and Kleppe, K. (1976) Phosphorylation of double-stranded DNAs by T4 polynucleotide kinase. *Biochemistry* **15,** 1858–1865.
20. Berkner, K. L. and Folk, W. R. (1977) Polynucleotide kinase exchange reaction. Quantitative assay for restriction endonuclease generated 5'-phosphoryl termini in DNAs. *J. Biol. Chem.* **252,** 3176–3184.
21. Richardson, C. C. (1981) Bacteriophage T4 polynucleotide kinase, in *The Enzymes,* 3rd ed., vol. 14 (Boyer, P. D., ed.), Academic, New York, pp. 299–314.
22. Lillehaug, J. R. and Kleppe, K. (1975) Kinetics and specificity of T4 polynucleotide kinase. *Biochemistry* **14,** 1221–1225.
23. *Biochemicals for Molecular Biology* (1990) Polynucleotide kinase. Boehringer Mannheim, Mannheim, Germany, p. 59.
24. Chaconas, G., van de Sande, J. H., and Church, R. B. (1975) End group labeling of RNA and double-stranded DNA by phosphate exchange catalyzed by bacteriophage T4 polynucleotide kinase. *Biochem. Biophys. Res. Commun.* **66,** 962–969.

25. Okazaki, R., Hirose, S., Okazaki, T., Ogawa, T., and Kurosawa, Y. (1975) Assay of RNA-linked nascent DNA pieces with polynucleotide kinase. *Biochem. Biophys. Res. Commun.* **62,** 1018–1024.
26. Lillehaug, J. R. and Kleppe, K. (1975) Effect of salts and polyamines on T4 polynucleotide kinase. *Biochemistry* **14,** 1225–1229.
27. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning*, vol. 1–3, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
28. Lerman, L. S. (1971) A transition to a compact form of DNA in polymer solutions. *Proc. Natl. Acad. Sci. USA* **68,** 1886–1890.
29. *New England Biolabs Catalogue* (1990,1991) T4 polynucleotide kinase. New England Biolabs, Inc., Beverly, MA, p. 55.
30. Weiss, B., Jacquemin-Sablon, A., Live, T. R., Fareed, G. C., and Richardson, C. C. (1968) Enzymatic breakage and joining of deoxyribonucleic acid. VI. Further purification and properties of polynucleotide ligase from *Escherichia coli* infected with bacteriophage T4. *J. Biol. Chem.* **243,** 4543–4555.
31. *Promega Protocols and Applications Guide* (2nd ed.) (1991) 5'-end labeling using T4 polynucleotide kinase. Promega Corp. Madison, WI, pp. 150,151, pp. 216,217.
32. Richardson, C. C. (1972) Polynucleotide kinase, in *Procedures in Nucleic Acids Research,* vol. 2 (Cantoni, G. L. and Davies, D. R., eds.), Harper and Row, New York, pp. 815–828.
33. Maxam A. M. and Gilbert, W. (1980) Sequencing of end-labeled DNA with base specific chemical cleavages. *Methods Enzymol.* **65,** 499–560.
34. Szostak, J. W., Stiles, J. I., Tye, B. K., Chiu, P., Sherman, F., and Wu, R. (1980) Hybridisation with synthetic oligonucleotides. *Methods Enzymol.* **68,** 419–428.
35. Agarwal, K., Brunstedt, J., and Noyes, B. E. (1981) A general method for detection and characterisation of an mRNA using an oligonucleotide probe. *J. Biol. Chem.* **256,** 1023–1028.
36. Donis-Keller, H. (1977) Mapping adenines, guanines, and pyrimidines in RNA. *Nucleic Acids Res.* **4,** 2527–2538.
37. Tu, C. P. D. and Wu, R. (1980) Sequence analysis of short DNA fragments. *Methods Enzymol.* **65,** 620–638.
38. Reddy, R., Henning, D., Epstein, P., and Busch, H. (1981) Primary and secondary structure of U2 snRNA. *Nucleic Acids Res.* **9,** 5645–5658.
39. Johnsrud, L. (1978) Contacts between *Escherichia coli* RNA polymerase and a lac operon promoter. *Proc. Natl. Acad. Sci. USA* **75,** 5314–5318.
40. Galas, D. and Schmitz, A. (1978) DNAse footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* **5,** 3157–3170.
41. Reddy, M. V., Gupta, R. C., and Randerath, K. (1981) ^{32}P-base analysis of DNA. *Anal. Biochem.* **117,** 271–279.
42. Silber R., Malathi, V. G., and Hurwitz, J. (1972) Purification and properties of bacteriophage T4 induced RNA ligase. *Proc. Natl. Acad. Sci. USA* **69,** 3009–3013.
43. Royer-Pokora, B., Gordon, L. K., and Haseltine, W. A. (1981) Use of exonuclease III to determine the site of stable lesions in defined sequences of DNA:

the cyclobutane pyrimidine dimer and *cis* and *trans* dichlorodiammine platinum II examples. *Nucleic Acids Res.* **9,** 4595–4609.

44. Harrison, B. and Zimmerman, S. B. (1986) T4 polynucleotide kinase: Molecular crowding increases the efficiency of reaction at DNA termini. *Anal. Biochem.* **158,** 307–315.
45. Geck, P. and Nasz, I. (1983) Concentrated digestible DNA after hydroxylapatite chromatography with cetylpyridinium bromide precipitation. *Anal. Biochem.* **135,** 264–268.
46. Stawinski, J., Hozumi, T., Narang, S. A., Bahl, C. P., and Wu, R. (1977) Arylsulfonyltetrazoles, new coupling reagents and further improvements in the triester method for the synthesis of deoxyribo-oligonucleotides. *Nucleic Acid Res.* **4,** 353–371.
47. Narang, S. A., Brousseau, R., Hsiung, H. M., and Michniewicz, J. T. (1980) Chemical synthesis of deoxy-oligonucleotides by the modified triester method. *Methods Enzymol.* **65,** 610–620.

Index

C

R