GENETIC ENGINEERING

J.G.Williams

R.K.Patient

GENETIC ENGINEERING

J.G.Williams

Imperial Cancer Research Fund, Clare Hall Laboratories, Blanch Lane, South Mimms, Potters Bar, Herts EN6 3LD, UK

R.K.Patient

Department of Biophysics, Cell and Molecular Biology, Kings' College, London WC2B 5RL, UK

IRL Press
Eynsham
Oxford
England

First published 1988

British Library Cataloguing in Publication Data

Genetic engineering.
1. Man. Genetic engineering. Cloning
I. Williams, Jeff II. Patient, Roger K.
III. Series
573.2'1

Library of Congress Cataloging in Publication Data

Williams, Jeff.
Genetic engineering/Jeff Williams. Roger K. Patient.
p. cm. – (In focus)
Includes bibliographies and index.
1. Genetic engineering. I. Patient, Roger K. II. Title.
III. Series: In focus (Oxford, England)
[DNLM: 1. Cloning, Molecular. 2. DNA, Recombinant. 3. Genetic Intervention. QH 442 W717g]
QH442.W55 1988
574.87'3282-dc19

ISBN 1 85221 071 0 (soft): $11.95

Typeset by Multiplex techniques ltd and printed by Information Printing Ltd, Oxford, England.

GENETIC ENGINEERING

IN FOCUS

Titles published in the series:

***Complement**

Enzyme Kinetics

Gene Structure and Transcription

Genetic Engineering

***Immune Recognition**

***Lymphokines**

Molecular Basis of Inherited Disease

Regulation of Enzyme Activity

***Published in association with the British Society of Immunology.**

Series editors

David Rickwood

Department of Biology, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, UK

David Male

Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, UK

Preface

The ability to purify specific pieces of prokaryotic and eukaryotic DNA has greatly accelerated progress in almost all fields of biology and is already beginning to find major medical and commercial applications. The aim of this book is to describe how DNA cloning is used to isolate and manipulate DNA. It is not possible in a book of this length to describe all of the procedures in use, however, our aim is to present the basic principles and to describe the most commonly used methods in detail emphasizing the latest methods and their applications.

J. Williams
K. Patient

Abbreviations

A	adenine
*Alu*I	restriction enzyme cutting at AGCT
ATP	adenosine triphosphate
*Bam*HI	restriction enzyme cutting at GGATCC
β-gal	β-galactosidase
bp	base pair
C	cytosine
cDNA	complementary DNA
CIP	calf intestinal phosphatase
cos	joined bacteriophage λ cohesive ends
Cys	cysteine
DNA	deoxyribonucleic acid
DNA pol I	DNA polymerase I
DNase	deoxyribonuclease
dC	deoxycytidylic acid
dG	deoxyguanidylic acid
dT	deoxythymidine
*Eco*RI	restriction enzyme cutting at GAATTC
*Eco*RII	restriction enzyme cutting at CC(A or T)GG
G	guanine
Gln	glutamine
Glu	glutamic acid
*Hin*dIII	restriction enzyme cutting at AAGCTT
IPTG	isopropylthiogalactoside
kb	kilobase pairs
*Kpn*I	restriction enzyme cutting at GGTACC
lacZ	β-galactosidase gene
MCS	multiple cloning site
Met	methionine
mRNA	messenger RNA
NTP	(unspecified) nucleoside triphosphate
oligo(dT)	oligodeoxythymidylate
phage	bacteriophage
RF	replicative form
RNA	ribonucleic acid

RNase	ribonuclease
*Sal*I	restriction enzyme cutting at GTCGAC
*Sau*3AI	restriction enzyme cutting at GATC
T	thymine
Trp	tryptophan
Tyr	tyrosine
Xgal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Contents

1

Introduction

1. General aspects of gene cloning

The common feature of all cloning experiments is the introduction into a host cell of a single DNA molecule which is able to replicate in parallel with the host cell genome but which remains in an episomal, non-integrated state. The simplest form of cloning might be considered to occur naturally, when a single bacteriophage (phage) particle infects a susceptible *Escherichia coli* host cell and a phage plaque is formed containing the progeny of the single infecting particle. Genetically engineered derivatives of bacteriophages have been constructed into which foreign segments of DNA can be introduced by enzymatic manipulation. When a complex mixture of DNA molecules is introduced into such a phage vector and a single plaque is selected after infection of *E.coli*, it will contain only one of the different 'target' molecules. Such a DNA molecule is said to be a recombinant. The power of the technique is illustrated by the fact that it is now possible to isolate a specific recombinant DNA from the 10^6 different fragments generated by breaking up a eukaryotic genome into clonable segments of DNA.

In general, recombinant DNA molecules are isolated by cloning in *E.coli* because there is such a wealth of experience in manipulating this bacterium genetically and because it has such a short generation time. This background of fundamental research has made it possible to construct elegantly engineered bacteriophage vectors designed to facilitate cloning. There are also extremely valuable vectors derived from bacterial drug resistance plasmids; these exist in wild-type *E.coli* cells as extrachromosomal, circular DNA molecules. Naturally occurring plasmids often encode genes for multiple drug resistance but the commonly used vectors contain only one or two such genes. Here the vector is introduced into *E.coli* directly, to transform the host cell to drug resistance. Each colony of bacteria which grows up is a clone containing the progeny of only one of the plasmid DNA molecules used in transformation.

By propagating the recombinant phage or recombinant plasmid in bacteria it is possible to prepare milligram amounts of DNA. This can be analysed to determine the precise organization and coding potential of the cloned gene. The gene can then

be used in many different ways. Thus a human globin gene might be used as a probe in nucleic acid hybridization to identify the specific mutation responsible for a hereditable disease, such as a thalassaemia. If the gene encodes a commercially valuable product it might be inserted into a prokaryotic or eukaryotic vector designed to maximize production of the specific protein. These, and the many other uses to which recombinant DNA technology can be put, all utilize enzymes which can be used to manipulate nucleic acids.

2

The enzymology of genetic engineering

1. Introduction

The basic substrate of molecular cloning is DNA and cloning is only possible because of the availability of enzymes able to make, break and join pieces of DNA. Before considering these enzymes individually, it is necessary to consider the basic structure of DNA and the way in which it is synthesized and degraded *in vivo*. A DNA chain is built up by the polymerization of nucleoside triphosphate molecules. These are composed of one of the four bases, guanine (G), cytosine (C), adenine (A) or thymine (T), attached to a deoxyribose ring at its 1′-carbon residue, with a triphosphate moiety on the 5′-position of the sugar ring (*Figure 2.1*). A DNA polymer is constructed by the sequential addition of nucleoside monophosphate residues which attach to the growing chain at the 3′-hydroxyl of the sugar ring of the preceding residue in the sequence. Thus DNA chains are said to grow in a 5′ to 3′ direction and the sequence of a DNA chain is always presented in this way, that is with the 5′-terminus residue, the first nucleotide to be incorporated during synthesis, as the first residue. The two antiparallel strands of a double-stranded DNA molecule are held together by complementary base-pairing between the C and G bases, and between A and T. The two strands can be separated (denatured) by heat treatment or exposure to high pH. They will rejoin (anneal) to one another when incubated under suitable conditions of temperature and ionic strength. This is known as nucleic acid hybridization and it is highly specific; a single strand of DNA will hybridize to its complementary sequence but not to unrelated sequences. This specificity allows nucleic acid hybridization to be used as a very powerful tool for detecting specific sequences in a complex mixture of nucleic acid molecules (1).

2. Enzymes which synthesize DNA

2.1 Terminal transferase

This enzyme, which is found in the mammalian thymus, adds residues to any free 3′-terminus. It is used to add complementary homopolymer tracts, for example poly

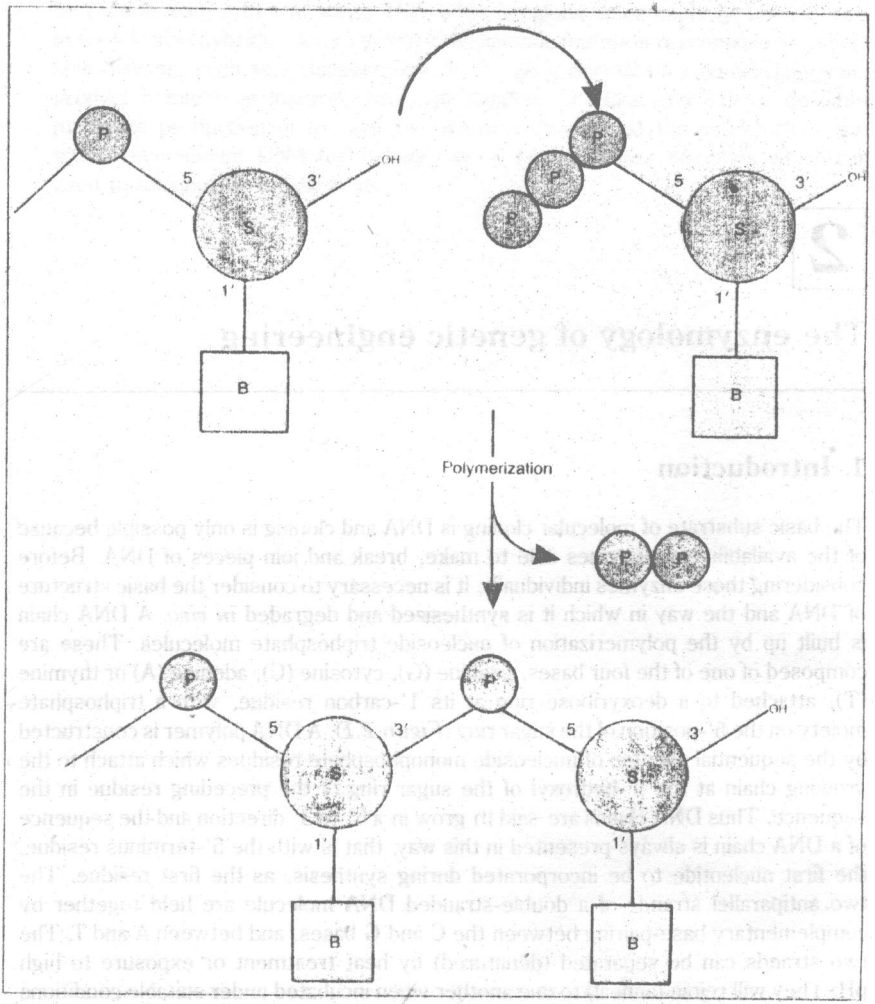

Figure 2.1. The synthesis of a DNA molecule. P, phosphate; S, sugar (deoxyribose); B, base.

deoxyguanidylic acid (dG) and poly deoxycytidylic acid (dC) to two DNA molecules which are to be joined together for cloning (2).

2.2 DNA polymerase I

A DNA chain is normally synthesized by the copying of a pre-existing DNA molecule and enzymes capable of such a reaction require as a substrate a template–primer complex (*Figure 2.2*). The *E.coli* enzyme DNA polymerase I will utilize such a substrate to synthesize a double-stranded DNA. In addition to its 5′- to 3′-polymerase activity, this enzyme has two intrinsic nuclease activities which are responsible for correcting errors introduced during DNA synthesis *in vivo*. These are exonuclease

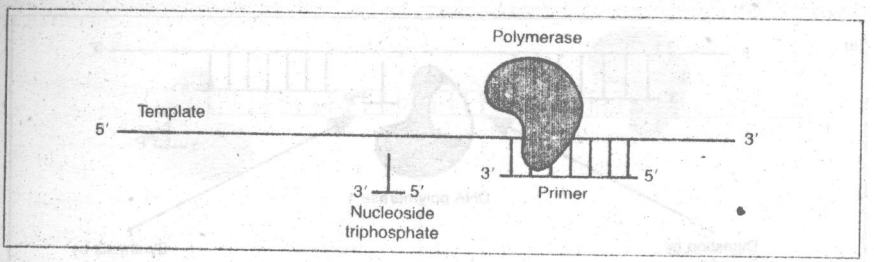

Figure 2.2. DNA synthesis on a template-primer complex.

activities, that is they progressively remove bases from the ends of the DNA. There are, within the single polypeptide chain of the enzyme, two separate domains, one of which acts to digest DNA in a 5′ to 3′ direction and the other in the opposite (3′ to 5′) direction.

One of the major uses of this enzyme is to generate radioactively labelled DNA probes. These are used in nucleic acid hybridization to detect complementary DNA sequences. In this process, which is termed nick translation (3), DNA is treated with a very low level of DNase I, an endonuclease which cleaves at more or less random points within one of the two strands of the DNA. The four nucleoside triphosphates, one or more of which contains a [^{32}P]phosphorus atom in the α-phosphate group (this phosphate group is retained after polymerization), and DNA polymerase I are then added. The 5′ to 3′ exonucleolytic action of this enzyme allows it to degrade the existing unlabelled strand and the polymerase activity allows it simultaneously to synthesize a new isotopically labelled copy of the complementary strand (*Figure 2.3a*).

The domain responsible for the 5′ to 3′ exonuclease activity of DNA polymerase I can be cleaved from the enzyme to leave a large polypeptide, known as the Klenow fragment, which retains the polymerase and 3′ to 5′ exonuclease activities only. This fragment has several very important uses because, in the presence of excess nucleoside triphosphates, its 3′ to 5′ exonuclease activity is supressed and it can act on a template–primer complex to generate a perfect double-stranded DNA copy of the template. It is, therefore, used to copy single-stranded DNA and one of its major uses is in the 'dideoxy method' of DNA sequence analysis (see Chapter 5). It can also be used as a method to generate highly labelled DNA as an alternative to nick translation (*Figure 2.3b*). If the DNA to be labelled is double stranded, it is first denatured, by heating, and then rapidly cooled to prevent re-annealing. A random mixture of short (10–20 residues) oligonucleotides, synthesized chemically using a mixture of all four nucleotide precursors, provides non-specific primers, with sufficient complementarity to initiate DNA synthesis (4). If a radiolabelled triphosphate is incorporated during copying, a high specific activity probe can be synthesized. Because it lacks the 5′ to 3′ exonuclease activity, the Klenow fragment is also used to fill in overhanging ends generated on double-stranded DNA by endonucleases which cleave the DNA leaving a projecting 5′-terminus. The latter function is sometimes performed using the DNA polymerase encoded by phage T4, which has very similar properties to the Klenow fragment.

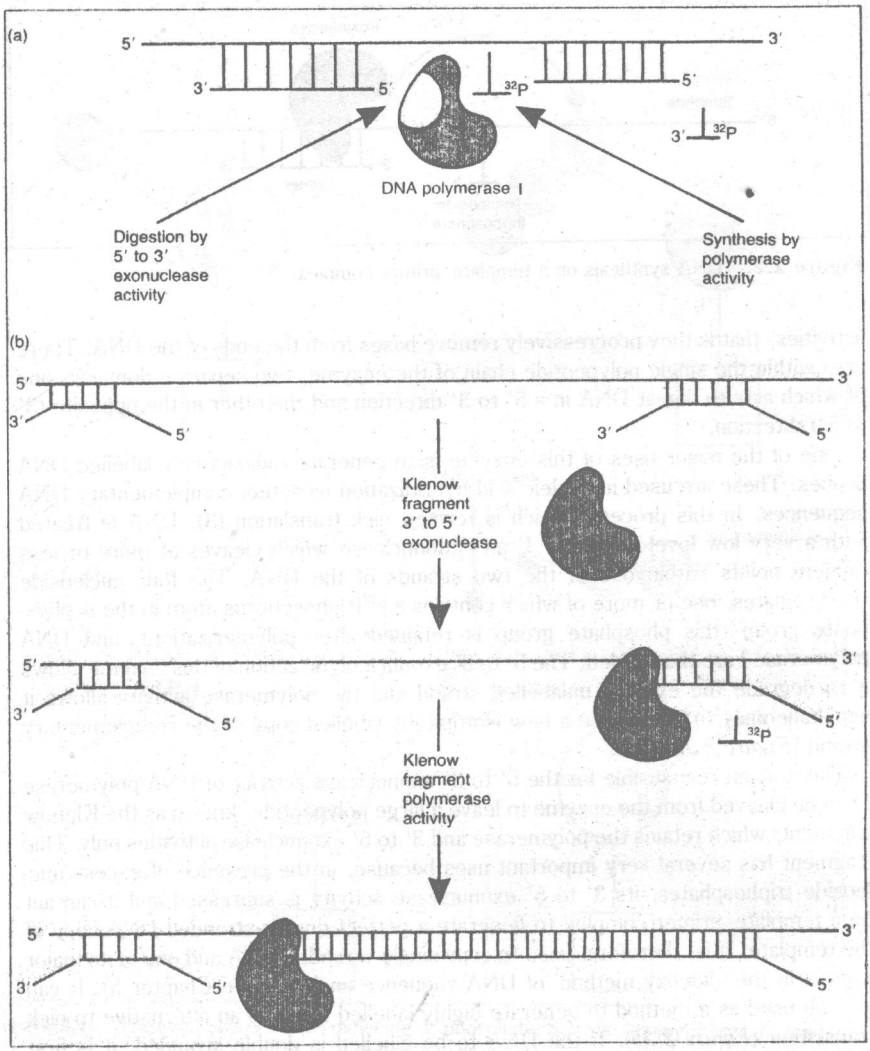

Figure 2.3. Isotopic labelling of DNA. (a) Nick translation, (b) random–primed synthesis.

2.3 Reverse transcriptase

This enzyme, which is found within the virion of RNA tumour viruses, will copy single-stranded RNA or DNA templates in the presence of a complementary primer. It has no associated exonuclease activities. Although occasionally used to copy DNA, its major use is to synthesize complementary DNA (cDNA) copies of polyadenylated eukaryotic mRNA sequences (see Chapter 4).

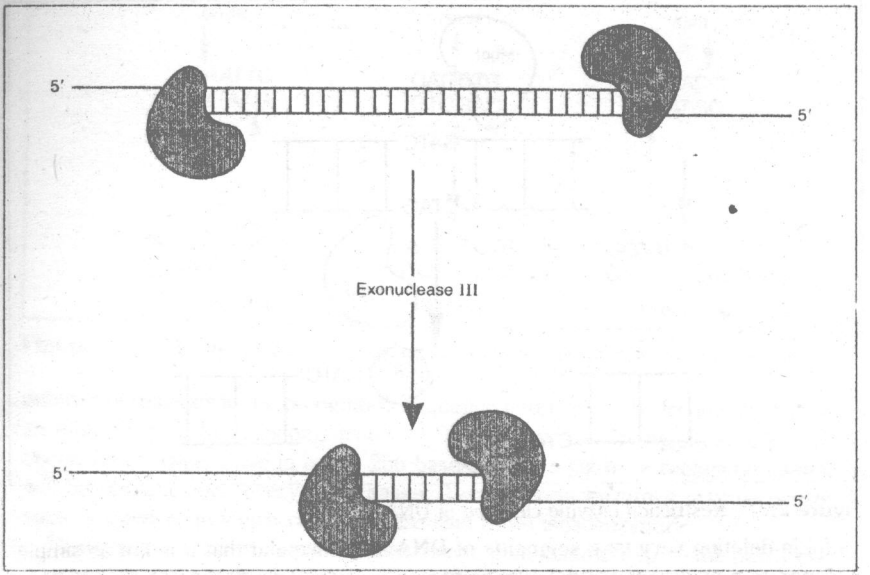

Figure 2.4. Degradation of DNA by exonuclease III.

3. Enzymes which degrade DNA

3.1 Nuclease S1

This enzyme from the fungus *Aspergillus* degrades single-stranded DNA in an endonucleolytic manner (5). It is used to remove projecting 3′-regions of single-stranded DNA in some cloning procedures and in the analysis of RNA transcripts in a hybridization technique known as S1 mapping (see Chapter 5).

3.2 Exonuclease III

This is an *E.coli* enzyme which progressively degrades one of the two strands of double-stranded DNA (6). It acts only on 3′-termini which must be part of a DNA double helix. Hence the enzyme will only initiate degradation upon flush-ended, double-stranded DNA, or DNA with a projecting 5′-terminus (*Figure 2.4*). This property has formed the basis of a widely used method of deleting DNA in a unidirectional fashion (see Chapter 5).

3.3 Bal 31

This is an endonuclease which simultaneously degrades both strands of double-stranded DNA (7). It is also used to delete regions of DNA sequence selectively. Its advantage over exonuclease III is that it deletes both strands of the DNA simultaneously. When a deletion 'series' is prepared using exonuclease III the 5′-strand of the DNA is left intact and must be degraded in a subsequent reaction using an enzyme such as nuclease S1. The major disadvantages of Bal 31 are that it is not as