Gene Cloning

The Mechanics of DNA Manipulation

David M. Glover

Gene Cloning

The Mechanics of DNA Manipulation

David M. Glover

Cancer Research Campaign
Eukaryotic Molecular Genetics Group
Department of Biochemistry
Imperial College of Science and Technology
London

London New York
CHAPMAN AND HALL

First published 1984 by
Chapman and Hall Ltd
11 New Fetter Lane, London EC4P 4EE
Published in the USA by
Chapman and Hall
733 Third Avenue, New York NY 10017

Printed in Great Britain at the
University Press, Cambridge

ISBN 0 412 26600 8 (HB)
ISBN 0 412 25430 1 (PB)

British Library Cataloguing in Publication Data

Glover, David M
Gene cloning.—(Outline studies in biology)
1. Molecular Cloning 2. Recombinant DNA
I Title II. Series
574.87'3282 QH442.2

ISBN 0-412-26600-8
ISBN 0-412-25430-1 Pbk

Library of Congress Cataloging in Publication Data

Glover, David M.
Gene cloning.

Rev. ed. of: Genetic engineering, cloning DNA. 2nd ed. 1980.
Bibliography: p.
Includes index.
1. Molecular cloning. 2. Recombinant DNA.
I. Glover, David M. Genetic engineering, cloning DNA. II. Title.
QH442.2.G56 1984 574.87'3282 84-9575
ISBN 0-412-26600-8
ISBN 0-412-25430-1 (pbk.)

Preface

This book was originally conceived in the form of a second edition of a volume published in 1980 in Chapman and Hall's 'Outline Studies in Biology' series and entitled *Genetic Engineering – Cloning DNA*. It very rapidly became apparent that with the impact of recombinant DNA techniques being felt in so many areas of biology, it was going to be difficult if not impossible to keep the book within the space confines of these little monographs. The stays were therefore loosened and the book expanded comfortably to its present size. I hope that this extra space has allowed me to clarify sections of the text that were 'heavy going' in the earlier version. The extra space has certainly allowed me to cover topics that were not mentioned at all in the earlier book. These are primarily to be found in Chapters 7 and 8, which cover the rapid advances that have been recently made in the use of plant and animal cells as hosts for recombinant DNA molecules. The development of other vectors has certainly not stood still over the past four years. This has necessitated a thorough revision of Chapters 3 and 4, which deal with bacteriophage and bacterial plasmid vectors. Numerous techniques for *in vitro* mutagenesis have now been tried and tested allowing me to give comprehensive coverage of this area in Chapter 2 along with the biochemical techniques used to construct recombinant DNA molecules. Readers with some background knowledge of the approaches to gene cloning will be able to go straight to a part of the book in which they are specifically interested. On the other hand, undergraduate readers with little background knowledge should benefit from studying the first chapter in which the basic principles of gene cloning are expounded.

The topics covered by the book are taught in a four week course, 'Genetic Engineering', given for final year undergraduates studying Molecular Biology at Imperial College. The book is intended for use in such advanced undergraduate study and also for graduate school coursework. I hope that it will be read as well by research workers wanting a general overview of the subject. When I first reviewed the subject of Gene Cloning in an article written in late 1975, I cited 38 papers which covered the whole field. Nowadays, it is difficult to find any paper dealing with the molecular biology of Nucleic Acids that does not use DNA cloning technology. Consequently it has been difficult to be selective in providing references. I hope that students will not feel overburdened by the references that I have given and that researchers will not feel neglected if their particular area has not been covered.

Preparation of the manuscript has taken considerable time and I am grateful to my research students and other colleagues for their tolerance of my involvement in the project. Finally I am grateful to Jean Beggs, Conrad Lichtenstein and Peter Rigby of Imperial College and Bill Brammar of Leicester University for their helpful comments on the manuscript at various stages of its evolution. Finally, I thank Martin Eilers, Manfred Frasch, Herbert Jäckle and Harry Saumweber of the Max Planck Institut für Entwicklungsbiologie in Tübingen for their help in reading the proofs.

Imperial College

David M. Glover

Cover picture

The photographs on the cover illustrate the procedure of colony hybridization explained in Section 1.2 (see also Fig. 1.1). The upper photograph shows a 'master' plate on which there are colonies of *Escherichia coli* containing recombinant plasmids spotted out to form a grid. The master plate has been used to 'print' out a number of replicas onto nitrocellulose filters. These have been incubated on nutrient agar allowing new colonies to grow in identical positions to those on the master plate. The nitrocellulose filters are later removed, and the DNA in the colonies denatured and immobilized on the filter using procedures described in Section 1.2. The filters are incubated with a radiolabeled nucleic acid probe which hybridizes to any complementary nucleic acid sequences in the colonies. The unhybridized radioactivity is washed away and the position of hybridization revealed by autoradiography. The lower photograph shows such an autoradiograph in which the probe has hybridized to DNA from two of the colonies. This particular experiment was carried out by Sarah Millar.

Contents

1 The principles of cloning DNA

Past progress in understanding the molecular biology of prokaryotic gene expression has relied heavily upon studies involving bacteriophage and bacterial plasmids. Of the bacteriophage themselves, the *E. coli* phage λ is perhaps the best characterized. The interaction of phage λ with the host cell is a particularly fruitful area of study, as here are a set of genes which can either direct cell lysis or become stably associated with the host chromosome in lysogeny. In the production of infectious phage from lysogens, the excision of the phage genome from the *E. coli* chromosome is usually precise. Occasionally, however, the excision is imperfect and results in a λ phage transductant which carries a segment of the bacterial chromosome that was adjacent to the phage attachment site. Such specialized transducing phage have been invaluable, providing the means to assay for specific messenger RNAs by nucleic acid hybridization or enabling the production of large amounts of particular gene products. Research on the bacterial plasmids has had a similar history. The discovery and rationalization of the mechanism whereby F factors promote bacterial conjugation was central to the development of *E. coli* genetics. Just as the imperfect excisions of phage λ from its lysogenic state can result in a circular phage genome carrying a segment of bacterial DNA, so the imperfect excision of an F plasmid from an Hfr strain results in an F′ plasmid which also carries a segment of bacterial DNA. Such F′ plasmids have been invaluable 'vectors' for carrying specific genes from one E. *coli* strain to another and have perhaps been most useful in the construction of merodiploid strains which have enabled the elucidation of the control circuits of many bacterial operons.

The principles of genetic engineering which are described in this book are analogous to these 'natural' events, but they overcome the limitation of an absolute dependence upon the *in vivo* recombinational mechanisms of the *E. coli* cell. The techniques for recombining DNA *in vitro* will be described in detail in Chapter 2. It is possible to insert DNA from any organism into a plasmid or viral replicon to form a chimaeric molecule. The host for this molecule can be a prokaryotic or a eukaryotic cell depending upon the replication origin present in the vector. The methods that were originally described for constructing recombinant DNA molecules *in vitro* were not straightforward and involved many enzymatic steps [1, 2] (see Section 2.1.3). Subsequently it was realized that certain restriction endonucleases generate cohesive ends when they cleave DNA. DNA molecules possessing these cohesive ends can be easily rejoined using DNA ligase [3] (see Section 2.1.2). The procedures for joining DNA molecules *in vitro* thereby became very much simpler and within the scope of many laboratories.

1.1 The debate on the safety of work with recombinant DNA

The techniques mentioned in the previous section, were originally applied to the joining of DNA from the tumour virus SV40 with bacterial plasmid DNA. These recombinants were not introduced into *E. coli* because a great deal of concern was voiced within the scientific community about the hypothetical hazards of introducing the DNA of a tumour virus into a bacterium capable of growing within the mammalian intestinal tract. Subsequently this concern was to spread to other experiments: perhaps there were genes within the mouse genome, for example, that could be dangerous when introduced into and expressed within *E. coli*. A meeting of scientists was called at Asilomar in California in 1975 to discuss these issues. An early assessment of the problems as they were then perceived can be found in the summary statement of the conference [4]. In the absence of evidence viewpoints became sharply polarized, and the whole issue provoked much debate and many reports from governmental and scientific bodies [5, 6]. At that time the publications from such bodies considerably outweighed scientific reports of experiments using recombinant DNA techniques. Fortunately that situation did not persist for long and publications which utilize recombinant DNA techniques now dominate the literature in the field of molecular biology. The main fear was that *E. coli* carrying a potentially dangerous cloned gene could accidentally escape the laboratory and successfully colonize the intestinal tracts of laboratory workers and precipitate some disastrous pandemic. The other point of view was that prokaryotic organisms within nature were

frequently in contact with and must therefore take up eukaryotic DNA from decaying plant and animal matter. It is then likely, given the enormity of the earth's population of microorganisms, that recombination analogous to the type that can now be carried out *in vitro* has already had the chance to occur and that the resulting 'recombinant organisms' have no selective advantage.

The concern about the biological safety of these experiments stimulated work on the development of biologically 'safe' host–vector systems. These utilized vectors that would only grow in certain strains of *E. coli* that could not survive outside the laboratory. Some of these developments will be encountered in Chapters 3 and 4, in which the bacteriophage and plasmid vectors of *E. coli* are described. The standard laboratory strains of *E. coli* have now been cultured for several decades and are very much enfeebled, and have low survival in the human gut. In virtually all cloning experiments, the recipient *E. coli* strain is deficient in its endogenous restriction system (see Section 2.1.1). This ensures that foreign DNA will not be degraded when it is introduced into the bacterium during cloning. These strains are consequently highly susceptible to bacteriophage infections. Furthermore the strains are often also deficient in one of the recombination pathways, usually *rec*A. This makes the bacteria deficient in their ability to promote homologous recombination, an essential factor if, for example, one is propagating tandemly arranged genes from a higher eukaryote. *Rec*A$^-$ strains are, however, extremely sensitive to ultraviolet light. It is possible to build in many more safeguards to ensure that a bacterium containing a potentially hazardous DNA molecule has a very low chance of survival outside the laboratory environment. One such strain was developed after the Asilomar conference and was available in the year in which the United States celebrated 200 years of independence and was appropriately christened χ1776 [7]. This strain contains mutations which make its growth absolutely dependent upon the presence of diaminopimelic acid and thymidine in the culture medium. Diaminopimelic acid (DAP) is an intermediate in the biosynthesis of lysine and is not found within the mammalian intestinal tract. The strain cannot therefore survive if it is ingested. It also shows increased sensitivity to ultraviolet irradiation as a result of removal of one of the genes responsible for DNA repair, *uvr*B, and it is very sensitive to both bile salts and ionic detergents. The strain is consequently difficult to handle in the laboratory unless scrupulously clean, detergent-free glassware is used. It is, however, possible to transform the strain with efficiencies comparable to those obtained with other commonly used recipient strains. Nowadays χ1776 is rarely used, since many of the hypothetical hazards of recombinant DNA work are now viewed from a more rational perspective and in

1979 many of the guidelines for work with recombinant DNA were relaxed. There is now some ten years experience of working with eukaryotic DNA cloned in bacterial cells. In the process an enormous amount has been learned about the eukaryotic genome which suggests that many of the early fears were ill-founded. Largely as a result of the exploitation of the recombinant DNA techniques, we now know that most eukaryotic genes have a pattern of chromosomal organization that precludes their expression in prokaryotic cells. In order to get expression of eukaryotic genes in *E. coli* it is usually necessary to clone either a DNA complement of the mRNA of interest or a chemically synthesized gene that lacks the intervening sequences commonly found in eukaryotic genes. Furthermore these must be linked to the correct prokaryotic signals for the initiation of transcription and translation (see Chapter 5).

1.2 General principles of cloning

The vectors that have been specifically designed to carry foreign DNA sequences have extended the traditional genetic approach developed with transducing phages and F plasmids. The techniques for cloning DNA were worked out using *E. coli* as the host organism. In Chapters 3 and 4 the phage and plasmid cloning vectors of *E. coli* will be described in detail. A vector should have its own replicon and thereby be capable of autonomous replication in the host cell; it should carry one or more selectable marker functions, in order to permit the recognition of cells carrying the parental form of the vector or a recombinant between the vector and foreign DNA sequences; finally, it should have sites for a variety of restriction endonucleases which cleave the molecule once or, at the most, twice in non-essential regions so that foreign DNA may be inserted into the vector or used to replace a segment of the vector. For the purposes of this introductory discussion, we will consider the general experimental approach for cloning DNA into the circular DNA of a plasmid vector or into the linear DNA of a vector derived from bacteriophage λ. This approach is illustrated for a plasmid vector in Fig. 1.1. The variations on this scheme applicable to cloning in λ vectors are described in outline later in this section and in detail in Chapter 3. The circular plasmid DNA must first be cleaved, preferably at a unique site to generate a linear molecule. One of the several techniques described in Chapter 2 is then used to join this linear vector molecule to the foreign segments of DNA that are to be cloned. A number of products can result from this reaction. In terms of the molecules that are subsequently capable of replicating in *E. coli* one has to consider both the formation of circular recombinant DNA molecules and the possibility that a significant proportion of circular

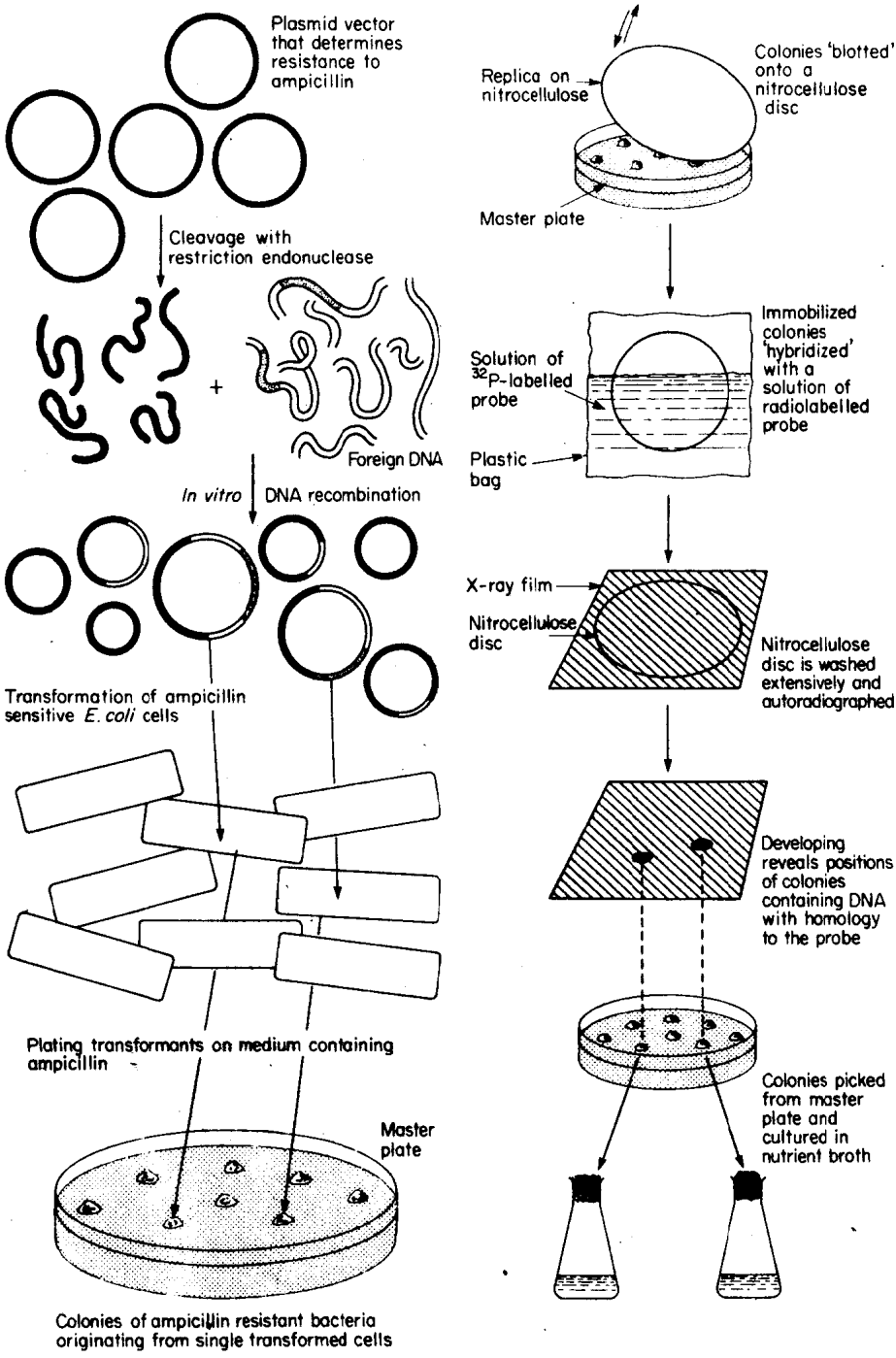

Figure 1.1 A general scheme for cloning DNA in a plasmid vector.

vector molecules can be regenerated. (Linear plasmids, recombinant or parental, transform *E. coli* at least an order of magnitude less efficiently than do circular molecules.) The products of the reaction are thus a highly heterogeneous mixture of recombinant molecules together with parental plasmids. This mixture is usually introduced directly into *E. coli* by a variation on the transformation procedure originally developed by Mandel and Higa [8]. In this procedure DNA is incubated with *E. coli* cells that have been made permeable by treatment with Ca^{2+} ions. The procedure was first used as a means of introducing λ DNA directly into *E. coli* and so acquired the term 'transfection' (a hybrid of transformation and infection). This term is still generally applied to the procedure. The efficiency of this process was such that about 10^5 transformants could be generated per μg of a covalently closed circular plasmid such as pBR322. The experimental parameters for transfection have now been more rigorously examined [9] and it is possible to achieve efficiencies of over 10^8 transformants per μg using the protocols that have emerged. DNA cloned in the plasmid vector is incubated with the calcium-treated cells and then plated out onto selective medium. Often the marker function on the plasmid determines resistance to an antibiotic. In this case the cells would be plated onto nutrient agar containing that antibiotic. The conditions of the transformation are such that only a single plasmid molecule enters a bacterial cell. The single transformed cell then grows to give a colony of cells on the plate. Each of these colonies contains a homogeneous population of identical – 'cloned' – plasmids derived from the single transforming plasmid.

The linear DNA of bacteriophage λ vectors also has to be cleaved with a restriction endonuclease before it can accept segments of foreign DNA. The foreign DNA is also usually cleaved with a restriction endonuclease so it can be joined to the vector by the action of DNA ligase. There are two major types of λ vector, which are described in Chapter 3. One type has a single cleavage site into which the foreign DNA is *inserted* (Section 3.1.2(b)). In the other, the vector has two cleavage sites for the restriction enzyme and DNA is inserted between these two sites to *replace* a segment of DNA in the vector (Section 3.1.2(c)). The heterogeneous population of molecules from this reaction mixture can then be introduced into bacteria by the transfection procedure outlined above. There is, however, an alternative strategy whereby λ DNA can be packaged *in vitro* into phage particles which can be used to infect the host cell (Section 3.1.3(c)). The insertion of foreign DNA does not affect the ability of the recombinant molecule to replicate in the host bacterium and to produce progeny recombinant phage particles. The infection of a single bacterial cell by a single recombinant molecule therefore results in the clonal propagation of the recombinant molecule. In this

way homogeneous populations of molecules can be prepared from phage preparations made initially by picking a single phage plaque from a bacterial lawn.

The next problem to be faced is how to identify which colony or which plaque contains the DNA sequence one wishes to isolate. The most commonly used screening approach requires that one has a pure or partially pure nucleic acid sequence which can be radiolabelled and used as a 'probe' to identify the corresponding recombinant DNA. The probe may be a partially pure mRNA, a chemically synthesized oligonucleotide or a related gene. The screening method of choice is the 'colony hybridization' technique developed by Grunstein and Hogness [10] suitable for use with plasmid vectors, or the analogous 'plaque hybridization' of Benton and Davis [11] for use with phage. In order to carry out colony hybridization the cells are first plated onto selectives plates. A replica of the colonies is made onto a nitrocellulose filter disc which is placed on the surface of a second plate (see Fig. 1.1). The colonies are allowed to grow on the master plate and the nitrocellulose disc. The nitrocellulose filter is then removed and placed onto blotting paper wet with 0.5 N NaOH solution. The alkali diffuses into the nitrocellulose, lyses the bacteria *in situ* and denatures their DNA. The filter is then neutralized using Tris buffer in the presence of high salt. The single stranded DNA binds to the nitrocellulose filter in the position originally occupied by the bacterial colony. The filter is then baked at 80° C, following which it is incubated with a solution containing the radiolabelled probe under conditions which favour nucleic acid hybridization. Unhybridized material is removed by extensive washing, thus allowing the identification of colonies containing sequences complementary to the probe by autoradiography. Colonies which give a positive autoradiographic signal can then be picked from the master plate and cultured in order to provide sufficient cells from which to make plasmid DNA. A virtually identical procedure is followed for making a replica of plaques generated in a bacterial lawn by bacteriophage λ recombinants. The filter onto which the plaques are transferred is, however, taken immediately to the alkali denaturation step.

In order to prepare plasmid DNA, cells are cultured in nutrient broth in flasks that are shaken vigorously at 37° C. The cells are pelleted by centrifugation and then treated with lysozyme in a buffered isotonic solution. A variety of detergent treatments can then be used to complete the breakage of the cells. These are carried out under conditions which cause bacterial chromosomal DNA and other cell debris to precipitate and plasmid DNA to be left in solution. Significant purification can therefore be achieved by medium speed centrifugation to pellet the chromosomal DNA and other debris.

Finally ethidium bromide is added to this supernatant together with caesium chloride to bring the solution to a density at which the DNA–dye complex will band. The DNA is then banded to equilibrium by ultracentrifugation. Ethidium bromide is an intercalating dye that will bind to both linear and circular DNA. There is, however, a topological restraint upon the amount of dye that can bind to covalently closed circular DNA. The contaminating chromosomal DNA is inevitably fragmented and this linear DNA, together with any nicked plasmid molecules, binds more ethidium and therefore bands at a lower density than the covalently closed circular DNA. Contaminating protein will float to the surface and RNA will pellet during the banding process. The band of supercoiled plasmid DNA can then be collected from the tube and the dye removed by extracting with an alcohol (see Fig. 1.2).

In order to prepare phage, a single plaque has to be picked and taken through a couple of steps of amplification. The first involves the lysis of a small culture of bacteria either in broth or on a single petri dish. The progeny phage from this step are titred and then used to infect a larger culture of cells in the early 'log phase' of their growth. The multiplicity of infection must be such that all the cells in the culture become infected and therefore all undergo lysis. Phage can be concentrated from the nutrient broth containing the lysed cells by precipitation with polyethylene glycol. This precipitate of phage and other cell debris is resuspended in caesium chloride solution and the phage particles purified by banding to equilibrium in the ultracentrifuge. A concentrated suspension of highly purified phage particles can be collected from such a density equilibrium gradient. The phage DNA is then extracted from the phage particles by deproteinization using phenol (see Fig. 1.3).

1.3 Strategies for gene cloning

1.3.1 Cloning by complementation

This approach has been most successful when the DNA to be cloned is from the same species as the host for the recombinant molecules. The approach may be considered as a direct extension of the techniques used in classical microbial genetics. In its simplest form, a 'library' of cloned segments of *E. coli* DNA (a heterogeneous population of recombinant DNA molecules representative of all the sequences in the *E. coli* genome) is introduced into the cells of an auxotrophic strain of *E. coli*. The recipient cells are plated on minimal medium lacking the substance required by the strain and those cells which have acquired a gene that complements the auxotrophic mutation will grow to produce colonies of transformed cells. The power of this

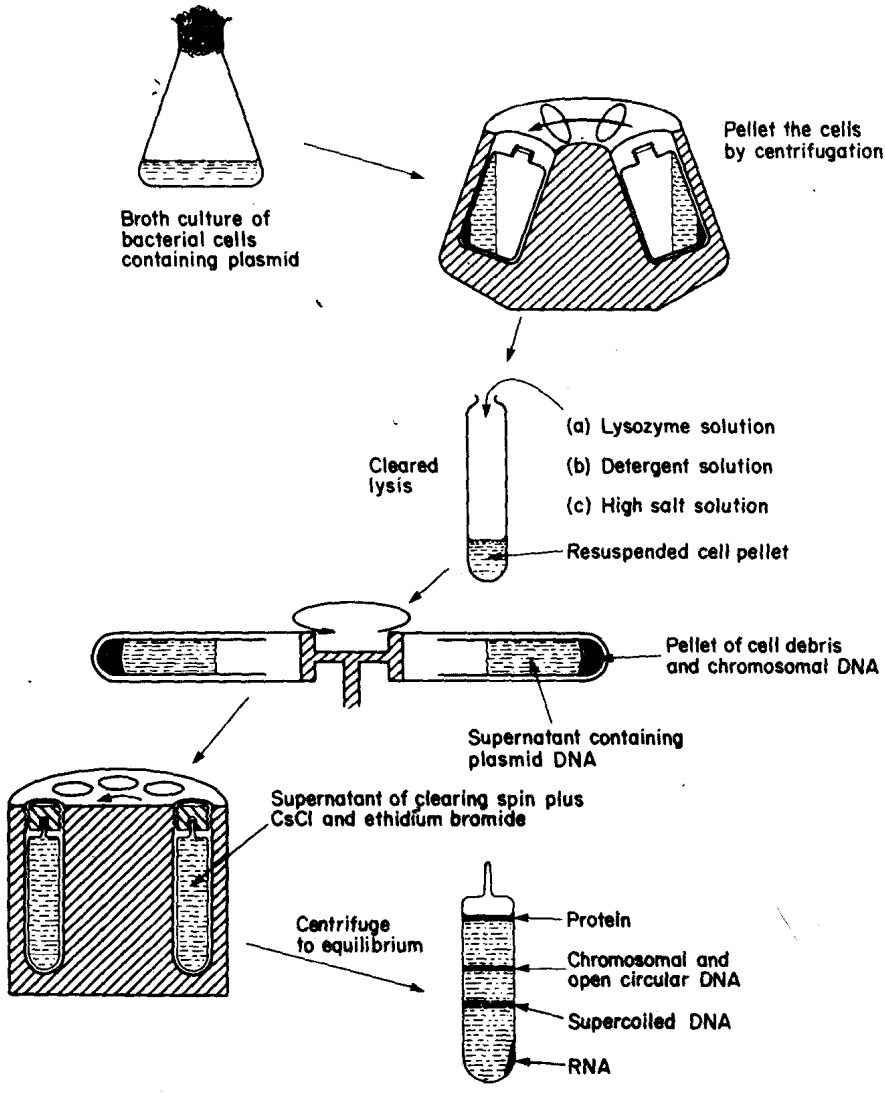

Figure 1.2 The preparation of plasmid DNA.

general approach was impressively demonstrated by Clarke and Carbon [12], who made libraries of random segments of *E. coli* DNA cloned into ColE1 by the dA:dT tailing technique (Section 2.1.3) and successfully isolated portions of several bacterial operons. This direct method of selection requires an auxotrophic strain that can be made competent for transformation. Some strains are, however, difficult to transform and so they developed an alternative approach whereby a

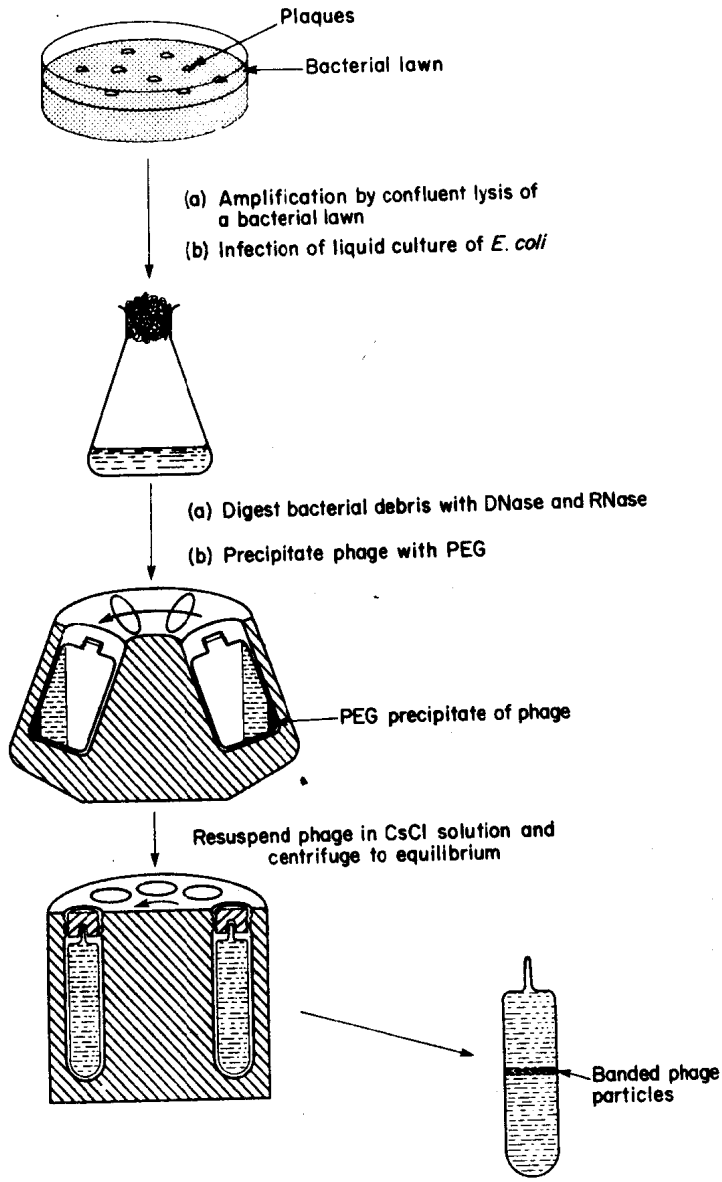

Figure 1.3 The preparation of bacteriophage λ.