INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University
Microfilms
International
300 N. Zeeb Road
Ann Arbor, MI 48106
Gayle, Richard Brownley, III

CONSTRUCTION AND CHARACTERIZATION OF ESCHERICHIA COLI PLASMIDS USEFUL IN THE MANIPULATION OF DNA

Rice University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

Ph.D. 1984
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs or pages √
2. Colored illustrations, paper or print
3. Photographs with dark background √
4. Illustrations are poor copy
5. Pages with black marks, not original copy
6. Print shows through as there is text on both sides of page
7. Indistinct, broken or small print on several pages
8. Print exceeds margin requirements
9. Tightly bound copy with print lost in spine
10. Computer printout pages with indistinct print
11. Page(s) __________ lacking when material received, and not available from school or author.
12. Page(s) __________ seem to be missing in numbering only as text follows.
13. Two pages numbered __________. Text follows.
14. Curling and wrinkled pages
15. Other

University Microfilms International
RICE UNIVERSITY

CONSTRUCTION AND CHARACTERIZATION OF ESCHERICHIA COLI PLASMIDS USEFUL IN THE MANIPULATION OF DNA

by

RICHARD BROWNLEY GAYLE III

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

DOCTOR OF PHILOSOPHY

APPROVED, THESIS COMMITTEE:

George N. Bennett
Assistant Professor of Biochemistry
Chairman

Kathleen Matthews
Associate Professor of Biochemistry

Frederick Rudolph
Associate Professor of Biochemistry

Susan Berget
Assistant Professor of Biochemistry

J. D. Prochnik
Assistant Professor of Biology

Houston, Texas
December, 1983
CONSTRUCTION AND CHARACTERIZATION OF ESCHERICHIA COLI
PLASMIDS USEFUL IN THE MANIPULATION OF DNA

Richard Brownley Gayle III

ABSTRACT

The restriction endonuclease \textit{MboII} recognizes a nonpalindromic, pentanucleotide DNA sequence and cleaves both strands of the DNA eight bases downstream, leaving a single 3' protruding base. \textit{MboII} is one of a group of restriction enzymes which has its site of cleavage removed from its recognition sequence. Insertions of DNA sequences into the \textit{MboII} cleavage site will not disrupt the recognition site allowing further additions to be performed. Several synthetic oligonucleotides containing \textit{MboII} sites have been used in this work to place \textit{MboII} sites in position for DNA manipulations. These oligonucleotides have been cloned into a plasmid vector which has no \textit{MboII} sites. The plasmids constructed with the synthetic oligonucleotides allow a DNA sequence to be inserted into the \textit{MboII} site, without disrupting the \textit{MboII} recognition sequence. Following isolation of the recombinant plasmids, further insertions have been performed, thus juxtaposing DNA sequences without the need for compatible restriction enzyme sites. In addition, a small DNA fragment has been constructed. This sequence consists of the \textit{lac} operator region, so that it may be easily screened, surrounded by two \textit{MboII} recognition sequences. This DNA fragment has
been randomly inserted into plasmid DNA using DNAase I. The MboII recognition sequences face outward and will cleave into surrounding DNA, removing up to 14 base pairs. This allows specific deletions to be performed, as well as insertion of DNA fragments into the cleaved DNA. Because of the mode of action of MboII, the entire sequence can be removed from the region of insertion. Some of the plasmids constructed in this thesis had unusual regions surrounding the origin of replication. This area was much smaller than had been previously reported necessary for stable replication. The stability of the plasmids in different strains was found to vary and they had a much higher copy number than pBR322 in exponentially growing cells, but did not amplify to the same extent as pBR322. The plasmids were found to be compatible with pBR322.
Acknowledgments

I would like to thank David Russell and Susan Manley for their help. They provided examples of what a scientist should be, in addition to providing their friendship. I would also like to thank Paul Miller for his friendship. He served as a proper example of a postdoctoral fellow. I would especially like to thank George Bennett. For over 4 years he has provided me with help and encouragement. He has given me insights into science and into the application of the scientific method. Along with this he has been a friend and a teacher. His presence in my life will never disappear, no matter where I am.

This research was supported by a research grant to Dr. G. N. Bennett from the National Institutes of Health (GM 26437) and by the Training Grant (GM 07833) from the National Institutes of Medical Sciences.
# Table of Contents

Chapter 1. Plasmid Replication and DNA Manipulation.

I. **INTRODUCTION** .................................................. 1
   - Plasmid Discovery and Characterization ........ 2

**GENETIC ORGANIZATION**
   - Conjugation and Mobilization .................. 9
   - Bactericinogenity and Immunity ............. 13

**PLASMID REPLICATION**
   - Incompatibility ......................................... 14
   - Modes of Replication and Segregation ........ 16
   - Host Enzymes Needed for Replication of Relaxing Plasmids ...... 21
   - DNA Regions Required for Plasmid DNA replication .......... 26
   - Control of Replication .............................. 28
   - Control of Replication in ColEl .................. 32

**USE OF PLASMIDS AS CLONING VECTORS** .......... 45

**CLONING STRATEGIES** ......................................... 51
   - Synthesis of Specific Sequences of DNA
     by Enzymatic Means ................................ 53
   - Chemical Synthesis of Specific DNA Sequences ........... 56
   - Enzymatic Procedures for DNA Manipulation ....... 69

**SUMMARY** .......................................................... 76

II. **EXPERIMENTAL PROCEDURES**

**MATERIALS** ..................................................... 77
   - Bacterial Strains ..................................... 77
   - Enzymes .................................................. 77
   - Plasmid DNA ............................................. 79
   - Bacterial Media ......................................... 79

**METHODS**
   - Rapid isolation of plasmid DNA .................. 80
   - Large scale isolation of plasmid DNA .......... 81
   - Gel electrophoresis .................................. 82
   - DNA fragment isolation ............................. 82
   - Enzyme reactions ..................................... 83
   - Transformations ....................................... 85
   - DNA sequencing ....................................... 85
   - Determination of plasmid copy number .......... 86
   - Plasmid stability ...................................... 87
   - Plasmid incompatibility ............................. 88
   - RNA polymerase binding to promoters .......... 88
   - lac operator selection ............................... 89
   - Selection of trp operon .............................. 90
   - DNAase I digestion of plasmid DNA ............ 91
Table of Contents - Continued

<table>
<thead>
<tr>
<th>Chapter 2. Construction of a plasmid cloning vector with a deletion in the origin of replication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTRODUCTION</strong> .......................................................................................................................... 92</td>
</tr>
<tr>
<td><strong>RESULTS</strong> ................................................................................................................................. 94</td>
</tr>
<tr>
<td>Ligation of replication origin and DHFR gene .............................................................................. 94</td>
</tr>
<tr>
<td>Plasmid stability ......................................................................................................................... 105</td>
</tr>
<tr>
<td>Studies of promoters present on pRBG156 ................................................................................... 110</td>
</tr>
<tr>
<td>Plasmid incompatibility .............................................................................................................. 116</td>
</tr>
<tr>
<td>Plasmid copy number .................................................................................................................. 118</td>
</tr>
<tr>
<td><strong>DISCUSSION</strong> ............................................................................................................................ 123</td>
</tr>
<tr>
<td>Chapter 3. Use of MboII restriction enzyme sites for DNA manipulations ...............................</td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong> .......................................................................................................................... 129</td>
</tr>
<tr>
<td><strong>RESULTS</strong> ................................................................................................................................. 129</td>
</tr>
<tr>
<td>Characterization of adaptor oligonucleotides .......................................................................... 133</td>
</tr>
<tr>
<td>Use of MboII sites in a plasmid-borne system ........................................................................... 150</td>
</tr>
<tr>
<td><strong>DISCUSSION</strong> .............................................................................................................................. 165</td>
</tr>
<tr>
<td>Chapter 4. Use of MboII &quot;cassettes&quot; for DNA manipulation .....................................................</td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong> .......................................................................................................................... 177</td>
</tr>
<tr>
<td><strong>RESULTS</strong> ................................................................................................................................. 177</td>
</tr>
<tr>
<td>Construction of the cassette ....................................................................................................... 186</td>
</tr>
<tr>
<td>Insertion of the cassette into DNA sequences .......................................................................... 192</td>
</tr>
<tr>
<td><strong>DISCUSSION</strong> .............................................................................................................................. 195</td>
</tr>
<tr>
<td>References ...................................................................................................................................... 201</td>
</tr>
</tbody>
</table>
List of Figures

1. Functional Maps of the Sex Factor F, the Resistance Factor R100, and the Bacteriogenic Plasmids, ColE1 and CloDF13.........................6
2. The Region Surrounding the Origin in Plasmid R1..31
3. RNA Transcripts Produced by the ColE1 Replicon...34
4. DNA Sequences of RNA I and RNA II.................38
5. Construction of pBR322................................50
6. Phosphate Diester Synthesis of Oligonucleotides..59
7. Phosphate Triester Synthesis of Oligonucleotides.62
8. Phosphite Triester Synthesis of Oligonucleotides.66
9. Types of Digestion of DNA by Nucleases.............71
10. Construction of pRBG156 and pRBG177.................97
11. Restriction Enzyme Mapping of pRBG156.............100
12. Nucleotide Sequence of pRBG156.....................102
13. Comparison of the Replication Regions of pBR322 and pRBG156.................................104
14. Comparison of the DHFR Promoter Regions of pMT100 and pRBG156........................107
15. Plasmid Stability.....................................112
16. RNA Polymerase Binding to pRBG156.................114
17. Plasmid Copy Number..........................120
18. Transfer of DNA Residues Using MboII Sites......132
19. Sequence of Synthetic Oligonucleotides.............135
20. Two-Dimensional Sequencing of Oligonucleotide 2.137
21. Identification of the 5' Terminal Residue ......139
22. Ligation of Bam Duplex II with and without a 3' Phosphate..........................143
23. Ligation and Cleavage of **Bam** Duplex III .......... 146
24. Ligation of **MboII** Fragments ......................... 149
25. Construction of pRBG150 and pRBG151 .................. 153
27. Construction and Restriction Mapping of pRBG211 .......... 161
28. Construction of pRBG380 and pRBG381 .................. 164
29. Insertion of DNA Fragments into pRBG380 and pRBG381 .......... 167
30. Construction of a Specific Sequence of DNA Using Restriction Enzymes .......... 171
31. Insertion of Synthetic DNA Sequences into **MboII** Cleavage Sites .......... 175
32. Cleavage Patterns of Different Forms of the **lac o-MboII** Cassette .......... 183
33. Use of the **lac o-MboII** Cassette for the Creation of Specific Deletions .......... 185
34. Construction of pRBG1 and pRBG10 ...................... 188
35. Restriction Mapping of pRBG10 ............................ 191
36. Time Course of DNAase I Digestion of pBR322 ............. 194
37. Restriction Mapping of **lac o-MboII** Cassette Inserts into pBR322 .......... 197
List of Tables

Page

1. Cleavage Sequences of Restriction Enzymes........74
2. Strains of E. coli and Their Genotypes..........78
3. Plasmid Stability.................................109
4. Plasmid Incompatibility..........................117
5. Plasmid Copy Number...............................122
GLOSSARY

copA - regulates replication of R1.
cryptic plasmid - a plasmid whose phenotype is unknown.
copB - regulates replication of R1.
Dam- - mutation which prevents methylation of DNA sequences downstream - in the 3' direction from a point.
dnaA - region whose protein product is required for replication.
dnaC - facilitates binding of replication proteins.
dnaE - same as polC.
dnaG - a rifampicin resistant RNA polymerase.
gyrB - beta subunit of DNA gyrase.
inc - region required for incompatibility properties.
mfaA - genes required for maintenance of F.
Mob- - mobilization deficient plasmids.
oriC - chromosomal origin of replication.
oriT - location of the specific nick on ColE1.
par - region involved in partitioning of the plasmid.
polA - DNA polymerase I.
polB - DNA polymerase II.
polC - subunit of DNA polymerase III.
repA - required for replication of R1.
tra - genes required for conjugation.
upstream - in the 5' direction from a point.
Abbreviations

ACH - acid casein hydrolysate
Ap - ampicillin
bp - base pairs
DHFR - dihydrofolate reductase
lac o - lactose operator
MOPS - morpholinepropanesulfonic acid
NTP - nucleotide triphosphate
Tp - trimethoprim
Tris-HCl - Tris(hydroxymethyl)aminomethane
trpPO - tryptophan operator-promoter
X-gal - 5-bromo-4-chloro-3-indolyl-beta-D-galactoside
Chapter 1. Plasmid Replication and DNA Manipulation

INTRODUCTION

The manipulation of DNA is an important feature of molecular biology. Cloning of DNA sequences permits application of a wide range of techniques which illuminate cellular processes. Experimental procedures used in the manipulation of DNA sequences include isolation of specific DNA sequences, ligation of DNA fragments, and insertion of the DNA sequences into a cloning vector. There are many procedures for DNA isolation which have been reviewed in Bolivar and Backman (1979). If the DNA sequence of interest is not easily isolated, oligonucleotide synthesis has been useful in generating the desired sequence (reviewed in Ohtsuka et al., 1982; Caruthers et al., 1983). This thesis describes a useful new procedure for manipulating DNA fragments. It is based upon the unique digestion properties of the restriction enzyme, MboII. This enzyme recognizes a specific pentanucleotide sequence and cleaves double-stranded DNA eight bases downstream, in the 3' direction, from this sequence (Brown et al., 1980). The separation of the recognition site for MboII from the cleavage site is exploited in the work described in this thesis.

In the course of the thesis work, several unusual plasmids were constructed. Plasmids are extremely useful as cloning vectors, and their replication properties have
been highly studied (reviewed in Veltkamp and Stuitje, 1981; Nordstrom, 1983). A plasmid which will be useful for DNA manipulations must contain a region which will support replication of the plasmid, along with a gene sequence which will allow the plasmid to be selected. Some of the plasmids described in this thesis are lacking a portion of the replication region which had been previously shown to be needed for stable replication. The replication properties of these constructed plasmids have been examined.

In this introduction I will discuss the current literature concerning plasmid replication. This will serve as a background for an examination of the control of plasmid replication. Similarly, the use of plasmids for DNA manipulation, along with some of the techniques used for recombinant DNA procedures, will then be discussed.

Plasmid Discovery and Characterization

Plasmids are semi-autonomous segments of DNA, which are capable of stable replication (Sherratt, 1974). While plasmids are found in eukaryotes, the plasmids found in prokaryotes have received the most attention. Bacterial plasmids often confer properties upon the host cell which give it a competitive advantage over cell without a plasmid. Some phenotypic properties conferred by bacterial plasmids are: gene transfer by conjugation, drug resistance, DNA restriction and modification systems, production of bacteriocins (substances which are toxic to other bacteria) and
production of antibiotics. Because of the extensive research on bacteria, much is known of the plasmids they harbor.

The plasmids found in *Escherichia coli* generally provide for three traits: conjugation, drug resistance, and bacteriocin production. Conjugation is a process which transfers the plasmid DNA from a donor cell to a recipient cell which does not contain the plasmid. The bacteriocins produced in *E. coli* strains are called colicins. These three categories are not mutually exclusive; some plasmids may exhibit more than one of these phenotypes. The plasmid is usually characterized by the first trait observed and not by any objective criteria. If it does not appear to produce any known phenotype, the plasmid is termed "cryptic".

The main feature of the conjugative plasmids is their ability to promote transfer of DNA, usually plasmid DNA, through cell-cell interactions. All conjugative systems have a series of genes which provide 3 features: presence of an extracellular organelle called a pilus which is essential for recognition and pair formation with a recipient cell; a system to provide for replication and transfer of plasmid DNA; and surface exclusion proteins to prevent the cell from being perceived as a recipient cell by another bacteria carrying the same plasmid. The resistance plasmids carry several genes which confer antibiotic resistance or other protection to hazardous chemicals in the environ-
ment. Bacteriogenic plasmids produce proteins which are toxic to bacteria which do not contain the plasmid, and produce proteins which give immunity to the cells carrying the plasmid. These latter two types of plasmids may or may not also promote conjugation.

The most studied of the conjugative plasmids is the sex factor, F (Figure 1). This plasmid is approximately 100 kilobases long and is found as a circular DNA in bacteria (Willetts and Skurray, 1980). F itself does not determine any phenotypes, but it possesses several insertion sequences which are able to integrate into the E. coli chromosome, producing Hfr strains. The Hfr strains are very useful in mapping the E. coli genome (Backman et al., 1976). During conjugation, the F plasmid DNA is passed to the donor cell, carrying along the host chromosome. By disrupting conjugation at specific times, researchers are able to look for the appearance of the desired phenotype in the recipient cell. This gives the "minutes" that gene is located from the site of F insertion. The integrated plasmid can also be excised from the chromosome, carrying along some host genes, resulting in an F' plasmid.

Resistance plasmids, or "R factors" were first found by Japanese investigators in the mid-1950s (Watanabe, 1963). These plasmids are usually capable of conjugation, often across species boundaries, which provides for their rapid dissemination throughout the environment. Many of the resistance plasmids have similar DNA regions which determine
Figure 1. Functional Maps of the Sex Factor F, the Resistance Factor R100 and the Bacteriocinogenic Plasmids, ColE1 and CloDF13. The size of the plasmids is given in base pairs. tra- genes needed for conjugation. The mobility regions of ColE1 and CloDF13 allow transfer if a conjugation plasmid is present in the same cell. F and R100 are from Willetts and Skurray (1981). ColE1 and CloDF13 are from Veltkamp and Stuitje (1981).
incompatibility and replication properties (Figure 1). R100 is a typical R plasmid, bearing resistance to mercuric ions, sulfonamides, streptomycin, chloramphenicol, fusidic acids, and tetracycline (Willetts and Skurray, 1980).

The bacteriocinogenic plasmids produce proteins which are toxic to bacteria. They also provide immunity to that protein for the bacteria harboring them. Colicins are proteins which are harmful to the coliform bacteria and the plasmids which produce them are colinogenic (Figure 1). These "colicin factors" were first discovered in 1925 by Gratia (Predicq, 1957). Gratia found that media in which *E. coli* V had been grown would not allow the growth of other *E. coli* strains. Spotting of this media on agar plates resulted in areas of inhibited growth of *E. coli* (Frick et al., 1981).

Besides production of toxic proteins, the colicinogenic plasmids often harbor other characteristics. Some of the plasmids are associated with virulence in *E. coli*. The ColV plasmid gives the bacteria increased resistance to the immune response of higher eukaryotes. Also the plasmid gives the host cell the ability to survive in an iron depleted environment (Williams and Warner, 1981). Iron sequestering is a mammalian response to bacterial infection, so the ColV plasmid would give the bacteria a survival advantage. Some of the proteins coded for in the plasmid also help with protection from antibodies (Timmis et al., 1981).
For many years the "colicin factors" were thought of as a type of lysogenic bacteriophage (Fredricq, 1957). Both bacteriophage and plasmids are very stably maintained. The original *E. coli* strain isolated in 1925 still contains the ColV plasmid (Fredricq, 1957). Colicinogenic plasmids and bacteriophage kill susceptible cells while providing immunity to the host cell. Damage to the plasmid-containing cell, such as ultraviolet irradiation or incubation with mutagens, results in induction and production of a substance which is toxic to susceptible bacteria. The factors could be transferred from one cell to another as could bacteriophage (Fredricq, 1957). This initial confusion was removed in the early 1960s, when the colicin factors were shown to be plasmids.

Some of the original confusion arose from the conjugative properties of the different colicinogenic plasmids. One group, such as ColV, carries all the genes necessary for conjugation, while others, such as ColE1, cannot fully provide for their own conjugation. They can be transferred, or mobilized, only if a conjugative plasmid is present in the same cell (Kahn and Helinski, 1964). Fredricq (1963) showed that some of the colicin factors were associated with F factors which were then known to be plasmids. In strains with both F factors and colicin factors, the F plasmid could be eliminated without loss of the colicin factor. ColE1 was shown to be nonchromosomal in an elegant experiment by Clowes (1963). The ColE1 plasmid was placed into five dif-
ferent Hfr strains which had the F factor integrated at
different positions on the *E. coli* chromosome. Conju-
gation was performed to a suitable recipient cell, and the
time at which the colicin factor began entering the cell
was found. It was observed that the transfer of the ColE1
factor started at 7-12 minutes. Since there was no single
gene at that distance in all five Hfrs, ColE1 could not be
on the chromosome. Purely genetic studies involving bac-
terial crosses of the ColE1 factor demonstrated that the
plasmid was not linked with any chromosomal genes (Nagel de

Mitomycin C had been known to induce production of
colicins and presumably the colicin factor (Fredricq, 1957).
In a ColE1-containing strain of *Proteus mirabilis*, which
had been induced with mitomycin C, a satellite band of DNA
was found which was not present in strains without the fac-
tor (DeWitt and Helinski, 1965). The DNA was further char-
acterized and found to be a circular molecule of double-
 stranded DNA which behaved as if it was supercoiled (Roth
and Helinski, 1967; Bazaral and Helinski, 1968). This
demonstrated the plasmid nature of the ColE1 factor.

**GENETIC ORGANIZATION**

**Conjugation and Mobilization**

Conjugation is a complicated process which is still not
fully understood. So far 19 genes (tra) which are re-
quired for proper conjugation have been mapped on the plas-
mids. These tra genes fall into 4 groups: pilus formation; stabilization of mating pairs; conjugal DNA metabolism; control of the other genes (Willetts and Skurray, 1980). Some of the genes, such as that for the surface exclusion protein, also prevent the bacteria from being attacked by antibodies or complement, allowing the bacteria to survive in serum. Elucidation of the exact function of each gene has been difficult. Mutants have been used to examine structural changes in pilus formation and mating pair formation. Attempts to clone the various genes have been hampered by the lack of many hexanucleotide restriction enzyme sites in the tra region (Willetts and Johnson, 1979). This seems natural considering the vast host range of these plasmids and the need for them to protect this region from restriction systems in many strains.

Work has shown that the 12 genes involved in pilus formation are usually interchangeable between similar plasmids (Willetts and Skurray, 1980). Many of the other genes involved in conjugation are complementable, especially between plasmids of the same incompatibility group. Of particular importance are the genes involved in conjugal DNA metabolism. These are the genes which allow the DNA to be transferred and properly replicated. The steps involved in conjugal DNA metabolism are complicated and still not entirely known. Several of the genes present in the tra region are necessary as are several bacterial gene products. Some of the plasmid-coded proteins are involved in
inducing a single-stranded nick at a specific location (oriT) in one strand of the plasmid DNA in the donor cell (Willets and Maule, 1979). This strand is then transferred to the recipient cell through the pilus. After nicking and transfer, host enzymes are used for synthesis of the complementary strands in the donor and recipient cells. Synthesis of the transferred strand is inhibited by rifampicin (rif; a RNA polymerase inhibitor), implying a role for an RNA primer in the synthesis of recipient plasmid DNA (Kingsman and Willett, 1978). The plasmid ColIb produces a "primase" which is transferred to the recipient cell, along with the plasmid DNA (Chatfield et al., 1982). This primase is not needed for vegetative replication and could be partially complemented by the dnaG gene product (a rif resistant RNA polymerase; Bouche et al., 1975). Conjugal DNA synthesis in a dnaB mutant was deficient in the absence of the plasmid primase. F does not code for its own primase but the dnaG primase may be utilized instead for conjugal DNA metabolism (Willetts and Skurray, 1980).

The role for other proteins involved in the transfer of plasmid DNA is not fully known. All the tra gene products must be present in order for DNA transfer to occur. DNA transfer can occur even if DNA synthesis in the donor cells (Kingsman and Willets, 1978) or in the recipient cells (Wilkins and Hollan, 1974) is inhibited. Inhibition of DNA gyrase with nalidixic acid prevents plasmid DNA from being transferred (Hane, 1971).
Most of the small, nonconjugative plasmids can be transferred to recipient cells if an F or F-like plasmid is present along with it. That is, the small plasmids do not contain the genes necessary for pilus formation, but are able to provide the needed structures for transfer. These plasmids are said to be mobilizable. ColE1 can be mobilized by many different kinds of conjugative plasmids. It was this property which was used in early work to show the plasmid nature of ColE1 (Kahn and Helinski, 1964; Fredricq, 1963; Clowes, 1963; Nagel de Zwaig and Puig, 1964). Gentle lysis of ColE1-bearing cells results in the isolation of a complex of several proteins covalently bound to a nick on the DNA (Clewell and Helinski, 1970; Blair et al., 1971). The production of this relaxation complex appears to be affected by the amount of glucose present. In the absence of glucose, 80% of the isolated plasmid DNA can be found in a complexed state, while with glucose present, only 30% is complexed (Clewell and Helinski, 1972). This implies a role for cAMP in the synthesis of complexed ColE1.

ColE1 mutants that do not form relaxation complexes are not mobilizable (Mob-)(Dougan et al., 1978). Using F tra mutants, it has been shown that ColE1 mobilization is independent of the F gene products which are involved in the nicking and recircularization reactions in conjugations (Willetts and Maule, 1979). Mob- mutants of ColE1 can be complemented by other similar plasmids indicating that gene products acting in trans are produced (Warren and Sher-
ratt, 1977). The region required for mobilization has been
determined although the protein products of this region
have not been isolated. Insertion of the mob genes into a
Mob- plasmid allow the new plasmid to be mobilized (Warren
et al., 1979). Mobilization of ColE1 requires all of
the tra genes except those involved in conjugal DNA metab-
olism. The details of mobilization is an area of intense
study.

Bacteriocinogenity and Immunity

The bacteriocins produced by plasmids have a variety
of effects. Cloacin DF13, a protein produced by a small,
nonconjugative plasmid isolated from Enterobacter
cloacae DF13, inhibits protein synthesis in sensitive
strains by a cleavage of a 49-nucleotide fragment from the
3' end of 16 S ribosomal RNA (DeGraff et al., 1973).
Colicin E1, produced by ColE1, disrupts the energized state
Cells carrying bacteriocinogenic plasmids also manufacture
substances which provide immunity to the toxin. The CloDF13
immunity protein inhibits the biological action of the
cloacin protein. Although the colicin E1 immunity protein
has not been purified, the regions on ColE1 required for
both colicin production and immunity have been located
(Veltkamp and Stuitje, 1981).

Normally, expression of the colicin genes is repressed
and bacteriocin production is only found in a small frac-
tion of the population. Agents which damage DNA metabolism, such as UV or mutagens, induce a large proportion of the cells to produce the toxin (Durkacz et al., 1974). It has been suggested that expression of the colicin gene is repressed by a protein sensitive to SOS repair functions (Nakazawa and Suzuki, 1977). Several bacteriocin genes have a \textit{lexA} binding site present which has been shown to control several SOS repair functions (van den Elzen et al., 1983). Induction of colicin production is a lethal event for a bacteriogenic cell (Ozeki et al., 1959). This lethality is not, however, due to the production of active colicin (Shaferman et al., 1979). Mini-plasmids of CloDF13 which contain only the origin of replication are still sensitive to mitomycin C (Veltkamp and Stuitje, 1981). Induction of ColE1 with mitomycin C results in a large increase in the amount of ColE1 plasmid DNA present in the cell (DeWitt and Helinski, 1965). Although a protein role in the lethality of induced cells has not been ruled out, it appears that the killing phenomenon may be a result of increased plasmid replication. The exact nature of this interaction remains to be elucidated.

\textbf{PLASMID REPLICATION}

\textbf{Incompatibility}

Two properties of plasmids, incompatibility and copy number, are often closely linked with the replication apparatus of a plasmid. This makes sense as they all are
involved with the ability of a plasmid to reproduce itself. Incompatibility is defined as the inability of two different plasmids to be stably maintained in the same bacterial cell without selective pressure (Staudenbauer, 1978). If two plasmids share similar replication regions, they are usually incompatible with one another. Often a lack of homology between plasmids is taken as an indication that the two plasmids will be compatible. Attempts to force two incompatible plasmids to remain in a cell, through the use of double drug selection, often results in a hybrid molecule (Bedbrook and Ausubel, 1976). A plasmid constructed by \textit{in vitro} ligation of two incompatible plasmids, pSC101 and ColE1, was itself incompatible with both of the parent plasmids (Cabello et al., 1976). Incompatibility functions are often carried on several places on a plasmid. In plasmids which use the \textit{E. coli} chromosomal origin of replication, incompatibility has been found to be determined by sequences near the origin. These plasmids have at least 3 different incompatibility determinants, of which at least one is a protein (Stuitje and Meijer, 1983). In RK2, a broad host range plasmid, one incompatibility (\textit{inc}) function is found near the origin. This function needs an active RK2 replicon, while another \textit{inc} region was discovered which did not need a RK2 replication origin present for incompatibility (Meyer and Hinds, 1982). In several plasmids there is evidence that a non-translated RNA is involved in the incompatibility characteristics. Moser and
Campbell (1983) described the insertion into pACYC184 of a small DNA fragment from pBR322 which produces a RNA that has been implicated in incompatibility. pACYC184 is usually stably replicated in the same cell with pBR322. With the addition of the DNA sequence pACYC184 is no longer compatible with pBR322 (Moser and Campbell, 1983).

Modes of replication and segregation

There are essentially two different modes of plasmid replication, stringent and relaxed. The actual definition of stringent and relaxed plasmids arises from the relationship between plasmid replication and protein synthesis. In stringent replication, plasmid replication is tied to chromosomal replication. Usually, plasmids which are stringent occur in very low copy number. F is an example of a stringent plasmid. In gently lysed cells, the F plasmid is found associated with the folded chromosome, apparently by a RNA link (Lane, 1981). This complex will dissociate upon treatment with rifampicin in vivo or with RNAase A in vitro. The mafA gene (for maintainence of F) on the E. coli chromosome is important in the stable replication of F. MafA mutants will receive and recircularize F plasmid DNA but the plasmid will not replicate (Lane, 1981). The mafA gene probably affects initiation of replication and seems to require RNA synthesis (Wada and Yura, 1982). Several plasmid-coded genes are also necessary for plasmid maintenance (Watson et al., 1982). Inser-
tion of Tn5 into certain regions of F results in a plasmid which cannot replicate in polA mutants, while normal F plasmids do not require either polA or polB (Lane, 1981). Several mini-F plasmids have been constructed which exhibit all the necessary replication properties of the entire F plasmid (Watson et al., 1982; Wada and Yura, 1982).

Two other highly studied stringent plasmids are pSC101 and the mini oriC. Mini oriC was constructed from the E. coli chromosomal origin region (Stuitje and Meijer, 1983) while pSC101 was isolated by Cohen et al. (1972). The oriC plasmids do not exhibit the same replication behavior as the E. coli chromosome. These plasmids seem to occur in 3-10 copies per cell and additional copies have no effect on chromosomal replication (that is, they display no inc). This implies that most of these plasmids do not contain all the needed sequences for chromosomal-like replication. Stuitje and Meijer (1983) have described a mini oriC plasmid which displays more chromosomal attributes. They localized 3 regions, including one which codes for a protein, that determine inc function. One of the other regions is required for bidirectional replication and also binds a membrane protein.

Although the details are not completely known, pSC101 replication exhibits behavior that has features similar to F or oriC plasmids. Both F and pSC101 have an AT rich region of about 80 bases adjacent to several direct repeats
(Churchward et al., 1983). Both pSC101 and oriC contain several 11 base pair (bp) sequences which are also repeated near the dnaA gene promoter, a product which both plasmids require for replication (Felton and Wright, 1979). Although the exact function of the dnaA protein is not known, purified dnaA protein will bind specifically to these sequences in pSC101 and oriC (Fuller and Kornberg, 1983). pSC101 has a 270 bp region next to the origin which is required for stable partitioning of the plasmid (par; Meacock and Cohen, 1981). This sequence is required in cis (i.e. it must be present on the plasmid which is being partitioned) and can rescue an unstable plasmid which does not contain the sequence if it is placed on the plasmid. F also has a par region which determines the segregation properties of the plasmid. This region codes for two apparent proteins which are both required for stable partitioning of the daughter plasmids (Austin and Wierzbicki, 1983). The similarities between the replication region of stringent plasmids implies a common mechanism to insure proper replication and segregation.

Plasmids which exhibit a relaxed mode of replication are often present in the cell in high copy numbers. These plasmids do not appear to contain an active par function, in contrast to the stringent plasmids. While these plasmids seem to be randomly partitioned, there is evidence for a definite segregation function. Many stringent plasmids are associated with the bacterial membrane and folded chromo-
some (Lane, 1981). Relaxed plasmids are found bound to the membrane, complexed with the folded chromosome and in a free state (Drygin et al., 1971; Kline et al., 1976).

Most of the nonreplicating ColE1 could be found in the cytoplasm while the membrane fraction was enriched for newly replicated molecules (Sherratt and Helinski, 1973). Early work with ColE1 had shown that the replication of any one molecule was random and did not depend on previous replication (Bazaral and Helinski, 1970). There are several hypotheses to explain segregation of relaxed plasmids. The random replication hypothesis assumes the template molecule is randomly replicated and inherited. The equipartition hypothesis states that all copies are randomly sorted into two equal populations and then partitioned. These two proposals are based upon the fact that no plasmidless cells were found for several generations in temperature-sensitive strains after the temperature shift. Using finer techniques, Hashimoto-Gotoh and Ishii (1982) found that plasmidless cells do develop immediately after the temperature shift. They propose a single-site model. One molecule is employed as a template for replication at any time. It is bound to the membrane and all other plasmids are inhibited from replicating. After replication, another molecule is chosen as a template at random. The two products of the last replication are actively transmitted to the daughter cells by the membrane. The other molecules are randomly segregated.
Stringent plasmids show a strict relationship between plasmid replication and protein synthesis. Several proteins have been shown to be required for stable maintenance of F (Watson, et al., 1982; Lane, 1981). Plasmid R1, another stringent plasmid, also produces a protein product which is required for replication (Light and Molin, 1982). The relaxed plasmids, such as ColE1, are able to replicate in the absence of protein synthesis. During amino acid starvation or during inhibition of protein synthesis with chloramphenicol, ColE1 can continue replicating, increasing in number several hundred fold (Clewell, 1972; Clewell and Helinski, 1972). A plasmid was constructed which had both the ColE1 and pSC101 plasmid origins (Timmis et al., 1974). pSC101 does not require DNA polymerase I but does require protein synthesis while ColE1 needs DNA polymerase I but not protein synthesis. In the absence of protein synthesis, the plasmid showed a high copy number while in the presence of a polA mutation, copy number was low. Another plasmid was constructed by fusing F with ColE1 (Tsutsui and Matsubara, 1981). In a temperature-sensitive polA mutant, the ColE1 replication origin was used at the permissive temperature. At the nonpermissive temperature, plasmid replication was shut down until the plasmid copy number was reduced to 1-2 copies per cell. Replication then started again using the normal F origin. These experiments demonstrate the difference between the two modes of replication.
Host Enzymes Needed for Replication of Relaxed Plasmids

The replication of relaxed plasmids has received the most study, as plasmid replication can be separated from host replication and tends to be somewhat simpler in its control processes. In ColE1 DNA replication, several host proteins are required. This was investigated using temperature-sensitive mutants for the replication enzymes in question. Collins et al. (1975) looked at the replication properties of ColE1 in several mutant strains of E. coli. These strains had temperature-sensitive mutations in dnaA (DNA binding protein necessary for chromosomal replication), dnaC (facilitates binding of dnaB to single-stranded DNA), dnaG (rifampicin-resistant RNA polymerase) and dnaE (same as polC; subunit of DNA polymerase III holoenzyme). ColE1 replication was found to be independent of dnaA but required the other proteins. More direct evidence of the bacterial proteins required for ColE1 replication came from in vitro replication extracts. Tomizawa et al. (1975) developed a cell extract which could easily replicate exogenous or endogenous ColE1 plasmid DNA. Extracts from dnaB and dnaC mutants would not replicate ColE1 plasmid DNA (Staudenbauer et al., 1978). These could be complemented by the addition of wild type dnaB or dnaC protein. Extracts from polA (DNA polymerase I) temperature-sensitive mutants did not allow replication (Staudenbauer, 1976). ColE1 could
not be transferred stably by conjugation into a polA mutant although other plasmids could (Kingsbury and Helinski, 1970). ColE1 was not maintained in a polA mutant which still retained the 5' to 3' exonuclease activity of DNA polymerase I (Itoh and Tomizawa, 1978). Purified DNA polymerase I could promote replication of ColE1 in extracts taken from this mutant, however, the large fragment of DNA polymerase I (Klenow fragment), which lacks the 5' to 3' exonuclease activity, could not complement the mutant (Itoh and Tomizawa, 1978). Extracts from polB (DNA polymerase II) mutants had no effect on plasmid replication in vitro, while extracts from either polA or polC (subunit of DNA polymerase III) mutants allowed no replication of ColE1 (Staudenbauer, 1976). It seems then that dnaB, dnaC, and dnaG, along with DNA polymerase I and III, are required for plasmid replication.

Several other enzymes have been shown to be required for ColE1 plasmid replication. Inhibition of RNA polymerase in vivo (Clewell et al., 1972) or in vitro (Staudenbauer, 1976; Sakakibara and Tomizawa, 1974) by rifampicin halts ColE1 plasmid replication. However, once replication had been initiated, no inhibition with rifampicin was found (Sakakibara and Tomizawa, 1974). The inhibition of DNA gyrase (responsible for placing negative supercoils in DNA) with novobiocin stops ColE1 DNA replication (Gellert et al., 1976). Mutants which are temperature sensitive for the beta subunit of DNA gyrase lose plasmids very
rapidly at the nonpermissive temperature. Cell-free extracts from a gyrb mutant were deficient in promoting replication of exogenous ColE1 DNA (Orr and Staudenbauer, 1981). This could be complemented with the addition of purified DNA gyrase. ColE1 plasmid DNA can be kept in a supercoiled state with purified DNA gyrase (Gellert et al., 1976). Addition of novobiocin to this mixture results in the relaxation of ColE1 DNA. A DNA gyrase prepared from a novobiocin-resistant strain, added to a mixture of ColE1 plasmid DNA and novobiocin, keeps the DNA supercoiled. A cell-free extract made from the novobiocin-resistant mutant allowed ColE1 plasmid DNA to replicate even in the presence of novobiocin (Gellert et al., 1976). Replication of ColE1 plasmid DNA requires supercoiled DNA (the product of DNA gyrase), rather than DNA gyrase itself (Gellert et al., 1976).

Another enzyme required for ColE1 plasmid DNA replication is RNAase H. This enzyme recognizes and cleaves a DNA-RNA heteroduplex. Hillenbrand and Staudenbauer (1982) isolated two fractions which were required for proper initiation of ColE1 plasmid DNA replication. The "priming fraction" contained DNA polymerase I, RNA polymerase, and DNA gyrase. The "discriminating fraction" contained RNAase H. The two fractions provided origin-specific replication of several amplifiable plasmids. Without the discriminating fraction, initiation of DNA synthesis occurred at random locations. DNA synthesis can be initiated on supercoiled
ColE1 plasmid DNA using purified DNA polymerase I, RNA polymerase, and RNAase H (Tomizawa and Itoh, 1982). The activities of these enzymes in plasmid DNA replication will be discussed below.

Replication of ColE1 does not require any proteins coded by plasmid. Incubation of supercoiled ColE1 plasmid DNA with cell extracts from strains that do not contain ColE1 still results in replication of the plasmid DNA (Tomizawa et al., 1975; Staudenbauer, 1976). In order to eliminate any possibility of plasmid-encoded proteins which are required in ColE1 replication, bacteriophage-ColE1 hybrids were constructed (Donoghue and Sharp, 1978). When these hybrids were allowed to infect chloramphenicol-treated cells (chloramphenicol inhibits protein synthesis), it was found that replication occurred only at the ColE1 replicon. Thus ColE1 replication can take place without the production of any plasmid proteins. This also indicates the absence of a plasmid-encoded protein that acts positively in the initiation of plasmid replication. Taken along with the fact that inhibition of protein synthesis results in an elevation of plasmid DNA replication (Clewell, 1972), these results best fit a model in which the control mechanism is negative.

There are two distinct steps in ColE1 plasmid DNA replication: initiation and elongation. Initiation requires RNA polymerase, DNA polymerase I, and RNAase H (Tomizawa and Itoh, 1982). In vivo (Inselburg, 1974) and in vitro
(Tomizawa et al., 1974) this results in the formation of the early replicative intermediate. This molecule is seen in the electron microscope as a double-stranded molecule with a "bubble" or D-loop present. Tomizawa (1975) looked at the properties of this intermediate in vitro. Synthesis began on one strand (H strand) and this required RNA polymerase, along with normal amounts of DNA polymerase I (Staudenbauer, 1978). Most of the DNA synthesized in this step is L-strand DNA (i.e. it is synthesized off of the H strand). If exogenous early intermediates are added to extracts containing rifampicin, elongation will still take place (Tomizawa, 1975). This later synthesis is by a discontinuous replication mechanism (Staudenbauer, 1974). This model proposes that RNA polymerase synthesizes an RNA primer which is extended by DNA polymerase I to form the early replicative intermediate (Staudenbauer, 1978). Further replication on both strands requires DNA polymerase III along with the other proteins needed for discontinuous DNA synthesis. In this stage, DNA polymerase I would only be used to remove RNA primers left in the plasmid DNA.

ColEl has been found to replicate by the Cairns mode in minicells, both in the absence and presence of chloramphenicol (Inselburg and Oka, 1975; Bolivar et al., 1977). Several investigations have shown that ColEl DNA replication begins at a unique site on the plasmid. ColEl plasmid DNA was found to contain a single EcoRI restriction enzyme site. Examination under the electron micro-
scope of ColE1 plasmids replicated in vitro and which had been cleaved with EcoRI showed the D-loops of the early intermediates, along with longer, double-stranded sequences (Tomizawa et al., 1974). The length from one end of the loop to the end of the molecule was always the same, whereas the other end of the D-loop changed. This observation was also made in vivo (Inselburg, 1974). This indicates that DNA replication of ColE1 begins at a unique site (the origin) and proceeds unidirectionally. Termination of replication takes place at a point where the replication fork meets the point of initiation, which is at or near the origin (Lovett et al., 1974).

DNA Regions Required for Plasmid DNA Replication

Attempts have been made to determine the minimum region required for autonomous plasmid replication by cloning fragments of a plasmid and looking for replication. This has been done for chromosomal origin, oriC (Stuitje and Meijer, 1983), RK2 (Figurski et al., 1982), F (Wada and Yura, 1982), and ColE1 (Hershfield et al., 1976). The oriC plasmids, while stable, do not behave in the same manner as E. coli chromosomes. They appear to replicate throughout the cell cycle and occur in varying copy numbers (Leonard et al., 1982). Apparently, not all the control regions for chromosomal replication are present near the chromosomal origin (Stuitje and Meijer, 1983). The mini-F plasmids do exhibit all the replication characteristics of
the entire F plasmid. These plasmids occur in a single copy per chromosome, are effectively partitioned and are affected by mafA mutants (Wada and Yura, 1982). The replication region of pSC101 (Cohen and Chang, 1973) has been explored by using the insertion of a transposable sequence into regions of the plasmid (Churchward et al., 1983). The area around the origin, about 1100 bp, contains all the functions needed for replication and partition of pSC101. Several repeated sequences were found which, when disrupted by the transposon, result in loss of stable replication. These repeats were found next to an AT rich region. Insertions here also destroyed stable replication. These regions are similar to those found in other replicons.

The most studied plasmid replicon is that of ColE1 and its derivatives. In vitro manipulations have shown that the region surrounding the origin of replication is the only DNA sequence required for replication. This entire region has been sequenced as have those of several similar plasmids (Ohmori and Tomizawa, 1979; Selzer et al., 1983). Oka et al. (1979) constructed a series of small ColE1 plasmids from DNA fragments of ColE1 and a DNA fragment carrying a gene conferring resistance to ampicillin. They were able to isolate a 436 bp fragment which was capable of replication, although not as stably as some larger fragments. This small fragment extends from 16 bp downstream from the origin to 420 bp upstream. Another region just adjacent to this fragment was also required for fully
stable replication (Oka et al., 1979). A 580 bp fragment from pMB1 (a plasmid closely related to ColE1), which extends further upstream from the origin, has been shown to carry the region needed for replication (Backman et al., 1979). These results suggest that the region upstream from the origin of replication is the most important for stable replication.

The general area of this region where the transition from the RNA primer to DNA takes place was located by Williams et al., (1982 a,b). In chloramphenicol-amplified cells, the RNA used in the primer is not removed. Apparently a repair property which normally removes incorporated RNA does not function. The RNA in one strand appeared to resemble a sequence of DNA near the origin. The exact location where initiation of DNA replication begins has been shown by in vivo (Bolivar et al., 1977a) and in vitro (Tomizawa et al., 1977) experiments to start within a string of 5 adenines. This region is highly conserved in several plasmids, indicating the importance of this area.

Control of Replication

Replication of plasmid DNA must be very carefully controlled. The plasmids may not be passed onto the daughter cells at too low a level of replication. At too high a level of replication the host cell may be killed. The elucidation of the control mechanisms of the different
plasmids has been slowly forthcoming. Control of the single copy plasmids such as F and pSC101 has been shown to be mediated by at least one protein. They both carry a partitioning region which is required for stable transmission of the plasmid to daughter cells. The manner by which the partition proteins function is unknown. In the case of F, a region on the E. coli chromosome is also needed.

The examination of replication control in some of the multiple copy plasmids has been more fruitful. These control mechanisms often contain both protein interactions and RNA-RNA interactions. An example is the plasmid R1 (Figure 2). This plasmid codes for a protein, repA, which is required for initiation of plasmid replication at the origin. This gene is adjacent to the origin and is transcribed towards the origin from a promoter several hundred bases upstream (Light and Molin, 1982). The transcription of the repA protein is repressed by another protein, copB. The gene for this protein is found adjacent to the repA gene and the copB protein acts as a repressor in a classical sense, binding to an operator near the repA promoter. In addition to this control mechanism, there is another which is present. There is a small 90 nucleotide RNA transcribed from the leader region of repA, called copA. This RNA is transcribed off of the opposite strand and in an opposite direction from repA transcription. This RNA has been shown to determine the incompatibility characteristics of R1 (Stougaard et al., 1981). This RNA acts upon the
Basic replicon of R1

CopA

CopB RepA ori
region downstream from the promoter of repA, in the same area of the leader where the RNA is transcribed. This RNA can form 2 stem-and-loop structures which seem to determine the characteristics of the RNA-RNA interactions. The action of this RNA inhibits expression of the repA protein. This use of both protein and RNA interactions to control aspects of plasmid replication has also been seen in ColE1 and its derivatives.

Control of Replication in ColE1

With the construction of mini-ColE1 derivatives work could progress on the control mechanisms involved. Since plasmid-encoded proteins were not thought to be involved in ColE1 plasmid DNA replication, other plasmid products were examined. Using in vitro replication systems, a small RNA (RNA I) was isolated which was the same size in all the derivatives examined (Figure 3; Morita and Oka, 1979). RNA I starts its transcription 445 bp upstream from the origin and proceeds away from the origin for 105 bases before terminating. This RNA was sequenced and found to be highly self-complementery, as demonstrated by the selective sensitivity of the molecule to T1 RNAase. Based on the partial T1 digest, Morita and Oka (1979) proposed a secondary structure for the molecule which had three stem-and-loop structures. There were no translation initiation codons in the RNA and termination codons occurred frequently, suggesting that the structure of the RNA is important, not
Figure 3. RNA Transcripts Produced by the ColE1 Replicon. 

**ini** - site of initiation of DNA replication. The primer RNA is transcribed towards the origin where it can form hybrids with the DNA. These heteroduplexes are required for stable replication of the plasmid. RNA I, which is transcribed off of the opposite strand from the primer RNA, acts to repress formation of the heteroduplexes.
protein products that it could produce.

Another RNA transcript has been shown to originate from the region upstream of the origin. Itoh and Tomizawa (1980) found that this RNA (the primer transcript; RNA II) was coded for starting 555 bp upstream from the origin (Figure 3). This primer transcript was transcribed towards the origin and the transcripts often extended past the origin. These transcripts formed hybrids with the DNA giving a substrate which RNAase H could recognize and cleave (Itoh and Tomizawa, 1980). Presumably the small RNA fragments released could act as primers for elongation by DNA polymerase I.

RNA I has been shown to be involved in the regulation of ColE1 plasmid DNA replication, along with maintenance of copy number and incompatibility. Addition of increasing amounts of purified RNA I to in vitro extracts results in decreasing amounts of plasmid replication (Tomizawa et al., 1981). It was found that in the presence of added RNA I, a larger number of primer transcripts extended past the origin but did not form a hybrid molecule with the template DNA. Primer formation was not affected by RNA I if a hybrid had already formed. A specific RNA I from a plasmid of another incompatibility group than ColE1 had no effect on plasmid DNA replication. Work has shown that the action of RNA I did not affect initiation and propagation of the primer transcript (Tomizawa et al., 1981). The total amount of RNA produced was not changed. Tomizawa and
Itoh (1982) demonstrated that the interaction of RNA I with the primer transcript caused transcriptional pausing. Therefore, the effect of RNA I is to inhibit the formation of the primer, presumably through interactions with RNA II.

Transcription of the primer molecule begins near a position where the transcript for RNA I ends. Since the two transcripts are initiated on different strands, this means that all of RNA I is complementary to the 5' end of the primer molecule. By both enzymatic methods (Morita and Oka, 1979) and by computer methods (Selzer et al., 1983) RNA I, and consequently the primer RNA, appears to form several stem-and-loop structures (Figure 4). Transcriptional control by stem-and-loop structures has been implicated in several systems (Kolter and Yanofsky, 1982). Secondary structure has been shown to be important in the RNA-RNA interactions which occur between RNA I and the primer RNA. Incorporation of inosine residues in the place of guanosine results in decreased stability of RNA secondary structures. In vitro experiments using inosine incorporation resulted in inhibition of the formation of an RNA primer at the origin (Tomizawa and Itoh, 1982). Substitution of inosine into RNA II 400 nucleotides upstream from the origin still resulted in inhibition of primer formation.

The use of plasmid mutations, either constructed in vivo or in vitro, has been very useful in isolating regions which are used in control of plasmid DNA replica-
Figure 4. DNA Sequences of RNA I and RNA II. The open arrows indicate RNA start sites for RNA I and RNAII. The solid arrow indicates the termination site for RNA I. The numbers indicate distance from the origin. The five palindromic sequences are shown by underlining (I-V). The overlap of one arm of structure IV with all of structure III can be seen.
tion. Most of these involve construction of ColE1 derivatives which then can be examined easily. Insertion of EcoRI linkers into the region which produces RNA I results in a plasmid with an increased copy number (Conrad and Campbell, 1979). Wong et al., (1982) described several ColE1-derived temperature-sensitive mutants. These plasmids carry a β-lactamase gene which was used to identify the mutant plasmids. A higher plasmid copy number results in resistance to higher amounts of ampicillin. At 31 °C, these plasmids exhibit the normal copy number but at 42 °C, the plasmid copy number increases 30-fold. The mutations were found to be due to basepair changes 5' to the start of RNA I (i.e. towards the origin). Since these changes were located between the -35 and -10 regions of the RNA I promoter, it was of interest to determine whether the temperature effect seen was due to changed promoter characteristics of the RNA I promoter or due to sequence changes in RNA II. The promoter of RNA I was cloned into the pK01 system of McKenney et al. (1981). Surprisingly, the mutant RNA I promoters were slightly stronger than the wild type RNA I promoter. No change due to temperature was seen with any of the mutant promoters (Wong et al., 1982). This indicates that the copy number effect of the base pair changes is due to RNA II and may be due to changed secondary structure.

An elegant series of experiments has also shown the importance of RNA I and its secondary structure. A high
copy number derivative of pMB1 was made which had an inser-
tion in the stem III of RNA I (pFH118; Moser and Campbell, 1983). The high copy number was suppressed if a 256 bp
fragment which contained RNA I and its promoter was cloned
in the mutant plasmid. This fragment does not contain RNA
II, implying that the structure of RNA is the determining
factor. The authors found a gene dosage effect of RNA I.
That is, a plasmid with 2 copies of the wild type RNA I
sequence had half the copy number of one with only one RNA
I sequence. If the RNA I sequence was cloned into pACYC184,
which is normally compatible with pFH118, the two plasmids
became incompatible. This demonstrates not only the copy
number control of RNA I but its incompatibility-determining
characteristics.

Base pair changes in the region upstream from the ori-
gin have been shown to affect the replication properties of
the plasmid. The proposed secondary structure of RNA I from
from several similar plasmids has three stem-and-loop struc-
tures (Morita and Oka, 1979; Selzer et al., 1983).
Plasmids with base changes in the stem or loop of one of
these structures have altered copy number and incompatibil-
ity characteristics. Several plasmids have been isolated
which have a single base pair change in the loop structure
and are no longer incompatible with the wild type plasmid
(Tomizawa and Itoh, 1981; Lacatena and Cesareni, 1981). A
single base pair change produces a plasmid which is now
recognized as a different plasmid by the wild type replica-
tion control system and will be stably replicated in the same cell with the original plasmid. Some of these new plasmids also showed lowered copy number relative to the wild type while others had drastic effects on stable plasmid replication (Lacatena and Cesareni, 1981). In the latter cases, the base changes occurred in the stem regions, implicating them in the control of replication. Another mutant plasmid had a single base change in structure I just prior to a string of U's in RNA I which is believed to represent a termination signal (Muesing et al., 1981). This plasmid produces little or no RNA I in vitro. Instead, much larger RNA molecules are found and the copy number is higher than in the original plasmid. These plasmids indicate that the regions corresponding to the stem-and-loop structures are very important in determining not only control of replication but plasmid copy number and incompatibility as well.

Other regions upstream from the origin also have roles in replication. A mutant of ColE1 which was defective in replication had a single base change 160 bp upstream from the origin (Naito and Uchida, 1980). There is an inverted repeat in this region and it was proposed that the base change destabilized the stem-and-loop structure that could form. The exact purpose of this region is still not known. In studying the transcription of RNA II, it was found that addition of RNA I enhanced pausing at a specific position (Tomizawa and Itoh, 1982). A small stem-and-loop structure
(Structure V) could be found at this area (Figure 4). It was suggested that the inhibition effect of RNA I, presumably through RNA-RNA interactions, allowed this structure to form and decreased the amount of primer transcript available for primer formation. In an examination of several similar but incompatible plasmid origin regions, many similarities were found which have been used to propose a mode of action of RNA I (Selzer et al., 1983). All the plasmids produced similar RNA I transcripts and primer transcripts. An analysis of the sequences showed that all could form several similar secondary structures. RNA I, and consequently RNA II, in each plasmid can fold into a configuration with 3 stem-and-loop structures. While the stems are almost the same, the loops have several base changes. This makes sense since these plasmids are compatible with one another and these loops have been shown to have an effect on incompatibility (Lacatena and Cesareni, 1981; Tomizawa and Itoh, 1981). The bases in the loops seem to be important in determining the rate of interaction of RNA I with the primer precursor or with other heterologous plasmids. The interaction of RNA I with the primer transcript inhibits replication, suggesting that it prevents RNA II from folding into a secondary structure needed for replication. Sequence analysis of RNA II revealed that a large stem-and-loop structure (Structure IV) could be formed by the primer transcript (Figure 4; Selzer et al., 1983). One side of the stem overlaps the third
stem-and-loop structure of RNA I and the other side overlaps structure V. Thus, if RNA I bound to the primer precursor, it would prevent the large structure IV from forming. Conversely, the formation of structure IV would prevent RNA I from completely interacting with RNA II. Evidence for this proposal, while not definite, is indicated by the pausing data (Tomizawa and Itoh, 1982). Structure V, which may form when RNA I interacts with RNA II and is located where pausing occurs, directly follows the sequence for the large structure (Structure IV). If this large structure is formed, structure V cannot form. If RNA I interacts with the primer, the large loop will not form and the smaller one, which provokes pausing, will. This could have some effect upon the ability of the primer transcript to form a hybrid with template DNA. If structure IV does form, structures III and V will not, allowing the primer transcript to form a hybrid with template DNA around the origin. This proposal, while not fully proved, does begin to explain some of the control properties exhibited by RNA I.

Just upstream from the promoter for the primer transcript, there is a palindrome which has also been implicated in DNA replication. This palindrome prevents any transcription from proceeding into the replication region, which could elevate the amount of primer transcript formed. Mutations which reduce the stability of the stem-and-loop structure formed by the palindrome allow transcription to pro-
ceed into the origin region (Selzer et al., 1983). These mutants are compatible with other ColE1-derived plasmids. This may occur because the RNA I produced cannot interact with the unusual product of the readthrough transcription. The main effect of this palindrome is to isolate the origin region from the rest of the plasmid.

In addition to the inhibitory effect of RNA I, ColE1 also appears to produce a small protein which acts to inhibit replication. Deletion of a HaeII fragment downstream from the origin results in a plasmid with a higher copy number (Twigg and Sherratt, 1980). The copy number is reduced to normal levels if a compatible plasmid is already present in the cell. This suggests that the copy number could be regulated by a product of this fragment. In vivo expression of a lacZ gene which uses the primer promoter for transcription is repressed by a plasmid which carries this HaeII region (Cesareni et al., 1982). This region was cloned by Som and Tomizawa (1983), and the properties of the product were examined. A 306 bp fragment was found which specified all the properties of the larger HaeII fragment. Plasmids with this 306 bp fragment reduced copy number in plasmids which contained sequences from the primer transcript. Both pBR322 and ColE1 have regions which perform the same inhibitory effect but the nucleotide sequences are different. Plasmids without this region have the same copy number as ColE1 which has been grown in the presence of chloramphenicol (Som and Tomizawa,
1983). This indicates that a protein is responsible, and a small 63 amino acid protein can be produced from this region. This protein has been isolated and its size agrees very well with the expected protein. Using the in vivo galk assay system, the effect of this protein on the promoter activity of the primer was examined. A plasmid was constructed which had the primer promoter positioned so that it could transcribe the galk gene. When a plasmid with the DNA region coding for the 63 amino acid protein was present in the cell, along with the primer promoter plasmids, the synthesis of galactose kinase was strongly repressed. The amount of repression depended on the length of the region present which was downstream from the primer promoter. The best repression was found when a region more than 50 bp downstream from the primer promoter was included (Som and Tomizawa, 1983). The purified protein was also found to enhance formation of hybrids between RNA I and the primer transcripts. This would serve to increase the inhibitory activity of RNA I. However, it must be noted that this protein is not required for stable plasmid replication, since the plasmid replicates without protein synthesis and derivatives can be made which replicate stably but do not contain the region coding for the protein.

USE OF PLASMIDS AS CLONING VECTORS

The properties of plasmids make them an ideal choice for manipulating DNA. Plasmids are easily isolated and car-
ry specific markers which allow for the selection of recombinant molecules. In many cases, the replication properties of the plasmids allow cloned DNA sequences to be amplified, making it easier to examine the properties of the cloned DNA and its products.

The first constructed plasmid was pSC101 (Cohen et al., 1973; Cohen and Chang, 1973). The resistance plasmid R6-5 was mechanically sheared and recircularized in vivo. A small, mobilizable plasmid was found that did not confer tetracycline resistance. It appears, however, that the plasmid isolated did not come from R6-5 but was a contaminating plasmid (Cohen and Chang, 1977). This plasmid, pSC101, was found to contain a single EcoRI site and to exist in low copy number. The first cloning experiments were performed on this plasmid in which Xenopus laevis ribosomal DNA was inserted into the EcoRI site (Morrow et al., 1974). This plasmid is compatible with ColE1-derived plasmids, thus allowing two plasmids with different cloned sequences to be present in one cell at a time.

The low copy number of pSC101 limited its usefulness for cloning purposes, while its compatibility properties were advantageous. In order to incorporate both high copy and compatibility with ColE1-derived plasmids, Chang and Cohen (1978) developed some plasmid cloning vectors based on the P15A replicon. P15A is a small, cryptic plasmid of approximately 2300 base pairs which does not code for any phenotypic markers, but which can be isolated by an indi-
rect selection system (Kretschmer et al., 1975). Using restriction enzyme digests and in vitro recombination, the replicon of P15A was joined to DNA fragments containing drug resistance genes. pACYC177 is resistant to ampicillin and kanamycin, while pACYC184 is resistant to chloramphenicol and tetracycline. These plasmids are amplifiable, are compatible with ColE1 derivatives, and contain several useful restriction enzyme sites.

By far the most used plasmid for cloning purposes is ColE1 and its derivatives. These plasmids occur in high copy number and are amplifiable several hundred fold. The ease of isolation along with the development of several extremely useful cloning vehicles has made this plasmid the most highly exploited plasmid. ColE1 was the first naturally occurring plasmid to be used as a cloning vehicle (Hershfield et al., 1974). DNA fragments from phage φ80 pt190, some of which contained the tryptophan operon (trp), were inserted into the EcoRI site of ColE1. The authors then looked for the appearance of a Trp+ phenotype in Trp− E. coli cells which had been transformed with the recombinant plasmids. Cells which were Trp+ contained a ColE1 plasmid with an insert in the EcoRI site. This plasmid, pVH15, did not produce any colicins but did have immunity to colicins, indicating that the EcoRI site is present in the gene responsible for colicin production. pVH15 has DNA derived from ColE1 and bacteriophage DNA, along with the trp operon. A spontaneous deletion of
this plasmid yielded a smaller plasmid which retained all
the phenotypes of the original plasmid. The vector
(mini-ColE1) pVH51 was then isolated and was found to be
quite small (Hershfield et al., 1976).

The most commonly used plasmid cloning vehicle is
pBR322. Betlach et al. (1976) isolated a large plasmid
which carried the genes for the EcoRI restriction and
modification system. This plasmid, pMB1, was essentially
the same as ColE1, except for the presence of the restric-
tion and modification genes. A series of smaller deriva-
tives of pMB1 were made which had additional drug resist-
tance genes (Bolivar et al., 1977b). A very small
derivative was constructed which just contained the origin
of replication region and the lac operator, which was used
for selection of the plasmid (Bolivar et al., 1977a). Throu-
gh a variety of \textit{in vitro} and \textit{in vivo} manipula-
tions, the cloning vehicle pBR322 was constructed (Figure
5; Bolivar et al., 1977c). The replication region of
this plasmid is derived from pMB1 and the plasmid contains
genes conferring resistance to ampicillin and tetracycline
to the cell along with several unique restriction enzyme
sites. The plasmid is amplifiable and is incompatible with
ColE1 or ColE1-derived plasmids. Having two selectable
markers with unique restriction enzyme sites present in
each allows a wide range of DNA insertions to be
performed.
Figure 5. Construction of pBR322. R7268 is the same as R1. pMB1 is a clinical isolate. The Tn3 transposon from R1drd19 was placed on pMB1 to form pMB3 (2). EcoRI* reduction of this plasmid produced pMB8, which only has colicin immunity (3). Into the unique EcoRI site of this plasmid was inserted the tetracycline resistance gene from pSC101 to form pMB9 (4). Tn3 was also placed on ColE1 to form pSP2124 (5). The Tn3 element was then transposed onto pMB9 to give pBR312 (6). EcoRI* cleavage and rearrangement resulted in pBR313 (7). Two separate fragments from this plasmid were then isolated and joined to form pBR322 (8).
CLONING STRATEGIES

In order to make use of the available cloning vectors, one must isolate the DNA of interest and prepare it so that it might be cloned. Some of the techniques used to do this are: use of reverse transcriptase to make cDNA copies of RNA, shearing of the DNA, chemical synthesis of the desired DNA fragment, and nuclease digestion of the DNA (Bolivar and Backman, 1979). The ends of the DNA can be engineered to be sticky or blunt ended, which facilitates the insertion of the DNA into the cloning vehicle.

The isolation of mRNA and its conversion into double-stranded DNA has been a useful tool in eukaryotic molecular biology. In general, total poly(A) mRNA is isolated from a cell by adsorption to an oligo(dT) support and enriched for the sequence of interest by size fractionation on sucrose gradients (Nagata et al., 1980) or polyacrylamide gels (Goodman and MacDonald, 1979). A DNA strand complementary to the isolated mRNA is made with reverse transcriptase. The mRNA is hydrolyzed with alkali and E. coli DNA polymerase I can then be used to make a double-stranded complementary DNA (cDNA)(Bolivar and Backman, 1979).

A common technique that is used to insert cDNA into a cloning vector, and one which can be applied to isolated DNA fragments, is the use of homopolymer addition of nucleotides to the 3' end of DNA (Nelson and Brutlag, 1979). Terminal transferase will add oligonucleotide tracts to single-stranded 3' termini of duplex DNA. In homopolymer addition,
a polypurine tract is placed on one DNA molecule and a complementary polypyrindine tract is placed on the other. These two DNA molecules will then hybridize through the homopolymer tracts, producing a recombinant molecule which is stable enough for transformation. Typically, a PstI site in the plasmid is used, since this restriction enzyme leaves a 3' end which is a substrate for terminal transferase. An oligo(dG) tract is produced with terminal transferase, after Pst cleavage of the plasmid. The cDNA molecule which is to be inserted then has an oligo(dG) tract placed at the 3' end. This cDNA is allowed to hybridize with the plasmid which has oligo(dG) tails. Transformation of the recombinant plasmid will result in in vivo repair of any single-stranded regions and, because of the choice in homopolymer additions, the PstI site will be reformed (Nelson and Brutlag, 1979). Okayama and Berg (1982) described the construction of a series of plasmids which simplify the cloning of cDNA. The essential feature of this system is the use of plasmids for the isolation of mRNA by oligo(dT) tracts on the plasmid, followed by the hybridization of the cDNA to the plasmid. An oligo(dT) tract is added to the 3' end of a restriction enzyme-cleaved plasmid and poly(A) mRNA is hybridized to these oligo(dT) tracts. The plasmid is then used as a primer for reverse transcription of the mRNA to produce a complementary copy of the mRNA. Following a few restriction enzyme cleavages and insertion of some restriction enzyme linkers, the entire
cDNA is cloned into the plasmid. This approach has vastly increased the number of full length transformants found (Okayama and Berg, 1982).

The shearing of DNA has been used to randomly produce DNA fragments which can then be cloned (Bolivar and Backman, 1979). This procedure does not depend on any inherent sequences or sizes of the DNA to be sheared and thus has been very useful in generating clones which contain DNA fragments from the entire genome. The main restriction of this technique is that the size of the sheared DNA fragments, usually greater than 10000 bp, is at the upper limit of the size of DNA which can stably be inserted into a plasmid. The randomness of the shearing allows this technique to be used to develop clone banks or genomic libraries. Clarke and Carbon (1975) first used this approach with $E. \text{coli}$ chromosomal DNA that was randomly sheared and, following homopolymer addition, was inserted into ColE1. The use of bacteriophage or cosmid vectors instead of plasmid vehicles is more common for the construction of genomic libraries because of the larger size of the insert which can be accommodated by the bacteriophage (Maniatis et al., 1978). This observation means that smaller number of recombinant vectors are needed in order to generate a library of the entire genome.

Synthesis of Specific Sequences of DNA by Enzymatic Means

Terminal deoxynucleotidyl transferase (Terminal trans-
ferase) is the least specific of the enzymatic methods and is used to create very simple sequences of DNA. This enzyme can add deoxynucleotides to the 3'-OH of DNA molecules (Bollum, 1974). The addition can be to single-stranded DNA with a 3'-OH end or to double-stranded DNA with a protruding 3'-OH end. In the presence of cobalt, double-stranded DNA which is either blunt-ended or has a recessed 3' end can be used (Roychoudhury et al., 1976). In most cases, this enzyme is used to add homopolymer tracts to the end of DNA molecules (Roychoudhury and Wu, 1980; Nelson and Brutlag, 1979). Typically, an oligo(dG) tract is added to either the vector or the DNA to be inserted and an oligo(dC) tract is added to the other. The two molecules can then be annealed to one another through the complementary G-C bases. If a PstI site is used in the vector and an oligo(dG) tract is added to it, the site will be reformd by this procedure (Nelson and Brutlag, 1979).

Another useful enzyme for the synthesis of specific oligonucleotides is polynucleotide phosphorylase. This enzyme catalyzes the sequential addition of a dNDP to another oligonucleotide (Gillam and Smith, 1980). The acceptor oligonucleotide must be at least 3 nucleotides long and conditions have been found where a major product is the addition of a mononucleotide (Gillam et al., 1974). While the reaction is irreversible, there can be several additions of the same mononucleotide. The addition of dADP and dCDP occurs much faster than does the addition
of dGDP or dTDP. This necessitates the isolation and characterization of the oligodeoxynucleotide at each step of the elongation. Typically each incubation takes 10-15 hours at 37 °C, with a yield of 10-70% (Gillam and Smith, 1980). The nature of the acceptor has an effect on the reaction. A stretch of purines is a good primer while a stretch of thymidines acts as a very poor acceptor (Trip and Smith, 1978). The fact that there is no real way to terminate the reaction, resulting in the need to purify at each step, has limited the usefulness of this procedure. A 13 base oligonucleotide has been constructed starting from a tetramer (Hutchison et al., 1978).

The most useful and specific of the enzymes is RNA ligase. This enzyme catalyzes the joining of a 5'-phosphate to a 3'-OH of single-stranded RNA or DNA (Brennan et al., 1983). RNA is a much better substrate than DNA but the DNA reaction will proceed under proper conditions. The DNA molecule with the 3'-OH is called the acceptor and must be 3 nucleotides or more in length (England et al., 1980). The DNA molecule with the 5'-P is called the donor and must be present in high concentrations. Usually, the donor molecule is a 5',3' substituted diphosphate. This results in a single addition of a donor molecule to the acceptor since the presence of the 3'-P terminates the reaction (Brennan et al., 1983). The 3' phosphate is then removed for further reactions. An advantage of RNA ligase over any other enzymatic method is the looseness of
its specificity for the donor molecule. DNA analogs can be just as easily inserted as can normal nucleotides. Also, oligonucleotides can be used as both donor and acceptor, in contrast to polynucleotide phosphorylase. The disadvantages of this system are the need for high concentrations of the donor molecules and the length of time needed for the reaction to take place. Usually, a 1-10 day incubation at 17°C is performed, resulting in yields from 10-70%. The RNA ligase must be very highly purified because of the high concentrations needed and the length of time for incubations.

Chemical Synthesis of Specific DNA Sequences

Chemical means have become the method of choice for synthesizing a specific sequence of DNA. These systems have been used to make restriction enzyme linkers, DNA probes, primers for DNA sequencing, and even complete genes. Site-specific, directed mutagenesis is now possible using synthetic DNA sequences. In 1955, a dinucleotide containing a normal 5'-3' linkage was chemically synthesized (Michelson and Todd, 1955). The first major use of chemical means for the synthesis of DNA sequences was performed by Khorana and his group. They employed a slightly different chemistry than did Michelson and Todd. This procedure is known as the phosphate diester method because the final product is a diester. Other methods used are the phosphate triester and phosphite triester approaches. Each has its
advantages and disadvantages which will be detailed below.

In order for any method to be useful, it must result in a high yield of product. To accomplish this, every reactive group present must be blocked, except for the desired ones. The first groups are the amino groups present on the base. The suitably blocked bases used for chemical synthesis are thymine, N-benzoyladenine, N-anisoylcytosine, and N-isobutyrylguanine (Brown et al., 1979). All three protecting groups can be removed in ammonium hydroxide. These protecting groups were first used by Khorana in the 1960's and are still used today.

The hydroxyl protecting groups used in the diester approach are a 5' acid-labile di-p-anisylphenylmethyl group and a 3'-base-labile acyl group (Figure 6; Caruthers et al., 1983). Using a proper condensing reagent, a 5' protected deoxynucleotide can be joined to a 3' protected deoxynucleotide. This was first done to produce a T-T dinucleotide, with 90-95% yield (Jacobs and Khorana, 1965). However, the yields rapidly decreased as the size of the deoxyoligonucleotide increased. A slightly different chemistry was used to produce longer oligonucleotides. Here, dinucleotides, trinucleotides or longer oligonucleotides are used in a condensation reaction with a 3'-hydroxyl component. The 5' phosphate is first converted to a cyanoethyl derivative and then condensed with the 3' acyl protected compound. The product can be purified and either the cyanoethyl or acyl group can be removed. This oligo-
Figure 5. Phosphate Diester Synthesis of Oligonucleotides. (MeO)₂Tr - di-p-anisylphenylmethyl. Ac - acetyl. B - protected base. A 5' protected nucleoside is condensed with a 3' protected nucleotide to give a dinucleoside monophosphate. This method results in a phosphate diester.
nucleotide can then be used in another round of reactions. The starting materials for this procedure are readily available and easily derivatized. The products can be easily purified and the yield ranges from 30-90% per condensation (Caruthers et al., 1983). The main drawback in this method is the unprotected nature of the phosphodiester internucleotide bonds which can result in undesirable side reactions.

The two most commonly used procedures are both triester approaches (i.e. the product of the condensation is a triester which can be converted to the phosphodiester). These methods block the highly reactive phosphate anions, thus preventing secondary reactions from taking place at these sites. This method also removes solubility problems and electrostatic repulsion effects found in the phosphate diester approach. The phosphate triester method, while first used by Michelson and Todd (1955), did not become useful until the mid-1960s. The initial reaction scheme involved attaching a 5' blocked deoxycytidine to polystyrene (Letsinger and Mahadevan, 1965). It was then condensed with cyanoethylphosphate and washed (Figure 7). This support-bound phosphate diester was activated with mesitylenesulphonyl chloride and the second deoxynucleotide was added to give a phosphate triester dinucleotide. Although yields were low, there were minimal side reactions.

There have been many improvements in the phosphate triester method. The chemistry can be done on polymer supports
Figure 6. Phosphate Triester Synthesis of Oligonucleotides. Tr - triphenylmethyl. P - polystyrene polymer. C - cytosine. This procedure results in a phosphate triester. Activation of the support bound phosphate diester, followed by the addition of thymine produced a dinucleoside monophosphate with a cyanoethyl protecting group which could be easily removed. from Caruthers et al. (1983).
(Ike et al., 1983) or in solution (Hsiung et al., 1983). Instead of simple addition of the mononucleotides, di- or tri-nucleotides can be used, thus speeding the construction of long oligonucleotides. The yields in solution are 90-95% while on polymer supports the yields are 80-85% per condensation. On a solid support, the condensation of the purines is much slower than for the pyrimidines, although the reaction is usually complete in 2 hours for all (Ike et al., 1983). This procedure is the method of choice for synthesizing milligram or gram quantities of an oligonucleotide. Drawbacks of this method include the problem of purification of the intermediate oligonucleotides and the need, at least for solid support schemes, for all 16 dinucleotides in order to quickly synthesize a sequence with significant yields.

The other major approach to chemically synthesizing a DNA sequence is the phosphite triester method. In this method, a nucleoside phosphochloridite is reacted with a nucleoside. The reaction is complete in 5 minutes, and the natural internucleotide bond is formed by oxidation (Letsinger et al., 1975). The yield of the dinucleotide is about 85%, while further extension gives comparable yields. Much work has been done recently to improve the stability of the compounds used in the reactions. In the place of phosphochloridites, phosphoramidites were substituted. It is not necessary to keep these derivatives in an oxygen free, inert environment (Dorper and Winnacker, 1983) and they can be
stored in acetonitrile for at least 40 days without decomposing (Caruthers et al., 1983). This method easily lends itself to solid support chemistry. In this procedure, a 5' protected nucleotide is bound to the support, such as silica gel, by a 3' ester linkage (Figure 8). The 5' group is removed and a 3' phosphoramide is reacted with the bound molecule. The reaction is very fast, usually complete in less than a minute. Unreacted 5' ends can be blocked by acylation with acetic anhydride so that they will not be substrates for further additions. The entire cycle for the addition of a single mononucleotide is less than 30 minutes. Because the synthesis is rapid and gives a high yield, the entire process can be done with mononucleotides.

Further modifications of these procedures have enhanced their use. Different 5' trityl-protecting groups give different colors upon removal. Thus, by having a slightly different protecting group for each nucleoside, the extent of the reaction can be monitored easily (Fisher and Caruthers, 1983). Both triester approaches can be automated (Alvarado-Urbina et al., 1981) however, this allows only one oligonucleotide to be synthesized at a time. Using syringes with a silica support, the cycle time for the phosphite method can be reduced to 13 minutes and several parallel reactions can be done at one time (Tanaka and Letsinger, 1982). Another procedure uses the nucleotide bound to cellulose filters (Frank et al., 1983). In-
Figure 7. Phosphite Triester Synthesis of Oligonucleotides. (Me)Tr = di-p-anisylphenylmethyl. Ac2CO = acetic anhydride. DMAP = 4-dimethylaminopyridine. THF = tetrahydrofuran. B = protected base. P = silica support. The aliphatic phosphoramidite (XXIV) is shown at the bottom. This procedure results in a dinucleoside monophosphate which is bound to a solid support at the 3' terminus. from Caruthers et al. (1983).
stead of adding the nucleotide derivatives to the support, the support is added to the nucleotide derivatives, which are in a petri dish. Thus, only four different reaction solutions need be used, one for each nucleotide, and simultaneous addition to several oligonucleotides is easily achieved by putting several filters in the same reaction solution.

Using these procedures entire genes have been synthesized. Small genes, such as those for somatostatin, insulin, and some tRNAs, were among the first to be produced (Itakura et al., 1977; Goeddel et al., 1979; Brown et al., 1979). The largest gene to be chemically synthesized to date is the 514 bp gene for human leukocyte interferon (Edge et al., 1981). This was constructed with 67 different overlapping oligonucleotides using the phosphotriester method. Small regulatory sequences have also been synthesized. Heynecker et al., (1976) describes the construction of a synthetic lac operator sequence which is functional in vivo. This lac operator sequence was cloned using synthetic linker molecules. Linker molecules are small oligonucleotides which contain the sequence for a restriction enzyme. Ligation of these onto the end of a DNA fragment allows that fragment to then be cloned into a particular restriction site (Heynecker et al., 1976). This gives added options for cloning.

Synthetic oligonucleotides have also been important in construction of primer sequences for Sanger dideoxy se-
sequencing. In the Sanger method, a short DNA sequence is hybridized to a longer DNA fragment. The short DNA sequence is used as a primer in a chain extension reaction with DNA polymerase I (Klenow fragment). Along with the dNTPs, a small amount of a deoxy-substituted nucleotide (ddNTP) is also added. This nucleotide will be inserted into the chain by the DNA polymerase I but will not act as a substrate for further additions. If four parallel reactions are performed, each with a different ddNTP, one will get a nested series of fragments which will provide the DNA sequence when electrophoresed on a polyacrylamide gel (Sanger et al., 1977). The difficult problem with this sequencing method is the isolation of a single strand of the sequence of interest. The M13 cloning system developed by Messing (1983) has overcome this obstacle. M13 is a single-stranded, filamentous bacteriophage which has a double-stranded replicative form. The replicative form can be isolated and DNA fragments can be inserted by restriction enzyme cleavage and ligation. The single-stranded bacteriophage produced by this recombinant replicative form will possess one strand of the inserted DNA. This can then be used in a extension reaction for Sanger deoxy sequencing (Messing, 1983) using a chemically synthesized oligonucleotide for a primer.

The M13 system is also useful in the site-specific mutagenesis of a cloned DNA sequence. A small oligonucleotide is chemically synthesized which is complementary to
the cloned DNA sequence in a recombinant M13 phage. This oligonucleotide has a single base pair mismatch which will be used to produce a mutation. If this oligonucleotide is used as a primer in an extension reaction with DNA polymerase I (Klenow fragment), a double-stranded molecule will be produced with one strand that has a single base change (Norris et al., 1983; Messing, 1983). This molecule can be cleaved with restriction enzymes and cloned into a plasmid or kept in M13. Following transformation, half of the resulting DNA molecules should have the mutation present.

**Enzymatic Procedures for DNA Manipulations**

Nuclease digestion of DNA is the most commonly used method to prepare the DNA for cloning. This digestion can be performed either by exonucleases or endonucleases. The most common exonucleases used are Exonuclease III, which catalyzes the 3' to 5' removal of mononucleotides; S1 nuclease, which digests single-stranded DNA; and Bal31, which combines properties of the other two (Figure 9). Often a fragment of interest is reacted with these enzymes in order to produce a shortened sequence which can then be cloned by a variety of methods. These enzymes are also used to generate deletions in cloned DNA fragments.

Endonucleases are more useful for most purposes since the DNA does not already have to be present in fragments. The endonucleases cleave DNA in either a random or a specific fashion. DNAse I is an example of the former. In the
Figure 9. Types of Digestion of DNA by Nucleases. The combined action of Exonuclease III and S1 nuclease results in similar DNA fragments as the action of Bal31 nuclease. DNAase I in the presence of manganese will randomly cleave double-stranded DNA. The fragments are almost blunt-ended and can be made flush by the action of T4 DNA polymerase. The different ends left by digestion with a restriction endonuclease are also shown.
A) Exonuclease III

\[ \begin{array}{c}
5' \quad A & B \\
A & B
\end{array} \xrightarrow{Mg^{++}}
\begin{array}{c}
5' \\
B
\end{array} \]

B) S1 Nuclease

\[ \begin{array}{c}
5' \quad A & B \\
B & C
\end{array} \xrightarrow{Zn^{++}}
\begin{array}{c}
5' \\
B
\end{array} \]

C) Bal31 Nuclease

\[ \begin{array}{c}
5' \quad A & B & C \\
A & B & C
\end{array} \xrightarrow{Ca^{++}}
\begin{array}{c}
5' \\
B
\end{array} \]

D) DNAase I

\[ \begin{array}{c}
5' \quad p & p \\
p & p & p & p & p
\end{array} \xrightarrow{Mn^{++}}
\begin{array}{c}
5' \\
p & p & p & p & p
\end{array} \]

E) Restriction Endonuclease

\[ \begin{array}{c}
5' \quad N & N & N & N & N \\
N & N & N
\end{array} \xrightarrow{Mg^{++}}
\begin{array}{c}
5' \quad N & N & N & N & N \\
N & N & N & N & N
\end{array} \]

- 5' sticky end
- 3' sticky end
- Blunt end
presence of manganese, this enzyme cleaves double-stranded DNA in a random fashion, leaving a complex mixture of double-stranded DNA (Figure 9). These fragments can be made flush-ended by T4 DNA polymerase and then cloned. Heffron and McCarthy (1980) inserted EcoRI linkers into DNAase-cleaved plasmid DNA. Insertion of an EcoRI linker into a gene sequence will inactivate the gene and the presence of the new restriction enzyme site allows the location to be easily determined.

Restriction enzymes cleave the DNA at specific points in the sequence. They were postulated by Wood to explain why certain mutants lost the ability to degrade foreign DNA and they were first isolated soon thereafter (Wood, 1966; Smith and Wilcox, 1970). There are three different species of restriction enzymes: Type I, Type II, and Type III. While all bind specific DNA sequences, they differ in their exact cofactor requirements and specificity of cleavage sites. Type I enzymes require ATP, magnesium, and S-adenosylmethionine (Arber, 1974; Boyer, 1971), while Type III enzymes only require ATP and magnesium but are stimulated by S-adenosylmethionine (Kanc and Piekaroqitz, 1978). Type I enzymes, after binding to the recognition site, cleave the DNA nonspecifically at some distance from the recognition site while Type III cleave specifically. The third kind, Type II, has proven to be the most useful for recombinant DNA work. These enzymes require only a divalent cation (usually magnesium), a proper pH, and a proper salt
concentration for maximal activity. Since their first discovery, the number of Type II restriction enzymes has grown to over 400 with 90 different specificities. Table 1 lists some of these enzymes commonly used, using the nomenclature first proposed by Smith and Nathans (1973).

Each restriction enzyme usually has a complementary modification enzyme, most often a methylase, which prevents the host's own DNA from being degraded. For example, EcoRI methylase transfers a methyl group from S-adenosyl-methionine to an adenine in the recognition site of EcoRI. This prevents cleavage by the EcoRI restriction enzyme (Greene et al., 1975; Dugiaczyk et al., 1974). As a consequence of this some restriction enzymes will not cleave DNA if one of the nucleotides in its recognition site is methylated, while others will still cleave the DNA even if the DNA is methylated. Some of these enzymes are also shown in Table 1. These differences in specificities of isoschizomers has been very useful in looking at methylation patterns in various types of DNA (Maniatis et al., 1978).

Type II restriction enzymes generally recognize a symmetric or "palindromic" sequence; that is, the DNA sequence of each strand is the same when read in a 5' to 3' direction. Most Type II enzymes recognize a four or six bp site. EcoRI is active as a dimer, binding to both strands of the DNA, with each monomer cleaving one strand (Modrich and Zabel, 1977). Ethidium bromide can be used to inhibit the
<table>
<thead>
<tr>
<th>Restriction Enzyme</th>
<th>Cleavage Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PfoI</strong></td>
<td>GGATGNNNNNNNNN</td>
</tr>
<tr>
<td></td>
<td>CCTACNNNNNNNNNN</td>
</tr>
<tr>
<td><strong>HpaI</strong></td>
<td>GACGCNNNNNN</td>
</tr>
<tr>
<td></td>
<td>CTGGCNNNNNNNNNN</td>
</tr>
<tr>
<td><strong>MboII</strong></td>
<td>GAAGANNNNNNNN</td>
</tr>
<tr>
<td></td>
<td>CTTCTNNNNNNNN</td>
</tr>
<tr>
<td><strong>MboI</strong> (Sau3A)</td>
<td>'GATC AluI</td>
</tr>
<tr>
<td></td>
<td>GATC TCGA</td>
</tr>
<tr>
<td><strong>BspRI</strong> (HaeIII)</td>
<td>G GCC CfoI(HhaI)</td>
</tr>
<tr>
<td></td>
<td>CCGG CGCG</td>
</tr>
<tr>
<td><strong>MspI</strong> (HpaII)</td>
<td>CGGG RsaI</td>
</tr>
<tr>
<td></td>
<td>GGGC TCGA</td>
</tr>
<tr>
<td><strong>HinfI</strong></td>
<td>G'ANTC</td>
</tr>
<tr>
<td></td>
<td>CTNAG</td>
</tr>
<tr>
<td><strong>HindIII</strong></td>
<td>AAGCTT</td>
</tr>
<tr>
<td></td>
<td>TTGCAA</td>
</tr>
<tr>
<td><strong>Aval</strong></td>
<td>CPyCGPuG</td>
</tr>
<tr>
<td></td>
<td>GPuGCPyC</td>
</tr>
<tr>
<td><strong>SmaI</strong></td>
<td>CCCGGG</td>
</tr>
<tr>
<td></td>
<td>GGGCCC</td>
</tr>
<tr>
<td><strong>BamHI</strong></td>
<td>GGATCC</td>
</tr>
<tr>
<td></td>
<td>CCTAGG</td>
</tr>
<tr>
<td><strong>EcoRI</strong></td>
<td>GAATTG</td>
</tr>
<tr>
<td></td>
<td>CTTAAG</td>
</tr>
<tr>
<td><strong>BclI</strong></td>
<td>TGATCA</td>
</tr>
<tr>
<td></td>
<td>ACTAGT</td>
</tr>
<tr>
<td><strong>XbaI</strong> (Sau3A)</td>
<td>TCTAGA</td>
</tr>
<tr>
<td></td>
<td>AGATCG</td>
</tr>
</tbody>
</table>

*-- requires unmethylated DNA. N- any nucleotide; Pu- purine Py- pyrimidine.
restriction enzyme activity somewhat, resulting in single-stranded nicks at the restriction enzyme site (Wallace et al., 1981).

All the restriction enzymes which recognize a palindromic DNA sequence cleave internally to that sequence. Depending on the exact site of cleavage, either a 5' sticky end, a 3' sticky end, or a blunt end can result. The self-complementarity of sticky end fragments has been very important. For example, EcoRI fragments, which have a 5' cohesive end, will only ligate to other EcoRI ends. On the other hand, any blunt-ended fragment will ligate to any other blunt-ended fragment. A 5' or 3' sticky end can be made flush by using T4 DNA polymerase in the presence of all 4 dNTPs. The 3' to 5' exonuclease activity will remove any overhanging 3' ends and the polymerase activity will fill in any sticky 5' end. These properties were first used by Cohen et al. (1973) and have been in wide use ever since.

There is a group of Type II restriction enzymes which recognize a penta-nucleotide DNA sequence that is not palindromic. All of these enzymes cleave the DNA outside of the recognition site. Examples in Table I are HgaI, FokI, and MboII. This unique feature of these enzymes has had several uses. HgaI, since it leaves such unusual sticky ends, has been used in some in vitro recombination experiments (Moses and Horiuchi, 1979). The 5 base sticky end can be any 5 nucleotides so that only very
specific fragments will ligate to one another. *MboII* has been used to remove several base pairs from a cloned sequence of DNA (Scarpulla et al., 1982).

**SUMMARY**

The work described in this thesis has involved plasmid constructions. These plasmids were designed to take advantage of the mode of action of the restriction enzyme, *MboII*. This enzyme has a recognition sequence which is located upstream from the site of cleavage. The actual cleavage site can be any sequence. The only requirement is that it must be eight bases downstream from the recognition sequence. The plasmids detailed in this thesis allow an *MboII* recognition site to be placed adjacent to virtually any DNA sequence. This permits a wide range of DNA manipulations to be performed which were difficult to achieve previously. The replication region of these plasmids is much smaller than had been previously reported. The replication properties of these plasmids were investigated. These investigations included: plasmid copy number, plasmid stability, incompatibility characteristics of the plasmids, and examination of the promoters found on the plasmids.
II. EXPERIMENTAL PROCEDURES

MATERIALS

Ampicillin, tetracycline-HCl, chloramphenicol, 5-methyl tryptophan, and trimethoprim were from Sigma. 5-bromo-4-chloro-3-indolyl-beta-galactoside (X-gal) was from Bethesda Research Laboratories. PEI cellulose plates were from Brinkman. Cellogel strips (3 X 57 cm) were obtained from Kalex.

Bacterial Strains

The strains used in this study and their genotypes are shown in Table 2. *E. coli* strain RR1 was used for lac operator selections. Strain W3110 trpA1417 was used for trp operator selection. GM119, a Dam- strain of *E. coli*, was used when unmethylated plasmid DNA was needed.

Enzymes

BspRI, an isoschizomer of HaeIII, was prepared as described in Beski *et al.* (1978). The other restriction enzymes utilized (Table 1) in this work were obtained from commercial sources. The reaction conditions were those recommended by the suppliers. T4 DNA ligase was prepared following the protocol of Panet *et al.* (1973), using the overproducing strain of Murray *et al.* (1979). *E. coli* RNA polymerase was prepared as detailed in Burgess and Jendrisask (1975) and Lowe *et al.* (1979).
### Table 2

Strains of *E. coli* and Their Genotypes

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>W3110</td>
<td><strong>trpA1417</strong></td>
<td>C. Yanofsky</td>
</tr>
<tr>
<td>W3110</td>
<td><strong>tna2</strong></td>
<td>&quot;</td>
</tr>
<tr>
<td>W3110</td>
<td><strong>tna2- trpR</strong></td>
<td>&quot;</td>
</tr>
<tr>
<td>W3110</td>
<td><strong>tna2- trpR recA56::Tn10</strong></td>
<td>&quot;</td>
</tr>
<tr>
<td>AB1157</td>
<td><strong>thr-1 leu-6 thi-1 lacY1 galK2 ara-14 xyl-5 mt1-1 proA2</strong></td>
<td>R. Moses</td>
</tr>
<tr>
<td></td>
<td><strong>his-4 argE2 str-31 tsx-33 sup-37</strong></td>
<td>&quot;</td>
</tr>
<tr>
<td>M76</td>
<td>recA derivative of AB1157</td>
<td>&quot;</td>
</tr>
<tr>
<td>RR1</td>
<td><strong>hsdS20(r&lt;sub&gt;B&lt;/sub&gt;, m&lt;sub&gt;B&lt;/sub&gt;) ara-14 proA2</strong></td>
<td>H. Boyer</td>
</tr>
<tr>
<td></td>
<td><strong>lacY1 galK2 rpsL20(Sm&lt;sup&gt;F&lt;/sup&gt;) xyl-5 mt1-1 supE44</strong></td>
<td>&quot;</td>
</tr>
<tr>
<td>HB101</td>
<td>recA derivative of RR1</td>
<td>&quot;</td>
</tr>
<tr>
<td>GM119</td>
<td><strong>dam-3 dcm-6 metB1 galK2 galT22</strong></td>
<td>M.G. Marinus</td>
</tr>
<tr>
<td></td>
<td><strong>lacY1 tsx-78 supE44 (thi-1 tonA31 mt1-1)</strong></td>
<td>&quot;</td>
</tr>
<tr>
<td>C600</td>
<td><strong>galK thr-1 leuB6 lacY1 tonA21</strong> supE44</td>
<td>M. Rosenberg</td>
</tr>
</tbody>
</table>

28
Alkaline phosphatase and snake venom phosphodiesterase were obtained from Boehringer Mannheim. T4 polynucleotide kinase lacking 3' phosphatase activity from New England Nuclear, while polynucleotide kinase possessing 3' phosphatase activity was purchased from New England Biolabs. DNA polymerase I (Klenow fragment) was from New England Biolabs and Boehringer Mannheim. T4 DNA polymerase was purchased from P-L Biochemicals. DNAase I was obtained from Boehringer Mannheim.

**Plasmid DNA**

pMT100 was obtained from G. Swift (Swift et al., 1981). Other trimethoprim-resistant plasmids were obtained from M. Fling (Fling and Elwell, 1980). pBR322, as sequenced by Sutcliffe (1979), was from the Plasmid Reference Center, Stanford, CA. pKK3535 was the gift of H. Noller (Brosius et al., 1981).

**Bacterial media**

Bacteria were plated out on nutrient agar plus cysteine (Cys-NA). Suitable antibiotics were added to select for bacterial strains carrying specific plasmids. Mueller-Hinton agar (BBL) was used for bacterial selection with trimethoprim as this media lacks many of the trimethoprim antagonists found in other media. Ampicillin (Ap) was added to a concentration of 30 mg/l and trimethoprim (Tp) to a concentration of 50 mg/l. For liquid cultures, the bac-
teria were grown in L-broth (Luria-Bertani medium) or M9CA medium (Maniatis et al., 1982).

METHODS

Rapid isolation of plasmid DNA

Small scale screenings of plasmid DNA were performed using the phenol lysis procedure (Klein et al., 1980). A single colony was used to inoculate a 5 ml L-broth culture. After incubation overnight at 37 °C, the culture was centrifuged and the cells were resuspended in 0.4 ml 50 mM Tris-HCl (pH 8). The solution was transferred to a 1.5 ml Eppendorf tube and 50 μl freshly prepared lysozyme (10 mg/ml) in 10 mM Tris-HCl (pH 8) was added. After 15 minute at room temperature, 0.5 ml phenol saturated with 100 mM Tris-HCl (pH 8) was added, the tubes were inverted 3 times and the samples were centrifuged in an Eppendorf centrifuge for 15 minutes. The aqueous phase was removed, placed in another Eppendorf tube, and extracted several times with ether. The DNA was precipitated by adding 1/10 volume 3 M sodium acetate and 2.5 volumes of 95% ethanol. The tubes were cooled to -70 °C for 10 minutes, followed by a 5 minute spin in the Eppendorf centrifuge. The DNA was redissolved in 0.3 mM sodium acetate and reprecipitated with ethanol, as above. The pellet was washed with ethanol, dried, and resuspended in TE (10 mM Tris-HCl, pH 8, 0.1 mM EDTA). The DNA was now suitable for restriction enzyme digests although 1 μl RNAase A (10 mg/ml) was added to the digestion
mix to allow smaller fragments to be visualized. Sufficient DNA was obtained for 3-4 digests.

**Large scale isolation of plasmid DNA**

Isolation of plasmid DNA was as described in Norgard (1981). M9 (600 mls) supplemented with 2.4 g acid hydrolyzed casein (ACH) was inoculated with 5 mls of an overnight L-broth culture. The cells were grown to log phase (70-100 Klett units) and 3 mls chloramphenicol (32 mg/ml dissolved in 95% ethanol) were added. After overnight amplification, the bacteria were harvested by centrifugation and a cleared lysate was prepared by the Brij lysis procedure (Clewell and Helinski, 1970). To the lysate was added 1/3 volume 5 M NaCl and 1/3 volume of 50% PEG 6000 (Sigma Chemical) (Humphreys et al., 1974). This precipitates both RNA and DNA. After 30 minutes at 4°C, the mixture was centrifuged, the pellet was resuspended in 5-10 mls of TE and 50 μl of heat-treated RNAase A (10mg/ml) were added. The solution was incubated at 43°C for 45 minutes, then extracted with 5 mls of phenol saturated with 100 mM Tris-HCl (pH 8). The aqueous supernatant was removed and passed over a A0.5M Bio-Gel column (BioRad laboratories). This step separates the plasmid DNA from contaminating RNA. The plasmid DNA fractions contained little chromosomal DNA as shown by electrophoresis on 0.8% agarose gels. The appropriate column fractions were collected and precipitated with ethanol.
Gel electrophoresis

Agarose gels (usually 0.8%) were used as described in McDowell et al. (1977). Native polyacrylamide gels (5%) followed the procedure of Maniatis et al. (1975). Denaturing sequencing gels were made according to Maxam and Gilbert (1977). Preparative gels for restriction fragments were formed in the same fashion as normal native polyacrylamide gels except that 3 mm spacers were used in constructing the gel sandwich. Staining of the gels with ethidium bromide and subsequent photography under UV illumination have been described (Maniatis et al., 1982). Detection of radioactive DNA using autoradiography and intensifying screens followed the procedures of Swenstrom and Shank (1978).

DNA fragment isolation

Restriction enzyme fragments from either native or preparative polyacrylamide gels were eluted from the gel using the procedure of Maxam and Gilbert (1977). The section of the gel containing the fragment was sliced into pieces and placed into 2 volumes of a Buffer X (0.5 M ammonium acetate, 0.01 M magnesium acetate, 0.01 % sodium dodecyl sulfate, 0.1 mM EDTA). This mix was incubated for 24 hours at 33 °C with shaking, filtered through siliconized glass wool to remove the gel slices, and precipitated with ethanol. In order to further remove residual acrylamide,
a phenol extraction was usually performed.

In the case of radioactively labeled oligonucleotides which had been eluted from denaturing gels, the liquid containing the fragment was passed over a DE-52 cellulose column. The DE-52 was equilibrated for 1-2 hours in TE and approximately 1 ml was poured into a column. The sample was brought up in a 10 times volume of water and loaded on the column. Following a 10 ml wash with water, the DNA was eluted with 4-6 ml of 30% triethylamine carbonate (prepared by bubbling carbon dioxide through the 2 phases of a 30% triethyl amine/water solution until a single phase forms). The pH of this solution is between 10 and 10.5. The collected samples were then lyophilized repeatedly to remove the volatile buffer.

**Enzyme reactions**

Restriction enzyme digests were performed under conditions recommended by the suppliers. DNA ligations were performed with T4 DNA ligase in 60 mM in Tris-HCl (pH 7.4), 6 mM magnesium chloride, 1 mM DTT, 0.4 mM ATP at 12 °C for 1-15 hours. In some cases, the ligation reaction was precipitated with ethanol and dissolved the residue in the appropriate restriction enzyme buffer. Restriction enzymes were often added to cleave ligated vectors that did not contain inserts. Removal of protruding 3’ ends was performed with T4 DNA polymerase in 3 mM Tris-acetate (pH 7.9), 6.6 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM DTT, 6 μM
each of the dNTP for 15 minutes at 37 °C. *E. coli* DNA polymerase I (Klenow fragment) was used to fill in protruding 5' ends. The DNA to be treated was suspended in 10 mM Tris-HCl, pH 7.9, 10 mM magnesium chloride, 1 mM DTT and 6 μM of the appropriate dNTP. DNA polymerase I (Klenow fragment) was added and the reaction was incubated at room temperature for 30 minutes. Following heat inactivation at 65 °C, the DNA was precipitated with ethanol. This reaction could also be used to label the 3' end of the DNA fragment if an appropriate alpha-^{32}P-NTP was used.

Oligodeoxyribonucleotides were labeled at the 5' end in one of two fashions. The DNA was resuspended in 60 mM Tris-HCl (pH 7.6), 10 mM magnesium chloride and 10 mM gamma-^{32}P-ATP. If the 3' phosphate present on the oligonucleotide was to remain, a T4 polynucleotide kinase which lacked a 3' phosphatase activity was then added. This enzyme was obtained from NEN and was prepared from the T4 mutant T4 PsetlamE10, which has no endogenous 3' phosphatase activity (Richardson, 1965). If the 3' phosphate was to be removed, T4 polynucleotide kinase from wild type phage-infected cells which contained a 3' phosphatase activity (New England Biolabs) was used. The reaction was incubated at 37 °C for 30 minutes followed by ethanol precipitation or DE-52 column purification. Cold ATP was used if there was no need for a radioactive oligonucleotide.
Transformations

Transformations followed the calcium chloride procedure (Kushner, 1978). Cells (50 mls) from the early exponential growth phase were centrifuged, washed in 10 mls of 100 mM MOPS (pH 6.5), 10 mM calcium chloride, and placed on ice for 30 minutes. The cells were pelleted and resuspended in 2 mls of the above buffer. Aliquots (0.2 mls) were placed in sterile tubes and DNA was added. The tubes were incubated on ice for 30 minutes, then were placed at 43°C for 2 minutes. The amount of DNA used varied due to the different transformation qualities of the different strains. Typically, 0.1 pmol of vector was needed for C600 while up to 1 pmol was used for GM119. L-broth (5 mls) was added and the transformed cells were then placed in a 37°C incubator for 1 hour before spreading (0.2 ml) on an appropriate agar plate for incubation overnight in a 37°C incubator.

DNA sequencing

Some of the oligonucleotides were sequenced using the 2 dimensional, mobility shift method. The labeled oligonucleotide to be sequenced was resuspended in 5 mM Tris-HCl (pH 7.6), 5 mM phosphate, 2 mM magnesium chloride, several units of snake venom phosphodiesterase were then added. Aliquots(1 μl) were removed at intervals, placed in a stop mix (10 μl of 2 mM EDTA in 0.1 mM ammonium hydroxide), and dried down. The sample was applied to a cellogel strip and electrophoresed at 4000 volts for 20 minutes in a
pyridine-acetate (pH 3.5) buffer. The strip was removed and the DNA was transferred to a PEI plate for homochromotography in a 3% homomix (Brownlee, 1972). The DNA could then be visualized by autoradiography. For the determination of the terminal residue, a complete digestion of the oligonucleotide was performed with snake venom phosphodiesterase. The products were spotted on some Whatman 3M paper, adjacent to appropriate standards, and electrophoresed in the pH 3.5 pyridine-acetate buffer. The location of the standard was observed by UV illumination, while the terminal residues were subjected to autoradiography.

Double-stranded DNA was sequenced using the methods of Maxam and Gilbert (1977). Using end-labeled DNA, base specific chemical reactions cause the scission of the DNA strand. Following polyacrylamide gel electrophoresis on denaturing gels and subsequent autoradiography, the DNA sequence can be read off of the film.

**Determination of plasmid copy number**

Cells containing plasmids were taken from drug plates and used to inoculate 50 mls L-broth. When the culture reached a Klett reading of 70, 20 mls were removed, the cells were pelleted, and stored for plasmid isolation. Chloramphenicol was added to the remaining cultures and the cells were amplified overnight, when another 20 mls were removed. Plasmid DNAs were isolated in the same manner as the large scale plasmid preparation detailed above, except
that the amounts of solutions used were decreased by a factor of 10. Aliquots of the plasmid DNA were cleaved with EcoRI and electrophoresed on 0.8% agarose gels. Visualization of the DNA after ethidium bromide staining permitted the examination of the relative copy number of each plasmid. A more quantitative inspection of the relative copy number of each plasmid was made by cleaving aliquots of the plasmid DNA with TaqI. The DNA fragments were electrophoresed on 5% polyacrylamide gels, visualized under UV after staining with ethidium bromide and photographed with Polaroid Type 55 film, which produces both a positive and a negative. The negative was scanned with a Kratos MD 3000 integrating densitometer and the intensity of similar sized bands was compared in order to determine relative copy number. This easily overcomes the difference in intensity of differently sized DNA fragments.

Plasmid stability

The strains used (Table 2) were transformed with the appropriate plasmid. After selection and restreaking, single colonies were removed and used to inoculate L-broth for overnight growth without selection. The cultures were diluted and plated on Mueller-Hinton agar plates, with and without Tp. The relative number of colonies on the two plates gives an indication of the stability of the plasmid in the strain. Several selected strains containing plasmids were used for a more detailed study of plasmid stabil-
ity. The cells were grown in L-broth and kept in exponential growth for over 40 generations. This was done by frequently diluting the culture into fresh media. At certain times aliquots were removed, serially diluted and plated on Mueller-Hinton agar plates with and without Tp. The generation time could then be determined for each strain. A comparison between the number of colonies on each type of plate reveals the stability (by retention of the drug resistance phenotype) of the plasmid over the time of growth studied.

**Plasmid incompatibility**

W3110 *trpA*1417 cells containing a resident plasmid were transformed with the appropriate donor plasmid, and plated on Mueller-Hinton agar plates containing both Ap and Tp. Several colonies were selected and grown in L-broth for 10-15 hours before being plated on Mueller-Hinton agar plates without drug. Single colonies were then replica plated onto Mueller-Hinton agar plates containing either Ap or Tp. After incubating the plates at 37 °C, the resistance of each colony to each drug was determined. Incompatibility of the two plasmids can be seen by the loss of one plasmid when selective pressure is removed.

**RNA polymerase binding to promoters**

This procedure has been previously described (Russell and Bennett, 1981). The appropriate plasmid DNA was cleaved
with CfoI (an isoschizomer for HhaI) or HinfI, the restriction enzyme was inactivated by heating at 65 °C for 10 minutes, and the DNA was precipitated with ethanol. The DNA (1.5 pmol of plasmid) was suspended in binding buffer (10 mM magnesium chloride, 100 mM KCl, 0.1 mM DTT in TE) and heated to 37 °C for 15 minutes. The binding reaction was stopped by adding 10 μl of a 1 mg/ml solution of heparin in binding buffer. After 1 minute, the solution was applied with low suction to a 25 mm nitrocellulose filter (Schleicher and Schuell) which had been prepared by soaking in 0.1 N KOH for 5 minutes and washing in binding buffer. After application of the binding mix, the filter was washed with 2 ml of binding buffer and placed in a 1.5 ml Eppendorf tube. The DNA was eluted from the filter by adding 1 ml Buffer X (0.5 M ammonium acetate, 0.01 M magnesium acetate, 0.01% sodium dodecyl sulfate, 0.1 mM EDTA) and incubating at 37 °C for 1 hour. The DNA was ethanol precipitated, taken up in TE, and electrophoresed on a 5% polyacrylamide gel. The flow-through from the nitrocellulose filter was collected on a 25 mm Whatman DE81 filter which had been placed beneath the nitrocellulose filter. This DNA was eluted as described above, precipitated and electrophoresed on a 5% polyacrylamide gel.

**lac operator selection**

Selection for the *lac* operator has been described (Heyneker *et al.*, 1976). Plasmids carrying the *lac*
operator sequence (lac o) will cause E. coli RR1 colonies to appear blue on agar plates which are supplemented with X-gal. Usually 0.2 g of X-gal was dissolved in 10 ml of N-N dimethyl formamide and 0.1 ml was spread on the agar plate. Colonies which do not have the lac o-containing plasmid will appear white.

**Selection of trp operator**

Selection of the trp operator (trp o) is based upon a similar scheme as lac o selection. In strains which do not produce a functional trp synthetase (trpA), low indole levels will not sustain cell growth unless the level of trpB activity is high. Trp repressor binding to the trp operator will keep synthesis of trpB low. This state is maintained by the presence of 5-methyltryptophan, which serves as a co-repressor. Thus, strains such as W3110 trpA1417 will grow on selective media only if it is trpR or operator constitutive (Bennett and Yanofsky, 1978). If a high copy number plasmid contains a trp operator sequence, it will partially titrate off the trp repressor, allowing transcription and translation of the trpB gene. W3110 trpA1417 cells with this plasmid will grow while cells without the plasmid, or with a plasmid that does not bear the trp operator sequence, will not grow. The media used was a glucose-minimal plates supplemented with 5-methyltryptophan (50 g/ml), indole (1.5 g/ml), ACH (0.2%), and the appropriate drug.
DNAse I digestion of plasmid DNA

The conditions needed to linearize plasmid DNA were determined empirically. Typically, 50 μg of plasmid DNA was placed in 50 μl of 5 mM Tris-HCl, pH 8, 0.05 mM EDTA, and 10 mM manganese chloride. DNAase I (5 μl of a 0.5 ng/ml solution) was then added and the reaction mix was held at 22 °C. Aliquots were removed and added to a stop mix of 25 mM EDTA. The time course samples were electrophoresed on 0.8% agarose gel in order to determine the time needed for linearization of the plasmid. For the insertion of a fragment following DNAase I treatment, the ends of the linearized plasmid DNA were made flush with T4 DNA polymerase. A 10-fold molar excess of insert to plasmid DNA was added and a ligation was performed, after which the ligation mixture was used to transform bacteria.
Chapter 2 Construction of a plasmid cloning vector with a deletion in the origin of replication

INTRODUCTION

The region surrounding the origin of replication of ColE1-type plasmids has been intensely studied (Nossal, 1983; Staudenbauer, 1977; Veltkamp and Stuitje, 1981). RNA polymerase, DNA polymerase I and ribonuclease H (RNAase H) have been shown to be sufficient for initiation of DNA replication (Itoh and Tomizawa, 1982). RNA polymerase begins transcription in vitro of a primer precursor (RNA II) 555 base pairs upstream from the origin of replication (Itoh and Tomizawa, 1980). Transcription proceeds towards the origin with most of the transcripts terminating past the origin. Some of the nascent transcripts form stable hybrids with the template DNA, providing a structure which can be cleaved with RNAase H. The cleaved RNA can then act as a primer for DNA replication starting at the origin (Itoh and Tomizawa, 1980).

The mechanism by which the hybrid is formed between the primer transcript and the template is unknown. Studies of plasmid replication have shown that a region about 500 base-pairs upstream from the origin is important for regulating plasmid copy number and for determining the incompatibility characteristics of the plasmid (Conrad and Campbell, 1979; Muesing et al., 1981; Tomizawa et al., 1981). There are two RNA transcripts produced from this area. One is the 5' terminus of RNA II (Itoh and Tomizawa, 92
1981) and the other is a small RNA about 100 bases long, designated RNA I (Morita and Oka, 1979). All of RNA I is complementary to the 5' end of RNA II. In vitro studies have shown that RNA I acts to inhibit primer formation (Tomizawa et al., 1981). This inhibition is also incompatibility group specific (Lacatena and Cesareni, 1981; Tomizawa and Itoh, 1981).

Both RNA I and RNA II can form internal stem-and-loop structures and it has been suggested that interactions between the two RNA molecules occur through the base-pairing of sequences in the single-stranded loop region (Tomizawa and Itoh, 1981). Changes of bases in the loops affect not only the inhibitory activity of RNA I but other aspects of RNA interactions such as incompatibility and rate of association between RNA I and RNA II (Tomizawa and Itoh, 1981; Lacatena and Cesareni, 1981). Incorporation of inosine in place of guanosine in the elongating transcript has a definite effect on primer formation even when the substitution is far upstream from the origin (Tomizawa and Itoh, 1982).

Several reports have sought to determine the smallest fragment of DNA which is capable of sustaining stable replication. These studies use CoIE1 or pBR322 which is derived from pMB1, a plasmid in the same incompatibility group as CoIE1. The smallest region determined by Backman et al. (1978) included the entire region containing the origin, RNA I and RNA II. Oka et al. (1979) produced a smaller plasmid which had a large deletion that removed all
of the RNA I region. While this could replicate it was reported to be unstable. Selzer et al. (1983) also report a construction which replaces the region 420 base pairs upstream from the origin with a sequence containing the promoter of RNA I. This plasmid was also observed to be unstable in the bacterial strains reported.

This chapter describes the construction and characterization of a small plasmid which has a CoIE1-type origin of replication. This plasmid contains a dihydrofolate reductase (DHFR) gene which confers resistance to trimethoprim (Tp), the E. coli tryptophan promoter-operator (trp P0), and several unique restriction enzyme sites. This plasmid is missing the region which produces RNA I, so studies on the stability of this plasmid along with its incompatibility characteristics and copy number were also examined.

RESULTS
Ligation of replication origin and DHFR gene

The trimethoprim resistant dihydrofolate reductase (DHFR) gene of pMT100 is on a EcoRI-PvuII fragment inserted into pBR322 (Swift et al., 1981). The protein is only 78 amino acids long, making it an attractive choice for a selectable marker. The entire region has been sequenced with probable promoters and coding regions identified. A 713 bp BspRI fragment from pMT100 extends 491 bp upstream from the coding sequence and contains all but the
last 4 nucleotides of the gene for the DHFR. The replication region of pBR322 can be isolated on a 436 bp BspRI fragment (from -420 to +16, using the numbering system of Selzer et al., 1983). This fragment contains the origin but does not contain the sequence for RNA I, the 5' end of the primer transcript or its promoter. The construction scheme is shown in Figure 10. pMT100 and pBR322 were each cleaved with BspRI and the DNA fragments were placed in one Eppendorf tube. The mixture was then ligated and transformed into E. coli strain C600. Following selection on Mueller-Hinton agar plates with Tp, several of the colonies were screened for plasmids. Since the sequences of all the fragments are known, the size of restriction fragments will provide information about the desired plasmid. Essentially, this involved looking for the smallest plasmid which had resistance to trimethoprim. One plasmid, pRBG57, was chosen for further study. This plasmid is small (about 1100 bp) which means that it has very few restriction enzyme sites. The plasmid was mapped with restriction enzymes and the orientation of the two BspRI fragments to each other was determined.

This plasmid was not very useful for cloning purposes as it did not have many unique cloning sites. This was alleviated by inserting another fragment into pRBG57 (Figure 10). A 390 bp BspRI fragment from pDR494 contains a 41 bp tryptophan promoter-operator (trpPO) sequence as well as a polylinker region containing several unique re-
Figure 10. Construction of pRBG156 and pRBG177. pMT100 was cleaved with BspRI and the blunt ends were joined with T4 DNA ligase. A plasmid, pRBG57, was chosen for further work. A 390 bp BspRI fragment (Russell, 1983) was inserted into the HincII site. The resultant plasmid, pRBG156, was selected by \textit{trp} operator activity (Russell and Bennett, 1982). The \textit{trpPO} region of pRBG156 was removed by cleaving with EcoRI, followed by ligation of the sticky ends. This plasmid was designated pRBG177. DNA sequences downstream from the origin, in the direction of DNA replication are given positive numbers. \textit{O}r - origin. \textit{B} - BspRI. \textit{E} - EcoRI. \textit{DHFR} - dihydrofolate reductase. \textit{trpPO} - the tryptophan promoter-operator region.
stricetion enzyme sites (Russell, 1983). This fragment was isolated and inserted into the \textit{HincII} site of pRBG57. Transformation into W3110 \texttt{trpA1417} allows selection for the \texttt{trp} operator sequence (Russell and Bennett, 1982). Following selection on minimal-ACH plates containing 5-methyltryptophan and indole, colonies were plated on Mueller-Hinton agar plates with Tp. Selected colonies were screened and one, pRBG156, was used for further work.

pRBG156 has been extensively mapped by restriction enzyme analysis and has been partially sequenced (Figure 11, 12). The expected sequences were found in the examined regions. An anomalous \texttt{Alul} site was found which may have resulted from a single base pair change. \texttt{CfoI} and \texttt{HinfI} digests allowed the orientation of the DHFR gene with respect to the origin of replication to be determined. The orientation of the \texttt{trpPO} and polylinker fragment was confirmed with \texttt{CfoI-BamHI} double digests. The presence of \texttt{EcoRI}, \texttt{HindIII}, \texttt{SmaI}, \texttt{AvaI}, and \texttt{HaeIII} sites has been confirmed. As can be seen in Figure 13, the region of the DHFR gene coding for the carboxy terminal region of the protein is fused onto the -35 region of the RNA I promoter region. The protein product of the DHFR gene should be six ammio acids larger than the normal protein. The -35 region of the RNA I promoter is still intact, however, the sequence for RNA I is no longer present. The entire promoter for RNA II is removed along with the 5' portion of the RNA II sequence. A region on ColE1 just upstream from the
**Figure 11.** Restriction Enzyme Mapping of pRBBG156.

pRBBG156 was digested with several different enzymes and electrophoresed on a 5% polyacrylamide gel. Size standards from pBR322 were used to determine the size of the DNA fragments. Since the DNA sequence was known for pMT100, it was relatively simple to orient the fragments. Lane 5 is of particular interest since it demonstrates the lack of any MboII sites on pRBBG156. The digestion pattern is the same as CfoI digestion alone. The restriction map determined by the digestions is shown next to the gel. pRBBG156 was digested with: Lane 1 - CfoI; Lane 2 - CfoI-BspRI;
Lane 3 - CfoI-BamHI; Lane 4 - CfoI-HindIII;
Lane 5 - CfoI-MboII; Lane 6 - BspRI; Lane 7 - RsaI;
Lane 8 - RsaI-HinfI; Lane 9 - HinfI; Lane 10 - MspI;
Lanes 11 and 12 are pBR322 digested with CfoI and BspRI, respectively. DHFR - dihydrofolate reductase. E - EcoRI.
P - PstI. S - SalI. Em - BamHI. Sm - SmaI.
Hd - HindIII. B - BspRI.
Figure 12. Nucleotide Sequence of pRBG156. Using the reported DNA sequences for pMT100 (Swift et al., 1981), as well as the sequence for the polylinker region (Russell, 1983), it was possible to generate the sequence for pRBG156. The trpPQ region is boxed and the orientation is indicated by the arrow. The -35 and -10 regions of the DHFR as proposed by Swift et al. (1981) is shown. The start of the DHFR gene is indicated by the arrow at position 880. The BspRI (B) site at the end of the DHFR is shown at position 1110. The origin (ori) is shown at position 1528. A presumptive AluI site (A) is shown at position 1324. The DNA regions which have been sequenced are shown by the overlining. This includes the junctions of both BspRI fragments. The numbering is the same as Figure 11.
<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
<th>110</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>120</td>
<td>130</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
<td>210</td>
<td>220</td>
<td>230</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- Table A and B are not clearly discernible due to the distortion in the image.
Figure 13. Comparison of the Replication Regions of pBR322 and pRBG156. (a) The general aspects of the replication regions of pBR322 and pRBG156 are shown. The direction of transcription for RNA I, RNA II (primer), and the dhfr gene are shown by arrows. The RNA I promoter region (-35 and -10) is presented, as is the location of the BspRI site in the RNA I promoter. (b) The nucleotide sequence around the RNA I promoter region in pBR322 and pRBG156. The location of the BspRI site is shown, along with the first nucleotide transcribed (+1). Blunt-end ligation of the end of the DHFR gene onto the BspRI site removes the -10 region from the promoter of RNA I.
(a) Primer $\rightarrow$ BspRI $\rightarrow$ Origin $\rightarrow$ RNAI $\rightarrow$ Replication Region (pBR322)

DHFR gene $\rightarrow$ BspRI $\rightarrow$ Origin $\rightarrow$ Replication Region (pRBGI56)

(b) $\text{BspRI}$

```
TCCTTCTAGTGCGGTTAGGCCCACCACCCCTTAAGGA
AGAAGATCACATCGGGCATCAATCCGGTGGTGAAATTT
```

pBR322

```
CCTG TGCGGACACTTTGAGTGGCCACCACTTCAAGAA
GGACACGCTCGTAAGTTGACACCCGGTGGTGAAATTT
```

pRBGI56

DHFR gene $\rightarrow$ RNAI
RNAII promoter has been shown to cause transcription termination (Itoh and Tomizawa, 1980). This structure is not present in pRBG156. Transcription of the DHFR gene proceeds in the same direction as the normal primer transcript so it is possible that readthrough transcripts of the DHFR gene could act as primer transcripts.

Insertion of the 390 bp fragment into the HincII site of pRBG57 changes the supposed -35 region of the DHFR promoter (Swift et al., 1981)(Figure 14). It was thought that if this insertion disrupted proper transcription of the DHFR gene, the trp promoter could replace the DHFR promoter. However, the trp promoter points away from the DHFR gene in pRBG156, as determined by mapping analysis. In an attempt to place the trp promoter in the same direction as the DHFR gene, pRBG156 was cleaved with EcoRI, which has sites on either side of the trpPO region. After ligation of the cleaved vector to itself, W3110 trpA1417 cells were transformed and plated out on Mueller-Hinton agar plates with Tp. Resultant colonies were screened and the only plasmids found were identical to pRBG156 or had the entire trpPO-polylinker region deleted. The latter plasmid was designated pRBG177 and was characterized by restriction enzyme mapping.

**Plasmid stability**

Similar plasmids to pRBG156 had been constructed (Oka et al., 1979), but they had been reported to be un-
Figure 14. Comparison of the DHFR Promoter Regions of pMT100 and pRBG156. The proposed promoter region of pMT100 is shown (Swift et al., 1981). The HincII site where the trpPO-polylinker region was inserted to produce pRBG156 is indicated. The two sequences are only different upstream from the HincII site.
pMT100 promoter region
\[-35\]
CGAACCGAGGTGACATAAGGCGCTTGGTTCGTAACACTGTAATGCAAGT
\[-10\]
HincII

pRBG156 promoter region
\[-35\]
AGTAAGCTTGAGCATAGCCCTTGTTCCGGTAAACCTGTAATGCAAGT
\[-10\]
stable. It was of interest to determine the stability of pRBG156, especially in different strains. In order to examine the stability of pRBG156, the plasmid was transformed into the appropriate strains and plated out on Mueller-Hinton agar with Tp (see table 2 for full strain genotypes). Colonies were grown overnight in L-broth without drug and then replated on agar plates with and without Tp in order to examine the stability of the plasmid in the strain (Table 3). A significant difference in the number of colonies on the two plates indicates an instability of the plasmid in that strain.

The results show that pRBG156 was extremely unstable in many commonly used laboratory strains of *E. coli*. Several strains, such as AB1157 and M76, did not yield any viable colonies on Mueller-Hinton agar plates containing Tp. Others showed a marked loss of pRBG156 when grown overnight, although complete loss of the plasmid was not observed. Individual mutant alleles may confer increased instability on pRBG156. The presence of the *recA* genotype results in increased instability of the plasmid. Many of the strains tested have mutations, such as *supE*, which might have some effect on plasmid stability. W3110 is essentially wild-type *E. coli* K-12, perhaps explaining the stability of pRBG156 when present in this strain.

A more detailed study of plasmid stability was then performed using several strains. Cells were grown in L-broth without drug selection and kept in log phase for
Table 3

Plasmid Stability

The plasmids were transformed into the appropriate strains and plated on Mueller-Hinton agar plates with Tp (50 mg/l). Resultant colonies were grown overnight in L-broth and re-plated on Mueller-Hinton agar plates with and without Tp to examine the stability of the plasmid.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Plasmid</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>W3110 trpA1417</td>
<td>pRBG156</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>pRBG177*</td>
<td></td>
</tr>
<tr>
<td>W3110 tna2</td>
<td>pRBG156</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>pRBG177</td>
<td>+</td>
</tr>
<tr>
<td>W3110 tna2 trpR</td>
<td>pRBG156</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>pRBG177</td>
<td>+</td>
</tr>
<tr>
<td>W3110 tna2 trpR recA::Tn10</td>
<td>pRBG156</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>pRBG177</td>
<td>--</td>
</tr>
<tr>
<td>AB1157</td>
<td>pRBG156</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>pRBG177</td>
<td>--</td>
</tr>
<tr>
<td>M76</td>
<td>pRBG156</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>pRBG177</td>
<td>--</td>
</tr>
<tr>
<td>RR1</td>
<td>pRBG156</td>
<td>-</td>
</tr>
<tr>
<td>HB101</td>
<td>pRBG156</td>
<td>-</td>
</tr>
<tr>
<td>GM119</td>
<td>pRBG156</td>
<td>-</td>
</tr>
<tr>
<td>C600</td>
<td>pRBG156</td>
<td>-</td>
</tr>
</tbody>
</table>

* pRBG177 is a derivative of pRBG156 which has had the trp P0 region has been removed. +: similar number of colonies on Mueller-Hinton plates with and without Tp. -: many more colonies on Mueller-Hinton plates without drug than on plates with Tp. --: no maintainable colonies were found after transformation.
over 40 generations (Figure 15). The stability of pRGB156 in the W3110 strains can readily be seen. While GM119 and RR1 do gradually lose pRGB156, the loss is not significant for at least 20 generations. pRGB156 in W3110 trpA1417 is stable over the entire time course of the experiment. The instability of pRGB156 in W3110 tna2 trpR is apparently due to the presence of the trpPO since removal of this region (pRGB177) restores the stability of the plasmid in this strain. Plasmid instability as a result of the trp promoter has been previously observed (Nichols and Yanofsky, 1983).

Studies of promoters present on pRGB156

The -35 region of the RNA I promoter is still present in pRGB156. The RNA I promoter is a strong promoter, as determined in the pkO system (Wong et al., 1982). In order to determine the presence or absence of a promoter which could fortuitously act to replace the RNAII promoter, RNA polymerase binding studies on pRGB156 were conducted. Plasmid DNA was cleaved with either CfoI or HinfI and added to a mix containing RNA polymerase. The samples were filtered through nitrocellulose filters. Any DNA fragments which bind RNA polymerase will be trapped on the filter.

As can be seen in Figure 16, the fragments from pRGB156 which contain the trp promoter (CfoI fragment 3 and HinfI fragment 3) are retained on the nitrocellulose
Figure 15. Plasmid Stability. Plasmids in different backgrounds were grown in L-broth and kept in exponential growth for at least 40 generations. Aliquots were diluted and plated at various times on Mueller-Hinton plates with and without Tp. The number surviving on the drug plates compared to the number on plates without drug gives the percentage Tp resistant colonies. The error bars are shown if a range in the percentage surviving was observed.

A) --- pRBC156 in GM119; --- pRBC156 in RR1; pRBC156 in W3110 trpA1417. B) --- pRBC156 in W3110 tna2; --- pRBC156 in W3110 tna2 tnaR; --- pRBC177 in W3110 tna2; --- pRBC177 in W3110 tna2 trpR.
Figure 16. RNA Polymerase Binding to pRBG156. After cleavage with a restriction enzyme and binding of RNA polymerase, the mixture was filtered through nitrocellulose filter. Flowthrough was collected on a DE81 cellulose filter. After elution, the DNA was electrophoresed on a 5% polyacrylamide gel. Lane 1 - CfoI digest of pRBG156 bound to nitrocellulose; Lane 2 - CfoI digest of pRBG156 flowthrough; Lane 3 - HinfI digest of pRBG156 bound to nitrocellulose; Lane 4 - HinfI digest of pRBG156 flowthrough; Lane 5 - HinfI digest of pBR322 bound to nitrocellulose; Lane 6 - HinfI digest of pBR322 flowthrough; Lane 7 - HinfI digest of pBR156; Lane 8 - HinfI digest of pBR322.
filter. The only other fragment from pRBG156 which binds RNA polymerase under these conditions is *Hinfl* fragment 2. Swift et al. (1981) attempted to locate the position required for DHFR expression by inserting linkers into pMT100 which would disrupt gene expression. They found two areas which resulted in loss of *Tp* resistance. One area was the structural gene itself, while the other appeared to contain the promoter. This region is found on *Hinfl* fragment 2. Zolg and Hanggi (1981) place the promoter much closer to the gene itself. No fragment from this area binds RNA polymerase under these conditions. This result supports the promoter location proposed by Swift et al. (1981), although this region is slightly modified in pRBG156 (Figure 14).

In a control binding experiment (Lanes 5 and 6), several fragments from pBR322 also bind RNA polymerase. *Hinfl* #2 contains the promoters for RNA I and RNA II. The conditions used were sufficient for RNA polymerase to bind these promoters in pBR322. The junction between the end of the DHFR gene and the -35 region of the RNA I promoter in pRBG156 is located on *CfoI* fragment 1 and *Hinfl* fragment 5. No promoter binding sites were found on these fragments of pRBG156. Under the conditions used, there does not appear to be any sequence in the origin that is acting as a promoter.
Plasmid incompatibility

When two plasmids can not be stably maintained together without supplying selective pressure for both, they are considered incompatible with one another (Staudenbauer, 1978). Since pRBG156 does not contain the region of the origin which determines incompatibility between plasmids (Tomizawa and Itoh, 1981), it was of interest to examine the compatibility of these derivatives with plasmids from the same ColE1 incompatibility group. pRBG156 and pBR322 were used in the incompatibility study since pBR322 could be selected with Ap while pRBG156 could be selected with Tp. W3110 trpA1417, the strain in which pRBG156 was shown to be stable, was used in all transformations. W3110 trp A1417 cells which already contained one plasmid were used as recipients in transformations with the other plasmid. This was done to examine any possible effects which might depend upon which plasmid was already present in the cell. After selection on double drug plates, colonies which contained both plasmids were grown overnight in L-broth and plated on Mueller-Hinton agar plates without drugs. The resultant colonies were replica plated on Mueller-Hinton agar plates with either Ap or Tp. The presence or absence of one plasmid gives an indication of the incompatibility properties of the plasmids.

The results (Table 4) show that when pRBG156 is the resident plasmid and pBR322 is the donor, there is no loss of either plasmid in overnight growth. Virtually all of the
### Table 4
Plasmid Incompatibility

<table>
<thead>
<tr>
<th>Resident Plasmid</th>
<th>Recipient Plasmid</th>
<th>Colonies Analyzed</th>
<th>( \theta ) of Ap(^R ) Colonies</th>
<th>( \theta ) of Tp(^R ) Colonies</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBR322</td>
<td>pRBG156</td>
<td>230</td>
<td>229</td>
<td>230</td>
</tr>
<tr>
<td>pRBG156</td>
<td>pBR322</td>
<td>196</td>
<td>196</td>
<td>195</td>
</tr>
</tbody>
</table>

W3100 \( \text{trpA1417} \) containing either pBR322 or pRBG156 was used as the resident plasmid in a transformation placing the other plasmid in the cell. After selection on Mueller-Hinton agar plates with Tp and Ap, several colonies were grown without selection in L-broth for 10-15 hours. The cells were then plated on Mueller-Hinton agar plates and the resultant colonies were replica plated on Mueller-Hinton plates with Tp or Ap to identify colonies containing pRBG156 or pBR322, respectively.
colonies that were replica plated had both plasmids. The same results were obtained when pBR322 was the resident plasmid and pRBG156 was the donor. Thus, pRBG156 is compatible with pBR322 and presumably any pBR322-derived plasmid.

**Plasmid copy number**

In order to determine the relative copy number of pRBG156, W3110 trpA1417 cells were grown with either pBR322, pRBG156 or both present. Aliquots were removed at two times: during exponential growth and after chloramphenicol amplification. The cells were lysed and the plasmid DNA was isolated. After cleavage of the DNA samples with EcoRI, the DNA was electrophoresed on 0.8% agarose gels. Cleavage with a restriction enzyme removes the complication of plasmid multimers and differences between supercoiled and relaxed forms of the plasmids. Figure 17 shows the results. In order to make sure that preparation artifacts were removed, the amount of DNA prepared from cells with both plasmids was compared with the amount of DNA from cells with only one plasmid present (Lanes 1 and 2) or the amount of DNA from cells with single plasmids which were mixed together prior to isolation of the plasmid DNA (Lane 3). The relative amounts of plasmid DNA present was quantified by cleaving the plasmid DNA with TaqI and electrophoresing the fragments on 5% polyacrylamide gels. A Polaroid photograph was taken of the ethidium bromide stained
Figure 17. Plasmid Copy Number. W3110 trpA1417 cells with pRBG156, pBR322 or both were grown in L-broth. Aliquots were removed in exponential growth and after over-night amplification. The plasmid DNA was prepared as detailed in the Methods. Aliquots of the DNA were cleaved with EcoRI and electrophoresed on a 0.8% agarose gel. Lane 1 - pBR322 (exponential growth); Lane 2 - pRBG156 (exponential growth); Lane 3 - Cells with pRBG156 mixed with cells containing pBR322 (exponential growth); Lane 4 - pRBG156 and pBR322 present in the same cell (exponential growth); Lane 5 - pBR322 (amplified); Lane 6 - pRBG156 (amplified); Lane 7 - Cells with pRBG156 mixed with cells containing pBR322 (amplified); Lane 8 - pRBG156 and pBR322 present in the same cell (amplified).
gel and the negative was scanned with a densitometer. TaqI was chosen as the restriction enzyme was so similar sized fragments could be compared between the two plasmids without producing overlapping sized fragments. Then, simple comparison of the intensity of two fragments on an ethidium stained gel gives an indication of relative copy number. These results are shown in Table 5.

As can be seen in Figure 17, the relative amount of each plasmid present in the cell is independent of the presence of the other plasmid. The relative amount of DNA isolated for each plasmid is the same whether the plasmids are both present in the same cell or whether only one plasmid is present in a cell and the cells are mixed prior to DNA isolation (Lanes 3 and 4). This also demonstrates the compatibility of the two plasmids since there is no loss of either plasmid. This experiment illustrates that pRBG156 and pBR322 amplify differently. In exponential growth, pRBG156 has a higher copy number than pBR322. After amplification, pBR322 increases 12-fold while pRBG156 copy number doubles. The reason for this is not clear.

Recombination between the two plasmids does not appear to be a major consequence of having both plasmids present in the same cell. Visualization of the plasmid DNAs from these plasmid preparations on agarose gels showed both plasmids to be present mostly as monomers (Figure 17). The lack of any large scale recombination is also shown by the copy number data. The relative amounts of the two plasmids
Table 5
Plasmid Copy Number

<table>
<thead>
<tr>
<th>Relative Copy Number</th>
<th>pBR322</th>
<th>pRBG156</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential growth</td>
<td>1</td>
<td>4.2</td>
</tr>
<tr>
<td>Amplified</td>
<td>12</td>
<td>7.6</td>
</tr>
</tbody>
</table>

After isolation of plasmid DNA as detailed in the Methods, the DNA was cleaved with Taq I and electrophoresed on a 5% polyacrylamide gel. A negative was made of the ethidium-stained gel and was scanned by a densitometer. Fragment bands were chosen which were unique for each plasmid and close in fragment size so that differences in intensities due to size differences were minimal.
change upon amplification but by very different amounts. The presence of a recombined hybrid plasmid would result in equal amounts of amplification since the two plasmids would be linked.

Preliminary studies indicate that pRBG177 may have a higher copy number than pRBG156. Interference between the *trp* promoter and a primer transcript is not present in pRBG177. This type of interference has been shown to lower copy number while placement of transcription terminators between the transcripts raises copy number (Stueber and Bujard, 1982).

**DISCUSSION**

A region upstream from the origin of replication controls both copy number and incompatibility in plasmids derived from ColE1 or pBR322. Several groups have investigated the minimum size fragment which will give stable replication. Backman *et al.* (1978) described a *FnuDII* fragment from pBR322 (from 570 bp upstream from the origin to 13 bp downstream) which was sufficient to give stable replication. This fragment contains all of RNA I and all but the promoter of RNA II. Other groups have constructed plasmids with similar origins to pRBG156 but they were reported to be unstable in the strains used while pRBG156 is stable in certain strains. Oka *et al.* (1979) constructed a plasmid, pA07, from ColE1 which contained a similar *HaeIII* origin fragment as in pRBG156 and included
a six kilobase fragment conferring ampicillin resistance. They found that it contained enough information for replication but was not as stable as ColE1. Kimura et al. (1982) details work with the same plasmid and although they describe it as stable, it appears that it is so only while drug is present. In an incompatibility experiment similar to the one used here, they transformed HB101 containing pML21 (a ColE1-derived kanamycin resistant plasmid) with pA07, bearing resistance to Ap. After selection on double drug plates and growth on rich media, they examined 100 colonies. Of these, only six had both plasmids and 94 had simply pML21. As soon as selective pressure was removed, pA07 seemed to be rapidly lost. pRBG156, on the other hand, is not rapidly lost in W3110. A similar incompatibility study with pRBG156 and pBR322 in W3110 trpA1417, resulted in 100% retention of both plasmids.

The plasmids pRBG156 and pRBG177 offer an unusual opportunity to explore requirements in plasmid replication. The region containing the origin of replication is smaller than previously determined to be capable of supporting stable replication. The stability of a plasmid considered here is determined by its ability to be retained in a cell when selection is removed. By this criterion, pRBG156 is unstable in most strains of E. coli K-12 examined. In several strains (AB1157, M76, and W3110 tna2 trpR rec A56::Tn10), the plasmid was apparently so unstable that viable colonies could not be selected after transformation,
although they could be readily transformed by pMT100. The plasmid was somewhat stable in RR1 and GM119, being retained in the cell for about 1.5 days before being lost. The W3110 strains (W3110 trpA1417, W3110 tna2) allowed pRBG156 to be stably replicated for well over forty generations. The difference in the retention of pRBG156 in the various strains indicates that the strain background has a definite contribution to plasmid stability. The reason for this strain difference is unknown, although the fact that the W3110 strains are wild type E. coli K-12 in essential genes might have some bearing.

The instability found for pRBG156 in W3110 tna2 trpR appears to be due to the presence of the trpPO region on the plasmid. Removal of this region, as in pRBG177, results in restored stability in W3110 tna2 trpR. Plasmid instability has been observed before in trpR strains when a trp promoter was on a CoLE1-derived plasmid (Nichols and Yanofsky, 1983). In pRBG156, the transcription from the trp promoter could interfere with plasmid replication. Stueber and Bujard (1982) stated that transcription into the origin towards the primer start lowers the copy number by interfering with plasmid replication. This is the direction of transcription from the trp promoter on pRBG156. pRBG177 has no trp promoter and appears to have a higher copy number from preliminary experiments. Plasmid stability and copy number differences in different strains have not been discussed or examined very
extensively.

pR Berg156 appears to occur in several-fold higher copy number in log phase cells than does pBR322. Since copy number is normally regulated by the presence of RNA I and since pR Berg156 does not contain RNA I, the copy number in pR Berg156 would seem to be solely determined by the amount of RNA II transcript and its relative ability to form a primer. pR Berg156 does not amplify to a large extent. The reasons for this are unknown. Another small plasmid, wVX, is derived from the FnuDII fragment in the origin. It does not contain the entire RNA II promoter and does not amplify well, while a derivative, pAN7 (which contains the usual RNA II promoter and the adjacent terminator) does amplify. Perhaps the normal RNA II promoter is necessary for amplification.

Replication of pR Berg156 most likely occurs because of readthrough transcription from the DHFR gene. In the normal ColE1 replicon, there is a palindrome centered 591 bp upstream from the origin which apparently terminates any transcription proceeding towards the origin. Itoh and Tomizawa (1980) showed that a β-lactamase transcript, transcribed off the same strand as RNA II, terminated near the palindrome. Mutations which affect the stability of the stem-and-loop structure formed by this palindrome allow the β-lactamase transcript to be extended past the normal termination site (Selzer et al., 1983). pR Berg156 does not contain this palindromic terminator, so presumably transcription can continue towards the origin. If it does, pri-
mer hybrids could be formed and processed by RNAase H, allowing plasmid replication to occur. Attempts to place a trp promoter fragment pointing in the same direction as the DHFR gene have failed. This result implies that too much transcription might result in an unstable plasmid. The amount of primer transcript required for stable plasmid replication may need to be maintained within a narrow range. The level of readthrough transcription from the DHFR gene in pRBG156 might allow sufficient primer to be formed so that stable replication can occur. Since there is no RNA I to inhibit primer formation, the level of transcription from the DHFR promoter alone may determine primer quantity. The possibility of altering or regulating the amount of transcription of the DHFR gene in these plasmids, and perhaps also the copy number, is being explored.

The only segments of pRBG156 which bind RNA polymerase under the conditions used are fragments containing the trpPO or the DHFR promoter. Swift et al. (1981) place the DHFR promoter 200 bp upstream from the start of the gene. Insertion of BamHI linkers in this area resulted in loss of trimethoprim resistance. Zolg and Hanggi (1981) place the promoter much closer to the amino terminus of the gene by sequence homologies to other promoters. The data presented demonstrate that a RNA polymerase binding site exists in the promoter region proposed by Swift et al. (1981). Although the sequence in this region is different in pRBG156 than in pMT100, it still acts to bind RNA poly-
merase. No binding site was observed in the region near the amino terminus of the gene. The promoter for DHFR seems to be located far upstream from the gene itself. There is an open reading frame in this region, but no function for it has been determined.

pRBG156 and its derivatives are unique vectors for DNA manipulations. These plasmids are quite small, which allows DNA fragments inserted into them to be easily analyzed by restriction mapping without the complicating presence of a large number of restriction fragments. Also cloning can be carried out using restriction enzyme sites not often available for cloning. The unique sites in the polylinker region, along with the few HaeIII sites, offer opportunities for cloning. The high copy number allows large amounts of plasmid to be purified. The incompatibility properties of pRBG156 and its derivatives means that two high copy number plasmids could be kept in one cell.
CHAPTER 3 Use of MboII Restriction Enzyme Sites for DNA Manipulations

INTRODUCTION

The use of restriction enzymes allows a wide range of DNA fragments and oligonucleotides to be manipulated in order to achieve a desired sequence. These manipulations might result in the isolation of a particular DNA sequence or in the insertion of a DNA fragment into a plasmid. While this results in many possibilities for sequence modifications, the very specificity of the restriction enzymes also limits the possible combinations which can be joined. In order to juxtapose two sequences, one must use restriction enzymes which give compatible sticky ends or use blunt-ended fragments. The former limits the general use of the restriction enzyme site as further additions are hampered. There are either two restriction enzyme sites present instead of the original unique site or the restriction site is no longer present. The use of blunt ended fragments usually destroys the restriction enzyme site, limiting further additions.

Scarpulla et al. (1982) detailed the use of a retrieving adaptor molecule. This synthetic adaptor contains an MboII recognition site and can be blunt-end ligated onto a DNA fragment. MboII recognizes a pentanucleotide sequence, GAAGA, and cleaves the DNA eight bases downstream, leaving a single 3' protruding end (Brown et al., 1980). MboII is one of a small group of restriction endonucleases
which have recognition sequences that are separated from the site of cleavage. Scarpulla et al. (1982) ligated this synthetic adaptor onto a DNA fragment. This DNA fragment had been cleaved with EcoRI and the sticky ends had been made flush with DNA polymerase I (Klenow fragment). After ligation of the retrieving adaptor molecule, cleavage with MboII removed several base-pairs from the ends, producing a DNA fragment which had no EcoRI ends. This fragment could now be used in further manipulations.

This chapter describes some similar MboII adaptor molecules and their use in a plasmid-borne system that allows virtually any two fragments to be joined. An example is shown in Figure 18. A sequence containing an MboII recognition site can be ligated onto another fragment and several DNA residues could be removed by cleavage with MboII. These could then be joined with another, similar sequence in order to juxtapose two sequences which could not be joined in any other fashion. These adaptor molecules were characterized by examining their ligation properties and by DNA sequencing methods. Several plasmids were constructed which allow a wide range of manipulations to be performed in addition to MboII cleavage. The advantage of plasmids in this scheme is the ease of plasmid isolation and the ability to amplify any cloned sequence.
Figure 18. Transfer of DNA Residues Using MboII sites. Ligation of an MboII recognition sequence, which is less than 8 bases upstream from the end of a fragment, onto another DNA fragment will position the MboII recognition sequence so that MboII cleavage will occur in the second fragment. This will remove several residues. Another similar reaction can be done with another fragment. The two DNA sequences which have MboII ends can now be joined if the ends are complementary. These reactions produce a new sequence whose existence does not depend on any pre-existing restriction enzyme sites.
RESULTS

Characterization of Adaptor Oligonucleotides

The basis for the MboII adaptor molecules are a group of oligonucleotides. The oligonucleotides used in this work are shown in Figure 19. They were constructed in the laboratory of Dr. P. T. Gilham at Purdue University. Some of these oligonucleotides have been characterized by 2-dimensional DNA sequencing (Figure 20; see Methods). The 5' end of an oligonucleotide was radioactively labeled with T4 polynucleotide kinase. The labeled oligonucleotide was partially digested with snake venom phosphodiesterase. The products are separated in the first dimension by electrophoresis in a pH 3.5 pyridine-acetate buffer and in the second by homochromotography. After autoradiography, the DNA sequence was determined by the shift in mobility of the digestion products, which can then be read off of the photographic plate. Only the 5' terminal residue can not be determined by this technique. The terminal residue can be determined by complete digestion of the 5' end-labeled oligonucleotides with snake venom phosphodiesterase, followed by electrophoresis in the pH 3.5 pyridine-acetate buffer. Using appropriate standards, the end residue can be determined by the distance traveled (Figure 21). The sequences could also be cloned (see below) and the presence of appropriate restriction enzyme sites in the oligonucleotides were used to confirm the DNA sequence. Some of the cloned oligonucleotides were sequenced by the chemical
Figure 19. Sequence of Synthetic Oligonucleotides. These oligonucleotides were synthesized in the lab of Dr. P.T. Gilham. Several have a 3' phosphate placed on them to prevent ligation at that end. The oligonucleotides can hybridize to form small double-stranded molecules. These small sequences have useful restriction enzyme sites which are detailed in the text.
Bam duplex II

5' AGAAGATCCG 3' oligo 5
3' pTCTTCTAGGC 5' oligo 3p

Bam duplex III

5' AGAAGAG 3' oligo 4
3' TCTTCTCCTAG 5' oligo 2

Bcl duplex II

5' AGAAGATCA 3' oligo 13
3' pTCTTCTAGT 5' oligo 15

Bcl duplex III

5' AGAAGAT 3' oligo 11
3' TCTTCTACTAG 5' oligo 16
Figure 20. Two-Dimensional Sequencing of Oligonucleotide 2. This oligonucleotide (see Figure 19) was labeled at its 5' end with T4 polynucleotide kinase. It was then digested with snake venom phosphodiesterase and aliquots were removed at various times. The products were spotted on a cellogel strip and electrophoresed in a pyridine-acetate (pH 3.5) buffer. The products were then transferred to a PEI cellulose plate and separated in the second dimension by homochromatography. The PEI plate was then used to expose a photographic plate. Each spot on the autoradiograph is one nucleotide longer and the shift in mobility is determined by which nucleotide is added. The sequence of oligonucleotide 2 is 3' TCTTCCTGTA. The terminal 5' residue cannot be determined in this procedure.
Figure 21. Identification of the 5' Terminal Residue. In order to determine the 5' end of the oligonucleotides, the 5' end-labeled oligonucleotides were digested to completion with snake venom phosphodiesterase. The products were spotted on Whatman 3M paper and electrophoresed with standards in a pyridine-acetate buffer (pH 3.5). The position of the standard was then determined by UV illumination and marked. Following autoradiography, the identity of the 5' terminal residue could be determined by comparison with the standard. In this example, the location of dCMP is noted. The oligonucleotides used, and the determined 5' residue, are: A - oligo 2 (G); B - oligo 3 (C); C - oligo 5 (A); D - oligo 4 (A).
scission method of Maxam and Gilbert (1978). None of the oligonucleotide sequences that were examined differed from the expected sequence.

Some of the oligonucleotides can hybridize with others to form DNA duplexes which have useful properties. Duplex II contains a cleavage site for MboI (GATC) between the recognition sequence for MboII (GAAGA) and its cleavage site (Figure 19). This presents the opportunity to dissociate the recognition sequence of MboII from its cleavage site by digesting the DNA with MboI. Bam Duplex II has a MboI site which contains the 3' side of a BamHI site (ie. GATCC) while BclI Duplex II contains the 3' side of a BclI site (GATCA). Cleavage of Bam Duplex II with MboI releases a DNA segment which has BamHI compatible ends. A BamHI site will be formed if this duplex is then joined to another fragment with BamHI sticky ends. An example would be Bam Duplex III which has the 5' side of a BamHI site present at its sticky end. The same description holds true for Bcl Duplex II except it will result in a BclI site being produced. The ability to separate the recognition site of MboII from its cleavage site means that sequences can be transferred from one fragment to another by ligating Duplex II onto a fragment, cleaving with MboII in order to remove several base pairs. Cleaving the released fragment with MboI allows a Duplex III to be ligated onto the sticky end. Under the right conditions, these bases can be joined with other such
fragments in order to build up a specific sequences of DNA (see Discussion).

Investigations into the restriction enzyme cleavage properties of Duplex II were made. One of the oligonucleotides (Oligo 3) of Bam Duplex II was phosphorylated with T4 polynucleotide kinase. This reaction was done to allow ligation to occur at only one end. The phosphorylated oligonucleotide was hybridized with the appropriate oligonucleotide (Oligo 5) to form Bam Duplex II. This duplex was then ligated to itself (Figure 22). If ligation was only occurring at the phosphorylated end, then only the dimeric form of Bam Duplex II should be formed. As can be seen, there are many more forms than this, indicating ligation at both ends of the duplex is occurring, even though only one end is phosphorylated. This makes the proper orientation of this duplex difficult to achieve. A SacII site will be formed if the proper ends ligate together. In Figure 22, most of the bands in the DNA ladder are still present after SacI digestion. In order for this procedure to be of use the MboII cleavage site must be pointing into a DNA fragment, not away from it. Ligation of the improper end of the Duplex II will complicate the technique. To overcome this, a 3' phosphate was left on one of the oligonucleotides (giving Oligo 3p for Bam Duplex II and Oligo 15 for Bcl Duplex II). This phosphate prevents any ligation at that end of the Duplex. A T4 polynucleotide kinase which has no 3' phosphatase activity was obtained
Figure 22. Ligation of Bam Duplex II with and without a 3' Phosphate. Bam Duplex II was formed with oligo 5 and either oligo 3 or oligo 3p. Only oligo 3 or 3p had a 5' phosphate. Oligo 3p contained a 3' phosphate which should hinder ligation at one end of the Duplex. The two Bam Duplexes were then ligated to themselves and the products were examined. If oligo 3p was used to form Bam Duplex II only the dimeric form of the Duplex is seen. A SacII site will only be formed if this Bam Duplex II ligates to itself about the 5' phosphorylated end. Cleavage with SacII does indeed digest the dimeric form of the Duplex. If oligo 3 is used, however, a ladder is seen made up of multimeric forms of the Duplex. SacII cleavage does not degrade most of these forms.
from NEN and used to phosphorylate the 5' end of the appropriate oligonucleotide. This enzyme had been isolated from a T4 mutant which did not have the endogenous 3' phosphatase activity (Richardson, 1965). Using this Bam Duplex II in a self-ligation reaction removes the multimeric forms (Figure 22). SacII digestion removes the major product of the reaction, which is presumably the dimeric form of the duplex. The 3' phosphate can be removed by bovine alkaline phosphatase or by using the same reaction that phosphorylates the 5' end but instead using a polynucleotide kinase which has a 3' phosphatase. This procedure allows a base-paired oligonucleotide to be properly orientated upon ligation.

Ligation of Duplex III did not present the same problem as did Duplex II. A radioactive 5' phosphate was placed on the sticky end of the Bam Duplex II by labeling Oligo 2, and the fragment was ligated to itself at the sticky end. Restriction digests were then performed in order to investigate the reaction. Figure 23 details the results. Only the dimeric form was formed in the ligation reaction. The dimeric form of Duplex III was cleaved with MboI, which produces the original duplex. Cleavage with XbaI, a site which will only be formed if the blunt ends of Bam Duplex III are joined, does not seem to cleave the dimer. Digestion with MboII produces an interesting pattern. Since both stands are labeled, there should be two fragments present after cleavage. On a denaturing gel, one
Figure 23. Ligation and Cleavage of Bam Duplex III.
Oligo 2 was labeled at the 5' end with T4 polynucleotide kinase. It was hybridized with oligo 4 to form Bam Duplex III. This duplex was ligated to itself and the products were examined on a 20% denaturing polyacrylamide gel. Only the dimeric form of the duplex is seen. MboI digestion restores the original oligonucleotide. There is no significant digestion with XbaI. The figure on the left details the digestion products found with MboII cleavage. The dimeric form of Bam Duplex III has opposing MboII sites which give rise to several bands.
Bam Duplex III

AGAAGAG
TCTTCTCCTAGp

lligation

MboII

AGAAGAGp*GATC CTCTTCT
TCTTCTC CTAGp*GAGAAGA

11 bases

MboII

18 bases-ligated
14 bases-single MboII cleavage
13 bases-single MboII cleavage
11 bases-unligated
9 bases-double MboII cleavage

UNLIG. LIG. MBO MBO XBA

I II I
will be 14 bases long and the other will be 13 bases long (Figure 23). The uncleaved dimeric form of the Duplex will be 18 bases and the unligated Duplex III will be 11 bases. As can be seen, there is a band present which is smaller than the original labeled oligonucleotide. This band apparently comes from the simultaneous cleavage of the dimer with MboII. In the dimer of Bam Duplex III there are two opposing MboII sites present. The cleavage site for each MboII site is in the recognition site of the other (Figure 23). Thus, if cleavage occurs from one site, it should prevent cleavage from the other site. However, a nine base fragment can be produced if cleavage occurs from both MboII sites after the enzyme has bound to both recognition sites. It seems likely that MboII can bind to the two recognition sites and cleave the DNA, even though in doing so it destroys both recognition sites for MboII.

Because cleavage with MboII leaves a single 3' overhanging base, it was of interest to determine that this base would be sufficient to allow proper hybridization of complementary sequences prior to ligation. This was investigated by ligating an eluted MboII fragment to itself (Figure 24). This fragment was isolated from pBR322. It had a single C residue at one end and a single G residue at the other so the ends were complementary. An internal BamHI site in the fragment allowed the orientation of the products to be determined. Proper ligation through the complementary ends, followed by digestion with BamHI, should
Figure 24. Ligation of MboII Fragments. An MboII fragment was isolated from pBR322. This fragment had complementary sticky ends, as shown in the figure on the left. There is an internal BamHI site in this fragment. This fragment was labeled with T4 polynucleotide kinase and ligated to itself. The length of ligation was either 6 hours or overnight. Following ligation, an aliquot was digested with BamHI. The products were electrophoresed on a 5% polyacrylamide gel and examined by autoradiography. As the figure on the left shows, ligation through the complementary ends will produce a 392 bp and 100 bp fragment, in addition to a 492 bp (which is the size of the original fragment). Aberrant ligation will result in either a 200 bp or a 792 bp fragment following BamHI cleavage. The control lane is unligated fragment before and after cleavage with BamHI. The sizes of the fragments are shown.
Ligation of Pbo II ends

1. Correct Base Pairing

2. Incorrect Base Pairing

Ligated Forms

BamHI FRAGMENTS

- 492

- 392

- 100
produce two fragments of 392 and 100 bp. Ligation through noncomplementary ends, followed by BamHI cleavage, will produce additional fragments of 784 or 200 bp (Figure 24). As can be seen, the major fragments formed are those expected if ligation occurred through complementary ends. Ligation through noncomplementary ends is not very frequent.

Use of MboII sites in a Plasmid-Borne System

Because MboII cleaves DNA in such a unique fashion, one can insert virtually any restriction enzyme fragments into the cleavage site without disrupting the recognition site. With most other restriction enzymes, insertion of a DNA fragment either destroys the restriction enzyme site or results in a duplication of the restriction enzyme site. This does not occur when MboII is used. Even after inserting a DNA fragment, MboII will still cleave the DNA because the recognition sequence is still located upstream from the site of insertion and there is no duplication of the MboII site. It is then possible to insert further sequences by cleaving with MboII, and repeating the procedure. This presents the opportunity to manipulate DNA sequences without the need to have appropriate restriction enzyme sites present. This can be illustrated by a plasmid-borne system that has been developed.

The plasmids used for this scheme must have a single MboII cleavage site, so that digestion with MboII will
not degrade the entire plasmid. Most commonly exploited plasmids contain several MboII sites. In order to overcome this, a plasmid was constructed which contained no MboII sites. These could then be added to the plasmid. The details of this construction are described in the previous chapter. Briefly, the region around the origin of replication in pBR322 does not contain a MboII site and is capable of providing for plasmid replication. A fragment from pMT100 contains a trimethoprim-resistant DHFR gene whose DNA sequence also lacks any MboII site (Swift et al., 1981). These two regions were ligated together to produce pRBG57. A polylinker region containing several unique restriction enzyme sites along with the tryptophan promoter-operator sequence (Russell, 1983) was then inserted into the HindII site of pRBG57 to yield pRBG156. This plasmid was used for all the other manipulations.

Two different methods were used to place a single MboII site in the plasmid. The simplest method involved inserting a DNA fragment which contained one MboII site into pRBG156 (Figure 25). pKK3535 was used as the source for fragments since its size allows a wide variety of suitable fragments to be used (Brosius et al., 1981). pKK3535 was digested with Sau3A (an isoschizomer of MboI) while pRBG156 was cut with BamHI. The DNA from each digest was mixed and then ligated together. The ends generated by a Sau3A digest are compatible with those left by BamHI cleavage. Following ligation, the DNA was di-
Figure 25. Construction of pRBG150 and pRBG151. pRBG156 was cleaved with BamHI. The resulting fragment has Sau3A ends. pKK3535 was digested with Sau3A and the fragments were inserted into the BamHI-cleaved pBR322. This will reform the BamHI site in many cases. The fragments of interest will not reform the BamHI site, so following ligation of the Sau3A fragments, a BamHI digestion was performed. This will linearize any plasmids with inserts that reform the BamHI site. After transformation, several recombinant plasmids were found which had MboII sites. pRBG150 has a 147 bp insert with a single MboII site. pRBG145 has two MboII sites. One is contained on aRsaI-SmaI fragment. This fragment was isolated and inserted into the SmaI site of pRBG156 to produce pRBG151. This plasmid has a 178 bp insert than contains a single MboII site.
a) cleave pRBGI56 with BamHI
b) insert Sau3A fragments from pKK3535
c) cleave with BamHI

a) isolate Smal-RegI fragment from inserted DNA
b) insert into Smal-cleaved pRBGI56

178 bp insert
gested with BamHI, since the fragments of interest from pKK3535 will destroy the BamHI site if they are inserted. The BamHI cleavage step will linearize any plasmid which contains a BamHI site, such as ligated vector, uncut plasmid, or any plasmid with an inserted fragment that produces a BamHI site. These linear molecules transform cells at a very low frequency. Plasmid DNA of the recombinant colonies was examined and some was found to have inserts. Because the entire sequence of pKK3535 is known, it was possible to map the inserts and identify the plasmids of interest. One of these, pRBG150, had a 147 bp insert which contained a single MboII site. This fragment is from 6955 to 7102, by the numbering of Brosius et al. (1981). The orientation of this insert was determined by CfoI-MboII double digests.

Another isolated plasmid, pRBG145, had a much larger 423 bp insert which contained two MboII sites (Figure 25). A smaller SmaI-RsaI fragment from this insert had a single MboII site. This SmaI-RsaI fragment was isolated from a 5% polyacrylamide gel (see Methods) and inserted into SmaI-cleaved pRBG156. This will reform the SmaI on only one side. Several recombinant plasmids were found to have this 178 bp fragment inserted, as judged by restriction enzyme mapping. pRBG151 contained a single MboII site pointing towards the SmaI site, in a similar orientation as pRBG150 (Figure 25). This insert contains the region from 1950 to 2128 of pKK3535 (Brosius et
al., 1981).

One drawback of pRBG150 and pRBG151 is the distance between the MboII and the polylinker region of the plasmid. Any fragment inserted at the MboII site will be fairly distant from any useful restriction enzyme site which might be needed and will also be joined to a large segment of unnecessary DNA. Consequently, there will be a large sequence of DNA carried along on any construction. To overcome this, a second series of plasmids were constructed which are more universal in their applicability.

The basis for their construction is the synthetic duplexes discussed earlier. These duplexes were cloned into plasmids and fragments from these plasmids were used in order to produce a vector with a single MboII site very close to the polylinker region. Figure 26 details the construction of the plasmids which are needed for these fragments. Bcl Duplex III was ligated to itself using the cohesive sticky end to form a dimeric form of the Duplex. This dimeric form was inserted into SmaI-cleaved pRBG156, followed by SmaI digestion. Insertion of the sequence will destroy the SmaI site, so this latter step will linearize any plasmid which still has a SmaI site. After plasmid DNA isolation, one of the plasmids, pRBG160, was found to contain the proper insert. pRBG160 has a BclI site between two opposing MboII sites (Figure 26). The presence of the BclI site was confirmed by cleavage of pRBG160 with both AvaI and BclI (Figure 27).
Figure 26. Construction of pRBG157, pRBG160, pRBG140 and pRBG159. Bcl Duplex III was ligated to itself using the BclI sticky ends and inserted into SmaI-cleaved pRBG156. The resulting plasmid, pRBG160, has a BclI site between two MboII sites (1). pRBG156 was cleaved with SmaI and a BclI linker (2) was inserted to produce pRBG147. This plasmid has both a BamHI site and a BclI site. pRBG140 was constructed by cleaving pRBG156 with BamHI and filling in the sticky ends with DNA polymerase I (Klenow fragment). Recircularizing the plasmid results in a vector that has no BamHI site (pRBG140). Bam Duplex III was ligated to itself using the sticky ends and inserted into SmaI-cleaved pRBG140. The resulting plasmid, pRBG159, has a BamHI site between two MboII sites (3).
**pRBGI40**

- Cleaved with *SmaI*
- *Sac3A* insert (3)

**pRBGI59**

- Cleaved with *SmaI*
- *Sac3A* inserts
- Cleaved with *BamHI* & *MboII*

**pRBGI56**

- Cleaved with *SmaI*

**pRBGI57**

- Cleaved with *BamHI*

**pRBGI60**

- Cleaved with *BamHI* & *SacII* & *MboII*

**Inserted Sequences**

1. AGAAGATGATCATCCTTCTTCTTCTACTAGTAGAAGA
   - *BclI* Duplex III
   - Self-ligated

2. CTGATCAAGACTAGT
   - *Bol* linker

3. AGAAGAGGATCCTCTTCTTCTCTCCTACCTAGGAGAAGA
   - *BamHI* Duplex III
A construction similar to pRBG160 was also performed with Bam Duplex III. In order to do this, however, the BamHI already present on pRBG156 had to be destroyed. This was done by cleaving pRBG156 with BamHI and filling in the sticky ends with DNA polymerase I (Klenow fragment). Circularization of these plasmids, followed by BamHI cleavage inorder to linearize vectors which still have a BamHI site, produced pRBG140. CfoI-BamHI double digests demonstrated the loss of the BamHI site. With the BamHI site inactivated, the dimeric form of Bam Duplex III was inserted into SmaI-cleaved pRBG140. Examination of the recombinant plasmids produced pRBG159 (Figure 26). This plasmid has a BamHI site present between 2 opposing MboII sites. These sites were confirmed by restriction enzyme analysis.

One more plasmid was needed in order to produce a useful vector which contains a single MboII site. This construction involved the placement of a BclI and a BamHI site on a plasmid, without the presence of any MboII sites. pRBG156 was chosen as the starting vector, as it already had a BamHI site, and did not have any MboII sites. This plasmid was cleaved with SmaI and a small linker molecule was inserted (Figure 26). This linker molecule destroyed the SmaI site and replaced it with a BclI site. Following transformation and plasmid DNA isolation, a plasmid which contained the BclI site and lacked a SmaI site was isolated. This plasmid, pRBG157,
was characterized by restriction enzyme mapping in order to confirm the presence of both a BclI and a BamHI site.

These three plasmids (pRBG157, pRBG159, and pRBG160) could be used in the construction of a vector with single MboII site. However, the close proximity of the BclI and and BamHI sites in pRBG157 hampered the construction. In order to alleviate this, a small fragment was inserted between these two sites (Figure 27). BspRI fragment 12 from pBR322 was isolated and cleaved with Sau3A. This releases a small 23 bp fragment which has a BamHI compatible end and a BclI compatible end. This fragment was inserted between the BamHI and BclI site of pRBG157. This produced pRBG211, which was characterized by restriction enzyme mapping (Figure 27).

The constructions used to produce a vector with a single MboII site resulted in two plasmids which had the MboII sites in different orientations. The details of these constructions are in Figure 28. pRBG159 was cleaved with AvaI, BamHI, and HindIII. A large DNA fragment is released with a single MboII site and possessing BamHI and AvaI ends. pRBG211 was cleaved with AvaI, BamHI and BspRI. This produces a small DNA fragment with AvaI and BamHI ends. These two fragments were joined by T4 DNA ligase and transformed into bacterial cells. pRBG380 was one of the resulting plasmids (Figure 28). It possesses an MboII site adjacent to the BamHI site, as well as a BclI site. In fact, the plasmid is
Figure 27. Construction and Restriction Mapping of pRBG211. pRBG157 was cleaved with both BamHI and BclI. This leaves Sau3A compatible ends. BspRI fragment 12 from pBR322 was cleaved with Sau3A, which releases a small 23 bp fragment. This fragment was inserted into the cleaved pRBG157 resulting in pRBG211. The fragment reformed both the BclI site and the BamHI site. The insert occurs in the 197 bp Cfo fragment 3 of pRBG157. An AvaI-BclI digest is shown in lane 4 for size comparison. The restriction enzymes used are in parentheses. Lane 1 - pRBG157(CfoI); Lane 2 - pRBG211 (CfoI); Lane 3 - pRBG211 (CfoI and BclI); Lane 4 - pRBG160 (AvaI and BclI); Lane 5 - pRBG211 (AvaI and BamHI).
virtually identical with pRBG211 except for the presence of the MboII site. Similarly, pRBG211 was cleaved with AvaI, BclI and HindIII. A DNA fragment was released with AvaI and BclI ends. pRBG160 was cleaved with AvaI, BclI and BspRI. A small DNA fragment was produced which contains an MboII site. Ligation of these two fragments resulted in the construction of pRBG381 (Figure 28). This plasmid contains a single MboII site adjacent to a BclI site. In pRBG380, the MboII site faces the DHFR gene, while in pRBG381, the MboII site points away from the DHFR gene.

The presence of a single MboII site on a plasmid permits the insertion of a DNA fragment into the plasmid at the MboII cleavage site without disrupting the MboII recognition site. The insertion into MboII-cleaved vector was demonstrated by using some BspRI DNA fragments. BspRI fragments 6 and 11 were isolated from pBR322. pRBG380 was cleaved with MboII and the ends were made flush with T4 DNA polymerase (see Methods). BspRI fragment 6 was added to the digestion mix and ligated into the vector. Following transformation, plasmid DNA was isolated and examined. Several plasmids had inserts which corresponded in size to BspRI fragment 6. The orientation of one of these plasmids, pRBG306, was determined by restriction enzyme mapping with CfoI and MspI. In a similar reaction, pRBG311 was produced by the insertion of BspRI fragment 11 and the orientation of the insert was deter-
Figure 28. Construction of pRBG380 and pRBG381. pRBG159 was cleaved with AvaI, BamHI and HindIII. A large DNA fragment with BamHI and AvaI ends is released. pRBG211 was cleaved with AvaI, BamHI, and BspRI. A small DNA fragment with AvaI and BamHI ends is released. These two DNA fragments were joined together to construct pRBG380. This plasmid has an MboII site pointing towards the BclI site. Similarly, pRBG211 was digested with AvaI, BclI and HindIII. A large fragment is released with AvaI and BclI ends. pRBG160 was cleaved with AvaI, BclI and BspRI. A small fragment is produced with AvaI and BclI ends. These two fragments were joined to form pRBG381. This plasmid has an MboII site pointing towards the BamHI site. B - BspRI; H - HindIII
A - AvaI; Bm - BamHI; Bc - BclI; MboII.
mined by restriction enzyme mapping (Figure 29).

The construction of pRBG306 and pRBG311 demonstrates the ability to juxtapose two DNA sequences using \textit{MboII}. Both of these plasmids have a 23 bp DNA sequence from pBR322, which is not produced by normal restriction enzyme cleavage, joined to a \textit{BspRI} fragment. Any fragment which can be made blunt-ended can be inserted into pRBG380 or pRBG381. The main difference between these two vectors is that the \textit{MboII} sites will be located on different sides of the insert. This could be useful in a scheme which will produce a specific sequence of DNA (see Discussion).

\textbf{DISCUSSION}

This chapter details the use of some synthetic DNA duplexes in the construction of several plasmids. The oligonucleotides which make up these duplexes were characterized by sequencing the DNA and by examining the restriction enzyme sites present. A 3'-phosphate was placed on some of these oligonucleotides in order to prevent improper ligation products. Usually, synthetic restriction enzyme linkers use only one oligonucleotide. This oligonucleotide will hybridize to itself and form a double-stranded linker that contains the restriction enzyme site. Since both ends of this linker are the same, it does not matter which end of the molecule is ligated onto the end of a DNA fragment. Because the \textit{MboII} recognition site is not palindromic, this approach will not work. The ends of each Duplex are
Figure 29. Insertion of DNA Fragments into pRBG380 and pRBG381. pRBG380 was cleaved with MboII and the ends were made flush with T4 DNA polymerase. BspRI fragment 6 from pBR322 was isolated and inserted into the MboII-cleaved plasmid. Following transformation, several plasmids were found which had the fragment inserted in either orientation. One of these pRBG306 was mapped further. Similarly, pRBG381 was cleaved with MboII and the ends were made blunt by T4 DNA polymerase. BspRI fragment 11 from pBR322 was inserted. Several recombinant plasmids were isolated and one, pRBG311, was chosen for further characterization.
1) cleave with MboII
2) make ends flush
T4 DNA polymerase

insert BsoRI fragment
#6 from pBR322

pRBG380

pRBG306

pRBG381

insert BsoRI fragment
#11 from pBR322

pRBG311

B BamHI
M MboII
Bc BstI

267 bp

124 bp
different. The use of the 3'-phosphate prohibits any ligation at the end of the Duplex which contains it. This makes sure that the Duplex will be orientated properly.

The plasmids constructed with these synthetic Duplexes have demonstrated the ability to join DNA fragments from different sources together without the need for convenient restriction enzyme sites to be present on the DNA fragments. The plasmids that have been constructed can also be used in a scheme which will produce virtually any sequence required. Instead of simply adding large restriction fragments to a growing sequence, small codon triplets could be added sequentially. This is presented in Figure 30. This procedure will effectively join codon triplets together in order to produce the desired sequence. To an isolated blunt-ended DNA fragment is ligated a Duplex II (Figure 30, step 1). This produces a fragment with MboII sites at either end. The Duplex II used will only be joined with the blunt-ended fragment so that the MboII cleavage sites will occur in the fragment itself. The choice of the Bam or Bcl Duplex II will depend upon which cleavage (BamHI or BclI) will be done at step 4. One of the constructed plasmids (either pRPG150 or pRPG151) is then cleaved with SmaI and the DNA fragment with a Duplex II on each end is inserted, giving the plasmid in step 2. This plasmid has three MboII sites, one from the original plasmid and two from Duplex II. The blunt-ended fragment used is chosen so that the protruding ends following MboII cleavage will be
different. The use of the 3'-phosphate prohibits any ligation at the end of the duplex which contains it. This makes sure that the duplex will be orientated properly.

The plasmids constructed with these synthetic duplexes have demonstrated the ability to join DNA fragments from different sources together without the need for convenient restriction enzyme sites to be present on the DNA fragments. The plasmids that have been constructed can also be used in a scheme which will produce virtually any sequence required. Instead of simply adding large restriction fragments to a growing sequence, small codon triplets could be added sequentially (Figure 30). This procedure will effectively join codon triplets together in order to produce the desired sequence. To an isolated blunt-ended DNA fragment is ligated a Duplex II (Figure 30, step 1). A DNA fragment with MboII sites at either end is produced. The Duplex II used will only be joined with the blunt-ended fragment so that the MboII cleavage sites will occur in the fragment itself. The choice of the Bam or Bcl Duplex II will depend upon which cleavage (BamHI or BclI) will be done at step 4. One of the constructed plasmids (either pRBG150 or pRBG151) is then cleaved with SmaI, and the DNA fragment with a Duplex II on each end is inserted, giving the plasmid in step 2. This plasmid has three MboII sites, one from the original plasmid and two from Duplex II. The blunt-ended fragment used is chosen so that the protruding ends following MboII cleavage will be
Figure 30. Construction of a Specific Sequence of DNA Using Restriction Enzymes. *Bam* or *Bcl* Duplex II is ligated onto the ends of a blunt-ended fragment which is then inserted into the SmaI site of a plasmid with only one MboII site (step 1). An MboII reduction can now be performed (step 2). This produces a plasmid with two opposing MboII sites, along with a three base pair segment from the blunt-ended fragment. This region can be removed by MboI cleavage (step 3). This fragment can be inserted into a plasmid with either a BamHI or BclI site (step 4), depending upon which Duplex II was used in step 1. This results in a plasmid with a single MboII site, which will cleave adjacent to the three base pair segment. This plasmid can be cleaved with MboII and joined to another MboII-cleaved plasmid which was similarly constructed (step 5). This produces a long linear DNA sequence, with the three base pair segments of each plasmid joined. Cleavage of this DNA sequence with AvaI, followed by a ligation reaction in order to circularize the DNA, will produce the plasmid in step 6. This plasmid has no MboII sites and now has joined six base pairs of DNA which derive from two different blunt-ended fragments.
complementary. Cleavage with MboII will now result in an MboII reduction yielding the plasmid in step 3. This plasmid has opposing MboII sites, but they cleave at the same point. A small, three base pair segment of the original blunt-ended fragment will now be present adjacent to the point of MboII cleavage. There are MboI sites present which will release a small fragment which contains a single MboII site, along with the three base pair segment, or codon (Figure 30, step 3). One of these MboI sequences derives from the Duplex II used in step 1. This MboI end will be compatible with BamHI or BclI, depending on which Duplex II was used. The isolated MboI fragment can then be inserted into the appropriate plasmid in step 4. The plasmid with the BamHI is pRBG156 while the one with the BclI site is pRBG157. The orientations of the inserted fragments are opposite in the two groups of plasmids produced (step 5). These orientations will be advantageous in later reactions.

Several parallel reactions can be performed in order to produce several vectors of the type shown in step 5. All of these plasmids will contain different sets of three base pair segments which will now be joined together in order to produce a DNA sequence. Cleavage with MboII can linearize each plasmid in step 5. The plasmids chosen for this step will have complementary MboII ends. Following ligation, a long segment of DNA will be produced which has the two codons joined together. Cleavage with MboII and
AvaI will degrade the original vectors into several fragments, but the long segment of DNA produced by the ligation of the two codons will have AvaI ends which can be joined to produce a viable plasmid. This final plasmid shown in step 6 now has a six base pair segment of two codons. It is similar to pRBG211 and MboII can also be placed on either side, as was done to produce pRBG380 and pRBG381. Further juxtapositions of codon triplets can now be performed.

The plasmids presented in this chapter also allow one to join virtually any sequence which can be isolated or synthesized to any other. Normally in the chemical synthesis of a large gene sequence, a series of overlapping single-stranded oligonucleotides are produced. These will hybridize together to give the desired double-stranded sequence; however, often the gene sequence is not exactly as in the wild type gene (Edge et al., 1981). The DNA sequence will be slightly different because of the need to reduce aberrant hybridization of similar oligonucleotides. Using one of these plasmids which has a single MboII site, one could assemble a synthetic gene without the need to produce overlapping sequences (Figure 31). The chemically synthesized sequence is produced so that two complementary oligonucleotides will hybridize leaving single 3' sticky ends. If these ends are complementary to those left by cleavage of pRBG380 with MboII, then the synthetic sequence could easily be inserted. Complementary MboII ends have been
Figure 31. Insertion of Synthetic DNA Sequences into MboII Cleavage Sites. Synthetic oligonucleotides can be hybridized together in order to construct a specific sequence of DNA. The oligonucleotides are produced so that they will be completely complementary, leaving single 3' protruding bases. A DNA sequence is cleaved with MboII. The DNA sequence used will give ends which are complementary to the ends of the synthetic DNA. The synthetic DNA can then be inserted and it will be able to be inserted in only one direction. Following isolation of the recombinant DNA, another synthetic DNA sequence, constructed in the same manner as the first, can be inserted. This can be repeated until the entire DNA sequence that is desired has been constructed.
12345678
GAAGANNNNNNNNnnnnnn
CTTCTNNNNNNNNnnnnnn
cleave with MboII

12345678
GAAGANNNNNNNNnnnnnn
CTTCTNNNNNNNNnnnnnn
hybridize synthetic
oligonucleotides and insert

12345678
GAAGANNNNNNNNnnnnnn
CTTCTNNNNNNNNnnnnnn
cleave with MboII

12345678
GAAGANNNNNNNNnnnnnn
CTTCTNNNNNNNNnnnnnn
insert synthetic
oligonucleotides

12345678
GAAGANNNNNNNNnnnnnn
CTTCTNNNNNNNNnnnnnn
repeat ...
shown in this chapter to be sufficient for proper ligation. The complementary ends will properly orient the synthetic nucleotide. After transformation and selection, another synthetic duplex can then be inserted into the recombinant plasmid by the same procedure. It would then be placed adjacent to the previous synthetic sequence and in the proper orientation. Thus, a large gene can very rapidly be constructed, with intermediate sequences also being cloned which could be useful in other constructions. There would not be any need to worry about improper hybridization of any of the synthetic oligonucleotides since no more than two of the oligonucleotides need be placed together at any time. Other restriction enzymes than MboII, such as FokI, which leave a larger sticky end, could also be used in a similar type scheme. This mode of digestion might be useful if longer protruding ends are needed.
CHAPTER 4 Use of MboII "cassettes" for DNA manipulation

INTRODUCTION

Useful manipulation of DNA often requires a restriction enzyme site at or near the sequence of choice. While there is often some type of restriction enzyme site present, it is usually not exactly situated for optimum use nor does it usually give compatible ends after cleavage. If there is no restriction enzyme site present at all, it is often difficult to engineer a suitable site. Several techniques have been used to overcome this by inserting DNA sequences which allow further manipulations to be performed.

Inactivation of a gene by insertion of a transposable sequence (transposons) into the gene has been used to examine sequences necessary for gene function. These transposons have selectable markers which allow easy isolation of prospective recombinants. This technique was used to identify regions of pSC101 which are necessary for replication function, as well as mapping gene function by inactivation (Churchward et al., 1983). The transposable sequences often have restriction enzyme sites but they are not generally useful for further manipulations.

A recent technique can overcome this disadvantage. A series of restriction enzyme sites called a restriction site mobilizing element (RSM) has been constructed and placed on a plasmid (Vieira and Messing, 1982). This RSM element has an aminoglycoside 3'-phosphotransferase gene
(APH) which confers resistance to kanamycin to the host cell. On either side of the APH gene are a symmetrical series of restriction enzyme sites. The entire RSM element can be removed from the plasmid by restriction enzyme digestion. Thus the gene and its surrounding restriction enzyme sites can be inserted into other DNA sequences, resulting in the introduction of kanamycin resistance to the bacterial cells. After isolation of the kanamycin resistance recombinant plasmids, subsequent restriction enzyme digests can remove the APH gene, leaving a small DNA segment containing several unique restriction sites behind it. This technique allows sequences to be inserted into genes which will inactivate the gene. The selectable phenotype of the APH protein produced by the RSM allows recombinants to be easily isolated. Subsequently, the restriction enzyme sites of the RSM element can be used for further DNA manipulations. The presence of the restriction enzyme sites does allow for more varied manipulations than a transposon but the sites can not be easily removed.

Sequences can be randomly inserted into a DNA sequence by the use of the nuclease, DNAase I. DNAase I, in the presence of manganese, will cleave both strands of the DNA molecule (Melgar and Goldthwaite, 1968). This cleavage does not leave blunt ends so T4 DNA polymerase, in the presence of all four dNTPs, is used to make the ends flush. Restriction enzyme linkers can then be inserted into the cleaved DNA, thus producing a useful restriction enzyme
site in the DNA. Conditions can be found which will produce approximately one double-stranded cleavage by DNase I per DNA molecule. This technique has been used to examine the structure of several genes and does often place a needed restriction enzyme site close to an area of interest (Heffron and McCarthy, 1978). Again, inactivation of a particular gene is often used to identify recombinant plasmids which contain an insert. However, gene inactivation makes it difficult for further manipulations which require a functional gene.

The disadvantage of all these techniques is that they leave sequences in the region of interest which can not be removed. Insertion often inactivates a gene sequence and limits further modifications. In order to overcome this drawback, a technique has been developed which is based upon the cleavage properties of one of an unusual group of restriction enzymes. Most restriction enzymes recognize a specific, palindromic DNA sequence and cleave both strands of the DNA internal to this sequence. A group of Type III restriction enzymes do not recognize a palindromic sequence. They recognize a pentanucleotide DNA sequence and, instead of cleaving in the sequence, they cleave the DNA downstream from the recognition site. The distance downstream varies with the enzyme. FokI cleaves the DNA thirteen bases downstream, leaving a four base 5' sticky end (Sugisaki and Kanazawa, 1981). Another restriction enzyme of this type, HgaI, has been used for in vitro recombination
experiments because of its unusual cleavage properties (Moses and Horiushi, 1979). The technique described in this chapter uses the restriction enzyme MboII. This enzyme recognizes a non-palindromic pentanucleotide sequence (GAAGA) and cleaves the DNA eight bases downstream, leaving a single 3' protruding base (Brown et al., 1980). In contrast to most other restriction enzymes, MboII has a recognition site which can be separated from the cleavage site. Because of this, the enzyme can cleave at a position which does not have to be in a specific sequence.

Using the unusual cleaving properties of MboII, a restriction enzyme "cassette" has been developed. This cassette can be inserted into any DNA sequence and can be completely removed at any time. The cassette has a genetic marker that is easily screened between two MboII recognition sites (Figure 31). The MboII sites point away from each other. In this procedure, DNAase I is used to randomly cleave the DNA, and the cassette is then inserted. The inserted sequence has several restriction enzyme sites, along with the lac operator (lac o) sequence (Heyneker et al., 1976). Growth of the appropriate recipient cells on agar plates containing X-gal makes identification of the lac o bearing recombinants possible. Plasmids which contain the lac o sequence will cause the lac repressor to dissociate from the chromosomal lac operon and bind to the plasmid-borne lac o sequence. The lac operon can now be expressed and lactose which has diffused into the
cell can be degraded. X-gal is a lactose derivative which will form a bright blue dye when it is hydrolyzed by β-galactosidase. Thus, cells which can degrade X-gal will become dark blue and only cells with a plasmid-borne lac o will be able to hydrolyze the X-gal.

Insertion of this lac o-MboII cassette produces similar results as the insertion of an RSM element. The same kind of manipulations can be performed. The lac o sequence can be removed, leaving behind a segment of DNA with several restriction enzyme sites. Where this technique differs is that the presence of the MboII recognition sites allows the entire cassette to be removed.

In addition to removal of the cassette, the placement of the MboII recognition sites make it possible to also remove several base-pairs from the DNA surrounding the cassette. The cassette can be isolated in several different sizes (Figure 32). The distance of the MboII recognition sites from the end of the cassette is offset by several base-pairs in each form. Using different sizes of cassettes, varying numbers of base-pairs can be removed from the region surrounding the site at which the cassette is inserted (Figure 33). Because the cleavage site can extend outward into the surrounding region, one can proceed with a variety of further manipulations in regions of DNA which are not accessible by other techniques.
Figure 31. Cleavage Patterns of Different Forms of the lac o-MboII cassette. The sequence of the cassette is shown, along with the restriction enzyme sites present. Depending on the exact method in which the cassette is isolated, the MboII recognition sequences are different distances from the ends of the cassette. This means that the cleavage site of MboII will extend different distances into the surrounding DNA. The numbers under each sequence are the number of residues between the recognition sequence for MboII and the end of the cassette. Only one end of the cassette is shown in 1 through 4 for clarity.
MboII-lac o cassette

BamHI    XbaI    EcoRI    MboII
GTGATCCTTCTCTAGAATTG lac operator GAATTCTAGAAGGATCTCT
CACCTAGGAGAAGATCCTAG CTAGACCTCTCTCTAGAG
MboII    EcoRI    XbaI    BamHI
87654321

1) Cleave with MboII. Make ends flush with T4 DNA polymerase.

TGGATCCTCTCTCTAGAATTG lac operator 0 0
ACCTAGGAGAAGATCCTAG
765 4 3 2 1

2) Cleave with BamHI. Fill-in ends with DNA polymerase I (Klenow fragment).

GATCCTCTCTCTAGAATTG lac operator 0 0
CTAGGAGAAGATCCTAG
54 3 2 1

3) Cleave with DpnI.

TCCTCTCTCTAGAATTG lac operator 0 0
AGGAGAAGATCCTAG
3 2 1

4. Cleave with BamHI. Digest with S1 nuclease.

CTCTCTCTAGAATTG lac operator 0 0
GAGAAGATCCTAG
1
Figure 33. Use of the MboII-lac o Cassette for the Creation of Specific Deletions. The manner in which the cassette is isolated will change the number of bases that the MboII cleavage site will occur in the surrounding DNA. Isolation of the cassette by cleavage with MboII will allow 2 base pairs to be removed. This can be done by inserting the cassette and cleaving with MboII. Isolation of the MboII-lac o cassette by digestion with BamHI, followed by S1 nuclease digestion to remove the sticky end, will release a fragment that can remove 14 base pairs.
1) Cassette isolated by cleavage with MboII, followed by treatment with T4 DNA polymerase to make the ends flush.

\[
\begin{align*}
&\text{1234567} \\
&\text{GAATTCTAGAAAGAGATCC} \text{CTTAAGATCCTTCCTAGG} \text{ANNNNNNNNN} \\
&\text{CTTAAGATCCTTCCTAGG} \text{ANNNNNNNNN} \\
&\text{123} \\
&1). \text{ cleavage with MboII} \\
&2). \text{ treatment with T4 DNA polymerase} \\
&3). \text{ ligation}
\end{align*}
\]

2) Cassette isolated by cleavage with BamHI. Sticky ends removed by S1 nuclease.

\[
\begin{align*}
&\text{987654321} \\
&\text{GAATTCTAGAAAGAG} \text{ANNNNNNNNNNNNNNNNNNNNN} \\
&\text{CTTAAGATCCTTCCTCN} \text{ANNNNNNNNNNNNNNNNNNNNN} \\
&\text{23456789} \\
&1). \text{ cleave with MboII} \\
&2). \text{ treatment with T4 DNA polymerase} \\
&3). \text{ ligation}
\end{align*}
\]
RESULTS

Construction of the cassette

The first step involved the insertion of the MboII sites into pBR322 (Figure 34). Oligonucleotide 4 (see Chapter 3) was phosphorylated and mixed with oligonucleotide 2 to form Bam Duplex III. The Duplex was then ligated to itself about the blunt end, resulting in a dimeric form of Bam Duplex III, with BamHI sticky ends and an internal XbaI site. The MboII sites present extend outward and the cleavage sites are past the end of the molecule. This dimeric form was inserted into the BamHI site of pBR322. DNA segments inserted into the BamHI site of pBR322 will inactivate the tetracycline resistance gene but the plasmid will still provide resistance to ampicillin. The recombinant plasmids were transformed into E. coli C600 and plated on agar plates with ampicillin. Colonies were then replica-plated onto agar plates with tetracycline or ampicillin. Tetracycline-sensitive, ampicillin-resistant colonies were screened to detect an insert into the BamHI site of pBR322. Suitable plasmids were then transformed into E. coli strain GM119, a Dam-strain, as XbaI needs an unmethylated DNA substrate for cleavage. The presence of an XbaI site was then confirmed. The resultant plasmid, pRGB1, contains an XbaI site, along with two BamHI and two additional MboII sites.

Next, the synthetic lac o sequence produced by
Figure 34. Construction of pRBG1 and pRBG10. Bam
Duplex III was ligated to itself using the blunt end and
inserted into the BamHI site of pBR322, producing pRBG1.
The lac o sequences (Heyneker et al., 1976) was
isolated, and the sticky ends were filled in with DNA
polymerase I (Klenow fragment). The resulting fragment was
inserted into the XbaI site of pRBG1 to form pRBG10.
cleave with BamHI
insert dimeric Bam Duplex III
GATCCCTTCTCTAGAAGAG
GAGAAAGATCTTCTCTTAG
MboI XbaI

AATTC-\text{\textit{loc}}-\text{G-CTTAA}
fill-in using Klenow

AATTC-\text{\textit{loc}}-\text{G-CTTAA}

cleave with XbaI
insert fill-in \textit{loc} fragment

B BamHI
X XbaI
E EcoRI
\text{\textit{loc}} \text{\textit{o}}
Heyneker et al. (1976), was inserted (Figure 34). The EcoRI fragment of this sequence was filled-in by DNA polymerase I (Klenow fragment), producing blunt ends (see Methods). Similarly, pRBG1 was cleaved with XbaI and the sticky ends were filled-in. The lac o fragment and cleaved pRBG1 were then ligated together. The recombinant plasmids were transformed into E. coli strain RR1 for selection of the lac o region on Cys-NA plates containing Ap and X-Gal. This strain was chosen because it contains a lacY mutation which reduces the background blue color of the bacterial colonies. Insertion of the lac o region into the XbaI site will reform both the EcoRI and XbaI sites, and the resulting plasmids will cause the cells to produce a bright blue color on the selection plates. This plasmid, pRBG10, was used as a source for the MboII-lac o cassettes (Figure 32). The presence of the MboII-lac o cassette was demonstrated by restriction enzyme mapping (Figure 35). EcoRI-XbaI double digests resulted in the expected fragment sizes. BamHI cleavage of pRBG10 releases a 60 bp fragment. A degradation of the 560 bp MboII to a 392bp and 168 bp fragments was seen upon MboII cleavage of pRBG10 (Some of the MboII bands in Figure 35 are not of the expected size because MboII will not cleave DNA if the final A residue in the recognition sequence is methylated).
Figure 35. Restriction Mapping of pRBG10. pRBG10 was digested with several restriction enzymes, and the products were electrophoresed on a 5% polyacrylamide gel. Restriction enzymes used are shown in parentheses. Lane 1 - pBR322 (CfoI). Lane 2 - pBR322 (BspRII). Lane 3 - pRBG10 (BspRII). Lane 4 - pBR322 (BamHI). Lane 5 - pRBG10 (EcoRI). Lane 6 - pRBG10 (MboII). Lane 7 - pBR322 (MboII). Lane 8 - RBG10 (EcoRI).
Insertion of the cassette into DNA sequences

To demonstrate the usefulness of this technique, the \textit{lac} o-MboII cassettes were randomly inserted into bacterial plasmids. pBR322 was cleaved with DNAase I, under conditions shown to yield linear plasmid DNA molecules (Figure 36). The conditions used resulted in random double-stranded cleavages of the plasmid DNA. The ends were made blunt-ended with T4 DNA polymerase (see Methods) in preparation for insertion of the cassettes. The \textit{lac} o-MboII sequence was isolated from pRBG10 by cleaving it with \textit{BamHI}. The resulting 60 bp fragment was eluted from a 5% polyacrylamide gel. The \textit{BamHI} sticky ends of an aliquot of this fragment were filled-in with DNA polymerase I (Klenow fragment) and the DNA fragment was placed in a ligation mix with the DNAase I-cleaved pBR322. Following overnight ligation, the mixture was used to transform \textit{E. coli} RR1. Recombinant cells were selected on Cys-NA plates with Ap or Tc, selecting for inserts into different regions of pBR322. Colonies which were blue after incubation were screened by cleavage with \textit{EcoRI}, which is present only once in pBR322 but should be present 3 times in any recombinant plasmid containing the \textit{lac} o-MboII cassette (the cassette has 2 \textit{EcoRI} sites). The \textit{EcoRI} digestion will result in linear plasmids in the case of pBR322, or in multiple fragments in the case of any plasmids with \textit{lac} o inserts. The size of the \textit{EcoRI} fragments will give an indication of the distance of the \textit{lac} o-MboII cassette from the \textit{EcoRI}
Figure 36. Time Course of DNAase I Digestion of pBR322. pBR322 was digested with DNAase I as described in the Methods. Aliquots were removed and placed in a stop mix of 0.25 M EDTA. The different time points were electrophoresed on a 0.8% agarose gel and the DNA was visualized by UV illumination following ethidium bromide staining. EcoRI cleaved pBR322 was also used as a control. From this experiment it was determined that digestion for 5-10 minutes produced the greatest proportion of linear plasmid, which results from a single cleavage by DNAase I.
site of pBR322 (Figure 37).

All blue colonies which were screened had an insert while all white colonies which were screened did not. Several of the plasmids had large deletions, and a complicated digestion pattern was observed. This pattern apparently resulted from insertion into a vector which had been produced by multiple cleavages with DNAase I. One plasmid was found to have a single lac o-MboII cassette inserted and to have no other size differences from pBR322. This plasmid, pRBG312, was mapped with restriction enzymes in order to localize the position of the insert (Fig. 37). The lac o-MboII cassette is 60 bp in size, so the restriction fragments which contain the cassette will increase by that amount. Comparing the various digests of pRBG312 and pBR322, and by accounting for the required change in fragment size, it was possible to determine the region of pBR322 which contained the insert. This region is between positions 1391 and 1415 by the numbering system of Sutcliffe (1978). The plasmid still has functional ampicillin- and tetracycline-resistance genes. The region where the cassette has been inserted is downstream from the tetracycline-resistance gene and is not known to code for any specific product.

DISCUSSION

The insertion of a lac o-MboII cassette into a DNA sequence allows one to perform a wide range of subsequent
Figure 37. Restriction Mapping of lac o-MboII Cassette Inserts into pBR322. The lac o-MboII cassette was isolated by BamHI cleavages of pRBG10. The sticky ends were filled in with DNA polymerase I (Klenow fragment) and inserted into DNAase I-cleaved pBR322. Following transformation into RRI and selection on plates containing X-gal, several recombinant plasmids were found. Cleavage of these plasmids with EcoRI demonstrated the presence of the cassette since there will be more than one fragment present. Subsequent cleavage with other enzymes isolated the region of pBR322 where the cassette was inserted. Some size standards are shown. The plasmid and the enzyme it is cleaved with are as follows: Lane 1 - pBR322 (BspRI); Lane 2 - pRBG312 (RsaI and EcoRI); Lane 3 - pRBG314 (RsaI); Lane 4 - pRBG312 (RsaI); Lane 5 - pRBG303 (RsaI); Lane 6 - pBR322 (EcoRI); Lane 7 - pRBG314 (EcoRI); Lane 8 - pRBG312 (EcoRI); Lane 9 - pRBG303 (EcoRI); Lane 10 - pRBG302 (EcoRI); Lane 11 - pBR322 (EcoRI).
modifications of the DNA sequence. Not only can the inserted cassette inactivate genes in which it is placed, but those insertions can be selected by the intense blue color of the colonies containing the recombinant plasmids on agar plates containing X-gal. The lac o region can be removed by digestion with BamHI, EcoRI, or XbaI, leaving one or more of the restriction enzyme sites. These sites can be used for any further modification by standard techniques. The advantage this system has over simple insertion of a restriction enzyme linker, besides the easier identification of inserted by color, is that the cassette can be entirely removed at a later time. The cassette contains MboII recognition sites which will cleave outward into the surrounding DNA. Depending on the manner in which the lac o-MboII cassette is isolated, the number of bases MboII will cleave into the surrounding DNA can be varied (Figure 32). Removal of the lac o cassette from pRBG10 by digestion with MboII, followed by reaction with T4 DNA polymerase, will produce a fragment which will remove two base-pairs the surrounding DNA, one on either side of the cassette. Filling-in the BamHI-cleaved fragment with DNA polymerase I (Klenow fragment), as was done in this chapter, will result in an inserted sequence which would remove six base-pairs. Similarly, DpnI-cleaved cassettes will remove ten base pairs and BamHI cleavage, followed by S1 nuclease digestion, will produce a cassette which will remove fourteen base-pairs. This variability of the
number of bases removed means that one can work on the regions outside the insert that would not necessarily be accessible under other procedures. Following cleavage of the constructed sequence with MboII, one can simply ligate the ends together, effectively producing a deletion in that area ranging from 2-14 base-pairs. Another cassette could be inserted and several more base pairs could be removed. The directed deletion of the DNA surrounding the initial insert can be performed with this technique.

Following MboII cleavage, one could also insert a chemically synthesized region of DNA or some other DNA fragment. These could be selected for by the loss of the lac o region, resulting in white colonies. The chemically synthesized sequence could have single 3' sticky ends which are complementary to the ends left by the MboII digest. The MboII generated ends may not be complementary to each other, so that they would not effectively ligate to one another. The only recircularized plasmid should result from the insertion of the chemically synthesized sequence. The nature of the MboII ends will allow this sequence to be inserted in the proper orientation, producing a DNA sequence with a desired insert in it. Thus, specific changes of any DNA sequence can be studied by this procedure. The cassette can be used to delete a desired region of DNA and a desired chemically synthesized sequence can be inserted. The use of a similar cassette using another restriction enzyme, such as FokI would allow larger sticky ends on
each side of the cassette to be generated. This feature might prove useful in reassembly of gene fragments. The ability to completely remove the lac o-MboII cassette, along with the ability to then insert any sequence of DNA into that area, opens a wide range of possibilities for DNA manipulations which have been difficult to achieve.
References


