Title
Regulatory and functional studies with the EcoRI endonuclease

Author
Yanofsky, Stephen David

Publication Date
1988

Peer reviewed|Thesis/dissertation

Regulatory and Functional Studies with the EcoRI Endonuclease

by

Stephen David Yanofsky

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

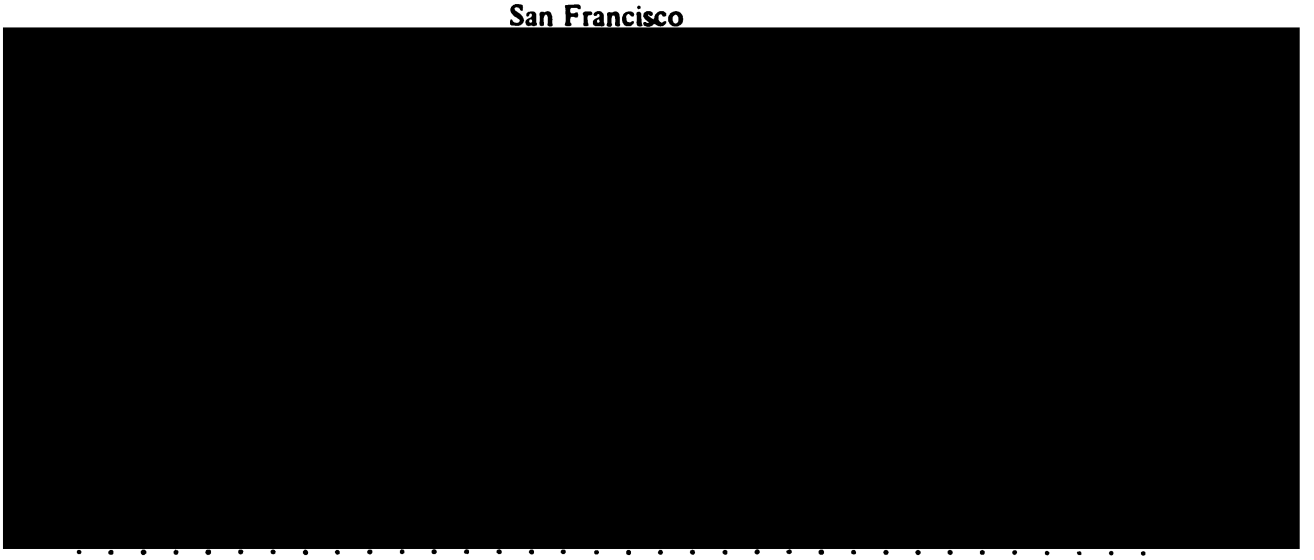
in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco



. . . . . .

Date University Librarian

Degree Conferred: . . . . . .

Acknowledgments

Writing this thesis has been a long and challenging process. While there were times where everything seemed to be going very smoothly, these were usually followed by some sort of stumbling block which had to be removed before I could push on. A number of people were instrumental in the process of removing these stumbling blocks. I would like to take this opportunity to thank all of the people who helped me both directly and indirectly in completing this seemingly endless endeavor.

First, I would like to express my appreciation to my wife, Paula, who has been both patient and supportive. I know that she must be quite tired of hearing that, "my thesis will be finished next week", as I have said over and over again. However, I grossly underestimated the large number of time consuming details required to actually finish this thesis.

I would especially like to thank my parents, who have been very supportive and who have provided me with encouragement during difficult times. Without their love and support I'm not sure I would have been able to get through many of the roadblocks that I encountered. I would also like to thank my father for reading an early draft of this thesis. His comments helped me to identify and correct many of the weak points in my writing style.

I would also like to thank Herb and Pat for their assistance and leadership when it was most needed and for standing behind me during difficult times. Without their support and confidence, this thesis

would never have been written. I found Herb's optimism to be particularly encouraging. I want to thank Pat for her critical analysis when it was needed. Although I often found the criticisms hard to take, her comments were most helpful.

I also want to thank Harold and Bruce for taking the time to be on my committee, for reading my thesis and especially for their helpful comments.

In addition, I would like to mention Swatch and Panda (our cats), who kept me company during the many nights that I sat in front of the computer.

Last, but not least, I need to thank myself for my persistence. This task often seemed insurmountable. However, by taking it a little bit at a time, it is now nearing completion.

REGULATORY AND FUNCTIONAL STUDIES WITH THE EcoRI ENDONUCLEASE

Stephen David Yanofsky

ABSTRACT

This study of the EcoRI restriction-modification system is divided into two parts: (1) investigations concerning regulation of the EcoRI genes, and (2), a genetic identification of sites important for EcoRI endonuclease function.

In part 1, plasmids were constructed which encoded a mutant endonuclease and a truncated, inactive methylase. Cultures containing these plasmids had greatly reduced endonuclease levels. When methylase was provided in trans, endonuclease levels were elevated 10- to 20-fold. One model postulated initial expression of methylase from a weak promoter located near the C-terminus of the endonuclease gene. Methylase would modify chromosomal DNA and subsequently activate transcription of a bi-cistronic message encoding endonuclease and methylase. Methylase-specific mRNA was expressed in vitro; in vivo expression from the methylase promoter was quantitated. To test the model, fusion plasmids were constructed placing β-galactosidase under EcoRI endonuclease control. Methylase failed to enhance β-galactosidase expression from these plasmids. Pulse-chase labellings were used to quantitate rates of endonuclease synthesis and turnover. Endonuclease synthesis rates were low in the methylase deficient

cultures. Endonuclease was stable in both the absence and presence of methylase. In order to rule out the possibility that residual endonuclease activity caused for the apparent activation, inactive (null) endonucleases were generated. Methylase failed to activate expression from plasmids encoding the null endonucleases. Thus, the previously observed activation was probably due to the decreased viablity of methylase deficient cells encoding a partially active endonuclease.

In part 2, a strong selection for an inactive (null) endonuclease was used to isolate several hundred endonuclease null mutants following in vitro hydroxylamine mutagenesis of plasmid DNA containing the endonuclease gene. The mutant proteins were analyzed on Western blots and compared to wild type endonuclease. DNA sequencing was performed to identify the substitution responsible for the null phenotype. The amino acid substitutions were examined relative to the 3 Å model of the EcoRI endonuclease-DNA complex. Mutations resulting in a null phenotype primarily involved substitutions of residues at either the protein-DNA interface or the dimer interface. Purified proteins were analyzed by HPLC and dimer interface substitutions were found to result in monomeric proteins.

Table of Contents

List of Tables

Chapter 1

Chapter 2

List of Figures

Chapter 1

Chapter 2

Chapter 1: Regulation of Expression of the EcoRI Genes

INTRODUCTION

Type II Restriction-Modification Systems

Type II restriction-modification systems contain two gene products, a restriction endonuclease and a modification methylase. To date, such systems have been identified in well over 100 different bacterial species. In fact, over 600 restriction enzymes have been identified (1-3). The reason for the existence of these systems is not known although their presence certainly protects host cells against foreign DNA. Cells containing restriction-modification systems are generally resistant to phage infection, i.e. phage replication is prevented as a result of endonucleolytic digestion of phage DNA. It has also been reported that these enzymes may promote site specific recombination *in vivo* (4). This process could permit a more rapid evolution of a species and thus confer a selective advantage.

The restriction and modification enzymes of type II systems recognize and interact with identical nucleotide sequences. The great majority of type II systems currently known contain enzymes which recognize 4-6 base pair palindromic DNA sequences. However, restriction enzymes have been identified which recognize longer or non-symmetrical sites (5-7).

For the majority of type II enzymes studied, the active form of the restriction endonuclease has been reported to be a dimer while the active form of the modification methylase is a monomer. Exceptions to this include *Bgl*I and *Bsp*I endonucleases which are claimed to be

monomeric (8, 9). Type II restriction endonucleases generally use Mg^{++} as a cofactor, and produce double stranded breaks within the recognition sequence. The sites of phosphodiester bond cleavage differ from enzyme to enzyme, being directly opposed (producing blunt ended products), or staggered (generating either 3' or 5' single stranded overhangs). The methylases employ S-adenosylmethionine (SAM) as a cofactor, producing N6-methyl adenine or 5-methyl cytosine by transferring a methyl group from SAM to the DNA substrate. DNA methylated at the recognition site can not be cleaved by the endonuclease. Continuous protection of host DNA from the corresponding restriction endonuclease is provided because the endonuclease is unable to cleave hemi-methylated recognition sites. Following semi-conservative DNA replication, hemi-methylated sites within DNA are recognized by the resident methylase and are fully methylated.

The availability of these enzymes has revolutionized biological research at the molecular level. These enzymes can be used in vitro to cleave DNA at specific sites, and thus are ideal tools for manipulation of DNA. Their use has played a major role in the rapid growth of the field of genetic engineering.

In addition to their widespread use in genetic engineering, restriction and modification enzymes provide model systems for investigation of sequence-specific interactions between protein and DNA. These enzymes recognize and interact with a small, but specific, nucleotide sequences and both the enzymes and DNA substrates are readily available in pure form for biochemical studies. The existence of multiple enzymes which recognize identical substrate sequences (isoschizomers) makes it possible to determine whether there are

multiple mechanisms or common features involved in the recognition and cleavage (or modification) of a particular nucleotide sequence. The enzymes of several restriction-modification systems are being studied with the ultimate goal of obtaining a more detailed understanding of the specific molecular features involved in protein-DNA interactions. The recent solution (to 3 Å resolution) of the crystal structure of the EcoRI endonuclease complexed with a 13 base DNA oligonucleotide has provided information concerning the specific interactions between this protein and its DNA substrate (10).

For a limited number of restriction-modification systems, the genes encoding the enzymes have been cloned (11-18) and the DNA sequences have been determined (17-23). Among the type II restriction-modification genes which have been cloned, the majority have been found to exist within their hosts on plasmids (14, 15, 18, 24, 25). However, some reside on bacterial chromosomes (16, 19, 26). The fact that these genes have been identified on bacterial chromosomes does not necessarily indicate that they evolved at chromosomal locations. Chromosomally located restriction-modification genes may represent examples of genes which have recently been recombined into the chromosome from bacterial plasmids or introduced via bacteriophage. It is also possible that these genes may have come from some other hosts. The base composition of several of the restriction-modification genes that have been sequenced differs from that of their host's DNA (17-19, 22). The base composition of the EcoRI and EcoRV genes is 65% A+T, which differs significantly from that of E. coli (50% A+T) (28) or plasmid ColE1 (50% A+T) (27). Similar differences are found for the A+T content of the BsuRI and PstI genes. These are 65% A+T, compared

with 57% A+T for B. subtilis DNA (77) and 59% for Providencia DNA (78). These differences suggest that these genes may have evolved in a different host organism.

The orientation of the genes, the direction of their transcription, and the protein sequences they encode have been deduced from DNA sequence information. Endonuclease and methylase genes generally have been found to be linked, but their relative orientations may differ. Examples are known in which the genes are arranged in tandem, in the same orientation, with endonuclease before methylase (17, 22, 23) or vice versa (20, 21). Examples also exist where the genes are opposed (encoded from opposite DNA strands), separated by a short spacer region (15, 18, 19) containing divergent promoters. These promoters can be close together, and may even overlap.

Establishment of Restriction-Modification Genes in a Host Cell

If restriction-modification systems are mobile, how do they become established in a new host cell? It would seem that transformation of a cell with a plasmid carrying these genes would be a lethal event, unless there was a mechanism providing for methylation of host DNA prior to the appearance of active endonuclease. The fact that methylase can recognize hemimethylated sites while endonuclease requires unmodified sites provides for maintenance of such systems. However, this does not explain establishment within a new host cell. Despite this apparent difficulty, plasmids or phage carrying restriction-modification systems have been reported to transform host cells effectively (13, 29, 30, S.Y. unpublished).

Why aren't these plasmids selected against? There are several possible explanations for this observation: 1) Methylase is active as a monomer while the endonuclease must dimerize in order to be active. Thus, the time required to form endonuclease dimers might be sufficient to permit methylase to modify host DNA. 2) The endonuclease may be sequestered away from chromosomal DNA within the cell, i.e. the endonuclease might be transported into the periplasmic space. Sequestering the endonuclease away from host DNA would provide a simple mechanism to permit establishment of these genes in a new host cell. 3) One or more regulatory mechanisms may have evolved to ensure methylase synthesis prior to endonuclease synthesis. Such a temporal expression could occur if methylase acts as a positive activator of endonuclease synthesis or activity. Regulation of this type could occur at the transcriptional, translational, or post-translational level. These and other possible mechanisms of control of expression of restriction-modification genes will be discussed subsequently.

Evidence has been obtained with the bacteriophage P1 restriction-modification system for expression of methylase prior to endonuclease upon transfection (30). These results suggest that there is temporal expression, allowing methylase to function before there is active endonuclease. Efficient transformation of E. coli HB101 cells with plasmids encoding the HhaII restriction and modification genes led Smith (29) to propose that methylase might act as a positive regulator of endonuclease expression. Such a regulatory process would permit mobile restriction-modification systems to become established.

Additional evidence consistent with some kind of regulatory process comes from studies on several other restriction-modification

systems. In the case of the PaeR7 system, cells can maintain plasmids which encode the PaeR7 endonuclease in the absence of a corresponding methylase (13). However, these cells do not restrict bacteriophage. High level expression of endonuclease appears to require an additional component (possibly methylase). In addition, the PstI methylase has been reported to be expressed before endonuclease (19).

Since there are several different genetic arrangements among the different restriction-modification systems which have been cloned, it is reasonable to assume that different mechanisms may have evolved to regulate their expression. PstI, EcoRV and PvuII are all transcribed divergently. Mechanisms have been proposed for both PstI and EcoRV to limit the rate of endonuclease expression. PaeR7 and HhaII are arranged in tandem, with methylase preceeding endonuclease. This arrangement seems to be most favorable for initial expression of methylase, but in the case of PaeR7, intracellular endonuclease levels are found to be much higher than methylase levels. EcoRI and BsuRI are arranged in tandem with endonuclease preceeding methylase. If these genes are expressed from a bicistronic message, this genetic arrangement would appear to favor endonuclease expression. Therefore it would seem especially important to have a mechanism to inhibit initial expression of these endonucleases.

The EcoRI endonuclease, and most type II restriction endonucleases, must dimerize to be active while the corresponding methylases are active as monomers. This might account partially for the ability of cells to survive introduction of a plasmid carrying these genes. However, the large number of potential sites within cellular DNA which methylase must protect from an endonuclease dimer

makes this unlikely as a sole mechanism to provide establishment. Evidence that the EcoRI endonuclease is dimeric at very low concentrations (33) makes it unlikely that the requirement for dimer formation greatly delays the appearance of active endonuclease. It is also interesting to note that plasmids carrying the BsuRI genes can transform cells (17). This endonuclease is active as a monomer, ruling out endonuclease dimerization as a regulatory mechanism for this system.

Sequestering the restriction endonuclease away from chromosomal DNA could provide a mechanism which would permit establishment of restriction-modification systems. This mechanism has been proposed for the PstI endonuclease, which has been found in the periplasmic space (11). Attempts to locate a higher molecular weight precursor for this endonuclease were unsuccessful and no apparent signal sequence is present at the amino terminus of this protein. It has been suggested that the periplasmic localization sequence may be internal, as in ovalbumin (32). However, in the case of EcoRI, sequestering of the endonuclease seems unlikely, since plasmids carrying the EcoRI endonuclease gene will not transform cells lacking methylase (31).

Regulation at the transcriptional level has been suggested for the PstI genes (19). The PstI endonuclease and methylase genes are transcribed divergently from opposite DNA strands with overlapping promoters. E. coli RNA polymerase has been shown to bind preferentially to the methylase promoter, resulting in both expression of methylase and inhibition of endonuclease expression. It has been suggested (19) that binding of RNA polymerase to the EcoRI methylase promoter might interfere with EcoRI endonuclease transcription.

However, the different genetic arrangement of the EcoRI genes necessitates utilization of a mechanism other than direct competition for RNA polymerase binding.

Regulation of endonuclease and methylase expression at the translational level has been proposed for several systems. These suggestions come from the fact that codon usage in several of the endonuclease and methylase genes differs significantly from that of host genes. Comparison of codon usage in highly expressed versus poorly expressed genes reveals that codon usage in restriction-modification genes resembles that of poorly expressed host genes (19). Rare codon usage may be due to the lack of any selective pressure to achieve high level expression of these genes (or selective pressure against high level expression). Examination of codon usage within the EcoRI genes reveals that several codons rarely used in E. coli are used frequently in the endonuclease and methylase genes. Expression of genes containing a large number of infrequently used codons might be inhibited by one of several possible mechanisms, including slower translation, ribosome stalling (which may lead to message degradation), or mistranslation (which may produce inactive or unstable proteins).

Translational control, based on the strength of the ribosome binding sites (Shine-Dalgarno sequences), has also been suggested for several restriction-modification systems. For example, the BsuRI methylase gene contains a much stronger Shine-Dalgarno sequence than the BsuRI endonuclease (17). The reverse is found for PaeR7, and intracellular levels of this endonuclease are found to be at least 10 times that of methylase (20). However, in the case of the EcoRV system, neither gene contains a good ribosome binding site (18).

Translational control by mRNA secondary structure has been proposed for the EcoRV system (18), where it has been suggested that mRNA secondary structure contributes to low level expression of endonuclease.

The EcoRI System

The subject of this study, the EcoRI system, contains enzymes which have been extensively characterized. The EcoRI endonuclease and methylase recognize the hexanucleotide sequence $\begin{matrix}5'\text{-GAATTC-}3' \\ 3'\text{-CTTAAG-}5'\end{matrix}$ within double stranded DNA. The endonuclease has been shown to cleave this sequence between the guanylic acid and adenylic acid residues, leaving a 5'-phosphate end and producing 4 base cohesive terminii (34). The EcoRI methylase modifies the central adenine in this sequence by transferring a methyl group from SAM, resulting in 6-methyl adenine (35). The aims of the work presented in this dissertation are to further our understanding of the processes which control expression of these genes, and to elucidate the features of the structure of the endonuclease that are responsible for its biological activity.

MATERIALS AND METHODS

Strains: all E. coli K-12

D1210 (31) recA13, proA2, lacY1, galK2, supE44, $lacI^Q$, ara14, hsdM, hsdR

MM294 (36) endI, pro, thi, hsdR

JM103 (37, 38) Δ(lac pro), thi, strA, supE, endA, sbcB, hsdR, proAB, $lacI^Q$, ZΔM15, /F'traD36

HB101 (39, 40) hsdS20, recA13, ara14, proA2, lacY1, galK2, rpsL20, xyl5, mtl1, supE44, $/F^-$

MC4100 (41) Δ(lac)U169, araD139, rpsL, relA, thiA, flbB, deoC, ptsF, rbsR, $/F^-$

GM48 (42) dam3, dcm6, gal, ara, lac, thr, leu, thi, tonA, tsx, $/F^-$

GM31 (42) dcm6, gal, ara, lac, xyl, thr, leu, thi, tonA, tsx, str, $/F^-$

Enzymes--all restriction enzymes were purchased from New England Biolabs with the exception of EcoRI, BamHI and PstI, which were purified in this laboratory according to the method of Greene et al. (43)

Media

L-broth (LB) (per liter)

10 g Bacto tryptone

5 g Yeast extract

10 g NaCl (pH 7.0)

for plates add 15 g Bacto agar

YT broth (YT) (per liter)

8 g Bacto tryptone

5 g Yeast extract

5 g NaCl (pH 7.0)

for plates add 15 g Bacto agar

for soft agar add 7.5 g Bacto agar

M9 (per liter)

10 ml 0.01 M $CaCl_2$

10 ml 0.1 M $MgSO_4$

20 ml 20% Casamino Acids (CAA)

H_2O to 880 ml autoclave

after autoclaving add 100 ml 10x M9 salts

for plates add 15 g Bacto agar

for M9 glucose add 20 ml 20% glucose

for M9 glycerol add 20 ml 20% glycerol

10x M9 salts (per liter)

$Na_2HPO_4.7H_2O$	132 g
KH_2PO_4	30 g
NaCl	5 g
NH_4Cl	10 g

Supplements: Thiamin 166 mg/l
Leucine 41 mg/l
Proline 166 mg/l

Isopropyl β-D-thio-galactopyranoside (IPTG) 1mM

O-nitrophenyl-β-D-galactoside (ONPG) 4 mg/ml in 100 mM $NaPO_4$ (pH 7.0)

5-bromo-4-chloro-3-indoyl-β-D-galactoside (Xgal)-- 20 mg/ml in Dimethylformamide--use 0.4 ml of this in media for 10 plates (200 ml)

Antibiotics: Kanamycin (kan) --- 50 µg/ml for plates
20 µg/ml in solution
Tetracycline (tet)- 20 µg/ml for plates
2 µg/ml in solution
Ampicillin (amp)---100 µg/ml for plates

Solutions:

666 6 mM Tris-HCl (pH 7.5 unless indicated otherwise)
6 mM $MgCl_2$
6 mM 2-mercaptoethanol

TMN 100 mM Tris-HCl (pH 7.5)

10 mM $MgCl_2$

100 mM NaCl

Cell suspension buffer (CSB)

20 mM KPO_4

300 mM NaCl (pH 7.0)

1 mM Ethylenediaminetetraacetic acid (EDTA)

7 mM 2-mercaptoethanol

Extract buffer (EB)

10 mM KPO_4

200 mM NaCl

1 mM EDTA (pH 7.0)

7 mM 2-mercaptoethanol

10-5-1 buffer

10 mM Tris-HCl (pH 7.5)

5 mM NaCl

1 mM EDTA

Lysis buffer

10 mM Tris-HCl (pH 7.0)

1 mM EDTA

1 mM N-ethylmaleimide (NEM)

+ 20 µg/ml phenylmethylsulfonyl fluoride (PMSF)

+ 1 mg/ml lysozyme (fresh)

DNase-RNase solution

10 mM Tris-HCl (pH 7.5)

10 mM $MgSO_4$

+ 10 µg/ml RNase

+ 50 µg/ml DNase

Triton buffer (TNT)

50 mM Tris-HCl (pH 7.5)

2 % Triton X-100 (Sigma)

1.0 M NaCl

Tris wash (TN)

50 mM Tris-HCl (pH 7.5)

100 mM NaCl

Z-buffer (per liter)

16.1 g $Na_2HPO_4.7H_2O$

5.5 g $NaH_2PO_4.H_2O$

0.75 g KCl

0.246 g $MgSO_4.7H_2O$

(pH 7.0)

add 2.7 ml 2-mercaptoethanol

Stop mix

5 % SDS, 25 % glycerol, 0.025 % bromphenol blue

Plasmids used

pMB1 (25), pPG31 (22), pMG31 (22), pMG31-6 (22), pBH20 (44), pBR322 (45, 46), pSC101-methylase (25), pAD1 (47), pMLB1010 (48), pMLB1034 (49), pACYC177 (50)

Plasmids constructed

Inv, Del, 322Inv, pSY21, pSY22, pSY23, pSY31, pSY32, pSY51, pSY60, pSY60T, pSY64, pSY64T, pSY71, pSY72, pSY73, pSY74, pMLB1010T, pMLB1034T, pSY81, pSY84

Plasmid descriptions (does not include plasmids for which maps are shown--descriptions of these plasmids are found in the figure legends).

pMB1--pMB1 contains the EcoRI endonuclease and methylase genes. This plasmid has been shown to heteroduplex with ColE1 with the exception of the 2300 base pair region encoding the endonuclease and methylase.

pPG31--contains the EcoRII fragment from pMB1 (carrying the EcoRI endonuclease and methylase genes) inserted into the EcoRI site of pBH20. This construction places both endonuclease and methylase under lac control. This plasmid encodes resistance to both ampicillin and tetracycline.

pMG31--pMG31 differs from pPG31 by the deletion of the 70 base pair HindIII fragment downstream from the methylase gene.

pSC101-methylase (pSC101-meth)--contains the EcoRI methylase gene from pMB1 (on a HindIII fragment) inserted into the HindIII site on pSC101. This plasmid is low copy number (6/cell), encodes tetracycline resistance and is compatible with the pMB1-derived, ColE1 type plasmids.

pBR322--This ColE1-related plasmid encodes resistance to ampicillin and tetracycline and is compatible with the pSC101 and pACYC derived plasmids. In addition, this plasmid is present at 20-30 copies/cell and is amplifiable with chloramphenicol.

pSY21--pSY21 contains the 955 base pair EcoRI-PstI fragment from the Del plasmid inserted into pBR322. This plasmid contains 342 base pairs of sequence information from the Del plasmid upstream of the endonuclease gene. The first 204 amino acids of the endonuclease protein are encoded. This partial endonuclease contains the S187 substitution. This plasmid also encodes resistance to tetracycline. pSY21 was constructed as follows: pBR322 and the Del plasmids were digested with EcoRI and PstI. These were ligated and transformed into E. coli 294. Tetracycline resistant transformants were selected and screened for ampicillin sensitive colonies which were white on x-gal plates. DNA was isolated, digested, and analyzed on agarose gels to determine the correct constructions.

pSY22--pSY22 was constructed from pSY21 by replacing the tetracycline resistance region with a 1201 base pair fragment encoding kanamycin resistance. pSY21 was digested with AvaI and ClaI and the ends filled

in using the DNA polymerase I Klenow fragment. This was ligated to a 1201 base pair fragment encoding kanamycin resistance, obtained from a pFB3 AvaII-HaeII fragment which had the ends filled in. Following ligation and transformation into E. coli 294, kanamycin resistant colonies were selected.

pSY23--pSY23 was constructed from pSY22 by insertion of the RpoBC transcriptional terminator (51) from plasmid pAD1 (47) into the EcoRI site. pSY22 was digested with EcoRI and the ends filled in using the DNA polymerase I Klenow fragment. BamHI linkers were ligated onto the RpoBC terminator fragment (pAD1 BamHI fragment digested with PstI followed by addition of BamHI linkers). These were mixed, ligated and transformed into E. coli 294.

pACYC177--This plasmid encodes resistance to ampicillin and kanamycin, is compatible with the ColE1 type plasmids and has been reported to be present at about 22 copies/cell. The location of the restriction sites shown on the pACYC177 map are not exact. They have been estimated from the circular map in Chang et al. (50).

pMLB1010--pMLB1010, the operon fusion vector, contains a complete β-galactosidase gene and encodes ampicillin resistance. The single EcoRI and BamHI sites approximately 800 base pairs upstream from the β-galactosidase gene are useful sites for insertion of DNA restriction fragments. Transcriptional activity from inserted sequences can be measured by quantitating levels of β-galactosidase.

pMLB1034--This gene fusion vector contains EcoRI, SmaI and BamHI sites at the N-terminus of the β-galactosidase gene. These can be utilized to construct in phase gene fusions which place β-galactosidase under both transcriptional and translational control of inserted sequences. This plasmid also encodes ampicillin resistance.

pSY81--pSY81 encodes a fused protein consisting of the amino terminal 204 amino acids of the EcoRI endonuclease (75%) fused in phase to β-galactosidase. This Pst gene fusion places β-galactosidase under both transcriptional and translational control of the EcoRI endonuclease promoter. This plasmid also encodes ampicillin resistance. This plasmid was generated from pMLB1034 and 322Inv. pMLB1034 was digested with SmaI and EcoRI. 322Inv was digested with PstI, treated with DNA polymerase I Klenow fragment and then digested with EcoRI. These were mixed, ligated and transformed into E. coli 294. Ampicillin resistant transfromants were selected. Colonies from a plate containing over 100 ampicillin resistant transformants were scraped off, DNA was isolated and transformed into E. coli MC4100. These were plated onto plates containint ampicillin and x-gal. DNA was isolated from blue colonies and analyzed on agarose gels. Transformants that produced the expected DNA fragments were termed pSY81.

pSY84--pSY84 is identical to pSY81 except that the 350 base pair EcoRI fragment from pSY23 encoding the RpoBC transcriptional terminator has been inserted at the EcoRI site. pSY84 was constructed by digesting pSY81 with EcoRI and treating this with calf intestinal phosphatase

(CIP). This was mixed with EcoRI digested pSY23. After ligation, DNA was transformed into E. coli 294 and ampicillin resistant transformants were selected. Transformants which produced blue colonies on x-gal plates and were kanamycin sensitive were analyzed further. DNA was isolated and the correct constructs determined.

pMLB1010T and pMLB1034T--These plasmids are identical to pMLB1010 and pMLB1034 except that they contain the 350 base pair EcoRI fragment from pSY23 containing the RpoBC transcriptional terminator sequence inserted at the EcoRI site.

pSY60T and pSY64T--These are identical to pSY60 and pSY64 except that they contain the RpoBC transcriptional terminator inserted at the EcoRI site.

Methods

Mini DNA preps (alkaline lysis mini screen) (80)

One milliliter cultures were grown overnight in YT broth at 37°C. Cells were harvested by centrifugation and the cell pellet frozen. The cell pellet was resuspended in 0.09 ml of solution 1 (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl (pH 8.0)) and allowed to sit for 5 minutes on ice. Lysozyme (0.01 ml of 5 mg/ml in solution 1) was added. After an additional 5 minutes on ice, 0.2 ml of solution 2 (0.2 N NaOH, 1% SDS) was added. After 5 minutes on ice 0.15 ml of ice cold 5 M KOAc pH 4.8 was added amd allowed to sit 5 minutes on ice. This was spun for 5 minutes in an Eppendorf centrifuge, and 0.4 ml of the supernatant was transferred into a fresh tube. Isopropanol (0.24 ml) was added, the solution was taken to -20°C and then centrifuged for 15 minutes in an Eppendorf centrifuge. The DNA was dried in a Savant spin-vac and resuspended in 0.02 ml of 10-5-1 buffer.

Preparation of competent cells

YT cultures (160 ml) were innoculated from 1 ml of an overnight culture. Cultures were grown at 37°C to an OD_{600} between 0.4-0.5 and then harvested by centrifugation in a Sorvall GSA rotor for 10 minutes at 7,000 rpm (4°C). Cell pellets were resuspended in 40 ml 0.1 M $MgCl_2$ and centrifuged in a Sorvall SS34 rotor for 10 minutes at 7,000 rpm (4°C). Cell pellets were resuspended in 40 ml 0.1 M $CaCl_2$ and allowed to sit for 20 minutes on ice. This was centrifuged in a SS34 rotor for

10 minutes at 7,000 rpm (4°C). Cells were resuspended in 8 ml 0.1 M $CaCl_2$ containing 17% glycerol and aliquoted into tubes containing 0.5 ml/tube. These aliquots were then frozen in liquid nitrogen and stored at -70°C. When prepared in this manner, cells can be used for up to 1 year.

Determination of viable cells

Cell cultures (5 ml) were grown overnight at 37°C in YT broth containing appropriate antibiotics. Overnight cultures were diluted 1/50 into fresh YT (or other indicated media) and grown at 37°C to an OD_{600} of 0.3-0.4. Cultures were diluted and plated onto M9 glycerol plates or M9 glycerol plates containing 1 mM IPTG. Plates were incubated overnight at 37°C and the number of colonies was determined.

Restriction test

To obtain a lawn of cells, 5 ml YT cultures were prepared for strains to be tested. Cell cultures (0.2 ml) were added to 5 ml soft agar, poured onto a YT plate, and incubated overnight at 37°C. As a control, a lawn was prepared from a non-restricting strain (i.e. E. coli 294). (Note: Use both unmodified, M^-, lambda vir (made in E. coli 294) and EcoRI modified, M^+, lambda vir (made in pMB1/294)). Dilute phage stock and spot 50, 500, and 50,000 of M^+ phage and equal numbers of M^- phage onto each plate. Expected result: For a restricting strain, the M^+ phage will produce plaques, while the M^- phage will yield a greatly reduced number of plaques. A non-restricting strain will contain equal numbers of plaques from the M^+ and M^- phage.

Test for the presence of methylase

Restricting and non-restricting lawns were prepared as follows: Restricting lawns were made by mixing 0.1 ml pPG30/D1210 cells with 3 ml soft agar and pouring this onto a YT plate. Non-restricting lawns were made by mixing 0.1 ml E. coli D1210 cells with 3 ml soft agar and pouring onto YT plates. The colony to be tested was picked with a wood applicator stick into 2 drops of LB, mixed, 100 phage lambda and 1 ml soft agar were added, and this was allowed to grow overnight at 37°C. Plaques were picked into 0.1 ml YT and 0.003 ml of this was spotted onto restricting and non-restricting lawns. If the original colony contained methylase, equal number of plaques will appear on both lawns. If the original colony was M^-, the number of plaques on the restricting lawn will be much less than the number of plaques on the non-restricting lawn.

Measurement of endonuclease levels

Growth of cells--Grow strains to be tested on YT plates containing the appropriate antibiotics and incubate plates overnight at 37°C. Innoculate 5 ml M9 glucose cultures from a single colony off of the YT plate and allow to grow overnight at 37°C. Innoculate 50 ml of the indicated media with 0.5-1.0 ml of the overnight culture. Grow at 37°C until the OD_{600} of cultures reaches 0.3. Harvest cells from 35 ml of culture by centrifugation for 10 minutes at 7,000 rpm in SS34 at 4°C and freeze the cell pellets. For induced cultures, add IPTG to 1 mM when cells reach an OD_{600} of 0.3, grow 1 hour and harvest cells as above. Note: check cells to be sure plasmid is retained as follows:

Plate cells from overnight and harvest cultures onto YT plates and allow to grow overnight. The following day, pick at least 20 colonies from each YT plate onto antibiotic plates and score for antibiotic resistance.

Preparation of cell extracts by sonication--Frozen cells were resuspended in 1 ml CSB and sonified 3-5 x 30 seconds keeping the temperature as close to $4^{\circ}C$ as possible. Sonication was monitored as follows: 0.01 ml aliquots were removed after each sonication step and added to 1.2 ml CSB. This was centrifuged 2 minutes in the microfuge and the OD_{280} of the supernatant was determined. When the OD_{280} was no longer increasing with additional sonication, sonicates were centrifuged 30-60 minutes in the SS34 rotor at 17,000 rpm at $4^{\circ}C$. The supernatant was poured into a fresh tube; this is the cell extract. Cell extracts were stored on ice and assayed as rapidly as possible.

Endonuclease assays--The following solutions are needed: Reaction cocktail (CT), 666 reaction buffer for assaying the S187 endonuclease, and TMN reaction buffer for assaying the wild type endonuclease. The reaction cocktail (CT) for 20 reactions is: Reaction buffer (0.04 ml of 10x), 0.02 ml of AvaI digested pBR322 (0.1 mg/ml) and 0.26 ml H_2O. Reactions containing 0.018 ml of CT were started by the addition of 0.002 ml of cell extract (or diluted cell extract). Reactions were halted at 2 minutes, 5 minutes, or 10 minutes by the addition of 0.003 ml of SDS stop mix. Samples were electrophoresed on a 1.2 % agarose gel at 200 volts until the bromphenol blue dye was near the bottom of the gel. Gels were stained with ethidium bromide, the bands visualized

with a short wave UV light and photographed. The extent of digestion of the pBR322 AvaI substrate was estimated visually from the photographs.

β-lactamase assays--β-lactamase assays were performed in 1 ml of 50 mM KPO_4 pH 7.0 containing 0.052 mg nitrocefin (83). Assays were carried out at room temperature, and were begun by the addition of 0.01 ml of cell extract (or diluted cell extract). The amount of β-lactamase activity was determined by monitoring the change in OD_{482} at time points of up to 10 minutes. (Note: The change in OD_{386} can also be used to follow breakdown of nitrocefin).

Calculation of endonuclease levels--Endonuclease levels are given in units of endonuclease activity per milligram protein. Therefore, activity measurements must be converted to units per ml and protein concentration must be determined as mg protein per ml. One unit of endonuclease activity is the amount of endonuclease required to cleave 1 pmole of phosphodiester bonds in 1 minute at 37°C. With pBR322 as substrate (M.W. 2.8×10^6 D.) at 0.1 mg/ml, the DNA concentration is 36 mM. Using 0.001 ml of this in a 0.025 ml reaction, the DNA concentration in the reaction is 1.44 nM, which equals 0.036 pmoles of DNA. Since pBR322 contains one EcoRI site, which contains 2 cleaved phosphodiester bonds, there are 0.072 pmoles of sites in the reaction mixture. Endonclease activity equals (pmoles of DNA x % cleaved x fraction of ml x dilution of extract) divided by time (in minutes). Endonuclease levels are obtained by dividing endonuclease activity by protein concentration in mg/ml.

Determination of methylase levels

Growth of cells and preparation of cell extracts--Strains to be tested were grown overnight at 37° on YT plates containing the appropriate antibiotics. Five milliliter M9 glucose cultures were innoclulated from a single colony and grown overnight at 37°C. One hundred milliliter cultures were innoculated from 1-2 ml of the overnight culture. These were grown at 37°C to mid log (OD_{600} of 0.3), the cultures were divided in half and IPTG was added to 1 mM to one of the halves. After an additional hour of growth, cells were harvested by centrifugation and the cell pellets were frozen. Extracts were prepared from these cells by the sonication method as described in determination of endonuclease levels.

Methylase assays--The following cocktail (CT) is sufficient for 20 reactions: 0.04 ml calf thymus DNA (sonified) (0.98 mg/ml), 0.05 ml 1.0 M Tris pH 8.0, 0.025 ml 200 mM EDTA, 0.04 ml ^3H S-adenosyl-methionine (SAM), 0.05 ml BSA (5 mg/ml), 0.295 ml H_2O. Each reaction contains 0.023 ml CT + 0.002 ml extract (or diluted extract). Reactions were performed at 37°C for 5 or 10 minutes. Reaction mixture (0.02 ml) was spotted onto DE81 filter circles (2.3 cm) (Whatman) and the filters were dropped into ice cold 50 mM Na_2HPO_4. This was shaken gently for 15 minutes. The Na_2HPO_4 buffer was changed and the wash continued for 15 minutes. This step was repeated and then followed by a wash in 70% ethanol, a wash in 95% ethanol and a wash with ether. The filters were then dried and counted in a scintillation counter.

Calculation of methylase levels--Methylase units equals pmoles CH_3/minute at 37°C. This is calculated as follows: (cpm-blank) x dilution divided by time (in minutes) x cpm/pmole (5,000). This is converted to units per ml of extract, and divided by the protein concentration (in mg/ml) to obtain methylase levels.

Generation and characterization of a temperature-sensitive methylase

Mutagenesis and selection--pPG31 DNA was mutagenized with hydroxylamine according to the method of Davis and Botstein (81). Samples were removed at 2 hours, 7 hours, 19 hours, 26 hours and 42 hours. These were dialyzed against 10-5-1 buffer and DNA was transformed into E. coli 294. Colonies appearing on plates containing tetracycline at 30°C were replica plated onto plates at 42°C. Those which grew at 30°C but not at 42°C were analyzed further.

Growth of cultures harboring altered methylases--Cultures were grown to early log phase at 30°C. A portion of each culture was transferred to a flask at 42°C. At intervals of 30 minutes, the OD_{600} of cultures was determined and a portion of each culture was diluted and plated to determine the number of viable cells.

β-galactosidase assays--(82) Cultures were grown to mid log phase and placed on ice. The OD_{600} was determined and 1 drop of toluene was added. This was mixed with a vortex mixer and shaken for 30 minutes at 37°C. Toluene treated cells (0.5 ml) were added to 0.5 ml of Z-buffer and 0.2 ml of ONPG (4 mg/ml). After 15 minutes at 37°C, 0.5 ml of

Na_2CO_3 was added and the OD_{420} and OD_{550} of the solution was determined. From these values, units of β-galactosidase were calculated as follows: Units of β-galactosidase equals 1000 x (OD_{420} -1.75(OD_{550})) divided by time (minutes) x volume x OD_{600}.

Rates of synthesis and turnover of endonuclease

This procedure was adapted from method of Hoyt et al. (49)

Cell growth and labelling--Dilute overnight culture 1/50 or 1/100 into labelling media (M9 glucose or M9 glycerol without CAA). Grow cultures to an OD_{600} of 0.3-0.4, remove 1 ml of culture and place in a prewarmed tube containing 100 µCi of ^{35}S-methionine. After 1 minute (or other indicated time), 300 µg of unlabelled methionine was added. Aliquots (0.15 ml) were removed at time points of 1 minute, 2 minutes, 5 minutes, 10 minutes, 30 minutes, and 60 minutes. These were added to 0.8 ml of stop cells. The cells were pelleted and then frozen.

Lysis--Cells were resuspended in 0.05 ml of lysis buffer and treated with 4 cycles of freezing and thawing. DNAse-RNAse buffer (0.1 ml) was added and the mixture was placed on ice for 30 minutes. After 1 minute at 25°C, this was again placed on ice and 0.0375 ml of 5 M NaCl was added. This was spun for 15 minutes at 4°C in a microfuge and 0.18 ml of the supernatant was transferred to a fresh tube. Following the addition of an equal volume of TNT, antibody was added (generally 0.003 ml of 1/10 dilution of anti-serum). This was allowed to sit overnight on ice.

Sample preparation--To the samples containing antibody, 0.05 ml of BRL immunoprecipitin was added. This was spun for 2 minutes, the pellet resuspended, washed with 0.5 ml TNT and pelleted again. After repeating this TNT wash, a wash with TN was performed. Samples were pelleted and then resuspend in SDS-loading buffer containing 200 ng of pure EcoRI endonuclease. Samples were boiled 4 minutes and loaded onto SDS-polyacrylamide gels. Gels were dried down with a Hoefer gel drier and the bands visualized on x-ray film. (Note: Anti-β-galactosidase was sometimes used as an internal control).

Preparation of stop cells--Five hundred milliliter cell cultures were grown overnight in LB and harvested by centrifugation. Cells were resuspended in M9 -CAA to an OD_{590} of 20.0. NEM was added to 3 mM and PMSF was added to 200 µg/ml. Cells were incubated 10 minutes at 37°C, placed on ice, and NaN_3 was added to 10 mM. An additional 200 µg/ml of PMSF was added along with an additional 3 mM NEM. Methionine was added to 300 micrograms/ml and cells were divided into 0.8 ml aliquots.

RESULTS

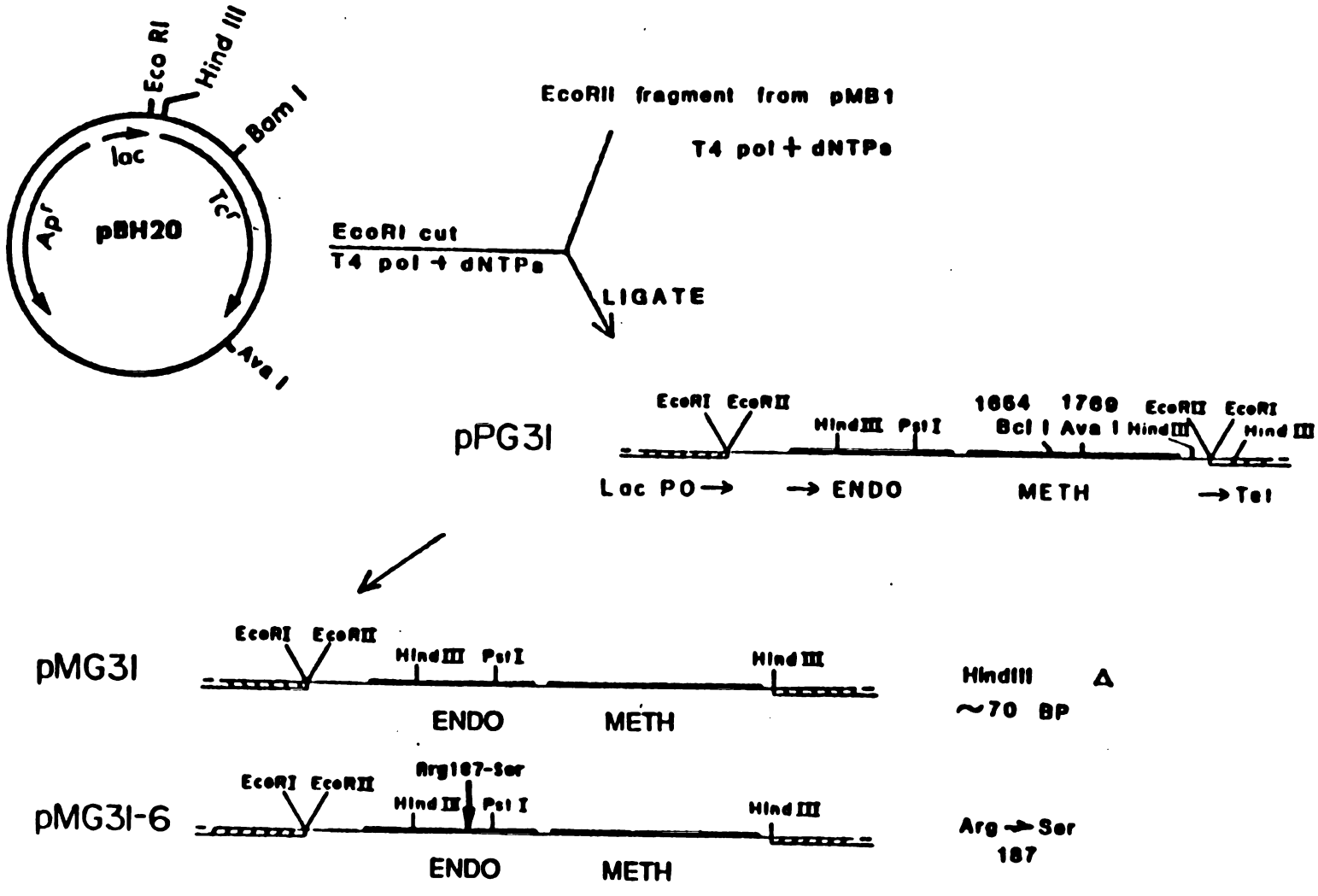
The EcoRI Genes

Characterization, cloning and sequencing. The EcoRI endonuclease and methylase genes reside naturally on plasmid pMB1 (25). In heteroduplex analysis with plasmid ColE1, pMB1 was shown to pair except for a 2300 base bair region containing the EcoRI endonuclease and methylase genes (25). The EcoRI genes from pMB1 were subcloned into the EcoRI site of plasmid pBH20 (44) (pBH20 is equivalent to pBR322 with the addition of a 200 base pair fragment containing the wild type lac promoter-operator sequence inserted at the EcoRI site). The resulting plasmid, pPG31 (Fig. 1), places the EcoRI endonuclease and methylase genes under lac repressor control (22). In plasmid pPG30 (22), the EcoRI genes are in the reverse orientation relative to that found in pPG31. Experiments with these plasmids (pPG30 and pPG31), have established the direction of transcription of the EcoRI genes.

The nucleotide sequence of the EcoRI genes was determined by Greene *et al*. (22) and by Newman *et al*. (23); the protein sequences were deduced from the DNA sequence. The endonuclease gene encodes a protein of 277 amino acids while the methylase gene encodes a 326 amino acid protein, corresponding to monomer molecular weights of 31 kd and 38 kd resprectively.

A mutant endonuclease, Arg187-Ser. A spontaneous mutation within the endonuclease gene, Arg187-Ser (S187), was identified by DNA sequencing (22). In order to study the S187 endonuclease, both mutant and wild

Fig. 1 EcoRI plasmids--The EcoRI genes were identified on plasmid pMB1. An EcoRII fragment containing the EcoRI genes was cloned into the EcoRI site of pBH20, resulting in pPG31. This places endonuclease and methylase under lac control. pPG30 contains the EcoRII fragment inserted in the opposite orientation. pMG31 differs from pPG31 by the deletion of a 70 base pair HindIII fragment downstream from the methylase gene. pMG31-6 is identical to pMG31 except that it encodes a mutant endonuclease, in which Arg187 has been replaced by Ser.


Eco RI
Hind III
Bam I
lac
pBH20
Ap^r
Tc^r
Ava I
EcoRII fragment from pMB1
T4 pol + dNTPs
EcoRI cut
T4 pol + dNTPs
LIGATE
pPG31
EcoRI EcoRII
HindIII PstI
1664 1769
BclI Ava I HindIII
EcoRII EcoRI
Hind III
Lac PO→
→ ENDO
METH
→Tet
pMG31
EcoRI EcoRII
HindIII PstI
HindIII
ENDO
METH
HindIII Δ
~70 BP
pMG31-6
EcoRI EcoRII
Arg187-Ser
HindIII PstI
HindIII
ENDO
METH
Arg → Ser
187

type sequences were cloned back into pPG31, resulting in plasmids designated pMG31-6 and pMG31, respectively. These plasmids differ from pPG31 in that they lack a 70 base pair HindIII fragment downstream from the methylase gene (Fig. 1).

Studies on endonuclease present in extracts of strains carrying pMG31-6 show that the S187 enzyme has low activity under standard EcoRI assay conditions (100 mM NaCl) (Table I). However, when extracts containing this endonuclease are assayed in low salt (<20 mM), the S187 endonuclease is as active as the wild type enzyme. The decreased activity of the S187 endonuclease under in vivo conditions is shown by the inability of strains containing this form of the enzyme to restrict phage lambda (Table I). The plating efficiency of phage λ on S187 containing cells (pMG31-6) is 0.5 compared with plating on the parent strain, while the plating efficiency of phage λ on cells containing pMG31 is 10^{-4} that on the parent strain. Studies on the ionic strength dependence of DNA binding have shown that the S187 substitution effects the electrostatic nature of the protein-DNA interaction without altering sequence specificity (79).

Alteration of the methylase gene. The availability of a mutant endonuclease containing limited in vivo activity made it possible to construct plasmids encoding this endonuclease and an inactive methylase. The methylase gene on plasmid pMG31-6 (Fig. 2) was altered by constructing plasmids with both an inversion and a deletion of approximately 50% of the methylase coding region. Analysis of DNA from transformants resulted in identification of a plasmid containing an inversion of the Bcl-Bam fragment, Inv (Fig. 3), and a plasmid

Table I. Properties of plasmids containing the EcoRI genes.

STRAIN	EFFICIENCY OF PLATING OF LAMBDA PHAGE		EcoRI ACTIVITY	
	UN-MODIFIED	EcoRI MODIFIED	METH (UNITS/MG)	ENDO (RELATIVE)
294	1.0	1.0	-	-
pPG31/294	2.7×10^{-4}	0.94	250	1.0
pMG31/294	4.7×10^{-3}	0.90	860	1.0
pMG31-6/294	0.52	1.0	1,090	0.003

Efficiency of plating and methylase assays were performed as described in Methods. Endonuclease assays were performed under "standard" EcoRI conditions, i.e. 100 mM Tris pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$.

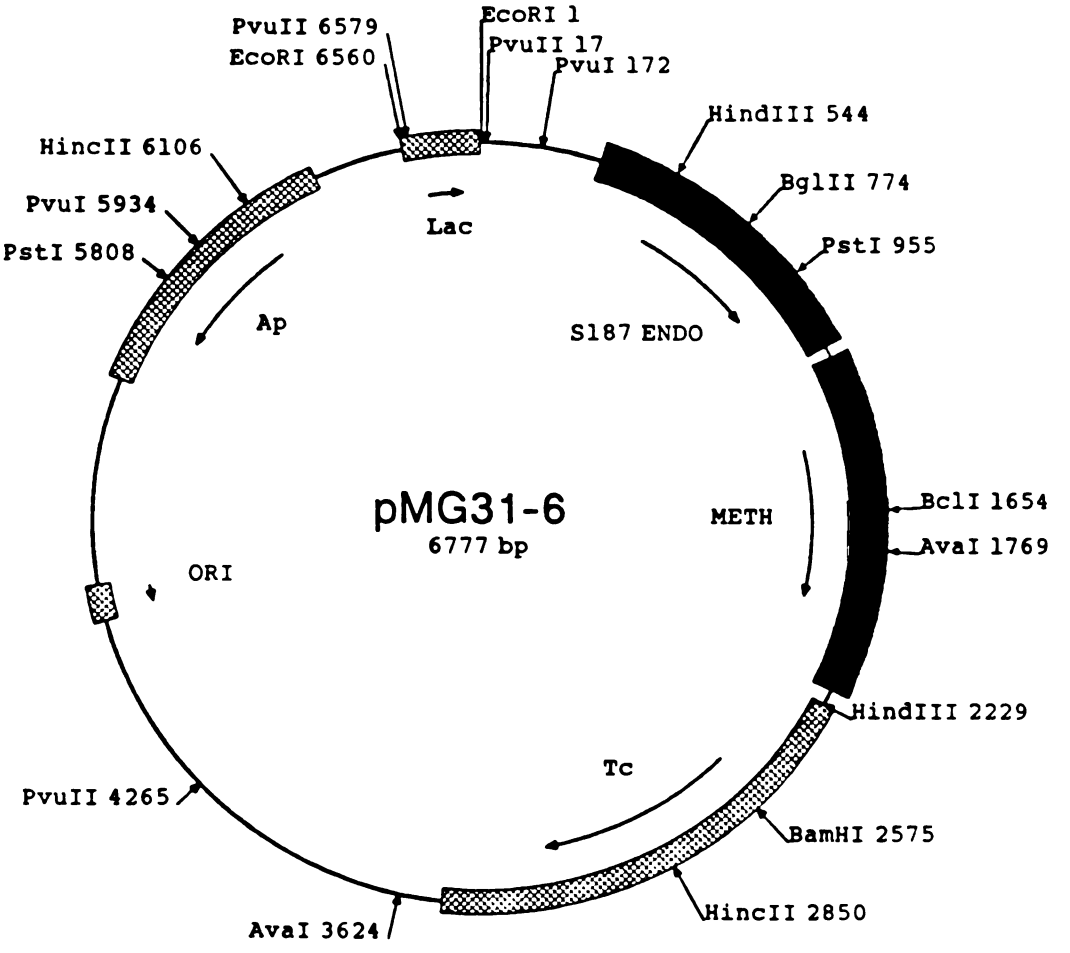


Figure 2. pMG31-6--pMG31-6 is identical to pMG31 except for the Arg187-Ser alteration in the endonuclease. This construction places both the S187 endonuclease and wild type methylase under control of the wild type lac promoter-operator sequence. This ColE1-type plasmid encodes resistance to ampicillin and low level tetracycline resistance.

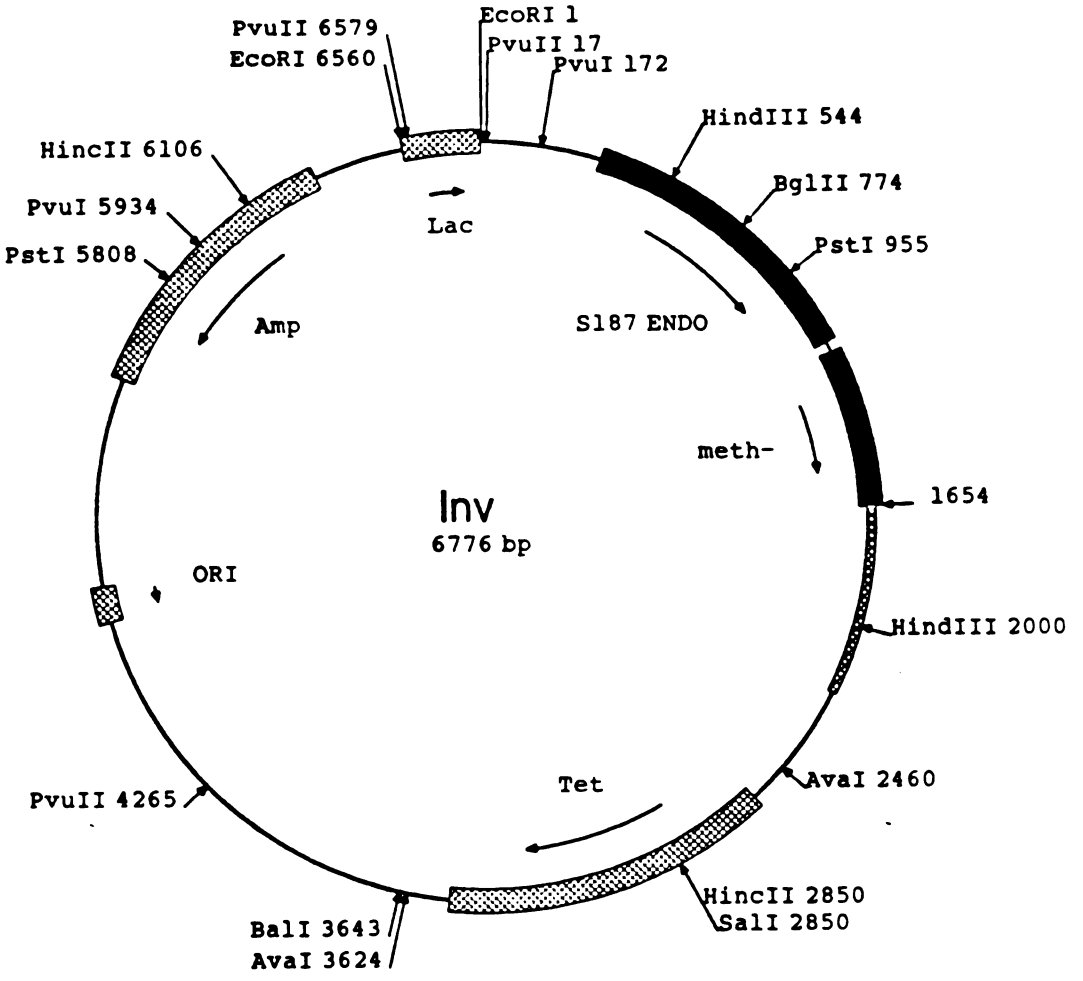


Figure 3. Inversion (Inv)--The Inv plasmid is derived from pMG31-6, encodes the S187 endonuclease under lac control and contains a truncated methylase gene. This plasmid encodes ampicillin resistance. pMG31-6 DNA was prepared from the dam$^-$ strain *E. coli* GM48 in order to be able to digest this DNA with BclI. DNA was digested with BclI and BamHI, ligated and redigested with BamHI to linearize any recircularized pMG31-6. After transformation into *E. coli* D1210, ampicillin resistant transformants were selected, plasmid DNA was isolated, and digests of this DNA were analyzed on agarose gels. The Inv plasmid has the orientation of the BclI-BamHI fragment (of pMG31-6) inverted.

containing a deletion of the Bcl-Bam fragment, Del (Fig. 4). E. coli D1210 cells harboring the Inv or Del plasmids survive, despite the absence of an active EcoRI methylase. However, cultures containing these plasmids grow more slowly than the parent strain.

From examination of the relevant DNA sequence of pMG31-6 it can be predicted that the Inv plasmid encodes a truncated gene product containing the N-terminal 152 amino acids of methylase while the Del plasmid encodes a 449 residue fused polypeptide, consisting of the N-terminal 152 amino acids of methylase fused in phase to the tet resistance protein of plasmid pMG31-6.

The ability of cells harboring pMG31-6 or the Inv plasmid to form colonies when grown in the absence and in the presence of the inducer IPTG was measured[1] (Table II). Cultures containg pMG31-6 plate equally well in both the absence and presence of IPTG, while cells containing the Inv plasmid do not survive IPTG induction. From Inv containing cultures, only 1 in 10^4 cells survive IPTG induction. DNA

[1] The initial Del isolate contained very low endonuclease levels (0.0013 that of pMG31-6). These were elevated to 0.03 upon IPTG induction. Del endonuclease levels were the same in both the presence and absence of methylase. The Del was later reconstructed and found to contain higher endonuclease levels. Strains containing this plasmid barely survive in the absence of methylase. It is now believed that the original Del plasmid is probably a double mutant, fortuitously picked up because its low endonuclease activity allows formation of a large colony, while cells containing the true Del plasmid (encoding the S187 endonuclease) grow very poorly in the absence of methylase and therefore produce small colonies.

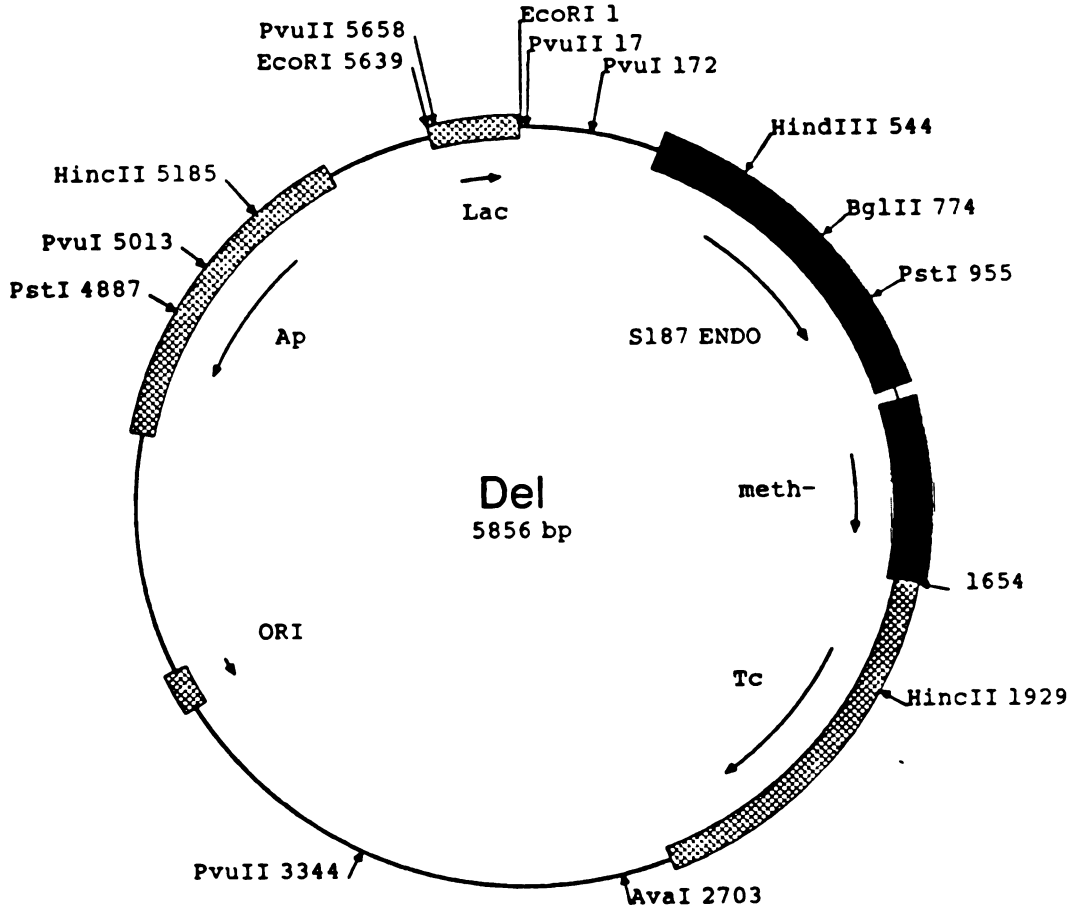


Figure 4. Deletion (Del)--The Del plasmid is derived from pMG31-6 and encodes the S187 endonuclease under lac control. This plasmid contains a truncated methylase gene and encodes resistance to ampicillin. This plasmid was constructed at the same time as the Inv plasmid and was isolated from the same pMG31-6 digest. The BclI-BamHI fragment on pMG31-6 has been deleted.

Table II. Viable cells in uninduced and induced *E. coli* D1210 cultures containing plasmids pMG31-6 or Inv.

PLASMID	-IPTG	+IPTG
pMG31-6	1.7×10^9	1.8×10^9
Inv	7.2×10^8	3.7×10^4

0.2 ml of 5 ml LB overnight cultures was used to innoculate 5 ml LB cultures. Cells were grown 3 hours at 37°C and plated onto M9 glycerol and M9 glycerol + IPTG plates. IPTG resistant survivors from the Inv/D1210 culture were picked onto plates containing ampicillin. 41/60 were ampicillin resistant, while the remaining 19 were sensitive to ampicillin. DNA was prepared from 6 ampicillin resistant and 6 ampicillin sensitive colonies. All 6 ampicillin sensitive colonies lacked any visible plasmid DNA. All 6 ampicillin resistant colonies contained plasmids which had IS1 inserted within the endonuclease gene.

isolated from these "survivors" was analyzed by restriction analysis. All of the survivors were found to contain IS1 insertions within the endonuclease gene. This experiment was also performed with cells containing the EcoRI methylase gene inserted in plasmid pSC101 (designated pSC101-methylase (25)). pSC101-methylase is compatible with the EcoRI plasmids. It encodes tet resistance and is present at about 6 copies per cell (52). With methylase provided from this plasmid, cultures harboring the Inv plasmid contained the same number of viable cells in the absence or presence of IPTG (Table III).

EcoRI endonuclease activity was quantitated from extracts of E. coli D1210 cells containing plasmids pMG31-6, Inv and Del grown in the presence and absence of IPTG. From both endonuclease activity and protein concentration data, endonuclease levels[2] were calculated (Table IV). Uninduced cultures containing the Inv plasmid had approximately 3% of wild type endonuclease levels, while Del plasmid containing cultures had approximately 25% of wild type levels. Upon IPTG induction, endonuclease levels expressed from the Inv plasmid were near wild type levels.

The unexpected result of this experiment was that cultures containing either the Inv or Del plasmids were found to have lower endonuclease levels than cultures harboring pMG31-6. This was unanticipated since both the Inv and Del plasmids were constructed by elimination of a portion of the methylase gene (located downstream from the endonuclease gene) and do not alter the endonuclease gene.

[2] Endonuclease levels are given in units of endonuclease activity per milligram of total cell protein.

Table III. Viable cells in *E. coli* D1210 cultures containing plasmids pMG31-6 or Inv in the absence or presence of the EcoRI methylase.

PLASMID	O/N	-IPTG	+IPTG
pMG31-6	3.7×10^9	1.6×10^8	2.0×10^8
pMG31-6 + 101-meth	3.3×10^9	1.1×10^8	2.4×10^8
Inv	1.3×10^9	1.0×10^8	6.5×10^4
Inv + 101-meth	2.2×10^9	8.0×10^7	2.2×10^8

10 ml M9 glucose cultures were innoculated and grown at 37°C to mid log. A portion of the cells were plated and IPTG was added to the culture. After 60 minutes of growth in the presence of IPTG, cells were plated to determine the number of viable cells. 101-meth = pSC101-meth

Table IV. Properties of *E. coli* D1210 cultures containing plasmids pMG31-6, Inv or Del

PLASMIDS	SURVIVAL -IPTG	SURVIVAL +IPTG	RELATIVE ENDO -IPTG	RELATIVE ENDO +IPTG
pMG31-6	+	+	1.0	10.0
Inv	+	-	0.03	8.5
Del[3]	+/-	-	0.25	ND
pMG31-6 + 101-meth	+	+	1.0	10.0
Inv + 101-meth	+	+	0.5	8.5
Del[3] + 101-meth	+	+	ND	ND

Survival was determined by ability to form colonies on plates. Endonuclease levels (relative to pMG31-6, uninduced) were determined as in Methods except that endonuclease assays were performed in 6 mM Tris pH7.5, 6 mM $MgCl_2$, 6 mM 2-mercaptoethanol (666 buffer). ND: Not determined
101-meth = pSC101-meth

[3] The initial Del isolate contained very low endonuclease levels (0.0013 that of pMG31-6). These were elevated to 0.03 upon IPTG induction. Del endonuclease levels were the same in both the presence and absence of methylase. The Del plasmid was later reconstructed and found to contain higher endonuclease levels. Strains containing this plasmid barely survive in the absence of methylase. It is now believed that the original Del plasmid is probably a double mutant, fortuitously picked up because its low endonuclease activity allows formation of a large colony, while cells containing the true Del plasmid (encoding the S187 endonuclease) grow very poorly in the absence of methylase and therefore produce small colonies.

Regulation by methylase--A model. To examine this observation further, the viability and endonuclease levels in cultures containing these plasmids were measured in both the absence and presence of the EcoRI methylase (Table IV). When present, methylase was provided on pSC101-methylase. Inv containing cultures have endonuclease levels 50% that of wild type when the EcoRI methylase is present. This represents more than a 10-fold increase over the levels found in the absence of methylase. The increase in endonuclease seen when methylase was present supports the hypothesis that methylase may act positively to regulate endonuclease expression.

A model for regulation of the EcoRI genes was proposed based on these observations (Fig. 5). Upon introduction of a plasmid carrying the EcoRI genes into a cell, methylase would be produced initially from a weak promoter located near the carboxy terminus of the endonuclease gene. Methylase would subsequently enhance both endonuclease and methylase expression by activating transcription from a promoter located upstream of the endonuclease gene. Transcription from this promoter would result in production of high levels of a bicistronic message. According to this model endonuclease levels would be low until methylase was present in amounts sufficient to protect host DNA. This model could explain how cells survive following transformation with plasmids carrying both the endonuclease and methylase genes. Initial expression of the methylase gene would permit modification of host DNA.

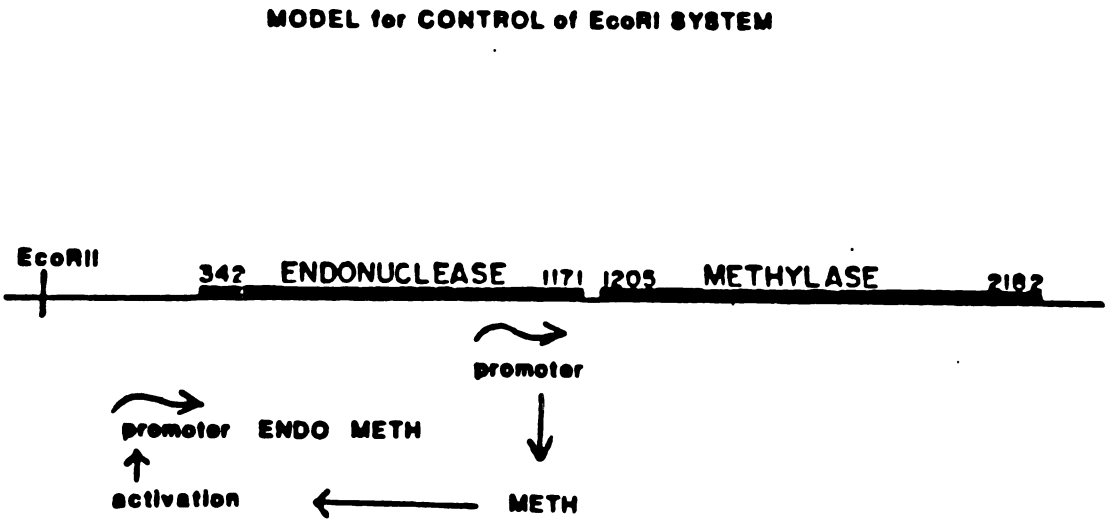


Fig. 5 Model for control of the EcoRI genes--Upon introduction of the EcoRI genes into a host cell, methylase is expressed initially from a weak promoter located near the C-terminus of the endonuclease gene. Methylase acts to protect host DNA and can subsequently activate transcription from a promoter upstream of the endonuclease gene. This produces high levels of a bicistronic message, encoding endonuclease and methylase.

Measurement of Endonuclease Levels

Endonuclease levels in extracts of methylase deficient cultures containing the Inv and Del plasmids were greatly reduced when compared with endonuclease levels expressed from plasmid pMG31-6. Thus, alteration of the methylase gene (which is downstream from the endonuclease gene in pMG31-6) greatly affects endonuclease levels. When methylase was provided (via pSC101-methylase), cultures containing the Inv or Del plasmids had endonuclease levels near that of pMG31-6 containing cultures. These data appeared to provide evidence for the positive regulation model of endonuclease expression.

Extracts of cells grown under a variety of conditions were assayed. A number of controls were performed to increase the accuracy of the endonuclease level determinations, and to rule out several less interesting explanations for the observed "activation". β-lactamase levels were assayed to determine relative plasmid copy mumber.

Endonuclease levels in E. coli D1210 expressed from pMG31-6. E. coli D1210 contains the lac repressor overproducer (lac I^Q) and is $recA^-$. Expression from the lac promoter is low during growth in glucose due to catabolite repression and the high levels of lac repressor resulting from the lac I^Q allele. In glycerol, catabolite repression is relieved, and higher expression levels are obtained. Endonuclease levels in E. coli D1210 are shown in Table V. Cultures containing pMG31-6 grown in glycerol had 18 units of endonuclease per mg protein. Upon IPTG induction this was increased more than 10-fold, to greater than 200. Endonuclease levels were not altered when additional

Table V. EcoRI endonuclease levels in cultures of *E. coli* D1210 containing plasmid pMG31-6.

PLASMIDS	-IPTG	+IPTG
pMG31-6	18	288
pMG31-6 + 101-meth	16	218

Endonuclease levels were determined as described in Methods. Endonuclease assays were performed in 666 buffer. 101-meth = pSC101-meth.

methylase was present in the cell, provided by the plasmid pSC101-methylase.

Endonuclease levels in E. coli D1210 expressed from the Inv and Del plasmids. Levels of endonuclease expressed from the Inv and Del plasmids in both the absence and presence of methylase were determined (Table VI). Cultures containing the Inv plasmid produced low levels of endonuclease. However, when methylase was provided (via either pSC101-methylase or pSY51) endonuclease levels increased to amounts similar to those expressed from pMG31-6. Extracts of cultures containing the Del plasmid (but no EcoRI methylase) had 4 units of endonuclease per mg protein. This level increased to more than 20 in the presence of methylase. These experiments clearly show that endonuclease levels in cultures containing the Inv and Del plasmids are low in the absence of methylase, and become elevated when methylase is provided.

Removal of the lac control region. In addition to the endogenous endonuclease promoter (from pMB1), the Inv and Del plasmids contain the lac promoter region 340 base-pairs upstream of the endonuclease gene. The presence of the lac promoter complicates interpretation of endonuclease activity data. Therefore, the lac control region was removed from the Inv plasmid, generating 322Inv (Fig. 6), a plasmid which is otherwise identical to the Inv plasmid. Table VII shows endonuclease levels expressed from plasmid 322Inv in E. coli D1210. As was seen with the Inv plasmid, there are low levels of endonuclease in cultures lacking the EcoRI methylase. Endonuclease levels increased to

Table VI. EcoRI endonuclease levels expressed from plasmids Inv or Del in *E. coli* D1210 cultures grown in M9 glycerol or M9 glucose.

PLASMIDS	M9 GLYCEROL	M9 GLUCOSE
Inv	2	0.3
Inv + pSC101-meth	15	8
Inv + pSY51	20	ND
Del	ND	4
Del + pSC101-meth	ND	26

Endonuclease levels were determined as described in Methods.
Endonuclease assays were performed in 666 buffer.
ND: Not determined.

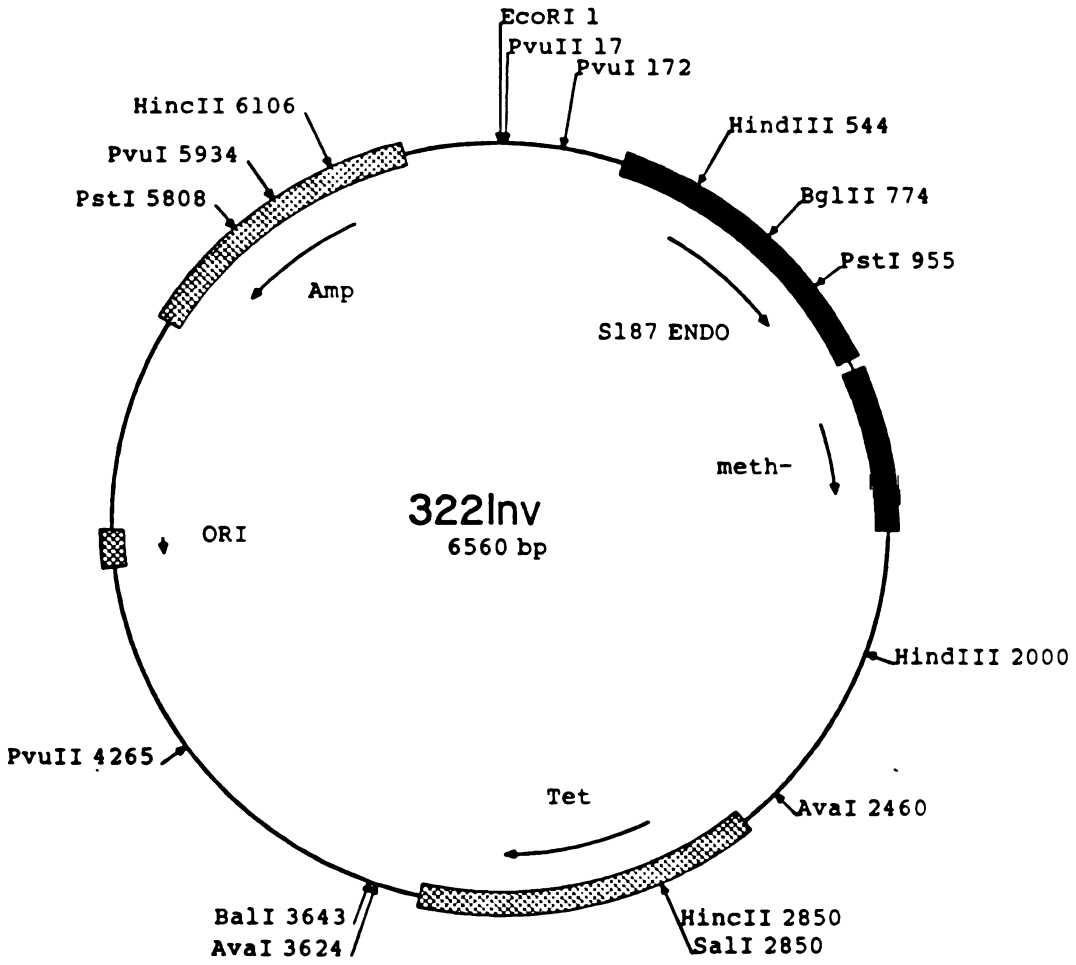


Figure 6. 322Inv--322Inv differs from the Inv plasmid in that the lac control region has been deleted. This plasmid was constructed by digesting plasmids pBR322 and Inv with SalI and EcoRI, ligating and then transforming either *E. coli* D1210 or *E. coli* D1210 containing pSY51. DNA was prepared from ampicillin resistant, tetracycline sensitive colonies and analyzed to confirm the correct constructs.

Table VII. EcoRI endonuclease levels in *E. coli* D1210 cultures grown in M9 glucose or M9 glycerol.

PLASMIDS	M9 GLYCEROL	M9 GLUCOSE
322Inv	2	3
322Inv + 101-meth	31	30
322Inv + pSY51	52	ND
322Inv + 101-meth + pSY51	48	ND

Plasmid 322Inv is identical to the Inv plasmid except that the lac control region has been deleted. pSY51 (pACYC-methylase) encodes the EcoRI methylase on a pACYC177 derivative. pACYC177 has been reported to be present as 22 copies/cell, higher that that reported for pSC101 (6 copies/cell). A map of pSY51 is provided in Figure 7. ND: Not determined. 101-meth = pSC101-meth. ß-lactamase levels were determined as a measure of levels of the endonuclease encoding plasmids. Cultures harboring 322Inv + pSY51 contained 87% of the ß-lactamase levels of cultures containing 322Inv + pSC101-meth. 322Inv cultures lacking methylase contained 83% as much ß-lactamase as pSC101-meth containing cultures.

15-50 with methylase present. This result was similar to that found for cultures containing the Inv plasmid and indicated that methylase elevation of endonuclease levels was independent of the lac control region.

Endonuclease levels in other E. coli strains. The endonuclease measurements described previously were all performed in E. coli D1210. While unlikely, the methylase-dependent increase in endonuclease levels might be a strain specific phenomenon. To address this issue, endonuclease encoding plasmids were transformed into E. coli HB101, another recA$^-$ strain. Results of endonuclease activity measurements in E. coli HB101 are shown in Table VIII. Endonuclease levels expressed from plasmid 322Inv in E. coli HB101 grown in M9 glycerol were 1.3 in the absence of methylase and 11-26 when methylase was present.

Plasmids were transformed into E. coli 294 and endonuclease levels were determined. E. coli 294 cultures grow more rapidly than either E. coli D1210 or E. coli HB101 cultures, since E. coli 294 contains the wild type recA allele. Table VIII shows the results of measurements performed on extracts of E. coli 294. Although there was variation from experiment to experiment, endonuclease levels expressed in the presence of methylase were consistently greater than those expressed when methylase was absent.

Endonuclease assays were also performed on extracts of cultures grown in minimal media in the absence of casamino acids (or at 30^{o}C). Under these poor growth conditions methylase deficient cells grew slowly, and the interesting observation was made that they contained high endonuclease levels.

Table VIII EcoRI endonuclease levels in cultures of *E. coli* HB101 or *E. coli* 294 grown in M9 glycerol or M9 glucose.

PLASMID/ STRAIN	GLYCEROL	GLUCOSE
322Inv/ HB101	1.3	ND
322Inv + 101meth/ HB101	26	ND
322Inv + pSY51/ HB101	11	ND
322Inv/ 294	4	2
322Inv[a]/ 294	15	7
322Inv[b]/ 294	ND	9
322Inv + 101-meth/ 294	25	26
322Inv + 101-meth[a]/ 294	25	41
322Inv + 101-meth[b]/ 294	ND	21

Endonuclease levels were determined as described in Methods. Endonuclease assays were performed in 666 buffer. Cells were grown in M9 glycerol or M9 glucose except where otherwise indicated. Cells grown in the absence of casamino acids were grown for in vivo labelling of proteins with ^{35}S-methionine. These cultures grew more slowly than those grown in the presence of casamino acids. a: Grown in M9 -CAA (with amino acid mix). b: Grown in M9 -CAA (without amino acid mix). ND: Not determined. ß-lactamase levels were determined as an internal control for the level of the EcoRI plasmids. These were as follows: 322Inv containing cultures of *E. coli* HB101 had 52% of the ß-lactamase levels found in cultures containing pSC101-meth. pSY51 containing cultures had 69% as much ß-lactamase as pSC101-meth containing cultures. In *E. coli* 294, cultures lacking methylase contained 91% of the ß-lactamase of methylase containing cultures. 101-meth = pSC101-meth

Methylase Expression

Previous results demonstrated that cultures containing the Inv plasmid had much higher EcoRI endonuclease levels in the presence of methylase than in the absence of methylase. For these experiments methylase was provided by transforming pSC101-methylase (a plasmid which is compatible with the ColE1-related EcoRI plasmids), into cells containing the Inv plasmid. pSC101-methylase contains the EcoRI methylase gene cloned into the HindIII site of pSC101, a low copy number plasmid reported to be present at about 6 copies per cell (52). This compares with 20-30 copies per cell for the ColE1-related EcoRI plasmids (53). The presence of methylase expressed from pSC101-methylase is sufficient to permit cells harboring either the Inv or Del plasmids to survive when S187 endonuclease levels are elevated by IPTG induction. However, pSC101-methylase produces lower levels of methylase than does pMB1 (about 15% as much). Survival of these cells following IPTG induction is due partly to the reduced in vivo activity of the S187 endonuclease. Cells carrying pSC101-methylase and a high copy plasmid encoding wild type EcoRI endonuclease grow poorly in the presence of IPTG.

Construction of methylase-encoding plasmids and determination of methylase expression levels. It was desirable to have the capability of obtaining higher levels of EcoRI methylase. To accomplish this, and to obtain regulated expression of the EcoRI methylase, several additional plasmids were designed. pSY51 (Fig. 7) was constructed by inserting the EcoRI methylase gene from pMB1 into pACYC177. pACYC177

is compatible with the ColE1-derived EcoRI plasmids and has been reported to be present at about 22 copies per cell (50), significantly higher than 6 copies per cell reported for pSC101. The lac control region from the Del plasmid was then inserted into pSY51 to generate pSY71 (Fig. 8). This construction restores the β-lactamase and endonuclease genes and places endonuclease and methylase under lac control. The β-lactamase gene was then removed, generating pSY72 (Fig. 9). To place the methylase directly under lac control, a large portion of the endonuclease gene was deleted, generating plasmids pSY73 (Fig. 10) and pSY74 (Fig.11). pSY71-pSY74 have the methylase gene under lac repressor control and therefore permit regulated methylase synthesis.

Levels of methylase expressed from these plasmids were determined. Table IX shows the methylase levels expressed in E. coli D1210[4] when grown in M9 glycerol. Levels of methylase expressed from plasmids pSC101-methylase, pSY51, pSY71, pSY72, pSY73 and pSY74 were measured in the absence and presence of IPTG. Cultures containing pSC101-methylase and pSY51, which do not contain the lac promoter-operator sequence, were found to have low methylase levels. These levels were

[4] The level of background methylation in E. coli D1210 was determined by quantitating methylase activity present in E. coli D1210 cultures which lacked any plasmid DNA. This background methylation, measured at 2.6 units/mg protein, has been subtracted from all results shown for E. coli D1210.

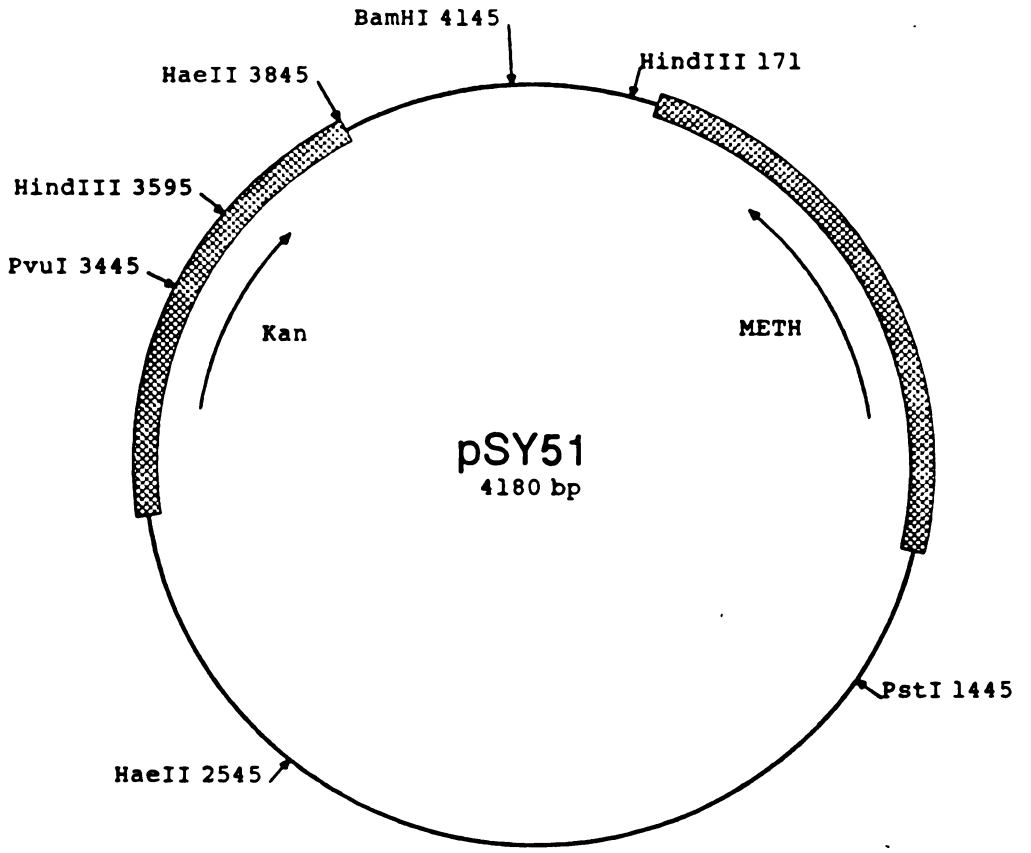


Figure 7. pSY51--pSY51 (pACYC-methylase) encodes the EcoRI methylase and resistance to kanamycin. This plasmid, which was derived from pACYC177, is compatible with the pMB1 derived plasmids. pSY51 was constructed by mixing a PstI and HaeIII digest of plasmid pMB1 with HincII and PstI digested pACYC177. These were ligated and transformed into *E. coli* 294. Kanamycin resistant transformants were screened for sensitivity to ampicillin. After performing DNA screens, methylation tests were conducted to identify transformants which produced EcoRI methylase. Three colonies were found to produce methylase. DNA screens revealed that 3 of these all contained a deletion of DNA, mapped to the region adjacent to the HincII site in pACYC177. These deletions probably resulted from HaeIII digestion of sites near the HincII site in the ampicillin resistance region of pACYC177. pSY51 was used because it is the smallest of the 3 plasmids recovered. The location of the restriction sites are not exact. They have been estimated from the circular map of pACYC177 in Chang et al. (50).

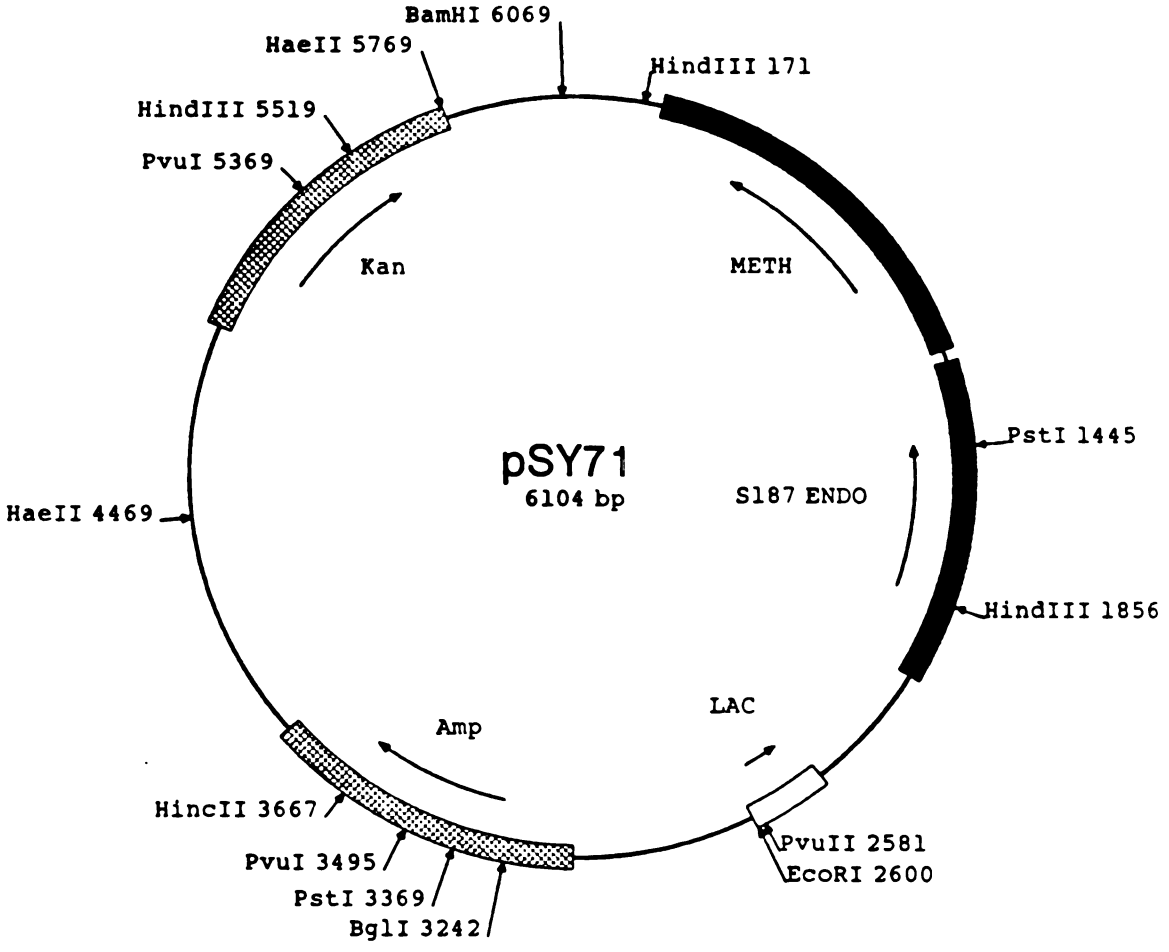


Figure 8. pSY71--pSY71 encodes both the S187 endonuclease and wild type methylase under lac control. This pACYC177 derivitive encodes resistance to ampicillin and kanamycin. pSY71 was constructed as follows: The Del plasmid was digested with PstI and ligated to PstI digested pSY51. Following transformation into *E. coli* 294, kanamycin resistant transformants were selected and screened on x-gal plates for blue colonies.

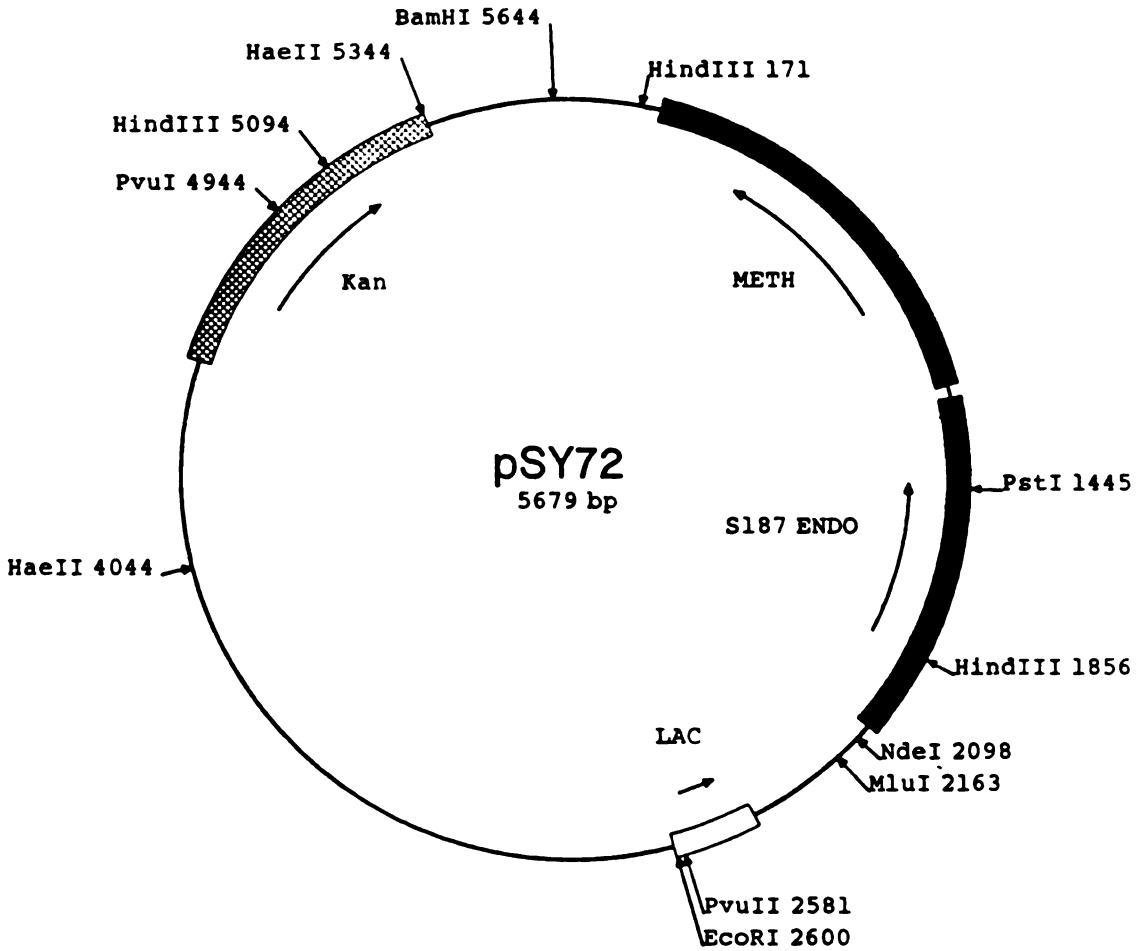


Figure 9. pSY72--This plasmid differs from pSY71 by the removal of the region encoding ampicillin resistance. pSY72 was constructed by digesting pSY71 with HincII and BglI. Overhang ends were removed with DNA polymerase I Klenow fragment. Ends were ligated, and this DNA was transformed into *E. coli* D1210. Kanamycin resistant transformants were selected and then screened for sensitivity to ampicillin.

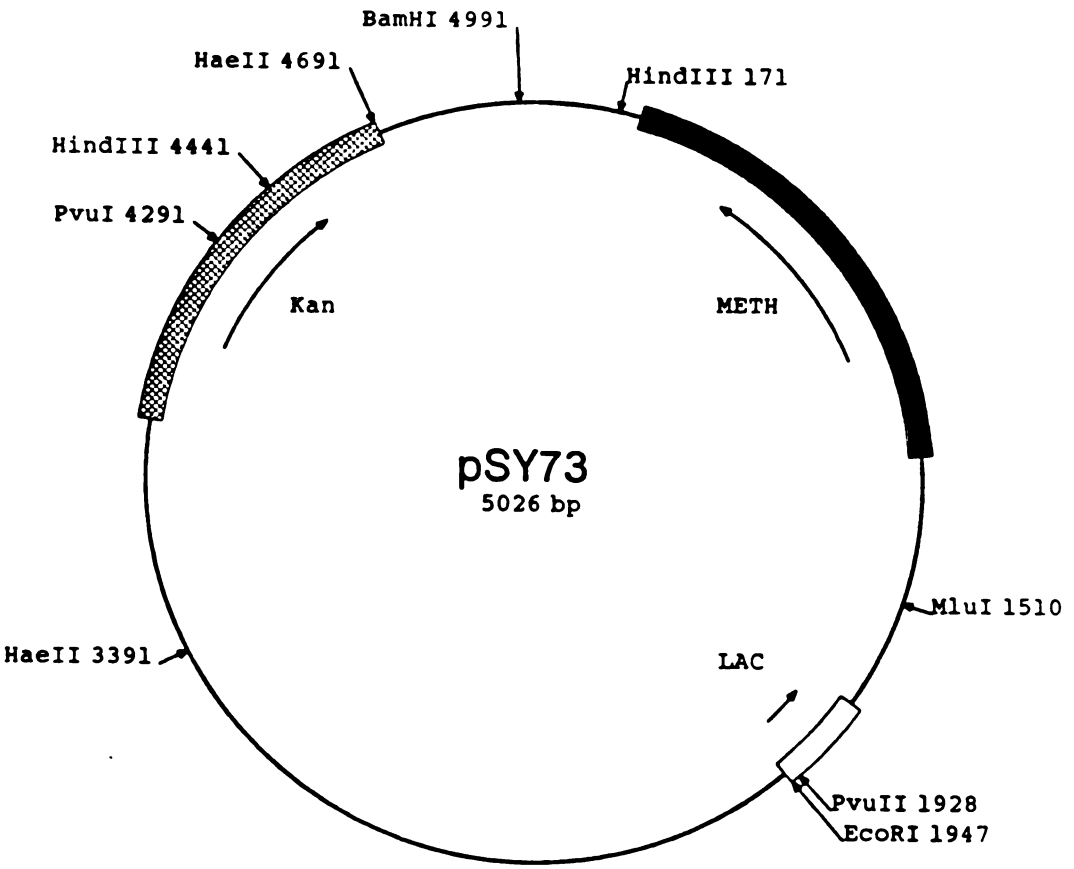


Figure 10. pSY73--pSY73 encodes the EcoRI methylase under lac repressor control and was derived from pSY72 by deleting most of the endonuclease gene. pSY72 was digested with NdeI and PstI. Ends were ligated and this DNA was transformed into *E. coli* D1210. Kanamycin resistant transformants were then selected. DNA isolated from these transformants was then analyzed on agarose gels to determine the correct constructions. This particular construction results from a deletion of 653 base pairs of the endonuclease gene on pSY72.

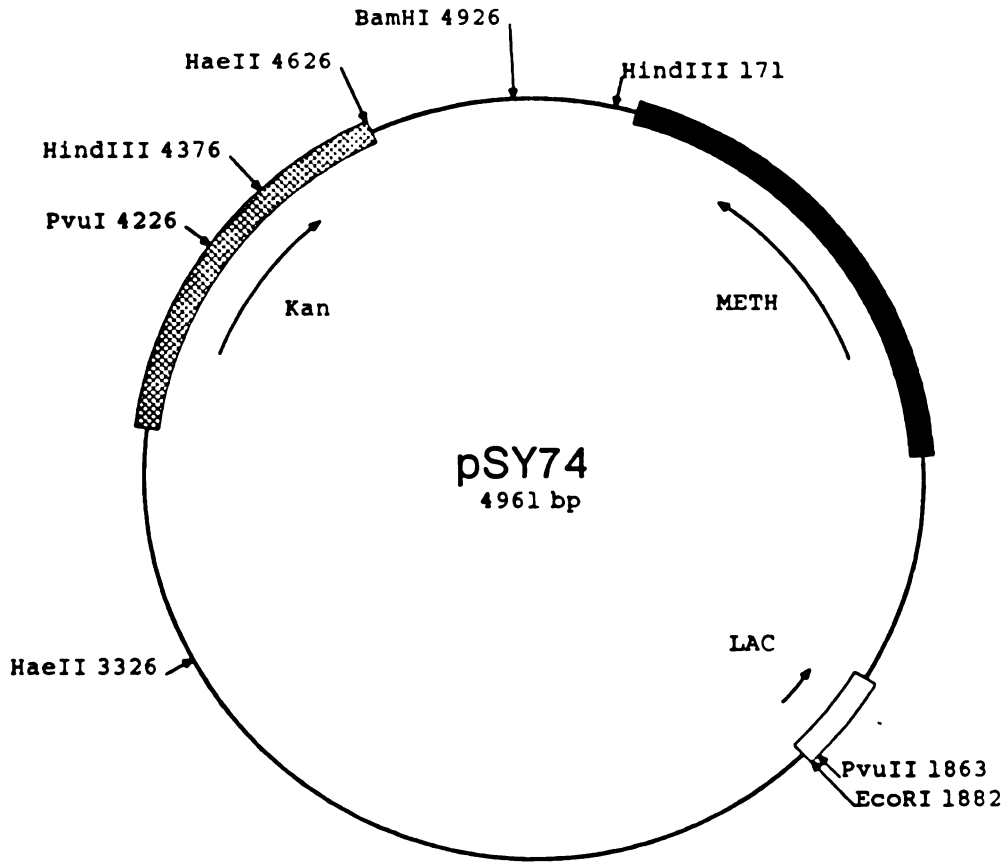


Figure 11. pSY74--pSY74 is similar to pSY73 but contains a deletion of a greater portion of the endonuclease gene than in pSY73. pSY74 was constructed by digesting pSY72 with MluI and PstI. Following ligation and transformation into *E. coli* D1210, kanamycin resistant transformants were selected. This results in a 718 base pair deletion from pSY72.

Table IX. EcoRI methylase levels in *E. coli* D1210 cultures grown in M9 glycerol

PLASMID	-IPTG	+IPTG
pSC101-meth	2	2
pSY51	6	6
pSY71	42	200
pSY72	37	301
pSY73	19	128
pSY74	10	98

The plasmids used in this experiment all encode the EcoRI methylase and are compatible with the endonuclease-encoding pBR322-derived plasmids. pSY51 and pSY71-74 are derived from pACYC177 and are all described in Figures 7-11.

not altered by IPTG induction. pSY71-74 containing cultures were found to contain significantly higher methylase levels. These levels were increased by IPTG induction.

A detailed analysis of methylase levels expressed from pSC101-methylase was performed. To examine the possibility that methylase levels varied as a function of the tetracycline concentration in the growth media, methylase levels were measured from extracts of cells grown at different tetracycline concentrations. Methylase levels in E. coli D1210 expressed from pSC101-methylase grown in M9 glycerol containing different tetracycline concentrations are shown in Table X. In this experiment pSC101-methylase was found to express about 3.5 units of methylase/mg protein. This level was independent of the tetracycline concentration.

Expression from the methylase promoter on pMG31-6. The existence of a separate promoter for methylase expression was proposed in our model of endonuclease and methylase expression (Fig. 5). Such a promoter could account for the initial expression of methylase upon transformation of the endonuclease and methylase genes into a recipient cell. Methylase expressed from this promoter could subsequently activate high level expression from a promoter upstream of the endonuclease gene, resulting in a bicistronic message. Data from in vitro transcription experiments using purified DNA fragments indicated that transcription could be initiated from a promoter located at the C-terminal end of the endonuclease gene (S.Y., F.S. unpulbished). Evidence for the existence of a promoter in this region has been reported by O'Connor and Humphreys (54). They isolated phage DNA from cells harboring a plasmid

Table X. EcoRI methylase levels in *E. coli* D1210 cultures grown in M9 glycerol

PLASMID	TET (μg/ml)	METH S.A.
pSC101-meth	0	3.6
pSC101-meth	2	3.8
pSC101-meth	5	3.5

Methylase levels (units/mg protein) were determined as described in Methods. Cells containing pSC101-meth were grown in the presence of 2 μg/ml tetracycline, 5 μg/ml tetracycline or in the absence of tetracycline.

which contained a transcriptional terminator inserted within the endonuclease gene. This DNA was resistant to EcoRI cleavage, presumably due to modification by the EcoRI methylase. Therefore methylase can be expressed from a promoter located near the C-terminus of the endonuclease gene. However, the actual level of methylase expression was not determined in these experiments.

To permit determination of in vivo expression levels from the proposed methylase promoter, the RpoBC transcriptional terminator (51) from plasmid pAD1 (47) was inserted at the BglII site in the middle of the endonuclease gene of pMG31-6. This terminator was inserted in both orientations. This effectively blocks transcription initiated upstream. Methylase expressed from the resulting constructs, pSY31 and pSY32 (Fig. 12), must therefore be transcribed from promoter sequences located between the terminator sequence and the start of the methylase gene. Efficient termination at the RpoBC terminator should result in methylase expression from pSY31 and pSY32 which is insensitive to IPTG induction of the lac promoter.

EcoRI methylase levels expressed from plasmids pMG31-6, pSY31 and pSY32 in E. coli 294 are shown in Table XI. pMG31-6 containing cultures were found to express over 600 units of methylase/mg protein while both pSY31 and pSY32 containing cultures express about 9 units/mg protein. Methylase levels produced from pSY31 and pSY32 were not elevated in the presence of IPTG, while pMG31-6 methylase levels increased slightly upon IPTG induction, to more than 800. These results demonstrate low level methylase expression from a presumed promoter sequence located downstream from the BglII site in pMG31-6, i.e. within the C-terminal portion of the endonuclease gene. The

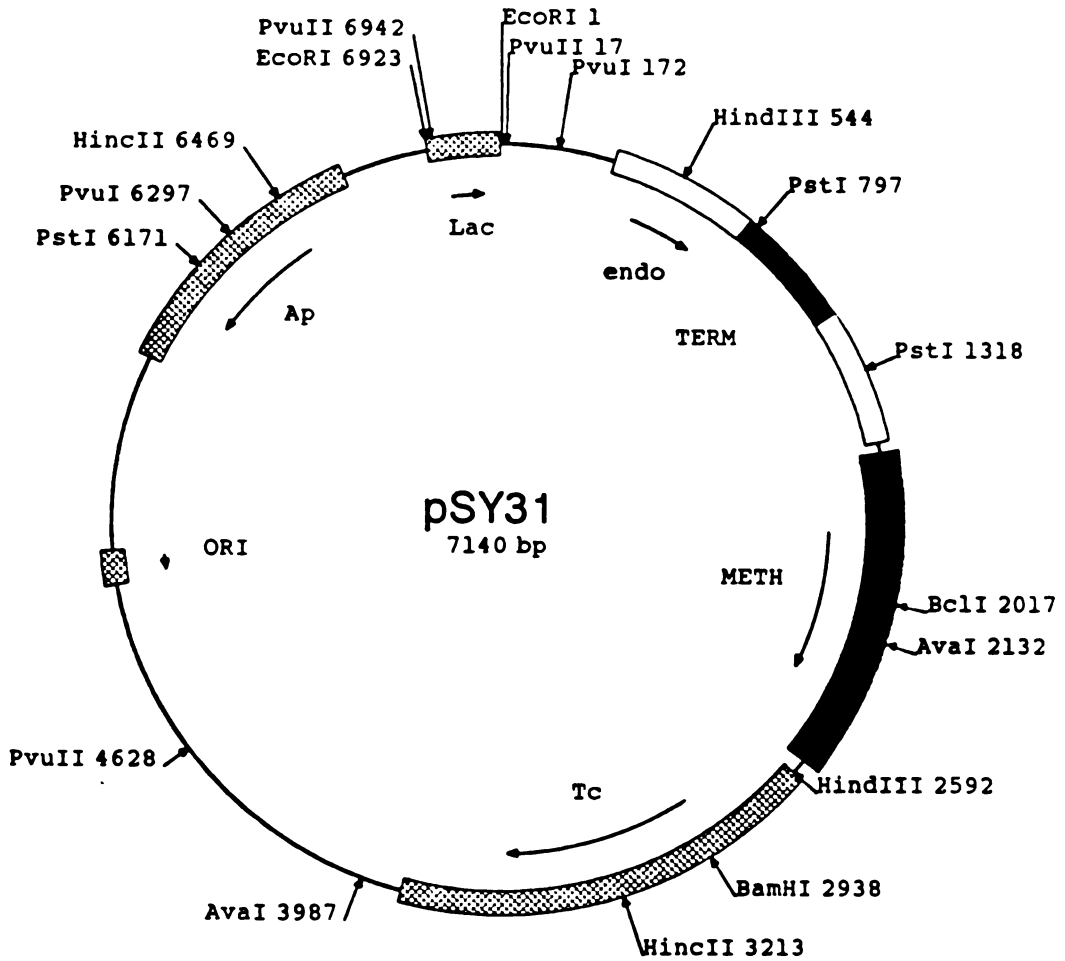


Figure 12. pSY31 and pSY32--These plasmids contain the RpoBC transcriptional terminator inserted into the BglII site of pMG31-6 (located within the EcoRI endonuclease gene). This terminates transcription originating upstream, i.e. from the lac or endonuclease promoters. These plasmids encode resistance to ampicillin and tetracycline. pSY31 and pSY32 were constructed from pMG31-6 as follows: BglII digested pMG31-6 was ligated to a 350 base pair BamHI fragment from pAD1 carrying the RpoBC terminator. This was then transformed into *E. coli* 294 and ampicillin resistant transformants were selected. From these transformants, DNA was isolated and alalyzed to determine the correct construct. pSY31 contains the transcriptional terminator inserted in its normal orientation, while pSY32 contains the terminator inserted in the opposite orientation.

Table XI. EcoRI methylase levels in uninduced and induced *E. coli* 294 cultures grown in M9 glycerol.

PLASMID	-IPTG	+IPTG
pMG31-6	624	873
pSY31	8.5	8.4
pSY32	9.2	9.0

Methylase levels (units/mg protein) were determined as described in Methods. pSY31 contains the RpoBC transcriptional terminator (which has been shown to terminate transcription in either orientation) inserted within the endonuclease gene in pMG31-6. pSY32 contains the terminator inserted in the reverse orientation.

insensitivity to IPTG induction serves as a control and indicates that the observed activity is not due to read through expression from the lac promoter.

Generation and characterization of a temperature sensitive methylase. Methylase encoding mRNA might play a role in regulation of endonuclease expression. Unstable RNA expressed from the Inv plasmid might be stabilized when pSC101-methylase is present. Endonuclease mRNA might be stabilized directly by the methylase protein, or through interaction with complementary anti-sense RNA expressed from pSC101-methylase. Generation of a temperature sensitive methylase protein permits determination of whether the methylase protein itself is directly involved in regulation of endonuclease expression. With a temperature sensitive methylase, it is possible to determine endonuclease levels expressed at temperatures at which methylase is active, or inactive.

To generate a temperature sensitive methylase, pPG31 DNA was mutagenized in vitro with hydroxylamine (55) and transformed into E. coli 294. From a large number of transformants, two were obtained which produced normal sized colonies at 30°C but did not produce colonies at 42°C. These were designated ts5 and ts8. Plasmid DNA was isolated from these temperature sensitive methylase candidates and transformed into E. coli D1210 cells which contained either pSY51 or pSC101-methylase. This was performed to determine whether the wild type methylase gene could complement the defect which resulted in inability to grow at 42°C. When either pSC101-methylase or pSY51 were present, cells containing ts5 and ts8 produced normal sized colonies at 42°C, demonstrating that wild type methylase could complement.

The growth properties of cells harboring these plasmids were studied in a temperature-shift experiment. Cells were grown at 30°C to early log phase, a portion of the cultures shifted to 42°C and both OD_{600} and viable cell counts were determined at intervals up to 4 hours. The results of this experiment are shown in Fig. 13. Cells harboring either wild type methylase or ts8 grew at 30°C and 42°C, as determined by OD_{600}. However, OD_{600} values for cultures containing ts5 levelled off after a shift to 42°C. When the number of viable cells in each culture were determined, it was found that cultures containing wild type methylase grew at 42°C. Cultures containing ts8 grew for 3 hours following a shift to 42°C. After 3 hours the number of viable cells decreased. Cultures containing ts5 grew slowly at 30°C and contained greatly reduced numbers of viable cells when grown at 42°C.

The in vitro activity of the mutant methylases was determined. Cells extracts were prepared as described in Methods and methylase assays were then performed at 30°C, 37°C and 42°C. The pPG31 extract, which contains the wild type EcoRI methylase, contained the same methylase levels at all 3 temperatures (Table XII). Extracts of cells harboring ts5 contained a low methylase activity at all temperatures. Surprisingly, extracts containing ts8 had a methylase activity 3 times that of wild type at 30°C. This value was reduced to 50% of wild type at 42°C, indicating that methylase expressed from ts8 was temperature sensitive. Although this methylase was temperature sensitive, ts8 extracts still contained significant levels of methylase activity at the restrictive temperature.

Our goal was to obtain a mutant methylase which retained normal levels of activity at 30°C but lacked activity at 42°C. With such a

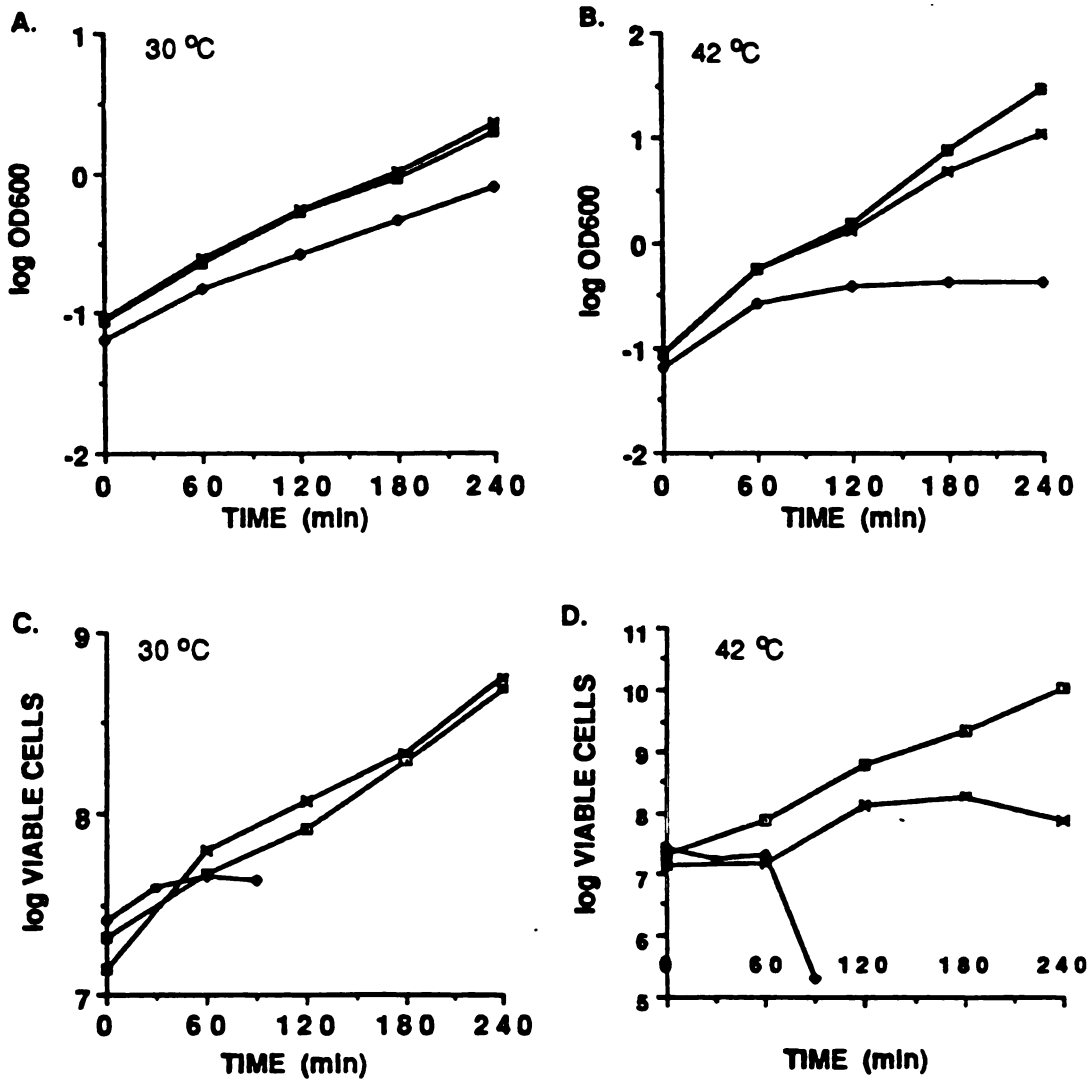


Fig. 13 Growth properties of cultures harboring plasmids encoding the wild type or mutant EcoRI methylase. Cultures were grown at 30°C to early log phase, split in half, and grown at 30°C or 42°C. Samples were removed at the indicated time points and the OD600 and number of viable cells were determined. -⊡- represents cultures encoding the wild type methylase. -◆- represents cultures containing a mutant methylase encoded by plasmid ts5. -✱- represents cultures containing plasmid ts8. A and B are the OD600 of cultures grown at 30°C and 42°C respectively. C and D are the number of viable cells in cultures grown at 30°C and 42°C respectively.

Table XII. Activity of wild type and mutant methylases at 30°C, 37°C and 42°C.

Plasmid/Strain	Meth S.A. at 30°C	37°C	42°C
pPG31/294	31.4	28.9	32.6
ts5/294	3.6	2.3	3.6
ts8/294	93	21	17.6

Plasmid pPG31 encodes a wild type EcoRI methylase. Plasmids ts5 and ts8, which encode mutant methylases, were isolated following hydroxylamine mutagenesis of plasmid pPG31. Cultures were grown and methylase assays were performed as described in Methods except that the methylase assays were done at the indicated temperatures.

mutant it would be possible to alter the amount of methylase activity in cells harboring the Inv plasmid by simply raising the temperature. However, neither of the methylase mutants described here meet these requirements. For this reason, no further experiments were performed with the temperature sensitive methylase mutants.

Use of β-galactosidase Fusion Plasmids

Fusion plasmids are useful tools in the study of regulation of gene expression. Endonuclease-β-galactosidase fusions were constructed to aid in determining whether methylase affected either transcription or translation of endonuclease. By placing β-galactosidase downstream from the EcoRI endonuclease control sequences, expression from the endonuclease promoter could be monitored by measuring β-galactosidase activity, which is easier to determine and quantitate than endonuclease activity. The effect of methylase on expression of a β-galactosidase gene fused to the endonuclease control region could then be determined.

Generation of operon fusion and gene fusion plasmids. pSY60 (Fig. 14) was constructed by inserting the EcoRI-BglII region from the Del plasmid into the operon fusion vector, pMLB1010. This operon fusion places the β-galactosidase gene downstream from the endonuclease control region, but does not produce a fused endonuclease-β-galactosidase polypeptide. Resulting β-galactosidase levels reflect transcriptional (but not translational) activity from the inserted endonuclease control sequences.

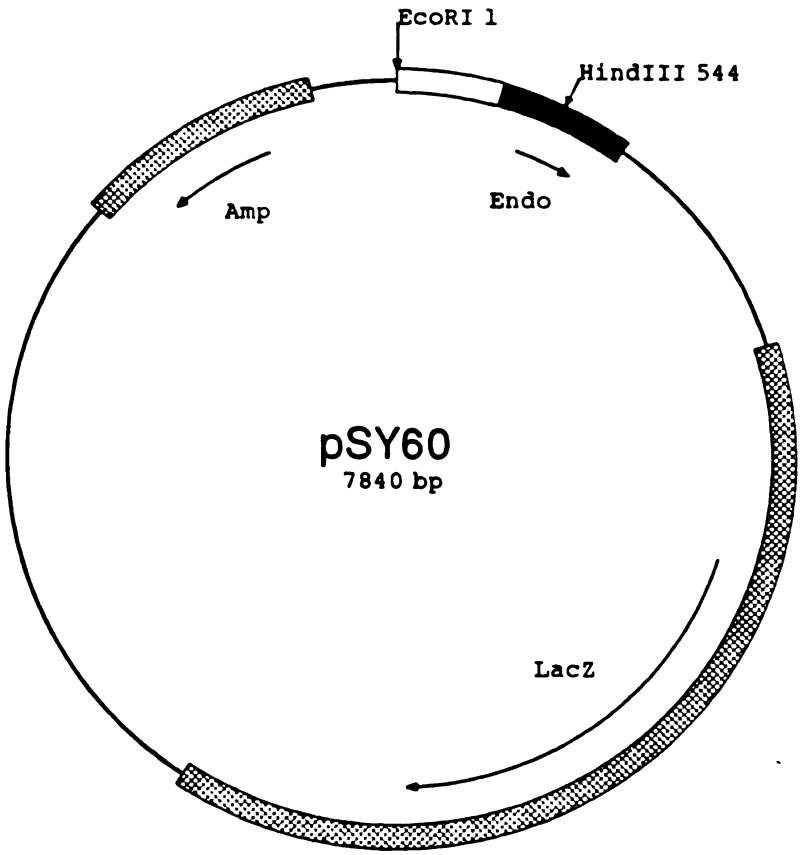


Figure 14. pSY60--pSY60 has the EcoRI-BglII fragment from pSY22 inserted into pMLB1010 upstream of the ß-galactosidase gene. This places ß-galactosidase under transcriptional control from endonuclease promoter sequences and expresses a truncated endonuclease protein (145 amino acids). pSY60 contains an intact ß-galactosidase gene and encodes resistance to ampicillin. pSY60 was constructed from pMLB1010 and pSY22. pMLB1010 was digested with BamHI and pSY22 was digested with BglII. These were mixed, digested with EcoRI and ligated. Following transformation into *E. coli* 294, ampicillin resistant transformants were selected.

In order to observe translational control, a gene fusion was constructed. The endonuclease gene was fused in phase to β-galactosidase by inserting the EcoRI-HindIII region from the Del plasmid into pMLB1034. The resulting plasmid, pSY64 (Fig. 15) places β-galactosidase under both transcriptional and translational control of EcoRI endonuclease signals.

β-galactosidase levels expressed in E. coli JM103. β-galactosidase levels were measured in cultures carrying the operon fusion, pSY60, or the gene fusion, pSY64. Initial experiments were performed E. coli JM103 (Table XIII). E. coli JM103 alone was found to produce no detectable β-galactosidase, while E. coli JM103 cultures carrying the plasmid pMLB1010 produced 700 units. Cultures harboring pSY60 contained 1100 units of β-galactosidase. The presence the EcoRI methylase did not significantly alter β-galactosidase levels.

β-galactosidase levels in E. coli D1210. The previously described experiments, in which methylase affected endonuclease levels expressed from the Inv plasmid, were performed in E. coli D1210. The fusion plasmids were therefore transformed into E. coli D1210 and β-galactosidase assays were performed (Table XIV). E. coli D1210 carrying plasmid pMLB1010 was found to contain 500 units of β-galactosidase. β-galactosidase levels expressed from cultures containing pSY60 were higher, about 700 units. The presence of methylase in the cell, encoded by either pSC101-methylase or pSY51, did not effect β-galactosidase levels. Cultures containing pMLB1034 contained no detectable β-galactosidase activity. Cultures containing

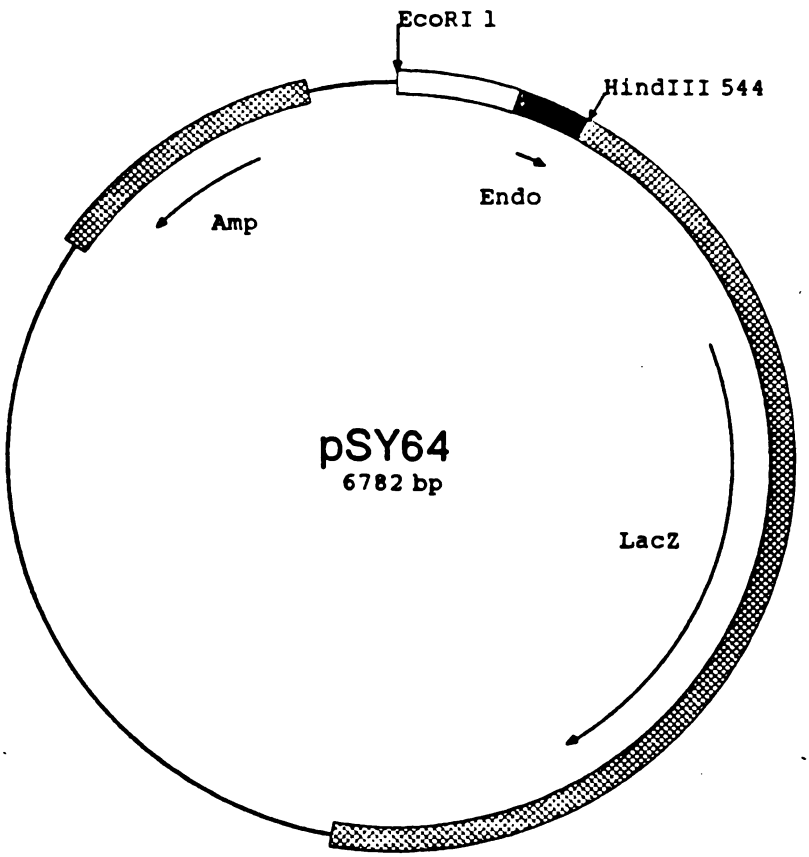


Figure 15. pSY64--pSY64 encodes a fused polypeptide consisting of the N-terminal 67 amino acids of the EcoRI endonuclease fused to amino acid 8 of ß-galactosidase. This plasmid contains 342 base pairs upstream of the endonuclease gene from pSY22. ß-galactosidase is placed under transcriptional and translational control of the EcoRI endonuclease. pSY64 was constructed by mixing BamHI digested pMLB1034 with HindIII digested pSY22. Ends were filled in with DNA polymerase I Klenow fragment. Following digestion with EcoRI and treatment with T4 DNA ligase, this was transformed into *E. coli* 294 and ampicillin resistant transformants were selected.

Table XIII. ß-galactosidase levels in *E. coli* JM103 cultures grown in LB.

PLASMID	B-GAL
none	<1
pMLB1010	695
pSY60	1,069
pSY60 + 101-meth	1,135
pSY60 + pSY51	1,036

ß-galactosidase assays were performed as described in Methods. Plasmid pMLB1010, the operon fusion vector, was used in the generation of pSY60. pSY60 places ß-galactosidase under transcriptional control of the endonuclease promoter.
101-meth = pSC101-meth

Table XIV. ß-galactosidase levels in *E. coli* D1210 cultures grown in LB.

PLASMID	B-GAL
pMLB1010	501
pSY60	681
pSY60 + pSC101-meth	591
pSY60 + pSY51	698
pMLB1034	<1
pSY64	1,181
pSY64 + pSC101-meth	1,256
pSY64 + pSY51	1,350

pMLB1010 and pSY60 are the operon fusion vector and operon fusion respectively. pMLB1034 is the protein fusion vector. pSY64 contains a gene fusion encoding the N-terminal 67 amino acids of the EcoRI endonuclease fused in phase to ß-galactosidase.

the gene fusion, pSY64, had about 1200 units of β-galactosidase. These levels were slightly higher with methylase present (1250 when methylase was provided via pSC101-methylase, 1350 with pSY51 present).

β-galactosidase levels in E. coli MC4100 cultures. The fusion plasmids were transformed into the lac$^-$ strain, E. coli MC4100. β-galactosidase levels were determined from cultures grown under different conditions (Table XV). When grown in LB, pSY64 containing cultures had about 1300 units of β-galactosidase, both in the absence and presence of methylase. When grown in M9 glycerol or M9 glucose, β-galactosidase levels were found to be elevated slightly in the presence of methylase. This was true for cultures containing the operon fusion (pSY60) and the gene fusion (pSY64).

Insertion of a transcriptional terminator. High β-galactosidase levels were observed in cultures containing the operon fusion vector, pMLB1010. This background complicated the interpretation of data on β-galactosidase expressed from the fusion plasmids. β-galactosidase expression from pMLB1010 results from promoter sequences upstream of the β-galactosidase gene. To minimize read through transcription from upstream of the endonuclease control region, the RpoBC transcriptional terminator was inserted into pMLB1010 and pMLB1034. This terminator was inserted to construct plasmids pSY60T, pSY64T, pMLB1010T and pMLB1034T. Following transformation of these plasmids into E. coli MC4100, β-galactosidase assays were performed (Table XVI). With the transcriptional terminator inserted into pMLB1010, a slight decrease in β-galactosidase levels was observed. β-galactosidase levels in

Table XV. ß-galactosidase levels in *E. coli* MC4100 grown in LB, M9 glucose or M9 glycerol.

PLASMIDS	LB	GLYCEROL	GLUCOSE
none	<1	ND	ND
pMLB1034	<1	<1	<1
pSY64	1,292	1,743	1,392
pSY64 + 101-meth	1,238	2,267	1,615
pSY64 + pSY51	1,278	2,109	1,587
pMLB1010	ND	906	ND
pSY60	ND	821	ND
pSY60 + 101-meth	ND	1,118	ND
pSY60 + pSY51	ND	1,195	ND

Cultures containing the designated plasmids were grown in LB, M9 glycerol or M9 glucose. ß-galactosidase levels were determined as described in Methods. Plasmids pMLB1010 and pMLB1034 are the operon fusion and gene fusion vectors respectively. Plasmid pSY60 places ß-galactosidase under transcriptional control of the EcoRI endonuclease. pSY64 encodes a fused polypeptide consisting of the N-terminal 67 residues of endonuclease fused to ß-galactosidase.
101-meth = pSC101-methylase. ND: Not determined

Table XVI. ß-galactosidase levels in *E. coli* MC4100 cultures grown in M9 glycerol or M9 glucose.

PLASMIDS	M9 GLUCOSE	M9 GLYCEROL -IPTG	M9 GLYCEROL +IPTG
none	60	ND	ND
pMLB1010	1,093	ND	ND
pMLB1010T	870	ND	ND
pSY60	ND	2,238	1,637
pSY60T	1,753	2,554	1,838
pSY60T + 101-meth	1,236	ND	ND
pSY60T + pSY51	1,340	3,375	2,594
pSY60T + pSY71	41	3,286	2,403
pSY71	ND	<1	20
pMLB1034	<1	ND	ND
pMLB1034T	ND	ND	<1
pSY64T	1,562	1,908	ND
pSY64T + 101-meth	1,842	1,889	ND
pSY74T + pSY51	1,784	1,811	ND

ß-galactosidase levels were determined as described in Methods. pMLB1010T and pMLB1034T are identical to pMLB1010 and pMLB1034 except that they contain a 350 base pair transcriptional terminator inserted into the EcoRI site. pSY60T and pSY64T also contain this terminator inserted at the same location 101-meth (pSC101-methylase) and pSY51 encode the EcoRI methylase. pSY71 encodes both the EcoRI endonuclease and methylase under lac repressor control. ND: Not determined.

cultures containing these plasmids were the same in the absence or presence of methylase.

Methylase was provided in cells containing the fusion plasmids by either pSC101-methylase or pSY51. To test the possibility that methylase activation of endonuclease expression required an intact endonuclease protein, plasmid pSY71 (which encodes both endonuclease and methylase) was transformed into cells containing the fusion plasmids. The presence of plasmid pSY71 did not cause an increase in β-galactosidase levels in cultures harboring the fusion plasmids (Table XVI).

Construction of a gene fusion containing 75% of the endonuclease gene. In plasmid pSY64, the β-galactosidase gene was fused at the HindIII site within the endonuclease gene. This generates a fusion protein containing 67 amino acids of endonuclease (the N-terminal 25% of the gene) fused in phase to β-galactosidase. β-galactosidase levels in cultures harboring this plasmid were not significantly altered by the presence of methylase. This result seemed inconsistent with methylase-dependent transcriptional activation as proposed in our model. However, it seemed possible that the sequences required for methylase action might be downstream of the HindIII site. To test this possibility, a fusion plasmid was constructed which linked three-fourths of the endonuclease gene to β-galactosidase. This plasmid, pSY81, fuses the β-galactosidase gene at the PstI site within the endonuclease gene, thereby encoding a fused polypeptide containing 204 amino acids (about 75%) of the endonuclease. Levels of β-galactosidase were measured from cultures harboring this plasmid (Table XVII).

Table XVII. ß-galactosidase activity in *E. coli* MC4100 cultures grown in M9 glucose or M9 glycerol.

PLASMIDS	M9 GLUCOSE -IPTG	M9 GLUCOSE +IPTG	M9 GLYCEROL
pSY81	157	153	406
pSY81 + 101-meth	292	ND	449
pSY81 + pSY51	195	ND	417
pSY81 + pSY72	215	209	ND
pSY81 + pSY73	341	300	ND
pSY51	7	ND	ND
pSY72	<1	3	ND

ß-galactosidase levels were determined as described in Methods. pSY81 encodes a fused polypeptide consisting of the N-terminal 204 amino acids of the EcoRI endonuclease (75% of the wild type protein) fused to ß-galactosidase at amino acid 8 in ß-galactosidase. pSC101-meth (101-meth) and pSY51 encode the EcoRI methylase. pSY72 and pSY73 encode the EcoRI methylase under lac repressor control.
ND: Not determined.

β-galactosidase levels in pSY81 containing cultures were slightly higher when methylase was present. Cultures grown in glycerol contained elevated β-galactosidase levels which were not altered by the presence of methylase. β-galactosidase levels in E. coli MC4100 cultures containing pSY81 (with and without pSY72 or pSY73) are shown in Table XVII. Methylase containing cultures had up to twice as much β-galactosidase as cultures lacking the methylase. Increased methylase levels (IPTG induction) did not further elevate β-galactosidase levels.

The RpoBC transcriptional terminator was inserted at the EcoRI site in pSY81, producing pSY84. β-galactosidase levels in cultures containing pSY84 in the absence and presence of methylase are shown in Table XVIII. Methylase containing cultures had somewhat higher β-galactosidase levels than methylase deficient cultures. It was also observed that β-galactosidase levels in cultures containing methylase were greater when antibiotics were provided in the growth medium, or when cultures were grown at 30°C.

Rates of Endonuclease Synthesis and Turnover

Data presented here demonstrate an increase in endonuclease expression from the Inv plasmid in the presence of methylase. This increase was not observed when the components of this system were rearranged onto fusion plasmids. Results of experiments performed with fusion plasmids seemed inconsistent with those obtained from the endonuclease activity determinations. To clarify this apparent contradiction, the rates of endonuclease synthesis were measured directly from cultures lacking or containing the EcoRI methylase.

Table XVIII. ß-galactosidase levels in *E. coli* MC4100 cultures grown in M9 glucose.

PLASMID	M9 Glu	M9 Glu + ANTIBIOTICS	M9 Glu at 30°C
pSY84	220	146	352
pSY84 + 101-meth	352	483	677
pSY84 + pSY51	298	292	448

ß-galactosidase levels were determined as described in Methods. Cultures were grown in M9 glucose at 37°C except where indicated otherwise. When indicated, pSY84 containing cultures were grown in the presence of 20 µg/ml kanamycin. 101-meth = pSC101-methylase

Pulse-chase labelling experiments were performed to permit measurement of endonuclease synthesis rates and to provide data on the rate of turnover of endonuclease within the cell.

Endonuclease expressed from plasmid 322Inv. The rate of endonuclease synthesis was measured in E. coli 294 and E. coli D1210 cultures harboring plasmid 322Inv (data not shown). Two to three times as much endonuclease was synthesized in E. coli 294 cultures containing the EcoRI methylase as in cultures lacking methylase. Extracts of these cells contained 2-3 fold higher endonuclease levels when methylase was present than in the absence of methylase. Unfortunately, E. coli D1210 cultures containing pSC101-methylase did not grow under these conditions, so nothing can be said about E. coli D1210. In the E. coli 294 cultures, the endonuclease protein was very stable, with no detectable decrease in concentration during the 60 minute chase period. A similar experiment was performed with E. coli 294 (Fig. 16). Here, the amount of endonuclease synthesized was 2-3 times higher in the presence of methylase, while endonuclease activity levels were 10 times higher with methylase present.

Endonuclease expression from plasmid pPG31. Plasmid pPG31 encodes wild type endonuclease and methylase under lac repressor control. Endonuclease expressed in cultures harboring pPG31 was monitored by labelling protein with ^{35}S-methionine. The stability of endonuclease was followed by observing levels of labelled protein during a 30 minute chase period. High levels of endonuclease were expressed from pPG31 and this endonuclease was stable during the 30 minute chase period.

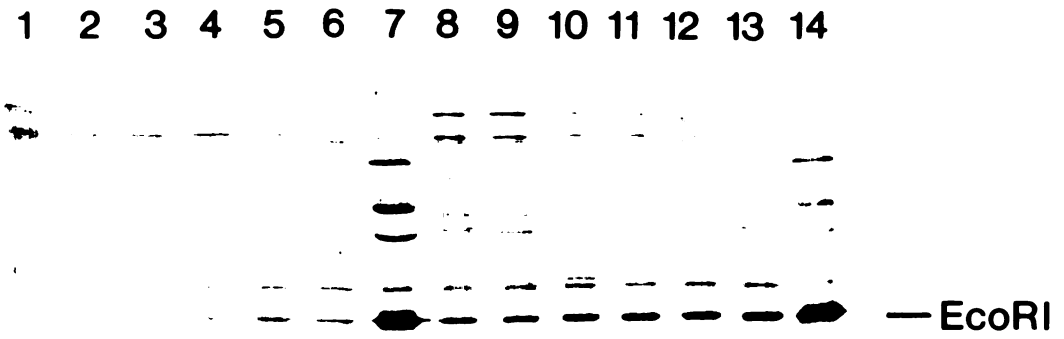


Fig. 16 EcoRI endonuclease produced from plasmid 322Inv in *E. coli* 294 cultures containing or lacking the EcoRI methylase. Cultures were labelled for 1-min. with ^{35}S-methionine, excess unlabelled methionine added, and aliquots removed at 1 min., 2 min., 5 min., 10 min., 30 min., and 60 min. Lanes 1-6: Cultures lacking methylase,1 min., 2 min., 5 min., 10 min., 30 min., 60 min., incubated with antibody to the EcoRI endonuclease. Lanes 8-13 are the same time points from cultures containing the EcoRI methylase. Lanes 7 and 14 are 60 min. samples treated with antibody to ß-lactamase.

Levels of endonuclease were greater when cultures were grown in glycerol than when cultures were grown in glucose. A similar experiment was performed with the addition of measurement of β-galactosidase levels. Since β-galactosidase is encoded by the bacterial chromosome (and not the resident plasmid), β-galactosidase levels should reflect overall cell concentration. Both endonuclease and β-galactosidase were stable during the 60 minute chase period (Fig. 17).

Generation and Analysis of Endonuclease Null Mutants

Strains harboring the S187 endonuclease were shown to have limited ability to restrict phage λ. While the S187 endonuclease is not very active *in vivo*, this enzyme does have some activity under *in vivo* conditions. This low level of activity could be harmful to cells harboring this enzyme. Differential viability in cultures containing methylase versus those lacking methylase would present a problem in interpretation of endonuclease activity data obtained with the Inv or Del plasmids. The variation in endonuclease levels in methylase deficient cultures might be due to such residual activity. To eliminate this complication, mutant endonucleases were generated which which were catalytically inactive (null mutants).

Generation and isolation of null mutants. Plasmid pKG2 (31) was the vector used in the generation of endonuclease mutants. This plasmid encodes wild type endonuclease under lac repressor control but does not encode the *Eco*RI methylase. In addition, pKG2 encodes resistance to

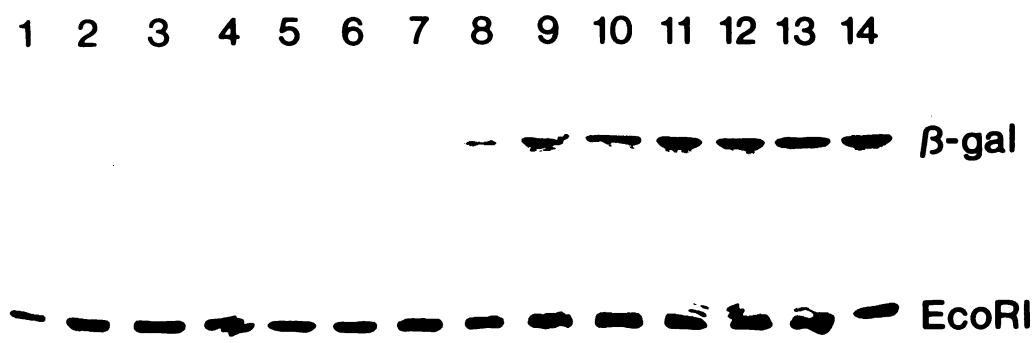


Fig. 17 EcoRI endonuclease and β-galactosidase expressed in *E. coli* 294 cultures containing plasmid pPG31. Cells were labelled for 30 sec. with ^{35}S-methionine, excess unlabelled methionine was added, and duplicate aliquots were removed at the indicated time points. Lanes 1-7: 1 min., 2 min., 5 min., 10 min., 30 min., and 60 min. aliquots incubated with antibody to the EcoRI endonuclease. Lanes 8-14: 1 min., 2 min., 5 min., 10 min., 30 min., 60 min. aliquots incubated with antibody to the EcoRI endonuclease and with antibody to ß-galactosidase.

kanamycin. pKG2 can be propagated in strains which contain methylase (generally methylase is provided via pSC101-methylase). pKG2 DNA can be isolated free of pSC101-methylase by purifying plasmid DNA, linearizing the pSC101-methylase plasmid with BamHI, and rebanding in CsCl. pKG2 was mutagenized *in vitro* with hydroxylamine (55) and transformed into *E. coli* D1210. Kanamycin resistant transformants were checked for viability in the presence of IPTG. Protein was isolated from these cells and analyzed on western blots. Approximately 50% of these contained protein which cross reacted with the EcoRI endonuclease and was the same molecular weight as the wild type endonuclease. These were termed western positive. The western positive mutants contained amounts of endonuclease cross-reacting protein roughly equivalent to that expressed from pKG2. The entire endonuclease gene was sequenced for each mutant (56) in order to determine the substitution responsible for the mutant phenotype.

Construction of the Inv plasmid containing a null endonuclease. From a large number of endonuclease mutants (see Chapter 2), two were selected for use in reconstruction of the Inv and Del plasmids. The objective was to regenerate the Inv and Del plasmids with null mutations in the endonuclease gene. With a null endonuclease lacking *in vivo* activity, cells harboring these plasmids should be insensitive to high levels of endonuclease. Levels of endonuclease in cultures containing these plasmids could then be determined from cultures lacking or containing methylase.

Two endonuclease mutants were used for this purpose: Glu144-Lys and Gly129-Glu. DNA from these mutants was isolated following growth

in the dam$^-$ strain, E. coli GM48. This was digested with MluI and BclI; pMG31-6 DNA was treated in a similar manner. These digests were mixed, treated with T4 DNA ligase, and transformed into E. coli GM48. DNA was isolated from several transformants and analyzed to determine the correct constructions. These were designated pMG31-K144 and pMG31-E129. The endonuclease gene on each of these was sequenced using dideoxy sequencing on double stranded plasmid DNA to confirm the constructions to the nucleotide level. The resulting plasmids are identical to pMG31-6 except that they encode the K144 or E129 endonuclease instead of the S187 endonuclease.

Inv and Del plasmids were then constructed from these plasmids. Plasmids pMG31-K144 and pMG31-E129 were digested with BclI and BamHI, ligated and transformed into E. coli D1210. The Gly129-Glu Inv was isolated and designated Inv-E129. Similarly, plasmids were identified which were designated Inv-K144 and Del-K144. These plasmids were then transformed in parallel into E. coli D1210 and E. coli D1210 containing plasmid pSC101-methylase.

Endonuclease levels expressed in the presence and absence of methylase. Endonuclease levels were determined from cultures which lacked or contained the EcoRI methylase. These cultures harbored the original Inv plasmid, the Inv-E129 plasmid, or the Inv-K144 plasmid. Endonuclease present in these cultures was determined by immunoprecipitation of protein from cell extracts (as described in Methods under measuremnet of endonuclease synthesis and turnover). This material was analyzed on Western blots (Fig. 18). The amounts of endonuclease in cultures containing the Inv plasmid were 10-fold higher

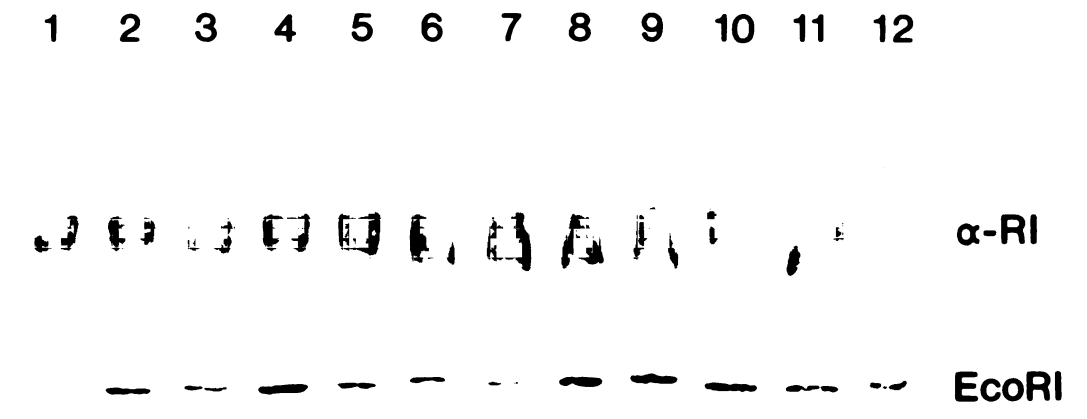


Fig. 18. Endonuclease produced in *E. coli* D1210 cultures lacking or containing the EcoRI methylase. Lane 1: Cultures containing the S187 endonuclease encoded by the Inv plasmid. Lane 2: Cultures encoding the E129 endonuclease from plasmid Inv-E129 Lane 3: Cultures containing the Inv-K144 plasmid Lanes 4-6 are from cultures containing the same endonuclease plasmids as lanes 1-3 and in addition contain the EcoRI methylase. Lanes 7-12 represent similar cultures grown in the presence of IPTG.

when cultures contained methylase. Cultures harboring plasmids encoding the mutant endonucleases, Inv-E129 and Inv-K144, contained the same endonuclease levels in the presence or absence of methylase. This result indicates that the difference in endonuclease levels in methylase containing versus methylase deficient cultures results from the limited activity of the S187 enzyme. With plasmids identical to the Inv, except for mutations in the endonuclease gene, methylase had no effect on endonuclease expression.

DISCUSSION

Cells harboring restriction-modification systems are resistant to foreign (invading) DNA. While the methylase of a type II restriction-modification system will recognize and modify hemimethylated DNA, the corresponding endonuclease only cleaves unmodified DNA. This simple mechamism allows survival of cells carrying these genes. While providing for maintenance of a restriction-modification system, this difference between the endonuclease and methylase would not be expected to aid in the establishment of a restriction-modification system within a new host cell. Introduction of restriction-modification enzymes might result in destruction of cellular DNA and lead to cell death, unless a mechanism existed to provide for methylation of host DNA. If genes encoding restriction and modification enzymes are mobile, regulatory mechanisms might have evolved to assure expression of methylase prior to endonuclease.

The fact that plasmid pPG31 transforms cells at a frequency equal to that obtained with transformation by other plasmids (S.Y. unpublished) demonstrates that there is no selection against plasmids carrying the EcoRI endonuclease and methylase genes. However, cells are unable to survive transformation by plasmids encoding only the endonuclease (31). A regulatory mechanism to guarantee EcoRI methylase expression before the appearance of active endonuclease would ensure that newly acquired EcoRI genes would not be deleterious to the cell.

A model was proposed as a testable hypothesis to explain regulated expression of the EcoRI genes (Fig. 5). Upon introduction of a plasmid carrying the EcoRI genes into a host cell, methylase would be expressed

initially from a weak promoter located near the carboxy-terminus of the endonuclease gene. This methylase would act to modify EcoRI sites within the bacterial DNA and could subsequently activate expression from a promoter upstream of the endonuclease gene. This promoter would produce a bicistronic message encoding endonuclease and methylase. This "positive control" of endonuclease expression would provide a mechanism to assure establishment of these genes within a new host cell. A similar regulatory process has previously been proposed for the HhaII genes (29).

Positive control is used to regulate the expression of a large number of bacterial genes (57). In particular, genes encoding proteins that participate in metabolic pathways and genes encoding proteins that must be temporally expressed are often under positive control. Regulation of this type is often accomplished through control of initiation of transcription. Several mechanisms are employed to mediate this type of regulation. These include: 1) recognition of a new promoter sequence by RNA polymerase, 2) replacing a subunit of RNA polymerase, or 3) production of a new RNA polymerase. For example, the E. coli CAP protein positively regulates expression of a number of operons (lac, ara, gal, mal and others) by binding to DNA adjacent to the promoter and permitting RNA polymerase to initiate transcription from these promoters (58). Similar mechanisms permit positive control of the λ-cI promoter (59). The EcoRI methylase might activate endonuclease expression through enhancement of endonuclease transcription by binding to (or methylating) a site upstream of the endonuclease gene.

Positive control can also be exerted after initiation of transcription. For example, proteins which act as antiterminators mediate positive control. Examples of this include the phage λ N and Q proteins (60). In addition, positive control of the argECBH genes by ppGpp has been reported to occur after transcription initiation (61).

The existence of a mutant EcoRI endonuclease (S187), which contained very low in vivo activity, made possible the study of endonuclease expression in methylase deficient cultures. An inversion (Inv) and a deletion (Del) of the C-terminal half of the methylase gene were constructed from the parent plasmid, pMG31-6. Cultures harboring the Inv and Del plasmids survive, even though they express endonuclease but lack methylase. However, these cultures cannot grow when endonuclease is over-expressed (IPTG induction). Measurement of intracellular endonuclease levels revealed that endonuclease levels were much lower in cultures containing the Inv or Del plasmids than in pMG31-6 containing cultures. With methylase provided via a compatible plasmid, endonuclease levels were elevated to near those expressed from pMG31-6. This result was suggestive of positive regulation of endonuclease activity by methylase. A number of variations on these experiments were performed in order to rule out alternative explanations. These included removal of the lac control region from the Inv plasmid, creating 322Inv. Cells harboring 322Inv were found to contain low levels of endonuclease. These endonuclease levels were found to be 10-20 fold higher when methylase was present. These experiments were also performed in a number of different E. coli strains, and similar results were obtained.

β-lactamase levels served as an internal control in these experiments. β-lactamase levels should reflect plasmid copy number, since β-lactamase is encoded by the same plasmids which carry the EcoRI genes. Endonuclease levels were 10-20 fold higher when the methylase was present, while β-lactamase levels were not significantly altered. However, β-lactamase levels were slightly lower (generally by about 10%), in extracts lacking methylase.

A major problem encountered with these experiments was that endonuclease levels fluctuated significantly from one experiment to the next. The largest fluctuations in endonuclease levels were in methylase deficient cultures. These values ranged from 0.3 to 14, with the high values being found in slow growing cultures. In particular, growth at 30°C or in poor growth media resulted in higher endonuclease levels in cultures lacking methylase. The reason for this is not known. Despite this problem, endonuclease levels were consistently higher when methylase was present.

Problems inherent in the endonuclease activity determinations made obtaining an alternative method of measuring endonuclease expression desirable. The S187 endonuclease, which has low *in vivo* activity, still does retain some activity *in vivo*. This residual activity is a primary concern, since it may be harmful to the host cell in the absence of methylase. In addition, other problems exist with these experiments. Preparation of cell extracts by sonication is not 100% reproducible, and can result in denaturation of some of the protein. This could have an effect on the activity level, since the activity level is given in units of endonuclease activity per milligram total cell protein. Endonuclease activity is also difficult to quantitate

when assays are performed on crude extracts. This is a problem especially when there are low levels of endonuclease activity as there are in the methylase deficient cultures.

These and other reasons led to a decision to study regulation of endonuclease expression from fusion plasmids. Fusion plasmids are very powerful tools in the investigation of regulation of gene expression (62, 63, 64). β-galactosidase gene fusions permit determination of expression levels through measurement of β-galactosidase levels, assays which are quantitative and relatively easy to perform. β-galactosidase fusions have been used in the investigation of a large number of regulatory processes. These include: a study of tnpR, a repressor of the Tn3 transposase (65), cloning of the ompR gene, which specifies a positive regulatory protein necessary for expression of ompC and ompF (66), and studies on the positive regulation of the maltose operon (67).

Through construction of the appropriate fusions, it is possible to distinguish regulation at the transcriptional level from regulation at the translational level. To accomplish this, both operon and gene fusions are used. Operon fusions leave the β-galactosidase gene and translation initiation signals intact. Expression of β-galactosidase results from transcription initiated at sequences fused upstream of the β-galactosidase gene. Gene fusions place β-galactosidase under both transcriptional and translational control of upstream sequences. This is accomplished by fusing the gene of interest in phase to β-galactosidase, resulting in expression of a hybrid protein which can be quantitated by performing β-galactosidase assays. Regulation at the transcriptional level will affect both operon and gene fusions while

translational control will only affect only the gene fusion. This approach has been used to demonstrate that a number of genes are regulated at the translational level (68-70).

Fusion plasmids were generated, placing β-galactosidase expression under EcoRI endonuclease control. Both operon fusion (pSY60) and gene fusion (pSY64) plasmids were constructed. Elevation of β-galactosidase levels expressed from pSY60 in the presence of the EcoRI methylase would indicate activation at the transcriptional level. Elevated expression from pSY64 could result from either transcriptional or translational control.

β-galactosidase levels in cultures containing these fusion plasmids were determined in both the absence and presence of the EcoRI methylase, and found to be similar. A slight increase in β-galactosidase activity was observed when methylase was present, in some cases a two-fold increase. Unfortunately, significant β-galactosidase levels were expressed from the vector alone (pMLB1010), complicating interpretation of these experiments. An attempt to eliminate this background expression was made by inserting a transcriptional terminator sequence upstream of the fusion site in both the vector and fusion constructions. However, even with the terminator sequence upstream, there was significant expression of β-galactosidase from the parent vector.

Additional fusion plasmids were constructed to permit experimental examination of alternative explanations for the observed results. For example, methylase might activate translation of the endonuclease gene through action at a site downstream from the fusion site in pSY64 (which contains about 1/3 of the endonuclease gene fused to

β-galactosidase). If this were the case, the results obtained with the Inv plasmid and the fusion plasmids would not be contradictory.

Bacteria employ a number of different control mechanisms to regulate gene expression. The data on endonuclease expression is consistent with possible regulation by several of these mechanisms. For example, completion of the bacteriophage λ-int gene translation is regulated by a site distal to the gene. Regulation of this type has been termed retroregulation (71, 72) and is exerted at the RNA level. Complementary RNA's (termed micRNA) have been shown to be involved in regulation of several genes (73), including the Tn10 transposase (70) and the E. coli outer membrane protein, ompF (69). In order to test such mechanisms for the EcoRI system, additional fusion plasmids were constructed.

A gene fusion in which a large portion (75%) of the endonuclease gene was fused in phase to β-galactosidase was constructed (pSY81). With methylase provided in trans, cultures harboring pSY81 produced higher β-galactosidase levels (2-3 fold) than those lacking methylase. However, this increase was nowhere near the 10-20 fold increase observed with the Inv plasmid.

To circumvent the problem of measuring gene expression levels through determination of enzyme activity, mRNA levels were measured from cells which lacked (or contained) the EcoRI methylase. These experiments permit direct determination of rates of transcription. mRNA levels were measured by hybridizing RNA (from extracts of cells containing or lacking methylase) to filter-bound DNA (data not shown). Low mRNA levels were found in the absence of methylase. These were 2-fold higher in the presence of methylase.

Attempts were also made to look for methylase-induced activation of endonuclease expression on northern blots using RNA prepared from extracts of cells containing the Inv plasmid with and without methylase. Endonuclease transcription might terminate at some point within the endonuclease gene when expressed in the absence of methylase. This "attenuation" might be relieved through the action of the methylase. However, meaningful results were not obtained from these northern blots.

Evidence for the existence of separate promoters for the endonuclease and methylase was obtained from in vitro transcription using purified DNA restriction fragments. An endonuclease promoter was mapped to approximately 40 base pairs upstream from the start of the endonuclease gene. Transcription was also found to originate near the end of the endonuclease gene (S.Y., Frank Stephenson, unpublished). Methylase, or methylated DNA, was provided in experiments measuring transcription from the endonuclease promoter, but no stimulation was observed.

Direct measurement of expression from the methylase promoter was obtained after inserting the RpoBC transcriptional terminator in the middle of the endonuclease gene (pSY31 and pSY32). Methylase expression from these plasmids must arise from promoter sequences located within the C-terminus of the endonuclease gene or in the 30 base pair intergenic region. Methylase levels expressed from pSY31 or pSY32 are approximately 10% of those expressed from pMG31-6 under conditions where the lac promoter is fully repressed. This data suggests that the expression from the methylase promoter is no more than 10% of the level of expression from the endonuclease promoter.

The experiments measuring mRNA levels and the experiments with fusion plasmids did not demonstrate any significant effect of methylase on endonuclease expression. These results appeared to contradict results of the activity measurements performed on cultures harboring the Inv or Del plasmids. In an attempt to further our understanding of the control of endonuclease expression, and to hopefully clarify this apparent contradiction, the rates of endonuclease synthesis and turnover were measured.

These experiments were performed in a manner similar to that used to demonstrate control of phage λ-cII protein levels by control of cII stability (74). Regulation of protein levels by protein turnover has been demonstrated for a number of bacterial genes (75, 76). The apparent enhancement of endonuclease activity by methylase might in fact result from enhancement of endonuclease stability.

To address this possibiltiy, rates of endonuclease synthesis were determined in pulse-chase experiments using cultures grown in the absesnce or presence of methylase. These experiments involved determination of rates of endonuclease synthesis, turnover (stability), and measurement of activity levels. To allow maximal incorporation of ^{35}S- methionine, cultures were grown in minimal media lacking cysteine and methionine. Under these conditions, cultures grew very slowly, especially those harboring the Inv plasmid in the absence of methylase. Endonuclease levels from cultures lacking methylase were found to be elevated under poor growth conditions. Activity levels were found to correlate with synthesis rates, a result which makes postranslational control unlikely. Chase times of up to 60 minutes showed no

deterioration in endonuclease protein levels, demonstrating that once synthesized, endonuclease is stable within the cell.

Levels of endonuclease were found to vary significantly from one experiment to the next in both the activity and synthesis/turnover determinations. Endonculease levels in cultures lacking methylase fluctuated to a greater extent than in cultures containing methylase. This result could be due to deleterious effects of the S187 endonuclease upon a host cell lacking methylase. Null endonuclease mutants were generated in order to eliminate this possibility. The goal was to obtain mutations within the endonuclease gene which result in a phenotypically null endonuclease (see Chapter 2). Using in vitro hydroxylamine mutagenesis, a large number of mutations were generated. A strong selection for an inactive endonuclease aided in the isolation of a large number of mutants. The locations and exact nature of the base substitutions which led to the null phenotype were determined by DNA sequencing. Two of these null mutants were chosen for use in reconstructing the original Inv and Del plasmids. By using plasmids encoding a null endonuclease, variation in expression resulting from the partial activity of the S187 endonuclease would be eliminated. Endonuclease levels expressed from the Inv plasmids encoding null endonucleases were found to be the same in both the presence and absence of methylase.

This result supports the conclusion that the apparent methylase induced activation of endonuclease expression from the Inv plasmid may have resulted from deleterious effects upon the host cell, due to the presence of a partially active endonuclease (S187) in cells lacking methylase. The following is an explanation of these results,

consistent with all of the presently available data. Methylase does not positively regulate endonuclease expression. The result obtained using the Inv plasmid comes from selective cell death when the S187 endonuclease present in cells lacking methylase. There may be considerable variation in endonuclease expresseion levels from cell to cell. While the great majority of cells in a given culture may contain very low endonuclease levels, a small number may contain extremely high endonuclease levels. In the absence of the methylase, cells containing high endonuclease levels will die. However, these cells can survive when methylase is present. This differential viability could result in cultures containing methylase appearing to have higher endonuclease levels than methylase deficient cultures.

While this model is consistent with the present data, it is not completely satisfying. It does not explain why there would be such a large variation in endonuclease expression within a culture, or what would cause a given cell to express either very high, or low, endonuclease levels. In addition, this model does not explain why cells can be transformed with plasmids carrying the endonuclease and methylase genes. In fact, this model does not provide a mechanism for regulation of the EcoRI genes.

The genetic arrangement of the EcoRI system would appear to necessistate some kind of regulated, temporal expression in order to assure viability of cells harboring these genes. Positive regulation of endonuclease expression by the EcoRI methylase remains an attractive model. Hoewever, the results presented here neither confirm, nor rule out this possibility. The exact nature of control of the EcoRI genes remains to be elucidated.

REFERENCES

1. Kessler, C., Newmaier, P.S. and Wolf, W. Recognition sequences of restriction endonucleases and methylases - a review. Gene 33:1-102, 1985

2. Modrich, P. and Roberts, R.J. Type-II restriction and modification enzymes. In: "Nucleases" Linn, S.M., Roberts, R.J. eds. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 1982:109-154

3. Roberts, R.J. In: Gene amplification and analysis, Volume 5: Restriction endonucleases and methylases. Chirikjian, J.G. ed., pp.1-50, Elsevier Science Publishing Co. Inc., New York, NY, 1987

4. Chang, S. and Cohen, S.N. In vivo site-specific genetic recombination promoted by the EcoRI restriction endonuclease. Proc. Natl. Acad. Sci. USA 74:4811-4815, 1977

5. Kleid, D., Humayun, Z., Jeffrey, A. and Ptashne, M. Novel properites of a restriction endonuclease isolated from Haemophilus parahaemolyticus. Proc. Natl. Acad. Sci. USA 73:293-297, 1976

6. Kroger, M., Hobom, G. Schutte, H. and Mayer, H. Eight new restriction endonucleases from Herpetosiphon giganteus-divergent evolution in a family of enzymes. Nucleic Acids Res. 12:3127-3141, 1984

7. Qiang, B-Q, Schildkraut, I. A type II restriction endonuclease with an eight nucleotide specificity from Streptomyces fimbriatus. Nucleic Acids Res. 12:4507-4515, 1984

8. Lee, Y.H. and Chirikjian, J.G. Sequence-specific endonuclease BglI: Modification of lysine and arginine residues of the homogeneous enzyme. J. Biol. Chem. 254:6838-6841, 1979

9. Koncz, C., Kiss, A. and Venetianer, P. Biochemical characterization of the restriction-modification system of Bacillus sphaericus. Eur. J. Biochem. 89:523-529, 1978

10. McClarin, J.A., Frederick, C.A., Wang, B-C., Greene, P., Boyer, H.W., Grable, J. and Rosenberg, J.M. Structure of the DNA-EcoRI endonuclease recognition complex at 3 Å resolution. Science 234:1526-1541, 1986

11. Walder, R.Y., Hartley, J.L., Donelson, J.E. and Walder, J.A. Cloning and expression of the PstI restriction-modification system in Escherichia coli. Proc. Natl. Acad. Sci. USA 78:1503-1507, 1981

12. Howard, K.A., Card, C., Benner, J.S., Callahan, H.L., Maunus, R., Silber, K., Wilson, G. and Brooks, J.E. Cloning the DdeI restriction-modification system using a two-step method. Nucleic Acids Res. 14:7939-7951, 1986

13. Gingeras, T.R. and Brooks, J.E. Cloned restriction/modification system from Pseudomonas aeruginosa. Proc. Natl. Acad. Sci. USA 80:402-406, 1983

14. Kosykh, V.G., Buryanov, Y.I. and Bayev, A.A. Molecular cloning of the EcoRII endonuclease and methylase genes. Molec. Gen. Genet. 178:717-718, 1980

15. Blumenthal, R.M., Gregory, S.A. and Cooperider, J.S. Cloning of a restriction-modification system from Proteus vulgaris and its use in analyzing a methylase-sensitive phenotype in Escherichia coli. J. Bact. 164:501-509, 1985

16. Mann, M.B., Rao, R.N. and Smith, H.O. Cloning of restriction and modification genes in E. coli: The HhaII system from Haemophilus haemolyticus. Gene 3:97-112, 1978

17. Kiss, A., Posfai, G., Keller, C.C., Venetianer, P. and Roberts, R.J. Nucleotide sequence of the BsuRI restriction-modification system. Nucleic Acids Res. 13:6403-6421, 1985

18. Bougueleret, L., Schwarzstein, M., Tsugita, A. and Zabeau, M. Characterization of the genes coding for the EcoRV restriction and modification system of Escherichia coli. Nucleic Acids Res. 12:3659-3676, 1984

19. Walder, R.Y., Walder, J.A., and Donelson, J.E. The organization and complete nucleotide sequence of the PstI restriction-modification system. J. Biol. Chem. 259:8015-8026, 1984

20. Theriault, G., Roy, P.H., Howard, K.A., Benner, J.S., Brooks, J.E., Waters, A.F. and Gingeras, T.R. Nucleotide sequence of the PaeR7 restriction/modification system and partial characterization of its protein products. Nucleic Acids Res. 13:8441-8461, 1985

21. Schoner, B., Kelly, S. and Smith H.O. The nucleotide sequence of the HhaII restriction and modification genes from Haemophilus haemolyticus. Gene 24:227-239 1983

22. Greene, P.J., Gupta, M., Boyer, H.W., Brown, W.E. and Rosenberg, J.M. Sequence analysis of the DNA encoding EcoRI endonuclease and methylase. J. Biol. Chem. 256:2143-2153, 1981

23. Newman, A.K., Rubin, R.A., Kim, S.H. and Modrich, P. DNA sequences of structural genes for EcoRI DNA restriction and modification genes. J. Biol. Chem. 256:2131-2139, 1981

24. Jacoby, G.A. and Sutton, L. Restriction and Modification Determined by a Pseudomonas R Plasmid. Plasmid 1:115-116, 1977

25. Betlach, M., Herschfield, V., Chow, L., Brown, W., Goodman, H.W., and Boyer, H.W. A restriction endonuclease analysis of the bacterial plasmid controlling the EcoRI restriction and modification of DNA. Federation Proc. 35:2037-2043, 1976

26. Trautner, T., Pawlek, B., Bron, S. and Anagnostopoulos, C. Restriction and Modification in B. subtilis: Biological Aspects. Molec. Gen. Genet. 131:181-191, 1974

27. Falkow, S. In: Infectious Multiple Drug Resistance (Lagnado, J.R. ed) pp.98-143, Pion Limited, London, 1975

28. Shapiro, H.S. CRC Handbook of Biochemistry, Selected Data for Molecular Biology (Sober, H.A. ed.) 2nd Ed. pp.H80-H88, The Chemical Rubber Co., Cleveland, 1970

29. Smith, H.O. Nucleotide sequence specificity of restriction endonucleases. Science 205:455-462, 1979

30. Arber, W. "DNA modification and restriction" In: Progress in nucleic acid research and molecular biology. 14:1-37, Academic Press, New York, 1974

31. Kuhn, I., Stephenson, F.H., Boyer, H.W. and Greene, P.J. Positive-selection vectors utilizing the lethality of the EcoRI endonuclease. Gene 44:253-263, 1986

32. Lingappa, V.R., Lingappa, J.R. and Blobel, G. Chicken ovalbumin contains an internal signal sequence. Nature 281:117-121, 1979

33. Modrich, P. and Zabel, D. EcoRI endonuclease: Physical and catalytic properties of the homogeneous enzyme. J. Biol. Chem. 251:5866-5874, 1976

34. Hedgpeth, J., Goodman, H.M. and Boyer, H.W. DNA nucleotide sequence restricted by the R1 endonuclease. Proc. Natl. Acac. Sci. USA 69:3448-3452, 1972

35. Dugaiczyk, A., Hedgpeth, J., Boyer, H.W. and Goodman, H.M. Physical identity of the SV40 deoxyribonucleic acid sequence recognized by the EcoRI restriction endonuclease and modification methylase. Biochem. 13:503-512, 1974

36. Backman, K., Ptashne, M. and Gilbert, W. Construction of plasmids carrying the cI gene of bacteriophage λ. Proc. Natl. Acad. Sci. USA 73:4174-4178, 1976

37. Messing, J.R., Crea, R. and Seeburg, P.H. A system for shotgun DNA sequencing. Nucleic Acids Res. 9:309-321, 1981

38. Messing, J. "New M13 Vectors for Cloning" In: Methods in Enzymol. 101:20-78, 1983

39. Boyer, H.W. and Roulland-Dussoix, D. A Complementation Analysis of the Restriction and Modification of DNA in Escherichia coli. J. Mol. Biol. 41:459-472, 1969

40. Bolivar, F. and Backman, K. Plasmids of Escherichia coli as Cloning Vectors. Methods in Enzymol. 68:245-267, 1979

41. Casadaban, M.J. Transposition and Fusion of the lac Genes to Selected Promoters in Escherichia coli Using Bacteriophage Lambda and Mu. J. Mol. Biol. 104:541-555, 1976

42. Marinus, M.G. Location of DNA Methylation Genes on the Escherichia coli K-12 Genetic Map. Molec. Gen. Genet. 127:47-55, 1973

43. Greene, P.J., Heynecker, H.L., Bolivar, F., Rodriguez, R.L., Betlach, M.C., Covarrubias, S.A., Backman, K., Russel, D., Tait, R., Boyer, H.W. A general method for the purification of restriction enzymes. Nucleic Acids Res. 5:2373-2380, 1978

44. Itakura, K., Hirose, T., Crea, R., Riggs, A.D., Heynecker, H.L., Bolivar, F. and Boyer, H.W. Expression in Escherichia coli of a Chemically Synthesized Gene for the Hormone Somatostatin. Science 198:1056-1063, 1977

45. Bolivar, F., Rodrigues, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. Construction and Characterization of New Cloning Vehicles: II. A Multipurpose Cloning System. Gene 2:95-113, 1977

46. Sutcliffe, J.G. Complete Nucleotide Sequence of the Escherichia coli Plasmid pBR322. Cold Spring Harbor Symp. on Quant. Biol. 43:77-90, 1979

47. Das, A., Crawford, I.P., and Yanofsky, C. Regulation of Tryptophan Operon Expression by Attenuation in Cell-free Extracts of Escherichia coli. J. Biol. Chem. 257:8795-8798, 1982

48. Travers, A.A., Lamond, A.I., Mace, H.A.F. and Berman, M.L. RNA Polymerase Interactions with the Upstream Region of the E. coli tyrT Promoter. Cell 35:265-273, 1983

49. Weinstock, G.M., Berman, M.L. and Silhavy, T.J. Chimeric genetics with β-galactosidase. pp.27-64 In: Gene amplification and analysis, Papas, T.S., Rosenberg, M. and Chirikjian, J.G. eds. Elsevier/North-Holland Publishing Co., Inc., New York, 1983

50. Chang, A.C.Y. and Cohen, S.N. Construction and Characterization of Amplifiable Multicopy DNA Cloning Vehicles Derived from the P15A Cryptic Miniplasmid. J. Bact. 134:1141-1156, 1978

51. Squires, C., Krainer, A., Barry, G., Shen, W-F. and Squires, C.L. Nucleotide sequence at the end of the RNA polymerase β' subunit (rpoC). Nucleic Acids Res. 9:6827-6840, 1981

52. Cabello, F., Timmis, K. and Cohen, S.N. Replication control in a composite plasmid constructed by in vitro linkage of two distinct replicons. Nature 259:285-290, 1976

53. Clewell, D.B., and Helinski D.R. Properties of a Supercoiled Deoxyribonucleic Acid-Protein Relaxation Complex and Strand Specificity of the Relaxation Event. Biochem. 9:4428-4440, 1970

54. O'Connor, C.D. and Humphreys, G.O. Expression of the EcoRI restriction-modification system and the construction of positive-selection cloning vectors. Gene 20:219-229, 1982

55. Davis, R.W., Botstein, D. and Roth, J.R. In: "Advanced Bacterial Genetics", Cold Spring Harbor, New York, Cold Spring Harbor Laboratory, 1980

56. Yanofsky, S.D., Love, R., McClarin, J.A., Rosenberg, J.M., Boyer, H.W. and Greene, P.J. Clustering of Null Mutations in the EcoRI Endonculease. Proteins: Structure, Function, and Genetics, Alan R. Liss, Inc., New York, Vol. 2, No. 4, in press

57. Raibaud, O. and Schwartz, M. Positive control of transcription initiation in bacteria. Ann. Rev. Genet. 18:173-206, 1984

58. de Crombrugghe, B., Busby, S. and Buc, H. Activation of transcription by the cyclic AMP receptor protein. Science 224:831-837, 1984

59. Hochschidd, A., Irwin, N. and Ptashne, M. Repressor structure and the mechanism of positive control. Cell 32:319-325, 1983

60. Friedman, D.I. and Gottesman, M. Lytic mode of lambda development. In:Lambda II, ed. Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.A. pp.21-51 Cold Spring Harbor, Cold Spring Harbor Lab., 1983

61. Zidwick, M.J., Kroshus, J. and Rogers, P. Positive control of expression of the argECBH gene cluster in vitro by guanosine 5'-diphosphate 3'-diphosphate. J. Bact. 159:647-651, 1984

62. Silhavy, T.J. and Beckwith, J.R. Uses of lac Fusions in the Study of Biological Problems. Microbiological Reviews 49:398-418, 1985

63. Casadaban, M.J. and Cohen, S.N. Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. J. Mol. Biol. 138:179-207, 1980

64. Casadaban, M.J., Chou, J. and Cohen, S.N. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: Escherichia coli plasmid vectors for the detection and cloning of translational initiation signals. J. Bact. 143:971-980, 1980

65. Chou, J., Casadaban, M.J., Lemaux, P.G. and Cohen, S.N. Identification and characterization of a self-regulated repressor of translocation of the Tn3 element. Proc. Natl. Acad. Sci. USA 76:4020-4024, 1979

66. Taylor, R.K., Hall, M.N., Enquist, L. and Silhavy, T.J. Identification of OmpR: a positive regulatory protein controlling expression of the major outer membrane matrix porin proteins of Escherichia coli K-12. J. Bacteriol. 147:255-258, 1981

67. Debarbouille, M., Shuman, H.A., Silhavy, T.J. and Schwartz, M. Dominant constitutive mutations in malT, the positive regulator gene of the maltose regulon in Escherichia coli. J. Mol. Biol. 124:359-371, 1978

68. Milura, A., Krueger, H., Itoh, S., deBoer, H.A. and Nomura, M. Growth rate dependent regulation of ribosome systhesis in E. coli: expression of the lacZ and galK genes fused to ribosomal promoters. Cell 25:773-782, 1981

69. Mizuno, T., Chou, M. and Inouye, M. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). Proc. Natl. Acad. Sci. USA 81:1966-1970, 1984

70. Simons, R.W. and Kleckner, N. Translational control of IS10 transposition. Cell 34:683-691, 1983

71. Guarneros, G., Montanez, C., Hernandez, T. and Court, D. Post-transcriptional control of bacteriophage λ-int gene expression from a site distal to the gene. Proc. Natl. Acad. Sci. USA 79:238-242, 1982

72. Schindler, D. and Echols, H. Retroregulation of the int gene of bacteriophage λ: Control of translation completion. Proc. Natl. Acad. Sci. USA 78:4475-4479, 1981

73. Coleman, J., Green, P.J. and Inouye, M. The use of RNAs complementary to specific mRNAs to regulate the expression of individual bacterial genes. Cell 37:429-436, 1984

74. Hoyt, A.M., Knight, D.M., Das, A., Miller, H.I. and Echols, H. Control of phage λ development by stability and synthesis of cII protein: Role of the viral cIII and host hflA, himA and himD genes. Cell 31:565-573, 1982

75. Simon, L.D., Randolph, B., Irwin, N. and Binkowski, G. Stabilization of proteins by a bacteriophage T4 gene cloned in

Escherichia coli. Proc. Natl. Acad. Sci. USA 80:2059-2062, 1983

76. Little, J.W. and Mount, D.W. The SOS regulatory system of Escherichia coli. Cell 29:11-22, 1982

77. Welker, N.E. and Campbell, L.L. Unrelatedness of Bacillus amyloliquefaciens and Bacillus subtilus. J. Bact. 94:1124-1130, 1967

78. Falkow, S., Ryman, I.R. and Washington, O. Deoxyribonucleic acid base composition of Proteus and Providence organisms. J. Bact. 83:1318-1321, 1962

79. Jen-Jacobson, L., Kurpiewski, M., Lesser, D., Grable, J., Boyer, H.W., Rosenberg, J.M. and Greene, P.J. Coordinate ion pair formation between EcoRI endonuclease and DNA. J. Biol. Chem. 258:14638-14646, 1983

80. Birnboim, H.C. and Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513-1523, 1979

81. Davis, R.W., Botstein, D. and Roth, J.R. In: "Advanced bacterial genetics", Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 1980

82. Miller, J.H. In: "Experiments in molecular genetics", Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 1972

83. O'Callaghan, C.H., Morris, A., Kirby, S.M. and Shingler, A.H. Novel method for detection of β-lactamases by using a chromogenic cephalosporin substrate. Antimicrobial agents and chemotherapy 1:283-288, 1972

Chapter 2: Clustering of Null Mutations in the EcoRI Endonuclease[5]

[5] The contents of this chapter have been accepted for publication in "Proteins: Structure, Function, and Genetics", Vol. 2, No. 4, Alan R. Liss, Inc., New York. The co-authors on the paper are: Robert Love, Judith A. McClarin, John M. Rosenberg, Herbert W. Boyer and Patricia J. Greene. Robert Love assisted in constructing and photographing the computer graphics pictures. Judith McClarin was responsible for the 3 Å model of the EcoRI-DNA complex.

ABSTRACT

EcoRI endonuclease mutants were isolated in a methylase deficient background following in vitro hydroxylamine mutagenesis of plasmid pKG2 (20). Mutants which survived high level endonuclease expression (IPTG induction) were termed null mutants. Sixty-two of 121 null mutants tested by western blot contained normal levels of endonuclease cross-reacting protein. The complete endonuclease gene was sequenced for 27 null mutants. This group was found to consist of 20 single base change missense mutations, 6 double mutations and 1 triple mutation. Ten of the 20 single mutations were clustered between residues 139 and 144. When examined with respect to the structure of the EcoRI-DNA complex (7), these alterations were found to fall predominantly into two classes: substitutions at the protein-DNA interface or substitutions at the protein-protein (dimer) interface. Protein from several of the mutants was purified and sized using HPLC. Wild-type EcoRI endonuclease and protein from three of the DNA interface mutations (Ala139-Thr, Gly140-Ser, Arg203-Gln) appeared to be dimeric, while protein from subunit interface mutations (Glu144-Lys, Glu152-Lys, Gly210-Arg) migrated as monomers.

INTRODUCTION

The ability of proteins to recognize and interact with specific nucleotide sequences plays an essential role in biological processes. The recognition process must provide a mechanism whereby a protein can discriminate a given nucleotide sequence from an excess of related (or unrelated) sequences. The interaction between the EcoRI endonuclease and its cognate DNA provides a model system for investigation of sequence specific protein-DNA interactions. The EcoRI endonuclease recognizes and cuts double-stranded DNA at the sequence GAATTC (1) with a high degree of specificity (2) and has been extensively characterized (3,4,5,6). The structure of the EcoRI endonuclease-DNA co-crystal has been determined to 3 Å resolution (7). From the crystal structure, the positions of amino acid residues within the endonuclease have been assigned relative to the DNA substrate. The structure reveals that the EcoRI endonuclease interacts with DNA in the major groove and the adjacent phosphodiester backbone leaving the minor groove exposed to solvent. In addition, the co-crystal reveals that the structure of the DNA is altered (kinked) (8) when complexed with endonuclease.

Studies on a number of prokaryotic DNA-binding regulatory proteins have yielded insight into the molecular basis for interactions between these proteins and their DNA recognition sequences. The structures of the λ-cI repressor (9), E.coli cap protein (10), trp repressor (11), and λ-cro repressor (12) have recently been determined in the absence of DNA. These proteins share a common secondary structure, a helix-turn-helix motif, which has been proposed to play a role in recognition of sequences within the major groove of DNA (13). Support for this model has recently been obtained from the structure of the

DNA-binding domain of bacteriophage 434 repressor complexed with DNA (14). This protein contains a helix-turn helix motif, with amino acid side chains from helix-3 forming hydrogen bonds with bases in the major groove of the DNA.

Although the helix-turn-helix motif is highly conserved among the regulatory proteins studied, EcoRI interacts with DNA in a different manner. Different mechanistic interactions may be required for restriction enzymes which, in addition to recognizing a short DNA sequence, function catalytically to cleave double stranded DNA. Since growing cells may not tolerate even an extremely low rate of incorrect double-strand cleavage, high specificity is essential.

Mutational studies on the trp repressor (15), λ-cro repressor (16) and other DNA-binding proteins (17-19) have helped to identify amino acid residues which participate in DNA binding. We report here the generation, isolation and initial characterization of a number of missense null mutants of the EcoRI endonuclease. The locations of several alterations have been determined by DNA sequencing and the possible effects of the amino acid substitutions in the null endonucleases have been examined in relation to the current 3 Å model of the EcoRI-DNA complex (7). Missense mutations could produce a null phenotype if the amino acid substitution produces a dramatic effect on either DNA binding, catalytic activity, protein stability, or subunit-subunit (dimerization) interactions. The majority of the null mutants described in this report result from amino acid substitutions located at the protein-DNA interface or the protein-protein interface.

MATERIALS AND METHODS

Strains: E.coli D1210 (20), hsdM, hsdR, recA13, ara-14, proA2, lacY1, lac I^Q, galK2, supE44

E.coli MM294 (21), hsdR, endI, pro, thi

Plasmids: pKG2 (20) contains the wild-type EcoRI endonuclease gene under control of the lac UV5 promoter and encodes resistance to kanamycin. Cells carrying pKG2 can survive if the EcoRI methylase gene is provided in trans on pSC101-methylase (3). In addition to encoding the EcoRI methylase, pSC101-methylase encodes resistance to tetracycline. pSC101-methylase is a low copy plasmid and is compatible with pKG2.

DNA isolation: pKG2 was isolated and purified by the method of Kuhn et al. (20). Plasmid DNA, isolated from cells harboring pKG2 and pSC101-methylase, was digested with BamHI to selectively cut the pSC101-methylase. pKG2 was then obtained free of pSC101-methylase by rebanding in CsCl. DNA was isolated for sequencing according to the method of Birnboim and Doly (22) and purified as supercoiled DNA by banding in CsCl.

Media: Cells were grown in YT broth (23). When used kanamycin was provided as 20 µg/ml in solution and 50 µg/ml for solid media. Tetracycline was added to 2 µg/ml for liquid media and 20 µg/ml for solid media. IPTG was used at 1 mM.

Selection: Endonuclease mutants were generated using the positive selection cloning vector, pKG2. In E.coli D1210 low levels of endonuclease are expressed from the endogeneous endonuclease promoter (S.Y., P.G. unpublished) on pKG2 in the absence of IPTG while endonuclease levels are elevated upon the addition of IPTG. Since

endonuclease expression cannot be turned off completely, pKG2 can transform cells which lack the EcoRI methylase only if the endonuclease gene, or its expression, has been inactivated. Plating under conditions of high endonuclease expression (IPTG induction) provides a strong selection for endonuclease null mutants. Mutagenized DNA was transformed into E.coli D1210 (or 294); kanamycin resistant transformants were selected and subsequently tested for survival in the presence of IPTG.

Mutagenesis: 1) Hydroxylamine mutagenesis was carried out by the method of Davis and Botstein (24). pKG2 (5 µg) was mutagenized in vitro in 2 ml containing 0.4 M hydroxylamine. At 24 hours, 36 hours and 48 hours, 0.2 ml aliquots were removed and dialyzed overnight against 10 mM Tris pH 7.5, 5 mM NaCl, 1 mM EDTA. This DNA was transformed subsequently into E.coli D1210 (or 294). When transformants were obtained from more than one time point, only those from the lesser time point were used. 2) 2-aminopurine mutagenesis was performed according to the method of Miller (23).

Western blots: One ml cultures from each mutant were grown in YT broth at 37°C to OD_{600} of 0.3. IPTG was added to 1 mM, the cultures grown 2 hours, and the cells harvested by centrifugation. The cell pellet was resuspended in SDS loading buffer, boiled 4 minutes and loaded onto a 12% SDS-polyacrylamide gel (25) and run at 35 mA until the bromphenol blue reached the bottom of the gel. Protein was electroblotted onto nitrocellulose (26) and probed with antibody raised against EcoRI endonuclease. After incubation with biotinylated goat anti-rabbit antibody and biotinylated alkaline phosphatase (Vector Labs), the bands

were visualized using the BCIP/NBT phosphatase substrate system (Kirkegaard and Perry Labs Inc.) (27).

DNA sequencing: The entire sequence of each mutant endonuclease gene was determined by Sanger dideoxy sequencing directly from supercoiled plasmid DNA according to the method of Chen and Seeburg (28). Four different 17 base oligonucleotides were used to routinely sequence the 830 bases encompassing the entire endonuclease gene. These were designed to anneal to the endonuclease sequence approximately every 220 bases. Briefly, 2 µg of supercoiled pKG2 DNA was denatured in alkali, neutralized, precipitated and mixed with 5 pmoles of primer. Sequencing reactions were carried out using all four ^{32}P-deoxynucleotide triphosphates (400 Ci/mmole, Amersham) as described by Greene et al. (4) with the following modifications: 1) Reactions were performed in micro titre plates to facilitate sequencing multiple mutants simultaneously. Aliquots were placed on the side of wells and reactions begun by centrifuging material to the bottom. 2) Samples were baked 15 minutes at 65°C in a dry oven prior to loading on the sequencing gel. 3) 6% acrylamide, 8 M urea, 60 cm wedge gels were used with the top thickness 0.18 mm and the bottom being 0.54 mm. 4) Gels were loaded using a filed down 33 gauge Hamilton syringe. 5) Gels were run under constant current at 23 mA cooled with an electric fan until the bromphenol blue reached the bottom. Initially, DNA sequencing was done from CsCl purified DNA, but later mini screen DNA (22) was used. Klenow was purchased from BRL.

Protein purification: Mutant endonuclease was purified from 4-8 g frozen E.coli cells which had been induced with IPTG for 3 hours. Cell extracts were prepared in EB (10 mM KPO_4, 1 mM EDTA, 6 mM

β-mercaptoethanol pH 6.8) + 0.2 M NaCl as described in Greene et al. (29). Protein was chromotographed on phosphocellulose and hydroxylapatite according to the method of Greene et al. (29). Mutant protein was concentrated to 0.5-1.0 mg/ml on a 1 ml mini PC or mini HA column. The elution profile of the mutant endonuclease was monitored by immunoprecipitation during the purification process. Several of the mutant proteins behaved quite differently from wild-type during purification, i.e. at practically every step in the purification of the Lys144 protein there appeared to be insoluble material which correlated with the loss of up to 90% of the protein on each column. When aggregation occurred, we were unable to solubilize the precipitated protein.

Immunoprecipitation: Samples were incubated at 4°C for 1 hour or more in EB + 0.2 M NaCl with an equal volume of TNT (2% Triton, 1 M NaCl, 50 mM Tris pH 7.5) plus 0.005 ml of anti-R1 serum which had been preabsorbed with D1210 cell paste. Following the addition of 0.02 ml immunoprecipitin (BRL), the mix was vortexed and centrifuged 2 minutes, the pellet washed one time with TNT, one time with TN (50 mM Tris, 100 mM NaCl pH 7.5) and resuspended in loading buffer. This was boiled 4 minutes, centrifuged 2 minutes and the supernatant was electrophoresed on a SDS polyacrylamide gel.

Protein sizing: One hundred to 200 µg of purified protein was injected onto a BIO SIL TSK-250 HPLC column equilibrated with EB + 0.8 M NaCl. Protein peaks were determined by monitoring the absorbance using a 274 nm filter. Correspondence of absorbance peaks and endonuclease was determined by immunopricipitation assays on fractions as described above. The HPLC column was calibrated using BioRad protein standard

and wild-type EcoRI. Analysis of wild type EcoRI endonuclease in EB + 0.2 M NaCl indicated that the elution profile varied as a function of protein concentration. When 100 µg or more EcoRI endonuclease was applied, the elution profile corresponded to 63 kd, the known molecular weight of the EcoRI endonuclease dimer (30). At lower endonuclease concentrations, the elution profile shifted and broadened (i.e. at 25 µg, wild type ran as a broad peak centered at 44 kd. It was found that running the column in high ionic strength buffer (i.e. EB + 0.8 M NaCl) minimized the concentration dependence of endonuclease migration. Therefore wild-type and mutant proteins examined here were applied in EB + 0.8 M NaCl.

RESULTS AND DISCUSSION

The lethality of a wild type restriction endonuclease in a cell lacking a companion methylase provides a powerful genetic selection for null mutations of the EcoRI endonuclease. Plasmid pKG2 bearing the cloned EcoRI gene under lac repressor control was mutagenized *in vitro* with hydroxylamine and transformed into *E.coli* D1210 under non-inducing conditions. Following IPTG induction, transformants were screened for survival. Eighty per cent of the transformants survived IPTG induction and were defined as null mutants. The 20% which did not survive IPTG induction were defined as leaky mutants. Sixty-two of the 121 null mutants screened by western blot were found to contain endonuclease cross-reacting protein (western positive). The amount of endonculease cross-reacting protein and the molecular weight of the mutant proteins were similar to that of wild-type endonuclease.

To identify alterations responsible for the mutant phenotypes, the entire endonuclease gene was sequenced by the dideoxy sequencing method using double stranded plasmid DNA as template. The complete endonuclease gene sequence of 27 null mutants was determined. Sequence analysis showed that the null mutants that produce endonuclease protein are primarily single base change missense mutations (20/27). The remainder include 6 double mutations and 1 triple mutation. The number of independent isolates of each mutation and results of DNA sequencing are summarized in Table I. We obtained duplicate mutations after a relatively small number of sites had been identified. The last seven mutants analyzed revealed only one new site. Therefore it appeared that we had identified a significant fraction of the sites at

Table I. Data on EcoRI Endonuclease Mutants

Location	Amino Acid Change	Number of Isolates	IPTG Resistant
Missense			
46	Leu-Phe	1	-
56	Arg-Gln	1	-
93*	Gly-Glu	1	+
96	Glu-Lys	1	+
129	Gly-Glu	2	-
139	Ala-Val	4	+
139	Ala-Thr	2	+
140	Gly-Ser	2	+
142	Ala-Thr	1	-
144	Glu-Lys	2	+
152	Glu-Lys	1	+
164	Pro-Ser	2	+
171*	Gly-Glu	1	+
203	Arg-Gln	1	+
210	Gly-Glu	1	-
210	Gly-Arg	1	+
259	Ser-Leu	1	-
262	Ser-Leu	2	+
Nonsense			
13	Gln-Stop	1	+
56	Arg-Stop	1	+
115	Gln-Stop	1	+
134	Gln-Stop	1	+
233	Gln-Stop	1	+
240	Gln-Stop	1	+
246	Trp-Stop	2	-
Others			
IS1 insertions		3*	
Double mutants		7	
Triple mutants		1	

* Isolated from 2-aminopurine mutagenized cultures

Table I Summary of data on EcoRI endonuclease mutants. The complete endonuclease gene has been sequenced for all mutants. All of the single missense mutants tested positive on Western blots. All nonsense mutants were negative on Western blots with the exception of the mutation Trp246-Stop which was isolated in a background strain containing the SupE44 allele.

which null mutants could be obtained using hydroxylamine as the mutagen.

In vivo mutagenesis with 2-aminopurine was performed to extend the range of null mutations. However, only 4 of 34 mutants analyzed following 2-aminopurine mutagenesis contained endonuclease cross-reacting protein. The endonuclease gene was sequenced from 3 of these; two arose from missense mutations while the third was due to an IS1 insertion (Table I). Both of the single base change nulls resulted from G-A transitions. Neither of these substitutions were recovered in the hydroxylamine induced mutations even though they both could have been generated by hydroxylamine. Two mutants were sequenced from the larger group of western negative mutants and both were found to have IS1 insertions within the endonuclease gene. In E. coli D1210 IS1 has been found to insert within the endonuclease gene at a rate of between 10^{-4} to 10^{-5} (S.Y., unpublished). The rate at which IS1 inserts within the endonuclease gene appears to be greater than the rate of 2-aminopurine induced mutagenesis and discouraged us from pursuing 2-aminopurine mutagenesis.

A histogram showing the distribution of mutations throughout the endonuclease gene is shown in Fig. 1. Hydroxylamine, which causes G-A and C-T transitions in DNA, is capable of generating 218 different amino acid substitutions within the endonuclease gene through mutations at 174 of the 277 codons. Despite the large number of potential mutable sites, changes to a null phenotype were recovered at only 12 locations with significant clustering occurring between residues 139 and 144. There were no mutations identified at residues 141 or 143; however these residues are not mutable with hydroxylamine.

Figure 1

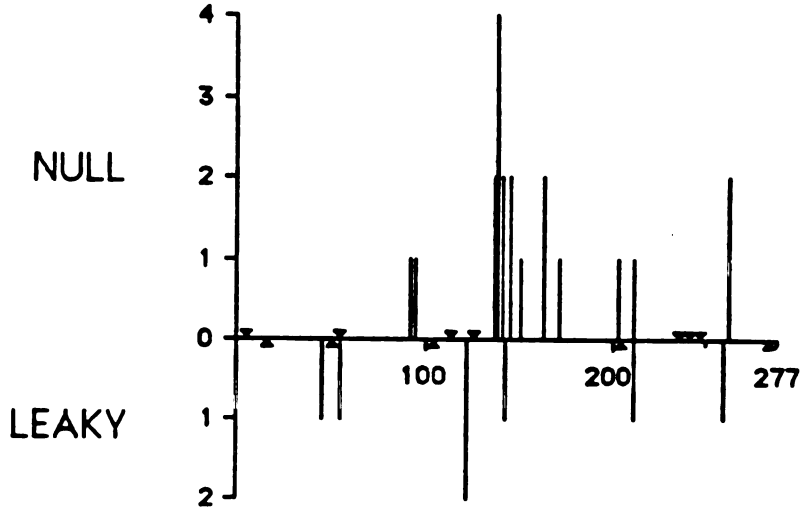


Figure 1 Histogram depicting both the location and number of isolates of each mutation relative to the primary sequence of the endonuclease. Bars above the line represent null mutants while those below the line depict leaky mutants. The small triangles represent sites at which nonsense codons could be generated by hydroxylamine; those above the line have been isolated, while those below the line have not.

DNA's from seven leaky mutants were sequenced and mutations were found at 6 loci (Fig. 1, Table I). Seven western-negative null mutants were also sequenced. Six of these contained mutations resulting in nonsense codons. The remaining mutant was the result of a double missense mutation, which apparently resulted in the production of a rapidly degraded protein. Both the leaky missense loci and the nonsense codons are distributed throughout the gene (Fig. 1, Table I). This data serves to emphasize the conclusion that the residues identified in the western-positive null mutant class are critical for enzymatic function and that the clustering is significant.

Results from mutational analysis of other proteins (31) demonstrate a large distribution of mutations which affect protein folding and intracellular stability. Mutations of this type have been shown to allow a protein to retain enzymatic activity. We have not recovered many mutations of this kind presumably because of the lethality of residual endonculease activity in cells lacking the EcoRI methylase.

The structure of the EcoRI-DNA co-crystal (7) consists of a complex formed in the absence of Mg++ between the EcoRI endonuclease and a 13 base oligonucleotide, TCGCGAATTCGCG (Fig. 2). The overall structure of the complex reveals an extensive interface between the protein and the major groove and phosphodiester backbone of the DNA. In addition there is an extensive protein-protein (dimer) interface. Hydrogen bonds formed between amino acids located in the N-terminal regions of α-helices (referred to as the inner and outer recognition helices) allow the protein to make specific interactions with nucleotide bases within the hexanucleotide recognition sequence.

Figure 2

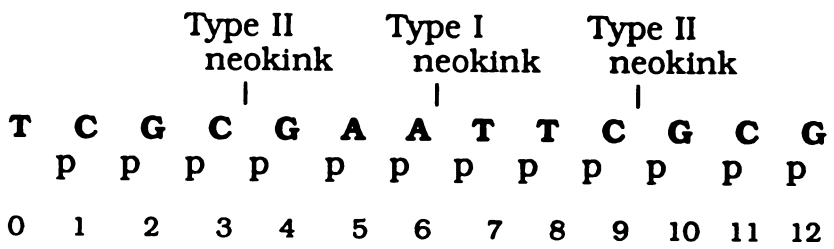


Figure 2 Sequence of the 13 base oligonucleotide used in formation of the EcoRI endonuclease-DNA complex. The phosphate numbering and the locations of the type I neokink and symmetrically related type II neokinks are shown.

Glu144 and Arg145 of opposite subunits interact with adjacent adenines in the recognition sequence forming bidentate hydrogen bonds which bridge the N6 and N7 moities. Arg200 from each outer recognition helix interacts with the guanine O6 and N7 moieties. The oligonucleotide used in formation of the complex is the same as the dodecamer solved by Dickerson et al. (32, 33) with the exception of the 5' unpaired thymine. The structure of the protein-bound DNA differs from that of the dodecamer and shows distortions (kinks) which distinguish it from unbound DNA, referred to as the type I neokink and type II neokink (Fig.2). The type I neokink effectively widens the major groove of the DNA allowing the endonuclease's recognition helices to fit into the major groove, while the type II neokink is believed to help position the scissile bond in the DNA relative to the protein.

Using the Evans and Sutherland PS300 computer graphics system, the locations of the null amino acid changes have been visualized in the structure of the endonuclease-DNA complex. The null mutants and amino acid substitutions are interpreted in terms of the current 3 Å model of the EcoRI protein-DNA structure. Two major groups are evident: protein-DNA interface mutations (Ala139-Val, Ala139-Thr, Gly140-Ser, Glu144-Lys, Arg203-Gln) and protein-protein (intersubunit) interface mutations (Glu144-Lys, Glu152-Lys, Gly210-Arg and possibly Ser262-Leu). The remaining mutations producing a null phenotype (Gly93-Glu, Glu96-Lys, Pro164-Ser and Gly171-Glu) identify amino acid residues critical for endonuclease function and will be discussed later.

DNA-protein interface

Several of the null mutants have amino acid substitutions of residues which lie at the protein-DNA interface in the wild type molecule. The proximity of residues 139, 140, 144 and 203 to the DNA is illustrated in Fig. 3. Fig. 4 shows a close-up view of the protein-DNA interface. Residues in the region of 139-144 define a critical portion of the protein-DNA interface. Residues 139-143 form a loop which packs close to the DNA, near the site of the type I neo-kink. Residues 139 and 140 are in a very tightly packed region close to phosphate 7 in the DNA backbone. Replacement of these residues by bulkier residues (Ala139-Val, Ala139-Thr, Gly140-Ser) could interfere with the correct positioning of the protein relative to the DNA. In addition, Ala139 is positioned close to His14. Replacement of Ala139 with a larger residue may interfere with its interaction with His14 and alter the position of the N-terminal residues. The N-terminal region has been shown to be required for catalytic activity (34).

Glu144 is the first amino acid of the inner recognition helix and forms hydrogen bonds to the N6 positions of adjacent adenines in the recognition sequence. Glu144 also participates in electrostatic interactions with Arg200 and Arg203. Replacement of Glu144 with Lys would eliminate the hydrogen bonds with the adenines in the DNA and the electrostatic contacts with Arg200 and Arg203.

Fig. 4 shows that Arg203 is positioned close to the DNA. This Arg makes an electrostatic contact with phoshate 3(C) and possibly with phosphate 4(G). Substitution with Gln at 203 would most likely

Figure 3 Computer graphics picture depicting one endonuclease subunit (purple) complexed with DNA (blue). The wild-type side chains are shown (yellow) for residues at the protein-DNA interface which have been substituted to generate null mutants. The substitutions are (top to bottom): Glu144-Lys, Arg203-Gln, Gly140-Ser, Ala139-Val and Ala139-Thr.

Figure 4 Computer graphics generated stereo close-up of the protein-DNA interface of one subunit showing protein (purple) and DNA (blue). Location of residues which have been substituted to generate null mutants are shown (top to bottom): Glu144-Lys, Arg203-Gln, Gly140-Ser, Ala139-Val and Ala139-Thr. This picture gives an end-on view of the outer recognition helix.

eliminate these phosphate contacts and would also affect the electrostatic interaction with Glu144. The importance of these interactions is exemplified by the null phenotype of this mutant.

Protein-protein (dimer) interface

Several of the null mutants result from mutations of amino acid residues which lie at (or near) the dimer interface. Since the active form of the enconuclease is a dimer (35), it should be possible to eliminate endonuclease activity by substitution of amino acids which interfere with dimer formation. Several null mutants resulting from substitution of amino acids at (or near) the dimer interface (Glu144-Lys, Glu152-Lys, Gly210-Arg) are shown in Fig. 5. In addition to being at the protein-DNA interface, Glu144 is in close contact with Arg145 across the subunit interface (not shown in Fig. 5). This electrostatic interaction may be crucial for dimer formation, and would be disrupted by the Lys substitution.

Glu152 is buried in the subunit interface. The closest residues across the subunit interface are Ala207 and Ala208 from the outer recognition helix and Phe271 and Leu274 from the C-terminus loop. Replacement of Glu152 with Lys would disrupt the hydrophobic intersubunit interactions.

Gly210 is located on the interior of the protein at the end of the outer recognition helix and is the site of a turn in the polypeptide backbone. Gly210 is near Leu217, Gln233 and Leu266 on the same subunit and very close to Asn155 across the subunit interface. Gln233 may hydrogen bond with Asn155 of the other subunit. Replacement of Gly210 with Arg is likely to disrupt the subunit interface between Asn155 and Gln233.

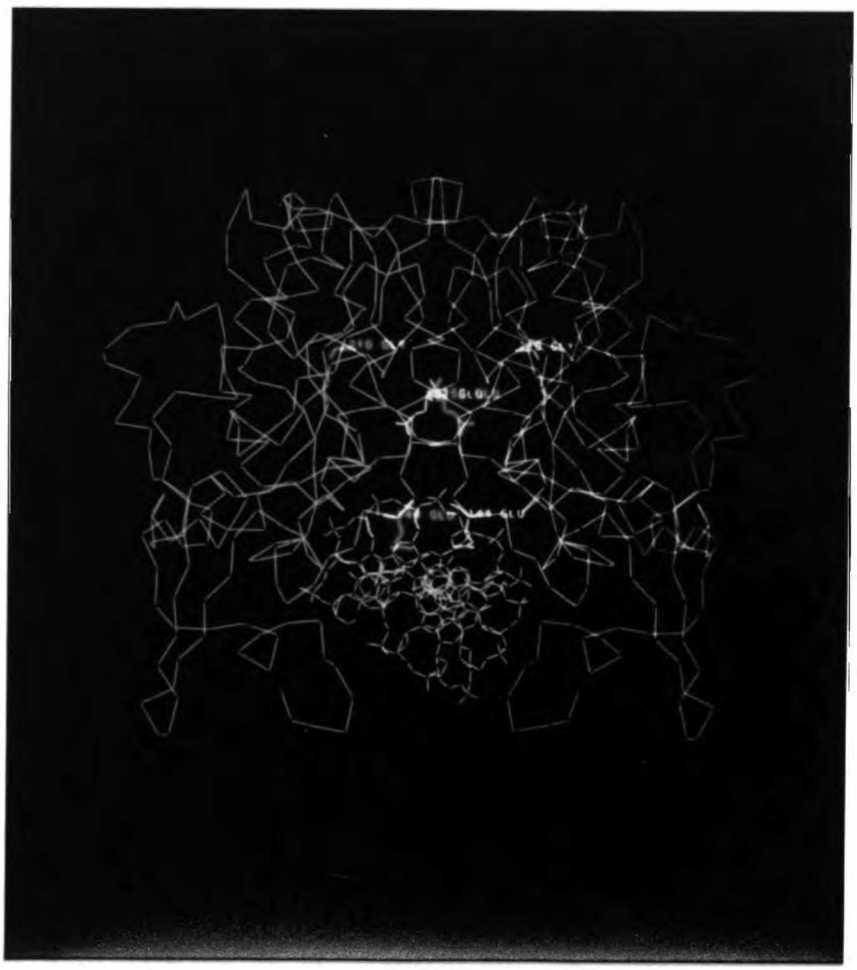


Figure 5 Computer graphics picture showing null mutations at the protein-protein (dimer) interface showing both endonuclease subunits (blue and purple) complexed with DNA (green). The symmetrically related wild-type side chains are shown (red side chains from subunit shown in blue, yellow side chains correspond to the subunit shown in purple). The substitutions are (top to bottom): Gly210-Arg, Glu152-Lys and Glu144-Lys.

Dimerization studies

In order to investigate the possibility that the null phenotype of the subunit interface mutations was due to the lack of dimer formation, null mutant proteins were purified to >90% homogeneity. Two-hundred micrograms to 1 milligram of protein was obtained for mutants Ala139-Thr, Gly140-Ser, Glu144-Lys, Glu152-Lys, Arg203-Gln and Gly210-Arg.

Purified protein was analyzed by HPLC using a BioSil TSK-250 column. Using conditions where the wild-type protein elutes as a 63 kD dimer, the Lys144 protein eluted at the position corresponding to a monomer (Table II), indicating that replacing Glu144 with Lys interferes with dimerization as predicted from the crystal structure. Protein from subunit interface mutants Glu152-Lys and Gly210-Arg was analyzed in a similar manner and these were also found to migrate as monomers. By contrast, protein from the DNA-interface mutations Ala139-Thr, Gly140-Ser and Arg203-Gln was found to migrate as dimers (Table II). Since the active form of the endonuclease is a dimer, this evidence shows that mutations which affect dimerization result in a phenotypically null endonuclease. Clearly, not all substitutions at the protein-protein interface would disrupt dimer formation. For example, substitution of Asp at residue 144 (36), which preserves the electrostatic environment and retains some enzymatic activity, does not alter the dimeric nature of the protein (Table II).

Other null mutations

Several of the null mutants do not immediately fall into the previous two groups, but possibly can be interpreted as indirectly affecting these interfaces. Pro164 is located in the interior of the

Table II. Molecular Weight Determination of Purified Proteins using BioSil TSK-250 HPLC Column.

Protein	Estimated Molecular Weight	Interpretation
EcoRI endonuclease	63,000	Dimer
Ala139-Thr	57,000	Dimer
Gly140-Ser	67,000	Dimer
Glu144-Lys	25,000	Monomer
Glu144-Asp	59,000	Dimer
Glu152-Lys	29,000	Monomer
Arg203-Gln	61,000	Dimer
Gly210-Arg	32,000	Monomer
EcoRI methylase	37,000	Monomer

protein away from either the protein-DNA or protein-protein interfaces. However, examination of the structure indicates that it may be crucial for maintaining the structure of the protein-DNA interface. Pro164 lies in the middle of β-strand 4, between Phe163 and Tyr165. The aromatic rings on these side chains are stacked up against Ile residues protruding from the inner recognition helix (Fig. 6) providing a potential structural framework for positioning this helix in the major groove of the DNA. The substitution of Ser might permit enough flexibility in the backbone of this β-sheet region to alter the positioning of the recognition helix.

Gly93 and Glu96 are located on the surface of the protein in a loop region which extends toward the DNA near the proposed Mg++ binding site and the site of catalysis. Glu96 is one of a cluster of negatively charged residues (including Asp133, Asp100, Asp99 and Asp26 on the surface of the protein). Replacement of Glu96 with Lys would alter the electrostatic environment in this region and could affect potential structural rearrangement upon Mg++ binding. The proximity of Gly93 to this cluster and the substitution with Glu could provide an adverse effect on this region.

Two of the mutants (Ser262-Leu, Gly171-Glu) are more difficult to interpret. At this time the following possibilities are presented. Ser262 is located on the surface of the molecule close to the dimer interface. Replacement of this polar group with the hydrophobic Leu may alter the folding of the C-terminal region (264-277), the residues of which are engaged in subunit-subunit and phosphate interactions.

Gly171 is located on the edge of the molecule facing the solvent and is at the end of the β-strand containing Pro164 (see above).

Figure 6 Computer graphics stereo pair depicting the relationship between Pro164 (yellow)and the inner recognition helix. The DNA (blue) and endonuclease amino acid residues 144-167 (purple) are shown. Also visible is the packing of Phe163 and Tyr165 up against Ile residues from the inner recognition helix.

Replacement of Gly171 with Glu could possibly affect the flexibility of this region of the β-strand postulated to relate to the protein-DNA interface.

Conclusions: With the aid of a strong genetic selection against active endonuclease, we have been able to identify a number of amino acid substitutions which result in an inactive endonuclease. The majority of these alterations fall at either the protein-DNA interface or at the protein-protein (dimer) interface in the crystal structure of the complex. We did not obtain any "catalytic" site mutants, i.e. enzyme which binds but can no longer cleave DNA. None of the null mutants bound DNA in preliminary binding assays performed on crude extracts. This could be due to limitations of the assay method (gel retardation assay). Binding and catalysis may not be separable, i.e. mutations at catalytic residues may alter DNA binding; alternatively, catalytically inactive proteins that can still bind DNA may be lethal.

The null mutations identified here have been interpreted with respect to the structure of the endonculease-DNA co-crystal (summarized in Table III). This analysis of the null mutants described here identifies important functional and structural regions of the EcoRI endonuclease and provides a framework for more detailed site-directed mutagenesis studies. The combination of a structural and a functional analysis provides the opportunity for a more thorough understanding of EcoRI endonculease action.

Acknowledgements: This work was supported by NIH grants GM-25729, GM-25671 and GM-33506. R. L. is a Damon Runyon-Walter Winchell fellow (#870).

Table III. Summary of Analysis of Mutant Endonucleases

Substitution	Region of EcoRI Affected
Gly93-Glu	surface (Mg++ channel?)
Glu96-Lys	surface (Mg++ channel?)
Ala139-Val	protein-DNA interface
Ala139-Thr	protein-DNA interface
Gly140-Ser	protein-DNA interface
Glu144-Lys	protein-DNA interface/subunit interface
Glu152-Lys	subunit interface
Pro164-Ser	DNA binding domain
Gly171-Glu	other
Arg203-Gln	protein-DNA interface
Gly210-Arg	subunit interface
Ser262-Leu	subunit interface

Table III Summary of substitutions resulting in a null endonuclease and effects of these alterations on the function of the endonuclease.

References

1. Hedgepeth, J., Goodman, H.M. and Boyer, H.W. DNA nucleotide sequence restricted by the RI endonuclease. Proc. Natl. Acad. Sci. USA 69:3448-3452, 1972

2. Polisky, B., Greene, P.J., Garfin, D., McCarthy, B.J., Goodman, H.M. and Boyer, H.W. Specificity of substrate recognition by the EcoRI restriction endonuclease. Proc. Natl. Acad. Sci. USA 72:3310-3314, 1975

3. Betlach, M., Hershfield, V., Chow, L., Brown, W., Goodman, H.M and Boyer, H.W. A restriction endonuclease analysis of the bacterial plasmid controlling the EcoRI restriction and modification of DNA. Federation Proc. 35:2037-2043, 1976

4. Greene, P.J., Gupta, M., Boyer, H.W., Brown, W.E. and Rosenberg, J.M. Sequence analysis of the DNA encoding the EcoRI endonuclease and methylase. J. Biol. Chem., 256:2143-2153, 1981

5. Newman, A.K., Rubin, R.A., Kim, S.H. and Modrich, P. DNA sequences of structural genes for EcoRI DNA restriction and modification enzymes. J. Biol. Chem. 256:2131-2139, 1981

6. Modrich, P. and Roberts, R.J. Type-II restriction and modification enzymes. In: "Nucleases" Linn, S.M., Roberts, R.J. eds. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 1982:109-154

7. McClarin, J.A, Frederick, C.A., Wang, B-C., Greene, P.J., Boyer, H.W., Grable, J. and Rosenberg, J.M. Structure of the DNA-EcoRI endonuclease recognition complex at 3 Å resolution. Science 234:1526-1541, 1986

8. Frederick, C.A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobsen, L., Wang, B-C., Greene, P.J., Boyer, H.W. and Rosenberg, J.M. Kinked DNA in crystalline complex with EcoRI endonuclease. Nature 309:327-331, 1984

9. Pabo, C.O. and Lewis, M. The operator-binding domain of λ-repressor: Structure and DNA recognition. Nature 298:443-447, 1982

10. McKay, D.B. and Steitz, T.A. Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. Nature 290:744-749, 1981

11. Schevitz, R.W., Otwinowski, Z., Joachimiak, A., Lawson, C.L. and Sigler, P.B. The three-dimensional structure of trp repressor. Nature 317:782-786, 1985

12. Anderson, W.F., Ohlendorf, D.H., Takeda, Y. and Matthews, B.W. Structure of the cro repressor from bacteriophage λ and its interaction with DNA. Nature 290:754-758, 1981

13.Sauer, R.T., Yocum, R.R., Doolittle, R.F., Lewis, M. and Pabo, C.O. Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. Nature 298:447-451, 1982

14.Anderson, J.E., Ptashne, M. and Harrison, S.C. Structure of the repressor-operator complex of bacteriophage 434. Nature 326:846-852, 1987

15.Kelley, R.L. and Yanofsky, C. Mutational studies with the trp repressor of Escherichia coli support the helix-turn-helix model of repressor recognition of operator DNA. Proc. Natl. Acad. Sci. USA 82:483-487, 1985

16.Eisenbeis, S.J., Nasoff, M.S., Noble, S.A., Bracco, L.P., Dodds, D.R. and Caruthers, M.H. Altered cro repressors from engineered mutagenesis of a synthetic cro gene. Proc. Natl. Acad. Sci. USA 82:1084-1088, 1985

17.Nelson, H.C.M., Hecht, M.H. and Sauer, R.T. Mutations defining the operator-binding sites of bacteriophage λ repressor. Cold Spring Harbor Symp. Quant. Biol. 47:441-449, 1982

18.Nelson, H.C.M. and Sauer, R.T. Lambda repressor mutations that increase the affinity and specificity of operator binding. Cell 42:549-558, 1985

19.Ebright, R.H. Evidence for a contact between glutamine-18 of lac

repressor and base pair 7 of lac operator. Proc. Natl. Acad. Sci. USA 83:303-307, 1986

20.Kuhn, I., Stephenson, F.H., Boyer, H.W. and Greene, P.J. Positive-selection vectors utilizing lethality of the EcoRI endonuclease. Gene 44:253-263, 1986

21.Backman, K., Ptashne, M. and Gilbert, W. Construction of plasmids carrying the cI gene of bacteriophage λ. Proc. Natl. Acad. Sci. USA 73:4174-4178, 1976

22.Birnboim, H.C. and Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513-1523, 1979

23.Miller, J.H. In: "Experiments in Molecular Genetics", Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 1972

24.Davis, R.W., Botstein, D. and Roth, J.R. In: "Advanced Bacterial Genetics", Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 1980

25.Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685, 1970

26.Towbin, H., Staehelin, T. and Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets:

Procedure and some applications. Proc. Natl. Acad. Sci. USA 76:4350-4354, 1979

27. Blake, M.S., Johnston, K.H., Russell-Jones, G.J. and Gotschlich, E.C. A rapid, sensitive method for detection of alkaline phosphatase-conjugated anti-antibody on western blots. Anal. Biochem. 136:175-179, 1984

28. Chen, E.Y. and Seeburg, P.H. Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. DNA 4:165-170, 1985

29. Greene, P.J., Heynecker, H.L., Bolivar, F., Rodriguez, R.L., Betlach, M.C., Covarrubias, S.A., Backman, K., Russel, D., Tait, R. and Boyer, H.W. A general method for the purification of restriction enzymes. Nucleic Acids Res. 5:2373-2380, 1978

30. Modrich, P. and Zabel, D. EcoRI endonuclease: Physical and catalytic properties of the homogeneous enzyme. J. Biol. Chem. 251:5866-5874, 1976

31. Pakula, A., Young, V.B. and Sauer, R.T. Bacteriophage λ-cro mutations: Effects on activity and intracellular degradation. Proc. Natl. Acad. Sci. USA 83:8829-8833, 1986

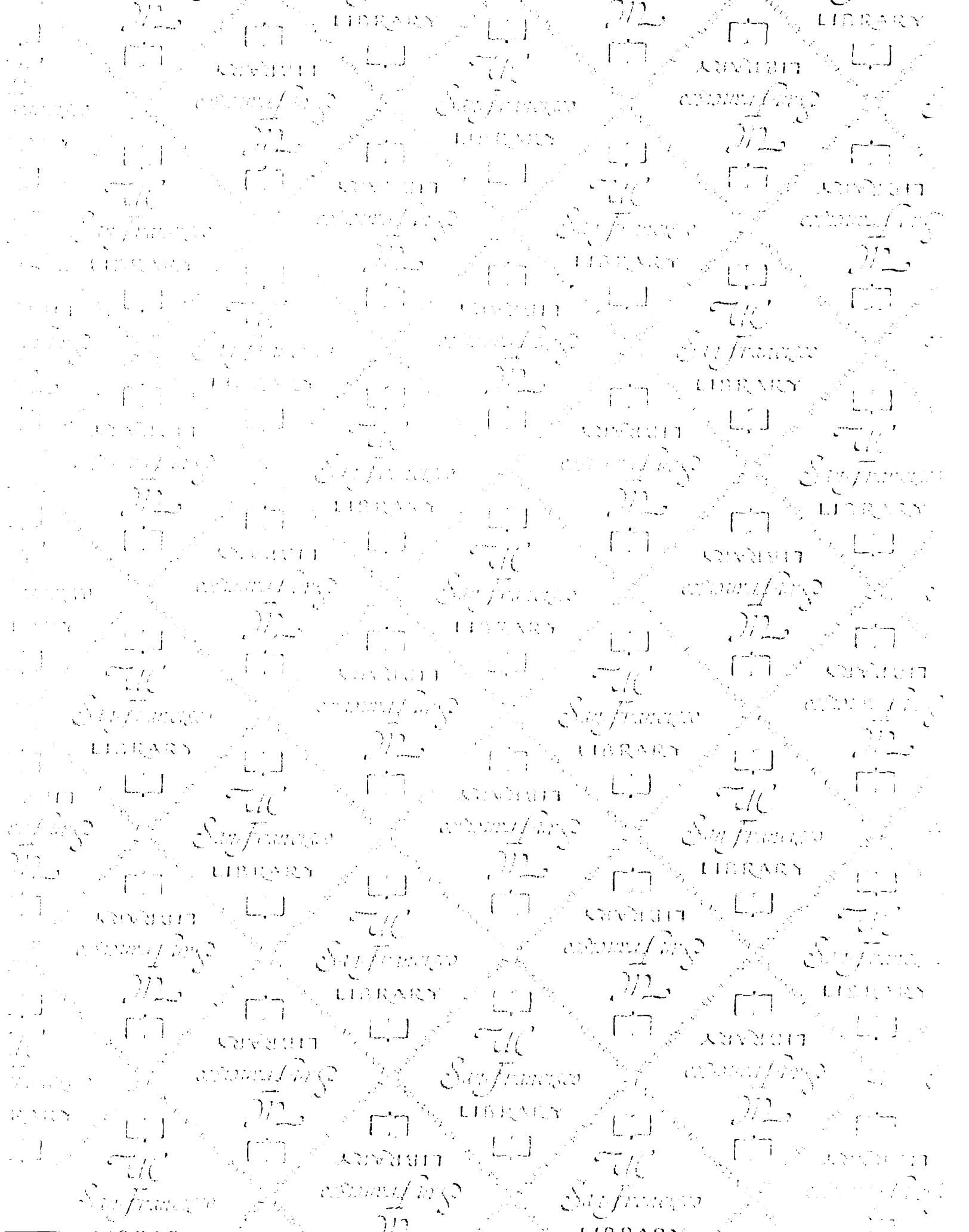
32. Dickerson, R.E. and Drew, H.R. Structure of a B-DNA dodecamer: Influence of base sequence on helix structure. J.Mol.Biol. 149:761-786, 1981

33.Drew, H.R., Wing, R.M., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. Structure of a B-DNA dodecamer: Conformation and dynamics. Proc. Natl. Acad. Sci. USA 78:2179-2183, 1981

34.Jen-Jacobson, L., Lesser, D. and Kurpiewski, M. The enfolding arms of EcoRI endonuclease: Role in DNA binding and cleavage. Cell 45:619-629, 1986

35.Jen-Jacobson, L., Kurpiewski, M., Lesser, D., Grable, J., Boyer, H.W., Rosenberg, J.M. and Greene, P.J. Coordinate ion pair formation between EcoRI endonuclease and DNA. J. Biol. Chem 258:14638-14646, 1983

36.Greene, P.J., Yanofsky, S., Reich, N., Day, J., Hager, P., Boyer, H.W., McClarin, J.A., Frederick, C.A., Wang, B-C., Grable, J., Love, R.L. and Rosenberg, J. Structure and function of the EcoRI endonuclease. UCLA Symposia on Molecular and Cellular Biology, New Series, Volume 69, Oxender, D. ed. Alan R. Liss, Inc., New York, NY, 1987 (in press)


San Francisco
LIBRARY