1989

On the Repair of DNA Breaks and the Specificity of the EcoRI Restriction Enzyme

Joseph Heitman

On the Repair of DNA Breaks

and the Specificity of the *Eco*RI Restriction Enzyme

A thesis submitted to the Faculty of The Rockefeller University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

by
Joseph Heitman

April 7, 1989
The Rockefeller University
New York, New York

"We shall not cease from exploration

And the end of all our exploring

Will be to arrive where we started

And know the place for the first time."

T. S. Eliot

from "Little Gidding", *Four Quartets*

Table of Contents

List of Figures

List of Tables

Acknowledgements

While I was an undergraduate at the University of Chicago, I started my scientific career as a synthetic organic chemist. I originally decided to come to Rockefeller because I wanted to study DNA-protein interactions by chemically synthesizing DNA-binding peptides. However a few months before leaving Chicago, I started a project with Malcolm Casadaban which seduced me away from chemistry and into genetics. I arrived only to find that my chosen laboratory was still entrenched in the chemical synthesis of peptides and had neither incubators nor shaking water baths. At this point Dr. Kaiser suggested I speak with Peter Model. This turned out to be the best advice anyone ever gave me. In the Model/Zinder laboratory I discovered a group of broad minded individuals primarily using genetics to dissect the function of proteins. The work described here would not have been possible without their advice and encouragement.

For creating what was a special place for me I thank my advisor, Peter Model. From my first unscheduled appearance, Peter strongly encouraged me to pursue my own independent project, was tolerant when I was stubborn and ignored even good advice and, through our daily conversations, left an impact on my scientific interests and convictions. I want to thank Norton Zinder for challenging me from the first day I arrived when he laughed at my project and told me he hoped I had twenty years to work on it, for pushing both me and my data further than we both thought we could go, and for strongly encouraging me at several points to pursue one project and neglect another. I would like to thank my three closest compatriots in the lab: Wilder Fulford, Benedicte Michel, and Ted Bargiello, for their insights and encouragement which helped bring this work to birth and to fruition. Marjorie Russel, Ken Horiuchi, and Nick Davis were very generous with their time to discuss results and suggest further experiments. I want to acknowledge the technical assistance of Tracy Ripmaster. Lastly, I want to thank my parents Millie and Bob, my sister Jen, my roommates Jeff and Brian, and mi novia Maria Elena for pulling me back from the scientific precipice and making life what it is at its best, an adventure.

Acknowledgements-Scientific

I want to thank the following people for discussions and encouragement, and for sharing their strains, plasmids, and unpublished data.

Kan Agarwal
Barbara Bachman
Francis Barany
Ashok Bhagwat
Bob Blumenthal
Joan Brooks
Malcolm Casadaban
John Clark
Bruce Duncan
Joe Fried
Susan Gottesman
Dick Gumport
Stan Hattman
Bill Holloman
Toni Kazic
Catherine Joyce
Antal Kiss
Marty Marinus
Paul Modrich
Lisa Raleigh
Helen Revel
Rich Roberts
Lucia Rothman-Denes
Ira Schildkraut
Allan Shatzman
Laura Sznyter
Michael Volkert
Joe Walder
Roxanne Walder
Graham Walker
Geoff Wilson
N. C. Yang

Abbreviations

bp	base pairs
D	dimer RFI form DNA
ds	double-stranded
e.o.p	efficiency of plating
EtBr	ethidium bromide
IPTG	isopropyl-β-D-thio-galactopyranoside
M	modification or methylase
m.o.i.	multiplicity of infection
OD	optical density
ONPG	o-nitrophenyl-β-D-galactopyranoside
R	restriction
RF	replicative form
RFI	closed circular supercoiled ds-DNA
RFII	nicked circular ds-DNA
RFIII	linear ds-DNA
RFIV	relaxed, closed circular ds-DNA
ss	single-stranded
ts (or TS)	temperature sensitive
wt (or WT)	wild type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Abstract

We would like to understand how proteins and enzymes interact with DNA. We describe here our studies of the Mrr methylation dependent restriction system, DNA single- and double-strand break repair in *E. coli*, and substrate recognition by the *Eco*RI endonuclease.

Many species of bacteria make restriction-modification systems to destroy foreign DNA that enters the cell. These systems usually consist of an endonuclease that cleaves a specific DNA sequence and a methylase which modifies the DNA to protect the host chromosome. We observed that when foreign site-specific methylases are expressed in *E. coli*, the SOS DNA repair response is induced. This DNA damage is inflicted by *E. coli* restriction enzymes that cleave adenine (Mrr) or cytosine (McrB) methylated DNA. The genes encoding four of the five known *E. coli* restriction systems lie clustered together, perhaps to coordinate or sequester the cellular defense system. Several of these restriction systems differ between *E. coli* species, suggesting that their action may establish species boundaries.

The *Eco*RI endonuclease cleaves DNA molecules at the sequence GAATTC. This enzyme is well characterized biochemically, the sequence of its gene is known, and the X-ray crystal structure of an *Eco*RI-DNA complex has been solved at 3 A resolution. (McClarin, *et al.*, 1986). *Eco*RI serves as a paradigm for other restriction enzymes and as a model of DNA-protein interactions.

We took a genetic approach to study the *Eco*RI endonuclease. We first asked if *Eco*RI DNA double-strand breaks are repaired in *E. coli*. To this end, a series of temperature-sensitive *Eco*RI endonuclease alleles were isolated. Temperature shifts with these alleles revealed that *in vivo* DNA scission induces the *E. coli* SOS DNA repair response. However, neither SOS induction nor recombination are required to repair these lesions. DNA ligase is required and may suffice to repair *Eco*RI breaks in the *E. coli* chromosome.

An *in vivo* DNA scission assay was devised based on the finding that DNA breaks induce the SOS response. SOS induction was monitored with strains carrying the lactose operon fused to an SOS inducible promoter. After DNA scission, these strains produce *3*-galactosidase and form blue colonies on X-Gal medium. With this blue colony phenotype as a screen, two approaches were taken to isolate *Eco*RI mutants altered or disrupted in substrate specificity. First, amino acids (E144, R145, R200)

implicated in substrate binding by the crystal structure were subjected to site-directed mutagenesis. Of 50 of the 60 possible substitutions, several alleles retain weak endonuclease activity which, *in vivo* and *in vitro*, is of wild-type specificity. Therefore the simple hydrogen bond model proposed from the crystal structure is insufficient to explain substrate recognition and additional interactions must participate in the substrate-enzyme complex.

In the second approach, mutants of an *Eco*RI TS allele were isolated which conditionally induce the SOS response and impair cell growth in spite of the normally protective methylase. *In vitro*, these mutant proteins exhibit enhanced cleavage activity at *Eco*RI* sites, sequences which differ by one nucleotide from the normal recognition site and are also cleaved by the wild-type enzyme under altered buffer conditions. Four of the five mutations of this type lie at the DNA-protein interface and may directly alter or disrupt substrate recognition. One other (H114Y) lies far from the binding and cleavage sites. This mutation falls three amino acids away from a previously described mutation, E111G (King, *et al.*, 1986, 1988), which severely impairs cleavage activity without altering DNA binding. These two mutations support a model whereby DNA scission by the *Eco*RI endonuclease is allosterically activated upon substrate binding: we suggest that the E111G mutation inhibits this conformational change while the H114Y mutation renders it more facile such that additional DNA sequences act as allosteric effectors and trigger cleavage.

Chapter 1

Introduction

I would like to know how proteins and DNA interact. Since several proteins which bind DNA are now understood in some detail, I concentrated instead on enzymes which interact with DNA. Although DNA-binding proteins and DNA-binding enzymes might be expected to share some features, we anticipated that important differences would arise from the need to coordinate binding with catalysis. We chose to study the *Eco*RI endonuclease for several reasons. Its gene was cloned and sequenced, its purification and biochemical characterization had been well described and, at the outset of this work, the crystal structure of an *Eco*RI-DNA complex was being solved.

Although *Eco*RI is most famous for its part as a reagent of molecular biology, its true role in life is to cleave DNA molecules at the sequence GAATTC and thereby to defend its host bacterium against invasion by foreign DNA molecules (Hedgpeth, *et al.*, 1972). Its partner in this endeavor, the *Eco*RI methylase, recognizes the same DNA sequence and alters it by transferring methyl groups from S-adenosyl methionine to the N^6-position of the central adenine on each strand (Rubin & Modrich, 1977). This chemical modification blocks scission by the endonuclease and protects the cellular DNA from destruction. Many species of bacteria elaborate similar enzymatic defense systems, called restriction-modification systems, which act as a sort of bacterial immune system to destroy foreign DNA that enters the cell (Roberts, 1976).

Restriction-Modification systems

Bertani and Weigle first described the phenomenon of restriction and modification (Bertani & Weigle, 1953). They observed that stocks of λ phage prepared from *E. coli* strain C grew poorly when propagated on a different strain, K-12. This effect

was named *restriction.* Rare phage that escaped restriction produced progeny which now grew with equal efficiency on either host. That these survivors of K-12 infection were not mutants could be demonstrated by passaging them again on the *E. coli* C host and showing that they regained sensitivity to restriction. In some way they were transiently altered or *modified* during growth on *E. coli* K-12. We now know that specific enzymes form the molecular basis underlying this epigenetic phenomenon (Arber, 1974; Roberts, 1976). Restriction results from the action of an endonuclease that recognizes a specific DNA sequence and cleaves there or elsewhere. Modification is the work of a methylase which recognizes the same DNA sequence and methylates it to protect the host chromosome from endonuclease action. Phage resistant to restriction (modified) occur when the methylase modifies each phage-borne recognition site prior to endonuclease action.

Based on the properties of the purified enzymes, restriction modification systems fall into three classes: type I, type II, and type III (Bickle, 1987). Type I systems are the most complex. They consist of three subunits (R=restriction, M=modification, and S=specificity) which form a pentameric (R_2M_2S) multifunctional enzyme, exhibit complex cofactor requirements, cleave at random sites distant from their recognition sites, unwind the DNA to form a loop between their sites of recognition and cleavage, and are inactive after a single round of DNA scission. Type III systems are the most mysterious: they consist of two subunits, R and M, which form a methylase of subunit M alone and a methylase-endonuclease complex containing both subunits. More curiously they require but do not hydrolyze ATP as a co-factor during DNA scission, cleave at unique sites but yield partial digests that do not go to completion, and modify only one DNA strand raising the conundrum of how the unmethylated daughter molecule produced after replication is protected. Type II systems are the simplest. The endonuclease is typically a homodimer which is distinct from and shows no sequence homology to its partner, a monomeric methylase. This difference in subunit arrangement may arise from the nature of the recognition sites. The en-

donuclease usually binds a palindromic DNA sequence and cleaves both strands symmetrically. The methylase's typical substrate is hemimethylated DNA that arises after replication and hence is asymmetric because it already bears one methyl group. Thus as a monomer, the methylase is well suited for its principal task of simply methylating the one remaining strand. Each enzyme requires a single cofactor: Mg^{+2} for scission and *S*-adenosylmethionine (SAM) for methylation.

Most restriction enzymes cleave unmethylated DNA and their action is blocked by DNA methylation. The first exception was the type II restriction enzyme *Dpn*I from *Streptococcus pneumoniae*, which cleaves the sequence GATC only when methylated at the N^6 position of both adenines (Lacks & Greenberg, 1977). While studying cloned restriction-modification systems from bacterial species other than *E. coli*, we and others (Blumenthal, *et al.*, 1985; Kiss, *et al.*, 1985; Raleigh & Wilson, 1986; Noyer-Weidner, *et al.*, 1986; Heitman & Model, 1987) observed that expression of the methylase genes alone was detrimental and in some cases lethal to *E. coli*. We showed that these foreign methylases induce a DNA repair system, the SOS response (see chapter 3). In contrast, several methylases indigenous to *E. coli* (*hsdM, dam, dcm*) produce the same chemical modifications as these noxious methylases yet do not induce the SOS response. Since this effect requires sequence specific methylation, it suggested that restriction rather than DNA repair might be the cause. Using a foreign methylase, we isolated an *E. coli* mutant in which SOS is not induced by adenine methylation and showed that the wild-type locus (*mrr*) behaves as a restriction system which cleaves adenine methylated DNA. Concurrently, Lisa Raleigh and co-workers demonstrated that cell lethality by cytosine methylases was due to the *mcrA* and *mcrB* loci, previously shown to encode restriction systems that cleave DNA containing 5-hydroxymethyl cytosine (Revel, 1983; Raleigh & Wilson, 1986). We showed that, as for the *mrr* locus, SOS induction by cytosine methylases was also attributable to restriction, in this case by the *mcrB* locus. These findings suggest that when foreign methylases are expressed in *E. coli* the chromosome becomes methylated

and subsequently suffers scission by the Mrr, McrA, or McrB restriction systems. SOS induction enhances cell growth under these conditions, suggesting that the products of SOS induced genes repair these DNA double-strand breaks.

Two other examples of methylation dependent restriction systems have been described (Sladek, *et al.*, 1986; MacNeil, 1988). These enzymes may be widespread in nature since first, they are parsimonious and require only a single enzyme and second, they would not have been detected by the usual screens for restriction enzymes. Whether or not these systems are related to the type I, II, or III restriction-modification systems awaits purification and characterization of the enzymes. By the present classification scheme, they already differ drastically by lacking methylation activity, and hence might be more appropriately termed type IV restriction systems.

DNA Repair

DNA repair in *E. coli* is carried out by large number of enzymes whose synthesis is either constitutive or induced by DNA damage. Two sets of inducible DNA repair genes have been described: the SOS response and the adaptive response (reviewed by Walker, 1984, 1985). Alkylating agents activate the adaptive response and the induced products mediate repair of alkyl-DNA lesions. The SOS response is a globally altered physiological state that arises after DNA damage (by UV light, mitomycin C, or nitrosoguanidine for example) or inhibition of DNA replication (i. e. nalidixic acid). SOS induction results from a cascade of events which activate a set of genes whose products slow cell division and repair DNA. After DNA damage, the RecBCD or RecF proteins (or other pathways) produce an intermediate (probably single-stranded DNA) that activates the RecA protein to a form called RecA*. RecA* then stimulates the autodigestion of the LexA protein, the common repressor of the SOS genes, inducing the SOS genes. Genes under control of the LexA repressor include *lexA* itself and also *recA* (recombination and SOS induction), *uvrABC* (excision repair), *umuCD* (error-prone repair), *recNQ* (Irino, *et al.*, 1986) and *ruv* (mediate the RecF

pathway of recombination with *recFJO*), *sfiA* (inhibitor of cell division), *himA* (subunit of IHF), *ssb* (single-strand DNA binding protein), and *dinABDF* (damage inducible genes of unknown function). Several other SOS induced functions, such as λ and the colicin antibiotics, are not present in all strains. In the case of λ lysogens, the λ *cI* repressor is subject to RecA* sponsored autodigestion. λ induction requires a higher dose of DNA damage compared to LexA regulated genes. This ensures that repair will be stimulated at low doses of DNA damage, whereas at higher doses λ will be induced and thereby escape the dying cell.

DNA double-strand breaks stimulate recombination in *E. coli* and yeast, and recombination is generally thought to repair DNA breaks (Resnick, 1976; Szostak, *et al.*, 1983; Thaler & Stahl, 1988). The double-strand break repair model suggests that an intact copy of the severed DNA acts as a template for recombinational repair (Szostak, *et al.*, 1983). This requires either a diploid genome or multiple copies of the chromosome in a haploid. Repair of DNA scissions has been studied previously using γ-ray lesions (Krasin & Hutchinson, 1977, 1981). In *E. coli*, prior induction of the SOS DNA repair response enhances the repair of γ-ray cleaved DNA. Two SOS proteins involved in recombination, RecA and RecN, and multiple copies of the genome are required to repair γ-ray DNA scissions, suggesting that recombination repairs these lesions. *RecA*$^-$ mutant strains are also more sensitive to scission by the λ terminase endonuclease, again implicating recombination in repair (Murialdo, 1988). Ligation may also repair some γ-ray breaks (Weibezahn & Coquerelle, 1981).

The model describing repair of γ-ray breaks may be compromised because γ-ray lesions include not only breaks but also nicks, base adducts, and oxidized proteins (Friedberg, 1985). After γ-ray treatment, single-strand nicks are 10-20 fold more abundant than breaks. Furthermore, DNA scission by γ-rays (and the radiomimetic bleomycin) generally shatters the sugar residue, releasing the attached base and leaving behind blunt termini with 5'-phosphoryl and 3'-phosphoryl or 3'-phosphoglycolate ends which cannot be ligated (Henner, *et al.*, 1983; Giloni, *et al.*, 1981). To repair

such a lesion by ligation would first require processing to produce a 3'-hydroxyl; even so, direct ligation would result in the loss of at least one basepair. Thus, complete restoration of the damaged sequence probably requires recombination.

In contrast, type II restriction endonucleases cleave DNA to yield staggered or blunt double-strand breaks with 3'-hydroxyl and 5'-phosphoryl termini (Roberts, 1976). Using a set of temperature sensitive *Eco*RI mutants (chapter 4), we find that although DNA scission by *Eco*RI induces the SOS response, neither this induction nor recombination are required for repair (chapter 5; Heitman, *et al.*, 1989a). DNA ligase is necessary and may be sufficient to repair *Eco*RI staggered breaks in the *E. coli* chromosome. Since *E. coli* DNA ligase cannot or poorly ligates blunt termini, different mechanisms may be required for their repair. We suggest that the nucleoid or higher order structure of the *E. coli* chromosome (Worcel & Burgi, 1972) holds the severed DNA such that each end is religated to its appropriate partner.

*Eco*RI scissions clearly activate both the λ Red and chi-mediated RecBCD recombination pathways (Stahl, *et al.*, 1983; Thaler, *et al.*, 1987a, 1987b). If ligation repairs *Eco*RI breaks, how do they stimulate recombination? Our findings suggest that these ends act as transient entry points for recombination machinery and are then resealed. By this means, recombination results from but does not repair the break. Because the RecBCD complex requires DNA ends for its recombination activity, we suggested that the cell may make an endonuclease to stimulate recombination. In fact, Mahan and Roth (1989) recently suggested that RecBCD enters the DNA at DNA double-strand breaks which occur during recombination.

DNA-Protein Interactions

DNA-binding proteins

The double helix model for DNA suggested that its complementary base paired structure was the key to faithful maintenance of the genome (Watson & Crick, 1953).

However, the mechanics who transcribe, modify and maintain this structure (by transcription, recombination, replication and repair) are DNA binding proteins. The ways they interact with DNA remain a central question of molecular biology.

Sequence comparisons of many site-specific DNA binding proteins has revealed two conserved structural motifs: the helix-turn-helix (Pabo & Sauer, 1984) (also encompassed within the homeodomain) and the zinc finger (Miller, *et al.*, 1985) (both reviewed in Schleif, 1988). Although genetic and biochemical evidence argue that the zinc finger domain binds DNA, there is as yet no crystal structure of a zinc finger containing protein. Genetic and X-ray crystallographic analysis reveal that the helix-turn-helix motif does in fact bind this class of dimeric proteins to DNA by projecting one α-helix from each subunit into adjacent major grooves of the DNA. Sequence discrimination results from hydrogen bonds, van der waals interactions, and avoidance of steric clash between functional groups of the protein and the nucleotides and backbone of the DNA.

High resolution structures of DNA-protein co-crystals of the phage 434 repressor (Anderson, *et al.*, 1987; Aggarwal, *et al.*, 1988) λ *cI* protein (Jordan & Pabo, 1988), and *E. coli trp* repressor (Otwinowski, *et al.*, 1988) suggest that amongst proteins of the helix-turn-helix family, the details of specific DNA sequence recognition differ. Within the λ *cI*-DNA complex, the DNA is of B-conformation and while some contacts to the sugar-phosphate backbone position the recognition helix, sequence discrimination is by hydrogen bonds and van der waals contacts to the nucleotides. In contrast, the DNA to which the 434 repressor binds is distorted and sequence discrimination by contacts to the nucleotides is augmented by the ability of specific DNA sequences to undergoe conformational deformations that allow further contacts to the backbone. Similarly, the DNA within the *trp* repressor-operator complex deviates from B-form. Furthermore, the protein makes no direct contacts to the base pairs of the DNA, but rather binds through solvent mediated hydrogen bonds to the bases and extensive contacts to the backbone. The specific predictions of these

DNA-protein interaction models have been well corroborated by genetic (Kelley & Yanofsky, 1985; Wharton & Ptashne, 1987; Benson, *et al.*, 1988; Bass, *et al.*, 1988; Benson & Youderian, 1989) and biochemical analysis (Koudelka, *et al.*, 1987; Nelson & Sauer, 1986; Klig, *et al.*, 1987). These findings suggest that while these DNA binding proteins share a common structural element which allows them to project an α-helix into the major groove, they take advantage of whatever contacts are available and differ significantly in the detailed interactions this helix makes with the DNA. In this respect DNA-protein interactions resemble antigen-antibody interactions.

Many DNA binding proteins do not carry either the helix-turn-helix or the zinc finger motif. These include sequence specific DNA binding proteins such as Mnt and Arc, which bind DNA with their N-terminal domains of unknown structure (Knight & Sauer, 1989), the non-specific DNA binding protein HU of *Bacillus stearothermophilus*, whose X-ray crystal structure reveals a lobster like structure which enwraps the DNA at a concave base between two anti-parallel β-ribbons (Tanaka, *et al.*, 1984), and the single-stranded DNA binding protein of phage f1, gene V, which carries aromatic residues that intercalate between the base pairs (Brayer & McPherson, 1984).

DNA-binding enzymes

While enzymes which interact with DNA may share some features of simple DNA binding proteins, they may employ different structural motifs to coordinate their binding with catalysis. Notably, a DNA-enzyme co-crystal structure of the *Eco*RI endonuclease shows that the DNA backbone is significantly distorted as is the case with the 434 and *trp* repressors (McClarin, *et al.*, 1986). However, the enzyme lacks the helix-turn-helix motif. Instead (as described below) the protein bears two α-helices on each subunit which project steeply into the major groove, an arrangement which may be necessary to accommodate a shorter DNA recognition site (6 BP versus 17-18 BP for the repressors). Bovine pancreatic DNase I binds through an exposed loop which

projects into the minor groove to make extensive contacts to the phosphodiester backbone which also significantly distort the DNA (Suck, *et al.*, 1988). The Klenow fragment of DNA polymerase I is made of two DNA binding domains that correspond to two different enzymatic activities (Ollis, *et al.*, 1985; Freemont, *et al.*, 1988). The 3'-5' exonuclease domain interacts with single-stranded DNA by intercalating aromatic residues, like gene V protein, which are superimposed upon a catalytic site of acidic and metal binding residues reminiscent of the active site of another enzyme, *Staphylococcal* nuclease. The polymerase DNA binding domain lies within a deep crevice covered by a flexible domain that may allow the enzyme to surround the DNA and project two of its α-helices into the major groove of the DNA in a way similar to the related helix-turn-helix motif. Thus DNA binding enzymes do appear to borrow some features of simple binding proteins.

The EcoRI restriction-modification system

Our ability to manipulate and study DNA molecules depends upon a special class of enzymes, the restriction enzymes, which recognize and cleave within specific DNA sequences with exceedingly high substrate fidelity. We have studied one such enzyme, the *Eco*RI endonuclease, both as a model of DNA-enzyme interactions and to understand how it recognizes its substrate with such precision. Our interest was in part motivated by the idea that if we understood how the protein recognizes its substrate, we might be able to design restriction enzymes with new specificities.

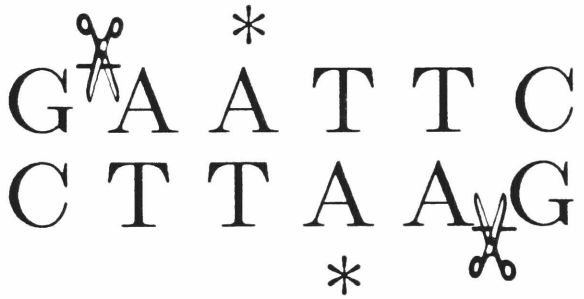
The *Eco*RI endonuclease is the paradigm for all the type II restriction enzymes. Although well characterized biochemically and crystallographically, when we began this work few mutants had been described (Jen-Jacobsen, *et al.*, 1983). We therefore took a genetic approach to dissect the mechanism of DNA recognition by the *Eco*RI enzyme.

The *Eco*RI endonuclease cleaves double-stranded DNA molecules at the sequence GAATTC and confers upon its host bacterial cell the ability to destroy foreign DNA

molecules (such as viruses) that enter the cell (Hedgpeth, *et al.*, 1972) (see figure 1-1). Its compatriot is the *Eco*RI methylase, which protects the host chromosome from destruction by adding methyl groups to the recognition site (Rubin & Modrich, 1977). The *Eco*RI endonuclease is a symmetric homodimer of known sequence which requires Mg^{+2} as a co-factor for enzymatic activity (Newman, *et al.*, 1981; Greene, *et al.*, 1981). Its reaction kinetics (Jack, *et al.*, 1982), DNA binding properties (Terry, *et al.*, 1983, 1985), activity with oligonucleotides containing modified nucleotides (Brennan, *et al.*, 1986a; McLaughlin, *et al.*, 1987), and stereochemistry of cleavage (Connolly, *et al.*, 1984) have been well described. They suggest that the enzyme initially binds DNA non-specifically, diffuses in a one-dimensional search to find its recognition site, interacts specifically with six nucleotides imbedded within a 10-12 nucleotide span covered by the enzyme, undergoes an allosteric conformational change induced by binding substrate (King, *et al.*, 1986, 1988; McClarin, *at al.*, 1986), and catalyzes strand scission by sequential cleavage with release of the doubly cleaved product as the rate limiting step. Cleavage proceeds with inversion of configuration at the scissile phosphodiester bond, indicative of an odd number of reaction steps which is consistant with cleavage by a hydroxyl ion rather than a covalent enzyme-DNA intermediate. Furthermore, it has been observed that under altered buffer conditions (higher pH, low salt, low ionic strength buffer, the presence of glycerol or organic solvents (DMSO), or with Mn^{+2} instead of Mg^{+2} as co-factor) the enzyme cleaves at additional sites, related to the wild-type substrate by one or more nucleotide substitutions. This activity is called *Eco*RI* activity (Polisky, *et al.*, 1975; Hsu & Berg, 1978; Malyguine, *et al.*, 1980; Woodbury, *et al.*, 1980; Gardner, *et al.*, 1982; Rosenberg & Greene, 1982).

In contrast to the endonuclease, the *Eco*RI methylase is a monomer which catalyzes methyl transfers in independent events (Rubin & Modrich, 1977). Although both enzymes bind the same DNA sequence by interactions in the major groove, they show no sequence homology. Methylation protection and nucleotide analog studies

Figure 1-1. The *Eco*RI Recognition Site. The *Eco*RI endonuclease and methylase recognize this palindromic hexanucleotide DNA sequence. DNA scission occurs on both strands between the guanine and the adenine at the scissor positions. Methylation is at the N^6-position of the central adenines, indicated by the astericks.


G A A T T C
C T T A A G

reveal that these enzymes do not bind this sequence in the same way. In some cases, they interact with different functional groups on the same nucleotide and in others, bind to opposite members of a basepair. (Modrich & Rubin, 1977; Lu, *et al*, 1981; Brennan, *et al.*, 1986b).

We now know about the endonuclease-DNA interaction in some detail. The crystal structure of an *Eco*RI endonuclease-DNA complex (solved to 3 A resolution in the absence of Mg^{+2} to prevent cleavage) reveals two α-helices in each monomer (the inner and outer helices) which carry amino acids that project into the major groove and interact with the purine nucleotides of the recognition site (Frederick, *et al.*, 1984; McClarin, *et al.*, 1986). Glutamic acid 144 and arginine 145 lie at the end of the inner α-helix and each, from opposite subunits, spans and makes two hydrogen bonds with the central adenines (gAAttc). Arginine 200 lies at the end of the outer helix and makes two hydrogen bonds to the outer guanine nucleotide (Gaattc). Contacts to the right and left halves of the palindromic substrate are symmetric. Thus for the dimer, a network of twelve hydrogen bonds form the substrate binding pocket.

Previously, conservative substitutions of these substrate binding amino acids have been described (glu144asp, arg145lys, arg200lys) (Wolfes, *et al.*, 1986; Greene, *et al.*, 1987). These mutations reduce enzymatic activity but do not alter substrate specificity. Since mutants bearing conservative amino acid substitutions may make contacts similar to the wild-type enzyme, we wished to test whether other amino acid substitutions at these three positions would alter the substrate specifity of the enzyme. To this end we developed a simple and sensitive *in vivo* plating assay for endonuclease activity.

Our earlier finding that *Eco*RI action induces the SOS response suggested a simple *in vivo* DNA scission assay. SOS induction can be conveniently monitored with strains in which the lactose operon has been fused to an SOS inducible promoter (Kenyon & Walker, 1980). After DNA damage, these strains make β-galactosidase,

which can be readily detected by the formation of blue colonies on indicator medium containing the chromogenic substrate X-gal. We have employed this blue colony phenotype to probe substrate recognition in two ways. First, the phenotypes were determined for site-directed *Eco*RI mutants altered in the substrate binding pocket (see chapter 6). These changes did not alter the substrate specificity of the enzyme *in vivo* or *in vitro*. Thus the enzyme-substrate interaction is not explicable by the simple hydrogen bond model predicted from the crystal structure. The enzyme must make additional contacts to recognize its substrate. These may be interactions with either the pyrimidines of the recognition site or the unusual conformation of the sugar-phosphate backbone observed in the co-crystal structure (Frederick, *et al.*, 1984).

We have also isolated mutants which cleave DNA in spite of the presence of the protective methylase (see chapter 7). *In vitro* these mutants exhibit enhanced activity at *Eco*RI* sites. Four of five mutants of this type result from amino acid changes at the DNA-protein interface, near to but distinct from the amino acids implicated in substrate binding by the crystal structure. The other lies in a region of the protein distant from both the binding and cleavage sites. This mutation lies three amino acids away from another interesting mutation (E111G) (King, *et al.*, 1986, 1988), which does not alter sequence specific DNA binding but completely blocks DNA scission. It has been suggested that *Eco*RI scission is allosterically activated by substrate binding, and that the E111G mutation disables the allosteric coupling between the binding and cleavage sites. We suggest that the H114Y mutation shifts the protein towards the allosterically activated conformation, allowing less than perfect substrates to act as allosteric effectors and trigger promiscuous cleavage by the mutant enzyme.

Restriction enzymes may differ from simple binding proteins because their shorter recognition sequences demand a more compact recognition domain. Additionally, binding proteins must often bind several related operators which differ in sequence,

while restriction enzymes require strict sequence recognition to avoid damaging the cellular DNA. Restriction enzymes attain their high fidelity through extensive contacts to the substrate and by coupling binding to cleavage through allosteric activation.

Chapter 2

Materials and Methods

I. Bacterial strains

The bacterial strains used in this study are described in Table 2-1. The majority are isogenic with strain K38 (originally S26 of A. Garen) whose complete genotype and ancestry are described more fully in Bachman (1987). Strains were constructed by P1 transduction using P1 Cm^r *clr*100 as previously described (Russel & Model, 1983). To more easily manipulate markers, we frequently employed nearby Tn*10* insertions (for example *srl300*::Tn*10* for *recA*), isolated tetracycline resistant transductants, and then scored for the desired marker phenotype. *RecA*, *recB*, and *lexA3* mutant strains were identified by their sensitivity to UV irradiation, and *recN* mutants by sensitivity to mitomycin C. In some cases strains harboring Tn*10* insertions were cured of tetracycline resistance by quinaldic acid selection (Maloy & Nunn, 1981). Tn*5* insertion mutagenesis was as described by Kleckner *et al.* (1975) using infection with λ^{467} (*c*I857 *b*221 *O*am29 *P*am80 *rex*::Tn*5*). Strains for assaying SOS induction carry the *dinD1*::Mu dI(Ap^r *lac*) marker (Kenyon & Walker, 1980). SOS induction was monitored in some cases by growing colonies on YT plates supplemented with 35 μg/ml X-gal and scoring for blue color intensity. X-gal plates were used while fresh ($<$ 3 days) as the intensity of blue color produced on this indicator medium decreases as the media ages. For β-galactosidase assays, bacteria were grown in K120 minimal media (Epstein & Kim, 1971) supplemented with 0.2% glucose and 0.4% casamino acids and assayed as described by Miller (1972).

Mu dI(Ap^r *lac*) fusions provide a convenient measure of gene regulation (Casadaban & Cohen, 1979). However, Mu d fusions confer temperature sensitivity, render host strains unstable due to secondary transposition, and undergo zygotic induction and transpose when transduced by phage P1. To circumvent these problems, the *dinD1*::Mu dI(Ap^r *lac*) fusion in GW1040 (Kenyon & Walker, 1980) was made tran-

Table. 2-1 *E. coli* Strains

Strain	Genotype	Comments, construction, or source
Strains obtained elsewhere		
AB2470	*recB21*	CGSC#2470 Howard-Flanders & Theriot, 1966
AR120	N99 cI^+ defective λ lysogen cI^+ Δ*gal nadA*::Tn*10* *lacZ* replaces λ left arm *T11* (prevents *cro* expression) Δ*cII-uvrB*	from A. Schatzman
ER21	*E. coli B su*$^-$ *DNase*$^-$	from L. Silver
GC3217	F$^-$ *tif-1 sfiA11 thr leu pro his arg* ilv^{ts} *gal str*	parent of the GW1000 strain series
GM33	F$^-$ *dam3 sup85(amber)*	from M. Marinus
GM1874	GW1040 *dam4 mutS456*	from M. Marinus (Craig, *et al.*, 1984)
GW1000	GC3217 pro^+ *lac*Δ*(U169)*	from G. Walker (Kenyon & Walker, 1980)
GW1040	GW1000 *dinD1::*Mu dI(Ap^r *lac*)	from G. Walker (Kenyon & Walker, 1980)
GW1060	GW1000 *uvrA215::*Mu dI(Ap^r *lac*)	from G. Walker (Kenyon & Walker, 1980)
HB101	*hsdS20* (r_B^-,m_B^-) *ara-14 proA2 lacY1 galK2 rpsL20* xyl-5 mtl-1 supE44 mrr_B $mcrB_B$ (λ^-)	Boyer & Roulland-Dussoix, *1969*[g]
JC13519	*lexA3 malE*::Tn*10*	
JM101	F' *traD36 proAB*$^+$ *lacI*^{q} *Z*Δ*M15/ supE thi* Δ*(lac-proAB)*	Yanisch-Perron *et al.*, 1985
JM103mutD	F' *traD36 proAB*$^+$ *lacI*^{q}*Z*Δ*M15/ lacpro thi strA supE endA sbcB15*	*J. Makris*
JM109	F' *traD36 proAB*$^+$ *lacI*^{q}*Z*Δ*M15/ recA1 endA1 gyrA96 thi hsdR17*	Yanisch-Perron *et al.*, 1985
K38	*HfrC phoA(Am)* (λ)	Russel & Model, 1983
K91	K38 (λ^-)	K38 cured of λ
K560	K38 Δ*lac*	
K561	K38 $lacI^q$	
K871	K561 *recA56 srl300::*Tn*10*	
K902	K38 *supE recA56 srl300*::Tn*10*	
K1019	K561 *ilvY864*::Tn*10 rep71*	
K1053	BW313 *dut ung thi1 relA spoT1/* F' *lysA*	Kunkel, 1985
KS391	*HfrH lacMS286* Π*80dIIlacBK1 thi*	Konrad, 1977
LE392	*supE supF hsdR*$^-$ *metB trpR lac Y tonA*	
MM294	F$^-$ *endA1 thi1 hsdR17 supE44*	from F. Barany
N1626	*lig4*	Gottesman *et al.*, 1973
N2603	lig^+	Gottesman *et al.*, 1973
N2604	*lig ts7*	Gottesman *at al.*, 1973
N3087	*tyrA*::Tn*10 IN(rrnD-rrnE)1*	CGSC #6662
SP254	F$^-$ *recN262*	Picksley, *et al.*, 1984 CGSC #6665
X2904	*thyA*::Tn*10*	

Strains constructed for this study

JH5	GW1040 Ts^+	GW1040 cured of temperature sensitivity
JH11	HB101 *recA*$^+$	HB101 P1 transduced to *recA*$^+$ and Tet^r from K590[c] and cured of Tet^r.
JH20	K91 *lexA3*	K91 P1 transduced to *lexA3* and Tet^r from JC13519.
JH27	K91 *recA56*	K91 P1 transduced to *recA56* and Tet^r and cured of Tet^r
JH30	K91 *lexA3 malE? lacI*^{q} *lacZ::*Tn*5*	K91 P1 transduced to *lexA3* and *malE::*Tn*10* from JC13519[a], cured of Tet^r, and transduced from GC2438[b].
JH35	JH5 *recA*$^+$ *srl300::*Tn*10*	JH5 P1 transduced to Tet^r and recA$^+$ from K590[c].
JH39	JH35 *srl*$^-$*? tet*^{S}	JH35 cured of Tet^r
JH43	JH39 */F' lacI*^{q} *lacZ::*Tn*5*	JH39 mated with GC2438[b].
JH56	JH39 *mcrB1 hsdR2*	JH39 P1 transduced to Tet^r from ER1351[e] and cured of Tet^r.
JH59	JH39 *recA56 srl*$^-$*?*	JH39 P1 transduced to *srl300::*Tn*10* and *recA56* from JC10240[d] (from A. J. Clark), and then cured of Tet^r.
JH63	JH56 /F' *lacI*^{q} *lacZ*::Tn*5*	JH56 mated with GC2438[b]
JH69	HB101 *recA*$^+$ *srl*$^-$*?*	*JH11 P1 transduced dinD1::*Mu dI(Ap^r *lac*) Ts^+
JH71	JH69 *zjj202*::Tn*10*	JH69 P1 transduced from ER1351[e].
JH74	SP254 *tyrA*::Tn*10*	
JH76	JH39 *mrr2*::Tn*5*	Tn*5* insertion that blocks SOS induction by *Hha*II methylase.
JH82	JH39 *mcrB1 hsdR2 zjj202*::Tn*10*	JH39 P1 transduced to Tet^r from ER1351[e].
JH83	JH39 *mrr2*::Tn*5 zjj202*::Tn*10*	JH76 P1 transduced to Tet^r from ER1351[e].
JH84	JH39 *zjj202*::Tn*10*	as JH83
JH94	AB2470 *thyA748::*Tn*10*	AB2470 P1 transduced to *thyA748::*Tn*10* from X2904 (from B. Bachman); still carries *recB21.*
JH100	K561 *recB21 thyA::*Tn*10*	K91 P1 transduced from JH94.
JH104	JH43 *recB21 thyA::*Tn*10*	JH43 P1 transduced from JH94.
JH107	K91 *lacZ::*Tn*5 lacI*^{q}	K91 P1 transduced from JH30
JH109	JH107 *dinD1::*Mu dI(Ap^r *lac*)	JH107 P1 transduced from
JH117	JH39 *recB21 thyA*::Tn*10*	JH39 P1 transduced from JH94.
JH119	JH76 *recB21 thyA::*Tn*10*	JH76 P1 transduced from JH94.

JH120	ER21 *lacZ*::Tn*5*	*ER21 P1 transduced from dinD1*::Mu dI(Ap^r *lac*) Ts$^+$ and JH5
JH121	JH39 *mcrB4*::Tn*10*	JH76 P1 transduced to Tet^r from ER1467[f]
JH122	JH39 *mcrB4*::Tn*10 hsdR2 mrr2*::Tn*5*	as JH121
JH123	JH39 *mcrB4*::Tn*10 hsdR2*	as JH121
JH124	JH39 *mrr2*::Tn5 *mcrB4*::Tn*10*	as JH121
JH125	AB1157 *rfa-550 lacZ*$^-$ *alkA51*::Mu dI(Ap^r *lac*)	spontaneous *lacZ*$^-$ derivative of MV1571 from M. Volkert.
JH129	JH39 *hsdR2 zjj202*::Tn*10*	JH76 P1 transduced to Tet^r from ER1351[e].
JH133	LE392 *mrr*$_K^+$ *zjj202*::Tn*10*	LE392 P1 transduced from JH71
JH134	LE392 *mrr*$_B$ *zjj202*::Tn*10*	LE392 P1 transduced from JH76
JH135	LE392 *mrr2*::Tn*5 hsdR*$^+$	LE392 P1 transduced from JH76[g]
JH137	K560 *dinD1*::Mu dI(Ap^r *lac*)	K560 P1 transduced to Amp^r from JH39.
JH138	JH137 *rep71 ilvY864*::Tn*10*	JH137 P1 transduced from K1019, selected for Tet^r, and screened for f1^r, f2^s.
JH139	GW1000 *dinD1*::Mu dI1734(Kan^r *lac*)	GW1000 infected with MudI 1734 as described in methods.
JH140	K560 *dinD1*::Mu dI1734(Kan^r *lac*)	K560 P1 transduced to Kan^r from JH139.
JH141	JH140 *rep71 ilvY864*::Tn*10*	JH140 P1 transduced from K1019 as for JH138.
JH144	K91 *recN262 tyrA16*::Tn*10*	K91 P1 transduced from JH74.
JH145	K91 *recB21 thyA*::Tn*10*	K91 P1 transduced from JH94.
JH152	JH69 (λ)	
JH153	MM294 *cI*$^+$ Δ*gal nadA*::Tn*10*	MM294 P1 transduced from AR120 with selection for Tet^r.
JH154	JH39 *lexA3 malE*::Tn*10*	JH39 P1 transduced from JC13519

[a] JC13519 carries *lexA3 malE*::Tn*10* (from A. J. Clark).

[b] Strain GC2438 (from S. Gottesman (Huisman, *et al.*, 1984)) served as the donor, either by P1 transduction or mating, of the *lacI*^{q} *lacZ*::Tn*5* markers carried on its F' episome.

[c] Strain K590 carries the *recA*$^+$ allele linked to a *srl300*::Tn*10* insertion (from M. Russel).

[d] Strain JC10240 carries the *recA56* allele linked to the *srl300*::Tn*10* insertion (from A. J. Clark).

[e] Strain ER1351 carries the *hsdR2* and *mcrB*$^-$ markers linked to a Tn*10* insertion at 99 minutes (*zjj202*::Tn*10*) (from E. Raleigh).

[f] Strain ER1467 (from E. Raleigh) carries the *hsdR2* and *mcrB4*::Tn*10* alleles.

[g] The *mcrB*$_B$ and *mrr*$_B$ alleles have weak or no activity.

sposition defective as follows. The Mu dI(Ap^r *lac*) phage encodes a temperature sensitive mu*C* repressor; at the non-permissive temperature, unrepressed replication and *kil* function are lethal to the host. We selected 42°C survivors of GW1040, and then screened for healthy colonies which still synthesized β-galactosidase in response to mitomycin C. One of these mutants, JH5, contains a defective Mu d fusion which transposes at a greatly reduced frequency at 42°C. Strain JH5 served as the parent in further strain constructions (see Table 2-1).

II. DNA manipulations and plasmid constructions

Plasmids used in this study are described in Table 2-2. Plasmid DNA was prepared by a non-denaturing lysozyme/triton method (Zinder & Boeke, 1982) or by alkaline lysis (Maniatis, *et al.*, 1982). DNA fragments were isolated from agarose gels and electroeluted. Other DNA manipulations were performed as described by Maniatis *et al.* (1982). *In vivo* methylase activity was quantified by measuring the extent to which plasmid-borne restriction sites were protected from cleavage by the cognate restriction endonuclease (or an isoschizomer known to be inhibited by the methylase). *Hha*II methylation was tested by *Hin*fI digestion, *Pae*R7 methylation with *Xho*I digestion, *Msp*I methylation with *Hpa*II digestion, *dam* methylation with *Mbo*I digestion, and *Bsu*RI, *M. Spr*, and *Bsp*RI methylation with *Hae*III digestion. All methylase clones, with the exception of uninduced pJH49 as described in the results, rendered the appropriate plasmid borne restriction sites completely resistant to cleavage.

Several plasmids conferred only ampicillin resistance. To study these plasmids in strains which carried a Mu dI(Ap^r *lac*) fusion, we inserted a different antibiotic resistance gene. Shapira *et al.* (1983) constructed plasmids in which the chloramphenicol or kanamycin resistance genes are flanked by polylinkers to form versatile gene cassettes. We used the chloramphenicol resistance gene cassette derived from pSKS114 to construct pJH1 by inserting the cassette into the *Pst*I site of pTP166 (Marinus, *et al.*, 1984). Similarly pJH51 was derived from pBamHI 2-50 (G. Wilson)

Table 2-2. Plasmids

Plasmid	Genotype/Description	Construction, reference, or source
misc.		
pGL101	*lac*UV5 Amp^r	Guarente *et al.*, 1980
pKC30	$\lambda P_L N^+$	Rosenberg *et al.*, 1983
pSKS101	Kan^r Amp^r	Shapira *et al.* 1983
pSKS114	Cam^r Amp^r	Shapira *et al.* 1983
Adenine Methylases		
pJC1	M.*Eco*RI Cam^r	from S-C. Cheng (Cheng & Modrich, 1983)
pJH1	PtacI-*dam* $Cama^r$	Cam^r cloned in the *Pst*I site in pTP166 (from M. Marinus) Marinus, *et al.*, 1984.
pJH40	*lac*UV5-M.*Hha*II Amp^r	M.*Hha*II cloned in the *Pvu*II site in pGL101.
pJH49	*lac*UV5-M.*Hha*II Cam^r	Cam^r cloned in the *Pst*I site of pJH40.
pJH56	M.*Pst*I Amp^r Cam^r	Cam^r cloned in the *Eco*RI site in pME101 (from R. Walder) Walder, *et al.*,1984.
pPAOM.177	M.*Pae*R7 Kan^r	from J. Brooks
pSK5	M.*Hha*II Tet^r	from H. Smith (Schoner *et al.*, 1983)
Cytosine Methylases		
pES2	M.*Bsp*RI Tet^r Amp^r	from A. Kiss (1983)
pHaeIII 1-1	M.*Hae*III Tet^r Amp^r	from G. Wilson
pJH51	M.*Bam*HI Amp^r Cam^r	Cam^r cloned in the *Sal*I site of pBamHI 2-50 (from G. Wilson).
pJH53	M.*Hha*I Cam^r	Cam^r cloned in the *Pst*I site of pHhaI 2-1 (from G. Wilson).
pMER3	M.*Msp*I Tet^r Amp^r	from R. Walder Walder, *et al.*, 1983.
pR215	M.*Eco*RII Tet^r	from A. Bhagwat
pSU11	M.*Bsu*RI Tet^r Amp^r	from A. Kiss (1983)
pSU21	M.*Spr* Tet^r	from A. Kiss (1983)
*Eco*RI RM Plasmids		
pAN4	*Eco*RI restriction-modification system	Newman *et al.*,1980
pCS497	*Eco*RI R^+ M^{ts} Amp^r	
pJH10	pAN4 R^+ M::Kan^r	Kan^r cassette clone in the *Bcl*I site of pAN4
pJH12	pJH10 (TS0) with *Eco*RI site filled in	
pJH13	*Sal*I fragment deleted from pJH12	
pJH14	TS0 mutation cloned into pAN4 ($R^{ts}M^+$)	
pJH15a	pJH10 with f1 IG region clockwise in *Cla*I site	
pJH15b	pJH10 with f1 IG region counter-clockwise in *Cla*I site	

pJH16a/b	pJC1 with f1 IG insertion at the *Cla*I site
pJH17	pJH15b with a *Bam*HI linker at the *Pvu*I site 5' to the endonuclease gene
pJH19	pJH17 with a *Xho*I linker at the *Sma*I site 3' to the endonuclease gene
pJH20	Δ of *Bam*HI to *Bgl*II segment of pJC1
pJH23	*Sal*I *Eco*RI endonuclease cassette in pUC7
pJH25	*Pvu*II *Eco*RI endonuclease cassette in the *Sma*I site of pUC18
pJH34	pCS497 with Kan^r cassette at *Pst*I site
pJH70	*Bam*HI to *Hin*dIII deletion of pJH19
pJH71	*Bam*HI to *Bgl*II deletion of pJH19
pJH72	*Bam*HI to *Pst*I deletion of pJH19
pJH74	*Bgl*II fragment of pJH15b deleted
pJH75	TS5 mutation in pJH15a
pJH76a/b	pAN4 with f1 IG insertion at the *Cla*I site
pJH80a/b	TS6 mutation in pJH76a/b
pJH81	TS6 mutation in pJH15a
pJH82	pAN4 *Bgl*II site filled in (R^- M^+)
pJH83	R200P mutation in pAN4
pJH84	R200S mutation in pAN4
pJH85	R200V mutation in pAN4
pJH86	R200K mutation in pAN4
pJH87	R200C mutation in pAN4
pJH89	pAN4 with a *Bgl*II-*Hpa*II fragment carrying λ P_L inserted at the *Nde*I site (filled in) just 5' to the *Eco*RI endonuclease gene that also destroys the *Bgl*II site on the end of the P_L bearing fragment
pJH98	R200P mutation in pJH89
pJH99	R200S mutation in pJH89
pJH100	R200V mutation in pJH89
pJH101	R200K mutation in pJH89
pJH102	R200C mutation in pJH89
pJH104	mutations R56Q, H114Y, R200K in pJH15b
pJH105	mutations R56Q, A138V, R200K in pJH15b
pJH115	mutations R56Q, A138T, R200K in pJH15b
pJH116	mutations R56Q, R200K in pJH15b
pJH113	mutations R56Q, V166I in pJH15b
pJH114	mutations R56Q, A138T in pJH15b
pJH117	mutations R56Q, R200C in pJH15b
pJH118	mutations R56Q, A138V, R200C in pJH15b
pJH119	mutations R56Q, H114Y, R200C in pJH15b
pJH120	mutations R56Q, A138T, R200C in pJH15b
pJH121	Δ of the *Bam*HI to *Pst*I fragment of pJC1
pJH122	pBR322 with the *Eco*RI site filled in
pJH123	pCS497 with a Cam^r cassette inserted at the *Bam*HI site
pJH124	pJH15b (RA2 mutation) *Eco*RI site destroyed
pJH125	pJH15b (TS6 mutation) *Eco*RI site destroyed

by inserting the cassette at the *Sal*I site, pJH53 from pHhaI 2-1 (G. Wilson) by inserting at the *Pst*I site, and pJH56 from pME101 (Walder, *et al.*, 1984) by inserting at the *Eco*RI site. pJH34 and pJH123 bear Kan^r (at the *PstI site) and* Cam^r *(at the* *Bam*HI site) resistance gene insertions in pCS497.

The *Hha*II methylase gene was cloned under control of the inducible *lac*UV5 promoter in plasmid pGL101 (Guarente, *et al.*, 1980) as follows. The 960 base pair *Eco*RI fragment from pSK5 (Schoner, *et al.*, 1983) carrying the *Hha*II methylase gene was isolated, filled in with the Klenow fragment of DNA polymerase I, and blunt-end ligated into pGL101 which had been digested with *Pvu*II and treated with calf intestinal phosphatase. The antibiotic resistance of the resulting plasmid (pJH40) was changed from ampicillin to chloramphenicol by inserting the chloramphenical resistance gene cassette from pSKS114 into the *Pst*I site of pJH40, yielding pJH49.

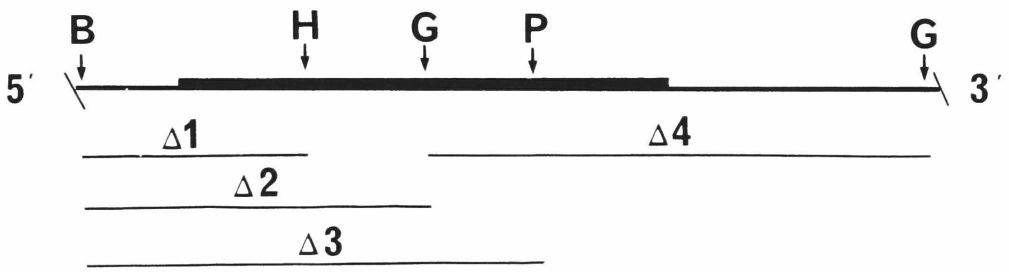
Plasmid pAN4 is a pBR322 derivative which encodes the *Eco*RI restriction-modification system and confers ampicillin resistance (Newman, *et al.*, 1981). The *Eco*RI methylase gene of pAN4 was inactivated by inserting a kanamycin resistance gene (*Bam*HI fragment) (Shapira, *et al.*, 1983) into the *Bcl*I site within the methylase gene (see Fig. 4-1) to yield plasmid pJH10. Plasmid pJC1 is a pACYC184 derivative which encodes the *Eco*RI methylase and chloramphenicol resistance (Cheng & Modrich, 1983). Plasmid pJH15a/b was constructed by ligating *Cla*I linkers onto a *Bgl*II fragment encompassing the f1 intergenic region (nucleotides 5486-5941, kindly provided by N. Davis) and inserting this fragment into the *Cla*I site of plasmid pJH10 (see Fig. 4-1). Plasmids pJH15a and b differ only in the orientation of the inserted f1 intergenic region. Replication from the f1 +strand origin is clockwise for pJH15a and counter-clockwise for pJH15b. pJH34 (R^+M^{ts}) was derived from pCS497 (carries a R^+M^{ts} version of the *Eco*RI restriction system, constructed by S. Chang and S. Cohen and provided by M. Casadaban) by inserting a *Bam*HI fragment carrying the kanamycin resistance gene (Shapira, *et al.*, 1983). Since plasmids pJH10, pJH15a, and pJH15b are pBR322 derivatives, they each carry one *Eco*RI site. In some cases

(TS0ΔRI (pJH12), TS6ΔRI (pJH125), and RA2ΔRI (pJH124)) this site was destroyed by treating plasmid DNA with *Eco*RI endonuclease, the Klenow fragment of the DNA polymerase I and deoxyribonucleotides to render the ends flush, and DNA ligase.

To construct the set of deletions of the *Eco*RI endonuclease gene shown in figure 2-1, we inserted a *Bam*HI linker into the *Pvu*I site upstream of the *Eco*RI endonuclease gene and a *Xho*I linker downstream at the *Sma*I site of plasmid pJH15b. The resulting plasmid (pJH19) was partially cleaved with *Hin*dIII, *Bgl*II, or *Pst*I (sites internal to the *Eco*RI gene), ligated to *Bam*HI linkers, digested with *Bam*HI and religated. To select for the desired deletion and against deletions with endpoints at sites other than within the endonuclease gene, we selected deleted plasmids which were viable in the absence of the methylase plasmid (endo$^-$) but still conferred resistance to both kanamycin and ampicillin. Screening a number of such colonies yielded the deletion plasmids pJH70 (deletes N-terminal one-quarter of the endonuclease gene to *Hin*dIII site), pJH71 (deletes N-terminal one-half to the *Bgl*II site), and pJH72 (deletes N-terminal three-quarters to the *Pst*I site). An additional pJH15b deletion (pJH74, Amp^r Kan^s) was constructed which spans from the *Bgl*II site within the endonuclease gene to the *Bgl*II site in the kanamycin resistance gene and thus removes the the C-terminal half of the endonuclease gene.

The *Eco*RI endonuclease gene of plasmid pJH10 (Heitman *et al.*, 1989) was excised as a 1.9 kb *Pvu*II fragment, and ligated into the *Sma*I site of plasmid pUC18 to yield plasmid pJH25, or alternatively ligated to *Sal*I linkers and ligated into the *Sal*I site of pUC7 to yield plasmid pJH23. The resulting plasmids carry the endonuclease gene flanked by convenient restriction sites. The endonuclease gene was excised as a *Bam*HI-*Sal*I cassette from pJH25 and ligated into DNA of the M13mp9 and J1 (=MP9 with the *Eco*RI site filled in) phages that had been linearized with *Bam*HI and *Sal*I and treated with calf intestinal phosphatase (CIP). Strain JM101/pJC1 (*lacI*^{q} *Eco*M$^+$ *supE*) was transfected with the resulting DNA and plated in top agar containing X-Gal (40 μg/ml and IPTG at 1 mM) on solid medium containing X-gal

Figure 2-1. Deletions of the *Eco*RI Endonuclease Gene. As described in the text, Δ1 (pJH70), Δ2 (pJH71), and Δ3 (pJH72) bear a *Bam*HI linker at the deletion endpoint while Δ4 retains its *Bgl*II site. These unique restriction sites were used to linearize the deletion plasmids for heteroduplex deletion mapping.


B
H
G
P
G
5′
3′
Δ1
Δ4
Δ2
Δ3

and IPTG. Phage that made white plaques were isolated. Phage J3-2 is the *Eco*RI endonuclease gene cloned in phage M13mp9, and J2-2 is cloned in phage J1.

Alternatively, the *Eco*RI endonuclease cassette was excised from plasmid pJH23 with *Bam*HI and ligated into the *Bam*HI site of M13mp18 phage DNA linearized with *Bam*HI and treated with CIP. Phage J9-9 bears the endonuclease gene under control of the *lac* promoter of phage M13mp18 while phage J9-8 carries the insert in the opposite orientation. The *Eco*RI site of phage J9-9 was destroyed by sequential treatement of phage RFI form DNA with *Eco*RI, the Klenow fragment of DNA polymerase I + dNTPs, and T4 DNA ligase to yield phage J11.

III. Random mutagenesis and mutant screens

Cells bearing plasmids pJH10 (or pJH15b) (R^+ M^-) and pJC1 (R^- M^+) were mutagenized either spontaneously or with nitrosoguanidine as described (Miller, 1972). The endonuclease and methylase plasmids were separated by treating a plasmid DNA mixture with *Bam*HI to linearize pJC1 but spare the endonuclease bearing plasmid. Since linear DNA transforms *E. coli* poorly, cells transformed by this DNA often receive pJH10 without pJC1. In practice, about 10% of the surviving colonies receive a mutant derivative of pJH10 while the remaining 90% are co-transformed by both plasmids.

Alternatively, the plasmids were separated by selectively packaging the endonuclease plasmid (pJH15b) with f1 helper phage. After cells were transduced with these lysates, we also observed a background of transductants harboring both plasmids. Since plasmids pJC1 and pJH10 (or pJH15a/b) carry regions of homology, we thought this might be attributable to packaging of interplasmid recombinants. This can not account for the entire background of pJC1 transductants, since we found that plasmid pJC1 could be packaged by f1 helper phage ($\sim 10^5$ TPs/ml compared to $\sim 10^{10}$ TPs/ml for pJH15b) from cells bearing only pJC1, even though this plasmid does not carry an f1 packaging signal and is not known to exist in a single-stranded

DNA form. In any event, following mutagenesis or using the host strain JH4 and helper phage R189 (described below), greater than 90% of transduced colonies received a plasmid carrying a mutant endonuclease.

To produce a mutagenized transducing particle lysate of pJH15b, cells harboring both pJH15b and pJC1 were mutagenized and then grown for one hour, diluted, infected with f1 helper phage, and grown overnight in soft agar. Plasmid transducing particles were separated from f1 helper phage by transduction of a *rep*$^-$ host, JH138, which supports plasmid but not phage replication. Alternatively, we employed an amber suppressing host (JH4) and a helper phage carrying a gene II non-sense mutation (R189), and subsequently transduced a non-suppressing host (K91) which does not allow growth of the mutant helper phage.

Temperature-sensitive (TS) endonuclease mutants were isolated as follows. The endonuclease bearing plasmid was mutagenized and separated from the methylase plasmid (as described above) and then introduced into cells lacking the methylase. Cells which acquire a wild-type copy of the endonuclease suffer DNA degradation and die. Survivors carry endonuclease mutants with reduced activity. By initially plating these survivors at 42°C and subsequently replica-plating to growth at 42°C and 30°C, we found that many (~ 30%) of the surviving colonies were cold-sensitive for growth. These cold-sensitive colonies harbor temperature-sensitive endonuclease mutants. In some cases we employed an SOS::*lacZ* host strain since many TS mutant endonucleases show residual activity and induce the SOS response at 42°C. The reduced activity endonuclease mutant RA2 was isolated by screening for a mutant that was viable in strain JH137 without the *Eco*RI methylase and which induced the SOS response at all temperatures.

*Eco*RI endonuclease mutants with enhanced *Eco*RI* activity were isolated by mutagenizing plasmid pJH15b carrying the TS6 allele *in vivo* with nitrosoguanidine and introducing the purified plasmid DNA into an SOS::*lacZ* fusion strain expressing the *Eco*RI methylase. We then screened for colonies which made blue colonies on X-gal

indicator media which were stable and TS for the blue colony phenotype.

Site-directed mutagenesis is described below.

IV. Mapping and sequence analysis

Mutations were mapped by either of three methods. Restriction fragments from wild-type and mutant plasmids were ligated and the ability of the hybrid to restrict λ phage at the non-permissive temperature (42°C) was determined after transformation. For star *Eco*RI mutants we scored induction of the SOS response. Alternatively, single-stranded mutant DNA was annealed to double-stranded restriction fragments from wild-type. After transformation, the frequency of marker rescue to wild-type (R^+) was determined by replica-plating to plates seeded with ~ 10^6 λ_{vir} per plate or by cross-streaking individual transformants against λ_{vir} that had been spread as a line on the plate with a 0.2 ml pipette and allowed to dry. The frequency of marker rescue varied from ~ 0.5 to ~ 10 %. Finally, in some cases we employed heteroduplex deletion mapping (Shortle, 1983) with a set of *Eco*RI deletions. For this method *Bam*HI linearized DNA of the *Eco*RI endonuclease deletion plasmids pJH70, pJH71, or pJH72, or *Bgl*II treated pJH74 was annealed to *Cla*I treated mutant DNA, and the transformants which result from this DNA were scored by cross-streaking against λ (for TS mutants) or plating on X-gal indicator media (for star mutants).

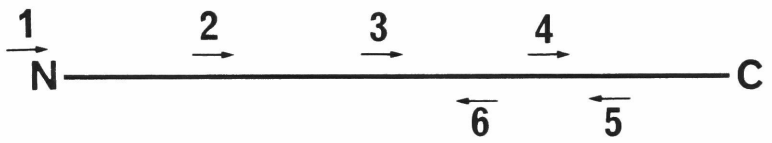
The implicated region for each mutation was sequenced by the dideoxy nucleotide method (Sanger, *et al.*, 1977). In several cases (TS6, TS9, RA2, R200K, R145K, R145K + Y193H, and all the *Eco*RI star mutants) the entire gene was sequenced. Sequencing templates were either single-stranded plasmid DNA isolated after helper phage infection or denatured (collapsed) double-stranded plasmid DNA. In one case (TS0) the relevant region was sub-cloned into M13mp18 to produce single-stranded template DNA for sequencing. Sequencing primers are listed with other oligonucleotides in table 2-3 and shown in figure 2-2. We employed six primers: four which were equally spaced to cover the coding strand (pJH15b packaged strand) and two for

Table 2-3. Oligonucleotides

Name	Sequence	Length	Use
JOHE1	AAATAACAGTGGAAACATGG	20	R1 primer 1
JOHE2	GTCGATCTAAXNNATTTAATATACC	25	R200X mutagen
JOHE3	GATAGTATAAAGAAAACAG	19	R1 primer 2 (replaced by 17)
JOHE4	GGGAAAAGAGGAGATCAAG	19	R1 primer 3
JOHE5	GGTTAGATCGACTAACTGC	19	R1 primer 4
JOHE6	GATTACTATTTATAGGC	17	R1 primer for R200
JOHE7	GATCTCGA	8	BglII or BamHI to XhoI adapter
JOHE8	GTCGATCTAACCTATTTAATATACC	25	R200R
JOHE9	CTTATGAGATCTXNNGATAGCATTACC	27	E144X mutagen
JOHE10	CTTATGAGAXNNTTCGATAGC	21	R145X mutagen
JOHE11	GTCGATCTAACGCATTTAATATACC	25	R200A mutagen
JOHE12	GTCGATCTAACTXATTTAATATACC	25	R200EQ mutagen
JOHE13	GTCGATCTAAATXATTTAATATACC	25	R200HD mutagen
JOHE14	GTCGATCTAAAOOATTTAATATACC	25	R200FYIN mutagen
JOHE15	GATTTAATGGTTGCTGG	17	A138V
(JOHE16	GTGATATTTTTTATGTTAACAAGG	24	HpaII site in IG)
JOHE17	GATAGTATAAAGAAAACAG	19	R1 primer 2 again
JOHE18	GTGGCTCTCAGAGAGC	16	R1 primer for E144,R145
JOHE19	CTTATGAGATCTTTCGATAGCATTACC	27	E144,R145 wild-type
JOHE20	GTCGATCTAACTTATTTAATATACC	25	R200K mutagen
JOHE21	GCTGCAGTTAGTTTATCTAA	20	R203K
JOHE22	GGTCTTGTTX1TGAGATATTTTCTG	25	I181VALS mutagen
JOHE23	CCTATTTAATX1ACCAGAATTATACTC	27	I197VALS mutagen
JOHE24	GTTGTTCAGTTX2CCTATTTGACTG	25	L10ITVA mutagen
JOHE25	GCATTACCAGCA2NCATTAAATCTTG	26	A138NDGSRHYC mutagen
JOHE26	GGAGTTTGGCTCCAGCAACAAG	22	R1 primer 7
JOHE27	TCGCGAATTCGCG	13	R1 site binding oligo

N=ATGC; X=GC; O=AT; 1=AC; 2=TC

Figure 2-2. Sequencing Primers for the *Eco*RI Endonuclease Gene. Short (16 - 20 bp) oligonucleotides complementary to regions of the endonuclease gene were synthesized on a fully automated Applied Biosystems 380A DNA synthesizer and used as sequencing primers without purification. Primers 1 (JOHE 1), 2 (JOHE3 or 17), 3 (JOHE4), and 4 (JOHE5) anneal to the non-coding strand (pJH15b packaged strand) and prime at nucleotide positions 338, 535, 745, and 959 respectively (Newman, *et al.*, 1981). Primers 5 (JOHE6) and 6 (JOHE18) anneal to the coding strand (pJH15a packaged strand) and prime at nucleotides 973 and 811.


1
2
3
4
N
C
6
5

the non-coding strand (pJH15a direction) which prime just before the sequences encoding amino acids R200 or E144/R145.

For sequencing double-stranded DNA, 1-2 μg of plasmid DNA prepared by the non-denaturing triton X-100/lysozyme lysis protocol of Zinder and Boeke (1982) was electrophoresed on a 0.6 % agarose gel, and the RFI band was excised under long wavelength UV irradiation and electo-eluted from the gel slice with an IBI electroelution apparatus. Although in some cases RFI form DNA is poorly recovered by electroelution (~ 10 % yield), this procedure removes RNA and nicked and linear plasmid DNA species which otherwise yield an unacceptable background to sequencing reactions. After ethanol precipitation, the DNA was base denatured with 20 μl of 0.2N NaOH for 10 min. at RT, neutralized by the addition of 20 μl distilled H_2O, 2 μl 2M NH_4AC, and 3 μl of 3M NaAc (pH=5.1), and ethanol precipitated. Typically, this amount of template was sufficient for 1-2 sequencing reactions. Sequencing gels were fixed with 5 % acetic acid and 5 % methanol for 15 min. at RT, dried on Whatman 3M filter paper, and exposed to Kodak XAR film for 3 to 6 days at RT. All nitrosoguanidine (NG) induced mutations were attributable to the known mutagenic pattern of NG (G to A transitions usually within GG or GA sequences) (Burns, *et al.*, 1987).

V. Site-directed mutagenesis

Uracil containing single-stranded pJH15a DNA, produced from the *dut*$^-$ *ung*$^-$ host K1053, was mutagenized as described by Kunkel (1985). Because we found that uracil containing templates lose activity upon prolonged storage, they were used within a few days after isolation. Reaction conditions were as described (Russel & Model, 1986) using a series of degenerate oligonucleotides (listed in table 2-3) in which the nucleotides of the codon for the amino acid we wished to mutate were replaced by a mixture of nucleotides: 5'-NNX-3', where N represents an equimolar amounts of all four nucleotides and X is one-half each of G and C. These oligonucleotide mixtures

are 32-fold degenerate and contain at least one codon for each of the 20 amino acids. Several additional oligonucleotides which were less degenerate and encoded specific desired changes (table 2-3) were used to isolate the remaining substitutions at the R200 position. Separate annealing reactions were carried out at a series of different temperatures (RT, 45 °C, and 65°C) (Goff, *et al.*, 1987). The reaction mixture was then introduced into a *dut*$^+$ *ung*$^+$ host (JH137) which destroys and thereby selects against both the parental template strand and the helper phage (which remains single-stranded since it is not complementary to the mutagenic primers).

After several rounds of purification by colony isolation, single-stranded DNA was isolated following infection with f1 helper phage R176. Dideoxy nucleotide DNA sequencing of random isolates revealed that in most cases, ~ 75% of colonies harbored a mutant endonuclease gene. While mutations resulting from single-mismatches predominated when oligonucleotides had been annealed at 65°C, with lower annealing temperatures a greater proportion arose from double and triple mismatches. In general, ~ 100 nucleotides spanning the mutated site were sequenced. In most cases mutations were isolated more than once independently. With only one exception, independent isolates and mutants encoding the same amino acid with different codons showed identical phenotypes. In the one case (R145K) where two isolates showed different phenotypes, sequencing the entire gene revealed that the more active allele carried a spontaneous suppressor mutation (R145K + Y193H). In all other cases to ensure the identified single nucleotide changes confer the observed phenotype, we subjected uracil containing mutant templates to site-directed reversion with an oligonucleotide bearing the wild-type codon at the mutated position. With high frequency (25-75%) every mutation regained the ability to restrict λ phage to the same extent as the wild-type enzyme. Thus we conclude that the observed phenotypes are attributable to single point mutations.

VI. Cellular extracts and protein purification

Cell extracts containing the *Eco*RI TS endonuclease were isolated by resuspending cell paste from 100 ml cultures (grown at 30°C) of strain K91/TS and pJC1 in a lysis buffer (100 mM NaCl / 50 mM Tris-HCl (pH= 7.4) / 10 mM $MgSO_4$ / bovine serum albumin at 1 mg/ml / 1 mM phenylmethylsulfonylfluoride) permissive for *Eco*RI activity and sonicating (6 X 20 sec) on ice. (Heitman, *et al.*, 1989). After centrifugation for 20 minutes in a microcentrifuge and 3 hr at 100,000 X g, the supernatant was mixed 1:1 with glycerol and stored at -20°C.

We initially tried to purify the site-directed *Eco*RI endonuclease mutants (chapter 6) expressed from the double plasmid system (pJH15a and pJC1) in strain JH137. These mutants have very weak endonuclease activity which proved extremely difficult to detect in the presence of cellular endonucleases. Instead, we over-expressed several mutants by first rejoining the mutant endonuclease gene and the wild-type methylase gene in plasmid pAN4. A fragment carrying the endonuclease mutation was then sub-cloned into a vector that carries both the the endonuclease and the methylase gene under control of the strong λ P_L promoter (pJH89). After thermal induction in a *cI857ts* host strain (K1140), these plasmids expressed the endonuclease and methylase enzymes to roughly 10% (5% each) of the total cellular protein. However, since these site-directed mutants are temperature-sensitive *in vivo*, this did not seem a prudent means of inducing expression. λ can also be induced by chemical treatments such as nalidixic acid (Mott, *et al.*, 1985). However, the *cI857* repressor is *ind*$^-$ and cannot be chemically induced. We obtained the *cI*$^+$ defective λ lysogen host strain AR120 (Rosenberg, *et al.*, 1983) and moved the defective λ lysogen into an *endA*$^-$ host (MM294) by P1 transduction with selection for a linked marker (*nadA*::Tn*10*). The resulting strain (JH153) effectively represses expression from λ P_L promoter, is chemically inducible, and makes little or no endogenous endonuclease. With this over-expression system, the R200K and the R200C mutant proteins were purified by measuring *Eco*RI specific endonuclease activity in column fractions. No endonuclease activity was detectable in the crude lysates of strains expressing these mutants, prob-

ably because the methylase is also present. However, after the first column a clear peak of endonuclease activity was observed. The wild-type and the R200K and R200C mutant enzymes were purified to homogeneity by phosphocellulose and hydroxylapatite chromatography, essentially as described (Cheng, *et al.*, 1984) except that the ammonium sulfate concentration step was usually omitted.

The wild-type and the TS6 enzymes and the star mutant enzymes described in chapter 7 were purified without over-expression from cells bearing both pJC1 and the mutant pJH15b plasmids.

VII. Enzyme assays and cleavage site mapping

A kinetic analysis of DNA scission by the wild-type enzyme and the R200K mutant enzyme was performed. Substrate concentration was varied from 0.56 to 11.8 nm with a constant amount of endonuclease (111 pM for the R200K mutant and 12 pM for the wild-type enzyme). Reactions were conducted at 30°C (because the R200K protein is at least TS *in vivo*) and portions were removed at time points up to one hour. These partial reaction products and linear DNA standards were displayed on 0.6% agarose gels containing 0.5 % EtBr. Gels were photographed under UV trans-illumination with Polaroid 665 positive-negative film. The negatives were scanned with a Joyce-Loebl microdensitometer and the amount of product determined by comparison to linear standards of known concentration determined by spectrophotometric absorption at 260 nm. The amount of product was plotted versus time and the initial reaction velocity determined for the linear portion of the curve. The reciprocal of the initial velocity was then plotted against the reciprocal of the substrate concentration to yield the kinetic constants K_m and V_{max}, which are its intercepts with the Y and X-axes respectively.

Two buffers were employed for restriction digests with the wild-type or mutant *Eco*RI endonucleases. Normal or preferred *Eco*RI buffer contains: 100 mM Tris-HCl at pH=7.5, 5 mM $MgCl_2$, 50 mM NaCl, and 100 μg/ml BSA. *Eco*RI* buffer condi-

tions contain: 25 mM Tris-HCl, pH=8.5, 2 mM $MgCl_2$, and 5% glycerol (from the enzyme storage buffer). To map the sites at which the *Eco*RI star mutants cleave f1, we first purified shorter restriction fragments of f1 (*Hha*I, *Hpa*II, *Hin*fI, or *Hae*III), by electrophoresis in 2% agarose gels followed by electoelution. These isolated fragments were subjected to scission by the mutant endonucleases and the sizes of the resulting fragments were calculated by comparing their migrational distance to the plot of $\log_{10}$(basepairs) versus distance migrated (in mm) for fragments of known length. By a number of such digests, each site of cleavage could be located to a 10-20 bp interval. Within each of these intervals was a consensus *Eco*RI* site based on the f1 sequence.

To determine if these *Eco*RI* sites are in fact the sites of scission, uniquely ^{32}P end-labelled f1 restriction fragments were subjected to Maxam and Gilbert sequencing reactions (Maxam & Gilbert, 1980; Eckert, 1987) and to cleavage by the wild-type and mutant enzymes. Approximately 0.5 ng (1000 - 2000 cpm) of a ^{32}P end-labelled restriction fragment (see below) and 10 - 20 ng of purified enzyme were incubated for 1 hour at 30°C. The reactions were terminated by incubation at 65°C for 15 minutes, followed by the addition of sodium acetate (to 0.3 M) and 10 μg carrier tRNA. The reactions were divided and ethanol precipitated. One-half was resuspended in 6 μl of 95% formamide, 0.1% xylene cyanol, and 0.3% bromophenol blue. These samples were displayed on 6% - 8M urea sequencing gels. The other half of the reactions were resuspended in 10 μl of 10% sucrose, 25 mM EDTA, 0.05% xylene cyanol, and 0.05% bromophenol blue. These samples were electrophoresed under non-denaturing conditions: 8 - 12% polyacryamide gels containing 1 X TBE buffer (Maniatis, *et al.*, 1982). The gels were dried to Whatman 3 mm paper and exposed to Kodak XAR film at -70°C with intensifying screens.

VIII. End-labelling f1 Restriction Fragments.

To obtain short f1 restriction fragments labelled at only one end, we treated 2.5 -

5 μg of CsCl purified f1 DNA with an excess of the first enzyme (100 - 200 units) for 2 - 4 hours at 37°C and then with 10 units of calf intestinal phosphatase (CIP) for an additional 1 hour in the same restriction enzyme buffer. The reaction mixture was incubated at 65°C and then extracted with chloroform and phenol and ethanol precipitated. The resulting restriction fragments were end-labelled with γ-^{32}P-ATP and polynucleotide kinase (5 units) with incubation at 37°C for one hour. Following labelling the unincorporated label was removed by centrifuging the kinase reactions though a short G65 spin column. The kinased fragments were chloroform and phenol extracted and ethanol precipitated. To yield fragments with only one labelled end, we treated with a second restriction enzyme and then separated the fragments on 2% agarose gels. Fragments of the desired sizes were excised under long-wave illumination, electro-eluted, and ethanol precipitated. For the 291 bp *Hin*fI - *Hae*III fragment, we first purified a mixture of the 345, 348, and 354 bp *Hin*fI fragments and subsequently treated with *Hae*III to unveil the 291 bp fragment. By this means we obtained the fragments listed later in table 7-2.

IX. Recombination Assay

To measure the frequency of recombination, we employed strain KS391 which bears two copies of the *lac* operon containing non-overlapping deletions (Konrad, 1977). This strain is *lac*$^-$ but can revert to *lac*$^+$ by recombination between the *lac* deletions. When plated on lactose-tetrazolium indicator plates (Miller, 1972), red colonies are formed which give rise to faster-growing white *lac*$^+$ papillae. The number of papillae is an indicator of the frequency of recombination. To avoid the bias of differing colony sizes (larger colonies yield more papillae), in some cases small discs of growth were produced by touching the tips of a multi-pronged applicator to bacterial cultures and then to the surface of an agar plate.

Chapter 3

The Methylation Dependent Restriction Systems Mrr and McrB Sponsor SOS Induction by Site-Specific Methylases

Introduction

This chapter is adapted from Heitman and Model (1987).

The DNA from many organisms is chemically modified by methylation. This post-replicational modification is the work of site-specific methylases, enzymes that transfer methyl groups from *S*-adenosyl methionine to either adenine (N^6-position) or cytosine (C-5 or N^4-position (Janulaitis, *et al.*, 1983)) at specific recognition sites (Razin, *et al.*, 1984). In general, DNA methylation supplements the information contained within the primary nucleotide sequence. In higher eukaryotes, DNA methylation may help regulate gene expression since cytosine methylation at CG dinucleotides is rare in expressed regions of the genome and prominant in inactive regions (Doerfler, 1983; Cedar, 1984, 1988). For unicellular eukaryotes both cytosine and adenine methylation have been described but their role is not known (Harrison, *et al.*, 1986). Prokaryotes also have both adenine and cytosine methylases. For example, the bacterium *Escherichia coli* makes several methylases including: the *dam* adenine methylase which participates in DNA mis-match repair, and the regulation of transcription, replication, and transposition; the *dcm* cytosine methylase whose function is inknown; and the HsdM, *Eco*RI, and *Eco*RII methylases which are part of restriction-modification systems (Marinus, 1984, 1987; Claverys & Lacks, 1986; Messer & Noyer-Wiedner, 1988).

Many species of bacteria make restriction and modification enzymes to destroy foreign DNA that enters the cell (Arber, 1974; Roberts, 1976). In general, two enzymes compose a restriction-modification system: an endonuclease or restriction en-

zyme that cleaves a specific DNA sequence, and a methylase that modifies the same sequence to block cleavage by the restriction enzyme and thereby protect the cellular DNA. Restriction-modification systems from various bacterial species have been cloned in *E. coli* (reviewed in Wilson, 1988). We have studied the *Hha*II system of *Haemophilus haemolyticus* (cloned by Mann, *et al.*, 1978), and found that *Hha*II methylase expression drastically inhibits growth of several standard *E. coli* K-12 lab strains. Here we show that this results from insults to the DNA which elicit the SOS DNA repair response. We find that, in general, methylases foreign to *E. coli* induce the SOS response. Recognition of adenine methylated DNA requires the product of a hitherto undescribed locus, which we name *mrr*, for *m*ethylated adenine *r*ecognition and *r*estriction. The *mrr* locus encodes a restriction system whose target is methylated DNA. We suggest that the *mrr* product is an endonuclease that cleaves adenine methylated DNA, and that DNA scission induces SOS. Cytosine methylation also induces the SOS response and this effect requires the *mcrB* locus which encodes an endonuclease specific for DNA containing 5-hydroxymethyl- (Revel, 1983), 5-methylcytosine (Raleigh & Wilson, 1986), or N^4-methylcytosine (Blumenthal, *et al.*, 1985). Three other examples of methylation-dependent restriction systems have been described (Lacks & Greenberg, 1977; Sladek, *et al.*, 1986; MacNeil, 1988).

So far, *E. coli* contains five restriction systems (*hsdRMS, mrr, mcrA, mcrB, mcrC*), four of which (*mcrB mcrC hsdRMS mrr*) map in a gene cluster at 98.5 minutes on the chromosome which we call the Ellis Island of *E. coli* to reflect its role in monitoring DNA immigration. These genes may be clustered to coordinate the cellular defense system in some way. Two well studied subspecies of *E. coli*, K-12 and B, differ in at least three of these four loci (*mcrB, hsdRMS,* and *mrr*). Mutations of this sort, which hinder or promote genetic exchange, may be common bacterial speciation events.

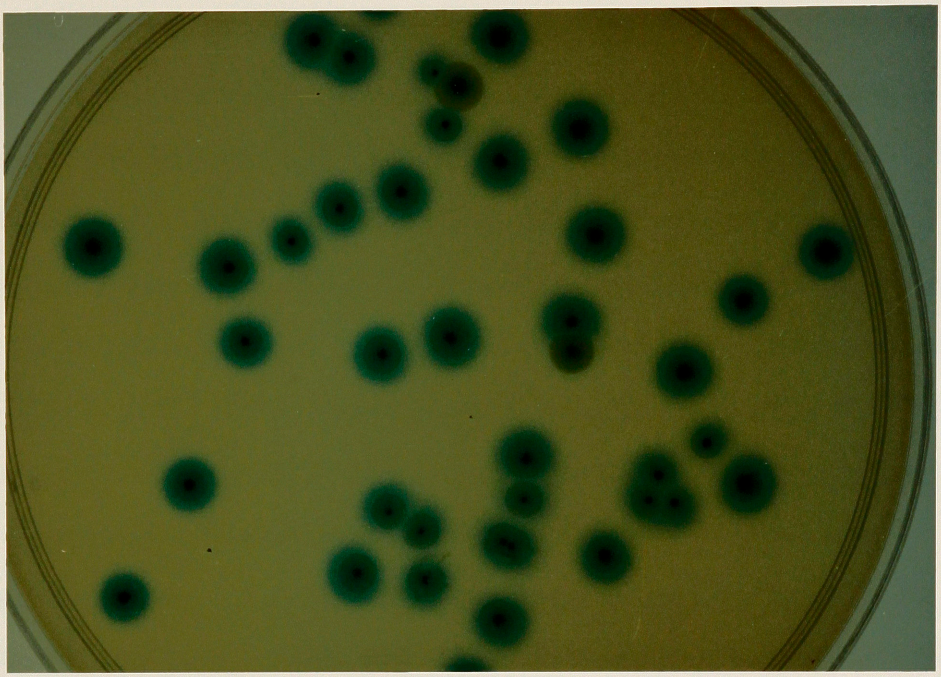
Results

HhaII Methylase Induces λ Lysogens

The *Hha*II restriction endonuclease and methylase are made by the bacterium *Haemophilus haemolyticus* and recognize the sequence GANTC (where N=G, A, T, or C)(Mann, *et al.*, 1978). Scission occurs between the guanine and the adenine, while methylation is at the N^6-position of the adenine. The *Hha*II restriction and modification genes have been cloned in *E. coli* (Mann, *et al.*, 1978). Subsequently, the two genes were cloned on separate compatible plasmids (Schoner, *et al.*, 1983). We obtained these plasmids with the aim of conducting a mutational analysis of the restriction and modification enzymes. However, when several standard laboratory strains (K38, JM101, JM109) were transformed with the plasmid pSK5 which expresses only the *Hha*II methylase ($G^{me}ANTC$), the resulting colonies were flat, translucent, mottled, and grew poorly, giving rise to faster-growing mutants. As shown in figure 3-1, when these transformants were plated on a medium containing X-gal, a chromogenic substrate which yields a blue dye when cleaved by β-galactosidase, and IPTG, a gratuitous inducer of the *lac* operon, the colonies of K38/pSK5 were ringed by blue haloes. Normally lac^+ strains do not leak β-galactosidase or the X-gal cleavage product and thus yield blue colonies with distinct borders on X-gal/IPTG plates. Therefore, the blue haloes induced by the *Hha*II methylase were indicative of cell lysis.

As strain K38 is a λ lysogen, we thought that this cell lysis might be due to λ induction. Culture supernatants of strain K38 expressing the *Hha*II methylase contained 10-fold more phage than did strain K38 alone. Furthermore, when the *Hha*II methylase plasmid pSK5 was introduced into strain K91, an isogenic derivative of K38 cured of λ, the resulting colonies were no longer ringed by blue haloes. However, these colonies still grew slowly and were translucent and mottled. Thus, besides causing cell lysis through induction of λ in a lysogenic strain, *Hha*II methylase exerts

Figure 3-1. *Hha*II Methylase Causes Cell Lysis. Colonies of strain K38/pSK5 were grown on X-gal (30 μg/ml) IPTG (1 mM) indicator plates containing tetracycline (25 μg /ml). The cellular enzyme β-galactosidase cleaves the chromogenic substrate X-gal to produce an insoluble blue dye. Usually neither β-galactosidase nor the blue dye diffuse beyond the colony borders. The blue haloes observed here are therefore indicative of cell lysis. The two colonies which lack blue haloes are spontaneous mutants which have been cured of λ or which carry *Hha*II or *mrr* (described below) mutations.



a further deleterious effect on the cell.

HhaII Methylase Induces the SOS Response

λ lysogens can be induced by DNA-damaging agents, such as mitomycin C or UV light, which induce the SOS response. Induction of λ and the SOS genes results from RecA stimulated auto-digestion of their respective repressors, *cI* and LexA. Kenyon and Walker (1980) constructed a family of strains in which SOS inducible promoters were fused to the *lac* operon using the Mu d(Ap^r *lac*) phage (Casadaban & Cohen, 1979). These strains make β-galactosidase in response to DNA damage and provide a convenient assay for SOS induction. Strain GW1060 carries a *uvrA*::Mu d(Ap^r *lac*) fusion. Transformation of this strain with plasmid pSK5 resulted in increased β-galactosidase expression (as measured by colony color intensity on X-Gal plates), indicating that the *Hha*II methylase does induce the SOS response. Strains GW1010, GW1030, GW1040, and GW1080 carry Mu d(Ap^r *lac*) fusions to other SOS-inducible promoters and gave similar results. By microscopic examination, cells expressing *Hha*II methylase were filamented, another characteristic of SOS induction.

The effect of mutations in various DNA repair genes upon methylation sponsored SOS induction was tested by introducing them into an appropriate derivative carrying a *dinD::lacZ* fusion. The resulting strains were transformed with plasmid pSK5 and plated on X-gal indicator media. Blue colonies were scored as induced for the SOS response, and white colonies were scored as uninduced. Mutations in either the *recA* (JH59) or *recB* (JH104) loci blocked SOS induction by *Hha*II methylation. In contrast, a mutation in the mis-match repair machinery (*mutS456*, strain GM1874) did not alter SOS induction by *Hha*II methylase, suggesting that methylation does not induce SOS by somehow disrupting mis-match repair. By the same criterion, *Hha*II methylase did not increase β-galactosidase expression in a strain (JH125) harboring a Mu d(Ap^r *lac*) fusion to one of the adaptive response loci, *alkA*, which is induced by other alkylating agents (data not shown) (Volkert & Nguyen, 1984; re-

viewed in Walker, 1985).

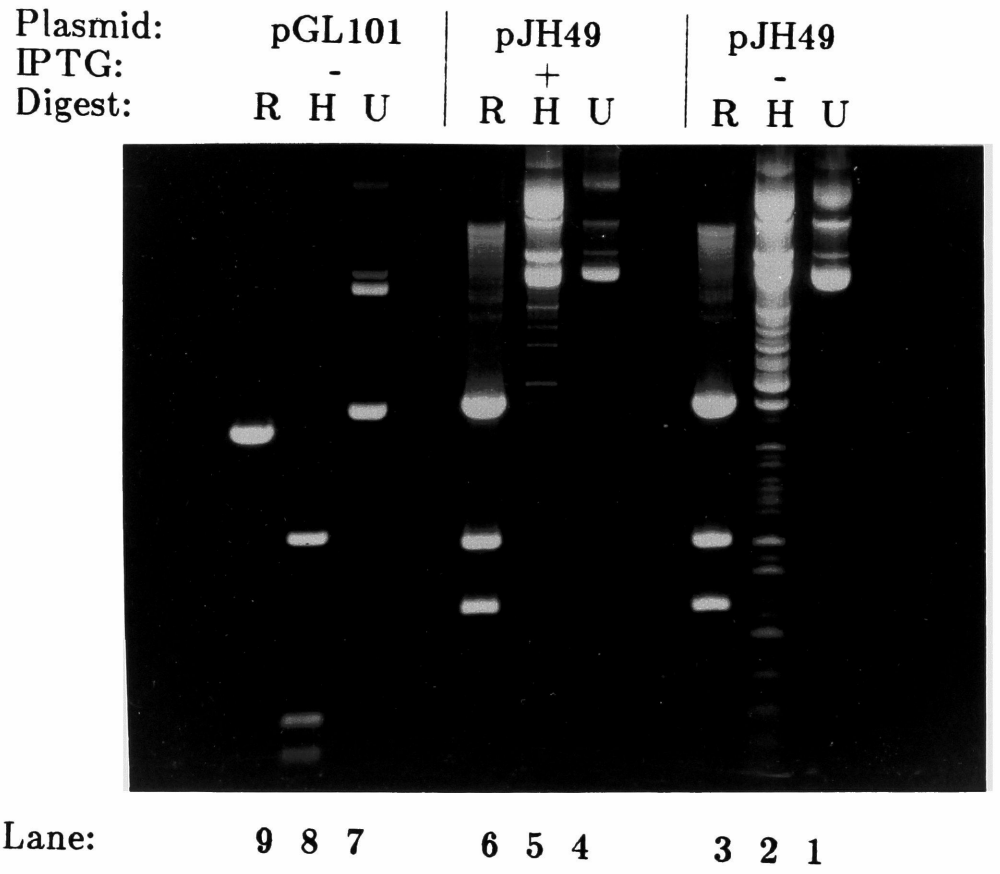
Cloning the HhaII Methylase Under the lacUV5 Promoter

Transformation with a plasmid (pSK5) expressing the *Hha*II methylase causes most *E. coli* strains to grow poorly and give rise to faster-growing mutants. To avoid this selective pressure, the methylase gene was placed under control of the *lac*UV5 promoter, yielding plasmid pJH49 (see materials and methods, Chapter 2). *Hha*II methylase expression was assayed by isolating plasmid DNA and testing its resistance to digestion by *Hin*fI, an isoschizomer of *Hha*II which is also inhibited by *Hha*II methylation (see fig. 3-2). pJH49 expressed a background level of enzyme sufficient to methylate roughly one-half of the plasmid *Hha*II sites under the conditions used (lane 2, fig. 3-2). Some transcription from the uninduced *lac*UV5 promoter probably accounts for the background level of expression from plasmid pJH49. IPTG induction rendered plasmid pJH49 almost completely resistant to *Hin*fI digestion (95% protection), indicating an increase in *Hha*II methylase (lane 5, fig. 3-2).

To assay SOS induction by this inducible plasmid, we constructed strain JH43, a derivative of the Kenyon and Walker strain GW1040 (Kenyon & Walker, 1980). GW1040 carries a Mu d(Ap^r *lac*) fusion to a locus induced by DNA damage, *dinD* (*din* stands for DNA-*d*amage *in*ducible). Although the *dinD* gene product is unknown, we chose this fusion because it has a low basal level of β-galactosidase and the ratio of its induced/uninduced β-galactosidase levels is high. Strain JH43 was derived from GW1040 as follows. The Mu d fusion was rendered transposition defective as described in materials and methods, Chapter 2. The resulting strain, JH5, was transduced with phage P1 grown on a *recA*$^+$ *srl300::*Tn*10* strain to replace the temperature sensitive *recA* allele (*tif-1*) originally carried by GW1040, cured of tetracycline resistance by selection for quinaldic acid resistance, and mated with GC2438 to transfer in the F' episome carrying *lacI*^{q}, to yield strain JH43.

The β-galactosidase levels in strain JH43 after mitomycin C treatment or expres-

Figure 3-2. *Hha*II methylase expression by plasmid pJH49. Plasmid pJH49 DNA was prepared from the $lacI^q$ strain K561 with and without IPTG induction of *Hha*II methylase expression. Samples were digested with *Hin*fI or *Eco*RI and run on a 0.6% agarose gel. pGL101 is the parent vector of pJH49, and here serves as a control. U=undigested sample, H=*Hin*f1, and R=*Eco*RI


Plasmid:
pGL101
pJH49
pJH49
IPTG:
-
+
-
Digest:
R H U
R H U
R H U
Lane:
9 8 7
6 5 4
3 2 1

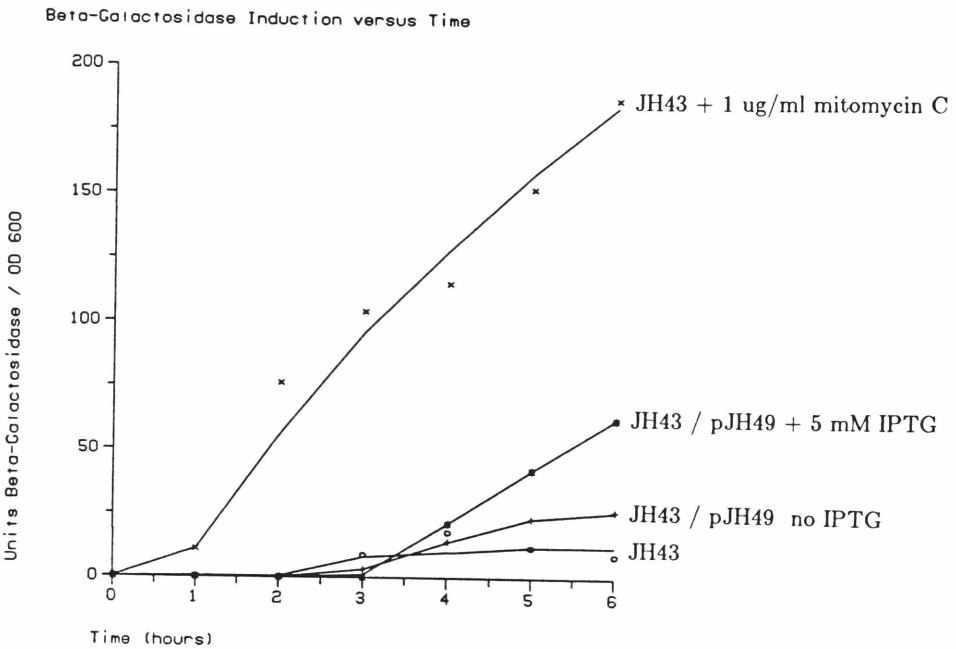
sion of *Hha*II methylase from pJH49 are shown in figure 3-3. Without IPTG induction, plasmid pJH49 made enough *Hha*II methylase to modify about one-half of the plasmid-borne *Hha*II sites, and to induce SOS to 3-fold higher than background. IPTG induction increased methylase expression so that nearly all plasmid-borne *Hha*II sites were methylated and, correspondingly, SOS induction increased to 7-fold higher than background. The increase in SOS induction after IPTG addition was somewhat modest (due to the high basal level of methylase expression), but nonetheless reflects the increased methylation at *Hha*II sites from 50% to essentially 100%.

In control experiments, a dose of mitomycin C of 1 μg/ml induced SOS 21-fold higher than background. Greater SOS induction by mitomycin C than by *Hha*II methylase may reflect a difference in the *nature* of the inducing lesions (since mitomycin C causes DNA-DNA crosslinks, whereas we suggest that *Hha*II methylation leads to DNA scission) or in their respective *concentrations* (a likely concern since mitomycin C can act at nearly any DNA sequence while *Hha*II induced lesions are probably no more frequent than *Hha*II sites). Other examples of partial or intermediate SOS induction have been described (Walker, 1984).

Does the Cell Need to Turn on SOS in Response to Methylation?

We next tested the effect of *Hha*II methylase on the growth of isogenic *recA*$^+$ (K561), and *recA*$^-$ (K871, *recA56* allele) strains. Growth (measured as OD_{600}) was mildly inhibited in strain K561, a λ lysogen (data not shown). Growth was greatly inhibited in the *recA56* strain: the culture doubling time increased by 40% within 6 hours after IPTG induction of methylase expression. However, methylase expression did not decrease colony forming efficiency. Thus methylation is bacteriostatic and not bacteriocidal. Strains carrying either a *lexA3* (JH30) or a *recB21* (JH100) allele gave results similar to the *recA56* strain. These three mutations all block SOS induction by methylation, suggesting that SOS is required to overcome some ill effect of methylation.

Figure 3-3. Kinetics of β-galactosidase induction in strain JH43 (*lacI*^{q} *dinD1::*Mu d(Ap^r *lac*)) and JH43/pJH49. Cells were grown at 30° C in K120 minimal media supplemented with 0.2% glucose and 0.4% casamino acids. Mitomycin C (1 μg/ml) or IPTG (5 mM) was added to cultures at t=0. Portions (100 μl) were removed hourly and β-galactosidase activity was assayed as described (Miller, 1972). Cell density was determined by measuring the OD_{600}. o, untreated JH43; x, JH43 + 1 ug/ml mitomycin C; +, JH43/pJH49; *, JH43/pJH49 + 5 mM IPTG.


Beta-Galactosidase Induction versus Time
Units Beta-Galactosidase / OD 600
200
150
100
50
0
0
1
2
3
4
5
6
Time (hours)
JH43 + 1 ug/ml mitomycin C
JH43 / pJH49 + 5 mM IPTG
JH43 / pJH49 no IPTG
JH43

Other Adenine Methylases Induce SOS

Strain K38 carrying the *Pst*I methylase (CTGC^{me}AG) plasmid pJH56 made mottled translucent colonies and grew poorly, giving rise to faster-growing mutants and blue haloes on X-gal/IPTG media. *Pst*I methylase induced 3 to 5-fold more β-galactosidase expression from the *dinD1*::Mu d(Ap^r *lac*) fusion carried by strain JH43, as measured by β-galactosidase assays (data not shown).

Similarly, when the *dam* methylase was over-expressed from the *tacI* promoter (pJH1), SOS was induced to 2 to 4-fold higher than background, as measured with the *dinD1*::Mu d(Ap^r *lac*) fusion. Similarly, the absence of *dam* methylase (*dam*$^-$ mutation) induces the SOS response to a similar degree (Craig, *et al.*, 1984). However, over- and under-expression of *dam* methylase are not analogous situations. The *dam*$^-$ mutation renders cells 2-aminopurine (2-AP) sensitive (Craig, *et al.*, 1984) and is inviable in combination with a *recA* mutation (Marinus & Morris, 1974). In contrast, we find that over-expression of *dam* does not increase 2-AP sensitivity or kill the *recA*$^-$ host K871 (data not shown). SOS induction by the *dam* mutation may be attributable to DNA lesions from aberrant mis-match repair while over-expression of *dam* may induce SOS by disrupting the regulation of DNA replication initiation from OriC (Messer, *et al.*, 1985).

In contrast, expression of the *Eco*RI methylase from plasmid pJC1 did not induce SOS. K38 harboring plasmid pJC1 grew normally, and strain JH43 did not make increased β-galactosidase in response to *Eco*RI methylation. Plasmid pPaoM.177 (Gingeras & Brooks, 1983) expressed the *Pae*R7 methylase (CTCG^{me}AG) at a level sufficient to fully methylate the single *Pae*R7 site on the plasmid, but did not induce SOS.

How was the Detrimental HhaII Methylase Originally Cloned?

The *Hha*II restriction-modification system (Mann, *et al.*, 1978) was originally

cloned in HB101, a hybrid strain containing regions of both the *E. coli* K-12 and *E. coli* B genomes (Boyer & Roulland-Dussoix, 1969). We have shown that methylation inhibits growth most strongly in *recA* mutant strains, and, surprisingly, HB101 carries a *recA* mutation. Nonetheless, we found that the viability of HB101 is apparently unaffected by the presence of pSK5. In JH69, a $recA^+$ HB101 derivative carrying the *dinD1::*Mu d(Ap^r *lac*) fusion, β-galactosidase levels were increased by mitomycin C but not by *Hha*II methylase.

HB101 is a genetic recombinant, and although most of its chromosome is from *E. coli* K-12, minutes 98 to 100 are derived from *E. coli* B (Boyer & Roulland-Dussoix, 1969). The ability of HB101 to tolerate the *Hha*II methylase is perhaps attributable to these *E. coli* B specific sequences. A *lacZ::*Tn*5* dinD1::Mu d(Ap^r *lac*) derivative (JH120) of the *E. coli* B strain ER21 was constructed. As with JH69, β-galactosidase levels were increased by mitomycin C but not by the *Hha*II methylase. When strains JH69 and JH120 were transduced with phage P1 grown on a K-12 donor strain carrying a transposon insertion at 99 minutes (*zjj::*Tn*10*, strain ER1351), a K-12 marker that was 70% linked to the transposon enabled both strains to induce SOS in response to *Hha*II methylation. This marker was also very tightly linked to the *hsdRMS* loci at 98.5 minutes, which encode the *E. coli* host restriction system. This locus is distinct from both the *hsdR* and *mcrB* genes, since mutations in either of these loci did not block SOS induction by adenine methylation. We call this locus *mrr*, for *m*ethylated adenine *r*ecognition and *r*estriction. We designate the *E. coli* K-12 allele mrr_K or simply *mrr*, and the *E. coli* B allele mrr_B.

Analysis of Tn5 Insertions that Block Induction of SOS by Methylation

We sought Tn*5* insertions in strain JH39 which would block SOS induction by *Hha*II methylase but not by mitomycin C. A random library of 100,000 Tn*5* insertions in strain JH39 were pooled, transformed with plasmid pJH49, and screened on X-gal/IPTG plates. Our reasoning was that transposon insertions in genes needed to

recognize adenine methylation would prevent induction of the Mu d(Ap^r *lac*) fusion and yield white colonies on X-gal/IPTG media. From a screen of 4600, 18 white or light blue colonies were obtained; 12 were totally white, grew slowly on X-gal/IPTG plates, and were presumably insertions in either *recA*, *recBC*, or the *dinD1::* Mu d(Ap^r *lac*) fusion. In contrast, 6 colonies were light blue and grew normally on X-gal/IPTG plates. Mitomycin C still induced SOS in all six. One insertion that co-transduced kanamycin resistance and the Mrr$^-$ phenotype at high frequency was further characterized (*mrr2::*Tn*5*, strain JH76). This insertion also blocked SOS induction by the *Pst*I methylase, an adenine methylase, but not by several cytosine methylases. Cells carrying the Tn*5* insertion restrict λ grown on strains which do not provide EcoK modification, hence the host *Eco*R$_K$ restriction system (*hsdRMS*) is unaffected by the *mrr2::*Tn*5* insertion. Construction of a *mrr2::*Tn*5 recB21* double mutant (JH119) revealed that the *mrr2::*Tn*5* insertion reversed the ill effects of adenine methylation in a *recBC*$^-$ strain. Similarly, the *mrr*$_B$ allele blocked the ill effects of adenine methylation in a *recA*$^-$ strain (HB101).

Mapping the mrr2::Tn5 Insertion

Preliminary mapping of the *mrr2::*Tn*5* insertion by P1 transduction mediated 3-factor crosses revealed that the Tn*5* insertion was 45 to 50% linked to a Tn*10* insertion at 99 minutes (*zjj::*Tn*10*) and very tightly linked to the *hsdR* and *mcrB* loci, both near 98.5 minutes. However, these crosses gave conflicting results as to the gene order (data not shown). To determine the gene order of this tightly linked cluster, we employed a Tn*10* insertion in the *mcrB* locus itself. The *mcrB4::*Tn*10* insertion was transferred by P1 transduction from the donor strain ER1467 (*mcrB4::*Tn*10 hsdR2 mrr*$^+$) into the recipient strain JH76 (*mcrB*$^+$ *hsd*$^+$ *mrr2::* Tn*5*) by selecting Tet^r transductants. From this cross, we find that the *mrr2::*Tn*5* marker is 82% linked to *mcrB4::*Tn*10* (85/104 Kan^S), and that *hsdR2* is 94% linked to *mcrB4::*Tn*10* (98/104 R$^-$) and 86% linked to *mrr2::*Tn*5* (89/104 Kan^{S}R$^-$ or Kan^{r}hsd$^+$). Therefore,

hsd lies closer to *mcrB* than does *mrr*, and *mrr* lies closer to *hsd* than to *mcrB*, establishing the gene order as *mcrB hsdSMR mrr* or *mrr hsdRMS mcrB*.

To further localize the *mrr* gene, a chromosomal *Eco*RI fragment carrying the *mrr2::*Tn*5* insertion was cloned into pBR322 by selection for kanamycin resistance. Since Tn*5* bears no *Eco*RI sites, the isolated fragment carries a portion of of chromosomal DNA (~ 2 kb) flanking the Tn*5* which, by restriction mapping, appears to lie adjacent to the *hsdR* locus (see fig. 3-4). We also obtained a set of λ clones that carry inserts encompassing the region of the *E. coli* chromosome between 98 and 99 minutes (Donna Daniels, personal communication). As discussed below, two of these λ transducing phage carry the intact *mrr* gene (18-115 and 22-25) while one does not (18-123). Phage 18-115 and 22-25 both carry regions of the chromosome that overlap the 2 kb *Eco*RI fragment to which the *mrr2::*Tn*5* insertion was localized. By comparing the resulting fine structure map of this region with the map of the whole *E. coli* chromosome (Kohara, *et al.*, 1987), we find that *mrr* lies just clockwise to *hsd* with the resulting gene order *mcrB hsdSMR mrr zjj::*Tn*10* (see figure 3-4).

The mrr Gene Product Restricts Methylated λ *in vivo*

We suspected that the *mrr* gene might encode an endonuclease that cleaves DNA containing N^6-methyladenine. If so, the *mrr* locus should act, *in vivo*, as a classical restriction-modification system (Arber, 1974). To test this, a λ^v stock was prepared from a strain making the *Hha*II methylase (JH83/pJH49, + 2 mM IPTG). This methylated phage made 10 to 20-fold more plaques on a mrr^- (JH76) than on a mrr^+ (JH39) host (see Table 3-1). In contrast, unmethylated phage plated with equal efficiency on mrr^+ and mrr^- strains. Thus, *Hha*II methylated λ is *restricted* on a mrr^+ host. Passaging these methylated phage on the mrr^- mutant strain JH83, to remove the methyl groups, yielded phage which plated with equal efficiency on mrr^+ and mrr^- strains. Methylated phage that survived restriction by Mrr (plaques picked from a JH39 lawn) also plated with equal efficiency. *Hha*II methylated phage are

Figure 3-4. Map location of the *mrr2::*Tn*5* insertion and fine structure of the *hsd* region. λ transducing phages 22-25, 18-115, and 18-123 carry cloned segments that map in or near the *hsd* region (Donna Daniels, personal communication). As described in the results, phage 22-25 and 18-115 carry the *mrr* gene. A limited restriction map of the *hsd* genes has been described (Sain & Murray, 1980). The more extensive and complete restriction map that appears here was constructed and generously provided by Donna Daniels (personal communication).

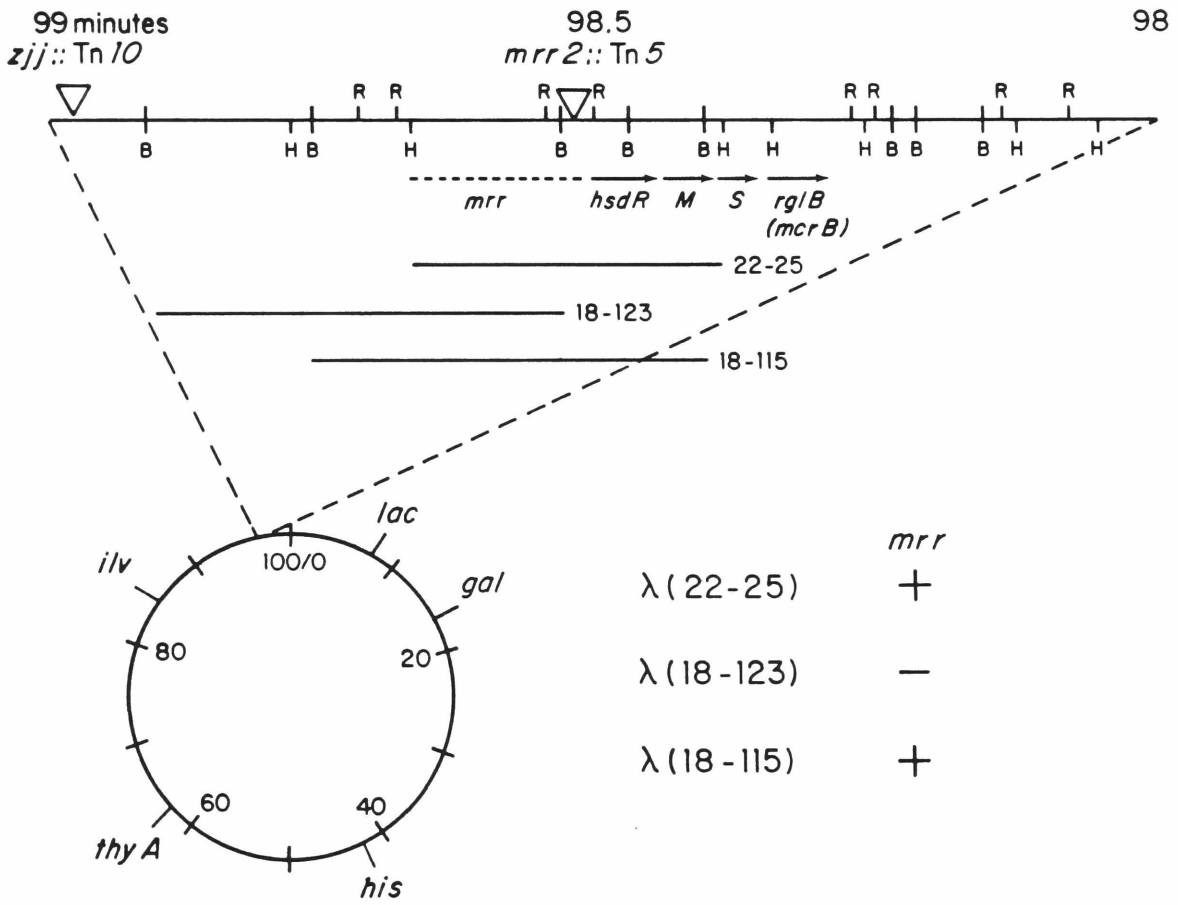

99 minutes
zjj::Tn10
98.5
mrr2::Tn5
98
R R R R R R R R
B H B H B B B H H H B B B H H
mrr
hsdR M S rglB
(mcrB)
22-25
18-123
18-115
lac
ilv
gal
100/0
80
20
60
40
thy A
his
mrr
λ(22-25) +
λ(18-123) —
λ(18-115) +

Table 3-1. Restriction of Methylated λ^v by Mrr

Efficiency of Plating Relative to JH76.			
of λ^v phage stock:		on host strain:	
first growth cycle	second growth cycle	JH39 (*mrr*$^+$)	*JH76 (mrr2::*Tn*5)*
JH83	-	1	1
JH83/pJH49 +IPTG[a]	-	0.05-0.1	1
JH83/pJH49 +IPTG	JH83[b]	1	1
JH83/pJH49 +IPTG	JH39[c]	1	1
JH83/pJH56[d]	-	0.1	1
JH83/pJH56	JH39[e]	1	1

[a] λ^v was methylated in vivo by growth on a strain expressing the *Hha*II methylase (JH83/pJH49 + IPTG).

[b] *Hha*II methylated λ^v was plated on strain JH83 to remove the methylation.

[c] *Hha*II methylated λ^v that survived growth on the *mrr*$^+$ strain JH39 were isolated and re-plated.

[d] λ^v was methylated in vivo by growth on a strain expressing the *Pst*I methylase (JH83/pJH56).

[e] Again, the methylated λ^v was passaged to remove the methylation.

therefore *modified* by passage on strains that do not express *Hha*II methylase. When methylated phage that survived Mrr restriction were methylated a second time, they once again became sensitive to Mrr restriction (data not shown). *Pst*I methylated λ gave similar results. Thus, the *mrr* gene behaves as a classical restriction-modification system that, rather than being inhibited by methylation, specifically restricts methylated DNA.

λ*pmrr Transducing Phage are Restricted on Hosts Expressing HhaII Methylase*

Our finding that methylated λ is restricted by Mrr suggested that a λp*mrr* transducing phage might be restricted on hosts making the *Hha*II methylase. We tested three λ transducing phage carrying segments of the *E. coli* chromosome to which *mrr* mapped (Donna Daniels, personal communication) for their ability to grow in the presence of *Hha*II methylase. Two of these, 22-25 and 18-115 (see figure 3-4), plated with greatly reduced efficiency (5×10^{-6}) on strain LE392 (*supE supF* host required to grow phage 22-25) making the *Hha*II methylase compared to strain LE392 alone. We suggest that, upon infection, *mrr* is expressed at high levels, the *Hha*II methylated host chromosome is degraded, and the subsequent cell death decreases or prevents the production of progeny phage. In addition, the infecting phage may also become *Hha*II methylated and suffer restriction by Mrr. λp*mrr* phage that survived growth in the presence of *Hha*II methylase plated with equal efficiency in the presence and absence of the methylase and presumably carry *mrr* alleles with reduced activity (data not shown). We note that strain LE392 is mrr^+ and that this contributes to the reduced plating efficiency of the λp*mrr* transducing phage since phage 22-25 and 18-115 were less inhibited (10^{-2}) by *Hha*II methylation in a mrr^- LE392 derivative.

Effects of Cytosine Methylases on E. coli

Many plasmids which express cytosine methylases cannot be introduced into stan-

dard *E. coli* strains (Kiss, *et al.*, 1985; Blumenthal, *et al.*, 1985; Raleigh & Wilson, 1986; Noyer-Weidner, *et al.*, 1986). The locus responsible for this effect, *mcrB*, encodes a restriction enzyme that cleaves DNA containing 5-hydroxymethyl or 5-methylcytosine (Revel, 1983; Raleigh & Wilson, 1986). We have found that after transformation by plasmids expressing cytosine methylases, poorly growing colonies arise that are induced for the SOS response. These transformants occur at low frequency (~ 0.1 to 1% of *mcrB*$^-$ strains) and upon restreaking or prolonged incubation give rise to faster-growing variants which are not induced for the SOS response. In most cases, the *mcrB* locus is required for this effect (see Table 3-2). In the same fashion as adenine methylation, *Hha*I specific cytosine methylation in an *mcrB*$^+$ background appeared to be lethal in combination with a *recA*$^-$ (JH59) or a *recB*$^-$ mutation (JH104), since no transformants could be obtained in the *recA*$^-$ or *recB*$^-$ background. The *mcrB*$^-$ mutation did not block SOS induction by the *Hha*II adenine methylase.

Discussion

By three separate criteria (λ prophage induction, cellular filamentation, and increased expression of Mu d(Ap^r *lac*) fusions to DNA damage inducible loci) we find that many site-specific methylases induce the SOS DNA repair response when expressed in *E. coli*. In general, methylases foreign to *E. coli* (*Hha*II, *Pst*I, *Msp*I, and others) induce the SOS response, while methylases indigenous to *E. coli* induce SOS to a lesser extent (*dam*) or not at all (*Eco*RI, *Eco*RII). Methylases that induce the SOS response also cause the cell to grow somewhat slowly. This slow growth may arise from unrepaired methylation dependent DNA lesions, or from SOS induction itself, since one of the SOS gene products is the SfiA protein, an inhibitor of cell division (Huisman, *et al.*, 1984). If one blocks SOS induction (using mutations in either *recA*, *recB*, or *lexA*), the cell grows more poorly in response to methylation. Clearly SOS induction counteracts some detrimental effect of methylation. On the other

Table 3-2. SOS Induction by Cytosine Methylases

Plasmid	Methylase	Specificity[a]	Induces SOS?[b]	
			in *mcrB*$^+$	in *mcrB*$^-$
pMER3	*Msp*I	^{me}CCGG	yes[c]	no
pHaeIII 1-1	*Hae*III	GG^{me}CC	yes[c]	no
pR215	*Eco*RII	C^{me}CA/TGG	no	no
pJH53	*Hha*I	G^{me}CGC	yes[c]	yes
pJH51	*Bam*HI	GGAT^{me}CC	yes	no
pSU11	*Bsu*RI	GG^{me}CC	yes[c]	no
pSU21	*M. Spr*	GGCC, CCGG, GGA/TCC	yes[c]	no
pES2	*Bsp*RI	GG^{me}CC	yes[c]	no

[a] ^{me}C=5-methyl cytosine

[b] SOS induction was assayed by plating transformants of strains JH43 and JH63 on X-Gal indicator medium. Blue colonies were scored as SOS induced, and white as uninduced.

[c] Transformation efficiency decreased relative to *mcrB*$^-$ strain; yields small colonies.

hand we have shown that certain mutations completely block the ill effects of methylation and circumvent the need for SOS induction. Thus a mutation in the *mcrB* locus allows normal cell growth without SOS induction in the presence of cytosine methylation, and likewise, a *mrr* mutation blocks the deleterious effects of adenine methylation. In the absence of McrB and Mrr activity, methylation itself has little or no effect on cell growth.

We initially considered that DNA methylases might induce the SOS response if the methylated bases are mutagenic. In fact, several types of DNA methylation (N^3-, N^7-, and O^6-methylguanine; N^3-methyladenine; O^2- and O^4- methylthymine; and O^2-methylcytidine) are mutagenic lesions which are repaired in most cases by glycosylases (Walker, 1985). We have shown that *Hha*II methylated DNA contains N^6-methyladenine and no other unusually modified adenine residues (Heitman, Model, Kuo, Phan, and Gehrke, unpublished results). N^6-methyladenine is not known to be a mutagen and we find that the *Hha*II methylase has little or no mutagenic effect (data not shown). The cytosine methylases we have studied produce 5-methylcytosine, which is weakly mutagenic because it spontaneously deaminates to thymine, producing GT mis-matches (Coulondre, *et al.*, 1978). However, we cannot attribute the SOS inductive effect to a direct mutagenic action of the modified base, because we find that several other adenine (*dam*, *Eco*RI) and cytosine (*Eco*RII) methylases do not induce SOS even though they produce the same chemical modification as methylases which do induce SOS. SOS induction therefore requires sequence specific methylation. It would be highly unusual for these modified bases to behave as a mutagenic lesion in one sequence but not another. The only example where this is known to occur is the sequence specific recognition of deaminated 5-methylcytosine residues (GT mis-matches) by the very short patch (VSP) mismatch repair system (Lieb, 1987). Even in this example the mutagenic base is not a sequence specific mutagen, but rather it is subjected to repair by a sequence specific protein. We therefore turned our attention to sequence specific recognition systems.

The *mcrB* locus (formerly called *rglB*) was originally described as a restriction system that cleaves DNA containing 5-hydroxymethylcytosine, which is present in unglucosylated T-even phages (Revel, 1983). Recently it was shown that *mcrB* also restricts DNA containing 5-methylcytosine (Raleigh & Wilson, 1986) and perhaps N^4-methylcytosine as well since the N^4-cytosine methylase *Pvu*II (Januliatis, *et al.*, 1983) is sensitive to McrB restriction (Blumenthal, *et al.*, 1985). We find that many cytosine methylases induce the SOS response, and that this effect requires *mcrB*. It seems likely that DNA double-strand breaks, caused by *mcrB* action, induce SOS. These findings suggest that the normal substrate for McrB is methylated DNA, which is ubiquitous. This answers the conundrum raised by the original description of McrB which was: why does the cell make a restriction system for 5-hydroxymethyl cytosine containing DNA since this substrate is only known to occur in T-even phage glycosylase mutants isolated in the laboratory.

I should point out that it was previously observed that *mcrB* restriction of unglucosylated T-even phage caused a weak induction of the SOS response (Dharmalingam & Golberg, 1980). In this case, it is thought that SOS induction occurs because the T-phage restriction products are degraded to single-stranded DNA that activates the RecA protein. We suggest that the more robust SOS induction we observe is mainly attibutable to DNA double-strand breaks in the chromosome and to a much lesser extent to restriction and degradation of the methylase expressing plasmid. This assertion is supported by the finding that the magnitude of SOS induction by at least one restriction system (*Eco*RI) is independent of plasmid-borne recognition sites.

In separate experiments, we have found that temperature-sensitive mutants of the *Eco*RI endonuclease induce the SOS response under conditions where *in vivo* DNA scission occurs (Heitman, *et al.*, 1989; Chapters 4 & 5). Therefore, DNA scission does induce the SOS response. The adenine methylases *dam*, *Hha*II, and *Pst*I also induce SOS. With *Hha*II and *Pst*I, this effect requires the *mrr* locus. SOS induction by *dam* over-expression is independent of *mrr* (unpublished results) and may arise from aber-

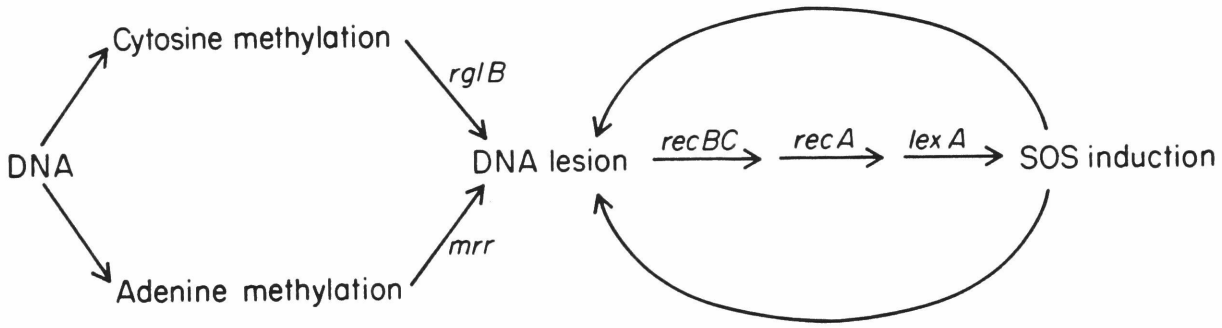
rant DNA replication since *dam* over-expression increases the rate of initiation from OriC (Messer, *et al.*, 1985). By analogy with McrB and *Eco*RI and based on our *in vivo* evidence that *mrr* functions as a restriction system, we suggest that the Mrr protein is an endonuclease that cleaves DNA containing N^6-methyladenine. A comparison of methylated sequences that result in restriction and SOS induction suggest that Mrr may recognize the methylated trinucleotides $G^{me}AC$ and $C^{me}AG$ (see figure 3-5). The notion that Mrr is an endonuclease is our simplest model. Alternatively, Mrr could be for example, a sequence-specific glycosylase which cleaves N^6-methyladenine to yield apurinic sites. Subsequent apurinic endonuclease action would convert these apurinic sites into nicks, and since the methylated sites are palindromic, the net result would be a DNA double-strand break. In fact, a perhaps analogous combination of uracil-N-glycosylase activity (ung^+) and apurinic endonuclease acts as a restriction system against DNA containing uracil (Duncan, 1985).

Our model for the pathway from methylation to SOS induction is shown in figure 3-6 which depicts McrB and Mrr introducing lesions in methylated DNA. We suggest that these lesions are DNA double-strand breaks which are processed by the RecBCD exonuclease or helicase activity to yield the SOS inducing signal, most likely single-stranded DNA. SOS functions then allow normal cell growth to continue. The role RecBCD plays here may be analogous to other known functions it performs including: degradation of restricted DNA (Simmon & Lederberg, 1972), activation of SOS by other inducing agents (McPartland, *et al.*, 1980), and recombinational activity at Chi sites (Stahl & Stahl, 1977). Some types of DNA double-strand breaks (γ-ray breaks, λ terminase) require recombination functions for repair, and are probably repaired by homologous recombination with an intact copy of the severed DNA (Stahl, *et al.*, 1983). If methylation dependent lesions are repaired in this way, the RecA and RecBCD proteins may be required not only to induce the SOS response, but also to participate directly in recombination mediated repair. Alternatively, SOS induction may serve only to delay cell division and allow other SOS-independent repair func-

Figure 3-5. Effects of Methylation on SOS induction. Putative recognition sites for Mrr and McrB (RglB) are shown. Methylase recognition specificities are from Smith & Kelly (1984).

	ADENINE METHYLASES		CYTOSINE METHYLASES	
Induce SOS	*Hha* II	G ^{me}ANTC	*Hha* I	G ^{me}CGC
	Pst I	C ^{me}AGCTG	*Hae* III	GG ^{me}CC
			Bsp RI	GG ^{me}CC
			Bsu RI	GG ^{me}CC
			Msp I	^{me}CCGG
Partially induce SOS	Dam (under tac control)	G ^{me}ATC	*Bam* HI	GGAT ^{me}CC
Do not induce SOS	*Eco* RI	GA ^{me}ATTC	*Eco* RII	C ^{me}C(A or T)GG
	Pae R7	CTCG ^{me}AG		
Recognition site		G ^{me}AC		G ^{me}C
		C ^{me}AG		(T ^{me}C)

Figure 3-6. Model of pathway from methylation to SOS induction. Mrr and McrB (RglB) are shown causing lesions in methylated DNA. We suggest that these are double-strand breaks which act as entry sites for the RecBCD complex. Subsequent RecBCD helicase or exonuclease action produces a signal (probably single-stranded DNA) that activates the RecA protein and induces the SOS response.


Cytosine methylation
rglB
DNA
DNA lesion
recBC
recA
lexA
SOS induction
mrr
Adenine methylation

tions more time to act.

In contrast, DNA breaks by the *Eco*RI enzyme induce the SOS response but do not require this induction for their repair (Heitman, *et al.*, 1989). Our observation that *mrr* restriction inhibits growth more severely in the absence of SOS induction, argues that in this case SOS induction has a protective affect. Since *Eco*RI produces staggered DNA breaks, the Mrr protein may instead cleave DNA to produce one of the types of DNA breaks which are known to require SOS and recombination for repair (3'-phosphate/lost base, cross-linked, protein bound, maybe blunt ends; see chapter 5 for a more detailed discussion). Alternatively, Mrr may nick DNA because we find that *Eco*RI mutants that nick DNA produce lesions that are repaired by SOS induction and recombination (chapter 8). Lastly, the *mrr* system may be subject to restriction alleviation, a phenomenon whereby SOS induction inhibits the activity of some (*hsd*) but not all (not *Eco*RI or *Eco*RII) restriction systems (Day, 1977).

Our findings that *Hha*II or *Pst*I methylated λ is restricted on mrr^+ strains, and λp*mrr* transducing phage are restricted on strains expressing the *Hha*II methylase clearly show that Mrr can restrict adenine methylated DNA. This raises two interesting points. First, the *mrr* locus maps adjacent to two other loci, *hsdRMS* and *mcrB*, which also encode restriction systems. This clustering may be functional, to ensure that *E. coli* infrequently divulges its means of marking its DNA, to increase the region of the chromosome that can be transferred without donating a potentially lethal gene, or to perhaps in some way coordinate the cellular defense system. Second, Mrr activity is found in *E. coli* K-12 but not in *E. coli* B. Similarly, the McrB activity in *E. coli* B is weaker and differs in specificity from that found in K-12 (Revel, 1983), and the *hsdRMS* loci have diverged to different specificities from a common ancestor (Gough & Murray, 1983). Mutations of this sort, which inhibit or promote genetic exchange, were perhaps responsible for the divergence of *E. coli* K-12 and B and may represent a common event in bacterial speciation.

Epilog

Since this chapter appeared (Heitman & Model, 1987) there have been a number of developments in the *mrr/mcr* story. Mutational analysis and cloning confirmed that the *mcr* and *rgl* systems are encoded by the same loci (Raleigh, *et al.*, 1989). However, the plot thickens because *mcrB* contains at least two complementation groups (Dila & Raleigh, 1989), the DNA sequence of *mcrB* reveals three open reading frames (Ross, *et al.*, 1989), and the cloned locus produces up to three polypeptides of disputed sizes (Ross, *et al.*, 1987; Noyer-Wiedner, personel communication; Sozhamannan & Dharmalingam, 1988). Thus the McrB enzyme may be a multisubunit enzyme with different subunits or subunit arrangements for each type of substrate (5-HOMeC, 5-MeC, and N^4-meC). The complex arrangement of the *mcrB* function is reminiscent of *hsdRMS*, and may explain why the McrB protein has been somewhat recalcitrant to purification. Other methylation-dependent restriction enzymes (McrA, Mrr) may be as complex.

In fact, several observations suggest that Mrr may be a multisubunit enzyme encoded by several genes. First is a consideration of the recognition specificity of Mrr, which includes $G^{me}AC$ and $C^{me}AG$. Since a GC or CG basepair displays an amino group in the minor groove while an AT or TA basepair does not, the enzyme could recognize the sequence G/C ^{me}A G/C (where G/C= G or C) by probing the minor groove. However, the sequence $G^{me}AG$ is not restricted when it occurs in the sequence TCTGAG. Thus the recognition of both $G^{me}AC$ and $C^{me}AG$ is not easily reconciled with one recognition domain. Since mutations can affect both specificities, one possible explanation is a multi-subunit enzyme with one common subunit and multiple recognition subunits.

In addition, a number of experimental observations suggest the *mrr* locus is an unusual enzyme or encoded by more than one gene. First, Lisa Raleigh has found that a previously described clone encompassing part of the *hsd* locus (pBG3, Sain &

Murray, 1980) complements the Mrr$^-$ phenotype of an *mcr hsd mrr* deletion strain and restores Mrr restriction of either *Hha*II or *Pst*I methylated λ. However, we find that this clone does not complement another *mrr*$^-$ allele (*mrr2::*Tn*5*, strain JH76) for phage restriction. These observations can be reconciled if the *mrr* locus is an operon in which the Tn*5* insertion but not the deletion has a polar effect on a downstream locus.

Second, Lisa Raleigh has fortuitously found a strain (GM2163 *dam*$^-$ *dcm*$^-$ *mcrB*$^-$ *mcrC*$^+$ *mrr*$^+$) which restricts *Pst*I but not *Hha*II methylated λ phage, thus splitting the *mrr* recognition specificities. This strain may contain a mutation affecting a subunit required for one specificity but not the other. Alternatively, since this strain carries a *dam*$^-$ mutation, its phenotype may be attributable to the recently described restriction alleviation phenomenon observed in *dam*$^-$ strains (Efinova, *et al.*, 1988). However, this explanation requires restriction alleviation to affect Mrr restriction of some methylated sequences (*Hha*II) but not others (*Pst*I).

Third, while determining the identity of the DNA base modified by the P15 methylase (CAGCAG), Tom Bickle (personel communication) observed an activity which could be the Mrr enzyme. They labelled a P15 site near the end of a DNA fragment with the P15 methylase and SAM (^{3}H), and then treated with DNA polymerase I to do wandering spot DNA sequencing and follow the position of the label. However, the commercial preparation of DNA polymerase I (Boehringer) was contaminated with an enzyme that released the radioactivity in a form which co-migrated on TLC with dAMP, and that this activity could be separated from the DNA polI by heparin agarose chromatography. If this represents Mrr activity, it suggests that the enzyme is some type of endonuclease. However, these observations are consistent with several mechanistic models including: incision on both sides of the base, or the combined action of a Mrr nick or break with the exonuclease activity of DNA polymerase I.

Lastly, we attempted to clone the *mrr* gene by screening a plasmid library for

members which would complement a *mrr*$^-$ host and restore induction of the SOS response by the *Hha*II methylase (borne on a plasmid compatible with that of the library). We isolated a ~ 10.5 kb *Bam*HI fragment from chromosomal DNA of the *E. coli* K-12 strain K38 which clearly complemented the mrr_B *dinD1*::Mu d(Ap^r *lac*) host JH69 for the SOS induction phenotype. However, the cloned gene did not restore restriction of either *Hha*II or *Pst*I methylated λ_{vir} to the mrr_B mutant or increase restriction of either methylated phage in a wild-type *mrr*$^+$ background. Furthermore, by restriction mapping this cloned fragment does not correspond to the map position of the *mrr* locus identified with the Tn*5* insertion and the mrr_B allele. Hence, this clone may identify a gene which acts as an extragenic suppressor when over-produced. We have been unable to locate this fragment on the restriction map of *E. coli* or λ.

Several models could explain these observations. First, Mrr could contain several subunits and a defect of one might be compensated by an increased amount of another subunit. For example, the mrr_B allele could produce a defective subunit which poorly binds the other subunit. Alternatively, restriction by Mrr could occur in two steps. The mrr_B allele could encode a partially active product, and the decreased number of DNA lesions it produces would require an increased level of some second function, encoded by the cloned complementing gene, that processes the lesions to induce the SOS response. For example, Mrr could be a glycosylase and the second function an apurinic endonuclease. Either model makes the testable prediction that the cloned gene would not complement a null *mrr* allele. Other approaches to define the function of this other gene would be to map it by insertional disruption, to determine the phenotype of strains lacking the gene product (are they *mrr*$^-$?), and to sequence the gene and characterize its encoded product.

Further characterization of the *mcrA, mcrB* and *mrr* genes and their products should enhance our understanding of both restriction-modification systems and the mechanisms whereby proteins recognize methylated DNA.

Chapter 4

Temperature Sensitive Mutants of the *Eco*RI Endonuclease

Introduction

The number of proteins for which both the primary amino acid sequence and the crystal structure have been determined is growing rapidly. However we still do not understand how proteins fold into their stable tertiary structures, or why they chose one particular structure. Algorithms allow one to deduce secondary structure from the known primary sequence with some degree of accuracy (Chou & Fasman, 1974, 1978; Pongor & Szalay, 1985; Richards & Kundrot, 1988), however tertiary structure can only be determined by X-ray crystallography. A genetic analysis of proteins with known three dimensional structures may help elucidate the interactions which contribute to protein stability and folding.

The structures of many proteins are plastic. For example, at several positions within λ repressor different amino acids are compatible with activity (Reidhaar-Olson & Sauer, 1988; Bowie and Sauer, 1989); β-lactamase and several proteins of phage f1 tolerate one or two amino acid insertions (Boeke, 1981; Barany, 1985); and mis-suppression of β-galactosidase nonsense mutants is often without consequence (Miller, 1979). In contrast, the amino acids of enzyme active sites can rarely be substituted by non-conservative amino acid changes. Somewhere in between lie residues which are important but not essential to protein function. Many should play a role in stabilizing the tertiary structure and can be identified by isolating mutants with partial or conditional activity.

A large collection of temperature-sensitive mutants have been previously described in a protein of known structure, the T4 phage lysozyme (Grutter, *et al.*, 1983; Alber, *et al.*, 1987a, 1987b). These mutations result from subtle to drastic amino acid changes that lie at positions throughout the primary and secondary structures, mainly in one lobe within the bilobed enzyme. These sites are correlated with

residues of low thermal mobility and solvent accessibility. *In vitro*, the purified mutant proteins exhibit dramatic reductions in thermal stability. In several other examples of TS mutants (the *S. cerevisiae* protein Cdc28 (Lorincz & Reed, 1986), the *E. coli lac* repressor (Miller, 1979), and the P22 tailspike protein (King, *et al.*, 1987; Fane & King, 1987)) the responsible mutations show a different pattern and fall at positions where the predicted secondary structure is in transition. However, in these cases the protein tertiary structures and behavior *in vitro* are unknown.

We have isolated temperature sensitive (TS) mutants of the *Eco*RI endonuclease, whose X-ray crystal structure is known to 3 A resolution (McClarin, *et al.*, 1986). We report here our findings that these TS mutations lie at positions within the X-ray crystal structure where the protein secondary structure is in transition: the ends of α-helices and β-sheets, turns and loops. Furthermore, when either of three amino acids lying at the ends of two α-helices was altered by site-directed mutagenesis, substitutions which were compatible with activity invariably conferred a temperature-sensitive phenotype. These mutant proteins are TS for function *in vivo* but not *in vitro*, and therefore may be defective in folding. Our observations are in contrast to earlier studies of phage T4 thermolabile lysozyme mutants (Grutter, *et al.*, 1983). We suggest that mutations which render protein folding TS may disrupt secondary structure transitions, while those that render function TS affect many additional positions. Protein folding may be orchestrated by protein segments which direct neighboring secondary structural elements to their correct tertiary configuration.

Results

Isolation of temperature sensitive EcoRI mutants

The *Eco*RI restriction-modification system consists of two proteins. The first is an endonuclease that cleaves DNA at the sequence GAATTC (Hedgpeth, *et al.*, 1972). The second is a methylase which recognizes the same sequence and transfers methyl groups from *S*-adenosyl methionine to the N^6 position of the central adenines on each

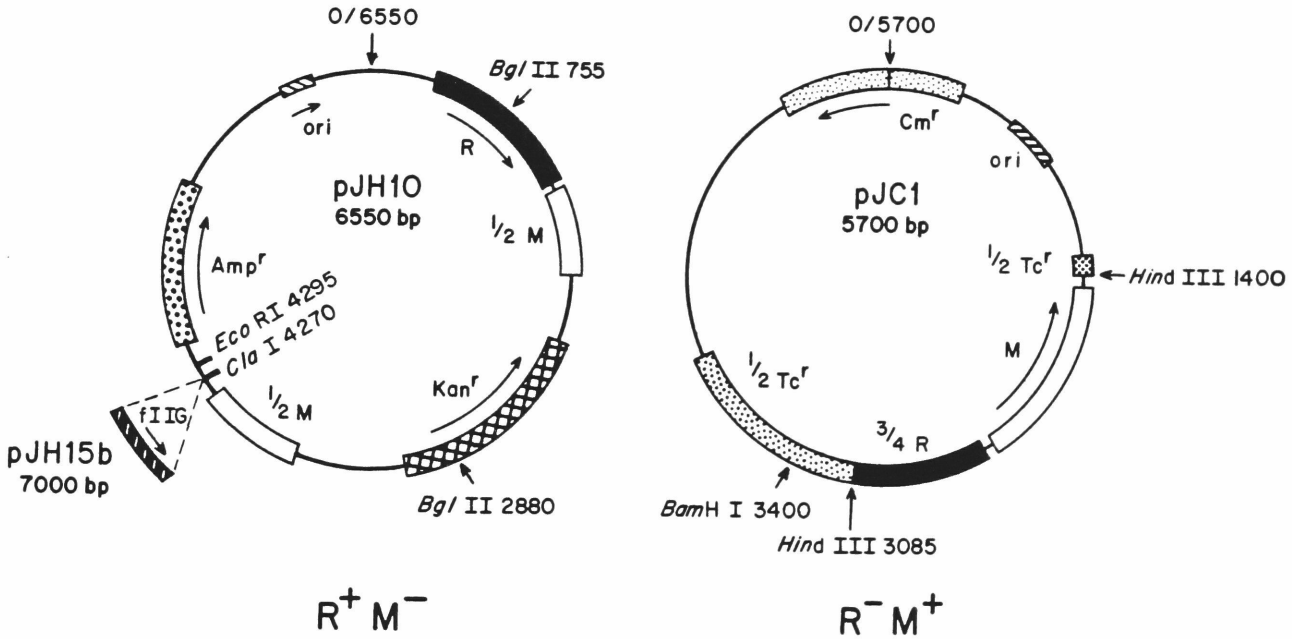
strand (Rubin & Modrich, 1977). Methylation blocks endonuclease action and thereby protects the host chromosome from destruction. By this means, unmethylated foreign DNA is destroyed while the cellular DNA is protected.

Normally the genes for the two proteins lie adjacent to one another and may constitute an operon (Newman, *et al., 1981; Greene, et al.*, 1981). To selectively mutagenize the endonuclease, the methylase and endonuclease genes were cloned on the separate compatible plasmids shown in Figure 4-1. The procedure for isolating *Eco*RI endonuclease mutants is described in detail in materials and methods (see also Heitman *et al.*, 1989). Cells bearing both the endonuclease plasmid pJH10 (or pJH15a/b) and the methylase plasmid pJC1 were mutagenized spontaneously or with nitrosoguanidine. The two plasmids were then separated by treating plasmid DNA with *Bam*HI to linearize pJC1 but spare pJH10, or by selectively packaging plasmid pJH15a/b into f1 transducing particles with f1 helper phage. The mutagenized endonuclease plasmid was then introduced into a strain lacking the *Eco*RI methylase. Cells that acquire a wild-type copy of the *Eco*RI endonuclease gene (R^+ M^-) suffer DNA destruction and die. Surviving colonies should therefore bear endonuclease mutants. By initially selecting colonies which grow at 42°C and subsequently replica-plating to 30°C and 42°C, we found that a high proportion (25-35%) of survivors were cold-sensitive for growth. In all cases the cold-sensitive phenotype was attributable to a plasmid-borne mutation mapping within the *Eco*RI endonuclease gene. Thus these colonies express temperature (or heat) sensitive *Eco*RI mutants which are partially or completely inactive at 42°C but become active and impair cell growth at lower temperatures. Since *Eco*RI endonuclease activity is lethal to the cell, these TS mutants confer cold-sensitive growth on the host.

Mapping and sequence analysis

To facilitate DNA sequencing, the mutations were mapped to one of the four quarters of the endonuclease gene. Three restriction sites (*Hin*dIII, *Bgl*II, and *Pst*I)

Figure 4-1. Compatible plasmid system. The *Eco*RI endonuclease and methylase genes were cloned on the separate compatible plasmids shown here and described in materials (see also Heitman, *et al.*, 1989). Plasmid pJH10 is a pBR322 derivative which confers ampicillin and kanamycin resistance, expresses the *Eco*RI endonuclease, and bears an interrupted version of the *Eco*RI methylase. The pJH10 derivatives pJH15a/b carry the f1 intergenic region which allows these plasmids to be packaged in a single-stranded DNA form in f1 transducing particles after their host cell is infected with f1 helper phage. Plasmid pJC1 is a pACYC184 derivative (kindly provided by Cheng & Modrich (1983)) which confers chloramphenicol resistance, expresses the *Eco*RI methylase, and is compatible with plasmids pJH10 or pJH15a/b. We note that the plasmids share a large region of homology including the entire methylase gene and three-quarters of the endonuclease gene. A deleted version of pJC1 (pJH20), lacking the region between its *Bam*HI site and the *Pst*I site in the endonuclease gene, was employed in some circumstances to reduce the region of homology between the two plasmids.


0/6550
Bgl II 755
ori
R
pJH10
6550 bp
1/2 M
Ampr
Eco RI 4295
Cla I 4270
fl IG
1/2 M
Kanr
pJH15b
7000 bp
Bgl II 2880
R+ M−
0/5700
Cmr
ori
pJC1
5700 bp
1/2 Tcr
Hind III 1400
M
1/2 Tcr
3/4 R
BamH I 3400
Hind III 3085
R− M+

are conveniently disposed at one-quarter, one-half, and three-quarters the length of the endonuclease gene and served as end-points. As described in detail in materials and methods, mutations were mapped by three methods: marker rescue using single-stranded mutant DNA annealed to restriction fragments of wild-type, ligation of mutant and wild-type restriction fragments, or heteroduplex deletion mapping (Shortle, 1983) with a set of known *Eco*RI deletions (see chapter 2, figure 2-1). Sequence analysis of the identified region revealed the amino acid substitutions shown on the crystal structure (McClarin, *at al.*, 1986) in Figure 4-2 (see also the summary in table 4-4).

To avoid multiple mutations induced by some chemical mutagens, most of these alleles were isolated spontaneously. Additionally, the entire gene was sequenced for several mutants (TS6, TS9, TS14). One nitrosoguanidine induced allele (TS14) contained two mutations (V97I and Y165F). By heteroduplex deletion mapping, the TS phenotype was attributable to the V97I mutation.

Interestingly, 11 of 12 TS mutations fall at positions in the crystal structure where the secondary structure is in transition. Thus mutations TS6, TS8, and TS9 lie at the ends of α-helices, TS7 and TS17 lie at the ends of β-sheets, and TS0, TS1, TS2, TS3, TS12, and TS14 lie in loops or turns. We note that one TS allele results from the substitution of a UGA stop codon at glycine 186. It is unlikely that the opal protein fragment is active, since deletions which remove similar regions are inactive and many TS mutations map to the distal region of the protein. This phenotype probably occurs because even non-suppressing strains are known to weakly suppress opal codons by inserting tryptophan.

Effects on cell growth of TS endonuclease expression

Table 4-1 summarizes the affects of these TS endonucleases on colony growth of strain K91. As the temperature of growth is reduced to activate the mutant enzymes, the colonies ruffle at their borders. With increasing activity, the colonies be-

Figure 4-2. Location of the TS *Eco*RI mutations on the crystal structure. TS *Eco*RI endonuclease mutants were isolated as described in results. The responsible mutations were mapped by either ligation of mutant and wild-type restriction fragments, marker rescue, or heteroduplex deletion mapping. The implicated region was then sequenced by the dideoxy chain termination method (Sanger, *et al.*, 1977). The protein structure shown here is derived from McClarin *et al.*, 1986.

thr261ile
TS0

met255ile
TS9

met157ile
TS8

cys218tyr
TS1

arg56gln
TS6

leu263phe
TS3

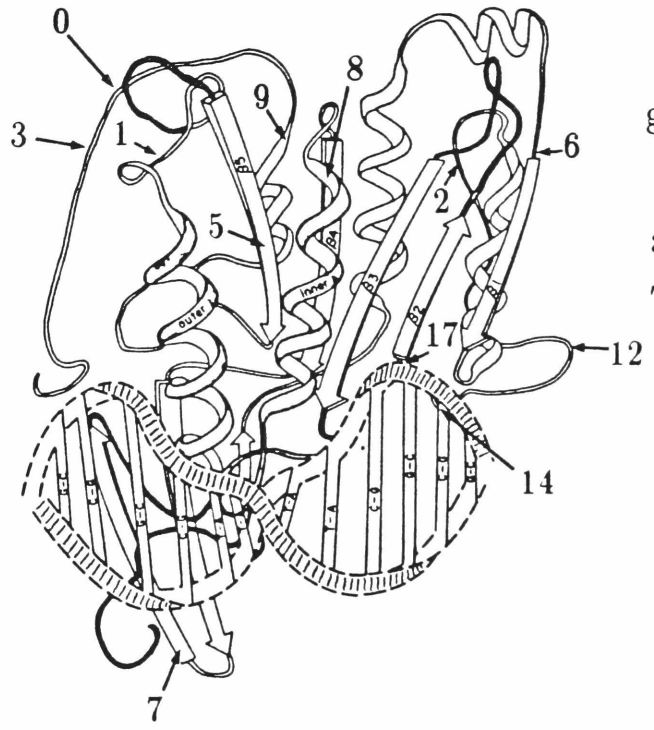
gly78asp
TS2

arg105lys
TS17

ala235glu
TS5

pro90ser
TS12

val97ile
TS14



gly186UGA
TS7

Table 4-1. Affect of *Eco*RI TS Mutants on Colony Growth of Strain K91

Allele	30°C	34°C	37°C	42°C
TS0	-	-	+++	++++
TS1	++	+++++	+++++	+++++
TS2	-	+++	+++++	+++++
TS3	++	+++++	+++++	+++++
TS5	-	+++++	+++++	+++++
TS6	-	+	+++	+++++
TS7	+++	+++++	+++++	+++++
TS8	++++	+++++	+++++	+++++
TS9	+	+++++	+++++	+++++
TS12	-	-	+	++
TS14	+	+	+++	++++
TS17	-	++++	+++++	+++++
TS0-RI(pJH12)	-	-	+++	++++
ΔRI(pJH71)	+++++	+++++	+++++	+++++
none	+++++	+++++	+++++	+++++

The term colony growth describes colony size, morphology, and efficiency of plating (E.O.P.).
Scoring is as follows:
+++++ = large, thick and round, EOP=0.5-1
++++ = medium-normal size colonies with ruffled edges EOP=0.1-0.5
+++ = medium size colonies that are thin and mottled EOP=0.01-0.1
++ = small colonies, very flat and mottled EOP=0.001-0.01
+ = tiny colonies, very flat and mottled EOP=0.001-0.01
- = no viable colonies EOP< 0.0001

come flat and translucent, and grow poorly, giving rise to faster-growing mutants. By colony forming efficiency, 0.1 - 1 % of cells survive to form a colony at this level of *in vivo* DNA scission. By microscopic examination, these cells are filamented. As described below, cellular filamentation occurs because DNA scission induces the SOS DNA repair response, and the product of one SOS induced gene, *sfiA*, inhibits cell division. At 30°C where the endonucleases are most active, some alleles are detrimental to cell growth but still allow colony formation (i.e. TS1, TS3) while others are fully lethal (i.e. TS0, TS2). The lethal effect was not attributable to restriction at the pJH10 *Eco*RI site and loss of the plasmid because a derivative lacking the *Eco*RI site was as lethal as its parent which carries one *Eco*RI site (see table 4-1, TS0-RI derivative). Expression of the *Eco*RI methylase from plasmid pJC1 completely blocked the lethal effects of these TS mutants.

λ *restriction by Eco*RI TS mutants

To determine the level of DNA scission that the cell can survive, we determined the activity of these TS endonucleases. Because the *Eco*RI methylase renders cells expressing a TS endonuclease viable, endonuclease activity can be assayed by simply measuring restriction of λ phage growth at several temperatures (30°C, 34°C, 37°C, and 42°C). As shown in Table 4-2, these mutants fall into two classes which were obtained with roughly equal frequency. Class I mutants (TS0, TS6, TS8, and TS9) exhibit increasing restriction activity with decreasing temperature of growth. For several Class I mutants (TS0, TS6), at the most permissive temperature (30°C) the mutant protein restricted λ growth to almost the same extent as cells producing the wild-type (WT) protein. In contrast, mutants of class II (TS1, TS2, TS3, TS5, and TS7) do not restrict λ at any temperature although several are fully lethal to the cell in the absence of the methylase. Thus, cell lethality is not strictly correlated with restriction of phage growth (see discussion). *Eco*RI modified λ plated with unit efficiency in all cases (data not shown).

Table 4-2. Restriction of λ_{vir} by *Eco*RI TS Mutants

Allele	30°C	34°C	37°C	42°C
TS0	7×10^{-5}	2×10^{-4}	2×10^{-4}	2×10^{-2}
TS1	1	1	1	1
TS2	1	1	1	1
TS3	1	1	1	1
TS5	1	1	1	1
TS6	4×10^{-4}	3×10^{-4}	6×10^{-3}	0.8
TS7	1	1	1	1
TS8	1×10^{-3}	3×10^{-2}	0.2	0.9
TS9	3×10^{-4}	3×10^{-2}	7×10^{-2}	0.5
WT	1×10^{-4}	4×10^{-5}	4×10^{-5}	4×10^{-5}
none	0.8	1.0	1.0	0.9

λ plating efficiency was measured on strain K91 (TS# and pJC1). Since pJC1 expresses the *Eco*RI methylase these strains are viable at all temperatures and allow one to measure the effects of the temperature sensitive restriction enzyme. Cells were grown at the indicated temperatures in λ broth supplemented with maltose, infected with λ, plated, and incubated overnight. The values listed above are the ratio of the phage titer on the strain carrying the *Eco*RI TS allele divided by the titer on K91 alone at 37°C. WT indicates the wild-type *Eco*RI restriction system K91 (pJH10 and pJC1) and none is strain K91 alone.

For several mutants substantial restriction activity is present at temperatures where the cell would remain viable without the methylase. For example, cells bearing the TS6 allele remain viable at 37°C even though they show reduced colony formation effiency (=0.01) and form smaller, flat, slowly growing colonies. When the methylase is present, cells bearing the TS6 mutant grow normally and restrict the growth of λ phage 200-fold. Since in general one observes 10-fold restriction per recognition site (Murray and Murray, 1974), 200-fold restriction reflects scission at roughly two sites out of the five *Eco*RI sites in λ. If the chromosome of the cell suffers a similar degree of scission, the cell must survive a substantial amount of DNA damage. Our earlier findings (Heitman, *et al.*, 1989; Chapter 5) suggest that they do so by actively repairing these DNA breaks. Because restriction of phage growth actually monitors the competition between the endonuclease and the methylase for phage-borne recognition sites, the level of restriction provides only a minimum estimate of the endonuclease action present in cells which lack the protective methylase. Furthermore, as discussed in more detail in Chapter 8, *Eco*RI endonuclease may not have equal access to λ and chromosomal DNA.

SOS induction by TS mutants

As will be described in more detail in Chapter 5, *in vivo* DNA scission by the *Eco*RI enzyme induces the SOS DNA repair response (see also Heitman, *et al.*, 1989). All of the TS alleles described here induce the SOS response, as monitored by increased β-galactosidase expression in the SOS::*lacZ* fusion strains JH137 or JH138 (described in Table 4-3). Although some TS mutants restrict λ (class I) and others do not (class II), under conditions where cell growth is impaired to similar extents mutants of both types induce comparable levels of β-galactosidase expression. Thus the cell lethality of these TS mutants correlates with their ability to induce the SOS response. Expression of the *Eco*RI methylase from the compatible plasmid pJC1 completely blocked SOS induction by these TS endonucleases.

Table 4-3. Colony Growth and SOS Induction by *Eco*RI TS Mutants in Strain JH138

Allele	30°C	34°C	37°C	42°C
TS0	-	-	+DB	+++MB
TS1	+DB	+++++LB	+++++LB	+++++LB
TS2	-	+++MB	+++++LB	+++++LB
TS3	-	+++MB	+++++LB	+++++LB
TS5	-	+++++LB	+++++LB	+++++LB
TS6	-	-	++DB	+++++LB-MB
TS7	-	+DB	+++DB	+++++LB
TS8	++DB	+++++LB	+++++LB	+++++LB
TS9	-	+++MB-DB	+++++LB	+++++LB
TS12	-	-	+DB	++MB-DB
TS14	-	-	+DB	++++MB
TS17	-	++MB-DB	++++LB-MB	+++++LB
ΔRI(pJH71)	+++++LB	+++++LB	+++++LB	+++++LB

The term colony growth describes colony size, morphology, and efficiency of plating (E.O.P.).
Scoring is as follows:
+++++ = large, thick and round, EOP=0.5-1
++++ = medium-normal size colonies with ruffled edges EOP=0.1-0.5
+++ = medium size colonies that are thin and mottled EOP=0.01-0.1
++ = small colonies, very flat and mottled EOP=0.001-0.01
+ = tiny colonies, very flat and mottled EOP=0.001-0.01
- = no viable colonies EOP< 0.0001
Colony color was determined on fresh (< 3 days old) indicator
medium containing 35 μg / ml X-gal.
LB=light blue colonies; MB=medium blue colonies; DB=dark blue colonies.
The background level of β-galactosidase expression in the
dinD1::Mu dI(Ap^r *lac*) strain JH138 is slightly higher
than in strain JH137 due to the *rep71* allele.

Site-directed mutants confer a TS phenotype

As part of an analysis of the *Eco*RI substrate binding we employed site-directed mutagenesis to isolate all 20 possible amino acids at position 200 (WT=arginine), 17 changes at position 144 (WT=glutamic acid) and 13 at position 145 (WT=arginine) (see Chapter 6 and methods for isolation procedure). These three amino acids lie at the ends of two α-helices which project into the major groove of the DNA and hydrogen bond to the nucleotides of the recognition site (McClarin, *et al.*, 1986). As will be described in detail elsewhere (see Chapter 6), we find that three mutants (arg200lys, glu144asp, and glu144cys) are lethal in the absence of the methylase in strain JH137. All three exhibit a TS phenotype in an HB101 derived strain (JH69) which is less sensitive to DNA scission than JH137. Several additional mutations at each position (arg200cys, ser or val; glu144ser or gly; arg145lys or cys) are not lethal to the cell but nonetheless confer a TS endonuclease phenotype as monitored by their ability to induce the SOS response and increase expression in the SOS::*lacZ* fusion strain JH138. All other changes at position 200 and many others at positions 144 and 145 confer a null phenotype. We conclude that mutations specifically engineered at points where the protein secondary structure is in transition also confer a TS phenotype.

Enzymatic activity of TS alleles

To determine if the *in vivo* actions of these mutants (cell lethality, SOS induction, phage restriction) correlate with activity, DNA cleavage was measured *in vivo* and *in vitro*. *In vivo* DNA scission will be described more fully in Chapter 5, but briefly: when cells containing plasmid pJH10 bearing the TS0, TS3, TS6, or TS9 allele, the plasmid is cleaved at its *Eco*RI site upon shift to temperatures permissive for endonuclease action (Heitman, *et al.*, 1989). In contrast, with the arg200lys mutant, we observed little or no plasmid cleavage *in vivo* (data not shown).

To determine activity *in vitro*, cellular extracts were prepared under conditions where the mutant proteins should have maximal activity (see Chapter 2, materials

and methods). DNA cleavage was measured at 30°C, the temperature at which the mutant endonuclease are most active *in vivo*. With plasmid pUC18 (or λ) DNA as substrate, *Eco*RI specific cleavage activity could be readily demonstrated for those mutants which restrict λ *in vivo* (class I). For example, figure 4-3 shows that extracts from cells expressing the TS0 mutant yield activity comparable to cells expressing the WT protein, while the class II mutants TS3 and TS8 exhibit greatly reduced activity (5% and 1-2% respectively). In fact for four of the five class II mutants which are unable to restrict λ, little or no endonuclease (<1% to 2% of WT) was detectable. The only exception was TS3, which showed ~ 5% of WT levels. Thus the level of endonuclease activity *in vitro* is correlated with the amount of λ restriction *in vivo*. None of the TS mutants were more active under *Eco*RI* buffer conditions (data not shown).

To determine if these mutant proteins exhibit a TS phenotype *in vitro*, extracts of the wild-type and several mutant enzymes were incubated for 1 hour at 42°C prior to DNA cleavage assays. No marked decrease in activity was detectable in subsequent cleavage reactions conducted at either 30°C or 42°C (to prevent any renaturation). Similar findings were observed with the purified wild-type and TS6 mutant enzymes. We conclude that the TS phenotype of these mutants is only manifested *in vivo* and suggest that the mutations may therefore affect folding of the protein to its stable tertiary structure.

Discussion

We have described a set of TS *Eco*RI endonuclease mutants. These mutants fall into two phenotypic classes: ones which restrict phage and exhibit endonuclease activity *in vitro* (class I), and others which do not restrict phage and have little or no *in vitro* activity. The cell lethal affect of these mutants correlates with the extent of SOS induction at intermediate temperatures but not with their ability to restrict λ or cleave DNA *in vitro*. Cell lethality is monitored in the absence of the methylase,

Figure 4-3. *In vitro* DNA cleavage by TS *Eco*RI mutants. Cell extracts were prepared from strains expressing an *Eco*RI TS mutant and the *Eco*RI methylase grown at 30°C. Extracts were incubated with 250 ng of RFI form pUC18 DNA for 20 minutes at 30°C and the reaction products were displayed by electrophoresis on a 0.6 % agarose gel containing 0.5 μg/ml EtBr. Lane a: untreated plasmid DNA; lane b: commercial *Eco*RI; lanes c, f, i, l: 1 μl of the indicated extract; lanes d, g, j, m: 1 μl of a 1:10 extract dilution; lanes e, h, k, n: 1 μl of a 1:100 extract dilution.

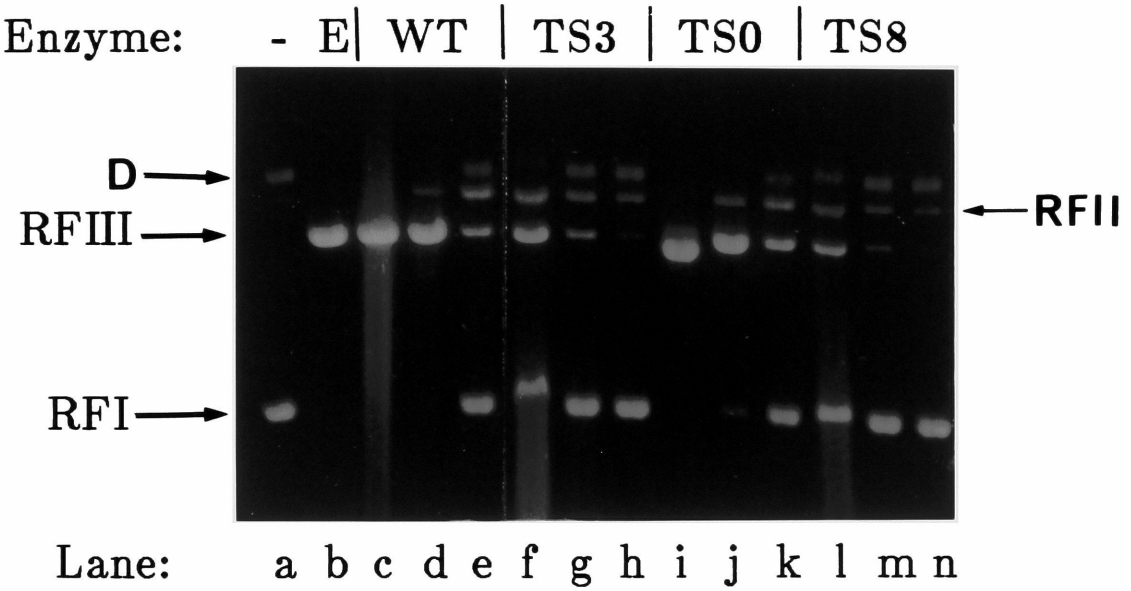

Enzyme: - E| WT | TS3 | TS0 | TS8
D
RFIII
RFI
RFII
Lane: a b c d e f g h i j k l m n

Table 4-4. Properties of *Eco*RI TS Mutants

Allele	Mutation	Sequence Change	Mutagen	lethal[a]	restricts[b]	active[c]
WT	-	-	-	yes	yes	100 %[d]
TS0	T261I	ACT...ATT	NG	yes	yes	100 %
TS1	C218Y	TGT...TAT	spon	no	no	1 - 2 %
TS2,TS10	G78D	GGC...GAC	spon,NG	yes	no	< 1 %
TS3,TS16	L263F	CTC...TTC	spon,NG	no	no	5 %
TS5	A235E	GCA...GAA	spon	yes	no	< 1 %
TS6	R56Q	CGA...CAA	spon	yes	yes	10 %[d]
TS7	G186UGA	GGA...TGA	spon	no	no	< 1 %
TS8	M157I	ATG...ATA	spon	no	yes	1 -2 %
TS9	M255I	ATG...ATT	spon	no	yes	10 %
TS12	P90S	CCT...TCT	NG	yes	ND	ND
TS14	V97I,Y165F	GTC.ATC,TAC.TTC	NG	no	ND	ND
TS17	R105K	AGA...AAA	NG	yes	ND	ND
SD6	R200S	AGG...AGC	SD	no	no	-
SD27	R200V	GTC(GTG)	SD	no	no	-[d]
SD39	R200K	AAG	SD	yes	no	-
SD64	R200C	TGC(TTG)	SD	no	no	-[d]
SD122	E144D	GAA...GAC	SD	yes	no	ND
SD131	E144G	GGC(GGG)	SD	no	no	ND
SD167	E144S	TCG	SD	no	no	ND
SD175	E144C	TGC	SD	yes	no	ND
SD195	R145K	AGA...AAG	SD	no	no	ND
SD204	R145C	TGC	SD	no	no	ND
SD207	R145UAG	TAG	SD	no	no	ND
SD227	R145K	AAG	SD	no	no	ND

[a] Affect of TS endonucleases expression in the absence of the methylase on colony growth of stain K91 (TS alleles) or the K91 related strain JH137 (site-directed alleles).

[b] As described in results and Table 4-2, strains expressing mutant endonucleases and the *Eco*RI methylase were tested for restriction of λ phage growth.

[c] DNA cleavage activity by cell extracts was measured *in vitro* with plasmid pUC18 or λ phage DNA as substrate.

[d] These mutant proteins have been purified to homogeneity and their cleavage activity determined *in vitro*. See results for a more detailed description.

while the measurement of phage restriction requires the presence of the methylase to protect the host cell. Thus λ restriction represents a competition between the endonuclease and methylase. If the class II mutants cleave DNA at a much reduced rate compared to the WT enzyme, the phage may be fully modified by the *Eco*RI methylase before the mutant endonuclease has time to act. In contrast, cell lethality was measured in the absence of the methylase. In this case, the chromosome of the cell is always unprotected and even a slow endonuclease may have ample time to kill the cell.

DNA sequencing revealed that the amino acid alterations responsible for the TS phenotype range from minor (i. e. valine to isoleucine) to drastic (i. e. alanine to glutamic acid) amino acid substitutions which lie at different positions within the protein primary structure. By comparison with the known three dimensional structure of the *Eco*RI enzyme (McClarin, *et al.*, 1986), these mutations in general lie at points where the secondary structure is in transition: the ends and beginnings of α-helices and β-sheets, loops, and turns. The only exception is the TS5 mutant, which bears an ala to glu substitution within a β-sheet. These findings are not attributable to mutagen specificity or mutagenic hotspots for the following reasons. We have isolated spontaneous point mutants which should arise from a variety of sources, and we rarely found the same allele twice. Furthermore, the significance of this pattern of distribution is supported by the similar phenotypes of a set of site-directed mutations that lie at the ends of two α-helices. In each of these cases, those amino acid substitutions that retained enzyme activity also rendered the protein temperature sensitive. We conclude that amino acid substitutions in regions of secondary structure transition frequently render the *Eco*RI endonuclease thermosensitive. For TS mutations lying at the ends of α-helices, we suggest that they may perturb the protein structure by disrupting those residues which mark the helix boundaries (Richardson & Richardson, 1988; Presta & Rose, 1988). Not all mutations at regions of secondary structure transition will render the protein temperature sensitive, since two mutants which exhibit

enhanced *Eco*RI* activity (Y193H and E192K; described in Chapter 7) carry substitutions at the end of a β-sheet but are not TS.

These findings raise the question of whether substitutions within helices and sheets would confer a null, silent, or reduced activity but not a temperature-sensitive phenotype. We have isolated two null mutants not described here. One is a substitution within a β-sheet (pro164leu) and the other is a nonsense mutation (Q233UAA). A set of null and partial activity mutants has been described by Yanofsky *et al.* (1987). They observed that null and decreased activity mutations often fall within helices and sheets. However, they did not check their mutants for a TS phenotype, and one of their partial activity mutations (arg 56 gln) is in fact a TS mutant (=TS6 described here).

To determine the *in vitro* activity which corresponds to the *in vivo* action of these TS *Eco*RI mutants, cellular extracts were prepared from cells expressing both the mutant endonuclease and the *Eco*RI methylase. These partially purified mutant enzymes cleave DNA *in vitro* at *Eco*RI sites with reduced activity compared to the wild-type protein. However, the activity of those mutants tested (TS0, TS6, and TS9) was not temperature sensitive *in vitro*. We suggest that these mutations confer a TS phenotype *in vivo* because they disrupt folding of the protein. Once synthesis and folding are complete, the mature protein is not sensitive to thermal denaturation *in vitro*.

In contrast to our findings, earlier studies of TS T4 phage lysozyme mutants found that many TS mutations often fell within α-helices and β-sheets (Grutter, *et al.*, 1983; Alber, *et al.*, 1987a, 1987b). These mutations render function TS *in vivo*, and markedly decrease the thermal stability of the folded protein *in vitro*. Furthermore the responsible amino acid substitutions fall at many positions within the primary and secondary structure of the protein but in general alter residues with low thermal mobility and solvent accessibility. Apparently these internal residues form interactions which restrict their movement and contribute to the thermal stability of

the native protein. TS mutants of another protein of known tertiary structure, gene V protein of bacteriophage f1, also fall both at many positions within the primary sequence, but their biochemical nature is unknown (Terwilliger, 1988). In contrast, TS mutants of several proteins of unknown tertiary structure (the Cdc28 protein of yeast (Lorincz & Reed (1986), *lac* repressor (Miller, 1979), and phage P22 tailspike protein (King, *et al.*, 1987; Fane & King, 1987)) often lie within regions where the predicted secondary structures are in transition. In the case of phage P22 tailspike protein, these TS mutations are often defective for protein folding. We find that TS mutations of *Eco*RI also lie at regions of secondary structure transition and that these mutant proteins are TS *in vivo* but not *in vitro.*

Our findings suggest that TS mutations that alter protein folding result from a different pattern of amino acid substitutions than TS mutations which render the native protein structure less stable. However, why do proteins differ with the frequency at which the two types of TS mutations occur? The explanation must be some feature of protein structure or function. Proteins can have either rigid or flexible structures. For example, proteins which undergo conformational changes require a flexible structure that can assume more than one final tertiary configuration. These proteins require hinges that permit their structures to interchange. Because secondary structural elements are rigid, protein loops, bends and turns are the most likely points of flexibility. If these structures mediate both folding and flexibility, they also have to exist in more than one conformation and therefore may be more sensitive to mutations. This predicts that proteins which undergo conformational changes may be more susceptible to a different pattern of TS mutations than enzymes or proteins which have more rigid tertiary configurations.

Our collection of TS mutants also bears on the structure and function of the *Eco*RI enzyme itself. For example, the R105K mutations lies at the DNA-protein interface and may alter an electrostatic interaction to the DNA backbone. This residue also lies very near the scissile phosphodiester bond and may form part of the enzyme

active site. The R56Q and G78D mutations lie in a region of the protein thought to allosterically couple binding with cleavage, based on mutations which separate binding from cutting (E111G, A, V, Q) (King, *et al.*, 1986, 1987) or allow promiscuous cleavage (H114Y) (see Chapter 7). Several mutations lie at the subunit-subunit interface (at least T261I and Y263L) and may affect dimerization or communication between subunits.

In addition to their implications for protein structure and function in general and the action of the *Eco*RI endonuclease in particular, this collection of TS *Eco*RI endonuclease mutants has proven useful to show that DNA double-strand breaks induce the *E. coli* SOS DNA repair response (see Chapter 5 and Heitman, *et al.*, 1989), to analyze the mechanism of double-strand break repair in *E. coli* (Heitman, *et al.*, 1989; Chapter 5), and to conditionally express *Eco*RI mutants which cleave DNA sites not modified by the *Eco*RI methylase (Chapter 7). These mutants may prove useful for studying DNA structure, chromosomal architecture, and DNA repair, both in *E. coli* and in other organisms.

Chapter 5

Repair of the *E. coli* Chromosome After *In Vivo* Scission by the *Eco*RI Endonuclease

Introduction

This chapter is adapted from Heitman, Zinder, & Model (1989).

Many bacteria make restriction and modification enzymes to degrade foreign DNA that enters the cell (Roberts, 1976). Generally, methylation protects the cellular DNA from cleavage, whereas unmethylated foreign DNA is destroyed. Two observations suggest that other mechanisms, in addition to methylation, ensure the integrity of the cellular DNA. First, *E. coli* can be transformed at normal efficiency by plasmids encoding a restriction-modification system. After DNA uptake, the cell must prevent or survive DNA damage until the chromosome is protected by methylation. Second, several restriction endonucleases have been cloned without their cognate methylases, and extracts from these strains contain endonuclease activity (Gingeras & Brooks, 1983; reviewed in Lunnen, *et al.*, 1988). Although it has been suggested that gene regulation must delay or reduce endonuclease expression to protect the cell in these situations (Mann, *et al.*, 1978; Walder, *et al.*, 1984; Bouguelerat, *et al.*, 1984; Kiss, *et al.*, 1985), this has been demonstrated only for the phage-borne P1 restriction system (Arber, 1974). Alternatively, the cell could repair DNA double-strand breaks.

DNA double-strand breaks stimulate recombination in *Escherichia coli* and yeast, and it is generally thought that recombination repairs DNA breaks (Thaler & Stahl, 1988; Szostak, *et al.*, 1983; Resnick, 1976). The repair of DNA scissions has been studied previously using γ-ray lesions (Krasin & Hutchinson, 1977, 1981). In *E. coli*, prior induction of the SOS DNA repair response enhances the repair of γ-ray cleaved DNA (Krasin & Hutchinson, 1981). Two SOS proteins involved in recombination, RecA and RecN, and multiple copies of the genome are required to repair DNA bro-

ken by γ-rays, suggesting that recombination repairs these lesions (Krasin & Hutchinson, 1977; Picksley, *et al.*, 1984). Ligation may also repair some γ-ray breaks (Weibezahn & Coquerelle, 1981).

A model for DNA scission repair based on γ-ray lesions is compromised since these lesions include breaks, nicks, base adducts and oxidized proteins (Friedberg, 1985). Furthermore, γ-rays (and the radiomimetic bleomycin) break DNA by shattering the sugar residues, releasing the attached base and leaving behind blunt termini with 5'-phosphoryl and 3'-phosphoryl or 3'-phosphoglycolate ends which cannot be ligated (Henner, *et al.*, 1983; Giloni, *et al.*, 1981). Even if these ends were repaired by phosphatase or exonuclease processing to a 3'-hydroxyl and subsequent ligation, at least one basepair of information would be lost.

In contrast, restriction endonucleases cleave DNA to yield staggered or blunt double-strand breaks with 3'-hydroxyls and 5'-phosphates as termini (Roberts, 1976). To address the repair of this type of DNA scission, we have delivered staggered double-stranded DNA breaks with the *Eco*RI endonuclease, which cleaves within the DNA sequence GAATTC (Hedgpeth, *et al.*, 1972). DNA double-strand breaks (including *Eco*RI breaks) clearly stimulate recombination by several pathways (M. Stahl, *et al.*, 1983; F. Stahl, *et al.*, 1986; Thaler, *et al.*, 1987a, 1987b). We find that when the *E. coli* chromosome suffers scission by the *Eco*RI endonuclease, the SOS DNA repair response is induced. However, strains blocked in SOS induction (*lexA3* mutant) or SOS induction and recombination (*recA56* and *recB21* mutants) are not more sensitive to DNA scission, while strains deficient in DNA ligase (*lig4*, *lig ts7* mutants) are 1000-fold more sensitive. We suggest that neither SOS induction nor recombination is required to repair *Eco*RI severed DNA. DNA ligase is required and may suffice to repair *Eco*RI mediated cleavage of the chromosome. Thus, although recombination is stimulated by DNA breaks, it need not be the means by which they are repaired. These findings further suggest that restriction-modification systems may not only inhibit but also promote genetic exchange (Arber, 1979), perhaps by

stimulating recombination in ways similar to *in vitro* cloning manipulations (Chang & Cohen, 1977).

Results

Isolation of EcoRI temperature sensitive alleles

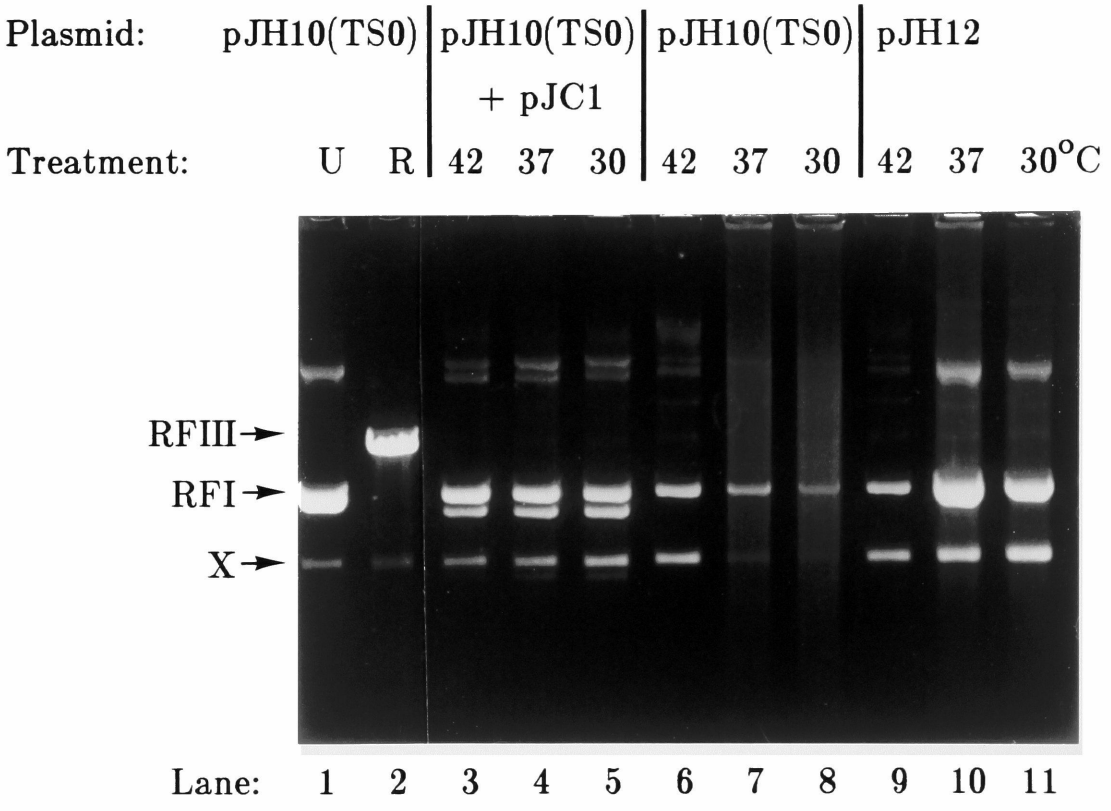
As described in Chapter 4, we isolated a series of temperature- sensitive conditional mutants of the *Eco*RI endonuclease. We note that the endonuclease activity of these mutants is temperature (heat) sensitive and, because they kill the cell, the phenotype of their host strain is cold-sensitive growth. Thus strains bearing these alleles form colonies at 42°C (permissive temperature for cell growth) but grow poorly and in some cases die at lower temperatures non-permissive for growth (see Chapter 4, table 4-1). Expression of the *Eco*RI methylase from the compatible plasmid pJC1 completely blocks the lethal endonuclease activity. Many of these TS *Eco*RI mutants restrict growth of λ phage when the methylase is present to ensure cell viability (see Chapter 4 table 4-2).

In Chapter 4 we concentrated on the implications of these TS *Eco*RI mutants for protein structure and *Eco*RI endonuclease function. Here we employ these mutants to deliver DNA double-strand breaks to the *E. coli* chromosome and study their mechanism of repair. We begin by considering the enzymatic action of the TS mutants.

In vivo DNA cleavage by the TS0 allele

The *Eco*RI TS mutants were isolated in plasmid pJH10, which carries one *Eco*RI recognition site. Thus plasmid pJH10 should be cleaved in cells bearing a TS *Eco*RI allele after shift to temperatures permissive for endonuclease action. As shown in Figure 5-1, when cells expressing the TS0 allele (T261I) were shifted from growth at 42°C to 37°C or 30°C, plasmid pJH10(TS0) DNA was degraded (lanes 7 and 8). However, the linear cleavage product was not seen even when DNA was prepared by

Figure 5-1. *In vivo* DNA scission by the *Eco*RI endonuclease. Strains JH11(pJH10(TS0) and pJC1), JH11(pJH10(TS0)), and JH11(pJH12) were grown at 42°C until OD_{600}=0.2, the culture was divided and portions were grown at 42°C, 37°C, and 30°C for 3 hours. Strains harboring both plasmid pJH10(TS0) and pJH12 (TSΔRI site) form normal colonies at 42°C, poorly growing colonies at 37°C, and die at 30°C. DNA was prepared by the alkaline lysis method (Maniatis, *et al.*, 1982). Lanes 1 and 2 are controls showing uncut (U) and *Eco*RI cut (R) pJH10(TS0). DNA from 3 to 5-fold more cells was loaded in lanes 7, 8, 10, and 11 (non-permissive temperatures) compared to lanes 6 and 9 (permissive temperature). The band labelled X often appears during the preparation of DNA by alkaline lysis and may be improperly renatured ("collapsed") plasmid. Plasmid pJH10 and pJH12 have one and no *Eco*RI sites respectively.


Plasmid:
pJH10(TS0)
pJH10(TS0) + pJC1
pJH10(TS0)
pJH12
Treatment:
U
R
42
37
30
42
37
30
42
37
30°C
RFIII
RFI
X
Lane:
1
2
3
4
5
6
7
8
9
10
11

a non-denaturing procedure (Zinder & Boeke, 1982) or from a *recB*$^-$ host which lacks exonuclease V, the major endonuclease of *E. coli* (data not shown). The cleavage and subsequent degradation occured *in vivo*, since plasmid pBR322 DNA added exogenously during plasmid preparation did not suffer degradation (data not shown). Chromosomal DNA was also degraded as DNA cleavage products larger than the linear plasmid were observed when DNA was prepared by either alkaline lysis (Fig. 5-1, lanes 7, 8, 10, and 11) or non-denaturing methods (data not shown). In contrast, a TS0 derivative lacking the *Eco*RI site (pJH12) remained supercoiled and intact (lanes 10 and 11) and likewise, expression of the *Eco*RI methylase from plasmid pJC1 conferred resistance to this *in vivo* degradation (lanes 4 and 5). Similar results were obtained with alleles TS3 (L263F), TS6 (R56Q), and TS9 (M255I) or with a filamentous phage (J9-9, an M13mp18 derivative) carrying an *Eco*RI recognition site and the *Eco*RI endonuclease gene expressed from the *lac* promoter (described in Appendix 1). In this latter case, both the RFI form of the phage and the chromosome were degraded when endonuclease expression was induced with IPTG after infection. This cleavage was not observed with a derivative lacking the *Eco*RI site or when cells expressing the *Eco*RI methylase were infected (data not shown). We conclude that DNA scission occured *in vivo* at *Eco*RI sites, and that the linear DNA molecules were degraded.

As a further measure of the cleavage activity of these mutant proteins, cell extracts were prepared under conditions where endonuclease activity should be maximal (see Chapter 2, Materials and Methods) and DNA cleavage of λ and pUC18 DNA was assayed *in vitro* (as shown in Chapter 4, figure 4-3). Cells expressing the TS0 mutant protein yielded *Eco*RI specific endonuclease with activity comparable to cells expressing the wild-type protein, while extracts containing the TS3, TS6, or TS9 mutant proteins showed decreased activity compared to wild-type (5, 10 and 10% activity respectively, data not shown). Furthermore, when purified to homogeneity the wild-type and TS6 mutant proteins did not differ in specific activity (2 - 4 X 10^6

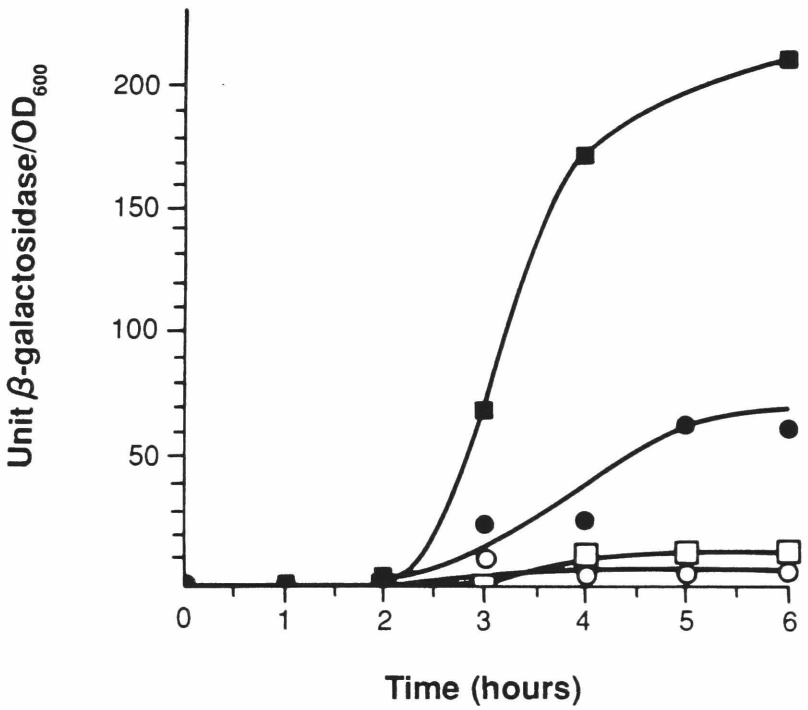
units/mg). We conclude that these *Eco*RI mutants display DNA scission activity *in vivo* and *in vitro* and proceed to an analysis of the *in vivo* consequences of this type of DNA lesion.

In vivo DNA scission induces the SOS response

Our earlier finding that Mrr or McrB restriction induces the SOS response suggested that DNA scission might, in general, induce SOS (Chapter 3; Heitman & Model, 1987). To test this, plasmid pJH10(TS0) was introduced into strain JH39, which carries the *lac* operon fused to the DNA damage inducible locus *dinD* (Kenyon & Walker, 1980). As shown in Figure 5-2, when strain JH39/TS0 was shifted from 42°C to 37°C, the β-galactosidase expression from the *dinD1*::Mu dI(Ap^r *lac*) fusion increased 6 to 8-fold above background. This induction was completely blocked by plasmid pJC1 which expresses the *Eco*RI methylase, since strain JH39/TS0 and pJC1 formed white colonies on X-Gal indicator medium at all temperatures. Temperature shifts of strain JH39 alone did not affect β-galactosidase levels. When strain JH39/TS0 was shifted from 42°C to 30°C, growth ceased rapidly with no increase in β-galactosidase expression. Apparently under these conditions the cell is killed before the SOS response is induced. Induction by DNA scission was observed with all the TS mutants. A plasmid encoding the wild-type endonuclease and a temperature sensitive *Eco*RI methylase (pJH34) gave similar results. In this case, strains harboring this plasmid are viable at 30°C but grow poorly at 34°C and die at 37°C or 42°C. Similarly, strain JH39/pJH34 made increased levels of β-galactosidase at 30°C (medium blue colonies on X-gal medium) and 34°C (dark blue, poorly growing colonies), indicative of SOS induction. Thus SOS induction reflects endonuclease activity and is not a special property of the TS endonucleases.

In control experiments, a dose of mitomycin C of 1 μg / ml induced the SOS response 20- to 22-fold above background, compared to the 6- to 8-fold induction with DNA scission by the *Eco*RI mutants. This may result from a difference in the

Figure 5-2. Kinetics of β-galactosidase induction in strain JH39 (*dinD1::*Mu dI(Ap^r *lac*)) and JH39(pJH10(TS0)). Cells were grown at 42°C in K120 minimal medium supplemented with 0.2% glucose and 0.4% Casamino Acids (Heitman & Model, 1987). Cultures were shifted from 42°C to 37°C or Mitomycin C was added to 1 μg/ml at T=0.5 hr. Samples (200 μl) were removed and β-galactosidase activity was assayed as described by Miller (1972). Symbols: O, JH39(pJH10(TS0)) at 42°C; ●, JH39(pJH10(TS0)) at 37°C;□, JH39 at 37°C;■, JH39 plus 1 μg/ml of Mitomycin C.


Unit β-galactosidase/OD600
200
150
100
50
0
1
2
3
4
5
6
Time (hours)

amount or in the nature of the inducing lesion. Thus we note that DNA scission by both the *Eco*RI or the Mrr restriction system (Heitman & Model, 1987) activates the SOS response to similar levels.

SOS induction is not required for DNA scission repair

We next asked whether the repair of DNA scission requires SOS induction. We tested SOS induction after DNA scission by the TS0, TS3, TS6, and TS9 alleles in *dinD1::lacZ* fusion strains carrying mutations known to block SOS induction by some (*recB21*, strain JH117) or all agents (*recA56*, strain JH59; *lexA3*, strain JH154). In contrast to the parent strain JH39, which on X-gal medium yielded blue colonies indicative of SOS induction at temperatures where DNA scission occured, the *recA56*, *lexA3*, and *recB21* mutant strains produced white colonies and thus do not induce the SOS response after DNA scission.

We also tested the sensitivity of mutant strains to DNA cleavage by the TS3 (L263F), TS6 (R56Q), or TS9 (M255I) *Eco*RI allele, expressed from plasmid pJH15b. Two isogenic strain series were constructed. One carries the Mu d fusion and enabled us to also assay SOS induction as described above (JH39 (wt), JH59 (*recA56*), JH117 (*recB21 thyA::*Tn*10*), JH154 (*recN*), and JH158 (*lexA3*). The other lacks any Mu d lysogen (K91 (wt), JH27 (*recA56*), JH20 (*lexA3*), JH145 (*recB21*), and JH144 (*recN262*)) because the Mu Gam protein can bind to and protect DNA ends (Akroyd & Symonds, 1982; Thaler, *et al.*, 1987c). In practice, we observed no protective effect of the Mu d lysogen when isogenic Mu d^+ and Mu d^- strains were subjected to DNA scission (see Chapter 8, figure 8-1). The *recN262* strain was included since the RecN protein plays a role in the repair of γ-ray induced DS-DNA breaks (Picksley, *et al.*, 1984). Sensitivity to DNA scission was measured qualitatively by determining colony forming efficiency and growth at 42°C, 37°C, 34°C, and 30°C (as in Table 4-1, Chapter 4). By this assay, the repair defective strains were no more sensitive to DNA scission than their isogenic wild-type parent.

Alternatively, sensitivity to DNA breaks was measured more quantitatively by determining colony forming efficiency at the permissive temperature (42°C) following incubation for increasing time at the non-permissive temperature (30°C) (as described in Fig. 5-3). Sensitivity was measured over a range of doses at which 10^{-4} to more than 50% of wild-type cells survive. As shown in Figure 5-3, the isogenic *recA*$^+$ and *recA*$^-$ strains JH39 and JH59 do not differ in sensitivity to DNA scission by the TS6 allele. Similar affects were observed with logarithmic and stationary phase cells. Similarly, Figure 5-4 shows that other SOS, recombination, or repair defective mutants (*recA* (JH27), *recB* (JH145), *recN* (JH144), and *lexA* (JH20)) were also no more sensitive to DNA scission than their isogenic wild-type parent (K91). Thus by both methods, DNA scission sensitivity is not increased by mutations that prevent SOS induction or recombination. SOS induction and cell lethality did not require a plasmid borne *Eco*RI site, as derivatives lacking the *Eco*RI site yielded similar results (data not shown).

DNA ligase mutants are sensitive to DNA scission

These findings suggested that *Eco*RI scissions are repaired by mechanisms other than SOS and recombination. To address this possibility, we tested strains deficient in DNA ligase activity for sensitivity to DNA scission. In this case, the host strains carry temperature-sensitive ligase alleles (*lig4* and *lig ts7*) (Gottesman, *et al.*, 1973; Konrad, *et al.*, 1973), so a plasmid encoding an *Eco*RI allele with weak activity at all temperatures (RA2=gly19 to asp) served to deliver DNA breaks. The *lig4* strain N1626 carrying the plasmid pJH15b(RA2) grew normally at 30°C. However, when incubated for increasing time at 42°C where the activity of the mutant DNA ligase is greatly reduced, the colony forming efficiency of strain N1626/RA2 was decreased up to 4000-fold compared to the *lig*$^+$ strain N2603/RA2, N1626 bearing the control plasmid pJH71, or N1626/RA2 also carrying the methylase plasmid pJC1 (see Fig. 5-5). When we tried to introduce the RA2 allele into a strain carrying the more severe

Figure 5-3. Sensitivity of Isogenic *recA*$^+$ and *recA*$^-$ Strains to DNA Scission. Cultures were grown at 42°C to stationary or logarithmic phase, serially diluted, and 20 μl portions were spotted on solid medium. Colony formation was challenged by prior incubation for increasing time at the non-permissive temperature for colony formation (30°C), followed by growth at the permissive temperature (42°C). The number of colonies was determined and the fraction surviving was calculated. Symbols: JH39/TS6 (*recA*$^+$) stationary phase (○); JH59/TS6 (*recA*$^-$) stationary phase (□); JH39/TS6 (*recA*$^+$) logarithmic phase (●); JH59/TS6 (*recA*$^-$) logarithmic phase (■).

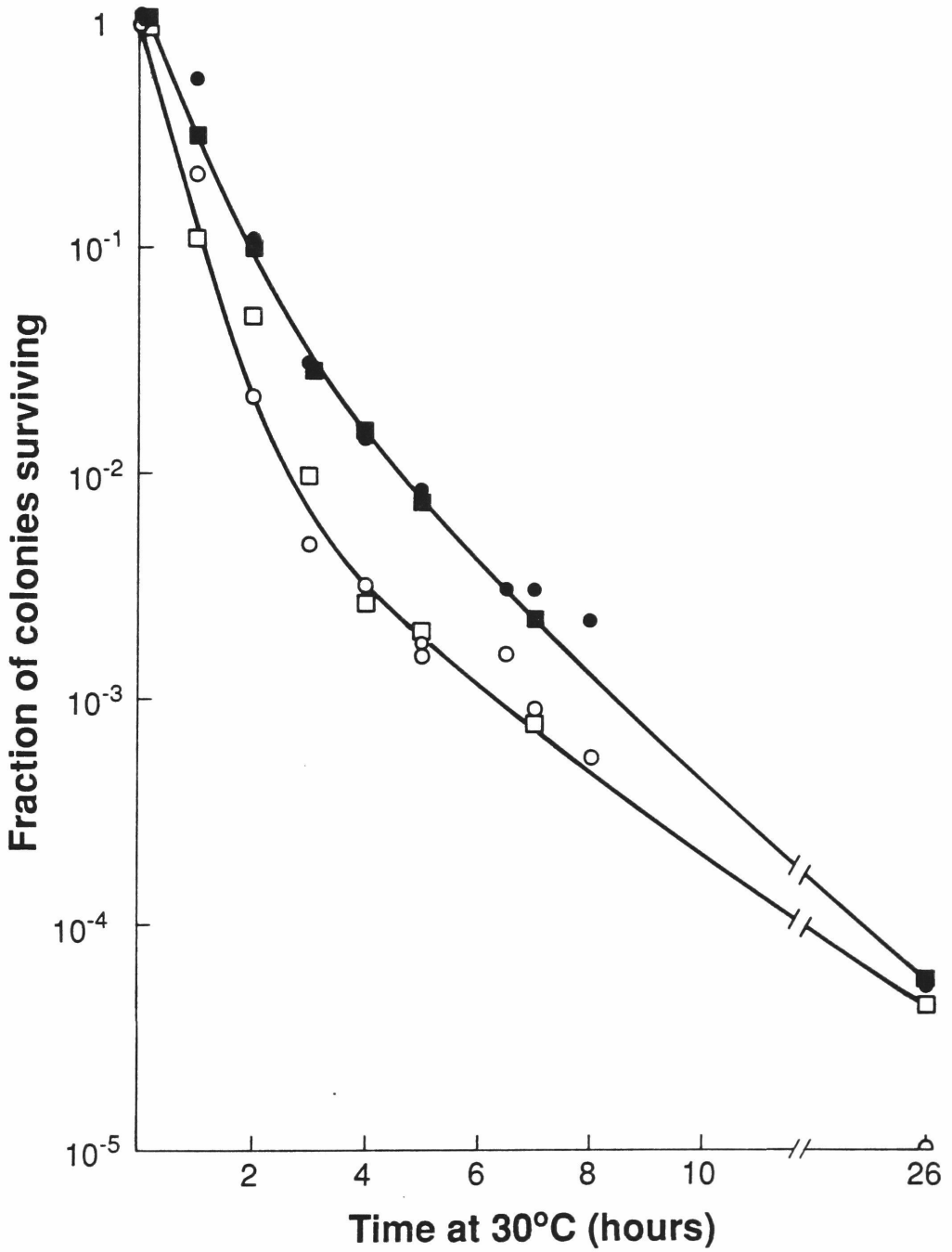
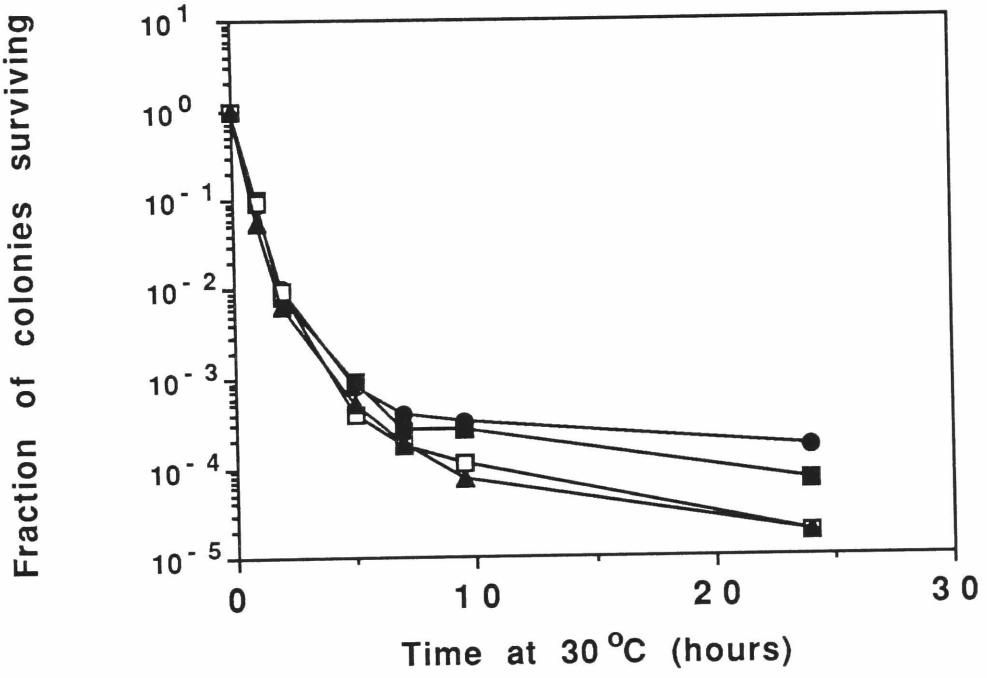

Fraction of colonies surviving
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
2
4
6
8
10
26
Time at 30°C (hours)

Figure 5-4. Sensitivity of Repair Defective Strains to DNA scission. Sensitivity to DNA scission was measured as in figure 5-3. Symbols: Strain K91/TS6 (wild-type) (●); K902/TS6 (*recA*$^-$) (■); JH20/TS6 (*lexA3*) (□); JH145/TS6 (*recB21*) (▲).


Fraction of colonies surviving
10 1
10 0
10 - 1
10 - 2
10 - 3
10 - 4
10 - 5
0
10
20
30
Time at 30 °C (hours)

lig ts7 allele (N2604), the transformation efficiency was reduced 1600- to 2400-fold compared to the control plasmid pJH71 or 1100-fold compared to transformation of the isogenic *lig*$^+$ strain N2603. Again, expression of the *Eco*RI methylase reversed the lethal effect of the RA2 allele. Similar results were obtained with a derivative (RA2ΔRI, pJH124) which lacks the plasmid borne *Eco*RI site. *RecA*$^+$ and *recA*$^-$ strains were transformed at equal efficiency by plamsid pJH15b (RA2) (data not shown).

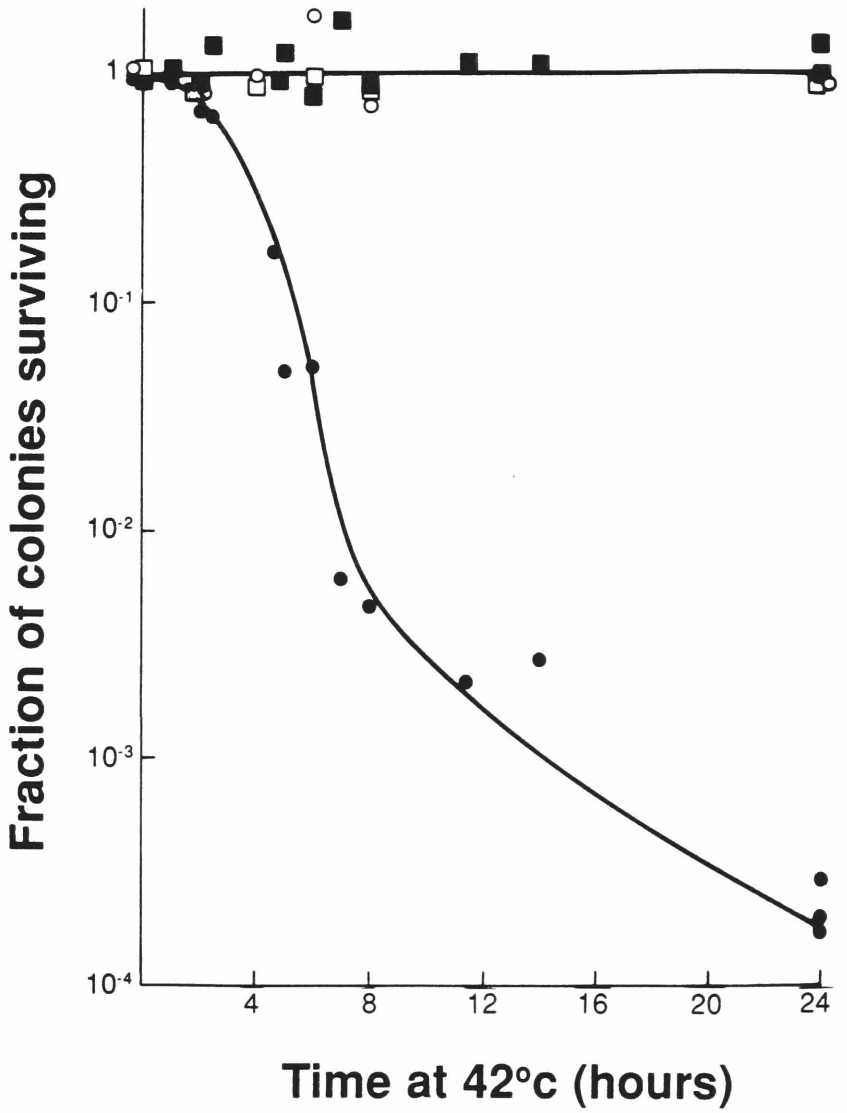
Discussion

We have isolated a set of temperature sensitive conditional alleles of the *Eco*RI endonuclease. Strains harboring these mutants are viable at 42°C, but grow poorly or die at lower temperatures. Under semi-permissive conditions, intracellular substrates with *Eco*RI recognition sites suffer cleavage, while those lacking sites or bearing modified *Eco*RI sites are spared. This *in vivo* DNA scission induces the SOS DNA repair response. Expression of the *Eco*RI methylase prevents both cell killing and SOS induction by these alleles. Neither cell killing nor SOS induction requires a plasmid borne *Eco*RI site. We conclude that *in vivo* DNA scission at *Eco*RI sites in the chromosome induces the SOS response and at higher doses kills the cell.

Some earlier observations bear on our findings. For example, Tn*10* transposition induces λ lysogens suggesting that DNA scission by transposase may act similarly (Roberts & Kleckner, 1988). In *dam*$^-$ mutants or after treatment with colicin E2 or bleomycin, DNA lesions including single and double strand breaks occur and the SOS response is induced (Craig, *et al.*, 1984; Smith & Oishi, 1978). In this latter case, the nature of the inducing lesion is not known. Previously it was demonstrated that *E. coli* repairs DNA double-strand breaks that occur in λ DNA after an incorporated radioactive phosphorus atom decays (Ogawa & Tomizawa, 1967; Tomizawa & Ogawa, 1967). The repaired breaks may actually be two nicks held together by the intervening sequence because 28% of all decay induced breaks remained annealed at 15°C and

Figure 5-5. Sensitivity of *lig4* mutant to DNA scission. Sensitivity to DNA scission was measured as in figure 5-3, however in this case the temperature shift was to inactivate the mutant DNA ligase rather than to activate the endonuclease. Here cultures were grown at 30°C, and colony formation was challenged by prior incubation for increasing time at 42°C, followed by growth at the permissive temperature (30°C). The number of colonies was determined and the fraction surviving was calculated. Symbols: ●, N1626(RA2); ■, N2603(RA2); ○, N1626(RA2 and pJC1);

□, N2603(RA2 and pJC1).


Fraction of colonies surviving
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
4
8
12
16
20
24
Time at 42°c (hours)

the same proportion (26%) of total breaks were repaired. Decay induced lesions in the *E. coli* chromosome required RecA for repair and were repaired much more efficiently than breaks borne by an extrachromosomal molecule (Tomizawa & Ogawa, 1968).

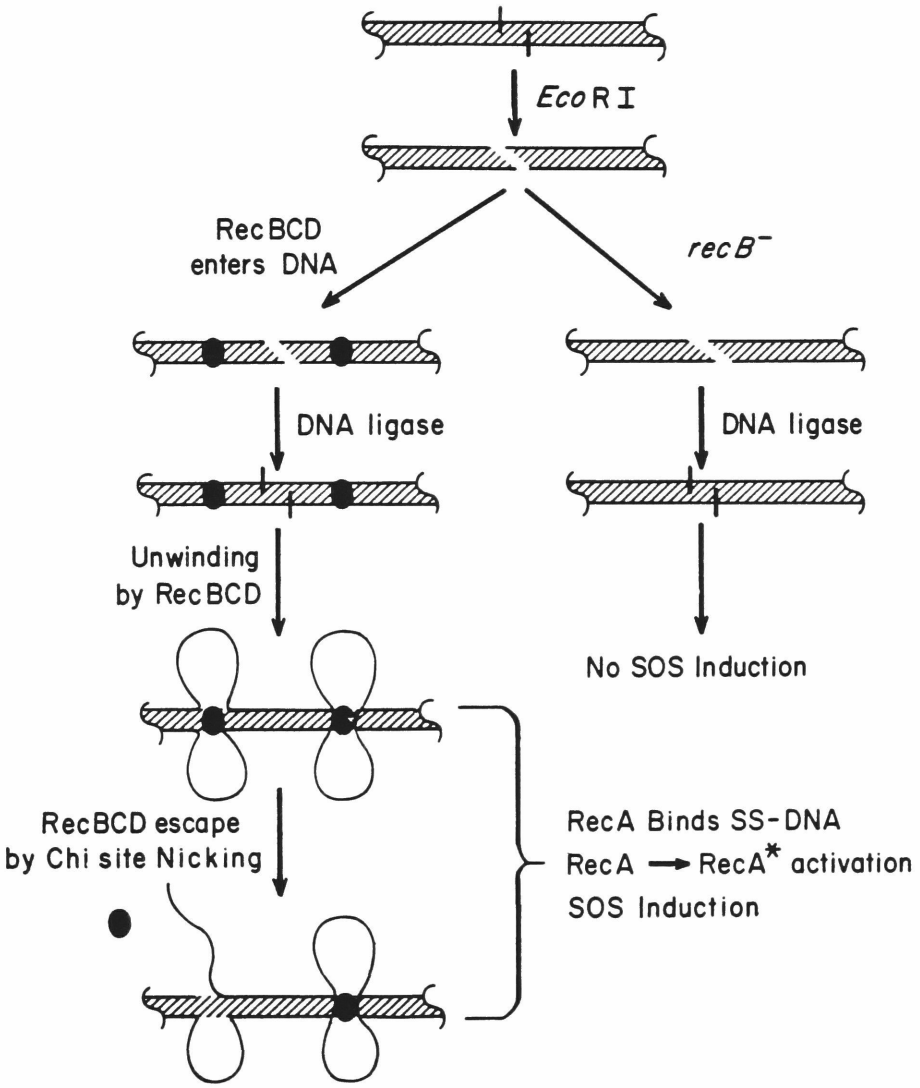
To determine the amount of *Eco*RI endonuclease which cells can survive, we assessed the relative activity of the temperature sensitive alleles at different temperatures by measuring restriction of phage λ growth. Restriction was measured in the presence of the *Eco*RI methylase since, in its absence, *in vivo* DNA induces cellular filamentation which interferes with λ infection. With the TS6 allele as an example, we find that when 1% of the wild-type restriction activity remains according to λ plating efficiency (at 37°C EOP=6 X 10^{-3} for TS6 and 4 X 10^{-5} for wild-type), in the absence of methylase poorly growing colonies are formed (E.O.P.=0.01 to 0.1). Phage λ bears five *Eco*RI recognition sites. Four of these result in 10-fold restriction per site, while the other site (sR2) is restricted only 3-fold (Murray & Murray, 1974). From the observed level of restriction, we can calculate the expected frequency of cleavage for a λ borne *Eco*RI site using the formula $(1\text{-}X)^n = Y$, where X=frequency of cleavage, n=the effective number of *Eco*RI restriction sites on λ (n=4.3) and Y=number of phage surviving restriction. We find that at ~ 1% the wild-type level of restriction (6 X 10^{-3}) where cells grow poorly, the expected frequency of cleavage is 0.7. If the *Eco*RI sites of the *E. coli* chromosome suffer a similar frequency of cutting, this predicts an average of 460 of the 659 known *Eco*RI sites (Kohara, *et al.*, 1987) are cleaved per chromosome. Since 1% of cells survive this level of DNA scission, either the cell rapidly repairs DNA breaks, or phage λ DNA is more sensitive to *Eco*RI cleavage than the *E. coli* chromosome. We note that this analysis may underestimate the extent of damage to the chromosome, since we measured restriction of λ in the presence of the methylase whereas cell lethality of the TS alleles was in the absence of the methylase.

Because *Eco*RI scission induces the SOS response, we wondered whether blocking

this induction would increase cell sensitivity to DNA cleavage. SOS induction results from a cascade of events which can be interrupted by mutations at several points (for a review see Walker, 1984, 1985). After DNA damage, the RecBCD or RecF proteins (or other pathways) produce an intermediate (probably single-stranded DNA) which activates the RecA protein to a form called RecA*. RecA* then stimulates the auto-digestion of the LexA protein, the common repressor of the SOS genes, to induce the SOS response. Strains carrying the *recA56* or *lexA3* alleles cannot cleave the LexA repressor to induce the SOS response. We found that DNA scission did not induce the SOS response in *recA* and *lexA3* mutant strains. However, the absence of SOS induction did not increase the sensitivity of these strains to DNA scission. Thus SOS induction is not required to repair DNA cleaved by *Eco*RI. In strains lacking the RecBCD helicase activity, SOS is induced by some agents (mitomycin C, UV) but not others (nalidixic acid) (McPartland, 1980). SOS induction after DNA scission proceeds via the RecBCD pathway, but a *recB21* mutant strain was not more sensitive to DNA scission. A model for RecBCD mediated SOS induction by *Eco*RI scission is presented in Figure 5-6. DNA breaks may simply allow entry of the RecBCD complex and then be resealed. While the RecBCD proteins exhibit potent exo- and endonuclease activity *in vitro*, under *in vivo* conditions RecBCD is more active as a helicase (Taylor & Smith, 1980). Translocation and unwinding by the trapped helicase would produce SS-DNA to activate the RecA protein. RecBCD could produce a gap or nick at chi sites to escape the DNA. In this way, the SOS response could be induced by DNA breaks but plays no part in their repair.

In addition to their role in SOS induction, the RecA and RecBCD proteins participate in homologous recombination. *RecA* and *recB* mutant strains are not more sensitive to *Eco*RI mediated DNA scission, suggesting that these breaks do not require recombination for repair. These results are in marked contrast to the observations that, compared to isogenic *recA*$^+$ strains, *recA* mutant strains are 10^4-fold more sensitive to γ-ray irradiation and other DNA damaging agents (UV light, mito-

Figure 5-6. Pathway for SOS Induction by DNA Scission. Here we see the RecBCD helicase entering a DNA double-strand break which is then repaired by DNA ligase. The trapped RecBCD complex moves along the DNA, unwinding and producing SS-DNA which binds to and activates the RecA protein to $RecA^*$. The $RecA^*$ protein then activates the SOS response. By this means, SOS induction by DNA scission need play no role in the repair of the inducing lesion. RecBCD may escape the DNA by creating a gap or nicking at chi sites.


EcoRI
RecBCD
enters DNA
recB⁻
DNA ligase
DNA ligase
Unwinding
by RecBCD
No SOS Induction
RecBCD escape
by Chi site Nicking
RecA Binds SS-DNA
RecA → RecA* activation
SOS Induction

mycin C) (Picksley, *et al.*, 1984) and 10^5-fold more sensitive to cleavage by another endonuclease, λ terminase (Murialdo, 1988). Our results are not attributable to different doses of DNA cleavage, since over a range of doses at which 10^{-4} to more than 50% of *recA*$^+$ cells survive, isogenic *recA*$^-$ cells are not more sensitive. Furthermore, a strain lacking the RecN protein, which plays a role in the repair of γ-ray induced DNA breaks (Picksley, *et al.*, 1984), was not more sensitive to *Eco*RI mediated DNA scission. These results may be explained by the difference in the nature of the breaks, since γ-ray cleavage releases a base and leaves non-ligatable termini (Henner, *et al.*, 1983; Giloni, *et al.*, 1981), while *Eco*RI scission produces staggered breaks with ligatable termini. In the case of λ terminase, the protein remains tightly bound to one end of the DNA after cleavage (Feiss, *et al.*, 1983). This lesion may also require recombination for repair.

Since DNA ligase is necessarily involved in DNA repair processes by playing an important role in rejoining DNA molecules, we expected it to play a role here. As described in Results and figure 5-5, conditional DNA ligase mutations conferred extreme sensitivity to DNA scission. Strains bearing the *lig4* allele exhibit reduced DNA ligase activity at 30°C (35% of wt levels) and 42°C (< 1% wt levels) with no loss in cell viability. The *lig ts7* allele more severely impairs ligase activity (5% of wt at 30°C) and does not support cell growth at 42°C (< 1% wt level) (Gottesman, *et al.*, 1973; Konrad, *et al.*, 1973). An *Eco*RI endonuclease allele with reduced activity (RA2) could not be introduced into a *lig ts7* strain and dramatically reduced the colony formation efficiency of a *lig4* strain. Thus the cell is extremely sensitive to DNA scission when ligase activity is reduced to 5% to < 1% of wild type levels. We conclude that DNA ligase is required and may be sufficient for the repair of these double-strand DNA breaks.

These findings raise the question: if DNA ligase repairs these DNA breaks that occur amongst the 659 *Eco*RI sites in the chromosome, how does DNA ligase know which ends to ligate together? We suggest that the tertiary or nucleoid structure of

the chromosome (Worcel & Burgi, 1972) may hold severed ends in proximity and promote their ligation. An interesting experimental prediction is that sensitivity to *Eco*RI breaks should be increased in mutant strains which lack unbiquitous DNA binding proteins (such as IHF or HU) thought to fold the *E. coli* chromosome (reviewed in Drlica & Rouviere-Yaniv, 1987; Pettijohn, 1988). It is known that the sedimentation properties of isolated *E. coli* nucleoids are not altered when up to 50 double-strand breaks are delivered by *in vitro* γ-irradiation (Lydersen & Pettijohn, 1977), arguing that some higher order structure is imposed upon the broken DNA. Alternatively, the cell may simply heal each break as fast as it is formed thus solving the problem of which ends to ligate together. In addition, some proteins (Mu Gam, RecA) are known to bind DNA ends and could assist DNA ligase (Akroyd & Symonds, 1986; Register & Griffith, 1986).

Plasmid-borne staggered or blunt DNA double-strand breaks are repaired when *E. coli* is transformed with linear DNA (Conley & Saunders, 1984; Winans, *et al.*, 1985; conley, *et al.*, 1986a, 1986b). Normally linear DNA transforms poorly because it is rapidly destroyed by exonucleases. Mutant strains which lack the RecBCD exonuclease and carry suppressor mutations which render them recombination proficient (*recBCD*$^-$ *sbcA* or *sbcBC* mutant strains for example) are efficiently transformed by linear DNA when the plasmid shares homology with the chromosome and can stably integrate (Winans, *et al.*, 1985). Although these mutant strains do not support efficient plasmid replication, under selective pressure they do maintain plasmids which do not share chromosomal homology (Kobayashi & Takahashi, 1988). Thus the finding that linear DNA integrates more frequently than it recircularizes suggests that plasmid borne double-strand breaks are usually repaired by recombination rather than ligation. Plasmids bearing a DNA double-strand gap are repaired by gene conversion using the RecE homologous recombination pathway which is active in these *recBC*$^-$ *sbcA* mutant strains (Kobayashi & Takahashi, 1988). In *recBCD*$^+$ hosts, linear plasmid DNA circularizes and yields transformants at 1% the frequency

observed with circular DNA. Transformation is largely dependent on recombination functions and deleted plasmids often occur (Conley & Saunders, 1984; Conley, *et al.*, 1986a, 1986b). Infecting phage that suffer *Eco*RI cleavage are restricted and do not form plaques. At a low frequency they can be repaired by recombination with an endogenous intact prophage (Stahl, *et al.*, 1977). These findings are in contrast to our observations that *Eco*RI breaks in the chromosome require ligation but not recombination for repair. This difference is further underscored by the observations that transformation by linear DNA (Conley & Saunders, 1984) and phage restriction (unpublished data) are not affected by DNA ligase mutations. Thus plasmid- or phage-borne DNA double-strand breaks are treated differently than chromosomal breaks. We suggest that small DNA molecules (plasmids and phage) lack the higher order structure of the chromosome. After DNA scission their DNA ends may be more exposed to nuclease degradation which prevents ligation and necessitates repair by other less effective means.

DNA double-strand breaks or ends stimulate recombination in *E. coli* and in yeast (Thaler & Stahl, 1988; Szostak, *et al.*, 1983) and sister chromatid exchanges in CHO cells (Morgan, *et al.*, 1988). In *E. coli* DNA breaks (including *Eco*RI breaks) activate the λ Red pathway, chi-mediated RecBCD recombination, and the RecE and RecF pathways (M. Stahl, *et al.*, 1983; F. Stahl, *et al.*, 1986; Thaler, *et al.*, 1987a, 1987b, 1987c). DNA that is transferred by transduction or conjugation is in a linear double- or single-stranded DNA form prior to recombination. Similarly, in yeast and certain mutant strains of *E. coli*, linear DNA readily transforms cells by recombining into the chromosome (Orr-Weaver, *et al.*, 1981; Winans, *et al.*, 1985). Furthermore, DNA scission has been proposed to initiate yeast meiotic recombination in general (Szostak, *et al.*, 1983), as it does in the specific case of mating type switching after HO endonuclease cleavage (Strathern, *et al.*, 1982). In fact, Sun *et al.* (1989) recently demonstrated that a DNA double-strand break occurs at the yeast *ARG4* concomitant with commitment to meiotic recombination.

It has been suggested that DNA breaks may promote exchange after exonuclease processing to a strand invasive substrate which increases recombination at the break, or by allowing recombinases entry to stimulate recombination away from the break (Thaler, *et al.*, 1987b). Our findings suggest that *Eco*RI DNA breaks act as transient entry points for the RecBCD complex and are then resealed. Because RecBCD requires DNA ends for its recombination activity, there may be an *Eco*RI cellular analog to provide RecBCD entry points. In fact, Mahan & Roth (1989) have recently argued that the recombination process itself may generate DNA double-strand breaks that act as RecBCD entry points.

DNA double-strand breaks are repaired and stimulate recombination in the yeast *Saccharyomyces cerevisiae* (Friedberg, 1988). Perhaps the most familiar example occurs when the HO endonuclease inflicts a DNA double-strand break at the MAT locus, provoking recombination between the silent and expressed cassettes to switch mating type (Strathern, *et al.*, 1982). DNA scission by the Flp recombinase behaves similarly (Jayaram, 1986). Another example (which was an important technical innovation as well) was the finding that DNA double strand breaks markedly enhance integration of a plasmid at a region of homology in the chromosome (Orr-Weaver, *et al.*, 1981). Furthermore, a mutation (*rad52*) has been described which increases sensitivity to γ-irradiation (Resnick, 1969). *rad52* strains are sensitive to a single γ-ray induced DNA break (Ho & Mortimer, 1973; Ho, 1975; Resnick & Martin, 1976), do not switch mating type and are in fact killed by the single break that occurs at MAT (Malone & Esposito, 1980; Weiffenbach & Haber, 1981), and are not transformed by linear DNA (Orr-Weaver, *et al.*, 1981). Since *RAD52* is required for mitotic and meiotic recombination (reviewed in Haynes & Kunz, 1981), these findings implicate recombination in the repair of DNA breaks. Furthermore, the double-strand break repair model suggests that not only are breaks repaired by recombination, but that DNA scission initiates genetic recombination (Szostak, *et al.*, 1983). However, the actual picture is now more complicated than this simple model predicts (Stahl, *et al.*,

1986).

Barnes and Rine (1985) studied the consequences of DNA breaks elicited in yeast by expression of the *Eco*RI endonuclease. They conditionally expressed the endonuclease from the tightly regulated *GAL1* promoter on a low copy number CEN plasmid. In this system, repression was sufficient to allow cell survival in the absence of the *Eco*RI methylase. They concluded that *rad52* mutants were much more sensitive than *RAD52* strains to *Eco*RI mediated breaks. However, these findings were somewhat compromised because the *Eco*RI expressing plasmid that was more lethal in *rad52* mutants also carried one *Eco*RI site. When they subsequently measured the rate at which *RAD52* and *rad52* mutant strains were killed by *Eco*RI expression from a plasmid without an *Eco*RI site, they found that *RAD52* and *rad52* strains **did not differ** in sensitivity to *Eco*RI breaks (Rine & Barnes, 1986). Although they do not state it, this finding clearly contradicts their earlier conclusion and demonstrates that *RAD52* is not required to repair *Eco*RI double-strand breaks in yeast chromosomes.

Where did they go wrong? The original observation was that *rad52* mutants are more sensitive than *RAD52* strains to *Eco*RI breaks when the *Eco*RI expressing plasmid bears an *Eco*RI site. They argued that *RAD52* strains survive these conditions because, in the process of repairing their broken plasmid, deletions occur that remove the *Eco*RI promoter. Thus the proper conclusion is that either *rad52* mutants cannot repair plasmid-borne *Eco*RI breaks or, they are defective in forming the deletions that allow *RAD52* cells to survive. These findings are of interest for two reasons. First, *Eco*RI breaks in the *E. coli* chromosome behave differently compared to breaks in a plasmid or phage. Secondly, we observed no difference in cell lethality when our *Eco*RI expressing has one or no *Eco*RI sites. Since we employed a high copy number plasmid (pBR322), copy number control mechanisms probably come into play before every plasmid suffers scission, and increased replication could effectively replace copies as they are destroyed. In contrast, Barnes and Rine used a low copy CEN plasmid, and every copy might have suffered scission soon after en-

donuclease expression and before increased replication could occur. We conclude that *Eco*RI breaks do not require recombination for repair in either *E. coli* or in yeast. In both organisms, chromosomal double-strand breaks are treated differently than breaks in a plasmid, possibly because the tertiary structure of these DNA molecules differs.

When the *Eco*RI endonuclease was expressed in CHO mammalian cells, the resulting breaks increased the frequency of sister chromatid exchanges (Morgan, *et al.*, 1988), Thus *Eco*RI scission stimulates recombination in a mammalian system, as it does in bacteria (Stahl, *et al.*, 1977; Thaler, *et al.*, 1987a, 1987b, 1987c). It is possible that recombination in CHO cells is enhanced by *Eco*RI breaks without directly playing a part in their repair, as is the case in bacteria (Heitman, *et al.*, 1989).

The work described in this chapter was in part motivated by two observations that suggest *E. coli* survives a low level of endonuclease expression. First, some endonuclease clones are not lethal to hosts which lack the protective methylase. Second, plasmids bearing the genes for restriction-modification systems transform host strains with unit efficiency even though some damage might be inflicted before the chromosome is fully methylated. Our finding that a weak activity *Eco*RI endonuclease mutant is inviable in combination with DNA ligase mutations argues that in this and similar situations (R^+ M^- clones), DNA repair is required for cell viability. In contrast, the transformation efficiency of a plasmid carrying the *Eco*RI restriction-modification system (pAN4) was not decreased in *lig* mutant strains (data not shown). This suggests that very little DNA damage occurs when a restriction-modification system is established in a new host cell. We therefore returned to the suggestion that endonuclease gene regulation might play a role in protecting the host cell after transformation. Inspection of the 5' non-coding region of the *Eco*RI endonuclease gene reveals a sequence with homology to the consensus LexA repressor binding site (CTtTataaaataACAG in the *Eco*RI gene compared to the consensus site CTGTatatatatACAG, Wertman & Mount (1985)). This suggests that *Eco*RI endonu-

clease expression could be repressed by the LexA protein immediately after transformation when the plasmid copy number is low. Further experiments will be required to address this model.

In conclusion, we have described a set of conditional alleles of the *Eco*RI restriction enzyme which allow one to deliver a dose of DNA double-strand breaks with defined staggered termini. These alleles should prove useful for further studies of DNA structure and repair, both in *E. coli* and, with appropriate expression systems, in other organisms. To produce breaks of other types, similar mutants could be isolated for enzymes that yield double-strand breaks with blunt or 3'-overhanging termini. Because blunt ends are poor substrates for *E. coli* DNA ligase, *E. coli* may be more sensitive to or employ different repair pathways for blunt DNA breaks. In the next two chapters, we exploit the finding that DNA scission induces expression of SOS::*lacZ* fusions to screen for and assay *Eco*RI endonuclease mutants. This SOS induction assay may prove generally useful for cloning, characterizing, and mutating endonucleases, methylases, and other enzymes (such as recombinases or topoisomerases) whose substrate is DNA.

Chapter 6

Substrate Recognition by the *Eco*RI Endonuclease

Introduction

X-ray crystallography revealed that three prokaryotic DNA-binding proteins (CAP from *E. coli*, and *cI* and *cro* of phage λ) share a super-secondary structural motif, the helix-turn-helix, which binds these dimeric proteins to adjacent major grooves of the DNA (McKay & Steitz, 1981; Anderson, *et al.*, 1981; Pabo & Lewis, 1982; Ohlendorf, *et al.*, 1982). Homologous sequences occur in many other prokaryotic and eukaryotic DNA-binding proteins that regulate transcription (Steitz, *et al.*, 1982; Anderson, *et al.*, 1982; Takeda, *et al.*, 1982; Laughon & Scott, 1984). The simple protein-DNA interaction models proposed initially have now grown more complicated with the solution of high resolution structures of DNA-protein co-crystals of the phage 434 repressor (Anderson, *et al.*, 1987; Aggarwal, *et al.*, 1988), λ *cI* protein (Jordan & Pabo, 1988), and the *E. coli trp* repressor (Otowinski, *et al.*, 1988). The specific details of DNA sequence recognition differ substantially amongst these helix-turn-helix proteins. For example, the *cI* repressor binds to B-form DNA whereas the 434 and *trp* repressors bind to DNA which is significantly distorted in structure.

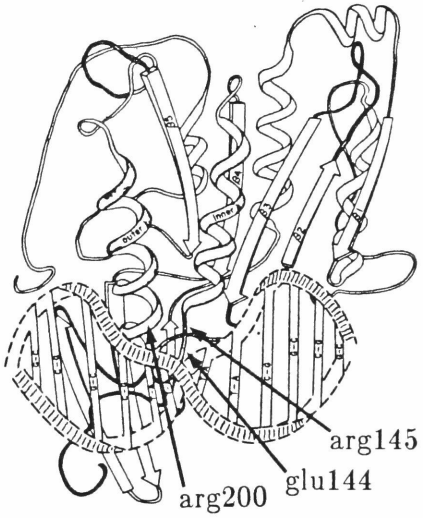
DNA-binding proteins and enzymes could employ similar strategies to interact with DNA. However, since even simple binding proteins differ dramatically in their specific DNA interactions, enzymes which bind DNA might be even more different. Additionally, enzymes need both catalytic and binding sites and their actions may be coordinated by structural motifs not observed for simple binding proteins. Notably, a DNA-enzyme co-crystal structure of the *Eco*RI endonuclease shows that the DNA backbone is significantly distorted (Frederick, *et al.*, 1984; McClarin, *et al.*, 1986), as is the case with the 434 and *trp* repressors. The enzyme also makes hydrogen bond contacts to the edges of the basepairs as do the 434 and λ *cI* repressors. However, the *Eco*RI endonuclease lacks the helix-turn-helix motif. Instead two α-helices from

each subunit of the symmetric dimer project steeply into the major groove of the DNA and carry amino acids that hydrogen bond to the edges of the nucleotides of the substrate (see figure 6-1). Glutamic acid 144 and arginine 145 lie at the end of the inner α-helix and each, from opposite subunits, spans and makes two hydrogen bonds with the central adenines (gAAttc). Arginine 200 lies at the end of the outer helix and makes two hydrogen bonds to the outer guanine nucleotide (Gaattc). Contacts to the right and left halves of the palindromic substrate are symmetric. Thus for the dimer, a total of twelve hydrogen bonds form the substrate binding pocket (shown in figure 6-1). Any nucleotide substitution within the recognition site disrupts one or more of these interactions. Thus this hydrogen bond network could enable the enzyme to uniquely recognize this DNA sequence.

One test of this model is to isolate mutants altered at the substrate recognition amino acids. Previously, conservative amino acids substitutions have been described (glu144asp, arg145lys, arg200lys) (Wolfes, *et al.*, 1986; Greene, *et al.*, 1987). These mutations reduce enzymatic activity but do not alter substrate specificity. Since mutants bearing conservative amino acid substitutions may make contacts similar to the wild-type enzyme, we wished to test whether other amino acids at these three positions would alter the substrate specificity of the enzyme. To this end we developed a simple and sensitive *in vivo* plating assay for endonuclease activity (see figure 6-2).

In Chapter 5 we described our findings that *Eco*RI staggered DNA breaks induce the SOS DNA repair response. This finding allows us to assay *in vivo* DNA scission by monitoring SOS induction. Here we employ this assay to rapidly assess the phenotypes of *Eco*RI mutants altered in the amino acids predicted to bind the substrate. *In vivo* and *in vitro*, we find that these mutant proteins are crippled in activity yet still bind to and cleave the wild-type substrate. The hydrogen bond network revealed by the crystal structure is insufficient to fully account for substrate recognition, and other amino acids must contact the DNA to help discern the substrate. The activity of these mutant proteins is enhanced by buffer conditions or mutations

Figure 6-1. Crystal Structure and Substrate Interaction Model for the *Eco*RI Endonuclease. On the left we see the 3 A resolution X-ray crystal structure of an *Eco*RI endonuclease-DNA complex. One monomer of the symmetric homodimer is omitted for clarity. Each monomer bears two α-helices (the inner and outer helices) which project steeply into the major groove of the DNA and deliver amino acids that bind the DNA. The three amino acids (glu 144, arg 145, and arg 200) predicted to interact with the nucleotides of the recognition site are indicated by the arrows. On the right side is a detailed model of the hydrogen bonds between the protein and the DNA. Arginine 200 makes two hydrogen bonds to the guanine of the recognition site (Gaattc) while glutamic acid 144 and arginine 145 (from opposite subunits) span and each make two hydrogen bonds to the central adenines (gAAttc). There are a total of six hydrogen bonds for one half of the recognition site. Since the contacts to the other half are symmetric, this model predicts a total of twelve hydrogen bonds. Any nucleotide substitutions within the recognition site would disrupt one or more of these interactions. Thus this hydrogen bond network could enable the enzyme to uniquely recognize this site. The crystal structure and model shown here are derived from McClarin *et al.* (1986).


arg145
glu144
arg200

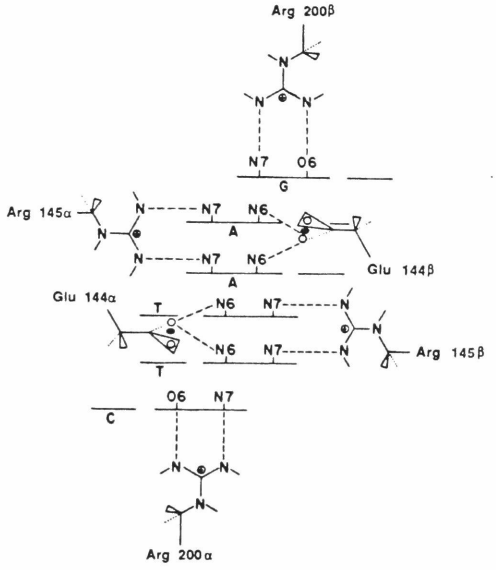
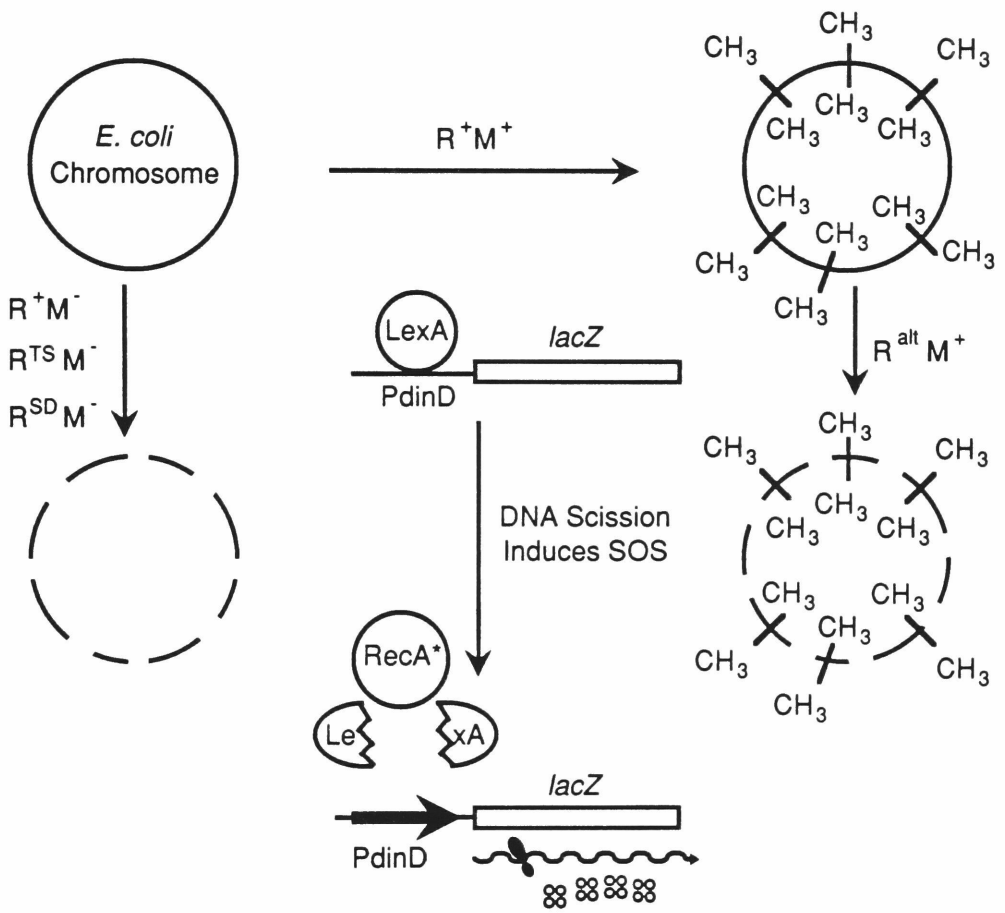

Arg 200β
N7
O6
G
Arg 145α
N7
N6
A
N7
N6
A
Glu 144β
Glu 144α
T
N6
N7
N6
N7
T
Arg 145β
O6
N7
C
Arg 200α

Figure 6-2. *In Vivo* DNA Scission Assay. As described in chapter 5, when the *E. coli* chromosome suffers DNA scission, the SOS DNA repair response is induced. Induction of the SOS response is a cascade whereby damaged DNA activates the RecA protein to a from called RecA*. RecA* then binds to and stimulates the auto-digestion of LexA, the transcriptional repressor of the SOS genes. To assay this induction, the *lacZ* gene was fused to an SOS promoter. Now DNA damage activates β-galactosidase expression which in turn causes the production of blue colonies on X-gal indicator media. We used this *in vivo* DNA scission assay to test the phenotypes of site-directed *Eco*RI mutants and also as a screen for mutants that evade the protection of the methylase. If a given *Eco*RI mutant induces the SOS response, we conclude that it exhibits endonuclease activity. If expression of the *Eco*RI methylase blocks this SOS induction, we conclude that the endonuclease is acting at the wild-type recognition site.


E. coli
Chromosome
R+M+
R+M-
RTS M-
RSD M-
LexA
lacZ
PdinD
DNA Scission
Induces SOS
RecA*
Le
xA
lacZ
PdinD
CH3
R alt M+

which we argue favor conformational changes that allosterically activate DNA cleavage. These substrate binding amino acids may initiate these conformational changes. We suggest that the *Eco*RI endonuclease achieves its high fidelity of substrate recognition by making extensive contacts to its substrate and by coupling binding and scission through allosteric activation.

Results

EcoRI mutants altered at amino acids 144, 145, or 200 retain activity

By site-directed mutagenesis using multiply degenerate oligonucleotides and the Kunkel method (1985), we rapidly generated many substitutions of amino acids 144, 145, and 200. The phenotypes of 50 of the 60 possible point mutants are summarized in table 6-1, and in more detail in tables 6-2 and 6-3. We employed two strains, JH137 and JH69, which allow us to monitor SOS induction since they carry the lactose operon fused to the DNA-damage inducible locus *dinD* (Kenyon & Walker, 1980). Strain JH69 is an HB101 derivative which is less sensitive to DNA scission than strain JH137. Like the wild-type enzyme, the three most active mutant proteins (E144D, E144C, and R200K) are lethal to strain JH137 in the absence of the protective methylase (see table 6-2). However, these three alleles (but not the wild-type enzyme) are viable at some growth temperatures in strain JH69, and under conditions where the cells remain viable these mutants clearly induce the SOS response (see table 6-3). By both the cell lethal phenotype and the SOS induction assay, we conclude that these three alleles exhibit endonuclease activity which is however somewhat reduced compared to the wild-type enzyme. We note that double mutants constructed from the most active single mutants (E144D + R200K, E144D + R200C, E144C + R200K, and E144C + R200C) were inactive. Furthermore, in one case where it appeared that another amino acid might substitute for the missing one (R203 for R200), altering the amino acid (R203 to K) in a wild-type endonuclease background resulted in a null phenotype. This suggests that R203 plays its own

Table 6-1. Phenotypes of Site-Directed *Eco*RI Endonuclease Mutants

Phenotype	Amino Acid		
	144	145	200
Wild-type	glu	arg	arg
Lethal without methylase	asp cys		lys
Induce SOS	ser gly	 lys cys	cys val ser
Null	lys trp thr met ala phe arg tyr val asn leu pro	glu thr phe ile met gly asp asn ser ala	trp gln pro asp leu asn gly his met ile thr phe ala tyr glu

Table 6-2. SOS Induction by E144Z, R145V, and R200X *Eco*RI Mutants in Strain JH137

Allele	Phenotype	M⁻ 42	37	34	30	30-42
WT	lethal w/o M	dead	dead	dead	dead	W-LB
R200K	lethal w/o M[a]	dead	dead	dead	dead	W
R200C	TS	LB(25)	MB	MB	DB,sick(100)	W
R200V	TS	W-LB(20)	LB	L-MB	MB(83)	W
R200S	TS	W-LB(12)	W-LB	LB	LB(41)	W
R200X	null	W	W	W	W	W
E144D	lethal w/o M[a]	dead	dead	dead	dead	W
E144C	TS	DB(240)	DB,sick(350)	dead	dead	W
E144S	TS	W-LB(24)	DB,sick	DB,sick	DB,sick(130)	W
E144G	TS	W(9)	W-LB	W-LB	W-LB(25)	W
E144X	null	W	W	W	W	W
R145K	TS	W-LB(28)	LB	LB	LB(61)	W
R145C	TS	W(13)	W-LB	W-LB	W-LB(30)	W
R145X	null	W	W	W	W	W

[a] The R200K and E144D alleles are lethal without methylase in strain JH137, but viable at high temperature in the HB101 derived strain JH69. In this strain background both alleles are temperature sensitive.

[b] Indicator plates contained 35 μg/ml X-gal. Under these conditions color intensity corresponded to the following units of β-galactosidase activity: W(white)= ~ 10 units, W-LB(white to light blue)= ~ 20 units, LB (light blue)= 25-60 units, MB(medium blue)= 60-100 units, and DB (dark blue) = 100+ units.

[c] the following alleles (R200X, E144X, R145X) showed no induction (null) of the SOS response: R200P, T, N, W, L, G, M, A, E, Q, D, H, I, F, and Y; E144K, T, A, R, V, L, W, M, F, Y, N, C, and P (E144Q, I, and H not tested); and R145E, F, M, D, S, T, I, G, N, and A (R145P, W, L, V, Q, H, and Y not tested).

w=white or faint blue, LB=light blue, MB=medium blue, DB= dark blue

Table 6-3. SOS Induction by *Eco*RI Mutants in Strain JH69

Allele	Phenotype	42	37	M⁻ 34	30	30-42
WT	lethal w/o M	dead	dead	dead	dead	W-LB
R200K	TS,lethal	LB-MB	DB,sick	dead	dead	W-LB
R200C	TS	LB	LB	LB	LB-MB	W-LB
E144D	TS	DB	DB,sick	DB,sick	DB,sick	W-LB
E144C	TS,lethal	LB-MB	DB,sick	DB,sick	dead	W-LB

w=white or faint blue, LB=light blue, MB=medium blue, DB= dark blue colonies on X-Gal indicator plates containing 35 μg / ml X-Gal.

The background level of *lac* expression is higher in strain JH69 ($lacZ^+$ $lacY^-$) than in JH137 ($\Delta lacZ$).

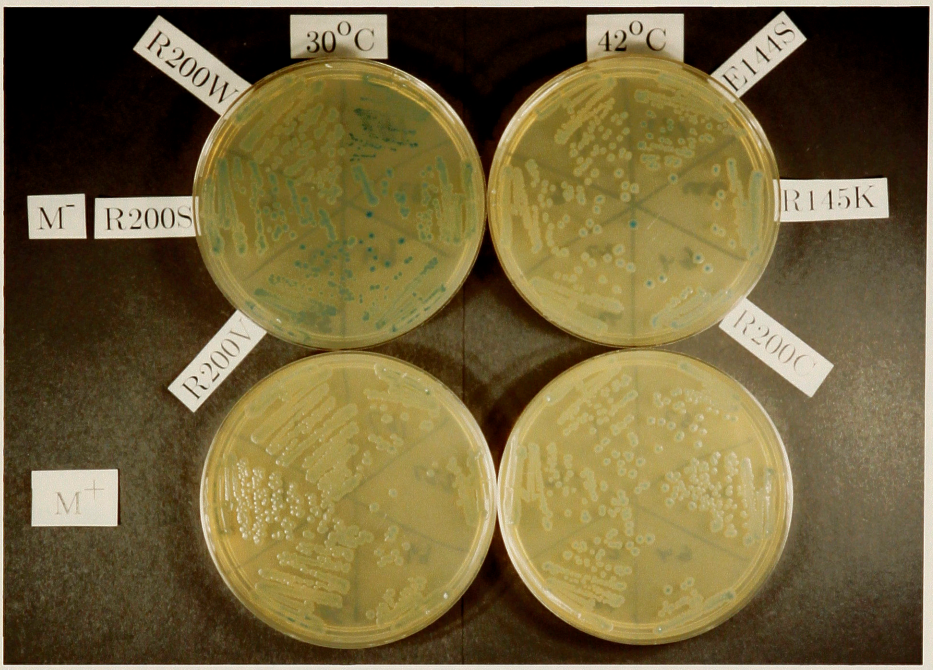
essential role and is therefore an unlikely candidate to substitute for the missing arginine in the R200K mutant.

Several additional alleles at each position are not lethal to strain JH137 or JH69 but do retain some endonuclease activity by the *in vivo* SOS induction assay (see tables 6-1, 6-2, and 6-3). Figure 6-3 shows colonies of the SOS::*lacZ* fusion strain JH137 expressing several of the arginine 200 mutants. The indicator medium contains X-gal, a chromogenic substrate which yields a blue dye when cleaved by β-galactosidase. As the *lacZ* gene is now controlled by a DNA-damage inducible promoter, a blue colony color is indicative of SOS induction, which occurs here in response to *in vivo* DNA scission. Mutants containing cysteine, serine, or valine at position 200 produce blue colonies. Because these mutants induce the SOS response, we conclude that they cleave DNA *in vivo*. Several other mutants shown here induce no more β-galactosidase expression than a deletion or non-sense *Eco*RI mutant and were classified as null mutants. By the same criteria, the serine and glycine mutants at position 144 and the lysine and cysteine substitutions at position 145 display some endonuclease activity. Every other amino acid substitution at position 200 and many other changes at positions 144 and 145 exhibit no SOS induction and hence result in a null phenotype. In summary, at each position every mutation reduces activity compared to the wild-type amino acid. The most conservative amino acid substitution exhibits greater activity than other changes, but several non-conservative changes are also partially active (R200CVS, E144CSG, and R145C). Furthermore, those mutant proteins which have activity display a temperature-sensitive phenotype. Since altering these three amino acids cripples enzyme activity, these residues play an important but not essential role.

Site-directed mutants retain wild-type substrate specificity

The endonuclease activity of the mutants was assayed in the presence of the *Eco*RI methylase as a measure of cleavage specificity (see table 6-2). Expression of the

Figure 6-3. Site-Directed *Eco*RI Mutants Induce β-Galactosidase Expression from an SOS::*lacZ* Fusion. This figure shows the SOS *in vivo* DNA scission assay in action. These colonies are expressing the indicated site-directed *Eco*RI endonuclease mutants. The media contains the chromogenic substrate X-gal and the host strain (JH137) bears an SOS::*lacZ* fusion. Thus blue colonies are experiencing DNA scission. The strains in the upper half lack the *Eco*RI methylase, and because the wild-type endonuclease would be lethal in this case, the mutant endonucleases are clearly reduced in activity. By their ability to induce the SOS response, several retain some endonuclease activity (R200S, R200V, R200C, R145K, and E144S). The R200W mutant is a null mutant because it induces no more β-galactosidase expression than a nonsense or deletion derivative of *Eco*RI. The strains in the lower half express the *Eco*RI methylase and this protects the cell from DNA scission by the mutant endonucleases (no SOS induction). Hence these site-directed *Eco*RI mutants act at the wild-type recognition site. These strains were grown at 30°C (left half) and 42°C (right half) to show that all the active site-directed mutants wxhibit a TS phenotype.


30°C
42°C
R200W
E144S
M⁻
R200S
R145K
R200V
R200C
M⁺

*Eco*RI methylase from a compatible plasmid (pJC1) completely blocks DNA cleavage by either the *Eco*RI site-directed mutants or the wild-type enzyme as measured by restoration of cell viability and the absence of SOS induction. Thus the mutant endonucleases act at the same sites as the wild-type methylase. As a control, we checked that expression of the *Eco*RI methylase did not block or decrease SOS induction by other DNA damaging agents such as nalidixic acid or mitomycin C (data not shown). One further consideration was that the *Eco*RI methylase might modify sites other than GAATTC, thereby masking an altered activity of the mutants. Under some *in vitro* conditions the *Eco*RI methylase modifies sites related to the wild-type by one or more nucleotide substitutions (*Eco*RI* sites) (Berkner & Folk, 1978; Woodbury, *et al.*, 1980a). However, we find no evidence for this activity *in vivo* (described in more detail in chapter 9). Furthermore, as will be described in Chapter 7, we isolated several *Eco*RI endonuclease mutants which clearly induce the SOS response and cleave DNA in spite of the normally protective methylase. Thus if the site-directed mutants had substantial activity at sites other than the wild-type substrate, this assay should have detected it. We conclude that the active site-directed *Eco*RI endonuclease mutants are not dramatically altered in substrate recognition.

Site-directed mutants do not restrict λ

As a further *in vivo* measure of the mutants enzyme activity, their competence to restrict phage growth was assessed. The wild-type *Eco*RI restriction system (R^+M^+) decreased the efficiency of λ plaque formation to 10^{-4} / phage. In contrast, λ phage plated with equal efficiency on strains expressing both a mutant endonuclease and the wild-type methylase. Thus, although some alleles have sufficient activity to kill the cell in the absence of the methylase, in competition with the methylase none restricts λ.

Site-directed mutants inhibit the wild-type enzyme in trans

Using the same λ restriction assay, we asked if expression of a mutant endonu-

clease would behave as a co-dominant negative allele (Herskowitz, 1987) and inhibit the action of the wild-type enzyme. For these purposes, the wild-type restriction system was cloned into a plasmid compatible with the one bearing the site-directed alleles. A *recA* host (to minimize interplasmid recombination) carrying both plasmids was tested for λ restriction. As a control, we employed either a deletion or an amber nonsense allele of the endonuclease.

As shown in Table 6-4, all of the site-directed mutants decrease restriction of λ by the wild-type enzyme from 20- to 1500-fold. There are two possible explanations for this effect: either the mutants act on the target DNA or on the wild-type enzyme. We find that the purified R200K and R200C mutant proteins do not inhibit DNA cleavage by the wild-type enzyme *in vitro* (data not shown). This suggests their action is not attributable to competitive inhibition by binding to the substrate. Alternatively, the mutant proteins could bind to the wild-type and form defective heterodimers. One of our mutants, E144K, is known to be defective in dimerization *in vitro* (Yanofsky, *et al.*, 1987). We therefore expected this mutant might not effectively inhibit the wild-type enzyme. In contrast we find that the E144K mutant (and the structurally similar E144R mutant) inhibits restriction to a greater degree than all the other mutants.

How can this be explained? The wild-type *Eco*RI endonuclease is a dimer with an extremely high association constant. Dimerization probably occurs very near the site of synthesis such that few wild-type and mutant monomers are actually free in the cell to form heterodimers. We suggest that the E144K (and possibly E144R) mutant self-dimerizes poorly and thus presents a larger pool of monomer subunits to compete with wild-type monomers for dimerization. Either this pool of mutant monomers is sufficiently large that it compensates for the dimerization defect, or the mutant is less impaired to dimerize with a wild-type partner than with itself. Since the glu144, arg145, and arg200 residues are proposed to form salt bridges with each other to stabilize the dimer interface (McClarin, *et al.*, 1986; Yanofsky, *et al.*, 1987; Greene, *et*

Table 6-4. *Eco*RI Mutants Inhibit Restriction by the Wild-Type Enzyme

	Amino Acid			
fold decreased	144	145	200	efficiency of restriction
1-4		(UAG)stop		10^{-4}
20-50	D,Y	all[a]	all[b]	
50-150	rest[c]			10^{-2}
150-300	T,P			
500-1500	K,R			10^{-1}

[a] all other 145 = KCEFMDSTIGNA [b] all other 200 = all amino acids except R (wild-type) [c] rest of 144 = AVLWMFNCSG

al., 1987), all of these site-directed mutants may be partially defective in dimerization. This could lead to the high levels of inhibition observed. We infer that all the mutant proteins are stably synthesized and folded properly, and that they inhibit the wild-type enzyme by forming defective heterodimers. Co-dominant negative inhibitory alleles have been described for the λ, *trp*, and *lac* repressors and attributed to defective heterodimer formation (Miller, 1978; Kelley & Yanofsky, 1985). This has been directly demonstrated for mutants of the *trp* repressor (Graddis, *et al.*, 1988).

Purification of mutant proteins

To test that the *in vivo* SOS induction assay accurately reflects the *in vitro* activity of the mutants, several were purified. Mutants were over-produced by placing them under the control of the strong λ promoter P_L. Expression from P_L is usually induced by thermal denaturation of a temperature-sensitive *cI* repressor (*cI857*). Since these *Eco*RI mutants are temperature-sensitive *in vivo*, it did not seem prudent to over-express them in this way. Instead, cultures were grown at 30°C and expression from P_L was induced with nalidixic acid (Mott, *et al.*, 1985). For this purpose, a host strain (JH153) was constructed which carried a defective cI^+ λ lysogen and an *endA* mutation to decrease endogenous endonuclease levels. Using this over-production system, the wild-type and the R200K and R200C mutant endonucleases were purified by monitoring cleavage activity, using phosphocellulose and hydroxylapatite chromatography, essentially as described for the wild-type enzyme (Cheng, *et al.*, 1984).

From restriction digests, we find that the R200K and R200C mutant proteins cleave substrates (λ, pBR322, pUC18) at the wild-type recognition site (GAATTC) with specific activity reduced by three to four orders of magnitude (see table 6-5). For example figure 6-4 shows cleavage of plasmid pUC18 DNA (one *Eco*RI site) by the wild-type enzyme and the R200K and R200C mutant proteins. The mutant proteins cleave at the wild-type recognition site since, when incubated with pUC18 DNA

Figure 6-4. The R200K and R200C Mutant *Eco*RI Endonucleases Cleave DNA at *Eco*RI Sites. To determine the *in vitro* cleavage specificity of these mutant enzymes, restriction digests with plasmid pUC18 were performed. Here 500 ng of CsCl purified RFI form plasmid DNA were incubated with the purified wild-type (15 ng = 30 units) or the mutant endonucleases (R200K 15ng; R200C, 5ng) for 1 hour at 30°C in normal *Eco*RI buffer conditions (lanes 2, 5 and 9) or *Eco*RI* buffer conditions (lanes 6, 7, 8, 10, 11, 12). To provide a reference for the site of cleavage, in some reactions a second enzyme was included to cleave the plasmid at one additional site. Thus the reactions shown in lanes 3, 7, and 11 also contained 5 units of *Sma*I and those in lanes 4, 8, and 12 contained 5 units of *Xmn*I. The reaction products were displayed on a 0.6% agarose gel containing 0.5 μg/ml EtBr.

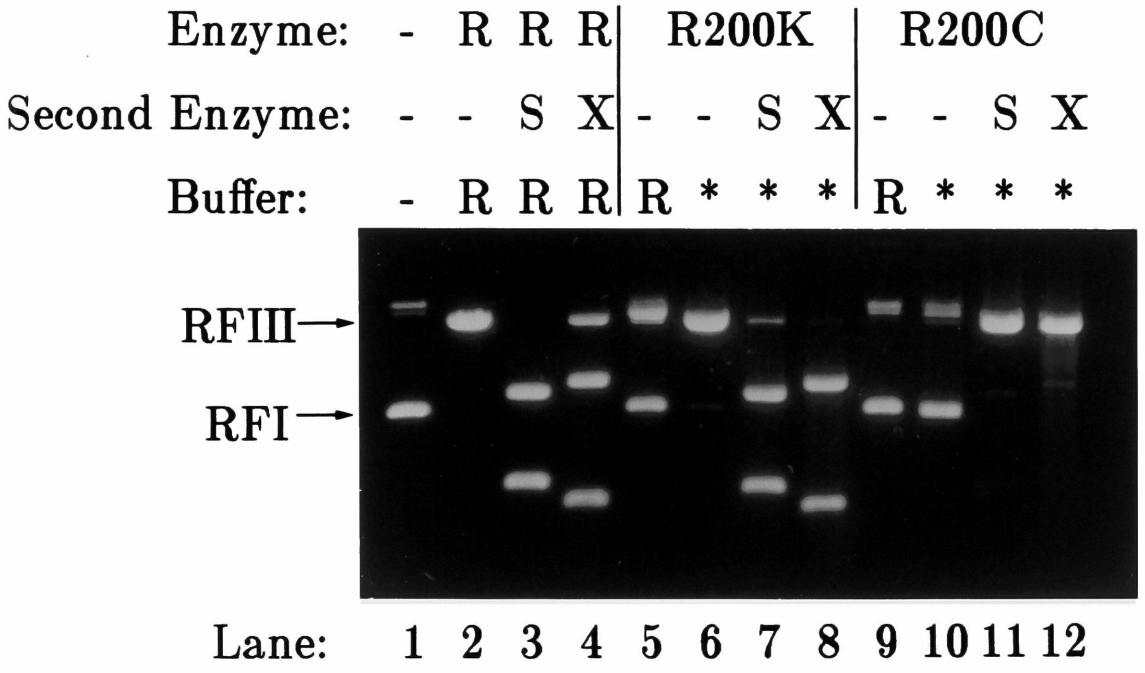

Enzyme: - R R R R200K R200C
Second Enzyme: - - S X - - S X - - S X
Buffer: - R R R R * * * R * * *
RFIII
RFI
Lane: 1 2 3 4 5 6 7 8 9 10 11 12

Table 6-5. Specific Activity and Kinetic Parameters of Enzyme Activity

Allele	*Eco*RI Buffer	Star Buffer	K_m (nm)	V_{max} (min^{-1})	RFII:RFIII
wild-type	1 - 2 $X10^6$ U/mg	ND	3.8	7.2	~ 1:4
R200K	9 X 10^2-1.8 X 10^3	7 X 10^4 - 1.1 X 10^5	2.7	0.4[a]	~ 2:1
R200C	2 -4 X 10^2	2 -4 X 10^3	ND	ND	~ 2:1

[a] for the RFII product under conditions of maximal cleavage activity (25 mM Tris-Cl pH=8, 2 mM $MgCl_2$, 5 % glycerol).

and a second enzyme which cleaves the plasmid once, two fragments are produced which co-migrate with restriction fragments produced by the wild-type enzyme (shown in figure 6-4 in *Eco*RI* reaction conditions, see below). Furthermore, substrates lacking *Eco*RI sites (f1, pBR322ΔRI) are not cleaved at detectable levels even under conditions where significant nicking by the wild-type enzyme occurs (data not shown). Thus these *Eco*RI mutants retain wild-type recognition specificity *in vivo* and *in vitro*.

Under certain buffer conditions (high pH, low ionic strength, presence of glycerol, organic solvents, or Mn^{+2} instead of Mg^{+2}) the wild-type *Eco*RI enzyme cleaves at additional non-canonical sites, each related to the wild-type recognition site by one or more nucleotide substitutions (i.e. AAATTC or GACTTC) (Polisky, *et al.*, 1975; Hsu & Berg, 1978; Woodbury, *et al.*, 1980b; Malyguine, *et al.*, 1980; Pribnow, *et al.*, 1981; Joyce & Grindley, unpublished manuscript; Gardner, *et al.*, 1982; Rosenberg & Greene, 1982). This is termed *Eco*RI* activity. As shown in figure 6-4, *Eco*RI* buffer conditions activate DNA scission by both the R200K and R200C mutant proteins (~ 100-fold for R200K and 5 to 10-fold for R200C). However this endonuclease activity retains its specificity for the wild-type recognition site, since a single linear DNA fragment is produced which, when cleaved with a second enzyme, yields two fragments that co-migrate with those produced by the wild-type enzyme (figure 6-4). In contrast, under these conditions the wild-type enzyme is no more active but does cleave additional sites. When incubated in normal or star buffer conditions, both the R200K and R200C mutant proteins produce a higher proportion of the nicked product compared to the wild-type protein. These mutant proteins may dissociate from the nicked intermediate more frequently than the wild-type, and may cleave DNA by sequential rather than concerted nicking. Alternatively, the proteins could slowly release the nicked product (see chapter 8).

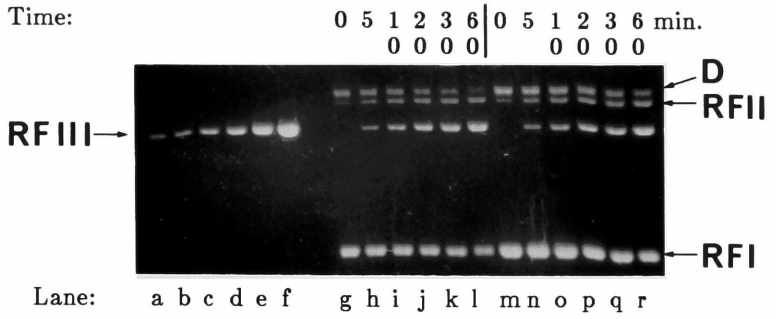
Kinetic analysis of DNA scission

A kinetic analysis was performed to determine why these mutants are reduced in specific activity. As shown in figure 6-5, different concentrations of DNA were subjected to cleavage with a limiting amount of enzyme (12 pM for the wild-type and 110 pM for the R200K mutant). Optimal reaction conditions for each enzyme were employed; thus normal *Eco*RI buffer conditions for the wild-type enzyme, and star buffer conditions for the R200K mutant enzyme. The initial velocity of the reaction was determined as described in detail in chapter 2, materials and methods. The rate of formation of the more prominent product, RFIII for wild-type and RFII for the R200K mutant protein, was plotted and the initial velocity determined in the linear portion of the curve. Figure 6-5 shows the resulting initial velocities plotted against substrate concentration in a double-reciprocal form. From the intercepts the K_m and V_{max} values were calculated for both the wild-type and the mutant enzyme (shown in table 6-5). We find that the decreased specific activity of the R200K mutant protein is attributable to a decrease in V_{max} rather than K_m. This is somewhat unusual for a mutation expected to affect the substrate binding pocket, and suggests that the binding and cleavage activities may be coupled.

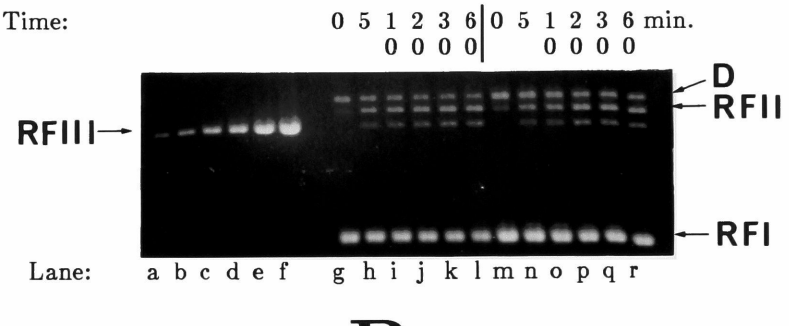
Mutations which enhance EcoRI activity suppress these site-directed mutants*

As will be described in Chapter 7, several *Eco*RI mutants with enhanced star activity were isolated. So that these potentially lethal mutants could be conditionally expressed, they were originally isolated as double mutants which also carry a temperature-sensitive mutation (R56Q, TS6). We wished to test if these enhanced star activity mutations would activate cleavage by the R200K and R200C mutants as do star buffer conditions. For this purpose, the mutants were recombined together *in vitro*. The TS6 and enhanced star activity mutations lie to the 5' side of a *Bgl*II site in the middle of the endonuclease gene, while the R200X mutations lie to its right. The mutations were combined together by ligating appropriate *Bgl*II restriction frag-

Figure 6-5. Restriction Digests for a Kinetic Analysis of DNA Scission by the Wild-Type and the R200K Mutant *Eco*RI Endonuclease. Wild-type (panel A, 77 pg enzyme) or the R200K mutant endonuclease (panel b, 690 pg enzyme) were incubated with 100 ng, 250 ng, 500 ng, 1 μg, or 2 μg of plasmid pUC18 DNA in a total reaction volume of 100 μl. The optimal reaction buffer for each enzyme was employed (normal *Eco*RI buffer for wild-type and *Eco*RI* buffer for the R200K mutant). 15 μl portions were removed at the indicated times, the reaction was terminated by the addition of 2 μl stop buffer (40% sucrose, 250 mM EDTA, 1% SDS, and 0.1% bromophenol blue). The reaction products were electrophoresed on a 0.6% agarose gel containing 0.5 μg/ml EtBr and photographed with transillumination on Polaroid 665 positive-negative film. Negatives were scanned by microdensitometry with a Joyce-Loebl densitometer and the amount of product determined by comparison with linear standards migrated on the same gel (lane a through f are 5, 10, 25, 50, 100, and 200 ng of RFIII (linear) form pUC18 DNA). Lanes g through l are derived from the reaction containing 1 μg total DNA and lanes m through r from the reaction containing 2 μg. Not shown here were additional reactions with 100 ng, 250 ng, or 500 ng total pUC18 DNA. D=dimer form of pUC18 DNA.

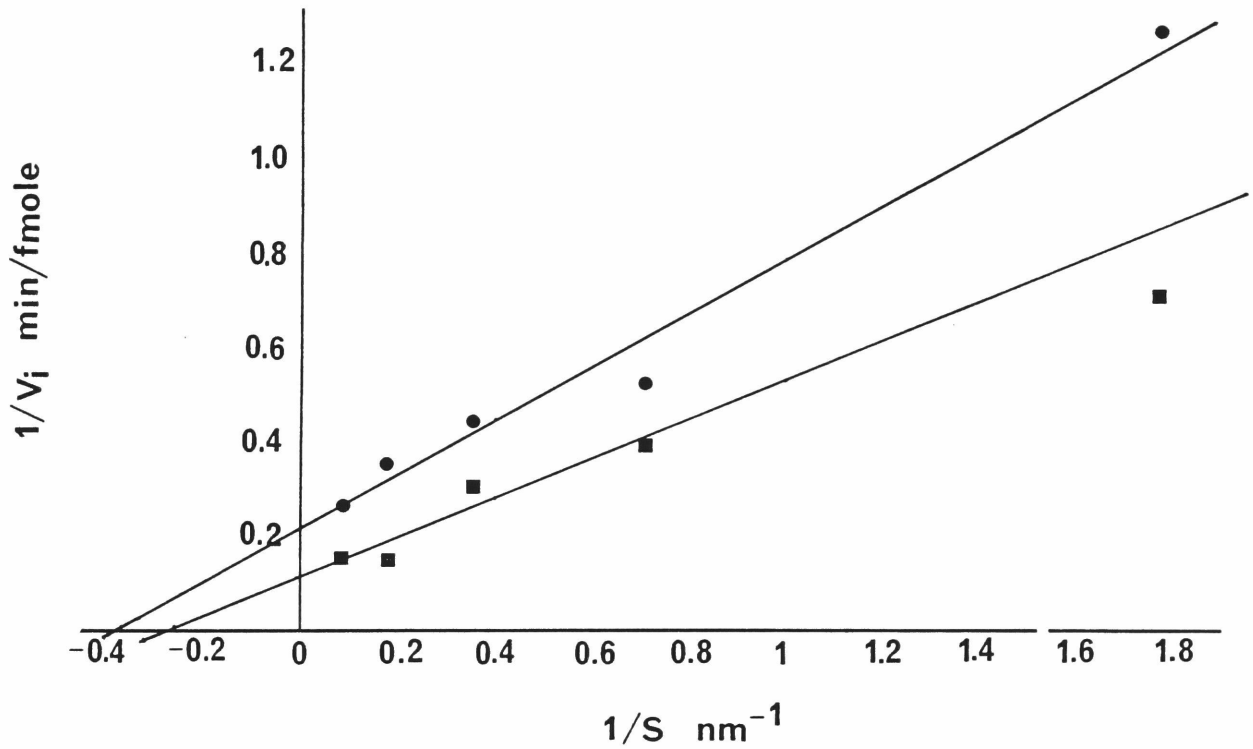

Time:
0 5 1 2 3 6 0 5 1 2 3 6 min.
0 0 0 0 0 0 0 0
D
RFII
RFIII
RFI
Lane:
a b c d e f g h i j k l m n o p q r

A


Time:
0 5 1 2 3 6 0 5 1 2 3 6 min.
0 0 0 0 0 0 0 0
D
RFII
RFIII
RFI
Lane:
a b c d e f g h i j k l m n o p q r

B

Figure 6-6. Kinetics of DNA Scission by the Wild-Type and the R200K Mutant *Eco*RI Endonuclease. From restriction digests (some of which are shown in Figure 6-5) with a constant amount of enzyme and varying substrate concentration, the amount of product formed was plotted versus time and the initial velocity of DNA scission was determined. For these calculations, the amount of the more predominant product was used (for the wild-type this was the RFIII product, while for the R200K mutant it was the RFII product). This graph shows the double reciprocal plot of V_i verus substrate concentration. The x- and y-intercepts yield the values for K_m and V_{max} respectively (listed in Table 6-5).


1.2
1.0
0.8
0.6
0.4
0.2
1/Vi min/fmole
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
1/S nm−1

ments. As described in table 6-6, we compared the R200K and R200C mutations in the TS6 mutant background (double mutants) to triple mutants which also carried the enhanced star activity mutation. By monitoring SOS induction we observe that the TS6 mutation reduces the activity of both the R200K and R200C mutants as would be expected from its TS phenotype. However in the presence of an enhanced star activity mutation, *in vivo* DNA scission by the R200C mutant is enhanced (see table 6-6). For example, the R200C+TS6 double mutant makes colonies at 30°C which are induced for the SOS response (dark blue colonies). When this mutant bears an additional mutation, for example either H114Y or A138T, the enzyme activity is increased and no colonies are produced at 30°C. Expression of the *Eco*RI methylase completely blocks this *in vivo* DNA scission, and enhanced cleavage therefore occurs exclusively at the wild-type recognition site. Thus DNA cleavage by the R200C mutant enzyme is increased *in vivo* by enhanced star activity mutations as it is *in vitro* by star buffer conditions. In contrast, although the R200K mutant is activated *in vitro* by star buffer, it is not activated *in vivo* by star activity mutations.

These findings suggest that these mutants differ in their substrate recognition mechanism as follows. For the wild-type enzyme, both star buffer and increased star activity mutations enable the enzyme to cleave at *Eco*RI* sites such as AAATTC. Since the A substitution within this site disrupts the enzyme's normal contact to the O_6 position of the first guanine by replacing it with an amino group, star buffer and star mutations must somehow substitute for this contact. Similarly, the R200K and R200C mutants are disrupted in their contacts to the outer guanine of the wild-type site (GAATTC). Star buffer can partially mitigate this defect for both enzymes, suggesting that these conditions provide a contact that both mutants lack. On the other hand, increased star activity mutations only activate the R200C mutant, suggesting that they mimic or provide a contact that the R200C mutant lacks but R200K retains. For example, the R200K mutant may make one or both of the hydrogen bonds that the arginine of the wild-type enzyme makes, while the R200C mutant may make

Table 6-6. Increased Star Activity Mutations Suppress the R200C Allele.

Allele	42	37	M^- 34	30	M^+ 30-42°C
R200K	dead	dead	dead	dead	W
R200K + R56Q	W-LB	MB	DB,sick	dead	W
R200K + R56Q + A138V	W-LB	MB	DB,sick	dead	W
R200K + R56Q + H114Y	W-LB	MB	DB,sick	dead	W
R200K + R56Q + A138T	W-LB	MB-DB	DB,sick	dead	W
R200C	LB	MB	MB	DB	W
R200C + R56Q	W-LB	W-LB	LB	DB	W
R200C + R56Q + A138V	W-LB	LB	MB-DB	DB,sick	W
R200C + R56Q + H114Y	W-LB	MB	DB,sick	dead	W
R200C + R56Q + A138T	W-LB	MB	DB,sick	dead	W

w=white or faint blue, LB=light blue, MB=medium blue, DB= dark blue colonies on X-Gal indicator plates containing 35 μg / ml X-Gal.

only one or no hydrogen bonds.

Discussion

We find that mutagenesis of amino acids implicated in substrate specificity of the *Eco*RI endonuclease cripples enzyme activity but does not alter substrate specificity *in vivo* or *in vitro*. Since the enzymatic activity of these mutants is dramatically reduced compared to the wild-type enzyme, these amino acids clearly play an important role in enzyme action as predicted by the crystal structure. With conservative amino acid substitutions (Wolfes, *et al.*, 1986; Greene, *et al.*, 1987), one can argue that the mutant makes similar but weaker contacts compared to the wild-type protein. However, we observe that several mutants with non-conservative amino acid changes still recognize the wild-type substrate. In these cases the wild-type amino acid has been substituted by a smaller amino acid which may simply be compatible with the existing protein structure and make no direct contact to the substrate. At the R200 position for example, substitution of lysine, cysteine, serine, and valine are compatible with enzyme activity. Since valine can make no hydrogen bonds, while in the wild-type protein arginine 200 makes 4 of the 12 hydrogen bonds to the substrate, the enzyme can recognize its substrate in the absence of these interactions. These findings are in agreement with studies that show *Eco*RI cleaves (at reduced rate) oligonucleotides which contain unusual nucleotides lacking functional groups thought to contact the enzyme (such as tubercidin which lacks the N_7 group of guanine) (Seela & Driller, 1986; Seela & Kehne, 1987). We conclude that in addition to the hydrogen bond interactions predicted from the crystal structure model, the enzyme must make other contacts to recognize its substrate.

Part of this sequence discrimination may arise simply from the ability of certain DNA sequences to adopt the unusual secondary structure seen in the crystal structure. Comparison of crystal structures of the DNA in the presence (Frederick, *et al.*, 1984; McClarin, *et al.*, 1986) and the absence of *Eco*RI (Wing, *et al.*, 1980; Dickerson

& Drew, 1981; Drew & Dickerson, 1981) reveals that although the DNA alone is slightly bent, significant distortions of the DNA structure occur when the protein is bound. For example, the major groove widens to allow the protein's recognition machinery access to the nucleotides and to accommodate this, the minor groove becomes more narrow. The four central AT base pairs may facilitate this structural perturbation since, unlike the 2-amino group of a GC basepair, AT basepairs have no functional groups that project into and inhibit narrowing of the minor groove. Furthermore, the central AT basepairs are displaced towards each other to facilitate the glu144/arg145 bridging interactions, the basepairs adjacent to the recognition site have an unusually high propeller twist, and the sugar-phosphate backbone is kinked both internal and external to the recognition site. *Eco*RI contacts to the sugar-phosphate backbone could facilitate or capture this structural perturbation and also serve to discriminate the substrate further. In the case of the 434 repressor, it has been argued that DNA secondary structure affects the complementary fit between protein and DNA by altering the position of contacted phosphates, functional groups in the major groove, and the two halves of the operator (Anderson, *et al.*, 1987). Sequence discrimination by the *trp* repressor is mediated almost exclusively by contacts to the DNA backbone (Otowinski, *et al.*, 1988). It would not be surprising if *Eco*RI took advantage of similar interactions as well.

In addition, other evidence suggests that the *Eco*RI enzyme may contact the pyrimidines of its recognition site. Studies with oligonucleotides containing modified nucleotides reveal that altering the cytosines or thymidines of the recognition site severely inhibits or blocks cleavage by the enzyme (Brennan, *et al.*, 1986a; McLaughlin, *et al.*, 1987). Substitution of 5-methyl cytosine for cytosine at the outer basepairs of the *Eco*RI recognition site renders the octanucleotide GGAATTCC inactive as a substrate (Brennan, *et al.*, 1986a). With longer substrates, the 5-methyl group slows the rate of cleavage by the enzyme about 100-fold (Tasseron-de Jong, *et al.*, 1988). Apparently longer substrates or different flanking sequences enable the enzyme to

tolerate the methylated substrate poorly. Alternatively, the enzyme may diffuse onto the site on longer substrates rather than bind directly as it must on short substrates. The effect of these modified nucleotides could be to distort the structure of the DNA or prevent it from assuming the kinked and bent conformation which may be required for cleavage. Alternatively the *Eco*RI enzyme may make a steric or hydrophobic contact to these cytosine residues. The presence of the 5-methyl group would then block binding or cleavage by the enzyme. This interaction would discriminate against recognition of thymidine with its 5-methyl group and against guanine and adenine, each with a nitrogen of the imidazole ring at this position. Similar arguments also support recognition of the thymidines, since substitution by uracil decreases, or for some substrates blocks, scission by the enzyme (Brennan, *et al.*, 1986a; McLaughlin, *et al.*, 1987). Thus the wild-type enzyme may contact both members of each basepair. In the E144X, R145X, and R200X mutants, recognition of the purine nucleotides is disrupted, but because base-pairing is complementary, contacts to the pyrimidines of the recognition site would allow the enzyme to still recognize the wild-type basepair. Examination of the *Eco*RI enzyme-DNA crystal structure reveals a number of hydrophobic amino acids in proximity to the substrate pyrimidines (I179, I181, V189, L191, I197, L198, L10). Since the present protein-DNA complex is a static intermediate in the kinetic pathway, conformational changes may occur during cleavage which bring these amino acids in closer proximity to bind substrate.

It has been suggested that two contacts per basepair is necessary and sufficient for a DNA binding protein to discriminate a basepair uniquely (Seeman, *et al.*, 1976). A model of substrate recognition by the *Eco*RI restriction enzyme, based on both the rate and specificity of star cleavage activity (Rosenberg & Greene, 1982) and the DNA-enzyme co-crystal structure (McClarin, *et al.*, 1986), suggests that the enzyme binds its substrate by making two hydrogen bonds per base pair. Most models of repressor-DNA interactions invoke two or only one specific contact per base pair. In some cases single hydrophobic or hydrogen bond interactions appear sufficient for

unique discrimination. For example, a mutant of the 434 repressor which recognizes an AT instead of a GC basepair has a hydrophobic amino acid substitution (Q43A) which is proposed to bind the 5-methyl group of the thymidine (Wharton & Ptashne, 1987). Our analysis of the substrate specificity of *Eco*RI mutants lacking specific functional groups which interact with the enzyme suggests that the wild-type *Eco*RI endonuclease effectively makes three contacts per basepair. In contrast to repressors where some binding of non-cognate sites can be tolerated and binding to closely related operators is required, restriction enzyme substrate recognition may utilize more contacts than seems necessary as a fail-safe mechanism to reduce cleavage of non-canonical substrates.

As will be described in more detail in chapter 7, it has been proposed that DNA cleavage by the *Eco*RI endonuclease is allosterically activated by substrate binding (McClarin, *et al.*, 1986; King, *et al.*, 1986, 1988). The idea that *Eco*RI is allosterically activated is based on the phenotypes of several mutant enzymes. The most notable are several (E111G, A, V, or Q) which bind substrate as tightly as (E111G, A, V) or more tightly than (E111Q) than the wild-type enzyme but are virtually unable to cleave DNA (King, *et al.*, 1986, 1988). The afflicted amino acid lies far from the DNA and may disrupt the allosteric conformation mechanism by blocking the required conformational change. As will be described in detail in chapter 7, we have isolated a series of mutants which cleave *Eco*RI* sites under normal buffer conditions. One of these (H114Y) lies three amino acids away from the E111X mutants which bind but do not cleave. Our mutation may behave in the opposite fashion and render the protein more easily allosterically activated such that other DNA sequences now act as allosteric effectors and trigger cleavage. Since our mutants which show enhanced star activity are most likely altered in allosteric activation, star buffer conditions may similarly provoke the allosteric conformational change.

The substrate binding site mutants described here are activated by star buffer conditions and in some cases by enhanced star activity mutations. Because their de-

fect is mitigated by conditions that we suggest favor the allosterically activated conformation, they may be defective in allosteric activation. This suggests that the residues missing in the mutant proteins play a role in activation. Since they normally bind to the substrate, one possibility is that the wild- type amino acids contact the substrate and initiate the allosteric conformational change. In this way these amino acids would serve a dual function to bind substrate and trigger DNA cleavage.

In conclusion, we suggest that the high fidelity of substrate recognition by the *Eco*RI endonuclease is the result of extensive contacts between the enzyme and its substrate, more than should be required on theoretical grounds to uniquely recognize a particular DNA sequence. Furthermore, cleavage by the enzyme may be allosterically activated upon substrate binding to provide a second check on substrate specificity and decrease the rate at which improper substrates suffer scission.

Chapter 7

Mutants of the *Eco*RI Endonuclease with Promiscuous Substrate Specificity

Introduction

The study of DNA molecules was revolutionized by the invention of recombinant DNA technology (Cohen, *et al.*, 1973). Our ability to manipulate DNA in this way depends upon several special types of enzymes. Central among these are the restriction enzymes, which recognize and cleave specific DNA sequences (Roberts, 1976). The best characterized example is the *Eco*RI endonuclease, whose substrate is the DNA sequence GAATTC (Hedgpeth, *et al.*, 1972). We have studied this enzyme as a model of DNA-protein interactions and to understand how it attains its high fidelity of substrate recognition.

The X-ray crystal structure of an *Eco*RI-DNA complex (Frederick, *et al.*, 1984; McClarin, *et al.*, 1986) reveals that this protein is not structurally homologous to DNA-binding proteins that carry either the helix-turn-helix (Pabo & Sauer, 1984) or zinc finger motifs (Miller, *et al.*, 1985). Instead, each monomer of the *Eco*RI dimer bears two α-helices which project steeply into the major groove of the DNA. These helices carry amino acids which hydrogen bond to the nucleotides of the recognition site and are thought to mediate substrate specificity (McClarin, *et al.*, 1986).

Earlier we found that *in vivo* DNA scission by the *Eco*RI endonuclease induces the SOS.DNA repair response (Heitman, *et al.*, 1989; Chapter 5). Armed with this *in vivo* assay that monitors DNA scission by induction of an SOS::*lacZ* fusion, we showed that alterations of the amino acids implicated in substrate binding by the crystal structure do not alter substrate recognition (Chapter 6). Thus the hydrogen bond network revealed by the crystal structure is not sufficient to explain substrate specificity of the *Eco*RI endonuclease, and additional interactions must direct the specific interaction.

Here we have isolated *Eco*RI endonuclease mutants which damage the DNA of their host cells and induce the SOS response, even though the normally protective *Eco*RI methylase is present. *In vitro* under normal *Eco*RI buffer conditions the purified mutant proteins preferentially cleave the wild-type substrate but also cleave additional sites which differ by one nucleotide from the normal substrate. These sites are also cleaved by the wild-type enzyme under altered buffer conditions. This is called *Eco*RI* activity (Polisky, *et al.*, 1975; Gardner, *et al.*, 1982; Rosenberg & Greene, 1982). Hereafter we refer to these mutants as star mutants to indicate that they exhibit *Eco*RI* activity under normal *Eco*RI buffer conditions.

Four of the five star mutants lie at the DNA-protein interface. These mutations may identify residues in the wild-type enzyme which normally interact with substrate and are unable to do so in their mutant form. Alternatively, the amino acid present in the mutant may perturb the DNA binding surface and impede substrate recognition. The other star mutation lies far from the DNA binding or cleavage sites. By analogy with another nearby mutation that disrupts DNA scission without affecting DNA binding (E111G, King, *et al.*, 1986, 1988), we suggest that this mutation lies in a region of the protein which mediates a conformational change that allosterically couples binding to scission. The H114Y mutation may shift the enzyme conformation towards its activated form or render this change more facile and allow other DNA sequences to act as allosteric effectors and trigger cleavage. The *Eco*RI enzyme appears to achieve its high fidelity of substrate recognition by making extensive contacts to the substrate (chapter 6) and also through an allosteric activation mechanism that acts as a second check of substrate identity.

Results

Previously we have shown that *in vivo* DNA scission by either the Mrr, McrB, or *Eco*RI restriction systems induces the SOS DNA repair response in *E. coli* (Heitman & Model, 1987; Heitman, *et al.*, 1989; chapters 3 and 5). Although we found that in-

duction of the SOS genes is not required for the repair of *Eco*RI staggered breaks (Heitman, *et al.*, 1989; Chapter 5), these observations nonetheless suggested that one could either screen for or assay endonuclease mutants by monitoring SOS induction. This assay measures endonuclease activity *in vivo* by its ability to sponsor increased β-galactosidase expression from an SOS::*lacZ* fusion. In chapter 6 this SOS induction assay was used to show that mutations which alter three amino acids of the substrate binding pocket of the *Eco*RI enzyme do not alter substrate specificity. Additional interactions must play a role in the specific DNA-protein interaction. Here we have screened for mutations that affect other amino acids involved in substrate recognition.

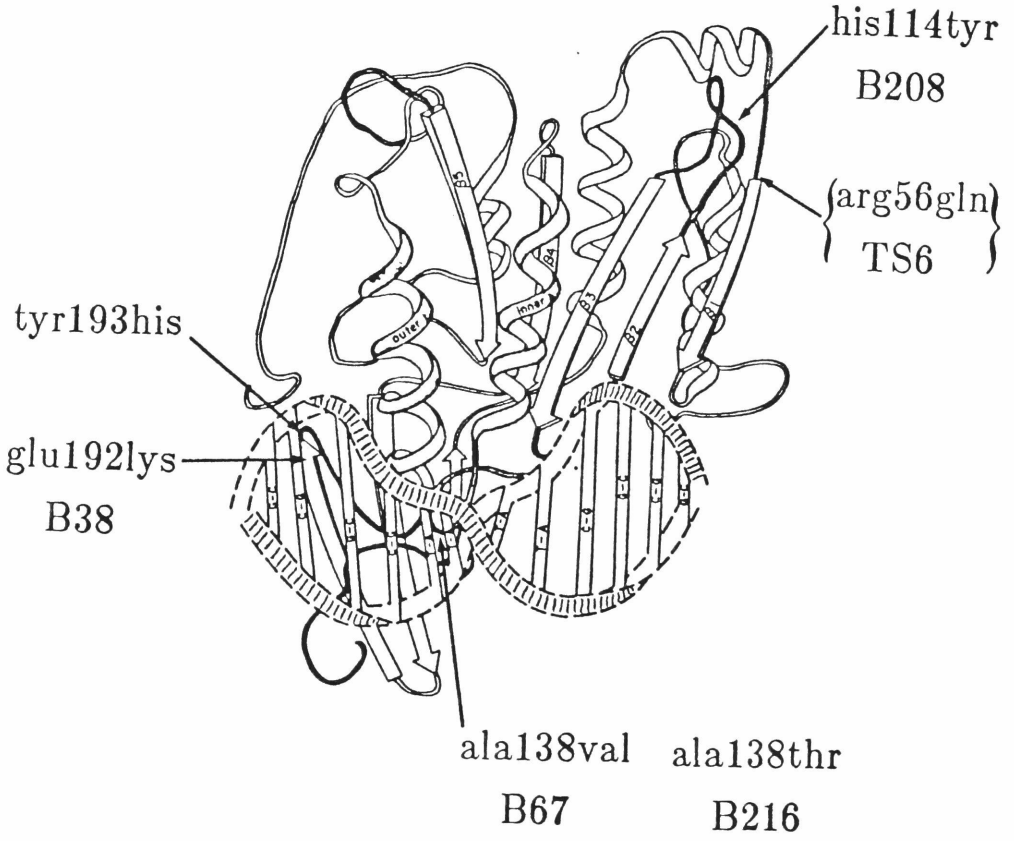
We screened for *Eco*RI mutants which cleave DNA and induce the SOS response in spite of the presence of the *Eco*RI methylase. We expected that some of these alleles would be disrupted or altered in substrate recognition. A plasmid bearing the endonuclease gene was mutagenized *in vivo* with nitrosoguanidine. Purified plasmid DNA was introduced into a strain that carries an SOS::*lacZ* fusion and a plasmid expressing the *Eco*RI methylase (pJC1), and the resulting transformants were plated on X-gal indicator medium to screen for SOS induced colonies. Our initial attempts were hampered by a high background of spontaneous blue colonies ($\sim$ 1%). These isolates were unstable and segregated blue and white colonies. Since the endonuclease and the methylase plasmids share homology, we thought this might be attributable to recombination between the two. In fact when the the region of homology was reduced, the frequency of blue isolates decreased. We suggest that these unstable blue colonies carry interplasmid recombinants that subsequently resolve and give a variegated phenotype.

To circumvent these difficulties and screen for potentially lethal alleles, we started with a genetically marked (temperature-sensitive) *Eco*RI endonuclease allele. By first isolating blue colonies and then screening amongst these for those in which the blue phenotype was stable and TS, we could quickly determine in which isolates the

SOS inducing phenotype was linked to *Eco*RI action. From a larger collection of *Eco*RI temperature sensitive mutants we chose one, TS6, because it is fully inactive at 42°C and nearly as active as wild-type at 30°C, yields substantial endonuclease activity *in vitro*, and the mutation it carries is near the N-terminus which was convenient for mapping additional mutations (Chapter 4). A plasmid bearing the TS6 mutant endonuclease gene (pJH15b) was mutagenized with either nitrosoguanidine or by growth in a *mutD* strain (JM103mutD). We then screened for plasmid transformants which produced blue colonies at temperatures partially and fully permissive for endonuclease action (37°C, 34 °C, 30°C). The screen at intermediate temperatures ensured that mutants which would be lethal at lower temperatures were not selected against. About 30 mutants were obtained by this means. One additional allele (B208) was found by replica-plating mutants ($\sim$ 20,000) and looking for ones unable to grow at 30°C. This task was facilitated by pre-incubating indicator plates for 5 hrs at 42°C, and then looking for blue ghost colonies after subsequent growth at 30°C.

The resulting mutants ($\sim$ 30) were all plasmid-linked and mapped within one of the four quarters of the endonuclease gene by heteroduplex deletion mapping. This region was then sequenced by the dideoxy chain termination method with denatured plasmid DNA as template. From this set of mutants four different amino acid substitutions were found. To confirm that no other extraneous mutations had occured, the entire gene was sequenced for four representative isolates. All contained the original TS6 mutation, arg 56 to gln. Three of the four had a single additional nucleotide substitution which resulted in the amino acid alterations shown on figure 7-1 (A138V, H114Y, E192K). The fourth carried two coding changes, A138T and V166I. This mutant was unusual and segregated two types of colonies during growth at 30°C: small, dark blue colonies and medium size, medium blue ones. Recombination between the pJH15a mutant plasmid and the wild-type endonuclease sequence present on plasmid pJC1 could account for this phenotypic variation. Since the

Figure 7-1. Positions of Amino Acid Substitutions that Allow the *Eco*RI Endonuclease to Evade the *Eco*RI Methylase. As described in the text, we screened for *Eco*RI endonuclease mutants that would damage the DNA of the host cell and induce the SOS response even though the *Eco*RI methylase was present. Because these are potentially lethal mutants, a temperature-sensitive (TS6) allele was used to conditionally express these mutants. Thus each of these mutants bears two amino acid substitutions: the R56Q temperature-sensitive mutation (shown in brackets) and the other indicated amino acid changes. The structure shown here is derived from McClarin *et al.*, 1986.


his114tyr
B208
arg56gln
TS6
tyr193his
glu192lys
B38
ala138val
B67
ala138thr
B216

A138V mutant had already been identified, it seemed likely that the SOS-inducing phenotype of the A138T + V166I double mutant was the result of the A138T substitution. In fact when we separated the two we found the A138T mutant became more active and was clearly responsible for the SOS-inducing phenotype.

Figure 7-2 shows a photograph of colonies expressing these *Eco*RI mutants. The phenotypes of these four mutants are also listed in table 7-1. The host strain carries an SOS::*lacZ* fusion and the media contains a chromogenic substrate (X-gal) which is cleaved by β-galactosidase to give a blue dye. Blue color is thus indicative of SOS induction, in this case as a response to *in vivo* DNA scission. This strain also expresses the *Eco*RI methylase and one of several *Eco*RI endonuclease mutants. These strains were grown at four temperatures; incubation at lower temperatures activates the mutant endonucleases. The upper left quadrant of each plate shows that the parent TS6 mutation does not make blue colonies at any temperature, indicating that its activity is completely blocked by the methylase. The other quadrants show three mutants which damage the DNA and activate SOS induction. At 42°C where the TS endonuclease is largely inactive, these mutants induce little or no SOS. As the temperature is decreased to activate endonuclease action, the SOS response is induced even though the methylase is present. In some way these mutant enzymes evade the protective action of the methylase and damage the DNA of the host cell. None of the mutants are fully lethal at 30°C, but the B208 (shown in figure 7-2) and B216 alleles are detrimental to cell growth. The methylase does protect the cell from at least some of the mutant endonuclease action, since in its absence these mutants are fully lethal at 30°C. This suggests that it is not the methylated site which has become the target of the endonuclease. Compared to the lethal effect of the TS6 parent, each mutant is somewhat reduced in lethal action. These *in vivo* observations suggest these mutants retain at least some cleavage activity at the WT site, and that the additional mutations they bear only slightly impair their activity compared to the parent TS6 mutant.

Figure 7-2. *Eco*RI Endonuclease Mutants that Induce SOS. As described in results, *Eco*RI endonuclease mutants were isolated which induce the SOS response in spite of the presence of the *Eco*RI methylase. These potentially lethal mutations were isolated in a TS endonuclease background (TS6) to allow conditional expression. This figure shows the SOS::*lacZ* fusion host strain JH137 expressing these mutants and the methylase. The media contains the chromogenic substrate X-gal, so a blue colony color is indicative of endonuclease activity. Growth was at 42°C (no endonuclease activity) or lower temperatures (30, 34, or 37°C) to activate the mutant endonuclease. The upper left panel shows the TS6 parent mutant which does not induce SOS at any temperature because its action is blocked by the methylase. In contrast, the other quadrants of the plates show mutants in which the endonuclease evades the protective action of the methylase to damage the DNA of the host and induce the SOS response.

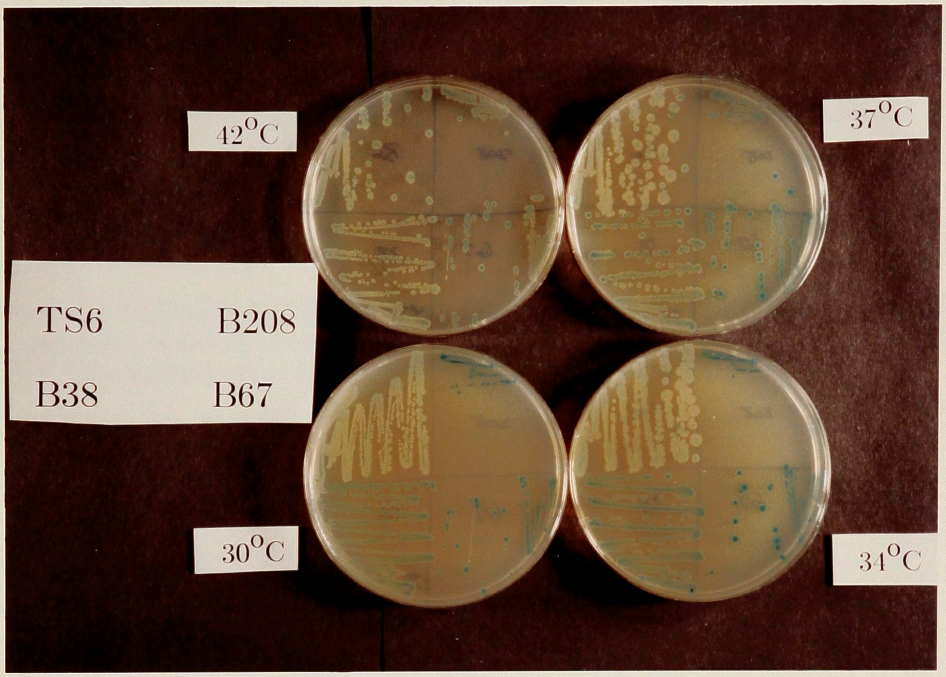

42°C
37°C
TS6
B208
B38
B67
30°C
34°C

Table 7-1. *Eco*RI Mutants which Induce SOS Even in the Presence of the Methylase

Allele:	WT	TS6	B38	B67	L2*	B208	L4**
Temp							
42°C	W-LB	W-LB	W-LB	LB	W-LB	W-LB	W-LB
37°C	W-LB	W-LB	LB	MB	LB	LB-MB	LB
34°C	W-LB	W-LB	LB-MB	MB	DB	DB,sick	MB-DB
30°C	W-LB	W-LB	LB	MB	dead	DB,sick	dead
Mutations:							
		R56Q	R56Q	R56Q	R56Q	R56Q	R56Q
			E192K		E192K		E192K
				A138V	A138V		
						H114Y	H114Y

*L2=B38 + B67

**L4=B38 + B208

SOS induction was assayed by measuring β-galactosidase production (using X-Gal indicator medium) from the *dinD1*::Mu dI(Ap^r *lac*) fusion borne by strain JH137. W=white colonies on X-gal (35 μg /ml) indicator media LB=light blue MB=medium blue DB=dark blue

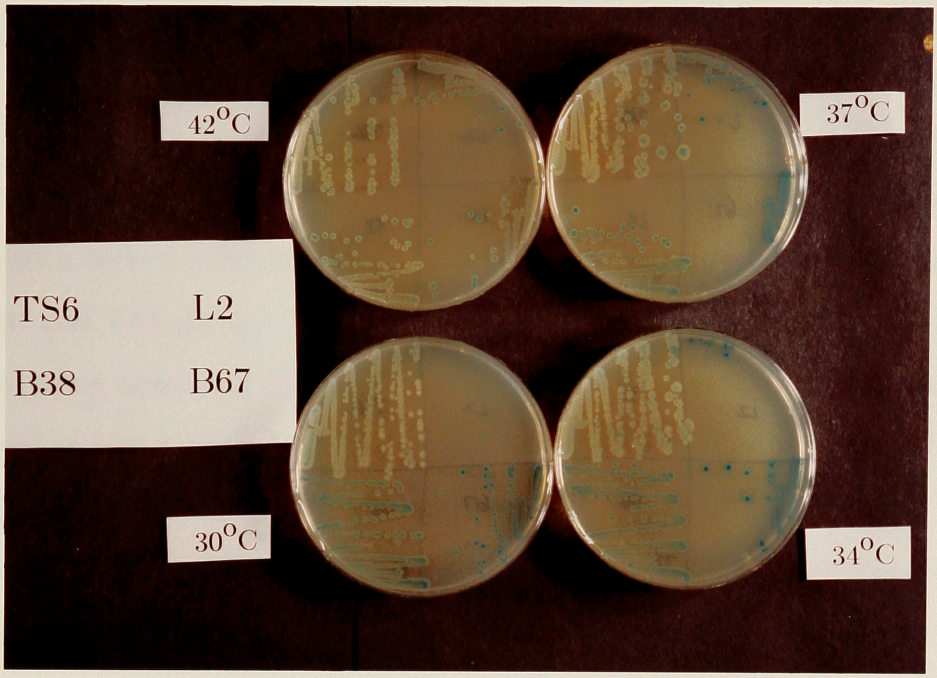
Since none of these mutants is fully lethal to the cell, we wondered if combining the mutations would enhance activity. Two triple mutants, L2=E192K, A138V, and R56Q and L4=E192K, H114Y, and R56Q were constructed (see table 7-1). As shown in figure 7-3, both of these triple mutants are lethal to the cell at 30°C and are clearly more active than their parents. Thus the phenotypes of the individual mutants are additive.

To determine if the TS mutation plays any role in the mutant phenotype, we separated two mutations (E192K and Y193H, described below) which bear the SOS inducing mutation in an otherwise wild-type background. The two alleles with the least activity were chosen to ensure that the constructs would be viable. In both cases the SOS inducing phenotype is independent of the TS mutation, and the isolated mutations induce the SOS response at all temperatures and thus do not themselves render the enzyme TS.

Purification of the mutant enzymes

We conclude from the preceding description that this set of *Eco*RI endonuclease mutants damages the cellular DNA even in the presence of the normally protective *Eco*RI methylase. To determine the *in vitro* cleavage activity which corresponds to this *in vivo* phenotype, the mutant enzymes were purified. The TS6 mutant was originally chosen to conditionally express these mutants because it yields endonuclease activity in cellular extracts. To determine if these mutants behave similarly, cultures of strain JH137 expressing an endonuclease mutant and the methylase were grown at 42°C until $OD_{600}=1$ and then shifted to 30°C for 8 hours. Cellular extracts from these cultures revealed *Eco*RI specific endonuclease activity for all of the mutants. Additionally, extracts of the L2 mutant clearly cleaved plasmid pACYC177 at one unique site even though this plasmid has no *Eco*RI sites. Larger cultures (2l) of each strain were grown by the same protocol and the mutant enzymes were purified by phosphocellulose and hydroxylapatite chromatography as described (Cheng, *et al.*,

Figure 7-3. Double Mutants are Fully Lethal. None of the original SOS-inducing endonuclease mutants were fully lethal to the host cell. To see if the activities of the single mutations are additive, some of the single mutations were recombined together *in vitro*. The host strain and plating conditions are as described in Figure 7-2. Again, the upper left quadrant shows the TS6 mutation. The lower two quadrants show two single mutants (E192K and A138V) which were recombined together to yield the L2 mutant shown in the upper right quadrant. (L2 is actually a triple mutant since it also bears the original TS6 mutation (R56Q).) As shown here, the L2 mutant is now fully lethal at 30°C were the activity of the mutant endonuclease is maximal. Thus the phenotypes of the single mutants are additive.


42°C
37°C
TS6
L2
B38
B67
30°C
34°C

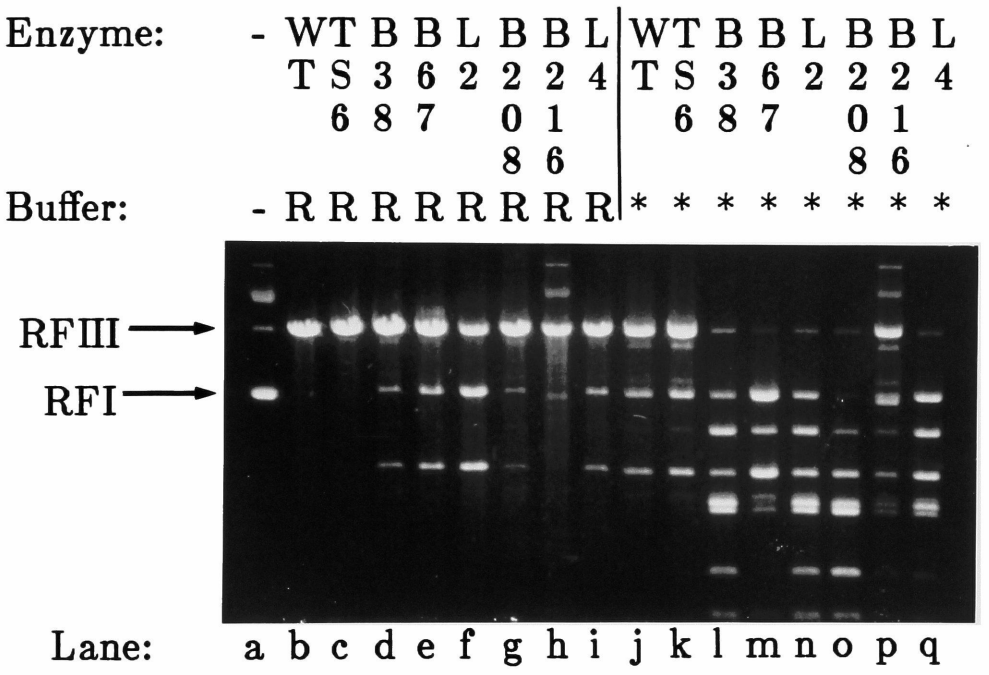
1985). Endonuclease activity in column fractions was assayed with λ DNA as substrate, or in the case of the L2 mutant, f1 DNA. The wild-type and mutant proteins eluted from both columns at similar salt concentrations, indicating that the structure of the mutant proteins is not markedly different from the wild-type enzyme. The purified mutant proteins are $> 50\%$ homogeneous, as estimated from coomassie blue stained SDS-polyacrylamide gels. By this means 100-200 μg of the wild-type enzyme, the TS6 parent, and the B38, B67, B208, B216, L2 mutants were obtained. The yield for the L4 triple mutant was low ($\sim$ 20μg), most likely because of the extremely lethal effect of this allele.

Cleavage specificity of the mutant enzymes

To determine the cleavage specificity of these purified endonucleases, DNA digests were performed with several substrates of known DNA sequence. As shown in figure 7-4, the wild-type and TS6 mutant enzymes cleave plasmid pBR322 uniquely at its sole *Eco*RI site to yield one linear fragment. In contrast, the *Eco*RI mutant proteins cleave the plasmid at two sites to produce two fragments. By appropriate double digests (data not shown), these mutant proteins preferentially cleave at the wild-type *Eco*RI recognition site. Cleavage occurs more slowly ($\sim$ 10-fold) at a second site to yield DNA fragments of 1600 and 2400 bp which co-migrate with DNA fragments observed after DNA scission by the wild-type enzyme under *Eco*RI* buffer conditions (lane j). This cleavage site maps to the position of a previously described *Eco*RI* site, gGAAGTCa (nucleotides 1637-1643) (Gardner, *et al.*, 1982).

When plasmid pBR322 DNA is incubated with these mutant enzymes under *Eco*RI* buffer conditions (lanes k to q) cleavage occurs at a number of additional sites. An identical pattern of fragments is observed with the wild-type enzyme after prolonged inducation under *Eco*RI* reaction conditions (not shown here). We conclude that these mutants exhibit enhanced cleavage activity at *Eco*RI* sites under reaction conditions where the wild-type enzyme exhibits no *Eco*RI* action. Under

Figure 7-4. Cleavage of pBR322 DNA by the Wild-type and Star Mutant Endonucleases. Plasmid pBR322 DNA (200 ng per reaction) was incubated with the purified wild-type or star mutant endonucleases for 1 hour at 30°C in either normal *Eco*RI buffer conditions (lanes b to i) or *Eco*RI* buffer conditions (lanes j to q). The cleavage products were then electrophoresed on a 0.6% agarose gel containing 0.5 μg/ml EtBr.


Enzyme: - WT WT B38 B67 L2 B2208 B2216 L4 WT WT B38 B67 L2 B2208 B2216 L4
Buffer: - R R R R R R R R * * * * * * * *
RFIII
RFI
Lane: a b c d e f g h i j k l m n o p q

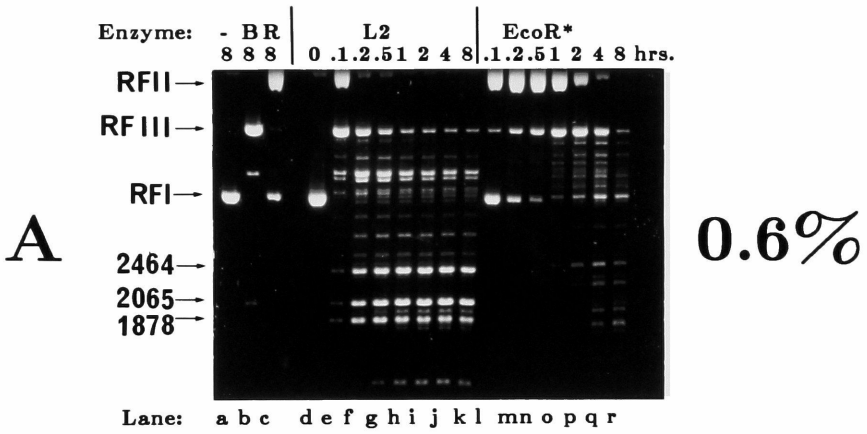
*Eco*RI* buffer conditions, the activity of the mutant proteins is increased further. From the pattern of partial digestion (lanes m and p) it appears that the B67 and B216 mutant enzymes cleave less readily at some *Eco*RI* sites, suggesting the mutants may differ in substrate preference.

Digestion of λ by these mutant proteins initially yields a pattern of fragments indistinguishable from those produced by the wild-type protein. With prolonged incubation a number of additional sites are cleaved to yield a pattern of fragments, many of which co-migrate with *Eco*RI* fragments. The L2 and L4 triple mutants showed less preference for scission at the wild-type site. Thus as with pBR322, the wild-type recognition sites present on λ remain the preferred substrates of these mutants, but under buffer conditions where the wild-type does not cleave *Eco*RI* sites, these mutants do. The patterns of λ partial cleavage were again somewhat different suggesting that the various mutants may differ in substrate preference. However, the large number of cleavage sites and the complexity of the restriction pattern hindered further characterization.

We turned to f1 DNA as a substrate because its sequence is known, it has no wild-type *Eco*RI sites, it contains more *Eco*RI* sites than pBR322 because it is larger and more AT rich, and it is smaller than λ and yields a less complex restriction pattern. Figure 7-5 shows digestions of f1 RFI form DNA with the wild-type *Eco*RI enzyme and the time course of partial digestion by the *Eco*RI* activity of the wild-type protein versus the L2 mutant (R56Q + A138V + E192K). Although f1 has no *Eco*RI sites, it is nonetheless nicked by the wild-type enzyme. Nicking occurs at several *Eco*RI* sites (see chapter 9). In contrast, the L2 mutant enzyme cleaves rapidly to yield three prominent fragments. Several additional sites are cleaved more slowly. These three cleavage sites map to within 20 bp of *Eco*RI* sites. A similar pattern of preferential cleavage was observed with the B67 (R56Q + A138V) and B216 (R56Q + A138T) star mutant enzymes. Compared to the star mutant proteins, DNA scission by the wild-type *Eco*RI* activity proceeds more slowly, does not produce the three

Figure 7-5. Restriction Digests of f1 DNA with the Star Mutant L2 or *Eco*RI* Activity. Panel A. Lane a shows the untreated RFI form f1 DNA. After treatment with *Bam*HI a linear species arises due to scission at the one *Bam*HI site (lane b). Although f1 bears no *Eco*RI sites, when incubated with the wild-type *Eco*RI endonuclease in its preferred buffer conditions a significant amount of DNA nicking occurs (see lane c and chapter 9). Each of these three lanes contains 250 ng f1 DNA. For the digests shown in lanes d through k, 2 μg of f1 DNA were incubated with 25 ng of the purified L2 mutant protein in 400μl of normal *Eco*RI buffer at 30°C. At the indicated times, 250 ng of the DNA cleavage products were removed and the reaction was terminated with stop buffer. For the reactions shown in lanes l through r, 2 μg f1 DNA was incubated with 62.5 ng of the wild-type enzyme in *Eco*RI* buffer conditions. These cleavage products were then displayed on a 0.6% agarose gel containing 0.5μg/ml EtBr.

Panel B. 250 ng of f1 DNA was incubated with the indicated enzymes for 8 hours at 30°C to produce a more complete restriction digest. These reactions were run on a 2% agarose gel. Several of the bands resulting from cleavage by the L2 mutant protein co-migrate with fragments produced by wild-type *Eco*RI* digestion. B=*Bam*HI; R=*Eco*RI; F=*Hin*fI; H=*Hpa*II.


Enzyme: - B R L2 EcoR*
8 8 8 0 .1 .2 .5 1 2 4 8 .1 .2 .5 1 2 4 8 hrs.
RFII
RF III
RFI
2464
2065
1878
Lane: a b c d e f g h i j k l m n o p q r

A

0.6%


enzyme: - B R F H L R
2 *
1596
819/829
771
680
608
486
350
Lane: 1 2 3 4 5 6 7

B

2%

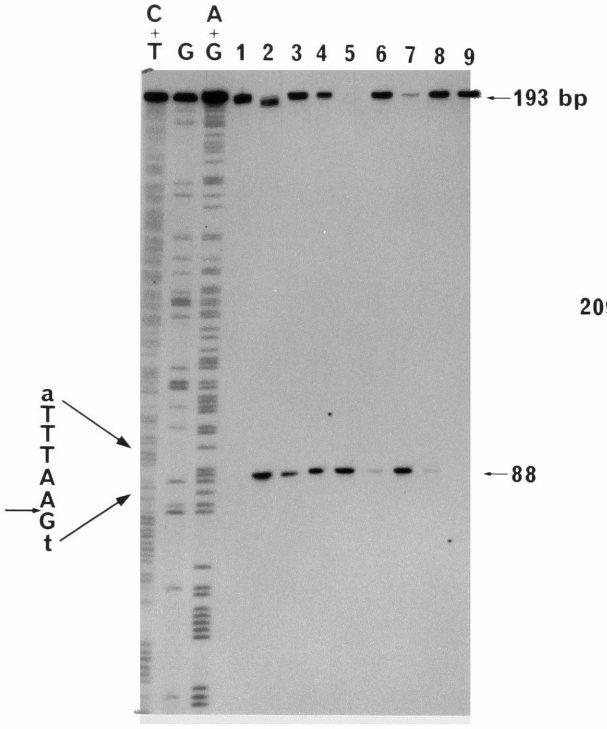
prominent fragments observed with these mutant proteins, and the pattern of partial digestion is much more complex. This suggests that the wild-type enzyme in *Eco*RI* conditions cleaves at more sites than the mutant enzyme and with lower specific activity. In contrast to the B67, B216, and L2 mutants, the B38 (R56Q + E192K) and B208 (R56Q + H114Y) star mutants yielded a partial pattern of digestion very similar to the wild-type *Eco*RI* activity, while the L4 mutant (R56Q + H114Y + E192K) yielded a pattern intermediate between the wild-type *Eco*RI* and the preferential cleavage activities.

By double digests and scission of small purified restriction fragments, we mapped these cleavage sites to within 10 to 20 bp. Each interval contains a site that corresponds to a consensus *Eco*RI* site, listed in table 7-2. To confirm these assignments, uniquely end-labelled f1 restriction fragments that span these sites were subjected to Maxam and Gilbert sequencing reactions and scission by the wild-type and mutant endonucleases. The reaction products were displayed on polyacrylamide sequencing gels and non-denaturing 8 or 12% polyacrylamide gels. One example is shown in figure 7-6. This 193 bp *Cla*I-*Aha*III fragment spans the preferred cleavage site at position 6125 (GAATTT). This fragment is cleaved by the wild-type enzyme in *Eco*RI* conditions or the mutant enzymes in normal buffer to yield an 88-bp cleavage product which migrates with or slightly slower than the fragment which corresponds to formic acid cleavage at the first adenine within the sequence GAATTT. Scission occurs between the guanine and the adenine within the *Eco*RI* site GAATTT, which differs from the wild-type recognition site in that a thymidine is substituted for the normal cytidine at the 3'-end. Scission produces a 3'-hydroxyl terminus because the enzymatic cleavage product migrates approximately one-half nucleotide slower than the Maxam and Gilbert cleavage product which bears a 3'-phosphoryl terminus. To determine if scission occurs on one or both strands, these reactions were electrophoresed on an 8%-polyacrylamide gel. Under these non-denaturing conditions, the fragments were still observed indicating that both strands

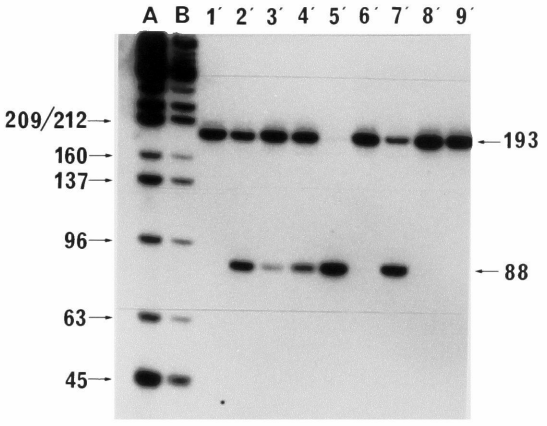
Table 7-2. DNA Sites Cleaved by *Eco*RI Star Mutant Endonucleases

Position	Sequence	Fragment	Comment
f1			
357	tGAAGTCt		
1499	gAAATTCa	*Hin*fI 1403* - *Hpa*II 1974 (521 bp)	
1674	aAAATTCa	*Hin*fI 1403* - *Hpa*II 1974 (521 bp)	
2182	tAAATTCa	*Hin*fI 2011* - *Hpa*II 2378 (367 bp)	strong site
2728	aAAATTCa	*Hae*III 2554 - *Hin*fI 2845* (291 bp)	
3789	aGAATTTg		cut poorly by A138V, A138T
4060	tAAATTCa	*Hpa*II 4018* - *Hae*III 5239 (1221 bp)	strong site
5912	gAAATTCa	*Ban*I 5677 - *Cla*I 6037* (360 bp)	Cut poorly by A138V, A138T
6125	tGAATTTa	*Cla*I 6037* - *Aha*III 6231 (194 bp)	strong site
pBR322			
4361	aGAATTCt		wild-type site
1637	gGAAGTCa		
pACYC177			
955	tGACTTCa		

Figure 7-6. Mapping Star Mutant Cleavage Sites to Single Nucleotide Resolution. As described in the results and in chapter 2 (materials and methods), uniquely end-labelled f1 restriction fragments were sequenced by the Maxam-Gilbert chemical method, subjected to DNA cleavage with the wt or star mutant endonucleases, and then electrophoresed under denaturing (panel A) or non-denaturing (panel B) conditions. Here we see a 193 bp *Cla*I-*Aha*III fragment which spans the recognition site at position 6125. Panel A shows the reaction products run on a 6% polyacrylamide - 8M urea sequencing gel. The left three lanes are the chemical sequencing reactions. Lane 9 shows the uncleaved sample, while lanes 1 through 8 are restriction digests with the wild-type or mutant enzymes. Scission occurs at one site to yield an 88-bp product, which is attributable to scission at the indicated position (arrow) in the sequence written to the left. Lanes: 1=WT enzyme in *Eco*RI normal buffer; 2=WT enzyme in star buffer; lanes 3-8 were all in normal *Eco*RI buffer; lane 3=B38; lane 4=B67; lane 5=L2; lane 6=B208; lane 7=B216; lane 8=L4. In panel B the same restriction digests were run on an 8% non-denaturing polyacrylamide gel. Lanes A and B are ^{32}P end-lagelled *Hin*fI f1 restriction fragments as size markers. Lanes 1' to 9' correspond to the same reactions shown in lanes 1 to 9 of panel A. The same 88-bp cleavage product also appears, indicating that DNA scission must have occured on both strands.


C + T
G
A + G
1 2 3 4 5 6 7 8 9
193 bp
88
a
T
T
T
A
A
G
t

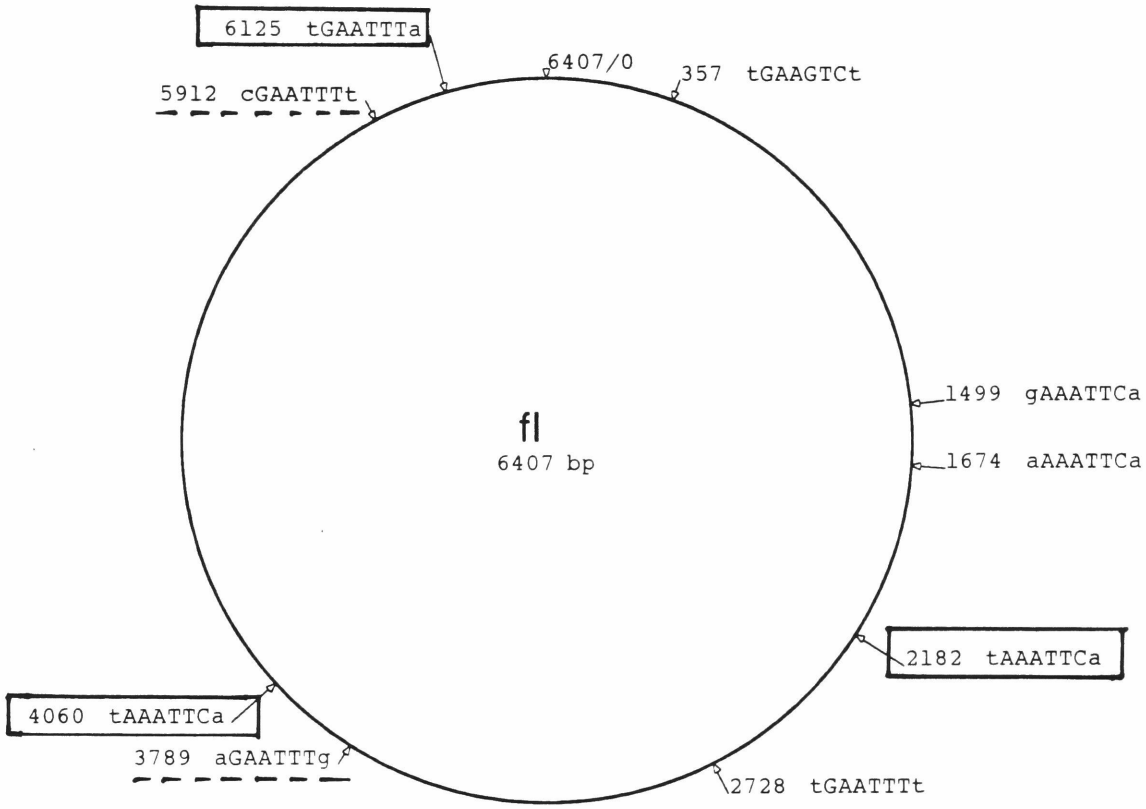
A


A B 1′ 2′ 3′ 4′ 5′ 6′ 7′ 8′ 9′
209/212
160
137
96
63
45
193
88

B

Figure 7-7. Sites of DNA Scission on the f1 Map. As described in the text, the sites at which the star mutant endonucleases cleave f1 were mapped and indicated here by arrows. These are all *Eco*RI* sites. The star mutant enzymes with a single star mutation exhibit two cleavage patterns. The B67 (A138V) and B216 (A138T) mutant enzymes prefentially cleave the boxed sites, cleave the sites with no special designation at intermediate rate, and do not cleave the dashed underlined sites. This yields the pattern of partial digestion seen in Figure 7-5 for the L2 mutant. In contrast, the B38 (E192K) and B208 (H114Y) mutant enzymes do not show this marked site preference and yield a pattern of digestion very similar to wild-type *Eco*RI* activity (see figure 7-5).

The L2 and L4 mutant star enzymes bear two star mutations and their cleavage pattern is intermediate between the two described above. The L2 mutant is a recombinant between the B67 and B38 mutants and while its activity most closely resembles the B67 parent (see figure 7-5), it also cleaves the dashed underlined sites like its B38 parent. The L4 mutant is a recombinant between the B38 and B208 mutants and although its two parents do not show the preferential cleavage pattern, this mutant enzyme cleaves f1 to yield the preferred pattern superimposed upon the wild-type *Eco*RI* pattern. These sites of scission are also listed in table 7-2.


6125 tGAATTTa
6407/0
357 tGAAGTCt
5912 cGAATTTt
1499 gAAATTCa
fl
6407 bp
1674 aAAATTCa
2182 tAAATTCa
4060 tAAATTCa
3789 aGAATTTg
2728 tGAATTTt

are cleaved. Additional cleavage sites were mapped in the same way and are shown on the f1 map in figure 7-6 and listed in table 7-2 (see also figure 9-1). The B67 and B216 mutants fail to cleave at two sites (3789, aGAATTTg; 5912, cGAATTTt) recognized by both the wild-type *Eco*RI* activity and the other mutant enzymes.

From a consideration of these cleavage sites we find that the mutant enzymes accept as substrate the wild-type site and *Eco*RI* sites containing certain substitutions: either an adenine instead of guanine at the first position (AAATTC) or the inverse of this site (GAATTT), or a substitution of the internal adenine by a cytosine (GACTTC) or its inverse (GAAGTC). The wild-type *Eco*RI* activity will tolerate other single base substitutions (see summary in Greene & Rosenberg, 1982). Additionally, the three f1 sites cleaved preferentially by some star mutant enzymes are all flanked by a thymidine at the 5'-end and an adenine at the 3'-end. The moderately cleaved sites have one or the other of these flanking nucleotides but not both, and the two poorest cleaved sites have neither. Although scission by the wild-type *Eco*RI enzyme shows some preference for sites flanked by AT basepairs (Thomas & Davis, 1975), we do not observe as marked an effect with wild-type *Eco*RI* activity compared to the star mutant enzymes (see figure 7-4).

Enhanced star activity mutations suppress EcoRI binding site mutations

In a study of the *Eco*RI substrate binding pocket, we generated a large number of site-directed mutants of amino acids E144, R145, and R200 (Chapter 6). In one case we found two isolates which both contained an R145K substitution but differed in phenotype. Both mutants were greatly reduced in activity because they survived in the absence of the methylase. On X-gal indicator medium to assay DNA scission by SOS induction, one isolate made light blue colonies and the other medium blue colonies. When the entire gene of both mutants was sequenced, an additional mutation was found in the more active allele. This spontaneous second-site suppressor

results from an amino acid substitution (Y193H) adjacent to one of the star activity mutations (E192K). To determine the phenotype of the Y193H mutation alone, the R145K mutation was changed to the wild-type amino acid (arginine) by site-directed reversion. The mutant bearing only the Y193H mutation exhibits a weak ability to induce the SOS response in cells expressing the *Eco*RI methylase. Although this mutant protein has not been characterized *in vitro*, based on its similar location and phenotype to the star mutants we suggest that the Y193H mutation represents a fifth star mutation.

This finding suggested that the star mutations might in general suppress other substate binding pocket mutations. By taking advantage of a convenient *Bgl*II site, we recombined together most of the possible substitutions of R200 with the A138V, H114Y, and A138T mutations. As shown in table 7-3, these enhanced star activity mutations suppress the decreased activity phenotype conferred by the R200C, V, and S mutations. Interestingly, the R200K mutant was not suppressed by any of the mutants. This was surprising because we previously found that the R200K mutant protein is clearly activated ~ 100-fold by *Eco*RI* buffer conditions. The enhanced star activity mutations must therefore differ in some way from the effect of star buffer conditions. Perhaps these mutations only afford the protein with a subset of conformational changes that *Eco*RI* conditions evoke. As we have argued elsewhere (Chapter 6), we suggest that the R200K allele may interact with substrate differently than the R200CVS mutants. For example, the R200K mutant may make one or both of the substrate contacts predicted for the WT amino acid (arginine), while the cysteine, valine, or serine substitutions may make only one or neither.

We also found that the star mutations rendered several inactive R200X mutants active. Thus for example, mutants containing methionine, threonine, isoleucine, or leucine at position 200 (wild-type=arginine) do not induce the SOS response and thus confer a null phenotype by this sensitive measure. In contrast, when these mutations were recombined with the star mutations, the cleavage defect of the R200X

Table 7-3. *Eco*RI Star Mutations Suppress Substrate Binding Site Mutations

Allele (R200X) X=		+R56Q	+R56Q +A138V	+R56Q +H144Y	+R56Q +A138T	Affect of Star Mutations
R	dead	dead	dead	dead	dead	increase star activity
K	dead	DB,sick	DB,sick	DB,sick	DB,sick	no affect
C	DB,sick	LB	DB,sick	dead	dead	increase wild-type activity
V	MB	MB	dead	dead	dead	increase wild-type activity
S	LB	W-LB	MB	MB	DB,sick	increase wild-type activity
T	W	W	LB	LB-MB	MB	render mutant active
M	W	W	MB	W-LB	MB	render mutant active
A	W	W	LB	W-LB	LB	render mutant active
I	W	W	LB	W-LB	LB	render mutant active
L	W	W	W	W	W	no affect
G	W	W	W	W	W	no affect
N	W	W	W	W	W	no affect
Q	W	W	W	W	W	no affect

Indicator media contained 35 μg/ml X-gal.
The host was strain JH137. Growth was at 30°C.
W=white colonies
LB=light blue colonies; MB=medium blue colonies;
DB=dark blue colonies.

mutants was suppressed and endonuclease was readily detectable by its ability to induce the SOS response. Since these mutants can be activated by the addition of additional mutations, we suggest that their original site-directed mutations confer a null phenotype because they render the mutant protein defective rather than unstable. We conclude that the enhanced star activity mutations act as suppressors of substrate pocket mutations. That mutations which enhance star action of the WT enzyme also activate scission of the WT substrate by mutants defective in the binding site further supports the notion that these mutations act by relaxing the enzyme's allosteric effector requirement. On the one hand they allow the wild-type enzyme to cleave at *Eco*RI* sites and on the other, allow a crippled enzyme to better cleave the wild-type substrate.

Discussion

We have described a genetic screen to isolate endonuclease mutants which cleave at DNA sites not protected by their respective methylase. This screen is based on the finding that DNA scission induces the SOS response and employs strains which bear the lactose operon fused to an SOS inducible promoter as an *in vivo* sensor of endonuclease action. To obtain mutants detrimental to cell growth requires a conditional expression system and here we chose to use a thermosensitive allele. With this SOS induction screen, *Eco*RI endonuclease mutants were isolated which damage the DNA of the host cell in spite of the presence of the *Eco*RI methylase. One additional allele was found as a spontaneous suppressor of a substrate binding pocket mutation.

These mutant proteins have been purified. *In vitro*, they preferentially cleave the wild-type recognition site with specific activity comparable or in some cases slightly reduced (10-fold) compared to the wild-type enzyme. When incubated in normal *Eco*RI* buffer conditions, these mutant proteins cleave additional sites which represent a preferred subset of *Eco*RI* sites. The cleavage activity of these mutants is enhanced by either recombining the individual mutations together or by incubating

the purified mutant enzymes in *Eco*RI* buffer conditions. In contrast, when incubated in its preferred buffer conditions the wild-type enzyme cleaves only the wild-type substrate. The wild-type enzyme does cleave *Eco*RI* sites under *Eco*RI* conditions, but with a slower rate and a different substrate preference compared to some of the star mutant enzymes.

The ability of these mutants to circumvent the protective action of the *Eco*RI methylase *in vivo* is thus attributable to increased *Eco*RI* activity. The fact that protein mutants can exhibit *Eco*RI* activity argues that *Eco*RI* buffer conditions must exert at least part of its effect on the structure or the action of the *Eco*RI enzyme. The effect of star mutations differs dramatically from star buffer conditions in several ways. When compared to *Eco*RI* activity with the wild-type enzyme, the mutant enzymes cleave *Eco*RI* sites faster and in some cases show a preference for a subset of *Eco*RI* sites with certain allowed substitutions and special flanking sequences. The fact that individual star mutations and star buffer conditions are synergistic argues that either each does something different that contributes to activity or that their effects are additive. We suggest that both the star mutations and star buffer conditions change the conformation of the protein to favor DNA scission. In this respect star buffer conditions can be thought of as denaturing the enzyme and affording it with a somewhat heterogeneous set of altered conformations. In contrast, those star mutations which favor scission at certain *Eco*RI* sites may allow only more restricted conformational changes.

By DNA sequencing we find that for 4 of the 5 mutants the responsible amino acid substitutions lie at the DNA-protein interface. Two of these introduce larger amino acids (valine or threonine) for A138 which lies in a loop of protein very near the DNA and immediately preceding two substrate recognition amino acids, glu144 and arg145. Because these are the two mutant enzymes which prefer certain flanking sequences, amino acid substitutions near the center of the substrate somehow effect recognition of the edges of the site. Perhaps flanking nucleotides modulate the ability

of the DNA to bend and kink in the fashion observed in the crystal structure. Two other star mutations, E192K and Y193H, are adjacent to each other and lie near the DNA backbone and the other substrate binding amino acid, arg200. Both place an additional positive charge near the backbone of the DNA. Because these four mutations all lie at the DNA interface, they may identify residues of the wild-type protein which bind to the substrate. Alternatively and less romantically, these residues may play no role in substrate recognition by the wild-type enzyme if the amino acid present in the mutants simply disrupts the DNA-protein interface.

The fifth star mutation, H114Y, lies distant from either the DNA recognition or cleavage site, and it is a consideration of this mutation which sheds the most light on the activity of these mutants. This mutation lies three amino acids away from another interesting set of mutations, E111X, isolated by King *et al.* (1986, 1988). The E111G, E111A, and E111V mutant proteins bind the wild-type recognition site with the same affinity as the wild-type enzyme but are markedly defective in DNA scission. Interestingly, the E111Q mutant protein binds 1000-fold more tightly than the WT enzyme but is still defective in scission. It has been proposed that these mutations disrupt an allosteric activation mechanism that couples *Eco*RI binding with cleavage (King, *et al.*, 1986, 1988; McClarin, *et al.*, 1986).

Allostery describes the phenomenon whereby the activity of an enzyme is modulated by the binding of a small molecule. Two allosteric models have been proposed: the rigid locking model in which the allosteric effector binds a site distant from the substrate binding site and stabilizes an alternative conformation of the enzyme (Monod, *et al.*, 1963), and the induced fit model in which binding by the allosteric effector induces a conformational change in the active site (Koshland & Neet, 1968). This may be mainly a semantic distinction; at the least, both models invoke a conformational change that alters the structure of the substrate binding site. They differ only in whether the allosteric effector plays a passive or an active role. For an enzyme such as *Eco*RI in which its substrate is also the proposed allosteric effector, this

distinction becomes even more blurred since the binding sites overlap.

The idea that *Eco*RI is allosterically activated is based on the phenotypes of several mutant enzymes. Mutants in which R187 is replaced by serine show decreased DNA scission and make only two phosphate contacts at pH=6 compared to eight for the wild-type enzyme (Jen-Jacobsen, *et al.*, 1983). Upon shift to pH=6, the R187S mutant protein now makes four more phosphate contacts, for a total of six. From the crystal structure we now know that R187 does not contact the DNA directly. Apparently some conformational change occurs which brings other amino acids into ionic contact with the DNA. Although the wild-type enzyme makes eight phosphate contacts at both pHs, Jen-Jacobsen *et al.* (1983) suggested that the R187S mutation disrupts a conformational change that occurs upon DNA binding. Proteolytic deletion derivatives of *Eco*RI that remove 12 to 29 N-terminal amino acids do not cleave DNA and form a much less stable protein-DNA complex (Jen-Jacobsen, *et al.*, 1986). This portion of the protein embraces the DNA in the crystal structure. After DNA binding, the flexible N-terminal arm wraps around the DNA to stabilize the interaction. The arm lies opposite the scissile phosphodiester bond and may anchor the DNA in the catalytic cleft of the other subunit to facilitate cleavage. However, by far the strongest evidence for allosteric activation is based on the isolation of mutants (E111X) which bind substrate as tightly as (E111G, A, V) or more tightly than (E111Q) the wild-type enzyme but are virtually unable to cleave DNA (King, *et al.*, 1986, 1988). The afflicted amino acid lies far from the DNA binding or cleavage site and it has been proposed that these mutations block allosteric activation by prohibiting the required conformational change. This amino acid forms a salt bridge with the other subunit and may therefore couple the recognition machinery of one subunit (the inner and outer α-helices) with the catalytic machinery (β-sheets 2 and 3) of the other subunit (P. Modrich, personal communication).

In contrast, the *Eco*RI mutants described here display promiscuous substrate specificity. One of these (H114Y) lies three amino acids away from the E111X mu-

tants which bind but do not cleave DNA. The H144Y mutation may behave in the opposite fashion to the E111X mutations and render the protein more easily allosterically activated such that other DNA sequences now act as allosteric effectors and trigger cleavage. Because the other promiscuously cleaving mutant enzymes also show increased *Eco*RI* activity, they may similarly disrupt or mimic the allosteric conformational change.

We find that mutations in the *Eco*RI substrate binding site are suppressed by star buffer conditions and in some cases by the enhanced star activity mutations. Thus their defect is mitigated by conditions that may favor the allosterically activated conformation. We suggest that the hydrogen bonds normally made by the wild-type amino acids at these positions could initiate the allosteric conformational change. In this way these amino acids would serve a dual function to bind substrate and trigger DNA cleavage.

We find that mutations which enhance *Eco*RI* activity appear to do so by affecting an allosteric activation mechanism. Many restriction enzymes and at least the *Eco*RI methylase exhibit star activity under altered buffer conditions (Gardner, *et al.*, 1982; George & Chirikjan, 1982; Nasri, *et al.*, 1985; Nasri & Thomas, 1986; Barany, 1987; Berkner & Folk, 1978; Woodsbury, *et al.*, 1980a, 1980b). We therefore suggest that restriction endonucleases and methylases may in general be allosterically activated upon substrate binding. Restriction enzymes may achieve their high fidelity of substrate recogniton through a two step mechanism, functionally analogous to the seive model for aminoacylation of tRNAs. In the first step the enzyme makes an extensive series of interactions which bind and discern the substrate (chapter 6). In the second, the substrate induces an allosteric conformational change which acts as a final check on the fidelity of DNA cleavage. By this means cleavage of non-canonical sites would be reduced to spare the host chromosome unnecesssary and perhaps lethal DNA damage.

Chapter 8

Site-Directed *Eco*RI Endonuclease Mutants Inflict Unusual DNA Lesions: The Repair of Nicked DNA Requires SOS Induction and Recombination in E. coli

Introduction

A DNA single-strand break is expected to be a much less deleterious DNA lesion than a double-strand break since at nicks the intact DNA strand holds the severed strand together. In addition, one would expect ligation to rapidly repair DNA nicks. Although DNA ligase is an essential enzyme, its activity is thought to be in gross excess because mutations which reduce ligase activity to 1% of wild-type levels do not alter cell viability (Gottesman, *et al.*, 1973; Konrad, *et al.*, 1973).

Certain *E. coli* mutations, including *lig* (Gottesman, *et al.*, 1973; Konrad, *et al.*, 1973), *dam* (Marinus & Morris, 1974), and *dut* (Konrad, 1977; Tye, *et al.*, 1977), result in an increased level of nicks or gaps in the cellular DNA. In *lig* mutants nicks persist longer during replication (Okazaki fragments), and are sealed more slowly during DNA repair and recombination. In *dam* and *dut* mutants strains, more DNA nicks and gaps occur because of increased mis-match repair or glycosylase (ung) and AP endonuclease action. In contrast to our expectation that these DNA lesions would be repaired by ligation or DNA polymerization and ligation with little consequence, these mutant strains are induced for the SOS response and filament (Peterson, *et al.*, 1988; Craig, *et al.*, 1984; Warner, *et al.*, 1981), require *recA* function for viability (Gottesman, *et al.*; Marinus & Morris, 1974; not tested for *dut*), and are hyper-recombinogenic for some pathways (Konrad, 1977; Marinus & Konrad, 1976; Tye, *et al.*, 1977) and hypo-recombinogenic for others (Zeig, *et al.*, 1978; Ennis, *et al.*, 1987). The lethal nature of these mutations may be attributable to rare DNA breaks which occur when two nicks or gaps are closely juxtaposed (Glickman & Radman, 1980; Radman & Wagner, 1984) or an excision repair tract crosses another repair tract or a

nick in the opposite strand (Wang & Smith, 1986) (see discussion). These lesions are not simple staggered DNA scissions because we have previously shown that *Eco*RI breaks are not dependent on SOS or recombination for repair (Heitman, *et al.*, 1989; chapter 5).

In Chapter 6 we constructed a set of site-directed *Eco*RI mutants altered in the substrate binding pocket. These site-directed mutants fortuitously displayed a TS phenotype. Although several of these mutants are lethal to the cell, they are unable to restrict phage growth, show little or no endonuclease activity in cellular extracts, and the purified proteins are greatly reduced in specific activity and produce a greater proportion of the nicked product compared to the wild-type enzyme. We therefore tested these mutants to see if the DNA lesion they inflict is repaired by the same pathway as a known staggered *Eco*RI break. We describe here the unusual finding that these lesions require recombination and SOS induction for repair. These mutant enzymes do not block DNA ligase by binding to their nicked product. We argue instead that some proportion of DNA nicks are not repaired by simple ligation and require SOS induced functions for repair. This requirement could arise if DNA nicks initiate DNA recombination or some other repair pathway before DNA ligase can act. In the absence of SOS and recombination functions, this recombinational or repair intermediate could be a lethal lesion.

Results

Site-directed mutants produce DNA lesions that require recombination and SOS repair

Site-directed *Eco*RI mutants altered at amino acids 144, 145, and 200 were described in Chapter 6. Several of these mutants exhibit TS endonuclease activity as assayed by SOS induction and cell lethality. *In vitro* the purified R200K and R200C mutant proteins cleave at wild-type recognition sites with reduced specific activity, and under conditions of partial digestion, produce a greater proportion of the nicked product compared to the wild-type enzyme. We suspected that their *in vivo* activity

might result from DNA nicking rather than scission. They do clearly induce the SOS response, and thus by this measure do not differ from enzymes that generate breaks. We therefore tested whether host strains defective in DNA repair are more sensitive to the DNA lesions these mutant endonucleases inflict.

As described in Chapter 5, an isogenic strain series (JH39 (wt), JH59 (*recA*), JH117 (*recB*), JH154 (*recN*), and JH158 (*lexA3*)) that carries an SOS::*lacZ* fusion was employed to monitor SOS induction. As for other *Eco*RI TS mutants, SOS induction by these site-directed mutants requires both *recA* and *recB* function and is absent in the *lexA3* strain. The sensitivity of these strains to DNA damage by the site-directed mutants was also measured. It was necessary to use this strain series since several of these mutants are lethal in the K91 strain background. We considered that the Mu Gam protein present in these strains might affect the results, since Gam is known to bind to and protect DNA ends under some conditions (Akyrod & Symonds, 1986; Thaler, *et al.*, 1987c). However as shown in figure 8-1, by colony forming efficiency the Mu d lysogen JH137 is no more resistant to DNA scission by the *Eco*RI TS6 allele than its isogenic parent K560.

The sensitivity of this strain series to scission by the R200K and E144C mutants is shown in figures 8-2 and 8-3. In contrast to the random *Eco*RI TS mutants (Chapter 5), these site-directed mutants are much more lethal in *recA*, *recB*, and *lexA3* mutant strains. These strains are defective in either SOS induction alone (*lexA3*) or SOS induction and recombination (*recA, recB*). Although the RecN protein is required for the repair of γ-ray breaks (Picksley, *et al.*, 1984), the *recN* mutant strain was much more resistant to this DNA damage than the other mutant strains. However, the *recN* mutant was demonstrably more sensitive than its isogenic parent. We conclude that these *Eco*RI mutants inflict DNA lesions that are not simply staggered breaks since they require SOS induction and perhaps also recombination for repair.

Figure 8-1. A Mu d Lysogen does not Confer Resistance to DNA Scission. Sensitivities of strains JH137 (Mu d^+) and K560 (Mu d^-) to DNA scission by the TS6 *Eco*RI endonuclease mutant were determined as described in Figure 5-3, chapter 5. Symbols: JH137/TS6 (●); K560/TS6 (■).

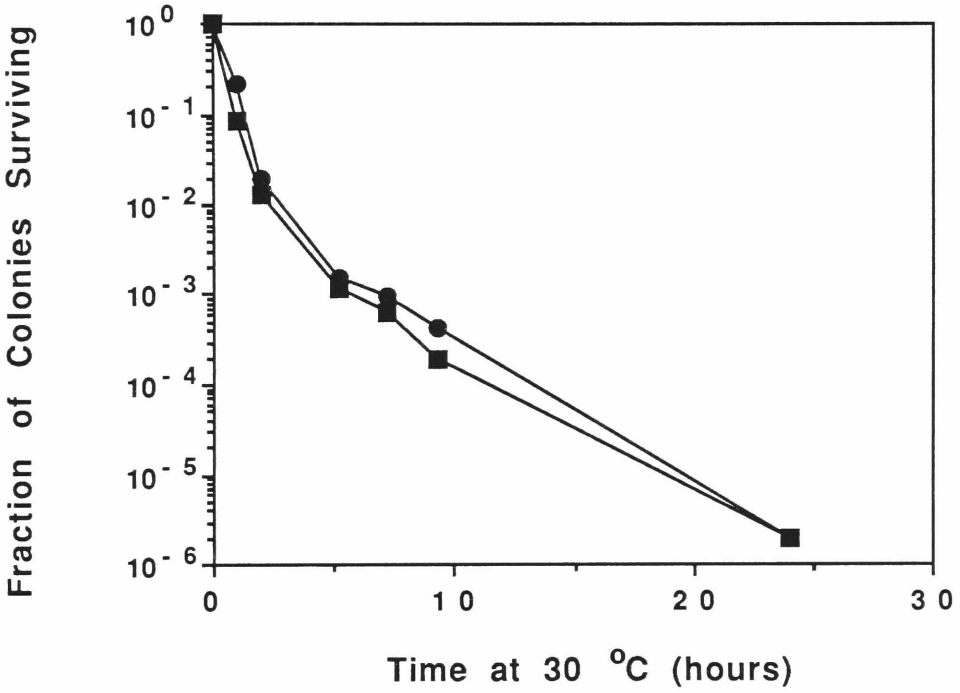

Fraction of Colonies Surviving
10 0
10 - 1
10 - 2
10 - 3
10 - 4
10 - 5
10 - 6
0
10
20
30
Time at 30 °C (hours)

Figure 8-2. Sensitivity of Mutant Strains to DNA Scission by the R200K Mutant Endonuclease. Sensitivity to DNA scission by the R200K mutant was as described in Figure 5-3. Symbols: JH39/SD39, wild-type (●); JH59/SD39, *recA56* (■); JH117/SD39, *recB21* (▲); JH154/SD39, *lexA3* (○); JH158/SD39, *recN* (□); SD39= the R200K site-directed *Eco*RI allele expressed from plasmid pJH15a.

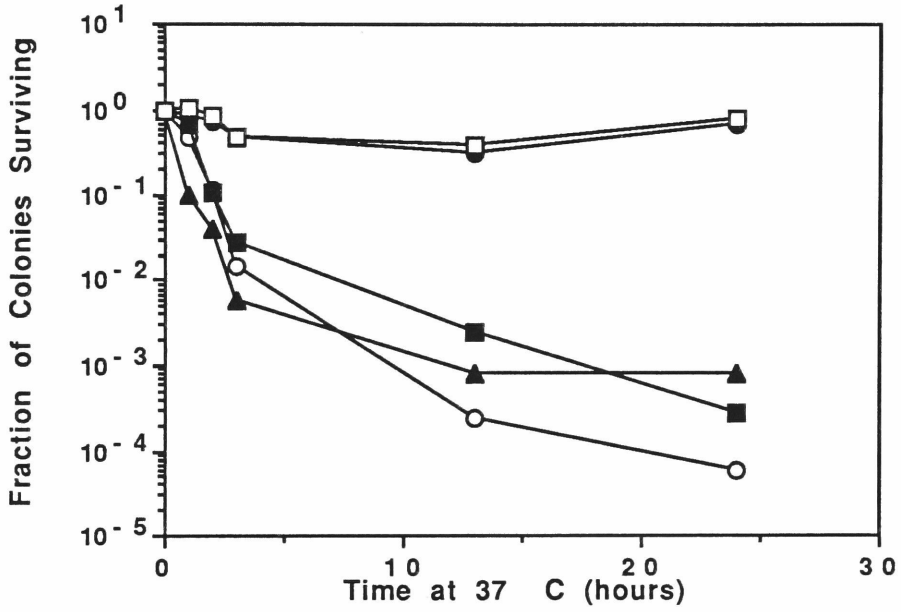
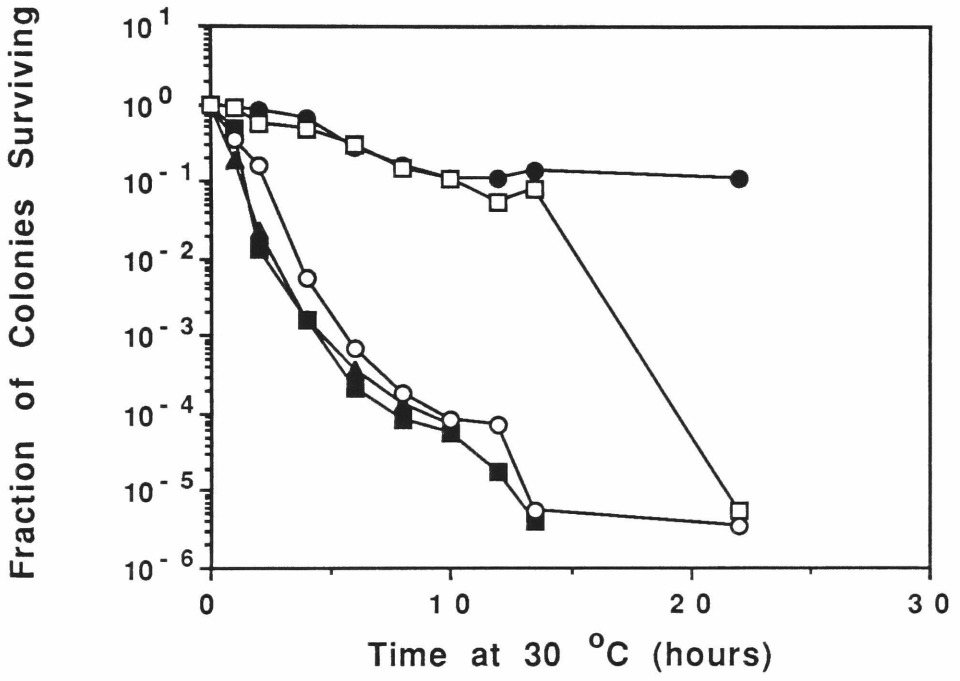

Fraction of Colonies Surviving
10 1
10 0
10 - 1
10 - 2
10 - 3
10 - 4
10 - 5
0
10
20
30
Time at 37 C (hours)

Figure 8-3. Sensitivity of Mutant Strains to DNA Scission by the E144C Mutant Endonuclease. Sensitivity to DNA scission by the E144C mutant was as described in Figure 5-3. Symbols: JH39/SD175, wild-type (●); JH59/SD175, *recA56* (■); JH117/SD175, *recB21* (▲); JH154/SD175, *lexA3* (○); JH158/SD175, *recN* (□); SD175= the E144C site-directed *Eco*RI allele expressed from plasmid pJH15a.


Fraction of Colonies Surviving
10 1
10 0
10 - 1
10 - 2
10 - 3
10 - 4
10 - 5
10 - 6
0
10
20
30
Time at 30 °C (hours)

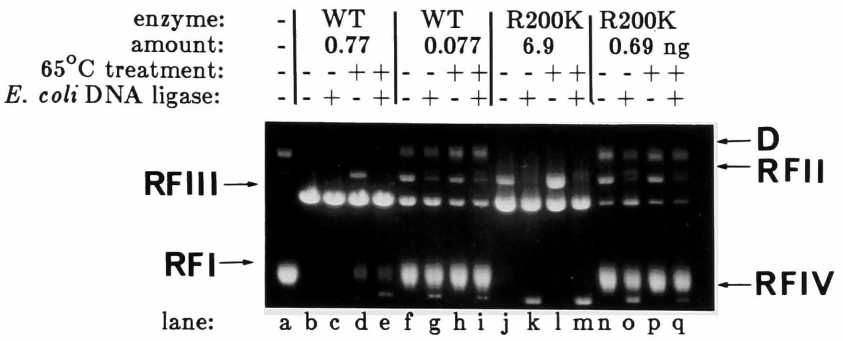
Site-directed mutants do not block DNA ligase

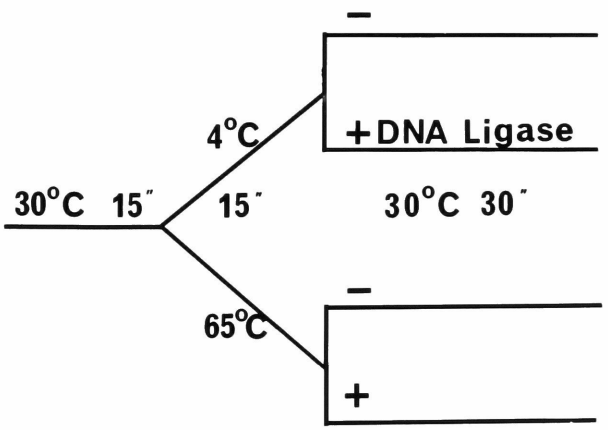
Our *in vitro* findings with the purified mutant proteins suggested that they nick the DNA. However, it is not clear why a nick should require SOS induction and recombination for repair when an *Eco*RI staggered break does not. One possibility is that the protein remains bound to the nicked or cleaved product. The protein might thereby mask the break from DNA ligase, necessitating repair by other means. This model was tested *in vitro* with *E. coli* DNA ligase and the purified R200K mutant protein. As shown in figure 8-4, limiting amounts of either wild-type *Eco*RI or the R200K mutant were incubated with 1 μg of pUC18 plasmid DNA. After 15 minutes incubation to generate a partial digestion, the reaction was divided and one-half was placed at 4°C while the other half was incubated at 65°C to heat denature and inactivate the endonucleases. Reactions were again divided and then supplemented with 26 μM DPN and 5 units of *E. coli* DNA ligase. After 30 minutes incubation at 30°C, the reaction products were displayed by electrophoresis on a 0.6% agarose gel containing 0.5 μg/ml EtBr. RFIV form DNA (covalent, relaxed circular DNA) appears after ligation as a DNA species that migrates slightly faster than RFI form DNA. However, there is no difference in the production of RFIV plasmid DNA in reactions where the mutant protein is present compared to reactions in which the enzyme was heat denatured (lanes k and m; lanes o and q). Thus by this assay, the R200K mutant does not markedly inhibit *E. coli* DNA ligase *in vitro*.

Discussion

In chapter 4 a set of TS *Eco*RI endonuclease mutants was described, several of which clearly produce DNA double-strand breaks *in vivo* and *in vitro*. These mutants enabled us to show that *Eco*RI scissions require ligation but not SOS function or recombination for repair (Heitman, *et al.*, 1989). In chapter 6 we isolated a set of site-directed mutants which fortuitously displayed a TS phenotype. In contrast to the random TS mutants, these site-directed TS alleles R200K and E144C (which do

Figure 8-4. The R200K Mutant Endonuclease does not Inhibit DNA Ligase. As described in more detail in results and shown in the figure, plasmid pUC18 DNA (1 μg) was incubated with the indicated amounts of the wild-type or R200K mutant enzyme for 15 minutes at 30°C. The reaction was divided and one-half was incubated at 65°C for 15 minutes to denature the endonucleases while the other half was kept at 4°C. These reactions were supplemented with DPN, divided again, and to one-half was added 5 units of *E. coli* DNA ligase (New England Biolabs). After a final incubation for 30 minutes at 30°C, the reaction products were displayed on a 0.6% agarose gel containing 0.5 μg/ml EtBr. The appearance of RFIV form DNA (covalently closed, relaxed DNA) is indicative of DNA ligase action.


enzyme: - WT WT R200K R200K
amount: - 0.77 0.077 6.9 0.69 ng
65°C treatment: - - + + - - + + - - + + - - + +
E. coli DNA ligase: - - + - + - + - + - + - + - + -
RFIII
RFI
D
RFII
RFIV
lane: a b c d e f g h i j k l m n o p q


30°C 15″
4°C
15′
65°C
−
+DNA Ligase
30°C 30′
−
+

not restrict λ and exhibit greatly reduced activity *in vitro*) are more lethal in strains defective in SOS induction or recombination. This suggests that the *in vivo* lesion may not be the same for the random and site-directed TS mutants. The unusual properties of these site-directed mutants may arise because their defects lie in the substrate binding pocket.

Paradoxically, the purified R200K mutant protein predominantly nicks DNA *in vitro*. Although these mutant proteins produce nicked and cleaved DNA *in vitro*, we attribute their more lethal phenotype in repair defective strains to nicking activity, because other *Eco*RI mutants that cleave the DNA require only ligation for repair. However, one would not expect a nick to be so lethal or to require recombination for repair. One possibility was that the site-directed mutants might slowly release either the cleaved or the nicked product. Such a lesion would be expected to require recombinational repair, since it is known that RecA function is needed to repair scissions by λ terminase, which remains bound to one end of the DNA after scission (Murialdo, 1988; Feiss, *et al.*, 1983). However, the purifed R200K mutant protein does not inhibit the action of *E. coli* DNA ligase at a nicked *Eco*RI site. Furthermore, Greene *et al.* (personal communication) find by single-turnover kinetic analysis that these site-directed mutant proteins release both the intact and the nicked substrate more rapidly than the wild-type enzyme. Unless the enzyme behavior is markedly different *in vivo*, we conclude that these mutants inhibit growth of repair defective strains by nicking DNA.

DNA nicks or short gaps occur in several other situations. After γ-ray irradiation, DNA contains both single and double strand breaks in a ratio of roughly 20 to 1 (Lydersen & Pettijohn, 1977). Although *recA*$^-$ strains are defective in the repair of γ-ray breaks, the sensitivity of *recA*$^-$ mutant strains to low doses of γ-rays might be attributable to a defect in the repair of nicks. Tomizawa and Ogawa (1967) studied the repair of λ phage that suffer DNA scission after the decay of incorporated radioactive phosphorus atoms. They found that 80% of nicks were rapidly repaired

(Ogawa & Tomizawa, 1967) and that roughly 40% (in some cases 80%) of lethal lesions were attributable to DNA double strand breaks (Tomizawa & Ogawa, 1967) These findings suggest that 20% of nicks are slowly repaired and may be lethal lesions. The repair of nicks in the chromosome has not been analyzed and could differ from repair of λ phage DNA.

A number of mutations are also known which increase the endogenous level of DNA nicks or strand gaps. These include: DNA ligase mutations which decrease the rate at which nicks are sealed (Gottesman, *et al.*, 1973; Konrad, *et al.*, 1973); dam methylase mutations which disable strand discrimination by mis-match repair enzymes and may increase their activity since an increased level of DNA single strand breaks occur (Marinus & Morris, 1974; Glickman & Radman, 1980; Radman & Wagner, 1984); and dUTPase enzyme mutations which result in increased cellular levels of dUTP, leading to more frequent incorporation of uracil into the DNA that is subsequently removed by Uracil-N-glycosylase, AP endonuclease, and exonuclease to to leave single-strand DNA gaps (Konrad, 1977; Tye, *et al.*, 1977). These mutations partially induce the SOS response (Craig, *et al.*, 1984; Peterson, *et al.*, 1988; Warner, *et al.*, 1981), are hyper- and hypo-recombinogenic (Marinus & Konrad, 1976; Konrad, 1977), and require *recA* function for viability (Marinus & Morris, 1974; Gottesman, *et al.*, 1973). *Dam* mutations are also inviable in combination with many other repair and recombination mutations including: *lexA3* (blocks SOS induction), *recB* and *recC* (SOS induction and recombination), *polA12* (TS for growth, DNA repair polymerase), and *recJ* and *ruv* mutations (two SOS induced members of the recF recombination pathway) (Marinus & Morris, 1974, 1975; Peterson, *et al.*, 1985). Thus some aspect of cellular physiology in these mutant strains induces the SOS response and requires repair and recombination functions for cell viability.

Although the *dam* mutation is pleiotropic and affects many cell functions (DNA repair, replication, gene expression, transposition) (Marinus, 1987), it shares some phenotypes with the *lig* and *dut* mutations perhaps because the three mutations give

rise to similar DNA lesions. It has been suggested previously that in *dam* mutants these lesions are rare DNA double-strand breaks that occur at juxtaposed nicks or colliding excision repair tracts (Glickman & Radman, 1980; Radman & Wagner, 1984). This situation arises in *dam* mutants because mis-match repair is no longer directed to excise only the newly replicated strand. However, it is not known if two excision repair tracts can collide or what transpires if they do. For the *lig* and *dut* mutants the rare double-strand break model is even less tenable. Although the Ung glycosylase can act as a restriction system and cleave uracil containing phage DNA (present in *dut* mutants), this requires more than one round of DNA replication to introduce uracil into both strands. In the *E. coli* chromosome of a *dut* mutant, uracil enters the DNA by two routes: misincorporation during DNA replication and rare spontaneous deamination of cytosine to uracil. The majority arises from misincorporation and thus lies predominantly in one strand of the DNA. Hence nicks and gaps should be rarely overlapping. This strand bias for lesions should be even more pronounced in *lig* mutants because in this case nicks should occur almost exclusively in the lagging strand at Okazaki fragments and thus rarely lie opposite another nick. Based on the phenotypes of *dam*, *dut*, *lig*, and the site-directed *Eco*RI mutants, we suggest that in some cases DNA nicks and gaps are unusual DNA lesions.

Why would DNA nicks require recombination and SOS induction for repair when we know DNA ligase can effectively seal nicks by itself *in vitro*. *In vivo* DNA nicks are substrates not only for DNA ligase but also for other enzymes which may effectively compete with DNA ligase. For example, DNA nicks can act as primers for DNA replication (Kelly, *et al.*, 1970; Masamune & Richardson, 1971), suffer exonuclease action to form gaps (Deutscher & Kornberg, 1969), and stimulate recombination (Meselson & Radding, 1975; Radding, *et al.*, 1982). If any of these competing activities reaches a DNA nick before DNA ligase, some replication, recombination, or repair process may ensue. In the absence of SOS and recombination activities, the recombination or repair intermediates might be lethal DNA lesions. The inciting

event may be conversion of DNA nicks to gaps which are no longer substrates for ligation and require recombination or DNA synthesis for repair. In *dam* and *dut* mutant strains, DNA nicks are necessarily converted to DNA gaps as part of the repair process. DNA nicks may also be enlarged to gaps in the *lig* and site-directed *Eco*RI mutants. Because DNA mis-match repair can initiate at random DNA nicks (Modrich, 1989) these *Eco*RI site-directed mutants may confuse the mis-match repair system and increase mutagenesis. The fact that DNA breaks behave differently than nicks suggests that breaks do not initiate this recombination or repair process and remain free as substrates for DNA ligase.

Since cell viability does not require RecA and SOS functions, DNA ligase must be sufficiently abundant to compete with other enzymes and rapidly seal nicks which occur naturally in the cell (due to normal mis-match repair and Okazaki fragments for example). The fact that cells survive with only 1% the normal level of DNA ligase has been promoted as an argument that DNA ligase is extremely more abundant than required for cell growth (Gottesman, *et al.*, 1973). However, these DNA ligase mutants do exhibit hyper-recombination and require RecA for viability. Thus although the cell requires only 1% of its DNA ligase for viability, the role of the other 99% may be to prevent hyper-recombination and DNA damage. We suggest that by this reasoning the cell does not make an over-abundance of ligase, and in fact may make just enough DNA ligase to effectively inhibit other DNA-metabolizing enzymes from acting at DNA nicks.

Chapter 9

The EcoRI Endonuclease Nicks DNA at EcoRI* Sites

Introduction

The *Eco*RI endonuclease preferentially cleaves its normal substrate (GAATTC) with exceedingly high fidelity (10^5-fold) compared to other DNA sequences. However, by altering its preferred buffer conditions the enzyme can be tricked into cleaving additional sites that usually differ from the normal site by one nucleotide substitution (Polisky, *et al.*, 1975; Rosenberg & Greene, 1982). This is called *Eco*RI* activity. Additionally, it has been reported that under normal buffer conditions the enzyme nicks (but does not cleave) the sequence GAATTA (Bishop, 1979). However, this finding has been disputed. If the *Eco*RI endonuclease can in fact nick *Eco*RI* sites, the enzyme might act at these sites in the chromosome of the host cell since they are not protected by the *Eco*RI methylase. Since DNA nicks stimulate recombination (Meselson & Radding, 1975; Konrad, 1977; Strathern, personal communication), the *Eco*RI restriction could enhance recombination and promote genetic exchange through nicking activity.

We show here that the *Eco*RI endonuclease clearly nicks DNA molecules (f1 RFI DNA) that lack *Eco*RI sites under normal buffer conditions. Nicking occurs at several *Eco*RI* sites. *In vivo* we observe that the *Eco*RI restriction-modification system induces the SOS response extremely weakly in comparison to SOS induction by mutations which cause DNA nicking (i.e. *dam*$^-$, *lig4*) (Craig, *et al.*, 1984; Peterson, *et al.*, 1988)). Furthermore, the *Eco*RI system stimulates recombination only slightly as monitored by the frequency of *lac*$^+$ progeny occuring in a *lac* deletion duplication strain (Konrad, 1977). We conclude that any effects of nicking by the *Eco*RI enzyme are minor.

Results

EcoRI endonuclease nicks a DNA molecule without EcoRI sites

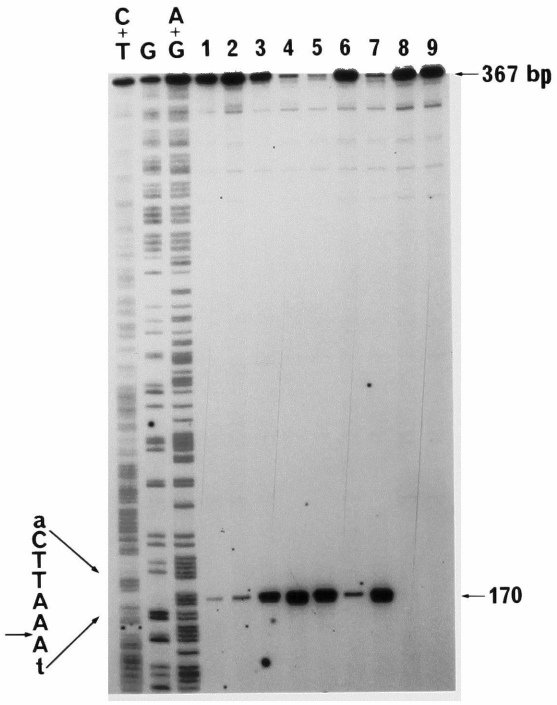
The filamentous phage f1 bears no *Eco*RI recognition sites. However, when incubated with *Eco*RI endonuclease in its normal preferred buffer conditions, the supercoiled phage DNA is relaxed to RFII form (see chapter 7, figure 7-7). Under these conditions, no detectable cleavage to linear (RFIII) molecules occurs. This nicking action was observed with several independent enzyme preparations, including different lots from the same supplier (NEB), from different suppliers (NEB vs. Pharmacia) or purified in our laboratory. Nicking required significantly more enzyme units than does cleavage of the normal substrate (~ 100-1000 units compared with 1 unit/μg DNA).

EcoRI nicks DNA at EcoRI sites*

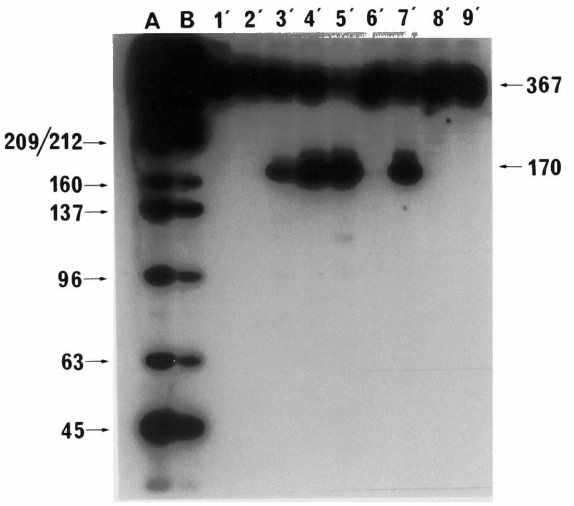
In chapter 7 we described a set of *Eco*RI endonuclease mutants which show enhanced cleavage activity at *Eco*RI* sites. To definitively map their sites of cleavage, short uniquely end-labelled f1 restriction fragments were subjected to both Maxam and Gilbert sequencing reactions and to scission by the mutant enzymes. Furthermore, we incubated each fragment with the wild-type enzyme in both its preferred buffer and *Eco*RI* buffer. The reaction products were displayed on 6% polyacrylamide/8 M urea denaturing sequencing gels to map the sites of cleavage (see figure 9-1). The cleavage products were electrophoresed in parallel on 8 - 12% polyacrylamide non-denaturing gels to determine if the cleavage products observed on denaturing gels result from single- or double-strand scission. Either the wild-type protein in star buffer conditions or the mutant proteins in normal buffer clearly produced DNA fragments attributable to double-strand DNA cleavage at *Eco*RI* sites. At a slower rate, the wild-type enzyme in its preferred buffer conditions produced fragments attributable to scission at star sites. However, these fragments were observed only under denaturing conditions and are therefore single-strand breaks. As listed in table 9-1, nicking occurs at several star sites but not at others. We conclude

Figure 9-1. The *Eco*RI Endonuclease Nicks DNA at *Eco*RI* Sites. A 367 bp *Hin*fI (2011) - *Hpa*II (2378) f1 fragment (labelled at the *Hin*fI end) was subjected to Maxam and Gilbert chemical DNA sequencing and scission by the wild-type or mutant *Eco*RI endonucleases. Reaction conditions and electrophoresis were as described in detail for figure 7-6, chapter 7. Again, panel A is a 6% polyacrylamide - 8 M urea denaturing sequencing gel. Lane 1 is the wild-type enzyme in normal *Eco*RI buffer, lane 2 is the wild-type enzyme in *Eco*RI* buffer conditions, and lane 9 is the untreated fragment. The other lanes (3-8) are restriction digests with the *Eco*RI star mutants described in chapter 7. DNA scission by either the wild-type or mutant enzymes occurs at the position of the arrow within the *Eco*RI* site shown to the left of the sequencing gel.

Panel B is a non-denaturing 8% polyacrylamide gel. The cleavage product observed in lane 1 of panel A is not present in lane1' of panel B, indicating that DNA scission by the wild-type *Eco*RI enzyme incubated in normal *Eco*RI buffer produces only DNA single-strand breaks.


C + T
G
A + G
1 2 3 4 5 6 7 8 9
367 bp
170
aCTTAAAt

A


A B 1′ 2′ 3′ 4′ 5′ 6′ 7′ 8′ 9′
367
170
209/212
160
137
96
63
45

B

Table 9-1. f1 Sites Nicked by Wild-Type *Eco*RI Endonuclease

Position	Sequence	Fragment	Nicked by *Eco*RI?
1499	gAAATTCa	HinfI 1403* - HpaII 1974 (521 bp)	moderately
1674	aAAATTCa	HinfI 1403* - HpaII 1974 (521 bp)	strongly
2182	tAAATTCa	HinfI 2011* - HpaII 2378 (367 bp)	weakly
2728	aAAATTCa	HaeIII 2554 - HinfI 2845* (291 bp)	no
4060	tAAATTCa	HpaII 4018* - HaeIII 5239 (1221 bp)	no
5912	gAAATTCa	BanI 5677 - ClaI 6037* (360 bp)	no
6125	tGAATTTa	ClaI 6037* - AhaIII 6231 (194 bp)	no

that the *Eco*RI endonuclease nicks *Eco*RI* sites under its normal buffer conditions and turn to a consideration of the *in vivo* consequences of this action.

EcoRI nicking has no marked physiological affect

Normally the *Eco*RI endonuclease does not cleave *Eco*RI sites in the chromosome of its host cell because they are modified and thereby protected by the *Eco*RI methylase. However, both the endonuclease and the methylase act at non-canonical sites, *Eco*RI* sites, when subjected to certain *in vitro* conditions (buffer or enzyme excess). Does this occur *in vivo*? Expression of the wild-type *Eco*RI restriction system induces the SOS response to only a very minor extent (faint blue colonies versus white colonies). Thus by this very sensitive measure of endonuclease action (see chapter 5), *Eco*RI produces few breaks *in vivo*. This suggest the endonuclease shows little star action *in vivo*. An alternative possibility was that the *Eco*RI methylase blocks action of the endonuclease at star sites. However, *Eco*RI* sites are not protected *in vivo*. Both a plasmid expressing the *Eco*RI methylase (pJC1) and f1 RFI DNA isolated from a host strain making the methylase are still sensitive to *Eco*RI* action *in vitro* (data not shown). We conclude that the *Eco*RI endonuclease and methylase show little or no star activity *in vivo*.

Since the *Eco*RI methylase does not protect star sites *in vivo*, our *in vitro* findings argue that the endonuclease should nick *Eco*RI* sites *in vivo*. Several *E. coli* mutations (*dam, lig, dut*) that increase the level of DNA nicks or single-strand gaps are known to partially induce the SOS response (Craig, *et al.*, 1984; Peterson, *et al.*, 1988; Warner, *et al.*, 1981). The extremely weak SOS induction observed with the wild-type *Eco*RI restriction system is substantially less than that found with *dam* or other mutants, but may nonetheless arise from *Eco*RI nicking the chromosome. Mutations which produce DNA nicks or gaps are also hyper-recombinogenic and increase the frequency of lac^+ papillae that arise in a strain whereby recombination between two different *lac* operon deletions restores a lac^+ phenotype (Konrad, 1977). The *Eco*RI

restriction system has only a slight stimulatory effect on recombination in this system (1.5-fold increase). Lastly, the wild-type *Eco*RI genes are viable in a *recA*$^-$ host, while *E. coli* or *Eco*RI endonuclease mutations that produce nicks are lethal to *recA*$^-$ hosts. These observations suggest that the *Eco*RI endonuclease makes few or no nicks in the host DNA.

Discussion

The *Eco*RI endonuclease and methylase exhibit exceedingly high fidelity towards their preferred substrate under optimal buffer conditions. However, both enzymes act at additional substrates when incubated under different buffer conditions. This is called *Eco*RI* activity (Polisky, *et al.*, 1975; Greene & Rosenberg, 1982). We thought that if star activity occurs *in vivo* (for example by the endonuclease), the *Eco*RI restriction sytem might promote genetic exchange by recombination. However, by monitoring SOS induction and the methylation state of *Eco*RI* sites, we find no evidence for star activity *in vivo.*

It has previously been reported that the *Eco*RI endonuclease nicks at least the GAATTA *Eco*RI* site *in vitro* (Polisky, 1979). Although this has been disputed, we show here that the *Eco*RI enzyme clearly nicks DNA which lacks *Eco*RI sites. These sites of single strand scission map to *Eco*RI* sites. We conclude that *Eco*RI nicks star sites *in vitro* without requiring star buffer conditions.

E. coli mutations known to produce DNA nicks or single-strand gaps (such as *dam, lig.*and *dut*) are hyper-recombinogenic, partially induce the SOS DNA repair response, and require at least *recA* function for viability (Craig, *et al.*, 1984; Gottesman, *et al*, 1973; Warner, *et al.*, 1981; Marinus & Morris, 1974; Konrad, 1977). Furthermore, as described in chapter 8, site-directed *Eco*RI endonuclease mutants which produce a greater proportion of the nicked product compared to the wild-type enzyme induce the SOS response and severely impair the growth of mutant hosts lacking *recA* or SOS functions. In contrast, the wild-type *Eco*RI restriction system

induces SOS only marginally, is not hyper-recombinogenic, and does not require *recA* for viability. Therefore we must conclude that although the *Eco*RI enzyme clearly nicks *Eco*RI* sites *in vitro*, either this nicking does not occur *in vivo* or it occurs at a level insufficient to perturb cellular physiology. It has been shown that some *in vivo* conditions (the presence of spermidine) reduce *Eco*RI* activity *in vitro* and could therefore enhance the fidelity of *Eco*RI activity *in vivo* (Pingoud, 1985). However, we cannot exclude the possibility that a very low level of nicking activity has, over long time spans, stimulated recombination enough to alter the organization of the genome.

One could imagine that an ancestor of the present *Eco*RI enzyme was a nicking enzyme whose role was to stimulate recombination. *Eco*RI may have subsequently evolved into a restriction system by learning to cleave both DNA strands concertedly. Restriction endonucleases which exhibit substantial nicking activity or cleave DNA by sequential nicking are better candidates for recombination stimulating enzymes.

Chapter 10

Conclusions and Speculations

The repair of single- and double-strand DNA breaks

Contrary to our expectation that DNA double-strand breaks should be particularly lethal DNA lesions, we find that *E. coli* has an effective repair system. Staggered DNA breaks made by the *Eco*RI endonuclease induce the SOS DNA repair response (chapter 5) and stimulate recombination (based on the work of others, see Stahl, *et al.*, 1977; Thaler, *et al.*, 1987a, 1987b, 1987c). However, these breaks require neither SOS induction nor recombination for repair. Strains deficient in DNA ligase activity are exceedingly sensitive to *Eco*RI breaks, suggesting that wild-type levels of ligase effectively repair *Eco*RI breaks. We suggest that the higher order structure of the *E. coli* chromosome holds the severed DNA to instruct DNA ligase which ends to ligate.

In contrast to DNA breaks in the chromosome, plasmid- or phage-borne DNA double-strand breaks are repaired at low efficiency, largely by a recombination dependent mechanism which often produces deletions (Stahl, *et al.*, 1977; Conley & Saunders, 1984). Furthermore these linear DNA molecules are frequently destroyed by exonucleases before repair can occur (Simmon & Lederberg, 1972). We suggest two models to explain why the fate of DNA breaks depends on their location. First, the nucleoid structure may hold severed chromosomal ends tethered together while plasmid DNA ends should be free to wander apart. DNA ligase may distinguish the two substrates by the relative proximity of their DNA ends. Second, plasmid or phage DNA should fully relax after scission while broken chromosomal DNA may be held in its supercoiled form by the local tertiary structure. When the RecBCD complex enters the DNA at these ends, it could discriminate between the two situations, preferentially degrading relaxed DNA and unwinding supercoiled DNA. In this way breaks in the chromosome would remain substrates for ligation while those inflicted on a plasmid or phage would target the molecule for destruction. This may be the

cell's mechanism for selectively degrading unwanted linear DNA fragments. Restriction-modification systems perhaps evolved to exploit the relative sensitivity of small DNA molecules to DNA double-strand breaks. They may also rely on the inherent resistance of the chromosome to protect the host cell from endonuclease action in some circumstances.

One would expect DNA ligase to rapidly repair DNA nicks. In contrast, we find that *Eco*RI endonuclease mutants that nick DNA *in vitro* produce DNA lesions *in vivo* which are repaired by SOS induction and recombination. Since nicks can be repaired by DNA ligase *in vitro*, the situation must be more complex *in vivo*. For example, recombination or repair may initiate at DNA nicks or gaps before ligase can act. Lethal DNA lesions may ensue if SOS and recombination functions are not available to metabolize this intermediate to completion.

SOS induction as an in vivo assay of DNA-protein interactions

We have developed an *in vivo* DNA scission assay based on our observation that *Eco*RI double-strand DNA breaks induce the *E. coli* SOS DNA repair response. SOS induction is monitored with strains that bear the *lacZ* gene under control of a DNA-damage inducible promoter. DNA damage activates β-galactosidase expression which is amenable to genetic screens and quantitative assay. This SOS induction assay should prove generally useful for studying enzyme-DNA interactions. Lesions other than breaks, such as nicks or bound protein, may also induce SOS and extend the useful range of this assay. The SOS induction assay has been or is being used to study: the *Eco*RI restriction enzyme (chapter 6 and 7), the Mrr and McrB methylation dependent restriction systems (chapter 3), DNA double strand break repair (chapter 5), the *Taq*I endonuclease (F. Barany), topoisomerase I (Y-C. Tse-Dinh), *dam* methylase (M. Marinus), T-even phage *dam* methylases (S. Hattman), *Msp*I methylase hybrids (J. Meyerton and R. Roberts), the T-even phage intron endonuclease (D. Bell-Pederson and M. Belfort), and the repair of blunt-end DNA breaks (J. Heitman

and B. Blumenthal).

Functions of Restriction-Modification Systems

As their name implies, many restriction-modification systems have been demonstrated to restrict phage growth. However, for most type II restriction systems this has not been tested due to the difficulties of working with unusual bacteria which may or may not have characterized phages. Forty or so restriction systems have been cloned and expressed in *E. coli* (reviewed in Wilson, 1988). A large number of these do not restrict phage growth although in several cases the enzyme is clearly well expressed. These enzymes may in fact serve some function other than restriction. One suggestion has been that DNA scission by restriction enzymes may stimulate recombination *in vivo*, perhaps similar to the ways we use them to stimulate recombination *in vitro* (Chang & Cohen, 1977; Arber, 1979). However, the evidence supporting this putative role is meager. Stahl *et al.* (1977) showed that *Eco*RI restriction stimulates recombination of an infecting λ phage with an intact prophage. However this effect is not dramatic (5-fold at best), perhaps because phage DNA is rapidly destroyed by exonuclease degradation after suffering scission.

We tested whether expression of the *Eco*RI restriction-modification system increases recombination. Since the *Eco*RI endonuclease can cleave or nick additional DNA sites *in vitro* (*Eco*RI* activity), we thought it might do so *in vivo* to stimulate recombination. We employed the host chromosome as substrate because it appears to be repaired differently than phage after DNA scission. However, neither SOS induction nor recombination were dramatically increased by the *Eco*RI system. One other possibility was that a restriction-modification system increases recombination transiently immediately after it enters a new host cell. Once the chromosome becomes protected by methylation, the now established restriction-modification system would no longer enhance recombination. This model is analogous to phages or transposons which are zygotically induced or activated when they enter a cell and then become

quiescent. However we found no increase in recombinants after transformation by a plasmid bearing the *Eco*RI system. We conclude that the role of the *Eco*RI restriction system is restriction and that its ability to effect recombination is minor.

The *Eco*RI enzyme predominantly makes DNA double-strand breaks. These lesions may not be very recombinogenic in *E. coli* since chromosomal breaks are ligated while phage-borne breaks sponsor exonuclease destruction. Because *Eco*RI mutants that nick DNA require recombination for repair, enzymes which nick DNA might more effectively stimulate recombination than enzymes that cleave DNA. The ancestral *Eco*RI protein might have nicked and cleaved DNA to serve two functions (restriction and recombination). Alternatively it may have only nicked DNA and later learned to cleave both strands concertedly. A careful examination of the *in vitro* properties of other "restriction" enzymes (nicking versus cleavage activity *in vitro*; concerted or independent strand scission) and the effects of their expression in *E. coli* may reveal enzymes whose role is to stimulate recombination. A good place to look would be at those restriction enzymes which clearly do not restrict phage.

Substrate recognition by the EcoRI endonuclease

What did we learn about the *Eco*RI substrate interaction from our mutagenic studies? The crystal structure model correctly predicts that several amino acids which hydrogen bond to the substrate play a role in the enzyme's function, since mutating these amino acids cripples the enzyme. However these amino acids are not the sole determinant of substrate specificity because these mutant proteins still cleave the wild-type recognition site. The enzyme must make additional contacts to discern its substrate. These extensive interactions contribute to the high substrate fidelity of the enzyme.

Using our SOS induction assay, we isolated *Eco*RI mutants that damage the cellular DNA in spite of the presence of the protective methylase. *In vitro*, the purified mutant proteins preferentially cleave the wild-type site. They also cleave some

*Eco*RI* sites under normal buffer conditions where the wild-type enzyme does not. By analogy with the nearby E111X mutations that block cleavage but not binding (King, *et al.*, 1986, 1988), we argue that these mutations support a model whereby DNA cleavage by *Eco*RI is allosterically activated upon substrate binding. This acts as a second check on the substrate's identity and further increases the fidelity of substrate recognition. DNA cleavage by *Eco*RI is thus a two-step mechanism which is analogous in some ways to the sieve mechanism for charging tRNAs.

Implications for other enzymes and proteins that bind DNA

Our mutations which allosterically activate *Eco*RI show increased *Eco*RI* activity. Many other endonucleases and at least the *Eco*RI methylase also exhibit star activity. Thus these enzymes may be allosterically activated by their substrates. This may a general mechanism to increase the fidelity of enzymes that interact with DNA.

With a few rare exceptions (*Eco*RI and *Eco*R124 and *Rsr*RI; P. Greene and T. Bickle, personal communication) restriction enzymes are not homologous and therefore share no common DNA recognition mechanism. In each case evolution has evolved a completely new enzyme to recognize a particular DNA sequence rather than simply tinkering with an existing model. We can therefore expect the substrate interactions of these enzymes to be as complex as for *Eco*RI. Furthermore, endonucleases and their cognate methylases share no homology even though they both recognize the same DNA sequence. Thus enzymes that cleave DNA may have to recognize DNA in a different way compared to enzymes that methylate DNA.

Restriction endonucleases and methylases do not carry either the helix-turn-helix or zinc finger motifs sported by so many transcription factors. There are now many examples of DNA-binding proteins which do not carry either motif as well. Even amongst the proteins of the helix-turn-helix class, each protein uses its recognition helix in a very different way. There is no simple code that governs which amino acids bind which nucleotides. Each protein appears to take advantage of whatever con-

tacts are available. Thus DNA-protein interactions are exceptionally complex and in this respect resemble antigen-antibody interactions. To understand the detailed interactions for a given protein-DNA interaction, we will probably have to solve the crystal structure for each complex one by one. Even when armed with a detailed structural model, we expect that it will prove difficult to alter the specificity of DNA-binding by simple and rational amino acid changes.

Appendix 1

Phage Trojanhorses:

A Conditional Expression System for Lethal Genes

Introduction

Many genes encode proteins required for cell growth. These genes can be studied genetically using conditional alleles, such as temperature (heat or cold) sensitive or nonsense mutations, which are active under some conditions but not others. Alternatively, one can employ regulated promoters (such as the *tac* promoter of de Boer *et al.* (1983)), to modulate expression of essential genes. Genes that encode cell lethal functions can also be studied conditionally. For example, the *in vivo* consequences of DNA double-strand breaks were studied in *E. coli* with temperature-sensitive mutants of the *Eco*RI endonuclease (Heitman, *et al.*, 1989). Nonetheless, these techniques have limitations which might be overcome by other conditional expression systems.

Cloned genes are borne in *E. coli* on plasmids or phages. Several genes of the phage f1 (gene I, gene IV, gene VIII) are lethal to the cell when over-expressed from plasmid pBR322 derivatives (Boeke, *et al.*, 1982; Horabin & Webster, 1986; Brissette, *et al.*, unpublished). These same genes are not lethal when borne by the phage, presumably because filamentous phage regulate their production to avoid killing the host (Fulford, *et al.*, 1986). However, certain phage mutations perturb this delicate balance and severely impair growth of the host cell. Although the host cell dies after infection, these phage produce progeny which go on to form extremely clear plaques. These observations suggested to us that the f1 phage might be an ideal vector system for studying lethal genes. Phage expressing a lethal function might kill their host but still yield plaques, allowing one to recover alleles of the lethal gene from phage DNA. This is in fact the case for some clear plaque forming P1 mutants which were later shown to express the P1 restriction enzyme but not the methylase (Rosner, 1973).

We describe here the construction of filamentous phage that express the *Eco*RI

endonuclease. This enzyme is lethal to cells which lack the *Eco*RI methylase and phage expressing *Eco*RI kill unprotected host cells to make clear plaques. When the host cell is protected by the *Eco*RI methylase, these phage make turbid plaques. Mutations which inactivate the *Eco*RI endonuclease gene were isolated by screening for phage that make turbid instead of clear plaques. This system may prove useful for studying other cell lethal functions.

Results

Construction of phage expressing the EcoRI endonuclease

The *Eco*RI endonuclease gene was cloned into the M13mp18 vector. Phage J9+ bears the *Eco*RI gene under the control of the *lac* promoter, and phage J9- carries the gene in the opposite orientation. In the $lacI^q$ host strain K561, expression from the wild-type *lac* promoter is regulated by both the *lac* repressor and the catabolite repression system whereby growth in glucose reduces cAMP levels and decreases transcriptional activation by the CAP-cAMP complex. Thus when cells are grown in glucose *lac* expression is induced by the addition of IPTG (a gratuitous inducer of *lac*) and can be increased further with cAMP. For our purposes, IPTG induction was sufficient. As described in table A-1, phage J9+ forms turbid plaques on strain K561. When *lac* expression from phage J9+ was induced by addition of 1 mM IPTG, the phage makes clear plaques. The M13mp18 phage itself makes turbid plaques with slightly clearer centers when induced with IPTG. This effect may be attributable to increased expression of the phage gene II protein, because its gene lies downstream from the *lac* promoter in M13mp18 and phage mutants that over-produce gene II are known to make somewhat clearer plaques (Michel & Zinder, 1989). When the host strain expressed the *Eco*RI methylase from plasmid pJC1, the *Eco*RI expressing phage J9+ produced turbid plaques (+IPTG). In contrast, phage J9- bears the endonuclease gene in the orientation opposite to the *lac* promoter and forms turbid plaques with or without IPTG induction. Neither phage made clearer plaques on host strains

Table A1-1. Phage Plaque Morphologies

Phage	Host IPTG	K561/pACYC184	K561/pJC1
MP18	-	T	T
	+	T*	T*
J9-9 (P_{lac}EcoRI)	-	T	T
	+	C	T
J11 (J9-9 without RI site)	-	T	T
	+	C	T
J9-8 (EcoRIP_{lac})	-	T	T
	+	T	T

T=turbid C=clear T*=turbid plaques with slightly clear centers

defective in DNA repair (*recA*$^-$, *recN*$^-$), in keeping with our findings that *Eco*RI scissions are not repaired by the products of these genes (Heitman, *et al.*, 1989).

The polylinker of phage M13mp18 contains an *Eco*RI recognition site. Thus phage J9+ bears both an *Eco*RI site and expresses the *Eco*RI endonuclease. Such a phage might be expected to impair its own growth by restriction. Derivatives which lack the *Eco*RI site (J11) make slightly larger plaques than phage J9+, but the affect is not very pronounced. We conclude that M13 can carry the *Eco*RI endonuclease gene under control of the *lac* promoter and, after infection of cells lacking the methylase, produce enough progeny before cell death to produce a plaque. Phage that bear an *Eco*RI site may survive restriction because they exist as single-stranded DNA during part of their life cycle.

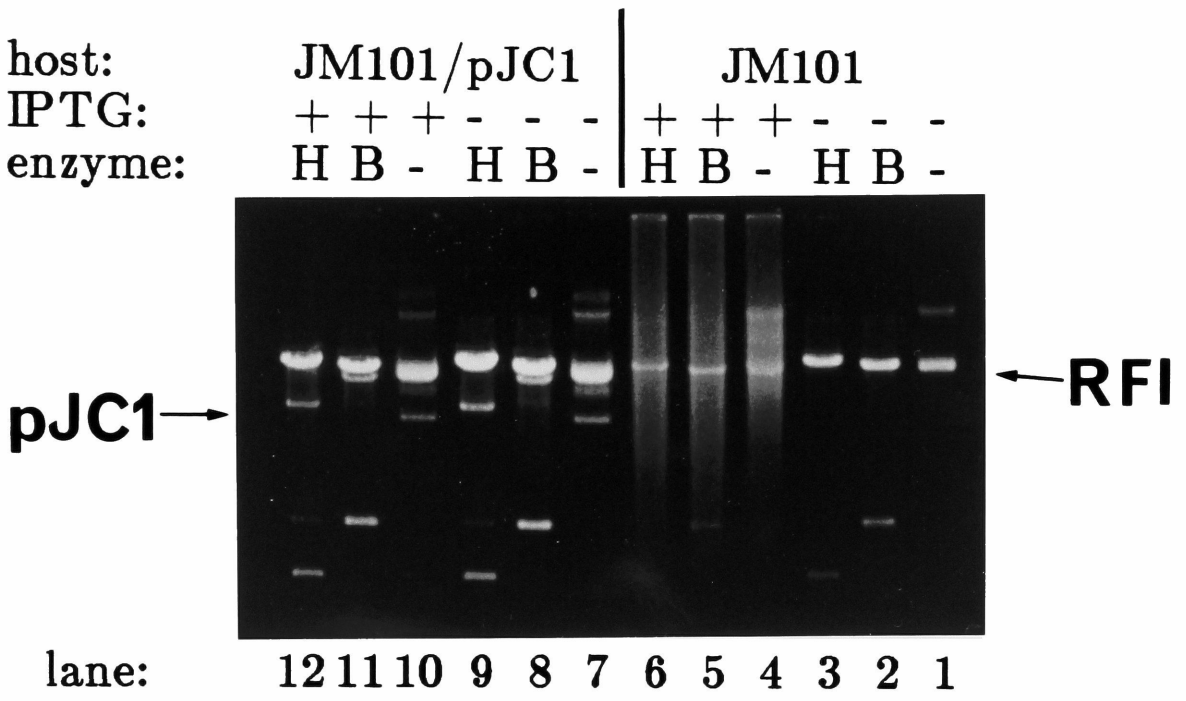
In vivo DNA scission by phage expressing EcoRI

To determine if the cell lethal phenotype of phage J9+ is attributable to DNA destruction, we followed the fate of the phage after infection. As shown in figure A1-1, when cells were infected with phage J9+, substantial phage DNA degradation occured after IPTG induction of *Eco*RI endonuclease expression (see lanes 4, 5, and 6). However, the linear DNA species was not observed. When DNA scission is produced by a temperature-sensitive *Eco*RI mutant expressed from a plasmid, we never observed the linear cleavage product under a variety of conditions (see chapter 5). Derivatives lacking the *Eco*RI site (J11) were immune to destruction although in these cases a smear of DNA, presumably chromosomal DNA, was observed (data not shown). Expression of the *Eco*RI methylase blocked *in vivo* DNA scission by either phage (figure A-1, lanes 10-12). We conclude that after infection, the *Eco*RI endonuclease is expressed, subsequently cleaves at *Eco*RI sites present *in vivo*, and the linear DNA fragments are rapidly destroyed by exonuclease action.

Site-directed EcoRI mutants

Since the *Eco*RI-expressing phage form clear plaques, one might isolate *Eco*RI

Figure A1-1. *In vivo* DNA scission after cells are infected by a phage expressing the *Eco*RI endonuclease. Strain JM101 or JM101/pJC1 (*Eco*RI M^+) was grown at 37°C to on OD_{600}=0.6 in FB media and infected with phage J9+ at an MOI=20. The culture was divided in two and 1 mM IPTG was added to one half. Tetracycline was added to 12.5μg/ml at 15 minutes after infection to block production of phage single-stranded DNA. After 1 hour, RF form DNA was prepared by the lysozyme/Triton X-100 procedure (Zinder & Boeke, 1982). Some samples were treated with *Bam*HI or *Hin*dIII endonuclease *in vitro* to show that lanes 7-12 contain the pJC1 methylase plasmid and in an attempt to test if the DNA smear in lanes 4-6 arises from scission at any specific sites.


host:
JM101/pJC1
JM101
IPTG:
+ + + - - - + + + - - -
enzyme:
H B - H B - H B - H B -
pJC1
RFI
lane:
12 11 10 9 8 7 6 5 4 3 2 1

mutants by screening for derivatives which make turbid plaques. To test this, phage J9+ was subjected to site-directed mutagenesis. As mutagen, we employed a 32-fold degenerate oligonucleotide mixture capable of generating each amino acid at position 200 (oligo JOHE 2, see table 2-3). Since this amino acid plays a crucial role in substrate binding (McClarin, *et al.*, 1986), we expected some amino acid substitutions to disrupt enzyme function. After mutagenesis 2% (12 of 600) of the plaques were turbid compared to < 0.1 % without mutagenesis. Six representative turbid plaque forming mutants were sequenced. As shown in table A-2, all six contain mutations at residue 200. The wild-type amino acid is arginine, while these mutants contain non-conservative amino acids, either tryptophan, leucine, serine, or methionine. These alleles have also been isolated in a plasmid background (see chapter 6). In this case, the tryptophan, leucine, and methionine mutants exhibit a null phenotype, whereas the serine mutant retains a low level of endonuclease action as assayed by its ability to induce the SOS response (Chapter 6). Thus by screening for turbid plaque mutants from the clear phage J9+, one can obtain *Eco*RI mutants with null or decreased activity. Alternatively, one can imagine screening for mutants on a host expressing the *Eco*RI methylase. In this case, phage J9+ forms turbid plaques. By isolating mutants which instead form clear plaques, one might find *Eco*RI mutants that cleave sites which the *Eco*RI methylase does not protect.

Discussion

Essential and cell lethal functions can both be studied by conditional expression. Methods for conditionally expressing genes include termperature (heat and cold) sensitive mutations, suppression of nonsense mutations, regulated promoters (*lac* (Guarente, *et al.*, 1980), *tac* (deBoer, *et al.*, 1983), P_L (Rosenberg, *et al.*, 1983)), and recombination mediated invertible constructs. These systems require the cell to survive at either the uninduced level (for lethal functions) or the fully induced level (for necessary functions). In some cases the uninduced level is high and may not be ap-

Table A1-2. Phenotype of *Eco*RI Site-Directed Mutants.

Allele	Phenotype	Mutation
WT	clear	R200
T6	turbid	R200L
T7	turbid	R200W
T8	turbid	R200S
T9	turbid	R200S
T10	turbid	R200M
T12	turbid	R200L

propriate for regulating the expression of very lethal functions. Here, we describe expression of cell lethal functions from filamentous phage vectors. This system does not require cell survival, since phage escape in spite of cell death and can be recovered as a plaque. Furthermore, f1 phage will continue to make plaques even when the rate of phage production has been reduced by 20-fold (Russel & Model, 1985).

Unlike lytic phage, filamentous phage continuously infect their host without killing it (Fulford,*et al.*, 1986; Model & Russel, 1988). In contrast, phage bearing the gene for the *Eco*RI endonuclease kill their host and form clear and in some cases smaller plaques. Mutations which inactivate the *Eco*RI endonuclease result in a turbid plaque phenotype. Similarly, when the host is protected by the *Eco*RI methylase these phage form turbid instead of clear plaques. This system should prove useful for further mutagenic studies. For example, one could isolate *Eco*RI mutants which evade the protective effect of the methylase.

In conclusion, we describe a phage based conditional expression system for studying proteins which are lethal to the cell. Even though these phage kill their host cell, they produce plaques from which viable phage can be recovered. The vector employed is the well described M13mp18 phage which has several attractive features: one can identify insertions by the absence of α-complementation (white versus blue plaques), in one orientation the inserted gene comes under the regulated control of the wild-type *lac* promoter (from which expression can be induced with both IPTG and cAMP), and lastly one can readily obtain single-stranded DNA for site-directed mutagenesis or sequencing.

References

Adams, R. L. P., & Burdon, R. H. (1985). Methylation and Its Relationship to Transcription, p. 115-161. *In Molecular Biology of DNA Methylation.* Springer-Verlag, New York.

Aggarwal, A. K., Rodgers D. W., Drottar, M., Ptashne, M. & Harrison, S. C. (1988). Recogniton of a DNA Operator by the Repressor of Phage 434: A View at High Resolution. *Science* **242**: 899-907.

Akroyd, J. & Symonds, N. (1986). Localization of the *gam* gene of bacteriophage Mu and characterisation of the gene product. *Gene* **49**: 273-282.

Alber, T., Dao-pin, S., Nye, J. A., Muchmore, D. C., & Matthews, B. W. (1987a). Temperature-Sensitive Mutations of Bacteriophage T4 Lysozyme Occur at Sites with Low Mobility and Low Solvent Accessibility in the Folded Protein. *Biochemistry* **26**: 3754-3758.

Alber, T. Dao-pin, S., Wilson, K., Wozniak, J. A., Cook, S. P., & Matthews, B. W. (1987b). Contributions of hydrogen bonds of Thr157 to the thermodynamic stability of phage T4 lysozyme. *Nature* **330**: 41-46.

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

Anderson, W. F., Takeda, Y., Ohlendorf, D. H., & Matthews, B. W. (1982). Proposed α-Helical Super-Secondary Structure Associated with Protein-DNA Recognition. *J. Mol. Biol.* **159**: 745-751.

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

Arber, W. (1974). DNA Modification and Restriction, p. 1-37. *In* Waldo E. Cohn (ed.), *Progress in Nucleic Acid Research and Molecular Biology.* Academic Press, NY, NY.

Arber, W. (1979). Promotion and Limitation of Genetic Exchange. *Science* **205**: 361-365.

Bachman, B. (1987). Derivation and Genotypes of Some Mutant Derivitives of *Escherichia coli* K-12, p. 1190-1219, *In* F. C. Neidhardt (ed.), *Escherichia coli* and *Salmonella typhimurium*, ASM, Washington, D. C.

Barany, F. (1985). Two-codon insertion mutagenesis of plasmid genes by using single-stranded hexameric oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* **82**: 4204-4206.

Barany, F. (1988). The *Taq*I star reaction: Strand preferences reveal hydrogen bond donor and acceptor sites in canonical sequence recognition. *Gene* **65**: 149-165.

Barnes, G. & Rine, J. (1985). Regulated expression of endonuclease *Eco*RI in *Saccharomyces cerevisiae*: Nuclear entry and biological consequences. *Proc. Natl. Acad. Sci. U. S. A.* **82**: 1354-1358.

Bass, S., Sorrells, V. & Youderian, P. (1988). Mutant Trp Repressors with New DNA-Binding Specificities. *Science* **242**: 240-245.

Benson, N., Sugiono, P., & Youderian, P. (1988). DNA Sequence Determinants of λ Repressor Binding *In Vivo*. *Genetics* **188**: 21-29.

Benson, N., & Youderian, P. (1989). Phage λ Cro Protein and *cI* Repressor use Two Different Patterns of Specific Protein-DNA Interactions to Achieve Sequence Specificity in vivo. *Genetics* **121**: 5-12.

Berkner, K. L. & Folk, W. R. (1978). Overmethylation of DNAs by the *Eco*RI Methylase. *Nucl. Acids Res.* **5**: 435-450.

Bertani, G. & Weigle, J. J. (1953). Host-Controlled variation in bacterial viruses. *J. of Bacteriol.* **65**: 113-121.

Bhagwat, A. S., Sohail, A., & Roberts, R. J. (1986). Cloning and Characterization of the *dcm* Locus of *Escherichia coli* K-12. *J. of Bacteriol.* **166**: 751-755.

Bhagwat, A. S., & Roberts, R. J. (1987). Genetic Analysis of the 5-Azacytidine Sensitivity of *Escherichia coli* K-12. *J. of Bacteriol.* 169: 1537-1546.

Bickle, T. (1987). DNA Restriction and Modification Systems, p. 692-969, *In*, F. C. Neidhardt (ed.), *Escherichia coli* and *Salmonella typhimurium*, ASM, Washington, D. C.

Bishop, J. O. (1979). A DNA Sequence Cleaved by Restriction Endonuclease R.*Eco*RI in Only One Strand. *J. Mol. Biol.* **128**: 545-559.

Blumenthal, R. M., Gregory, S. A., & Cooperider, J. S. (1985). Cloning of a Restriction-Modification System from *Proteus vulgaris* and Its Use in Analyzing a Methylase-Sensitive Phenotype in *Escherichia coli*. *J. of Bacteriol.* **164**:501-509.

Boeke, J. D. (1981). One and Two Codon Insertion Mutants of Bacteriophage f1. *Mol. Gen. Genet.* **181**: 288-291.

Boeke, J. D., Model, P., & Zinder, N. D. (1982). Effects of Bacteriophage f1 Gene III Protein on the Host Cell Membrane. *Mol. Gen. Genet.* **186**: 185-192.

Bougueleret, L., Schwarzstein, M., Tsugita, A., & Zabeau, M. (1984). Characterization of the genes coding for the *Eco*RV restriction and modification system of *Escherichia coli*. *Nucl. Acids Res.* **12**: 3659-3675.

Bowie, J. U., and Sauer, R. T. (1989). Identifying determinants of folding activity for a protein of unknown structure. *Proc. Natl. Acad. Sci. U. S. A.* **86**: 2152-2156.

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

Brayer, G. D., & McPherson, A. (1984). Mechanism of DNA Binding to the Gene V Protein of Bacteriophage fd. *Biochemistry* **23**: 340-349.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986a). The Effects of Base Analogue Substitutions on the Cleavage by the *Eco*RI Restriction Endonuclease of Octadeoxyribonucleotides Containing Modified *Eco*RI Recognition Sequences. *J. Biol.*

Chem. **261**: 7270-7278.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986b). The Effects of Base Analogue Substitutions on the Methylation by the *Eco*RI Modification Methylase of Octadeoxyribonucleotides Containing Modified *Eco*RI Recognition Sequences. *J. Biol. Chem.* **261**: 7279-7286.

Burns, P. A., Gordon, A. J. E., & Glickman, B. W. (1987). Influence of Neighboring Base Sequence on N-methyl-N'-nitro-N-nitrosoguanidine Mutagenesis in the *lacI* gene of *Escherichia coli. J. Mol. Biol.* **194**: 385-390.

Casadaban, M. J., & Cohen, S. N. (1979). Lactose genes fused to exogenous promoters in one step using a Mu-*lac* bacteriophage: *In vivo* probe for transcriptional control sequences. *Proc. Natl. Acad. Sci. U. S. A.* **76**: 4530-4533.

Cedar, H. (1984). DNA Methylation and Gene Expression, p. 147-164. *In* A. Razin, H. Cedar, and A. D. Riggs (ed.), *DNA Methylation, Biochemistry and Biological Significance.* Springer-Verlag, New York.

Cedar, H. (1988). DNA Methylation and Gene Activity. *Cell* **53**: 3-4.

Chang, S. & Cohen, S. N. (1977). *In vivo* site-specific genetic recombination by the *Eco*RI endonuclease. *Proc. Natl. Acad. Sci. U. S. A.* **74**: 4811-4815.

Cheng, S-C., & Modrich, P. (1983). Positive-selection Cloning Vehicle Useful for Overproduction of Hybrid Proteins. *J. of Bacteriol.* **154**: 1005-1008.

Cheng, S-C., Kim, R., King, K., Kim, K., & Modrich, P. (1984). Isolation of Gram Quantities of *Eco*RI Restriction and Modification Enzymes from an Over-Producing Strain. *J. Biol. Chem* **259**: 11571-11575.

Chou, P. Y. & Fasman, G. D. (1974). Prediction of protein conformation. *Biochemistry* **13**: 222-245.

Chou, P. Y. & Fasman, G. D. (1978). Prediction of secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* **47**: 45-147.

Claverys, J-P., & Lacks, S. A. (1986). Heteroduplex Deoxyribonucleic Acid Base Mismatch Repair in Bacteria. *Micro. Rev.* **50**:133-165.

Cohen, S. N., Chang, A. C. Y., Boyer, H. W., & Helling, R. B. (1973). Construction of Biologically Functional Bacterial Plasmids *In Vitro. Proc. Natl. Acad. Sci. U. S. A.* **70**: 3240-3244.

Conley, E. C., & Saunders, J. R. (1984). Recombination-dependent recircularization of linearized pBR322 plasmid DNA following transformation of *Escherichia coli. Mol. Gen. Genet.* **194**: 211-218.

Conley, E. C., Saunders, V. A., & Saunders, J. R. (1986a). Deletion and rearrangement of plasmid DNA during transformation of *Escherichia coli* with linear plasmid molecules. *Nucl. Acids Res.* **14**: 8905-8917.

Conley, E. C., Saunders, V. A., Jackson, V., & Saunders, J. R. (1986b). Mechanism of intramolecular recyclization and deletion formation following transformation of *Escherichia coli* with linearized plasmid DNA. *Nucl. Acids Res.* **14**: 8919-8932.

Connolly, B. A., Eckstein, F., & Pingoud, A. (1984). The Stereochemical Course of the Restriction Endonuclease *Eco*RI Catalyzed Reaction. *J. Biol. Chem.* **259**: 10760-10763.

Coulondre, C., Miller, J. H., Farabough, P. J., & Gilbert, W. (1978). Molecular basis of base substitution hotspots in *Escherichia coli. Nature* **274**: 775-780.

Cox, E. C. (1976). Bacterial Mutator Genes and the Control of Spontaneous Mutation. *Ann. Rev. Genet.* **10**: 135-156.

Cox, E. C. & Horner, D. L. (1982). Dominant Mutators in *Escherichia coli. Genetics* **100**: 7-18.

Craig, R. J., Arraj, J. A., & Marinus, M. G. (1984). Induction of damage inducible (SOS) repair in *dam* mutants of *Escherichia coli* exposed to 2-aminopurine. *Mol. Gen. Genet.* **194**: 539-540.

Davis, N. G., & Model, P. (1985). An Artificial Anchor Domain: Hydrophobicity Suffices to Stop Transfer. *Cell* **41**: 607-614.

Day, R. S. III (1977). UV-Induced Alleviation of K-Specific Restriction of Bacteriophage λ. *J. Virol.* **21**: 1249-1251.

deBoer, H. A., Comstock, L. J. & Vasser, M. (1983). The *tac* promoter: A functional hybrid derived from the *trp* and *lac* promoters. *Proc. Natl. Acad. Sci. U. S. A.* **80**: 21-25.

Deutscher, M. P. & Kornberg, A. (1969). Enzymatic Synthesis of DNA. XXIX. Hydrolysis of DNA from the 5' terminus by an exonuclease function of DNA polymerase. *J. Biol. Chem.* **244**: 3029-3037.

Dharmalingam, K. & Goldberg, E. B. (1980). Restriction in vivo V. Induction of SOS Functions in *Escherichia coli* by Restricted T4 Phage DNA, and Alleviation of Restriction by SOS Functions. *Mol. Gen. Genet.* **178**: 51-58.

Dickerson, R. E., & Drew, H. R. (1981). Structure of a B-DNA Dodecamer II. Influence of Base Sequence on Helix Structure. *J. Mol. Biol.* **149**: 761-786.

Dila, D. & Raleigh, E. A. (1988). Genetic dissection of the methylcytosine-specific restriction system *mcrB* of *Escherichia coli* K-12. *Gene* **74**: 23-24.

Doerfler, W. (1983). DNA Methylation and Gene Activity. *Ann. Rev. Biochem.***52**: 93-124.

Drew, H. R. & Dickerson, R. E. (1981). Structure of a B-DNA Dodecamer III. Geometry of Hydration. *J. Mol. Biol.* **151**: 535-556.

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

Drlica, K. & Rouvierre-Yaniv, J. (1987). Histone-like Proteins of Bacteria. *Microbiol. Rev.* **51**: 301-319.

Duncan, B. K. (1985). Isolation of Insertion, Deletion, and Nonsense Mutations of the

Uracil-DNA Glycosylase (*ung*) Gene of *Escherichia coli* K-12. *J. of Bacteriol.* **164**: 689-695.

Eckert, R. L. (1987). Sequencing by the Chemical Method, p. 7.5.1-7.5.9, *In* F. M. Ausubel (ed.), *Current Protocols in Molecular Biology,* Greene Publishing Associates, New York.

Efimova, E. P., Delver, E. P. & Belogurov, A. A. (1988). Alleviation of type I restriction in adenine methylase (*dam*) mutants of *Escherichia coli. Mol. Gen. Genet.* **214**: 313-316.

Ennis, D. G., Amundsen, S. K., & Smith, G. R. (1987). Genetic functions promoting homologous recombination in *Escherichia coli*: a study of inversions in phage λ. *Genetics* **115**: 11-24.

Epstein, W., & Kim, B. S. (1971). Potassium Transport Loci in *Escherichia coli* K-12. *J. of Bacteriol.* **108**: 639-644.

Evans, R. M. & Hollenberg, S. M. (1988). Zinc Fingers: Gilt by Association. *Cell* **52**: 1-3.

Fane, B. & King, J. (1987). Identification of Sites Influencing the Folding and Subunit Assembly of the P22 Tailspike Polypeptide Chain Using Nonsense Mutations. *Genetics* **117**: 157-171.

Feiss, M., Kobayashi, W. & Widner, W. (1983). Separate sites for binding and nicking of bacteriophage λ DNA by terminase. *Proc. Natl. Acad. Sci. U. S. A.* **80**: 955-959.

Frederick, C. A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobsen, L., Wang, B-C., Greene, P., Boyer, H. W., & Rosenberg, J. M. (1984). Kinked DNA in Crystalline Complex with *Eco*RI Endonuclease. *Nature* **309**: 327-331.

Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R., Steitz, T. A. (1988). Cocrystal Structure of an editing complex of Klenow fragment with DNA. *Proc. Natl. Acad. Sci. U. S. A.* **85**: 8924-8928.

Friedberg, E. C. (1985). *DNA Repair,* W. H. Freeman, New York.

Friedberg, E. C. (1988). Deoxyribonucleic Acid Repair in the Yeast *Saccharomyces cerevisiae. Microbiol. Rev.* **52**: 70-102.

Fulford, W., Russel, M. & Model, P. (1986). Aspects of the Growth and Regulation of the Filamentous Phages. *Prog. in Nucl. Acid Res.* **33**: 141-168.

Gardner, R. C., Howarth, A. J., Messing, J. & Shepherd, R. J. (1982). Cloning and Sequencing of Restriction Fragments Generated by *Eco*R*. *DNA* **1**: 109-115.

George, J. & Chirikjian, J. G. (1982). Sequence-specific endonuclease *Bam*HI: Relaxation of sequence recognition. *Proc. Natl. Acad. Sci. U. S. A.* **79**: 2432-2436.

Giloni, L., Takeshita, M., Johnson, F., Ideu, C., & Grollman, A. P. (1981). Bleomycin-induced Strand-Scission of DNA. *J. Biol. Chem.* **256**: 8608-8615.

Gingeras, T. R. & Brooks, J. E. (1983). Cloned restriction/modification system from *Pseudomonas aeruginosa. Proc. Natl. Acad. Sci. U. S. A.* **80**: 402-406.

Glickman, B., van den Elsen, P., & Radman, M. (1978). Induced Mutagenesis in *dam*⁻ Mutants of *Escherichia coli*: A Role for 6-methyladenine Residues in Mutation Avoidance. *Mol. Gen. Genet.* **163**: 307-312.

Glickman, B. W., & Radman, M. (1980). *Escherichia coli* mutator mutants deficient in methylation-instructed DNA mismatch repair. *Proc. Natl. Acad. Sci. U.S.A.* **77**: 1063-1067.

Goff, S., Short-Russell, S. R., & Dice, J. F. (1987). Efficient Saturation Mutagenesis of a Pentapeptide Coding Sequence Using Mixed Oligonucleotides. *DNA* **6**: 381-388.

Gottesman, M. M., Hicks, M. L. & Gellert, M. (1973). Genetics and Function of DNA Ligase in *Escherichia coli*. *J. Mol. Biol.* **77**: 531-547.

Gough, J. A., & Murray, N. E. (1983). Sequence Diversity among Related Genes for Recognition of Specific Targets in DNA Molecules. *J. Mol. Biol.* **166**: 1-19.

Graddis, T. J., Klig, L. S., Yanofsky, C. & Oxender, D. L. (1988). Formation of Heterodimers Between Wild-Type and Mutant *trp* Aporepressor Polypeptides of *Escherichia coli*. *Proteins: Struc., Func., & Genetics* **4**: 173-181.

Greene, P. J., Gupta, M., Boyer, H. W., Brown, W. E., & Rosenberg, J. M. (1981). Sequence analysis of the DNA encoding the *Eco*RI endonuclease and methylase. *J. Biol. Chem.* **256**: 2143-2153.

Greene, P. J., Yanofsky, S., Reich, N., Day, J., Hager, P., & Boyer, H. W. (1987). Structure and Function of the *Eco*RI Endonuclease, p. 3-7, *In* D. L. Oxender (ed.), *Protein structure, Folding and Design 2.* Alan R. Liss, New York.

Grutter, M. G., Weaver, L. H., Gray, T. M., & Matthews, B. W. (1983). Structure, Function, and Evolution of the Lysozyme from Bacteriophage T4, p. 356-360, *In* C. K. Mathews, E. M. Kutter, G. Mosig, & P. B. Berget (eds.), *Bacteriophage T4*, American Society for Microbiology, Washington, D. C.

Guarente, L., Lauer, G., Roberts, T. M., & Ptashne, M. (1980). Improved Methods for Maximizing Expression of a Cloned Gene: a Bacterium That Synthesizes Rabbit β-Globin. *Cell* **20**: 543-553.

Harrison, G. S., Findly, R. C., & Karrer, K. M. (1986). Site-Specific Methylation of Adenine in the Nuclear Genome of a Eucaryote, *Tetrahymena thermophila*. *Mol. Cell. Biol.* **6**: 2364-2370.

Hattman, S., Brooks, J. E., & Masurekar, M. (1978). Sequence Specificity of the P1 Modification Methylase (M.*Eco* P1) and the DNA Methylase (M.*Eco* dam) Controlled by the *Escherichia coli dam* gene. *J. Mol. Biol.* **126**: 367-380.

Hattman, S. (1983). DNA Modification: Methylation, p. 152-155. *In* C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget (ed.), *Bacteriophage T4*. American Society for Microbiology, Washington, D. C.

Haynes, R. H. & Kunz, B. A. (1981). DNA repair and mutagenesis in yeast, p. 371-414, *In* J. Strathern, E. W. Jones, & J. R. Broach (eds.), *The molecular biology of the yeast Saccharomyces. Life cycle and Inheritance*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Hedgpeth, J., Goodman, H. M. & Boyer, H. W. (1972). DNA Nucleotide Sequence Restricted by the RI Endonuclease. *Proc. Natl. Acad. Sci. U. S. A.* **69**, 3448-3452.

Heitman, J. & Model, P. (1987). Site-Specific Methylases Induce the SOS DNA Repair Response in *Escherichia coli*. *J. Bacteriol.* **169**, 3243-3250.

Heitman, J., Zinder, N. D., & Model, P. (1989). Repair of the *Escherichia coli* chromosome after *in vivo* scission by the *Eco*RI endonuclease. *Proc. Natl. Acad. Sci. U. S. A.* **86**: 2281-2285.

Henner, W. D., Rodriguez, L. O., Hecht, S. M. & Haseltine, W. A. (1983). γ-Ray Induced Deoxyribonucleic Acid Strand Breaks. *J. Biol. Chem.* **258**: 711-713.

Herskowitz, I. (1987). Functional inactivation of genes by dominant negative mutations. *Nature* **329**: 219-222.

Ho, K. S. Y. & Mortimer, R. K. (1973). Induction of Dominant Lethality by X-rays in a Radiosensitive Strain of Yeast. *Mutat. Res.* **20**: 45-51.

Ho, K. S. Y. (1975). Induction of DNA double strand breaks by X-rays in a radiation sensitive strain of the yeast *Saccharomyces cerevisiae*. *Mutat. Res.* **30**: 327-334.

Howard-Flanders, P., & Theriot, L. (1966). Mutants of *Escherichia coli* K-12 Defective in DNA Repair and in Genetic Recombination. *Genetics* **53**: 1137-1150.

Hsu, M-t., & Berg, P. (1978). Altering the Specificity of Restriction Endonucleases: Effect of Replacing Mg^{+2} with Mn^{+2}. *Biochemistry* **17**: 131-138.

Huisman, O., D'Ari, R. & Gottesman, S. (1984). Cell-division control in *Escherichia coli*: Specific induction of the SOS function SfiA protein is sufficient to block septation. *Proc. Natl. Acad. Sci. U. S. A.* **81**: 4490-4494.

Irino, N., Nakayama, K., & Nakayama, H. (1986). The *recQ* of *Escherichia coli* K-12: primary structure and evidence for SOS regulation. *Mol. Gen. Genet.* **205**: 298-304.

Janulaitis, A., Klimasaukas, S., Petrusyte, M., & Butkus, V. (1983). Cytosine modification in DNA by *Bcn*I methylase yields N^4-methylcytosine. *FEBS Lett.* **161**: 131-134.

Jayaram, J. (1986). Double-strand-break repair in Yeast Results in Conversion Events that Resemble Mating-type Switching, p. 19-24, *In* A. Klar & J. Strathern (eds.), *Mechanisms of Yeast Recombination*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Jen-Jacobsen, L., Lesser, D., & Kurpiewski, M. (1986). The Enfolding Arms of *Eco*RI Endonuclease: Role in DNA Binding and Cleavage. *Cell* **45**: 619-629.

Jen-Jacobsen, L., Kurpiewski, M., Lesser, D., Grable, J., Boyer, H. W., Rosenberg, J. M., & Greene, P. J. (1983). Coordinate Ion Pair Formation between *Eco*RI Endonuclease and DNA. *J. Biol. Chem.* **258**: 14638-14646.

Jordan, S. R. & Pabo, C. O. (1988). Structure of the Lambda Complex at 2.5 A Resolution: Details of the Repressor-Operator Interactions. *Science* **242**: 893-899.

Joyce, C. M. & Grindley, N. D. F. Relaxed Specificity cleavage by endonucleases *Eco*RI and *Bst*I as a model for protein-DNA recognition, unpublished manuscript.

Kelley, R. L. & Yanofsky, C. (1985). 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. U. S. A.* **82**: 483-487.

Kelly, R. B., Cozzarelli, N. R., Deutscher, M. P., Lehman, I. R., & Kornberg, A. (1970). Enzymatic synthesis of deoxyribonucleic acid XXXII. Replication of duplex deoxyribonucleic acid by polymerase at a single strand break. *J. Biol. Chem.* **245**: 39-45.

Kenyon, C. J., & Walker, G. C. (1980). DNA-Damaging agents stimulate gene expression at specific loci in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **77**: 2819-2823.

King, J., Haase, A. & Yu, M.-H. (1987). Temperature sensitive mutations affecting kinetic steps in protein folding pathways, p. 109-121, *In* D. Oxender and C. F. Fox (eds.), *Protein Engineering*, Alan R. Liss, New York.

King, K., Wright, D., & Modrich, P. (1986). Mutationally Altered forms of *Eco*RI Endonuclease Selectively Defective in DNA Cleavage. *Fed. Proc.* **45**: 1913 (abstract 2523).

King, K., Wright, D. J., & Modrich, P. (1988). Glutamate 111 of *Eco*RI Endonuclease is Required for DNA Cleavage Activation, p.79, *In* K. Brew *et al.* (ed.), ICSU Short Reports vol. 8: Advances in Gene Technology: Protein Engineering and Production, IRL Press, Washington, D. C.

Kiss, A. & Baldauf, F. (1983). Molecular cloning and expression in *Escherichia coli* of two modification methylase genes of *Bacillus subtilis*. *Gene* **21**: 111-119.

Kiss, A., Posfai, G., Keller, C. C., Venetianer, P., & Roberts, R. J. (1985). Nucleotide Sequence of the *Bsu*RI restriction-modification system. *Nucl. Acids Res.* **13**: 6403-6421.

Kleckner, N., Chan, R. K., Tye, B-K., & Botstein, D. (1975). Mutagenesis by Insertion of a Drug-resistance Element Carrying an Inverted Repetition. *J. Mol. Biol.* **97**: 561-575.

Klig, L. S., Crawford, I. P., & Yanofsky, C. (1987). Analysis of *trp* repressor-operator interaction by filter binding. *Nucl. Acids Res.* **15**: 5339-5351.

Knight, K. L. & Sauer, R. T. (1989). DNA binding specificity of the Arc and Mnt repressors is determined by a short region of N-terminal residues. *Proc. Natl. Acad. Sci. U. S. A.* **86**: 797-801.

Kobayashi, I. & Takahashi, N. (1988). Double-Stranded Gap Repair of DNA by Gene Conversion in *Escherichia coli*. *Genetics* **119**: 751-757.

Kohara, Y., Akiyama, K., & Isono, K. (1987). The Physical Map of the Whole *E. coli* Chromosome: Application of a New Strategy for Rapid Analysis and Sorting of a Large Genomic Library. *Cell* **50**: 494-508.

Konrad, E. B., Modrich, P. & Lehman, I. R. (1973). Genetic and Enzymatic Characterization of a Conditional Lethal Mutant of *Escherichia coli* K-12 with a

Temperature-sensitive DNA ligase. *J. Mol. Biol.* **77**: 519-529.

Konrad, E. B. (1977). Method for the Isolation of *Escherichia coli* Mutants with Enhanced Recombination Between Chromosomal Duplications. *J. of Bacteriol.* **130**: 167-172.

Koshland, D. E., Jr., & Neet, K. E. (1968). The Catalytic and Regulatory Properties of Enzymes. *Ann. Rev. Biochem.* **37**: 359-410.

Koudelka, G. B., Harrison, S. C., & Ptashne, M. (1987). Effect of non-contacted bases on the affinity of 434 operator for 434 repressor and Cro. *Nature* **326**: 886-888.

Krasin, F. & Hutchinson, F. (1977). Repair of DNA Double-strand breaks in *Escherichia coli* which Requires recA Function and the Presence of a Duplicate Genome. *J. Mol. Biol.* **116**: 81-98.

Krasin, F. & Hutchinson, F. (1981). Repair of DNA Double-strand breaks in *Escherichia coli* cells requires synthesis of proteins that can be induced by UV light. *Proc. Natl. Acad. Sci. U. S. A.* **78**: 3450-3453.

Kunkel, T. A. (1985). Rapid and efficient site-specific mutagensis without phenotypic selection. *Proc. Natl. Acad. Sci. U. S. A.* **82**: 488-492.

Lackey, D., Krauss, S. W., & Linn, S. (1982). Isolation of an altered from of DNA polymerase I from *Escherichia coli* cells induced for *recA/lexA* functions. *Proc. Natl. Acad. Sci. U. S. A.* **79**: 330-334.

Lacks, S. & Greenberg, B. (1977). Complementary Specificity of Restriction Endonuclease of *Diplococcus pneumoniae* with Respect to DNA Methylation. *J. Mol. Biol.* **114**: 153-168.

Laughon, A. & Scott, M. P. (1984). Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. *Nature* **310**: 25-31.

Lieb, M. (1987). Bacterial Genes *mutL, mutS,* and *dam* Participate in Repair of Mismatches at 5-Methylcytosine Sites. *J. Bacteriol* **169**: 5241-5246.

Lorincz, A. T., & Reed, S. I. (1986). Sequence Analysis of Temperature-Sensitive Mutations in the *Saccharomyces cerevisiae* Gene CDC28. *Mol. Cell. Biol.* **6**: 4099-4103.

Lovett, S. T., & Clark, A. J. (1983). Genetic Analysis of Regulation of the RecF Pathway of Recombination in *Escherichia coli. J. of Bacteriol.* **153**: 1471-1478.

Lu, A-L., Jack, W. E., & Modrich, P. (1981). DNA Determinants Important in Sequence Recognition by *Eco*RI Endonuclease. *J. Biol. Chem.* **256**: 13200-13206.

Lu., A-L. & Chang, D-Y. (1988). A Novel Nucleotide Excision Repair for the Conversion of an A/G Mismatch to C/G Base Pair in *Escherichia coli. Cell* **54**: 805-812.

Lunnen, K. D., Barsomian, J. M., Camp, R. R., Card, C. O., Chen, S-Z., Croft, R., Looney, M. C., Meda, M. M., Moran, L. S., Nwankwo, D. O., Slatko, B. E., VanCott, E. M., & Wilson, G. G. (1988). Cloning type-II restriction and modification genes. *Gene* **74**: 25-32.

Lydersen, B. K. & Pettijohn, D. E. (1977). Interactions Stabilizing DNA Tertiary

Structure in the *Escherichia coli* Chromosome Investigated with Ionizing Radiation. *Chromosoma* **62**: 199-215.

MacNeil, D. J. (1988). Characterization of a Unique Methyl-Specific Restriction System in *Streptomyces avermitilis*. *J. of Bacteriol.* **170**: 5607-5612.

Mahan, M. J. & Roth, J. R. (1989). Role of *recBC* Function in Formation of Chromosomal Rearrangements: A Two-Step Model for Recombination. *Genetics* **121**: 433-443.

Malone, R. E. & Esposito, R. E. (1980). The *RAD52* gene is required for homothallic interconversion of mating types and spontaneous mitotic recombination in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **77**: 503-507.

Maloy, S. R. & Nunn, W. D. (1981). Selection for Loss of Tetracycline Resistance by *Escherichia coli*. *J. of Bacteriol.* **145**:1110-1112.

Malyguine, E., Vannier, P. & Yot, P. (1980). Alteration of the Specificity of Restriction Endonucleases in the Presence of Organic Solvents. *Gene* **8**: 163-177.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982). *In* Molecular Cloning, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

Mann, M. B., Rao, R. N., & Smith, H. O. (1978). Cloning of Restriction and Modification Genes in *E. coli*: the *Hha*II System from *Haemophilus haemolyticus*. *Gene* **3**: 97-112.

Marinus, M. G., & Morris, N. R. (1973). Isolation of Deoxyribonucleic Acid Methylase Mutants of *Escherichia coli* K-12. *J. of Bacteriol.* **114**: 1143-1150.

Marinus, M. G. & Morris, N. R. (1974). Biological Function for 6-Methyladenine Residues in the DNA of *Escherichia coli* K-12. *J. Mol. Biol.* **85**: 309-322.

Marinus, M. G. & Morris, N. R. (1975). Pleiotropic Effects of a DNA Adenine Methylation Mutation (*dam3*) in *Escherichia coli* K-12. *Mutat. Res.* **28**: 15-26.

Marinus, M. G., Poteete, A. & Arraj, J. A. (1984). Correlation of DNA adenine methylase activity with spontaneous mutability in *Escherichia coli* K-12. *Gene* **28**: 123-125.

Marinus, M. G. (1984). Methylation of procaryotic DNA, p. 81-109, *In* A Razin, M. Cedar, & A. D. Riggs (eds.), *DNA Methylation*, Springer Verlag, NY.

Marinus, M. G. (1987). Methylation of DNA, p. 697-702, *In* F. C. Neidhardt (ed.), *Escherichia coli and Salmonella typhimurium*, American Society of Microbiology, Washington, D. C.

Masamune, Y. & Richardson, C. C. (1971). Strand displacement during DNA synthesis at single strand breaks. *J. Biol. Chem.* **246**: 2692-2701.

Maxam, A. M. & Gilbert, W. (1980). Sequencing End-Labelled DNA with Base-Specific Chemical Cleavages. *Meth. Enz.* **65**: 499-560.

McClarin, J. A., Frederick, C. A., Wang, B-C., Greene, P., Boyer, H. W., Grable, J., & Rosenberg, J. M. (1988). Structure of the DNA-*Eco*RI Endonuclease Recognition

Complex at 3 A Resolution. *Science* **234**: 1526-1541.

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

McLaughlin, L. W., Benseler, F. B., Graeser, E., Piel, N., & Scholtissek, S. (1987). Effects of Functional Group Changes in the *Eco*RI Recognition Site on the Cleavage Reaction Catalyzed by the *Eco*RI Endonuclease. *Biochemistry* **26**: 7238-7245.

McPartland, A., Green, L. & Echols H. (1980). Control of *recA* gene RNA in *E. coli*: regulatory and signal genes. *Cell* **20**: 731-737.

Meselson, M. S. & Radding, C. M. (1975). A General Model for Genetic Recombination. *Proc. Natl. Acad. Sci. U. S. A.* **72**: 358-361.

Messer, W., Bellekes, U., & Lothar, H. (1985). Effect of *dam* Methylation on the activity of the *E. coli* replication origin, *oriC*. *EMBO J.* **5**: 1327-1332.

Messer, W. & Noyer-Weidner, M. (1988). Timing and Targeting: The Biological Functions of Dam Methylation in *E. coli*. *Cell* **54**: 735-737.

Michel, B. & Zinder, N. D. (1989). Translational repression in the bacteriophage f1: Characterization of the gene V protein target on the gene II mRNA. *Proc. Natl. Acad. Sci. U. S. A.* , in press.

Miller, J., McLachlen, A. D., & Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIa from *Xenopus* oocytes. *EMBO J.* **4**: 1609-1614.

Miller, J. H. (1972). Assay of β-galactosidase, p. 352-359. *In, Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

Miller, J. H. (1978). The *lacI* Gene: Its Role in *lac* Operon Control and Its Use as a Genetic System, p. 31-88, *In* J. H. Miller & W. S. Reznikoff (eds.), *The Operon*, Cold Spring Harbor Press, NY.

Miller, J. H. (1979). Genetic Studies of the *lac* Repressor: XI. On Aspects of *lac* Repressor Structure Suggested by Genetic Experiments. *J. Mol. Biol.* **131**: 249-258.

Modrich, P. & Rubin, R. A. (1977). Role of the 2-Amino group of Deoxyguanosine in Sequence Recognition by *Eco*RI Restriction and Modification Enzymes. *J. Biol. Chem.* **252**: 7273-7278.

Modrich, P. (1989). Methyl-directed DNA Mismatch Correction. *J. Biol. Chem.* **264**: 6597-6600.

Monod, J., Changeux, J-P., & Jacob, F. (1963). Allosteric Proteins and Cellular Control Systems. *J. Mol. Biol.* **6**: 306-329.

Morgan, W. F., Fero, M. L., Land, M. C., & Winegar, R. A. (1988). Inducible Expression and Cytogenetic Effects of the *Eco*RI Restriction Endonuclease in Chinese Hamster Ovary Cells. *Mol. Cell. Biol.* **8**: 4204-4211.

Mott, J. E., Grant, R. A., Ho, Y-S., & Platt, T. (1985). Maximizing gene expression from plasmid vectors containing the λ P_L promoter: Strategies for over-producing transcription termination factor ρ. *Proc. Natl. Acad. Sci. U. S. A.* **82**: 88-92.

Murialdo, H. (1988). Lethal Effect of λ Terminase in recombination deficient *Escherichia coli*. *Mol. Gen. Genet.* **213**, 42-49.

Murray, N. E. & Murray, K. (1974). Manipulation of restriction targets in phage λ to form receptor chromosomes for DNA fragments. *Nature* **251**: 476-481.

Nasri, M., Sayadi, S. & Thomas, D. (1985). Relaxation of *Pvu*II recognition sequence. *FEBS* **185**: 101-104.

Nasri, M. & Thomas, D. (1986). Relaxation fo recognition sequence of specific endonuclease *Hin*dIII. *Nucl. Acids Res.* **14**: 811-821.

Nathans, D. & Smith, H. O. (1975). Restriction Endonucleases in the Analysis and Restructuring of DNA Molecules. *Ann. Rev. Biochem.* **44**: 273-293.

Nathans, D. (1979). Restriction Endonucleases, Simian Virus 40, and the New Genetics. *Science* **206**: 903-909.

Nelson, H. C. M. & Sauer, R. T. (1986). Interaction of Mutant λ Repressors with Operator and Non-Operator DNA. *J. Mol. Biol.* **192**: 27-38.

Newman, A. K., Rubin, R. A., Kim, S-H. & Modrich, P. (1981). DNA Sequences of Structural Genes for *Eco*RI DNA Restriction and Modification Enzymes. *J. Biol. Chem.* **256**: 2131-2139.

Noyer-Weidner, M., Diaz, R., & Reiners, L. (1986). Cytosine-specific DNA modification interferes with plasmid establishment in *Escherichia coli* K-12: Involvement of *rglB*. *Mol. Gen. Genet.* **205**: 469-475.

Ogawa, H. & Tomizawa, J. I. (1967). Breakage of Polynucleotide Strands by Disintegration of Radiophosphorus Atoms in DNA Molecules and their Repair I. *J. Mol. Biol.* **30**: 1-7.

Ohlendorf, D. H., Anderson, W. F., Fisher, R. G., Takeda, Y., & Matthews, B. W. (1982). The molecular basis of DNA-protein recognition inferred from the structure of *cro* repressor. *Nature* **298**: 718-723.

Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G. & Steitz, T. A. (1985). Structure of the large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* **313**: 762-766.

Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. (1981). Yeast Transformation: A model system for the study of recombination. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 6354-6358.

Otwinowski, Z., Schevitz, R. W., Zhang, R-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F. & Sigler, P. B. (1988). Crystal Structure of *trp* repressor/operator complex at atomic resolution. *Nature* **335**: 321-329.

Pabo, C. O. & Lewis, M. (1982). The operator-binding domain of λ repressor: structure and DNA recognition. *Nature* **298**: 443-447.

Pabo, C. O. & Sauer, R. T. (1984). Protein-DNA Recognition. *Ann. Rev. of Biochem.* **53**: 293-321.

Peterson, K. R., Wertman, K. F., Mount, D. W., & Marinus, M. G. (1985). Viability of *Escherichia coli* K-12 DNA adenine methylase (*dam*) mutants requires increased expression of specific genes in the SOS regulon. *Mol. Gen. Genet.* **201**: 14-19.

Peterson, K. R., Ossana, N., Thliveris, A. T., Ennis, D. G., & Mount, D. W. (1988). Derepression of Specific Genes Promotes DNA Repair and Mutagenesis in *Escherichia coli*. *J. of Bacteriol.* **170**: 1-4.

Pettijohn, D. E. (1988). Histonelike Proteins and Bacterial Chromosome Structure. *J. Biol. Chem.* **263**: 12793-12796.

Picksley, S. M., Attfield, P. V., & Lloyd, R. G. (1984). Repair of DNA double-stramd breaks in *Escherichia coli* K-12 requires a functional recN product. *Mol. Gen. Genet.* **195**: 267-274.

Pingoud, A. (1985). Spermidine increases the accuracy of type II restriction endonucleases. *Eur. J. Biochem.* **147**: 105-109.

Polisky, B., Greene, P., Garfin, D. E., McCarthy, B. J., Goodman, H. M. & Boyer, H. W. (1975). Specificity of Substrate Recognition by the *Eco*RI restriction endonuclease. *Proc. Natl. Acad. Sci. U. S. A.* **72**: 3310-3314.

Pongor, S. & Szalay, A. A. (1985). Prediction of homology and divergence in the secondary structure of polypeptides. *Proc. Natl. Acad. Sci. U. S. A.* **82**: 366-370.

Posfai, G., Baldauf, F., Erdai, S., Posfai, J., Venetianer, P. & Kiss, A. (1984). Structure of the gene coding for the sequence-specific DNA-methyltransferase of the *B. subtilis* phage SPR. *Nucl. Acids Res.* **12**: 9039-9049.

Presta, L. G. & Rose, G. D. (1988). Helix Signals in Proteins. *Science* **240**: 1632-1641.

Pribnow, D., Sigurdson, D. C., Gold, L., Singer, B. S., Napoli, C., Brosius, J., Dull, T. J., Noller, H. F. (1981). rII Cistrons of Bacteriophage T4: DNA Sequence Around the Intercistronic Divide and Positions of Genetic Landmarks. *J. Mol. Biol.* **149**: 337-376.

Pukkila, P. J., Peterson, J., Herman, G., Modrich, P. & Meselson, M. (1983). Effects of High Levels of DNA Adenine Methylation on Methyl-Directed Mismatch Repair in *Escherichia coli*. *Genetics.* **104**: 571-582.

Radman, M., Wagner Jr., R. E., Glickman, B. W., & Meselson, M. (1980). DNA Methylation, mismatch correction and genetic stability, p. 121-130. *In* M. Alacevic (ed.), *Progress in Environmental Mutagenesis.* Elsevier/North Holland, Biomedical Press, Amsterdam.

Radman, M. & Wagner, R. (1984). Effects of DNA methylation on mismatch repair, mutagenesis, and recombination in *Escherichia coli*. *Curr. Top. Microbiol. Immunol.* **108**: 23-28.

Raleigh, E. A., & Wilson, G. (1986). *E. coli* K-12 restricts DNA containing 5-methylcytosine. *Proc. Natl. Acad. Sci. U. S. A.* **83**: 9070-9074.

Raleigh, E. A., Trimarchi, R. & Revel, H. (1989). Genetic and physical mapping of the *mcrA (rglB)* and *mcrB (rglA)* loci of *Escherichia coli* K-12, *Genetics*, in press.

Ravetch, J. V., Horiuchi, K. & Zinder, N. D. (1978). Nucleotide sequence of the recognition site for the restriction-modification enzyme of *Escherichia coli* B. *Proc. Natl. Acad. Sci. U. S. A.* **75**: 2266-2270.

Razin, A., Cedar, H., & Riggs, A. D. (1984). Introduction and General Overview, p. 1-10. *In* A. Razin, H. Cedar, and A. D. Riggs (ed.), *DNA Methylation*, Biochemistry and Biological Significance. Springer-Verlag, New York.

Register III, J. C. & Griffith, J. (1986). RecA protein filaments can juxtapose DNA ends: An activity that may reflect a function in DNA repair. *Proc. Natl. Acad. Sci. U. S. A.* **83**: 624-628.

Reidhaar-Olson, J. F. & Sauer, R. T. (1988). Combinatorial Cassette Mutagenesis as a Probe of the Informational Content of Protein Sequences. *Science* **241**: 53-57.

Resnick, M. A. (1969). Genetic Control of Radiation Sensitivity in *Saccharomyces cerevisiae*. *Genetics.* **62**: 519-531.

Resnick, M. A. (1975). The Repair of Double-Strand Breaks in Chromosomal DNA of Yeast, p. 549-556, *In* P. C. Hanawalt (ed.), *Molecular Mechanisms for Repair of DNA*, Plenum Press, New York.

Resnick, M. A. (1976). The Repair of Double-strand Breaks in DNA: A Model Involving Recombination. *J. Theor. Biol.* **59**: 97-106.

Resnick, M. A. & Martin, P. (1976). The repair of double strand breaks in the nuclear DNA of *Saccharomyces cerevisiae* and its genetic control. *Mol. Gen. Genet.* **143**: 119-129.

Revel, H. R. (1983). DNA Modification: Glucosylation, p. 156-165. *In* C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget (ed.), *Bacteriophage T4*. American Society for Microbiology, Washington, D. C.

Richards, F. M. & Kundrot, C. E. (1988). Identification of Structural Motifs From Protein Coordinate Data: Secondary Structure and First-Level Supersecondary Structure. *Proteins: Struc., Func. and Genet.* **3**: 71-84.

Richardson, J. S. & Richardson, D. C. (1988). Amino Acid Preferences for Specific Locations at the Ends of α-Helices. *Science* **240**: 1648-1652.

Rine, J. & Barnes, G. (1986). Entry of a Procaryotic Endonuclease into the Nucleus of *Saccharomyces cerevisiae*, p. 395-413. *In* J. Hicks (ed.), *Yeast Cell Biology*, Alan R. Liss, New York.

Roberts, D. & Kleckner, N. (1988). Tn*10* Transposition promotes RecA-dependent induction of a λ prophage. *Proc. Natl. Acad. Sci. USA* **85**: 6037-6041.

Roberts, R. J. (1976). Restriction Endonucleases. *CRC Critical Rev. in Biochem.* **4**: 123-164.

Rosenberg, J. M. & Greene, P. (1982). *Eco*RI* Specificity and Hydrogen Bonding. *DNA* **1**: 117-124.

Rosenberg, M., Ho, Y-S., & Shatzman, A. (1983). The Use of pKC30 and Its Derivatives for Controlled Expression of Genes. *Meth. Enz.* **101**: 123-138.

Rosner, J. L. (1973). Modification-Deficient Mutants of Bacteriophage P1 I: Restriction by P1 Cryptic Lysogens. *Virology* **52**: 213-222.

Ross, T. K., Achberger, E. C., & Braymer, H. D. (1987). Characterization of the *Escherichia coli* Modified Cytosine Restriction (*mcrB*) Gene. *Gene* **61**: 277-289.

Ross, T. K., Achberger, E. C., & Braymer, H. D. (1989). Nucleotide Sequence of the *mcrB* Region of *Escherichia coli* K-12 and Evidence for Two Independent Translational Initiation Sites at the *mcrB* Locus. *J. of Bacteriol.* **171**: 1974-1981.

Rubin, R. A. & Modrich, P. (1977). *Eco*RI Methylase. *J. Biol. Chem.* **252**: 7265-7272.

Russel, M. & Model, P. (1983). A Bacterial Gene, *fip*, Required for Filamentous Bacteriophage f1 Assembly. *J. of Bacteriol.* **154**: 1064-1076.

Russel, M. & Model, P. (1985). Direct Cloning of the *trxB* Gene that Encodes Thioredoxin Reductase. *J. of Bacteriol.* **163**: 238-242.

Russel, M., Kidd, S. & Kelley, M. R. (1986). An improved filamentous helper phage for generating single-stranded plasmid DNA. *Gene* **45**: 333-338.

Russel, M. & Model, P. (1986). The Role of Thioredoxin in Filamentous Phage Assembly. *J. Biol. Chem.* **261**: 14997-15005.

Sain, B. & Murray, N. E. (1980). The *hsd* (Host Specificity) Genes of *E. coli* K-12. *Molec. Gen. Genet.* **180**: 35-46.

Samson, L., & Cairns, J. (1977). A new pathway for DNA repair in *Escherichia coli*. *Nature* **267**: 281-283.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**: 5463-5467.

Schlief, R. (1988). DNA Binding by Proteins. *Science* **241**: 1182-1187.

Schoner, B., Kelly, S. & Smith, H. O. (1983). The nucleotide sequence of the *Hha*II restriction and modification genes from *Haemophilus haemolyticus*. *Gene* **24**: 227-236.

Seela, F. & Driller, H. (1986). Palindromic oligonucleotides containing 7-deaza-2'-deoxyguanosine: solid phase synthesis of d(pG GAATTCC) octamers and recognition by the endodeoxyribonuclease *Eco*RI. *Nucl. Acids Res.* **14**: 2319-2332.

Seela, F. & Kehne, A. (1987). Palindromic Octa- and Dodecanucleotides containing 2'-Deoxytubercidin: Synthesis, Hairpin Formation, and Recognition by the Endodeoxyribonuclease *Eco*RI. *Biochemistry* **26**: 2232-2238.

Seeman, N. C., Rosenberg, J. M. & Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. U. S. A.* **73**: 804-808.

Shapira, S. K., Chou, J., Richard, F. V., & Casadaban, M. J. (1983). New versatile plasmid vectors for expression of hybrid proteins coded by a cloned gene fused to *lacZ* gene sequences encoding an enzymatically active carboxy-terminal portion of β-galactosidase. *Gene* **25**: 71-82.

Shimatake, H. & Rosenberg, M. (1981). Purified λ regulatory protein *c*II positively activates promoters for lysogenic development. *Nature* **292**: 128-132.

Simmon, V. F. & Lederberg, S. (1972). Degradation of bacteriophage lambda deoxyribonucleic acid after restriction by *Escherichia coli* K-12. *J. Bacteriol.* **112**: 161-169.

Shortle, D. (1983). A genetic system for analysis of Staphyloccal nuclease. *Gene* **22**: 181-189.

Sladek, T. L., Nowak, J. A., & Maniloff J. (1986). Mycoplasma Restriction: Identification of a New Type of Restriction Specificity for DNA Containing 5-Methylcytosine. *J. of Bacteriol.* **165**: 219-225.

Smith, C. L. & Oishi, M. (1978). Early Events and Mechanisms in the induction of bacterial SOS functions: Analysis of the phage repressor inactivation process *in vivo*. *Proc. Natl. Acad. Sci. U. S. A.* **75**: 1657-1661.

Smith, D. W., A. M. Garland, G. Herman, R. E. Enns, T. A. Baker, & Zyskind, J. W. (1985). Importance of state of methylation of *oriC* GATC sites in initiation of DNA replication in *Escherichia coli*. *EMBO J.* **4**: 1319-1326.

Smith, G. R. (1988). Homologous Recombination in Procaryotes. *Microbiol. Rev.* **52**: 1-28.

Smith, H. O. (1979). Nucleotide Sequence Specificity of Restriction Endonucleases. *Science* **205**: 455-462.

Smith, H. O. & Kelly, S. V. (1984). Methylases of the type II restriction-,odification systems, p. 39-71. *In* A. Razin, H. Cedar, and A. D. Rigges (eds.), *DNA Methylation, biochemistry and biological significance*, Springer-Verlag, New York.

Sozhamannan, S. & Dharmalingam, K. (1988). Molecular cloning and identification of protein products of *rglB* locus of *Escherichia coli*. *Gene* **74**: 51-52.

Stahl, M. M., Kobayashi, I., Stahl, F. W. & Huntington, S. K. (1983). Activation of Chi, a recombinator, by the action of an endonuclease at a distant site. *Proc. Natl. Acad. Sci. U. S. A.* **80**: 2310-2313.

Stahl, F. W., Thaler, D. S., Kolodkin, A., Rosenberg, S., & Sampson, E. (1986). Double-chain Breaks: Thinking about Them in Phage and Fungi, p. 69-73, *In* A. Klar & J. Strathern (eds.), *Mechanisms of Yeast Recombination*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Steitz, T. A., Ohlendorf, D. H., McKay, D. B., Anderson, W. F., & Matthews, B. W. (1982). Structural Similarity in the DNA-binding domains of catabolite gene activator and *cro* repressor protein. *Proc. Natl. Acad. Sci. U. S. A.* **79**: 3097-3100.

Strathern, J. N., Klar, A. J. S., Hicks, J. B., Abraham, J. A., Ivy, J. M., Nasmyth, K. A. & McGill, C. (1982). Homothallic Switching of Yeast Mating Type Cassettes is Initiated by a Double-Stranded Cut at the MAT Locus. *Cell* **31**: 183-192.

Suck, D., Lahm, A., & Oefner, C. (1988). Structure refined to 2 A of a DNA octanucleotide complex with DNaseI. *Nature* **332**: 464-468.

Sun, H., Treco, D., Schultes, N. P., & Szostak, J. W. (1989). Double-strand breaks at

an initiation site for meiotic gene conversion. *Nature* **338**: 87-90.

Szomolanyi, E., Kiss, A. & Venetianer, P. (1980). Cloning the modification methylase gene of *Bacillus sphaericus* R in *Escherichia coli*. *Gene* **10**: 219-225.

Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., & Stahl, F. W. (1983). The Double-Strand Break Repair Model for Recombination. *Cell* **33**: 25-35.

Takeda, Y., Ohlendorf, D. H., Anderson, W. F., & Matthews, B. W. (1982). DNA Binding Proteins. *Science* **221**: 1020-1025.

Tanaka, I., Appelt, K., Dijk, J., White, S. W., & Wilson, K. S. (1984). 3-A resolution structure of a protein with histone-like properties in prokaryotes. *Nature* **310**: 376-381.

Tasseron-de Jong, J., Aker, J., den Dulk, H., van de Putte, P., & Giphart-Gassler, M. (1988). Cytosine Methylation in the *Eco*RI site in the active and inactive herpes virus thymidine kinase promoter, in press.

Taylor, A. & Smith, G. R. (1980). Unwinding and Rewinding of DNA by the RecBC Enzyme. *Cell* **22**: 447-457.

Terry, B. J., Jack, W. E., Rubin, R. A., & Modrich, P. (1983). Thermodynamic Parameters Governing Interaction of *Eco*RI Endonuclease with Specific and Nonspecific DNA Sequences. *J. Biol. Chem.* **258**: 9820-9825.

Terry, B. J., Jack, W. E., & Modrich, P. (1985). Facilitated Diffusion during Catalysis by *Eco*RI Endonuclease. *J. Biol. Chem.* **260**: 13130-13137.

Terwilliger, T. C., Fulford, W. D., & Zabin, H. B. (1988). A genetic selection for temperature-sensitive variants of the gene V protein of bacteriophage f1. *Nucl. Acids. Res.* **16**: 9027-9039.

Tessman, E. S. & Peterson, P. (1985). Plaque Color Method for Rapid Isolation of Novel *recA* Mutants of *Escherichia coli* K-12: New Classes of Protease-Constitutive *recA* Mutants. *J. of Bacteriol.* **163**: 677-687.

Thaler, D. S., Stahl, M. M. & Stahl, F. W. (1987a). Tests of the Double-Strand-Break Repair Model for Red-Mediated Recombination of Phage λ and Plasmid λdv. *Genetics* **116**: 501-511.

Thaler, D. S., Stahl, M. M. & Stahl, F. W. (1987b). Double-chain-cut Sites are Recombination Hotspots in the Red Pathway of phage λ. *J. Mol. Biol.* **195**: 75-87.

Thaler, D. S., Stahl, M. M. & Stahl, F. W. (1987c). Evidence that the normal route of replication-allowed Red-mediated recombination involves double chain ends. *EMBO J.* **6**: 3171-3176.

Thaler, D. S. & Stahl, F. W. (1988). DNA Double-chain Breaks in Recombination of Phage λ and of Yeast. *Ann. Rev. Genet.* **22**: 169-197.

Thomas, M. & Davis, R. W. (1975). Studies on the Cleavage of Bacteriophage Lambda DNA with *Eco*RI Restriction Endonuclease. *J. Mol. Biol.* **91**: 315-328.

Thoms, B. & Wackernagel, W. (1982). UV-Induced Alleviation of λ Restriction in

Escherichia coli K-12: Kinetics of Induction and Specificity of this SOS Function. *Mol. Gen. Genet.* **186**: 111-117.

Tomizawa, J. I. & Ogawa, H. (1968). Breakage of Polynucleotide Strands by Disintegration of Radiophosphorus Atoms in DNA Molecules and their Repair II. *J. Mol. Biol.* **30**: 7-15.

Tomizawa, J. I. & Ogawa, H. (1969). Breakage of DNA in rec^+ and rec^- bacteria by disintegration of radiophosphorus atoms in DNA and possible cause of pleiotropic effects of *recA* mutation. *Cold Spring Harbor Symp. Quant. Biol.* **33**: 243-251.

Tye, B-K., Nyman, P-O., Lehman, I. R., Hochhauser, S., & Weiss, B. (1977). Transient accumulation of Okazaki fragments as a result of uracil incorporation into nascent DNA. *Proc. Natl. Acad. Sci. U. S. A.* **74**: 154-157.

Volkert, M. V. & Nguyen, D. C. (1984). Induction of specific *E. coli* genes by sublethal treatments with alkylating agents. *Proc. Natl. Acad. Sci. U. S. A.* **81**: 4110-4114.

Walder, R. Y., Hartley, J. L., Donelson, J. E., & Walder, J. A. (1981). Cloning and Expression of the *Pst*I restriction-modification system in *Escherichia coli. Proc. Natl. Acad. Sci. U. S. A.* **78**: 1503-1507.

Walder, R. Y., Langtimm, C. J., Chatterjee, R., & Walder, J. A. (1983). Cloning of the *Msp*I Modification Enzyme. *J. of Biol. Chem.* **258**: 1235-1241.

Walder, R. Y., Walder, J. A., & Donelson, J. E. (1984). The Organization and Complete Nucleotide Sequence of the *Pst*I Restriction-Modification System. *J. of Biol. Chem.* **259**: 8015-8026.

Walker, G. C. (1984). Mutagenesis and Inducible Responses to Deoxyribonucleic Acid Damage in *Escherichia coli. Microbiol. Rev.* **48**: 60-93.

Walker, G. C. (1985). Inducible DNA Repair Systems. *Ann. Rev. Biochem.* **54**: 425-457.

Wang, T-C. V. & Smith, K. C. (1986a). Inviability of *dam recA* and *dam recB* cells of *Escherichia coli* is correlated with their Inability to Repair DNA Double-Strand Breaks Produced by Mismatch Repair. *J. of Bacteriol.* **165**: 1023-1025.

Warner, H. R., Duncan, B. K., Garret, C., & Neuhard, J. (1981). Synthesis and Metabolism of Uracil-Containing Deoxyribonucleic Acid in *Escherichia coli. J. Bacteriol* **145**: 687-695.

Wang, T. V. & Smith, K. C. (1986b). Postreplicational formation and repair of DNA double-strand breaks in UV-irradiated *Escherichia coli uvrB* Cells. *Mutat. Res.* **165**: 39-44.

Watson, J. D., & Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxynucleic Acids. *Nature* **171**: 737-738.

Weibezahn, K. F. & Coquerelle, T. (1981). Radiation induced DNA Double strand breaks are repaired by ligation and recombination processes. *Nucl. Acids Res.* **9**: 3139-3150.

Weiffenbach, & Haber, J. (1981). Homothallic mating type switching generates lethal chromosome breaks in *rad52* strains of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **1**: 522-534.

Wertman, K. F. & Mount, D. W. (1985). Nucleotide Sequence Binding Specificity of the LexA Repressor of *Escherichia coli* K-12. *J. of Bacteriol.* **163**: 376-384.

West, S. C. & Howard-Flanders, P. (1984). Duplex-Duplex Interactions Catalyzed by RecA Protein Allow Strand Exchanges to Pass Double-Strand Breaks in DNA. *Cell* **37**: 683-691.

Wharton, R. P. & Ptashne, M. (1987). A new-specificity mutant of 434 repressor that defines an amino acid-base pair contact. *Nature* **320**: 888-891.

Wilson, G. G. (1988). Cloned restriction-modification systems - a review. *Gene* **74**: 281-289.

Winans, S. C., Elledge, S. J., Krueger, J. H., & Walker, G. C. (1985). Site-Directed Insertion and Deletion mutagenesis with Cloned Fragments in *Escherichia coli*. *J. of Bacteriol.* **161**: 1219-1221.

Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980). Crystal Structure analysis of a complete turn of B-DNA. *Nature* **287**: 755-758.

Wolfes, H., Alves, J., Fleiss, A., Geiger, R. & Pingoud, A. (1986). Site-directed mutagenesis experiments suggest that Glu 111, Glu 114, and Arg 145 are essential for endonucleolytic activity of *Eco*RI. *Nucl. Acids Res.* **14**: 9063-9080.

Woodbury Jr., C. P., Downey, R. L., von Hippel, P. H. (1980a). DNA Site Recognition and Overmethylation by the *Eco*RI Methylase. *J. Biol. Chem.* **255**: 11526-11533.

Woodbury Jr., C. P., Hagenbuchle, O., & von Hippel, P. H. (1980b). DNA Site Recognition and Reduced Specificity of the *Eco*RI Endonuclease. *J. Biol. Chem.* **255**: 11534-11546.

Worcel, A. & Burgi, E. (1972). On the Structure of the Folded Chromosome of *Escherichia coli*. *J. Mol. Biol.* **71**, 127-147.

Yanisch-Perron, C., Vieira, J., & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.

Yanofsky, S. D., Love, R., McClarin, J. A., Rosenberg, J. M., Boyer, H. W., & Greene, P. J. (1987). Clustering of Null Mutations in the *Eco*RI Endonuclease. *Proteins: Struc., Func., and Genetics* **2**: 273-282.

Zeig, J., Maples, V. F., & Kushner, S. R. (1978). Recombination levels of *Escherichia coli* K-12 Mutants Deficient in Various Replication, Recombination, or Repair Genes. *J. of Bacteriol.* **134**: 958-966.

Zinder, N. D. & Boeke, J. D. (1982). The filamentous phage (Ff) as vectors for recombinant DNA-a review. *Gene* **19**: 1-10.

End