The Development of a Phenotypic Assay for Monitoring the Mutagenic Effects of 5-methylcytosine Deamination in *Escherichia coli*

Louisa Petropoulos

A Thesis

in

The Department

of

Biology

Presented in Partial Fulfilment of the Requirements
for the Degree of Master of Science at
Concordia University
Montreal, Quebec, Canada

January 1994

ISBN 0-315-97680-2

Canada

ABSTRACT

The Development of a Phenotypic Assay for Monitoring the Mutagenic Effects of 5-Methylcytosine Deamination in *Escherichia coli*

Louisa Petropoulos

A phenotypic assay was developed which monitors C to T transition mutations at the second C within CCAGG sequences in *Escherichia coli*. The second C of CCAGG sequences is methylated at the 5 carbon site. Deamination at methylated cytosines occurs at a frequency three times higher than unmethylated cytosines (JONES *et al.*, 1992). The cell system has evolved a specific repair mechanism, very short patch (VSP) repair, which can counteract such mutations.

An assay was developed from the construction of a mutant *lacZ* gene which can assess the function of VSP repair in *E. coli*. The substitution of serine (TCA) by proline (CCA) at site 462, within the *lacZ* gene, resulted in the introduction of the CCAGG sequence, the recognition site for the *dcm* methylase. The introduction of this sequence resulted in a Lac^- phenotype which reverts to Lac^+ via a G·T or C·A mismatch (G-C to A-T transition). The mismatch may result from deamination of the second (methylated) cytosine within the sequence CCAGG. VSP repair normally corrects the G·T mismatch back to a C-G base pair. The assay system was used to evaluate the function of the *dcm* and *vsr* genes

cloned from CC401, a potential VSP repair mutant. Secondly, it was found that deamination of 5-methylcytosines may not be due to a spontaneous event. Finally, a new mutator phenotype was isolated which may be due to an alteration in cytosine methylase activity.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude and thanks to an excellent supervisor, Dr. C. Cupples who has always been there to offer guidance and support. It has not only been an invaluable learning experience but a pleasure to work for Dr. Cupples.

I would like to thank my committee, Dr. M. Herrington and Dr. P. Gulick for helpful advice and their time for evaluating the thesis.

I would also like to thank all my co-workers on the 12^{th} floor who make it a pleasure to come in to work every-day.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

A tale has it that Caesar, intending to grant amnesty to one of his army officers, issued the order "Execute not, liberate". The message passed was "Execute, not liberate" and the man lost his life (RADMAN and WAGNER, 1988).

As a comma altered the message Caesar intended, a single base can alter the DNA message with potentially lethal consequences.

DNA is vulnerable to damage. This damage occurs continuously due to environmental factors such as chemicals and UV light. Damage also occurs during DNA recombination and replication where it is estimated that DNA polymerase I inserts one incorrect base for every 10^4 to 10^5 bases replicated (TOPAL and FRESCO, 1976). Unrepaired, the damage results in mutation.

Fortunately, mutations are rare, one error per 10^9 replicated nucleotides, due to the DNA's maintenance ability which involves recognizing the damaged portion and repairing it (DRAKE, 1969; SCHAAPER and RADMAN, 1989) . However, what is generally not realized is that different proteins work together in repairing specific DNA damage.

The maintenance of DNA fidelity in *Escherichia coli* is attributed to highly specialized enzymes which repair specific DNA mutations. Nucloeotide mismatches are repaired by

enzymatic systems which are classified into two groups. The first, is the methyl-directed repair system, also referred to as the "long patch repair system" and "dam-directed repair". The second, is the methyl-independent repair system which includes the VSP repair system and the mutY directed repair system.

1) METHYL-DIRECTED REPAIR

The methyl-directed repair pathway functions in the correction of mutations generated through errors of DNA replication and recombination (WAGNER and MESSELSON, 1976; GRILLEY *et al.*, 1990). A variety of mispairs are processed in a strand-specific manner involving the excision and repair of DNA tracts of a kilobase or more (GRILLEY *et al.*, 1990) hence the name "long patch repair". The *dam* gene codes for an adenine methylase responsible for methylating the A within the d(GATC) sequences. Following DNA replication in *E. coli*, the double stranded DNA is temporarily found in a hemimethylated state. The lag time in which only the adenine bases of the d(GATC) sequences found within the parental strand are methylated, directs corrections to the newly synthesized strand (PUKKILA *et al.*, 1983; WAGNER *et al.*, 1984).

Loss of methylation at the d(GATC) site results in the inability to discriminate the template from the daughter strand. As a result, 50% of the time, the daughter strand

has been favored as the template which results in cleavage of the correct base thus introducing a mutation. When both strands are methylated, very little repair occurs.

The genes *mut*H, *mut*L, *mut*S and *mut*U are required for mutation recognition and initiation of repair (LU *et al.*, 1984; LAHUE, 1989). The *mut*S gene product is involved in mismatch recognition and binding. The *mut*H gene product generates an incision 5' of the G of the d(GATC) sequence on the unmethylated strand (MODRICH, 1987). The mutL protein has no detectable activity. It is postulated that the mutL protein interacts with the mutS protein/mismatch complex in an ATP dependent manner (AU *et al.*, 1989; SANCAR and HEARST, 1993). That is, the mutL protein would bind to the mutS protein/mismatch complex and activate the latent ATPase of the mutS protein which would enable the mutH protein to incise the unmethylated strand of the GATC sequence (SANCAR and HEARST, 1993). The product of the *mut*U gene, helicase II, unwinds the DNA with the aid of the *ssb* gene product. The *ssb* gene product is involved in stabilizing the unwound DNA. The *sbc*B gene product, exonuclease I, hydrolyzes the DNA from the nick created 5' of the G by the mutH protein up to and including the DNA mismatch in a 3' to 5' direction. Finally, the gap within the DNA is resynthesized by the DNA polymerase III holoenzyme.

2) METHYL-INDEPENDENT REPAIR

The methyl-independent repair systems are characterized by the ability to recognize the mismatched base irrespective of d(GATC) strand methylation. Furthermore, the repair system corrects a particular type of mismatch in a stand-specific reaction that involves excision-repair tracts of only a few nucleotides (GRILLEY *et al.*, 1990). In *E. coli*, two such repair systems are known: *mut*Y and very short patch (VSP) repair.

3) METHYL-INDEPENDENT POST REPLICATIVE REPAIR

Similar to the methyl-directed repair system, mutY-mediated repair is involved in post DNA replication repair. However, the repair system is a *mut*H, *mut*L, *mut*S and *mut*U independent system. The *mut*Y gene product is an adenine glycosylase that hydrolyzes the glycosyl bond linking the mispaired adenine to deoxyribose of the G·A mispair (LU and CHANG, 1988; AU *et al.*, 1989). Hydrolysis of the adenine generates a site sensitive to apurine/apyrimidine (AP) endonuclease action (AU *et al.*, 1989). The AP endonuclease II hydrolyzes the phosphodiester 5' to an abasic site to generate 3'-hydroxyl and 5'-phosphate termini (KANE and LINN, 1981). Following DNA strand cleavage by AP endonuclease, DNA polymerase I introduces the correct nucleotide base, cytosine. Finally, DNA ligase is involved in catalyz-

ing the terminal step to restore the phosphodiester backbone.

4) METHYL-INDEPENDENT NON-REPLICATIVE REPAIR

i) VSP REPAIR IN *ESCHERICHIA COLI*

VSP repair is a mutH independent system specialized in the repair of G·T mismatches which occur in quiescent cells. The G·T mismatch is thought to be the result of spontaneous deamination of a 5-methylcytosine which results in its conversion into another naturally occurring nucleotide, thymine.

It was long believed that mutations arising during the quiescent state of the cell were unrepairable until recombination studies performed by M. Lieb (1983), found that the mutation *amb*6 in the *cI* (repressor) gene of bacteriophage lambda recombined with other *cI* mutations much more frequently than predicted by the physical map distances. Furthermore, there was an absence of co-repair of *amb*6 and *cI* mutations when *amb*6 was crossed with adjacent *cI* mutations. The lack of co-repair indicated that the repair tracts did not extend farther than about 20 base pairs from either side of the *amb*6 mismatch. The *amb*6 mutation contains the sequence CTAGG at bp 98-102 (SAUER, 1978; LIEB, 1983) where the wild type sequence is CCAGG. Upon recombination, the T·G mismatch was produced where a very short patch was excised that included the T and corrected the G·T to G-C.

In *E. coli*, only cytosines located within the second position within the 5'-CC(A/T)GG-3' sequence are methylated at the 5 carbon site. Methylated cytosines have been shown to exhibit a high mutation frequency (COULONDRE *et al.*, 1978). Methylation of cytosines within CCAGG sequences was found to be due to the *dcm* gene product, cytosine methylase, and repair of G·T mismatches due to 5-methylcytosine deamination by the *vsr* gene product, the vsr endonuclease (BHAGWAT *et al.*, 1988b; SOHAIL *et al.*, 1990). The *dcm* and *vsr* genes exist in an operon located at 43 minutes within the *E. coli* chromosome. The *dcm* gene is 2472 base pairs long where 19 base pairs at the 3' end are also part of the 1419 base pair coding region of *vsr*. Furthermore, the *vsr* gene is in a reading frame different from that of the *dcm* gene (SOHAIL *et al.*, 1990).

The vsr endonuclease has been found to be capable of performing two functions: 1) recognizing G·T mismatches only within the CC(A/T)GG sequence and 2) cleaving the T from G·T mismatches. *In vitro* studies have demonstrated that the vsr gene product is a strand and sequence specific DNA mismatch endonuclease (HENNECKE *et al.*, 1991).

Purified vsr endonuclease was found to be all that is required for G·T mismatch repair. *In vitro* the G·T mismatch was introduced within the cytosine methylase target sequence CTAGG/GGTCC at the center of a DNA heteroduplex. The purified vsr endonuclease was incubated with heteroduplex

producing an incision 5' of the mismatched T (HENNECKE *et al.*,1991). Following DNA strand cleavage, exonuclease I removes the T and DNA polymerase I introduces the cytosine base. Finally, the phosphodiester backbone is repaired by DNA ligase.

Changing the recognition sequence, such as CTAGG to CTGAG was found to abolish endonucleolytic cleavage of the mispaired T by the vsr endonuclease. Also, mismatches other than G·T were not repaired (HENNECKE *et al.*, 1991). The presence of a 5-methyl group on the inner cytosine residue of the substrate sequence is not required; the vsr endonuclease introduces a nick on the strand containing the mismatched T residue (HENNECKE *et al.*, 1991).

ii) VSP REPAIR IN HUMAN CELLS

In human systems, methylation of cytosines occurs at CpG sequences. The mutagenic effects of spontaneous deamination of 5-methylcytosines are counteracted by a mechanism similar to that of mutY-directed repair in *E. coli*. A DNA glycosylase is responsible for cleaving the glycosidic bond of the thymine at G·T mismatches (JIRICNY *et al.*, 1988). Repair of G·T mismatches occurs irrespective of site methylation or flanking sequences (BROWN and JIRICNY, 1987). Hence, in contrast to VSP repair in E. coli, human VSP repair corrects any G·T mismatch specifically back to G-C. The excision of the mispaired thymine generates a single nucleotide gap. The apyrimidinic (AP) site opposite guanine

is cleaved 3' from the baseless site by AP endonuclease I whereby DNA polymerase I then introduces cytosine.

5) COOPERATIVE REPAIR OF MISMATCHES

The maintenance of DNA fidelity involves the cooperative work of various proteins at each stage of DNA replication. The repair enzymes are involved in 1) the selection of the correct base by the DNA polymerase 2) the removal of a noncomplementary bases 3) the correction of misincorporated bases by the methyl-directed postreplicative repair and methyl-independent repair systems (AKIYAMA *et al.*, 1989).

Organisms that use oxygen need to defend themselves against oxydative damage of nucleic acids and other cellular molecules (MICHAELS *et al.*, 1992). One lesion caused by oxygen is the conversion of guanine to 7,8-dihydro-8-oxyguanine (GO). The incorporation of GO during DNA synthesis can result in the misincorporation of adenine opposite the damaged guanine hence, lead to a G-C to T-A transversion (MICHAELS *et al.*, 1992; MICHAELS *et al.*, 1991; CABRERA *et al.*, 1988).

The first mutation prevention measure is the elimination of GO bases from the nucleotide pool by the product of the *mut*T gene. The mutT protein hydrolyzes 8-oxo-dGTP to 8-oxo-dGMP (MICHAELS *et al.*, 1992).

Should mutT fail to eliminate GO from the nucleotide pool resulting in incorporation into the DNA, an adenine or

a cytosine may be paired with the GO base. The resulting A/GO or C/GO mispairs are not efficiently removed by the proofreading system of polymerase III. A C/GO mispair is corrected by the *mut*M gene product. The mutM glycosylase removes the GO base and subsequent repair restores the original C/G base pair.

If an A/GO base pair is formed, the *mut*Y gene product, adenine glycosylase removes the A base where a C is then paired with the GO base by DNA polymerase I. Hence, the C/GO lesion can then be acted on by the mutM-mediated repair pathway.

6) THE ROLE OF CYTOSINE METHYLATION

In the mammalian system, cytosine methylation at CpG regions has been correlated with gene control. Experimental evidence suggests that the loss of methylated cytosines from promoter regions leads to the activation of genes (CEDAR, 1988). Methylation of promoter regions has been shown to inhibit transcription of cloned genes, both in vivo and in vitro (BEN-HATTAR and JIRICNY, 1988; WATT and MOLLOY, 1988; IGUCH-ARIGA and SCHAFFNER, 1989). Prevention of gene methylation during DNA replication by 5-azacytidine, an inhibitor of cytosine methylation, has resulted in gene copies which are actively transcribed in daughter cells (JONES *et al.*, 1983).

In *E. coli*, adenine methylase deficient cells which do not methylate adenine sites within GATC sequences exhibit

increased spontaneous mutagenesis and sensitivity to mutagens and UV light (MARINUS and MORRIS, 1975). However, *E. coli* cells deficient in cytosine methylation which do not methylate the second C within CCAGG sequences, demonstrate no ill-effects.

The increased mutability of 5-methylcytosines relative to cytosine is thought to be due to the fact that deamination of a 5-methylcytosine leads to a thymine which is a natural component of DNA. Deamination of cytosine leads to uracil and removed by uracil glycosylase and deamination of adenine to hypoxanthine is readily removed by hypoxanthine glycosylase.

The rate of 5-methylcytosine deamination is 3.5 fold higher than that of cytosine. The half life of cytosine is 30 000 years where that of 5-methylcytosine is about 10 000 years. Hence, in 6×10^7 5-methylcytosine residues per diploid human genome, there are approximately 8 deamination events per day per diploid genome (JONES *et al.*, 1992). Although it is not clear why 5-methylcytosines have an increased mutability, their deamination in the human germline has been suggested to account for 30-40% of all point mutations.

Cancer is a genetic disease in which cells have sustained mutations in genes responsible for control of cellular division (JONES *et al.*, 1992). One human gene, p53, is a tumor suppressor gene that has been found to be altered in

approximately half of all human tumors at CpG sites (JONES *et al.*, 1992).

7) THE LACZ ASSAY SYSTEM

An assay system developed by Cupples and co-workers, was used for selecting strains with a high frequency of C to T transitions at 5-methylcytosines. The assay involves the use of the *lacZ* gene which codes for the enzyme ß-galactosidase. The enzyme is responsible for cleaving lactose into glucose and galactose. Studies have indicated that ß-galactosidase is tolerant of missense mutations (unpublished data cited in CUPPLES and MILLER, 1988). However, specific alterations made within the enzymatically essential residues, such as Gly-461 or Tyr-503, have been shown to result in an inactive enzyme (HERRCHEN and LEGLER, 1984; SINNOT and SMITH, 1978; FOWLER *et al.*, 1978; CUPPLES and MILLER, 1988).

The requirement of Gly-461 for enzymatic activity of ß-galactosidase, allowed for the introduction of the CCAGG sequence into the *lacZ* gene in such a way as to provide a phenotypic marker of VSP repair. The alteration of the wild type codons 460 to 462 from AAT GAG TCA (Asn Glu Ser) to AAC CAG GGG (Asn Gln Gly) resulted in a Lac^- phenotype due to the Gly to Gln substitution at site 461 (RUIZ *et al.*, 1993). A deamination event results in the alteration of the codon at site 461 from CAG to TAG (Figure 1). The TAG is a termi-

nation codon where if left unrepaired, as is the case in VSP repair deficient cells, premature termination of ß-galactosidase synthesis results in a Lac^- cell. However, whether the cell has a functional VSP repair system or not, a Lac^- phenotype is the result. To differentiate between CAG and TAG, a suppressor plasmid was introduced. The suppressor plasmid carries a gene for a tRNA which recognizes the TAG codon and inserts the wild type glutamic acid at position 461. Cells with a functional VSP repair system would recognize and correct the T·G mismatch thus, maintaining the Lac^- phenotype. However, cells deficient in VSP repair allow for the T·G mismatch to remain. The resultant TAG codon is then recognized by the suppressor tRNA where the wild type glutamic acid is inserted; this reinstates the enzymatic activity of ß-galactosidase, thus the Lac^+ phenotype of the cell. Therefore, cells with an active VSP repair system have a low Lac^+ reversion frequency and cells with a defective VSP repair system have a high Lac^+ reversion frequency (Figure 1). The assay was used to identify a potential new mutator strain, CC401 with a high frequency of C-G to T-A mutations at the second C of CCAGG sequences (RUIZ *et al.*, 1993). However, the above assay has a few drawbacks. One drawback is that the suppressor is on a high copy number plasmid which seems to affect cell viability (RUIZ *et al.*, 1993) Another drawback is that the presence of the suppressor plasmid makes it difficult to introduce a second plasmid for

the purpose of studying the effects of other genes (RUIZ *et al.*, 1993).

This study describes the development of the modified assay system which eliminates the need for a suppressor. The assay was then used to 1) assess whether the *dcm* and *vsr* genes of the CC401 mutator are functional; 2) isolate a new VSP repair mutator with altered cytosine methylase activity.

MATERIALS AND METHODS

The methods are divided into three sections. The first outlines the procedure used in the development of the pro462 assay. The second section involves testing the VSP assay system. The final section describes the procedures used in the isolation of a *dcm* mutant using the VSP assay.

PART I. THE DEVELOPMENT OF THE VSR ASSAY

A mutant *lacZ* gene containing a proline (CCA) in place of the wild type serine (TCA) at site 462 was constructed. The CCA codon was introduced by site-directed mutagenesis using the single stranded f1-Z phage as a template. The mutant *lacZ* gene was then subcloned into plasmid pBR329. The mutant *lacZ* gene was then introduced into an F'episome via recombination with pBR329. The ß-galactosidase levels produced by the mutant *lacZ* gene in the episome were measured; the mutant *lacZ* gene results in a Lac^- phenotype.

All bacterial strains, bacteriophage and plasmids used in this study are described below.

TABLE 1:

BACTERIAL STRAINS

STRAIN	GENOTYPE
CC401[1]	*ara* Δ(*lac pro*)*gyrA argE-am rpoB thi*; F$^+$: *proA,B*; *lacZ*$^-$, *lacY*, A; pGFIB (GluA derivative)
CC112[1]	*ara* Δ(*lac pro*)*gyrA argE-am rpoB thi*; F+: *proA,B*; *lacZ*-, *lacY*, A; pGFIB (GluA)
CC215[2]	Δ(*gpt-lac*)5 derivative of GM31
CC216[3]	Δ(*gpt-lac*)5 deivative of RP4182
GM31[4]	*dcm*-6 *thr*-1 *ara*-14 *leuB6 tonA31 tsx*-78 *glnv44 galT22 hisG4 rpsL136 xyk*-5 *mtl*-1 *thi*-1
JM105[5]	*supE endA sbcB15 hsdR4 rpsL thi* Δ(*gpt-lac*)5
JM801[6]	*supE endA sbcB15 hsdR4 rpsL thi* Δ(*gpt-lac*)5
P90C[7]	*ara* Δ(*gpt-lac*)5
RP4182[4]	(*flaD-flaP*) DE4 *trp gal rpsL* deleted for *dcm* and *supD*
S90C[6]	*ara* Δ(*gpt-lac*)5 *thi rpsL*
DS 22-28[6]	*ara* Δ(*gpt-lac*)5 *thi rpsL*; F': *proA, proB*

[1]RUIZ *et al.*, 1993.

[2]PETROPOULOS *et al.*, 1993.

[3]This study, unpublished.

[4]MARINUS, 1973.

[5]CUPPLES and MILLER, 1988.

[6]CUPPLES and MILLER, 1989.

[7]SOHAIL *et al.*, 1990.

TABLE 2:

LIST OF PRIMERS

PRIMER	NUCLEOTIDE SEQUENCE
PRO462	5'-TAGCGCCGTGGCCTGGTTCA-3'
LEU462	5'-TAGCGCCGTGGCCTGATTCA-3'
VSR:1806	5'-ACAAATCCGGACGTCCG-3'
VSR:1951	5'-CAGTCGCTTGCAGGAACT-3'
VSR:2145	5'-TAGTTGCGCCAGTTATTCAGGACGCATCAA-3'
VSR:2262	5'-GCGGATCCGGACTGGGTGGAGAAACAC-3'
Z-COOH-6	5'-GTCTGGTGTCAAAAATAATAAGAATTCTGCAGT-3'

TABLE 3:

LIST OF PLASMIDS

PLASMID	SIGNIFICANT GENES AND MARKERS
pBR322	amp^r, tet^r
pBR329	amp^r, tet^r, cm^r
pACYC184	tet^r, cm^r
pDCM21	*dcm* gene cloned into tet gene of pBR322
pDVW	*dcm* and *vsr* genes from CC112 cloned into *Bam*HI/*Sph*I site of pBR322; amp^r, tet^s, cm^r
pDVM	*dcm* and *vsr* genes from CC401 cloned into *Bam*HI/*Sph*I site of pBR322; amp^r, tet^s, cm^r
pDW	Same as pDVW but *vsr* gene deleted by cleavage with *Aat*II
pDM	Same as pDVM but *vsr* gene deleted by cleavage with *Aat*II
$pDVW_M$	*dcm* and *vsr* genes from CC112 cloned into *Bam*HI/*Sph*I site of pACYC184; cm^r, tet^s; chemically mutagenized with NH_2OH
$pDVW_W$	*dcm* and *vsr* genes from CC112 cloned into *Bam*HI/*Sph*I site of pACYC184; cm^r, tet^s
pRP215	EcoRII methylase gene cloned into *cm* coding gene of pACYC184; tet^r, cm^s

cm= chloramphenicol
amp= ampicillin
tet= tetracyline

PHAGE

F1-Z is a filamentous phage which contains the complete *lacZ* gene. The construction of a complete *lacZ* gene was made possible by removing the 3' internal *Eco*R1 site of the *lacZ* gene and constructing two new *Eco*R1 sites on either side of the *lacZ* gene (CUPPLES and MILLER; 1988). The 3.4kb *lacZ* gene was then cloned into R229, a derivative of f1 phage. The advantage of using the f1 phage is that the genome is a single stranded DNA which can be used as a template for site-directed mutagenesis and sequencing. F1-Z was maintained in the JM801 strain.

PLASMID

The vector used to carry the mutated *lacZ* gene was pBR329. The 4.2 kb plasmid contains three antibiotic resistance coding genes (chloramphenicol, ampicillin, and tetracycline). The mutated *lacZ* gene was cloned into the unique *Eco*R1 site located within the chloramphenicol resistance coding gene. The plasmid with the mutated *lacZ* was maintained in S90C.

EPISOME

The 50 kb F' episome is *proA, proB, lacZ, lacY, lacA* and *lacI-*. The episome was maintained within the P90C strain.

SITE-DIRECTED MUTAGENESIS

Site-directed mutagenesis was done by a variation of the technique of ZOLLER and SMITH (1982) using two oligonucleotides as primers. The oligonucleotides were synthesized with the use of the Applied Biosystems 391 PCR-Mate DNA synthesizer. One oligonucleotide formed a heteroduplex with the *lacZ* gene in the f1-Z phage, due to a base mismatch. The second oligonucleotide, which served as a helper in DNA synthesis, was complementary to a region 3' of the *lacZ* gene. The oligonucleotides served as primers for DNA polymerase in synthesizing the complementary f1-Z strand in vitro and introducing the missense mutation.

i) PHOSPHORYLATION OF THE OLIGONUCLEOTIDES

Each oligonucleotide (2 μg) was added to 4 μl of kinase buffer (350 mM Tris pH 7.4, 50 mM $MgCl_2$ and 25 mM dithiothreitol) where it was phosphorylated by 0.5 μl of polynucleotide kinase. The mixture was allowed to incubate for one hour at 37°C. Termination of the reaction involved heating the mixture to 65°C for 10 minutes where the polynucleotide kinase was inactivated.

ii) MUTAGENESIS

The phosphorylated oligonucleotides (pro462 and Z-COOH-6; Table 2) were annealed to the single stranded f1-Z phage DNA where they served as primers for DNA polymerase to

synthesize a complementary strand *in vitro*. Hence, 0.1 μg of helper oligonucleotide and 0.5 μg of f1-Z template were added to 1.0 μl of annealing buffer (200 mM Tris pH 7.4, 100 mM $MgCl_2$ and 500 mM NaCl). The mixture was then heated to a temperature between 90°C and 100°C for 5 minutes. This was followed by a 45 minute cooling period to room temperature where the oligonucleotides were allowed to anneal to the f1-Z template. To the mixture was added, 40 μl of extension buffer (20 mM Tris pH 7.4, 10 mM $MgCl_2$, 20 mM dithiothreitol and 1 mM ATP), 0.5 μl of each 10 mM dNTP (dATP, dGTP, dCTP, dTTP), 1 unit of ligase, 2.5 units of Klenow fragments of DNA polymerase. The synthesis of the complementary strand was allowed to proceed at room temperature for 1 hour.

COMPETENT CELLS (MANIATIS *et al.*, 1982)

A cell that is able to take up DNA is said to be competent. Hence, to encourage competence, S90C cells were grown in 35 ml of LB to log phase (OD between 0.2 and 0.25). The cells were then cooled on ice for 10 minutes and centrifuged for 5 minutes at 4,000xg at 4°C. After aspiration of the supernatant, the pellet was resuspended with 15 ml of cold 50 mM $CaCl_2$. The cell suspension was kept on ice for 30 minutes followed by centrifugation and resuspension in 3 ml of cold 50 mM $CaCl_2$.

CELL TRANSFORMATION (MANIATIS *et al.*, 1982)

Competent JM801 cells were transformed with the hybrid replicative form of f1-Z. The hybrid f1-Z is composed of one strand with a wild type *lacZ* gene and the complementary DNA strand containing the mutant *lacZ* gene. Hence, 0.1 μg of DNA was added to 300 μl competent JM801 cells and incubated on ice for 30 minutes. The mixture was then heat shocked for 1.5 minutes at 42°C. The cells were then cooled briefly, followed by the addition of 200 μl of log phase JM801 cells, 50 μl of 2% X-gal and 3 ml of top agar. The mixture was plated on LB plates.

ISOLATION OF MUTAGENIZED *LACZ* DNA FROM PLAQUES

i) SINGLE STRANDED F1-Z DNA

A single plaque was picked, put into 100 μl of the log phase JM801 bacterial strain and incubated at 37°C for 10 minutes. Following the incubation, 2 ml of LB was added and the culture was left to grow for 4.5 hours to overnight with aeration at 37°C. From the culture, 1.5 ml was centrifuged at 16,000xg for 1 minute at 4°C. The supernatant was transferred to a new tube where 200 μl of PEG/NaCl (20% polyethylene glycol 600, 2.5 M NaCl) was added, vortexed and left at room temperature for 15 minutes. Another round of centrifugation (16,000xg) at room temperature for 5 minutes

produced a small pellet which was suspended in 100 μl of TE (10 mM Tris. pH 7.4 and 1 mM EDTA). To remove phage protein coat, 100 μl of phenol was added. The mixture was vortexed and centrifuged (16,000xg for 3 minutes at room temperature) and the aqueous layer saved. To precipitate the DNA out of the supernatant, 10 μl of 3.0 M sodium acetate and 200 μl of cold 95% ethanol was added then stored overnight at -20°C. The pellet obtained from centrifugation (16,000xg for 15 minutes at 4°C) was dried and resuspended in 25 μl of distilled water.

ii) REPLICATIVE FORM F1-Z DNA (PLASMID)

A single f1-Z phage plaque was isolated and incubated with 100 μl of log phase JM801 cells at 37°C for 10 minutes without aeration. Following incubation, 2 ml of LB was added and the cell culture was grown for 4.5 hours to overnight with aeration. The culture was centrifuged (16,000xg at 4°C) for 1 minute; the pellet was resuspended in 100 μl of Solution I (50 mM glucose, 25 mM Tris (pH 8.0) and 10 mM EDTA). The mixture was left at room temperature for 5 minutes. Following, was the addition of 200 μl of Solution II (0.2 M NaOH and 1% sodium dodecylsulfate), and a gentle mixing; the suspension was then placed on ice for 5 minutes. After the addition of 150 μl of Solution III (5 M

potassium acetate, pH 4.8) the suspension was placed on ice for a period of 5 minutes. The denatured proteins were allowed to precipitate out of solution. After centrifugation (16,000xg at 4°C for 1 minute) the plasmid containing supernatant was isolated and 400 µl of 95% ethanol added to the recovered supernatant to precipitate the plasmid. A final centrifugation (16,000xg at 4°C for 1 minute) concentrated the DNA into a pellet which was then resuspended in 50 µl of double distilled water.

SEQUENCING OF THE MUTANT F1-Z

DNA was isolated from single stranded f1 phage with the mutant *lacZ* gene from plaques as outlined above. The DNA was sequenced with the T7 DNA polymerase kit and [^{35}S]dATP with the Sanger dideoxy sequencing technique following instructions in the kit (Pharmacia). The DNA sample was electrophoresed on an 8% acrylamide sequencing gel with a voltage of 1025 V.

SUBCLONING OF THE LACZ GENE INTO pBR329

Double stranded f1-Z phage DNA and pBR329 containing 0.08 units of ribonuclease, were digested with 1 unit of *Eco*R1 in to 2 µl of a buffer (10 mM Tris. pH 7.4, 100 mM NaCl, 7 mM $MgCl_2$ and 1 mM ß-mercaptoethanol) and incubated at 37°C for 1.5 hours. After digestion of the double stranded DNA from f1-Z phage and pBR329 with *Eco*R1, the

the enzyme was heat inactivated at 65°C for 30 minutes. Digested pBR329 and f1-Z were added at a ratio of 1:2 respectively to a buffer (30 mM Tris pH 7.8, 10 mM $MgCl_2$, 10 mM DDT, 0.5 mM ATP, 0.1 mg/ml BSA) and ligated using 3 units of T4 DNA ligase.

TRANSFORMATION OF P90C WITH THE LACZ-pBR329 LIGATE

For transformation of P90C cells, 10 μg of the mutant *lacZ*-pBR329 ligation mixture was mixed with 200 μl of competent P90C cells. These cells were plated on minimal glucose plates containing X-gal and ampicillin then incubated at 37°C for 36 hours. The P90C cells containing the mutant *lacZ* in pBR329 produced pale blue colonies.

TRANSFERRING THE MUTANT LACZ FROM PLASMID TO EPISOME

First, the aim was to transfer a $lacI^-$ $lacZ^+$ episome into the strain which contained the plasmid with the mutant *lacZ* gene. This was achieved by conjugating stationary phase P90C with log phase cells containing the episome in a ratio of 2:1 respectively and incubating them for 1 hour at 37°C without aeration. The progeny were selected on minimal glucose plates with tetracycline.

Cells which grow on minimal glucose plates with tetracycline contain both the plasmid and the episome. To allow for recombination to occur between the plasmid and the episome, the cells were grown overnight in 2 ml of LB and 2 μl of tetracycline (15 mg/ml).

In order to separate the episome from the plasmid once the desired recombination had occurred, log phase cells containing the plasmid and the episome were conjugated with stationary phase S90C cells at a ratio of 1:2 respectively. The cells were incubated at 37°C without aeration for 1 hour to permit the transfer of the episome into S90C. The S90C cells containing the episome were selected on minimal glucose plates containing streptomycin and X-gal. Cells containing an episome with the mutant *lacZ* gene were pale blue.

MAPPING OF THE MUTATION IN *LACZ* ON THE EPISOME

To verify that the Lac$^-$ phenotype is not due to secondary mutations as a result of recombination, such as point mutations or deletions outside the pro462 region, the *lacZ* gene was mapped. Mapping of the *lacZ* gene was done using deletion strains 22 to 28. The deletion strains (DS) contained progressively larger deletions within the *lacZ* gene where DS 22 contained the shortest deletion which included codon 462 (up to codon 460) and DS 28, the longest (up to codon 503). As such, DS with deletions 5' to codon 463 would recombine with the mutant *lacZ* gene to reinstate the wild type *lacZ* gene hence, a Lac$^+$ phenotype. Furthermore, DS which have deletions in codon regions 462 to 460 would remain as a Lac$^-$ phenotype. The DS were conjugated with S90C

cells containing the mutant episome. The episome was transferred to deletion mapping strains by the plate mating technique (MILLER, 1972). The cells with the mutant *lacZ* episome were gridded onto LB plates. The procedure involved growing the deletion strains overnight in LB then plating them on minimal lactose plates. The LB plates were replica plated onto the minimal lactose plates and incubated at 37°C for 48 hours to allow for recombination between the episome and deletion strain episome to occur. Deletion strains which underwent recombination within the *lacZ* gene of the episome resulted in a Lac^+ pherotype.

ß-GALACTOSIDASE ASSAY IN WHOLE CELLS (MILLER, 1972)

The cells with episomes which contained the mutant *lacZ* gene were assayed for the amount of ß-galactosidase produced. The cells were grown overnight in minimal glucose medium at 37°C. The overnight culture was diluted to 10^{-1} in 1x MinA salt solution (60 mM K_2HPO_4, 33 mM KH_2PO_4, 7.6 mM $(NH_4)_2SO_4$, 1,7 mM $Na_3C_6H_5O_7 \cdot 2H_2O$) where 0.1 ml was added to 0.9 ml of Z buffer (0.06 M Na_2HPO_4 $7H_2O$, 0.04 M NaH_2PO_4, 0.01 M KCl, 0.001 M $MgSO_4$ $7H_2O$, 0.05 M ß-mercaptoethanol). This was followed by the addition of 2 drops of chloroform and 1 drop of 0.1% SDS. Once the mixture was vortexed for 10 seconds, it was transferred to a 28°C water bath. The

reaction started and the time was noted upon the addition of 0.2 ml of ONPG (0-nitrophenyl-ß-D-galactoside). Once a light yellow color was observed, the reaction was terminated with 0.5 ml of 1 M sodium carbonate. At this point, the final time was recorded. The absorbance of the reaction mixture was measured on a LKB Biochrom Ultraspec II spectrophotometer at OD_{420}. The absorbance of the diluted overnight culture was measured at OD_{600}.

The units of ß-galactosidase were calculated as follows:

$$\text{units of ß-gal.} = \frac{OD_{420} \times 10\ 000}{OD_{600} \times \text{TIME (MINUTES)}}$$

PART II. TESTING THE *LACZ* ASSAY (MILLER,1972)

The ability of the assay system to discriminate between functional VSP repair and non-functional VSP repair was assessed. This was done by introducing the episome containing the mutant *lac*Z gene into known wild type and mutant VSP repair backgrounds. The assay detects a mutation event as a Lac^- to Lac^+ reversion. Each Lac^+ revertant in the population forms a single colony on minimal lactose plates.

CLONING DCM/VSR IN pBR322

i) AMPLIFICATION OF DCM/VSR GENES USING PCR (RUIZ, 1992)

Genomic DNA from the CC112 strain was used as a template in the amplification of the dcm/vsr genes. The genes of interest were PCR amplified as follows:

7 μl of CC112 chromosomal DNA (266 ng/μl)

1 μl of primer VSR:2262 (386 ng/μl)

1 μl of primer DCM:-007 (368 ng/μl)

10 μl of 2 mM dNTP's

70 μl of sterile water

The reaction mixture was heated to 96°C for 3 minutes. Following heating, the mixture was cooled to 75°C for a period of 2 minutes during which 10 μl of 10x *Taq* buffer and 1 μl of *Taq* polymerase (Bio/Can) was added. The reaction was allowed to proceed for 35 cycles . Each cycle involved a 20 second denaturation period at 95°C. The denaturation period was followed by a 1 minute annealing period at 60°C. Finally, an extension reaction was allowed to proceed for 45 seconds at 72°C. Following PCR amplification, the 2.2 kb *dcm/vsr* gene fragment was purified using the conditions specified in the Gene Clean kit (Bio 101 Inc.). This was done to remove unincorporated dNTP and primer which may interfere with the cloning process.

ii) DIGESTION AND LIGATION OF VECTOR AND INSERT

Both pBR322 and the *dcm/vsr* fragment were digested with *Sph*I and *Bam*HI for 1.5 hours at 37°C. The enzymes were heat inactivated at 65°C for 30 minutes. The vector and insert were combined at a ratio of 1:3 and allowed to ligate for 6 hours at 16°C. The ligation mixture was then used to transform competent S90C cells which were then plated on LB with ampicillin.

iii) SCREENING FOR TRANSFORMANTS CONTAINING THE INSERT

Cells that grew on LB containing ampicillin were replica plated onto LB plates containing tetracycline. A successful ligation of the insert into the vector would disrupt the tetracycline resistance coding region thus cells will be ampicillin resistant but tetracycline sensitive. Thus, cells which exhibited tetracycline sensitivity were identified and the plasmid isolated. The plasmids were digested with *Eco*R1 and analyzed on a 1% agarose gel. The appearance of a 6.4 kb fragment indicated an insert was present. The 6.4 kb fragment was further analyzed by *Sph*I and *Bam*HI digestion and by agarose gel analysis to verify the presence of both the 4.2 kb pBR322 vector and the 2.2 kb *dcm/vsr* vector.

CONSTRUCTION OF STRAIN WITH FUNCTIONAL VSP REPAIR

CC215 is a Δ(*gpt-lac*)5 derivative of GM31 (*dcm*-6 *thr*-1 *ara*-14 *leu*B6 *ton*A31 *lac*Y1 *tsx*-78 *gln*V44 *gal*T22 *his*G4 *rps*L136 *xyk*-5 *mtl*-1 *thi*-1) (MARINUS, 1973; Table 1). The CC215 strain is *dcm*$^-$ and *vsr*$^-$, due to a termination codon within the *dcm* gene (*dcm*-6) which inactivates both *dcm* and *vsr*. By introducing pDVW (dcm and vsr from CC112; Table 3) into CC215, the cell is provided with the methylase and VSP repair system; effects on the Lac$^+$ reversion frequency were assayed with the use of the F' episome containing the CCA (proline) codon at site 462 of the *lac*Z gene.

CONSTRUCTION OF A STRAIN DEFICIENT IN VSP REPAIR

CC215 had the genes *dcm* (pDW and pDCM21; Table 3) and *vsr* (pDCM28) introduced independently into the strain and the rate of mutation monitored by the F'episome containing the mutant *lac*Z gene.

LACTOSE REVERSION FREQUENCY

The lactose reversion allows for a direct measure of the mutation frequency. The cell spontaneously reverts from a Lac$^-$ phenotype to a Lac$^+$ phenotype through specific mutational events which occur on the modified *lac*Z gene

within the F'episome. Each Lac+ reversion event is a mutation event and is exhibited as a single colony on a ML plate.

Cell strains containing the episome with the mutant *lacZ* were grown overnight in 2 ml of LB at 37°C with aeration. The overnight culture was diluted to 10^{-6} and 200 μl were plated on LB plates. Also, 200 μl of undiluted cells were plated on lactose plates. The colonies on each plate were counted and the reversion frequency determined as follows:

$$\#\ \text{of Lac}^+\ \text{revertants}/10^8\ \text{cells} = \frac{\#\ \text{of colonies on ML}}{\#\ \text{of colonies on LB}}$$

SEQUENCING

To determine what was responsible for the Lac^+ phenotype, a portion of the *lacZ* gene from cells that had reverted to a Lac^+ phenotype, was sequenced using the Sanger dideoxy method.

COLONY HYBRIDIZATION

Colony hybridization permitted the screening of a large number of colonies in a short period of time and thus replaced sequencing for determining the frequency with which a CCAGG to CTAGG change was responsible for a Lac^+ phenotype.

The transition results in the alteration of codon 462 within the *lacZ* gene from a proline (CCA) to leucine (TCA). Hence, a 14 base oligonucleotide (LEU462; Table 2) which contains the leucine codon was constructed. However, to prevent false positives, cellular debris was removed as described below.

1) COLONY LIFTS

Lac^+ revertant cells were gridded on an LB plate and grown overnight at 37°C. The LB plate with the gridded cells was allowed to cool for half an hour at 4°C before using. Upon cooling, 0.22 μm nylon membrane (60mm diameter, Millipore) was placed on the plate for a few seconds. The membrane was then lifted and placed colony side-up on 3 MM Whatman paper saturated with High Salt Denaturation Solution (0.5 M NaOH, 1.5 M NaCl) for 5 minutes. The membrane was blotted briefly and placed on 3 MM Whatman paper saturated with Neutralizing Solution (0.5 M HCl·Tris pH 7.4, 1.5 M NaCl) for 5 minutes. The membrane was then baked in a vacuum oven for 1 hour at 80°C.

2) PROTEASE TREATMENT OF COLONY LIFTS

The baked membrane was placed in 25 ml of 1xTBS for 10 minutes with gentle shaking. A 25 ml solution of 1xTBS with 10 mM $MgCl_2$ and 200 μg/ml of proteinase K was prepared and

preheated to 37°C. The membrane was incubated in the proteinase K solution for 1 hour at 37°C. Following incubation, the membrane was removed and rinsed in 25 ml of fresh 1xTBS for another 10 minutes then allowed to air dry.

3) CHLOROFORM TREATMENT

Following protease treatment, the colony lifts were placed in 25 ml of 95% ethanol for 2 minutes with gentle shaking. The membrane was then placed in a Buchner funnel where 40 ml of chloroform was filtered through then air dried.

4) PREHYBRIDIZATION

The membrane was immersed in 10 ml of hybridization solution (4xSSC, 5xDenhart's, 0.2 mg/ml Salmon Sperm DNA, 0,1% SDS; MANIATIS *et al*, 1982) and incubated for 1 hour at 40°C.

5) HYBRIDIZATION

A 14 nucleotide probe (LEU462) was end labeled with [γ-^{32}P] ATP. In a microfuge tube, 1.0 μl (10 pmole) of the LEU462 nucleotide, 2.0 μl of 10 x bacteriophage T_4 polynucleotide kinase buffer, 1.0 μl of T_4 polynucleotide kinase, 6.0 μl of [γ-^{32}P] ATP and 10 μl of double distilled H_2O

were added (SAMBROOK *et al.*, 1982). The mixture was incubated for 45 minutes at 37°C to allow for the labeling reaction to proceed. Following labeling, the T_4 polynucleotide kinase was inactivated by heating the mixture for 10 minutes at 68°C. For each membrane, 2 pmole of radiolabeled probe was used. The probe was denatured by boiling for 5 minutes and added to the hybridization solution. Hybridization was allowed to proceed for 4 hours at 40°C.

6) WASHING MEMBRANE

The hybridization solution was discarded and 50 ml of 6xSSC was added. The membrane was washed for 15 minutes at 45°C. The 50 ml of 6xSSC solution was discarded and 30 ml of fresh 6xSSC was added. The membrane was incubated at 43°C for 10 minutes. Again, the solution was discarded the membrane air dried. The membrane was exposed overnight on X-ray film.

PART III. THE ISOLATION OF A NEW HIGH FREQUENCY LAC$^+$ REVERTER

The purpose of the following experiments was to isolate a mutant with high Lac$^+$ reversion due to a structural /functional alteration of the *dcm* or *vsr* gene. The mutation was introduced randomly by hydroxylamine mutagenesis. The mutation was localized to the *dcm* gene by complementation and sequencing.

A) HYDROXYLAMINE (NH_2OH) MUTAGENESIS

Hydroxylamine binds specifically to cytosine residues and causes mispairing with adenine residues resulting in G-C to A-T transitions.

The wild type *dcm/vsr* genes cloned into the tetracycline resistance coding gene of pACYC184 ($pDVW_W$; Table 3), was used as the target of hydroxylamine mutagenesis. In a microfuge tube, 30 μg of pDVW was suspended in a volume of 20 μl of H_2O, 20 μl of 1 M NH_2OH, and 20 μl of H_2O was added. The 60 μl solution was first placed on ice for 30 minutes. The sample was transferred from ice to a 75°C water bath. After 0, 15, 3 and 45 minutes of incubation, 15 μl of solution was removed and placed on ice. Each sample was then used to transform competent JM801 cells.

B) ISOLATION OF SUCCESSFULLY NH_2OH MUTAGENIZED PLASMID

Cells transformed with NH_2OH mutagenized plasmid were plated on Maconkey Lactose Agar containing chloramphenicol and incubated at 37°C for 2 days. Cells which exhibited high levels of papillation (>3 papillae) were isolated and the Lac^+ reversion frequency was determined. Those which demonstrated elevated reversion rates compared to the wild type, were isolated and the plasmid transferred to a new host. The second transfer with the same plasmid was to

ensure that the elevated Lac^+ reversion frequency was due to a mutation in pDVW and not a cellular mutation. Thus, the frequency of Lac^+ reversion for cells containing the NH_2OH mutagenized plasmid, $pDVW_M$ (Table 3), was again determined and compared to those containing the wild type, $pDVW_W$. To verify that the *dcm* gene of $pDVW_M$ product was producing an active cytosine methylase after NH_2OH mutagenesis, an *Eco*RII digestion of genomic DNA was done from CC216 + F'pro 462 harboring $pDVW_M$.

RESULTS

PART I. DEVELOPMENT OF THE ASSAY

The introduction of the CCAGG sequence involved the alteration of the wild type codon sequence 462-463 in the *lacZ* gene from TCA GGA (Ser, Gly) to CCA GGA (Pro, Gly). This modification was accomplished through site-directed mutagenesis using single-stranded f1-Z phage DNA as a template. F1-Z phage is a filamentous f1 phage derivative which has a modified *lacZ* gene inserted into its unique *Eco*R1 site. The *lacZ* gene is modified in such a way as to eliminate an internal *Eco*R1 site and introduce to flanking *Eco*R1 sites (CUPPLES & MILLER, 1988). The new *lacZ* construct is constitutively expressed in the f1 phage due to the absence of a complete *lacI* gene. A mutagenic oligonucleotide containing proline at site 462 was used to insert the missense codon into the *lacZ* gene. A second oligonucleotide served as an aid during site-directed mutagenesis, for the in vitro synthesis of the mutagenized strand.

The mutagenized DNA was used to transform JM801. The f1-Z phage containing the wild-type Ser-462 produced dark blue plaques on LB plates containing X-gal. F1-Z phage which were successfully mutagenized in the proposed *lacZ* region produced very pale blue plaques on X-gal containing

plates. Upon purifying mutant plaques, double stranded and single stranded DNA was isolated.

In order to verify that the lack of ß-galactosidase activity was due to the introduction of the proline (CCA) at site 462 and not due to any other mutation, three experimental procedures were performed. First, to see if the *lacZ* gene was inactivated due to any large deletions within the gene, a digest and electrophoresis of the double stranded f1-Z DNA was done. The mutant f1-Z DNA was digested with *Eco*R1 and electrophoresed on a 1% agarose gel. Two bands of 6.5 kb and 3.4 kb, corresponding to the full size f1 phage DNA and *lacZ* gene respectively, were obtained.

However, since secondary missense mutations were a possibility, a reverse site-directed mutagenesis was done. The reverse mutagenesis involved using the mutant f1-Z phage as a template and an oligonucleotide complementary to the mutant *lacZ* gene except at site 462 where the wild type amino acid, serine (TCA) was present. By reintroducing the wild-type amino acid, the wild-type phenotype should be reinstated, and active ß-galactosidase produced. Therefore, the mutants of interest produced blue plaques on X-gal containing plates, indicating that the missense mutation originally introduced was the only mutation within the *lacZ* gene.

Finally, to confirm the actual presence of the proline codon (CCA) at site 462, single stranded DNA isolated from

the mutant was sequenced. The sequence indicated that indeed the proline codon was present.

SUBCLONING OF THE MUTANT LACZ GENE INTO pBR329

The mutant *lacZ* gene was isolated by *Eco*R1 digestion and inserted into the unique *Eco*R1 site of pBR329. P90C cells containing the recombinant plasmid were plated on media con taining ampicillin and X-gal; pale blue colonies were produced by the plasmids containing the *lacZ* gene. To verify that the *lacZ* gene was present within the *Eco*R1 site of the pBR329, the double stranded DNA isolated from pale blue colonies was digested and electrophoresed on a 1% agarose gel. Plasmid digested with *Eco*R1 produced two bands; the 4.2 kb and 3.4 kb bands corresponded to pBR329 and the *lacZ* gene respectively.

TRANSFER OF THE MUTANT LACZ GENE ONTO AN EPISOME

Recombination between pBR329 containing the mutant *lacZ* gene and the episome which resulted in the episome acquiring the mutant *lacZ*, were selected for their ability to grow on minimal glucose plates containing X-gal without a supplement of proline. Since the wild type episome is *proA, proB, lacZ, lacY, lacA,* and $lacI^{-}$, dark blue colonies form on X-gal plates. However, recombinants containing the mutant *lacZ* gene produced a pale blue colony color in the absence

of the plasmid version. The episome contains a strong promoter (P^+) which allows for the production of ß-galactosidase levels 15 folds higher than the L8 promoter found on f1-Z. Hence, episomes which retained the P^+ promoter while acquiring the proline-462 mutation were identified on Maconkey Lactose agar plates. Episomes with the P^+ promoter produced dark red colonies whereas the weak promoter resulted in pink colonies. The units of ß-galactosidase produced by the episome which contained the P^+ promoter and the mutant *lacZ* were measured. It was found that the mutant episome produced an average of 0.4% of active ß-galactosidase compared to the wild type episome (Table 5).

CODONS AT SITE 462 THAT PRODUCE A LAC^+ PHENOTYPE

When the wild type serine (TCA) 462 codon was altered to a proline (CCA) and the codon at 463 being a glycine (GGC), the CCAGG sequence was introduced. Altering, the serine to a proline, the cell was rendered Lac^-. However, reversion to a Lac^+ phenotype could be due to other base substitutions at site 462. Thus, it was necessary to determine which base substitutions produce a Lac^+ phenotype and which a Lac^- (Table 4). This was done by introducing the F'pro 462 episome into S90C, a wild-type strain. A lactose reversion test was done and Lac^+ cells obtained were sequenced. Codons which resulted in the Lac^+ phenotype were:

TCA (Ser), CTA (Leu), CAA (Gln) and GCA (Ala). However, codons ACA (Thr) and CGA (Arg) were not found.

ß-GALACTOSIDASE ACTIVITY OF THE MISSENSE-462 MUTANTS

To investigate the effect of amino acid substitution at site 462 on the catalytic activity of ß-galactosidase, an *in vitro* assay was performed (Table 5). As one would expect, a different level of ß-galactosidase activity was observed for each amino acid substitution. Furthermore, the amino acid present was responsible for a characteristic and reproducible level of ß-galactosidase activity. Hence, the amino acid present at site 462 could quickly be identified by its ß-galactosidase activity as opposed to sequencing.

TABLE 4:

AMINO ACIDS FOUND AT SITE 462 OF LAC$^+$ REVERTANTS

AMINO ACID	AMINO ACID FOUND AT SITE 462 OF LAC$^+$ REVERTANT
Ser (TCA)	+
Leu (CTA)	+
Gln (CAA)	+
Ala (GCA)	+
Thr (ACA)	NF
Arg (CGA)	NF

NF= Not Found

TABLE 5:

THE EFFECTS OF AMINO ACID SUBSTITUTION AT SITE 462 ON ß-GALACTOSIDASE ACTIVITY

CODON	ß-GALACTOSIDASE ACTIVITY*
TCA[a]	7323 ±410
CTA	2029 ±326
GCA	6093 ±630
CAA	3225 ±169
CCA	27 ±11

* The average activity was determined after n=4 trials

[a] wild-type codon present at site 462

± is the SEM (Standard Error Mean)

PART II. TESTING THE ASSAY SYSTEM

In order for VSP repair to occur, the *vsr* gene must be functional. Cytosine methylase, the *dcm* gene product, is responsible for attaching a methyl group at the 5-carbon site of the second cytosine of the CCAGG sequence. The *vsr* gene product is an endonuclease responsible for removing the T in G·T mismatches. The *dcm* gene alone or in combination with *vsr* were introduced into *dcm-/vsr-* strains to test the following hypotheses: 1) The *dcm* gene alone will result in a high Lac^+ reversion frequency due to lack of repair of deaminated 5-methylcytosines. 2) The introduction of both *dcm* and *vsr* genes would result in a low Lac^+ reversion frequency (wild-type levels) due to vsr endonuclease mediated repair.

EFFECTS OF ADDITION OF CLONED dcm AND/OR vsr GENES ON LAC^+ REVERSION FREQUENCIES IN CC215

The effects the *dcm* and *vsr* genes have on the mutation frequency when introduced into the CC215 strain independently and in combination, were investigated (Table 6). Cells containing only *dcm* (pDCM21) revert to Lac^+ at a high frequency as do cells which contain the *Eco*RII methylase gene (pR215). The *Eco*RII methylase, like the dcm methylase, methylates the second C of the sequence CCAGG. Cells containing both dcm and *vsr* (pDCM21 and pDCM28) produced low

numbers of Lac^+ revertants, comparable to cells containing the control plasmids (pBR322 and pACYC184). Four separate point mutations in the proline 462 codon (CCA) produced a Lac^+ phenotype. In order to be certain that the increased number of Lac^+ revertants observed in strains lacking the vsr gene was due to a C to T mutation at 5-methylcytosine, colony hybridization was used to identify Lac^+ mutants with a leucine (CTA) codon at site 462. In all cases >92% of the colonies contained the CTA codon . The remaining colonies were Lac^+ due to other mutations. These mutations account for the Lac^+ colonies which arose from cells containing both *dcm* and *vsr*.

Confident that the assay monitors the desired mutation, it was used to determine whether the vsr gene from a potential VSP repair mutator is functional. CC401 is a mutator which exhibits an elevated level of C to T mutations at the second C of CCAGG sequences (Ruiz *et al.*, 1993). Since this phenotype is compatible with a defect in VSP repair, the possibility that the strain is vsr^- was investigated. Table 6 indicates that the vsr endonuclease from CC401 is functional. Cells with *dcm* and *vsr* from either CC401 or CC112 have a low mutation rate.

TABLE 6:

EFFECT OF ADDITION OF CLONED dcm AND/OR vsr GENES ON LAC$^+$ REVERSION FREQUENCIES IN CC215

PLASMID	CLONED GENE	# OF LAC$^+$/10^{8}CELLS[a]	
pR215[b]	*Eco*RII methylase	118 ±37	(n=4)
pDCM21	*dcm*	83 ±10	(n=6)
pDCM21 + pDCM28	*dcm* + *vsr*	4 ±0.9	(n=3)
pDW	*dcm*[c]	85 ±25	(n=6)
pDVW	*dcm/vsr*[c]	6 ±1.8	(n=6)
pDM	*dcm*[d]	33 ±11	(n=6)
pDVM	*dcm/vsr*[d]	5 ±1.2	(n=5)
pBR322	none	4 ±0.9	(n=6)
pACYC184	none	1.2±0.3	(n=3)

[a] Values are means ±SEM

[b] pR215 is a derivative of pACYC184; other plasmids are derivatives of pBR322

[c] Cloned from CC112

[d] Cloned from CC401

TABLE 7:

THE EFFECT OF ADDITION OF CLONED dcm AND/OR vsr ON LAC$^+$ REVERSION FREQUENCIES IN CC216

PLASMID	GENE CLONED	# OF LAC$^+$/10^8 CELLS
pR215	*EcoR*II methylase	0.15 ±0.07
pDCM21	*dcm*	1.0 ±0.5
pDW	*dcm*	2.7 ±0.5
pDVW	*dcm* + *vsr*	0.1 ±0.05
pDM	*dcm*	0.1 ±0.05
pDVM	*dcm* + *vsr*	0.3 ±0.15
pBR322	none	0.1 ±0.05
pACYC184	none	2.5 ±0.75

*Data is average n=2 trials; values are means ±SEM

EFFECTS OF THE ADDITION OF CLONED dcm AND/OR vsr GENES ON LAC$^+$ FREQUENCIES IN CC216

To determine whether the genes *dcm* and *vsr* in pDM and pDVM appeared to be functional due to a recombination event with the gencmic genes from CC215, the experiment was repeated using CC216. CC216 is a Δ(*gpt-lac*)5 derivative of RP4182. The strain contains a deletion which extends from approximately 43.78 minutes to 44 minutes which includes the *dcm* and *vsr* genes. Surprisingly, the results were not as expected (Table 7). The frequency of Lac$^+$ reversion with a methylase gene alone (pDCM21, pR215, pDW and pDM) was no higher than with both the *dcm* and *vsr* (pDVM and pDVW) nor that of the control plasmids (pBR322 and pACYC184). The strain appeared to behave as though no cytosine methylation was occurring. To determine if a lack of methylation was the case, genomic DNA from CC216 harboring the *dcm* gene (pDW and pDM) was isolated and subjected to *Eco*RII digestion. Interestingly, no digestion was observed indicating that indeed the *dcm* gene was properly methylating the second C within the CCAGG sequence. The DNA from the CC216 strain was sensitive to *Eco*RII digestion. An episome which contains a leucine (CTA) at site 462 due to a G-C to A-T transition mutation, was introduced into CC216. As the results in Table 6 indicate, CC216 harboring the F'leu462 and the

dcm gene did grow on lactose where CC216 with F'pro462 did not grow (Table 8). Hence, CC216 is capable of transporting lactose into the cell and utilizing the sugar as a carbon source when provided with an ß-galactosidase. The inability of CC216 with F'pro462 to revert to a Lac^+ phenotype when provided with the *dcm* gene and F'pro462, indicates that another protein may be required. A possible explanation for the lack of mutation at 5-methylcytosine sites is that the region deleted between 43 minutes and 44 minutes contains a vsp repair gene. The protein product of the gene may behave as expected in the presence of *dcm* and *vsr*. However, when *vsr* is absent the protein may have mutagenic consequences in the presence of only *dcm* as further described in the discussion.

TABLE 8:

TESTING THE ABILITY OF CC216 TO GROW ON LACTOSE WITH VARIOUS EPISOMES

STRAIN	GENE PRESENT ON PLASMID	GROWTH ON LACTOSE
RP4182	-	+
CC216 + F'pro462	-	-
CC216 + F'pro462 + pDW	*dcm*	-
CC216 + F'leu462	-	+

TABLE 9:

LACTOSE REVERSION FREQUENCY OF HYDROXYLAMINE MUTAGENIZED pDVW

STRAIN	PLASMID	# LAC^+ CELLS/10^8 CELLS
CC215	pACYC184	1.0 ±0.5 (n=2)
CC215	$pDVW_W$	1.25 ±0.5 (n=2)
CC215	$pDVW_M$	52.3 ±10 (n=4)

n= number of trials

±Standard Error Mean (SEM)

PART III. THE ISOLATION OF A NEW HIGH FREQEUENCY LAC$^+$ REVERTER

Hydroxylamine was used as a chemical mutagen to randomly introduce missense mutations within the *dcm* gene. The random introduction of missense mutations would permit a study of the effects of amino acid substitution on the function of cytosine methylase. The *in vitro* hydroxylamine mutagenesis of pDVW$_W$ and transformation into CC215 resulted in the isolation of a high papillator and high Lac$^+$ reverter (Table 9). The pDVW$_W$ plasmid contains the *dcm* and *vsr* genes. To localize the mutation, the vsr gene was sequenced and found to be wild type. Hence, the mutation responsible for the high Lac$^+$ reversion phenotype may be due to an alteration in cytosine methylase activity. The isolation and *Eco*RII digestion of CC215 genomic DNA revealed that the *dcm* gene is functional.

DISCUSSION

In *Esherichia coli*, VSP repair is involved in correcting T·G mismatches which arise upon deamination of 5-methylcytosine sites. However, not all cytosines in *E. coli* are methylated. The product of the *dcm* gene, cytosine methylase, adds a methyl group at the 5 carbon of the second cytosine found within the CCAGG sequence. The methylated cytosine within the CCAGG region is a mutation hot spot (COULONDRE *et al.*,1978; SCHAAPER *et al.*, 1986) . The G·T mismatches that arise are repaired by the *vsr* gene product, an endonuclease. Through an unresolved mechanism, the *vsr* endonuclease recognizes the T as the faulty nucleotide and introduces a nick 5' of the T.

An assay system was developed which specifically monitors G-C to A-T transitions at 5-methylcytosines. First, the assay was useful in determining that a potential VSP mutator strain, CC401, contains a mutation outside the *dcm* and *vsr* gene region. Second, the assay was used in isolating a plasmid which confers a mutator phenotype which can be attributed to a possible alteration in the cytosine methylase function.

PART I: DEVELOPMENT OF THE ASSAY

The assay system developed involves the use of the *lacZ* gene which codes for the enzyme ß-galactosidase. The enzyme cleaves lactose into glucose and galactose and thus enables the organism to use lactose as a carbon source.

A mutant *lacZ* gene was constructed which contains the CCAGG sequence at sites 462-463. This was done by altering the wild type sequence TCA GGA (Ser, Gly) to CCA GGA (Pro, Gly). Only 0.4% of the wild type ß-galactosidase activity was produced with the substitution of serine for proline at site 462 (Table 5). The level of ß-galactosidase produced by the mutant *lacZ* gene was insufficient for growth on lactose and hence, rendered the cell Lac^- (Table 8).

The CCAGG sequence was recognized by the methylase of the *dcm* gene where the second cytosine was methylated. Deamination of the 5-methylcytosine resulted in a thymine hence, the sequence CTAGG. Furthermore, 5-methylcytosine deamination alters the codon at site 462 from a CCA to CTA thus the amino acid from proline to leucine. The insertion of leucine at position 462 in ß-galactosidase reverted the cell back to Lac^+ (Figure 1).

PART II: TESTING THE ASSAY

EFFECT OF ADDITION OF CLONED dcm AND/OR vsr GENES ON LAC$^+$ FREQUENCIES IN CC215

In the presence of the *dcm* gene alone, a high frequency of G-C to A-T transition mutations was observed (Table 6). The mutation frequency decreased to background levels when both the *dcm* gene and *vsr* gene were present. First, the results demonstrated that the assay can specifically monitor C-G to T-A transitions. Second, a potential VSP mutator, $pDVW_M$ is *vsr*$^+$.

The VSP mutator strain, CC401, exhibits a high frequency of C-G to T-A transitions at CCAGG sequences. The mutation was found to map at 43 minutes as does the *vsr* gene (RUIZ, *et al.*, 1993). It had been found that *in vitro*, the *vsr* gene product is the only protein necessary for VSP repair (SOHAIL *et al.*, 1990). The *dcm* and *vsr* genes (pDVM) isolated from the mutant behaved similarly to that of the wild-type (pDVW; Table 6) . Hence, the *dcm* and *vsr* genes from the mutator strain are functional. This indicated that another gene may also play a role in G·T mismatch repair.

The unknown gene product may work in association with the vsr endonuclease in G·T mismatch recognition and/or initiation of mismatch repair.

EFFECT OF ADDITION OF CLONED dcm AND/OR vsr GENES ON LAC$^+$ REVERSION FREQUENCIES IN CC216

The results of the Lac$^+$ revrsion frequency experiment lends support to the idea that a new gene may be involved in VSP repair. The CC216 strain is deleted in tregion between 43 minutes and 44 minutes which includes the *dcm* and *vsr* genes. In comparison, the CC215 strain contains only a point mutation in *dcm* which results in a *dcm*$^-$/ *vsr*$^-$ phenotype. Yet, the introduction of a cytosine methylase gene alone (pDW, pDM, pDCM21, pR215) into CC216, resulted in a Lac$^+$ reversion frequency no higher that background levels (Table 7). The cytosine methylase of the *dcm* gene was introducing a methyl group at the second cytosine of CCAGG sequences of CC216 genomic DNA, demonstrated by the inability of *Eco*RII to cut. Interestingly, it has previously been demonstrated that 5-methylcytosines are 2.5 times more likely than cytosines to spontaneously deaminate (JONES *et al.*, 1992). Yet, the results obtained using CC216 do not demonstrate any 5-methylcytosine instability.

VSP repair is assumed to have an antimutagenic function in bacteria that contain 5-methylcytosines. Thus, the persistence of mutation hotspots at 5-methylcytosine sites despite overexpression of *vsr* suggests that events other than cytosine deamination might be responsible for some of

the mutagencity of the modified base (LIEB, 1991).

The two results suggest that *dcm* and *vsr* alone do not explain 5-methylcytosine hotspots are 1) the CC401 strain exibits a high frequency of Lac^+ reversion yet *dcm* and *vsr* are functional; 2) The CC216 strain does not exhibit a 5-methylcytosine hotspot despit the presence of the cytosine methylase.

The proposed hypothesis provides an explanation for the unexpected results obtained with CC216. The deletion of DNA between 43 minutes and 44 minutes could also have deleted a gene involved in 5-methylcytosine deamination. The gene product may influence 5-methylcytosine deamination indirectly by interacting with the cytosine methylase. In the presence of the *vsr* endonuclease, the effects of the cytosine methylase and the unknown protein are counteracted. On the other hand, the unknown protein may interact directly to deaminate 5-methylcytosines. Although the presence of a deaminase in elevated levels would be exhibited as a vsr^- phenotype, the activity of the vsr endonuclease should remain unaffected. Evidence supporting the hypothesis is supplied indirectly through *in vitro* studies of G·T mismatch repair. A heteroduplex containing a G·T mismatch was incubated with RP4182 cell extract where G·T to G-C repair was highly favored (YASHAR and MODRICH, personal communication). This indicates that the ability of RP4182 (thus CC216) to repair G·T mismatches is unaffected.

CC401 AND CC216, SINGLE GENE MUTATION OR DIFFERENT?

Phenotypically, CC401 exhibits a high Lac^+ reversion frequency hence, has a high G-C to A-T transition frequency. Yet, CC401 contains functional *dcm* and *vsr* genes (Table 6). On the other hand, CC216 exhibits a wild-type mutation frequency even in the absence of the vsr gene (Table 7). If a single gene is involved in 5-methylcytosine deamination, then the high mutation rate observed in CC401 may be due to overexpression of the gene compared to a complete absence of the gene in CC216. This possibility can be tested in vitro using a heteroduplex, a homoduplex as a control and cell extracts.

By creating a radiolabeled heteroduplex with the sequence CTAGG/GGTCC at the center, the T·G mismatch is introduced. Cell extracts obtained from CC401 and CC216 would be independently and simultaneously incubated with the heteroduplex.

A) HYPOTHESIS 1: GENE ASSOCIATES WITH VSR ENDONUCLEASE

In the wild type situation, cell extract from a functional vsp repair system, CC112, when incubated with the heteroduplex, would result in cleavage of the T at the 5' end. As such, two bands should be obtained. One band would represent the DNA strand containing the GGTCC sequence. The

second band would be half the size of the original strand due to cleavage of the T at the 5' end within the CTAGG sequence.

Cell extract obtained from CC401 and CC216, when incubated independently with the heteroduplex should not result in 5' cleavage of the T at the T·G mismatch due to a missing component necessary for vsp repair. If the phenotypes observed for CC401 and CC216 are due to mutation within the same gene, mixing of both cell extracts would not complement the mutation hence, no cleavage of the heteroduplex would occur. If different genes are involved, the cell extract of one strain would supply the missing component of the other hence, complementation would occur resulting in cleavage of the heteroduplex at the 5'end of the T.

B) HYPOTHESIS 2: GENE IS A DEAMINASE

If the phenotype observed with CC401 is due to overexpression of a deaminase vsp repair would remain unaffected. Thus, incubating CC401 cell extract with the G·T heteroduplex would result in cleavage 5' of the T. On the other hand, incubation of CC401 cell extract with methylated C within the CCAGG containing homoduplex would result in 5-methylcytosine deamination. The resultant G·T mismatch would be detected by resistance to *Msp*I cleavage.

In the case of CC216, the absence of deaminase would have no effect on the homoduplex hence, remain *Msp*I sensitive.

Although there is evidence to support the existence of a deaminase, no evolutionary advantage has been proposed (STEINBERG and GORMAN, 1993).

Besides the *vsr* gene which map at 43 minutes, it appears that one if not two new genes involved in VSP repair also maps at 43 minutes. It is not unreasonable to have more than one VSP repair gene mapping at 43 minutes considering that one minute is approximately 100 kb of DNA. Of the known genes found to map at 43 minute, *sbc*B which maps at 43.925 minutes, codes for exonuclease I (BACHMANN, 1990). However, the *sbc*B gene product participates in the methyl-directed repair mechanism. The *sbc*B gene product is involved in suppressing the decrease in genetic recombination associated with *rec*B or *rec*C mutations. Also, *sbc*B suppresses the DNA-damage sensitivity of *rec*B and *rec*C mutants (MOLINEUX and GEFTER, 1975; KUSHNER et al., 1972).

PART III. THE ISOLATION OF A NEW HIGH FREQUENCY LAC$^+$ REVERTER

In vitro hydroxylamine mutagenesis of pDVW$_M$ (*dcm* and *vsr* cloned into pACYC184) resulted in a mutator where the increased frequency of G-C to A-T transition mutations was due to an alteration of function in the cytosine methlyase,

not the *vsr* endonuclease (Table 9). Results suggest that although both genes are functional, the *vsr* gene contains the wild type sequence. Therefore, the hydroxylamine may have introduced a missense mutation within the *dcm* gene, resulting in an altered methylase activity. *Eco*RII digestion of genomic DNA isolated from the CC215 strain harboring the $pDVW_M$ plasmid indicated that the ability of the cytosine methylase to methylate the second C of CCAGG sequences was not affected.

The increase in G-C to A-T transition mutations observed may be due to an increase in cytosine methylase activity thereby, overwhelming the repair system of the vsr endonuclease. To determine if enhanced cytosine methylase activity is responsible for the mutator phenotype the rate of transfer of tritium from [methyl-^{3}H]S-adenosyl methionine into poly d(IC)-poly d(IC), a synthetic DNA substrate which has high methyl-accepting activity, would be measured (PEDRALI-NOY and WEISSBACH, 1986; LI *et al.*, 1992).

The assay has proved useful in assessing vsp repair *in vivo* thus, eliminating possible artifacts attributed to in vitro experimentation. Past experiments supported the hypothesis that the *dcm* and *vsr* genes (stimulated by MutL and MutS) are all that is required for VSP repair (SOHAIL *et al.*, 1990). The experimental results presented by RUIZ *et al.*, 1993, indicate that at least one new gene which maps at

43 minutes, is essential for vsp repair. Similarly, data presented here support the conclusions drawn by RUIZ *et al.*, 1993.

Interestingly, the high deamination frequency associated with 5-methylcytosines, had been attributed to structural instability (COULONDRE *et al.*, 1978). However, no such instability was observed in CC216 whose genomic DNA was being methylated at the second C within CCAGG sequences. Although the mechanism of 5-methylcytosine deamination was not elucidated, results obtained using CC216 indicate that deamination may be due to an enzymatic process and not to structural instability. Deamination may be the action of a specific deaminase or protein product that directly or indirectly associates with the vsr endonuclease.

However, deamination of 5-methylcytosines is not exclusive to *E. coli*. In mammalian systems, 5-methylcytosines are also mutation hotspots. Besides the obvious need to maintain DNA sequence fidelity, cytosine methylation has been correlated with gene silencing. Hence, 5-methlycytosines are potentially important in development and regulation. The elucidation of vsp repair in a simple organism like *E. coli* can provide insight into the same mechanism in more complex organisms such as humans.

FIGURE 1: The *lacZ* assay system used to monitor C to T mutations at the second C of CCAGG. Thin lines represent the *lacZ* gene; open boxes represent ß-galactosidase. The sequences shown are codons 460-462. Deamination of the 5-methylcytosine resulted in the conversion of a glutamine to an amber codon at site 461. If no VSP repair occurred an amber suppressor introduced a glutamic acid at site 461. Thus, a complete and active ß-galactosidase was translated reverting the cell to a Lac^+ phenotype.

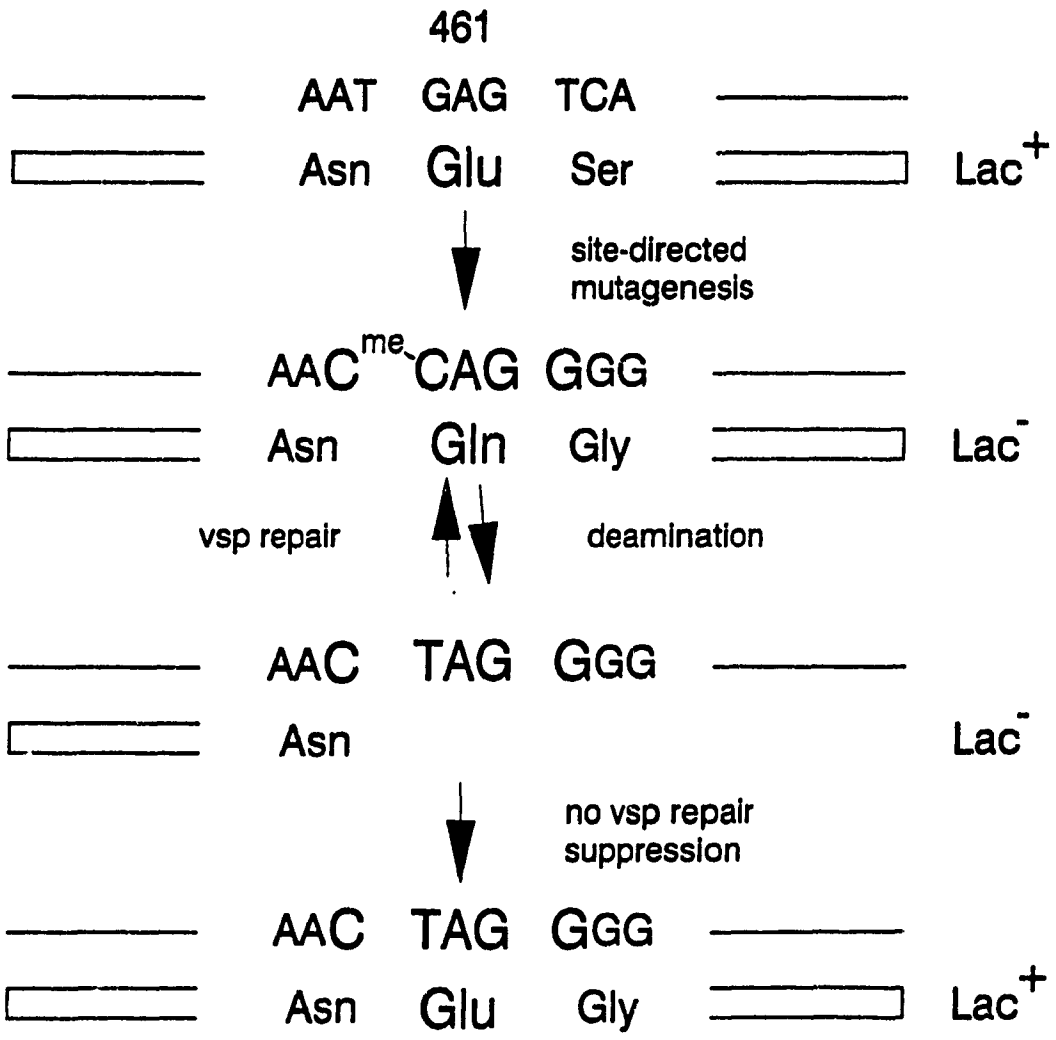
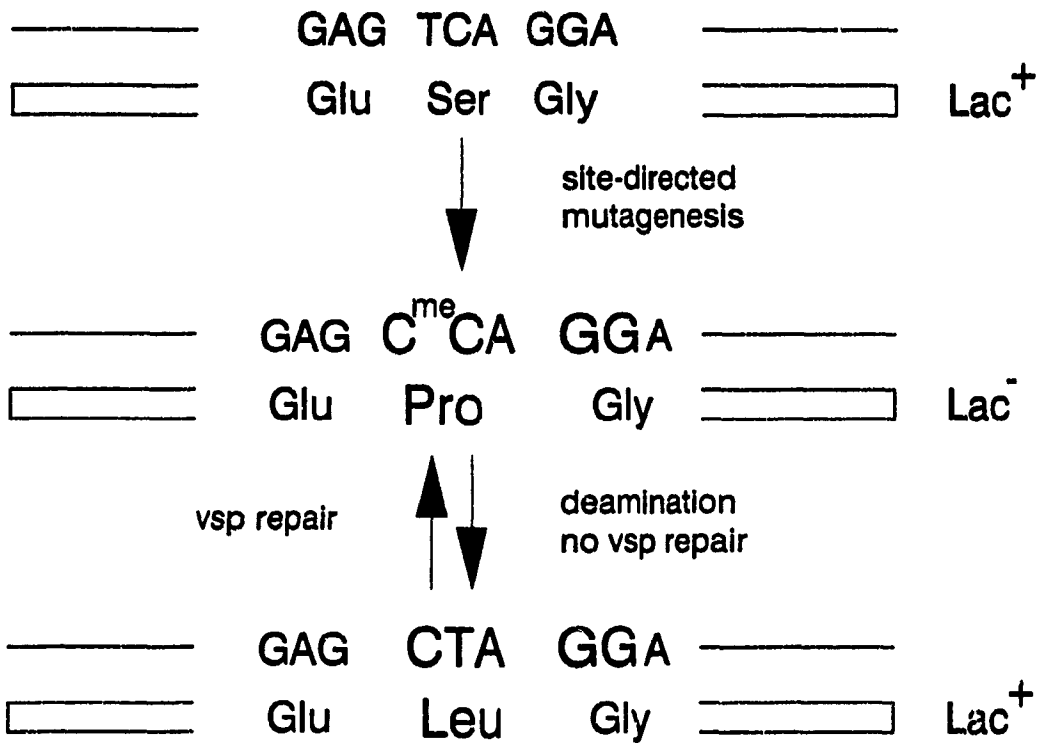

461
AAT GAG TCA
Asn Glu Ser
Lac+
site-directed
mutagenesis
AAC me CAG GGG
Asn Gln Gly
Lac-
vsp repair
deamination
AAC TAG GGG
Asn
Lac-
no vsp repair
suppression
AAC TAG GGG
Asn Glu Gly
Lac+

FIGURE 2: The modified lacZ assay involved sequences 462 and 463. Deamination of the 5-methylcytosine results in the alteration of codon 462 from a proline (CCA) to a leucine (CTA). If no VSP repair occurred, the introduction of a leucine which results in an active ß-galactosidase and a Lac^+ phenotype.

MODIFIED ASSAY SYSTEM



REFERENCES

AKIYAMA, M., H. MAKI, M. SEKIGUCHI and T. HORIUCHI 1989. A specific role of MutT protein: To prevent dC-dA mispairing in DNA replication. Proc. Natl. Acad. Sci. USA. 86: 3949-3952.

AU, K.G., M. CABRERA, J.H. MILLER, and P. MODRICH 1988. *Escherichia coli mutY* gene product is required for specific A·G to C·G mismatch correction. Proc. Natl. Acad. Sci. USA 85: 9163-9166.

AU, K.G., S. CLARK, J.H. MILLER, and P. MODRICH 1989. *Escherichia coli mutY* gene encodes an adenine glycosylase active on G·A mispairs. Proc. Natl. Acad. Sci. USA 86: 8877-8881.

AU, K.G., K. WELSH and P. MODRICH 1992. Initiation of methyl-directed mismatch repair. J. Biol. Chem. 267 (17): 12142-12148.

BACHMANN, B.J. 1990. Linkage map of *Escherichia coli* K-12, edition 8. Microbiological Reviews, 54 (2): 130-197.

BADER, D.E., M. RING and HUBER 1988. Site-directed mutagenic replacement of Glu461 in ß-galactosidase (Escherichia coli): Evidence that Glu461 is important for activity. Biochem. Res. Commun. 153: 301-306.

BHAGWAT, A.S., A. SOHAIL and M. LIEB 1988b. A new gene involved in mismatch correction in *Escherichia coli*, Gene 74: 155-156.

BEN-HATTAR,J. and J. JIRICNY 1988. Methylation of single CpG dinucleotides within a promoter element of the herpes simplex virus *tk* gene reduces its transcription *in vivo*. Gene 65: 219-227.

BESTOR, T.H. and A. COXON 1993. The pros and cons of DNA methylation. Curr. Biol. 3 (6): 384-386.

BOYES, J. and A. BIRD 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell. 64: 1123-1134.

BROWN, T.C. and J. JIRICNY 1987. A specific mismatch repair event protects mammalian cells from loss of 5-methylcytosine. Cell. 50: 945-950.

CABRERA, M., Y. NGHIEM, and J.H. MILLER 1988. *Mut*M, a second mutator locus in *Escherichia coli* that generated G-C to T-A transversions. J. Bacteriol. 170 (11): 5405-5407.

CEDAR, H. 1988. DNA methylation and gene activity. Cell. 53: 3-4.

CLAVERYS, J-P and S.A. LACKS 1986. Heteroduplex Deoxyribonucleic acid base mismatch repair in bacteria. Micro. Rev. 50: 133-165.

COULONDRE, C. and J.H. MILLER, P.J. FARABAUGH and W. GILBERT 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. Annu. Rev. Genet. 10: 135-156.

CUPPLES, C.G. and J.H. MILLER 1988. Effects of amino acid substitutions at the active site in *Escherichia coli* ß-galactosidase. Genetics 120: 637-644.

CUPPLES, C.G. and J.H. MILLER 1989. A set of *lacZ* mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. Proc. Natl. Acad. Sci. USA 86: 5345-5349.

EHRLICH, M., X-Y. ZHANG and N.M. INAMDAR 1990. Spontaneous deamination of cytosine and 5-methylcytosine residues in DNA and replacement of 5-methylcytosine residues with cytosine residues. Mut. Res. 238: 277-286.

FOWLER, A.V., I. ZABIN, M.L. SINNOTT and P.J. SMITH 1978. Methionine 500, the site of covalent attachment of an active site-directed reagent of ß-galactosidase. J. Biol. Chem. 253: 5283-5285.

GRILLEY, M., J. HOLMES, B. YASHAR, and P. MODRICH 1990. Mechanisms of DNA-mismatch correction. Mutation Research 236: 253-267.

HENNECKE, F., H. KOLMAR, K. BRUNDL, and H-J. FRITZ 1991. The *vsr* gene product of *E. coli* K12 is a strand and sequence specific DNA mismatch endonuclease. Nature 353: 776-778.

HERRCHEN, M. and LEGLER, 1984. Identification of an essential carboxylate group at the active site of *lacZ* ß-galactosidase from *Escherichia coli*. Eur. J. Biochem. 138: 527-531.

HERMAN, G.E., and P. MODRICH 1981. *Escherichia coli* K12 clones that overproduce *dam*-methylase are hypermutable. J. Bacteriol. 145: 644-646.

HOLLIDAY, R., and G.W. GRIGG 1993. DNA methylation and mutation. Mutation Res. 285: 61-67.

IGUCHI-ARIGA, S.M. and W. SCHAFFNER 1989. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. Genes Dev. 3: 612-619.

JIRICNY, J., M. HUGHES, N. CORMAN, and B.B. RUDKIN 1988. A human 200-kDa protein binds selectively to DNA fragments containing G·T mismatches. Proc. Natl. Acad. Sci. USA 85: 8860-8864.

JONES, P.A. S.M. TAYLOR and V.L. WILSON 1983. Inhibition of DNA methylation by 5-azacytidine. Recent Results Cancer Res. 84: 202-211.

JONES, M., R. WAGNER, and M. RADMAN 1987a. Mismatch repair of deaminated 5-methylcytosine. J. Mol. Biol. 194: 155-159.

JONES, M., R. WAGNER, and M. RADMAN 1987b. Mismatch repair and recombinaation in *E. coli*. Cell 50: 621-626.

JONES, P.A., W.M. RIDEOUT (III), J-C. SHEN, C.H. SPRUCK and Y.C. TSAI. 1992. Methylation, mutation and cancer. BioEssays 14 (1): 33-36.

KRAMER, B., W. KRAMER, and H-J FRITZ 1984. Different base/base mismatches are corrected wit different efficiencies by the methyl-directed DNA mismatch-repair system of *E. coli*. Cell 38: 879-887.

KUSHNER, S.R., H. NAGAISHI and A.J. CLARK 1972. Indirect expression of *recB* and *recC* mutations by exonuclease I deficiency. Proc. Natl. Acad. Sci. USA 69: 1366-1370.

LAHUE, R.S., K.G. AU and P. MODRICH 1989. DNA mismatch correction in defined system. Science, 245: 160-164.

LI, E., T.H. BESTOR and R. JAENISCH 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 69: 915-926.

LIEB, M. 1983. Specific mismatch correction in bacteriophage lambda crosses by very short patch repair. Mol. Gen. Genet. 191: 118-125.

LIEB, M. 1987. Recombination in the lambda repressor gene: Evidence that very short patch (VSP) mismatch correction restores a specific sequence. Mol. Gen. Genet. 199: 465-470.

LIEB, M. 1991. Spontaneous mutation at a 5-methylcytosine hotspot is prevented by very short patch (VSP) mismatch repair. Genetics 128: 23-27.

LU, A-L. and D-Y. CHANG 1988. A novel nucleotide excision repair for the conversion of an A/G mismatch to C/G base pair in *E. coli*. Cell. 54: 805-812.

LU, A-L., S. CLARK, and P. MODRICH 1983. Methyl-directed repair of DNA base-pair mismatches *in vitro*. Proc. Natl. Acad. Sci. USA 80: 4639-4643.

LU, A-L., K. WELSH, S.J. SU and P. MODRICH 1984. Repair of DNA base pair mismatches in extracts of *Escherichia coli*. Cold Spring Harbor Symp. Quant. Biol. 49: 589-596.

MAKI, H. and M. SEKIGUCHI 1992. MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. Nature 355: 273-275.

MANIATIS, T., E.F. FRITSCH and J. SAMBROOK 1982. *Molecular Cloning*: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

MARINUS, M.G. 1973. Location of DNA methylation genes on the *Escherichia coli* K12 genetic map, Nol. Gen. Genet., 127: 47-55.

MARINUS, M.G. and N.R. MORRIS 1975. Pleiotropic effects of a DNA adenine methylation mutation (*dam-3*)in *Escherichia coli* K12. Mutat. Res. 28: 15-26.

MICHAELS, M.L., C. CRUZ, A.P. GROLLMAN and J.H. MILLER 1992. Evidence that MutY and MutM combine to prevent mutations by an oxidatively damaged form of guanine in DNA. Proc. Natl. Acad. Sci. USA 89: 7022-7025.

MILLER, J.H. 1972. *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

MILLER, J.K. and W.M. BARNES 1986. Colony probing as an alternative to standard sequencing as a means of direct analysis of chromosomal DNA to determine the spectrum of single-base changes in regions of known sequence. Proc. Natl. Acad. Sci. USA 83: 1026-1030.

MODRICH, P. 1987. DNA mismatch correction. Ann. Rev. Biochem. 56: 435-466.

MOLINEUX, I.J. and M.L. GEFTER 1975. Properties of the *Escherichia coli* DNA-binding protein interaction with nucleolytic enzymes and DNA. J. Mol. Biol. 98: 811-825.

NGHIEM, Y., M. CABRERA, C.G. CUPPLES, and J.H. MILLER 1988. The *mutY* gene: A mutator locus in *Escherichia coli* that generates G-C to T-A transversions. Proc. Natl. Acad. Sci. USA 85: 2709-2713.

PEDRALI-NOY, G. and A. WEISSBACH 1986. Mammalian DNA methyltransferases prefer poly (dI-dC) as substrate. J.Biol. Chem. 261: 7600-7602.

PETROPOULOS, L. J.J. VIDMAR, E. PASSI and C.G. CUPPLES 1993. A simple assay for monitoring the mutagenic effects of 5-methylcytosine deamination in *Escherichia coli*. Mutat. Res., in press.

PUKKILA, P.J., J. PETERSON, G. HERMAN, P. MODRICH, and M. MESELSON 1983. Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in *Escherichia coli*. Genetics 104: 571-582.

RADMAN, M. and R. WAGNER. 1986. Mismatch repair in *Escherichia coli*. Ann. Rev. Genet. 20: 523-538.

RADMAN M., and R. WAGNER 1988. The High Fidelity of DNA Duplication. Scientific American *August*: 40-46.

RAZIN, A. and A.D. RIGGS 1980. DNA methylation and gene function. Science. 210: 604-610.

RUIZ, S.M. 1992. The characterization of an *Escherichia coli* strain deficient in VSP mismatch repair. M.Sc. Thesis.

RUIZ, S.M., S. LETOURNEAU and C.G. CUPPLES 1993. Isolation and characterization of an *Escherichia coli* strain with a high frequency of C-to-T mutations at 5-methylcytosines. J. Bacteriol. 175: 4985-4989.

SANCAR, A. and J.E. HEARST 1993. Molecular Matchmakers. Science 259: 1415-1420.

SCHAAPER, R.M., B.N. DANFORTH and B.W. GLICKMAN 1986. Mechanism of spontaneous mutagenesis, an analysis of the spectrum of spontaneous mutations in *Escherichia coli lacI* gene. J. Mol. Biol. 189: 273-284.

SCHAAPER, R.M. and R.L. DUNN 1987. Spectra of spontaneous mutations in *Escherichia coli* strains defective in mismatch correction: The nature of *in vivo* DNA replication errors. Proc. Natl. Acad. Sci. USA. 84: 6220-6224.

SINNOTT, M.L. 1978. Ion-pairs and catalysis by the *lacZ* ß-galactosidase of *Escherichia coli*. FEBS Lett. 94: 1-9.

SOHAIL, A., M. LIEB, M. DAR, and A.S. BHAGWAT 1990. A gene required for very short patch repair in *Escherichia coli* is adjacent to the DNA cytosine methylase gene. J. Bacteriol. 172 (8): 4214-4221.

STEINBERG, R.A. and K.B. GORMAN 1993. Linked spontaneous C-G to T-A mutations at CpG sites in the gene for protein kinase regulatory subunit. Mol. and Cell. Biol. 12 (2): 767-772.

SU, S-S., R.S. LAHUE, K.G. AU, and P. MODRICH 1988. Mispair specificity of methyl-directed DNA mismatch correction *in vitro*. J. Biol. Chem. 263 (14): 6829-6835.

TOPAL, M.D. and J.R. FRESCO, 1976. Complementary base pairing and the origin of substitution mutations. Nature 263: 285-289.

WAGNER, R., C. DOHET, M. JONES, M-P DOUTRIAUX, F. HUTCHINSON and M. RADMAN. 1984. Involvement of Escherichia coli mismatch repair in DNA replication and recombination. Cold Spring Harbor Sympl Quant. Biol. 49: 611-615.

WAGNER, R. and M. MESSELSON 1976. Repair tracts in mismatched DNA heteroduplexes. Proc. Natl. Acad. Sci. USA. 73:

WATT, F. and P.L. MOLLOY 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter Genes. Dev. 2: 1136-1143.

WIEBAUER, K., and J. JIRICNY 1989. *In vitro* correction of G·T mispairs to G-C pairs in nuclear extracts from human cells. Nature 339: 234-236.

WIEBAUER, K., and J. JIRICNY 1990. Mismatch-specific thymine DNA glycosylase and DNA polymerase ß mediate the correction of G-T mispairs in nuclear extracts from human cells. Proc. Natl. Acad. Sci. USA. 87: 5842-5845.

ZOLLER P. and M. SMITH 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vector. Meth. Enzymol. 100: 468-500.