NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Construction and characterization of mutational targets in the lacZ gene of *Escherichia coli*

Joseph John Vidmar

A Thesis

in

The Department

of

Biology

Presented in Partial Fulfillment of the Requirements

for the degree of Master of Science at

Concordia University

Montreal, Quebec, Canada

November 1992

ISBN 0-315-84639-9

Canada

ACKNOWLEDGMENTS

There are many people who have helped me during my stay at Concordia. First and foremost, I would like to thank my supervisor Dr. C.G. Cupples, for her help, understanding, guidance, and patience.

I would like to also thank my committee members, Dr. A. Tsang, Dr. M.J. Kornblatt, Dr. R. Germanario, and Dr. E. Newman for their insight and knowledge.

For my friends who have given me plenty of support in making the research and thesis enjoyable of fun, Louisa, Marc, John (both of them), Erica, Alain, Mike and Tim. And, also, the friendship of Catherine Soos.

I would also have to mention the friendship and moral support of Claudine Charlery, and my parents, who pushed me when I needed a push.

TABLE OF CONTENTS

LIST OF FIGURES

Number Page

LIST OF TABLES

INTRODUCTION

The main focus of this thesis was the development and characterization of specific assay systems for the analysis of repair systems in *Escherichia coli*. The use of these assay systems can elucidate the mechanism of prevention and/or repair of mutations. Great energy has been focused on the understanding of how cells keep the fidelity of DNA replication high, approximately one mistake in 10^{10}-10^{11} bases replicated. In the past, the initial characterization of repair systems was tedious. The isolation of a mutator (a strain deficient in a repair system) was quite difficult, and usually depended on brute force. Once a mutator was found, it was characterized according to the type of mutation it caused *in vitro* using heteroduplex analysis (e.g. Lu et al. 1983) or *in vivo* by repeated sequencing of a specific gene (e.g. Schaaper and Dunn, 1987). Cupples and Miller (1989) developed an assay system, whereby the isolation and characterization of repair deficient strains is carried out in a simple phenotypic manner. The assay utilizes site 461 of the lacZ gene (encodes β-galactosidase). By manipulating the DNA sequence at this site, they were able to monitor base substitution events.

The present work is an expansion of the lacZ phenotypic assay in the development of novel target sequences and further analysis of previously constructed target sequences. The project is discussed in three parts:

a) the construction and preliminary analysis of a potential target sequence, site ser-462 in the lacZ gene (through amino acid substitution);

b) the effects of repair system interactions on base substitutions at the original target site, 461;

c) the construction and characterization of $(GC)_n$ dinucleotide frameshift targets at site 504.

DNA replication

DNA replication is the process by which an organism makes a copy of its genome. It is important that DNA replication is faithful to the original DNA sequence, in order that the subsequent daughter cell has an exact copy of the parental genome. This is necessary since DNA carries the genetic material which is the blueprint of the cell and its biological functions.

In bacteria, DNA replication reviewed by Suzuki et al. 1987) commences at a unique origin. In *E. coli* this origin is oriC locus. OriC locus, is 245 base pairs in size, and is located at 87 minutes on the *E. coli* genome map. At this locus, a complex of dnaB and dnaC proteins bind to the site, where they bend and open the double helix of DNA. The template DNA strands are then stabilized by the additional binding of ssb proteins. This allows the binding of the DNA replication machinery. DNA replication starts at this site and proceeds in two opposite directions at the same velocity, forming two replication forks. At each of the DNA replication forks both strands serve as templates for DNA synthesis. The template strands are anti-parallel, one going in the 5' to 3' direction while the other goes in the 3' to 5' direction. DNA synthesis only occurs in the 5' to 3' direction, therefore two types of strand synthesis occurs, leading and lagging. The leading strand synthesis is continuous (no breaks in replication), while lagging strand synthesis is discontinuous. Both types of synthesis are accomplished by DNA polymerase III holoenzyme, but lagging strand synthesis also needs the addition of the enzymes DNA polymerase I and DNA ligase.

DNA replication can result in many types of errors, from base substitutions to frameshifts. Base substitutions occur, when there is a mispairing of the template base with a non complementry base. If this mismatch is not repaired, on the subsequent round of DNA replication, the mutation will be fixed in the DNA sequence. There are two types base substitutions events transitions and transversion. Transition events occur when a

purine is replaced by another purine, or a pyrimidine is replace by another pyrimidine. Transversion events occur when a purine is replaced by a pyrimidine, or a pyrimidine is replaced by a purine. Frameshifts occur by the addition or deletion of nucleotides in the DNA sequence. It has been hypothesized (Streisinger et al 1966) that these mutagenic events occur by a strand slippage mechanism. The addition event occurs when the DNA polymerase inserts a non-pairing base (addition event). The deletion event occurs when the DNA polymerase does not insert a base opposite a template base, hence the template base becomes non-pairing (deletion event). Both these events if not repaired, will result in frameshift mutations on the subsequent round of replication.

Mechanisms of DNA replication correction

Since DNA carries the information needed for continued life, it is important for the passage of this information that it be faithfully copied. Organisms have evolved many strategies to insure fidelity of replication. In *E.coli* these include: 1) removal of mutagenic base analogs from the nucleotide pool, mutT, 2) insertion of the proper base at the replication fork followed by proof reading by the ε subunit of DNA polymerase, dnaQ, 3) long excision repair where a long stretch of DNA is removed and resynthesized, mutHLS, and 4) short patch repair, mutG, mutY, and mutM. The end result is a decrease in the frequency of mutations and increased accuracy of replication. DNA polymerase make one mistake per 10^4-10^5 bases replicated (Loeb et al. 1986). Proof reading by the DNA polymerase ε subunit (coded by dnaQ or mutD) reduces the error rate to 10^{-7}-10^{-8}, an increase in fidelity of 1000 fold. The combined activity of the other repair systems increases fidelity a further 10^3. Therefore the cell ensures that mutations resulting from DNA replication are extremely rare.

Dam-directed mismatch repair

The dam-directed mismatch repair system, is part of long patch repair, which corrects for a variety of replication errors. This system is strand specific, and involves the excision of up to a kilobase or more of DNA (reviewed by Grilley et al. 1990, Claverys and Lacks, 1986; Modrich, 1987). There are five major genes which code for proteins involved in this system. They are dam, mutS, mutL, mutU, and mutH.

After replication there are mismatches present in the DNA. The mutS protein is a mismatch recognition protein. The binding of this protein to a mismatch will activate a set of interactions with other proteins involved in this repair. The mutL protein will bind to both the mutS protein and to the mismatch. This will signal the mutH protein to bind to a palindromic sequence GATC/CΓ_{me}AG up to several kilobases from the mismatch, and cleave the phosphodiester bond at the 5' P of the G on the non-methylated strand. The mutU protein, with the assistance of ssb protein, will unwind the DNA past the mismatch. The displaced strand is resynthesized by DNA polymerase I. The ability of this system to distinguish the correct base from the mispaired base is dependent on under methylation of the newly synthesized strand.

It has been shown that methyl-directed mismatch repair deficient stains have an increase in the mutational frequency *in vivo* for transition events G:C to A:T and A:T to G:C (Cupples and Miller, 1989), and +1 and -1 frameshifts at monomeric sequences (Cupples et al. 1990).

MutT repair

Mutator strain mutT was isolated by Treffers et al. (1954). It has been found that the mutT strain increases the mutation rate specifically for A:T to C:G transversions (Yanofsky et al. 1966). The mutT protein was found to code for a GTPase (Akiyama et

al. 1989). MutT protein also has the ability to hydrolyze the mutagenic analog 8-oxo-7,8-dihydro-2'deoxyguanosine (8-oxodGTP), (Maki and Sekiguchi, 1992). This analog arises from the specific spontaneous oxidation of dGTP. This base has the ability to base pair with adenine or cytosine (Shibutani et al. 1991), and the base pairing with adenine will cause an A:T to C:G transversion.

MutY repair

The mutY mutator was isolated by Nghiem et al. (1988). Its phenotype displayed an increase in the mutational frequency of G.C to T:A transversions. The mutY protein codes for an adenine glycoylase that excises the adenine base from A-G mismatches (Au et al. 1989). This protein has the ability to bind to an A-G mismatch, and cleave the adenine base out leaving an apurinic site. This apurinic site is then processed by an AP endonuclease which removes the sugar phosphate backbone (Lu and Chang, 1988), leaving a small tract for resynthesis. The tract is then resysnthesized by DNA polymerase I.

MutM repair

MutM mutator was isolated by Cabrera et al. (1988). The phenotype it displays is an increase in the mutational frequency of the G:C to T:A transversion. The sequence of the mutM gene is identical to that of fpg. The mutM/fpg gene codes for an endonuclease. The target of the protein is 8-oxoG contained in the DNA (Tchou et al. 1991). Guanines in the DNA can spontaneously oxidize, forming 8-oxoG. This C - 8-oxoG pair may lead to a G:C to T:A transversion event, if the mutagenic analog base pairs with adenine at the next round of DNA replication.

Lac operon

The lac operon was the first genetic regulation system elucidated (reviewed by Miller and Reznikoff, 1978). It comprises three structural genes lacZ, lacY, lacA, and one regulatory gene lacI. These genes code for the enzymes ß-galactosidase, lactose permease, thiogalactose transacetylase, and the lac repressor, respectively. There are two sites of regulation upstream of the lacZ gene. This region comprises the operator and the promoter. What makes this set of genes an operon is that all three structural genes are regulated together and transcribed into a poly-cistronic mRNA. The lac operon is controlled at two levels; one is dependent on lactose being present in the cell, and the other is the lack of glucose. In the cell, the amount of glucose regulates the amount of cAMP (reviewed by Suzuki et al. 1989). If there is a low amount of glucose present, then there is an increase of cAMP. The cAMP will bind to the CAP protein, this molecule will then bind upstream from the lacZ gene and let the transcription of the operon occur. The opposite will occur if glucose is present. Then there is a decrease in the amount of cAMP, hence transcription will not occur. The second level is controlled by lactose. The lac repressor will bind to the operator region, in the absence of lactose. When lactose is present in the cell ß-galactosidase transforms lactose into allolactose. Allolactose can then bind to the lac repressor, changing the lac repressor's conformation. This change in conformation makes the lac repressor unable to bind the operator region, so that transcription of the operon can occur.

ß-galactosidase

ß-galactosidase is coded in E. coli by the structural gene lacZ contained in the lac operon. This disaccharidase catalyses the conversion of lactose into glucose and galactose. It is a tetrameric enzyme, which is made up of four identical subunits. The

monomers are composed of 1023 amino acids. There has been a lot of work in deducing the mechanism of enzyme activity, and the identity of amino acids involved in the reaction. ß-galactosidase gene catalyses the hydrolysis and transgalactolysis of lactose (Huber et al. 1976)

It has been proposed by Sinnott (1978) that the ß-galactosidase mechanism is analogous to that of lysozyme. There are two residues which are involved, one which acts as a general acid, and one group which is involved in the stabilizing of a gactosyl intermediate, allowing H_2O to react.

It has been proposed that tyr-503 is the general acid, which would donate a proton to the glycosidic oxygen (Ring et al. 1985, 1988). Amino acid substitutions at tyr-503, through missense codons and suppression of nonsense codons, decreased the activity of ß-galactosidae to 1% of the wild type (Cupples et al. 1988). Another site which is thought to be important is glu-461; when this site was substituted with gln there was a greater than 99% decrease in activity (Bader et al. 1988). The site was further characterized by the introduction of an additional 12 substitutions (Cupples and Miller, 1988), in which all substitutions showed less than 10% of wild type activity; all substitutions exhibited a Lac^- phenotype. The study by Cupples et al. (1990b) characterized the roles of glu-461 in the catalytic mechanism of ß-galactosidase. First, glu-461 is involved in the binding of the substrate to the enzyme. Second, the negative charge of the R-group is involved in the stabilization of the carbonium ion galactosyl-transition state of the intermediate. Third, it plays a role in the degalactosylation step of the reaction.

The characterization of asn-460 by amino acid substitution (Passi and Cupples, unpublished) found that the size of the amino acid at that site is more important that the charge. They hypothesized that the function of the amino acid at this site is only important in giving a shape to the active site.

As mentioned above, the development of site-directed mutagenesis has given us a powerful tool in the characterization of amino acids important in the catalytic activity of an enzyme. It has also given us the ability to manipulate the DNA sequence in such a way that target sequences for monitoring DNA repair can be constructed. The system in principle works, by the introduction of a specific change in the sequence, the cell becomes Lac$^-$, and only a specific mutation will revert it to a Lac$^+$ phenotype. Therefore the mutation frequency is proportional to the Lac$^+$ reversion rate.

I) The construction and preliminary analysis of a potential target sequence in the lacZ gene at site 462

As an example of this type of manipulation, site 461, because of its specificity, was used to develop a system for the monitoring of base substitution mutations (Cupples and Miller, 1989). ß-galactosidase is functional only when residue 461 is a glutamic acid. A series of six lacZ mutations was constructed, to monitor the two transition and four transversion mutations. They were constructed by changing one nucleotide at the first or second nucleotide position of the glu-461 codon. By doing this the amino acid is changed, thus rendering the cell Lac$^-$. If a mutational event occurs which changes the codon back to one coding for the amino acid glutamic acid, the cell becomes Lac$^+$. As an example of the power of this assay, the strain CC104 which contains the lacZ mutation which reverts to Lac$^+$ by a G:C to T:A transversion was used in the isolation of a strain with a high G:C to T:A mutation rate. This in turn led to the characterization of the mutY gene (Nghiem et al. 1988), part of a system involved in the repair of A:G mismatches. Another example of this type of experimental design is the construction of stain CC112, which was used in the isolation of the mutG mutator (Ruiz et al submitted) The DNA sequence of site 460, 461, and 462 was changed from AAT-GAG-TCA to

AAC-CAG-GGG. This stretch of DNA now contains the CCAGG sequence, which can be used to monitor VSP repair (Lieb, 1983). The strain is Lac⁻ because it contains a glycine codon at site 462, hence the strain is unable to use lactose as a carbon source. The second cytosine in the sequence is methylated by dcm methylase forming C_{me}CAGG. If deamination occurs at the methylated cytosine, it is transformed into a thymine. This causes a T-G mismatch; if this is not repaired, a C:G to T:A transition occurs. After replication the mutation is fixed into the DNA and the sequence becomes CTAGG. Hence, an amber codon is introduced at site 461. In the isolation of a mutG mutator, the parental strain contained a plasmid carrying a suppressor tRNA, which introduces a glutamic acid at the amber codon (TAG). Therefore cells with CTAGG are Lac⁺. Cells with defects in VSP repair have high rates of conversion of CAG to TAG, and as a result have a high rate of Lac⁻ to Lac⁺ reversion. Isolation and characterization of one strain which exhibited high Lac⁺ reversion led to characterization of a new gene involved in VSP repair. MutG is involved in the removal of the mispaired T in a T-G mismatch. The list of repair genes isolated from this assay system has been growing, for example mutM (Cabrera et al. 1988), mutC, and mutA (Micheals et al. 1990).

With the characterization of site 460 and 461 in β-galactosidase, we decided to characterize site 462 through amino acid substitution to test its potential as a mutation target sequence. Serine 462 was changed to an amber codon (TAG). There are a series of twelve suppressor strains classical suppressor strains, isolated by there ability to introduce an amino acid at a stop codon, (Miller, 1972) and artificial suppressor strains, constructed by the use of gene synthesis, (Normanly et al. 1986, 1990)). These strains carry transfer RNA's (tRNA) that will introduce an amino acid at an amber (stop) codon. We used the suppressor strains to substitute the serine at site 462 with eleven amino acids, glutamine, tyrosine, lysine, leucine, glycine, alanine, cysteine, histidine, phenylalanine, and proline. We then measured the β-galactosidase activity (Miller, 1972) and analyzed the ability of the cells containing the mutant β-galactosidase proteins to utilize lactose as the sole

carbon source. With this characterization, we could determine which DNA sequences would cause the cell to be Lac$^+$. Since 460 and 461 are quite specific in what amino acid substitution they accept, we thought 462 would be as well. If we include site 462, this will generate a stretch of 9 nucleotides, where we know which DNA sequences result in a Lac$^-$ phenotype. By manipulating the DNA sequences, we could construct Lac- lacZ mutants, which revert back to the Lac$^+$ phenotype by specific mutations

II) The effects of repair system interactions on base substitutions at site 461

The study of DNA repair involves the characterization of repair deficient mutants (mutators). The characterization consists of dissecting what mutations are caused by the mutator. This causes a problem in that other (intact) repair systems may correct the excess mutations or the excess mutations swamp may (intact) repair systems (Schaaper and Radman, 1989). This leads to two possibilities, 1) that the mutation rate observed is lower than the actual, or 2) that some types of mutations which arise are not specific for the repair system. We decided on a new approach. We studied the effect of double and triple mutators, looking at how they affect specific mutagenic events. We carried out the characterization of the effect of the interaction of mutT, mutY, and dam-directed mismatch (mutHLS) repair on three different base substitution events, G:C to A:T transitions, A:T to C:G transversions, and G:C to T:A transversions in *E. coli*.

The focus of this part of the thesis is to dissect the interaction of mutT, mutS, and mutY on specific DNA target sequences. The targets are a series of lacZ mutants (tester episomes CC101 to CC106), which monitor specific base substitution events (table 1).

Table 1 **Base substitutions monitored by CC101-CC106**

Strain	Amino Acid	Event
CC101	amber	A:T to C:G
CC102	gly	G:C to A:T
CC103	gln	C:G to G:C
CC104	ala	G:C to T:A
CC105	val	T:A to A:T
CC106	lys	A:T to G:C

III) The construction of a $(GC)_{12}$ target which monitors $(GC)_8$ frameshift event and characterization of three $(GC)_n$ dinucleotide frameshift targets at site 504

Frameshifts are detrimental to organisms. They cause damage by altering the reading frame of the DNA sequence. This alteration is due to the addition or deletion of bases. The bases in a gene are read by RNA polymerase in a specific way. The gene has a fixed starting point (start codon), and then proceeds reading in groups of three nucleotides (codons). The DNA information is said to be in frame. If there is a frameshift all the information downstream of this event will be unreadable (out of frame with the trinucleotide codon). Hence the DNA will code for a non-functional protein. There have been many models on how spontaneous frameshifts occur. The most common site at which a frameshift occurs is at monomeric runs, leading Streisinger et al. (1966) to propose the strand slippage model (figure 1). In this model, there is a misalignment of a base at monomeric runs and at dinucleotide runs. This misalignment occurs due to the base(s) looping out during DNA replication. The looped out base(s) is (are) non-pairing. The difference between the occurrence of an addition or deletion, depends on which strand contains the looped out base (template or newly-synthesized).

The focus of this part of the thesis deals with how *E.coli* prevents spontaneous frameshifts mutations at target sequences of runs of GC dinucleotides. It was necessary to construct CC127, and compare it with two other lacZ mutants CC109 (Cupples et al. 1990a), and CC126 (Passi and Cupples, unpublished). These three lacZ mutants monitor three different mutagenic events at runs of (GC) dinucleotides.

1) CC109: deletion of a (GC) dinucleotide at a run of $(GC)_5$

2) CC126: addition of a (GC) dinucleotide at a run of $(GC)_4$

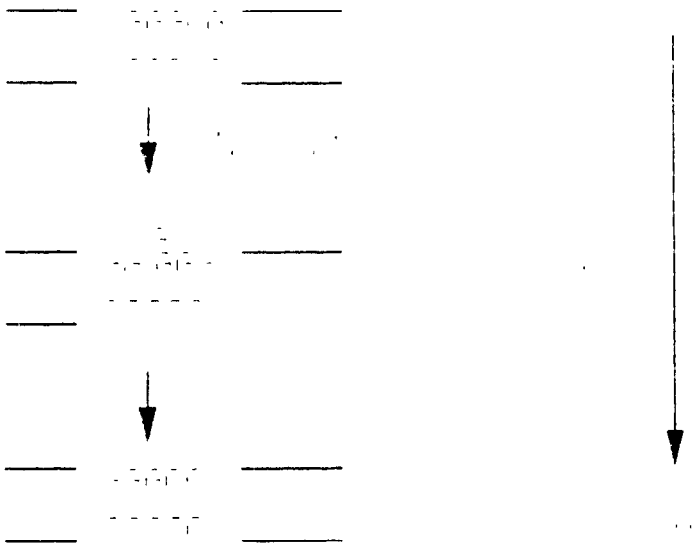
3) CC127: deletion of a $(GC)_8$ dinucleotides at run of $(GC)_{12}$

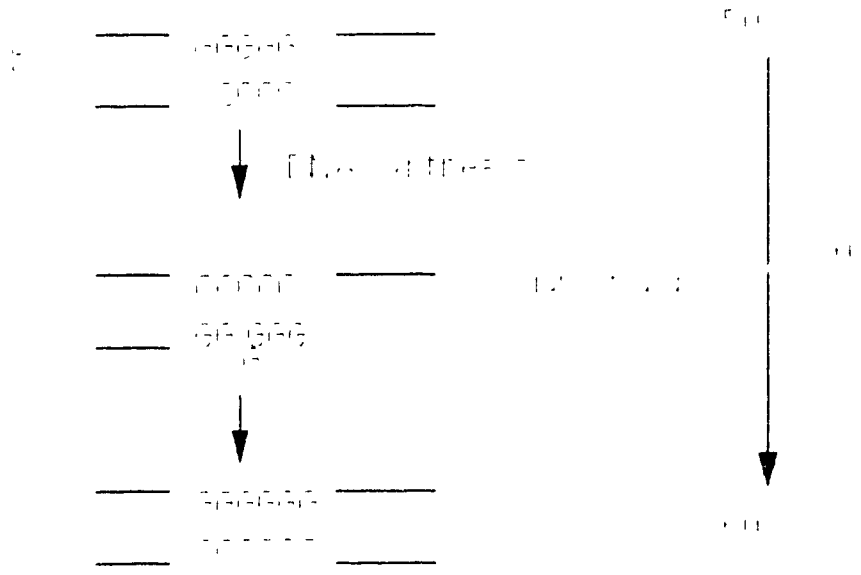
Figure 1:

Strand slippage model: model for

a) minus one base frameshifts

b) plus one base frameshifts.





We compared the spontaneous frameshift frequency of the CC127 to CC109 and CC126. To further study the mechanism, it was necessary to characterize the effect of chemical mutagens, DNA supercoiling, and methyl-directed mismatch repair on spontaneous frameshift frequency of CC109 and CC127.

This system is based on work by Cupples et al. (1990a). They developed a lacZ phenotypic assay, which monitors specific frameshift mutations. This system takes advantage of the active site residues of ß-galactosidase (glu-461 and tyr-503). Five lacZ mutants, CC107-CC111 were constructed by altering sequences around tyr-503 and glu-461, to monitor +1G, -1G, -1(GC), +1A, and -1A, respectively (table 2). In the study of Cupples et al. (1990), they analyzed the effect of chemical mutagens on the mutational frequency of the frameshift events on all five lacZ mutants, and also looked at the effect of dam-directed mismatch repair on one base frameshifts. It was found that ethyl methanesulfonate (EMS) and 2-aminopurine (2-AP) which were not thought to induce frameshifts, would in fact induce them. They hypothesized that these chemical mutagens saturated or inactivated dam-directed repair, so that frameshifts would escape repair.

Fuchs et al. (1988) hypothesizes that frameshift mutations arising from hot spots (where a large frequency of mutational events occur) are due two possible mechanisms.

1) Mutations arise as errors made when replication proceeds through a sequence which can form an unusual DNA structure.

2) Mutations arise before replication due to unusual DNA structures.

Fuchs and his coworkers constructed a series of plasmids that contained various lengths of $(GC)_n$ dinucleotides (n= 8, 9, 12, and 13). These repeats were inserted in the early part of the lacZ gene. These insertions conferred a Lac^- phenotype. If a deletion or addition event occurs that puts the lacZ gene in the proper reading frame, the strain would display a Lac^+ phenotype (except in pUC-$(GC)_{13}$ where the phenotype changes from Lac^+ to Lac^-). They found, by increasing the number of (GC) repeats, that there was an increased mutational frequency.

Table 2: **Frameshifts monitored by CC107-CC111**

Strain	Site	Target sequence	Event
CC107	491	run of six G	+1(G)
CC108	491	run of six G	-1(G)
CC109	504	run of five (GC)	-1(GC)
CC110	510	run of six A	+1(A)
CC111	510	run of seven A	-1(A)

DNA supercoiling

Supercoiling is the further twisting of the double helical DNA molecule. Supercoiling is important to the cell for many reasons. One of the most important is the compacting of the DNA in the cell (Watson et al. 1987). The cell also needs to untwist supercoiled DNA so that it can replicate, transcribe, and recombine (reviewed by Gellert, 1981). Genes that are involved in the twisting and untwisting of supercoiled DNA are called toposomerases. The toposomerases are a group of enzymes which have the ability to break and rejoin DNA phosphodiester bonds (these are bonds between the nucleotides 3'-OH and 5'-P, which form the backbone of DNA). The two genes of interest are topA (Sternglanz et al. 1981) and gyrB Dinardo et al. 1982). TopA codes for the enzyme toposomerase I. This protein catalyzes the reaction that relaxes the negative supercoiling of DNA, while gyrB codes for the enzyme which catalyzes the reaction of putting negative supercoils in the DNA. We decided to analyze the effect of torsional strain on the spontaneous frameshift events.

Chemical mutagens

Many chemical mutagens have already been extensively characterized (Miller, 1983). Understanding the effect of chemical mutagens on a specific type of mutational event enables us to understand the actual mechanisms by which the mutations occur. We have looked at the effect of four different chemical mutagens on the $(GC)_n$ targets. These four mutagens are 2-aminopurine (2-AP), ethylmethanesulfonate (EMS), 5-azacytidine and ICR-191.

2-AP is a base analog. It can be incorporated into the DNA by pairing with either a thymine or a cytosine. If incorporated opposite a thymine it may mispair with cytosine the next round of replication. This will cause A:T to G:C transitions if not repaired. If it

mispairs with a cytosine, and is not repaired before the next round of replication it may pair with a thymine, resulting in a G:C to A:T transition occurs. EMS is an alkylating agent which causes a specific transition mutation. EMS alkylates predominately at the O^6 position of guanine forming O^6-alkylguanine (Friedberg, 1985). O^6-alkylguanine mispairs with thymine, resulting in G:C to A:T transitions. EMS also can alkylate at other O positions of purines which result in the following modified bases. O^1-alkyladenine, O^3-alkyladenine, O^7-alkyladenine, O^3-Alkylguanine, and O^7 alkylguanine. 2-AP and EMS also cause frameshift mutations, possibly by saturating the dam-directed mismatch repair system so that spontaneous frameshift mutations are not repaired (Cupples and Miller, 1989). ICR-191 causes frameshifts by stabilizing looped-out bases. DNA polymerase bypasses looped out bases resulting in frameshift mutations. 5-AC causes G:C to C:G transversions, but its mechanism of action is unknown.

The understanding of how the repair systems prevent and/or repair mutations is quite important. The work presented in this thesis deals with the preliminary analysis of potential target sequence, the effect of repair system interactions on base substitutions, and the characterization of $(GC)_n$ target sequence. The ultimate goal is to add knowledge to the pool of information on DNA repair.

Materials and Methods

All molecular techniques were performed as stated in Molecular cloning a laboratory manual, Maniatis et al. (1982). All bacterial genetic methods were performed as stated in Experiments in molecular genetics, Miller, (1972). All manipulations of the lacZ mutants were performed as stated in Cupples and Miller (1989).

I) Bacterial, plasmid and phage strains

Escherichia coli strains are listed in table 3. F' episomes (CC101-111, CC126-127), and pBR329 and its derivatives were maintained in s90c. Plasmids (pGFIB-1) carrying the suppressor tRNA's were maintained in XAC. Phage f1-Z#3 (modified R229) and its derivatives were maintained in JM801. Plasmids are listed in table 4.

II) Maintenance of bacterial strains

All strains were maintained at 4°C on minimal glucose plates containing the appropriate supplements (table 5); plasmid in s90c were stored on LB plates containing the appropriate antibiotic (table 5).

III) Media

Two types of media were used, LB and minimal media A with the appropriate carbon source (glucose, lactose or galactose). LB medium (Miller, 1972) contains 10 g Bacto tryptone, 5 g Bacto yeast extract and 10 g sodium chloride in one liter of medium. Minimal medium A (Miller, 1972) contains 105 g K_2HPO_4, 45 g KH_2PO_4, 10 g $(NH_4)_2SO_4$, 5 g of sodium citrate*$2H_2O$ per liter of medium. Appropriate supplements,

Table 3: **List of Strains**

Name	sex	genotype
S90c[a]	F-	ara Δ(lacpro) thi rpsL
P90c[a]	F-	ara Δ(lacpro) thi
JV101[b]	F-	ara Δ(lacpro) thi rpsL leu::Tn10
JV102[b]	F-	ara Δ(lacpro) thi rpsL mutY::Tn10
JV103[b]	F-	ara Δ(lacpro) thi rpsL mutT::Is1
JV104[b]	F-	ara Δ(lacpro) thi rpsL partial ΔmutY
JTT1[c]	F-	pyrF strA195 gal25
RS2[c]	F-	pyrF strA195 gal25 topA
SD7[c]	F-	pyrF strA195 gal25 topA gyrB
CSH63[a]	Hfr	val[r] thi Δ(lacpro)
JV105[c]	F-	pyrF strA195 gal25 val[r] thi Δ(lacpro)
JV106[c]	F-	pyrF strA195 gal25 val[r] thi Δ(lacpro) topI
JV107[c]	F-	pyrF strA195 gal25 val[r] thi Δ(lacpro) topA gyrB
JM108[a]	F'	ara Δ(lacpro) thi rpsL, F'kan[r]
JM802[a]	F'	ara Δ(lacpro) thi rpsL, F'lacI proA$^+$ proB$^+$
CSH13-20a[a]	F'	Δ(lacpro) supE thi, F'lacZ proA$^+$ proB$^+$ with progressively larger deletions in lacZ

Table 3 continued: **List of Strains**

Name	sex	genotype
XA101[a]	F⁻	ara Δ(lacpro) gyrA metB argE⁻am rpoB supD thi
XA102[a]	F⁻	ara Δ(lacpro) gyrA metB argE⁻am rpoB supE thi
XA103[a]	F⁻	ara Δ(lacpro) gyrA metB argE⁻am rpoB supF thi
XA105[a]	F⁻	ara Δ(lacpro) gyrA metB argE⁻am rpoB supG thi
XA96[d]	F⁻	ara Δ(lacpro) gyrA metB argE⁻am rpoB supP thi
XAC[d]	F⁻	ara Δ(lacpro) gyrA argE⁻am rpoB thi
JV108[b]	F⁻	ara Δ(lacpro) thi rpsL mutY::Tn10 mutT::Is3
JV109[b]	F⁻	ara Δ(lacpro) thi rpsL mutS::Tn10
JV110[b]	F⁻	ara Δ(lacpro) thi rpsL mutS::Tn5
JV111[b]	F⁻	ara Δ(lacpro) thi rpsL mutY::Tn10 mutS::Tn5
JV112[b]	F⁻	ara Δ(lacpro) thi rpsL mutT::Is3 mutS::Tn10
JV113[b]	F⁻	ara Δ(lacpro) thi rpsL mutT::Is3 mutY::Tn10 mutS::Tn10
CC101[f]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA⁺ proB⁺ with amber codon at site 461
CC102[f]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA⁺ proB⁺ with glycine codon at site 461
CC103[f]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA⁺ proB⁺ with glutamine codon at site 461

Table 3 continued: **List of Strains**

Name	sex	genotype
CC104[f]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with alanine codon at site 461
CC105[f]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with valine codon at site 461
CC106[f]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with lysine codon at site 461
CC107[g]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with a additional A nucleotide at site 522
CC108[g]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with an addition G nucleotide at site 488
CC109[g]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with an additional GC dinucleotide at site 504
CC110[g]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with a deletion of one A nucleotide at site 522
CC111[g]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with a deletion of one G nucleotide at site 488

Table 3 continued: **List of Strains**

Name	sex	genotype
CC126[i]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with a deletion of GC dinucleotide at site 504
CC127[b]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with an additional 8 GC dinucleotide at site 504
CC128[b]	F'	ara Δ(lacpro) thi rpsL F'lacZ proA$^+$ proB$^+$ with an amber codon at site 462
AB1157[h]	F$^-$	thi-1 leu-6 thr-1 lacY1 galK2 ara15 xyl-5 mtl-1 kdgK51 proA2 his4 argE3 str31 supE44
GW3732[h]	F$^-$	thi-1 leu-6 thr-1 lacY1 galK2 ara15 xyl-5 mtl-1 kdgK51 proA2 his4 argE3 str31 supE44 mutS::Tn5
GW3731[h]	F$^-$	thi-1 leu-6 thr-1 lacY1 galK2 ara15 xyl-5 mtl-1 kdgK51 proA2 his4 argE3 str31 supE44 mutS::Tn10
PT108[i]	F$^-$	ara Δ(lacpro) thi rpsL mutY::Tn10

Table 3 continued: **List of Strains**

Name	sex	genotype
NR9082[h]	F⁻	ara Δ(lacpro) thi azi[r] mutT::ls1
NR9418[h]	F⁻	ara Δ(lacpro) thi leu::Tn10

a: Miller, (1972). b: this study c: Sternglanz et al. (1981).

d: Coulondre and Miller, (1977). e: Kleina et al. (1990).

f: Cupples and Miller, (1989). g: Cupples et al. (1990).

h: Schaaper et al. (1989) i: Michaels et al. (1990).

j: Passi and Cupples, (unpublished).

Table 4 **Plasmids**

Plasmid	Vector	Insert	Features
p473EV[a]	pBR329	EcoRV	Contains the mutY gene
pCClac127[b]	pBR329	EcoRI	Contains the lacZ gene which has has an additional $(GC)_8$ dinucleotides at site 504
pCClac128[b]	pBR329	EcoRI	Contains the lacZ gene which has an amber codon at site 462
pGHB-1[c]	pEMBL8+		Contains synthetic suppressor tRNA
pCClac4[d]	pBR329	EcoRI	Contains the lacZ gene under the control of P^+ promoter.
pCClacZI[b]	pBR329	EcoRI	Contains a modified lacZ gene that has an additional (GC) dinucleotide at site 504

a: Michaels et al. (1990). b: this study. c: Normanly et al. (1986).

d. Cupples and Miller, (1988).

Table 5: **Compounds added to plates.**

Compound	Concentration	Method of sterilization
L-leucine	20	autoclave
L-methionine	50	autoclave
L-proline	100	autoclave
Uracil	20	autoclave
L-valine	50	autoclave
P-gal	500	autoclave
X-gal	40	stock solution made in dimethyl formamide
Ampicillin	100	filter sterilize
Chloramphenicol	20	stock solution made in 100% ethanol
Kanamycin	50	filter sterilize
Nalidixic acid	30	filter sterilize
Rifampicin	100	stock solution made in 100 % methanol
Streptomycin	100	filter sterilize
Tetracyline	15	stock solution made in 50 % ethanol

0.02% magnesium sulfate, 0.005% thiamine (B_1) and 0.25 % carbon source (glucose, lactose or galactose) were added to the media. In the making of plates 12 g of Sigma agar was added to one liter of media. Maconkey lactose plates contained 50 g Maconkey lactose agar per liter of medium.

IV) Strain construction

Strains JV101 to JV103, and JV108 to JV113 were constructed through P1 mediated transductions (table 6). JV105-107 were constructed through the interrupted mating of CSH63 with JTT1, RS2, and SD7. Interrupted mating was preformed in order that the lac-pro deletion could be introduced into JJT1, SD7, and RS2 strains, while conserving parental genotypes for the topA (28 min) and gyrB (83 min) toporisomerase genes. The mating time was 8 min, after which the progeny were plated on minimal glucose plates containing 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal), proline, uracil, streptomycin and high concentration (50 µg/ml) of L-valine. The CSH63 strain was selected against by streptomycin, while JJT1, SD2, and RS7 were selected against by high valine concentration. The progeny were screened for inability to use galactose as a sole carbon source (gal25 mutation), and for uracil auxotrophy (pyrF mutation linked to topA).

V) Conjugation

E.coli strains are divided into two groups F^+ or F^-. This separation is determined by the cell's ability to form pili and on its conjugation properties. The F^+ (male) cell will donate chromosomal or episomal DNA through a protein filament, the F pilus. The F^- (female) cell will receive the DNA and as a result becomes male.

Table 6. **Summary of P1CM Mediated Transductions used in Strain Constructions**

Cross	Donor	Recipient	Selected Phenotype	Relevant Genotype	Constructed Strain
1	NR9416	S90c	tet^I	leu::Tn10 $mutT^+$	JV101
2	TT108	S90c	tet^r	mutY::Tn10	JV102
3	NR9082	JV101	leu^+	mutT::Is1 leu^+	JV103
4	TT108	JV103	tet^I	mutY::Tn10	JV108
5	GW3731	S90c	tet^I	mutS::Tn10	JV109
6	GW3732	S90c	kan^I	mutS::Tn5	JV110
7	GW3732	JV102	kan^r	mutS::Tn5	JV111
8	GW3731	JV103	tet^r	mutS::Tn10	JV112
9	GW3732	JV108	kan^r	mutS::Tn5	JV113

a) Episomal transfer

F' and F⁻ cells were grown in 2 ml LB broth at 37° C overnight with aeration. The overnight culture of F' cells was subcultured 1:100 in fresh LB broth at 37°C with aeration for 1 hr. The subculture of F' cells and stationary F⁻ cells were then mixed at a 2:1 ratio, and incubated at 37 C for 1 hr. An aliquot of 10 µl was removed and plated on selective media.

b) Interrupted conjugation (Htr)

Interrupted conjugation was performed in a manner similiar to episomal transfer. However, with interrupted conjugation, the amount of DNA transferred is dependent on the amount of time conjugation is allowed to proceed. Conjugation was interrupted by vigorously vortexing the tube. Hence, a specific amount of chromosomal DNA is allowed to transfer.

c) Plate mating

The recipient strain (F⁻) was grown on a non-selective plate (LB plate) and incubated at 37 C for 16 hr. The donor strain (F') was grown in LB overnight at 37°C with aeration, then it was subcultured 1:100 in LB and grown for 2 hr at 37°C with aeration. An aliquot of the subcultured donor (0.2 ml) was then spread onto the selective plates. The recipient stain was replica plated onto the selective plates (containing F' cells).

VI) P1 transduction

a) Making of lysogen

A heavy inoculum of the host strain was cross streaked with 100 µl temperture senstitive P1cm lysate on LB Cm plates, and incubated at 28 C (permissive) for 48 hr. The resultant colonies should be lysogens (carrying the P1cm phage). The lysogen was purified and verified by temperature sensitivity (lysis of the lysogen should occur at 37 C).

b) Making of lysate

The lysogen was inoculated in superbroth and incubated overnight at 28-31 C with aeration. The lysogen was subcultured in a flask containing 5 ml of superbroth at a dilution of 1:50, and grown at 28 C with aeration until an OD_{600} of 0.15-0.2 was reached. The subculture was then temperature shifted to 37 C until lysis of the culture occurred. The lysate was then transferred to a 15 ml polyethylene tube containing 5 ml of chloroform, and centrifuged at 5,000 RPM (IEC). The supernatant (P1 lysate) was transferred to a screw top glass tube containing 1 ml of chloroform. The P1 lysate was then stored at 4 C.

c) P1 transduction

The recipient strain was grown at 37 C overnight with aeration. The culture was transferred to a microfuge tube, and was then centrifuged for 1 min. at 16,000xg. The supernatant fluid was discarded, and the pellet was resuspended in 2 ml of MC buffer (10 mM $CaCl_2$, 1 mM $MgCl_2$). The resuspended pellet was then incubated at 37 C with

aeration for 20 min. Five culture tubes were set up designated a through e. Tube a, b, c, and d contained 100 µl of recipient cells, with b, c, d, and e containing the P1 phage at varying concentrations. Tubes a and e are controls. All tubes were incubated at 37 C with aeration for 30 min, then 0.1 ml of sodium citrate was added to stop phage growth. The tubes were further incubated at RT for 1 hr. One ml of LB was then added to the tubes. The mixture was then incubated at 37 C for 1 hr. After the incubation, the cells were plated on selective media to obtain the P1 transductants.

VII) Competent cells

A single colony was inoculated into 2 ml of LB broth, and grown overnight at 37 C with aeration. An aliquot of 0.3 ml was added to 35 ml of LB broth in a 250 ml flask, and grown at 37 C with aeration until it reached an OD_{550} of 0.2-0.25. The flask was then cooled on ice for 10 min. The culture was then decanted into a screw cap centrifuge tube, and centrifuged at 6,000xg for five min at 4 C. The supernatant fluid was then aspirated off, and the pellet resuspended in 15 ml of cold 50 mM $CaCl_2$ and left on ice for 30 min. The cells were centrifuged (same conditions as previously stated), the supernatant fluid was again aspirated off, and the pellet was resuspended in 3 ml $CaCl_2$. The cells are now competent and able to be used for transformation.

VIII) Transformation

DNA (0.1 ng) was added to the competent cells (200 µl for plasmid and 300 µl for phage transformation, RF form). The cells and DNA were mixed very gently in a microfuge tube and incubated on ice for 30 min. The microfuge tube and its contents were then transferred to a 42 C water bath and incubated for 1.5 min.

For plasmid transformation, 0.8 ml of LB broth was added to the microfuge tube and incubated at 37 C for 30-60 min (depending on antibiotic resistance conferred by the plasmid). The resulting cells were then centrifuged, the supernatant was removed, and the cells were resuspended in 0.2 ml of LB broth. These cells were then plated on the appropriate media.

For phage transformation, the competent cells and DNA were briefly cooled on ice, and then were transfered to a small glass culture tube. An aliquot of 200 ul of log phase cells from the specific host strains, 50 µl of 2% X-gal, and 3 ml of melted top agar (55 C) was added. This mixture was then vortexed and poured onto LB plates.

IX) Isolation of f1 phage particles and ssDNA template

A single colony of the appropriate host strain was inoculated in 2 ml LB broth, and incubated overnight at 37 C with aeration. The saturated culture was subsequently subcultured at a dilution of 1:25 and grown to early log phase. A single plaque was picked and placed in a microfuge tube containing 100 µl of 1X minA salts. It was then vortexed briskly for 30 sec. Alternatively a 100 µl of 10^{-2} dilution of a purified phage preparation was used. An aliquot of 200 µl of the log phase host strain was added, and incubated at RT for 10 min. The cells and f1 phage were transferred to a small culture tube containing 2 ml of LB broth, and incubated at 37°C with aeration for 4.5-12 hr.

After the incubation, 1.5 ml of the culture was transferred to a microfuge tube, and centrifuged for 1 min at 16,000xg. The supernatant was then transferred to a new microfuge tube (the pellet was either discarded or used for the isolation of dsDNA RF-DNA). To the microfuge tube 200 µl of PEG/NaCl (20% polyethylene glycol, 2.5 M NaCl) was added. The solution was mixed well, and left to stand at RT for 30 min. The

microfuge tube was then centrifuge at 16,000xg for 5 min. The supernatant was removed, but due to the viscosity of PEG/NaCl, the microfuge tube was centrifuged for 15 sec, and the remaining supernatant was removed with a Gilman Pipetman. The pellet (f1 phage particles) was then dissolved in 100 µl of TE buffer, and stored at 4 °C as a purified phage preparation.

For ssDNA isolation, an equal volume of hydrated phenol was added, and vigorously vortexed. The microfuge tube was centrifuged for 3 min at 13,000 RPM. The top layer was then transferred to a new microfuge tube, containing 200 µl of cold ethanol and 10 µl of 3 M sodium acetate. This solution was mixed well and incubated overnight at -20 °C. The next day the microfuge tube was centrifuge at 13,000 RPM for 15 min at 4 °C. The supernatant was removed and the pellet was then dried in the Speed Vac for 30 min. The dried pellet (ssDNA) was resuspended in 25 µl of H_2O, for immediate use or stored at -20 °C.

X) Site directed mutagenesis

a) Phosphorylation of the oligonucleotide

Each oligonucleotide (2 µg) was phosphorylated by 5 units of polynucleotide kinase for 1 hr at 37 °C. The reaction buffer contains 70 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 5 mM DTT, and 750 µM ATP. The mixture was then heated to 65°C to inactive the nase.

b) Mutagenesis

SsDNA f1 template (0.5 µg), mutagenic oligonucleotide (0.1 µg), and helper oligonucleotide (0.1 µg) were mixed in the annealing buffer (20 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 50 mM NaCl). This mixture was incubated at 90°C for 5 min and slowly cooled to RT (45 min) to permit hybridization between the ssDNA template and

complementary oligonucleotide.

Synthesis of circular dsDNA was induced by combining, 2.5 units of Klenow fragment of DNA polymerase, 1 unit of T4 DNA ligase, and 1 mM of each of the dNTP in reaction buffer (27 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 10 mM NaCl, 16 mM DTT, and 1 mM ATP). The mutagenesis mixture was then incubated for 1 hr at RT. The DNA was then used for transformation of competent cells.

XI) Ds DNA preparations of RF form of f1 and plasmid

In the isolation of plasmid DNA, a single colony containing the plasmid of interest was inoculated into 2 ml of LB broth with appropriate antibiotic, and incubated at 37°C with aeration overnight. In the isolation of RF form of f1 phage, an isolated plaque was added to log phase host cells, and incubated for 10 min at 37°C, then 2 ml of LB broth was added and the culture was grown at 37°C with aeration for ~4.5-12 hr.

The culture was then transferred to a microfuge tube and centrifuged at 16,000xg for 1 minute; the supernatant was then aspirated off. The pellet was resuspended in 100 µl of solution I (50 mM glucose, 25 mM Tris-HCl (pH 8), and 10 mM EDTA) and left to stand at RT for 5 min. Then 200 µl of solution II (0.2 N NaOH, and 1% SDS) was added and incubated on ice for 5 min. Then 150 µl of solution III was added and incubated on ice for 5 min. This mixture was then centrifuged at 16,000xg. The supernatant was transferred to a new microfuge tube and 400µl of hydrated phenol was added. This mixture was then vortexed for 1 min, and was centrifuged at 16,000xg for 3 min. The upper layer was carefully removed and transferred to a new microfuge tube, containing 800 µl of cold ethanol. It was then centrifuged at 16,000xg for 1 min, then the supernatant was discarded. The pellet was dried in the Speed Vac for 30 min. The dried pellet was then resuspended in 50 µl of H_2O.

XII) Single stranded DNA sequencing

Ss DNA isolated from f1-Z#3 and the mutants derived from it, were sequenced by the Sanger dideoxy method. T7 polymerase kit (Pharmacia), and ^{35}S-ATP (Dupont) was used for sequencing reactions. The subsequent sequencing reactions were electrophoresed at 70 W on an 8% polyacrylamide/ 6.8 urea denaturing gel. The DNA sequences were analyzed by autoradiography.

XIII) Agarose gel electrophoresis

All DNA samples (100 ng) were electrophoresed at 120 V in a 1% agarose gel containing 0.5 µg/ml ethidium bromide at 120 V. The electrophoresis buffer (TBE, pH 8.0) contained 0.089 M Tris-base, 0.088M boric acid, and 0.02 M EDTA. The loading dye contained 4% sucrose, 0.005% bromophenol blue, and 8% ficoll. For analysis of results UV illumination (Fisher 312nm Transilluminator) was used.

XIV) Restriction enzyme digests

DsDNA f1 phage or plasmid (100 ng) were digested at 37°C for 1 hr. Resistriction endonucleases were added in excess for digestion reactions. Also 0.08 units of ribonuclease was added to remove RNA.

BamHI buffer contained 6 mM Tris-HCl (pH 7.5), 100 mM NaCl, 6 mM $MgCl_2$, and 1 mM DTT. EcoRI buffer contained 10 mM Tris-HCl (pH 7.4), 100 mM NaCl, 10 mM $MgCl_2$, and 1 mM ß-mercaptoethanol. HincII buffer contained 10 mM Tris-HCl pH 7.5), 100 mM NaCl, and 7 mM $MgCl_2$. HindIII buffer contained 6 mM Tris-HCl (pH 7.5), 50 mM NaCl, 6 mM $MgCl_2$, and 1 mM DTT.

XV) Construction of lacZ mutants

Site-directed mutagenesis was used on f1-Z#3 ssDNA template, with mutagenic oligonucleotides to construct lacZ mutants. These circular dsDNA molecules were then used to transformed JM802 (host strain). The plaques produced were of two types, white (Lac^-)and blue (Lac^+). The white plaques were further purified and ssDNA isolated. This ssDNA was then electrophoresed to verify whether it was Lac^- due to a large deletion in the lacZ gene. If it remained the same size as the control, further analysis was performed. The ssDNA template was then sequenced to verify if the mutagenic sequence was incorporated into the lacZ gene. For corresponding changes fixed into the lacZ sequence, the RF form of the f1 (200 ng) and pBR329 (20 ng) were digested with EcoRI. EcoRI was then heat-inactivated at 65 C for 20 min. The lacZ (mutant) insert was cloned into the unique EcoRI site of pBR329 (the EcoRI site is in the $chloramphenicol^R$ gene). The putative pBR329-lacZ mutants were then introduced into competent p90c, and the cells were plated onto LB tetracycline plates (pBR329 confers resistance to tetracycline). After 14 hrs, the transformants were replica plated onto LB chloramphenicol plates. The transformants that were $chloramphenicol^S$ were streak purified and plasmid DNA was isolated. The plasmid DNA was digested with EcoRI, or HincII and electrophoresed on 1% agarose gel, to verify presence of the lacZ gene

XVI) Transfer of lacZ⁻ gene from the plasmid to the episome

Stationary phase p90c containing pBR329-lacZ mutant insert, were conjugated with log phase JM108 at a ratio of 2:1 and incubated at 37 C for 1 hr. An aliquot of 100 µl was then plated onto selective media on which only the progeny could survive (minimal glucose with tetracycline). The progeny, containing the pBR329-lacZ mutant and wild type lacZ gene on the episome, was then purified. An isolated colony was then inoculated into 2 ml LB broth with tetracycline and incubated at 37° C without aeration for 6 hr, in order that a recombination event between the plasmid and the episome could occur. This culture was then mated to stationary s90c culture, at a 2:1 ratio at 37°C for 1 hr. An aliquot of 100 µl was plated onto minimal glucose containing streptomycin and X-gal. Two types of progeny resulted from this conjugation, blue, and white colonies. The blue colonies contain the wild type lacZ gene on the episome (no recombination event occurred) and the white colonies contain an episome in which a recombination event may have occurred. The white colonies where further purified and analyzed by plate mating with the lacZ deletion strains (CSH 13-20a) on minimal lactose plates. The colonies that showed the proper growth on these plates are the putative lacZ mutants strains (the progeny that receive the wild type sequence of the mutated area of the putative lacZ mutants from the deletion strains), and were used for further study.

XVII) Spontaneous Lac$^+$ reversion test

The lactose reversion test is a method used to test the frequency of mutations at a specific site. An overnight culture of the specific lacZ mutant was grown at 37°C to stationary phase (approximately $10^{9\text{-}10}$ cells/ml). The cells were then plated onto minimal lactose plates and LB plates at different dilutions. The total number of cells were calculated from the LB plates and the number of revertants were calculated from

minimal lactose plates. The frequency is then calculated as number of Lac^+ cells per 10^8 cells.

XVIII) Checking of lacZ mutants and revertants by f1 rescue

A saturated culture inoculated with the lacZ mutation located on the episome was prepared, and a subculture of this strain was grown until log phase. An aliquot of 200 µl of this culture was then infected with f1-Z#3 (100 µl of 10^{-2}). After 10 min at RT the infected cells were transferred to a culture tube containing 2 ml of LB broth, and incubated at 37°C for 5-6 hr with aeration (so that recombination between the wt lacZ of the phage and the mutant lacZ of the episome can occur). The f1 phage particles were isolated, and then used to infect JM802. Infected cells were plated in LB top agar containing X-gal on LB plates. The white (Lac^-) plaques that arose were then purified, and the ssDNA isolated. The ssDNA was then analyzed by electrophoresis and ssDNA sequencing.

For checking the spontaneous Lac^+ revertants, the same procedure was carried out with certain modifications 1) f1 phage used was f1-Z#3 $lacZ^-_{amb461}$ instead of f1-Z#3, 2) blue (Lac^+) plaques were analyzed instead of white plaques.

XIX) Curing of episomes

An aliquot containing 1000-2000 cells of strains containing F' episomes (carrying the $proA^+$, $proB^+$, and the reverted $lacZ^+$ genes) was inoculated into 2 ml LB broth containing 75 µg/ml acridine orange. The culture was grown at 37°C with aeration, in the dark, until saturation. Aliquots of 100 µl of 10^{-4} and 10^{-5} dilutions of the culture were plated on the appropriate medium (minimal glucose containing proline and X-gal). The white colonies ($lacZ^-$) were then screened for proline auxotrophy

XX) β-Galactosidase assay in whole cells

The F' lacZ$^-_{amb462}$ mutant was transferred to the 12 tRNA suppressor strains and β-galactosidase activity was measured. The strains to be assayed were grown overnight in minimal glucose medium. They were diluted to 10^{-1} in 1X minA salts solution, and an OD_{600} was taken. A sample of 0.1 ml was further diluted (10^{-1}) in 0.9 ml of Z buffer (60 mM $Na_2HPO_4*7H_2O$, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4*7H_2O$, 50 mM β-mercaptoethanol) in a small culture tube, followed by the addition of 2 drops of chloroform and one drop of 0.1% SDS. The mixture was vortexed for 10 sec, and then transferred to a 28 C water bath. The addition of 0.2 ml of 4 mg/ml solution of ONPG (o-nitrophenol-β-D-galactose) to each tube started the reaction. The reaction was timed and stopped when a yellow colour developed, by the addition of 0.5 ml of 1 M sodium carbonate. An OD_{420} was measured, to determine the amount of o-nitrophenyl produced. The amount of β-galactosidase activity was calculated as follows:

$$\text{unit} = \frac{OD_{420} \times 10{,}000}{OD_{600} \times \text{time (min.)}}$$

XXI) Chemical mutagenesis

Strain CC109 and CC127 were treated with the chemical mutagens 2-aminopurine (2-AP), ethylmethanesulfonate (EMS), 5-azacytidine and ICR-191. 2-AP, and EMS mutagenesis was done as in Coulondre and Miller (1977). ICR-191 and 5-AC mutagenesis was done as in Calos and Miller (1981).

RESULTS

Insertion of amber and frameshift mutations into the lacZ gene cloned into f1-Z#3

F1-Z#3 is a filamentous phage containing a modified lacZ gene. The lacZ gene has two EcoRI sites flanking it; one is located at the 3' end of the gene and the other is located in the lacI gene. This piece of DNA was inserted into the unique EcoRI site of the f1 phage, R229 (Boeke 1981). The genotype of the f1-Z#3 is $lacI^-$, $lacZ^+$. Filamentous phage contain single stranded DNA. The removal of the protein coat from the phage leaves the ssDNA, and this ssDNA is used as the template for site-directed mutagenesis. The construction of two mutant lacZ genes, one with the introduction of an amber codon at site 462, and the other with an additional eight GC' dinucleotides at position 504, used the mutagenic oligonucleotides and a helper oligonucleotide as shown in table 7. After mutagenesis, competent JM802 cells were transformed with the circular dsDNA, and cells were analyzed for their ability to cleave X-gal. SsDNA was isolated, from the Lac^- f1-Z#3 mutant phage electrophoresed on a 1% agarose gel. The ssDNA from f1-Z#3 phage mutants that showed the same molecular weight as the wild type was sequenced, to confirm that the gene contained the desired mutation.

The dsDNA isolated from the mutant f1-Z#3 phage, was then digested with EcoRI, and cloned into a unique EcoRI site in pBR329. The EcoRI site is within the gene that confers chloramphenicol resistance. Competent p90c cells were transformed with the recombinant pBR329, and plated onto LB tetracylin plates. The Tet^r and Cm^s transformants were further purified. To verify whether or not the transformants contained the EcoRI fragment, the isolated plasmid DNA was digested separately with EcoRI and HincII. The digested DNA was electrophoresed on a 1% agarose gel. Digestion with EcoRI resulted in 2 bands with sizes of 4.2 kb and 3.4 kb. The 3.4 kb band corresponded to the lacZ insert and the 4.2 kb band to the vector. The digestion

Table 7: **Mutagenic and secondary oligonucleotides**

name	DNA sequence
Wild type sequence	5' GGGGAATGAATCAGGCCACGCC 3'
Amb 462 oligo	5' GGGGAATGAATAGGGCCACGCC 3'
wild type sequence	5' TCTTCATCCACGCG CGCGTACATGC 5'
$(GC)_8$ oligo	5' TCTTCATCCACGCG $(GC)_8$CGCGTACATGC 3'
Z-COOH-2 2° oligo.	3' AGTCGACTCGCGGCCAGCG 5'

with HincII resulted in 5 bands, 2.3 kb, 1.89 kb, 1.8 kb, 0.98 kb, and 0.62 kb, or 2.3 kb, 1.8 kb, 1.59 kb, 1.28 kb, and 0.62 kb. The two different groups are dependent on the orientation of the EcoRI insert. The plasmids containing the correct insert were named pCClac127 ($lacZ_{(GC)8}$) and pCClac128 ($lacZ_{amb462}$).

The lacZ mutants carried on the plasmids pCClac127 and pCClac128 were recombined with wild type lacZ carried on an F' episome. The episome contains the wild type proA, proB, lacZ, lacY, and lacA genes. The resulting Lac$^+$ mutants were then analyzed by plate mating with the lacZ deletion strains (CSH 13-20a). CSH 13-20a carry increasingly larger deletions of the lacZ gene. The ability to grow on lactose is dependent on the deletion strain carrying wild type lacZ sequence which spans the mutated area of CC127 and CC128. Growth did occur on appropriate deletion strains (strains which carry the wild type lacZ sequence at the mutated area of CC127 and CC128), hence the mutation was approximately located near codon 462 (for CC128) and 504 (for CC127) respectively. This assay does not have high resolution, since the deletion end points are not precisely known. The resulting Lac$^+$ progeny were analyzed to determine which promoter is regulating the expression of the lacZ gene. The promoter on the f1-Z#3 is the L8, while the promoter on the episome is the P$^+$. The P$^+$ promoter has eight times more activity than the L8. This was accomplished by plating them onto MacConky lactose plates and looking at the intensity of the colour, (low production of ß-galactosidase pink halo, high production deep purple). The cells that developed the purple halo were P$^+$ and therefore were used for further analyses.

By this procedure we constructed two lacZ mutants carried on F' episomes CC127 and CC128. CC127 episome contains eight additional GC dinucleotides at site 504 in the lacZ gene, while CC128 episome has an amber codon at site 462.

Checking the lacZ sequence on CC127 and CC128 episomes

The lacZ mutants on episome CC127 and CC128 were recombined with the lacZ$^+$ gene on f1-Z#3. SsDNA from the f1 phage displaying a Lac$^-$ phenotype was sequenced. The sequences were then compared to the original oligonucleotide sequence used in the site directed mutagenesis. Also, five spontaneous Lac+ revertents resulting from lacZ mutant on episome CC127 were recombined with f1-$Z_{amb\ 461}$. SsDNA from the f1 phage displaying a Lac+ phenotype was sequenced. The sequence corresponded to the wild type sequence lacZ gene in all five cases.

PROTEIN STRUCTURE AND FUNCTION

Amino acid substitution of site 462

The lacZ 462 amber mutant carried on an F' episome was transferred by conjugation into 12 suppressor strains, 6 classical and 6 artificial. These suppressor strains carry modified tRNA's which insert different amino acids at an amber codon. This suppression allows the formation of a full length protein and specifically inserts an amino acid at the amber codon at 462 site of the lacZ gene. The resulting progeny were plated on minimal glucose with X-gal. The resulting information is a preliminary indication of ß-galactosidase activity. X-gal is a chromogenic compound. When it is cleaved by ß-galactosidase, a blue color is produced, therefore the amount of x-gal cleaved is dependent on the activity of the enzyme, shown in table 8. The progeny were then tested for the ability to grow on lactose, also shown in table 8. The ß-galactosidase activity was calculated using the whole cell assay. The wild type lacZ, from strain JM102 was used as a control.

Table 8: **The suppresion of amber 462**

Amino Acid	X-gal	Lactose	Units of β-galactosidase (ONPG)	% Units of wild type
Cys	db	+	1251.8	54.2
Phe	db	+	1133.3	49.1
Ala	db	+	914.3	39.6
Glu	db	+	814.3	35.3
His	db	+	778.6	33.7
Tyr	db	+	754.8	32.7
Leu	db	+	668.2	28.9
Ser	db	+	667.0	28.8
Gly	db	+	553.6	23.9
Gln	db	+	491.5	21.3
Lys	lb	+	5.54	0.2
Pro	w	-	10.3	0.4
wt (Ser)	db	+	2309	100

intensity of color for the cleavage of x-gal, db: dark blue, lb. light blue, w: white

+: able to grow on lactose, - :unable to grow on lactose

The data demonstrates that the site can accommodate a great number of amino acid substitutions (11 of 12 amino acids substitutions resulted in a Lac$^+$ phenotype). The substitution of serine (the original amino acid) resulted in a 28.9 % activity compared to the wild type (no suppression). The difference in activity is due to the efficiency of the tRNA suppressor (Cupples and Miller, 1988). The only substitution that resulted in a Lac$^-$ phenotype was the insertion of proline at site 462. This finding was subsequently confirmed by the construction and analysis of a proline 462 missense mutation (Petropoulos, personal communication). The substitution of lysine showed a low activity in the whole cell assay (0.2% vs. 0.4% proline, due to the low efficiency of the lysine tRNA suppressor Cupples and Miller, 1988), but the strain was Lac$^+$.

FRAMESHIFT PREVENTION IN *E.COLI*

Spontaneous Reversion Frequency of LacZ mutants CC109, CC126, and CC127

These LacZ mutants monitor three different spontaneous mutagenic events. CC109 monitors the deletion of a single GC dinucleotide at a run of $(GC)_5$. CC126 monitors the addition of a single GC dinucleotide at a run of $(GC)_4$. CC127 monitors the deletion of eight GC dinucleotides at a run of $(GC)_{12}$. These mutagenic events change the phenotype from the Lac$^-$ to a Lac$^+$ (ability to grow on lactose as a sole carbon source). The spontaneous reversion frequencies were calculated as the number of Lac$^+$ per 10^8 viable cells and are shown in table 9. Table 9 depicts the differences in spontaneous reversion frequencies among the different LacZ mutants carried on F' episomes, from 0.7 ± 0.34 in CC126, to 131.1 ± 57.7 in CC109, and to 817.4 ± 225.7 in CC127.

Table 9: **Spontaneous reversion frequency of CC109, CC126, and CC127**

tester episome	event monitored	Reversion frequency $Lac^+/10^8$ viable cells
CC109	$(GC)^-$	131 ± 57.7
CC126	$(GC)^+$	0.7 ± 0.37
CC127	$(GC)_8^-$	817 ± 255.7

Effect of Chemical Mutagens on the Lac$^+$ Reversion Frequency on CC109, and CC127

Four chemical mutagens, 2-aminopurine (2-AP), 5-azacytidine (5-AC), ethyl methanesulfonate (EMS), and ICR-191 were tested for their ability to induce specific frameshifts on strains CC109, and CC127. To ensure that the conditions of mutagenesis were correct, other strains known to respond to each mutagen (increase in the Lac$^+$ reversion frequency) were employed as controls. The responding strains are the following: CC102 (G:C to A:T) for EMS and 2-AP, CC107 $(G)_6$ to $(G)_7$) for ICR-191, and CC103 (G:C to C:G). Tables 10 to 16 show the effect of the mutagens on the Lac$^+$ reversion frequency of CC109 and CC127 tester episomes.

Table 10 and 11 show a comparison of the reversion rates of CC109 and CC127 with CC102 in response to 2-AP. Treatment increases Lac$^+$ reversion frequency about 100 fold, in strain CC102. The strain CC102 monitors C:G to A:T transitions, therefore the chemical mutagen treatment was effective. The Lac$^+$ reversion frequency due to the $(GC)^-$ event (monitored by strain CC109) increased approximately 30-fold in response to mutagen treatment, but the treatment had no effect on the $(GC)_8^-$ event (CC127).

Tables 12 and 13 show a comparison of the reversion frequencies of CC109 and CC127 with CC103 in respose to 5-AC. Treatment increased the Lac$^+$ frequency (of strain CC103). The strain CC103 monitors G:C to C:G transversions. 5-AC treatment increases the Lac$^+$ frequency approximately 6-fold for strain CC109 which reverts to a Lac$^+$ phenotype by a $(GC)^-$ event, and 140-fold for strain CC127 which reverts to a Lac$^+$ phenotype by a $(GC)_8^-$ event.

Tables 14 and 15 display the Lac$^+$ reversion frequencies of strains CC109 and CC127 compared with CC102 in response to EMS. Treatment increased the Lac$^+$ reversion frequency of strain CC102 by 20,000 fold. Strain CC102 is reverted to a Lac$^+$ phenotype by a G:C to A:T transversion. EMS had no effect on the Lac$^+$ frequency for strain CC109 event, but showed a 20-fold increase for strain CC127.

Table 10: **The rate of reversion of CC109 after treatment with 2-AP**

	CC109 (Lac^+ revertants/ 10^8 viable cells)	CC102 (Lac^+ revertants/ 10^8 viable cells)
Spontaneous rate	133.3	6.4
2-AP 50 µg/ml	500.0	26.4
100 µg/ml	610.5	131.5
500 µg/ml	1736.1	524.7
700 µg/ml	3688.5	486.3
1000 µg/ml	2177.8	470.7

Table 11 The rate of reversion of CC127 after treatment with 2-AP

	CC127 (Lac$^+$ revertants/ 10^8 viable cells)	CC102 (Lac$^+$ revertants/ 10^8 viable cells)
Spontaneous rate	1883.4	< 1
2-AP 50 μg/ ml	1391.0	88.6
100 μg/ml	1694.0	281.0
500 μg/ml	1572.0	366.9
700 μg/ml	1459.0	365.0
1000 μg/ml	1404.2	608.0

Table 12: **The rate of reversion of CC109 after treatment with 5-AC**

	CC109 (Lac$^+$ revertants/ 10^8 viable cells)	CC103 Lac$^+$ revertants/ 10^8 viable cells)
Spontaneous rate	188.2	3.1
5-AC 5 ug/ml	506.2	250.0
10 ug/ml	1000.0	489.8
50 ug/ml	1256.4	3333.3

Table 13 **The rate of reversion of CC127 after treatment with 5-AC**

	CC127 (Lac^+ revertants/ 10^8 viable cells)	CC103 (Lac^+ revertants/ 10^8 viable cells)
Spontaneous rate	1427	< 1
5-AC 5 µg/ml	4391	1131
10 µg/ml	14730	3590
50 µg/ml	2.0×10^5	6036
70 µg/ml	3.0×10^5	6486
100 µg/ml	2.0×10^5	9375

Table 14: **The rate of reversion of CC109 after treatment with EMS**

	CC109 Lac$^+$ revertants/ 10^8 viable cells)	CC103 (Lac$^+$ revertants/ 10^8 viable cells)
Spontaneous rate	72.6	1.08
15 min	48.7	3106.5
30 min	61.9	20547
45 min	57.5	22259
60 min	97.8	24936

Table 15 **The rate of reversion of CC127 after treatment with EMS**

	CC127 (Lac$^+$ revertants/ 10^8 viable cells)	CC703 (Lac$^+$ revertants/ 10^8 viable cells)
Spontaneous rate	3.1×10^4	< 1
15 min	2.7×10^5	4378
30 min	5.1×10^5	28322
45 min	7.1×10^5	25000
60 min	8.0×10^4	5000

ICR-191 increased the Lac$^+$ reversion frequency 67-fold for strain CC107. Strain CC107 is known to revert to a Lac$^+$ phenotype by a +1G addition. ICR-191 increases the Lac$^+$ reversion frequency of strain CC127 which is know to reverts to a Lac$^+$ phenotype by a $(GC)_8^-$ event (Table 16).

Effect of mutations in <u>topA</u> and <u>gyrB</u> on reversion frequency of CC109, CC126, and CC127

The <u>lacZ</u> mutants, monitoring the $(GC)_7^+$, $(GC)^+$ and $(GC)_8^-$ frameshifts were transferred into strains JV105, JV106, and JV107. These strains have different supercoiling densities due to chromosomal mutations for <u>topA</u> and <u>gyrB</u>. JV105 is wild type for <u>topA</u> and <u>gyrB</u>, while JV106 is <u>topA</u>$^-$ (increase in negative supercoiling), and JV107 is <u>topA</u>$^-$/ <u>gyrB</u>$^-$ (decrease in negative supercoiling). The Lac$^+$ reversion frequency was calculated for strains containing CC109, CC126, and CC127 <u>lacZ</u> mutations to determine the effect of supercoiling on frameshifts at $(GC)_n$ sequences. The reversion frequencies are shown in Table 17.

The effect of supercoiling on the Lac$^+$ reversion frequencies of the lacZ mutants which revert to a Lac$^+$ phenotype by $(GC)^-$, $(GC)^+$ and $(GC)_8^-$ events were characterized and the following was found. CC109 ($(GC)^-$ event) showed a 2 fold increase in the Lac$^+$ reversion frequency under increased negative supercoiling, but decreased negative supercoiling had no effect. Supercoiling density had no effect on the Lac$^+$ reversion frequency, as shown by the fact that the reversion frequency of CC126 ($(GC)^+$) remained at 0.5 Lac$^+$ revertants per 10^8 viable cells. The Lac$^+$ reversion frequency of CC127 ($(GC)_8^-$) decreased approximately 3-fold under both altered supercoiling densities.

Stains JV108, JV109, and JV110 were transformed with the plasmids pCClacZ1 and pCClac127. The plasmid pCClac127 contains the <u>lacZ</u> mutation which monitors the

Table 16: **The rate of reversion of CC127 after treatment with ICR-191**

	CC127 (Lac^+ revertants/ 10^8 viable cells)	CC107 (Lac^+ revertants/ 10^8 viable cells)
Spontaneous rate	1038.7	87.7
ICR-191 2.5 μg/ml	1071.1	2317.8
5 μg/ml	5694.4	5838.4
10 μg/ml	8333.3	5225.8
20 μg/ml	5617.0	320.0

Table 17: **The effect of supercoiling on spontaneous reversion frequency of CC109, CC126, and CC127, tester episomes**

strain	genotype	reversion frequency Lac$^+$/10^8 viable cells		
		CC109	CC127	CC126
JV105	wt	32.4 ± 11.9	608 ± 230.9	0.5
JV106	topA$^-$	62.0 ± 14.8	288.8 ± 101.8	0.5
JV107	topA$^-$, gyrB$^-$.	36.8 ± 13.3	220 ± 106.9	0.5

(GC)$^-$ event, and plasmid pCClac127 contains a lacZ which monitors the (GC)$_8^-$ event. All three strains contain an F' episome which carries the proA$^+$ proB$^+$ Δ(lacZ) lacY$^+$. The episome is necessary since it carries the lacY gene (lacY codes for lactose permease, a lactose transporter) so that the Lac$^+$ reversion test can be carried out. The Lac$^+$ reversion frequencies were calculated and are shown on table 18.

The effect of supercoiling on the Lac$^+$ reversion frequencies of the (GC)$^-$, and (GC)$_8^-$ events on the plasmids were characterized and the following was found: the (GC)$^-$ event monitored by plasmid pCClac127, showed a 2.5 fold increase in the under increased or decreased negative supercoiling. The (GC)$_8^-$ event, monitored by plasmid pCClac127, was decreased approximately 3-fold under both the decreased and increase negative supercoiling conditions.

The effect of dam-directed mismatch repair on CC109, and CC127

The mutS gene codes for a mismatch recognition protein involved in dam-directed mismatch repair (methyl-dependent). This repair system works following DNA replication, since DNA is in a hemi-methylated form (newly synthesized DNA is non-methylated). When the repair system encounters a mismatch it has the ability to distinguish the correct base (methylated strand) from the mispaired base (non-methylated strand). There are a number of other genes that are involved in this repair system; they are mutL, mutU, mutH, and dam. In concert they can correct most transition mismatches and single unpaired bases in the template or in the new strand. JV110 strain carries the mutS::Tn10. Three lacZ mutations carried on F' episomes were introduced by conjugation, CC109, CC127, and CC102 (monitors G:C to A:T transitions). Table 19 shows the rate of reversion of F' episomes CC109, CC127, and CC102 in the mutS deficient background. CC102 monitors G:C to A:T transition mutations which are specifically prevented by mutHLS repair. The mutation frequency of CC102 increased

Table 18: **The effect of supercoiling on spontaneous reversion frequency of CC109 (pCClacZZ1), and CC127 (pCClac127) tester plasmids**

strain	genotype	reversion frequency $Lac^+/10^8$ viable cells	
		CC109	CC127
JV105	wt	77.1 ± 22.4	830 ± 228.6
JV106	<u>topA</u>$^-$	264.7 ± 19.2	215 ± 68.6
JV107	<u>topA</u>$^-$, <u>gyrB</u>$^-$	304.5 ± 77.9	371 ± 57.8

Table 19 **The effect of mutS on the reversion frequency of CC109 and CC127**

Strain	Genotype	Reversion frequency Lac$^+$/10^8 viable cells	
		CC109	CC127
90c	wt	131.1 ± 57.7	817.4 ± 255.7
JV109	mutS::Tn10	287.2 ± 29.2	2322.8 ± 880.2

60-fold in the mutS⁻ strain, while CC109 and CC127 showed a 2.0 and 2.8 fold increase, respectively.

REPAIR SYSTEM INTERACTION IN *E.COLI*

In this series of experiments, the interaction of repair systems on specific base substitution events are examined. Seven repair deficient mutants were constructed for this purpose. There are three which are single mutants, three which are double mutants and one which is a triple mutant. These seven repair deficient strains are derived from three repair systems : mutT, mutY and dam-directed mismatch repair.

The seven repair deficient strains were constructed using P1 transductions. The strains in question are the following:

1) mutY⁻,

2) mutT⁻,

3) mutS⁻,

4) mutT⁻/mutY⁻,

5) mutT⁻/mutS⁻,

6) mutY⁻/mutS⁻,

7) mutT⁻/mutY⁻/mutS⁻.

We then introduced into these backgrounds the six tester episomes (lacZ mutants) CC101-106, which monitor six base substitution events. The analysis was carried out in two parts. Initially the mutational spectrum of each repair deficient strain was characterized, (how each specifically effects each of the six base substitution events), and secondly a new series of experiments which looked directly at the interaction of the repair systems on the occurrence of specific mutations.

The lacZ mutants (tester episomes) CC101-106 have different point mutations at site 461 of lacZ gene which render β-galactosidase non-functional (the cell cannot utilize lactose as a carbon source). If a specific base substitution occurs that restores the GAG (wild type sequence) codon at site 461, the cell will revert to Lac$^+$ phenotype, and the mutational frequency can be calculated.

Mutational spectra

The initial characterization was carried out to ensure that each mutator locus was generating its characteristic mutational spectrum.

As shown in the Table 21, the lack of mutY repair causes G:C to T:A transversions. LacZ mutant CC104 is used to monitor this event since it is known to revert to a Lac$^+$ phenotype by this specific transversion event. The lacZ mutant CC104 showed a 60 fold increase in the Lac$^+$ reversion frequency in the mutY deficient strain when compared to the wild type strain (s90c, table 20). Lack of mutY repair had no effect on the other base substitution events as shown by the low level of Lac$^+$ reversion with other lacZ mutants. This agrees with the function of the protein coded by mutY which is involved in the prevention of G:C to T:A transversion events.

The mutational spectrum of the strain defective in mutT repair (table 21) shows a dramatic increase in the A:T to C:G transversions as monitored by lacZ mutation CC101. In comparison to the wild type background, there was a 10,000-fold increase in the Lac$^+$ reversion frequency. This result is in accordance with the known phenotype conferred when this gene is defective.

In the mutational spectrum (table 21) of the mutS$^-$ strain (deficient in dam-directed mismatch repair) there was an elevation in Lac$^+$ reversion frequency in the lacZ mutant CC102, which is known to revert to a Lac$^+$ phenotype by a G:C to A:T transition event, and a slight increase in the A:T to G:C transition event, monitored by lacZ mutant

Table 20: **Mutational frequency of s90c**

Event	s90c (Lac$^+$ revertants / 10^8 viable cells)
A:T to C.G	0.5 ± 0.2
G:C to A:T	<0.5
C:G to G:C	0.7 ± 0.7
G.C to T:A	0.9 ± 0.8
T.A to A:T	1.7 ± 0.8
A:T to G:C	<0.5

Table 21: **Mutational frequency of single mutators**

Event	Strain		
	mutL⁻	mutY⁻	mutS⁻
	(Lac⁺ revertants / 10^8 viable cells)		
A:T to C:G	9315 ± 3200	3.2 ± 1.1	5.6 ± 4.4
G:C to A:T	1.4 ± 0.5	0.9 ± 0.8	31.2 ± 5.28
C:G to G:C	<0.5	<0.5	<0.5
G:C to T:A	1.8 ± 1.0	45.6 ± 5.2	7.6 ± 3.4
T:A to A:T	<0.5	1.1 ± 1.3	<0.5
A:T to G:C	<0.5	<0.5	8.3 ± 3.7

CC106. The increase was 40 and 16 fold, respectively

The mutational spectra in the double mutants were characterized. In table 22, the mutational spectrum mutY/mutT double mutant was analyzed. This double mutant shows an increase in the Lac$^+$ reversion frequency in the lacZ mutants CC104 and CC101. There was no increase in the Lac$^+$ reversion frequency for the other lacZ mutants, which monitor other base substitution events

The mutational spectrum of the mutT/mutS double mutant shows an increase in Lac$^+$ reversion frequency in the strains carrying the lacZ mutants CC101 and CC102 (table 22). This complies with the phenotypes of both mutators, mutT increases A:T to C:G transversions and mutS$^-$ increases G:C to A:T transitions

The last double mutant looked at was mutY/mutS. This strain shows an increase in the Lac$^+$ reversion frequency for the lacZ mutants CC102 and CC104 (table 22). CC102 monitors G:C to A:T transitions, and the increase was due to the lack of dam directed mismatch repair, while CC104 monitors the G:C to T:A transversions, which is due to the lack of mutY repair.

The mutational spectrum of the triple mutant mutT/mutY/mutS is shown in table 23. It shows an increase in the Lac$^+$ reversion frequencies in both transversion events and G:C to A:T transition event, monitored by lacZ mutants CC101, CC104, and CC102

Analysis of repair system interaction on base substitution events, G:C to A:T, G:C to T:A, and A:T to C:G

In this series of experiments we wanted to analyse the specific interactions of the repair deficient strains. This is due to the fact that the interaction of repair systems may increase or decrease the relative mutational frequencies. We decided to look at three types of base substitution events:

Table 22. **Mutational frequency of double mutators**

Event	Strains		
	mutT⁻/mutY⁻	mutT⁻/mutS⁻	mutY⁻/mutS⁻
	(Lac⁺ revertants / 10^8 viable cells)		
A.T to C.G	776.2 ± 903.4	3388 ± 2219	<0.5
G.C to A.T	0.7 ± 1.4	28.7 ± 6.6	58.5 ±9.2
C.G to G.C	<0.5	<0.5	<0.5
G.C to T.A	19.4 ± 10.2	1.5 ± 0.7	31.4 ± 12.9
T.A to A.T	0.6 ± 0.8	0.6 ± 0.7	<0.5
A.T to G.C	<0.5	7.5 ± 6.8	15.2 ± 8.2

Table 23. **Mutational frequency of the triple mutator**

Event	mutT⁻/mutY⁻/mutS⁻ (Lac⁺ revertants / 10^8 viable cells)
A:T to C:G	3075.8 ± 784.6
G:C to A:T	36.1 ± 8.8
C:G to G:C	<0.5
G:C to T:A	50.9 ± 23.1
T:A to A:T	<0.5
A:T to G:C	<0.5

1) A:T to C:G transversion (CC101)

2) G:C to A:T transition (CC102)

3) G:C to T:A transversion (CC104)

We decided on these three base substitution events since they are known to be affected by the repair systems mutT, mutY and dam-directed mismatch repair, since characterization of the mutational spectra of the repair deficient backgrounds confirmed their effect. We wanted to analyze the interaction and dissect the effect of each mutator has on each specific base substitution event.

The first mutational event characterized was the G:C to A:T transition. This event has been shown to be prevented by dam-directed mismatch repair. LacZ mutant CC102 monitors this mutational event. We analyzed the mutational frequency of this event in all seven repair deficient strains. There was an increase in mutational frequency in all the mutants strains containing mutS$^-$, approximately 30 Lac$^+$ per 10^8 viable cells (figure 2), with vary little variation. This indicates that the G:C to A:T event is normally prevented by the mutHLS system and that mutT or mutY have no direct or indirect effect.

The second base substitution event characterized was the G:C to T:A transversion. This event has been shown to be prevented by the mutY repair system through processing of the A-G mismatches. This transversion event was monitored by lacZ mutant CC104. A dramatic increase was observed in all the strains which are mutY$^-$ of 40-50 Lac$^+$ per 10^8 viable cells with little variation (figure 3). The cells that carried the mutY$^+$ showed no increase, therfore mutT and mutS do not play a role in the prevention of this base substitution event.

The third mutational event characterized was the A:T to C:G transversion. It was analyzed because it is prevented by the mutT repair system. This event was by far the most exciting. This event is monitored by the CC101 tester episome (figure 4). This mutational event was not affected in mutY$^-$, mutS$^-$, and mutY$^-$/mutS$^-$ strains; the mutational frequency remained at wild type levels. In contrast, in a mutT deficient

Figure 2:

Characterization of the effect of deficiencies in one or more repair systems on the G:C to A:T mutations monitored by CC102. The mutational frequency is expressed as the number of Lac^+ revertants/ 10^8 viable cells.

CC102
GC → AT

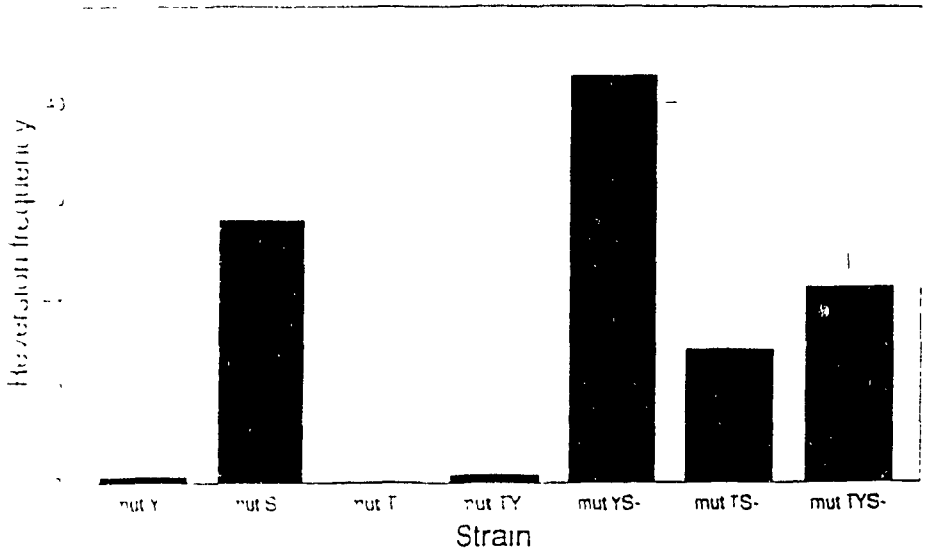


Figure 3:

Characterization of the effect of deficiencies in one or more repair systems on the G:C to T:A mutations monitored by CC104. The mutational frequency is expressed as the number of Lac^+ revertants/ 10^8 viable cells.

CC104
CG →AT

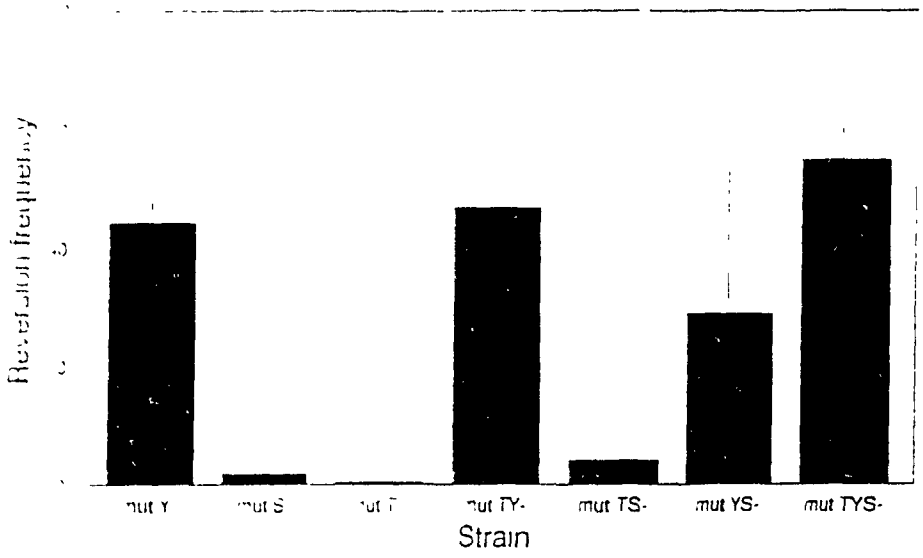
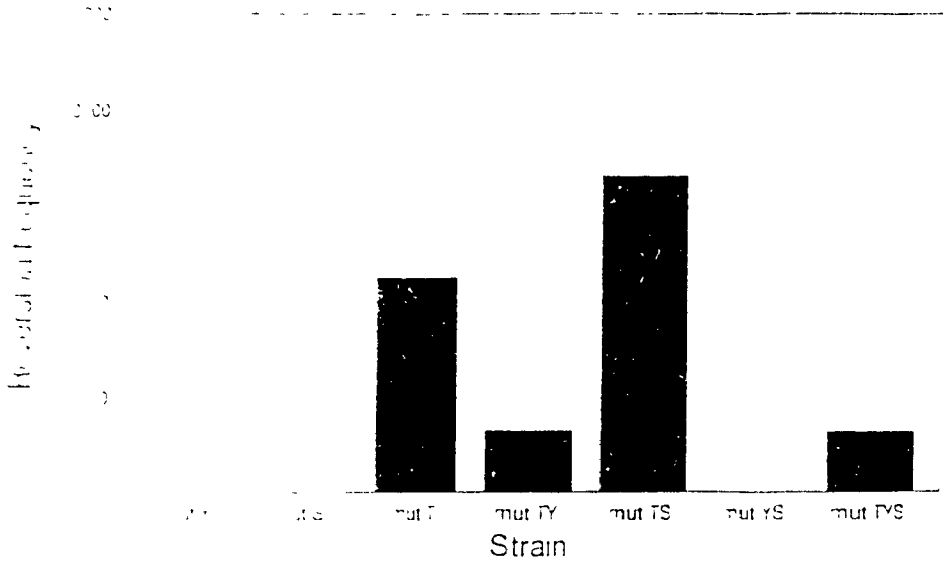


Figure 4:

Characterization of the effect of deficiencies in one or more repair systems on the A:T to C:G transversions event, monitored by CC101. The mutational frequency is expressed as the number of Lac^+ revertants/ 10^8 viable cells

CC101
AT →CG



background. mutY and mutS repair systems had an effect on this transversion event MutT⁻ strains carrying a functional mutY gene product had a 2.5 fold higher mutational frequency when compared to mutT mutY⁻. The mutT strain carrying mutS versus mutS⁺ also showed a higher mutational frequency (approximately 2.5 fold). From the data it appears as though mutY repair is mutagenic in a mutT background and that mutS may play a small role in the prevention of the A:T to C:G transversions.

To further characterize the role of mutY and mutT on the LacZ mutant CC101, we transformed two repair deficient backgrounds containing the tester episome CC101, mutT⁻/mutY⁺ and mutT⁻/mutY⁻ with p473EV (pBR329 derivative with the cloned mutY gene) and pBR329. We found mutT mutY⁺ with p472EV or with pBR329 and mutT⁻/mutY⁻ carrying p473EV have a higher reversion frequency (6.2×10^{-4}, 1.1×10^{-4}, and 7×10^{-5}, respectively) compared with mutT⁻/mutY⁻ carrying pBR329 ($<5.1 \times 10^{-6}$). The mutY gene product increases the Lac⁺ reversion frequency. This is further evidence that mutY is mutagenic (increases A:T to C:G tranversions) in a mutT deficient background.

DISCUSSION

The objective of this thesis was the construction and characterization of potential target sequences in the lacZ gene (*E.coli*) for monitoring specific mutations. This project was divided into three sections, as follows

I) the characterization of site 462 for potential target sequences.

II) construction and analysis of frameshift target sequence $(GC)_{12}$ which monitors a deletion event of $(GC)_8$ dinucleotides.

III) analysis of repair system interactions on three target sequences, which monitor,

a) A.T to C.G

b) G.C to A.T

c) G.C to T.A

I) Site 462 and its potential as a target sequence

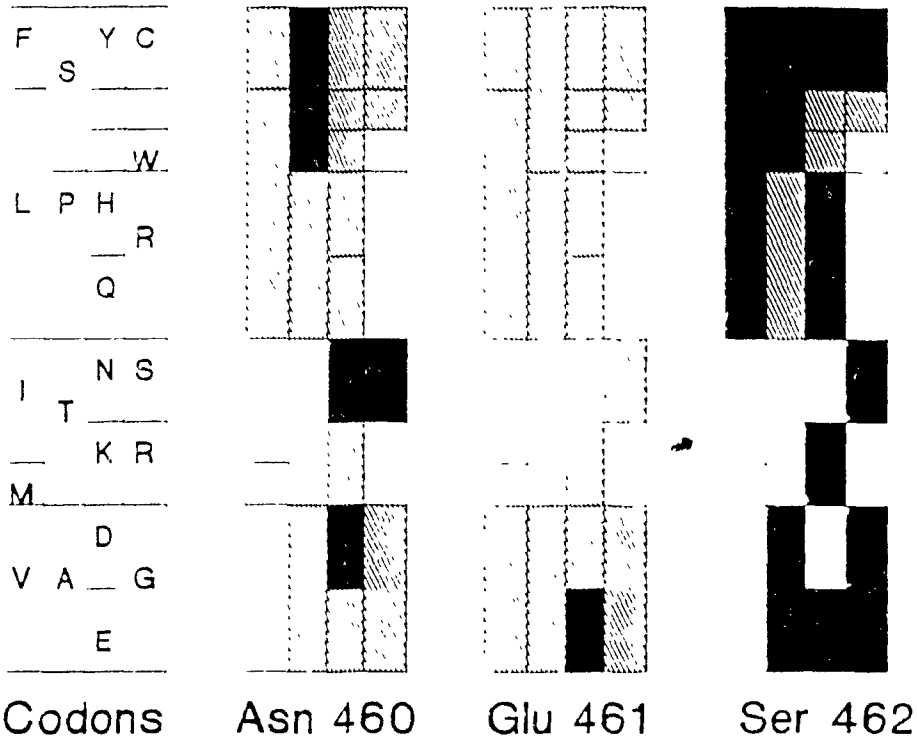
Since site 460 and 461 of the lacZ gene have been characterized to determine which amino acids will result in a Lac^+ phenotype (Passi and Cupples, unpublished, Cupples and Miller, 1988), we decided to characterize site 462. Site 462 codes for the amino acid serine. We introduced an amber codon (TAG) at site 462 through site directed mutagenesis, and then by using suppressor tRNAs we introduced 12 different amino acids. Eleven of the twelve amino acids introduced at the site 462 resulted in a Lac^+ phenotype. Only the proline substitution resulted in a Lac^- phenotype. It has been found that amino acid glu-461 is a catalytic residue in β-galactosidase which is involved in β-cleavage of lactose into glucose and galactose. It was found that all amino acid substitutions at this site formed a non functional β-galactosidase protein (Cupples and Miller, 1988). The

work of Passi and Cupples (unpublished) found that site asn-460 seems to be important in forming a three-dimensional shape to the catalytic site. In the amino acid substitutions at site 460, only the size of amino acid was important and not the charge for β galactosidase activity. The reason for the low β-galactosidase activity for the proline substitution at site-462, which resulted in a Lac^- phenotype is probably that proline is an imino acid and not an amino acid. The proline side chain is bound to both the nitrogen and the α-carbon molecule; this cyclic structure influences the protein structure. This amino acid is known to place kinks in the tertiary structure of the protein. The proline substitution may cause an added kink in the tertiary structure of the β-galactosidase protein, and may move the glu-461 out of place, so that it cannot interact with the substrate.

With the characterization of sites 460, 461, and 462, we can now look at the possible manipulations of this sequence (figure 5) to construct novel target sequences. The proline codon is a CCN and the adjacent codon is GGC. This allows us to construct a target sequence CCAGG, for monitoring VSP repair (Lieb, 1983). Petropoulos and Cupples (personnal communication) used site directed mutagenesis to insert the CCA proline codon at site 462. This was found to be target sequence for VSP repair. The target sequence was mutated preferentially to a CTAGG sequence in both VSP and VSP^+ strain. The CTA codon codes for leucine which confers a Lac^+ phenotype. Using a VSP^- strain (mutG) they found an increase in the mutational frequency of approximately 100 fold compared to the wild type strain. The development of this assay system allows a simpler method of monitoring VSP repair status in comparison of the original assay system (Ruiz and Cupples, submitted). The original assay system (strain CC112) which was used in the isolation of mutG, contains a suppressor tRNA gene on a plasmid. The parental strain was the XA203 which contains the tRNA suppressor gene glnA. This gene codes for a tRNA which will introduce a glutamic acid at the amber codon. Therefore the phenotypic assay (CC112) is not as

Figure 5:

Amino acids, and codon sequences at site 460, 461 and 462 of the lacZ gene and resulting phenotype. Solid: Lac^+. Hatched: Lac^- and Blank: untested. The letters denote the amino acids (standard symbols) and their accompanying codons.


F Y C
S
W
L P H
R
Q
I N S
T
K R
M
D
V A G
E
Codons
Asn 460
Glu 461
Ser 462

direct lacZ+ reversion of the C.G to T.A transition requires amber suppression. Also the suppressor tRNA gene is located on a plasmid, which causes difficulties when introducing other plasmids which contain the same origin of replication.

II) $(GC)_n$ Frameshifts

The lacZ mutant CC127 was constructed to study the mechanism of frameshift mutations, specifically how they occur and/or how they are prevented. This work deals specifically with frameshifts occuring at $(GC)_n$ dinucleotide runs. Passi and Cupples (unpublished), and Cupples and Miller (1989) constructed two lacZ mutants CC126 and CC109, which monitor a (GC) dinucleotide addition event at a run of $(GC)_4$ and a (GC) dinucleotide deletion event at a run of $(GC)_5$, respectively. The spontaneous reversion frequencies of these lacZ mutants were compared to the spontaneous reversion rate of lacZ mutant CC127 that monitors 8(GC) dinucleotide deletion at a run of $(GC)_{12}$. The spontaneous reversion frequency differed drastically for the mutational events, 0.7 ($(GC)^+$), 131.1 ($(GC)^-$), and 817.4 ($(GC)_8^-$) The comparison between CC126 and CC109 indicates that the addition frameshift occurs at a very low frequency versus the deletion. If we look at the strand slippage model (Streisinger et al. 1966) the occurrance of the looped out bases should equally happen on either strand (template or newly synthesized). If the slippage model does apply to both the $(GC)^+$ and $(GC)^-$ events, we expect that the resulting frequencies should remain in the same magnitude. To further support this model, Cupples et al. (1990) have shown that the +1G and the -1G at runs of G's events occur at a Lac^+ reversion frequency of 30 and 54, and that the +1A and the -1A at runs of A's events occur at a Lac^+ reversion frequency of 4.8 and 14 Lac^+ revertants per 10^8 viable cells. All comparisons are in the same order of magnitude. We suggest two possibilities why this did not occur with the $(GC)^+$ and $(GC)^-$ events. First, the slippage model proposes that the nucleotide assumes a position during

polymerization that neither interferes with nor instructs the incorporation of its neighboring base. This can occur with 1N, which will result in additions and deletions, as shown in the Bebenek and Kunkel (1990). But do two nucleotides assume a position where they do not interfere with the DNA polymerase? The second consideration is that the target sequences are not the same size (4 versus 5 GC repeats for CC126 and CC109, respectively), and this might explain the dramatic differences in spontaneous reversion frequency. The question arises, why is the mutational rate of $(GC)^+$ so low? Is it that the $(GC)^+$ event is stringently repaired or prevented by repair systems or is it that the premutagenic events just do not occur. We decided to screen CC126 for mutators, strains that would show an increase in the mutational frequency. 60,000 colonies from cultures treated with EMS were screened using the papillation assay (Nghiem et al. 1989) to no avail, no mutators were found. So we propose that the premutagenic event does not occur at a high frequency.

If we now compare the mutational frequency of CC127 ($(GC)_8^-$) to CC109 ($(GC)_5^-$) there is an increase of approximately 7-fold. Fuchs et al. (1988) also found that increasing the target sequence size led to an increase in the mutational frequency. The pUC-$(GC)_n$ assay system monitored both (GC) additions and deletions. Each tester plasmid could monitor more than one specific $(GC)_n$ event (both addition and deletions (n= +1,-2,-4,-6, etc.)). Fuchs found that the addition event occured rarely. Fuchs et al. (1988) hypothesised that the frameshifts are probably due to unusual DNA structures formed during DNA replication or afterwards. To test this hypothesis we constructed two topoisomerase mutants that would change supercoiling density in the cell (one increases, the other decreases, negative supercoiling). By changing the environment of the DNA, we may be able to change the degree of the formation of these unusual DNA structures (hairpins, cruciforms or Z-DNA). We initially tested the Lac^+ reversion frequency of the lacZ mutants in F' episome and then in a plasmid. We decided to test the effect of supercoiling on the tester plasmids due to the smaller size (a 7 kb versus 50

kb), since it might show a greater effect. There was no specific pattern in the mutational frequency observed (e.g. increase negative supercoiling, increase in the mutational frequency). The changes observed could be due to changes of gene expresion, and DNA replication, which have been shown to be affected by supercoiling.

The effect of chemical mutagenesis was characterized on the lacZ mutants CC109 and CC127. The characterization of mutagenic events monitored by CC126 $((GC)^+)$ was shown previously to be unaffected by treatment with 2-AP, EMS and ICR-191, (Passi and Cupples, unpublished).

Chemical mutagen ICR-191 is of interest because it has been shown to cause frameshift mutations, possibly by stabilizing a looped out intermediate (a non-paired base). Cupples et al. (1990a) found an 6-fold increase in the rate of reversion of CC109 $((GC)^-)$ with ICR-191 treatment. This is not as dramatic as the increased Lac^+ reversion shown in episome CC107 (+1G event), 1000 fold. ICR-191 treatment increased the reversion frequency of CC127 $((GC)_8^-)$ by 8-fold, hence it seemed ICR-191 has the same effect on CC109 and 127 by an increasing the rate of reversion of only 6 and 8 fold, respectively. It seems that the $(GC)_n$ target is not specific for the mutagen ICR-191.

The chemical mutagen 2-AP, works by saturating mutHLS repair, so that unpaired bases will escape repair and lead to addition or deletions. 2-AP treatment had no effect on CC127, but did increase the mutational frequency of $(GC)^-$ by 27-fold. Cupples et al. (1990a) found 2-AP increased the +1G event 300 fold. In comparing these to the 27-fold increase experienced by CC109, the $(GC)_n$ target is relatively unaffected by 2-AP treatment. EMS is an alkylating agent that will alkylate guanine to O^6-alkylguanine. O^6-alkylguanine leads to direct mispairing with thymine (G:C to A:T transition); like 2-AP it may swamp mutHLS repair and therefore allow unpaired bases to form frameshifts. EMS showed little effect on the $(GC)_n$ targets: no effect on CC109 and only a 20-fold increase on CC127. One problem with CC127 is that it has such a high spontaneous reversion rate, that the treatment with the EMS increases the base line spontaneous

mutational frequency by 10 fold. EMS treatment of lacZ mutant CC102 showed a 20,000 fold increase in the mutational event of the G:C to A:T transition. The twenty fold increase monitored by CC127 is quite small in comparison for this powerful mutagen.

The manner in which 5-azacytidine causes mutations is unknown. 5-AC increased the mutation rate 6-fold of the event monitored by CC109, and increased the reversion event monitored by CC127 by 140 fold, it is not in the same magnitude as the mutational frequency increased by 1000 fold increase seen in the G:C to C:G transversion event monitored by CC103, but it was dramatic.

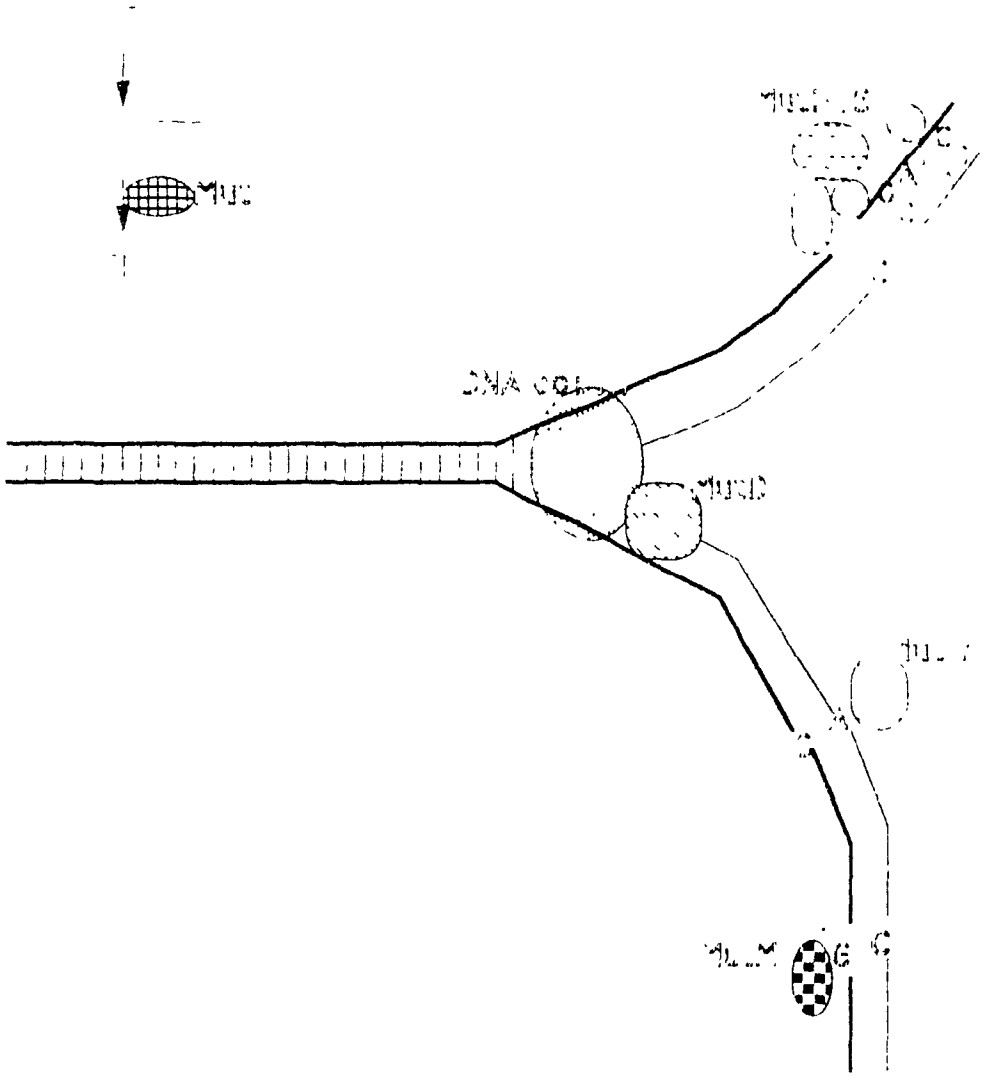
In summary, the effects of the chemical mutagens leads us to believe that, I) (GC) and the $(GC)g^-$, were not the direct targets of any of the mutagens tested, II) that there is a slight indirect effect, which may be due to the stress the cell is dealing with, III) there is indirect evidence that mutHLS repair plays no role in the prevention of these frameshift events. The effect of dam-directed mismatch repair on CC109 and CC127 was determined by introducing the episomes by conjugation into a mutS deficient background. The lack of repair caused a 2-fold and 3-fold increase in the rate of reversion for the $(GC)^-$ and $(GC)g^-$ events, respectively. In contrast, the CC102 episome (which monitors the G:C to A:T transition) shows a 60-fold increase in the rate of reversion. Hence we show here that mutHLS repair is not involved in the prevention of these two mutational events.

III) The interaction of repair systems on specific target sequences

In this section of the study we looked at the interaction of the repair systems mutT, mutY and mutHLS (dam-directed mismatch repair) on specific base substitution events A:T to C:G transversions, G:C to A:T transitions, and G:C to T:A transversions. The repair systems are depicted in figure 6, shown is their arrangement around the DNA

Figure 6:

Distribution of repair systems in relationship to the DNA replication fork in E.coli: mutT, mutD, mutHLS, mutY, and mutM. MutT functions before replication. MutD functions with the DNA polymerase III. MutHLS, mutY, and mutM function after DNA replication



polymerase. We constructed a series of seven strains of single, double and triple mutators which are deficient in every combination of the above mentioned repair systems.

The analysis of the G:C to A:T transversion event found that it was prevented solely by the action of mutHLS repair. Of the seven strains tested only the mutS deficient strain showed an increased in the Lac$^+$ reversion frequency (approximately 60-fold in comparison to the strains carrying the mutS$^+$). MutT and mutY repair systems played no role in prevention or repair of this mutagenic lesion. This agrees with previous studies which have shown that mutT (Schaaper et al. 1989) and mutY (Nghiem et al. 1988) do not to affect G:C to A:T transition events.

The analysis of the A:T to C:G transversion was the most interesting by far. We have shown that mutY$^+$ has a mutagenic effect in a mutT$^-$ background, on the A:T to C:G transversion event. MutY$^+$ compared to mutY$^-$ increased the event 3.5-fold, and the reintroduction the mutY gene (on a plasmid) in a mutT$^-$/mutY$^-$ strain increases mutagenic activity still further. Other groups have previously mentioned an overall decrease in mutational rate in mutT$^-$/mutY$^-$ strains compared to mutT$^-$/mutY$^+$ of approximately 2 fold (Schaaper et al. 1989), but have not determined the relative contributions of the repair systems involved. We are hypothesising that the effect of mutT and mutY on the A:T to C:G transversion event is related directly to when these repair system function relative to the DNA replication fork. MutT functions before the replication fork, while mutY works after. Also the confirmation that mutY is mutagenic in a mutT$^-$ background. MutT protein codes for a nucleoside triphosphatase which specifically degrades 8-oxo-7,8-2'-dGTP (8-oxodGTP). In comparing dGTP to 8-oxodGTP (mutagenic analog), dGTP does not pair with A to any significant extent, while 8-oxodG can pair equally well with A or C (Shibutani et al. 1991). Therefore, in a mutT$^-$ strain the insertion of this mutagenic analog during replication would cause an increase in T:A to G:C transversions. We have shown a 10,000 fold increase in the

mutational frequency. It has been hypothesised that mutY$^+$ is mutagenic in a mutT background (Au et al. 1988). MutY codes for an adenine glycosalase which removes the A from A-G mismatches, and has no apparent innate strand specificity. After replication the mutY protein will seek out A-G mismatches and repair them. It has been hyposthesised that the A-G mismatch corrected by mutY is dependant on the conformation of the bases (e.g. anti-syn conformation of the base by the rotation around the glycosidic bond in comparison to the sugar phosphate backbone) (Au et al. 1989). The A-*G (*G = 8-oxoG) mismatch formed in a mutT$^-$ will be repaired by mutY protein to form C:*G, at the next round of replication; the resulting progeny will carry one C·G and C:*G base pair at the mismatch site. Alternatively, if mutY is deficient only one of the two resulting progeny will contain a C·G base pair while the other will be the original A:T base pair (figure 7).

We also studied the effect of mutS and mutT on the G:C to A:T mutation. In the past, studies utilized heteroduplex analysis to determine that mutHLS system has no effect on the repair of A-G mismatches. Maki and Sekiguchi (1992) found that the substrate for mutT protein hydrolysis is 8-oxodG (*G). The question arises, is the A-*G mismatch a target for mutHLS repair? We found that mutT$^-$/mutS$^-$ in comparison to mutT$^-$/mutS$^+$ showed an increase of 25 % in mutational frequency. What might occur is that both mutS and mutY are attempting to repair the mismatch. If we compare mutT/mutY$^+$/mutS$^+$ and mutT$^-$/mutY$^+$/mutS$^-$ there was a an increase in the mutational frequency from 11,241 to 16,555. We feel that mutHLS repair system is removing the G analog *G from the A-*G mismatches restoring it back to an A·T base pair. MutY therefore does not have a mismatch to process, so it can not increase the mutational frequency, by processing the A-*G lesion to C-*G (figure 8). Schaaper and Dunn (1989), using an in vivo assay (trp$^-$ reversion assay) to monitor the A·T to C·G transversion event, found that mutT$^-$/mutS$^-$ strain in comparison to a mutT/mutS$^+$ strain that some sites there would be an increase in the mutational frequency, while other sites

Figure 7:

The effect of mutY in a mutL deficient background

a) The wild type sequence

b) The mispairing of 8-oxodGTP with A, due to the mutL deficiency

c) Functional mutY will process the mismatch as if the A is the mispaired base, therefore, fixing the mutation into the DNA

d) MutY$^-$ background will not process the mismatch therefore the mutation gets fixed into the DNA on the next round of DNA replication

(8-oxo-7,8-dihydro-2'deoxyguanosine ^{8}G)

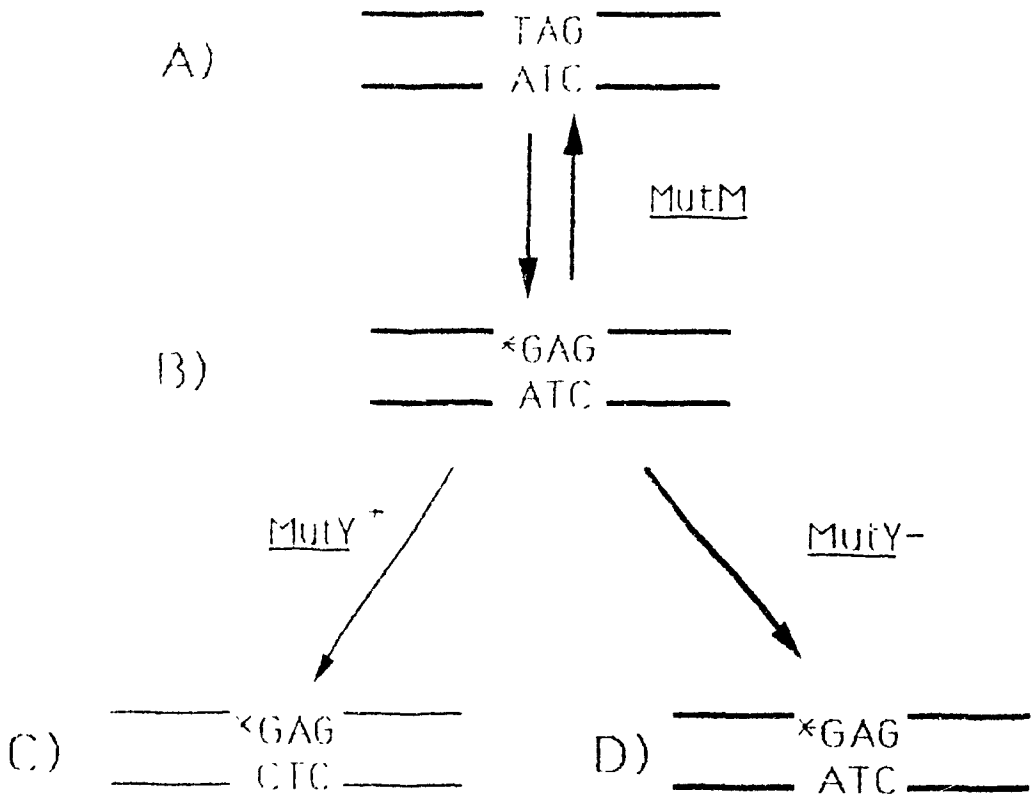

A)
TAG
ATC
MutM
B)
*GAG
ATC
MutY+
MutY-
C)
*GAG
CTC
D)
*GAG
ATC

Figure 8:

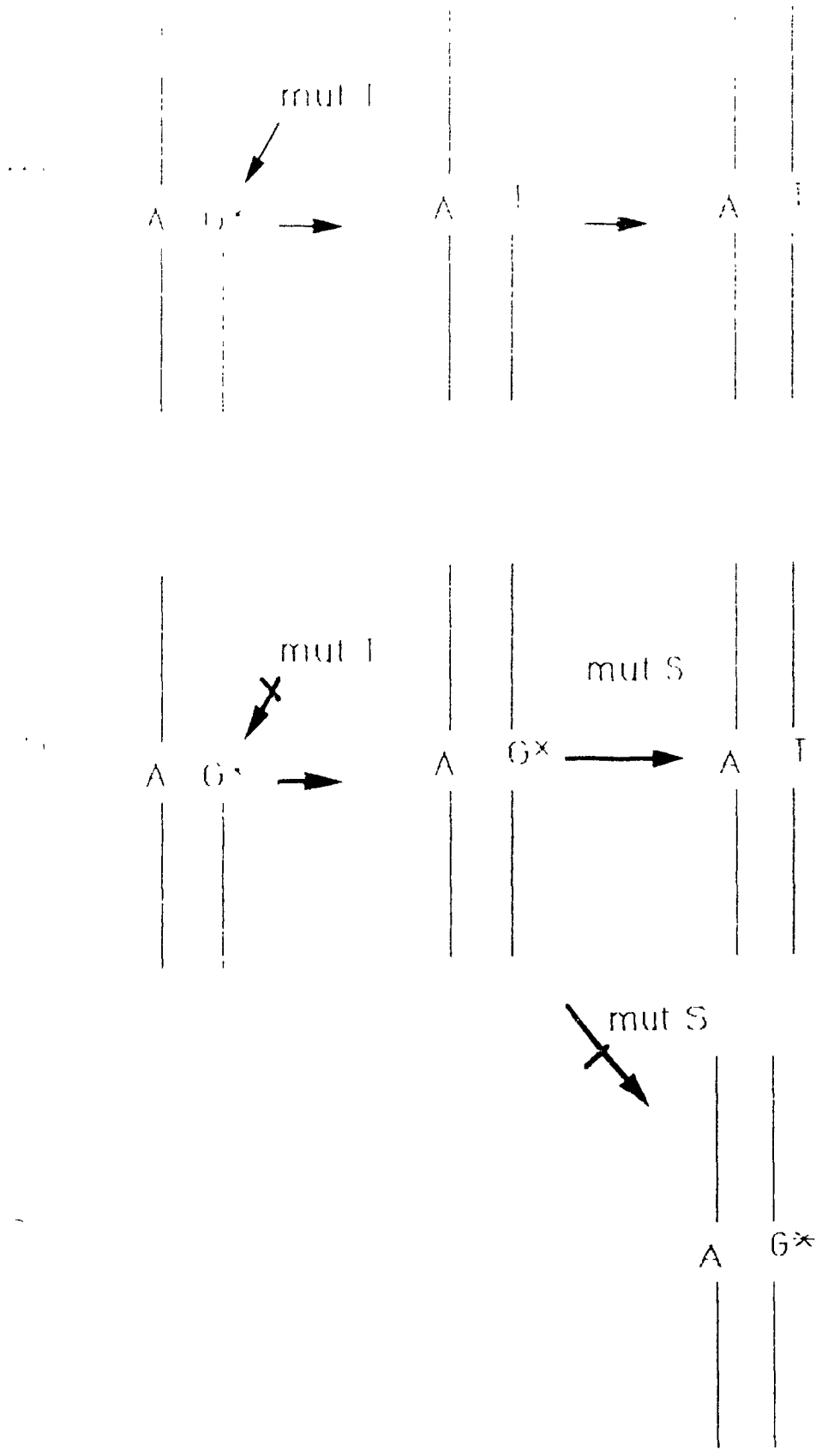
The effect of mutS in a mutL deficient background

a) The effect of wild type mutL, the insertion of 8-oxodGTP does not occur

b) The mispairing of 8-oxodGTP with A, processing of the mismatch by dam directed mismatch repair. The removal of 8-oxoG, and resynthesis of the DNA strand

c) The effect of mutS deficient strain no processing of the mismatch, the mutation will get fixed in the DNA sequence on the next round of DNA replication

(8-oxo-7,8-dihydro-2deoxyguanosine ^{8}G)


mut I
A
A
A
mut I
mut S
A
A
A
mut S
A

displayed decreases. We feel that site 461 is affected by the repair of mutS, and in turn decreasing the effect of mutY. The effect of repair system interaction on G:C to T:A tranversion. There was no effect of the mutT or mutHLS repair systems on the mutational frequency with all mutators carrying the deficient mutY gene (80 Lac$^+$ revertants per 10^9 viable cells). We felt that the mutT$^-$ mutator would have increased the mutational frequency of mutY$^-$, due to the insertion of 8-oxodGTP opposite a dC. This insertion may lead to an 8-oxoG mispairing with an A during the next round of replication, therefore increasing the mutational frequency of the G:C to T:A transversion. This did not occur probably due to the action of fpg/mutM gene which codes for 8-hydroxyguanine endonuclease, which effeciently removes 8-oxoG from the DNA strand when paired with dC, dG or dT, but not with dA. 8-hydroxyguanine endonuclease would remove the 8-oxoG from the 8-oxoG:C pair therefore making it unavailable to base pair with dA (figure 9).

In this study, we have tried to understand how mutations occur, and how they are prevented and/or repaired. There are many steps involved in this process. First, We developed and characterized a potential target sequences, site 462 of the lacZ gene. Second, we constructed and characterized $((GC)_n)$ target sequences. Third, we analysed the interaction of repair systems (mutY, mutT, and mutHLS) by looking at there effect on mutation frequencies of three different base substitution events. By using the Lac$^+$ phenotypic assay we were able to increase our understanding of mechanisms of repair in *Escherichia coli*.

Figure 9:

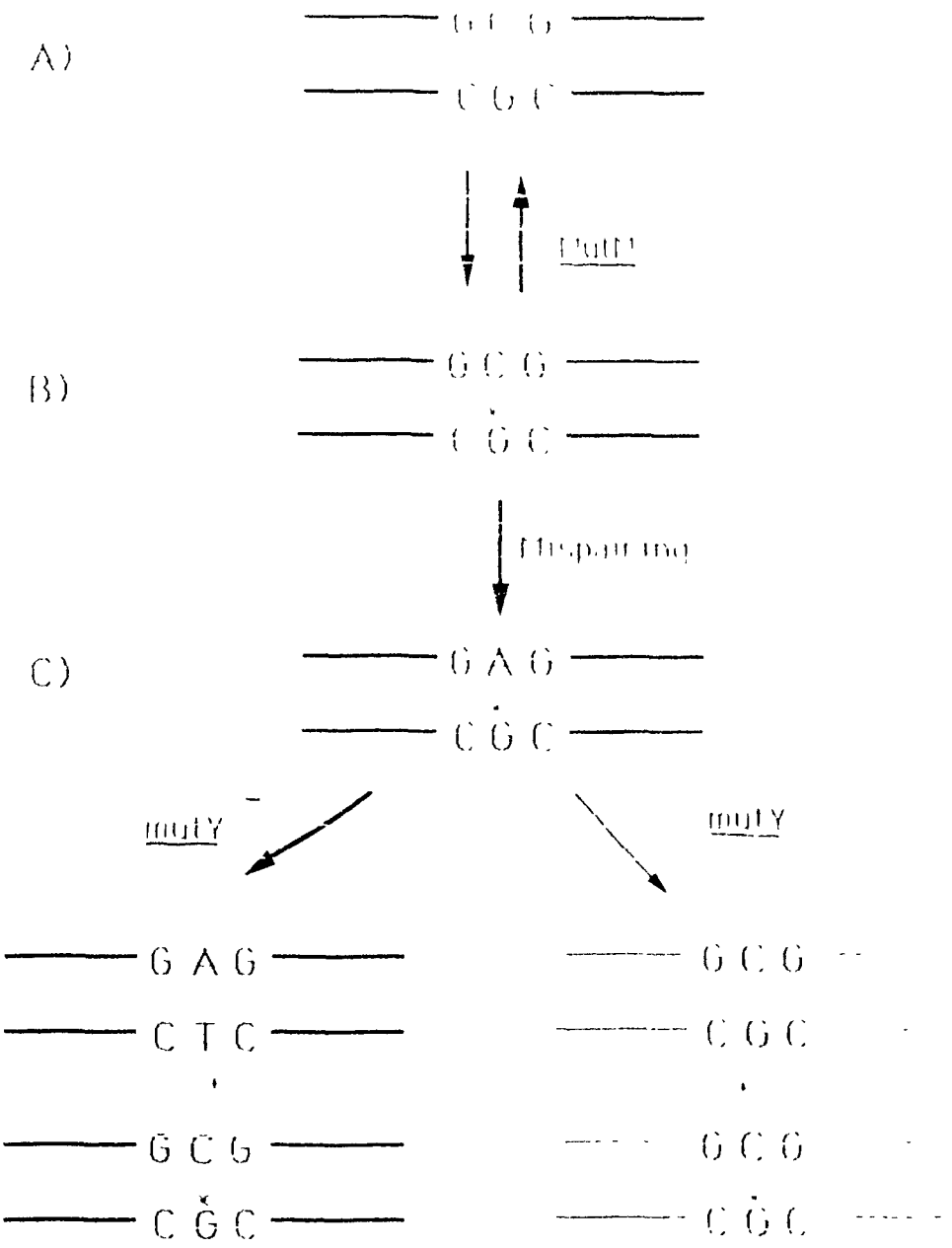
The effect of mutI deficiency in a mutY⁻ background and the effect of mutM in countering the mutI effect.

a) Wild type sequence

b) The pairing of 8 oxodGTP with cytosine in a mutM⁺ background. 8-oxoG will be remove by mutM protein and replaced by guanine nucleotide

c) In a mutM deficiency, at subsequent round of replication an adenine nucleotide may mispair with 8-oxoG. Hence a there will be a transversion event may occur depending on if it is in a wild type or mutant mutY background.

(8 oxo 7,8 dihydro- 2'deoxyguanosine, °G)


A)
C G C
B)
G C G
C G C
C)
G A G
C G C
mutY
mutY
G A G
C T C
G C G
C G C
G C G
C G C

Bibliography

Akiyama, M., H. Maki, M. Sekiguchi and T. Horiuchi (1989) A specific role of mutL protein: To prevent dG-dA mispairing in DNA replication. Proc. Natl. Acad. Sci. U.S.A. 86: 3949 - 3952.

Au, K.G., M. Cabrera, J.H. Miller and P. Modrich (1988) The *Escherichia coli* mutY gene product is required for the specific A-G to C-G mismatch correction. Proc. Natl. Acad. Sci. U.S.A. 85: 9163 - 9166.

Au, K.G., S. Clark, J.H. Miller and P. Modrich (1989) The *Escherichia coli* mutY gene encodes an adenine glycosylase active on A-G mispairs. Proc. Natl. Acad. Sci. U.S.A. 86: 8877 - 8881.

Bader, D.E., M. Ring, and R.E. Huber (1988) Site- directed mutagenic replacement of glu 461 with gln in β- galactosidase (*E.coli*): evidence that glu-461 is important for activity. Biochem. Biophys. Res. Commun. 153: 301 - 306.

Bebenek, K., and T.A. Kunkel (1990) Frameshift errors initiated by nucleotide misincorporation. Proc. Natl. Acad. Sci. U.S.A. 87: 4946 - 4950.

Boeke, T.D., (1981) One and two codon insertion mutants of bacteriophage f1. Mol. Gen. Genet. 181: 288 - 291.

Cabrera, M., Y. Nghiem, and J.H. Miller (1988) mutM, a second mutator locus in *Escherichia coli* that generates G:C to T:A transversions. J. Bact. 170: 5404 - 5407.

Calos, M.P., and J.H. Miller (1981) Genetic and sequence analysis of frameshift mutations induced by ICR- 191. J. Mol. Biol. 153: 39 - 66.

Claverys, J-P., and S.A. Lacks (1986) Heteroduplex deoxynucleic acid base mismatch repair in bacteria. Microbiol. Reviews 50: 133 - 165.

Coulondre, C., and J.H. Miller (1977) Genetic studies of the lac repressor. III. Additional correlation of mutational site with specific amino acid residues. J. Mol. Biol. 117: 525 - 576.

Cupples, C.G., and J.H. Miller (1988) Effects of amino acid substitutions at the active site in *Escherichia coli* β-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. U.S.A. 86: 5345 - 5349.

Cupples, G.C., M. Cabrera, C. Cruz, and J.H. Miller (1990) A set of lacZ mutations in *Escherichia coli* that allow rapid detection of specific frameshift mutations. Genetics 125: 275 - 280.

DiNardo, S., K.A. Voelkel and R. Sternglanz (1982) *Escherichia coli* DNA topoisomerase I mutants have compensatory mutations in DNA gyrase genes. Cell 31: 43 - 51.

Friedberg, E.C., (1985) DNA repair. W.H. Freeman and Company, New York, New York.

Fuchs, R.P., A.-M. Freund, M. Brichara and N. Koffel Schwartz (1988) DNA structure and mutation hot spots. DNA Replication and Repair. American Society for Microbiology, Washington, D.C.

Gellert, M., (1981) DNA topoisomerases. Ann. Rev. Biochem. 50: 878 - 910.

Grilley, M., J. Holmes, B. Yasher and P. Modrich (1990) Mechanisms of DNA mismatch correction. Mut. Res. 236: 253 - 267.

Huber, R.E., G. Kruz, and K. Wallenfels (1976) A quantification of the factors which affect the hydrolase and transgalactosylase activities of β-galactosidase (*E. coli*) on lactose. Biochem. 15: 1994 - 2001.

Kleina, L.G., J.M. Masson, J. Normanly, J. Abelson and J.H. Miller (1990) Construction of *Escherichia coli* amber suppressor tRNA genes. II. Synthesis of additional tRNA genes and improvement of suppressor efficiency. J. Mol. Biol. 213: 705 - 715.

Lieb, M. (1983) Specific mismatch correction in bacteriophage lambda crosses by very short patch repair. Mol. Gen. Genet. 190 : 465 - 470.

Leong, P. K., H.C. Hsia, and J.H. Miller (1986) Analysis of spontaneous base substitutions generated in mismatch repair-deficient strains of *E. coli*. J. of Bact. 166 : 112 - 116.

Lu, A. L., and D.Y. Chang (1988) A novel nucleotide excision repair for the conversion of 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 mismatch in vitro. Proc. Natl. Acad. Sci. U.S.A. 80 : 4639 - 4643.

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 Habor Laboratory, Cold Spring Harbor, N.Y.

Michaels, M.L., C. Cruz and J.H. Miller (1990) MutA and MutC: Two mutator loci in *Escherichia coli* that stimulate tranversions. Proc. Natl. Acad. Sci. U.S.A 87 : 9215 - 9219.

Miller, J.H. (1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Miller, J.H. (1983) Mutational specificity in bacteria. Ann. Rev. Genet. 17 :215 - 238.

Miller, J.H. and W.S. Reznikoff (1978) The Operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Modrich, P. (1987) DNA mismatch correction. Ann. Rev. Biochem. 56 : 435 - 466.

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 A:T transversions. Proc. Natl. Acad. Sci. U.S.A. 85 : 2709 - 2712.

Normanly, J., L.G. Kleina, J.M. Masson, J. Abelson and J.H. Miller (1990) Construction of *Escherichia coli* amber suppressor tRNA genes. III. Determination of tRNA specificity. J. Mol. Biol. 213 : 719 - 726.

Normanly, J., J.M. Masson, L.G. Kleina, J. Abelson and J.H. Miller (1986) Construction of two *Escherichia coli* amber suppressor genes: tRNA^{Phe} and tRNA^{Cys}. Proc. Natl. Acad. Sci. U.S.A. 83 : 6548 - 6552.

Ring, M., J.M. Armitage, and R.J. Huber (1985) m-Fluorotyrosine substitution in β-galactosidase: evidence for the existence of a catalytically active tyrosine. Biochem. Biophys. Res. Commun. 131 : 675 - 680.

Ring, M., D.E. Bader, and R.L. Huber (1988) Site directed mutagenesis of β-galactosidase (*E.coli*) reveals that Tyr-503 is essential for activity. Biochem. Biophys. Res. Commun. 152 : 1050 - 1055.

Schaaper, R.M., B.I. Bond, and R.G. Fowler (1989) A·T to C·G transversions and their prevention by the *Escherichia coli* mutT and mutHLS pathways. Mol. Gen. Genet. 89 : 256 - 262.

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. U.S.A. 84 : 6220 - 6224.

Schaaper, R.M., and M. Radman (1989) The extreme mutator effect of *Escherichia coli* mutD results from the saturation of mismatch repair by excessive DNA replication errors. EMBO J. 8 : 3511 - 3516.

Shibutani, S., M. Takeshita and A.P. Grollman (1991) Insertion of specific bases during DNA synthesis past the oxidation-damage base 8-oxodG. Nature 349 : 431 - 434.

Sinnott, M.L. (1978) Ions, ion-pairs and catalysis by lacZ β-galactosidase of *Escherichia coli*. FEBS Lett. 94 : 1 - 9.

Sternglanz, R., S. DiNardo, K. A. Voelkel, Y. Nishimura, K. Becherer, L. Zumstein, and J.C. Wang (1981) Mutations in genes coding for *Escherichia coli* DNA topoisomerase affect transcription and transposition. Proc. Natl. Acad. Sci. U.S.A. 78: 2747-2751.

Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi, and M. Inouye (1966) Frameshift mutations and the genetic code. Cold Spring Harbor Symp. Quant. Biol. 31: 77-84.

Suzuki, D.T., A.J.F. Griffiths, J.H. Miller, and R.C. Lewontin (1986) An introduction to genetic analysis. W.H. Freeman and company, New York, New York.

Stryer, L. (1975) Biochemistry. Freeman and Company, N.Y.

Tchou, J., H. Kasai, S. Shibutani, M.-H. Chung, J. Laval, A.P. Grollman and S. Nishimura (1991) 8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. Proc. Natl. Acad. Sci. U.S.A. 88: 4690-4694.

Treffers, H.P., V. Spinelli, and N.O. Belser (1954) A factor (or mutator gene) influencing mutation rates in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. 40: 1064-1071.

Watson, J.D., N.H. Hopkins, J.W. Roberts, J.A. Steitz, and A.M. Weiner (1987) Molecular biology of the gene, Volume I. Benjamin Cummings, Menlo Park, California.

Yanofsky, C., F.C. Cox, and V.D. Horn (1966) The unusual mutagenic specificity of an *Escherichia coli* mutator gene. Proc. Natl. Acad. Sci. U.S.A. 55: 274-281.