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 a désirer, surtout si les pages originales ont été dactylogra phiées 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

Molecular basis of frameshift mutations
in *Escherichia coli*

Erica Passi

A Thesis

in

The Department

of

Biology

Presented in Partial Fulfilment of the Requirements
for the degree of Master of Science at
Concordia University
Montréal, Québec, Canada

July 1991

ISBN 0-315-73676-3

Canada

ABSTRACT

Molecular basis of frameshift mutations
in *Escherichia coli*

Erica Passi

The goal of this research was to understand how frameshift
mutations are generated. The approach was to develop a
simple phenotypic assay to monitor the occurence of specif-
ic frameshift mutations. The assay involved introducing
frameshift target sites into *lacZ* of *Escherichia coli* in
such a way so as to inactivate the gene and produce a Lac$^-$
cell. A specific frameshift mutation was required to
revert the cell back to Lac$^+$ phenotype. The frequency of
Lac$^+$ reversion was monitored in different genetic back-
grounds and following treatment with chemical mutagens. A
stretch of DNA downstream of tyrosine at position 503 which
has repeating CG dinucleotides was used to study the mecha-
nism of frameshift mutations. The research indicated that
the addition of CG bases does not occur by a strand slip-
page mechanism in this stretch of DNA.

In the course of identifying suitable frameshift
targets in *lacZ* it was found that site 460, which codes for
asparagine, is very important for the activity of β-
galactosidase and is not tolerant of amino acid substitu-
tions. Thirteen amino acids were substituted at this site
using two different approaches. An amber codon introduced
at the site was suppressed by 12 nonsense suppressor
strains which introduced 12 amino acids individually at

site 460, namely ser, cys, ala, glu, gly, leu, lys, phe,
tyr, gln, his and pro. Three missense mutants which had
asp, ser and gln at site 460 were constructed. The kinetic
analysis of the activity of the mutant proteins indicated
that the size and shape of asn-460 is more important than
charge for its role.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

2-AP	2-aminopurine
b	blue
db	dark blue
ds	double-stranded DNA
DS	deletion strains
E. coli	*Escherichia coli*
EMS	Ethyl methanesulfonate
lb	light blue
LB	Luria broth
ONPG	o-nitrophenyl-ß-D-galactopyranoside
Pgal	phenyl-ß-D-galactopyranoside
PNPG	p-nitrophenyl-ß-D-galactopyranoside
RF	replicative form
ss	single-stranded DNA
w	white
Xgal	5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside

INTRODUCTION

The initial goal of this study was to understand the mechanism of frameshift mutations. The approach was to design a phenotypic analysis in order to monitor the occurence of frameshift mutations. Frameshift mutations which were introduced into *lacZ* by site-directed mutagenesis *in vitro* inactivated the gene and resulted in Lac⁻ phenotype. Only specific spontaneous frameshift mutations *in vivo* could restore the wild-type Lac⁺ phenotype. Two types of targets were selected to create frameshifts : runs of G's and runs of CG dinucleotides. The first target was upstream of glu-461 and the second one was downstream of tyr-503 in the *lacZ* gene of *E. coli*.

Glu-461 and tyr-503 are essential for the catalytic activity of β-galactosidase. In order to ensure that frameshifts occur at the chosen target only and not at other sites, I decided to specifically change the DNA sequence adjacent to the two essential residues. Any frameshift at the first target could go back to the wild-type phenotype only after bringing glu-461 back into frame. A run of 9 G's before the site 461 would result in Lac⁻ phenotype (fig 1). The cell could revert back to the Lac⁺ phenotype only by the spontaneous deletion of a G which would bring glu-461 back into frame.

A run of 9 G's could be put before the site 461 only if asn-460 (AAT) could be substituted by gly (GGG) without

Figure 1

Construction of a frameshift target:

The construction of a +G frameshift target is shown if asn-460 allowed the substitution of asn by gly.

GENOTYPE   PHENOTYPE

458   459   460   461   Lac$^+$
Leu   Gly   Asn   Glu   (wild-type sequence)

CTG   GGG   AAT   GAA

CTG   GGG   GGG   GAA   Lac$^+$
                                      (asn-460 -> gly-460 substitution)

CTG   GGG   GGG   GGA   A   Lac$^-$
                                      (Addition of a G by
                                      site-directed mutagenesis)

CTG   GGG   GGG   GAA   Lac$^+$
                                      (Spontaneous deletion of a G to
                                      go back to Lac$^+$ phenotype)

significant loss of activity of ß-galactosidase. Thus, before creating frameshift targets upstream of glu-461, it was essential to study the residue asn-460 since frameshift targets could be made only if asn-460 allowed amino acid substitutions. Investigation of asn-460 showed that it was very important for the activity of ß-galactosidase and was not tolerant of amino acid substitutions. Thus, frameshift targets could not be created upstream of glu-461. A frameshift target was, however, created downstream of tyr-503. In this thesis I report the following :

(I) role of asparagine at position 460 in ß-galactosidase of *Escherichia coli*.

(II) mechanism of generation of frameshift mutations at a run of CG dinucleotides adjacent to tyr-503.

(I) <u>Role</u> <u>of</u> <u>asparagine</u> <u>at</u> <u>position</u> <u>460</u>

Nearly all chemical reactions in biological systems are catalyzed by specific macromolecules called enzymes. With the exception of a few RNA molecules, all of the known enzymes are proteins. Enzymes are catalysts that can increase reaction rates by up to 10^{12} to 10^{15}-fold.

<u>Structure</u> <u>of</u> <u>proteins</u>

Amino acids are the basic structural unit of proteins. There are twenty amino acids which vary in size, shape, charge, hydrogen bonding capacity or chemical reactivity. All proteins are constructed from the same set of twenty amino acids. An amino acid consists of a central

4

alpha carbon atom, an amino group, a carboxyl group, a hydrogen atom and a distinctive R-group bonded to a carbon atom. An R-group is referred to as a side-chain. The 20 amino acids, along with their charges, are presented in table 1.

Each protein has a precisely defined amino acid sequence. F. Sanger (1953) was the first to determine the amino acid sequence of a protein i.e., insulin (Stryer, 1975).

Each protein is made of a specific number of amino acids. In a protein, the amino acids are linked together in a linear array called a polypeptide chain. In a protein the amino group of one amino acid is linked to the carboxyl group of another, with the loss of a water molecule, giving rise to a peptide bond. The nitrogen, carbon and oxygen atoms share electrons which makes the peptide bond resistant to twisting. The amino acid composition of the protein affects both its folding and the conformation that can be taken by the protein backbone. The linkage between the connected amino acids is quite rigid. The proteins can fold only by rotation about the bonds to the alpha carbon. The amino acid sequence is called the primary structure of a protein. Most of the backbone of a protein can be divided into regions of secondary structures, which are distinct segments with characteristic shapes. The secondary structures fall into two main categories : alpha helices and

Table 1:

The 20 amino acids with their respective charge

Amino acid	charge*
Alanine	uncharged
Valine	uncharged
Leucine	uncharged
Isoleucine	uncharged
Proline	uncharged
Phenylalanine	uncharged
Tryptophan	uncharged
Methionine	uncharged
Glycine	uncharged
Serine	uncharged
Threonine	uncharged
Cysteine	uncharged
Tyrosine	uncharged
Asparagine	uncharged
Glutamine	uncharged
Aspartate	negative
Glutamate	negative
Arginine	positive
Histidine	positive
Lysine	positive

* The charge of the negatively and positively charged amino acids
mentioned above is that at pH 6.0.

beta strands or sheets. The alpha helix is a rod-like structure. The tightly coiled polypeptide main chain forms the inner part of the rod and the side-chains extend outward in a helical array. The β pleated sheet is a sheet rather than a rod. The polypeptide chain in the β sheet is almost fully extended rather than being tightly coiled. Most proteins have compact, globular shapes due to frequent reversals of the direction of their polypeptide chains which are called β-turns. The helices and sheets are assembled to give the complete three-dimensional conformation of a single polypeptide strand, which is called the tertiary structure. The remarkable range of functions mediated by the proteins result from the diversity and versatility of these twenty building blocks. These amino acids are used to create the intricate, well defined three-dimensional structures that enable the proteins to carry out so many biological processes. A stretched-out or randomly-arranged polypeptide chain is usually devoid of any biological activity.

Protein synthesis

The amino acid sequence of the proteins is genetically determined. The precise sequence of bases in DNA carries the genetic information. A set of three bases in DNA which codes for the structure of an amino acid s called a codon. The code is read in blocks of three bases from a fixed starting point. The code is degenerate since

7

there are 20 amino acids and 61 codons that code for them.
DNA is not the direct template for protein synthesis;
rather, messenger RNA (mRNA) is the template for protein
synthesis. Transfer RNA (tRNA) carries amino acids in an
activated form to the ribosome for peptide bond formation
in a sequence determined by the mRNA template. There is
one, or more than one, kind of tRNA for each of the 20
amino acids. Transfer RNA contains an amino acid attach-
ment site and a template recognition site, which is a
sequence of three bases called the anticodon. The antico-
don on tRNA recognizes a complementary sequence of three
bases on mRNA, called the codon. There are no tRNAs to
recognize the stop codons (TAG, TAA and TGA). Ribosomal
RNA (rRNA) along with ribosomal proteins makes up the
ribosomes.

Function analysis of proteins

Knowledge of the role of specific amino acids in the
protein sequence is very important for the following rea-
sons; first, this helps in elucidating the mechanism of
action of some proteins, e.g., the catalytic mechanism of
an enzyme; second, the analysis of the relations between
the amino acid sequences and the three-dimensional struc-
tures of the proteins uncovers the rules that govern the
folding of polypeptide chains and the three-dimensional
structure that determines a protein's biological function.
Function of a protein depends significantly on its confor-

mation.

Proteins are frequently characterized by correlating specific changes in the amino acid sequence with effects on function. This is done by manipulating the DNA sequence of the protein-coding gene. The classical approach to this problem has been to obtain mutations by selecting for organisms having new properties as compared to the wild-type. This can be accomplished by using random mutagenesis i.e., by exposure to chemical mutagens e.g., nitrosoguanidine, ICR191, 2-aminopurine (2-AP), ethyl methanesulphonate (EMS), 5-bromouracil, nitrous acid and hydroxylamine or by physical agents e.g., ultraviolet light. This classical mutagenesis has been used widely but it suffers from serious disadvantages. First, the methods of mutagenesis severely constrain the kinds of mutations that are obtained. For example, EMS causes mainly GC -> AT transition mutations and ICR191 causes frameshift mutations like GGG -> GG (review by Miller, 1983). Second, since the entire organism is subjected to the mutagen, mutations occuring in the gene of interest are relatively rare. Third, since mutations are identified by virtue of their phenotype, it is impossible to obtain "mutants" that behave indistinguishably from the wild-type; these can be important for determining parts of the protein that are not important for function.

Some of the above mentioned problems can be overcome

9

by the use of recombinant DNA technology. Mutations are first generated in cloned segments of DNA by using a variety of chemical or enzymatic methods. Once generated, the mutant DNAs are subjected to DNA sequence analyses and then the protein is analyzed for the specific function of interest either *In vivo* or *in vitro*. The approach used for altering the nucleotide sequence of cloned segments of DNA in my study is called site-directed mutagenesis (Zoller and Smith, 1982). This approach involves the synthesis of an oligonucleotide with the desired mutation, using a DNA synthesizer. The oligonucleotide is complementary to the sequence of the gene except for the mutated region. The gene in which one wishes to introduce a mutation is cloned in a filamentous phage vector e.g., M-13 phage vector or its derivatives. The oligonucleotide is annealed to the single-stranded phage template containing the gene of interest. DsDNA is synthesized *in vitro* from the oligonucleotide primer and the appropriate strain is transformed with this dsDNA. Plaques formed can be screened for the desired mutation, by phenotype or genotype.

The original DNA sequence can also be altered by linker mutagenesis which employs the use of DNAaseI to administer random double-strand breaks followed by repair of the ends by Klenow fragment. The linkers can then be ligated to these ends resulting in change of the DNA sequence (Maniatis *et al.*, 1982).

10

The *lac* operon

In *E. coli* a cluster of three genes namely *lacZ*, *lacY*, *lacA* constitute the lactose or *lac* operon. The first two gene products are required for the utilization of lactose as a carbon source. *LacZ* codes for the enzyme ß-galactosidase which catalyzes the cleavage of lactose into galactose and glucose (review by Miller and Reznikoff, 1978). *LacY* directs the synthesis of a permease. Lactose permease is a hydrophobic, cytoplasmic membrane protein that catalyzes the coupled translocation of a single ß-galactoside with a single H^+. Several charged residues of the enzyme are important for the substrate recognition and the transport process. When any of the four amino acids, arg-302, lys-319, his-322 and glu-325 are substituted by neutral amino acids, marked loss of activity or altered sugar recognition results (Kaback, 1988; Püttner et al., 1986; Carrasco et al., 1986; King and Wilson, 1990). Glu-325 and his-322 are in close proximity and probably form a salt bridge. Asp-237 and lys-358 are closely placed in the three-dimensional space, possibly forming a "charge neu-tralizing" salt bridge (King et al., 1991). Sequential truncation of the lactose permease at position 396-401 leads to progressive loss of activity and stability (McKenna et al., 1991). The third gene, *lacA*, codes for thioga-lactoside transacetylase. The *in vivo* role of this enzyme is not clear, though it is not essential either for cell

11

growth or for lactose metabolism. These three genes are transcribed into a single poly-cistronic messenger RNA molecule and are thus said to constitute an operon.

Structure of ß-galactosidase

ß-Galactosidase from *E. coli* is a tetramer of four identical subunits. The molecular weight of each subunit is equal to 116,250 daltons. The monomer contains 1023 amino acid residues in a single polypeptide chain. Despite the fact that each of the subunits has an independent catalytic site, only the tetramer is biologically active (Zabin and Fowler, 1978).

Function of ß-galactosidase

ß-galactosidase is a disaccharidase which catalyzes the hydrolysis and transgalactosylis of ß-D-galactopyranosides. Lactose (galactosyl-ß-D-(1->4)-glucopyranose) is known to be the natural substrate of ß-galactosidase (ß-D-galactoside galactohydrolase, EC 3.2.1.23) of *E. coli*. At low lactose concentrations the rate of galactose production is equal to the rate of glucose production. The rate of galactose production relative to glucose, however, drops dramatically at lactose concentrations higher than 0.05 M and production of allolactose, trisaccharides and tetrasaccharides begins (Huber *et al.*, 1976). The transgalactosylase activity of ß-galactosidase on lactose results in the formation of allolactose (galactosyl-ß-D-(1->6)-glucopyranose) (Burstein *et al.*, 1965, Jobe and Bourgeois,

12

1972). The allolactose so produced, rather than lactose itself, is the "natural inducer" of the *lac* operon (Müller-Hill *et al.*, 1964; Jobe and Bourgeois, 1972).

The mechanism of action of ß-galactosidase has not been firmly established, but it has been suggested that it functions in a manner somewhat analogous to lysozyme (Sinnott, 1978). Thus, it has been proposed that ß-galactosidase has a group which stabilizes a galactosyl intermediate, allowing water to react and a group which acts as an acid (donating a proton to the glycosidic oxygen).

Catalytically-essential residues of ß-galactosidase

Several researchers have proposed that a carboxyl group at the active-site of ß-galactosidase covalently stabilizes a transition state carbonium ion form of galactose. Tenu *et al.* (1971) studied the enzyme activity as a function of pH (ranging fron 5.16 to 10.0) in the presence and absence of Mg^{2+}. The activity of both types of enzymes was seen to be controled by at least one unprotonated group which ionizes in the acidic range. This group has a pK less than 6.0 in both types of enzymes. It was thus suggested that this group is a carboxylate. Sinnott and Souchard (1973) suggested that a glycosyl-cation carboxylate ion-pair of this enzyme is formed. This was further supported by Sinnott and Withers (1974). Sinnott (1978) confirmed that the function of the counterionic carboxylate is to capture the highly reactive glycosyl cation and

13

enable the aglycone to diffuse away. Amino groups dramati-
cally improve the capacity of sugars and alcohols to inter-
act with free ß-galactosidase in competition with the
natural substrate (Huber and Gaunt, 1982). The improvement
varies from a factor of about 2 for aminolyxose over lyxose
to a factor of about 300 in the case of 1-aminogalactopyra-
nose as compared with galactopyranose but, in general, the
amino inhibitors are more efficient by a factor of between
10 and 30 as compared to their non-amino derivatives.
Basic N-substituted ß-D-galactosylamines are bound to the
active-site of ß-D-galactosidase 10^2-10^4-fold more tightly
than non-basic, neutral or positively-charged ß-D-galacto-
syl derivatives of closely related structures (Legler and
Herrchen, 1983). Since, the negative charge is unlikely to
be of any importance in binding the neutral ß-galactosidase
substrate, this charge was suggested to be critical for
stabilizing a positively-charged reaction intermediate. In
1984, Herrchen and Legler used an irreversible active-site-
directed inhibitor ([^{3}H] conduritol-C-cis-epoxide) to
identify glutamic acid at 461 as a residue with a carboxyl
group that might be involved at the active-site of ß-galac-
tosidase by the isolation and partial sequence analysis of
a radioactive octapeptide from the cyanogen bromide and
pepsin fragments of the labeled enzyme.

The active-site residue with the negative charge
forms a transient covalent bond with the transition state

14

carbonium ion form of galactose besides stabilizing it (Sinnott and Souchard,1973). Rosenberg and Kirsch (1981) used kinetic slope effect (KIE) method to study the transition-state structure and suggested that the formation of the covalent galactosyl enzyme is the predominant route for most substrates. Withers et al. (1988) inactivated β-galactosidase by 2-deoxy-2-fluoro-β-D-galactosyl flouride due to the accumulation of the enzyme as a covalent glycosyl enzyme intermediate.

It is known that the binding site is rather small since effectors as large as tri and tetrasaccharides are not able to interact with the enzyme (Deschavanne et al., 1978).

Bader et al. (1988) replaced glutamic acid 461 by glutamine using site-directed mutagenesis. Kinetic studies on the purified mutant enzyme showed that it had <0.4% of the wild-type activity. This confirmed that the negative charge on glu-461 is important for the activity. Cupples and Miller (1988) introduced 13 different amino acids at site 461. Enzyme activity in all these mutants is reduced to less than 10% of the wild-type. Glu-461 is required for each part of the action of β-galactosidase (Cupples et al., 1990b). It is important for binding of the substrate. It is involved in the acid catalytic assistance component of galactosylation. It is also essential for the degalactosylation step of the reaction.

Met-502 was labeled by active-site-directed inhibitors N-bromoacetyl-β-D-galactosylamine (Naider et al., 1972) and β-galactosylmethyl-p-nitrophenyltriazene (Fowler et al., 1978). Replacement of met-502 with norleucine gives an enzyme which is not inactivated by N-bromoacetyl-β-D-galactosylamine which shows that this residue is not catalytically essential. Met-502 is very near to the active-site (Naider et al., 1972; Fowler et al., 1978). Thus, it was suggested that if a galactosyl enzyme is formed with the participation of an acid catalytic group, a prime candidate for such a group could be the adjacent residue, tyr-503 (Sinnott, 1978; Sinnott and Smith, 1978; Fowler et al., 1978; Herrchen and Legler, 1984). Direct convincing evidence was provided by Cupples and Miller in 1988 who substituted tyr-503 by 13 different amino acids resulting in a decrease of the enzyme activity to less than 1% in crude preparations. It is thought that tyr-503 acts as a general acid catalyst in the galactosylation step of the reaction of β-galactosidase (Ring et al., 1985, 1988).

Factors Important for the activity of β-galactosidase

The study of the effects of Na^+ and Mg^{2+} on β-galactosidase indicate that sodium ion enhances the affinity of the enzyme for the substrate as well as the maximum rate of hydrolysis at substrate saturation (Neville and Ling, 1967). It has also been shown that Na^+ binds to the free enzyme but less readily to the substrate-bound enzyme

16

(Neville and Ling, 1967; Hill and Huber 1974). A divalent ion, either Mg^{2+} or Mn^{2+}, is required for maximal activity of ß-galactosidase. The negatively charged side-chain of glu-461 is important for divalent cation binding to ß-galactosidase (Edwards et al., 1990). The pH optimum of the enzyme is about 7.2.

Mechanism of action of ß-galactosidase

Tyr-503 acts as an acid catalyst in the "galactosylation" step (breakage of the glycosidic bond). Glu-461 stabilizes a positively charged carbonium ion form of galactose which collapses to form a transitory covalent bond with the carboxyl group of glu-461. Reaction of the enzyme-bound carbonium ion form of galactose takes place with water in the "degalactosylation" step. Tyr-503, acting as a base, probably activates the water for this reaction.

Effect of amino acid substitutions in ß-galactosidase

The large size of the individual polypeptide chains of ß-galactosidase may account for the fact that the enzyme is very tolerant of amino acid substitutions in the primary sequence. Langridge (1974) selected nitrosoguanidine-induced mutants which had less than 50% of the wild-type ß-galactosidase activity and showed that 72% of the point mutations were nonsense rather than missense mutations. Similarly, out of 1,257 mutants of E. coli isolated after nitrosoguanidine treatment on lactose-tetrazolium medium,

17

missense mutants were found to be only 5% of the point mutants. This indicates that most amino-acid substitutions donot decrease enzyme activity sufficiently to prevent growth on lactose (Langridge and Campbell, 1969). Welply et al. (1981) substituted tyrosine for the wild-type amino acids at positions 17, 23, 36 and 41. These positions occur close to or within a proposed dimer-dimer interface of this tetrameric enzyme (review by Zabin, 1982). The amino acid substitutions did not have significant effect on the mutant enzymes. Cupples and Miller (unpublished data) used ethylmethane sulphonate to individually convert 43 glutamine (CAG) and 39 tryptophane (TGG) residues through-out lacZ to amber (TAG) codons. Twelve different amino acids were substituted at these 82 sites in response to the amber codons by using nonsense suppressor strains. Eighty-one of these residues could accept any substitution without significant effect on the enzyme activity. These studies suggest that most missense mutations have little effect on ß-galactosidase activity.

Amino acid substitutions at site 460

In this thesis, I report the effect of making 13 amino acid substitutions at asn-460. Two approaches were employed to achieve this. First, the asn-460 codon (AAT) was changed to an amber codon (TAG) using site-directed mutagenesis. This nonsense mutation was then individually suppressed using 12 nonsense suppressor strains resulting

18

In 12 amino acids at the position 460. The suppressor strains have tRNAs with base changes in their anticodons which can recognize the amber codon and put known amino acids in response to it during translation. The classical suppressor strains have mutant tRNA genes in their chromosomes while artificial suppressor strains have synthetic mutant tRNA genes cloned into plasmids. The amino acids substituted were serine, glutamine, tyrosine, lysine, leucine, glycine using classical suppressor strains and glutamic acid, alanine, cysteine, histidine, phenylalanine, proline using artificial suppressor strains (Normanly et al., 1986, 1990; Kleina et al., 1990). Second, the codon asn-460 (AAT) was changed to asp (GAT), ser (AGT), gln (CAG) using site-directed mutagenesis. Units of ß-galactosidase were measured (Miller, 1972). The proteins were extracted from the three missense mutants and I^-P^+ and were subjected to partial purification. In order to check that the difference in ß-galactosidase activity was not due to the degradation of proteins but due to the point mutations introduced, the proteins were run on a denaturing SDS-polyacrylamide gel. Kinetic analysis was done on the mutant and the wild-type enzymes to determine the Km and Vmax of the enzymes using ONPG and PNPG as substrates.

This enabled me to examine the effect of amino acid substitutions at asn-460 and to determine the role of asn

at this position.

(II) <u>Mechanism</u> <u>of</u> <u>generation</u> <u>of</u> <u>frameshift</u> <u>mutations</u>

The second issue that this thesis addresses is the mechanism of frameshift mutations.

Frameshift mutations result from the alterations of DNA sequence due to the addition or deletion of bases. Alterations of DNA which result in the addition or removal of one or several bases can have drastic consequences for the protein whose structure is encoded by those sequences of DNA. Since the code is read in blocks of three bases from a fixed starting point, the addition or removal of bases from the DNA causes the translation process to lose the proper reading frame.

A model was proposed (Streisinger et al., 1966) to explain how frameshift mutations can occur at runs of repeating bases e.g., G's and A's and runs of dinucleotides e.g., TA and GT. This model is called the "strand slippage model". According to this model, the addition of a base takes place by the looping out of a base from the strand which is being synthesized and the deletion of a base takes place by the looping out of a base from the template strand (figure 2).

Frameshifts often occur at monotonous runs of A's, G's and stretches of alternating G's and C's. Frameshifts occuring at runs of A's and G's are often +/- 1 base frame-shift mutations (e.g., +A, -G) whereas those occuring at

20

Figure 2

Strand slippage model:

The " strand slippage model" proposed by Streisinger (1966) is represented. The addition of a base takes place by the looping out of a base from the strand which is being synthesized and the deletion of a base occurs by the looping out of a base from the template strand.

21

ADDITION OF A BASE

WILD TYPE

```
CTG  GGG
GAC  CCC
```

REPLICATION INTERMEDIATE

```
                  G
5' --------CTG  G GG ----> 3'
3' --------GAC  CCC ------ 5'   TEMPLATE
```

+1 FRAMESHIFT MUTATION

```
CTG  GGG  G
GAC  CCC  C
```

DELETION OF A BASE

WILD TYPE

```
CTG  GGG  A
GAC  CCC  T
```

REPLICATION INTERMEDIATE

```
5' --------CTG  GGA ------> 3'
3' --------GAC  C C  T -- 5'   TEMPLATE
                  C
```

-1 FRAMESHIFT MUTATION

```
CTG  GGA
GAC  CCT
```

22

runs of alternating G's and C's are +/- 2 bases frameshift mutations (e.g., +CG, -CG).

A stretch of alternating G's and C's can form Z DNA. Z DNA is defined as a left-handed helical structure in which there are 12 bp/helical turns and the successive nucleotide residues alternate between the syn and the anti conformations. In the case of a stretch of alternating G's and C's, G's are in the syn conformation and C's remain in the anti conformation while B DNA is a right-handed helical structure in which there are 10 bp/helical turn and all the bases are in the anti conformation. Freund et $al.$ (1989) cloned various stretches of GC dinucleotides in pUC8. The topoisomers of the plasmids were analyzed on a two-dimensional agarose gel and it was seen that pUC-(GC)$_6$ did not have a region of Z DNA whereas pUC-(GC)$_{12}$ had the whole (GC)$_{12}$ insert in Z conformation.

In this study, I constructed the mutant EP5, which has the deletion of a CG dinucleotide from a stretch of 5 CG dinucleotides and requires the addition of a CG dinucleotide to go back to the wild-type phenotype, in order to study the mechanism of frameshift mutations. It is not known whether this stretch of CG dinucleotides forms Z DNA; this work is in progress. Reversion frequency of EP5 to Lac$^+$ phenotype was compared to that of CC109 (Cupples et $al.$, 1990a). CC109 has the addition of a CG dinucleotide leading to a stretch of 5 CG dinucleotides and requires the

23

deletion of a CG dinucleotide to go back to the wild-type sequence. Reverion frequency was also determined as a result of the treatment with mutagens i.e., ICR191, ethyl methanesulfonate, 2-aminopurine and in *recA*, *mutS* backgrounds.

MATERIALS AND METHODS

BACTERIAL STRAINS

E. coli strains are listed in Table 2. Phage f1-Z was maintained in JM801, plasmid pBR329 in p90c, plasmid pGFIB-1 in XAC and F'episomes in s90c.

PHAGE

F1-Z phage contains the *lacZ* gene cloned into R229 which is an f1 filamentous phage derivative containing a unique *EcoRI* site. The internal *EcoRI* site of *lacZ* near the 3' end of the gene was removed and new *EcoRI* sites were inserted on either side of the gene (Cupples and Miller, 1988). The size of f1-Z phage is equal to 9.8 kb which includes *lacZ* which is equal to 3.4 kb.

PLASMID

The size of the cloning vector pBR329 used in this study is equal to 4.2 kb. It has genes coding for chloramphenicol, tetracycline and ampicillin resistance. The unique *EcoRI* site, used for cloning the mutated *lacZ* genes from f1-Z lies in the chloramphenicol-resistance gene.

The suppressor tRNA genes are cloned into the plasmid pGFIB-1 (Normanly *et al.*, 1986, 1990; Kleina *et al.*, 1990).

EPISOME

The episome has *proA*, *proB*, *lacZ*, *lacY*, *lacA* and is *lacI⁻*. Its size is equal to approximately 50 kb. The cells with the episomes containing wild-type *lacZ* gene

give dark blue colonies on minimal glucose plates with Xgal. The episomes with mutant *lacZ* result in either pale blue or white colonies on minimal glucose plates with Xgal.

MAINTENANCE OF BACTERIAL STRAINS

All strains were checked for their phenotype and were stored on appropriate media e.g., classical suppressor strains (XA101-105, XA96) were stored on minimal glucose agar with proline and methionine and artificial suppressor strains (with tRNA genes cloned into pGFIB-1, in the strain XAC) were kept on minimal glucose agar with proline and ampicillin. The suppressor strains were grown in 2 ml LB for 5-6 hours at 37°C without aeration. One hundred ul of the 2 ml culture was then added to 2 ml minimal glucose medium with 0.01% proline, 0.10 mg/ml methionine, in the case of the classical suppressors and 2 ml minimal glucose medium with 0.01% proline, 0.10 mg/ml ampicillin in the case of the artificial suppressor strains. The other strains were grown directly in the appropriate media.

MEDIA

LB medium (Miller, 1972) contains 10 g Bacto tryptone, 5 g Bacto yeast extract and 10 g sodium chloride per liter of medium. LB agar plates have 12 g of Sigma agar in addition to the above constituents. LB top agar has 6.0 g Sigma agar instead of 12 g.

10 X minimal medium A (Miller, 1972) consists of 105 g K_2HPO_4, 45.0 g KH_2PO_4, 10.0 g $(NH_4)_2SO_4$, 5.0 g of sodium

26

citrate.$2H_2O$ per liter of medium.

Minimal glucose and lactose agar plates (Miller, 1972) contain 12 g Sigma agar in addition to 10.5 g K_2HPO_4, 4.5 g KH_2PO_4, 1.0 g $(NH_4)_2SO_4$, 0.5 g sodium.citrate.$2H_2O$, 0.2% glucose or lactose, 0.02% magnesium sulphate, 0.005% thiamine (B_1) per liter of medium, which were added after autoclaving agar and water.

Compounds added to the selection plates are listed in table 3 along with their final concentrations in the media (mg/liter of media) and the methods of sterilization before adding to the media.

Lactose Maconkey plates contain 50 g lactose Maconkey agar per liter of medium.

TABLE 2:

E. coli strains

Strain	Genotype	Source reference
JM801	F⁺Kan, ara Δ(lacproB) thi rpsL	a
CJ236	F⁺,dut ung relA thi pCJ105(Cmʳ)	b
P90c	F⁻, ara Δ(lacproB) thi	a
XAC	F⁻,ara Δ(lacpro) gyrA argE-am rpoB thi	c,d
S90c	F⁻, ara Δ(lacproB) thi rpsL	a
XA101	F⁻, ara Δ(lacpro) gyrA metB argE-am rpoB supD thi	c,d
XA102	F⁻, ara Δ(lacpro) gyrA metB argE-am rpoB supE thi	e
XA103	F⁻, ara Δ(lacpro) gyrA metB argE-am rpoB supF thi	c,d
XA105	F⁻, ara Δ(lacpro) gyrA metB argE-am rpoB supG thi	c,d
XA96	F⁻, ara Δ(lacpro) gyrA metB argE-am rpoB supP thi	c,d
GlyU	F⁻, ara Δ(lacpro) gyrA metB argE-am rpoB supt-RNA thi	a
DS	F⁺lacZ, Δ(lacpro) supE thi	a

28

DS (1-40) contain progressively

larger deletions in *lacZ*

I-P+ F^+lacI^-lacZ, ara $\Delta(lacpro)$ a

thi rpsL

lacZ is wild-type

CC102 F^+lacI^-lacZ, ara $\setminus(lacpro)$ j

thi rpsL

lacZ has glycine missense

mutation at 461

CC107 F^+lacI^-lacZ, ara $\Delta(lacpro)$ k

thi rpsL

lacZ has a -G mutation at 490

CC108 F^+lacI^-lacZ, ara $\Delta(lacpro)$ k

thi rpsL

lacZ has a +G mutation at 490

CC109 F^+lacI^-lacZ, ara $\Delta(lacpro)$ k

thi rpsL

lacZ has a +GC mutation at 505

EP1 F^+lacI^-lacZ,ara $\Delta(lacpro)$ i

thi rpsL

lacZ has an amber mutation at 460

EP2 F^+lacI^-lacZ, ara $\Delta(lacpro)$ i

thi rpsL

lacZ has aspartate missense

mutation at 460

EP3 F^+lacI^-lacZ, ara $\Delta(lacpro)$ i

thi rpsL

lacZ has serine missense

mutation at 460

EP4 F^+*lacI⁻lacZ, ara* Δ(*lacpro*)   l

thi rpsL

lacZ has glutamine missense

mutation at 460

EP5 F^+*lacI⁻lacZ, ara* Δ(*lacpro*)   l

thi rpsL

lacZ has a −GC mutation at 506

MC601 F^-*,ara* Δ(*lacproB*)   f

thi recA::Tn10

GW3731 F^-, *AB1157*   g

mutS215::Tn10

AB1157 *thr−1 leu−6 thi−1 lacY1 galK2*   h

ara14 xyl−5 mtl−1

kdgK51 proA2 his4 argE3

str31 tsx33 supE44

mutS F^-*,ara* Δ(*lacpro*)   l

thi mutS::Tn10

source reference : a, Miller. b, Kunkel *et al.*, (1987). c, Coulondre and Miller (1977). d, Miller and Albertini (1983). e, Kleina *et al.* (1990). f, Cabrera and Miller. g, Walker. h, Walker. i, Cupples, j, Cupples and Miller, (1989) k, Cupples *et al.* (1990) l, this study.

TABLE 3:

Compounds added to the selection plates

Compound	Concentration (mg/liter of media)	Method of sterilization
Proline	100	autoclave
Methionine	50	autoclave
Pgal	500	autoclave
Xgal	40	stock solution made in dimethyl formamide
Ampicillin	100	filter sterilize
Chloramphenicol	20	stock solution made in 100% ethanol, *
Kanamycin	50	filter sterilize
Streptomycin	100	filter sterilize
Rifampicin	100	stock solution made in 100% methanol, *
Tetracycline	15	stock solution made in 50% ethanol, *
Nalidixic acid	30	* high concentration of antibiotic inhibits growth

OLIGONUCLEOTIDE SYNTHESIS AND PURIFICATION

Oligonucleotides were synthesized in order to put the desired nonsense, missense or frameshift mutations in the *lacZ* gene cloned in the phage f1-Z, using site-directed mutagenesis (Zoller and Smith, 1982). Each missense oligonucleotide was 24 bases long and the -GC oligonucleotide was 18 bases long. These were complementary to the *lacZ* gene on the phage, except for the mutated region. These were made on an Applied Biosystems 391 PCR-Mate DNA synthesizer. Four A_{260} units of oligonucleotide (80 ug) were purified by electrophoresis through a 20% acrylamide/7M urea gel, visualized by UV shadowing, excised and eluted in 0.3 M NaCl/10 mM Tris-HCl, pH 7.4/1 mM EDTA/1% phenol, at 37°C for 18 hrs. The DNA was precipitated with 4% glycogen, 10mM $MgCl_2$ and 0.3 M sodium acetate, washed with ethanol, dried and resuspended in distilled water.

SITE-DIRECTED MUTAGENESIS

(A) Phosphorylation of oligonucleotides

Each oligonucleotide (2 ug) was phosphorylated in 70 mM Tris-HCl, pH 7.4/10 mM $MgCl_2$/5 mM dithiothreitol/751 uM ATP, with 5 units of polynucleotide kinase for 1 hour at 37°C. The mixture was then heated to 65°C for 10 minutes to inactivate the enzyme.

(B) Preparation of single-stranded phage DNA

An isolated plaque of f1-Z, added to 100 ul of log phase cells (CJ236) was incubated at 37°C, for 10 min,

32

without aeration, followed by the addition of 2 ml LB and incubation at 37°C for 4.5 hours to overnight with aeration; this was centrifuged at 16,000xg for 5 min and 200 ul PEG/NaCl (20% polyethylene glycol 6000, 2.5 M NaCl) was added to the supernatent; this was kept at room temperature for 5 minutes followed by centrifugation (16,000xg, 5 min). The pellet so obtained was suspended in 100 ul TE (10 mM Tris, pH 7.4, 1 mM EDTA) and 100 ul phenol; this was centrifuged (16,000xg for 3 min) to separate the supernatent, to which 10 ul 3.0 M sodium acetate and 200 ul cold 95% ethanol was added; this was kept at -20°C overnight and centrifuged at 16,000xg for 15 min, at 4°C, to isolate the pellet. The pellet was dissolved in 25 ul distilled water giving about 0.5 ug/ul.

(C) Mutagenesis

Ss f1-Z template (0.5 ug), 0.1 ug mutagenic oligonucleotide, 0.1 ug helper oligonucleotide (complementary to *lacZ*, in a region upstream of mutagenic oligonucleotide at the 3' end of *lacZ*) were mixed in 20 mM Tris, pH 7.4, 10 mM $MgCl_2$, 50 mM NaCl and heated to 90°C for 5 min followed by cooling to room temperature for a period of 45 min, in order to hybridize oligonucleotides to the template. DsDNA was synthesized by adding 40 ul extension buffer (20 mM Tris, pH 7.4, 10 mM $MgCl_2$, 20 mM dithiothreitol, 1 mM ATP), 0.5 ul of each of the four 10 mM dNTP's i.e, dATP, dCTP, dGTP, dTTP, 1 unit of ligase, 2.5 units of Klenow fragment

33

to the mutagenesis mixture after cooling to room tempera-
ture. This was incubated at room temperature for 1 hour.

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

The appropriate cells were grown to late log phase
(O.D. = 0.2-0.25) in 35 ml LB. Cells were suspended in 15
ml cold 50 mM $CaCl_2$ after centrifugation (4000xg, 5 min at
4°C). These were kept on ice for 30 min followed by cen-
trifugation. Cells were gently suspended in 3 ml cold 50 mM
$CaCl_2$.

TRANSFORMATION (Maniatis *et al.*, 1982)

DNA (0.1 ug) was added to the competent cells (300
ul if transforming with phage and 200 ul if transforming
with plasmid) and incubated on ice for 30 min followed by
incubation at 42°C for 2 min. In case of transformation
with phage, the mixture was cooled briefly on ice, 200 ul
log phase JM801, 50 ul 2% Xgal, 3 ml top agar at 50°C were
added and this was poured onto LB plates after briefly
vortexing. In case of transformation with plasmid DNA, the
mixture was cooled briefly on ice, 800 ul LB was added and
the mixture was incubated at 37°C for 30 min to 1 hour.
Cells were plated on appropriate antibiotic containing
plates.

PURIFICATION OF MUTANT PLAQUES FOR FURTHER ANALYSIS

One plaque was put in 1 ml 1 X minA and vortexed
vigourously for 10 min. This was diluted to 10^{-6} in 1 X
minA. One hundred ul of log-phase cells (JM801) were added

to 100 ul of 10^{-6} dilution of the phage.   This was incubated at 37° C for 10 min without aeration. Fifty ul of Xgal, 3 ml top agar were added and the mixture was poured on LB plates. Single-stranded and dsDNA were isolated from these mutant plaques for restriction enzyme digests, sequencing and subcloning into pBR329.

SsDNA TEMPLATE SEQUENCING

Ss f1-Z template with mutated lacZ genes, isolated from the mutant plaques was sequenced by the Sanger dideoxy sequencing technique using T7 DNA polymerase sequencing kit (Bio/Can) and ^{35}SdATP.   The DNA was electrophoresed on a 8% acrylamide/6.8 M urea gel and analyzed by autoradiography.

PREPARATION OF ds PHAGE DNA (REPLICATIVE FORM) AND PLASMID DNA (Maniatis et al., 1982)

In the case of the RF, an isolated plaque of f1-Z, added to 100 ul log phase cells (JM801), was incubated at 37°C for 10 min, without aeration, followed by the addition of 2 ml LB and incubation at 37°C for 4.5 hours to overnight with aeration. This was centrifuged (16,000xg, 2 min). In the case of the plasmid pBR329, a 2 ml overnight culture of p90c with the plasmid, grown overnight with tetracycline (15 mg/ml) was centrifuged. The cells in both cases were then suspended in 100 ul solution I (50 mM glucose, 25 mM Tris, pH 8.0, 10 mM EDTA); this was kept at room temperature for 5 min, 200 ul solution II (0.2 M NaOH,

35

1% sodium dodecyl sulphate) was added and this was left on ice for 5 min. Solution III (3 M potassium and 5 M acetate, pH 4.8) (150 ul) was added followed by incubation on ice for 5 min; this was centrifuged at 16,000xg at 4˚C for 1 min and 400 ul phenol was added to the supernatent; this was vortexed, centrifuged and 800 ul 95% ethanol was added to the supernatent. After 1 min this was centrifuged at 4˚C for 1 min and the pellet was isolated and suspended in 50 ul double distilled water giving approximately 20 ng/ul.

DIGESTION WITH RESTRICTION ENZYMES

DsDNA (100 ng) from the phage and the plasmid (pBR329, pBR329 with *lacZ*) were digested with excess *EcoRI*, 0.08 units of ribonuclease in 10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM ß-mercaptoethonol at 37˚C for 1 hour.

The plasmid DNA (100 ng) was digested with excess *HincII*, 0.08 units of ribonuclease in 10 mM Tris-HCl, pH 7.5, 100 mM NaCl, 7 mM $MgCl_2$ at 37˚C for 1 hour.

AGAROSE GELS

Approximately 100 ng of the digested ds phage DNA and plasmid DNA were electrophoresed on a 1% agarose gel (1 g agarose, 5 ul of 10 ug/ul ethidium bromide in 0.089 M Tris-base, 0.088 M boric acid and 0.02 M EDTA, pH 8.0) with 4% sucrose, 0.005% bromophenol blue and 0.8% ficoll. The bands were seen with UV illumination.

LIGATION OF lacZ WITH PBR329

PBR329 (20 ng) digested with EcoRI, was ligated to 200 ng f1-Z, which had been digested with EcoRI, by 3 units of T4 DNA ligase in 30 mM Tris-HCl, pH 7.8, 10 mM MgCl$_2$, 10 mM DTT, 0.5 mM ATP, 0.1 mg/ml BSA.

TRANSFORMATION AND SCREENING FOR CLONES

The ligated DNA (10 ng) was used to transform 200 ul of competent p90c. The cells were plated on LB plates with tetracycline. The plates were incubated at 37°C for 24 hours. The colonies were grided on a LB plate with tetracycline followed by incubation at 37°C for 24 hours. The plates were replicated on to a chloramphenicol plate. Since the unique EcoRI site of pBR329 lies in the chloramphenicol-resistance gene of the plasmid, insertion of DNA into the EcoRI site results in chloramphenicol sensitivity. The DNA inserted in the EcoRI site of pBR329 could be either lacZ or the remaining part of the ds DNA of f1-Z. The clones which were tetracycline-resistant but chloraphenicol-sensitive were selected for the isolation of plasmid DNA and further analysis. The plasmid DNA was isolated from the chloramphenicol sensitive clones as described earlier. The plasmid DNA was digested with EcoRI and HincII and was electrophoresed on a 1% agarose gel.

TRANSFERRING MUTATIONS FROM PLASMID TO EPISOME BY GENETIC RECOMBINATION

Stationary-phase p90c with pBR329 carrying mutant

lacZ were mated with log-phase I⁻P⁺ in the ratio of 2:1 by incubating at 37°C for 1 hour without aeration. The progeny were selected on minimal glucose plates with tetracycline. Cells were grown overnight in LB medium with tetracycline to allow recombination to take place between the plasmid and episomal *lacZ* genes. Log-phase cells were then mated with stationary-phase s90c in order to transfer the episomes into s90c leaving the plasmids behind. The progeny were selcted on minimal glucose plates with streptomycin and Xgal. Cells with a mutation in *lacZ* gene gave white or pale blue colonies.

<u>PLATE MATING TECHNIQUES</u> (Miller, 1972)

(A) Mapping

Episomes with mutant *lacZ* were mapped using DS 1-40 by griding EP 1-5 on LB plates. EP1 contains amber mutation at site 460; EP2-4 contain asp, ser, gln missense mutations at site 460 respectively, and EP5 contains -GC mutation at 506. The LB plates were replicated on to ML plates with 200 ul of overnight-culture of deletion strains spread on each plate and incubating them at 37°C for 24-48 hours. Recombination takes place between the *lacZ* genes of the mutant and the deletion strains. The 40 deletion strains have progressive deletions of the *lacZ* gene from the 5' end of the gene . Mutations 3' of the deletion end point can give Lac⁺ phenotype following recombination between the mutant and the deletion strains leading to

38

growth of the recombinants on minimal lactose plates. The position of the mutation is defined by the last deletion strain which leads to growth on minimal lactose plates i.e., the strain with the largest *lacZ* deletion which gives Lac+ phenotype following recombination.

(B) Suppression

Episomes carrying *lacZ* gene with the amber mutation at 460 were transfered into 12 nonsense suppressor strains by griding EP1 on LB plates. The LB plates were replicated on to minimal glucose plates containing methionine and rifampicin with 200 ul overnight-culture of suppressor strains spread on them and incubating for 24 hours. These plates were replicated onto minimal glucose plates with methionine, nalidixic acid and Xgal.

ASSAY OF β-GALACTOSIDASE IN WHOLE CELLS (Miller, 1972)

Overnight-culture was diluted 1:10 and O.D. was measured at 600 nm to determine the number of cells present in the culture. An aliquot (0.1 ml) of diluted overnight-culture was added to 0.9 ml Z buffer, pH 7.0 (0.06 M $Na_2HPO_4 \cdot 7H_2O$, 0.04 M $NaH_2PO_4 \cdot H_2O$, 0.01 M KCl, 0.001 M $MgSO_4 \cdot 7H_2O$, 0.05 M β-mercaptoethanol). Two drops of chloroform and 1 drop of 0.1% SDS were added, vortexed for 10 seconds and 0.2 ml 4 mg/ml ONPG (o-nitrophenyl-β-D-galacto-side) was added. The mixture was incubated at 28°C until a medium-yellow colour developed. Sodium carbonate (0.5 ml 1M) was added to stop the reaction and the time was record-

39

ed. O.D. was measured at 420 nm to determine the amount of O-nitrophenol produced. The units of ß-galactosidase activity were calculated using the following formula :

$$\text{Units of ß-galactosidase} = \frac{O.D._{420} \times 10000}{O.D._{600} \times \text{time}}$$

EXTRACTION OF PROTEINS

EP2-4 and I^-P^+ were grown to an O.D. of 0.8 in 2 liters LB. The cultures were centrifuged (6000xg, 10 min) and cells were suspended in sonication buffer, pH 7.4 (20 mM Tris-HCl, 10 mM $MgCl_2$, 0.04% NaN_3) (1 g cells/5 ml buffer). The cells were broken using sonication. The supernatant was isolated after centrifugation at 12,000xg for 30 min. The amount of total protein was measured using a Bio-rad protein assay kit and the units of ß-galactosidase were measured using ONPG.

PARTIAL PURIFICATION OF ß-GALACTOSIDASE

Streptomycin sulphate (5% w/v) was added to the supernatent in the above step. This was stirred slowly at 4˙C for 2 hours, centrifuged at 16,000xg for 30 min and ammonium sulphate was then added to the supernatant to bring it to 25% saturation maintaining pH between 7.0-7.2. This mixture was stirred at 4˙C for 30 min, centrifuged at 16,000xg for 30 min and the supernatant was brought to 43% saturation with ammonium sulphate. This was again stirred for 30 min and centrifuged at 16,000xg for 30 min and the

pellet was dissolved in 4 ml suspension buffer (80 mM Tris, pH 7.5, 1 mM $MgCl_2$, 1 mM ß-mercaptoethanol, 0.1 mM EDTA, 0.04% NaN_3)

SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE)

Fifty ug of partially purified and crude protein extracts, in suspension buffer were electrophoresed on a denaturing SDS-polyacrylamide gel (10% lower gel and 3% stacking gel) with equal volume of 0.625 M ß-mercaptoethanol, 0.002% bromophenol blue, 4.04% SDS, 53.19 mM Tris-HCl, 17.72% glycerol (Silhavy, et al., 1984). The gel was stained with coomassie brilliant blue (R250) (0.24% coomassie blue in 25% isopropanol, 10% acetic acid). The gel was then destained with 5% methanol, 7% acetic acid.

KINETIC ANALYSIS

All four proteins were stored in the suspension buffer which contained 1 mM ß-mercaptoethanol and 0.04% NaN_3 in order to prevent bacterial growth and to keep the enzymes reduced. Azide and ß-mercaptoethanol are potent nucleophiles which can replace water in the reaction, thus the addition of these reagents can affect the rate of the reaction. Therefore, before assaying the enzymes, they were always desalted into the assay buffer, pH 7.0 (30 mM TES, 140 mM NaCl, 1 mM $MgSO_4$) using Bio-rad 10 DG desalting columns (without substrate).

The amount of total protein was measured using a Bio-rad protein assay kit. ß-Galactosidase was assayed

41

using two substrates ONPG and PNPG. Different amounts of substrates were added to the assay buffer and O.D.$_{420}$ was measured as a function of time at 25°C after the addition of the enzyme. The wild-type enzyme being highly reactive was always diluted 1:10, while mutant enzymes were used undiluted.

CHEMICAL MUTAGENESIS

Strain EP5 which contains lacZ with -GC mutation at 506 was treated with chemical mutagens ICR191, EMS and 2-AP. Reversion to Lac$^+$ phenotype on minimal lactose plates and survival on LB plates was measured after 24 hours incubation at 37°C.

(A) ICR191 mutagenesis (Calos and Miller,1981)

Twenty ul of 1:100 diluted overnight culture of EP5 was added to a set of five tubes with 2 ml minimal glucose media containing 0, 2.5, 5.0, 10.0, 20.0 ug/ml ICR191. CC107 strain was used as a control and was given similar treatment. ICR191 being light sensitive, the tubes were kept covered with aluminium foil at all times. These tubes were incubated at 37°C, with aeration overnight. The cultures were diluted to 10^{-6} in 1 X minA. 100 ul of different dilutions were plated on minimal lactose plates and 10^{-6} dilution was plated on LB plates.

(B) EMS mutagenesis (Coulondre and Miller, 1977)

Log-phase EP5, CC102 cells grown in LB were washed with 1 x minA and were suspended in 1 ml of 1 x minA.

These were treated with 30 ul EMS for varying periods of time i.e., 0, 15, 30, 60 min by the addition of EMS and vigourous shaking. Cells were washed with 1 X minA and suspended in the same volume, followed by dilution in 1 x minA to 10^{-6}. Cells were plated on minimal lactose and LB plates.

(C) 2-AP mutagenesis (Coulondre and Miller, 1977)

Twenty ul of 1:100 diluted overnight culture of EP5 was added to a set of six tubes with LB containing 0, 50, 100, 500, 700, 1000 ug/ml 2-AP. These were incubated overnight, at 37°C, with aeration. CC102 was used as a control and was given similar treatment. The cultures were diluted 10^{-6}-fold in 1Xmin A and were plated on minimal lactose and LB plates.

RECOVERY AND SEQUENCING OF LAC[+] REVERTANTS

LacZ genes from spontaneous Lac[+] revertants of EP5 strain were transfered to f1-Z phage carrying amber muta-tion at site 460, for sequencing, as follows. Plaques from phage carrying lacZ with amber mutation at site 460 are white on plates containing Xgal. Lac[+] revertants of EP5 strain were grown in rich medium to mid-log phase. Samples of 100 ul of cells were mixed with 100 ul of phage diluted 10^{-6}-fold in 1 X minA. Cells and phage were incubated at 37°C for 10 min to allow attachment. Two milliliters of LB were added and the cultures were incubated overnight with aeration to allow recombination between the phage and the

43

episomal *lacZ* genes. The cells were pelleted, and the phage in the supernatent were used to infect JM801. Phage containing *lacZ* from the episome produced bright blue plaques on plates containing Xgal. Since the amber-460 mutation cannot revert to Lac$^+$ by a single base change, blue plaques were the result of recombination with the episome. SsDNA was recovered and sequenced as described earlier.

EP5 IN RECA AND MUTS MUTATOR BACKGROUNDS

(A) *recA*

Episomes with -CG mutation from EP5, -G mutation from CC108 and +CG mutation from CC109 were transferred into *recA* background in MC601 by mating EP5, CC108 and CC109 with stationary phase Mc601. The progeny were selected on minimal glucose plates with tetracycline

(B) *mutS*

Episomes with -CG mutation from EP5 and gly-461 missense mutation from CC102 were transferred into *mutS* mutator background by mating EP5, CC102 with mutS. The progeny were selected on minimal glucose plates with tetracycline.

44

RESULTS

INSERTION OF AMBER, MISSENSE AND FRAMESHFT MUTATIONS IN LACZ GENE CLONED INTO F1-Z

F1-Z phage is $lacI^-/lacZ^+$ and has two advantages over the M13 series for these experiments. First, ß-galactosidase is constitutively expressed. Second, the entire lacZ gene is on a single fragment unlike vectors which require alpha complementation. Single-stranded f1-Z was used as a template for site-directed mutagenesis. Oligonucleotides were used to substitute an amber codon at position 460, missense mutations (namely, aspartate, serine, glutamine) at site 460, alanine at 506 and to have a -GC mutation at 506 position in lacZ. The oligonucleotides which were used are presented in table 4. The underlined regions represent the mutated regions. The helper oligonucleotide Z-COOH-2 which was used as a primer in site-directed mutagenesis to assist the in vitro synthesis of dsDNA, is also presented in table 4.

Strain JM801 was transformed with the dsDNA synthesized in vitro using f1-Z single-stranded template, mutagenic oligonucleotides and the helper oligonucleotide. F1-Z phage containing wild-type lacZ gene form bright blue plaques on plates containing Xgal while f1-Z phage containing lacZ gene with amber-460, gln-460 and -GC mutation at 506 formed white plaques and f1-Z phage containing lacZ with asp-460 and ser-460 missense mutations formed pale

Table 4:

The mutagenic and secondary primers

Wild type	3' GGTCGCTGGGGAATGAATCAGGCC 5'
Amber 460	5' GGCCTGATTCCTACCCCAGCGACC 3'
Asp 460	5' GGCCTGATTCATCCCCCAGCGACCA 3'
Ser 460	5' GGCCTGATTCACTCCCCAGCGACC 3'
Gln 460	5' GGCCTGATTCCTGCCCCAGCGACC 3'
Wild type	3' CGCGCGCGTGGATGAAGAC 5'
-GC 506	5' GGTCTTCATCCGCGCGCG 3'
Ala 506	5' GGTCTTCATCCGCGCGCGCG 3'
Z-COOH-2	3' AGTCGACTCGCGGCCAGCG 5'

blue plaques on plates containing Xgal. Mutant plaques were purified. This was done in order to avoid mixing the mutant and wild-type phage as the plaques were in close proximity on some plates. DsDNA and ssDNA was isolated from the mutant phage.

DsDNA was digested with *EcoRI* and was electrphoresed on a 1% agarose gel in order to check that there were no large deletions in the *lacZ* genes. Two bands of the size of 6.5 kb and 3.4 kb corresponding to f1-Z phage and *lacZ* respectively were obtained in the case of each mutant.

In order to check that there were no secondary mutations, the amber, missense and frameshift mutations were reverted back to wild-type sequence by site-directed mutagenesis, using a wild-type oligonucleotide in the mutated region. DsDNA was synthesized *in vitro* using the wild-type oligonucleotide, ss f1-Z phage template from the mutant phage and helper oligonucleotide (Z-COOH-2). Blue plaques were obtained on plates containing Xgal indicating that the mutation being inserted was the only mutation in *lacZ* in each case.

SsDNA from the mutant phage was sequenced. This confirmed that the mutations were present.

SUBCLONING OF MUTANT *LACZ* GENES FROM F1-Z INTO PBR329

Mutant *lacZ* genes were subcloned from f1-Z phage into the plasmid pBR329. Since the *lacZ* gene in f1-Z phage contains *EcoRI* sites at both ends, dsDNA from the mutant

47

phage were digested with EcoRI. PBR329 was also digeted
with EcoRI. Mutant lacZ genes were individually cloned
into the unique EcoRI site of pBR329.

PBR329 with mutant lacZ genes individually cloned
into it, was digested with EcoRI and HincII. The digested
DNA was electrophoresed on a 1% agarose gel. This was done
to ensure that the entire lacZ gene was cloned into pBR329.
Two bands of the size of 4.2 kb and 3.4 kb corresponding to
pBR329 and lacZ respectively were obtained on digestion
with EcoRI. There are 3 HincII sites in lacZ gene and 2
sites in pBR329; therefore, five bands of the size of 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 were obtained depending on the
orientation of lacZ in pBR329, on digestion with HincII as
shown in fig 3.

RECOMBINATION BETWEEN LACZ IN PBR329 AND F' EPISOME.

Each of the five mutations was inserted into the
lacZ gene of the episome by recombination between pBR329
and the episome. The episome used has proA, proB, lacZ,
lacY, lacA and is lacI⁻. A set of 40 deletion strains was
used to check that each mutation mapped to the correct
deletion interval. The deletion strain with the biggest
deletion of lacZ which was able to grow on lactose after
recombination with the four mutants at the site 460 was
deletion strain 24. The deletion strain 24 has a deletion
of approximately 460 N-terminal amino acids. The exact

48

Figure 3

Restriction digests of pBR329:

One hundred ng of PBR329 and pBR329 with mutated *lacZ* gene digested with *EcoRI* and *HincII* were electrophoresed on a 1% agarose gel. Lanes 1 and 2 contain molecular weight DNA size markers. Lane 3 contains pBR329 digested with *EcoRI* showing a single band of 4.2 kb. Lanes 4-8 contain pBR329 with mutated *lacZ* genes carrying amber-460, asp-460, ser-460, gln-460 and -CG mutations respectively digested with *EcoRI*. These show 2 bands of the size of 4.2 kb and 3.4 kb corresponding to pBR329 and *lacZ* respectively. Lane 9 contains pBR329 digested with *HincII*. It shows 2 bands of the size of 2.3 kb and 1.9 kb since there are 2 *HincII* sites in pBR329. Lanes 10 and 13 contain plasmids with asp-460 and -CG mutations. These show 5 bands of the size of 2.3, 1.89, 1.8, 0.98 and 0.62 kb. Lanes 11 and 12 contain plasmids with ser-460 and gln-460 mutations. These show 5 bands of the size of 2.3, 1.8, 1.59, 1.28 and 0.62 kb.
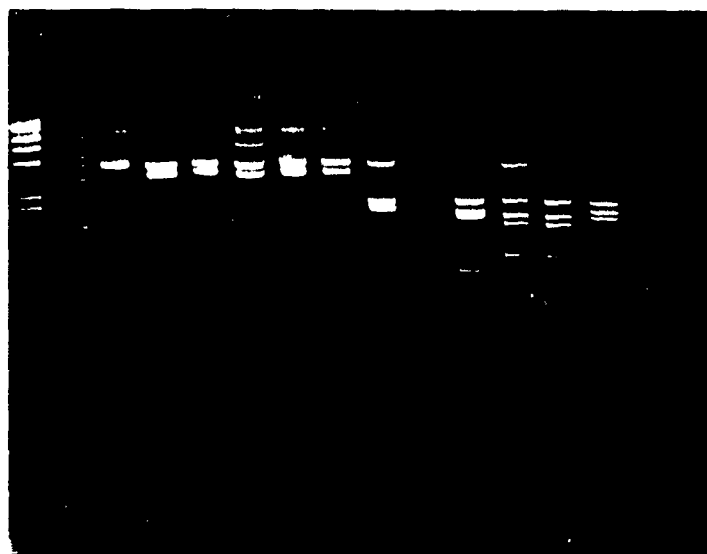
sizes of the deletions of the *lacZ* in the deletion strains have not been determined. The -GC mutation at the position 506 mapped to the deletion interval 26. The deletion strain 26 has a deletion of approximately 506 N-terminal amino acids. The episome with the wild-type *lacZ* gives bright blue colonies on plates containing Xgal. The episome containing *lacZ* with asp-460 mutation gave blue colonies, ser-460 resulted in pale blue colonies and amber-460, gln-460 and -GC mutation resulted in white colonies on plates containing Xgal.

The *lacZ* cloned in f1-Z has a weak promoter (L8), whereas the *lacZ* on the episome has a strong promoter (P^+). Recombinants containing the P^+ promoter have approximately 15-fold higher levels of ß-galactosidase. Thus, the episomes with the Lac^+ revertants of the mutant *lacZ* genes were checked on Maconkey agar plates to ensure that they had the P^+ promoter. The episomes with the P^+ promoter give dark red colonies, whereas the ones with the L8 promoter give pale red colonies. The episomes with the P^+ were used in the further studies.

(I) EFFECTS OF AMINO ACID SUBSTITUTIONS AT ASN-460

(A) Suppression of nonsense mutation

The episome with the amber-460 mutation was transferred into 12 nonsense suppressor strains. Amber-460 was thus, suppressed individually by 12 nonsense suppressor strains, resulting in 12 mutant ß-galactosidases with

51

different amino acids at the position 460. Table 5 shows
the intensity of blue colour of the 12 mutants on plates
containing Xgal, Lac$^{+/-}$ phenotype on minimal plates con-
taining lactose as the only carbon source and the units of
ß-galactosidase activity produced by the mutants with
twelve individual amino acid substitutions at the position
460. The units of ß-galactosidase activity produced by I$^-$
P$^+$ with wild-type *lacZ* gene are also presented. The mutant
in which serine was substituted at position 460 was the
only one which could grow on lactose as the only carbon
source.

(B) Missense mutations

Three amino acids were substituted at 460 by insert-
ing missense mutations, namely asp, ser and gln. Two of
these amino acids were also introduced by nonsense suppres-
sion i.e., ser, gln. Table 6 shows the intensity of blue
colour of the three mutants on plates containing Xgal,
Lac$^{+/-}$ phenotype on minimal plates containing lactose as
the only carbon source and the units of ß-galactosidase
activity as measured by the hydrolysis of ONPG. The mis-
sense mutants, asp and ser could grow on lactose as the
only carbon source. The activity of the missense and the
suppressed nonsense mutant with serine substituted at 460,
differed by a factor of 1.7. The activity of the missense
and the suppressed nonsense mutant with gln substituted at
site   460   differed   by   a   factor   of   1.4.

Table 5:

The analysis of the 12 suppressed nonsense mutants

Amino acid	Xgal	Lactose	Units of β-galactosidase (ONPG)	% units of wild type
Asn (wt)	db	+	5500	100
Ser	lb	+	34.3	0.623
Cys	lb	−	9.11	0.165
Ala	lb	−	5.18	0.094
Glu	lb	−	3.08	0.056
Gly	lb	−	1.62	0.029
Leu	lb	−	1.26	0.023
Lys	w	−	1.00	0.018
Phe	w	−	0.82	0.014
Tyr	w	−	0.74	0.013
Gln	w	−	0.69	0.012
His	w	−	0.66	0.012
Pro	w	−	0.43	0.007

Table 6:

The analysis of the 3 missense mutants

Amino acid	Xgal	Lactose	Units of ß-galactosidase (ONPG)	% units of wild type
Asn (wt)	db	+	5500	100
Asp	b	+	119	2.17
Ser	lb	+	61.3	1.11
Gln	w	−	1.00	0.01

The enzyme activity decreased drastically as a result of all the 13 amino acid substitutions at asn-460.

PARTIAL PURIFICATION OF ß-GALACTOSIDASE FROM THE 3 MISSENSE MUTANTS

ß-Galactosidase from the three missense mutants and I^-P^+ was partially purified. The proteins were electrophoresed on a denaturing SDS-polyacrylamide gel, as shown in fig 4. The amount of ß-galactosidase in all 3 missense mutants and I^-P^+ was similar. The ß-galactosidase band was identified by using molecular weight size markers and by comparing with extract from the strain s90c which has the deletion of *lacZ* (data not shown). Extract from s90c did not have the band corresponding to ß-galactosidase. Thus the drastic decrease in enzyme activity can not be accounted for by the degradation of the mutant proteins.

KINETIC ANALYSIS OF MUTANT ENZYMES

Table 7 shows the kinetic values obtained for the partially purified ß-galactosidase extracted from the missense mutants and the wild-type strain. ONPG amd PNPG were used as substrates. Both Km and Vmax values of the three mutant enzymes were much lower than that of the wild-type enzyme.

SUBSTITUTION OF VAL-506 BY ALA-506

In order to create the mutant EP5 which had the deletion of a CG dinucleotide from a run of CG dinucleotides, it was useful to substitute val-506 by ala-506. The

55

figure 4

Proteins electrophoresed on a SDS-polyacrylamide gel:

50 ug of proteins were electrophoresed on a denaturing SDS-polyacrylamide gel. Lane 1 contains the molecular size markers. Lanes 2-5 contain partially purified proteins from I^-P^+, asp-460, ser-460, gln-460 respectively. Lanes 6-9 contain crude proteins from I^-P^+, asp-460, ser-460 and gln-460 respectively. * indicates the band corresponding to ß-galactosidase.

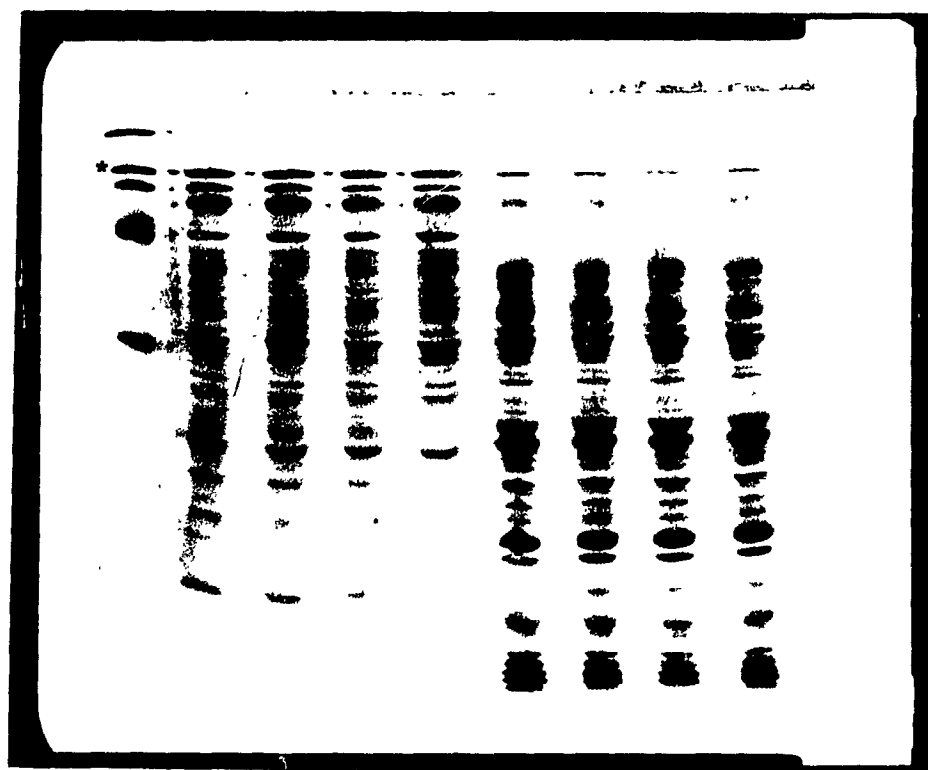Table 7:

The kinetic analysis of the three missense mutants

Amino acid	Km (ONPG) mM/mg	Km (PNPG) mM/mg	Vm (ONPG)	Vm (PNPG)
Asn (wt)	30.0	6.19	5.00	3.09
Asp	1.14	0.11	0.14	0.20
Ser	1.39	0.14	0.04	0.01
Gln	0.62	0.07	0.01	0.005

substitution of val-506 by ala is useful in order to have the frameshift target consisting of 5 CG dinucleotides as compared to the wild-type sequence with val-506 which results in a target of only 4 CG dinucleotides (figure 5). Bigger stretches of CG dinucleotides tend to form Z DNA. Position 506 allows the substitution of valine by alanine without affecting the Lac[+] phenotype. This was checked by reverting a mutant lacZ gene with -CG mutation in f1-Z phage to lac[+] phenotype by using an oligonucleotide with alanine at 506 (table 4). DsDNA was synthesized using a ss f1-Z template with lacZ carrying the -CG mutation, an oligonucleotide with alanine at 506 and a Z-COOH-2 helper oligonucleotide. Blue plaques were formed on plates containing Xgal. Blue plaques could be formed only if the site 506 allowed the substitution of valine by alanine. The mutant EP5 had a stretch of 4 CG dinucleotides. It could revert back to Lac[+] phenotype by the addition of a CG dinucleotide resulting in a stretch of 5 CG dinucleotides which restores the wild-type phenotype but not the wild-type sequence.

(ii) CHEMICAL MUTAGENESIS OF EP5

EP5 which has the deletion of CG bases at 506 in the lacZ gene and requires the addition of CG bases to revert back to Lac[+] phenotype was treated with chemical mutagens, ICR191, 2-AP and EMS, to test whether the rate of frameshift mutations was affected as a result of the treatment

59

Figure 5

Sequence of EP5 and CC109:

The wild-type sequence with val-506 is compared to the sequence with ala-506. The sequence of the mutant EP5 and its Lac⁺ revertant after the addition of CG bases is shown. Sequence of the mutant CC109 which has the addition of CG bases is compared to that of its revertant which has the deletion of CG bases.

SEQUENCE OF EP5

503	504	505	506	507	
Tyr	Ala	Arg	Val	Asp	
TAC	GCG	CGC	GTG	GAT	Lac$^+$ (wild-type sequence)
TAC	GCG	CGC	_GCG_	GAT	Lac$^+$ (Val-506 -> Ala-506)
TAC	GCG	CGC	GGA	T	Lac$^-$ (deletion of a CG dinucleotide by site-directed mutagenesis)
TAC	GCG	CGC	GCG	GAT	Lac$^+$ (+CG to go back to wild-type phenotype)

SEQUENCE OF CC109

503	504	505	506		
Tyr	Ala	Arg	Val		
TAC	GCG	CGC	GTG		Lac$^+$ (wild-type sequence)
TAC	GCG	CGC	GCG	TG	Lac$^-$ (Addition of a CG dinucleotide by site-directed mutagenesis)
TAC	GCG	CGC	GTG		Lac$^+$ (-CG to go back to wild-type sequence)

61

with chemical mutagens. ICR191 causes frameshift mutations by stabilizing the looped out base or bases, which is an intermediate structure according to the strand slippage model proposed by Streisinger (review by Miller, 1983). 2-AP and EMS, on the other hand, cause frameshifts indirectly by saturating the *dam* directed mismatch repair system (Cupples et al., 1990a). Tables 8, 9, 10 show the spontaneous rate of reversion to Lac$^+$ phenotype and the rate of reversion after the treatment with chemical mutagens, in terms of the number of revertants to Lac$^+$ phenotype on minimal lactose plates/10^8 cells. The rate of reversion did not increase as a result of the treatment with chemical mutagens.

Table 11 shows a comparison of the spontaneous rate of reversion to Lac$^+$ phenotype of EP5 and CC109. CC109 requires the deletion of CG bases to go back to the wild-type sequence unlike EP5 which requires the addition of CG mutation to go back to the wild-type phenotype. There was a difference of greater than 100-fold in the rate of reversion of the two strains.

The rate of reversion of EP5 to Lac$^+$ phenotype was tested in the *recA* background. The results are presented in table 12. CC109 and CC108 were used as controls in the *recA* background. The rate of reversion to Lac$^+$ phenotype of CC109 which requires the deletion of CG bases to go back to Lac$^+$ phenotype was increased greater than 5-fold in the

62

Table 8:

The rate of reversion of EP5 after treatment with ICR191

	EP5 (Lac$^+$ revertants/ 10^8 viable cells)	CC107 (Lac$^+$ revertants/ 10^8 viable cells)
Spontaneous rate	0.5	0
ICR191 20 ug/ml	0	25,222
10 ug/ml	0	3,066
5.0 ug/ml	0	1,921
2.5 ug/ml	0.5	592
0 ug/ml	0.6	0

Table 9:

The rate of reversion of EP5 after treatment with EMS

	EP5 (Lac$^+$ revertants/ 10^8 viable cells)	CC102 (Lac$^+$ revertants/ 10^8 viable cells)
Spontaneous rate	0.6	0
EMS 60'	0	6500
30'	0	3438
15'	0	207
0'	0.5	2

Table 10:

The rate of reversion of EP5 after treatment with 2-AP

	EP5 (Lac$^+$ revertants/ 10^8 viable cells)	CC102 (Lac$^+$ revertants/ 10^8 viable cells)
Spontaneous rate	0.5	0
2-AP 1000 ug/ml	0	323
700 ug/ml	0	342
500 ug/ml	0	405
100 ug/ml	0.4	193
50 ug/ml	0	44
0 ug/ml	0.6	0

Table 11:

The spontaneous rate of reversion of EP5 and CC109

	EP5 (Lac$^+$ revertants/ 10^8 viable cells)	CC109 (Lac$^+$ revertants/ 10^8 viable cells)
Spontaneous rate	0.6	95

Table 12:

The rate of reversion of EP5 in recA background

Strain	Lac^{+} revertants/10^{8} viable cells
EP5	0.6
EP5/recA::Tn10	0
CC108	7.18
CC108/recA::Tn10	4.34
CC109	92.0
CC109/recA::Tn10	482

recA background while the rate of reversion of CC108 which requires the deletion of a G base to go back to wild-type was not affected significantly in the *recA* background. The rate of reversion of EP5 to Lac$^+$ phenotype was not affected in the *recA* background.

The rate of reversion of EP5 was also tested in the *mutS* mutator background. The results are presented in table 13. CC102 was used as a control in the *mutS* mutator background. CC102 requires G.C -> A.T base substitution to go back to Lac$^+$ phenotype. Since *mutS* is part of the mismatch repair system, there should be an increase in the rate of reversion of CC102 to Lac$^+$ phenotype in the *mutS* background. The results presented here show an increase of 6-fold in the *mutS* background. The rate of reversion to Lac$^+$ phenotype of EP5 did not increase in the *mutS* mutator background.

SEQUENCING OF LAC$^+$ REVERTANTS

The spontaneous Lac$^+$ revertants obtained from EP5 were sequenced after transferring *lacZ* gene from Lac$^+$ revertants into f1-Z which contains *lacZ* with an amber mutation at site 460. The Lac$^+$ revertants had the addition of CG bases resulting in a stretch of 5 CG dinucleotides (figure 5).

Table 13:

The rate of reversion of EP5 in mutS background

Strain	Lac^+ revertants/10^8 viable cells
EP5	0.5
EP5/mutS::Tn10	0.5
CC102	0.5
CC102/mutS::Tn10	30.0

DISCUSSION

(I) Analysis of the role of asparagine at position 460

ß-Galactosidase is tolerant of amino acid substitutions at most sites in the primary sequence (Langridge, 1974; Langridge and Campbell, 1969; Welply, 1981; Cupples and Miller, Unpublished data). Most missense mutations inserted randomly throughout lacZ gene have little effect on ß-galactosidase activity.

However, glu-461 is essential for the activity of ß-galactosidase (Herrchen and Legler, 1984; Bader et al., 1988; Cupples and Miller, 1988). Glu-461 was identified to be important for the formation of a transient covalent bond with the transition state carbonium ion form of galactose and stabilizing the carbonium ion form of galactose (Sinnott and Souchard, 1973; Rosenberg and Kirsch, 1981; Withers et al., 1988; Cupples et al.,1990b). It is known to be involved in the acid catalytic assistance component of galactosylation and in the degalactosylation step of the reaction (Cupples et al., 1990b).

Tyr-503 is also essential for the activity of ß-galactosidase, as it is the general acid catalyst in the galactosylation step of the reaction of ß-galactosidase (Sinnott, 1978; Sinott and Smith, 1978; Fowler et al., 1978; Herrchen and Legler, 1984; Cupples and Miller, 1988, Ring et al., 1985, 1988).

I have identified asn-460 to be very important for

the activity of ß-galactosidase. Asn-460 is adjacent to the essential residue glu-461. Two experimental approaches were used to study asn-460.

(A) Suppression of nonsense mutation at 460

An amber codon was introduced at the site 460 using site-directed mutagenesis. The episome carrying *lacZ* with the amber mutation at 460 was introduced into 12 nonsense suppressor strains. The nonsense mutation was thus individually suppressed by 12 nonsense suppressors producing 12 mutants which had different amino acids at the site 460. The units of ß-galactosidase produced by the 12 suppressed mutants, as measured by the hydrolysis of ONPG ranged from 0.007% to 0.623% of the wild-type. Ser-460 was the only mutant which could grow on lactose as the only carbon source. It gave 34.3 units of ß-galactosidase (table 5). All the other 11 mutants were phenotypically Lac⁻. This data indicated that asn-460 is not tolerant of amino acid substitutions. Further analysis of this site was essential in order to confirm that asn-460 is essential for the activity of ß-galactosidase.

(B) Missense mutations at asn-460

Nonsense suppressor strains have an efficiency ranging from 8% to 100% (Kleina et al., 1990, Normanly et al., 1990). The efficiency of suppression further depends on the position of the amber codon and the sequence of bases adjacent to it (Kleina et al., 1990, Normanly et al.,

71

1990). A suppressor strain which could insert asn in response to amber (460) could be used as a control but such a suppressor strain has not been constructed. Thus to study the site 460, I introduced 3 missense mutations, asp, ser and gln at the site 460. To facilitate the understanding of the importance of a particular residue, it is important to substitute a number of amino acids at that site. The pattern of changes in the activity as a result of the substitutions can be useful to determine the role of the wild-type residue. The 3 missense mutations at the site 460 were chosen on the basis of the results of suppression of nonsense mutation at 460 and taking into account the biochemical aspects like the structure and charge of the amino acids. Ser and gln were chosen since the suppressed nonsense mutant, ser-460 produced the maximum units of ß-galactosidase out of the 12 suppressed nonsense mutants and gln-460 was one of the mutants which had very little ß-galactosidase activity. The substitution of asn by asp changes the charge on the side-chain from neutral to negative without significantly affecting the size of the side-chain. Ser and asn have a different structure, however both amino acids have a similar size and are uncharged. Gln is bigger in size than asn due to an extra CH_2 group. Gln is also uncharged. The 3 missense mutants enabled us to understand whether the shape or the charge of asn at 460 is crucial for its role. Asp-460 and ser-460 could grow on

lactose as the only carbon source. Asp-460 produced the maximum units of ß-galactosidase. Since asp and asn have similar size but different charge, this result indicates that the charge of asn-460 is not as important as its size for its function. The structure of asn-460 gives a shape to the active site of ß-galactosidase which is essential for its activity. Ser-460 produced 61.3 units of ß-galactosidase as compared to asp-460 which produced 119 units of ß-galactosidase. The ability of ser-460 to grow on lactose as the only carbon source could be due to the similar size of ser and asn, though they have different structures. Gln-460 which is bigger in size as compared to asn produced the least number of units of ß-galactosidase (1.00) and could not grow on lactose as the only carbon source. This again indicates that the shape and size of asn-460 are critical for its function.

KINETIC ANALYSIS OF THE MISSENSE MUTANTS

The three missense and wild-type enzymes were partially purified. The difference in enzyme activity is due to the point mutations introduced and not due to the degradation of the mutant enzymes This was established by electrophoresing equal amounts of total proteins on a denaturing SDS-polyacrylamide gel. The intensity of the ß-galactosidase band was similar in all the four lanes with the wild-type and the 3 missense mutant enzymes. The band corresponding to ß-galactosidase was identified by using

molecular weight protein size markers and protein extract from the strain s90c which has the deletion of the *lacZ* gene. The extract from s90c did not have the ß-galactosidase band.

The Km of the mutant enzymes ranged from 2.06% to 4.63% (with ONPG as substrate) and from 1.13% to 2.26% (with PNPG as substrate) of the wild-type ß-galactosidase. The Km values of the mutant enzymes were lower than that of the wild-type enzyme which suggests that the mutant enzymes bind both ONPG and PNPG more strongly than the wild-type ß-galactosidase. Since the Km value is not necessarily indicative of the substrate binding, it is required to measure the rate constant of each step of the catalytic reaction of ß-galactosidase to further investigate how the binding is affected. Km value is partly related to the dissociation constant for the enzyme-substrate complex. A decrease in Km thus suggests that there is a decrease of the dissociation of the enzyme-substrate complex.

The Vm values of the mutant enzymes ranged from 0.2% to 2.8% (with ONPG as substrate) and from 0.16% to 6.47% (with PNPG as substrate) of the wild-type ß-galactosidase. A decrease in the Vm values of the mutant enzymes is indicative of the fact that the mutant enzymes have a much lower catalytic efficiency as compared to the wild-type ß-galactosidase for both the substrates.

This data strongly suggests that asn-460 is very

important for the activity of ß-galactosidase.

Three missense mutants were constructed in this study. The analysis of the ß-galactosidase activity of more missense mutants at site 460 can be helpful to study the role of asn-460 since the pattern of changes in the enzyme activity of a large number of missense mutants can give useful information. In order to establish the exact role of asn-460 the rate constant of each step of the catalytic reaction of ß-galactosidase is required. The structure of ß-galactosidase is undergoing analysis by X-ray crystallogrphy. The results presented here should be interpreted in the light of the knowledge of the structure of ß-galactosidase. In order to know more about the active site of ß-galactosidase, it will be useful to know whether gly-459 is also important for the activity of the enzyme since it is adjacent to asn-460. It is known that ser-462 which is adjacent to glu-461 is not important for the activity of the enzyme (Vidmar and Cupples, unpublished data). This can give information about the size of the active site of the enzyme.

(II) MECHANISM OF FRAMESHIFT MUTATIONS

The mutant EP5 was constructed in order to study the mechanism of frameshift mutations in a stretch of DNA downstream of tyr-503 which contains repeating CG dinucleotides. EP5 requires the addition of a CG dinucleotide to go back to Lac[+] phenotype. The spontaneous rate of rever-

sion of EP5 to Lac$^+$ phenotype was compared to that of CC109 which requires the deletion of a CG dinucleotide to go back to Lac$^+$ phenotype. The mutant EP5 has a stretch of 4 CG dinucleotides and its reversion to Lac$^+$ phenotype leads to a stretch of 5 CG dinucleotides. The mutant CC109 has a stretch of 5 CG dinucleotides and its Lac$^+$ revertant has a stretch of 4 CG dinucleotides. There is a difference of greater than 100-fold between the spontaneous rate of reversion to Lac$^+$ phenotype of EP5 and CC109. This result suggests that the addition and deletion of CG bases in this particular target does not occur by the same mechanism, contrary to the model proposed by Streisinger *et al.* (1966). The low rate of reversion of EP5 by the addition of CG bases is in agreement with the results of Freund *et al.*. They cloned a stretch of 12 GC dinucleotides in pUC8 which leads to Lac$^-$ phenotype. The stretch of 12 GC dinucleotides formed Z DNA. Fifty independent blue colony-forming mutants from pUC-(GC)$_{12}$ were analyzed. It was seen that the addition of GC bases was a rare event and represented only 8% of the mutants whereas the deletion of 4, 10 and 16 bp occurred at relative frequencies of 10%, 67% and 14% respectively.

CHEMICAL MUTAGENESIS OF EP5

ICR191 is thought to cause frameshift mutations by stabilizing the looped out base(s) by stacking between them. Most of the ICR191 induced mutations (97.9%) studied

In the *lacI* gene were the addition or deletion of a single G base from GGG sequences and in GGGG sequence in one case. The +1 frameshift was preferred by a ratio of 10:1 over the -1 frameshift at some sites whereas at other sites the -1 frameshift occured at a ratio of 2:1 over the +1 frameshift (Calos and Miller, 1981). Ames *et al.*, 1973 and McCann *et al.*, 1975 have shown that ICR191 induces the deletion of a CG dinucleotide in a $(CG)_4$ sequence at 30% of the rate of induction of the deletion of a G base from a GGGG sequence. ICR191 (20 ug/ml) leads to a greater than 6-fold increase in the rate of reversion of CC109 to Lac[+] phenotype as compared to an increase of greater than 1000-fold in the rate of reversion of CC107 which requires the addition of a G base to go back to the wild-type (Cupples *et al.*, 1990a). ICR191 has no effect on the rate of reversion of EP5 to Lac[+] phenotype. This suggests the absence of a looped out intermediate in the addition of CG bases in EP5 to go back to the Lac[+] phenotype.

EMS and 2-AP cause frameshift mutations indirectly by saturating the *dam* directed mismatch repair system (Cupples *et al.*, 1990a). This allows replication errors leading to frameshifts to escape repair. *MutS* is part of the *dam* directed mismatch repair system. The mismatch repair system is involved in the repair of frameshifts. The loss of the system leads to an increase in frameshifts (Glickman, 1979; Schaaper and Dunn, 1987; Cupples *et al.*,

77

1990a). The *E. coli* mismatch repair system can act on newly synthesized DNA strands to remove replication errors involving the insertion or deletion of a single base (Dohet *et al.*, 1986). This was suggested on the basis of results obtained by transfecting *E. coli* with DNA hetroduplexes with a single unpaired base prepared by annealing separated strands of lambda phage. If the addition of CG bases in EP5 occurs by the looping out mechanism, treatment with EMS, 2-AP and *mutS* mutator background should lead to an increase in the rate of reversion of EP5 to Lac$^+$ phenotype by the addition of CG bases. However, treatment with EMS, 2-AP and *mutS* mutator backgrounds have no effect on the rate of reversion of EP5. 2-AP (700 ug/ml) results in greater than 9-fold increase in the rate of reversion of CC109 (Cupples *et al.*, 1990a). Thus, this data suggests that the addition of CG bases in EP5 does not occur through a looped out intermediate.

RecA plays a vital role in the cell due to its role in SOS response and recombination. RecA protein has been shown to have 5 to 7-fold greater affinity to bind to Z DNA than to B DNA (Blaho and Wells, 1987). RecA-Z DNA complexes are more stable than the ones with B DNA. N-2-acetylaminofluorene (AAF) is shown to induce frameshift mutations by two pathways that are under different genetic controls. First, it causes the deletion of a G base (GGGG -> GGG) by a UmuDC and RecA* protein dependent pathway.

78

RecA* is the activated form of the protein RecA. Second, it causes the deletion of CG bases (GGCGCC -> GGCC) by a LexA controlled pathway that is UmuDC and RecA* independent. A non-activated form of the RecA protein acts as an inhibitor of this pathway (Koffel-Schwartz and Fuchs, 1989). In order to determine the role of recA in the addition of CG bases in EP5 the rate of reversion of EP5 in recA backgound was tested. RecA background leads to a greater than 5-fold increase in the rate of reversion of CC109 to Lac$^+$ phenotype. Due to the effect of recA on CC109 and its role in causing frameshifts, I predicted that recA backgroud would lead to an increase in the rate of reversion of EP5 to a Lac$^+$ phenotype. However, recA backgroud had no effect on the rate of reversion of EP5 to a Lac$^+$ phenotype. Therefore, the addition of CG bases occurs by a recA-independent pathway.

The results presented here indicate that the addition of CG bases in EP5 does not occur through a looped out intermediate. More than one model is required to explain frameshift mutations at all kinds of targets. It is possible that the CG dinucleotide from the template strand is excised from this target instead of forming a looped out intermediate. The excision would be followed by the ligation of the template strand. The addition of the CG dinucleotide in the primer strand would be a rare event in that case. The results presented here for the spontane-

ous reversion of EP5 to Lac⁺ phenotype, reversion of EP5
after chemical mutagenesis with ICR191, EMS and 2-aminopu-
rine and in recA and mutS backgrounds support this hypothe-
sis.

A nitrocellulose filter binding assay could be used
to determine whether the protein RecA binds preferentially
to the targets in EP5 and CC109. Direct visualization of
recA binding to the target in plasmid DNA could be obtained
by using electron microscopy. The construction of more
targets which require the addition of different dinucleo-
tides (e.g GT, TA) to go back to wild-type can be useful to
determine whether the addition of 2 bases is a rare event
in general or the addition of CG bases is a rare event only
in some targets.

<u>REFRENCES</u>

Ames, B.N., F.D. Lee, and W.E. Durston (1973) An improved bacterial test system for the detection and classification of mutagens and carcinogens. Proc Natl. Acad. Sci. U.S.A. <u>70</u> : 782 - 786.

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. <u>153</u> : 301 - 306.

Blaho, J.A., and R.D. Wells (1987) Left-handed Z-DNA binding by the *recA* protein of *Escherichia coli*. J. Biol. Chem. <u>262</u> : 6082-6088.

Burstein, C., M. Cohn, A. Kepes, et J. Monod (1965) Rôle du lactose et de ses produits métaboliques dans l'induction de l'opéron lactose chez *Escherichia coli*. Biochim. Biophys. Acta. <u>95</u> : 634 - 639.

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

Carrasco, N., L.M. Antes, M.S. Poonian, and H.R. Kaback (1986) Lac permease of *E. coli*: histidine-322 and glutamic acid-325 may be components of a charge-relay system. Biochemistry <u>25</u> : 4486-4488.

Coulondre, C., and J.H. Miller (1977) Genetic studies of the *lac* repressor. III. Additional correlation of

mutational sites 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* ß-galactosidase. Genetics 120 : 637 - 644.

Cupples, C.G., and J.H. Miller (1989) A set of *lacZ* mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. Proc. Natl. Acad. Sci. U.S.A. 86 : 5345 - 5349

Cupples, C.G., 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.

Cupples, C.G., J.H. Miller, and R.E. Huber (1990) Determination of the roles of glu-461 in ß-galactosidase (*Escherichia coli*) using site-specific mutagenesis. J. Biol. Chem. 265 : 5512 - 5518.

Deschavanne, P.J., O.M. Viratelle, and J.M. Yon (1978) Conformational adaptability of the active site of ß-galactosidase. J. Biol. Chem. 253 : 833 - 837.

Dohet, C., R. Wagner and M. Radman (1986) Methyl-directed repair of frameshift mutations in heteroduplex DNA. Proc. Natl. Acad. Sci. U.S.A. 83 : 3395 - 3397.

Edward, R.A., C.G. Cupples, and R.E. Huber (1990) Site specific mutants of ß-galactosidase show that tyr-503 is unimportant in Mg^{2+} binding but that glu-461 is very

important and may be a ligand to Mg^{2+}. Biochem. Biophys. Res. Commun. <u>171</u> : 33 - 37.

Fowler, A.V., and I. Zabin (1977) The amino acid sequence of ß-galactosidase of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. <u>74</u> : 1507 - 1510.

Fowler, A.V., I. Zabin, M.L. Sinnott and P.J. Smith (1978) Methionine 500, the site of covalent attachment of an active site-directed reagent of ß-galactosidase. J. Biol. Chem. <u>253</u> : 5283 - 5285

Freund, A., M. Bichara, and R.P.P. Fuchs (1989) Z-DNA forming sequences are spontaneous deletion hot spots. Proc. Natl. Acad. Sci. U.S.A. <u>86</u> : 7465 - 7469.

Glickman, B.W. (1979) Spontaneous mutagenesis in *Escherichia coli* strains lacking 6-methyladenine residues in their DNA: An altered mutational spectrum in *dam*⁻ mutants. Mutat. Res. <u>61</u> : 153 - 162.

Herrchen, M., and G. Legler (1984) Identification of an essential carboxylate group at the active site of *lacZ* ß-galactosidase from *Escherichia coli*. Eur. J. Biochem. <u>138</u> : 527 - 531.

Hill, J.A., and R.E. Huber (1974) The mechanism of Na$^+$ activation of *E. coli* ß-galactosidase and the inhibitory effect of high concentrations of Mg^{2+} on this activation. Int. J. Biochem. <u>5</u> : 773 - 779.

Huber, R.E., G. Kurz, and K. Wallenfels (1976) A quantitation of the factors which affect the hydrolase and

transgalactosylase activities of ß-galactosidase (*E. coli*) on lactose. Biochemistry 15 : 1994 - 2001.

Huber, R.E., and M.T. Gaunt (1982) The inhibition of ß-galactsidase (*Escherichia coli*) by amino sugars and amino alcohols. Can. J. Biochem. 60 : 608-612.

Jobe, A., and S. Bourgeois (1972) *lac* repressor-operator interaction. VI. The natural inducer of the *lac* operon. J. Mol. Biol. 69 : 397 - 408.

Kaback, H.R. (1988) Site-directed mutagenesis and ion-gradient driven active transport: On the path of the proton. Annu. Rev. Physiol. 50 : 243 - 256.

Koffel-Schwartz, N., and R.P.P. Fuchs (1989) Genetic control of AAF-induced mutagenesis at alternating GC sequences: An additional role for RecA. Mol. Gen. Genet. 215 : 306 - 311.

King, S.C., C.L. Hansen and T.H. Wilson (1991) The interaction between aspartic acid 237 and lysine 358 in the lactose carrier of *Escherichia coli*. Biochim. Biophys. Acta. 1062 : 177 - 186.

King, S.C., and T.H. Wilson (1990) Sensitivity of efflux-driven carrier turnover to external pH in mutants of the *Escherichia coli* lactose carrier that have tyrosine or phenylalanine substituted for histidine-322. J. Biol. Chem. 265 : 3153-3160.

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 - 717.

Kunkel, T.A., J.D. Roberts, and R.A. Zakour (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods in Enzymology 154 : 367 - 382.

Langridge, J. (1974) Genetic and enzymatic experiments relating to the quaternary structure of ß-galactosidase. Aust. J. Biol. Sci. 27 : 309 - 320.

Langridge, J., and J.H. Campbell (1969) Classification and intragenic position of mutations in the ß-galactosidase gene of *Escherichi coli*. Mol. Gen. Genet. 103 : 339 - 347.

Legler, G., and M. Herrchen (1983) N-Substituted D-galactosylamines as probes for the active site of ß-D-galactosidase from *Escherichia coli*. Carbohydrate Res. 116 : 95 - 103.

Maniatis, T., E.F. Fritsch, and J. Sambrook (1982) Molecular cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y.

McCann, J., N.E. Spingarn, J. Kobori, and B.N. Ames (1975) Detection of carcinogens as mutagens: bacterial tester strains with R factor plasmids. Proc. Natl. Acad. Sci. U.S.A. 72 : 979 - 983.

McKenna, E., D. Hardy, J.C. Pastore, and H.R. Kaback

(1991) Sequencial truncation of the lactose permease over a three-amino acid sequence near the carboxyl terminus leads to progressive loss of activity and stability. Proc. Natl. Acad. Sci. U.S.A. 88 : 2969 - 2973.

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.

Miller, J.H., and A.M. Albertini (1983) Effects of surrounding sequence on the suppression of nonsense codons. J. Mol. Biol. 164 : 59 - 71.

Müller-Hill, B., H.V. Rickenberg, and K. Wallenfeis (1964) Specificity of the induction of the enzymes of the *lac* operon in *Escherichia coli*. J. Mol. Biol. 10 : 303 - 318.

Naider, F., Z. Bohak, and J. Yariv (1972) Reversible alkylation of a methionyl residue near the active site of ß-galactosidase. Biochemistry 11 : 3202 - 3207.

Neville, M.C. and G.N. Ling (1967) Synergistic activation of ß-galactosidase by Na^+ by Cs^+. Arch. Biochem. Biophys. 118 : 596 - 610.

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) Constuction of two *Escherichia coli* amber suppressor genes: tRNA Phe and tRNA cys. Proc. Natl. Acad. Sci. U.S.A.   83 : 6548 - 6552.

Püttner, I.B., H.K. Sarkar, M.S. Poonian, and H.R. Kaback (1986) *Lac* permease of *Escherichia coli*: Histidine-205 and histidine-322 play different roles in Lactose/H$^+$ symport. Biochemistry 25 : 4483 - 4485.

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

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

Rosenberg, S., and J.F. Kirsch (1981) Oxygen-18 leaving group kinetic isotope effects on the hydrolysis of nitrophenyl glycosides. 1. ß-galactosidase-catalyzed hydrolysis. Biochemistry 20 : 3189 - 3196.

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

Schaapper, 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 replica-
tion errors. Proc. Natl. Acad. Sci. U.S.A. 84 : 6220 -
6224.

Silhavy,T.J., M.L.Berman, and L.W.Enquist (1984)
Experiments with gene fusions. Cold Spring Harbor Laborato-
ry, Cold Spring Harbor, N.Y.

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

Sinnott, M.L., and I.J.L. Souchard (1973) The mecha-
nism of action of β-galactosidase. Biochem. J. 133 : 89 -
98.

Sinnott, M.L., and P.J. Smith (1978) Affinity label-
ling with a deaminatively generated carbonium ion. Biochem.
J. 175 : 528 - 538.

Sinnott, M.L., and S.G. Withers (1974) The β-galac-
tosidase catalysed hydrolyses of β-D-galactopyranosyl pyri
dinium salts. Biochem. J. 143 : 751 - 762.

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.

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

Tenu, J.P., O.M. Viratelle, J. Garnier and J. Yon

(1971) pH dependence of the activity of ß-galactosidase from *Escherichia coli.* Eur. J. Biochem. 20 : 363 - 370.

Welply, J.K., A.V. Fowler, and I. Zabin (1981) ß-Galactosidase alpha-complementation. J. Biol. Chem. 256 : 6811 - 6816.

Withers, S.G., K. Rupitz, and I.P. Street (1988) 2-Deoxy-2-fluoro-D-glycosyl fluorides; a new class of specific mechanism-based glycosidase inhibitors. J. Biol. Chem. 263 : 7929 - 7932.

Zabin, I. (1982) A model of protein-protein interaction. Mol. Cell. Biochem. 49: 87 - 96.

Zabin,I., and A.V.Fowler (1978) The Operon. edited by J. H. Miller and W. S. Reznikoff. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Zoller, M.J. and M. Smith (1982) Nucleic Acids Research 10 number 20: 6487-6500