CHAPTER 1

GENERAL INTRODUCTION
Section A: Introduction to mutations

1.1 Introduction

One of the major reasons that so much effort has been devoted to the analysis of DNA damage and to understanding the causes of mutagenesis, is an awareness that such damage can lead to considerable subsequent physiological consequences in any organism. Mutations have the capacity to change an organism's ability for growth, survival, development, and reproduction. As John Drake (64) has put it, "Mutation fascinates because of its three faces: the variability it generates that conditions all evolutionary change, the disease it generates that consumes a substantial proportion of our resources, and the means it offers for dissecting all facets of biological phenomena". With respect to human health, an interest in understanding the molecular basis of mutagenesis has been strongly stimulated because of its natural relevance to the studies on the understanding of the origin of genetic diseases, and also by the discovery that mutations occurring in somatic cells are important in the development of cancers.

In this section, a number of terms (specifically pertaining to the subject of this thesis) that are used when discussing mutations and their consequences have been defined. In addition, some concepts regarding the nature and types of mutations, methods (assay systems) of detecting them, causes and pathways of spontaneous mutagenesis, and identification of the repair mechanisms etc. are discussed.

1.2 Mutations and mutants

A mutation is a heritable change in the sequence of an organism's genome; an organism that carries one or more mutations in its genome is referred to as a mutant. The genetic information that an organism encodes in its genome is referred to as its genotype, and the ensemble of observable characteristics of an organism is referred to as its phenotype. When comparing two organisms that differ by one or more mutations, the "normal" or "parental" organism is often referred to as having a wild-type phenotype. The consequence of a mutation can vary to different extents: it may not alter the phenotype at all, it may change one or more aspects of the phenotype, or it may even lead to inviability of the organism. Different mutations that are situated at the same genetic locus are said to be alleles of each other.

A mutation that changes the phenotype from wild-type to a mutant phenotype is a forward mutation, whereas a mutation that restores the wild-type phenotype in an organism
which had previously suffered a forward mutation is said to be a *reversion* mutation. If a condition can be set in which all the members of the population die or fail to grow except for the desired class of mutants, it is called as *selection*. If all the members of the population are examined to identify the mutants having the desired phenotype, then it is referred to as *screening*.

The process by which mutations are produced is referred to as *mutagenesis*. A *mutagen* is an agent that leads to an increase in the frequency of occurrence of mutations. Mutagenesis that occurs without treatment of the organism with an exogenous mutagen is referred to as *spontaneous mutagenesis*. Spontaneous mutations can occur because of uncorrected replication errors, or can arise as a consequence of lesions (i.e., damage to the nitrogen bases by spontaneous deamination, depurination, alkylation or oxidation) that are introduced in DNA during the normal life of the cell. In addition, spontaneous mutations can also arise as a consequence of other types of mechanisms such as insertions of transposable genetic elements and rearrangements caused by them, or by recombination between homologous or partially homologous sequences. *Induced mutagenesis* results from the treatment of an organism with a chemical or a physical mutagen. This chapter focuses mostly on the occurrence and pathways of spontaneous mutagenesis.

The *mutant frequency* refers to the proportion of mutants in a population. The *mutation rate* is a measure of the number of mutations giving a particular scorable phenotype per DNA replication; sometimes this can be expressed as mutations per base pair per replication. The *genomic mutation rate* expresses mutations per genome per DNA replication, which has been calculated as ~ 0.003 mutations per genome per DNA replication and shown by Drake (63) to be constant for diverse DNA-based microbes.

### 1.3 Nature of mutations

For convenience, mutations can be classified into categories by specifying whether they are point mutations or not. *Point mutations* are mutations that result from (i) *base substitutions*, i.e., substitution of one base pair for another or (ii) the addition/deletion of a small number of base pairs (*frameshifts*). The base substitutions can be of two kinds, *transitions* which involve the substitution of one purine for another or one pyrimidine for another on the same strand, or *transversions*, that involve the substitution of a purine for a pyrimidine or vice versa on the same strand. Thus, there are two kinds of transitions (\(A:T \rightarrow G:C\) or \(G:C \rightarrow A:T\)) and four of transversions (\(A:T \rightarrow T:A\), \(A:T \rightarrow C:G\), \(C:G \rightarrow G:C\) or \(G:C \rightarrow T:A\)).
The effects of mutations that occur within the protein-coding regions allow the substitution mutations to be classified into nonsense, missense, and silent mutations. 

**Nonsense** mutations are mutations that change an amino acid codon to one of the three stop codons: TAG (amber), TAA (ochre), or TGA (opal), resulting most frequently in premature termination of protein synthesis and hence loss of activity of that particular gene product. Nonsense mutations are useful in genetic studies, because their phenotypes can be suppressed by a second type of mutation termed a nonsense suppressor. A **nonsense suppressor** is a mutation occurring within a gene encoding a tRNA and alters the anticodon of that tRNA so that it can recognize a nonsense codon. **Missense** mutations are base substitution mutations that change the codon for one amino acid to that for another; if such an amino acid change in the gene product results in no phenotypic consequence, the mutation is referred to as neutral. **Silent** mutations represent synonymous codon changes that result in no alteration of the amino acid sequence.

In contrast to base substitutions, **frameshift mutations** alter the translational reading frame, most often resulting in proteins which will not have normal activity. Depending on the number of base pairs inserted or deleted, they are classified as +1, +2, −1 or −2 frameshifts. The basis for frameshift mutagenesis was postulated by Streisinger et al. (222) in which they proposed that frameshift mutations arise when misalignments between the primer and template occur during DNA synthesis, particularly in repetitive sequences such as homopolymeric tracts (198).

The other class of mutations involves more extensive changes in DNA sequence. These include deletions, insertions, duplications, and inversions that very often arise as a result of replication slippage or recombination between partially homologous sequences within the genome or are mediated by the action of transposable elements.

### 1.4 Methods to detect mutagenesis

Mutations are inherently very rare genetic events. Therefore, to study mutagenesis and to understand the molecular mechanisms responsible for it, there was a need to develop assay systems that permit convenient detection and analysis of mutations. A wide variety of detection systems that use either reversions or forward mutations have been developed over the years. Earlier, phage resistance was extensively used to measure mutation rates (12, 154) and subsequently, a streptomycin-dependent strain that could revert to streptomycin-independence was used to examine various chemicals for mutagenic activity (59). One general and easy reversion assay uses auxotrophic mutations (e.g., *trpA*, *tyrA*, or *hisG* are
commonly used alleles, but in principle any auxotrophic mutation that can revert, can be
used) that can revert to prototrophy. One important system was developed by Yanofsky and
coworkers (13) in which strains carrying specific alleles of the trpA locus could be used in
reversion assays to monitor all the possible base substitution events.

The best known use of reversion systems has been the work of Bruce Ames and his
colleagues, who constructed a set of histidine-requiring strains of Salmonella enterica
(previously known as S. typhimurium) (also carrying many other mutations that improve the
mutagen detection potential) and this system is widely used to monitor the mutagenicity
and/or carcinogenicity of various compounds in the eponymous Ames assay (5).

Strains with mutations in lac genes that can revert to Lac+ are also commonly used to
study mutational mechanisms. A recently developed lac reversion system is based on the
specific requirement for a glutamic acid residue at position 461 in E.coli β-galactosidase, and
consists of a set of strains that can determine each of the six base substitutions by monitoring
the number of Lac+ colonies generated (54). Systems that detect exclusively frameshift
mutations are also available (55). Nonsense mutations in lac have also been used in a number
of ways to detect mutagenic specificity (169). Another system that has been developed
recently is based on the reversion of the galK2 allele (186).

Most of the systems described above are examples of reversion systems that employ
direct selection of revertants. The advantage of reversion systems is that they are simple to
use and are ideal for screening large numbers of mutagens or conditions that may influence
mutation frequency. However, the disadvantage with reversion systems is that they are biased
for the detection of a few specific events at selected sites. In principle, a forward mutational
system will not have this limitation and can give rise to an unbiased collection of mutants as
it monitors the mutations of a wild-type gene to a nonfunctional variant. An extensively used
forward mutational assay developed by Miller and his colleagues involves the use of lacI
repressor gene (170). Some other commonly used forward mutational assays are: λ cI+ to λ
cI− or phage-sensitivity to phage-resistance.

There are some mutational assays, e.g., valine<sup>5</sup> to valine<sup>+</sup> or antibiotic-sensitivity to
antibiotic-resistance (such as resistance to rifampicin, streptomycin and nalidixic acid etc.)
which are intermediate to the reversion and the forward mutation assays. In particular, the
selection for antibiotic resistance (e.g., Rif<sup>+</sup> to Rif<sup>−</sup>, Str<sup>+</sup> to Str<sup>−</sup> or Nan<sup>5</sup> to Nan<sup>+</sup>) is a very
routinely used assay in most of the laboratories to measure the spontaneous mutation
frequency due to the ease of the technique, even though this assay is biased for sampling exclusively point mutations.

1.5 **Generation of spontaneous mutations**

As the genetic material is double-stranded DNA, the generation of spontaneous mutation generally requires two steps, since the alteration in each of the two strands does not occur simultaneously. In the initial step, the sequence on one strand is altered (for example, by mispairing during the DNA replication or by the presence of lesions caused by DNA damage); in the next step, during the semiconservative DNA replication, the alteration 'instructs' a corresponding change in the second strand, and mutation is now said to be **fixed**.

Not all the alterations that occur in the first step will be manifested as mutations as most of them would be repaired by the various cellular repair mechanisms before the fixation. Because mutations are rare events, (~$10^{-10}$-$10^{-9}$/bp/cell/generation is the typical mutation frequency observed in wild-type *Escherichia coli*) the molecular origin of a mutation that arises spontaneously in a chromosome is almost impossible to specify and direct biochemical studies are not feasible. Therefore, to obtain clues, albeit indirect, to understand the mechanisms of mutagenesis, genetic studies have been extensively used.

The great power of the *E. coli* system to study mutations is its amenability for use in genetic dissections of various physiological functions (that is, the third 'face' in Drake's famous quotation). In particular, the analyses of mutator/antimutator mutations (that alter the spontaneous mutation frequency) have permitted the inference of various mechanisms by which organisms produce or prevent mutations. **Mutators** are strains with elevated spontaneous mutation rates, whereas **antimutators** are strains with lowered mutation rates, relative to the wild-type. Both mutators and antimutators reveal the pathways of mutation generation or avoidance. However, the two types of mutants answer different aspects of this problem. Mutators have defects in pathways that cells use to prevent mutations, and their study therefore discloses the potential sources of mutations (i.e., mutations that could have occurred, but do not). In contrast, antimutators can reveal the actual sources of spontaneous mutations, because their action prevents mutations that normally occur in wild-type cells when all the other error correction pathways are functional.

1.6 **Identification of spontaneous mutators**

Strains with higher mutation rates than wild-type were observed in *Drosophila* species over 50 years ago (189) and were first described in *Escherichia coli* in 1954 by Treffers (233). These very early mutator variants were discovered among laboratory strains,
but later on many procedures were designed to detect mutators in a population. Initially, mutagenized cells were screened for increased rate of either reversion of auxotrophic markers or of mutation to antibiotic resistance (119, 122). The use of strains that were capable of giving rise to visible mutant sectoring in single colonies also aided in the identification of mutators (169).

The detection of mutator strains was greatly facilitated by the use of papillation techniques, which dates back to as early as 1907 when Massini employed them to characterize the properties of *Escherichia coli mutabile*, a Lac\(^-\) strain that reverted to Lac\(^+\). The principle of papillation is very simple and can be illustrated with the example of Lac\(^-\) to Lac\(^+\) reversions. Cells of a Lac\(^-\) strain are plated to grow as colonies on medium containing a utilizable carbon source along with lactose as the second carbon source. After colony growth has ceased following local depletion of the former, only Lac\(^+\) mutant cells within each colony that can use the unutilized lactose can grow further and form microcolonies or papillae on the surface of the primary colonies. Lactose indicator medium (e.g., MacConkey-lactose) has frequently been used for Lac\(^-\) strains to study Lac\(^+\) papillation, and its use has led to the isolation of numerous mutator strains (38, 167, 184).

Selection methods have also generated mutators. In these cases, mutagenized cells are subjected to a selection, such as resistance to streptomycin or reversion of auxotrophic mutations (52, 122). A small fraction of all spontaneous Str\(^r\) mutants were shown to have arisen in mutator strains (52), and when six auxotrophic markers that revert by frameshifts were used in successive selections after mutagenesis, virtually all the surviving colonies were frameshift mutators (122). Natural bacterial populations were also shown to contain mutators at a high frequency (140, 228).

The identification of antimutator strains has received much less attention than has mutators, because of the rarity of their occurrence. Antimutator alleles of T4 DNA polymerase were isolated in late 1960's by Drake and were shown to reduce A:T→G:C transition mutations (65). Antimutator mutants with increased DNA replication fidelity were also isolated in *Escherichia coli* using the *galK2* papillation screen (76).

Extensive analysis of mutators and antimutators by both genetic and biochemical means facilitated the discovery of numerous DNA repair pathways and identification of the causes of spontaneous mutagenesis. Several of these mutational repair mechanisms are described in the following sections.
All the mutators and antimutators that have been detected in *E.coli* are listed, and their salient features are summarized in Tables 1.1 and 1.2, respectively.

1.7 DNA repair and error avoidance systems

In most bacteria including *E.coli*, the fidelity of DNA replication is determined by three main processes: base selection by the polymerase subunit of DNA polymerase III, editing by the associated 3'-5' exonuclease activity (proofreading), and post-replicative DNA mismatch repair (reviewed in 67). In addition to the above mechanisms, the organisms elaborate many other specific repair systems that correct the spontaneous lesions (e.g., damage caused by spontaneous deamination, depurination, alkylation or oxidation of bases) occurring in DNA during its normal growth/metabolism. With all these fidelity and repair mechanisms operating both during and after DNA replication, the mutation rate typically approaches $\sim 10^{-10}$ per base pair per round of replication.

(A) Base selection pathway

During the process of chromosomal DNA replication in *E.coli*, base selection and incorporation is performed by *dnaE* gene product, the α (i.e., polymerase) subunit of DNA polymerase III holoenzyme. The mutations in *dnaE* (that affect the base selection; null mutations in the gene are not viable, as it is essential for DNA replication) fall into two classes, mutators that show increased mutation frequency or antimutators that exhibit lowered mutation rates (75, 76). The genetic analysis of these mutants and also the measurements of the base insertion specificities by the mutant polymerases, demonstrated that the base selection in the polymerization reaction is an important determinant in maintaining the fidelity of DNA replication.

(B) Editing by 3'→5' exonuclease

The concept of exonucleolytic editing was introduced by Brutlag and Kornberg in 1972, who showed that the 3'-exonuclease of DNA polymerase I removed incorrectly paired terminal nucleotides more rapidly than it did correctly paired nucleotides (34). In the mid-1970's mutations in the gene *dnaQ* (originally termed *mutD*) that encodes the epsilon subunit of DNA polymerase III (*mutD5* and *dnaQ49*) that result in a very strong mutator phenotype were identified (58, reviewed in 52). The characterization of these mutators and analysis of the mutational spectrum in these mutants along with extensive biochemical studies of the wild-type and mutant proteins (68) defined the essential proofreading function (that is, of
<table>
<thead>
<tr>
<th>Mutator mutation</th>
<th>Function of the wild-type gene</th>
<th>Spectrum of mutations generated</th>
</tr>
</thead>
<tbody>
<tr>
<td>dnaE</td>
<td>The polymerase subunit of DNA polymerase III</td>
<td>All base substitutions and frameshifts</td>
</tr>
<tr>
<td>dnaQ (mutD)</td>
<td>The editing 3'→5' exonuclease subunit of DNA polymerase III</td>
<td>All base substitutions, (mostly transversions), and frameshifts</td>
</tr>
<tr>
<td>mutH</td>
<td>An endonuclease involved in methyl-mismatch repair</td>
<td>Transitions and frameshifts</td>
</tr>
<tr>
<td>mutL</td>
<td>Involved in methyl-mismatch repair</td>
<td>Transitions and frameshifts</td>
</tr>
<tr>
<td>mutS</td>
<td>Mismatch binding protein in methyl-mismatch repair</td>
<td>Transitions and frameshifts</td>
</tr>
<tr>
<td>uvrD (mutU)</td>
<td>Encodes helicase II, involved in methyl-mismatch repair</td>
<td>Transitions and frameshifts</td>
</tr>
<tr>
<td>dam</td>
<td>Encodes DNA adenine methyltransferase</td>
<td>Transitions and frameshifts</td>
</tr>
<tr>
<td>mutT</td>
<td>Codes for 8-oxo dGTPase</td>
<td>A:T→C:G transversions</td>
</tr>
<tr>
<td>mutY</td>
<td>Encodes an adenine glycosylase</td>
<td>G:C→T:A transversions</td>
</tr>
<tr>
<td>mutM</td>
<td>Encodes Fapy glycosylase</td>
<td>G:C→T:A transversions</td>
</tr>
<tr>
<td>mutA</td>
<td>Encodes glyV tRNA</td>
<td>A:T→T:A transversions</td>
</tr>
<tr>
<td>mutC</td>
<td>Encodes glyW tRNA</td>
<td>A:T→T:A transversions</td>
</tr>
<tr>
<td>ung</td>
<td>Encodes uracil DNA glycosylase</td>
<td>G:C→A:T transitions</td>
</tr>
<tr>
<td>vsr</td>
<td>Codes for Vsr endonuclease</td>
<td>G:C→A:T transitions at 5meC sites</td>
</tr>
<tr>
<td>oxyR</td>
<td>Positive regulator of hydrogen peroxide response</td>
<td>-</td>
</tr>
<tr>
<td>sodA and sodB</td>
<td>Encode superoxide dismutases, MnSOD and FeSOD, respectively</td>
<td>-</td>
</tr>
<tr>
<td>fur</td>
<td>A negative regulator of iron assimilation</td>
<td>-</td>
</tr>
<tr>
<td>xthA and nfo</td>
<td>Encode exonuclease III and endonuclease IV, respectively</td>
<td>-</td>
</tr>
<tr>
<td>nth and nei</td>
<td>Encode endonucleases III and VIII, respectively</td>
<td>G:C→A:T transitions</td>
</tr>
<tr>
<td>ada and ogt</td>
<td>Encode methyl transferases that remove methyl groups from methylated bases</td>
<td>G:C→A:T and A:T→G:C transitions</td>
</tr>
<tr>
<td>miaA</td>
<td>Encodes an isopentenylpyrophosphate transferase, that modifies tRNAs</td>
<td>G:C→T:A transversions</td>
</tr>
<tr>
<td>recA (constitutive)</td>
<td>Codes for RecA, a recombination protein</td>
<td>A:T→T:A transversions</td>
</tr>
<tr>
<td>polA</td>
<td>Codes for DNA polymerase I</td>
<td>Frameshifts and deletions</td>
</tr>
<tr>
<td>ndk</td>
<td>Encodes nucleoside diphosphate kinase</td>
<td>-</td>
</tr>
<tr>
<td>dut</td>
<td>Encodes deoxyuridine triphosphatase</td>
<td>-</td>
</tr>
<tr>
<td>dinB+ (multicopy)</td>
<td>Encodes DNA polymerase IV</td>
<td>Base substitutions and single nucleotide deletions</td>
</tr>
<tr>
<td>Locus/gene</td>
<td>spectrum of mutations generated</td>
<td>Function of the wild-type gene</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>dnaE</td>
<td>A:T→T:A transversions</td>
<td>Polymerase subunit of DNA polymerase III</td>
</tr>
<tr>
<td>Multicopy mutL*</td>
<td></td>
<td>Involved in methyl-mismatch repair system</td>
</tr>
<tr>
<td>Multicopy mutY*</td>
<td></td>
<td>Encodes an adenine DNA glycosylase</td>
</tr>
<tr>
<td>Multicopy oxyS*</td>
<td></td>
<td>Encodes a regulator RNA molecule involved in oxidative stress</td>
</tr>
<tr>
<td>Multicopy sodA*</td>
<td></td>
<td>Codes for a Mn-dependent superoxide dismutase (MnSOD)</td>
</tr>
<tr>
<td>Multicopy dps*</td>
<td>G:C→T:A transversions</td>
<td>Encodes a stationary-phase specific DNA-binding protein</td>
</tr>
<tr>
<td>Multicopy yggX*</td>
<td></td>
<td>May be a protein that protects iron-sulphur clusters?</td>
</tr>
<tr>
<td>dinB</td>
<td></td>
<td>Encodes an inducible DNA polymerase IV</td>
</tr>
<tr>
<td>umuC</td>
<td>A:T→T:A transversions</td>
<td>Encodes a subunit of DNA polymerase V</td>
</tr>
</tbody>
</table>

* The multicopy effects are observed in *Salmonella enterica.*
excising out the incorrectly paired terminal nucleotide from the newly synthesized strand) of this subunit.

(C) Post-replication mismatch repair pathways

Bacteria and eukaryotic cells possess several distinct mismatch repair pathways. A very significant and well understood pathway, termed methyl-directed mismatch repair pathway or long-patch mismatch correction system, operates immediately after DNA replication. This pathway is characterized by broad mismatch specificity and is responsible for correcting DNA replication errors (both base substitutions and frameshifts) and also for processing recombination intermediates that contain mismatched base pairs. Short-patch systems, which exhibit restricted mismatch specificity and whose primary function may be processing of particular chemically or enzymatically damaged base pairs, are discussed later.

1. Methyl-directed mismatch repair system

This is one of the best understood repair systems in E.coli and is also known to occur in many other organisms (reviewed in 174 and 175). The existence of mismatch repair functions was predicted by Robin Holliday in 1964, as a correction system for DNA base pair mismatches within recombination intermediates that lead to gene conversion events (123). In the same year, Evelyn Witkin also independently suggested the existence of mismatch repair on the basis of experiments on bacterial mutagenesis (244). In 1976, Wagner and Meselson suggested that daughter strand-specific processing of mismatches in semi-conservatively replicated DNA could serve to eliminate DNA biosynthetic errors and thus contribute to the overall fidelity of chromosome replication (234). Although the existence of a mismatch repair system was first indicated by the presence of high negative interference in crosses with phage lambda and by mismatch heteroduplex experiments (234, 242), the major breakthrough came from the studies of mutators resulting from defects in the mutH, mutL, mutS, and uvrD genes (reviewed in 52 and 195).

Combined genetic and biochemical analyses have implicated ten proteins in methyl-directed mismatch correction: the products of the mutH, mutL, mutS genes, DNA helicase II (the product of the mutU or uvrD gene), single-stranded DNA binding protein (SSB), exonuclease I, exonuclease VII, RecJ exonuclease, the replicative DNA polymerase III holoenzyme, and DNA ligase (149). The homogeneous preparations of these proteins were sufficient to reconstitute in vitro a methyl-directed reaction with the same specificity as that observed in vivo (138). Of the eight possible base-base mispairs, the G-T and A-C transition mismatches and the G-G and A-A transversion mispairs are typically well corrected, while
the T-T, C-T, and G-A transversion pairing errors are weak substrates, and the C-C mispair is totally refractory to the correction. The specificity of this system is not restricted to base-base mismatches, because efficient correction of mismatches involving insertion/deletion of up to four nucleotides is also seen both in vivo and in vitro.

(i) Mechanism of methyl-directed mismatch correction

Analysis of the mechanism of the methyl-directed reaction has been facilitated by the development of the biochemical assay that permits mismatch correction to be scored in a cell-free system (149). Repair of the mismatch is specifically targeted to the newly synthesized daughter strand, and this strand specificity for the repair is provided by patterns of adenine methylation in d(GATC) sequences in the genome. The d(GATC) sequence is a palindrome and is ordinarily methylated at the A residue on both strands by the Dam methylase, product of the \( \text{dam} \) gene. In newly replicated DNA, the sequence is transiently in a hemi-methylated state and it is the absence of methylation on the new strand that serves as the signal to target the process of mismatch correction to this strand. A single d(GATC) sequence in the vicinity of the mismatch is sufficient to direct the mismatch repair. Optimally, the distance separating the GATC sequence from the mismatch should be less than two kb.

The repair is initiated by the binding of MutS to the mismatch. Such binding is dependent on ATP and promotes the formation of \( \alpha \)-shaped loop structures, which are predicted to be intermediates in mismatch correction. The \( \text{mutH} \) gene product is responsible for d(GATC) site recognition and for the incision in the unmethylated strand of hemimethylated DNA 5' to the G of the tetranucleotide sequence. In vitro, MutH has been shown to exhibit a weak, \( \text{Mg}^{2+} \)-dependent d(GATC) endonuclease activity that becomes activated during the assembly of a repair complex on the heteroduplex. In contrast to MutS and MutH, no simple activity has been attributed to MutL, but one possibility that has been suggested is that MutL functions as a protein-protein interface between MutS and MutH.

The excision reaction in the repair pathway is dependent on the strand break by the MutH and subsequent action of the exonucleases in the formation of excision tracts. It has been shown that methyl-directed system has a bidirectional capability, since the mismatches are corrected efficiently irrespective of their orientation to the d(GATC) sites in the unmethylated strand. Different exonucleases are required for repair of the two hemimethylated configurations of a heteroduplex containing a single d(GATC) site. When the unmethylated d(GATC) sequence resides 3' to the mismatch, a 3'→5' hydrolytic activity provided by exonuclease I is necessary, whereas correction of the mismatch in alternate
configuration requires a 5'→3' exonuclease that is provided by either exonuclease VII or RecJ exonuclease. However, it has been observed that the triple mutant lacking all the three exonucleases does not exhibit a mutator phenotype, suggesting the ability of at least one other redundant exonuclease activity to participate in the mismatch correction pathway (116). Recently, a new exonuclease (designated as Exo X) with 3'→5' exonucleolytic activity has been identified and it has been shown that a mutant defective in all the four exonucleases exhibits mutator activity; the cell-free extract of this quadruple mutant is unable to support the mismatch reaction in vitro (35).

The role of DNA helicase II in the methyl mismatch repair pathway has not been clearly defined, but it is an obvious candidate for bidirectional unwinding of the incised strand, to render it susceptible to single-strand hydrolytic activity of the exonucleases. Even though it has been shown that polarity of UvrD helicase in unwinding DNA is from 3' to 5', yet this single enzyme is sufficient to unwind DNA from both directions in the mismatch reaction. The last step of the methyl-directed reaction is gap repair synthesis by DNA polymerase III holoenzyme and ligation of the repair product by DNA ligase (174).

(ii) Role of methyl-directed repair in replication fidelity

The strand-specific nature and the broad mismatch specificity of the methyl-directed pathway are consistent with the Wagner-Meselson proposal that this type of system could function in the postreplication elimination of DNA biosynthetic errors. This idea has received wide support from both genetic and biochemical experiments, and it is now clear that this system contributes in a major way to the fidelity of chromosome replication.

Mutations in mutH, -L, -S or uvrD result in a 100- to 1000-fold increase in spontaneous mutability, with the elevated mutation rate comparable to the in vitro error rate of the replicative DNA polymerase holoenzyme. Mutations accruing in these mutator strains are mostly transitions and single-nucleotide deletions, with few transversions; this mutational spectrum is consistent with the misincorporation propensity of the DNA polymerase III holoenzyme (211). Mutations leading to a deficiency of Dam methylase, the enzyme responsible for d(GATC) methylation, also result in elevated mutation rates, primarily because of the loss of strand discrimination in targeting of mismatch repair (96).

In addition to their mutator phenotype under normal growth conditions, strains defective in methyl-mismatch repair are hypersensitive to base-substitution mutagens such as 5-bromouracil and 2-aminopurine, and to frameshift mutagens like 9-aminoacridine. Since these base analogues and frameshift mutagens are known to exert their action at the level of
DNA synthesis, these findings are also indicative of the importance of methyl-directed repair of lesions arising during replication (97).

2. Oxidative damage repair

Reactive oxygen species can damage DNA and result in increased mutations, and this in turn has been linked by some workers to carcinogenesis and aging (4, 45). These oxygen free radicals are generated by normal cellular metabolism as well as by exogenous sources such as ionizing radiation or by exposure to oxidizing agents. Reactive oxygen species that damage DNA include superoxide radical, OH radical, and singlet oxygen. Oxidative damage to DNA can result in a variety of alterations such as strand breaks, deamination, depurination, ring-opened purines, thymine or cytosine glycols, and also in the generation of 2'-deoxy-7, 8-dihydro-8-oxoguanosine, abbreviated as 8-oxodG (166). Of these, 8-oxodG is implicated in causing a significantly high proportion of spontaneous mutations; the basis for this increased mutation frequency and the repair mechanisms of 8-oxodG lesions are discussed in the section below.

(i) The GO system for oxidative damage repair

The isolation and characterization of appropriate mutator strains have mainly been responsible for the discovery and elucidation of a three-component system that prevents mutations caused by the oxidative lesion 8-oxodG, or the "GO" lesion. The GO system is an error avoidance "short patch" repair pathway devoted to enhancing the fidelity of DNA replication and is composed of at least three proteins, MutT, MutY, and MutM. Unlike other systems involved in the repair of oxidative DNA damage that are discussed in the subsequent sections, the GO system appears to be equally important during both aerobic and anaerobic growth of E.coli.

Among the many mutators, mutT has drawn particular attention because it was the very first mutator found in E.coli, by Treffers (233) in 1954. The mutT gene was cloned and the purified protein was shown to have a specific 8-oxodGTP hydrolyzing activity, thus preventing incorporation of 8-oxodG into DNA by degradation of the cytoplasmic pools of 8-oxodGTP (160). Cells lacking the MutT protein are strong mutators (that specifically suffer A:T→C:G transversions, [248]), indicating the potent mutagenic activity of the GO lesion. MutT has also been shown recently to have an 8-oxo-rGTPase activity that leads to hydrolysis of the cytoplasmic pools of 8-oxo-rGTP: this observation suggests an additional role for MutT in preventing the transcriptional errors (227).
The other two mutators involved in this pathway, *mutY* and *mutM*, were detected in Lac\(^+\) papillation assays as mutations that specifically increased G:C→T:A transversions (38, 184). Cloning, sequencing and identification of the enzymatic activity after protein purification revealed that *mutM* encodes a formamidopyrimidine (or Fapy) DNA glycosylase that removes 8-oxoG from 8-oxoG:C pair in double-stranded DNA (38). The *mutY* gene encodes an adenine glycosylase, which removes A from G:A and 8-oxoG:A mispairs (7).

The GO repair pathway is believed to function in the following manner. Oxidation of guanine in the DNA would form an 8-oxoG:C pair which would potentially induce G:C→T:A transversion if the 8-oxodG is not removed by MutM protein. This is because DNA polymerase III inserts either A or C residue with equal efficiency against an unrepaired 8-oxodG base in the template strand during DNA replication. Even so, in most such instances, where 8-oxoG has not been removed from 8-oxoG:C by the MutM protein, the resulting 8-oxoG:A pair in newly replicated DNA is reversed back to 8-oxoG:C by the action of MutY protein which removes the adenine from the 8-oxoG:A mispair. These concerted actions of MutM and MutY explain why defects in either of these genes result in mutators with increased G:C→T:A transversions. MutT acts at a different level in the GO repair pathway, by reducing the cytoplasmic pool of 8-oxodGTP and hence reducing the likelihood of misincorporation of this nucleotide against an A residue in the template; hence *mutT* mutators exhibit an increased frequency of A:T→C:G transversions.

Although the mutator effects are moderate for *mutM* and *mutY* single mutants, the double mutants lacking both *mutM* and *mutY* activities have very high mutation rates; this mutator strength reflects not only the high levels of spontaneous 8-oxodG formation in DNA but also the need for the concerted action of MutY and MutM proteins in elimination of the GO lesions in cells. The mutator strains *mutM* and *mutY*, but not *mutT*, have comparable mutation rates after growth under aerobic or anaerobic conditions, suggesting that endogenous processes are sufficient to produce the active oxygen species that lead to GO lesions in these strains (91, 166).

(ii) Other pathways of oxidative damage repair

There are many inducible enzyme systems that can neutralize reactive oxygen species before they damage DNA. The inducible enzymes include two superoxide dismutases that convert superoxides to hydrogen peroxide and two catalases that convert hydrogen peroxide to water. The regulation of these enzymes is well studied and they fall into one of three
regulons controlled, respectively, by \( \text{oxyR} \) (a gene that positively regulates hydrogen peroxide-induced response), \( \text{soxR} \) (a positive regulator of superoxide-induced genes), or \( \text{rpoS} \) (encoding \( \sigma^E \), a stationary-phase specific sigma factor) (reviewed in 60 and 71).

Earlier work has suggested that deletions of the \( \text{oxyR} \) locus, a positive regulator for a number of enzymes including a catalase (HPI) and a detoxifying enzyme of hydroperoxides (Ahp), result in both a mutator phenotype (which is more pronounced in aerobic conditions) and increased sensitivity to oxidizing agents (100). However, Yamamura et al. (247) showed recently that mutation in \( \text{oxyR} \) does not increase the spontaneous mutation frequency in two forward mutational assays. The hydrogen peroxide response has been shown by Altuvia et al. (3) to also be regulated by a small stable RNA called \( \text{oxyS} \), in addition to \( \text{oxyR} \); the \( \text{oxyS} \) gene in multiple copies confers an antimutator activity, and they propose that a defense system, possibly methyl-directed mismatch repair, regulated directly or indirectly by the \( \text{oxyS} \) RNA, is contributing to the protection against oxidative DNA damage (3).

Mutations in \( \text{sodA} \) gene, encoding a manganese superoxide dismutase MnSOD, confer a weak mutator phenotype (72). Conversely, mutation in \( \text{sodB} \), encoding an iron superoxide dismutase FeSOD, does not result in a mutator phenotype, but the \( \text{sodA sodB} \) double mutant exhibits a moderate mutator phenotype. The enhanced mutagenesis in \( \text{sodA sodB} \) strain is dependent on the presence of oxygen, (i.e., is seen only in aerobic growth) thus confirming the role of molecular oxygen in spontaneous mutagenesis (72). This observation suggests (i) that oxygen causes the generation of superoxide radicals, (ii) that the SodA and SodB enzymes can detoxify these radicals in wild-type cells, and (iii) that in the absence of these enzymes the effect of superoxide radicals is manifested as increased spontaneous mutagenesis.

Touati et al. (232) demonstrated that mutations in \( \text{fur} \) gene also increase spontaneous mutagenesis to a marginal extent. The Fur protein is a regulator of iron metabolism, and iron assimilation is constitutive in \( \text{fur} \) mutants. The mutator activity in this mutant was observed to be oxygen-dependent, permitting them to predict that iron overload leads to increased mutations in aerobic conditions. However, an earlier study (48) had reported that iron limitation increases the reversion frequency of the \( \text{trpA46} \) allele. Studies with M13 phage have shown that 5-hydroxy-2'-deoxyctydine (5-OH-dC), the oxidative damage product of C residues in DNA, is also mutagenic, leading to \( \text{C} \rightarrow \text{T} \) transitions (74).
**D) Repair of damage caused by depurination/depyrimidination**

Apurinic/apyrimidinic (AP) lesions are abasic sites in DNA that result from the cleavage of the N-glycosidic bond between the base and the deoxyribose sugar, leaving the DNA phosphodiester backbone intact. The spontaneous loss of purines and pyrimidines from DNA is the major source of AP sites. In addition, AP sites can be generated by the action of DNA glycosylases, that excise damaged bases generated by either oxidation or alkylation.

These AP sites are expected to be mutagenic as there is lack of information (base) in the template strand for DNA polymerase III to incorporate the correct base; it has been shown that under these conditions polymerase III has a preference to incorporate an adenine across an AP site during DNA replication (220). However, these AP sites are generally repaired by the sequential actions of an AP endonuclease, which initiates the repair by cleaving the phosphodiester bond in DNA, an exonuclease that subsequently removes the AP site, and finally DNA polymerase I that performs the repair synthesis. If AP sites are mutagenic, then mutants lacking AP endonuclease might be expected to exhibit a mutator phenotype.

The three-letter nomenclature for designating genes encoding various biochemically characterized exonucleases and endonucleases is as follows; the first letter x or n denotes exonuclease or endonuclease respectively, and the second and third denote the first two letters of the numeral; for example, xon encodes exonuclease I and nfi encodes endonuclease V. An E.coli strain, mutant for either exonuclease III (Xth, a misnomer, because it is also a major AP endonuclease in the cell) or endonuclease IV (encoded by nfo, that is positively regulated by SoxR) does not show mutator activity whereas a double mutant of xth and nfo exhibits only a weak/moderate mutator activity, suggesting the possible presence of redundant AP endonucleases in this pathway (136).

It has been shown that E.coli also possesses two DNA glycosylase/apurinic lyase activities with overlapping specificities, endonuclease III and endonuclease VIII, that recognize and remove oxidized pyrimidines from DNA (127). Both enzymes exhibit DNA glycosylase and also AP endonuclease activities. The mutant defective in endonuclease III (encoded by nth) shows a very weak mutator activity, while the endonuclease VIII (encoded by nei) mutant is not distinguishable from the wild-type. However, the nth nei double mutant exhibits a moderate/strong (20-fold) spontaneous mutator phenotype (with mutational specificity of G:C→A:T transitions), suggesting a role for oxidized pyrimidines in spontaneous mutagenesis (127).
(E) Deamination repair

Deamination of the bases (cytosine, adenine, guanine; also 5-methyl cytosine) containing exocyclic amino groups occurs spontaneously in pH- and temperature-dependent reactions of DNA. Some of these products of deamination can give rise to mutations since they miscode during DNA replication. The fact that the formation of uracil by cytosine deamination is mutagenic and biologically significant has been revealed by the mutator phenotype (66) of the cells lacking uracil DNA glycosylase (i.e., ung mutants are mutators for G:C→A:T transitions). The ung mutants fail to remove uracil from DNA that results from the spontaneous deamination of cytosine, with the C→U conversion leading to a mutation because uracil will pair with adenine in the next round of replication.

Spontaneous deamination of 5-methyl cytosine (5meC) in DNA can result in the formation of thymine. In *E. coli*, 5meC residues occur at CC(A/T)GG sequences; the second C in such a sequence is the target for methylation by Dcm methylase. When 5meC is deaminated a T:G mispair will result, and if it is not repaired a G:C→A:T transition would occur. Cells possess a "very short patch" or VSP repair pathway that excises T' from the T:G mispair. Mutants lacking VSP repair (i.e., Vsr endonuclease encoded by vsr gene) are mutators, but only for G:C→A:T transitions at 5meC sites (146). Interestingly, it has been shown that Vsr overexpression also causes a mutator phenotype, the most likely explanation for which is that the methyl-mismatch repair enzymes are inhibited (or saturated) under these conditions (155).

(F) Alkylation damage repair

The endogenous methylation of DNA bases also results in spontaneous mutations if not repaired by one of several methylation repair enzymes. The bases which are shown to be methylated most frequently are guanine and thymine, generating O6-methyl guanine and O4-methyl thymine, respectively. Both the methylated bases are mutagenic lesions because they can potentially mispair during DNA replication and generate, respectively, G:C→A:T or A:T→G:C transitions. The ada ogt genes encode redundant Ada and Ogt DNA methyl transferases, each of which transfers methyl groups from O6-methyl guanine and O4-methyl thymine, directly restoring them to guanine and thymine, respectively. Although mutants lacking either Ada or Ogt functions do not have an increased spontaneous mutation rate, the double mutant ada ogt is a mutator, confirming that either one is required to repair the lesions caused by methylation of endogenous metabolites (156).
1.8 Other pathways of mutagenesis

(A) Mutator tRNAs

Michaels et al. (167) identified two mutators \textit{mut}A and \textit{mut}C that generate predominantly G:C→T:A and A:T→T:A transversions. Both mutants were shown to have alterations in the anticodon of glycine tRNA (i.e., tRNAs encoded by \textit{gly}V and \textit{gly}W, respectively) resulting in mutant tRNAs that insert glycine at aspartate codons during translation (217). The authors postulated that the mutator tRNA effect may perhaps be exerted indirectly, by generating a mistranslated polymerase (i.e., the ε subunit of DNA polymerase III) deficient in proofreading (217). Later, it has been shown by Ren et al. (197) that the expression of \textit{mut}A mutator phenotype requires recombination functions such as those of RecA, -B, -C, or RuvABC, but not those of RecD, -F, -O, and -R; they hypothesize that the \textit{rec}BCD-dependent homologous recombination system is a component of the signal pathway that activates an error-prone DNA polymerase in \textit{mut}A cells, however, the mechanism of the mutator phenotype of \textit{mut}A or \textit{mut}C remains unclear.

Mutations in the \textit{mia}A gene (encoding a tRNA Δ²-isopentenylpyrophosphate transferase involved in modifications of certain tRNAs) confer a weak mutator effect, which is specific for the G:C→T:A transversion (48). However, the mechanism behind this mutator phenotype is poorly understood. It has recently been shown that the \textit{mia}A mutator activity is also dependent on homologous recombination functions, as observed for \textit{mut}A and \textit{mut}C mutators (250).

(B) Concentrations of dNTP pools

The imbalance in the cellular concentration of dNTP pools may also cause increased mutagenesis. It has been shown that a mutant defective in nucleoside diphosphate kinase (encoded by \textit{ndk}) that controls the cytoplasmic concentration of nucleoside triphosphates, exhibits a considerable mutator activity (150).

The incorporation of uracil into DNA may also occur as a consequence of increased concentration of dUTP pools and it may be mutagenic, as the mutants deficient in deoxyuridine triphosphatase (the enzyme that converts dUTP into dUMP) display a mutator phenotype (121). The mutator phenotype of \textit{mut}T mutants deficient in 8-oxodGTPase activity has been described in a previous section.
(C) SOS-dependent spontaneous mutagenic pathways

RecA730 belongs to a class of mutant RecA proteins that results in constitutive SOS activation even in the absence of exogenous DNA damage. RecA protein is involved in homologous recombination and also in the activation of SOS response (along with a repressor encoded by lexA) following DNA damage (235). E.coli strains carrying recA730 allele exhibit a strong spontaneous mutator activity (recA441 is another allele that shows both constitutive SOS activation and mutator activity at higher temperatures) (171, 239). The mutational spectrum analysis of recA730 strain shows that A:T→T:A transversions occur most frequently; this spectrum is similar to that of dnaQ, mutA, and mutC mutators, thus raising the possibility that the proofreading functions of wild-type DnaQ is not occurring during DNA replication in strains expressing RecA730 protein. The mutator activity conferred by the recA730 mutation is abolished in a ΔumuDC strain, indicating that functionally active UmuD'C proteins are required for their generation (239). Expression of UmuD'C is dependent on SOS activation, and this complex has been shown to have an error-prone DNA polymerase activity, designated Pol V.

It has also been shown that mutants either defective in RecA or UmuC proteins, carrying recA56 or umuC36 alleles, respectively, exhibit an antimutator phenotype during aerobic growth (15). These mutations reduce the spontaneous base substitution reversion frequency of several trpA alleles, although, no antimutator effect is observed using Rif forward mutation assay. The analysis of mutational spectrum in trpA revertants indicated that A:T→T:A transversion frequency is decreased (15), and this result is in accordance with those described for the recA730 allele.

Bates et al. (9) demonstrated that mutations in DNA polymerase I (encoded by polA) confer both chronic SOS-induction and a 7- to 10-fold higher mutation frequency as measured by two forward mutational assays, (i.e., mutations to valine and rifampicin resistance). The mutator activity of polA was abolished in cells carrying an additional umuC mutation, indicating that the mutator activity is through such constitutive SOS induction (as umuC is induced in SOS-damage and is responsible for most of the SOS-dependent mutagenesis).

It was shown in an early study that dinB (an SOS-inducible gene), when present in multiple copies, confers a mutator phenotype (129). Subsequently, it was demonstrated that dinB codes for DNA polymerase IV, and that mutations in dinB also exhibit a weak
antimutator activity in assays for Rif\(^{R}\) to Rif\(^{S}\) and for reversions of frameshifts and base substitutions in \(lac\) (221).

(D) Translational errors

Boe (21) has predicted that translational errors can also cause increased mutation rates. He showed that (i) streptomycin (in sublethal concentrations) can induce the formation of auxotrophic mutants in streptomycin-sensitive cells, but not in \(rpsL\) mutants resistant to streptomycin, and (ii) strains having hyper-accurate ribosomes show reduced mutation rates. However, his experiments did not rule out alternative explanations for the altered mutagenesis rates observed in these conditions.

(E) \(ppGpp\) mediated alteration of mutation frequency

The synthesis of \(ppGpp\), a signal nucleotide, is primarily controlled by the RelA protein, and the effect of \(ppGpp\) on mutation rates has also been studied (245). The reversion rates of two auxotrophic mutations were higher in \(relA^{+}\) strains than in \(relA^{-}\) strains (245). It has been proposed that mutation rates are dependent upon rates of transcription and upon the genes that regulate the level of \(ppGpp\) in the cell (246).

(F) Other antimutator alleles

It has been shown by Kolter and coworkers that Dps, a stationary-phase specific DNA-binding protein is an antimutator, if it is overexpressed in exponential phase of growth, specifically reducing the G:C→T:A mutation frequency (162). Interestingly, a mutation in \(dps\) does not confer mutator activity, and the basis for the antimutator activity of Dps is not clear. Very recently, it has been shown that YggX protein, when overexpressed confers an antimutator activity to the cells of \(S.\ enterica\) (99).

1.9 Spontaneous DNA rearrangements

Spontaneous DNA rearrangements can constitute a significant fraction of spontaneous mutations and, unlike point mutations, involve extensive changes in the DNA sequence. These rearrangements can broadly be classified into RecA-dependent (i.e., mediated by RecA protein that is involved in homologous recombination) or RecA-independent (also called as illegitimate recombination events).

(i) RecA-dependent rearrangements

The rearrangements that are RecA-dependent involve large homologies that are usually provided in the cell by the presence of repeated sequences. The sources for these repeat sequences include duplicated copies of any of the following: structural genes, accessory genetic elements, sequence motifs, or insertion sequences (8). Recombination
involving these repeats can lead to complex rearrangements, depending on the position and orientation of repeats (i.e., whether in 'direct' or 'invert' orientation) and their presence on the same replicon or on different replicons.

Insertion sequences and composite transposable elements are the major sources for the rearrangements because they are (i) normal constituents of many bacterial chromosomes and plasmids and (ii) usually occur in multiple copies. The rearrangements occur when recombination between these homologous sequences result in unequal exchanges of the material during the resolution of the recombination intermediates. The commonly occurring alterations are deletions, large duplications, and inversions.

(ii) RecA-independent rearrangements

These rearrangements most often occur due to misalignments during the DNA replication and are thus RecA-independent. They occur largely between short regions of sequence homology and usually include deletion or duplication of the sequence in between the repeated sequences. The rearrangements between repetitive sequence elements also contribute to the genomic instability of an organism and may generate mutations (2). Several molecular mechanisms have been proposed for the occurrence of RecA-independent rearrangements: simple slipped misalignment, sister-chromosome slipped misalignment, and single-strand annealing and all of them appear to occur during the process of DNA replication (36).

(iii) Insertions and rearrangements mediated by transposable elements

As it has been mentioned in the earlier section, transposable elements provide homology for the occurrence of DNA rearrangements leading to deletions, duplications, and inversions. In addition, they can also cause random insertion mutations by transposase-mediated transpositions within the same or another replicon. The transposition processes are independent of homologous recombination systems encoded by the host and are dependent on transposon-encoded and certain other host-encoded functions (53, 61).

1.10 Origin of spontaneous mutations in wild-type cells

Although studies on mutators provide insights into the repair mechanisms, the question remains: what is the cause for mutations observed in a wild-type strain when all the repair mechanisms are functioning normally? One possible explanation would be that they are simply the sum of the uncorrected lesions and errors that either escape through or are blind to the various repair systems. An example of this leakage may perhaps be seen by the presence of 5meC hotspots in the wild-type strain even though the VSP repair system (which
Fig. 1.1 Possible sources of mutations. Each box represents an example of a potential mechanism by which mutations can arise. [modified from Ref. 210]
normally corrects these errors) is operational in such strains. Another possibility is that at least some of them represent the consequences of basal levels of activity of error-prone DNA polymerases such as UmuD'C and DinB. A third possibility is that these mutations arise in "transient mutator" cells in which transcriptional or translational errors result in a mutator polypeptide e.g., a non-functional proof-reading subunit of DNA polymerase III, even though the corresponding gene is not mutated (185).

The antimutator studies answer these questions to some extent, for e.g., the dnaE antimutator reduces the mutation frequency ~2-fold for certain markers; this 2-fold reduction in the level of spontaneous mutations will imply that 50% of all mutations occurring in the wild-type result from some attribute of the wild-type dnaE allele (75, 76). Furthermore, recA56, umuC36, and dinB alleles also behave as weak antimutators, suggesting a possibility of the role of error-prone polymerases such as UmuD'C and DinB in generating the spontaneous mutations in wild-type cells. Nonetheless, it is generally accepted that the question regarding the mechanisms responsible for the occurrence of mutations in wild-type cells is still open.

The emphasis of this thesis work is largely on the (i) occurrence of spontaneous mutations in nondividing cells and the mechanisms of their generation, and (ii) precise excisions of transposon Tn10. In the next section an overview on "the mechanisms of spontaneous mutation in nondividing cells" is presented and the subsequent section deals with "transposon Tn10 and its associated excision events".

Section B: Mechanisms of spontaneous mutation in nondividing cells

1.11 Introduction to adaptive mutations

A fundamental precept of neo-Darwinian evolutionary theory is the notion that, in biology, mutation and selection occur independently of one another. In other words, the probability of occurrence of a mutation is independent of selection pressure or its anticipated utility. It is commonly believed that mutations occur randomly, that is, without regard to their immediate utility or fitness. The randomness of spontaneous mutations and their independence from selection pressure was inferred from the elegant experiments reported in 1943 by Luria and Delbruck (154). Their studies were based on the distribution of bacteriophage T1 resistant mutants (that is, fluctuation tests) in Escherichia coli (154), and they concluded that mutations to phage resistance occur during growth, independently of the
action of virus. Their observations were further confirmed in the early 1950's by studies such as redistribution experiments by Newcombe (183), replica plating tests by Lederberg and Lederberg (142), and sib-selection experiments by Cavalli-Sforza and Lederberg (44).

In late 1950's, Francis Ryan (204) examined reversions to prototrophy in a His− mutant of *Escherichia coli*, but he noticed that His+ revertants continued to arise for a long period, in a medium devoid of histidine (i.e., in starved or stationary-phase populations). He attempted to account for the continued appearance of these mutants, and showed in a series of papers that they did not result from slow growth, phenotypic lag or cell turnover (182, 205). Left with no explanation other than that DNA synthesis might occur in the absence of cell division, he tried to measure the DNA turnover but failed to document the requisite amount of DNA replication in nondividing cells (206). His final conclusion was that a small amount of DNA must be synthesized in stationary-phase cells, (but by some process that increased the error rate) and that mutations could occur in such cells (206).

Notwithstanding Ryan's standing and the rigor of his work, however, and perhaps in part because of his untimely demise, the idea that spontaneous mutations could occur in nondividing cells was not pursued by others. Instead, the overwhelming impact of the classic Luria-Delbruck experiment, the constant use of growing cells for studies on mutagenesis, and the initial assumptions that all spontaneous mutations originate as errors made during genomic replication, together led to an axiom that all mutations arise randomly in growing bacterial populations.

In 1988, John Cairns and his collaborators (40) challenged the notions (i) that the occurrence of mutations is confined to dividing cells, and (ii) more importantly, that their occurrence is independent of their immediate utility; these authors argued that the crucial experiments of Luria and Delbruck upon which the conventional wisdom was based had employed a lethal selection, and therefore would have precluded the recovery of mutations that might occur after the imposition of selection (i.e., post-plating mutations). Cairns et al. (40) instead used a nonlethal selection to study reversions (i.e., *Lac− → Lac+*) and observed that *Lac+* revertants of an *Escherichia coli* strain carrying an *lacZ*(Amber) mutation on an F' plasmid, continue to arise on lactose minimal agar plates for several days, in addition to the pre-existing *Lac+* mutations which form colonies within the first two days. The authors also presented the results of fluctuation tests showing deviations from the Luria-Delbruck distribution (expected from pre-existing mutations) towards the Poisson distribution (expected of post-plating mutations).
The authors also demonstrated that the occurrence of these post-plating Lac+ revertants was (i) dependent on the presence of lactose (i.e., as a selective agent), and (ii) not associated with a general increase in the frequency of mutations at other (unselected) loci, implying that this was an example of 'adaptive' or 'directed' mutagenesis (i.e., only such mutations arise which are immediately beneficial to the organism). During the same time, evidence in support of directed mutagenesis also came from the studies of Shapiro (213) and Hall (101) using other assay systems in *Escherichia coli*. These are described in the later section of this chapter.

The claims for the directed mutation have generated lot of interest and controversy in the field of both mutagenesis and evolutionary biology because they were taken by some workers to reflect neo-Lamarckian views and led to vigorous exchanges in learned journals (79, 80, 143, 219). The results of Cairns et al. had implications in two aspects: one is the occurrence of mutations in starved or stationary-phase populations in which active DNA turnover is not expected to occur (that was observed earlier by Ryan) and the other is the phenomenon of "directedness" of the late-arising mutants. The subject of this thesis is primarily on the occurrence of spontaneous mutations in nondividing cells, and hence the "directed" aspect of the mutations is not discussed in detail. However, the directed mutation controversy was resolved quite early (85) and it was conclusively established that the late-arising mutations do occur even if there is no selection pressure (81).

The late-arising or post-plating mutations were called with various names such as 'directed', 'Cairnsian', 'selection-induced' or 'adaptive' and now at this juncture, many workers prefer to call them either as 'adaptive mutations' (even though they are not adaptive in the strict sense) or as 'stationary-phase mutations' (reviewed in 83).

1.12 Mutational systems employed to study adaptive mutations

Different workers used various mutational systems to study either the phenomenon and/or the mechanism of adaptive mutagenesis (81, 83). One particular system that has enjoyed a very high profile and has been intensively studied by many groups during the past 10 years is that employing strain FC40 (39). Though this system is not a typical one, the study of adaptive mutation in FC40 has revealed many new mechanisms by which mutations can possibly arise in stationary-phase cells during selection. This system is described in detail in the following section.
1.12.1 Adaptive mutation in *Escherichia coli* strain FC40

Since its description in 1991 (39), FC40 has become the most popular *E.coli* strain for the study of adaptive mutation. Strain FC40 is deleted for the chromosomal *lac* operon and carries a Lac⁻ allele on an F' episome. This strain cannot utilize lactose, but yields a large number of Lac⁺ revertants on appropriate selection plates. The episomal *lac* allele is an in-frame fusion of the *lacI* (repressor) gene to the *lacZ* gene (encoding β-galactosidase) such that the last four codons of *lacI* and the first 23 codons of *lacZ* along with all of the intervening *lac* promoter-operator region, have been deleted (178). Hence, transcription of the *lacI-lacZ* fusion gene is initiated constitutively from *lacI* promoter (that makes 10-fold more LacI repressor protein). This allele, termed *lacI33::lacZ*, also has an ICR191-induced +1 frameshift at the 320th codon of *lacI*, changing CCC to CCCC so that translation of *lacZ* is rendered out of frame relative to the proximal part of *lacI* (41). Any reversions (by addition or deletion mutations) that could restore the reading frame give rise to a functional β-galactosidase fusion protein and such mutations can be selected as Lac⁺ reversion events.

The advantage with the FC40 system is the abundance of adaptive Lac⁺ mutations that arise, which makes the phenomenon easy to study and permits artifacts to be easily identified. Two days after plating FC40 cells on minimal lactose plates, the pre-existing Lac⁺ revertants (that had occurred during growth of the population) would appear and their numbers would exhibit the classical fluctuation described by Luria-Delbruck; then steadily over a period of 10 days, the adaptive or post-plating revertants (~1 per 10⁷ cells per day) would arise. The characteristics of adaptive mutation in FC40 have been studied by various groups and are summarized below.

1. It has been shown that the late-arising Lac⁺ mutants are not artifacts of slow growth, phenotypic lag or turnover of the stressed or stationary-phase populations (39, 84).

2. The mutational spectrum of the pre-existing and the adaptive Lac⁺ mutants was studied (87, 202) and it was shown that the former represent a mixture of deletions, duplications, and frameshifts whereas the latter consist almost exclusively of −1 base-pair frameshifts in runs of iterated bases. The spectrum seen in the adaptive mutants is similar to that seen in mismatch repair deficient mutants, indicating that physiological mismatch repair deficiency may occur in stationary-phase populations.

3. Recombination functions, specifically the activities of the RecABCD pathway for double-strand break repair, are required for adaptive but not growth-dependent Lac⁺
reversions. Mutations in recA, or -B eliminate the adaptive reversions completely whereas mutations in recD increase such reversions (117).

4. Two enzyme systems responsible for the branch migration of recombination intermediates (RuvABC or RecG) appear to have opposite roles in adaptive Lac+ mutation, whereas both contribute equally to normal recombination (89, 115).

5. The DNA damage response (SOS) induced functions are required for adaptive Lac+ mutagenesis: Ind- mutations in lexA (e.g., lexA3 allele) decrease the adaptive reversions 3- to 5-fold (164). Mutation in recF (that encodes a DNA-binding recombination protein that is not SOS-induced) also decreases adaptive reversions 3- to 5-fold (164). RecF protein plays a poorly defined role in recombination but is required for SOS induction by some DNA damaging agents. The role of RecF in adaptive reversion is at present not very clear.

6. Most importantly, the high rate of adaptive reversion to Lac+ as also all the features described above, apply only when the lacI33 allele is located on a conjugation-proficient F' episome (92, 194). If the same allele is located on the chromosome, then adaptive reversion to Lac+ occurs at about 1/100th the rate and is no longer RecA-dependent (83).

7. Defects in conjugal functions (i.e., mutations like traD or traI, that abolish the F'-transfer) cause a 10-fold reduction in adaptive mutation (88). Nevertheless, the actual physical episome transfer is not required for adaptive mutation (as long as the episome is conjugation-proficient).

8. The adaptive reversions are increased in a ΔpolB strain, suggesting a role for DNA polymerase II in episomal DNA synthesis (70). In particular, the allele of polB in which the proofreading exonuclease is inactivated (polBex1) shows a ~6-fold increase in adaptive reversions, whereas it has no effect on growth-phase reversions (78).

9. The adaptive reversions are attributed to the errors made by the normal replicative DNA polymerase III (114), as the antimutator allele of dnaE decreases adaptive reversions as well as growth-phase reversions.

10. The adaptive reversions are decreased ~85% in a strain (carrying dinB mutation) that is defective in an error-prone DNA polymerase IV (163). Most likely, DinB is responsible for causing the SOS-dependent component of adaptive reversions because dinB single mutant shows an equivalent reduction in adaptive mutagenesis as the lexA3 mutant (which is SOS-noninducible) or the lexA3 dinB double mutant (163).

11. The methyl-directed mismatch repair pathway, which corrects mismatches in newly synthesized (hemi-methylated) DNA strand, corrects most of the errors that could lead
to adaptive Lac+ mutations. Mutations that render the mismatch repair pathway defective (e.g., mutS or mutL), increase the frequency of both adaptive and growth-phase mutations (85).

12. In biochemical and genetic experiments, Harris et al. (113) have shown that (i) overexpression of either mutS+ or mutH+ does not reduce the adaptive reversion frequency; (ii) on the other hand, the adaptive reversion frequency is reduced ~4-fold, and growth-phase mutation frequency is not altered, when a multicopy mutL+ plasmid is introduced into the FC40 strain, implying that deficiency of MutL in stationary-phase cells causes the occurrence of adaptive mutations; and (iii) MutL protein becomes limiting in stationary-phase cells. However, Foster (82) has argued that the effect of MutL overexpression on mutation frequency is nonspecific and that both growth-dependent and adaptive mutations are reduced under these conditions. Hence, the question whether deficiency of mismatch repair enzymes is responsible for the occurrence of adaptive mutations still remains open.

13. The adaptive Lac+ mutants also show high frequency of unselected mutations at other loci both on the F' and chromosome, compared to the Lac- population that is subjected to similar starvation conditions (200, 231). These observations have led to the suggestion for the existence of a 'transient hypermutator' state in a subpopulation of cells, a model proposed originally by Hall (102), and is further described below.

14. Andersson and Roth (6) have suggested that adaptive mutations in FC40 result from growth-dependent mutations in the spontaneously amplified regions of the lac allele. Their hypothesis is that the lacI33 allele is 'leaky' and that in cells in which there has been spontaneous tandem amplification of this locus, there is sufficient β-galactosidase produced to permit slow growth to occur, and that this sets the stage for growth-dependent mutation to Lac+ to occur in one of the amplified copies. The authors have transferred the F' plasmid of FC40 into Salmonella enterica and showed the heterogeneity of the adaptive Lac+ mutants with regard to the lac locus. Consistent with this hypothesis, Galitski and Roth reported the absence of adaptive reversions of a non-leaky lac allele (93). Other workers have also reported the 'instability' phenotype in some of the adaptive revertants in strain FC40 (190).

15. Contrary to the observation of Andersson and Roth (6), Hastings et al. have shown recently that adaptive amplification of the mutant lac sequences itself (without the need for growth-dependent mutations in them) can contribute to the occurrence of a component of adaptive reversions in FC40 (118).
1.12.2 Other mutational assay systems employed for studies on adaptive mutation

A variety of other mutational assays have been employed by different groups for studies on adaptive or stationary-phase mutation, but as mentioned above, none has been as intensive as those on FC40 and clues on mechanisms are also less abundant. Some of these systems are described below.

(A) Reversion of episomal lac alleles

1. The original strain used by Cairns et al. (40) for studying adaptive reversions carried an F' plasmid-borne amber mutation at codon 17 of lacZ. This strain was shown to give rise to the post-plating Lac⁺ revertants on minimal glucose plates and formed the basis for the discovery of adaptive mutations.

2. Cupples and Miller (54) have constructed a set of strains with F'-borne mutations in lacZ that alter an essential glutamic acid residue. Each allele reverts only by one specific bp alteration, and together this set can identify all six base changes and five different frameshifts. Two studies (106, 156) have determined the reversion rates of these strains to Lac⁺ during lactose selection, and both groups concluded that adaptive reversions of G:C→T:A and G:C→A:T are particularly prominent. Hall (106) also tested the lac frameshift alleles and found that during lactose selection -G, +G, and -A mutations occurred, but that the +A did not occur.

3. Mackay et al. (156) studied the effect of ada and ogt mutations (that code, respectively, for Ada and Ogt DNA alkyl transferases involved in the detoxification of alkylated bases) in the strain set of Cupples and Miller. They have shown that ada and ogt double mutations significantly increased the rate of adaptive reversions, whereas the increase of growth-phase reversions was only marginal.

(B) Reversion of chromosomal lac alleles

1. The lacI33 allele, when located on the chromosome, shows 100-fold less adaptive mutation rate, and such reversions are not affected by the recombination functions RecABCD, RecG or RuvABC. Under these conditions the reversion rate is also not increased by supplying conjugal functions in trans (83).

2. Galitski and Roth (93) constructed 30 different strains of S. enterica carrying various mutant lacZ alleles on a defective Mu element inserted in the chromosomal hisC gene. About ten strains produced Lac⁺ revertants during lactose selection; these included some but not all strains with nonsense mutations and single-base frameshifts; none of the strains carrying insertion mutations yielded adaptive Lac⁺ mutations. The authors showed that
all the strains that revert during lactose selection carry 'leaky' alleles (i.e., strains that produce basal levels of β-galactosidase) whereas those that do not revert are tight non-leaky alleles. Based on these results, they suggested that the leakiness accompanied by gene amplification is important for adaptive mutagenesis (93).

(C) **Mu-mediated Ara-Lac^+ fusions**

This system was the first in which adaptive mutation was described, by Shapiro in 1984 (213). His strain carried the regulatory region of the arabinose operon and the carboxy-terminal region of the lacZ gene separated by a defective Mu bacteriophage. The strain is both Lac^- and Ara-. Cells with deletions that fuse lacZ in frame with araB can utilize lactose, if arabinose is also present in the medium as an inducer [these were called as Lac(Ara)^+, and the prefusion strain as Lac(Ara^-)]. Shapiro found that after plating Lac(Ara^-) cells on lactose-arabinose minimal medium, the first Lac(Ara)^+ mutants appear after 5 days and then the number of colonies increase steadily for about 20 days. He suggested that these revertants could be adaptive in nature and would appear only in the presence of lactose. Subsequently, several studies (86, 139, 159, 172) demonstrated, however, that Ara-Lac^+ fusion formation could even be induced by aerobic starvation in the absence of lactose, and that they are not therefore adaptive in a true sense.

(D) **Activation of cryptic operons**

(i) **Evolved β-galactosidase (ebg) operon**

The ebg operon of *E.coli* is cryptic, and has the potential to encode a second β-galactosidase whose existence was detected following lactose selection in ΔlacZ strains. The operon includes an ebgR gene that specifies a repressor that controls expression of the downstream structural genes ebgACB (111). The ebgA and -C genes encode α- and -β subunits of Ebg enzyme, respectively, while the function of ebgB gene product is not known. The natural substrate of Ebg enzyme is not known, but it has a very weak β-galactosidase activity towards lactose. If a strain is deleted for lacZ but can make lactose permease (product of lacY gene), then mutations in the ebg operon can allow for growth on lactose (Lac) or related sugars such as lactulose (Lu). Various mutations in ebgA change the substrate specificity of the Ebg enzyme. For example, Class I mutations (Lac^+ Lu^-) allow good growth on lactose but only poor growth on lactulose, whereas Class II mutations (Lac^+ Lu^+) allow growth on both lactose and lactulose. However, for these substrates to support growth, the ebgR gene must also be inactivated.
Hall has shown that when a strain that is $ebgA^{\text{II}}$ but is wild-type for $ebgR$, was incubated on lactulose, $ebgR$ mutants allowing growth on lactulose appear continuously over a period of 10 days (107). He compared the mutational spectra between the early and late arising $ebgR$ mutants, which revealed that 6% of the early mutations and 39% of the late mutations were due to insertions of IS30 in $ebgR$. If the IS30 insertions were eliminated from the analysis, then the early and late mutational spectra did not differ (109, 110). He suggested that the transposition of IS30 appears to be responsive to some conditions like prolonged starvation or slow metabolism, a hypothesis supported by the independent work of Arber and coworkers (180) on aging stab cultures of E. coli.

A screen for genes that decrease adaptive mutation of $ebgR$ yielded mutations in $phoP$ and $phoQ$, a two-component regulatory system that responds to a variety of environmental stress signals (109). It is possible that PhoPQ regulates IS30 transposition, directly or indirectly.

Hall (107) also investigated the specificity of $ebgA$ mutations in the selective conditions. When $ebgR$ $ebgA^+$ cells were incubated on lactulose, Lu$^+$ colonies due to Class II mutations appeared slowly, but new Lac$^+$ Lu$^+$ (i.e., Class I mutations) did not appear among the population, even though they would have appeared if they had been incubated on lactose.

(ii) β-glucoside ($bgl$) operon

The $bgl$ operon encodes genes for utilization of β-glucosides such as salicin or arbutin. The operon is cryptic in E. coli K-12 strains because of an interaction of a DNA-binding protein (H-NS) with sequences upstream of the promoter ($bglR$ site).

Starting with a strain that was $bglR$ (silent) and that also had an insertion of IS103 in $bglF$ (structural gene for salicin utilization), Hall showed that after 10 days of incubation on salicin MacConkey plates, Sal$^+$ papillae appear on Sal$^-$ colonies and continue to appear until 60% of the colonies had papillae. To become Sal$^+$, two rare events had to occur: (i) precise excision of the IS103 element from $bglF$, and (ii) mutation of $bglR^-$ to $bglR^+$; Hall concluded that these events had occurred at an enormously higher frequency compared to that normally expected for double mutations (101). On the other hand, Mittler and Lenski have shown (173) that the single excision event alone (of IS103) provides an advantage to the cells within that colony to grow slowly on salicin, reaching a population large enough to give rise subsequently to $bglR^+$ mutants. Hence, they argue that these double mutations did not occur simultaneously, but sequentially.
Mutations in several genes have been shown to activate the wild-type bgl operon, but the most common spontaneous mutation is insertion of IS1 or IS5 elements into bglR (105). Hall showed that adaptive Bgl+ mutants appear continuously during incubation on minimal arbutin plates, and he compared the mutational spectra of the early and late-arising Bgl+ mutants. The proportion of Bgl+ mutants due to IS insertions decreased from 98% in the former to 79% in the latter, while the number due to mutations in hns increased from 0% to 21% in the late-arising mutants (108).

(E) Reversion of amino acid auxotrophies

Reversions of auxotrophic mutations to prototrophy were extensively used as assay systems to study the adaptive or stationary-phase mutations by various groups.

(i) As mentioned above, the earliest work was that of Ryan, where he used His- strains to study the reversions to His+, and made significant contributions to the field of spontaneous mutations in nondividing cells (204).

(ii) Hall has used strains with missense mutations in either trpA or trpB. When plated on minimal medium with limiting amount of tryptophan, these strains form small colonies within 3 days and from these colonies, Trp+ papillae arose over a period of 10 days (these papillae are considered to be adaptive in nature). Hall also reported that Trp+ papillae occur at a very high frequency in a population of doubly mutant trpA trpB cells (103). However, this turned out to be an artifact as the two reversion mutations were shown most likely to have occurred sequentially and not simultaneously (104).

(iii) Bridges used reversion of amino acid auxotrophies to investigate the occurrence of stationary-phase mutations. Most of the strains employed by him are E.coli B/r derivatives carrying ochre mutations in trpE or tyrA, or a missense mutation in trpA (22, 32). He has shown that stationary-phase reversions occur during prolonged incubations on the plates unsupplemented with the auxotrophic requirement (22, 32). However, the population dynamics of these strains on the selective plates was complex, and he observed slow growth on these plates (27, 28, 31).

Bridges has observed that, the phenomenon of stationary-phase mutagenesis was not affected by defects in nucleotide excision repair, DNA polymerase I, reverse transcriptase, transcription repair coupling factor (encoded by mfd), SOS functions, or RecA (22, 23, 25, 26). On the other hand, mutations in mutY (encoding a DNA adenine glycosylase) and mutT (encoding an 8-oxoGTPase) increased stationary-phase mutation 10- and 100-fold, respectively (24, 27). Although mutations in mutM (encoding a DNA glycosylase that
removes 8-oxoG) did not increase the stationary-phase mutations, the frequency was greatly stimulated when combined with mutY mutation (32). All the three proteins are known to participate in a pathway that serves to reduce the oxidative damage (specifically the 8-oxoG induced lesions) in cells. Furthermore, overproduction of MutY or MutM reduced the late reversion frequency about twofold (32). So, in this assay system, oxidative damage appears to be involved in producing the lesions/damage responsible for the occurrence of late-arising revertants.

Defects in RpoS (encoding the stationary-phase specific sigma factor) did not affect stationary-phase reversion of trpE or trpA, which is curious because RpoS regulates genes encoding proteins (like Dps and Xth) that protect against oxidative damage in stationary phase (30). However, production of carotenoids which are oxygen scavengers, reduced late reversion of trpE and trpA about twofold (30).

(iv) Jayaraman has studied the post-plating mutations of thr-1 and leuB6 alleles and shown that RecA-dependent, late-arising mutations occur during starvation (126). He has shown that both the alleles are leaky and are capable of giving rise to higher number of post-plating mutations in the presence of either streptomycin (an agent known to increase the translational errors and thus, leakiness) or a mutation (termed adi or ppm), that increases the translational errors. He was thus the first to suggest a correlation between leakiness of genetic markers and adaptive mutagenesis, a point subsequently supported by the findings of Roth and coworkers (93).

(v) Benov and Fridovich (11) demonstrated adaptive mutagenesis using the argE3 reversion system and also showed that its frequency is decreased during anaerobic growth, thus implicating a role for molecular oxygen in this phenomenon. A strain carrying mutations in both sodA and sodB (which encode manganese- and iron-dependent superoxide dismutases, respectively) exhibited a ~4-fold increased adaptive mutagenesis when compared to the wild-type, and this increase was also oxygen-dependent.

(vi) Godoy et al. (98) have demonstrated the occurrence of adaptive reversions of a chromosomal trpE7999 allele in a strain also carrying lacI33::lacZ allele on an F' plasmid. They have shown that a feature of the Trp+ mutability is the accumulation of Trp+ and Lac+ revertants with additional unselected mutations on the chromosome, most of which are not due to heritable mutators. These results have hence supported the suggestions for the presence of transient hypermutator cells in the stationary-phase populations.
1.12.3 Forward mutational assay systems

(i) Taddei et al. (226) measured the mutagenesis frequency in resting bacterial populations using Rif$^+$→Rif$^-$ mutational assay system. They have demonstrated that bacteria in aging colonies (on agar plates) exhibit SOS induction (which is cAMP-dependent) and increased mutagenesis.

(ii) Jayaraman (125) also demonstrated the occurrence of post-plating mutations using an assay of valine-sensitivity to valine-resistance. He has shown that mutations in genes of the methyl mismatch repair pathway, mutS and mutL, increase the post-plating mutation frequency to a great extent, but the increase seen in mutS mutant is RecA- and LexA-dependent whereas the increase in mutL strain is RecA-independent. This result is interesting, because both MutS and MutL are known to function in the same pathway and the reason for this differential activity is not evident.

(iii) Recently, Bhagwat and coworkers (176) have shown that mutations in mug (encoding a DNA glycosylase whose function is not yet clearly defined), increase the spontaneous Rif$^-$ frequency ~3-fold in cultures that are incubated for more than 72 h, although no effect was observed in 24 h incubated cultures. Based on these results, the authors suggest that Mug is an antimutator specific to stationary-phase populations.

1.13 Examples of adaptive mutagenesis in other microorganisms

(A) Mutations in the histidine operon of Salmonella enterica

Two of the hisG alleles (hisG46, a missense mutation and hisG428, an ochre mutation) have been used extensively for adaptive mutation studies in S. enterica. The adaptiveness of the reversions, and its RecA-independence was shown by Gizatullin and Babynin (95).

In an earlier study, Prival and Cebula (191) investigated the genetic alterations that occur during the reversion of the above alleles, in cells that were grown either in the presence or absence of histidine (prolonged starvation). Analysis of the histidine-independent revertants of hisG428 that arose during the growth in the presence of histidine showed that both intragenic and extragenic suppressor mutations occurred whereas the frequency of such extragenic suppressors was greatly reduced among the late-arising His$^+$ revertants. Moreover, DNA sequence analysis revealed striking differences in the distribution of particular transversions at the hisG428 locus in late-arising revertants as compared to those that had arisen in the presence of histidine, thus suggesting that there may be different mutagenesis mechanisms operating in starvation conditions.
(B) Examples of adaptive mutagenesis in \textit{Pseudomonas putida}

To date, three studies of \textit{P. putida} have documented the occurrence of stress- and substrate-associated genetic changes (128, 141, 216). The genetic events involve cryptification or decryptification of catabolic pathways, and appear to be influenced by factors associated with both starvation and the specific substrates of the pathways.

(C) Stationary-phase mutations in \textit{Lactobacillus plantarum}

When \textit{Lactobacillus plantarum} ATCC 8014 was maintained in liquid culture for very long periods of time, Thomson et al. observed (230) that mutations to antibiotic resistance (such as resistance to rifampicin, streptomycin, and sodium fusidate) occurred with very high frequency, even in the absence of measurable cell division.

1.14 A model to explain the mechanism of adaptive mutation in FC40

Even though the FC40 system is not a typical one, it has been extensively studied and has led to the elucidation of new mutational pathways. The current model for the occurrence of adaptive reversions in FC40 is explained below and the salient points are also illustrated in Fig. 1.2 [reproduced from (201)].

In this model, nicking at the conjugal origin of DNA replication on the episome (i.e., \textit{oriT}) initiates recombination. Kuzminov (137) proposed that a double-strand end is created when a replication fork initiated at the vegetative origin (i.e., \textit{oriV}) of the episome collapses at the nick at \textit{oriT}. The exonuclease and helicase activities of RecBCD create an invasive 3' end that initiates recombination by RecA, accompanied by the fork reassembly. This recombinational replication may be error-prone (these intermediates may also generate the substrates for SOS-induction and this may explain the role of DinB) because of the activity of DNA polymerases (163) or by the transient limitation of methyl mismatch repair proteins, particularly MutL (113). A suggestion that the occurrence of the adaptive mutations might be caused by the slow mismatch repair activity due to the limitation of these proteins in stationary-phase cells was proposed by Stahl (219) as early as 1988 and later by Boe (20).

The new replication fork differs from a normal fork because it resembles a four-stranded recombination intermediate (Holliday junction). Translocation of the Holliday junction towards the fork by RuvABC proteins may stabilize the 3' ends for priming the replication, whereas unwinding of the fork in the opposite direction by RecG protein may destroy the intermediates of DNA replication. This may explain the differential activity of these two junction resolvases.
Fig. 1.2  Model for the occurrence of adaptive mutations in FC40 system, (reproduced from Ref. 201)
It has been shown by Hastings et al. (118) that the growth of adaptive Lac\(^+\) revertants can occur not only following point mutations but also by amplification of Lac\(^-\) sequences. These authors propose that adaptive mutations can possibly occur by at least two parallel pathways: (i) point mutations occurring due to polymerase errors along with the transient mismatch repair inhibition, and (ii) adaptive amplification of the \(lac\) sequences.

1.15 Alternative models to explain the occurrence of adaptive mutations

(A) Gene amplification

Andersson and Roth (6) proposed that RecA-dependent amplification of the Lac\(^-\) allele occurs during lactose selection and that any growth-dependent Lac\(^+\) mutations which occur during the DNA amplification would be selectively retained. A role for RecBCD can be explained by assuming that the amplified arrays are produced by rolling-circle replication initiated by the double-strand end repair.

(B) Transient mutation

To explain the specificity of adaptive mutation, Hall (102) proposed that during selection most cells in the population do not mutate, but that a subpopulation experiences a transiently high mutation rate, with the cells in this subpopulation dying unless a useful mutation occurs. A specific prediction of this 'hypermutable state model' is that nonselected mutations should occur at a higher frequency in the cells that bear adaptive mutations than in cells that do not.

In studies on the FC40 system, it has been shown that adaptive Lac\(^+\) revertants are associated with unselected mutations at other loci on both F' and chromosome, when compared to (i) the stressed Lac\(^-\) population, and (ii) unstressed Lac\(^-\) populations (200, 231). However, in their experiments, the comparison was not made with the growth-phase dependent Lac\(^+\) revertants, thus it is still premature to say that adaptive Lac\(^+\) mutants arise as a consequence of transient hypermutable state in a subpopulation of stressed cells.

(C) Transcription: as a cause of adaptive mutation

Davis has proposed that transcription can be mutagenic and can cause spontaneous mutations in nondividing cells (57).

(D) Methylation-mediated adaptive mutations

Mahajan and coworkers have proposed a mechanism wherein adaptive mutations can probably arise only in the neighbourhood of a methylatable cytosine (16).
1.16 Current understanding of the chromosomal adaptive mutations

All the models that explain the adaptive reversions of strain FC40 had to consider an obligatory role for RecA protein in the process, since mutations in recA totally abolish the adaptive reversions. However, unlike the F'-system, the chromosomal adaptive reversions in several assays were found to be RecA-independent (22, 106), suggesting that the adaptive mutational pathway(s) in one system may be totally different from that in the other.

It is known that spontaneous damage to DNA (such as alkylation, oxidation or deamination) occurs during the normal life of the cell, by the actions of a variety of molecules that are generated by the metabolic processes. Therefore, it is reasonable to consider the possibility that such damage could be responsible for the occurrence of stationary-phase mutations (reviewed in 29).

(A) Endogenous alkylating agents

There is good evidence to implicate endogenous alkylation damage in the formation of stationary-phase mutations. By measuring His$^-$→His$^+$ reversion frequency in a mutant strain lacking both the Ada and Ogt methyl transferases that detoxify the alkylated bases, Rebeck and Samson (196) were the first to show that endogenous agent(s) exists in nondividing cells that generate alkylation damage and mutagenesis. This suggestion was later confirmed by other studies (85). A further clue to the mechanism and the origin of the endogenous alkylating agent has come from the observation of Taverna and Sedgwick (229) that stationary-phase mutation is abolished in moa bacteria defective in nitrosation of amides and related compounds. The moa gene encodes a molybdopterin cofactor required for the function of several reductases.

(B) Oxidation products of guanine

It is now clear that 8-oxoG, a product of oxidation damage, can lead to mutations in starved bacteria; the rate of stationary-phase mutations is considerably elevated in mutant strains (in particular mutT, mutY, and also mutY mutM double) that are defective in processing of 8-oxoG lesions (32). Overproduction of MutY or MutM protein reduces the mutation frequency by ~50%, indicating that 8-oxoG is formed and is involved in causing mutations even in wild-type cells (32).

(C) Active oxygen species (AOS)

The observation that adaptive mutagenesis is decreased in anaerobic conditions (11), implicated the involvement of AOS in this phenomenon. However, the exact nature of the AOS is not very clear, although a deficiency in superoxide dismutase (SOD) substantially
increases the stationary-phase mutation frequency (11). Moreover, when a mimic of SOD is added to the wild-type strain, the mutation rate was decreased by two-fold indicating that there is enough superoxide produced even under these conditions for it to contribute to the occurrence of stationary-phase mutations. Similarly, when the concentration of carotenoids (which are specific scavengers of AOS) was increased in E.coli cells by introduction of a multicopy plasmid bearing carotenoid biosynthetic genes from Erwinia herbicola, the stationary-phase mutation frequency was shown to be decreased in two different assay systems (30, 33).

(D) Other DNA lesions

Even though not very clearly delineated, there are instances described of increased stationary-phase mutagenesis mediated by the SOS-dependent functions. The work of Taddei et al. has shown that aging bacterial colonies exhibit SOS-induction and also an increase in mutation rate to rifampicin resistance (226).

(E) The role of DNA synthesis and leakiness of alleles

In many instances in which 'resting', 'stationary-phase' or 'nondividing' bacteria have been studied, at least in mutation experiments, the adjectives are misnomers because there is slow metabolism or DNA turnover in these populations (20, 27, 28). Examination of stationary-phase reversions of a number of auxotrophic mutations with varying inherent degrees of leakiness almost leads to the general conclusion that 'tight' alleles show little or no reversions, even though the frequency is not directly related to the magnitude of the leakiness (93). DNA synthesis resulting from leakiness certainly appears to promote the occurrence of mutations in stationary-phase.

(F) Transient mutators

Rosche and Foster (200) showed that hypermutators also appear to occur when Lac+ selection is employed with the lacI33 allele present on the chromosome (although the prevalent model predicts that DSB repair, mediated by the RecABCD system, contributes to the origin of hypermutators in F'-carrying cells). However, it is not very clear whether hypermutators also contribute to the generation of adaptive mutations (chromosomal) in RecA-independent assay systems.

Taken together, the evidence suggests that the occurrence of mutations in starved or stationary-phase populations may not be due to a single pathway/mechanism, but to an amalgam of many different pathways.
1.17 Questions of interest

Although there is plenty of evidence to support the models that were proposed for adaptive mutagenesis, there are still many interesting questions about how mutations arise in cells that are not proliferating or are very slowly growing. For example, questions remain to be answered such as (i) the amount of residual DNA synthesis occurring in resting cells and the factors responsible for the initiation of DNA synthesis, (ii) the role of DNA lesions in causing the adaptive mutations, (iii) the role of transcription and translation in mutagenesis, and (iv) the role of hypermutators (transient or heritable) in the generation of mutations.

1.18 Conclusions

The research on adaptive mutation has revealed a plethora of mutagenic mechanisms that appear to have been selected in evolution. The recombinational mechanism used by F' episomes could be an important source of spontaneous mutation in cells that are not undergoing genomic replication.

Conjugative plasmids like F and related plasmids frequently recombine into and out of the chromosome, and during this process can acquire chromosomal genes that would undergo high mutation rates. When these conjugative plasmids are transferred by horizontal transfer, then the diverged alleles could spread to other organisms. Finally, a subpopulation of nutritionally deprived cells entering into a state of transient hypermutation could be a source of multiple variant alleles and provide a mechanism for rapid adaptive evolution under adverse conditions.

Section C: Transposon Tn10 and its associated excision events

1.19 Introduction

Insertion elements (IS elements) and transposons, including the transposable bacteriophage Mu, are so called because they have the ability to transpose and to insert themselves, as autonomous genetic elements, fairly randomly within the genomes of their hosts. A variety of translocatable genetic elements have been identified in almost all classes of organisms. The transposition of each element is mediated by a protein termed as transposase that is encoded by a gene included within the structure of the transposable element and that acts at the transposon ends which typically are a pair of inverted repeats. Both in vivo and in vitro approaches have been used to study the mechanisms of transposition of several transposons. Transposition may be either nonreplicative (conservative) or
replicative, and the latter may also include a pathway for cointegrate formation and resolution of these intermediates (53).

A characteristic feature of transposition is the generation of host sequence duplication at the site of insertion, of a short stretch of directly repeated sequences, immediately flanking the ends of the transposable element. The length of the duplicated host sequence is distinctive for any particular transposon, although some transposons may generate variable length of duplication sequences (224).

1.20 Organization of transposon Tn10

Transposon Tn10 is one of the well characterized examples of transposable elements (131). It encodes tetracycline resistance determinant and was originally isolated in 1962 from a conjugative drug resistance plasmid R100, found in an enterobacterium Shigella flexneri (238). The transposon is 9147 bp in length and has ~1400-bp inverted repeats (IS10 elements) at its ends (Fig. 5.1). The two flanking IS10 elements are very nearly, but not completely, identical to each other. Although both IS10 elements structurally cooperate to mediate transposition of the intervening material, IS10-Right codes for a functional transposase, whereas IS10-Left is defective for trans-acting transposition functions. The intervening 6500 bp of material between the two IS10 elements in Tn10 includes the 2500-bp tetracycline-resistance determinant and three open reading frames corresponding to a glutamate permease, a putative repressor of heavy metal resistance operons, and a hypothetical protein of Bacillus subtilis (46).

Tn10 has the potential to undergo transposition as a genetically discrete and functionally intact unit, with a transposition frequency of about $10^{-7}$ per element per cell per generation. The transposase gene of Tn10 (that is, IS10-Right) is weakly expressed due to inefficient translation of the transposase mRNA, hence accounting for the low transposition frequency (131). The transposition of Tn10 appears to operate through a nonreplicative or simple cut- and -paste mechanism (131).

Transposition of Tn10 to a new target site is generally associated with duplication of a 9-bp target sequence at the position of insertion, so that the inserted material is always sandwiched between these 9-bp direct repeats. If the Tn10 insertion occurs within a structural gene, it normally abolishes the function of the gene; additionally, if the gene is promoter-proximal within an operon, it usually causes polarity, that is, it reduces the expression of the promoter-distal genes in that operon because of termination of transcription within the Tn10
element. However, in some instances, the TnJ0 insertions also turn on the expression of adjacent downstream genes (131).

1.21 Rearrangements mediated by transposon TnJ0

Genetic and physical analyses have revealed that, in addition to transposition, TnJ0 mediates a variety of DNA recombination events, particularly deletions and inversions. Most of these rearrangements are independent of general homologous recombination functions, but require the transposase enzyme functions. They include simple insertions and replicon fusions (intermolecular events), or adjacent deletions and inversions (intramolecular events). In addition, TnJ0 also mediates RecA-dependent rearrangements as it provides considerable regions of portable homology when present in two or more copies on one or more replicons.

1.22 Excision events associated with TnJ0

Apart from transpositions and rearrangements, the TnJ0 element is associated with three different excision events involving specific transposon sequences (90). As the emphasis of this thesis work is mostly on the excisions mediated by TnJ0-derived transposable elements, these events are discussed in detail below.

1. Precise Excision (PE): As described above, insertion of TnJ0 into a new target site results in duplication of a 9-bp target DNA sequence and integration of the transposon between the resulting 9-bp direct repeats. PE is a deletion event that removes the transposon and one copy of the direct repeat, thus exactly reconstructing the target sequence to its original wild-type sequence. PE is detected genetically as true reversion of a TnJ0 insertion mutation. The frequency of PE varies from one insertion site to another but is usually less than $10^{-8}$ per element per generation.

2. Nearly Precise Excision (NPE): NPE is a deletion event similar in nature to PE but involving short direct repeat sequences internal to TnJ0, one near each end. Deletion of material between these repeats results in excision of all but 50 bp of TnJ0, so that the target gene remains nonfunctional (the residual sequence, not being an integral multiple of 3, results in translation going out of frame), but there is relief of polarity on the expression of downstream genes in the operon. The frequency of NPE is approximately 100-fold higher than that of PE, that is, about $10^{-6}$ per element per generation.

3. Precise Excision of Nearly Precise Excision (PENPE): The 50-bp sequence remaining after NPE contains sequences from each end of TnJ0 plus the flanking 9-bp direct repeats. This material can be further excised to give complete restoration of the wild-
Another locus, called *uup*, was also identified as a *tex* locus on the *E.coli* linkage map at 21 min, mutations in which increased the frequency of PE of Tn10 and another transposon Tn5 (124). However, the *uup* locus had not been molecularly characterized, and neither its function nor its mechanism of action was known.

Conditions that increase or stabilize the amount of single-stranded DNA also appear to enhance the PE frequency (56). The PE frequency is markedly increased when the transposon is present on an F' plasmid rather than on the chromosome. This stimulation appears to be the result of conjugal transfer of the F' (which involves single-stranded intermediates) as the mutants defective in the conjugal functions show decreased PE frequency (225). In a related context, cells deficient in integration host factor, IHF (encoded by *himA*) were reported to exhibit a 1000-fold decreased frequency of PE (168). However, in their study, the target mutation was located on an F' episome, and it is now known that mutations in *himA* abolish the conjugal transfer. Therefore, it appears that the decreased PE frequency in the *himA* mutant might be an indirect effect mediated through defective conjugal transfer of the fertility factor F.

The frequency of the PE of Tn10 was also found to be significantly increased by treatment of cells with UV or mitomycin C (145). This increase was shown to be RecA- and LexA-dependent, implicating the involvement of SOS-induced functions (47).

### 1.22.2 The use of mini-Tn10 derivatives

Smaller derivatives of Tn10 that lack the transposase gene and are thus defective in the autonomous transposition were constructed by Way et al. (240). These mini-Tn10 derivatives carry perfect invert repeats of the outer 70 bp of IS10-right element on either side of the tetracycline-resistance determinant (Fig. 5.1), whereas the same regions in native Tn10 represent imperfect inverted repeats. These elements are also capable of exhibiting all the three kinds of excision events that have been described in the above section.

### 1.22.3 Mechanism of precise excision

The mechanism by which precise excisions occur is not very clear. Foster et al. (90) had provided early evidence that whereas PE and NPE may occur by related pathways, PENPE appears to occur by a different mechanism. One model has been that PE occurs by a RecA-independent replication slippage event across the pair of direct repeats flanking the insertion and that the inverted repeats at the ends of the element facilitate the process. The inverted repeats may participate in the formation of intrastrand stem-loop structures, although alternative structures involving duplex DNA are also possible.
The former possibility is supported by the findings that the frequencies of PE and NPE are increased under conditions where the single-stranded template (which would more readily be able to form the stem-loop structure) is expected to be abundant, such as in the presence of an M13 ori sequence on the template (56) or during Tra-dependent synthesis of single-stranded DNA during conjugal transfer of an F plasmid (225). An in vitro model that mimics PE and that is mediated by replication slippage has also been reported (42).

### 1.23 Other kinds of DNA rearrangements

In addition to the rearrangements mediated by transposable genetic elements, a variety of other RecA-independent DNA rearrangements occur in *Escherichia coli*. Recombination between tandemly repeated sequences (i.e., either an expansion or deletion of unit lengths of a tandem repeat sequence) is one such source of DNA alteration (37, 177). Such rearrangements are of particular interest because some of them are found in human genetic diseases. In *E. coli* they occur independent of RecA protein, even though the repeated sequences provide large tracts of homology that are good substrates for RecA.

One particular system was developed by Michel and coworkers (17) to study recombination between tandemly repeated sequences (specifically, tandem repeat deletion or TRD) in the lacZ gene of *E. coli* chromosome. In this construct, the lacZ gene is inactivated by the introduction of a 624-bp internal in-frame tandem repeat such that deletion of one repeat would restore a Lac+ phenotype. Using this assay, the authors have shown that mutations in (i) *dnaE* (encoding the polymerase subunit of DNA polymerase III), (ii) *rep* (encoding an helicase implicated in replication), and (iii) *uvrD* (which codes for an helicase II that is involved in nucleotide excision repair and methyl mismatch repair) increase the frequency of TRD events (17, 18).

In addition, Lovett and others (37, 209) have also developed different assay systems, with constructs that are located either on the plasmid or on the chromosome, to study deletions or expansions of tandem repeats. They have identified various DNA replication and repair mutations that increased the TRD frequency, and have suggested that replication slippage occurring during DNA replication is responsible for recombination events involving tandem repeats.

### Role of replication in RecA-independent rearrangements

Molecular and genetic analyses of these rearrangements have provided several clues to the mechanisms by which they occur. There is evidence for three mechanisms for RecA-independent sequence rearrangements: simple replication slippage, sister-chromosome
exchange-associated slippage, and single-strand annealing. Replication plays a critical role in the first two slipped misalignment mechanisms, and difficulties in replication appear to trigger rearrangements via all these mechanisms (36).

**Section D: Objectives of the present work**

There was an established notion that spontaneous mutations occur randomly during the exponential growth of bacterial populations, that is, during active DNA replication and cell division (154). However, in 1988, a suggestion was made by Cairns et al. (40) that spontaneous mutations can also occur in nondividing cells and that these mutations are adaptive in nature, i.e., only such mutations occur which are immediately beneficial for growth of such cells. This idea had generated considerable debate and interest in the field of mutation research.

The primary aim in initiating this work was to establish the occurrence of spontaneous mutations in stationary-phase cells of *Escherichia coli*, and to understand the mechanisms by which they might be generated. The major objectives of this work were:

(i) To develop and standardize a novel conditional-lethal strategy to study the occurrence of spontaneous mutations in nondividing cells. In this strategy, the revertants for a particular mutation that arise during the exponential growth of a population would be selectively eliminated with the aid of a conditional-lethal mutation so that revertants that arise exclusively in the stationary-phase populations could then be analyzed.

(ii) To identify the mutations (i.e., mutators or antimutators) that alter the mutation frequency in nondividing cells, in order to understand the mechanisms of occurrence of such mutations.

In Chapter 3, the development and use of the conditional-lethal strategy for studying the occurrence of spontaneous mutations in nondividing cells are described. The *tex* mutations that are described in Chapter 5 were utilized in this part of the study to examine the precise excision frequency of Tnl/OdKan in nondividing cells.

To delineate the pathways of spontaneous mutagenesis in stationary phase, an attempt was made to identify the mutations that alter the mutation frequency in dividing and/or nondividing cell populations, and these results are presented in Chapter 4.
In order to study and to understand the occurrence of spontaneous precise excisions of Tn10 in nondividing cells, new mutator mutations (*tex*) that increase the PE frequency of a mini-Tn10 element (Tn10dKan) were isolated and characterized. In Chapter 5, the isolation of *tex* mutations and their characterization is described.

During the above study, a mutation in a genetic locus designated *uup*, was also identified as a *tex* mutation that increased the PE frequency of Tn10. As this locus had not been characterized in earlier work, detailed genetic and molecular studies were performed to understand the function and the mechanism of action of the genes in this locus, and these results are presented in Chapter 6.