A new direction for directed mutation?
Zhongge Zhang, Ming Ren Yen, Milton H. Saier Jr.
Division of Biological Sciences, University of California, San Diego, La Jolla, CA, USA

Abstract

Directed mutation is a controversial process that allows mutations to occur at higher frequencies when they are beneficial. Here we review evidence for transposon-mediated directed mutation. crp deletion mutants (Glp-) of Escherichia coli (E. coli) mutate specifically to glycerol utilization (Glp+) at rates that are enhanced by glycerol or the loss of the glycerol repressor (GlpR), and depressed by glucose or gfpR overexpression. Of the four tandem GfpR-binding sites (O1-O4), O4 specifically controls glpFK expression while O1 controls mutation rate. Mutation is due to insertion of the IS5 transposon into a specific site upstream of the glpFK promoter. Mutational control by GlpR is independent of the selection and assay procedures, and IS5 insertion into other gene activation sites is unaffected by the presence of glycerol or the loss of GfpR. The results establish an example of transposon-mediated directed mutation, identify the protein responsible for its regulation, and define essential features of the mechanism involved. We discuss this phenomenon from an evolutionary standpoint and provide examples of analogous switch mechanisms that may or may not be directed.

Introduction

Sixty-nine years ago, Luria and Delbrück reported that bacterial mutations from virus sensitivity to resistance arose randomly. They generalized their results, concluding that genetic mutations occur in the absence of and independently of selection. It has since become a basic principle of genetics that the likelihood of a particular mutation occurs independently of its phenotypic consequences. The concept of directed mutation, enunciated by John Cairns and his coworkers and defined as genetic change that is specifically induced by the stress conditions that the mutation relieves, challenges this principle. The topic of directed mutation is controversial, and its existence, as defined above, has been altogether questioned.

Studies of the lactose (lac) and leucine (leu) operons of E. coli

The most studied potential example of directed mutation involves the Escherichia coli (E. coli) lactose (lac) operon frameshift reversal system, in which cells carrying a lac +1 frameshift allele on an F episome mutate to Lac+ at a rate that is elevated by lactose during starvation. A subsequent genome-wide analysis showed that such Lac+ point mutations (i.e., -1 deletions) may not be directed, although they are adaptive, since higher frequencies of mutation also occur simultaneously in other chromosomal genes unrelated to lactose metabolism. Two mechanisms were proposed to explain the increased rates of reverision to Lac+, one involving double-strand break repair mediated by error-prone DNA polymerase IV, the other involving amplification of the leaky lac frameshift allele, thereby amplifying the low activity of the mutant allele. The first model was supported in a follow-up study. Both lac allele amplification and Lac+ point mutations were induced by lactose starvation and regulated by the stationary phase regulator, RpoS. These two processes have been shown to be two independent outcomes of genetic change during starvation, and both of them are adaptive processes.

Amino acid biosynthetic auxotrophs of E. coli can be mutated to grow in the absence of the required amino acids. One such auxotroph, which contains a C-to-T point mutation at nucleotide 857 in the leuB gene (leading to a S286L change of the protein), reverts to Leu+ when the Leu- cells are incubated under leucine-limited conditions. Thirty six out of 53 Leu+ revertants or pseudorevertants were found to harbor a single nucleotide substitution that resulted in alteration of the 286th residue from leucine back to serine, or to valine or methionine. These Leu+ reversion mutations proved to occur in response to leucine starvation. The increased rate of Leu+ mutation is believed to have resulted from increased rates of transcription elicited by increased cytoplasmic concentrations of guanosine tetraphosphate (ppGpp), an alarmone whose synthesis is increased in response to leucine starvation. Transcription-promoted Leu+ mutation is probably an example of adaptive mutation, and could be a potential example of directed mutation since most (but not all) of the identified mutations seemed to be directed to the leuB gene whose product is needed for the Leu+ phenotype. However the involvement of ppGpp suggests that this response may also be pleiotropic (affecting other operons under ppGpp control) rather than being operon specific.

Part of the justifiable skepticism concerning directed versus adaptive mutation of the lac system in E. coli resulted from experiments that supposedly demonstrated this phenomenon, but were subsequently shown to be explainable by classical genetics. Mutation rates vary with environmental conditions and genetic background (e.g., the presence of mutator genes). However, this does not render the mutation directed. To convincingly establish the principle of directed mutation, it is necessary to demonstrate the specificity of the phenomenon, identify the proteins involved, and characterize the mechanism responsible.

Transposon hopping as a means of gene activation

Transposons such as Insertion Sequence 5 (IS5) provide the host organism with benefits such as the opportunity for genetic changes that might, for example, relieve a stress condition or allow utilization of a nutrient not otherwise usable. These jumping genes can translocate by at least two distinct mechanisms, one involving replication and the other being replication independent. More than 500 transposons have been identified to date, and they are commonly found in the genomes of both prokaryotes and eukaryotes. They are believed to be primary causes of DNA rearrangements including chromosomal inversions and deletions, all of which could be beneficial under specific stress conditions. The hopping of transposable elements, transposons, can activate or inactivate critical
genes or operons when inserted into appropriate chromosomal loci. Transposon-mediated mutations occurring under stress conditions (e.g., starvation or the presence of a toxin) can be beneficial to the host organisms. For example, the products of the E. coli nfsAB operon convert nitroaromatic compounds to toxic nitro-anion free radicals via nitroreduction. When E. coli is exposed to these compounds in agar plates, resistant colonies appear during inoculation, and all such mutants arise due to insertion of IS1 or IS5. Activation of the normally cryptic β-glucosidase (bgI) catabolic operon and the ade gene, encoding an adenosine deaminase in E. coli, can be accomplished by insertion of either IS1 or IS5 proximal to these promoters. IS5 has also been found to activate the fucose (Fuc)/propandiol (Ppd) fucAO promoter along with the flagellar motility fibril master switch promoter. These studies, while not shown to be directed, provided the background for recent studies on the E. coli glpFK operon in which IS3-mediated activation occurs in a directed fashion by a well-defined mechanism.

**Structure of the E. coli glp regulon**

The E. coli glp regulon consists of five operons, two of which (glpFK and glpD) are required for aerobic growth on glycerol. The glpFK operon encodes the glycerol transport facilitator, GlpF, allowing rapid entry of glycerol into the cell, and glycerol kinase, GlpK, converting glycerol to glycerol-3-phosphate. The glpD gene encodes the aerobic glycerol-3-P dehydrogenase that oxidizes glycerol-3-P to glycolytic intermediate, dihydroyacetone phosphate. Both operons are subject to negative control by the DNA-binding glp regulon repressor, GlpR, which also binds glycerol-3-phosphate, the inducer of the glp regulon. The glpFK operon, but not glpD, is additionally subject to positive regulation by the cyclic AMP receptor protein, Crp, complexed with cAMP. The glpFK regulatory region contains four GlpR binding sites, O1-O4, and two Crp binding sites which overlap O2 and O3 (Figure 1). The strong Crp dependency of glpFK transcription is reflected by the fact that crp mutant cells are unable to utilize glycerol.

**A novel mechanism of directed mutation?**

In the absence of Crp, the glpFK promoter can be activated by IS5 when this genetic element inserts upstream of the promoter. High level expression of the resultant activated operon is nearly constitutive and relies on the DNA phase between the inserted IS5 and the promoter. A short region of 177bp (IB) at the 3’ end of IS5, which contains an IHF binding site and A-tracts that generate a permanent bend, is necessary and sufficient for such activation. Moreover, IS5 or IB, when inserted into appropriate sites, can activate other Crp-controlled promoters such as the lac promoter of E. coli.

This insertional event appears to represent a genuine example of directed mutation. Mutation is mediated by IS5 insertion at a specific site upstream of the glpFK promoter. Mutation of a crp deletion mutant to Glp+ occurs with a ten fold higher frequency when glycerol is present or GlpR is lacking, but overexpression of glpR greatly depresses the mutation rate. Frequencies of IS5 insertion to other sites that activate dissimilar promoters were unaffected by glycerol or the absence of GlpR. GlpR therefore seems to provide two distinct biological functions, one, recognized previously, to control gene expression by binding to the downstream operator, O4, and the other, to control the IS3-dependent mutation rate by binding to the upstream operator, O1. This is the first example of transposon-mediated directed mutation where the molecular explanation, involving a DNA-binding protein, has been provided. Mechanistic details will be presented below.

**Glp+ mutations in a crp genetic background**

E. coli cells that lack Crp can not utilize glycerol as the sole carbon and energy source since expression of the glpFK operon is essential for glycerol uptake and metabolism. However, when crp cells are incubated on solid glycerol minimal medium, Glp+ colonies appear after prolonged incubation. Dozens of crp Glp+ mutant colonies from different plates and arising on different days were purified and tested for growth on glycerol in defined liquid medium. Glp+ colonies appear first after 3 days although act and crp Glp+ cells formed visible colonies in <2 days. New colonies continued to appear at increasing frequencies thereafter. When the same crp cells were plated as before, but variable numbers of the cells from a crp Glp+ strain were included with the cells before plating, colonies appeared from the crp Glp+ cells within two days, and new Glp+ mutants arose at the same frequency as before (Figure 2B). This experiment was repeated many times using independently isolated crp Glp+ mutant strains, and they all behaved similarly. Thus, the Glp+ mutants arising from crp cells on glycerol minimal medium plates were not present in the cell culture before plating, and the potential presence of a growth inhibitor could not account for the results. The ratio of Glp+ colonies (independent mutations that occur after plating) to the total population increased with time. When sorbitol (another sugar crp cells can not utilize) replaced glycerol, the frequency of appearance of Glp+ mutations was much lower than that on glycerol plates but much higher than that on sorbitol plates. The lag phase for the crp Glp+ strain was shorter, and its growth rate was slightly greater than that of wild type (wt) E. coli. Using Biolog plates, these two strains (parental crp Glp+ and crp Glp+ cells) were tested for oxidation of other biological carbon, nitrogen, phosphorous and sulfur sources, but no obvious differences were observed. The mutation that enables crp cells to utilize glycerol evidently did not affect other phenotypes.

The relative frequencies of Glp+ mutation were determined with glycerol, sorbitol or glucose as the sole carbon source on minimal agar plates. On glycerol plates, colonies appeared first after 3 days although act and crp Glp+ cells formed visible colonies in <2 days. New colonies continued to appear at increasing frequencies thereafter. When the same crp cells were plated as before, but variable numbers of the cells from a crp Glp+ strain were included with the cells before plating, colonies appeared from the crp Glp+ cells within two days, and new Glp+ mutants arose at the same frequency as before (Figure 2B). This experiment was repeated many times using independently isolated crp Glp+ mutant strains, and they all behaved similarly. Thus, the Glp+ mutants arising from crp cells on glycerol minimal medium plates were not present in the cell culture before plating, and the potential presence of a growth inhibitor could not account for the results. The ratio of Glp+ colonies (independent mutations that occur after plating) to the total population increased with time. When sorbitol (another sugar crp cells can not utilize) replaced glycerol, the frequency of appearance of Glp+ mutations was much lower than that on glycerol plates but much higher than that on sorbitol plates. The lag phase for the crp Glp+ strain was shorter, and its growth rate was slightly greater than that of wild type (wt) E. coli. Using Biolog plates, these two strains (parental crp Glp+ and crp Glp+ cells) were tested for oxidation of other biological carbon, nitrogen, phosphorous and sulfur sources, but no obvious differences were observed. The mutation that enables crp cells to utilize glycerol evidently did not affect other phenotypes.

The relative frequencies of Glp+ mutation were determined with glycerol, sorbitol or glucose as the sole carbon source on minimal agar plates. On glycerol plates, colonies appeared first after 3 days although act and crp Glp+ cells formed visible colonies in <2 days. New colonies continued to appear at increasing frequencies thereafter. When the same crp cells were plated as before, but variable numbers of the cells from a crp Glp+ strain were included with the cells before plating, colonies appeared from the crp Glp+ cells within two days, and new Glp+ mutants arose at the same frequency as before (Figure 2B). This experiment was repeated many times using independently isolated crp Glp+ mutant strains, and they all behaved similarly. Thus, the Glp+ mutants arising from crp cells on glycerol minimal medium plates were not present in the cell culture before plating, and the potential presence of a growth inhibitor could not account for the results. The ratio of Glp+ colonies (independent mutations that occur after plating) to the total population increased with time. When sorbitol (another sugar crp cells can not utilize) replaced glycerol, the frequency of appearance of Glp+ mutations was much lower than that on glycerol plates but much higher than that on sorbitol plates. The lag phase for the crp Glp+ strain was shorter, and its growth rate was slightly greater than that of wild type (wt) E. coli. Using Biolog plates, these two strains (parental crp Glp+ and crp Glp+ cells) were tested for oxidation of other biological carbon, nitrogen, phosphorous and sulfur sources, but no obvious differences were observed. The mutation that enables crp cells to utilize glycerol evidently did not affect other phenotypes.

The relative frequencies of Glp+ mutation were determined with glycerol, sorbitol or glucose as the sole carbon source on minimal agar plates. On glycerol plates, colonies appeared first after 3 days although act and crp Glp+ cells formed visible colonies in <2 days. New colonies continued to appear at increasing frequencies thereafter. When the same crp cells were plated as before, but variable numbers of the cells from a crp Glp+ strain were included with the cells before plating, colonies appeared from the crp Glp+ cells within two days, and new Glp+ mutants arose at the same frequency as before (Figure 2B). This experiment was repeated many times using independently isolated crp Glp+ mutant strains, and they all behaved similarly. Thus, the Glp+ mutants arising from crp cells on glycerol minimal medium plates were not present in the cell culture before plating, and the potential presence of a growth inhibitor could not account for the results. The ratio of Glp+ colonies (independent mutations that occur after plating) to the total population increased with time. When sorbitol (another sugar crp cells can not utilize) replaced glycerol, the frequency of appearance of Glp+ mutations was much lower than that on glycerol plates but much higher than that on sorbitol plates. The lag phase for the crp Glp+ strain was shorter, and its growth rate was slightly greater than that of wild type (wt) E. coli. Using Biolog plates, these two strains (parental crp Glp+ and crp Glp+ cells) were tested for oxidation of other biological carbon, nitrogen, phosphorous and sulfur sources, but no obvious differences were observed. The mutation that enables crp cells to utilize glycerol evidently did not affect other phenotypes.
substantially higher than that on glucose plates where mutation rates were negligible.

**ISS upstream of PglpFK in the crp Glp+ mutants**

The glpFK regulatory region from 116 out of 116 independently isolated Glp+ colonies contained a ~1.1 kb insert that proved to be ISS, located between the 127th nucleotide and the 126th nucleotide upstream of the glpFK transcriptional start site. It was always orientated with the 3' end proximally upstream of the promoter (Figure 1). The four base ISS recognition sequence, CTAA, was repeated immediately adjacent to ISS as expected.35,49 No other mutation in addition to the ISS insertion could be detected, and other DNA elements inserted at the same site did not activate glpFK-expression. Furthermore, insertion of an 85bp fragment either within or downstream of the CTAA ISS target site essentially abolished (~2% frequency) appearances of Glp+ mutants, and insertion of the same element upstream of this tetrancleotide target sequence reduced the insertion frequency to about 50%. These observations showed that (i) the CTAA insertion site upstream of PglpFK is required for ISS hopping to this region, (ii) an appropriate location of the CTAA element upstream of PglpFK is essential for glpFK operon activation, and (iii) the adjacent sequence upstream of the CTAA target site is important to maximize the ISS insertion rate.

**Dependence of the Glp+ mutation rate on the glycerol repressor, GlpR**

Glycerol is phosphorylated by GlpK to glycerol-3-phosphate which binds to and releases the GlpR repressor from its operators.43,46 When GlpR dissociates from its operators, a conformational change might be transmitted through the DNA, promoting insertion of ISS at the CTAA site upstream of PglpFK.

To test this possibility, the glpR gene was deleted, and the frequencies of appearance of Glp+ mutations in the crp single mutant and the crp glpR double mutant were measured in the absence and presence of glycerol. When the crp and crp glpR strains were plated without glycerol, individual Glp+ colonies from the crp glpR double mutant formed on the plates with a 10-fold higher frequency than from the crp mutant. In the presence of glycerol, similar mutation frequencies were observed for the crp and crp glpR strains, and these rates were similar to those observed for the crp glpR double mutant in the absence of glycerol. These experiments demonstrated that deletion of glpR is equivalent to inclusion of excess glycerol in the growth medium. GlpR therefore mediates the response to glycerol.49

To further demonstrate that GlpR binding to the DNA inhibits the appearance of Glp+ mutations, the glpR gene was cloned into a vector and expressed under the control of an inducible promoter. When cells expressing the glpR gene at high levels were quantitated for the appearance of Glp+ mutants, the frequencies of these mutations decreased below the background rate observed for the crp strain. It thus became clear that GlpR controls transposition of ISS to the site upstream of the glpFK promoter. But what about other sites on the E. coli chromosome? Examination of three other operons known to be activated by ISS insertion, the lacO1, lacO2, flhDC, and bglGFB operons,38,39 revealed that neither the presence of glycerol nor the loss of GlpR influenced the ISS hopping rates to these sites. It was therefore concluded that GlpR is site specific, only influencing ISS-mediated mutation rate by transposition of ISS to the activating site upstream of the glpFK promoter.

**GlpR operators differentially control glpFK expression and Glp+ mutation rate**

DNA footprinting identified four GlpR binding sites, O1–O4, in the upstream glpFK operon regulatory region45 (Figure 1). The far upstream site (O1) and the far downstream site (O4) were mutated, and the effects on glpFK expression and mutation rate were compared.45 For glpFK expression measurements, a lacZ reporter gene fusion construct was used (Figure 3A), and for measurement of the frequency of mutation to Glp+, a standard cell counting procedure was used (Figure 3B). Mutation of O4 increased glpFK operon expression about 5-fold although mutation of O1 was almost without effect (Figure 3A). In contrast, loss of O1 yielded a sevenfold increase in mutation frequency although loss of O4 had only a 2-fold effect (Figure 3B). Thus, while O1 primarily controls mutation rate, O4 primarily controls operon expression. Mutation rate is therefore not a function of glpFK expression level, and GlpR regulates expression and mutation rate independently.
IS5 insertion rates using the chloramphenicol (Cm) resistance gene (cat) for selection

To confirm the effects of glycerol and GlpR on IS5 insertion upstream of the glpFK promoter and to show that control of mutation rate by GlpR is independent of the Glp+ phenotype, the chromosomal glpFK operon was replaced with a chloramphenicol resistance (cat) structural gene so that cat was expressed solely from the glpFK promoter (i.e., PglpFK-cat at the glpFK locus). crp and crp glpR cells were both sensitive to Cm at < 25 μg/mL, but IS5 insertion as reported above rendered these cells resistant to Cm at >50 μg/mL. Using this chromosomal PglpFK-cat construct, IS5 insertion assays were performed by incubating cells on LB agar plates with 50 μg Cm per mL. Cm resistant (Cmr) colonies arose on both crp and crp glpR plates, but the rate of appearance of Cmr colonies relative to the total populations was about 20 times higher for the crp glpR cells than for the crp cells (Figure 4). Furthermore, when glpR was overexpressed, the rate at which Cmr colonies arose decreased dramatically compared to the same cells carrying the empty plasmid. Sequencing analysis showed that 20 out of 20 independently isolated Cmr mutants from either crp cells or crp glpR cells carried IS5 in the usual location and orientation. These results support the conclusions that (i) GlpR represses the appearance of Glp+ mutations in the absence of glycerol; (ii) IS5 insertion upstream of PglpFK is the sole cause of the Glp+ phenotype; and (iii) regulation of mutation rate still occurs when a phenotype unrelated to glycerol metabolism is used to measure mutation rate.

RecA independence of Glp+ mutation

The increased mutation rate in response to the loss of GlpR binding might have resulted from increased gene dosage accompanying homologous recombination-dependent partial chromosomal duplications. Such duplications are RecA-dependent. The dependency of Glp+ mutation on RecA was therefore examined. When cells were incubated on LB agar plates, and individual colonies were examined for Glp+ and total populations, only a 15% apparent decrease in mutation frequency was observed in crp recA double mutant cells compared to crp cells. Similarly, when cells were incubated on minimal glycerol agar plates, introduction of the recA mutation decreased the Glp+ mutation frequency by only 15%. It is therefore clear that the effect of glycerol or the binding of GlpR to its glpFK operators is not dependent on RecA and therefore is not dependent on homologous recombination for partial chromosomal duplication.

The mechanism of IS5-mediated activation of the glpFK promoter

The results summarized above lead to the conclusion that IS5 insertion into a single chromosomal locus in front of the glpFK promoter occurs under starvation conditions with high frequency when glycerol is present and glucose is absent. These are the same conditions that allow these IS5 insertion mutations to be beneficial by relieving the starvation stress. However, the experiments described above do not address the mechanism by which IS5 activates the glpFK promoter.

The promoter activation effect of IS5 proved to be solely due to a short (177bp) region at the 3’ end of IS5. This region is both necessary and sufficient for full promoter activation. It harbors a permanent bend, due to the presence of appropriately spaced A-tracts and an overlapping binding site specific for the genome shaping histone-like protein, Integration Host Factor (IHF). Both of these elements proved to be required for full promoter activation. When each was eliminated by mutation, about 50% of the activation was lost, and their effects were additive. When both were lost, no activation was observed. In support of the conclusion that DNA bending provides the basis for activation by both the A-tracts and IHF binding to its site within the 177bp sequence at the end of IS5, phasing of the DNA proved to be important. Thus in B-DNA a helical turn is about 10bp long, and when a 10bp segment was inserted, there was only a slight loss of glpFK expression, activation, and when a 5bp sequence was inserted at the same location, activation was totally lost.

The mechanism thus appears to involve DNA bending. Presumably the upstream region of the DNA interacts with the transcriptional initiation complex to activate the glpFK promoter that is normally activated by the cyclic AMP-Crp complex in wild type cells. Additionally, it was shown that the E. coli lactose (lac) operon could be activated by the 177bp IB fragment in a crp genetic background. These observations suggest that this newly demonstrated activation mechanism could be applicable to many catabolite controlled operons in a variety of bacteria. It would
not be unexpected if other operons, not under Crp control, will prove to be subject to this ISS-mediated mechanism as well. Moreover, other transposons may have similar capacities to use an analogous mechanism or a variant of it for promoter activation. Further experimentation will be required to establish or refute these possibilities.

Precise excision of ISS

Some transposable elements have been shown to excise both precisely and imprecisely, with precise excision usually occurring at much lower rates than imprecise excision.62,63 In one study, ISS was found to be excised imprecisely, causing deletion of adjacent regions.64 Until recently, no information had been available regarding precise excision of ISS.

The fus regulon for L-fucose (Fuc) uptake and metabolism consists of two divergent operons, fucPIK and fucAO (Figure 5). fucPIK encodes a permease (P) for fucose uptake, an isomerase (I) that converts fucose to fuculose, and a kinase (K) that phosphorylates fuculose to fuculose-1-P; fucA encodes an aldolase that cleaves fuculose-1-P to lactaldehyde and dihydroxyacetone-P. Under aerobic conditions, lactaldehyde is converted to L-1,2-propanediol (Ppd) by FucO, and Ppd is utilized in the TCA cycle. Under anaerobic conditions, lactaldehyde is converted to 1,2-propanediol by FucO, and Ppd is available as a source of carbon for growth of E. coli from these sites when IS5 hops into the critical site within the operon, which is more vulnerable to mutation.19,20 The results serve to dissociate the two functional components of GlpR, which binds to its four operators (O1–O4) in front of the glpFK operon, is displaced from these sites when α-glycerol phosphate, derived from glycerol by phosphorylation, is bound (Figure 1). GlpR clearly controls both glpFK expression and the Glp+ mutation frequency. O4, which overlaps the −10 region, primarily influences gene expression, O2 and O3, which overlap the two Crp binding sites and the −35 promoter region, presumably antagonize activation by the cyclic AMP-Crp complex in wild type cells, and O1 primarily influences ISS hopping into the specific CTAa site, 126.5 base pairs upstream of the glpFK transcriptional start site, 37 bps upstream of O1 (Figure 1). The results serve to dissociate the two functions of GlpR. The mutational inhibitory mechanism, dependent on the binding of GlpR to O1, may be direct, involving interaction of DNA-bound GlpR with the transpososome, or indirect, involving changes in DNA conforma-

tion (e.g., supercoiling, secondary structure, etc).

Adaptive Lac+ mutations that arose during lactose selection appeared not to be directed specifically to the gene in which mutation relieved the stress.5,10 However, in the case of Glp+ mutations, it was shown that (i) each of the five glp operons (glpFK) is activated; (ii) the presence of glucol or the loss of GlpR did not affect ISS hopping into three other gene activating sites in the genome; and (iii) there are no observable differences between Glp+ mutants and the parental crp cells in the utilization of other carbon, nitrogen, sulfur, and phosphate sources.63 These facts strongly suggest that mutation to Glp+ is directed specifically to the glpFK operon, while other ISS-mediated gene activating events do not occur at altered rates when glucol is present or the glpR gene is deleted. These results are therefore fully compatible with the most rigorous definition of true directed mutation.

Several mechanisms have been proposed to account for increased rates of mutation in E. coli cells under stress conditions.65 Two of these (a point mutagenesis mechanism and an adaptive gene amplification mechanism) have been proposed to promote lac frameshift reversion.65 The point mutagenesis mechanism includes DNA double-strand breaks and their subsequent repair, executed by error-prone DNA polymerase IV that is induced under stress conditions such as starvation. In growing cells, double-strand break repair is mediated by a high-fidelity polymerase, such as PolII.11,12 The gene amplification mechanism involves increasing the copy number of the lac +1 allele to up to 50 tandem repeats per cell.27,12,17,66 Each copy of the allele accounts for 1 to 2% of the wild type level of β-galactosidase. Therefore, a sufficient level of expression of the leaky allele causes the Lac+ phenotype.67,12,16

Another mechanism for increasing mutation rate involves the guanosine tetraphosphate (ppGpp)-mediated elevation of transcription of amino acid metabolic genes in response to nutritional stress, thereby enhancing the amount of local single-stranded DNA, which is more vulnerable to mutation.19,20 The most studied example of transcription-elicited mutation is the Leu+ reversion of an E. coli leuB auxotroph during leucine starvation.5,19,18 A fraction (~25%) of the Leu+ mutants proved not to be true revertants, and they contained one or more mutations other than a nucleotide substitution in the leuB gene.18 Because this mechanism is dependent on ppGpp, it is not likely to be specific for leuB.

These reported mechanisms appear entirely different from the Glp+ mutation mechanism reviewed here. In the case of Glp+ mutations,
(i) no evidence suggests an involvement of DNA double-strand breaks or Pol IV-mediated error-prone repair; (ii) homologous recombination-mediated gene amplification proved not to be responsible for the elevated mutation rate induced by the presence of glycerol or the loss of GlpR; (iii) the Glp+ mutation rate is not a function of the gfpFK expression level, and (iv) binding of a local protein (GlpR) to one of its four operators, O1, depresses the Glp+ mutation rates in the absence of glycerol.69

In several cases, transposon insertion frequencies have been shown to increase in the presence of specific carbon sources.67,68 A series of host proteins have been shown to increase or decrease transposition of transposons in E. coli.33-35,69 Many of these proteins are either nucleoid structuring proteins such as H-NS, HU, IHF and Fis50 or proteins related to DNA recombination and repair such as RecG, Dcl, and DinD.70,71. Often, increased transposition occurs in response to nutritional stress.71 Host DNA structuring proteins have been reported to be involved in assembly of the transpososome and in target selection.72-74 The exact roles of these DNA recombination and repair proteins in transposition have not been elucidated.71

In some cases, the DNA adenine methylase (Dam) affects transposition, probably due to its effect on transposase transcription and activity.75,76 Similar observations have been made in yeast, in which multiple host factors are involved in transposition of retrotransposons such as Ty1 and Ty3.77,78 The E. coli RpoS stationary phase sigma factor has also been shown to influence transposition rates,79,80 and unpublished observations. It is not surprising that these host proteins affect the overall transposition rates for transposons in E. coli.33-35,69,80

The mechanism recently identified for PtpF transposition provides relief from starvation and therefore could have been selected for through evolutionary time. Our results suggest that the DNA loop directly contacts RNA polymerase as the means for transcriptional activation (Figure 6). It appears to be a genuine example of directed mutation, since mutation is directed to a specific operon and occurs at a greater rate under conditions that allow benefit to the organism (i.e., in the presence of glycerol and the absence of glucose). The fact that mutation rate is influenced by the presence of glycerol in a process mediated by the glycerol repressor provides a mechanistic explanation for IS5-mediated directed mutational control. It leads us to propose that GlpR has two functions, one in controlling gfp regulon gene expression as recognized previously, and the other in controlling the conformational state of the upstream DNA so as to influence the rate of IS5 insertion. These mechanisms, illustrated in Figure 7, may provide a partial explanation for the presence of four GlpR binding sites in the control region of the gfpFK operon. To what extent these novel transposon-mediated adaptive evolutionary processes will prove to be applicable to other situations and organisms, poses intriguing questions for further study.

Odegard and Schatz81 have reviewed evidence for a 10+ fold increase in mutation rate in the human immune system, relative to spontaneous rates of somatic cell mutation, a phenomenon termed Somatic Hypermutation (SHM). SHM occurs in the variable regions of immunoglobulin (IgG) genes with a rate of up to 10-3 mutations/bp/cell division. In one case, point mutations arise in a specific tetrancleotide hot spot where primary sequence plays an essential role.81 While our studies have focused on mutations occurring by transposon (IS5) insertion in E. coli, the IgG system in animals involves the introduction of point mutations. Reversibility is likely to be more important in the bacterial system than in the animal system; furthermore, it is only in the bacterial system that reversibility has been demonstrated.82 Interestingly, both processes are dependent on a specific tetrancleotide hot spot sequence.

The enzyme promoting the introduction of point mutations in human IgG genes is the activation-induced cytidine deaminase.81 By contrast, the enzyme that recognizes the tetrancleotide hot spot, CTAA, in E. coli is known to be the Insl5A transposase.49 In neither case does the enzyme know what is needed. However, in the latter case, the downstream binding of GlpR to its O1 binding site in part determines the frequency of mutation. Since the long term benefit to the organism cannot be denied, it is intuitively clear that the occurrence of directed mutation could be a previously unrecognized property of living organisms. It is presently impossible to say how widespread the phenomenon of directed mutation is likely to be.

Many switching mechanisms have been documented in microorganisms. These include flagellar phase variation in Salmonella, controlled by inversion of a DNA segment in the chromosomese82,83 and fimbral (fim) switching in E. coli, based on DNA methylation and binding of the Leucine-Responsive Protein (LRP) to specific sites in the fim control region.84,85 Over 100 genes in Neisseria species appear to be subject to switching, including the well characterized Neisserial opa and fim loci which change antigenic properties of cell surface proteins by two different mechanisms.86-88 A well-characterized eukaryotic switch mechanism involves sex determination in homothallic yeast strains.89-91 In both cases, the ability to alternate between two or more states is beneficial to the organism, and no one questions the conclusion that these mechanisms have evolved to provide survival advantages.

The same can be argued for directed mutation. Having alternatives (i.e., low versus high frequency of a mutational event, depending on physiological need) clearly would have survival value and therefore could have evolved through natural selection. We anticipate that there will be multiple mechanisms by which this occurs. The involvement of GlpR in control of the rate of IS5 hopping upstream of the O1 GlpR binding site may merely represent the first example where the mechanism of transposon-mediated directed mutation has been elucidated.

Our preliminary results with E. coli suggest that while the bgl and flhDC operons, both activated by ISS insertion, are not subject to directed mutation, selection for Ppd+ Fuc– mutations in the fuc operon of E. coli may be. This last system therefore provides a potential alternative system for examining and under-

---

**Figure 6.** Model for the activation of gfpFK expression by the upstream IS5 in E. coli K12. IB: IS5 permanent bend/IHF binding region in the downstream part of IS5. α, β, β', G, the four dissimilar subunits of E. coli RNA polymerase. The -10 and -35 represent the two sigma-70 (σ70) binding sites in the promoter region to which σ70 binds for open complex formation.

**Figure 7.** Schematic diagram illustrating GlpR-mediated control of (right) gfpFK transcription and (left) the rate of IS5 hopping (directed mutation) into the CTAA site upstream of the gfpFK promoter. With GlpR bound to its operators (O1-O4), transcription and IS5 hopping both occur at low rates. When GlpR is not bound to its operators, both transcriptional initiation and IS5 hopping increase about 10-fold. Binding of GlpR to operator O1 blocks ISS insertion, while binding of GlpR to operator O4 blocks transcription as indicated. Reproduced from reference 49.
standing transposon-mediated directed muta-
tion. Further studies will be required to iden-
tify other examples of directed mutation to
define their regulatory mechanisms and esti-
mate their importance to organismal evolu-
tion.

References

1. Luria SE, Delbruck M. Mutations of bacte-
ria from virus sensitivity to virus resistance. Genetics 1943;28:491-511.
2. Cairns J, Overbaugh J, Miller S. The origin of
3. Rosenberg SM. Evolving responsively:
adaptive mutation. Nat Rev Genet 2001;2:
504-15.
4. Foster PL. Mechanisms of stationary phase
5. Wright BE. Stress-directed adaptive muta-
6. Foster PL. Nonadaptive mutations occur
on the F episome during adaptive muta-
7. Hendrickson H, Slecha ES, Berghthorsson
that "directed" adaptive mutation and general hypermutability result from
growth with a selected gene amplification.
Origin of mutations under selection: the
adaptive mutation controversy. Annu Rev
9. Cairns J, Foster PL. Adaptive reversion of a
frameshift mutation in Escherichia coli.
Genome-wide hypermutation in a subpop-
ulation of stationary-phase cells underlies
recombination-dependent adaptive muta-
11. Harris RS, Longerich S, Rosenberg SM.
Recombination in adaptive mutation.
Science 1994;264:258-60.
12. Andersson DI, Slecha ES, Roth JR. Evidence
that gene amplification underlies adaptive mutability of the bacterial lac
13. Ponder RG, Fonville NC, Rosenberg SM. A
switch from high-fidelity to error-prone
DNA double-strand break repair underlies
19:791-804.
14. Layton JC, Foster PL. Error-prone DNA
polymerase IV is controlled by the stress-
response sigma factor, RpoS, in Escherichia coli. Mol Microbiol 2003;50:
549-61.
15. Lombardo MJ, Aponyi I, Rosenberg SM.
General stress response regulator RpoS in
adaptive mutation and amplification in
Escherichia coli. Genetics 2004;166:669-
80.
and point mutation are independent mecha-
nisms: evidence for various stress-
On the mechanism of gene amplification
induced under stress in Escherichia coli.
18. Wright BE, Minnick MF. Reversion rates in
a leuB auxotroph of Escherichia coli K-12 correlate with ppGpp levels during exponen-
Increased transcription rates correlate with increased reversion rates in leuB and
20. Wright BE, Longacre A, Reimers JM.
Hypermutation in derepressed operons of
Escherichia coli K12. Proc Natl Acad Sci U
S A 1999;96:5089-94.
Neidhardt FC, et al., editors. Escherichia
coli and Salmonella typhimurium: Cellular
and Molecular Biology. Washington, DC:
22. Roth JR, Andersson DI. Amplification-
mutagenesis—how growth under selection
contributes to the origin of genetic diver-
sity and explains the phenomenon of adap-
tive mutation. Res Microbiol 2004;155:42-
51.
23. Denamur E, Matic I. Evolution of mutation
rates in bacteria. Mol Microbiol 2006;60:
820-7.
24. Foster PL. Stress responses and genetic
variation in bacteria. Mutat Res 2005;569:
3-11.
25. Ahmed A. Evidence for replicative transpo-
sion of Tn5 and Tn9. J Mol Biol 1986;191:
75-84.
26. Dale JW. Movable genes. In: Dale JW, edi-
tor. Molecular Genetics of Bacteria.
England: John Wiley & Sons; 1998. pp. 187-
210.
27. Weinert TA, Schaus NA, Grindley ND.
Insertion sequence duplication in transposi-
tional recombination. Science 1983;222:
755-65.
28. Hallet B, Sherratt DJ. Transposition and
site-specific recombination: adapting DNA
cut-and-paste mechanisms to a variety of
gene rearrangements. FEMS Microbiol Rev
29. Craig NL. Mobile DNA II. Washington, DC:
30. Louarn JM, Bouche JP, Legendre F, Louarn
J, Patte J. Characterization and properties
of very large inversions of the E. coli chro-
mosome along the origin-to-terminus axis.
Mol Gen Genet 1985;201:467-76.
31. Reif HJ, Saedler H. I51 is involved in dele-
tion formation in the gal region of E. coli
32. Schneider D, Dupercy E, Courbes A, Lenski
RE, Blot M. Long-term experimen-
tal evolution in Escherichia coli. IX.
Characterization of insertion sequence-
mediated mutations and rearrangements.
33. Chandler M, Mahillon J. Insertion
sequences revised. In: Craig NL, Craig R,
Gellert M, Lambowitz AM, editors. Mobile
p. 305-66.
34. Galas DJ. Bacterial insertion sequences.
In: Berg D, Howe M, editors. Mobile DNA.
Washington, DC: American Society for
35. Mahillon J, Chandler M. Insertion
62:725-74.
36. McCalla DR. Mechanism of action of anti-
bacterial agents. Hahn FE, editor. Berlin,
Germany: Springer-Verlag; 1979.
Oxygen-insensitive nitroreductases:
analysis of the roles of nfsA and nfsB in
development of resistance to 5-nitrofuran
derivatives in Escherichia coli. J Bacteriol
38. Reynolds AE, Felton J, Wright A. Insertion
of DNA activates the cryptic bgl operon in
39. Schnetz R, Rak B, ISS: a mobile enhancer
transcription in Escherichia coli. Proc
40. Petersen C, Moller LB, Valentin-Hansen P.
The cryptic adenine deaminase gene of
Escherichia coli. Silencing by the nucleoid-associated DNA-binding protein,
H-NS, and activation by insertion elements.
41. Chen YM, Lu Z, Lin EC. Constitutive activa-
tion of the fucAO operon and silencing of the divergently transcribed fucPAB oper-
on by an ISS element in Escherichia coli
mutants selected for growth on L-1,2-
propanediol. J Bacteriol 1989;171:6097-
105.
42. Barker CS, Pruss BM, Matsumura P.
Increased motility of Escherichia coli by
insertion sequence element integration into
the regulatory region of the nfsB operon.
43. Lin EC. Glycerol dissimilation and its reg-
ulation in bacteria. Annu Rev Microbiol
44. Lin ECC. Dissimilatory pathways for sug-
ars, polyols, and carboxylates. In:


81. Jordan PW, Snyder LA, Saunders NJ.


