UNIVERSITY OF CALIFORNIA SAN DIEGO

Evidence of Directed Mutation in the *Escherichia coli bgl* Operon by Insertion of IS*5* and IS*1*

A Thesis submitted in partial satisfaction of the requirements for the degree

Master of Science

In

Biology

by

Kingswell Zhou

Committee in Charge:
Professor Milton Saier, Jr., Chair
Professor Matthew Daugherty
Professor Rachel Dutton

2019

The Thesis of Kingswell Zhou is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____
Chair

University of California San Diego

2019

TABLE OF CONTENTS

Signature Page…… iii

Table of Contents…… iv

List of Tables…… v

List of Figures……... vi

Acknowledgements…… vii

Abstract…… viii

Introduction……... 1

Materials and Methods…… 3

Results……... 8

Discussion…… 11

Tables…… 16

Figures……... 19

References…… 25

LIST OF TABLES

Table 1. List of primers used in Bgl^+ genotyping…… 16

Table 2. List of oligos used in strain construction…….. 17

Table 3. List of strains……. 18

LIST OF FIGURES

Figure 1. Increased Bgl^+ mutation frequency of *BW25113G*…… 19

Figure 2. Correlation of Bgl^+ mutation frequency with level of *bglG* expression 20

Figure 3. The level of *bglG* expression affects…… 21
the composition of Bgl^+ mutations.

Figure 4. Other carbon sources do not affect Bgl^+ mutation frequency…….. 22

Figure 5. *bglG* expression does not affect the…… 23
frequency of IS*5* insertions in fuc*AO*.

Figure 6. *bgl* operon activity increases in response to *bglG* expression…… 24

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Milton Saier, Jr. for his outstanding support and guidance as the chair of my committee. His scientific prowess has been a primary inspiration for my work.

I would also like to acknowledge the members of the Saier Lab, specifically Zhongge Zhang, for supervising and aiding my research. I hope my work proves worthy of the high standard held in my lab.

ABSTRACT OF THE THESIS

Evidence of Directed Mutation in the *Escherichia coli bgl* Operon by Insertion of IS*5* and IS*1*

by

Kingswell Zhou

Master of Science in Biology

University of California San Diego, 2019

Professor Milton Saier, Jr., Chair

The cryptic *bgl* operon in *E. coli* can be activated via mutation under starvation conditions to enable the uptake and metabolism of β-glucosides as a carbon source for growth and division. I found that the expression of one of the

genes in this operon, *bglG*, leads to an increase in operon mutation frequency that confers the Bgl^+ (β-glucoside utilizing) phenotype. I compared the mutation rates of a single copy, constitutive *bglG* expressing BW25113 *E. coli* strain to a wild type BW25113 strain on various agar plate mutation assays and used colony PCR to analyze differences in mutation frequencies. Then, I performed mutation assays on cells grown on a good carbon source (glycerol) and incubated on a poor carbon source (propanediol) to show that these differences arising from *bglG* expression are isolated to the *bgl* operon. To identify a plausible mechanism, I generated native promoter *bglG-lacZ* fusions as a reporter and analyzed promoter activity under different conditions. I found that *bglG* expression specifically increases *bgl*-activating mutations and that this effect is dependent on the presence of β-glucosides.

Introduction

Directed mutation describes the phenomenon in which an organism selects mutations favorable for growth or survival. In systems where directed mutation occurs, beneficial mutations occur at greater frequencies given certain environmental conditions, while harmful mutations occur at lower frequencies. Mutations in *Escherichia coli* have been used to illustrate this principle. Studies of operons such as the *glpFK* and *flhDC* operons have demonstrated the role of directed mutation in both carbon source utilization and organismal motility, showing that the principle of directed mutation is broad in application (Humayun et al., 2017). Insertion sequence (IS) element insertion, or transposition of an IS element to a specific site within the *E. coli* genome, was shown to be elevated in response to stress induced by factors such as starvation, toxicity, and limited motility. The focus of my studies on the *bgl* operon suggests another system that illustrates the principle of directed mutation.

The *bgl* operon in *E. coli* is a cryptic operon that encodes the proteins responsible for uptake and utilization of β-glucosides as a carbon source for growth (Schnetz et al., 1987). Three genes are present in this operon: *bglG*, *bglF*, and *bglB*. While the BglB protein catalyzes hydrolysis of phospho-β-glucosides, the PTS membrane enzyme responsible for uptake of β-glucosides is encoded by *bglF*. BglF catalyzes simultaneous uptake and phosphorylation of β-glucosides (Chen et al., 1997). Transcription is arrested by RNA stem loops,

designated rho-independent terminators, which flank the *bglG* DNA and RNA sequences (Schnetz & Rak, 1988). BglG is activated by phosphorylation by HPr to form a homodimer, which binds the RNA near these stem loop structures to destabilize the 2˚ structure, allowing transcription to continue. Dimerization of BglG is prevented in the absence of β-glucosides via phosphorylation by BglF. During uptake, β-glucosides occupy the phosphorylation site on BglF, favoring the activation of BglG. Phosphorylation of BglG by HPr serves as a mechanism of catabolite repression by alternative rapidly utilized sugars, such as glucose, while phosphorylation by BglF serves as an induction mechanism in response to the presence of a β-glucoside.

Most promoter-activating mutations involve transposon activities in which a mobile insertion element, IS*5* or IS*1*, inserts upstream of the promoter (Schnetz and Rak, 1992). Activation may also occur via insertion of other insertion elements such as IS*1*, IS*2*, and IS*10* at much lower frequencies. Mutations in other operons conferring the Bgl^+ phenotype are caused by transposition as well (Hall and Xu, 1991; Parker and Hall; 1989). Spontaneous transposition to the upstream regions of the *bgl* promoter occurs at higher frequencies than expected for such a limited region, which suggests that a regulating mechanism exists (Reynolds et al., 1981). “Targeted” mutation by transposition has been the major theme in the study of directed mutation. Regarding the *bgl* system, activation via transposition has been speculated to be a molecular switch for a population as opposed to an isolated mutational event.

In order to see if a gene product from within the *bgl* operon would promote mutation, I investigated the effect of single copy *bglG* expression on Bgl^+ mutation frequency for both insertion and non-insertion mutations. Since *bgl* is cryptic, I define this single copy expression system as *bglG* overexpression. The *bglG* gene was chosen because it was reported that BglG was the target for glucose-mediated *bgl* operon repression (Gulati and Mahadevan, 2000). This gave insight into the contribution of targeted transposition to increased mutation frequency at the molecular level. I tested this effect using two different β-glucosides, arbutin and salicin, which are both substrates of BglF and BglB, to demonstrate this phenomenon's dependence on environmental factors and to better understand the mechanism of directed mutation in this system.

Materials and Methods

Construction of the *bglG* expressing strain: chromosomal, P*tet*-driven *bglG*

Since overexpression of *bglG* is toxic to cells, P*tet* driven *bglG* cannot be cloned into a regular integration plasmid such as pKD13 before being moved to the chromosome. To make a chromosomal copy of P*tet* driven *bglG*, the region containing the kanamycin (km) resistance gene (km^r), the *rrn*B terminator (*rrnB*T) and P*tet* (that is, km^r:*rrnB*T:P*tet*) was amplified from pKDT (pKD) using primers Ints-P1 and Ptet-R (Table 2). *bglG* was amplified from BW25113 chromosomal DNA using primers bglG-F and IntS-bglG-P2 (Table 2). Please note that the 37

bp region in the 3' end of the "*km*^{r}:*rrnB*T:P*tet*" fragment is the same as the first 37 bp region of the *bglG* fragment since the reverse primer Ptet-R and the forward primer bglG-F shared this 37 bp overlapping region in their 5' ends. After gel purification, the *km*^{r}:*rrnB*T:P*tet* fragment and the *bglG* gene were fused together by standard fusion PCR using both PCR products in equal ratios as templates and IntS-P1 and IntS-bglG-P2 as primers. The resultant fused product, "*km*^{r}:*rrnB*T:P*tet-bglG*", was gel purified and then electroporated into the chromosome of BW25113 cells (expressing the λ-Red recombinase) to replace the *intS* gene using the Lamada-Red methods as described in (Datsenko and Wanner, 2000). After being confirmed by colony PCR and DNA sequencing, one correct clone was cultured in LB at 40˚C overnight to remove the Lamada-red plasmid. The final strain carrying the "*km*^{r}:*rrnB*T:P*tet-bglG*" construct at the *intS* locus was named BW25113*G* (Table 3).

To titrate the expression of *bglG,* the "*km*^{r}:*rrnB*T:P*tet-bglG*" construct was transferred from BW25113*G* by P1 phage transduction to the recipient strain BW-RI (Table 3), i.e., BW25113 constitutively expressing *tetR*. A P1 transductant was purified on a LB+Km plate, yielding strain BW-RI*G* (Table 3).

Construction of *Pbgl-bglG-lacZ* transcriptional fusion

Using primers Pbgl-Xho-F/bglG-BamH-R (Table 2), the region carrying the *bgl* promoter and the first gene in the operon, *bglG*, including the 2nd terminator (the 335th nucleotide upstream of the translational start codon of *bglG* to the 9th

codon of *bglF*) was amplified from BW25113 genomic DNA. A stop codon TAA was added to the end of *bglF*'s 9th codon. The PCR products, which carried the *bglGFB* promoter, the first terminator, *bglG*, the *bglG/bglF* intergenic region (containing the 2nd terminator) and the first 9 codons of *bglF* plus a stop codon in the 3' end, were digested with *Xho*I and *BamH*I and then inserted into the same sites of the plasmid pKDT (Klumpp et al., 2009), yielding the plasmids pKDT_P*bgl-bglG*. Present in this plasmid, the "*km*:*rrnB*T:P*bgl-bglG*" DNA fragment was amplified using primers Pbgl-Z1/bglG-Z2 and then integrated into the chromosome of default strain EQ42 (Klumpp et al., 2009) to replace the region containing *lacI* and the entire *lacZ* promoter. After being confirmed by colony PCR and DNA sequencing, the *PbgI-bglG-lacZ* transcriptional fusion was transferred to BW25113 and BW25113*G* by P1 phage transduction, yielding strains BW25113_*Z* and BW25113*G*_*Z,* respectively (Table 3).

Strains were liquid-cultured in LB or M9 minimal media with 0.5% salicin, arbutin, glycerol, or propanediol (PPD) as the carbon source at 37˚C or 30˚C. Solid media culturing was performed on 1.5% agar plates. Kanamycin (Km) or ampicillin (Amp) was added to media at 25 μg/ml and 100 μg/ml, respectively, as needed.

The working M9 salts concentration in both solid and liquid media consisted of: 34 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, and 9.4 mM NH_4Cl.

Comparison of Bgl^+ mutation frequency between BW25113 and BW25113*G*

Bgl^+ mutation rates were determined on minimal salicin and arbutin agar plates. Cells were grown overnight in LB, washed, and diluted with M9 salts. 2×10^7 cells were inoculated onto each plate and then examined every 24 hours for the appearance of arbutin- or salicin-positive colonies. The plates were incubated at 30˚C throughout the assay. To determine the presence of insertion elements proximal to the *bgl* promoter, colony PCR was performed with primers Pbgl-F2 and Pbgl-R2.

Bgl^+ mutation assay with induction of BW-RI*G*

Mutation was measured on minimal salicin agar plates with chlortetracycline (CLTC, 40 ng/ml) or anhydrotetracycline (ATC, 6 or 25 ng/ml) as inducers. BW-RI*G* was cultured as above, and 10^8 cells were inoculated onto the appropriate plates. IS mutants were confirmed as above.

Measurement of the effect of glycerol or propanediol (PPD) on Bgl^+ mutation frequency

Overnight LB cultures were washed with an M9 salts solution and plated onto minimal glycerol and minimal PPD agar plates. Plates were maintained at 30˚ C. To measure Bgl^+ mutation frequencies, a 1 cm coring tool was used to obtain cores from each plate. Each core was washed with M9 salts, with 100 μL being applied to minimal salicin plates at 37˚ C for Bgl^+ mutant screening.

Colonies that appeared from 48 to 60 hours were counted as mutants. A dilution of each wash was applied to LB plates for cell population count.

PPD$^+$ mutation assay with BW25113 and BW25113*G*

Propanediol-positive (PPD$^+$) mutation was measured on minimal PPD agar plates (Zhang et al., 2010). Plates were inoculated with 10^8 cells and incubated in 30˚C. Plates were examined every 48 hours for the appearance of PPD$^+$ mutants, which were verified for *fucAO*-proximal IS5 insertion by colony PCR using primers.

lacZ reporter measurements with *Pbgl-bglG-lacZ* reporter strains

All culturing for activity assays was performed in liquid media at 37˚C. Cells were cultured overnight in LB as a pre-culture, and then inoculated into fresh LB media at an OD_{600} of 0.020 for harvesting or minimal media containing glycerol only or both glycerol and salicin as minimal media pre-cultures. The minimal media pre-cultures were grown for 24 hours, and then inoculated into identical fresh minimal media at an OD_{600} of 0.020 for harvesting during linear growth. β-galactosidase activity was measured using the method of Miller (Miller, 1972).

Results

bglG expression increases the frequency of both IS and non-IS Bgl^+ mutations

At the end of an 11-day incubation on M9 salicin or arbutin agar plates, a large increase in mutation frequency, expressed in terms of total mutants per total cell number was observed (Figure 1 A & B). Total mutation frequency increased by 19 and 48-fold in response to *bglG* expression on arbutin and salicin, respectively. IS mutation frequency increased by 10 and 16-fold, respectively, in like manner. The percent of IS mutants relative to total mutants was found to be 53% and 33% for BW25113*G* on arbutin and salicin, respectively. However, the percentage of IS mutants from the wild-type was found to be approximately 95% on both arbutin and salicin.

An increase in Bgl^+ mutation frequency is dependent on the expression level of *bglG*

Higher levels of *bglG* expression exhibited higher Bgl^+ mutation frequencies. The *ptet-bglG* construct was transferred to a BW25113 strain expressing a *tet* promoter repressor. Mutant colonies were recorded on minimal medium agar plates containing salicin and different concentrations of chlortetracycline (CLTC) or anhydrotetracycline (ATC) as inducer. Appearance of mutant colonies was measured at no (salicin only), low (40 ng/mL CLTC), medium (80 ng/mL CLTC), and high induction of *tet* (25 ng/mL ATC). A

significantly greater number of mutants was observed on higher induction compared to lower induction after 6 days (Figure 2).

bglG expression increases the composition of non-IS mutants in Bgl^+ mutants

Greater induction of *bglG* expression yielded a greater percentage of non-IS mutants out of the total Bgl^+ mutants (Figure 3). Low and no induction yielded mutant compositions more similar to that of the wild type. Medium and high induction resulted in approximately a threefold and tenfold increase in percentage of non-IS mutants, respectively. Nevertheless, both the IS- and non-IS- type mutation rates increased markedly due to overexpression of *bglG*.

Enhancement of Bgl^+ mutation frequency by *bglG* is not seen in the absence of arbutin and salicin

In addition to the level of *bglG* expression, the presence of arbutin and salicin is essential for the enhancement of Bgl^+ mutation by *bglG*. In the presence of glycerol as the only carbon source, *bglG* expressing cells did not exhibit a significant increase in mutation frequency compared to the wild type (Figure 4A). A similar pattern was seen in the presence of a poor carbon source, PPD (Figure 4B). The highest number of mutants counted for a given sample was 23 mutants/10^8 cells and 13 mutants/10^8 cells for the *bglG* expressing and wild type strains, respectively. This difference is lower than two-fold and does not coincide with the large increase seen in the presence of arbutin or salicin.

Effect on the insertion frequency by *bglG* overexpression is specific to the *bgl* operon

To verify that the Bgl^+ mutations are targeted, that is, that this is a true example of directed mutation, enhancement of insertion by *bglG* must be observed only for the *bgl* operon. Transposition of insertion element IS5 activates the *fucAO* operon in *E. coli*, enabling growth on PPD as a sole carbon source (Chen et al., 1989). To show that increased mutation frequency caused by the expression of *bglG* is specific to the *bgl* operon, I tracked the growth of mutant colonies on minimal agar plates containing PPD, comparing BW25113 and BW25113*G*. No significant difference in PPD^+ mutation frequency was observed (Figure 5).

bglG overexpression increases *bgl* operon activity expression

To examine a possible cause of mutation frequency enhancement by high level *bglG* expression, reporter activity for *bgl* was measured. BW25113 and BW25113*G* constructs harboring *lacZ* fusions to the *phoU-bglGFB* intergenic region spanning to the second terminator sequence following *bglG* were assayed for operon expression (Figure 6). BW25113G showed an increase in *bgl* operon activity. β-galactosidase activity increased 20-, 13-, and 8-fold for BW25113*G* compared to wild-type constructs cultured in LB, minimal glycerol, and minimal glycerol plus salicin cultures, respectively.

Discussion

Dependence of increased Bgl$^+$ mutation rate on the presence of β-glucosides

A large-fold increase in the Bgl$^+$ mutation rate, in comparison to the wild type, was observed for *bglG* expression in the presence of arbutin or salicin as the sole carbon source (Figure 1). Specific mutation to the Bgl$^+$ phenotype is the dominant effect of *bglG* expression, as indicated by the lack of change in mutation for *fuc*, a carbon utilization system with similar activation by insertion (Figure 5). Furthermore, incubation on PPD or glycerol did not increase the Bgl$^+$ mutation rate in response to *bglG* expression (Figure 4). This indicates that the enhancement of Bgl$^+$ mutation frequency is dependent on the presence of a β-glucoside and not merely on starvation.

Enhancement of non-insertional mutations

Two classes of Bgl$^+$ mutation were observed, one with an IS element upstream of the *bgl* operon and one without such an insertion. While overexpression of *bglG* increases the frequency of both types of mutations, a larger increase in the non-insertional type was observed (Figure 2). The contribution of non-insertional mutants to the total mutant composition increased with the level of *bglG* expression. These non-IS mutants comprise only a small percent of the Bgl$^+$ mutants originating from the wild type. It is possible that *bglG* overexpression allows the cell to accumulate other types of *bgl*-activating

mutations in a starvation-induced stress environment. Studies have shown that spontaneously arising Bgl$^+$ mutations differ in different genetic backgrounds (Moorthy & Mahadevan, 2002). Mutation by insertion was dominant in salicin-utilizing mutants for an *rpoS*$^+$ background but not for an *rpoS*$^-$ background. In the same study, disruption of *rpoS* also increased the rate of Bgl$^+$ mutation. It is possible that increased production of BglG serves as an additional sensor of starvation, facilitating the change in mutation rate in the nutritional stress response. This upstream feedback to regulatory factors may be an indirect mechanism of enhanced non-IS Bgl$^+$ mutations.

Increased *bgl* operon activity in response to *bglG* overexpression and non-IS mutants

Reporter assays suggested a potential mechanism by which *bglG* overexpression might confer increased mutation rate. A 10-fold increase in *bgl* promoter activity was observed for a non-mutated (Bgl$^-$) *bglG* overexpressing strain compared to that of a non-overexpressing wild type strain in minimal growth conditions (Figure 6). Although no measurement exceeded 10 activity units, the slight increase in the basal level of *bgl* expression could have allowed individual cells to survive starvation until mutations arise. However, cell population assays were factored into the analysis of mutation frequency, which demonstrated that differences in mutation frequency were not caused by differences in survival. Furthermore, increased operon activity could have caused

slight unwinding of the three-dimensional DNA structure. Since IS*5* and IS*1* have high specificity for the region upstream of the *bgl* promoter, a small change in the DNA structure could increase the frequency of IS insertion substantially (Humayun et al., 2017).

Mapping of non-IS *bgl* mutations

So far, the *bgl*-activating mutation(s) in the non-IS mutants have not been mapped. PCR and sequencing analyses of the *chb* and *asc* operons as well as the *hns* and *bglJ* genes, all of which, when mutated, could possibly increase the frequency of *bgl* activating mutations, did not reveal mutations in these systems among purified non-IS Bgl^+ mutants. Point mutations were not detected in sequencing of the *bgl* operon among these mutants. Deletion of the *bgl* operon in these mutants resulted in Bgl^- strains, which means that the activating mutation is not within the *bgl* system but nevertheless relies on *bgl* activity. However, these mutants exhibited slower growth rates in both complex and minimal media (unpublished data). This suggests that the activating mutation may affect a global regulatory factor, such as an RNA polymerase subunit or another global transcription factor. Studies have shown that increased mutation rate is associated with slower growth rates for *E. coli* challenged by a nutritional deficiency (Nishimura et al., 2017). Because our mutation assay provides this kind of environment, it is possible that the mutation affecting growth rate may be the activating mutation itself or a factor that promotes the occurrence of the *bgl*

activating mutation. Still, as seen from the assays conducted on glycerol and propanediol, these mutations do not arise from decreased growth rate alone.

Understanding the enhancement of insertions by *bglG* expression

Since the non-IS Bgl^+ mutation has not been mapped, I cannot implicate any mechanism for the enhancement of this type of mutation. However, I can suggest two mechanisms by which *bglG* overexpression enhances IS insertion. H-NS is a global transcriptional silencer that also functions to compact DNA. Evidence has suggested that H-NS has two DNA binding sites within *bgl*: one downstream of the promoter proximal to the 3' end of *bglG*, and another within 100 bases upstream of the promoter (Dole et al., 2004). Binding of H-NS to the DNA may be counteracted by increased co-transcriptional translation of mRNA proximal to the H-NS binding site. The binding of H-NS repression complexes may function in promoter occlusion to prevent RNA polymerase binding or may assist in the formation of DNA-protein bridges with other H-NS complexes to trap the polymerase and/or prevent transcriptional elongation (Rangarajan & Schnetz, 2018). Since repression of *bgl* transcriptional elongation by H-NS has been demonstrated by previous studies, the DNA bridge complex may be present in *bgl*, supported by the existence of two H-NS binding sites in the operon. Measurement of operon activity showed that increased *bglG* expression leads to marginally increased read-through (Figure 6) and should contribute to de-repression of the *bglG* site by H-NS. Since H-NS binding is encouraged by

internal bends, dissolution of the DNA bridge complex may free additional H-NS from the upstream site (Ussery et al., 1994). Removal of H-NS from the upstream region may be sufficient to unmask the target region for IS insertion and therefore enhance mutation by insertion. Another possible consequence of H-NS removal at the *bglG* site is that the upstream H-NS site could alter the DNA bend angle to a phase that is more favorable for insertion.

The second possible mechanism involves an indirect effect of increased *bglG* read-through. Transcriptional elongation results in increased negative superhelicity of DNA opposite the direction of read-through (Gilbert and Allan, 2014). Spontaneous separation of DNA strands may occur in response to increased negative superhelicity of DNA structure, which is characterized as superhelical stress-induced DNA duplex destabilization (SIDD) (Humayun et al., 2017). Targeted IS-mediated gene activation is promoted by a SIDD structure present in the system, and strengthening of SIDD structures should result in increased IS insertion efficiency (Humayun et al., 2017). The SIDD zone in *bgl* matches the region of common insertion sites. It may be that enhanced read-through caused by *bglG* expression strengthens the negative superhelical character of the upstream SIDD, thus improving insertional efficiency.

Table 1. List of primers used in Bgl$^+$ genotyping

Name	Purpose	Analysis	Sequence (5' to 3')
Pbgl-F2 Pbgl-R2	Detection of insertion mutations near *bgl* promoter; forward and reverse primers.	Colony PCR, gel	TGGCGATGAGCTGGATAAACTGCTGTCAGTTCATGACTGCTCAAGGCATAC
bglF-ver-F2 bglF-ver-R2	Amplification of *bglF* region in non-IS Bgl$^+$ mutants; forward and reverse primers	Colony PCR, gel	GCAAAGCGGTAGAGGGCAAGTTATGACACCTTCCACCTGATTGGCAGCTGTTGC
bglF-ver-F3	Sequencing *bglF*	Sequencing	GTCACATTGTTGAATTACTCGTCATCG
bglF-ver-F4	Sequencing *bglF*	Sequencing	TCAACGGGAATTGATTTCACCGTCTG
Cel-ver-F Cel-ver-R	Detection of insertion mutation upstream *chb* promoter; forward and reverse primers	Colony PCR, gel	ACCTACTTTGTCGCGTTAGATGATACAAATGCTTCAATAATGACCGGAACTTC
chbR-ver-F chbR-ver-R	Detection of point mutation in *chbR*	Sequencing	ACAGACGAAACTGATTGAAGGCGATGTACAGAATACTGATATCTGGCATATC
ascG-ver-F ascG-ver-R	Detection of insertion mutation in *ascG*	Colony PCR, gel	GATTGTGCAAGGACAATCGGCAAACGGCCGATACATTCACAAATCATGCCAG
yjjQ-ver-F bglJ-ver-R	Detection of insertion mutation upstream of *bglJ*	Colony PCR, gel	ATTGGTTGTGTGATTTTGCATTCTGTGACTCCTGACTGGTATAGGTATGGAG
hns-ver-F hns-ver-R	Detection of mutations around *hns* promoter	Colony PCR, gel, sequencing	CTCACGTGCTGCGAAATCATCGGTGATCCGTATCGGTGTTATCCACGAAAC

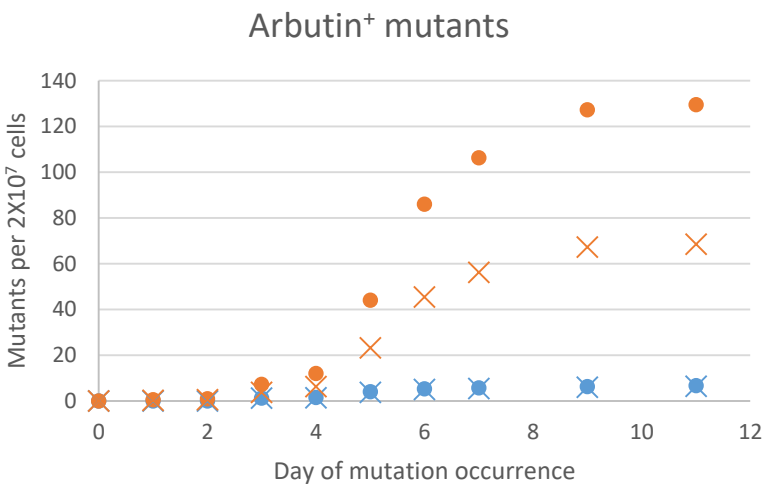
Table 2. List of oligos used in strain construction

Name	Purpose	Sequence (5' to 3')
IntS-P1	Amplification of *km*^{r}:*rrnB*T:P*tet* from plasmid PKDT and fusion of *km*^{r}:*rrnB*T:P*tet* to *bglG*	AGATTTACAGTTCGTCATGGTTCGCTTCAGATCGTTGACAGCCGCACTCCATGTGTAGGCTGGAGCTGCTTC
Ptet-R	Amplification of *km*^{r}:*rrnB*T:P*tet* from plasmid PKDT	GAATTTTGGTGATTTGCATGTTCATGGTACCTTTCTCCTCTTTAATGAATTC
bglG-F	Amplification of *bglG* from BW25113 chromosome	GAAAGGTACCATGAACATGCAAATCACCAAAATTCTC
IntS-bglG-P2	Amplification of *bglG* from BW25113 chromosome and fusion of *km*^{r}:*rrnB*T:P*tet to bglG*	GATAGTTGTTAAGGTCGCTCACTCCACCTTCTCATCAAGCCAGTCCGCCCTTCAGTGTTCTTTGCGCACGCGCTC
IntS-ver-R2	Verification of integration of *km*^{r}:*rrnB*T:P*tet-bglG* into BW25113	AAAGGAATGAAGTCTATCATCCAAGTC
Pbgl-Xho-F	Amplification of *bgl* operon fragment for *lacZ* fusions	ATACTCGAGTGGCGATGAGCTGGATAAACTGCTG
bglG-BamH-R	Amplification of *bgl* operon fragment for *lacZ* fusions	TTAGGATCCTTAGACTATTTTTCTGGCTAACTCCGTC
Pbgl-Z1	Amplification of *km*:*rrnB*T:P*bgl-bglG* from plasmid pKDT_P*bgl-bglG*	GCATTTACGTTGACACCATCGAATGGCGCAAAACCTTTCGCGGTATGTGTAGGCTGGAGCTGCTTC
bglG-Z2	Amplification of *km*:*rrnB*T:P*bgl-bglG* from plasmid pKDT_P*bgl-bglG*	CGACGGCCAGTGAATCCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTAGACTATTTTTCTGGCTAACTCC
pKDT Ptet	Plasmid template used for cloning *km*^{r}:*rrnB*T-fragments	

Table 3. List of strains

Strain	Abbreviation/Designation
K-12 BW25113	BW25113 or “wild type”
BW25113 P*tet-bglG*	BW25113*G*
K-12 BW25113 expressing *tet* repressor	BW-RI
BW-RI P*tet-bglG*	BWRI*G*
BW25113 P*bgl-bglG-lacZ*	BW25113_*Z*
BW25113 P*tet-bglG* P*bgl-bglG-lacZ*	BW25113*G*_*Z*
EQ42	EQ42

A



B

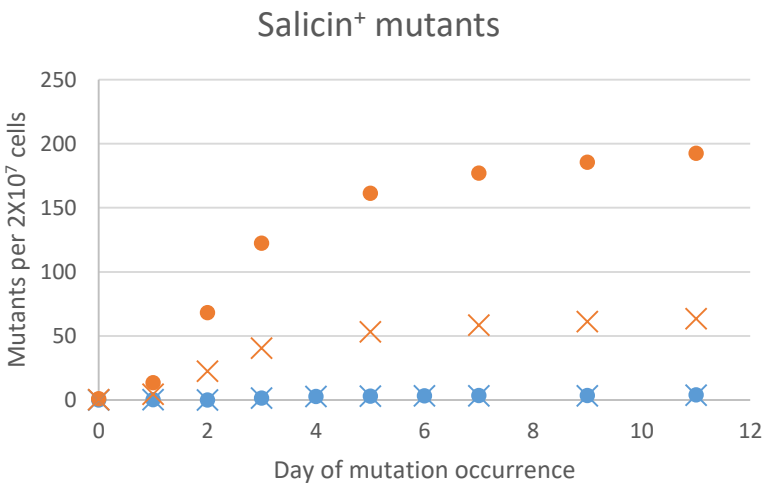


Figure 1. Increased Bgl$^+$ mutation frequency of *BW25113G*. Wild-type and BW25113*G* parent cells were cultured in LB. The overnight cultures were washed with M9 salts. 2×10^7 cells were plated onto M9 minimal agar plates containing 0.5% arbutin or 0.5% salicin. The plates were incubated at 30˚ C for 11 days, during which the appearance of mutant colonies was recorded. Verification of IS mutants was performed by PCR. **A, B** Total mutation frequency on arbutin and salicin, respectively, of wild type (blue circles) and BW25113*G* (orange circles). The number of IS mutants is indicated by X marks. Since pre-existing mutants appear at day two, “day of mutation occurrence” refers to two days prior to the appearance of novel mutant colonies.

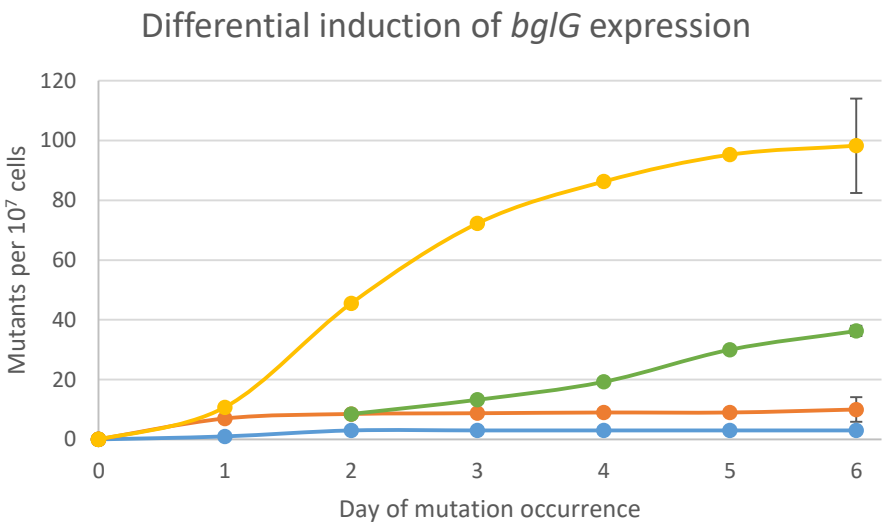


Figure 2. Correlation of Bgl$^+$ mutation frequency with level of *bglG* expression. BW-RIG was cultured in LB, washed, and diluted accordingly. 10^7 cells were inoculated onto salicin plates with varying levels of induction. Bgl$^+$ mutants were counted on no induction, (salicin only, blue), low induction (40 ng/mL CLTC, orange), medium induction (6 ng/mL ATC, green), and high induction (25 ng/mL ATC, yellow) of the *tet* promoter that drives *bglG* expression. At 6 days, medium induction yielded significantly more mutants than did low induction but significantly less mutants than did high induction.

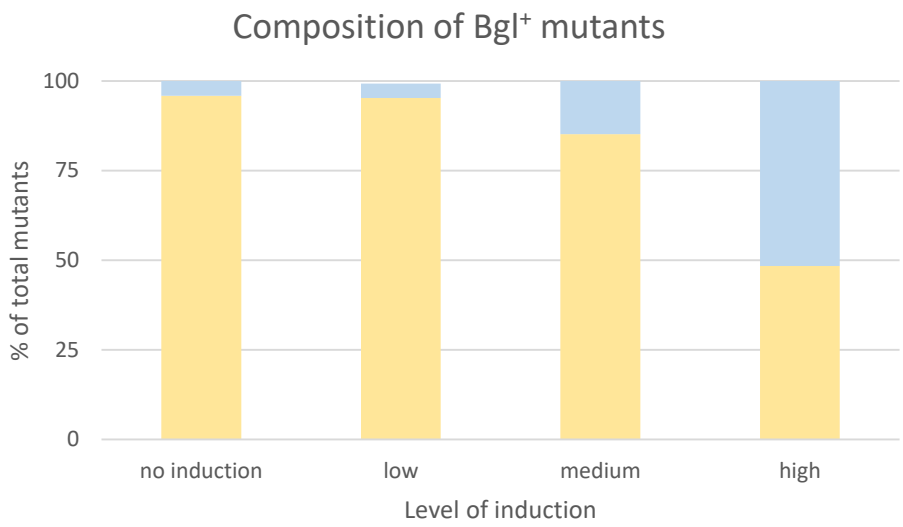


Figure 3. The level of *bglG* expression affects the composition of Bgl$^+$ mutants. Colonies from salicin plates (Fig. 2) were picked and verified by colony PCR for IS insertions. The ratio of the number of Bgl$^+$ non-IS mutants (top) to the total number of Bgl$^+$ mutants increases concurrently with the level of induction. $n_{Bw25113}$ = 48; $n_{40\ ng/mL\ CLTC}$ = 49; $n_{6\ ng/mL\ ATC}$ = 81; $n_{25\ ng/mL\ ATC}$ = 62.

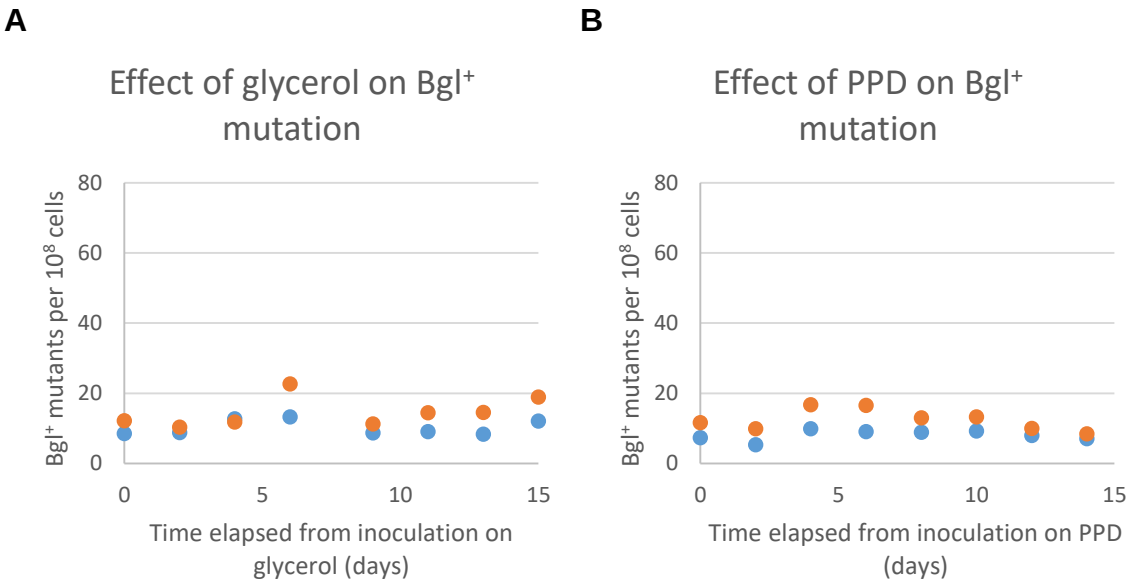


Figure 4. Other carbon sources do not affect Bgl$^+$ mutation frequency. Overnight BW25113 and BW25113*G* cultures grown in LB were washed with M9 salts solution, applied to minimal glycerol **(A)** and PPD **(B)** agar plates, and incubated at 30˚ C. Every 48 or 72 hours, cells from 1 cm cores obtained from each plate were washed with 1 mL M9 salts solution, after which 100 µL of the wash was applied to minimal salicin plates for Bgl$^+$ mutant screening and appropriate dilutions were applied to LB plates for cell count. No significant difference in Bgl$^+$ mutation frequency between the wild type (blue) and *bglG* expression (orange) was observed. Each data point represents the number of mutants, normalized to the population, that was observed for a core assayed on a given day.

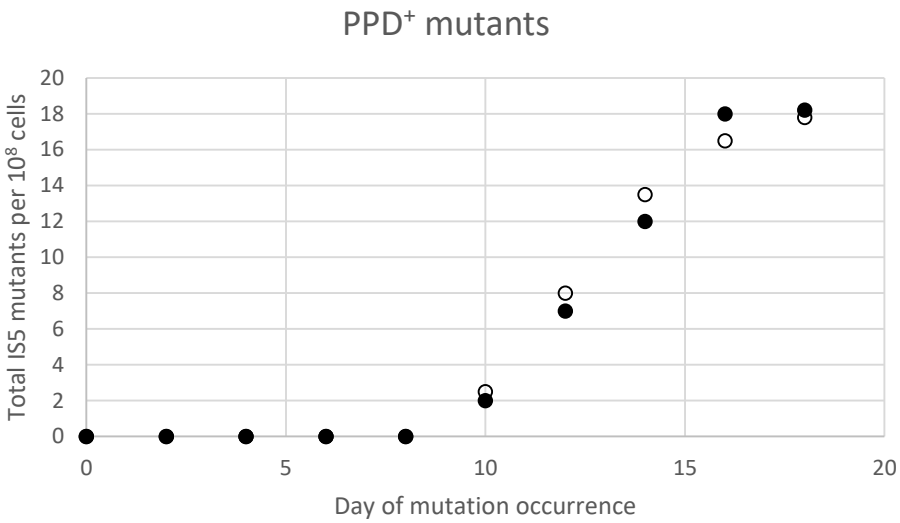


Figure 5. *bglG* expression does not affect the frequency of IS*5* insertions in fuc*AO*.
10^8 wild type (solid circles) and BW25113*G* (open circles) cells were applied to minimal propanediol plates and incubated at 30˚ C. Colonies were counted and totaled every 48 hours, and *fucAO* IS*5* mutants were confirmed by colony PCR. No significant difference was observed in the PPD^+ mutation frequency.

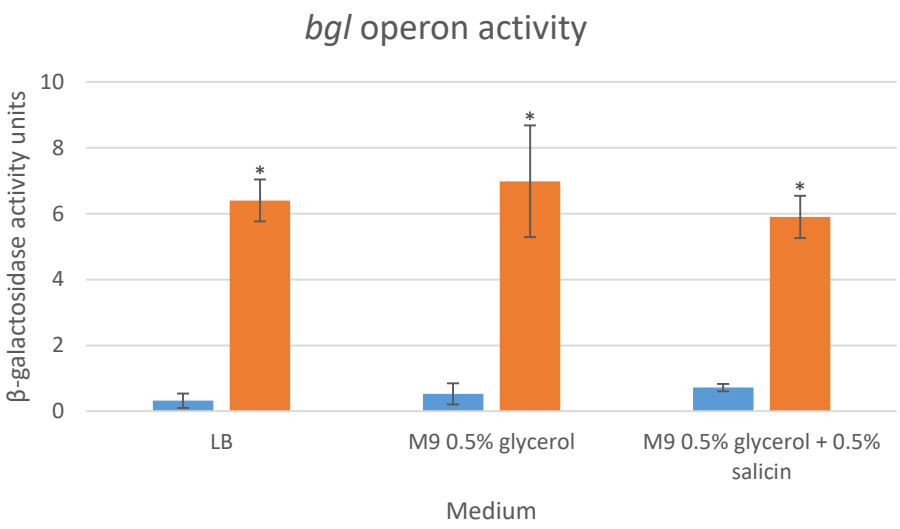


Figure 6. *bgl* operon activity increases in response to *bglG* expression. The wild-type sequence of the upstream intergenic region of *bgl*, along with *bglG* and its flanking terminator sequences, was fused to *lacZ* to serve as a reporter of native operon activity. This fusion was transferred to both wild-type (blue) and BW25113*G* (orange) strains. To measure β-galactosidase activity, the strains were first grown in an LB pre-culture for 6 hours and transferred to their respective minimal media pre-cultures to grow for another 24 hours. They were then transferred to their respective experimental media at an initial OD_{600} of 0.020. Samples were harvested at four time points within the OD_{600} range from 0.2 to 1.3 for the activity assay. * $p < 0.05$.

References

Chen Q, Arents JC, Bader R, Postma PW, Amster-Choder O. BglF, the sensor of the *E. coli bgl* system, uses the same site to phosphorylate both a sugar and a regulatory protein. The EMBO Journal. 1997 Apr; 16(15):4617-4627.

Chen YM, Lu Z, Lin EC. Constitutive activation of the *fucAO* operon and silencing of the divergently transcribed *fucPIK* operon by an IS*5* element in *Escherichia coli* mutants selected for growth on L-1,2-propanediol. Journal of Bacteriology. 1989 Nov; 171(11):6097-105.

Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. PNAS USA. 2000 June; 97:6640-5.

Dole S, Nagarajavel V, Schnetz K. The histone-like nucleoid structuring protein H-NS represses the *Escherichia coli bgl* operon downstream of the promoter. Molecular Microbiology. 2004 March; 52(2):589-600.

Gilbert N, Allan J. Supercoiling in DNA and chromatin. Current Opinion in Genetics and Development. 2014 Apr; 25:15-21.

Gulati A, Mahadevan S. Mechanism of catabolite repression in the *bgl* operon of *Escherichia coli*: involvement of the anti-terminator BglG, CRP-cAMP and EIIAGlc in mediating glucose effect downstream of transcription initiation. Genes to Cells. 2001 December; 5(4):239-250.

Hall BG, Xu L. Nucleotide Sequence, Function, Activation, and Evolution of the Cryptic *asc* Operon of *Escherichia coli* K12. Molecular Biology and Evolution. 1992 July; 9(4):688-706.

Humayun MZ, Zhang Z, Butcher AM, Moshayedi A, Saier, Jr., MH. Hopping into a hot seat: Role of DNA structural features on IS5-mediated gene activation and inactivation under stress. PLoS One. 2017 June; 12(6):e0180156.

Klumpp S, Zhang Z, Hwa T. Growth rate-dependent global effects on gene expression in bacteria. Cell. 2009 Dec; 139:1366–75.

Miller JH (1972). Experiments in Molecular Genetics. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Moorthy S, Mahadevan S. Differential spectrums of mutations that activate the *Escherichia coli bgl* operon in an *rpoS* genetic background. Journal of Bacteriology. 2002 July; 184 (14):4033-8.

Nishimura I, Kurokawa M, Liu L, Ying BW. Coordinated Changes in Mutation and Growth Rates Induced by Genome Reduction. MBio. 2017 July; 8(4):e00676-17.

Parker LL, Hall BG. Mechanisms of Activation of the Cryptic *cel* Operon of *Escherichia coli* K12. Genetics. 1990 Mar; 124(3):473-82.

Rangarajan AA, Schnetz K. Interference of transcription across H-NS binding sites and repression by H-NS. Molecular Microbiology. 2018 May; 108(3):226-239.

Schnetz K, Rak B. IS5: A mobile enhancer of transcription in *Escherichia coli*. PNAS USA. 1992 Feb; 89:1244-1248.

Schnetz K, Rak B. Regulation of the *bgl* operon of *Escherichia coli* by transcriptional antitermination. The EMBO Journal. 1988 June; 7(10):3271-3277.

Schnetz K, Tolocyzki C, Rak B. P-Glucoside (*bgl*) Operon of *Escherichia coli* K-12: Nucleotide Sequence, Genetic Organization, and Possible Evolutionary Relationship to Regulatory Components of Two *Bacillus subtilis* Genes. Journal of Bacteriology. 1987 June; 169(6):2579-2590.

Ussery DW, Hinton JC, Jordi BJ, Granum PE, Seirafi A, Stephen RJ, Tupper AE, Berridge G, Sidebotham JM, Higgins CF. The chromatin-associated protein H-NS. Biochimie. 1994; 76(10-11):968-80.

Zhang Z, Ming RY, Saier, Jr., MH. Precise Excision of IS*5* from the Intergenic Region between the *fucPIK* and the *fucAO* Operons and Mutational Control of *fucPIK* Operon Expression in *Escherichia coli*. Journal of Bacteriology. 2010 Apr; 192(7):2013-2019.