UNIVERSITY OF CALIFORNIA SAN DIEGO

The Effects of DNA-Binding Proteins on Insertion Sequence Element Transposition Upstream of the *bgl* Operon in *Escherichia coli*

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

in

Biology

by

Peter Kopkowski

Committee in charge:

Professor Milton H. Saier, Jr., Chair
Professor James Golden
Professor Justin Meyer

2022

The Thesis of Peter Kopkowski is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my wife and mother for supporting me throughout my academic career. I would also like to express my gratitude for Dr. Milton Saier, Dr. Zhongge Zhang, Dr. Arturo Medrano-Soto, Dennis Tran, Katie Lam, and Harry Zhou, without whom I would have been lost far more often.

The information contained in this thesis, in full, is being prepared for submission for publication. Kopkowski, Peter; Zhang, Zhongge; Saier, Milton H., 2022. The thesis author was the primary investigator and author of this paper.

ABSTRACT OF THE THESIS

The Effects of DNA-Binding Proteins on Insertion Sequence Element Transposition Upstream of the *bgl* Operon in *Escherichia coli*

by

Peter Kopkowski

Master of Science in Biology

University of California San Diego, 2022

Professor Milton H. Saier, Jr., Chair

The *bglGFB* operon in *Escherichia coli* K-12 strain BW25113, encoding the proteins necessary for the uptake and metabolism of β-glucosides, is normally not expressed. Insertion of either IS1 or IS5 upstream of the *bgl* promoter activates expression of the *bgl* operon when the cells are starving in the presence of a β-glucoside, drastically increasing transcription and allowing the cell to survive using this carbon source. Details surrounding the exact mechanism and regulation of the IS insertional event remain unclear. In this work, the role of several DNA-binding

proteins influencing the rates of insertion upstream of *bgl* are examined via mutation assays and protocols measuring transcription. Both cAMP-activated Crp and IHF exert a positive effect on insertional Bgl^+ mutations when present in the cell. The mechanism of how these two proteins cause this effect are discussed. Our results characterize IHF's effect in conjunction with other mutations, show that IHF's effect on IS insertion rate into *bgl* also affects other operons, and indicate that it likely exerts its effect by binding to and altering the DNA conformation of IS1 and IS5 in their native locations, rather than by directly influencing transposase gene expression. By contrast, the cyclic AMP-CRP complex acts directly on the *bgl* operon by binding to its site upstream of the *bgl* promoter.

INTRODUCTION

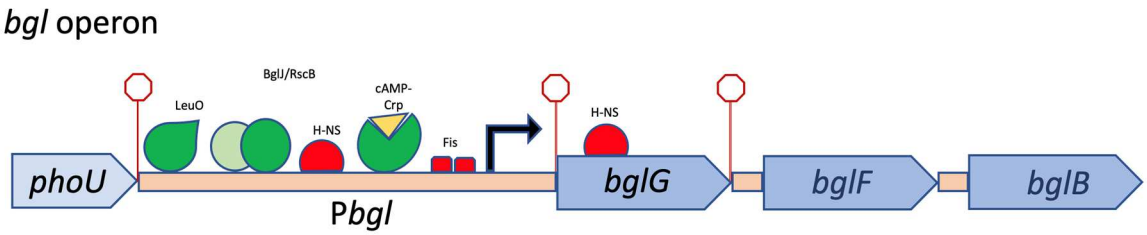
1.1. The *bgl* Operon, Adaptive Mutation, and Project Goals

Since their discovery, transposons have been studied in prokaryotic and eukaryotic models for their ability to insert at variable locations within an organism's genome [1]. In some cases, transposition takes place upstream of or within a structural gene, which may cause a change in protein expression for the former and loss of gene function for the latter [2-4].

One of the best known and characterized examples of this phenomenon is the *bglGFB* operon, which is not expressed in wild-type (WT) *Escherichia coli* K-12 strain BW25113. Binding of the histone-like nucleoid structuring (H-NS) protein at two sites flanking the otherwise active *bgl* promoter is the most important factor in silencing transcription of the *bgl* operon by a strong repression mechanism [5-7]. Data recently published by our group indicated that the upstream and downstream H-NS binding sites exhibit synergy with each other and may create a repression loop that blocks access of RNA polymerase to the *bgl* promoter (P*bgl*) [7, 8].

The first gene in the operon, *bglG*, contains the downstream H-NS binding site within its coding region and is itself flanked by two *rho*-independent terminators, limiting the amount of the RNA transcript that is made of *bglG* as well as the two downstream genes, *bglF* and *bglB* (Scheme 1) [9, 10]. BglG is a homodimer that binds to its own transcript and prevents early termination, allowing transcription to continue and promoting expression of the entire operon [11]. BglG also has other functions including the positive regulation of insertional and noninsertional Bgl^+

mutations, although how it accomplishes these functions has yet to be elucidated [12]. The *bglF* gene immediately follows *bglG*'s downstream terminator and encodes a membrane-integrated protein responsible for the uptake and concomitant phosphorylation of β-glucosides via a phosphotransferase (PTS)-dependent mechanism [13]. Thus, BglF passes a phosphoryl group from HPr or a BglG monomer to the incoming β-glucoside concomitant with transport [14], marking it for hydrolysis of the aglycone from the glucose-phosphate moiety by BglB, the product of the third gene in the operon [15], and this sugar- P feeds directly into glycolysis. Since phosphorylated BglG is in equilibrium with phosphorylated BglF, transfer of the phosphoryl group from BglG to BglF allows the dephosphorylated BglG to dimerize and enter its active anti-termination state upon its additional phosphorylation on a distinct histidyl residue in BglG by HPr [16, 17].



Scheme 1- The *bgl* operon with relevant DNA-binding protein sites and structural features.
The *Pbgl* intergenic region is where IS insertion occurs, upstream of the promoter. Several DNA-binding proteins are known to bind in the promoter region and influence transcription. A major focus of this work is to determine whether protein binding in this region influences the rate of insertion of IS1 and/or IS5.

While repression of *bgl* by H-NS is strong, preventing almost 100% of the maximal transcription rate, transposition of an Insertion Sequence (IS) element (IS1 or IS5) may occur upstream of the *bgl* operon promoter [18]. As noted above, most of IS elements that insert into this area are either IS1 or IS5, and both elements insert in

either orientation in a ~200 bp area [18]. This insertional event eliminates H-NS repression of *bgl* operon expression [19, 20]. This event increases *bgl* operon expression several-hundredfold [7] and allows cells to grow using β-glucosides as the sole carbon source [21], which we define as possessing a 'Bgl$^+$' phenotype.

Two details about this insertional event remain of great interest. First, Bgl$^+$ mutations of any type only occur when the *E. coli* cell is starving in the presence of a β-glucoside, which it can freely take in and use after *bgl* operon activation via mutation or IS insertion. Secondly, the rate of insertion into *bgl* is much higher than the random mutation rate for *E. coli* [22]. Since mutations are currently considered to be random events rather than directed by the environment, paradigms such as that of *bgl* deserve further study so that regulation of these operons can be better understood and incorporated into today's accepted models of mutation and evolution.

As noted above, *bgl* is one of several operons in *E. coli* that are preferential sites for insertion of IS elements. These sites are commonly found in the operon's promoter region, have high Gibbs free energy signatures, and exhibit an increase in IS insertion frequency when bacterial cells are experiencing a specific type of stress [23]. These so-called superhelical stress-induced duplex destabilization (SIDD) sites are under study by our group and others to determine whether insertion into a SIDD site is directed by environmental conditions, as well as to know whether SIDDs evolved specifically to allow IS element-mediated operon activation [23].

Since the hydrogen bonds between individual base pairs are relatively weak in a SIDD site, the opening of the DNA may be what allows IS insertion to occur there

with increased frequency. This would support previous unpublished findings that H-NS lowers the rate of insertion into the *bgl* operon (Z.Z., unpublished data). When H-NS binds at or near the SIDD site, it may prevent it from opening to a conformation that allows for easy insertion by transposons. It is therefore important to identify what roles, if any, DNA-binding proteins have upon transposition into the sequences surrounding their insertion sites. Towards this goal, we studied several DNA-binding proteins to better understand their roles in regulating the insertion of IS elements and the phenomena of adaptive and directed mutations in general. Adding to the body of knowledge of how transposons and their movements are regulated is vital for an understanding of the "mobilomes" present in almost every life form on Earth.

In this work, we show evidence that two DNA-binding proteins affect the rate of IS insertion into *bgl*: the cAMP-Crp complex and integration host factor (IHF). Both proteins are positive regulators of insertion; that is, their presences and abundances maintain or increase the frequencies of IS-related Bgl^+ mutations observed in WT cells. The effect of IHF is not specific to *bgl* but applies to several other operons into which IS elements insert in response to environmental stresses. Our results suggest that binding of IHF to known sites in IS1 and IS5 have an effect on DNA conformation, transpososome formation and/or the energetics at the site of transposition.

1.2. DNA Binding Proteins Relevant to this Research

Integration host factor (IHF) is a heterodimeric global histone-like DNA-binding protein involved in several cellular functions including transcriptional regulation, DNA

recombination, and chromosome compaction [24, 25]. After binding to the minor groove of the double helix, it significantly bends the DNA at least 160° per dimer [26]. Under starvation conditions, IHF concentrations increase 4- to 8-fold in *E. coli* [27, 28] and has been implicated in the induction of several proteins related to carbon starvation [29]. Also of interest is IHF's role in the transposition of Mu prophage in *Pseudomonas putida*. *In vitro* studies of IHF binding to the Mu promoter suggest that IHF assists in formation of a transposase complex that may facilitate excision via supercoiling relief [30, 31]. *In vivo*, IHF is not required for Mu transposition to occur but has a dual effect on Mu transposase transcription. IHF binds to the end of the Mu transposase promoter region and activates transcription of the transposase while simultaneously relieving repression by H-NS at a downstream site [32]. Several groups have shown that a Gly to Glu mutation at the 62nd amino acid in IhfA causes loss of IHF's ability to bind to its DNA consensus sequence [33, 34] but does not prevent dimerization with IhfB.

Crp is a DNA-binding protein that relies upon cyclic AMP to become activated [35]. It is a global transcriptional regulator that affects expression of almost 200 genes in *E. coli* [36]. Most applicable to this work, cAMP-Crp is generally involved in the positive regulation of genes concerned with the metabolism of carbon sources [37, 38] and is a necessary factor in the activation of *bgl*. Even if repression by H-NS is relieved, *bgl* remains transcriptionally inactive and unable to grow using only β-glucosides unless activated Crp is present [39, 40].

RESULTS

2.1. Deletion of ihfA Decreases the Bgl$^+$ Mutation Rate

The first phase of our research was to determine which genes, if any, demonstrated a clear effect on the frequency of IS insertion into *bgl* upon deletion. Figure 1A and 1B show data gathered from several single deletion strains over a 10-day period. The full list of deletion strains tested is tabulated under Materials and Methods (Section 4.1). Genes were selected for study based on the DNA-binding abilities of their gene products and were loosely separated into two groups: the global histone-like DNA-binding proteins and operon-specific DNA-binding proteins with known binding sites within the *bgl* control region. Of the deletion strains tested, Δ*ihfA* alone was selected for further study, as it caused the greatest change in the frequency of Bgl$^+$ mutant appearances compared to WT (Figure 1A and 1B). Further confirmatory Bgl$^+$ mutation assays were performed between WT and Δ*ihfA*, showing again that the Δ*ihfA* mutant gave rise to only one fifth the number of Bgl$^+$ mutants as WT, and only one seventh or less as many insertional colonies as WT (Figure 1C and 1D). An IS+ Δ*ihfA* strain was constructed and was Bgl$^+$, showing that Δ*ihfA* acts specifically to lower the IS insertion frequency and that it does not hinder growth of Bgl$^+$ mutants. These observations led us to consider the mechanism of IHF's role in IS insertion, as well its specificity for the *bgl* operon.

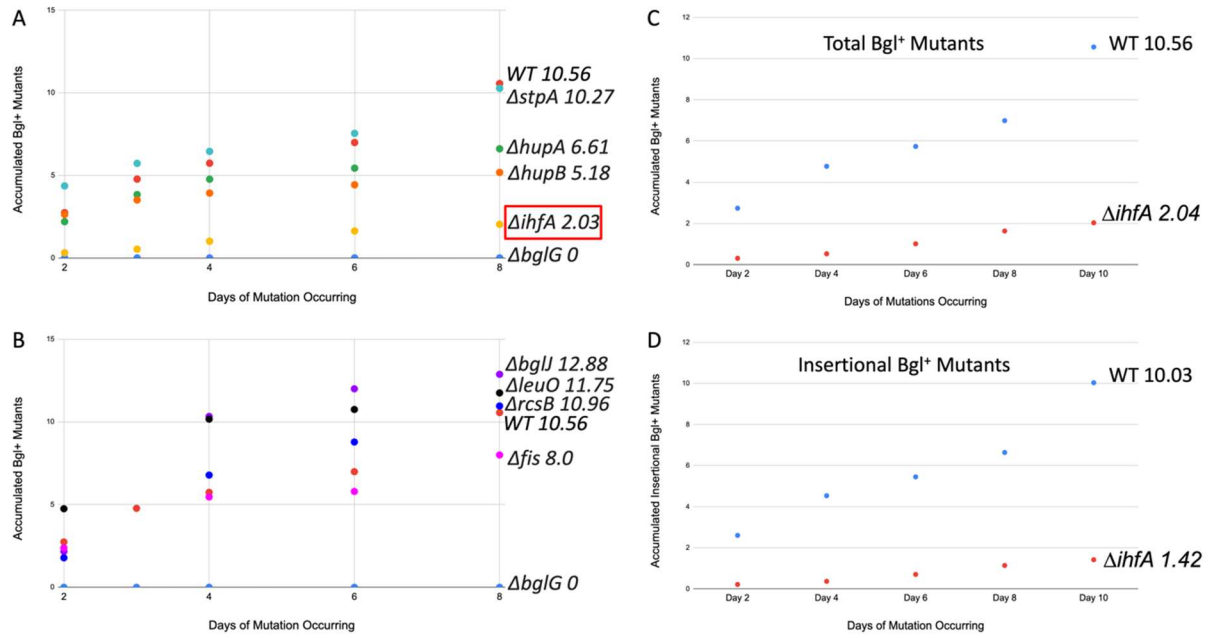


Figure 1- Bgl⁺ mutation frequencies due to deletion of genes encoding DNA-binding proteins.

A and B) Mutation assays were carried out on M9 + 0.5% salicin plates as described in Materials and Methods. **C and D)** Further mutation assays performed between the WT and the Δ*ihfA* mutant. Colony PCR using primers flanking the region of *bgl* insertion was performed to differentiate between insertional and non-insertional colonies.

2.2. Further characterization of an ihfA deletion mutant

The Δ*ihfA* background was transferred to P*tet bglG* and *lacI*$_q$ *bglG* strains, which overexpress the gene encoding the antiterminator protein BglG. Our objective was to observe how IHF's previously observed effect would interact with increased BglG levels, which has been previously shown to increase the rate of Bgl⁺ insertional and noninsertional mutations [12]. The results presented in Figure 2 show that the stimulatory effect of P*tet-bglG* on the insertional frequency counteracts Δ*ihfA's* negative effect on mutation rate. However, the insertional frequency is similar between *lacI*$_q$ *bglG* strains, although in the Δ*ihfA* background, the appearance of Bgl⁺ colonies fell by about 20%.

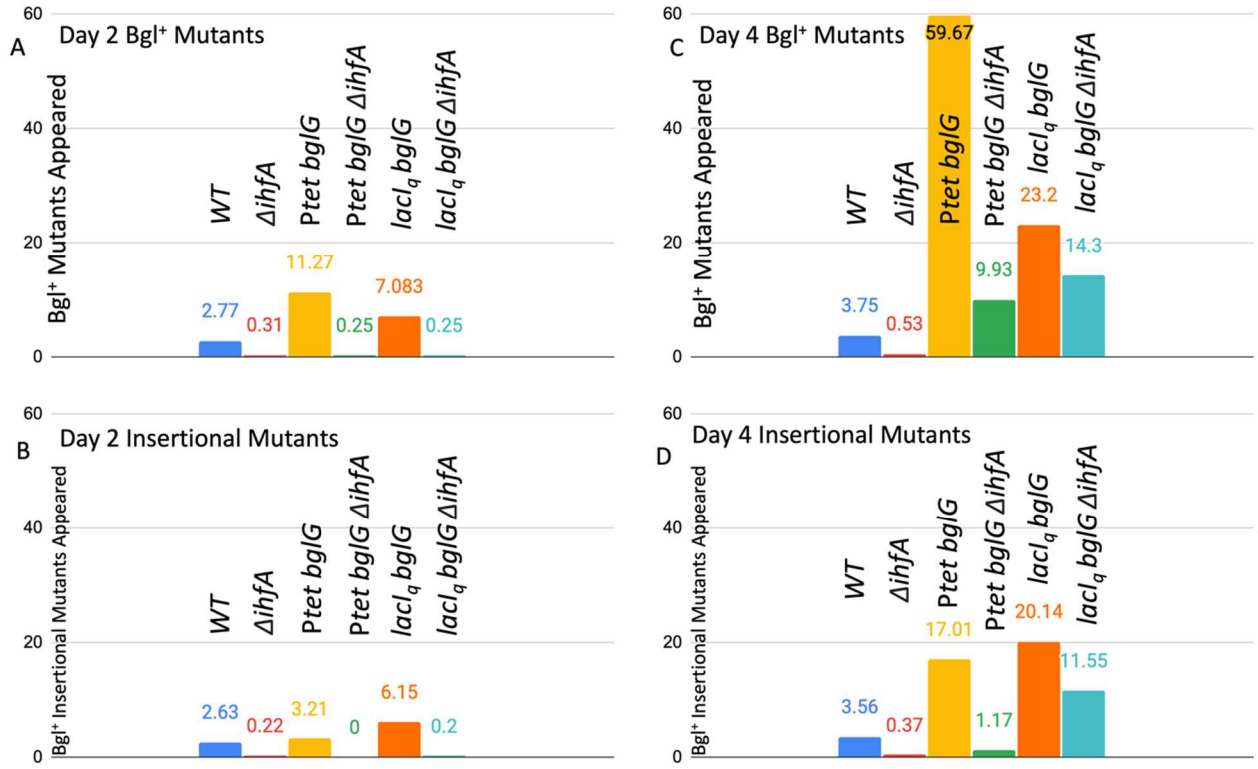


Figure 2- The effect of Δ*ihfA* combined with increased *bglG* expression.
A and B) The M9-salicin utilization assay as previously described, with two time points shown. Here we used WT, P*tet-bglG* and *lacI*$_q$ *bglG*. This experiment serves as a control with which to compare the Δ*ihfA* strains. Colony PCR was used to differentiate insertional and noninsertional Bgl$^+$ mutants. **C and D)** The M9 salicin assay was as previously described, with two time points shown. The three strains used in A) and B) with Δ*ihfA* were assayed. Colony PCR was used to differentiate between insertional and non-insertional Bgl$^+$ mutants.

2.3. Characterization of P*tet-driven ihfA/B expression*

Since Δ*ihfA* had a significant effect on IS insertion upstream of the *bgl* promoter, we decided to determine if increasing the presence of both IhfA and IhfB subunits would have the opposite effect. *PBPA* (P*tet-ihfA* + P*tet-ihfB*), a strain expressing both IHF subunits using the constitutive P*tet* promoter, was therefore constructed and subjected to a mutation assay as previously described. Interestingly, an approximately 2-fold decrease in both total and insertional mutants was observed

after 10 days (Figure 3A). We therefore considered several possible reasons for why this occurred (see the Discussion, Section 3.1).

PBPA-RI was constructed, containing the *tetR* gene, which encodes the repressor protein for P*tet* which is most effective in the absence of its inducer, any one of several tetracycline derivatives. Our goal was to see if different levels of P*tet* induction would reveal a similar trend as that already observed when using *PBPA*. Figure 3B shows the results of a mutation assay as described, but with the addition of 5 uM or 30 uM of the *Tet* inducer anhydro-tetracycline (aTC). The level of mutants was similar, even following maximal aTC induction (30 uM), supporting the results of the previous *PBPA* assay.

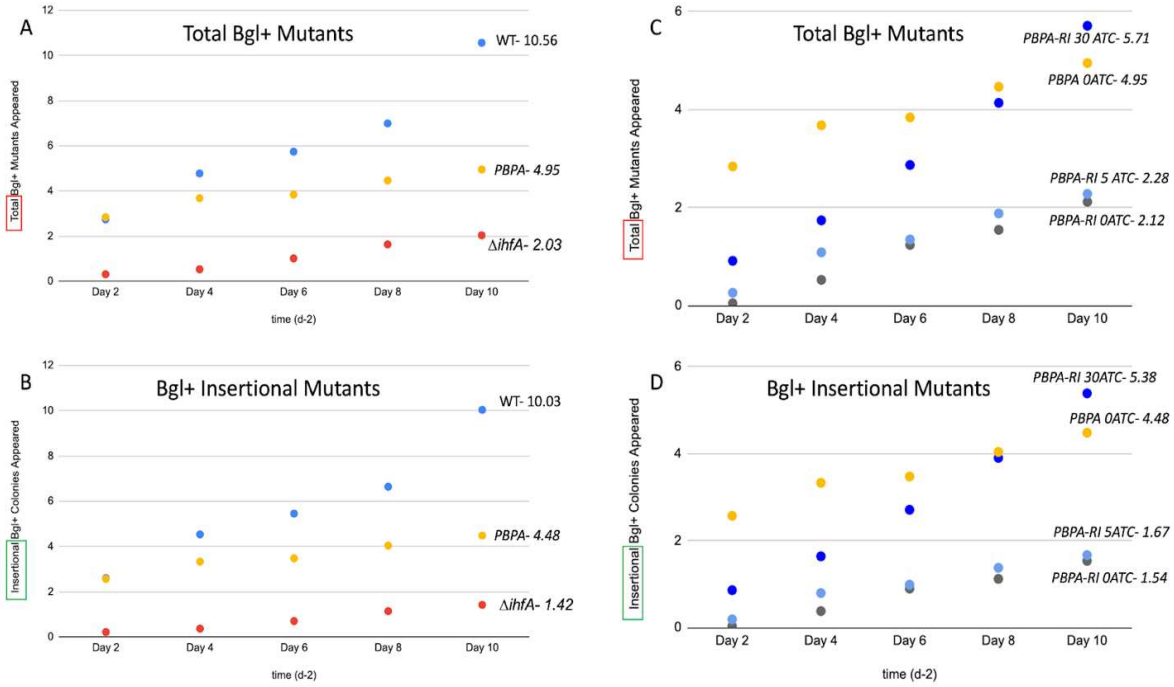


Figure 3- Overexpression of *ihfA* and *ihfB* slightly lowers the IS insertional rate into *bgl*.
A) and B) Bgl^+ mutation assays and cPCR verification were performed as previously described using WT, Δ*ihfA*, and *PBPA*. A: plots of total mutants; B: plots only of insertional mutants. **C) and D)** Mutation assays were performed as previously described on M9 salicin plates containing either 0 uM, 5 uM, or 30 uM ATC over a 12-day period. C shows total Bgl^+ mutants, while D shows only insertional Bgl^+ mutants.

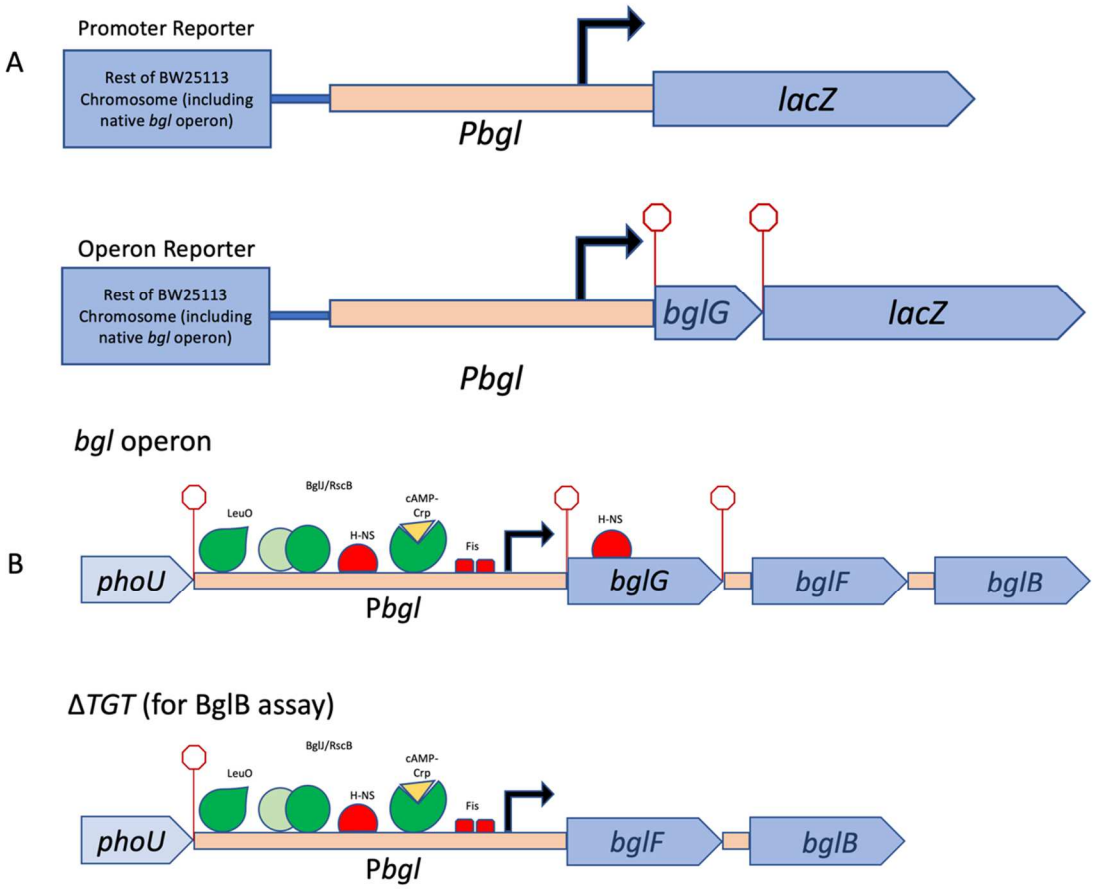
2.4. Changing IHF levels has no significant effect on transcription of the bgl operon

Our next goal was to establish a mechanism by which IHF exerts its effect on *bgl* operon expression. If Δ*ihfA*'s effect on the frequency of Bgl^+ insertional mutations is due to its inability to bind to *bgl*'s upstream region, then *lacZ* assays should show a difference in transcriptional activity when IHF is or is not present. Figure 4A shows the constructs devised for this purpose. The native *bgl* operon was left intact in order to maintain normal induction by β-glucosides. However, a second, altered *bgl* promoter and its upstream region was positioned in front of the native *lacZ* gene, instead of the usual *lacZ* promoter. This new construct can measure the promoter activity of this secondary *bgl* operon (see Scheme 2). Another similar construct which also contains the *bglG* gene was made in order to test the operon activity of *bgl* (see Scheme 2). Figure 4B contains the *lacZ* activity results, which demonstrated a 2-fold difference between Δ*ihfA* and WT in the promoter assay, but no significant overall change in operon activity was observed.

We repeated these experiments using a BglB assay, which functions similarly in practice to the LacZ assay. However, this assay uses PNP-Glucoside (PNPG) as its substrate and does not rely on constructs involving the *lacZ* gene. Rather, the BglB assay measures *bgl* operon transcription directly, and the strains used in this experiment confirmed the *LacZ* results as expected.

We suspected that any change that was to be seen in operon transcriptional activity was not detectable because the LacZ and BglB assay values were low (Figure 4C). To investigate this possibility, the Δ*TGT* strains were constructed, which

are similar to those originally used for the BglB assay, with an important change: the Δ*TGT* strains lack *bglG* and both flanking terminators in their native positions. With the loss of transcriptional termination as well as the loss of the downstream H-NS binding site, the base Δ*TGT* has a BglB assay activity value of approximately 50 units, which is about 200 times higher than the WT BglB assay value (Figure 4C and 4D). In order to see if any change in these strains was observable, BglB assays were run again using all of the above strains in a Δ*TGT* background. No large change was observed across the board, leading us to reject the notion that IHF lacks significant specific binding to the *bgl* operon.



Scheme 2- Construction of strains used to measure *bgl* transcriptional activity.
A) A separate *bgl* operon was placed in front of *lacZ* at its native location, leaving the native *bgl* operon unchanged to allow normal induction by the substrate. This *bgl* reporter contained either the *bgl* promoter region alone to measure promoter activity, or the promoter region followed by *bglG* to measure operon activity. **B)** The native *bgl* operon was used for Figure 4C, but Δ*TGT* was constructed (Figure 4B) in order to increase overall transcriptional activity and see any potential changes more clearly. In the Δ*TGT* strain, the *bglG* gene and both terminators were missing downstream of the *bgl* promoter, but the rest of the operon remained as before.

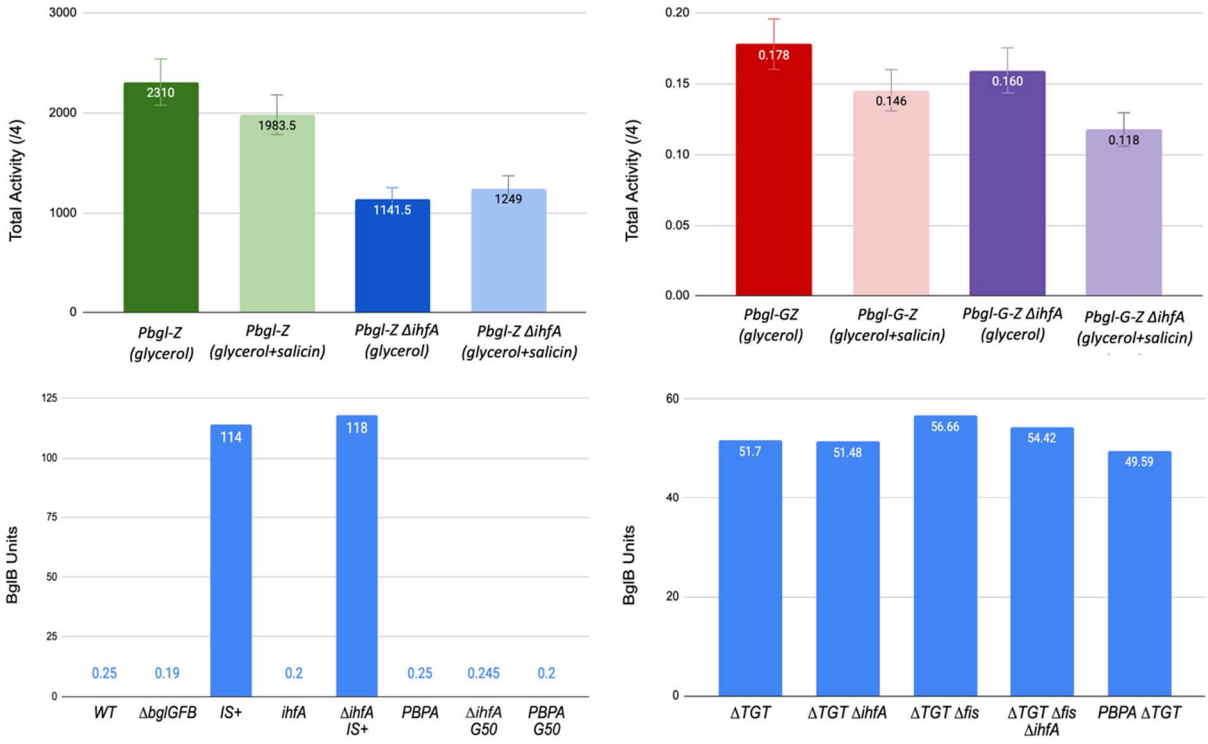


Figure 4- Changing IHF levels does not have a significant effect on *bgl* operon transcriptional activity.
LacZ and BglB assays carried out as previously described (Materials and Methods 4.4) **A)** *lacZ* assay, promoter activity. **B)** *lacZ* assay, operon activity. **C)** BglB assay. **D)** BglB assay using Δ*TGT strains*.

2.5 The effects of ihfA deletion on expression of several IS-activated operons

Since our *lacZ* assay data did not reveal a substantial change in transcriptional activity in the context of the *bgl* operon, we hypothesized that Δ*ihfA* does not affect *bgl* directly but exerts its effect on insertion upstream of *bgl* by a non-specific means. If IHF is important for IS insertion generally, then it may be involved in the upstream process of IS excision from other locations in the genome. In this situation, other operons into which IS1 and/or IS5 insert would also experience a change in insertion frequency upon deletion of *ihfA*. To explore this possibility, a Δ*ihfA* strain was tested alongside an isogenic WT strain using the mutation assay for several operons where preferential IS insertion into their respective SIDD sites have

been previously documented [4, 23, 41-44]. Agar plates were made with the intention of promoting insertion into the *flhDC*, *glpFK*, and *fucPIK* operons as well as the *bglFGB* operon, and were inoculated with a set number of cells, specific to the paradigm tested. The results for the three additional operons are presented in Figure 5, and all three showed a significant decrease in insertion frequency as observed for *bgl*.

In these experiments, Δ*ihfA* was compared with WT, just as was the case for *bgl*. However, the parameters of each experiment differed depending on the specific operon being tested. Interestingly, all the operons examined showed a significant decrease in the IS+ mutant appearance rate, suggesting that the dependency on IHF affects several operons into which insertions occur, and consequently, it may affect excision or transcription of the IS1 or IS5 element.

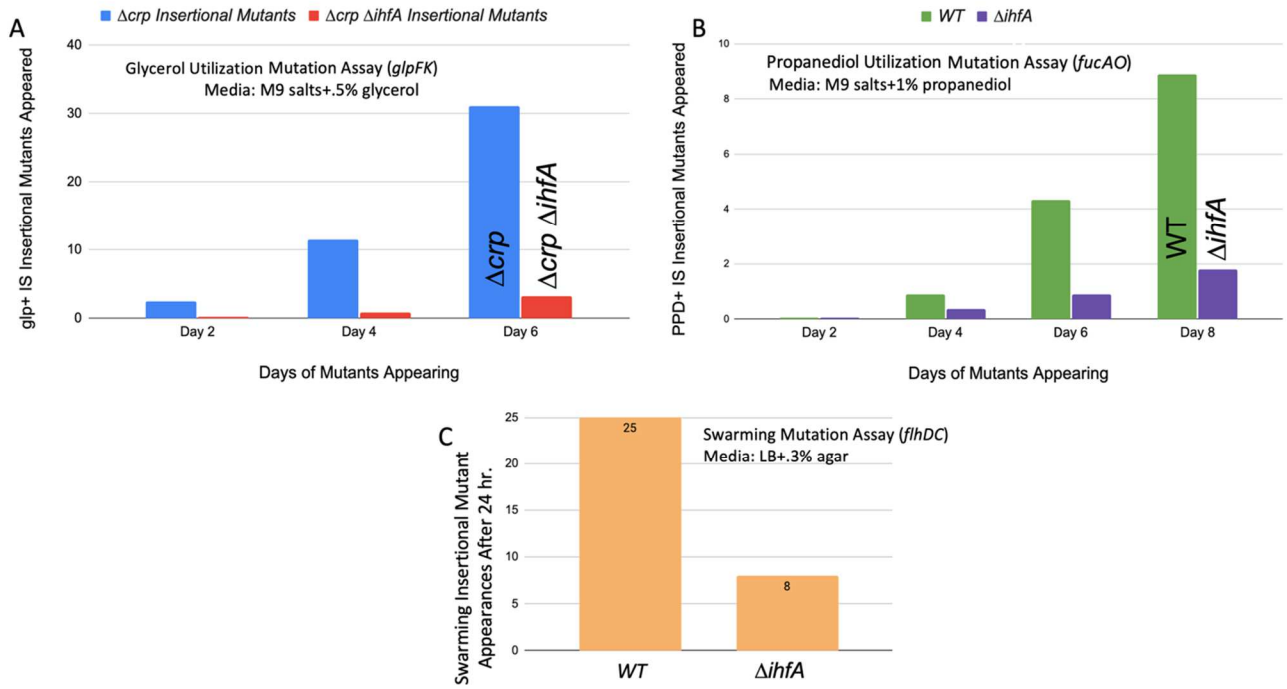


Figure 5- Deletion of *ihfA* affects IS1/5 insertion into other operons.
Mutation assays as described in Materials and Methods 4.2 (additional mutation assays). ***A)*** *A Δcrp mutant and a Δcrp ΔihfA mutant were subjected to the mutation assay on M9 + 0.5% glycerol media as previously described.* ***B)*** *WT and ΔihfA strains were subjected to the mutation assay on 1% propanediol media as previously described* ***(Materials and Methods 4.4). C)*** *WT and ΔihfA were subjected to the mutation assay on 0.3% LB agar media as previously described* ***(Materials and Methods 4.4).***

2.6 Examining the mechanism of the non-specific effect of IHF on IS1/IS5 transposition by loss of the ihfA DNA-binding function and lacZ measurements of IS1/IS5 Promoters

Since our view of IHF's effect proved to be relevant to other operons, the question remained: What is the mechanism of the IHF effect on transposition? First, we tested whether the DNA-binding ability to IHF was the cause of the observed multi-operon effect. A mutation assay was performed comparing WT and Δ*ihfA* to two other strains. These strains have P*tet*-driven *ihfA* at the *intS* locus, but the *ihfA* gene contains a G-E substitution at the 62nd position (referred to as *ihfA*G62E). The product of this *ihfA* mutant loses its DNA-binding function but is still able to dimerize

with IhfB [33, 34]. The *ihfAG62E* construct was placed into WT as well as the strain lacking *ihfA* at its original position (Δ*ihfA)*. If binding to DNA was important for the IHF effect, then Δ*ihfA*+*ihfA*G62E would show similar growth as the normal Δ*ihfA*, while the *ihfA*G62E mutant strain would have both binding and nonbinding IhfA and therefore would not experience as drastic of a change. Figure 6 shows the results of a *bgl* mutation assay, conducted as previously described. This strain grew according to expectation.

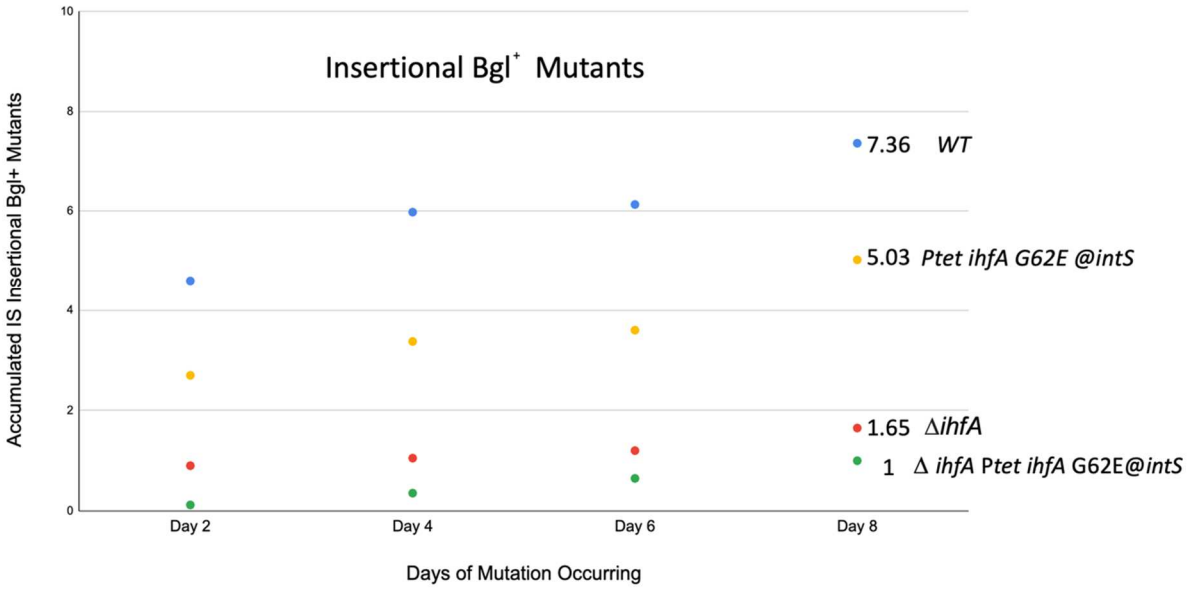


Figure 6- Binding of IHF is necessary to effectively exert its positive effect on IS insertion upstream of the *bgl* operon.
Mutation assays were carried out as previously described; cPCR was performed to verify insertional colonies.

2.7 IHF has no significant effect on transcription of IS1 or IS5

Research conducted by other laboratory groups has demonstrated that IHF binding sites are present at both ends of IS1 and on one end of IS5 (hereafter referred to as IS5B; IS5A is the end without an IHF binding site) [45-47]. We therefore decided to test whether the presence and/or binding of IHF causes a

change in the level of transcription of either or both IS elements. This was accomplished by placing the promoter regions for IS1 and both ends of IS5 (containing the Inverted Repeats to which IHF binds) directly before the native *lacZ* gene. This construct was placed in the WT, Δ*ihfA*, P*tet*-driven *ihfA* @*ints*, and P*tet*-driven *ihfA*G62E @*intS* (containing the G-E substitution at the 62nd position that abolishes the DNA-binding ability of IHF). The results for each reporter are presented in Figure 7. Overall, no significant change was observed, regardless of the genetic background. Since IHF is known to bind to either end of IS1 and to IS5B, these results strongly suggest that IHF exerts its effect via direct binding to IS1 and IS5, but not by altering the IS1 or IS5 expression level.

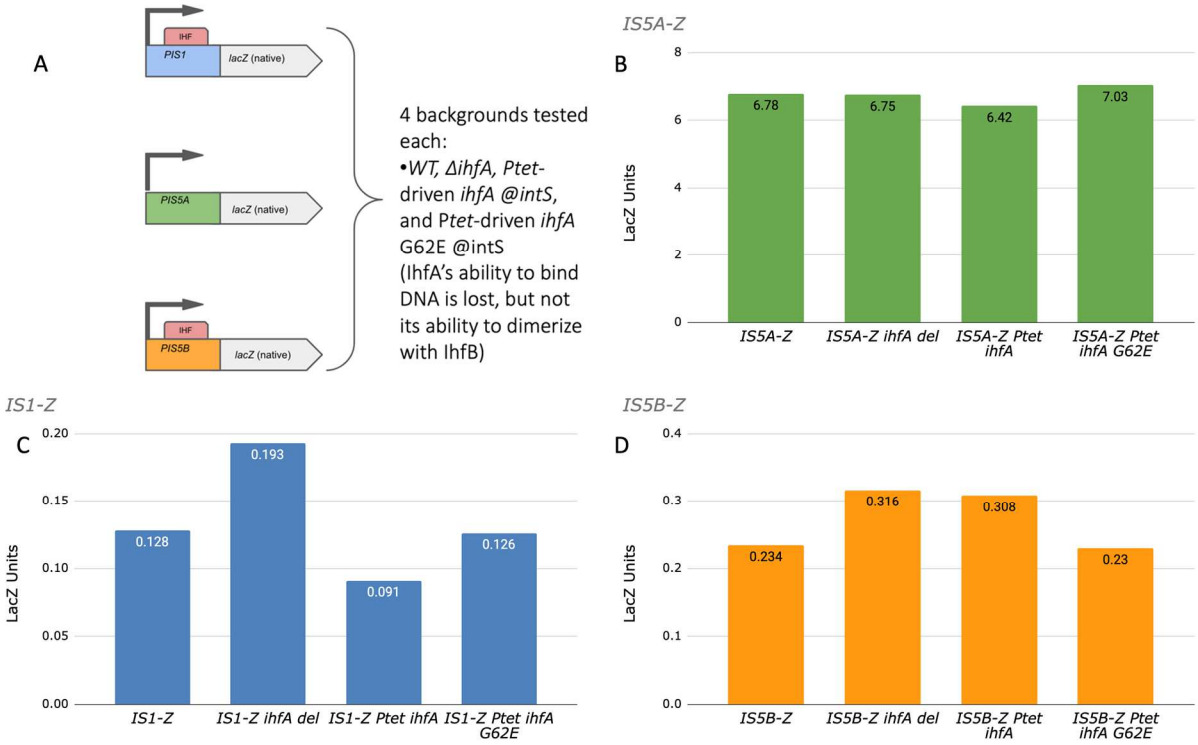


Figure 7- IHF does not influence transcriptional activity of IS1 or IS5.
A) A graph showing the construction of each *lacZ* reporter, which uses the promoter region of the 5' end of IS1 and both ends of IS5 (PIS5A and PIS5B). **B)** *lacZ* activity levels for PIS1-Z activity in backgrounds modulating the presence of IhfA. **C)** *lacZ* activity levels for PIS5A-Z activity in backgrounds modulating the presence of IhfA. **D)** *lacZ* activity levels for PIS5B-Z activity in backgrounds modulating the presence of IHF.

2.8 cAMP-Crp is a positive regulator of IS insertion when present upstream of Pbgl

In order to test whether Crp (when activated by cAMP) affects the rate of insertional mutations in the upstream region, two types of mutation assays were conducted. Simply testing a Δ*crp* deletion mutant is not an option here, as Crp is required for transcription of the *bgl* operon. If insertion takes place in a *crp* deletion mutant, we would be unable to observe the activity since the colony would not grow.

Therefore, we first decided to perform mutation assays using WT on normal M9 salicin plates and compare mutation rates to those growing on plates with an excess of cAMP. The results of this experiment are shown in Figure 8A. They demonstrate a slight increase in mutant appearance of the WT when excess cAMP is provided, but no increase in the appearance of insertional mutants. Next, we compared these WT results to those of a *cyaA* deletion mutant. *cyaA* encodes adenylate cyclase, and without it, the cells are unable to produce cAMP and must rely on extracellular cAMP provided. In Figure 8B, we observed a more than 3-fold decrease in the *cyaA* mutant growing in the presence of excess cAMP as compared with cells lacking cAMP. Interestingly, the number of total colonies was similar to those observed for the wild-type strain, suggesting that the lack of *cyaA* has a significant effect on IS insertion specifically.

The next set of experiments used *cpdA* to control the levels of intracellular cAMP. *cpdA* encodes a phosphodiesterase which degrades cAMP to 5'-AMP, and in its absence, cAMP levels increase. Similarly, if *cpdA* is expressed at higher levels, the amount of cAMP, and therefore the amount of active Crp, should decrease. A *cpdA* deletion mutant was constructed (Δ*cpdA)*, and a pZA31 plasmid containing

P*tet-cpdA* was electroporated into WT cells, yielding strain P*tet-cpdA.* For an isogenic control, separate WT and Δ*cpdA* strains were given an 'empty' pZA31 plasmid with a P*tet* promoter, yielding the absence of a functional gene product (pZA31 P*tet* Rf). We thus produced strains WT Rf and Δ*cpdA Rf*. All three strains were subjected to the *bgl* mutation assay as previously described. The results are presented in Figure 8C. The Δ*cpdA* mutant showed a 2-fold increase in insertional mutants compared to the WT strain, and the *cpdA* overexpression strain showed a similar number of total colonies as WT, but only half as many insertional mutants. Together, these data show that Crp, when activated by cAMP, is a positive regulator of insertion upstream of the *bgl* operon. In the absence of cAMP via deletion of *cyaA* or via increased degradation of cAMP, the insertion rate decreased 2-3-fold.

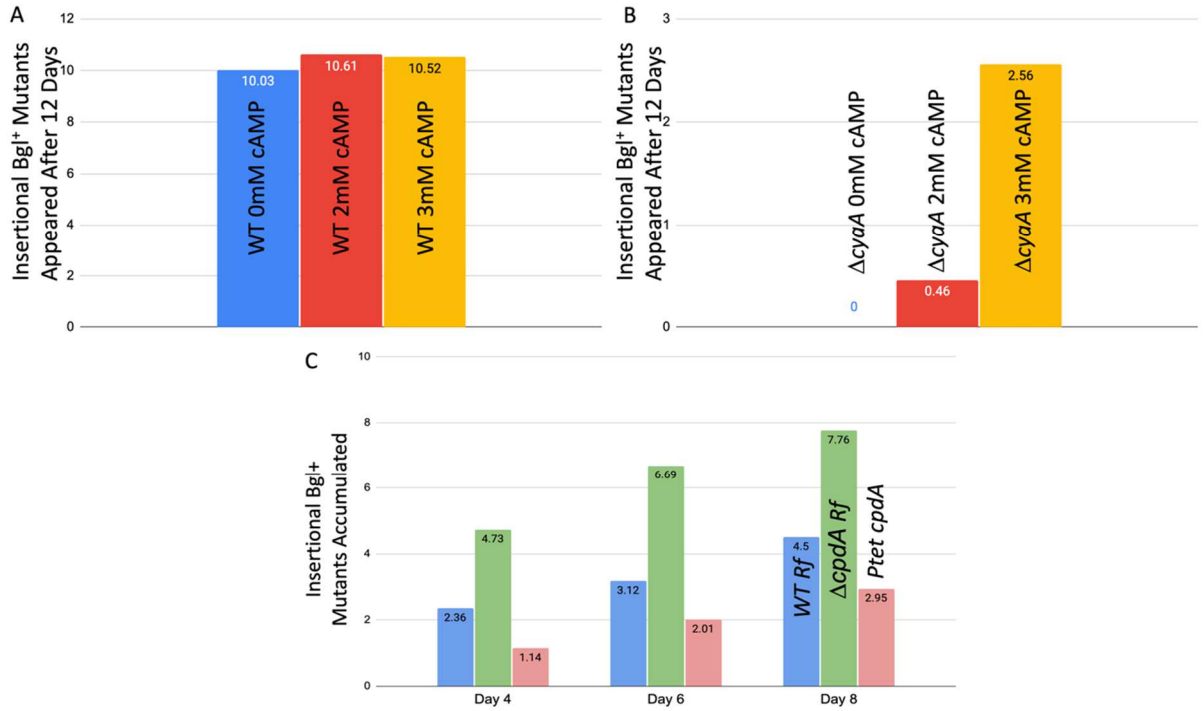


Figure 8- cAMP-Crp is a positive regulator of IS insertion into *bgl*.
A) and **B)** The results of a Bgl$^+$ mutation assay as previously described; **A)** shows total Bgl$^+$ mutants, while **B)** shows insertional mutants only. **C)** cPCR verification results displaying the average insertion frequencies for each strain tested. **D)** The results of a Bgl$^+$ mutation assay comparing WT on media containing 0 cAMP, WT on media containing 2.7 mM cAMP, and Δ*cyaA* on media containing 2.7 mM cAMP. Total Bgl$^+$ mutants and insertional mutants are indicated separately.

2.9- ihfA deletion or ihfA/B constitutive expression with differential cAMP-cpdA levels

In order to observe how loss of the *ihfA* gene affects mutation frequencies with differing cAMP-Crp complex levels, the three strains mentioned in section 1.8 were transferred into the Δ*ihfA* genetic background. The results of this set of mutation assays are shown in Figures 9A and 9B. Two of the three Δ*ihfA* strains performed as expected, with mutation rates equal to or lower than that observed for the normal Δ*ihfA* mutant. Strain 2 (Δ*ihfA*+Δ*cpdA*) displayed a 2-fold increase relative to the normal Δ*ihfA* strain, although the rate remained significantly below that observed for the WT strain.

To observe how *ihfA/ihfB* overexpression interacts with differing cAMP-Crp complex levels as in section 1.9, the three strains mentioned in section 1.8 were transferred into the *PBPA* genetic background. The results of this set of mutation assays are shown in Figures 9C and 9D. Strains 4 and 5 grew essentially the same as *PBPA*, but Strain 6 (P*tet cpdA* + *PBPA*) was quite low, demonstrating a more than 10-fold decrease in IS insertion rate compared to that observed for its counterpart lacking *cpdA* overexpression, PBPA Rf (Figure 9D). The synergy between P*tet ihfA/B* and P*tet cpdA* is surprising and deserves more study (see Discussion section 3.1).

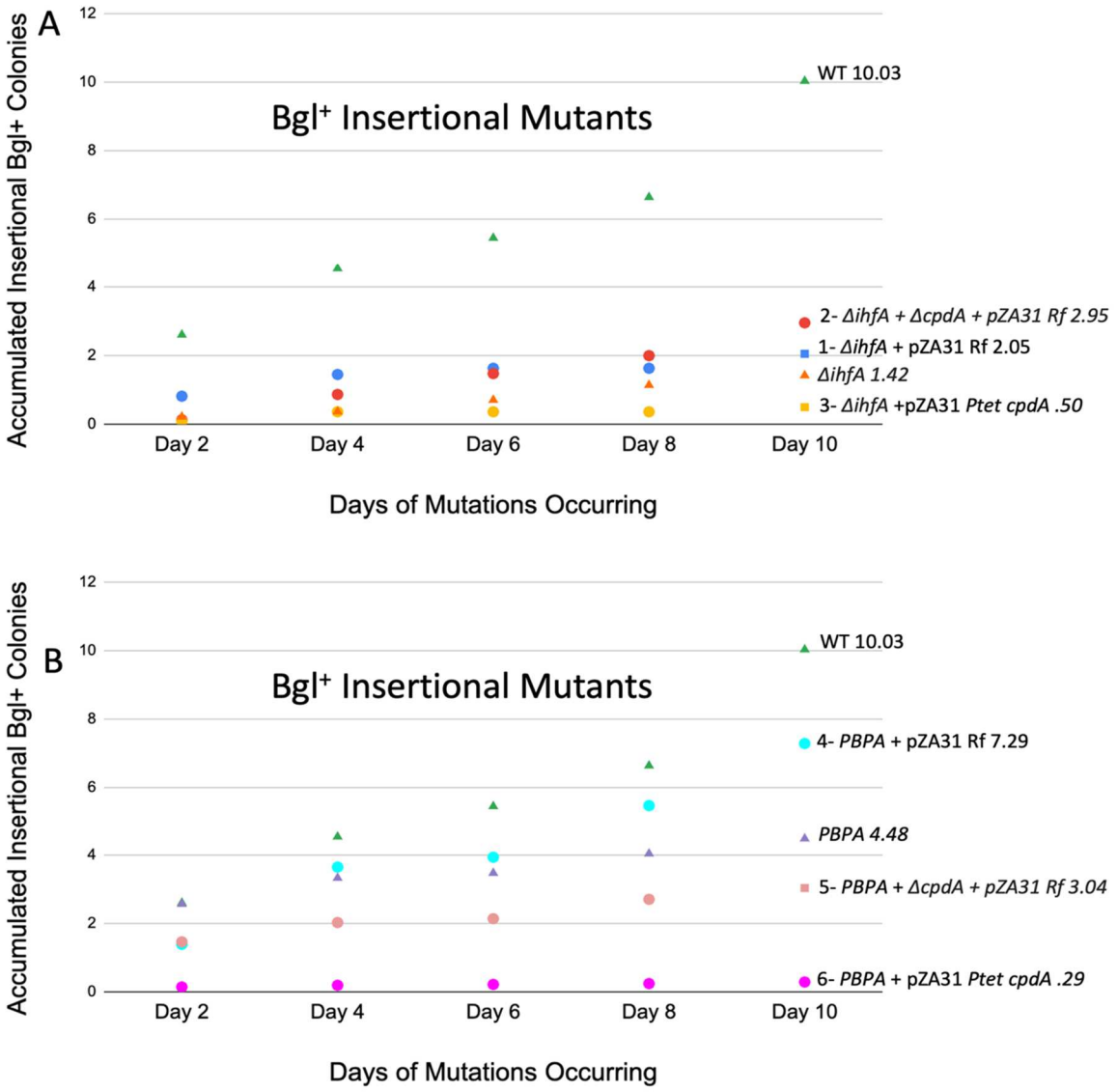


Figure 9- Effects of *ihfA* deletion/P*tet*-driven expression with *cpdA* deletion/ Overexpression on the IS insertion frequency upstream of *bgl.*

Mutation assays were performed on solid M9 salicin media over a 12-day period as previously described. Colony PCR was used to differentiate insertional and noninsertional Bgl$^+$ mutants. **A)** shows the insertional Bgl$^+$ mutations of 3 Δ*ihfA* strains combined with all 3 of the *cpdA* modulation mutations shown in Figure 8C. These mutations were combined with the three background mutants used for the *cpdA* mutation assays. **B)** shows the insertional Bgl$^+$ mutations of 3 *PBPA* strains combined with all 3 of the *cpdA* modulation mutations described in Figure 8C.

DISCUSSION

3.1- The effect of PBPA, causing a decrease in the IS insertional rate

P*tet*-driven expression of the *two* IHF subunits yielded a 2-fold decrease in the frequency of Bgl^+ mutants compared to the wild-type strain. Similarly, the double mutant P*tet cpdA* + *PBPA* displayed fewer Bgl^+ mutant colonies than expected, although a decrease relative to *PBPA* was hypothesized due to the influence of *cpdA.* There are at least two possibilities as to why this may have occurred. The first possibility is that the P*tet* promoter may not be as strong as the *ihfA/B* native promoters. As a global protein with several functions in the cell and with genes in locations adjacent to other important genes, promoters for both subunits may be very strong [48]. With a lower subunit availability compared to WT, our data may suggest the effects of lower gene expression rather than more as we had anticipated. Our data on *PBPA-RI* as titrated by varying levels of ATC inducer supports this suggestion, as even at maximal induction levels, the effect is the same.

Another possibility is that we are in fact expressing IHF at levels higher than WT, but that this has an overall negative effect on mutation frequency. For example, these increased levels of expression may be toxic to the cell, or the increased binding of IHF to the multitude of its recognized binding sites may have pleiotropic effects that lead to lower growth rates, and consequently less insertion.

A final possibility is related to the fact that IHF levels increase greatly when the cell is starving [27, 28]. It may be that this increase is regulated by transcription factors affecting the native *ihfA/ ihfB* promoters under starvation conditions such as

during the mutation assay. If we change the promoters to P*tet*, we may well have inadvertently removed the cell's normal ability to respond to starvation by upregulating *ihfA* and *ifhB* expression, which would also bring IHF levels to a lower-than-expected value.

In either case, another potential way that overexpression of both IHF subunits could be attained would be to express *ihfA* and *ihfB* on a plasmid with a high copy number, as for the *cpdA* experiments described in this work. IHF would be abundant regardless of the promoter used to drive the subunits, allowing for more confidence concerning the amount of IHF in the cell relative to that in the WT cell.

3.2 Effect of cyaA deletion on the IS1/5 insertion rate

cAMP-activated Crp, which is known to bind in *bgl*'s promoter region and positively regulate transcription, was shown to also positively regulate the rate of IS element transposition upstream of *bgl*. Results of a mutation assay comparing WT with an adenylate cyclase deletion mutant (Δ*cyaA*) showed a decrease in the incidence of insertional colonies, although the overall Bgl^+ colony counts remained comparable. Similarly, in backgrounds overexpressing or deleting the cAMP-degrading phosphodiesterase gene, *cpdA*, the rate of insertion decreased or increased, respectively. Thus, these results reveal a new role for Crp within the context of *bgl* as a regulator of transposition as well as transcription. Interestingly, previous data from our group have suggested that in the *glpFK* operon, cAMP-Crp has a negative effect on upstream IS insertion. It appears that while Crp maintains its status as a positive transcriptional regulator in the vast majority of operons which it

influences, its role with respect to mutational changes in the *bgl* operon may be responsive to the sequences and/or DNA structures surrounding its binding site.

Crp has a known binding site in *bgl*'s promoter region, and this site is close to or overlapping the projected SIDD site into which IS elements insert [23]. This novel effect of cAMP-activated Crp presented here likely results from the increased energetic favorability for insertion to take place due to its binding. However, advances in computational programs that analyze energetics would be helpful to confirm this, as the current programs do not account for the binding of proteins to the DNA.

3.4 Strength of the IS1 promoter region

The IS1-*lacZ* reporter activity proved to be low, even when compared to the two orientations of the IS5 reporter. After a thorough examination of the sequence of the construct, no mutation was found, leading us to consider other possibilities as to why IS1 exerted such a low operon transcriptional activity compared to IS5. We submit that either the growth conditions were not sufficient to increase IS1's promoter to higher levels, or the region upstream of the IS1-*lacZ* gene construct is lacking a site for a postulated positive transcription factor for the IS1 promoter. If so, it must be present in the other native locations in the genome where IS1 exists. Both considerations may independently or together explain why IS1's reporter showed such low activity.

3.5 Effect of IHF via DNA conformation of the IS element to allow for efficient transposase binding

As stated in Results section 2.7, both ends of IS1 and at least one end of IS5 contain IHF binding sites [45-47]. However, LacZ assays measuring the activities of the IS1 and IS5 promoters showed no clear change when the level of IHF binding was altered. Despite this, our other data indicate a non-specific effect of IHF on transposition of IS1 and IS5. Thus, the effect exerted by IHF on these IS elements must be of a nature other than transcriptional. Both IS1 and IS5 have DDE catalytic regions [49, 50], and may therefore act similarly for the purposes of replicative excision from the *E. coli* genome. Our IHF mutation assay data supports this; if IHF only affected one element's transposition and not the other, then only a 2-fold change at best would be expected, and all the insertional colonies would contain the same element, instead of the roughly equal levels of IS1 and IS5 insertion that has been classically observed (Schnetz & Rak, 1991; P.K., unpublished data). Work on other transposons with DDE activity has suggested that IHF plays an important, but not essential, role in the efficient transposition of Tn10 and the Mu bacteriophage [51-53]. The role of IHF involves assisting in the conformation of the transposon via binding so that the transposase protein can effectively bind to the transposon ends and then dimerize (or tetramerize in the case of Mu). It has been shown that in the absence of IHF, DNA supercoiling has a compensatory effect, allowing for transposition to occur, even when an important conformational factor is not present [54]. Interestingly, Tn10 *in vitro* experiments mimicking *in vivo* supercoiling levels observed a ~10-fold decrease in transposition frequency (called the 'basal level' of

excision) when IHF is absent, which is what we observed in the *ihfA* deletion mutant (Δ*ihfA*) for the Bgl^+ mutation assay [54] (Figure 1A). The conclusions of this paper utilize *in vivo* models to support past findings in relation to other transposons and draw further similarities between IS elements and other transposons in the DDE transposon family.

MATERIALS AND METHODS

4.1 List of Strains and Oligos

(k) indicates resistance to kanamycin; (c) indicates resistance to chloramphenicol

Table 1- List of strains used for mutation assays.

A. Mutation Assays
1. *BW25113* aka WT- *Wild-type for all experiments related to this work* [55]
2. ΔstpA - deletion of stpA (k) [8]
3. ΔhupA - deletion of hupA (k) (this work)
4. ΔhupB - deletion of hupB (k) (this work)
5. ΔihfA - deletion of ihfA (k) (this work)
6. ΔbglG - deletion of bglG (k) [8]
7. ΔbglJ - deletion of bglJ (k) [8]
8. ΔleuO - deletion of leuO (k) [8]
9. ΔrcsB - deletion of rcsB (k) [8]
10. Δfis - deletion of fis (k) [8]
11. PBPA - P*tet*-driven expression of *ihfB and ihfA (k) (this work)*
12. PBPA-RI - as A11, with the tetR gene (k) (this work)
13. P*tet-bglG-* P*tet-driven bglG at the intS locus, bglG overexpression (k); the native bgl operon is fully present* [12]
14. $lacI_q$ bglG- $lacI_q$ -driven bglG; increased bglG expression, but not as much as P*tet bglG (k); native bgl operon is fully present* [8]
15. P*tet-bglG ΔihfA- as A13, with ΔihfA (this work)*
16. $lacI_q$ -bglG ΔihfA- as A14, with ΔihfA (this work)
17. Δcrp- deletion of crp (k) [3]
18. ΔihfA Δcrp- deletion of ihfA and crp (k) (c) (this work)
19. P*tet ihfA G62E@intS-* P*tet-driven ihfA containing a point mutation changing the glycine at the protein's 62nd position to a glutamate, removing ihfA's ability to bind to DNA but not its ability to dimerize with ihfB; this construct is at the intS locus, leaving the native ihfA gene intact (k) (this work)*
20. P*tet ihfA G62E@intS ΔihfA- as A17, but the native ihfA gene has now been removed (k) (c) (this work)*
21. WT Rf- WT, but with pZA31 granting chloramphenicol resistance; has a P*tet promoter transcribing a small, noncoding transcript or random fragment, hence Rf (GET MORE INFO FROM Z AKA CONFIRM THIS IS CORRECT) (c) (this work)*
22. ΔcpdA Rf- cpdA deletion, with the same pZA31 Rf as described in A19 (c) (this work)
23. P*tet cpdA- WT, but with pZA31 granting chloramphenicol resistance and* P*tet-driven cpdA; cpdA overexpression combined with pZA31's copy number (~15/cell,* 9092630*) (c) (this work)*
24. ΔihfA Rf aka Strain 1; *ihfA deletion with pZA31 granting chloramphenicol resistance;* P*tet-driven random fragment (k) (c) (this work)*
25. ΔcpdA ΔihfA Rf aka Strain 2; *ihfA and cpdA deletion; pZA31 granting chloramphenicol resistance;* P*tet-driven random fragment (k) (c) (this work)*
26. ΔihfA P*tet cpdA* aka Strain 3; *ihfA deletion with pZA31 granting chloramphenicol resistance and* P*tet-driven cpdA (k) (c) (this work)*
27. PBPA Rf aka Strain 4; P*tet-driven ihfA and ihfA; pZA31 granting chloramphenicol resistance;* P*tet-driven random fragment (k) (c) (this work)*
28. PBPA ΔcpdA Rf aka Strain 5; *as strain A26, but with deletion of cpdA (k) (c) (this work)*
29. PBPA P*tet cpdA* aka Strain 6; *as strain 25, but with pZA31 granting chloramphenicol resistance and* P*tet-driven cpdA (k) (c) (this work)*

Table 2- List of strains used for transcriptional activity measurements and other experiments.

A. LacZ Reporter Strains
1. Pbgl-Z [8]
2. Pbgl-G-Z [8]
3. Pbgl-Z ∆ihfA- as B1, with ihfA deletion (this work)
4. Pbgl-G-Z ∆ihfA- as B2, with ihfA deletion (this work)
5. IS1-Z- the native lacZ promoter is replaced with the promoter region contained within IS1, including the Inverted Repeat (IR) where IHF binds (k) (this work)
6. IS1-Z ∆ihfA- as B5, with ihfA deletion (k) (c) (this work)
7. IS1-Z P*tet ihfA @intS- as B5, with an additional* P*tet-driven ihfA gene at the intS locus (k) (this work)*
8. IS1-Z P*tet ihfA G62E@intS- as B7, but the* P*tet-driven ihfA gene contains the G62E mutation (k) (this work)*
9. IS5A-Z- the native lacZ promoter is replaced with the promoter region contained within IS5; specifically, the promoter region and IR of the classic orientation of IS5A, with which it usually inserts into bgl and other operons (IS THIS RIGHT? SOURCE?); this promoter region does not have IHF binding (k) (this work)
10. IS5A-Z ∆ihfA- as B9, with ihfA deletion (k) (c) (this work)
11. IS5A-Z P*tet ihfA @intS- as B9, with an additional* P*tet-driven ihfA gene at the intS locus (k) (this work)*
12. IS5A-Z P*tet ihfA G62E @intS- as B11, but the* P*tet-driven ihfA gene contains the G62E mutation (k) (this work)*
13. IS5B-Z- the native lacZ promoter is replaced with the promoter region contained within IS5, including the IR in the reverse orientation as IS5A-Z; this region does contain an IHF binding site (k) (this work)
14. IS5B-Z ∆ihfA- as B13, with ihfA deletion (k) (c) (this work)
15. IS5B-Z P*tet ihfA @intS- as B13, with an additional* P*tet-driven ihfA gene at the intS locus (k) (this work)*
16. IS5B-Z P*tet ihfA G62E @intS- as B15, but the* P*tet-driven ihfA gene contains the G62E mutation (k) (this work)*
B. Strains used for BglB Assays
1. *WT- A1* [55]
2. *∆bglGFB- the entire bgl operon has been deleted downstream of bgl's native promoter* [7]
3. *IS+- as WT, with an activating insertion of IS1? IS5? upstream of bgl's promoter* [8]
4. *∆ihfA- A5, ihfA deletion (k)* (this work)
5. *IS+ ∆ihfA- as C3, with ihfA deletion (k) (this work)*
6. *PBPA- A11;* P*tet-driven expression of ihfB and ihfA (k) (this work)*
7. *∆ihfA G50- ihfA deletion, with bglG's terminators removed and the bglG gene truncated (k) (c) (this work); (bglG50: [7])*
8. *PBPA G50- as A11, but with the G50 mutation as described in C7 (k) (this work) (this work)*
9. *∆TGT- native bgl operon with both terminators and the entire bglG gene removed (k)* [7]
10. *∆TGT ∆ihfA- as C9, with ihfA deletion (k) (c) (this work)*
11. *∆TGT ∆fis- as C9, with fis deletion (k) (this work)*
12. *∆TGT ∆fis ∆ihfA- as c9 with ihfA and fis deletion (k) (c) (this work)*
13. *PBPA ∆TGT- as A11, with bglG and both terminators deleted from the native bgl promoter (k) (this work)*
C. Other Experiments
1. IS+- *C3; as WT, with an activating insertion of IS1? IS5? upstream of bgl's promoter* [8]
2. IS+ *∆ihfA- C5; as C3, with ihfA deletion (k) (this work)*

Table 3- Oligonucleotides used in this study.

Name	Sequence	Use
cpdA-Kpn-F	ataggtaccatggaaagcctgttaacccttcctctggctg	Cloning *cpdA* into pZA31P*tet*
cpdA-Bam-R	aatggatcctcagtagccttctgaagcggtatcaggttg	Cloning *cpdA* into pZA31P*tet*
cpdA-ver-F	aagcgcgtgtttattggtgagcaatg	Verification of *cpdA* cloning
cpdA-ver-R	actgcggaatgatcatctcaacgtcag	Verification of *cpdA* cloning
Ptet-ihfA-P1	aatgtgtagaggcattaaaagagcgattccaggcatcattg agggattgaatgtgtaggctggagctgcttc	Construction of P*tet* driving *ihfA* at the *ihfA* locus
Ptet-ihfA-P2	cccaagcttatcaaacagatattctgacatttcagcttttgtaa gcgccatggtacctttctcctctttaatgaattc	Construction of P*tet* driving *ihfA* at the *ihfA* locus
ihfA-ver-R	tcgtctttgggcgaagcgttttcgac	Verification of P*tet-ihfA* at the *ihfA* locus
Ptet-ihfB-P1	gtttcgtcctgtaatcaagcactaagggcggctacggccgcc cttaatcaatgtgtaggctggagctgcttc	Construction of P*tet* driving *ihfB* at the *ihfB* locus
Ptet-ihfB-P2	gaatgtgcgattgctgggtggcaagtctttctatcaattctgact tggtcatggtacctttctcctctttaatgaattc	Construction of P*tet* driving *ihfB* at the *ihfB* locus
ihfB-ver-R	tccagttctactttatcgccagtcttc	Verification of P*tet-ihfB* at the *ihfB* locus
PtetM2-Xho-F	atactcgagactctatcattgatagagtttg	Cloning P*tet-ihfA* and P*tet-ihfA*.G62E into pKDT
ihfA-Bam-R	ataggatcctggtatccgttctgctgaagtgtcatg	Cloning P*tet-ihfA* and P*tet-ihfA*.G62E into pKDT
ihfA-R	cgttttcgggttacgttccgggcgttgattcttatc	Construction of *ihfA*G62E
ihfA-F	gaatcaacgcccggaacgtaacccgaaaacg	Construction of *ihfA*G62E
intS-P1	agatttacagttcgtcatggttcgcttcagatcgttgacagccg cactccatgtgtaggctggagctgcttc	Construction of P*tet-ihfA and* P*tet-ihfA*G62E at the *intS* locus
ihfA2-P2	agttgttaaggtcgctcactccaccttctcatcaagccagtcc gcccaccattactcgtctttgggcgaagcgttttc	Construction of P*tet-ihfA and* P*tet-ihfA*G62E at the *intS* locus
intS-ver-R2	aaaggaatgaagtctatcatccaagtc	Verification of P*tet-ihfA and* P*tet-ihfA*G62E at the *intS* locus
PIS1-Xh-F	ttactcgaggtgatgctgccaacttactgatttagtgtatg	Cloning IS1 transposase promoter into pKDT
PIS1-Bm-R	ataggatccatccaacgccattcatggccatatc	Cloning IS1 transposase promoter into pKDT
PIS5a-Xh-F	atactcgaggaaggtgcgaacaagtccctgatatgag	Cloning *Ins5A* promoter into pKDT
PIS5a-Bm-R	aatggatccaatgctgcatgcagtgaatgcgtagc	Cloning *Ins5A* promoter into pKDT
PIS5b-Xh-F	atactcgaggaaggtgcgaataagcggggaaattcttc	Cloning *Ins5CB* operon promoter into pKDT
PIS5b-Bm-R	attggatccatttcgcatcatcaagcgacagttc	Cloning *Ins5CB* operon promoter into pKDT
PIS-Z-P1	agccgcgctttcttcttcgcgacgttcgttaacaacaacttcta atttttcatcaaagggaaaactgtccatatgc	Construction of IS1 and IS5 promoter *lacZ* reporter at the *lac* locus
PIS1-Z-P2	gtaaaacgacggccagtgaatccgtaatcatggtcatagct gtttcctgtgtgttaggagggacagctgatagaaacagaag	Construction of IS1 transposase promoter *lacZ* reporter at the *lac* locus
PIS5a-Z-P2	gtaaaacgacggccagtgaatccgtaatcatggtcatagct gtttcctgtgtgttattcactgtcggcgaaggtaagttgatg	Construction of *Ins5A* promoter *lacZ* reporter at the *lac* locus
PIS5b-Z-P2	gtaaaacgacggccagtgaatccgtaatcatggtcatagct gtttcctgtgtgttacgataaccaactggcgatgttattcac	Construction of *Ins5CB* operon promoter *lacZ* reporter at the *lac* locus

4.2 Construction of Strains Used

Construction of deletion mutants

CGSC strains JW3964-1, JW0430-3, JW1702-1 and JW3000-1 (*E. coli* Genetic Stock Center, Yale Univ.) carry the deletion mutations of *hupA, hupB, ihfA* and *cpdA*, respectively. For each of these mutants, a kanamycin resistance (km^r) gene was substituted for the target gene. These mutations were individually transferred to strain BW25113 (wild-type/WT) [55] by P1 transduction, and the km^r gene was subsequently flipped out by pCP20 [55], yielding deletion mutant strains Δ*hupA,* Δ*hupB,* Δ*ihfA* and Δ*cpdA*, respectively (Table S1). Using P1 transduction, the *ihfA* mutation was transferred into the *crp* deletion strain Δ*crp* [3], yielding the Δ*ihfA*Δ*crp* double mutant. The same *ihfA* mutation was transferred into strain Δ*cpdA*, yielding the Δ*cpdA* Δ*ihfA* double mutant.

Construction of cpdA overexpression plasmid

The *cpdA* gene was amplified from BW25113 chromosomal DNA using oligos cpdA-Kpn-F and cpdA-Bam-R (Table S2). The PCR products were gel purified, digested with *KpnI*/*BamH*I and then ligated into the same sites of pZA31P*tet* [56], yielding pZA31P*tet-cpdA*. pZA31P*tet-cpdA* was transformed into the WT strain, yielding P*tet cpdA* (Table S1, A21). Plasmid pZA31P*tet*-RF [57] carries a random fragment (Rf) and was used as a negative control.

Construction of P*tet driving ihfA, ihfB, and ihfAG62E on the chromosome*

Using plasmid pKDT:P*tet* [58] as a template, the cassette *km*^{r}:*rrnB*T:P*tet*, containing the *km*^{r} gene, the *rrnB* terminator (*rrnB*T) and the P*tet* promoter, was amplified using the primer pair Ptet-ihfA-P1 and Ptet-ihfA-P2 (Table S2). Using lambda-red system [55], the PCR products were integrated into the chromosome of BW25113 to replace the "CCT" nucleotides immediately upstream of the *ihfA* translational start point. Chromosomal integration was confirmed, first by colony PCR, and subsequently by DNA sequencing. This yielded strain P*tet ihfA*, in which the strong *tet* promoter drives expression of the *ihfA* gene. Similarly, the *km*^{r}:*rrnB*T:P*tet* cassette amplified by Ptet-ihfB-P1 and Ptet-ihfB-P2 (Table S2) from pKDT:P*tet* was substituted for the *ihfB*'s upstream promoter region (-46 to -1 relative to the *ihfB*'s translational start point), yielding strain P*tet ihfB*.

To make an *ihfA*/*ihfB* double overexpression strain, the *km*^{r} gene was first flipped out from strain P*tet_ihfB* by pCP20. The cassette of *km*^{r}:*rrnB*T:P*tet* driving *ihfA* gene from strain P*tet_ihfA* was then transferred to Km-sensitive P*tet_ihfB* by P1-transduction, yielding strain P*tet_ihfB Ptet_ihfA (PBPA,* Table S1 A11*)*, in which both *ihfA* and *ihfB* are simultaneously driven by P*tet*. To titrate expression of *ihfA* and *ihfB*, the transcription unit including a constitutively expressed *tetR* gene and a spectinomycin resistance (sp^r) marker was transferred to *PBPA* by P1-transduction as mentioned above, yielding strain *PBPA RI.*

To further increase *ihfA* expression, a second copy of P*tet* driving *ihfA* was inserted to another chromosomal location. To do so, the *km*^{r}:*rrnB*T:P*tet-ihfA*

expression cassette was amplified from the gnomic DNA of strain P*tet_ihfA* using primers intS1-P1 and ihfA2-P2. The products were integrated into the *intS* site to replace the region between -229 and +1101 (relative to the *intS* translational initiation site). The chromosomal integration was confirmed by colony PCR and subsequently by DNA sequencing, yielding strain P*tet-ihfA* @*intS*.

The Glycine residue at position 62 is required for IhfA to bind to DNA [33, 34]. To reduce or abolish IhfA's DNA-binding ability, this residue was changed to a glutamate residue using fusion PCR. The first part (5' region) of *ihfA* was amplified from BW25113 gnomic DNA using ihfA-Kpn-F and ihfA.G62E-R (carrying the G62E alteration). The second part (3' region) was amplified using ihfA.G62E-F (carrying the G62E alteration and overlapping ihfA.G62E-R) and ihfA-Bam-R. Both products were gel purified and fused together using primers ihfA-Kpn-F and ihfA-Bam-R. The fused products (that is, *ihfA*.G62E) were gel purified, digested with *Kpn*I and *BamH*I, and then ligated into the same sites of pKDT_P*tet*, yielding pKDT_P*tet-ihfA*.G62E. The cassette *km*^{r}:*rrnB*T:P*tet-ihfA*.G62E was amplified using intS-P1 and ihfA-P2, gel purified and then integrated into the *intS* site as for P*tet-ihfA* @*intS*. The chromosomal integration was confirmed by colony PCR and subsequently by DNA sequencing, yielding strain P*tet-ihfA* G62E, which still maintains the native *ihfA* but constitutively expresses the modified *ihfA*.G62E at the *intS* locus. This cassette was transferred to an *ihfA* deletion background by P1 transduction, yielding Δ*ihfA* P*tet-ihfA* G62E.

Construction of the IS1 and IS5 promoter lacZ reporters

The promoter region (-55 to + 30 relative to the *insA* translation intiation site) driving the IS1 transposase gene *insAB1* was amplified using IS1p-Xho-F and IS1p-Bam-R from BW25113 genomic DNA. This region contains all upstream region (including the left-hand IR) and the first 10 residues of *insA* (plus a stop codon TAA). The amplified products were digested with *Xho*I and *BamH*I and cloned into the same sites of the defaut integration vector, pKDT [58], yielding pKDT-P_{IS1}. The region carrying the *km*^{r}*, rrnB*T and P_{IS1} (*km*^{r}*:rrnBT*:P_{IS1}) was PCR amplified using oligos IS1-Z-P1 and IS1-Z-P2 (Table S2) and then integrated into the chromosomal default strain EQ42 [58] to replace the *lacI* gene and the *lacZ* promoter. The resultant strain carries the *km*^{r}*:rrnB*T:P_{IS1} cassette followed by *lacZ's* ribosomal binding site (RBS) and the *lacZ* structural gene within the *lac* locus. After being confirmed by PCR and sequencing, the promoter reporter, P_{IS1}, driving *lacZ* expression (that is, P_{IS1}*-lacZ*) was transferred into BW25113 and various genetic backgrounds by P1 transduction. This yielded the IS1 promoter reporter strains BW_P_{IS1}*-Z, ΔihfA*_P_{IS1}*-Z,* P*tet_ihfA*_P_{IS1}*-Z*, and P*tet_ihfA G62E*_P_{IS1}*-Z,* respectively.

IS5 carries three open reading frames (*ins5A*, *ins5B* and *ins5C*). *ins5A* encodes the main transposase and it is transcribed from its own promoter located close to the left-hand IR while the divergent *ins5B* and *ins5C* genes form an that is driven by another promoter located close to the right-hand IR [59]. The *ins5A* promoter region (-68 to +30 relative to *ins5A* translational intiation site) and the *ins5CB* promoter region (-207 to +30 relative to ins5C translational initiation site) were individually cloned into pKDT, yielding pKDT_P_{IS5A} and P_{IS5CB}, respectively. As

for P_{IS1} described above, P_{IS5A} and P_{IS5CB} each carries the first 10 amino acids from the target gene followed by a stop codon TAA. These promoter were integrated into the *lac* site on BW25113 chromosome as for P_{IS1}-*Z*, yielding strains BW_P_{IS5A}-*Z* and BW_P_{IS5B}-*Z*, respectively. Using P1 transductions, these new IS5 promoter *lacZ* reporters were transferred to BW25113 and various genetic backgrounds by P1 transduction. This yielded the IS5A promoter reporter strains BW_P_{IS5A}-*Z, ΔihfA*_P_{IS5A}-*Z,* P*tet_ihfA*_P_{IS5A}-*Z*, and P*tet_ihfA G62E*_P_{IS5A}-*Z,* as well as the IS5B promoter reporter strains BW_P_{IS5B}-*Z, ΔihfA*_P_{IS5B}-*Z,* P*tet_ihfA*_P_{IS5B}-*Z*, and P*tet_ihfA G62E*_P_{IS5B}-*Z*.

4. 3 β-glucosidase assays

E. coli reporter strains were cultured in 4 mL of LB contained in glass test tubes (1.5 cm in diameter × 15 cm in length) with shaking at 37 °C for 8 h. An amount of 30 μL of LB cultures were used to inoculate 3 mL of M63 minimal media in smaller glass tubes (1.2 cm × 12 cm), and the tubes were shaken at 37 °C overnight. The carbon sources were 0.5% glycerol, 0.5% salicin, or both. The tubes were rotated at 250 rpm and 37 °C, and cell densities (OD600) were measured with a Bio-Rad spectrophotometer. During the exponential growth phase, four samples were collected in the range of OD600 from 0.1 to 1. The samples (roughly 0.3 mL for promoter reporter strains (0.9 for the IS1/5 promoters), and 0.6 mL for operon reporter strains were immediately frozen at −20 °C prior to β-galactosidase assays.

To measure β-galactosidase activities in *bgl* promoter reporter strains, 0.8 mL of Z-buffer containing β-mercaptoethanol (2.7 μL/mL) and sodium dodecyl sulfate (SDS)

(0.005%) was mixed with 0.2 mL of sample and 25 µL of CHCl3 in test tubes. Alternatively, for *bgl* operon reporter strains, 0.5 mL of Z-buffer was mixed with 0.5 mL of the sample. The tubes were vortexed twice (each time for 10 s at a constant speed) and incubated in a 37 °C water bath until temperature equilibration. A 0.2 mL aliquot of O-nitrophenyl galactoside (ONPG) substrate (4 mg/mL) was then added to each test tube. When a yellow color developed, the reaction was stopped by adding 0.5 mL of 1 M Na_2CO_3 followed by vortexing. Reaction mixtures were centrifuged (15,000 rpm, 3 min), and the absorbance values of the supernatants were measured at 420 nm and 550 nm. A control tube was run in parallel using M63 salts instead of the test sample. β-galactosidase activity (Miller units) = [(OD420−1.75 × OD550)/(sample volume in mL × time in min)] × 1000 [60]. For a given test strain, the slope of OD600 values versus β-galactosidase activities was referred to as the promoter activity or the operon activity.

4.4 Mutation Assays

4.4A. Bgl^+ Mutation Assays

Bgl^+ mutation assays were performed on minimal M9 agar plates with 0.5% of a β-glucoside (salicin) as the sole carbon source. Strains to be tested (from single fresh colonies) were cultured in LB liquid medium for approximately 7 h at 37 °C, washed twice with carbon source-free M9 salts (M9) and applied onto plates (2×10^7 cells/plate). The plates were then incubated in a 30 °C incubator and were examined every two days for the appearance of Bgl^+ colonies, with each colony representing a new Bgl^+ mutation. On these β-glucoside minimal agar plates, any colonies

appearing by day 2 were considered to be from Bgl$^+$ cells initially applied onto the plates. They were therefore subtracted from the subsequent measurements.

To determine the effects of other carbon sources on Bgl$^+$ mutations, mutation assays were performed on minimal M9 agar plates with 0.25% glycerol or 1% propanediol as the sole carbon source. To determine the total populations, the cells were washed off the minimal M9 agar plates at relevant time points, serially diluted, and plated onto LB agar plates. To determine Bgl$^+$ populations, appropriate dilutions were applied on M9 + salicin agar plates. The frequencies of Bgl$^+$ mutations were determined as described above for Bgl$^+$ on M9 + salicin agar plates.

4.4B. Glp$^+$ Mutation Assays

Glp$^+$ mutation assays were conducted on glycerol M9 minimal agar plates as described in Zhang and Saier [3]. Strains Δ*crp* and Δ*crp* Δ*ihfA* were used for the mutation assays.

4.4C. Swarming Mutation Assays

Using the wild-type and Δ*ihfA* strains, the swarming mutation assays for the appearance of hyper-swarming mutants (outgrowing subpopulations from the inoculated cells) were carried out using the method of Barker et al., 2004 [61]. Briefly, overnight cell cultures in LB media were washed once with M9 salts and diluted to an OD600 of 1.0prior to use. 2 µL of the cell suspensions were streaked across the centers of LB semisolid (0.3% agar) plates (diameter = 9 cm) using a plastic transfer loop. The plates were incubated at 30 °C. The swarming mutants, represented by

outgrowths of motile subpopulations from the streaked cells, were counted. The mutation frequency was normalized as outgrowths (mutations) per 9-cm cell streak. Insertional mutants were verified via colony PCR.

4.4D. Propanediol (PPD) Growth Mutation Assay

The assay for PPD^+ mutations was conducted by applying cell suspensions from fresh overnight cultures onto propanediol (1%) ± salicin (0.5%) M9 minimal agar plates (~10^8 cells/plate). The wild-type and P*tet*-*G* strains were used for M9 + salicin plate assays while Δ*bglB* and Δ*bglB* P*tet*-*G* were used for M9 + PPD + salicin plate assays. Growth positive mutations and total populations were determined as described above under "Bgl^+ mutation assays".

The information contained in this thesis, in full, is being prepared for submission for publication. Kopkowski, Peter; Zhang, Zhongge; Saier, Milton H., 2022. The thesis author was the primary investigator and author of this paper.

REFERENCES

1. Mc, CB. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci U S A*. **1950**. 36(6):344-55. DOI: 10.1073/pnas.36.6.344. PMID: 15430309.

2. Zhang, Z and Saier, MH, Jr. A novel mechanism of transposon-mediated gene activation. *PLoS Genet*. **2009**. 5(10):e1000689. DOI: 10.1371/journal.pgen.1000689. PMID: 19834539.

3. Zhang, Z and Saier, MH, Jr. A mechanism of transposon-mediated directed mutation. *Mol Microbiol*. **2009**. 74(1):29-43. DOI: 10.1111/j.1365-2958.2009.06831.x. PMID: 19682247.

4. Whiteway, J. Koziarz, P. Veall, J. Sandhu, N. Kumar, P. Hoecher, B. and Lambert, IB. Oxygen-insensitive nitroreductases: analysis of the roles of nfsA and nfsB in development of resistance to 5-nitrofuran derivatives in *Escherichia coli. J Bacteriol*. **1998**. 180(21):5529-39. DOI: 10.1128/JB.180.21.5529-5539.1998. PMID: 9791100.

5. Dole, S. Nagarajavel, V. and Schnetz, K. The histone-like nucleoid structuring protein H-NS represses the *Escherichia coli bgl* operon downstream of the promoter. *Mol Microbiol*. **2004**. 52(2):589-600. DOI: 10.1111/j.1365-2958.2004.04001.x. PMID: 15066043.

6. Schnetz, K. Silencing of *Escherichia coli bgl* promoter by flanking sequence elements. *EMBO J*. **1995**. 14(11):2545-50. DOI: 10.1002/j.1460-2075.1995.tb07252.x. PMID: 7781607.

7. Lam, KJK. Zhang, Z. and Saier, MH. Histone-like Nucleoid Structuring (H-NS) Protein Silences the beta-glucoside (*bgl*) Utilization Operon in *Escherichia coli* by Forming a DNA Loop. **2022**. (under review).

8. Tran, D. Zhang, Z. Lam, KJK. and Saier, MH, Jr. Effects of Global and Specific DNA-Binding Proteins on Transcriptional Regulation of the *E. coli bgl* Operon. *Int J Mol Sci*. **2022**. 23(18). DOI: 10.3390/ijms231810343. PMID: 36142257.

9. Schnetz, K and Rak, B. Regulation of the *bgl* operon of *Escherichia coli* by transcriptional antitermination. *EMBO J*. **1988**. 7(10):3271-7. DOI: 10.1002/j.1460-2075.1988.tb03194.x. PMID: 2846278.

10. Mahadevan, S and Wright, A. A bacterial gene involved in transcription antitermination: regulation at a rho-independent terminator in the *bgl* operon of *E. coli. Cell*. **1987**. 50(3):485-94. DOI: 10.1016/0092-8674(87)90502-2. PMID: 3301003.

11. Amster-Choder, O and Wright, A. Modulation of the dimerization of a transcriptional antiterminator protein by phosphorylation. *Science*. **1992**. 257(5075):1395-8. DOI: 10.1126/science.1382312. PMID: 1382312.

12. Zhang, Z. Zhou, K. Tran, D. and Saier, M. Insertion Sequence (IS) Element-Mediated Activating Mutations of the Cryptic Aromatic beta-Glucoside Utilization (BglGFB) Operon Are Promoted by the Anti-Terminator Protein (BglG) in *Escherichia coli. Int J Mol Sci*. **2022**. 23(3). DOI: 10.3390/ijms23031505. PMID: 35163427.

13. Fox, CF and Wilson, G. The role of a phosphoenolpyruvate-dependent kinase system in beta-glucoside catabolism in *Escherichia coli. Proc Natl Acad Sci U S A*. **1968**. 59(3):988-95. DOI: 10.1073/pnas.59.3.988. PMID: 4870648.

14. Chen, Q. Arents, JC. Bader, R. Postma, PW. and 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. *EMBO J*. **1997**. 16(15):4617-27. DOI: 10.1093/emboj/16.15.4617. PMID: 9303306.

15. Prasad, I. Young, B. and Schaefler, S. Genetic determination of the constitutive biosynthesis of phospho-glucosidase A in *Escherichia coli* K-12. *J Bacteriol*. **1973**. 114(3):909-15. DOI: 10.1128/jb.114.3.909-915.1973. PMID: 4576407.

16. Gorke, B and Rak, B. Catabolite control of *Escherichia coli* regulatory protein BglG activity by antagonistically acting phosphorylations. *EMBO J*. **1999**. 18(12):3370-9. DOI: 10.1093/emboj/18.12.3370. PMID: 10369677.

17. Rothe, FM. Bahr, T. Stulke, J. Rak, B. and Gorke, B. Activation of *Escherichia coli* antiterminator BglG requires its phosphorylation. *Proc Natl Acad Sci U S A*. **2012**. 109(39):15906-11. DOI: 10.1073/pnas.1210443109. PMID: 22984181.

18. Reynolds, AE. Felton, J. and Wright, A. Insertion of DNA activates the cryptic *bgl* operon in *E. coli* K12. *Nature*. **1981**. 293(5834):625-9. DOI: 10.1038/293625a0. PMID: 6270569.

19. Singh, J. Mukerji, M. and Mahadevan, S. Transcriptional activation of the *Escherichia coli bgl* operon: negative regulation by DNA structural elements near the promoter. *Mol Microbiol*. **1995**. 17(6):1085-92. DOI: 10.1111/j.1365-2958.1995.mmi_17061085.x. PMID: 8594328.

20. Lopilato, J and Wright, A, *Mechanisms of activation of the cryptic bgl operon of Escherichia coli K-12*, in *The Bacterial Chromosome*, K Drilca and M Riley, Editors. 1990, American Society for Microbiology: Washington DC. p. 435-444.

21. Prasad, I and Schaefler, S. Regulation of the beta-glucoside system in *Escherchia coli* K-12. *J Bacteriol*. **1974**. 120(2):638-50. DOI: 10.1128/jb.120.2.638-650.1974. PMID: 4616943.

22. Hall, BG. Activation of the *bgl* operon by adaptive mutation. *Mol Biol Evol*. **1998**. 15(1):1-5. DOI: 10.1093/oxfordjournals.molbev.a025842. PMID: 9491599.

23. Humayun, MZ. Zhang, Z. Butcher, AM. Moshayedi, A. and Saier, MH, Jr. Hopping into a hot seat: Role of DNA structural features on IS5-mediated gene activation and inactivation under stress. *PLoS One*. **2017**. 12(6):e0180156. DOI: 10.1371/journal.pone.0180156. PMID: 28666002.

24. Goosen, N and van de Putte, P. The regulation of transcription initiation by integration host factor. *Mol Microbiol*. **1995**. 16(1):1-7. DOI: 10.1111/j.1365-2958.1995.tb02386.x. PMID: 7651128.

25. Engelhorn, M and Geiselmann, J. Maximal transcriptional activation by the IHF protein of *Escherichia coli* depends on optimal DNA bending by the activator. *Mol Microbiol*. **1998**. 30(2):431-41. DOI: 10.1046/j.1365-2958.1998.01078.x. PMID: 9791186.

26. Sugimura, S and Crothers, DM. Stepwise binding and bending of DNA by *Escherichia coli* integration host factor. *Proc Natl Acad Sci U S A*. **2006**. 103(49):18510-4. DOI: 10.1073/pnas.0608337103. PMID: 17116862.

27. Bushman, W. Thompson, JF. Vargas, L. and Landy, A. Control of directionality in lambda site specific recombination. *Science*. **1985**. 230(4728):906-11. DOI: 10.1126/science.2932798. PMID: 2932798.

28. Ditto, MD. Roberts, D. and Weisberg, RA. Growth phase variation of integration host factor level in *Escherichia coli. J Bacteriol*. **1994**. 176(12):3738-48. DOI: 10.1128/jb.176.12.3738-3748.1994. PMID: 8206852.

29. Santos, I. Martin de Dios, R. Barrios, V. Barcia, F. Pey, J. Ojeda, JL. and Sanchez, E. [Anomalous origin of the right coronary artery from the left sinus of Valsalva. Apropos of 2 cases]. *Rev Esp Cardiol*. **1991**. 44(9):618-21. PMID: 1775707.

30. Surette, MG. Lavoie, BD. and Chaconas, G. Action at a distance in Mu DNA transposition: an enhancer-like element is the site of action of supercoiling relief activity by integration host factor (IHF). *EMBO J*. **1989**. 8(11):3483-9. DOI: 10.1002/j.1460-2075.1989.tb08513.x. PMID: 2555166.

31. Allison, RG and Chaconas, G. Role of the A protein-binding sites in the in vitro transposition of mu DNA. A complex circuit of interactions involving the mu ends and the transpositional enhancer. *J Biol Chem*. **1992**. 267(28):19963-70. PMID: 1328189.

32. van Ulsen, P. Hillebrand, M. Zulianello, L. van de Putte, P. and Goosen, N. Integration host factor alleviates the H-NS-mediated repression of the early promoter of bacteriophage Mu. *Mol Microbiol*. **1996**. 21(3):567-78. DOI: 10.1111/j.1365-2958.1996.tb02565.x. PMID: 9082117.

33. Granston, AE and Nash, HA. Characterization of a set of integration host factor mutants deficient for DNA binding. *J Mol Biol*. **1993**. 234(1):45-59. DOI: 10.1006/jmbi.1993.1562. PMID: 8230206.

34. Hales, LM. Gumport, RI. and Gardner, JF. Mutants of *Escherichia coli* integration host factor: DNA-binding and recombination properties. *Biochimie*. **1994**. 76(10-11):1030-40. DOI: 10.1016/0300-9084(94)90027-2. PMID: 7748924.

35. Kolb, A. Busby, S. Buc, H. Garges, S. and Adhya, S. Transcriptional regulation by cAMP and its receptor protein. *Annu Rev Biochem*. **1993**. 62:749-95. DOI: 10.1146/annurev.bi.62.070193.003533. PMID: 8394684.

36. Fic, E. Bonarek, P. Gorecki, A. Kedracka-Krok, S. Mikolajczak, J. Polit, A. Tworzydlo, M. Dziedzicka-Wasylewska, M. and Wasylewski, Z. cAMP receptor protein from *Escherichia coli* as a model of signal transduction in proteins--a review. *J Mol Microbiol Biotechnol*. **2009**. 17(1):1-11. DOI: 10.1159/000178014. PMID: 19033675.

37. Gorke, B and Stulke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol*. **2008**. 6(8):613-24. DOI: 10.1038/nrmicro1932. PMID: 18628769.

38. Deutscher, J. The mechanisms of carbon catabolite repression in bacteria. *Curr Opin Microbiol*. **2008**. 11(2):87-93. DOI: 10.1016/j.mib.2008.02.007. PMID: 18359269.

39. Gulati, A and 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 Cells*. **2000**. 5(4):239-50. DOI: 10.1046/j.1365-2443.2000.00322.x. PMID: 10792463.

40. Mukerji, M and Mahadevan, S. Characterization of the negative elements involved in silencing the *bgl* operon of *Escherichia coli*: possible roles for DNA

gyrase, H-NS, and CRP-cAMP in regulation. *Mol Microbiol*. **1997**. 24(3):617-27. DOI: 10.1046/j.1365-2958.1997.3621725.x. PMID: 9179854.

41. Chen, YM. Lu, Z. and Lin, EC. Constitutive activation of the *fucAO* operon and silencing of the divergently transcribed *fucPIK* operon by an IS5 element in *Escherichia coli* mutants selected for growth on L-1,2-propanediol. *J Bacteriol*. **1989**. 171(11):6097-105. DOI: 10.1128/jb.171.11.6097-6105.1989. PMID: 2553671.

42. Zhang, Z. Yen, MR. and Saier, MH, Jr. Precise excision of IS5 from the intergenic region between the *fucPIK* and the *fucAO* operons and mutational control of *fucPIK* operon expression in Escherichia coli. *J Bacteriol*. **2010**. 192(7):2013-9. DOI: 10.1128/JB.01085-09. PMID: 20097855.

43. Zhang, Z. Kukita, C. Humayun, MZ. and Saier, MH. Environment-directed activation of the *Escherichia coliflhDC* operon by transposons. *Microbiology (Reading)*. **2017**. 163(4):554-569. DOI: 10.1099/mic.0.000426. PMID: 28100305.

44. McCalla, DR. Kaiser, C. and Green, MH. Genetics of nitrofurazone resistance in *Escherichia coli. J Bacteriol*. **1978**. 133(1):10-6. DOI: 10.1128/jb.133.1.10-16.1978. PMID: 338576.

45. Prentki, P. Chandler, M. and Galas, DJ. *Escherichia coli* integration host factor bends the DNA at the ends of IS1 and in an insertion hotspot with multiple IHF binding sites. *EMBO J*. **1987**. 6(8):2479-87. DOI: 10.1002/j.1460-2075.1987.tb02529.x. PMID: 2822395.

46. Gamas, P. Chandler, MG. Prentki, P. and Galas, DJ. *Escherichia coli integration* host factor binds specifically to the ends of the insertion sequence IS1 and to its major insertion hot-spot in pBR322. *J Mol Biol*. **1987**. 195(2):261-72. DOI: 10.1016/0022-2836(87)90648-6. PMID: 2821273.

47. Muramatsu, S. Kato, M. Kohara, Y. and Mizuno, T. Insertion sequence IS5 contains a sharply curved DNA structure at its terminus. *Mol Gen Genet*. **1988**. 214(3):433-8. DOI: 10.1007/BF00330477. PMID: 2851094.

48. Pozdeev, G. Beckett, MC. Mogre, A. Thomson, NR. and Dorman, CJ. Reciprocally rewiring and repositioning the Integration Host Factor (IHF) subunit genes in *Salmonella enterica* serovar Typhimurium: impacts on physiology and virulence. *Microb Genom*. **2022**. 8(2). DOI: 10.1099/mgen.0.000768. PMID: 35166652.

49. Ohta, S. Tsuchida, K. Choi, S. Sekine, Y. Shiga, Y. and Ohtsubo, E. Presence of a characteristic D-D-E motif in IS1 transposase. *J Bacteriol*. **2002**. 184(22):6146-54. DOI: 10.1128/JB.184.22.6146-6154.2002. PMID: 12399484.

50. Curcio, MJ and Derbyshire, KM. The outs and ins of transposition: from mu to kangaroo. *Nat Rev Mol Cell Biol*. **2003**. 4(11):865-77. DOI: 10.1038/nrm1241. PMID: 14682279.

51. Swinger, KK and Rice, PA. IHF and HU: flexible architects of bent DNA. *Curr Opin Struct Biol*. **2004**. 14(1):28-35. DOI: 10.1016/j.sbi.2003.12.003. PMID: 15102446.

52. Kobryn, K. Watson, MA. Allison, RG. and Chaconas, G. The Mu three-site synapse: a strained assembly platform in which delivery of the L1 transposase binding site triggers catalytic commitment. *Mol Cell*. **2002**. 10(3):659-69. DOI: 10.1016/s1097-2765(02)00596-8. PMID: 12408832.

53. Gueguen, E. Rousseau, P. Duval-Valentin, G. and Chandler, M. The transpososome: control of transposition at the level of catalysis. *Trends Microbiol*. **2005**. 13(11):543-9. DOI: 10.1016/j.tim.2005.09.002. PMID: 16181782.

54. Chalmers, R. Guhathakurta, A. Benjamin, H. and Kleckner, N. IHF modulation of Tn10 transposition: sensory transduction of supercoiling status via a proposed protein/DNA molecular spring. *Cell*. **1998**. 93(5):897-908. DOI: 10.1016/s0092-8674(00)81449-x. PMID: 9630232.

55. Datsenko, KA and Wanner, BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*. **2000**. 97(12):6640-5. DOI: 10.1073/pnas.120163297. PMID: 10829079.

56. Lutz, R and Bujard, H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res*. **1997**. 25(6):1203-10. DOI: 10.1093/nar/25.6.1203. PMID: 9092630.

57. Levine, E. Zhang, Z. Kuhlman, T. and Hwa, T. Quantitative characteristics of gene regulation by small RNA. *PLoS Biol*. **2007**. 5(9):e229. DOI: 10.1371/journal.pbio.0050229. PMID: 17713988.

58. Klumpp, S. Zhang, Z. and Hwa, T. Growth rate-dependent global effects on gene expression in bacteria. *Cell*. **2009**. 139(7):1366-75. DOI: 10.1016/j.cell.2009.12.001. PMID: 20064380.

59. Sawers, RG. Transcript analysis of *Escherichia coli* K-12 insertion element IS5*. FEMS Microbiol Lett*. **2005**. 244(2):397-401. DOI: 10.1016/j.femsle.2005.02.019. PMID: 15766797.

60. Miller, F. Glycopeptides of human immunoglobulins. 3. The use and preparation of specific glycosidases*. Immunochemistry*. **1972**. 9(3):217-28. PMID: 4338321.

61. Barker, CS. Pruss, BM. and Matsumura, P. Increased motility of *Escherichia coli* by insertion sequence element integration into the regulatory region of the *flhD* operon*. J Bacteriol*. **2004**. 186(22):7529-37. DOI: 10.1128/JB.186.22.7529-7537.2004. PMID: 15516564.