UNIVERSITY OF CALIFORNIA SAN DIEGO

Histone-like Nucleoid Structuring (H-NS) Protein Silences the beta-glucoside (*bgl*) Utilization Operon by Formation of a DNA Loop in *Escherichia coli*

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

in

Biology

by

Katie Jing Kay Lam

Committee in charge:

Professor Milton H. Saier, Jr., Chair
Professor James W. Golden
Professor Katherine L. Petrie

2022

The thesis of Katie Jing Kay Lam 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 TABLES

LIST OF FIGURES

ABSTRACT OF THE THESIS

Histone-like Nucleoid Structuring (H-NS) Protein Silences the beta-glucoside (*bgl)* Utilization Operon by Formation of a DNA Loop in *Escherichia coli*

by

Katie Jing Kay Lam

Master of Science in Biology

University of California San Diego, 2022

Professor Milton H. Saier, Jr., Chair

The β-glucoside (*bgl*) utilization operon in *Escherichia coli* encodes proteins necessary for uptake and hydrolysis of aromatic β-glucosides. The *bgl* operon is silent in wild-type cells due to strong repression by the histone-like nucleoid structuring (H-NS) protein, which binds to the upstream *bgl* regulatory region and within the *bglG* gene. In addition to its DNA-binding domain, H-NS contains a dimerization/oligomerization domain, allowing formation of oligomeric nucleoprotein complexes which are important for gene silencing. In this study, we demonstrate that silencing of the *bgl* operon is due to the self-oligomerization of the dimeric H-NS protein, which brings the two H-NS binding sites within the *bgl* operon together, leading to

formation of a DNA loop that largely blocks the operon transcription. Our results show that the *bgl* operon can be activated, or partially activated, by preventing H-NS self-oligomerization, and by reducing H-NS binding to either of its two binding sites within the operon. H-NS binding alone with no looping is not capable of efficiently repressing the operon, and other regulators are not required for operon repression. This indicates that formation of the H-NS-mediated DNA loop is necessary and sufficient to silence the *bgl* operon expression. In addition, we have found that formation of the DNA loop in the *bgl* operon is flexible and is independent of the distance between and the DNA phasing of the two H-NS binding sites in the DNA helix. Overall, our results revealed that H-NS-mediated DNA looping is robust and strongly represses transcription of the *bgl* operon.

1. Introduction

Due to limited resources available in the natural environment, bacteria are sometimes under environmental stress and in competition with other microorganisms for survival. In order to better adapt to resource-limited environments and increase their fitness, one of the mechanisms evolved in bacteria is the activation of cryptic genes to utilize substrates available in their habitats. Cryptic genes are not expressed in the nutrient-sufficient environment but can be fully functional once activated by mutations like transposition of Insertion Sequence (IS) Elements. A known example of cryptic genes is the β-glucoside (*bgl*) utilization operon in *Escherichia coli* (*E. coli*) (Hall *et al*., 1983).

1.1 The *bgl* operon

The *bgl* operon in *E. coli* encodes necessary proteins for the uptake and utilization of aromatic β-glucosides such as salicin and arbutin (Schnetz and Rak, 1988). The operon is cryptic and silent in wild-type *E. coli* strains such that cells are not able to grow using these β-glucosides as sole carbon sources. The silence of the *bgl* operon may provide survival benefits to the cells, as some toxic β-glucosides such as cyanogenic glucosides are present in the environment (Gleadow & Møller, 2014). In spite of that, the cryptic *bgl* operon can be activated by mutations, which then become inducible in the presence of β-glucosides, so that *E. coli* cells can utilize β-glucosides for growth, leading to the Bgl positive (Bgl^+) phenotype (Schaefler, 1967). One of the most common types of mutations giving rise to the Bgl^+ phenotype is the transposition of IS elements like IS1 and IS5 into the *bgl* regulatory region (Reynolds *et al*., 1986). It has been shown that a superhelix stress-induced DNA duplex destabilization (SIDD) region is located within the *bgl* regulatory region, where IS insertions occur frequently (Humayun *et al*., 2017). In

addition, the operon can be activated by non-insertional mutations in the *bgl* regulatory region or outside the *bgl* operon, which blocks the silencing of the operon as well (Dole *et al*., 2004; Giel *et al*., 1996; Reynolds *et al*., 1986; Madhusudan *et al*., 2005; Ueguchi *et al*., 1998; Venkatesh *et al*., 2010).

The *bgl* operon includes three genes, *bglG*, *bglF* and *bglB*, with two rho-independent terminators flanking *bglG* (Schnetz and Rak, 1988). Under normal conditions, transcription is terminated at the two terminators due to the formation of hairpin-like secondary structures. However, in the presence of β-glucosides as inducers, the product of the first gene *bglG* acts as a transcriptional antiterminator which binds to an upstream region that partially overlaps with each terminator. This prevents the formation of the hairpin-like structure of the two terminators which relieves the termination and allows transcription of the operon (Schnetz and Rak, 1988). The second gene, *bglF*, encodes enzyme II (EII) of the phosphoenolpyruvate:sugar phosphotransferase system (PTS), which is responsible for the transport and phosphorylation of β-glucosides (Chen *et al*., 1997). The BglF protein can also phosphorylate and dephosphorylate BglG, and it acts as a negative regulator of BglG antitermination (Schnetz and Rak, 1988). In the presence of β-glucosides, BglG is dephosphorylated by EII and phosphorylated at a different site by a phosphocarrier protein (HPr) of the PTS, which leads to full activity of BglG antitermination (Görke and Rak, 1999). The third gene, *bglB*, encodes a phospho-β-glucosidase which hydrolyzes phospho-β-glucosides that are transported and phosphorylated by BglF for cell metabolism (Schnetz *et al*., 1987). Even though the two terminators terminate transcription and BglG antitermination is inhibited by BglF under normal conditions, it is still unclear if the terminators contribute to and play a significant role in the silencing of the *bgl* operon.

1.2 Regulation of the *bgl* operon

There are several regulators reported to be involved in the upregulation of the *bgl* operon, such as the cyclic-AMP receptor protein (Crp), LeuO and BglJ (Reynolds *et al*., 1986). Crp is a global regulator which once bound to cyclic-AMP (cAMP), upregulates the *bgl* operon by binding to the regulatory region (Madan *et al*., 2005). A single mutation in the Crp binding site created more effective Crp-cAMP binding with the *bgl* regulatory region which leads to the transcriptional activation of the operon (Reynolds *et al*., 1986). In addition, LeuO and BglJ are transcriptional regulators which also bind to the *bgl* regulatory region and increased the expression of the *bgl* operon (Giel *et al*., 1996; Madhusudan *et al*., 2005; Ueguchi *et al*., 1998). BglJ forms heterodimers with another transcriptional regulatory protein RcsB, which binds to the *bgl* regulatory region and activates the operon (Venkatesh *et al*., 2010). Both LeuO and BglJ relieve the repression of the *bgl* operon through the *bgl* regulatory region but have no effect on the region downstream of the promoter (Madhusudan *et al*., 2005).

As mentioned above, the *bgl* operon is silent in wild-type *E. coli* cells. A major repressor of the *bgl* operon is histone-like nucleoid structuring protein (H-NS), which is involved in chromosome organization and acts as a global negative regulator (Wang *et al*., 2011). It has been found that H-NS preferentially binds to AT-rich regions and to curved DNA (Lang *et al*., 2007). H-NS proteins consists of a dimerization/oligomerization domain at the N-terminus and a DNA-binding domain at the C-terminus (Shindo *et al*., 1995; Ueguchi *et al*., 1997), which can form higher order oligomers of up to 20-mers (Smyth *et al*., 2000). H-NS-mediated nucleoprotein complexes are considered to be important in gene silencing. This protein acts by bridging DNA duplexes through H-NS oligomerization, thus allowing the formation of DNA loops (Dame *et al*., 2006). The formation of such nucleoprotein complexes would prevent transcription by either

(1) trapping RNA polymerase, or (2) blocking RNA polymerase from binding to the promoter region (Dame *et al*., 2002). However, the changing of leucine to proline at the 30th codon (L30P) within the N-terminal domain causes H-NS to lose its ability to form higher order oligomers while retaining its ability to form dimers (Ueguchi *et al*., 1997; Stella *et al*., 2005).

In relation to the *bgl* operon, H-NS binds close to the Crp-cAMP binding site within the regulatory region, as well as a downstream region within the *bglG* gene (Dole *et al*., 2004; Reynolds *et al*., 1986). However, the exact H-NS binding sites in the *bgl* operon are still unknown. In addition, it is still unclear what the exact mechanism by which H-NS silences the *bgl* operon is, although it is known that the L30P mutation in H-NS allows de-repression of the operon (Ueguchi *et al*., 1997). Also, when H-NS loses the ability to bind DNA, another repressor of the *bgl* operon, StpA can function as a DNA-binding adapter of H-NS to repress the operon (Free *et al*., 1998). However, StpA itself is a weak repressor of the *bgl* operon and is negatively regulated by H-NS (Wolf *et al*., 2006). Therefore, H-NS is still the major repressor of the *bgl* operon in wild-type *E. coli* cells.

1.3 Purpose of the study

As proposed in previous studies, H-NS binds to DNA and forms bridges that regulate gene expression. It is likely that the same mechanism applies to the repression of the *bgl* operon by H-NS. In addition to H-NS binding to the *bgl* regulatory region and within the *bglG* gene, the oligomerization of H-NS allows formation of a bridge which brings the two H-NS binding sites together. A DNA loop is then formed, so that the transcription of the *bgl* operon is blocked, and the operon is silent in wild-type *E. coli* cells (**Figure 1.1**). In this study, I provide evidence that H-NS mediated looping is required for the full silencing of the *bgl* operon. This proposed mechanism of *bgl* operon repression by formation of an H-NS loop was investigated in-depth.

I examined the effect of preventing the formation of the H-NS-mediated DNA loop on the *bgl* operon activity, as well as the ability of *E. coli* cells to grow on β-glucosides. In order to do so, the formation of the DNA loop was altered using the following approaches: (1) preventing H-NS self-oligomerization (**Figure 1.2A**), (2) preventing or reducing H-NS binding to the *bgl* regulatory region (**Figure 1.2B**), (3) preventing or reducing H-NS binding to *bglG* (**Figure 1.2C**, (4) changing the distance between the two H-NS binding sites (**Figure 1.2D**), and (5) changing the phase of the two H-NS binding sites (**Figure 1.2E**). First, to prevent H-NS oligomerization, the mutant with L30P mutation in H-NS was used to abolish the H-NS oligomerization ability while retaining its DNA-binding ability as mentioned above (Ueguchi *et al*., 1997; Tran *et al*., manuscript in preparation). Second, to prevent or reduce H-NS binding to the *bgl* regulatory region, three different methods were used: (1) overproducing positive regulators of the *bgl* operon, including LeuO, BglJ and RcsB, (2) creating a stronger Crp-cAMP binding site within the *bgl* regulatory region (Reynolds *et al*., 1986), (3) inserting IS elements into the *bgl* regulatory region, and (4) removing the presumptive H-NS binding site in the *bgl* regulatory region. Third,

to prevent H-NS binding to *bglG*, (1) the H-NS binding site was replaced by TetR operators, (2) the H-NS binding site was removed using a truncated *bglG* gene, and (3) the entire *bglG* gene and the two *bgl* terminators were removed. Next, the distance between the two H-NS binding sites at the *bgl* operon was changed by insertions within *bglG*. Lastly, the phasing of the H-NS binding sites was changed by either a 5- or 10-base pair insertion within *bglG*. I hoped to provide a new perspective on the mechanism of how the cryptic *bgl* operon is repressed by H-NS in wild-type *E. coli* cells.

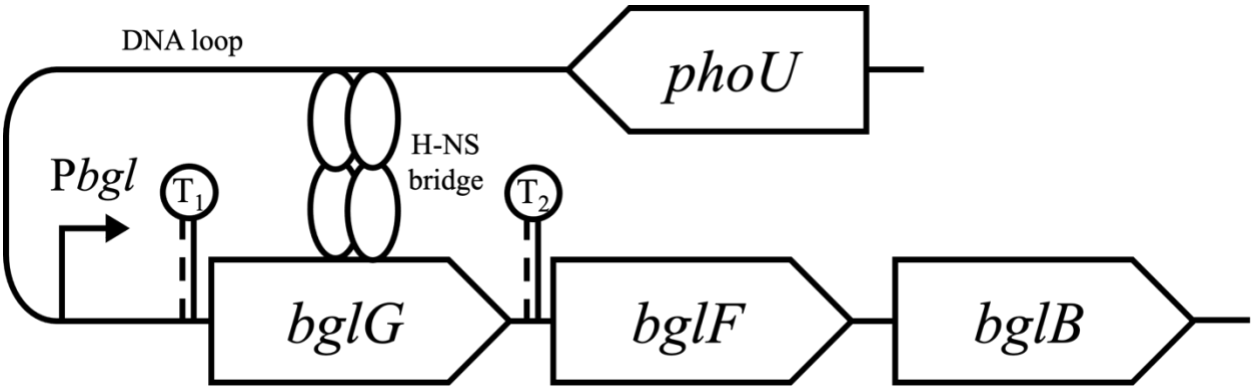


Figure 1.1 Proposed mechanism of *bgl* operon repression by H-NS. H-NS binds to the *bgl* regulatory region and within the *bglG* gene. Due to its ability of oligomerization, a H-NS bridge together with a DNA loop are formed, which prevents transcription of the *bgl* operon.

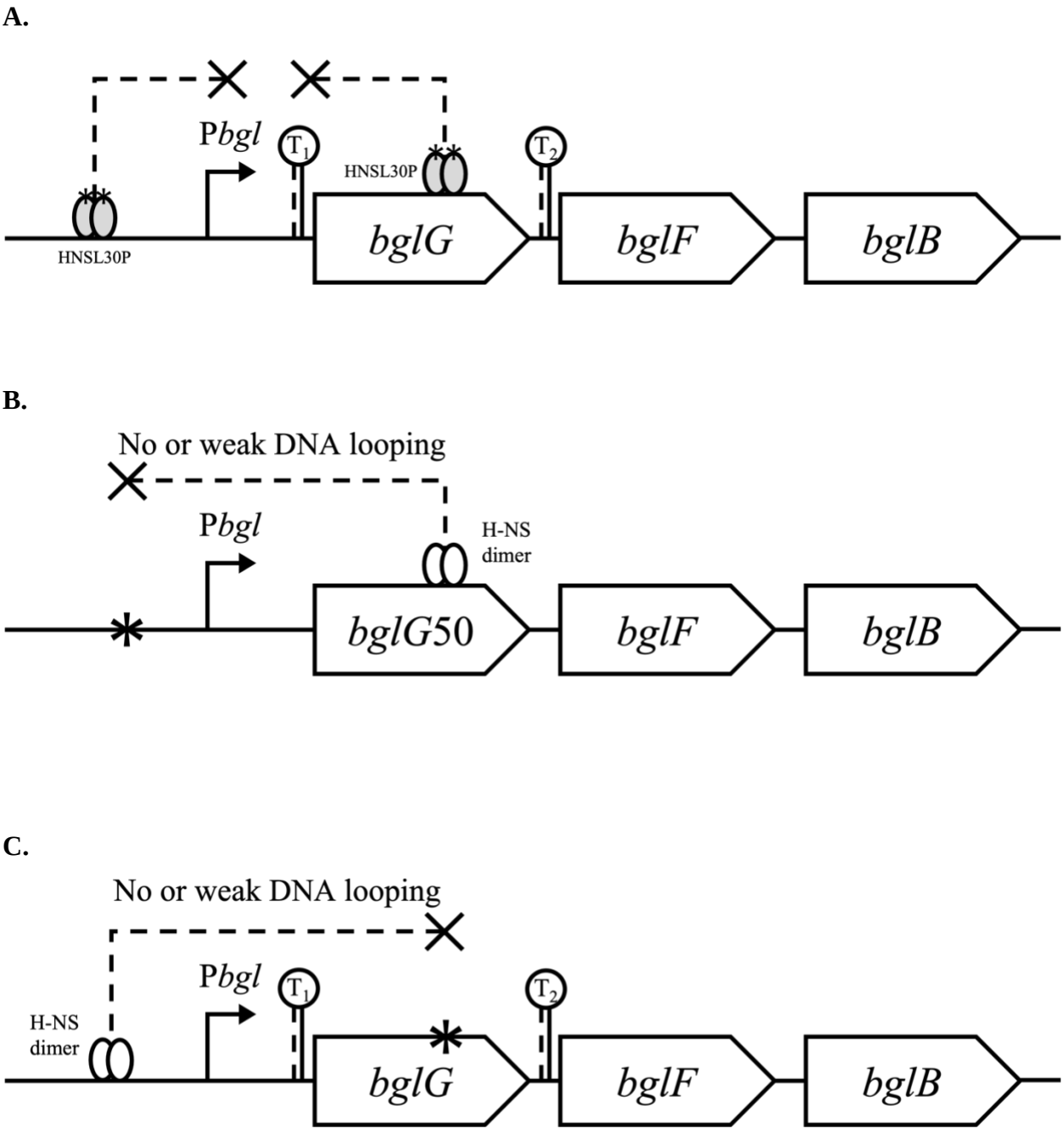
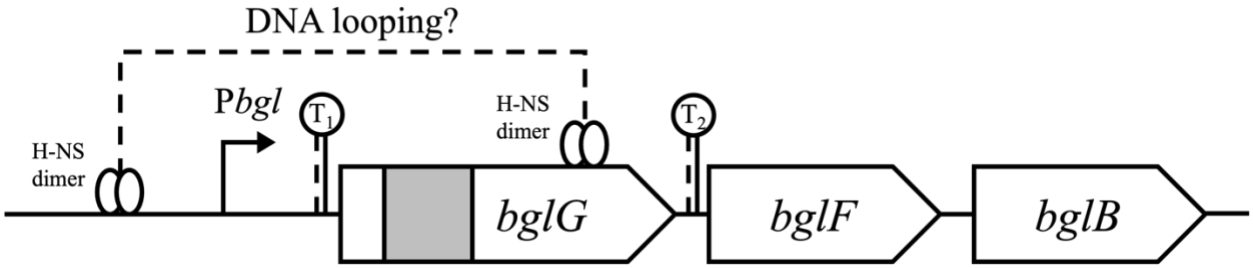


Figure 1.2 Different approaches to alter the formation of the DNA loop. (**A**) Preventing H-NS self-oligomerization. (**B**) Preventing or reducing H-NS binding to the *bgl* regulatory region. (**C**) Preventing or reducing H-NS binding within the *bglG* gene. (**D**) Changing distance between two H-NS binding sites. (**E**) Changing the phase of the H-NS binding sites.

D.



E.

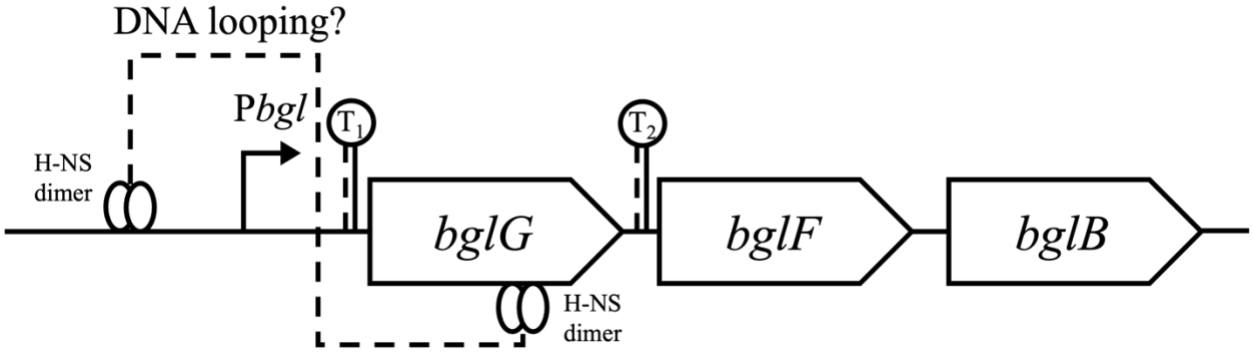


Figure 1.2 (continued) Different approaches to alter the formation of the DNA loop. (**A**) Preventing H-NS self-oligomerization. (**B**) Preventing or reducing H-NS binding to the *bgl* regulatory region. (**C**) Preventing or reducing H-NS binding within the *bglG* gene. (**D**) Changing distance between two H-NS binding sites. (**E**) Changing the phase of the H-NS binding sites.

2. Materials & Methods

2.1 Strains & operon reporters

2.1.1 Strains

Wild-type strain

E. coli K12 strain BW25113 was used as the wild-type in this study. All other strains are derived from this strain. BW-RI is another wild-type strain used in this study, which constitutively produced the TetR protein, the repressor of the *tet* promoter (Levine *et al*., 2007).

Construction of mutant strains Δ*hns*, Δ*stpA*, ΔP*bglG* and ΔP*bglGFB*

The *hns* deletion in strain JW1225-2 and the *stpA* deletion in strain JW2644-3 (*E. coli* Genetic Stock Center, Yale Univ.), in which a kanamycin resistance (km^r) gene was substituted for the target gene, were transferred into BW25113 by P1-transduction (Thomason *et al*., 2007). The km^r gene that replaces the *hns* gene or the *stpA* gene was individually flipped out by transformation of pCP20 (Datsenko & Wanner, 2000), yielding strains Δ*hns* and Δ*stpA*, respectively. The *stpA* deletion was subsequently transferred into strain Δ*hns* by P1 transduction, yielding the double mutant, Δ*hns*Δ*stpA*.

To remove the *bgl* promoter and the *bglG* gene, the region containing the *bgl* promoter and the *bglG* gene plus the flanking terminators was replaced by a *cat* gene amplified from pKD3 (Datsenko & Wanner, 2000) by oligos Pbglcat-P1 and bglFcat-P2 (**Table 2.1**). The *cat* gene was flipped out by pCP20, yielding strain ΔP*bglG*, in which both P*bgl* and *bglG* plus both terminators were replaced by an 85-bp FRT scar. This strain was then used as a beginning strain to make several other strains.

To make strain ΔP*bglGFB* that is deleted for the entire *bglGFB* operon and its regulatory region, the *km*^{r} gene was amplified from pKD4 (Datsenko & Wanner, 2000) using oligos PhoU1-P1 and bglB2-P2 (**Table 2.1**) and was then substituted for the region from -277 relative to the *bglG* translation initiation site to -6 relative to the stop codon of *bglB*. The *km*^{r} gene was deleted by pCP20, yielding strain ΔP*bglGFB*.

Removal of both terminators flanking the *bglG* gene

In the *bgl* operon, the first gene, *bglG*, is flanked by two terminators. To determine if these terminators are important in H-NS mediated *bgl* operon silencing, they were deleted on the chromosome. To do this, the strain ΔP*bglG*, in which the *bgl* promoter region, *bglG* gene and the two *bgl* terminators were replaced by an 85-base pair FRT scar, was used as the beginning strain.

Present in plasmid pKES50 (Dole *et al*., 2004), there is a *lacZ* gene (with its own Shine-Dalgarno sequence) fused to a region carrying the *bgl* promoter (P*bgl*) and the modified *bglG* gene, *bglG*50 (the first two ATG codons in *bglG* each changed to GCG) without both flanking terminators. The P*bgl-bglG*50 DNA region together with the downstream *lacZ* Shine-Dalgarno sequence (between *bglG*50 and *lacZ*) was amplified from the plasmid pKES50 using oligos Pbgl-Xh-Fbg and bglG-Bm-Rm (**Table 2.1**). The PCR products were gel purified, digested with *Xho*I and *BamH*I, and then ligated into the same sites of pKDT (Klumpp *et al*., 2009) with "T" referring to the *rrnB* terminator (*rrnB*T), yielding pKDT_P*bgl*-G50, in which the native *bgl* promoter driving a modified terminator-less *bglG* gene (i.e., *bglG*50). The cassette containing the *km*^{r} gene, the *rrnB*T and the P*bgl*-G50 (*km*^{r}:*rrnB*T:P*bgl*-G50) was amplified from pKDT_P*bgl*-G50 using oligos Pbgl.G50-P1 and bglF.G50-P2 (**Table 2.1**). The PCR products were gel purified and electroporated into ΔP*bglG* cells expressing the red proteins to replace the 85-bp

FRT scar. This yielded strain G50, in which the native *bgl* promoter drives a terminator-less *bglG* (i.e., *bglG50*) followed by *bglF* and *bglB*. The intergenic region (aatagcttcacaggaaacagct) between *bglG* and *bglF* is the same as one between *bglG* and *lacZ* in pKES50 (Dole *et al*., 2004), and it carries the *lacZ* ribosome binding site (RBS) for *bglF*.

The strain G50 was used as the recipient strain in P1-transduction in the making of other new strains. First, the transcription unit including *tetR* and a spectinomycin resistance (sp^r) marker was transferred to G50 by P1-transduction as mentioned above, yielding strain RIG50. Also, *crp* replaced by the km^r gene was transferred to G50 by P1-transduction, yielding Δ*crp*G50 (Zhang & Saier, 2009).

To make strain G51, plasmid pKES51 (Dole *et al*., 2004) was used as the template to construct plasmid pKDT_P*bgl*-G51, which is essentially the same as pKDT_P*bgl*-G50 except that the presence of a stronger Crp-cAMP operator (O^*_{Crp}) in P*bgl* resulted from a C to T substitution at position -134 relative to the *bglG* translation initiation site (Reynolds *et al*., 1986). Using similar methods as described above for G50, the cassette of P*bgl*_O^*_{Crp} driving *bglG* was integrated upstream of *bglF* in ΔP*bglG* cells, yielding strain G51. This strain is the same as G50 except that it carries a stronger Crp-cAMP binding site at the *bgl* promoter region. A new strain Δ*crp*G51 (*crp* deletion in the G51 background) , was then made by transferring the km^r gene, replacing the *crp* gene to the strain G51 using P1-transduction (Zhang & Saier, 2009).

Construction of the HNSL30P strain using TetA-SacB positive/negative selection

H-NS is the major repressor of the *bgl* operon. H-NS proteins usually exist in oligomeric forms that contribute to its biological activity (Dame *et al*., 2002). The N-terminal domain is responsible for H-NS oligomerization, and the leucine residue at position 30 is essential for such

a process (Suzuki *et al*., 2016; Ueguchi *et al*., 1997). To test the possible looping mechanism by which H-NS silences the *bgl* operon, a two-step recombineering protocol based on TetA-SacB positive-selection and counter-selection (Li *et al*., 2013) was used to change the leucine codon CTG (+88 to +90 relative to the *hns* translation initiation site) to a proline codon CCT in the *hns* gene. The nucleotides "TG" (+89 to +90 relative to the translation start site) was first replaced by the *tetA-sacB* cassette that was amplified from the chromosomal DNA of strain T-SACK (Li *et al*., 2013) using chimeric oligos HNS-AB-F and HNS-AB-R (**Table 2.1**). The replacement of "TG" by *tetA-sacB* in several tetracycline (Tc) resistant mutants was confirmed by colony PCR and sequencing. A 100-bp single strand DNA fragment, which covers the region from +38 to +138 relative to the translation start site within the *hns* gene with "CT" replacing "TG" in the middle of the fragment, was synthesized. This fragment was amplified using oligos HNS-F and HNS-R (**Table 2.1**) and the PCR products were electroporated into the cells of a Tc resistant mutant expressing lambda-Red proteins encoded by pKD46 (Datsenko & Wanner, 2000). After one-hour incubation, the electroporated cells were applied to TetA/SacB counter-selection agar plates (containing 6% sucrose and 24 mg fusaric acid per liter). After incubation at 42^{o}C for about two days, about 10 colonies resistant to sucrose and fusaric acid were purified on LB agar plates and tested for both sensitivity to Tc and resistance to sucrose. Several Tc-sensitive/sucrose-resistant colonies were confirmed for the replacement of the tetA-*sacB* cassette by "CT" by PCR and DNA sequencing. The resultant altered strain was named HNSL30P, in which the 30th codon is changed from leucine to proline. The modified H-NS proteins can still bind to the DNA but have lost the oligomerization capability (Ueguchi *et al*., 1997).

Changing the H-NS binding site upstream of the *bgl* promoter

H-NS binds to two locations in the *bgl* operon, in the *bgl* regulatory region and within the *bglG* gene, even though the exact binding sites are unknown (Dole *et al*., 2004; Madhusudan *et al*., 2005). To abolish or reduce H-NS binding to the *bgl* regulatory region, the region (located between positions -207 to -585 relative to the *bglG* translational start site) carrying the possible H-NS binding site in G50, was first replaced by a km^r gene amplified from pKD4 using oligos Pbglcat-P1 and Pbgl2-P2 (**Table 2.1**). The km^r gene was flipped out by pCP20, yielding strain G50ΔO_{HNS}-P*bgl*.

IS insertional mutation is one of the approaches in abolishing or reducing H-NS binding to the *bgl* regulatory region. To obtain IS insertional mutants, wild-type BW25113 cells were grown in LB medium at 30°C for at least 6 hours. The cells were washed twice in 1x M9 salts without any carbon source, and around 10^8 cells were inoculated onto M9 minimal agar plates with 0.5% salicin. The plates were incubated at 30°C and were examined daily for any colony growth, indicating mutations leading to a Bgl^+ phenotype (Zhang *et al*., 2022). The colonies were picked for PCR using Pbgl-F2 and Pbgl-R2 (**Table 2.1**) to verify the presence of any insertional mutations at the *bgl* regulatory region. The type, location, and orientation of the IS elements inserted at the *bgl* regulatory region were identified by DNA sequencing.

The IS insertional mutant, IS5r(-207.5) with IS5 inserted in reverse orientation at -207.5 relative to the *bglG* transcriptional initiation site, was used as the positive control and the recipient strain for P1-transduction. The deletion of *stpA* and the deletion of *hns* were individually transferred to IS5r(-207.5) by P1-transduction, yielding strains IS5r(-207.5)Δ*stpA* and IS5r(-207.5)Δ*hns*.

Construction of P*tet* driving *bglJ, rcsB* and *leuO* on the chromosome

Using plasmid pKDT:Ptet (Klumpp *et al*., 2009) as 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-bglJ-P1 and Ptet-bglJ-P2 (**Table 2.1**). The PCR products were integrated into the chromosome of *E. coli* K12 strain BW25113 to replace the *yjjQ* promoter (from the -73 to the +1 relative to the translational start point of *yjjQ*). Chromosomal integration was confirmed first by colony PCR and subsequently by DNA sequencing. This yielded strain P*tet-bglJ*, in which the strong *tet* promoter drives the *yjjQ*/*bglJ* operon. Similarly, the "*km*^{r}:*rrnB*T:P*tet*" cassette amplified by Ptet-leuO-P1 and Ptet-leuO-P2 (**Table 2.1**) was substituted for the *leuO* promoter region (from -151 to -1 relative to the translational start point of *leuO*), yielding strain P*tet-leuO*. The same cassette amplified by Ptet-rcsB-P1 and Ptet-rcsB-P2 (**Table 2.1**) was substituted for the *rcsD*/*rcsB* intergenic region (from -12 to -1 relative to the translational start point of *rcsB*), yielding strain P*tet-rcsB*.

To make other strains in the Ptet-bglJ background, the *km*^{r} gene was first flipped out by pCP20, yielding Ptet-bglJ-km^S, which is sensitive to kanamycin. The *rcsB* gene, driven by a strong *tet* promoter with a *km*^{r} marker in P*tet-rcsB*, was transferred to Ptet-*bglJ*-km^S by P1-transduction, yielding the double overexpression strain P*tet-bglJ*/*rcsB*. The *rcsB* mutation was transferred to P*tet-bglJ*-km^S as well, yielding strain P*tet-bglJ*Δ*rcsB* with deletion of *rcsB* and overexpression of *bglJ*.

Similar approaches were used for making strains with overexpression of positive regulators of the *bgl* operon in the G50 background. As mentioned above, the PCR products amplified using different primer pairs were integrated into the chromosome of G50 instead of BW25113, yielding strains G50P*tet-bglJ*, G50P*tet-leuO* and G50P*tet-rcsB*. The *km*^{r} gene in

G50P*tet-bglJ* was deleted using pCP20, yielding G50P*tet-bglJ*-km^S which is sensitive to kanamycin. Using P1-transduction with G50P*tet-bglJ*-km^S as the recipient strain, G50P*tet-bglJ*/*rcsB* (double overexpression of *bglJ* and *rcsB* in G50) and G50P*tet-bglJ*Δ*rcsB* (deletion of *rcsB* and overexpression of *bglJ* in G50) were made.

Changing H-NS binding to the *bglG* gene

In vitro protein/DNA binding assays showed that H-NS proteins are capable of binding to a region within the *bglG* gene, which is +479 to +606 relative to the *bglG* translation initiation site (Dole *et al*., 2004). To reduce or abolish H-NS binding to this region (i.e., the presumptive H-NS operator), part of this region in strain G50 was replaced by one or two operators of TetR using the TetA-SacB positive and negative selection system (Li *et al*., 2013). Briefly, to add one TetR operator, the 19 nucleotides with sequences aaatgctgcaattaataaa, between +544 and +564 relative to the *bglG* translation initiation site, was replaced by one TetR operator with sequence tccctatcagtgatagaga (Lutz & Bujard, 1997), yielding G50.1R. Similarly, a region of 40 nucleotides with sequence gtgtcacgcagttaatgcgcgaaatgctgcaattaataaa (between +523 and +564) was replaced by two identical TetR operators with sequences tccctatcagtgatagagaCGtccctatcagtgatagaga (two TetR operators separated by two nucleotides, CG, in the center), which yielded G50.2R. The two strains, G50.1R and G50.2R, were used as recipient strains in the P1-transduction for transferring *tetR* with a *km*^{r} marker into their chromosomes, yielding strains RIG50.1R and RIG50.2R, respectively.

To completely abolish any possible binding of H-NS to the *bglG* gene, the entire *bglG* gene plus both flanked terminators was deleted. First, the region from -78 to +946 relative to the *bglG* translation start point, containing the first terminator, the *bglG* gene and the second

terminator, was replaced by a *km*^{r} gene amplified from pKD4 (Datsenko & Wanner, 2000) using oligos G50-P1 and G50-P2 (**Table 2.1**). The *km*^{r} gene was then flipped out using pCP20, yielding strain Δ*bglG*T_1T_2, in which the *bglG* gene with two terminators was replaced by an 85-bp fragment.

Changing the distance between the two H-NS binding sites

To see if the stability of H-NS mediated looping in the *bgl* operon is size dependent, the distance between the two H-NS binding sites was increased. A 1kb fragment carrying a *cat* gene from pKD3 was inserted into the beginning of the *bglG* gene of BW25113, replacing the small region between +209 and +220 relative to the *bglG* translation initiation site. In the resulting strain, a 1017 bp DNA fragment carrying the *cat* gene with its own promoter and two FRT sites was inserted into the *bglG* gene between the two H-NS binding sites, oriented in an opposite direction as the *bglG* gene. This yielded strain *cat*-G, in which the distance between these two H-NS binding sites was increased by a 1 kb. The *cat* gene was then flipped out using pCP20, yielding strain 85bp-G, in which an 85-bp fragment is inserted between the two H-NS binding sites. Similarly, a *cat* gene was inserted into the same location of the *bglG* gene in the G50 background, yielding strain *cat*-G50, with an 1kb distance increase between the two H-NS binding sites.

Additions of 5 bp and 10 bp fragments into the H-NS-produced DNA loop

DNA is a double-stranded helix with 10–10.5 bps per turn. To determine if the DNA looping within the *bgl* operon, formed via H-NS protein oligomerization, is DNA phase dependent or not, a 5-bp fragment with nucleotides ATAGA and a 10-bp fragment with

nucleotides ATAGAATAGA were inserted individually into the *bglG* gene between +108 and +109 relative to the *bglG* translation initiation site in strain G50.

Briefly these two small fragments (5 bps and 10 bps) were first integrated into the designated locus (i.e., between +108 and +109) within the *bglG* gene in plasmid pKDT_P*bgl*-G50 using fusion PCR, yielding plasmids pKDT_P*bgl*-G50.5bp and pKDT_P*bgl*-G50.10bp, respectively. Using these plasmids as templates, the "*km*^{r}:*rrnB*T:P*bgl*-G50.P5" cassette and "*km*^{r}:*rrnB*T:P*bgl*-G50.P10" cassette were amplified using oligos G50-cat-P1 and bglF.G50-P2 (**Table 2.1**). The PCR products were electroporated into ΔP*bglG* cells to substitute for the 85-bp scar that replaced the *bgl* regulatory region and the *bglG* gene together with the two flanked terminators. Several Km resistant colonies were confirmed by colony PCR and DNA sequencing to carry a 5-bp insertion or a 10-bp insertion at the expected location within the *bglG* gene, yielding strains G50.P5 and G50.P10. Compared to the original G50 strain, the phasing of those two H-NS binding sites is almost unchanged in strain G50.P10. However, the phasing should be almost essentially opposite in strain G50.P5.

Table 2.1 List of oligonucleotides used in this study

Primers	Sequence	Purpose
Pbglcat-P1	taagtctggagtcgctgggccgtcataccatccagatgctgcac gacgtgctgtgtaggctggagctgcttcg	Construction of ΔP*bglG* and G50ΔO$_{HNS}$-P*bgl*
bglFcat-P2	caatgttatctgcgcccccgactcctgcgactatttttctggctaac tccgtcatatgaatatcctccttagttc	Construction of ΔP*bglG*
bglF-ver-R4	tggttacctatgaccacctgaaactg	Verification for ΔP*bglG*
PhoU1-P1	gatgagctggataaactgctggcggggaaagatagcgacaaat aattcaccagtgtgtaggctggagctgcttcg	Construction of ΔP*bglGFB*
bglB2-P2	gaattaacgcatcgcatccagactgttctgaatgcgacgataatta aggtgcatatgaatatcctccttagttc	Construction of ΔP*bglGFB*
bglH-ver-R	gttgttccaccgttaatgattcag	Verification for ΔP*bglGFB*
Pbgl-Xh-Fbg	atactcgagcggatggacattgacgaagcggtacg	Construction of pKDT_P*bgl*-G50, G50.1R, G50.2R, G50.P5 and G50.P10
bglG-Bm-Rm	ttaggatccatagctgtttcctgtgaagctatttcagtg	Construction of pKDT_P*bgl*-G50, G50.1R, G50.2R, G50.P5 and G50.P10
Pbgl.G50-P1	taagtctggagtcgctgggccgtcataccatccagatgctgcac gacgtgctgtgtaggctggagctgcttcg	Construction of G50 and G51, G50.P5 and G50.P10
bglF.G50-P2	caatgttatctgcgcccccgactcctgcgactatttttctggctaac tccgtcatagctgtttcctgtgaagctatttcagtg	Construction of G50, G51, G50.P5 and G50.P10
HNS-AB-F	ctcttcgtgcgcaggcaagagaatgtacacttgaaacgctggaa gaaatgctcctaatttttgttgacactctatc	Construction of HNSL30P
HNS-AB-R	agccgcgctttcttcttcgcgacgttcgttaacaacaacttctaattt ttcatcaaagggaaaactgtccatatgc	Construction of HNSL30P
HNS-100	cttcgtgcgcaggcaagagaatgtacacttgaaacgctggaaga aatgcctgaaaaattagaagttgttgttaacgaacgtcgcgaaga agaaagcgcgg	Construction of HNSL30P
HNS-F	cttaaaattctgaacaacatccgtactcttcgtgcgcaggcaaga gaatg	Construction of HNSL30P
HNS-R	gagtgcgctcttcaacttcagcagcagccgcgctttcttcttcgcg acgttc	Construction of HNSL30P
HNS-ver-R	agattattgcttgatcaggaaatcgtc	Verification of HNSL30P
Pbgl2-P2	gaaaatagcaatgcgctattgataaaaatatgaccatgctcgcag ttattaacatatgaatatcctccttagttc	Construction of G50ΔO$_{HNS}$-P*bgl*
Ptet-bglJ-P1	tgatatgaaagtgaatgctaaggataatttattcgcttaatctattaa tttgtgtaggctggagctgcttc	Construction of P*tet-bglJ*
Ptet-bglJ-P2	cataacaggtattttactgataacaattccatttttgcagcatcctgg caacatggtacctttctcctctttaatgaattc	Construction of P*tet-bglJ*
bglJ-ver-R	aacagaacgtggatcttcactttcac	Verification of P*tet-bglJ*
Ptet-leuO-P1	ttatggattattatgctgtggtaaatgactcattccacggcaatgga ttctgtgtaggctggagctgcttc	Construction of P*tet-leuO*

Table 2.1 (continued) List of oligonucleotides used in this study

Primers	**Sequence**	**Purpose**
Ptet-leuO-P2	ggtttgcttaactccgccgtctctggatgatctgtttgtacctctgg catggtacctttctcctctttaatgaattc	Construction of P*tet-leuO*
leuO-ver-R	tgttttgctcctgcatcacggcatcg	Verification of P*tet-leuO*
Ptet-rcsB-P1	gaaaaatacatcagcgacattgacagttatgtcaagagcttgctg tagcaagtgtgtaggctggagctgcttc	Construction of P*tet-rcsB*
Ptet-rcsB-P2	gaacaagactatcggatggtcatcggcaataattacgttcatattg ttcatggtacctttctcctctttaatgaattc	Construction of P*tet-rcsB*
rcsB-ver-R	tacttatcgccaggcatggagagatcg	Verification of P*tet-rcsB*
Ptet-Xh-F	atactcgagactctatcattgatagag	Cloning of P*tet*
Ohns-F1	gttaatgcgcgtccctatcagtgatagagaatttcagttcagcctta attaccaggaag	Construction of G50.1R
Ohns-R1	ctctatcactgatagggacgcgcattaactgcgtgacacctgcaa c	Construction of G50.1R
bglG.2R-F	gaatttcagttcagccttaattaccaggaagaaagcttg	Construction of G50.2R
bglG.2R-R	gactgcaacatcctccatatttccgctc	Construction of G50.2R
2R-F	gtgcccaaatgagcggaaatatggaggatgttgcagtc	Construction of G50.2R
G50-P1	cctggatgttcgttataaaaaccattaataaatgactggattgttact gctgtgtaggctggagctgcttc	Construction of Δ*bglG*T_1T_2
G50-P2	gactatttttctggctaactccgtcataacttgccctctaccgctttg cggattccggggatccgtcgacctg	Construction of Δ*bglG*T_1T_2
G50-cat-P1	cagtcatgaactgaacgggcgattaagcgaactcttaagtcatat tcctcttgtgtaggctggagctgcttcg	Construction of *cat*-G, *cat*-G50, G50.P5 and G50.P10
G50-cat-P2	gtaattttcccaagcgctcctgcgctaaagagataatacgatcac aggttgccatatgaatatcctccttagttc	Construction of *cat*-G and *cat*-G50
G50-pha5-F	ctttcaaaaacgcatagagctggcgaaagaattaactcaagtgg	Construction of G50.P5
G50-pha5-R	gagttaattctttcgccagctctatgcgtttttgaaagccaattcc	Construction of G50.P5
G50-pha10-F	gctttcaaaaacgcatagaatagagctggcgaaagaattaactc aagtgg	Construction of G50.P10
G50-pha10-R	ctttcgccagctctattctatgcgtttttgaaagccaattccgc	Construction of G50.P10
Pu1n-P1n	gcatttacgttgacaccatcgaatggcgcaaaacctttcgcggta tgtgtaggctggagctgcttc	Construction of all *lacZ* reporters
Pbgl.dT-GZ-P2	cgacggccagtgaatccgtaatcatggtcatagctgtttcctgtgt gaaatttcagtgttctttgcgcacgcgctctat	Construction of G50-Z and G51-Z
G50tn-Z-P2	gttgtaaaacgacggccagtgaatccgtaatcatggtcatagctg tttcctgtgaagcttagcccacttcatctttcggtaactg	Construction of G50T-Z
Pbgl-F2	tggcgatgagctggataaactgctg	Amplification of *bgl* regulatory region
Pbgl-R2	tcagttcatgactgctcaaggcatac	Amplification of *bgl* regulatory region

Table 2.2 List of plasmids used in this study

Plasmid	**Descriptions**	**Reference or source**
pCP20	For removing *km*^{r} gene; ampicillin and chloramphenicol resistance	Datsenko & Wanner, 2000
pKD3	Template for FRT-flanked *cat*	Datsenko & Wanner, 2000
pKD4	Template for FRT-flanked *km*^{r}	Datsenko & Wanner, 2000
pKES50	P*bgl* driving a modified *bglG* gene (*bglG*50) without the two terminators	Dole *et al*., 2004
pKDT	The *rrnB* terminator (*rrnBT*) in pKD13	Klumpp *et al*., 2009
pKDT_P*bgl*-G50	P*bgl* driving *bglG*50 in pKDT	This study
pKES51	pKES50 with a point mutation in the Crp-cAMP operator	Dole *et al*., 2004
pKDT_P*bgl*-G51	pKDT_P*bgl*-G50 with a point mutation in the Crp-cAMP operator	This study
pKD46	Providing lambda-Red proteins for chromosomal integration	Datsenko & Wanner, 2000
pKDT:Ptet	The *rrnBT* followed by P*tet*;	Klumpp *et al*., 2009
pKDT_P*bgl*-G50.5bp	pKDT_P*bgl*-G50 with a 5 bp insertion in *bglG50*	This study
pKDT_P*bgl*-G50.10bp	pKDT_P*bgl*-G50 with a 10 bp insertion in *bglG50*	This study

Table 2.3 List of strains used in this study

Strains	**Descriptions**
BW25113	Wild-type
BW-RI	BW25113 constitutively expressing *tetR*
Δ*hns*	*hns* deletion in BW25113
Δ*stpA*	*stpA* deletion in BW25113
Δ*hns*Δ*stpA*	*hns* and *stpA* double deletions in BW25113
Δ*PbglG*	Deletion of P*bgl*, *bglG* and two terminators flanking *bglG*
Δ*PbglGFB*	Deletion of the entire *bgl* operon, including the regulatory region, *bglG*, *bglF*, and *bglB*
G50	Deletion of the two terminators flanking *bglG*
RIG50	G50 constitutively expressing *tetR*
Δ*crp*G50	G50 with deletion of *crp*
G51	G50 with a single mutation at the Crp binding site at the *bgl* regulatory region
Δ*crp*G51	G51 with deletion of *crp*
HNSL30P	Mutation of Leucine to Proline at 30th codon of *hns*
G50ΔO_{HNS}-Pbgl	Removal of H-NS binding site at the *bgl* regulatory region in G50
IS1r(-236.5)	IS1 element inserted in reverse orientation at position -236.5 in relation to the *bglG* translational initiation site
IS1r(-231.5)	IS1 element inserted in reverse orientation at position -231.5 in relation to the *bglG* translational initiation site
IS1d(-230.5)	IS1 element inserted in direct orientation at position -230.5 in relation to the *bglG* translational initiation site
IS5r(-224.5)	IS5 element inserted in reverse orientation at position -224.5 in relation to the *bglG* translational initiation site
IS5d(-207.5)	IS5 element inserted in direct orientation at position -207.5 in relation to the *bglG* translational initiation site
IS5r(-207.5)	IS5 element inserted in reverse orientation at position -207.5 in relation to the *bglG* translational initiation site
IS5d(-206.5)	IS5 element inserted in direct orientation at position -206.5 in relation to the *bglG* translational initiation site
IS2d(-206.5)	IS2 element inserted in direct orientation at position -206.5 in relation to the *bglG* translational initiation site
IS2r(-205.5)	IS2 element inserted in reverse orientation at position -205.5 in relation to the *bglG* translational initiation site
IS5r(-207.5)Δ*stpA*	Removal of *stpA*; Insertion of IS5 element in reverse orientation in the *bgl* regulatory region
IS5r(-207.5)Δ*hns*	Removal of *hns*; Insertion of the IS5 element in reverse orientation in the *bgl* regulatory region
P*tet-bglJ*	Overexpression of *bglJ* in BW25113
P*tet-bglJ*-km^S	Overexpression of *bglJ* in BW25113; sensitive to kanamycin
P*tet-leuO*	Overexpression of *leuO* in BW25113
P*tet-rcsB*	Overexpression of *rcsB* in BW25113
P*tet-bglJ*/*rcsB*	Overexpression of *bglJ* and *rcsB* in BW25113

Table 2.3 (continued) List of strains used in this study

Strains	**Descriptions**
P*tet-bglJ*Δ*rcsB*	Deletion of *rcsB* and overexpression of *bglJ* in BW25113
G50P*tet-bglJ*	Overexpression of *bglJ* in G50
G50P*tet-bglJ*-km^S	Overexpression of *bglJ* in G50; sensitive to kanamycin
G50P*tet-leuO*	Overexpression of *leuO* in G50
G50P*tet-rcsB*	Overexpression of *rcsB* in G50
G50P*tet-bglJ*/*rcsB*	Overexpression of *bglJ* and *rcsB* in G50
G50P*tet-bglJ*Δ*rcsB*	G50 with deletion of *rcsB* and overexpression of *bglJ*
G50.1R	G50 with H-NS binding site within *bglG* replaced by one TetR binding site
RIG50.1R	G50.1R constitutively expressing *tetR*
G50.2R	G50 with H-NS binding site within *bglG* replaced by two TetR binding sites
RIG50.2R	G50.2R constitutively expressing *tetR*
Δ*bglG*T$_1$T$_2$	Removal of *bglG* and two terminators in BW25113
cat-G	1 kb insertion of the *cat* gene in the beginning of *bglG*
cat-G50	G50 with a 1 kb insertion of the *cat* gene within *bglG*
85bp-G	Removal of the *cat* gene from *cat*-G50, leaving an 85-bp scar within *bglG*
G50.P5	G50 with 5 base pairs inserted within *bglG*
G50.P10	G50 with 10 base pairs inserted within *bglG*

2.1.2 Operon reporters

Operon reporters were created for β-galactosidase (LacZ) assay in this study. The original *bgl* operon was intact in the operon reporters. The *bgl* promoter followed by the *bglG* gene with or without the two terminators, were fused with the *lacZ* gene at the original *lacZ* location (**Figure 2.1**). The new operon reporter construct replaced the promoter of *lacI*, the *lacI* gene and the promoter of *lacZ*.

A.

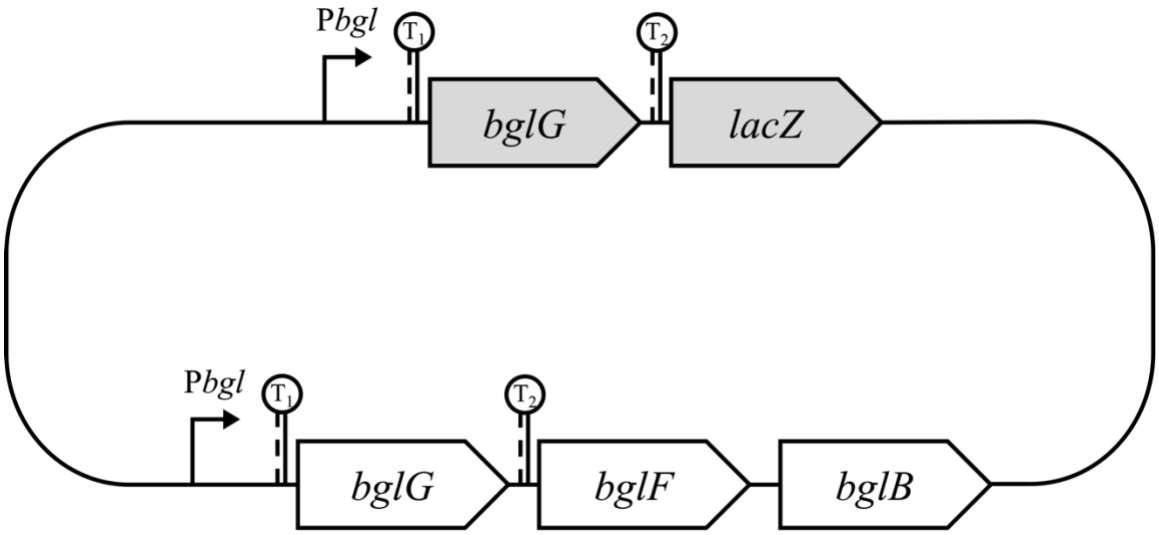


Figure 2.1 Structure of operon reporters in the chromosome. (**A**) BW25113_Z (operon reporter of P*bgl-bglG-lacZ* at the original *lacZ* location in wild-type BW25113). (**B**) G50-Z (operon reporter of P*bgl-bglG-lacZ* without the two *bgl* terminators, at the original *lacZ* location in wild-type BW25113)

B.

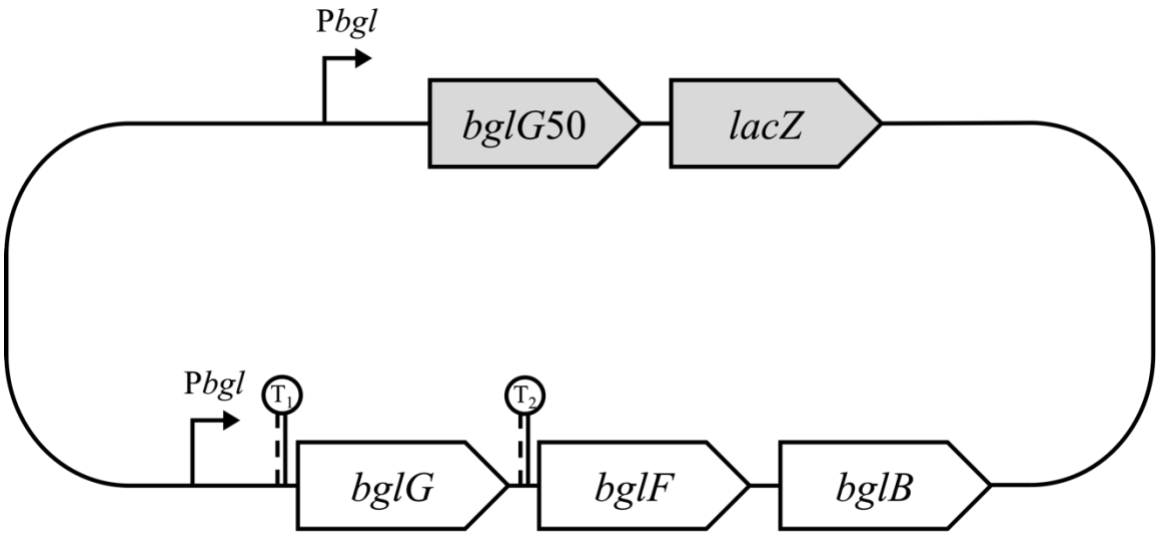


Figure 2.1 (continued) Structure of operon reporters in the chromosome. (**A**) BW25113_Z (operon reporter of P*bgl-bglG-lacZ* at the original *lacZ* location in wild-type BW25113). (**B**) G50-Z (operon reporter of P*bgl-bglG-lacZ* without the two *bgl* terminators, at the original *lacZ* location in wild-type BW25113)

Construction of the *bgl* operon reporters

Strain BW25113_Z (Zhang et al., 2022) carries a *bgl* operon LacZ reporter, which is composed of the *bglGFB* promoter, the first terminator, *bglG*, the *bglG*/*bglF* intergenic region (containing the second terminator) and the first 9 codons of *bglF*, plus a stop codon at the 3' end immediately upstream of the *lacZ* gene at the *lac* locus (**Figure 2.1A**). A *km*^{r} gene followed by a *rrnB* terminator is located upstream of this reporter.

Using the same protocol for strain BW25113_Z construction (Zhang *et al*., 2022), the cassettes of "*km*^{r}:*rrnB*T:P*bgl-bglG50*" in plasmid pKDT_P*bgl*-G50 and "*km*^{r}:*rrnB*T:P*bgl-bglG50*" in plasmid pKDT_P*bgl*-G51 were individually integrated into the chromosome of the default strain MGIBKY (Klumpp *et al*., 2009) to replace the *lacI* gene and the *lacI*/*lacZ* intergenic region, yielding strains G50-Z and G51-Z. The two operon reporters can be

transferred into other backgrounds by P1-transduction after the *km*^{r} gene is removed by pCP20, yielding G50-Z-km^{S} and G51-Z-km^{S} respectively. For example, Δ*crp*G50-Z and Δ*crp*G51-Z were made by transferring the *km*^{r} gene, replacing the *crp* gene to the strains G50-Z-km^{S} and G51-Z-km^{S} respectively.

In addition, positive regulatory genes of the *bgl* operon were overexpressed in the G50-Z background. Using G50-Z-km^{S}, The strong *tet* promoter driving *bglJ*, *leuO* and *rcsB*, each with a *km*^{r} marker, were transferred to G50-Z-km^{S} by P1-transduction, yielding operon reporters G50-Z-P*tet-bglJ*, G50-Z-P*tet-leuO* and G50-Z-P*tet-rcsB*, respectively. The *km*^{r} gene in the operon reporter, G50-Z-P*tet-bglJ*, was flipped out by pCP20, yielding G50-Z-P*tet-bglJ*-km^{S}, which was then used to make other operon reporters. For example, *rcsB* driven by the *tet* promoter with a *km*^{r} marker and the *km*^{r} gene replacing *rcsB*, were transferred to G50-Z-P*tet-bglJ*-km^{S} by P1-transduction, yielding operon reporters G50-Z-P*tet-bglJ*/*rcsB* and G50-Z-P*tet-bglJ*Δ*rcsB*, respectively.

As described above, H-NS proteins are capable of binding to a region within the *bglG* gene, which is +479 to +606 relative to the translation initiation site. To examine the effect of this region on the β-galactosidase (LacZ) activity, part of this region of the *bglG*50 gene in plasmid pKDT_P*bgl*-G50 was replaced by one or two operators of TetR using fusion PCR, yielding pKDT_P*bgl*-G50.1R and pKDT_P*bgl*-G50.2R, respectively. The cassette of "*km*^{r}:*rrnB*T:P*bgl-bglG50*.1R" in pKDT_P*bgl-G50*.1R and the cassette of "*km*^{r}:*rrnB*T:P*bgl-bglG50*.2R" in pKDT_P*bgl-G50*.2R were transferred to the same location on the chromosome as for G50-Z , yielding strains G50.1R-Z and G50.2R-Z, respectively. The two operon reporters are the same as G50-Z, except that part of the H-NS binding site within *bglG* was replaced by one or two TetR operators. Using G50.1R-Z and G50.2R-Z as recipient strains, two new operon

reporters were made by transferring *tetR* with a *km*^{r} marker using P1-transduction, yielding RIG50.1R and RIG50.2R respectively.

To further prevent H-NS binding to *bglG*50, a new operon reporter similar to G50-Z, carrying the *bgl* promoter and only the 5' region of *bglG*50 with no 3' region was made (G50T-Z). The cassette of "*km*^{r}:*rrnB*T:P*bgl-bglG*50T" (*bglG*50T refers to the region from +1 to +465 relative to the *bglG* translation start start), amplified from pKST_P*bgl*-G50 by oligos P1n-P1n and G50tn-Z-P2 (**Table 2.1**), was fused to the *lacZ* gene as for other operon reporters, yielding operon report G50T-Z. There is a stop codon (TAA) added to the end of *bglG*50T and the RBS for *lacZ* remains the same as for G50-Z.

Table 2.4 List of operon reporters used in this study

Operon reporters	**Descriptions**
BW25113_Z	The *bgl* promoter followed by *bglG* driving *lacZ* gene expression
G50-Z	The *bgl* promoter followed by *bglG* without the two terminators, driving *lacZ* gene expression
G51-Z	G50-Z with a single mutation at the Crp binding site at the regulatory region of the *bgl* operon reporter
G50-Z-km^S	G50-Z; kanamycin sensitive
G51-Z-km^S	G51-Z; kanamycin sensitive
Δ*crp*G50-Z	G50-Z with the removal of *crp*
Δ*crp*G51-Z	G51-Z with the removal of *crp*
G50-Z-P*tet-bglJ*	G50-Z with overexpression of *bglJ*
G50-Z-P*tet-bglJ*-km^S	G50-Z with overexpression of *bglJ*; kanamycin sensitive
G50-Z-P*tet-leuO*	G50-Z with overexpression of *leuO*
G50-Z-P*tet-rcsB*	G50-Z with overexpression of *rcsB*
G50-Z-P*tet-bglJ*/*rcsB*	G50-Z with overexpression of *bglJ* and *rcsB*
G50-Z-P*tet-bglJ*Δ*rcsB*	G50-Z with deletion of *rcsB* and overexpression of *bglJ*
G50.1R-Z	The *bgl* promoter followed by *bglG* without the two terminators, and H-NS binding site within *bglG* replaced by one TetR binding site, driving *lacZ* gene expression
G50.2R-Z	The *bgl* promoter followed by *bglG* without the two terminators, and the H-NS binding site within *bglG* replaced by two TetR binding site, driving *lacZ* gene expression
RIG50.1R-Z	G50.1R-Z with the presence of *tetR*
RIG50.2R-Z	G50.2R-Z with the presence of *tetR*
G50T-Z	The *bgl* promoter followed by a truncated *bglG*, with the H-NS binding site and the two terminators removed, driving *lacZ* gene expression

2.2 Determining the Bgl^+/Bgl^- phenotype of strains

E. coli cells were grown on MacConkey indicator plates with 0.5% salicin at 37°C overnight. Visible red color was shown for colonies with a Bgl^+ phenotype, indicating their ability to ferment salicin. Otherwise, colonies with a Bgl^- phenotype remained white in color. In addition, cells were grown on M9 minimal agar plates with 0.5% salicin as the sole carbon source at 37°C for 2 days to further confirm their phenotype. To be a Bgl^+ phenotype, cells should form visible colonies during the 2-day incubation. Otherwise, the cells are considered as Bgl^-. In this study, we define a Bgl^+ phenotype as the appearance of red colonies on MacConkey + salicin plates plus the formation of visible colonies on M9 + salicin plates.

2.3 β-galactosidase (LacZ) activity assay (Miller, 1972)

Sample preparation:

Reporter strains were grown in LB medium at 37°C with shaking. After at least 6 hours, 20 µL of the culture was transferred to 3 mL of 1x M63 minimal media with 0.5% glycerol. Then it was left overnight for growth at 37°C with shaking. An appropriate amount of the overnight culture was inoculated into 5 mL 1x M63 minimal medium with 0.5% glycerol, with a starting OD_{600} of 0.03. The cells were grown at 37°C with shaking. Four samples were collected during the exponential growth phase, when the OD_{600} reaches around 0.2, 0.4, 0.6 and 0.8. Collected samples could be kept frozen at -20°C prior to subsequent measurement.

β-galactosidase (LacZ) assay:

200 µL of samples and 800 µL of Z-buffer were mixed with 25 µL of chloroform by vortexing 10 secs twice. The samples were warmed in a 37°C water bath. To start the reaction, 200 µL of o-nitrophenyl-ß-D-galactopyranoside (ß-ONPG) was added to each sample. After a visible yellow color appeared, 0.5 mL of 1M sodium carbonate was added and vortexed to stop the reaction. The reaction mixture was diluted, if appropriate, and centrifuged with a speed of 15,000 rpm for 2.5 mins. The absorbance of the reaction mixture was measured at 420 nm and 550 nm. The LacZ activities of each sample were calculated using the formula below:

$$\frac{1000 \times (OD_{420\,nm} - 1.75 \times OD_{550\,nm}) \times Dilution\ factor}{Time\ of\ reaction\ (min) \times Volume\ of\ sample\ (mL)}$$

The activity of the reporter stain was determined by the slope of the LacZ activity versus cell density of the four samples collected at different times.

2.4 β-glucosidase (BglB) activity assay (Schnetz & Rak, 1988; Dole *et al*., 2002)

Sample preparation:

The cells were grown in LB medium at 37°C with shaking. After at least 6 hours, 20 µL of the culture was transferred to 1x M63 minimal media with 0.66% casamino acid, then it was left overnight for growth at 37°C with shaking. An appropriate amount of the overnight culture was inoculated into 6 mL of 1x M63 minimal medium with 0.66% casamino acids and 0.5% salicin, at a starting OD_{600} of 0.025. The cells were grown at 37°C with shaking. Five samples with 0.8 mL each were collected during the late exponential growth phase when the *bgl* operon was fully induced with an OD_{600} above 1.5. The samples were centrifuged in a Sorval centrifuge at a speed of 5,500 rpm for 2.5 mins. The supernatant was removed, and the cells were suspended with 1 mL of Z-buffer with 50 µL/mL chloramphenicol.

β-glucosidase (BglB) assay:

The samples were warmed in 37ºC water baths. To start the reaction, 200 µL of p-nitrophenyl-β-D-glucoside (PNPG, 8 mg/mL) were added to the cell suspension in Z-buffer. After a visible yellow color appeared, 0.5 mL of 1M sodium carbonate was added and vortexed to stop the reaction. The reaction mixture was diluted, if appropriate, and centrifuged in a Sorval centrifuge at a speed of 15,000 rpm for 2.5 mins. The absorbance of the reaction mixture was measured at 420 nm and 550 nm. The BglB activity of each sample was calculated using the formula:

$$\frac{1000 \times (OD_{420\,nm} - 1.75 \times OD_{550\,nm}) \times Dilution\ factor}{Time\ of\ reaction\ (min) \times Volume\ of\ sample\ (mL)}$$

The activity of the strain was determined by averaging the BglB activities of the samples measured.

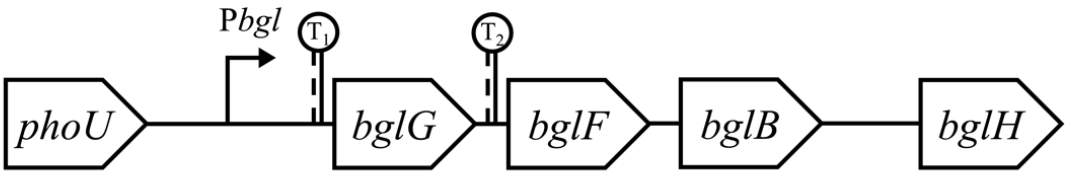
3. Results

3.1 Determination of background β-glucosidase (BglB) activities in *E. coli* cells

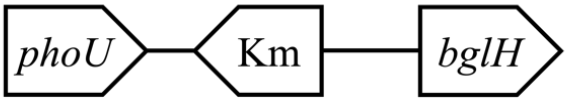
The *bglGFB* operon is silent in wild-type *E. coli* strains. Under our growth conditions, the wild-type strain, BW25113, exhibits only 0.24 units of β-glucosidase (BglB) activity. To see if this low β-glucosidase (BglB) activity is encoded within the *bgl* operon, the entire *bgl* operon, including the *bgl* regulatory region, *bglG*, *bglF* and *bglB*, was deleted from the wild-type strain, yielding strain Δ*PbglGFB* (**Figure 3.1A**). The resultant strain exhibits 0.165 units of β-glucosidase (BglB) activity that is attributed to other proteins since the *bglB* gene is deleted from this strain (**Figure 3.1B**). By subtracting the background activity, the actual β-glucosidase (BglB) activity of wild-type BW25113 was 0.075 units. The β-glucosidase (BglB) activities observed for the deletion strain (the background activity of 0.165 units) was subtracted for values obtained for the remainder of this study. In addition, both the wild-type BW25113 and Δ*PbglGFB* showed a Bgl^- phenotype on both MacConkey agar plates with salicin and M9 agar plates with salicin.

A.

(i) BW25113



(ii) ΔP*bglGFB*



B.

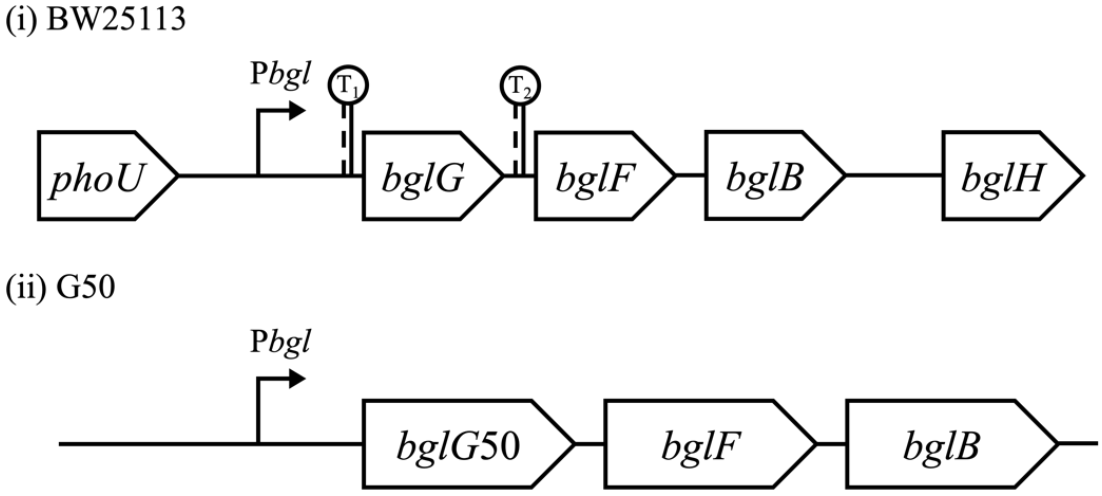
Strains	Initial β-glucosidase (BglB) activities	β-glucosidase (BglB) activities	Bgl Phenotype	
			MacConkey + salicin	M9 + salicin
BW25113	0.240	0.075	-	-
Δ*PbglGFB*	0.165	0	-	-

Figure 3.1 The operon activity of Δ*PbglGFB*. (**A**) Structure of (**i**) wild-type BW25113 and (**ii**) Δ*PbglGFB* (Km gene replacing the entire *bgl* operon), as well as (**B**) their β-glucosidase (BglB) activities and Bgl phenotypes.

3.2 Significance of the two *bgl* terminators for the utilization of β-glucosides in *E. coli*

To determine if the two terminators flanking both sides of *bglG* are required for the utilization of aromatic β-glucosides in *E. coli*, both terminators were removed from the *bgl* operon, yielding strain G50 (**Figure 3.2A**). The β-glucoside activity of the strain was similar to the wild-type, although with an increase. The strain also showed a Bgl$^-$ phenotype (**Figure 3.2B**), indicating that the *bgl* operon remains silent in the absence of both terminators flanking the *bglG* gene.

A.



B.

Strains	β-glucosidase (BglB) activities	Bgl Phenotype	
		MacConkey + salicin	M9 + salicin
BW25113	0.08	-	-
G50	0.25	-	-

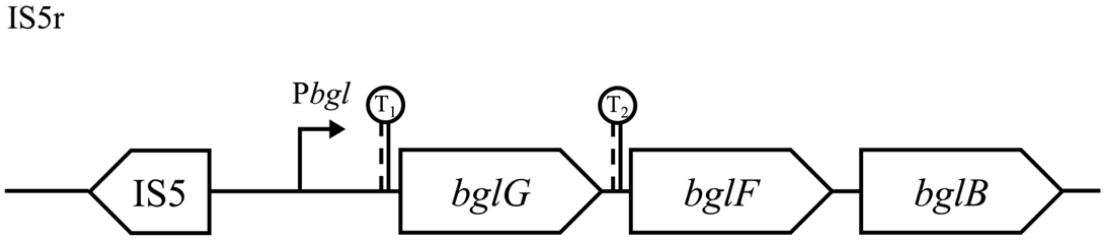
Figure 3.2 The operon activity of G50. (**A**) Structure of (**i**) wild-type (BW25113) and (**ii**) G50 (with the two *bgl* terminators removed), as well as (**B**) their β-glucosidase (BglB) activities and Bgl phenotypes.

3.3 Activation of the *bgl* operon by removal of the *bgl* repressors

H-NS is a major repressor of the *bgl* operon and StpA can serve as a molecular replacement of H-NS in repressing the expression of the *bgl* operon (Free *et al*., 1998). To investigate the mechanism of repression of the *bgl* operon by the products of the *hns* and *stpA* genes, these negative regulatory genes were deleted from wild-type BW25113. Their operon activities, as indicated by β-glucosidase (BglB) activities and growth phenotypes, were determined as compared to that of the IS5 insertional mutant in the *bgl* regulatory region (**Figure 3.3**). Compared to the wild-type, there was only a slight increase in the β-glucosidase (BglB) activity in the Δ*stpA* mutant strain, but the Δ*hns* mutant showed a 700-fold increase in activity. This confirms that H-NS is the primary repressor that silences the *bgl* operon. Moreover, the activity of the Δ*hns*Δ*stpA* double mutant was further increased by 2-fold in comparison to the Δ*hns* mutant, which was similar to that of the IS5 insertional mutant (IS5r(-207.5)). These results show that StpA does not repress the *bgl* operon in the presence of H-NS. However, in the absence of H-NS, StpA moderately represses the operon.

In addition, the *hns* and *stpA* mutations were individually transferred to IS5 insertional mutant cells with a Bgl^+ phenotype. These mutations do not further elevate the *bgl* operon activities in the IS5 insertional mutants (**Figure 3.3B**). The β-glucosidase (BglB) activity of the Δ*stpA* mutant (IS5r(-207.5)Δ*stpA*) was similar to that of the IS5 insertional mutant alone, while the Δ*hns* mutant (IS5r(-207.5)Δ*hns*) was slightly higher. In regard to their phenotypes, except for the Δ*stpA* mutant, all the other mutants showed a Bgl^+ phenotype.

A.



B.

Strains	β-glucosidase (BglB) activities	Bgl Phenotype	
		MacConkey + salicin	M9 + salicin
BW25113	0.08	-	-
Δ*stpA*	0.13	-	-
Δ*hns*	53	+++	+
Δ*hns*Δ*stpA*	108 ± 7	+++	+
IS5r(-207.5)	104 ± 5	+++	+++
IS5r(-207.5)Δ*stpA*	107 ± 4	+++	+++
IS5r(-207.5)Δ*hns*	116 ± 9	+++	++

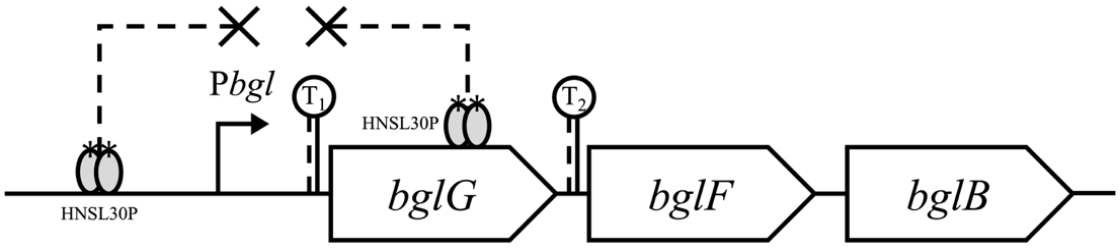
Figure 3.3 The operon activities upon deletion of *bgl* operon repressors. (**A**) Structure of IS5r-(-207.5) (IS5 element inserted in the *bgl* regulatory region). (**B**) The *bgl* operon repressors, H-NS and StpA, were removed separately from the wild-type strain and the IS5 insertional mutant. Their β-glucosidase (BglB) activities were measured, and their associated standard deviation values are shown. Their phenotypes were determined on MacConkey + salicin plates and M9 + salicin plates. The *hns* mutants (Δ*hns* and Δ*hns*Δ*stpA*) have severe growth defects, leading to a diminished Bgl^+ phenotype on M9 + salicin plates.

3.4 Breaking the H-NS loop at the *bgl* operon

3.4.1 Preventing H-NS self-oligomerization

To prevent the oligomerization of H-NS without altering H-NS binding to the *bgl* regulatory region and the *bglG* gene, a missense mutation was introduced into the *hns* gene leading to a change of leucine to proline at the 30th codon, yielding strain HNSL30P (**Figure 3.4A**) (Ueguchi *et al*., 1997). This derivative of H-NS is still capable of dimerizing, thereby binding to the target DNA. However, this mutant H-NS has lost the ability to oligomerize (Stella *et al*., 2005). Therefore, it is conceivable that this modified H-NS is unable to form nucleoprotein complexes involving DNA bridges and DNA looping. The operon activity of the strain increased over 580 times compared to wild-type even though it was lower than that of the strain Δ*hns* (**Figure 3.4B**). It also showed a Bgl^{+} phenotype on both MacConkey + salicin plates and M9 + salicin plates.

A.



B.

Strains	β-glucosidase (BglB) activities	Bgl Phenotype	
		MacConkey + salicin	M9 + salicin
BW25113	0.075	-	-
Δ*hns*	53 ± 5	+++	+
HNSL30P	44 ± 4	+++	+

Figure 3.4 The operon activity of HNSL30P. (**A**) The *bgl* operon of strain HNSL30P. The H-NS dimer is able to bind to the *bgl* regulatory region and *bglG*, but it has lost the capacity for self-oligomerization. (**B**) The β-glucosidase (BglB) activity with the associated standard deviation values, and the phenotype of the strain HNSL30P on MacConkey + salicin plates and M9 + salicin plates are reported. The *hns* mutants (Δ*hns* and HNSL30P) have severe growth defects, leading to a diminished Bgl^+ phenotype on M9 + salicin plates.

3.4.2 Preventing or reducing H-NS binding to the upstream site of the *bgl* regulatory region

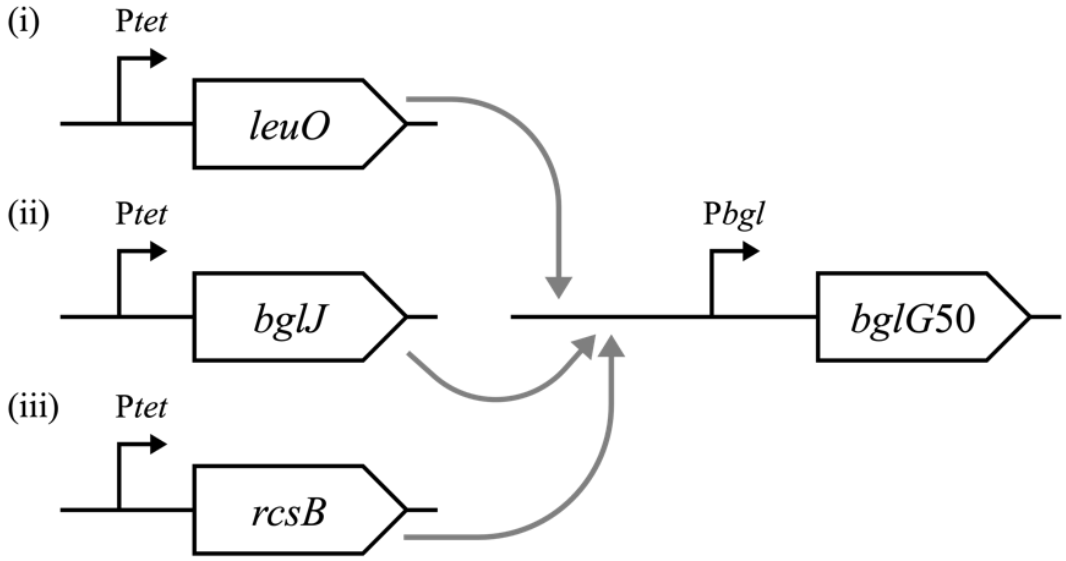
3.4.2.1 Overproduction of positive regulators of the *bgl* operon

Positive regulatory genes of the *bgl* operon, including *bglJ*, *leuO* and *rcsB*, were overexpressed in the background of the operon reporter G50-Z, so that the positive regulators could outcompete H-NS binding to the *bgl* regulatory region (**Figure 3.5A Bi**). In comparison to G50-Z, the β-galactosidase (LacZ) activities of operon reporters increased by approximately 4 times and 14 times when *leuO* and *bglJ* were overexpressed respectively (**Figure 3.5C**). The operon activity only increased 2 times when both *bglJ* and *rcsB* were overexpressed. However, the operon activity even dropped slightly when *rcsB* was overexpressed. Also, the operon activity remained the same as G50-Z when *bglJ* was overexpressed and *rcsB* was deleted.

Similarly, positive regulatory genes were overexpressed in the G50 background, in which both terminators flanking the *bglG* gene were removed, and the β-glucosidase (BglB) activities were measured (**Figure 3.5A Bii**). The data obtained were consistent with the results from the β-galactosidase (LacZ) measurements (**Figure 3.5D**). The β-glucosidase (BglB) activities increased significantly compared to G50 when *leuO* or *bglJ* was overexpressed individually. The operon activity doubled when *bglJ* and *rcsB* were overexpressed, but it dropped slightly when *rcsB* alone was overexpressed, as well as when *bglJ* was overexpressed and *rcsB* was deleted.

The overexpression strains in both wild-type and G50 backgrounds were tested for their growth phenotypes. All of them were still Bgl^- and were incapable of growing on salicin plates. However, colonies of *leuO*, and *bglJ* overexpression strains appeared to be slightly red on MacConkey + salicin plates although these strains cannot utilize salicin as the sole carbon source. These growth phenotypes are consistent with low levels of the operon activities.

A.



B.

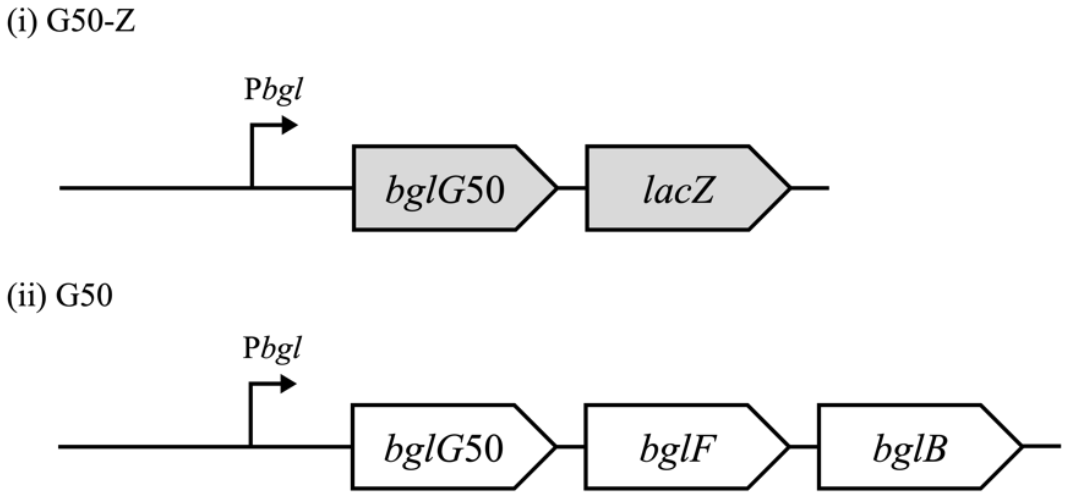
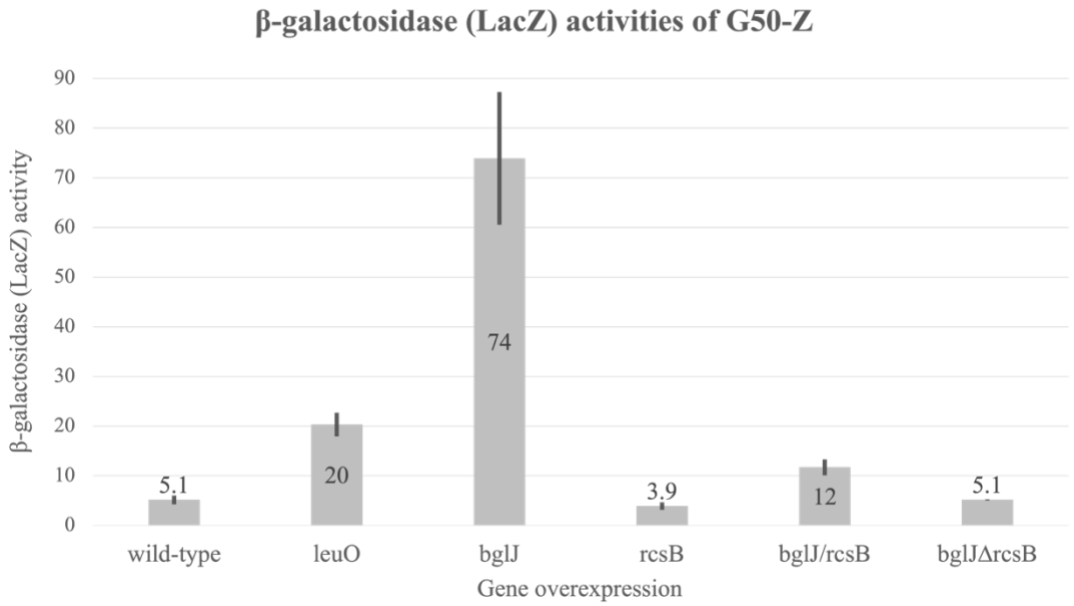


Figure 3.5 The operon activities resulting from overexpression of *bgl* regulatory genes. (**A**) Overexpression of the *bgl* positive regulatory genes using the strong *tet* promoter driving (**i**) *leuO*, (**ii**) *bglJ* and (**iii**) *rcsB*. (**B**) Structure of (**i**) operon reporter G50-Z (wild type *bgl* promoter and *bglG* gene fused to the *lacZ* gene, with removal of the two terminators) and (**ii**) G50 (with the two *bgl* terminators removed). (**C**) The β-galactosidase (LacZ) activities of operon reporters (G50-Z) with overexpression of *bgl* positive regulatory genes. (**D**) The β-glucosidase (BglB) activities of strains with overexpression of *bgl* positive regulatory genes in the G50 background. Error bars represent standard deviations. (**E**) The Bgl phenotypes with overexpression of *bgl* positive regulatory genes in wild-type and G50 backgrounds.

C.



D.

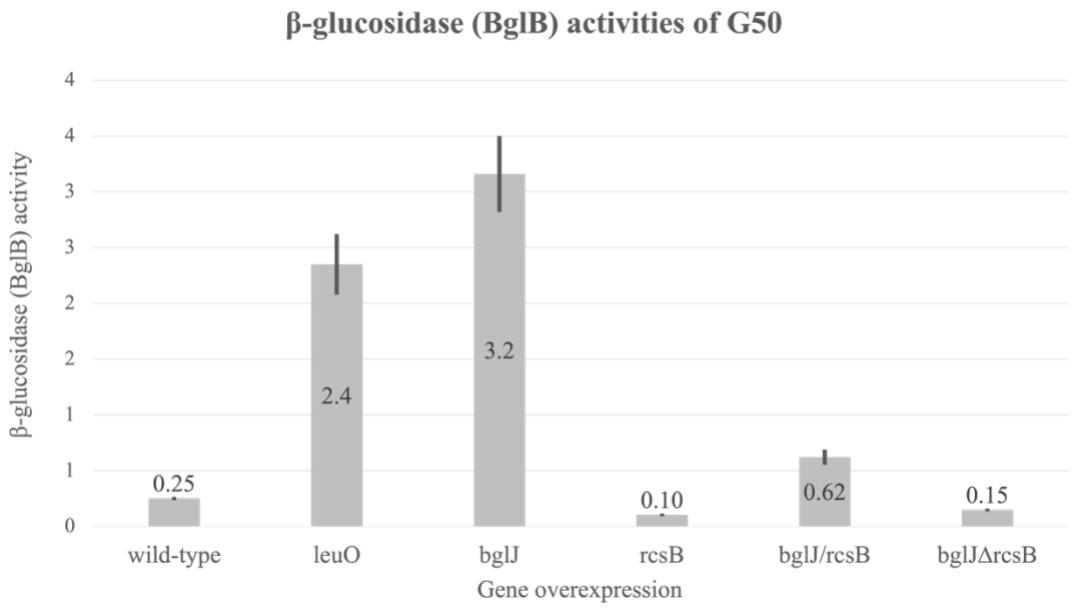


Figure 3.5 (continued) The operon activities resulting from overexpression of *bgl* regulatory genes. (**A**) Overexpression of the *bgl* positive regulatory genes using the strong *tet* promoter driving (**i**) *leuO*, (**ii**) *bglJ* and (**iii**) *rcsB*. (**B**) Structure of (**i**) operon reporter G50-Z (wild type *bgl* promoter and *bglG* gene fused to the *lacZ* gene, with removal of the two terminators) and (**ii**) G50 (with the two *bgl* terminators removed). (**C**) The β-galactosidase (LacZ) activities of operon reporters (G50-Z) with overexpression of *bgl* positive regulatory genes. (**D**) The β-glucosidase (BglB) activities of strains with overexpression of *bgl* positive regulatory genes in the G50 background. Error bars represent standard deviations. (**E**) The Bgl phenotypes with overexpression of *bgl* positive regulatory genes in wild-type and G50 backgrounds.

E.

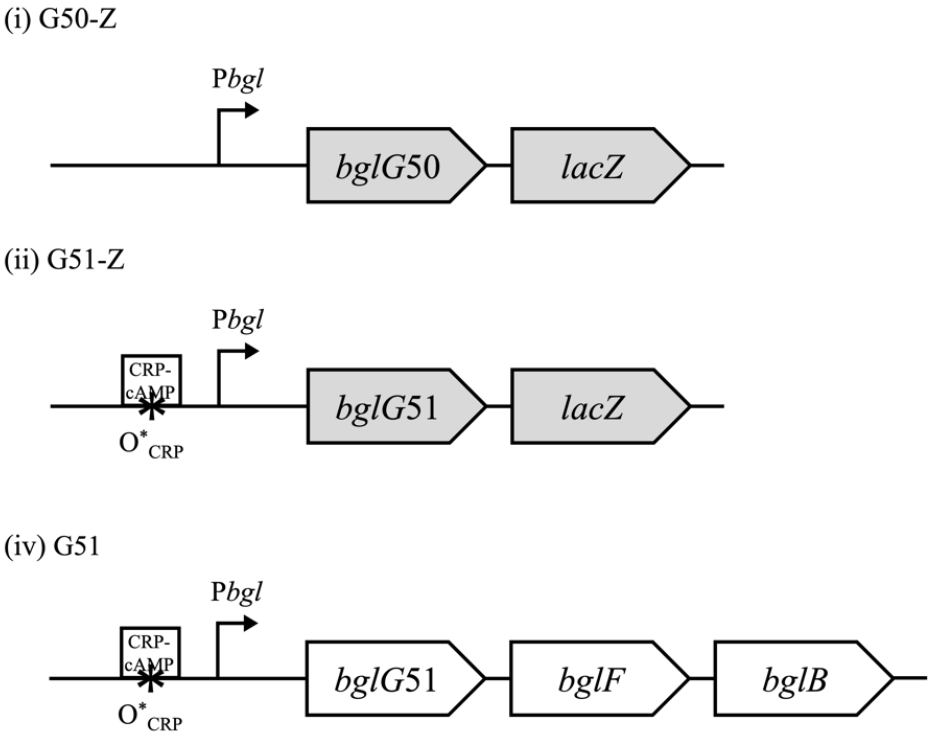
Strains	Bgl Phenotype	
	MacConkey + salicin	**M9 + salicin**
G50	-	-
Overexpression of *bglJ*	+/-	-
Overexpression of *leuO*	+/-	-
Overexpression of *rcsB*	-	-
Overexpression of *bglJ* and *rcsB*	-	-
Overexpression of *bglJ* + Deletion of *rcsB*	-	-

Figure 3.5 (continued) The operon activities resulting from overexpression of *bgl* regulatory genes. (**A**) Overexpression of the *bgl* positive regulatory genes using the strong *tet* promoter driving (**i**) *leuO*, (**ii**) *bglJ* and (**iii**) *rcsB*. (**B**) Structure of (**i**) operon reporter G50-Z (wild type *bgl* promoter and *bglG* gene fused to the *lacZ* gene, with removal of the two terminators) and (**ii**) G50 (with the two *bgl* terminators removed). (**C**) The β-galactosidase (LacZ) activities of operon reporters (G50-Z) with overexpression of *bgl* positive regulatory genes. (**D**) The β-glucosidase (BglB) activities of strains with overexpression of *bgl* positive regulatory genes in the G50 background. Error bars represent standard deviations. (**E**) The Bgl phenotypes with overexpression of *bgl* positive regulatory genes in wild-type and G50 backgrounds.

3.4.2.2 Stronger Crp-cAMP binding to the *bgl* regulatory region

A point mutation in the Crp binding site was introduced into the operon reporter strain without the two *bgl* terminators (G50-Z), yielding the new operon reporter G51-Z (**Figure 3.6A**). The new operon reporter has stronger Crp-cAMP binding to the *bgl* regulatory region. The β-galactosidase (LacZ) activity of the new reporter showed a significant increase of more than 70 times (**Figure 3.6B**). To verify that the great increase in operon activity was due to stronger Crp-cAMP binding to the *bgl* regulatory region, *crp* was removed from the new operon reporter (Δ*crp*G51-Z). In the absence of *crp*, the β-galactosidase (LacZ) activity dropped to the level observed for G50-Z without the point mutation. This confirmed that the stronger binding of Crp to the *bgl* regulatory region caused the great increase in activity of the operon reporter G51-Z. Despite the great increase in operon activity, the strain G51 showed a Bgl^- phenotype (**Figure 3.6C**).

A.



B.

Operon reporters	**β-galactosidase (LacZ) activities**	
	Crp present	Crp absent
G50-Z	5.1	1.3
G51-Z	375	1.2

C.

Strains	**Bgl Phenotype**			
	MacConkey + salicin		**M9 + salicin**	
	Crp present	Crp absent	Crp present	Crp absent
G50	-	-	-	-
G51	-	-	-	-

Figure 3.6 The operon activity of G51-Z. (**A**) Structure of (**i**) operon reporter G50-Z (wild type *bgl* promoter and the *bglG* gene fused to the *lacZ* gene, with removal of the two terminators), (**ii**) operon reporter G51-Z (G50-Z with a point mutation in the Crp binding site), and (**iii**) strain G51 (G50 with a single mutation in the Crp binding site within the *bgl* regulatory region). (**B**) The β-galactosidase (LacZ) activities of the operon reporters were measured, in the presence and absence of Crp. (**C**) Their phenotypes were also determined on MacConkey + salicin plates and M9 + salicin plates.

3.4.2.3 IS insertional mutations within the *bgl* regulatory region

All the IS insertional mutants with insertion in the *bgl* regulatory region showed Bgl^+ phenotypes on MacConkey + salicin plates as well as M9 + salicin plates. To further investigate the mechanism of the activation of the *bgl* operon by IS insertions in relation to the H-NS loop, the (1) type, (2) location and (3) orientation of the IS elements inserted at the *bgl* regulatory region were examined. The *bgl* regulatory regions of the insertional mutants were sequenced and analyzed with their associated β-glucosidase (BglB) activities.

Locations of IS insertional elements in the *bgl* regulatory region

All of the IS insertional mutants, including IS1, IS2 and IS5, were inserted from positions -255 to -205 relative to the *bglG* translational initiation site (**Figure 3.7**). IS1 elements were inserted between positions -237 to -229 relative to the *bglG* translational site, IS5 elements were inserted between positions -225 to -205, and IS2 elements were inserted between positions -207 to -204.

Orientation of IS insertional elements at the *bgl* regulatory region

The same type of IS elements that are inserted at the same location (IS5d and IS5r at position -207.5) or 1 base pair apart (ISd and IS1r at position -230.5 and -231.5 respectively; IS2d and IS2r at position -206.5 and -205.5 respectively) at the *bgl* regulatory region, but in different orientations (direct and reverse orientations are indicated with “d” and “r”, respectively) were selected (**Figure 3.8A**), and their β-glucosidase (BglB) activities were measured. IS1 elements, inserted in opposite orientations, showed an over 4-fold difference in *bgl* operon

activity, while both IS2 and IS5 elements showed an over 2 times difference in activity when the same element was inserted in opposite orientations (**Figure 3.8BC**).

These results indicate that insertions of IS1, IS2 and IS5 all activated the *bgl* operon, leading to a Bgl^+ phenotype. However, the activation levels by IS insertions are not closely correlated with the type of IS element, the insertional positions, or the IS element orientations.

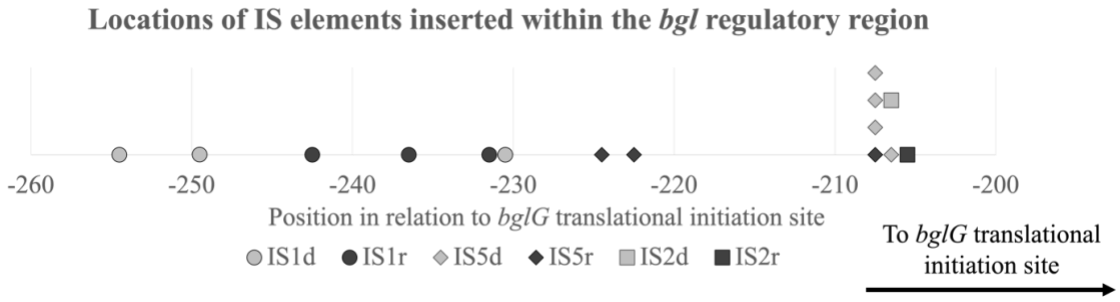


Figure 3.7 Locations of IS insertions in the *bgl* regulatory region. IS elements, including IS1, IS2 and IS5, were inserted within the *bgl* regulatory region in relation to the *bglG* translational initiation site.

A.

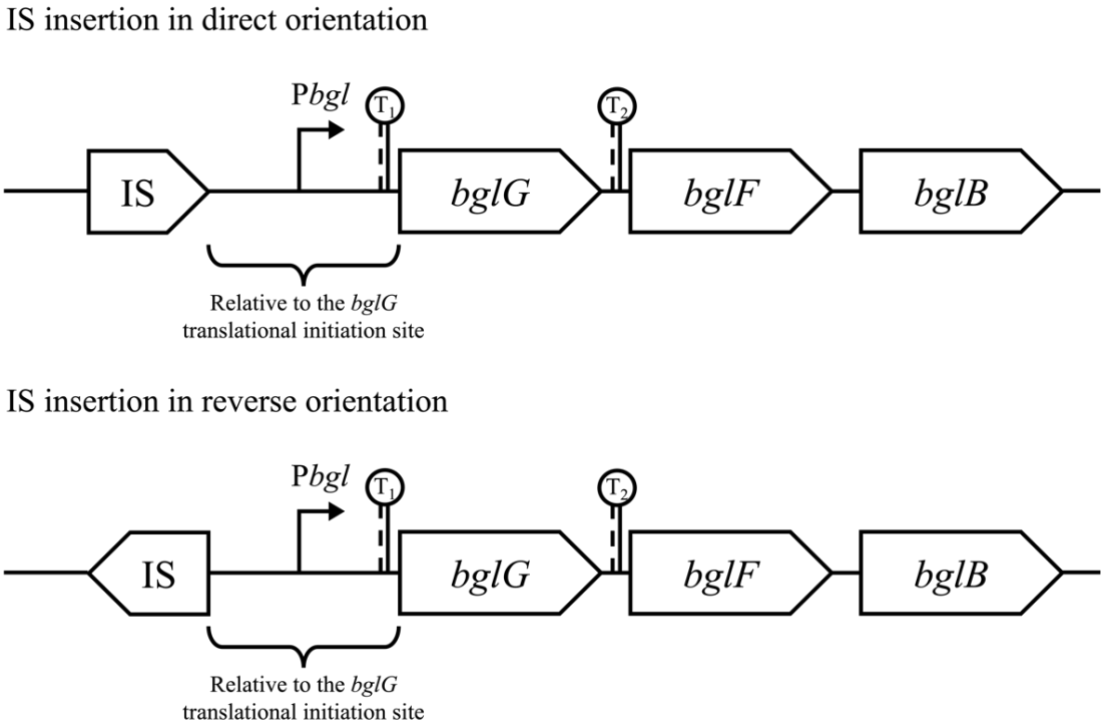
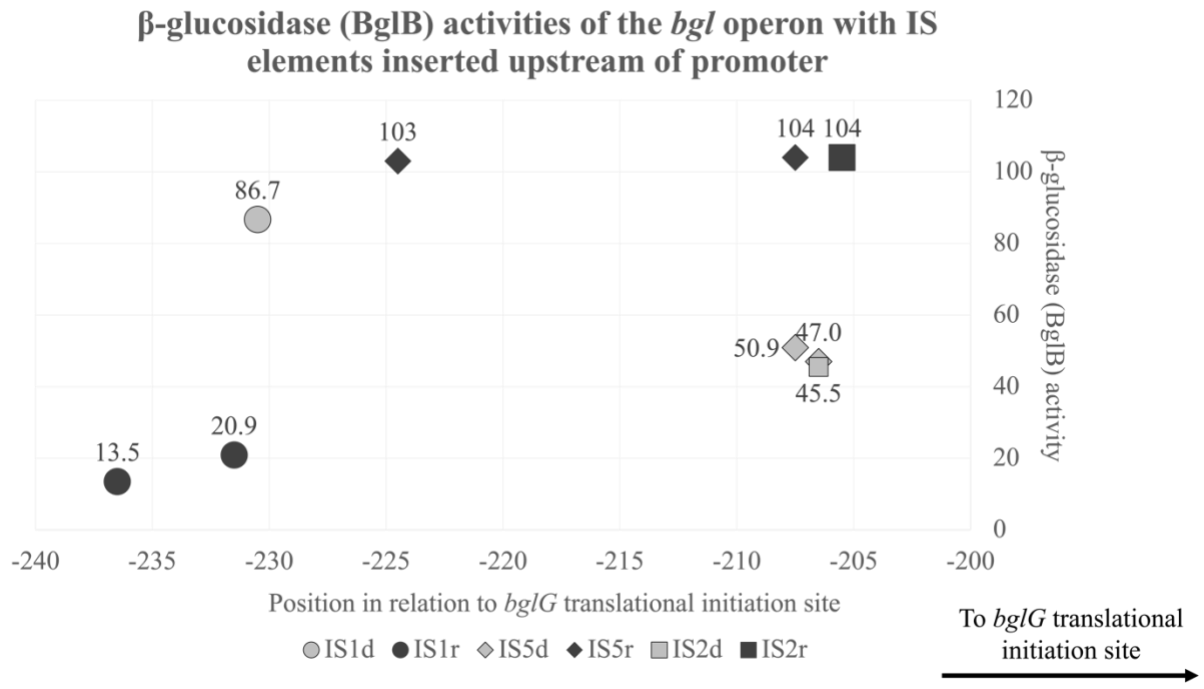


Figure 3.8 The operon activities with IS insertions within the *bgl* regulatory region. (**A**) Structure of IS elements inserted in different orientations within the *bgl* regulatory region. (**B**) The β-glucosidase (BglB) activities resulting from IS element insertions (IS1, IS2 and IS5), inserted in different orientations (direct or reverse). (**C**) Data from Figure 3.8B are presented for easier viewing.

B.



C.

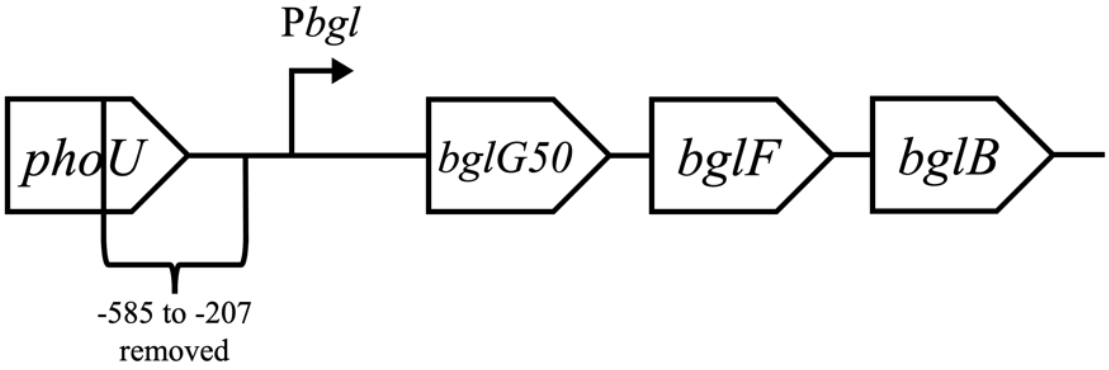
IS element	Orientation	Location (relative to *bglG* translational initiation site	β-glucosidase (BglB) activities
IS1	Direct	-230.5	87
	Reverse	-231.5	21
IS5	Direct	-207.5	51
	Reverse	-207.5	104
IS2	Direct	-206.5	45
	Reverse	-205.5	104

Figure 3.8 (continued) The operon activities with IS insertions within the *bgl* regulatory region. (**A**) Structure of IS elements inserted in different orientations within the *bgl* regulatory region. (**B**) The β-glucosidase (BglB) activities resulting from IS element insertions (IS1, IS2 and IS5), inserted in different orientations (direct or reverse). (**C**) Data from Figure 3.8B are presented for easier viewing.

3.4.2.4 Removal of the presumptive H-NS operator within the *bgl* regulatory region

To abolish H-NS binding to the *bgl* regulatory region completely, the region upstream of the Crp operator (-585 to -207 relative to the *bglG* translational start site), including the presumptive H-NS binding site, was removed from the regulatory region, yielding strain G50ΔO$_{HNS}$-P*bgl* (**Figure 3.9A**). The β-glucosidase (BglB) activity of G50ΔO$_{HNS}$-P*bgl* increased almost 90 times compared to G50 (**Figure 3.9B**). The strain also showed a Bgl^{+} phenotype.

A.



B.

Strains	β-glucosidase (BglB) activities	Bgl Phenotype	
		MacConkey + salicin	M9 + salicin
BW25113	0.08	-	-
G50ΔO$_{HNS}$-P*bgl*	22	+	+

Figure 3.9 The operon activity of G50ΔO$_{HNS}$-P*bgl*. (**A**) Structure of G50ΔO$_{HNS}$-P*bgl* (G50 with removal of the H-NS binding site within the *bgl* regulatory region), as well as (**B**) its β-glucosidase (BglB) activity.

3.4.3 Preventing or reducing H-NS binding to the downstream site within the *bglG* gene

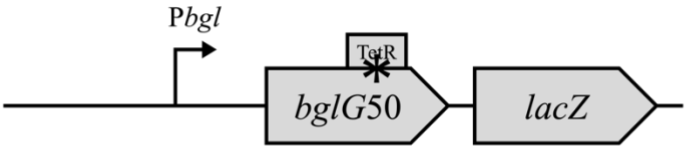
3.4.3.1 Replacing H-NS binding site within *bglG* with TetR operators

Previous studies have shown that the H-NS binding site is within positions +468 to +606, relative to the *bglG* translation initiation site (Dole *et al*., 2004). To reduce H-NS binding to this site, two new operon reporters were created by replacing part of the H-NS binding sites with either one or two TetR operators, yielding operon reporters G50.1R-Z and G50.2R-Z, respectively (**Figure 3.10Ai-ii**). Both operon reporters showed an increase in their β-galactosidase (LacZ) activities compared to G50-Z, with approximately 1.7 and 2 times increases in G50.1R-Z and G50.2R-Z, respectively (**Figure 3.10B**). In the presence of the TetR protein, their activities increased even more in both operon reporters, with an over 4.5 times increase, compared to G50-Z.

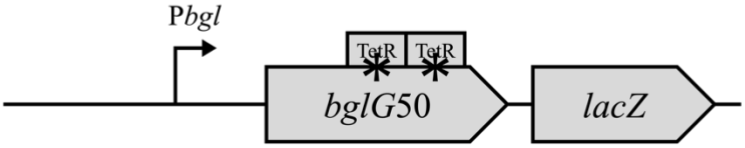
The above approach was also used to make the two new strains, G50.1R and G50.2R, with part of the H-NS binding sites within the *bglG* gene replaced by either one or two TetR operators, respectively (**Figure 3.10Aiii-iv**). Both strains showed similar β-glucosidase activities as G50 when the TetR protein was absent (**Figure 3.10C**). Consistent with the results from the β-galactosidase (LacZ) assays, both strains showed increases in activities in the presence of the TetR protein, even though the increases were not much compared to G50. On the other hand, both strains showed a Bgl^- phenotype whether or not the TetR proteins were present, as shown on MacConkey + salicin plates, as well as M9 + salicin plates.

A.

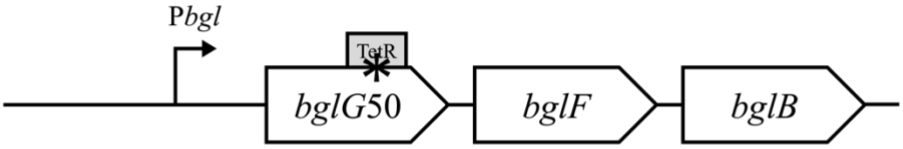
(i) G50.1R-Z



(ii) G50.2R-Z



(iii) G50.1R



(iv) G50.2R

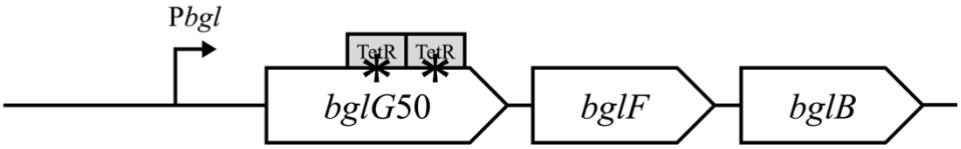
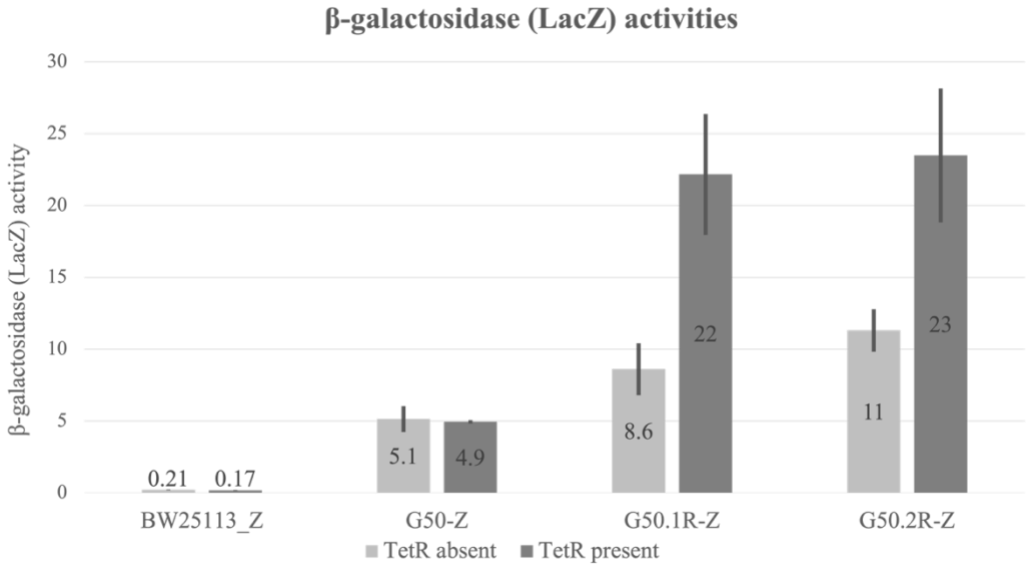


Figure 3.10 The operon activities resulting from the H-NS binding site within *bglG* replaced by TetR operators. (**A**) Structure of (**i**) operon reporter G50.1R (G50-Z with the H-NS binding site within *bglG* replaced by one TetR binding site), (**ii**) operon reporter G50.2R (G50-Z with H-NS binding site within *bglG* replaced by two TetR binding sites), (**iii**) strain G50.1R (G50 with the H-NS binding site within *bglG* replaced by one TetR binding site), and (**iv**) strain G50.2R (G50 with the H-NS binding site within *bglG* replaced by two TetR binding sites). (**B**) The β-galactosidase (LacZ) activities of the operon reporters G50.1R-Z and G50.2R-Z in the presence and absence of the TetR protein. (**C**) The β-glucosidase (BglB) activities of the strains G50.1R and G50.2R in the presence and absence of the TetR protein. Error bars represent standard deviations.

B.



C.

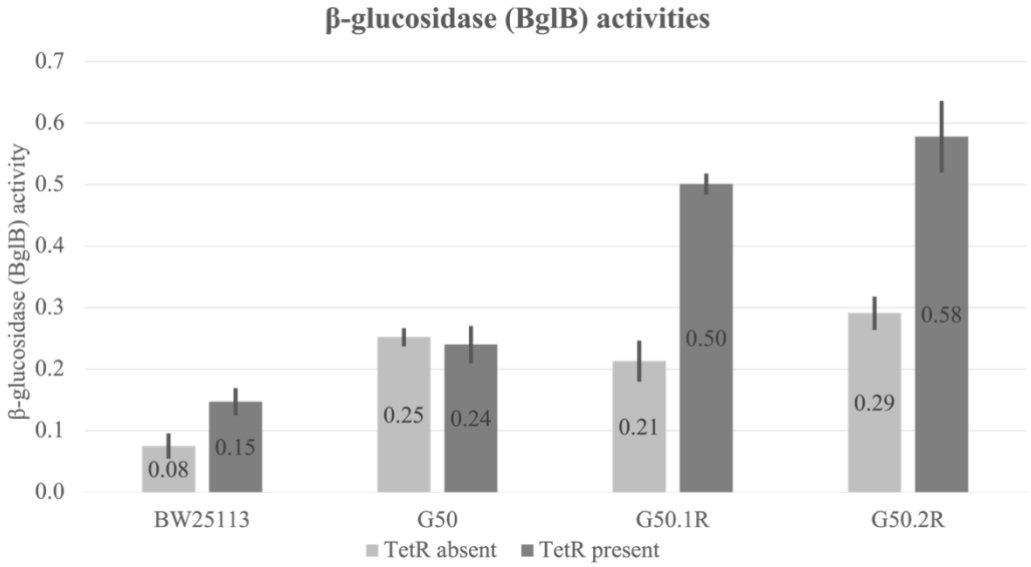
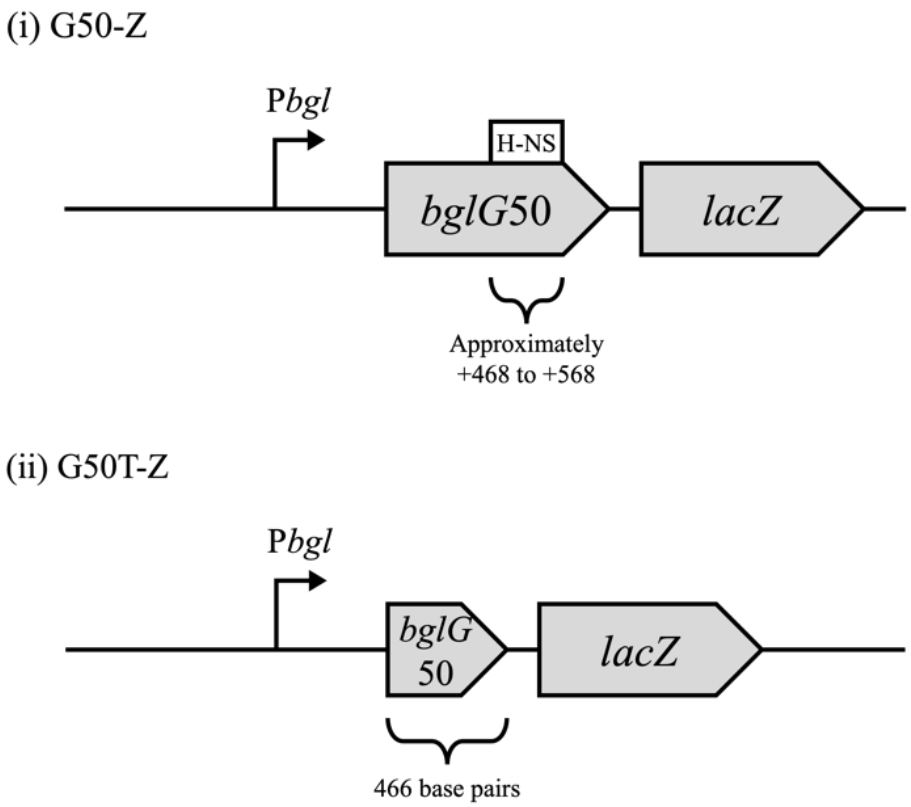


Figure 3.10 (continued) The operon activities resulting from the H-NS binding site within *bglG* replaced by TetR operators. (**A**) Structure of (**i**) operon reporter G50.1R (G50-Z with the H-NS binding site within *bglG* replaced by one TetR binding site), (**ii**) operon reporter G50.2R (G50-Z with H-NS binding site within *bglG* replaced by two TetR binding sites), (**iii**) strain G50.1R (G50 with the H-NS binding site within *bglG* replaced by one TetR binding site), and (**iv**) strain G50.2R (G50 with the H-NS binding site within *bglG* replaced by two TetR binding sites). (**B**) The β-galactosidase (LacZ) activities of the operon reporters G50.1R-Z and G50.2R-Z in the presence and absence of the TetR protein. (**C**) The β-glucosidase (BglB) activities of the strains G50.1R and G50.2R in the presence and absence of the TetR protein. Error bars represent standard deviations.

3.4.3.2 Removal of the H-NS binding site with a truncated *bglG* gene

As mentioned above, the presumptive H-NS binding site in *bglG* is within region +468 to +606 relative to the *bglG* translational initiation site. To further reduce H-NS binding to the *bglG* gene, a new operon reporter G50T-Z, with a modified *bglG50* deleted for the region starting from position +466 to +568 in relation to the *bglG* translational initiation site, was created (**Figure 3.11A**). The operon reporter showed more than a 5-fold increase in its β-galactosidase (LacZ) activity compared to the operon reporter, G50-Z (**Figure 3.11B**).

A.



B.

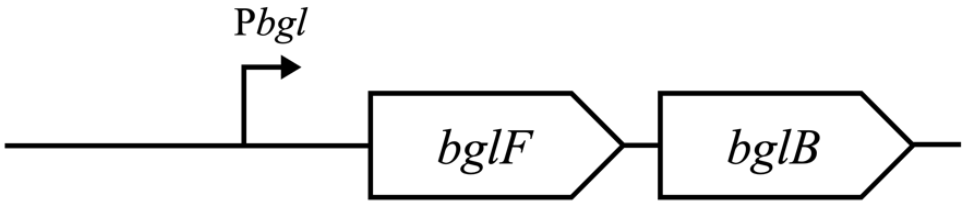
Operon reporter	β-galactosidase (LacZ) activities
G50-Z	5.1
G50T-Z	27

Figure 3.11 The operon activity of G50T-Z. (**A**) Structure of (**i**) operon reporter G50-Z and (**ii**) operon reporter G50T-Z (operon reporter G50-Z with truncated *bglG* starting from position +466 in relation to the *bglG* translational initiation site, where the H-NS binding site is located as present originally), as well as (**B**) their β-galactosidase (LacZ) activities.

3.4.3.3 Removal of *bglG* gene and the two *bgl* terminators

As the exact H-NS binding sites within the *bgl* operon are still unknown, the entire *bglG* gene and the two terminators were removed to completely prevent H-NS binding to *bglG* (Δ*bglG*T_1T_2) (**Figure 3.12A**). The β-glucosidase (BglB) activity of the new strain increased dramatically by 200-fold in comparison to G50, while showing a Bgl^+ phenotype (**Figure 3.12B**).

A.



B.

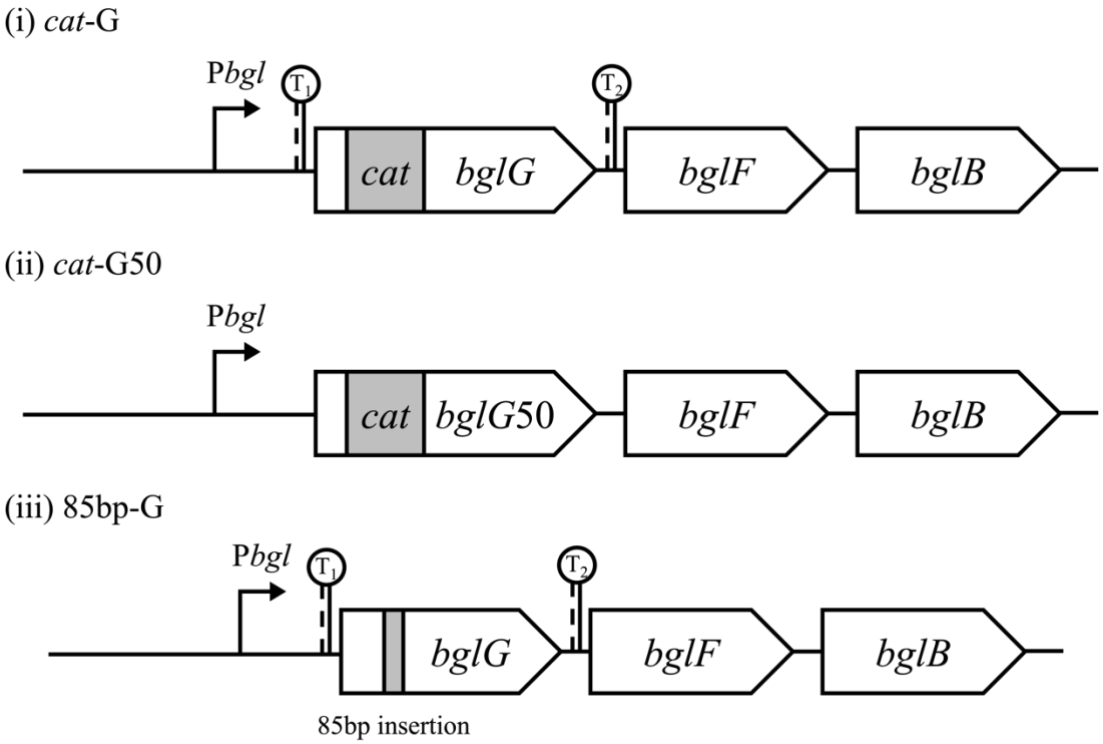
Strains	β-glucosidase (BglB) activities	Bgl Phenotype	
		MacConkey + salicin	M9 + salicin
BW25113	0.08	-	-
G50	0.25	-	-
Δ*bglG*T_1T_2	50	+	+++

Figure 3.12 The operon activity of Δ*bglG*T_1T_2. (**A**) Structure of the strain Δ*bglG*T_1T_2 (removal of the entire *bglG* gene and the two flanking terminators), as well as (**B**) its β-glucosidase (BglB) activity. The Bgl phenotype is also recorded.

3.4.4 Changing the distance between the two H-NS binding sites

To examine the flexibility of H-NS loop formation, the distance between the two H-NS binding sites at the *bgl* operon was changed by: (1) inserting a 1 kb *cat* gene within *bglG* in the wild-type, yielding strain *cat*-G, (2) inserting a 1 kb *cat* gene within *bglG* in G50, yielding strain *cat*-G50, and (3) removing *cat* from *cat*-G, leaving a 85 bp "scar" within *bglG*, yielding strain 85bp-G (**Figure 3.13A**). The β-glucosidase activities of the above three strains did not show a significant increase compared to the wild-type (**Figure 3.13B**). The strains also showed Bgl^- phenotypes on MacConkey + salicin plates as well as M9 + salicin plates. These results demonstrate that the insertions of 85 to 1 kb fragments do not affect strong repression of the *bgl* operon by H-NS, indicating the H-NS mediated DNA looping is robust.

A.



B.

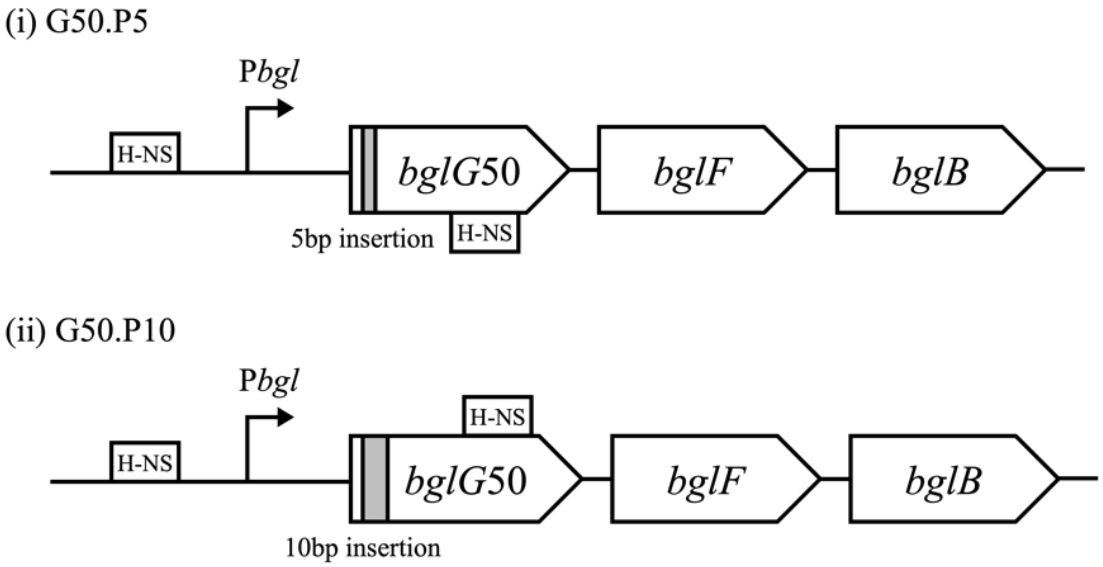
Strains	β-glucosidase (BglB) activities	Bgl Phenotype	
		MacConkey + salicin	M9 + salicin
BW25113	0.08	-	-
G50	0.25	-	-
cat-G	0.07	-	-
cat-G50	0.39	-	-
85bp-G	0.10	-	-

Figure 3.13 The operon activities with changes in distance between the H-NS binding sites. (**A**) Structures of (**i**) *cat*-G (1 kilobase pair of *cat* insertion within *bglG*), (**ii**) *cat*-G50 (1 kilobase pair of *cat* insertion within *bglG*, and removal of the two *bgl* terminators), and (**iii**) 85bp-G (85 base pair insertion within *bglG*), as well as (**B**) their β-glucosidase (BglB) activities and Bgl phenotypes.

3.4.5 Changing the phasing of the H-NS binding sites

The DNA double helix is around 10.5 base pairs per turn. Two small DNA fragments of 5 bp and 10 bp were individually inserted within the *bglG* gene between the two H-NS binding sites, yielding the two strains G50.P5 and G50.P10 (**Figure 3.14A**) (Levitt, 1978). The two H-NS binding sites appear to be on the same phase as in the wild type in strain G50.P10, but in strain G50.P5, these binding sites should be of the opposite phasing. Both strains showed similar β-glucosidase activities compared to G50, and they are still Bgl^- (**Figure 3.14B**). These results demonstrated that H-NS silencing of the *bgl* operon is DNA phase independent.

A.



B.

Strains	β-glucosidase (BglB) activities	Bgl Phenotype	
		MacConkey + salicin	M9 + salicin
G50	0.25	-	-
G50.P5	0.19	-	-
G50.P10	0.20	-	-

Figure 3.14 The operon activities with change in phasing of the H-NS binding sites. (**A**) Structure of (**i**) G50.P5 (a 5 base pair insertion within *bglG* to change the phase of the two H-NS binding sites within the bgl operon, with removal of the two *bgl* terminators), and (**ii**) G50.P10 (a 10 base pair insertion within *bglG*, and removal of the two *bgl* terminators), as well as (**B**) their β-glucosidase (BglB) activities and Bgl phenotypes.

4. Discussion

4.1 Defining the Bgl$^+$ phenotype

To test for the Bgl phenotype of an *E. coli* strain, some previous studies only used either MacConkey indicator plates with salicin or bromothymol blue indicator plates with salicin. Using MacConkey indicator plates with salicin as an example, a red color will be visible on colonies with a Bgl$^+$ phenotype. These indicator plates have been used in previous studies for mutagenesis screens, which involved the introduction of mutations in *E. coli* cells and subsequent screenings for the Bgl$^+$ phenotype (Giel *et al*., 1996; Venkatesh *et al*., 2010). This approach is effective in locating and identifying the genes responsible for the Bgl$^+$ phenotype. However, the use of these indicator plates does not represent an actual condition for the cells to grow using β-glucosides as the sole carbon source, as all cells are able to grow on the indicator plates. In fact, some strains that had been claimed to have a Bgl$^+$ phenotype in previous studies, were not able to grow on M9 minimal agar plates with salicin. These strains include those that overproduce BglJ alone, LeuO alone, and those that overproduce both BglJ and RcsB. Even though visible red color could be seen at the centers of their colonies on MacConkey agar plates with salicin, the appearance of the red colonies was not as strong as several other strains with the Bgl$^+$ phenotype, like mutants with IS5 insertions within the *bgl* regulatory region, and the results were not fully consistent in all trials. Phenotypic results on MacConkey indicator plates with salicin were sometimes inconclusive. As a result, strains considered to have a Bgl$^+$ phenotype in this study had to fulfill two requirements: (1) visible red color could be seen in colonies on MacConkey agar plates with salicin after 1 day of incubation at 37°C, and (2) growth on M9 agar plates with salicin after 2 days of incubation at 37°C.

4.2 Effect of Rho-independent terminators on the *bgl* operon

After the *bgl* operon is activated by either IS insertional mutations or non-insertional mutations, operon expression remains silent in the absence of a β-glucoside. This is due to the two terminators flanking the *bglG* gene (the first gene of the operon), which abolish transcription due to the formation of hairpin-like anti-terminator structures (Schnetz & Rak, 1988). To investigate the extent of the two terminators in regulating *bgl* operon expression, the two terminators were removed from the operon (G50). However, operon activity increased only slightly compared to wild-type, and it showed a Bgl^- phenotype (**Figure 3.2B**). When the *bglG* gene was overexpressed in strain G50, the cells still exhibited a minimal level of operon activity as well as a Bgl^- phenotype (Lam *et al*., unpublished data). This indicated that the two terminators flanking the two sides of *bglG* do not have an obvious effect in regulating the *bgl* operon in wild-type *E. coli* cells. Therefore, we concluded that the silent state of the *bgl* operon is not due to the presence of these two terminators, confirming that H-NS alone is sufficient to silence the operon.

4.3 Effect of the *bgl* repressors on *bgl* operon expression

H-NS is known to be the major repressor of the *bgl* operon, and StpA is a weak repressor of the *bgl* operon that can function as a DNA-binding adapter of H-NS if the H-NS oligomerization domain at the N-terminus of the protein is present (Free *et al*., 1998; Wolf *et al*., 2006). My results were consistent with previous findings, showing that StpA is a weak repressor of the *bgl* operon indicating that StpA does not exert an effect on *bgl* regulation when H-NS is present (**Figure 3.3B**). However, the results also showed that *bgl* operon expression increased two-fold, from 53 units in the absence of H-NS alone (Δ*hns*), to 108 units in the double mutant of Δ*hns*Δ*stpA* (**Figure 3.3B**). This indicates that StpA can repress the *bgl* operon expression to a moderate degree, even though H-NS is absent. This result is inconsistent with the suggestion that StpA regulates the *bgl* operon expression in the presence of the H-NS N-terminal protein domain (Wolf *et al*., 2006). As H-NS negatively regulates *stpA*, more StpA can be produced when H-NS is absent. These results suggest that in the absence of H-NS, StpA represses *bgl* operon expression either directly by binding to the *bgl* operon regulatory region or indirectly by affecting other factors involved in *bgl* operon regulation.

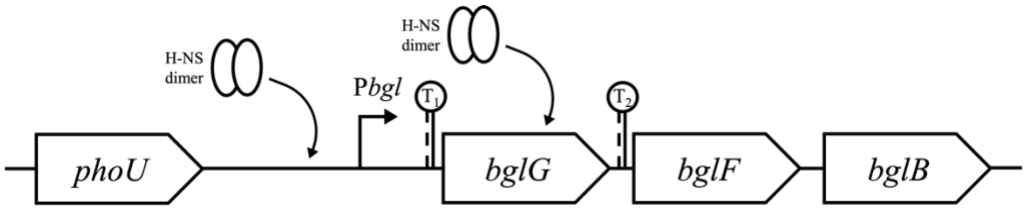
It is known that H-NS binds to both the *bgl* regulatory region and the 3' end of the *bglG* gene, and insertion of IS elements in the *bgl* regulatory region activates the *bgl* operon, leading to the Bgl^+ phenotype. It is conceivable that IS insertions could prevent bindings of H-NS and StpA to the *bgl* regulatory region. My data showed that with an IS5 insertion, the deletion of *stpA* did not further increase the *bgl* operon activity (**Figure 3.3B**), indicating that IS5 insertion completely prevents StpA binding to the *bgl* operon. On the other hand, with IS5 insertion, the deletion of *hns* further increased the *bgl* operon activity (**Figure 3.3B**). It is possible that H-NS is still able to bind weakly to the *bgl* regulatory region despite the IS5 insertion. In addition, since

IS insertions usually lead to higher operon activities than the deletion of the *hns* gene, it is likely that IS insertions create new promoters or attract new regulators, activating the *bgl* operon. The mechanism of IS5 activation of the *bgl* operon will be discussed further below.

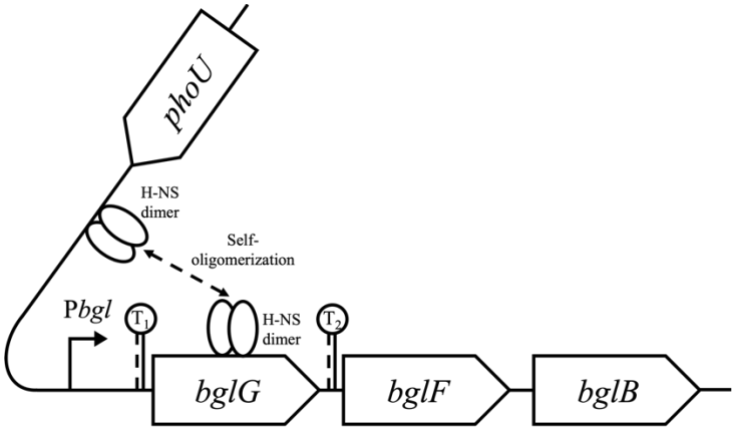
4.4 Formation of an H-NS-mediated DNA loop within the *bgl* operon

H-NS is characterized by self-dimerization and oligomerization via its N-terminal domain, leading to formation of higher-order oligomers, up to a 20-mer (Smyth *et al*., 2000; Ueguchi *et al*., 1997). The protein is able to bind to DNA using its C-terminal DNA-binding domain, and H-NS-mediated nucleoprotein complexes are considered to be important for gene silencing (Shindo *et al*., 1995). In relation to the *bgl* operon, H-NS is the major repressor which binds to the *bgl* regulatory region and at the 3' end of the *bglG* gene (Dole *et al*., 2004). It has also been found that H-NS binding to both locations is synergistic in repressing the *bgl* operon, as the repression is much high greater when H-NS binds to both locations than when it binds to either one (Nagarajavel *et al*., 2007). Based on the above information, it is possible that silencing of the *bgl* operon is due to the presence of an H-NS-mediated DNA looping that is formed via H-NS binding to the two locations within the *bgl* operon and bridging two duplexes together through H-NS self-oligomerization (**Figure 4.1**). This proposal has not been verified experimentally. To verify this mechanism, I examined formation of the H-NS-promoted DNA loop, focusing its effect on expression of the *bgl* operon, by weakening or abolishing the formation of the DNA loop.

A.



B.



C.

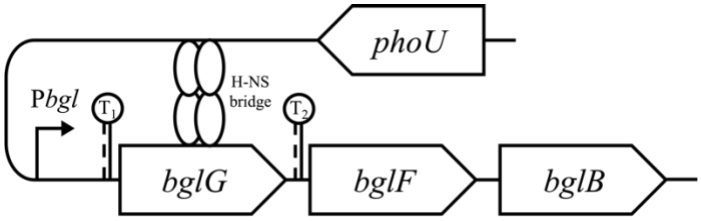


Figure 4.1 Proposed mechanism for repression of the *bgl* operon by H-NS. (**A**) H-NS dimers bind to the regulatory region of the *bgl* operon and the 3' end of the *bglG* gene. (**B**) H-NS oligomers are formed by self-oligomerization, bringing the two binding sites together. (**C**) An H-NS-mediated DNA loop is formed, which is largely responsible for the repression.

4.4.1 Effect of H-NS oligomerization on *bgl* operon expression

The change of leucine to proline at the 30th codon within the N-terminal dimerization domain in the *hns* gene causes H-NS to lose its ability in oligomerization (Ueguchi *et al*., 1997). It has been shown that H-NS with the L30P mutation is able to form dimers but not higher-order oligomers (Stella *et al*., 2005). Furthermore, this mutated portein retains its capacity for DNA binding, thereby repressing gene transcription (Stella *et al*., 2005). Recent work in our lab found that this mutated protein is even more inhibitory than the wild type H-NS when bound to the *bgl* promoter region, probably due to its greater binding affinity (Tran *et al*., manuscript in preparation). My data showed that *bgl* operon expression of strain HNSL30P, in which H-NS is unable to oligomerize, increased dramatically, from 0.08 units in wild-type cells to 44 units in the mutant genetic background, while showing a Bgl$^+$ phenotype (**Figure 3.4B**). Compared to the operon activity of 53 units for strain Δ*hns* (no H-NS protein present), HNSL30P showed a slightly lower operon expression level of 44 units. This could be due to the fact that the HNS_L30P protein partially inhibits transcription of the operon by binding to the regulatory region and terminates transcription by binding within the *bglG* gene (**Figure 3.3A**). These data suggest that (1) binding to two locations of the *bgl* operon alone is not sufficient to abolish *bgl* operon transcription, and (2) the formation of the H-NS loop via oligomerization is necessary and sufficient for the maximal level of transcriptional repression of the *bgl* operon.

4.4.2 H-NS binding to the *bgl* operon regulatory region

In addition to changing the H-NS oligomerization capability, another way to prevent or weaken the H-NS loop formation in the *bgl* operon is to alter H-NS binding to the two binding sites. To reduce H-NS binding to the *bgl* regulatory region, the following approaches were taken: (1) overproducing positive regulators of the *bgl* operon, including LeuO and the BglJ-RcsB heterodimers, (2) creating a strong CRP-cAMP binding site within the *bgl* regulatory region to outcompete H-NS binding, and (3) inserting IS elements into the *bgl* regulatory region (Madhusudan *et al*., 2005; Reynolds *et al*., 1986; Ueguchi *et al*., 1998; Venkatesh *et al*., 2010). Overall, the *bgl* operon transcriptional activity can be increased by reducing H-NS binding to the *bgl* regulatory region, even though not all of the strains thus obtained were able to grow on aromatic β-glucosides.

4.4.2.1 Overproduction of positive regulators of the *bgl* operon

The overexpression of positive regulatory genes of the *bgl* operon, including *leuO* and *bglJ,* showed increases in *bgl* operon activities compared to the control strain (**Figure 3.5CD**). This is consistent with the notion that reducing H-NS binding to the promoter region would (partially) eliminate DNA looping, thereby enhancing operon transcription. As BglJ and RcsB form heterodimers when regulating the *bgl* operon (Venkatesh *et al*., 2010), BglJ might be the limiting factor in the regulation of the operon with RcsB. On the other hand, even though the operon activity increases when both *bglJ* and *rcsB* were overexpressed, the increase was not as much as that when *bglJ* was overexpressed alone (**Figure 3.5CD**). Similar trends were obtained when using both β-galactosidase (LacZ) and β-glucosidase (BglB) assays, indicating the validity of the data. At the same time, a growth defect has been observed in strains with operon reporters when *rcsB* was overexpressed. As RcsB is a global regulator, overexpression of *rcsB* might cause pleiotropic effects on cell growth and *bgl* operon transcription. Even though the overexpression of positive regulatory genes increases the *bgl* operon activity, all strains showed a Bgl^- phenotype (**Figure 3.5E**). It is possible that H-NS is only partially outcompeted by the positive regulators and is still able to bind to the *bgl* regulatory region to form a H-NS-mediated DNA loop, causing repression of *bgl* operon transcription.

4.4.2.2 Stronger Crp-cAMP binding to the *bgl* regulatory region

A point mutation at the Crp-cAMP operator creates stronger Crp-cAMP binding to the *bgl* regulatory region (Reynolds *et al*., 1986). The greater binding of Crp-cAMP increased *bgl* operon expression significantly as shown when using G51-Z (**Figure 3.6B**). It was confirmed that Crp-cAMP binding to the location of the point mutation in the *bgl* regulatory region caused a greater increase in operon transcription, as compared with the deletion of *crp* in G51-Z (Δ*crp*G51-Z). This strain showed low operon expression similar to that in the control operon reporter, G50-Z. Despite the substantial increase in operon activity when Crp-cAMP binds more strongly to the *bgl* regulatory region, the strain, G51, still showed a Bgl^- phenotype. Similar to the result obtained when positive regulatory genes of the *bgl* operon were overexpressed, a stronger Crp-cAMP binding to the *bgl* regulatory region may only be able to partially outcompete H-NS binding. Therefore, H-NS can still bind to the *bgl* regulatory region so that an H-NS loop can be formed to repress the *bgl* operon.

4.4.2.3 IS insertions block H-NS binding to the *bgl* regulatory region

It has been well established that IS elements inserted within the *bgl* regulatory region activate the operon and give rise to the Bgl^+ phenotype. Most IS elements are inserted within the superhelix stress-induced DNA duplex destabilization (SIDD) region upstream of the *bgl* promoter., in which this region can open up into single-stranded DNA loops (Humayun *et al*., 2017). My data were consistent with previous studies that all of the IS insertional mutants examined in this study had insertions within the SIDD region. The IS elements were inserted between positions -255 to -205 relative to the *bglG* translational initiation site (**Figure 3.7**), while the SIDD region is within positions -304 to -165 (Humayun *et al*., 2017). Among the IS element insertions, IS1 elements tend to insert at locations further away from the *bgl* promoter, followed by IS5 elements, while IS2 elements are inserted more closely to the *bgl* promoter (**Figure 3.7**). Despite this observation, more data will be needed to verify that there is tendency for each type of IS element to insert at certain locations within the *bgl* operon regulatory region.

The orientation of the IS elements seems to play a role in influencing expression of the *bgl* operon. For the same type of elements, the insertions in different orientations showed two-fold (IS2 and IS5) or even four-fold (IS1) differences (**Figure 3.8BC**). In addition, trends concerning *bgl* operon expression can be seen among IS elements inserted in direct or reverse orientations (**Figure 3.8B**). The *bgl* operon transcription appeared to be greater when the IS elements were inserted closer to the promoter when they were in the reverse orientation. In contrast, the activity was higher when the IS elements were inserted further away from the promoter when they were in the direct orientation. This observation needs to be verified with more data, as other factors need to be considered, such as the type of IS elements. Nevertheless, the *bgl* operon expression appears to be dependent on the type, location and orientation of the IS

element inserted in the regulatory region. The activation of the *bgl* operon by IS element insertions might be due to (1) the abolition of H-NS binding to the regulatory region, preventing formation of the H-NS-induced DNA loop, (2) generation of a new promoter, allowing transcription of the operon, and/or (3) new regulators, binding to the *bgl* operon control region.

4.4.3 H-NS binding to the *bgl*G gene

An alternative way to investigate the formation of the H-NS loop within the *bgl* operon is to reduce H-NS binding to the *bglG* gene. It has been found that H-NS binds to a region approximately 600 to 700 base pairs downstream of the *bgl* promoter, which is around positions +479 to +606 relative to the *bglG* translational initiation site (Dole *et al*., 2004). First, the H-NS binding site within the *bglG* gene was replaced by one or two TetR operators, which showed increases in *bgl* operon activities when compared to wild-type (**Figure 3.10B**). The *bgl* operon activities increased even more when TetR was present. This showed that H-NS binding to the *bglG* gene can be reduced. However, despite the presence of TetR, all of the strains with one or two TetR operators replacing the H-NS binding site within the *bglG* gene showed a Bgl^- phenotype. It is possible that H-NS binding can only be reduced partially when TetR is absent, as the exact H-NS binding site is still unknown. On the other hand, in the presence of TetR, even though H-NS binding to *bglG* can be reduced further, TetR proteins might act as roadblocks, preventing transcription of the *bgl* operon, leading to the Bgl^- phenotype.

4.4.4 Removal of H-NS binding individually to the two binding sites

Considering that the above approaches might not have been effective in preventing H-NS binding to either *bgl* operon sites (one in the regulatory region and one within the *bglG* gene), each presumptive H-NS binding site was removed (G50ΔO_{HNS}-P*bgl* and Δ*bglG*T_1T_2, respectively). The removal of presumptive H-NS binding sites in either of these two regions led to >100-fold increase in *bgl* operon activities as well as the Bgl^+ phenotypes (**Figure 3.9 3.12**). This fact indicates that H-NS binding to both locations is necessary to fully repress the *bgl* operon, which is consistent with previous studies, suggesting that H-NS represses *bgl* operon expression by binding to both locations (Nagarajavel *et al*., 2007). Without H-NS binding to either location, the H-NS loop cannot be formed to repress the *bgl* operon, causing an increase in operon expression and the Bgl^+ phenotype.

4.5 Characterizing the H-NS-promoted DNA loop within the *bgl* operon

To characterize the H-NS-mediated DNA loop in the *bgl* operon, the flexibility of the distance between the two H-NS binding sites was first examined. The results showed that there was no influence on the *bgl* operon activity, irrespective of whether 85 bp or 1 kb was inserted between the two H-NS binding sites (**Figure 3.13B**). This indicates that the formation of the H-NS loop is flexible with respect to loop size.

In addition, the effect of phasing of the H-NS binding sites on the H-NS loop was examined. As the DNA helix is around 10.5 base pairs per turn, 5 base pairs were inserted within the *bglG* gene to increase the distance between the two H-NS binding sites in the *bgl* operon by half a turn (G50.P5) (**Figure 3.14Ai**) (Levitt, 1978). The H-NS binding sites would be positioned on opposite sides of the DNA helix if the original binding sites are on the same side, or vice versa. It is important to note that the exact locations and phasing of the H-NS binding sites are still unknown. My data suggest that the *bgl* operon cannot be activated by a phase change between the two H-NS binding sites within the *bgl* operon by a half (G50.P5) or a full turn (G50.P10) (**Figure 3.14B**). This shows that changing the phasing or size of the two H-NS binding sites has no influence on the *bgl* operon activity. The formation of the H-NS loop at the *bgl* operon is flexible and independent of the phasing of the H-NS binding sites on the DNA helix.

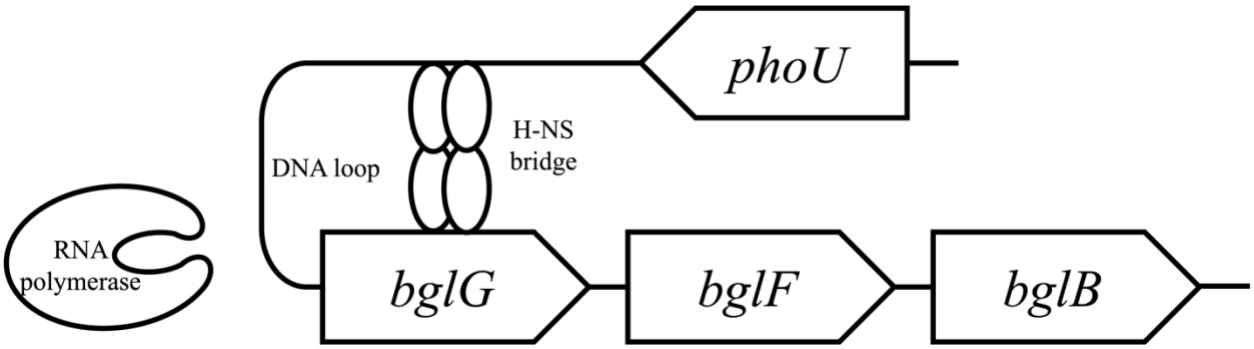
4.6 Silencing of the *bgl* operon by H-NS-mediated DNA looping

Overall, it seems like the *bgl* operon can be activated or at least shows an increase in expression when the formation of the H-NS-mediated DNA loop is weakened by abolishing the ability of H-NS to bind to the operon or to self-oligomerize. All of my results support the proposed mechanism that the formation of an H-NS-mediated DNA loop represses the *bgl* operon (**Figure 4.1**). H-NS binds to the *bgl* regulatory region and to a region within the *bglG* gene, and the subsequent self-oligomerization of H-NS proteins brings the two binding sites together, forming an H-NS bridge and a DNA loop. The H-NS-mediated DNA loop efficiently prevents the transcription of the *bgl* operon by (1) preventing the binding of RNA polymerase to the *bgl* operon (**Figure 4.2A**), or (2) trapping RNA polymerase within the two H-NS binding sites (**Figure 4.2B**) (Dame *et al*., 2002).

Repression of the *bgl* operon by H-NS-mediated DNA looping is so strong that other factors, such as other DNA-binding proteins and the two terminators flanking the *bglG* gene, do not contribute much to the operon regulation prior to its activation. For example, the DNA loop prevents the transcription of the *bgl* operon before the formation of the hairpin-like termination structures. Also, StpA, a repressor of the *bgl* operon, only plays a role in repressing the *bgl* operon when H-NS is absent (**Figure 3.3**). In parallel, positive regulators of the *bgl* operon, such as LeuO, BglJ-RcsB heterodimers and Crp-cAMP, would be expected to exert an effect to upregulate the operon only when the DNA loop is absent. In addition, the formation of the H-NS-mediated DNA loop within the *bgl* operon is robust as the loop can hardly be removed. The changing of the distance and the phasing of the two H-NS binding sites does not affect the formation of the DNA loop (**Figure 3.13 3.14**), indicating that its formation is independent of distance and DNA phase.

In conclusion, I have provided evidence that the silencing of the *bgl* operon by H-NS binding alone is not sufficient to silence the operon, but is instead dictated by an H-NS-mediated DNA looping mechanism. The two terminators flanking the *bglG* gene and the other repressor, StpA, only play a minor role in silencing the *bgl* operon. The known activating mutations, such as IS insertions, overexpression or DNA binding of other regulators, activate the *bgl* operon most likely by breaking or partially breaking the DNA loop. The looping structure is robust and is independent of the distance and the DNA phasing. Finally, the formation of H-NS mediated looping efficiently prevents or abolishes transcription by either excluding Crp-cAMP and RNA polymerase from the promoter region or trapping the polymerase on the promoter.

A.



B.

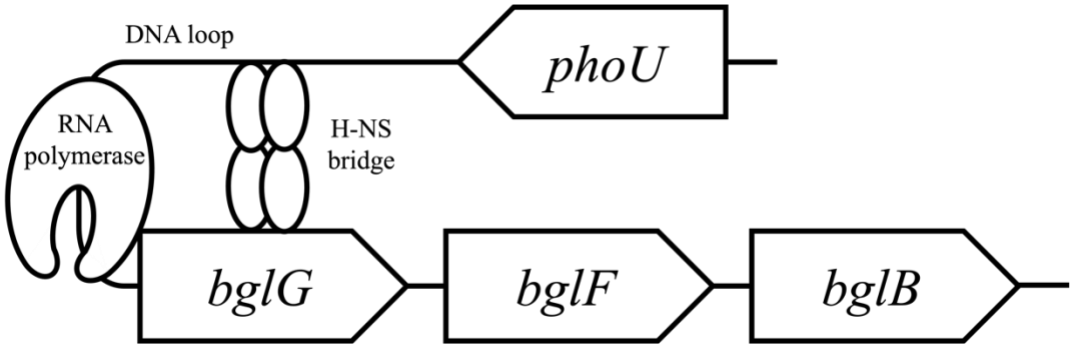


Figure 4.2 A model for the H-NS-mediated DNA looping, preventing transcription of the *bgl* operon. A DNA loop is formed with the bridging of H-NS between its two binding sites, first, within the *bgl* regulatory region and second, within the *bglG* gene. The *bgl* promoter and the two terminators flanking the *bglG* gene are not shown here for simplification. (**A**) Formation of the H-NS-mediated DNA loop prevents RNA polymerase from binding to the *bgl* operon, thus preventing transcription of the operon. (**B**) Formation of the H-NS-mediated DNA loop traps RNA polymerase between the two H-NS binding sites, thus preventing transcription of the operon.

5. References

Chen, Q., Arents, J. C., Bader, R., Postma, P. W., & Amster-Choder, O. (1997). 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*, *16*(15), 4617–4627. https://doi.org/10.1093/emboj/16.15.4617

Dame, R. T., Noom, M. C., & Wuite, G. J. (2006). Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature*, 444(7117), 387–390. https://doi.org/10.1038/nature05283

Dame, R. T., Wyman, C., Wurm, R., Wagner, R., & Goosen, N. (2002). Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at the rrnB P1. *The Journal of biological chemistry*, *277*(3), 2146–2150. https://doi.org/10.1074/jbc.C100603200

Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12), 6640–6645. https://doi.org/10.1073/pnas.120163297

Dole, S., Klingen, Y., Nagarajavel, V., & Schnetz, K. (2004). The protease Lon and the RNA-binding protein Hfq reduce silencing of the Escherichia coli bgl operon by H-NS. *Journal of bacteriology*, 186(9), 2708–2716. https://doi.org/10.1128/JB.186.9.2708-2716.2004

Dole, S., Kühn, S., & Schnetz, K. (2002). Post-transcriptional enhancement of Escherichia coli bgl operon silencing by limitation of BglG-mediated antitermination at low transcription rates. *Molecular microbiology*, 43(1), 217–226. https://doi.org/10.1046/j.1365-2958.2002.02734.x

Dole, S., Nagarajavel, V., & Schnetz, K. (2004). The histone-like nucleoid structuring protein H-NS represses the Escherichia coli bgl operon downstream of the promoter. *Molecular microbiology*, *52*(2), 589–600. https://doi.org/10.1111/j.1365-2958.2004.04001.x

Free, A., Williams, R. M., & Dorman, C. J. (1998). The StpA protein functions as a molecular adapter to mediate repression of the bgl operon by truncated H-NS in Escherichia coli. *Journal of bacteriology*, *180*(4), 994–997. https://doi.org/10.1128/JB.180.4.994-997.1998

Giel, M., Desnoyer, M., & Lopilato, J. (1996). A mutation in a new gene, bglJ, activates the bgl operon in Escherichia coli K-12. *Genetics*, *143*(2), 627–635. https://doi.org/10.1093/genetics/143.2.627

Gleadow, R. M., & Møller, B. L. (2014). Cyanogenic glycosides: synthesis, physiology, and phenotypic plasticity. *Annual review of plant biology*, 65, 155–185. https://doi.org/10.1146/annurev-arplant-050213-040027

Görke, B., & Rak, B. (1999). Catabolite control of Escherichia coli regulatory protein BglG activity by antagonistically acting phosphorylations. *The EMBO journal*, *18*(12), 3370–3379. https://doi.org/10.1093/emboj/18.12.3370

Hall, B. G., Yokoyama, S., & Calhoun, D. H. (1983). Role of cryptic genes in microbial evolution. *Molecular biology and evolution*, 1(1), 109–124. https://doi.org/10.1093/oxfordjournals.molbev.a040300

Humayun, M. Z., Zhang, Z., Butcher, A. M., Moshayedi, A., & Saier, M. H., Jr (2017). Hopping into a hot seat: Role of DNA structural features on IS5-mediated gene activation and inactivation under stress. *PloS one*, *12*(6), e0180156. https://doi.org/10.1371/journal.pone.0180156

Klumpp, S., Zhang, Z., & Hwa, T. (2009). Growth rate-dependent global effects on gene expression in bacteria. *Cell*, 139(7), 1366–1375. https://doi.org/10.1016/j.cell.2009.12.001

Lang, B., Blot, N., Bouffartigues, E., Buckle, M., Geertz, M., Gualerzi, C. O., Mavathur, R., Muskhelishvili, G., Pon, C. L., Rimsky, S., Stella, S., Babu, M. M., & Travers, A. (2007). High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic acids research*, 35(18), 6330–6337. https://doi.org/10.1093/nar/gkm712

Levine, E., Zhang, Z., Kuhlman, T., & Hwa, T. (2007). Quantitative characteristics of gene regulation by small RNA. *PLoS biology*, 5(9), e229. https://doi.org/10.1371/journal.pbio.0050229

Levitt M. (1978). How many base-pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations. *Proceedings of the National Academy of Sciences of the United States of America*, 75(2), 640–644. https://doi.org/10.1073/pnas.75.2.640

Li, X. T., Thomason, L. C., Sawitzke, J. A., Costantino, N., & Court, D. L. (2013). Positive and negative selection using the tetA-sacB cassette: recombineering and P1 transduction in Escherichia coli. *Nucleic acids research*, 41(22), e204. https://doi.org/10.1093/nar/gkt1075

Lutz, R., & Bujard, H. (1997). 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 research*, 25(6), 1203–1210. https://doi.org/10.1093/nar/25.6.1203

Madan, R., Kolter, R., & Mahadevan, S. (2005). Mutations that activate the silent bgl operon of Escherichia coli confer a growth advantage in stationary phase. *Journal of bacteriology*, 187(23), 7912–7917. https://doi.org/10.1128/JB.187.23.7912-7917.2005

Madhusudan, S., Paukner, A., Klingen, Y., & Schnetz, K. (2005). Independent regulation of H-NS-mediated silencing of the bgl operon at two levels: upstream by BglJ and LeuO and downstream by DnaKJ. *Microbiology (Reading, England)*, *151*(Pt 10), 3349–3359. https://doi.org/10.1099/mic.0.28080-0

Miller, J. H. (1972). *Experiments in molecular genetics*. Cold Spring Harbor Laboratory Press.

Nagarajavel, V., Madhusudan, S., Dole, S., Rahmouni, A. R., & Schnetz, K. (2007). Repression by binding of H-NS within the transcription unit. *The Journal of biological chemistry*, 282(32), 23622–23630. https://doi.org/10.1074/jbc.M702753200

Reynolds, A. E., Mahadevan, S., LeGrice, S. F., & Wright, A. (1986). Enhancement of bacterial gene expression by insertion elements or by mutation in a CAP-cAMP binding site. *Journal of molecular biology*, *191*(1), 85–95. https://doi.org/10.1016/0022-2836(86)90424-9

Schaefler S. (1967). Inducible system for the utilization of beta-glucosides in Escherichia coli. I. Active transport and utilization of beta-glucosides. *Journal of bacteriology*, *93*(1), 254–263. https://doi.org/10.1128/jb.93.1.254-263.1967

Schnetz, K., & Rak, B. (1988). Regulation of the bgl operon of Escherichia coli by transcriptional antitermination. *The EMBO journal*, *7*(10), 3271–3277.

Schnetz, K., Toloczyki, C., & Rak, B. (1987). Beta-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*, *169*(6), 2579–2590. https://doi.org/10.1128/jb.169.6.2579-2590.1987

Shindo, H., Iwaki, T., Ieda, R., Kurumizaka, H., Ueguchi, C., Mizuno, T., Morikawa, S., Nakamura, H., & Kuboniwa, H. (1995). Solution structure of the DNA binding domain of a nucleoid-associated protein, H-NS, from Escherichia coli. *FEBS letters*, 360(2), 125–131. https://doi.org/10.1016/0014-5793(95)00079-o

Smyth, C. P., Lundbäck, T., Renzoni, D., Siligardi, G., Beavil, R., Layton, M., Sidebotham, J. M., Hinton, J. C., Driscoll, P. C., Higgins, C. F., & Ladbury, J. E. (2000). Oligomerization of the chromatin-structuring protein H-NS. *Molecular microbiology*, 36(4), 962–972. https://doi.org/10.1046/j.1365-2958.2000.01917.x

Stella, S., Spurio, R., Falconi, M., Pon, C. L., & Gualerzi, C. O. (2005). Nature and mechanism of the in vivo oligomerization of nucleoid protein H-NS. *The EMBO journal*, 24(16), 2896–2905. https://doi.org/10.1038/sj.emboj.7600754

Suzuki-Minakuchi, C., Kawazuma, K., Matsuzawa, J., Vasileva, D., Fujimoto, Z., Terada, T., Okada, K., & Nojiri, H. (2016). Structural similarities and differences in H-NS family proteins revealed by the N-terminal structure of TurB in Pseudomonas putida KT2440. *FEBS letters*, 590(20), 3583–3594. https://doi.org/10.1002/1873-3468.12425

Thomason, L. C., Costantino, N., & Court, D. L. (2007). E. coli genome manipulation by P1 transduction. *Current protocols in molecular biology*, Chapter 1, 1.17.1–1.17.8. https://doi.org/10.1002/0471142727.mb0117s79

Tran, D., Zhang, Z., & Saier, M. H., Jr (2022). The effects of global and specific DNA-binding proteins on transcriptional regulation of the bgl operon. Manuscript in preparation.

Ueguchi, C., Ohta, T., Seto, C., Suzuki, T., & Mizuno, T. (1998). The leuO gene product has a latent ability to relieve bgl silencing in Escherichia coli. *Journal of bacteriology*, *180*(1), 190–193. https://doi.org/10.1128/JB.180.1.190-193.1998

Ueguchi, C., Seto, C., Suzuki, T., & Mizuno, T. (1997). Clarification of the dimerization domain and its functional significance for the Escherichia coli nucleoid protein H-NS. *Journal of molecular biology*, 274(2), 145–151. https://doi.org/10.1006/jmbi.1997.1381

Venkatesh, G. R., Kembou Koungni, F. C., Paukner, A., Stratmann, T., Blissenbach, B., & Schnetz, K. (2010). BglJ-RcsB heterodimers relieve repression of the Escherichia coli bgl operon by H-NS. *Journal of bacteriology*, *192*(24), 6456–6464. https://doi.org/10.1128/JB.00807-10

Wang, W., Li, G. W., Chen, C., Xie, X. S., & Zhuang, X. (2011). Chromosome organization by a nucleoid-associated protein in live bacteria. *Science (New York, N.Y.)*, 333(6048), 1445–1449. https://doi.org/10.1126/science.1204697

Wolf, T., Janzen, W., Blum, C., & Schnetz, K. (2006). Differential dependence of StpA on H-NS in autoregulation of stpA and in regulation of bgl. *Journal of bacteriology*, *188*(19), 6728–6738. https://doi.org/10.1128/JB.00586-06

Zhang, Z., Zhou, K., Tran, D., & Saier, M. (2022). Insertion Sequence (IS) Element-Mediated Activating Mutations of the Cryptic Aromatic β-Glucoside Utilization (BglGFB) Operon Are Promoted by the Anti-Terminator Protein (BglG) in Escherichia coli. *International journal of molecular sciences*, 23(3), 1505. https://doi.org/10.3390/ijms23031505

Zhang, Z., & Saier, M. H., Jr (2009). A mechanism of transposon-mediated directed mutation. *Molecular microbiology*, 74(1), 29–43. https://doi.org/10.1111/j.1365-2958.2009.06831.x