Volume 10:1-10

UJEMI+

The Undergraduate Journal of Experimental Microbiology & Immunology (+Peer Reviewed)

Promoter Mapping Analysis Using a GFP Reporter Plasmid Suggests that the *lac* Promoter May Drive *brka* Expression on Plasmid pDO6935 in *Escherichia coli*

Nidhin Biju, Sharisse Chan, Dahyeon Hong, Renz Po

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY Autotransporters (AT) are a class of bacterial proteins that play a crucial role in bacterial virulence and pathogenesis. Understanding these proteins is essential to advance research in infectious disease and identify possible targets for therapeutic interventions. *Bordetella pertussis* autotransporter, BrkA belongs to the AT-1 subfamily of autotransporter proteins in Gram negative bacteria. Since its discovery, plasmids have been created to study *brkA* expression in various systems, pDO6935 being one of them. The mechanism by which *brkA* is expressed in pDO6935 is yet to be fully elucidated. We used a promoter trapping method to map the promoter regions that drive *brkA* expression. An analysis of pDO6935 gene sequence revealed the presence of a *lac* operon upstream of *brkA*. The *lac* operon is known to exhibit basal expression in the absence of lactose or the presence of a repressor. We studied whether the leaky expression caused by the *lac* operon may result in the expression of *brkA* in an *Escherichia coli* system. In our study, we utilized pSPPH21, a promoterless reporter plasmid that was designed to contain green fluorescent protein (GFP). Using this vector, we inserted regions of putative promoter sequences upstream of *brkA* in pDO6935. Fluorescence imaging and quantification on a plate reader suggested that the *lac* operon may be driving *brkA* expression in pDO6935. We further investigated this argument upon treatment with glucose, a known catabolite repressor of the *lac* operon. Glucose treatment resulted in repressed levels of GFP, providing additional evidence to our findings.

INTRODUCTION

Bordetella resistance to killing gene A (*brkA*) is a *Bordetella pertussis* autotransporter that belongs to the autotransporter-1 (AT-1) subfamily of autotransporter proteins in Gram-negative bacteria (1). BrkA has been shown to play a role in bacterial adhesion and virulence by conferring resistance to complement-mediated immunity (2). Since its discovery by Fernandez et al. (1), several plasmids have been created and used to transform *brkA* into *Escherichia coli*, including pDO6935 (3). While the sensor kinase system *Bordetella* virulence gene (Bvg) is believed to regulate *brkA* expression in *B. pertussis* (4), the specific promoter region responsible for driving *brkA* transcription in *E. coli* has not been identified. An analysis of the pDO6935 gene sequence using the SnapGene program revealed the presence of a *lac* operon upstream of *brkA*. One fundamental principle about the *lac* operon is that it has a "leaky" expression, meaning that the transcriptional control is not 100% efficient and exhibits basal expression in *E. coli* (5). This basal expression could potentially turn on *brkA*, which is located downstream of the *lac* operon. These observations suggest that the promoter region that may drive *brkA* expression in pDO6935 is likely the same promoter as one used by *lac*. As such, we hypothesize that the *lac* operon is causing transcription of

Published Online: September 2024

Citation: Biju, Chan, Hong, Po. 2024. Promoter mapping analysis using a GFP reporter plasmid suggests that the *lac* promoter may drive *brkA* expression on plasmid pDO6935 in *Escherichia coli*. UJEMI+ 10:1-10

Editor: Ronja Kothe, University of British Columbia

Address correspondence to:
https://jemi.microbiology.ubc.ca/

brkA in pDO6935. By systematically testing the sequences upstream of *brkA*, we expect to map the region that contains the promoter and initiates transcription of the *brkA* expression.

Here, we used the substrate-induced gene expression (SIGEX) reporter plasmid pSPPH21 created by the 2022 UBC International Genetically Engineered Machine (iGEM) team. This reporter plasmid contains various multiple cloning sites upstream of a promoterless green fluorescent protein (GFP) gene, to identify the sequence causing *brkA* transcription (6). We created five amplicons from pDO6935, all of which contain sequences upstream of its *brkA* gene and downstream of the origin of chromosomal replication (oriC), using primers that add the restriction sites EcoRI or XbaI to either end. We individually ligated these amplicons into pSPPH21 upstream of its promoterless GFP and transformed the cloned plasmids into DH5α *E. coli*. GFP gene expression levels were quantified and assessed to investigate regions that may be driving *brkA* expression in pDO6935. Treatment with isopropyl β-D-1-thiogalactopyranoside (IPTG), a known *lac* promoter inducer, and glucose, a catabolite repressor, were also conducted to support our analysis (7). We were able to demonstrate expression of GFP through fluorescent imaging of colonies and quantification of fluorescence intensities with a spectrophotometer. Our findings report that the *lac* operon may be driving *brkA* transcription in pDO6935.

METHODS AND MATERIALS

Bacterial strains and plasmids. All bacterial strains were obtained from the inventory of The University of British Columbia's Microbiology & Immunology Department. The pDO6935 plasmid was designed by Oliver et al. (3). The pSPPH21 vector with no promoter was constructed by Abrishamkar et al. (6), and the pSPPH21 vector with a promoter was constructed by Gawol et al. (8). The pSB1C3 vector with a promoterless GFP which we originally intended to use before switching to pSPPH21 was constructed by the UBC iGEM team (6). Starter plates were streaked from cryopreserved bacteria stocks.

Competent cell preparation. Overnight cultures of DH5α *E. coli* were subcultured until an optical density of 0.4 at 600 nm (OD_{600}) was achieved. The culture was then transferred to an ice bath for 15 minutes to arrest growth, followed by centrifugation at 4°C at 4000 rpm for 10 minutes. The resulting cell pellet was resuspended in ice-cold 0.1 M $CaCl_2$, and the centrifugation step was repeated. Finally, the competent cells were aliquoted and stored in 0.1 M $CaCl_2$ with 10% glycerol at -70°C until further use.

Luria-Bertani (LB) and chloramphenicol media preparation. To prepare LB broth, 10 g tryptone, 5 g yeast extract, and 10 g NaCl were added to 1 L of distilled water and autoclaved. To make LB broth containing chloramphenicol, the antibiotic was added to make a final concentration of 20 µg/mL prior to each use. To prepare LB agar plates, 10 g tryptone, 5 g yeast extract, 10 g NaCl, and 15 g agar were added to 1 L of distilled water and autoclaved. Roughly 20 mL of the solution was poured onto each plastic petri dish. For LB agar plates with chloramphenicol, the solution was allowed to dissipate heat until cool enough to hold comfortably before adding chloramphenicol to a final concentration of 20 µg/mL.

Primer design. Using SnapGene, the pDO6935 sequence downstream of the origin of replication and upstream of *brkA* was selected and sectioned into 5 regions based on the *lac* operon. A total of 6 primers were designed to amplify the 5 constructs for testing our hypothesis (Tables S1-S2, Figure S1). Primers were each designed to have 24-25 annealed bases, an added 5' sequence for restriction enzyme digestion, and a melting temperature of 58-61ºC. Restriction sites were based on the originally planned vector pSB1C3 (with GFP), thus primers were designed with either EcoRI or XbaI restriction sites. Primer designs were simulated on SnapGene to ensure products were correct.

Plasmid isolation and quantification. Plasmid isolation was conducted using the alkaline lysis technique through the EZ-10 Spin Column Plasmid DNA Miniprep Kit (BioBasic; Cat.# BS614). The quantification and evaluation of DNA concentration and purity were performed using a NanoDrop2000® Spectrophotometer (ThermoFisher) at 260 nm.

PCR amplification of promoter regions. Plasmid was isolated from an overnight culture of *E. coli* pDO6935. For the PCR of each construct, 10 µL of 5X SuperFi™ Buffer (Invitrogen; Cat.# 12351-010), 1 µL of 10 mM dNTP mix (Invitrogen; Cat.# 18427-013), 10 µL of 5X SuperFi™ GC Enhancer (Invitrogen; Cat.# 12351-010), 0.5 µL of 2U/µL Platinum™ SuperFi™ DNA Polymerase (Invitrogen; Cat.# 12351-010), 2.5 µL of 10 µM forward primer, 2.5 µL of 10 µM reverse primer, and 10 ng of template DNA were added to a PCR tube. Each tube was topped up with nuclease-free water to a final volume of 50 µL. After capping, tubes were briefly mixed and centrifuged. Using a Bio-Rad T100 Thermal Cycler, the reaction tubes were subjected to an initial denaturation at 98ºC for 30s. Then, 35 cycles of denaturation (98ºC for 10s), annealing (57ºC for 10s), and extension (72ºC for 60s). For the final extension, the temperature was at 72ºC for 5 minutes. Amplicon products were isolated according to the manufacturer's instructions for the GeneJET PCR Purification Kit (Thermo Scientific™; Cat.# K0701).

Agarose gel electrophoresis. 10X TAE buffer was prepared with 24.25 g tris-acetate, 5.7 mL glacial acetic acid, and 10 mL of 0.5 M EDTA added with distilled water to a final volume of 500 mL. 1.5 g of agarose was dissolved in 100 mL of 1X TAE using 30-second intervals in a microwave. Once the agar was completely dissolved, the flask was allowed to dissipate heat until cool enough to hold comfortably. 5 µL of 20,000X RedSafe (FroggaBio; Cat.# 21141) was added and mixed, then the mixture was poured into the gel-casting platform with the gel comb in place. The gel was left at room temperature to solidify before being immersed in 1X TAE buffer. DNA samples were prepared by adding 10µL of samples with 2 µL of 6X DNA loading dye (New England BioLabs; Cat.# B7024S). The prepared DNA samples and an appropriate DNA ladder were carefully pipetted into the wells of the agar gel. The gel ran at 100V for roughly 1.5 hours. The gel was viewed under the Bio-Rad ChemiDocMP imaging system using the GelRed setting.

Digestion and Ligation of inserts into pSPPH21. 1 µg of the amplified promoter regions was digested with 20 units of EcoRI-HF (New England Biolabs; Cat.# R3101), rCutSmart buffer, and nuclease free water at 37°C for 15 minutes. For pSPPH21, a sequential digest was performed. Following digestion of pSPPH21 with EcoRI-HF, purification of DNA was completed using the GeneJET PCR Purification Kit (Thermo Scientific™; Cat.# K0701). 10 units of NruI (New England Biolabs; Cat.# R3192) and NEBuffer r3.1 were then added to the purified product and incubated at 37°C for 15 minutes. A final purification step was completed following the final digestion reaction. Ligation was then performed using a 3:1 and 7:1 molar ratio of insert to vector and using T3 DNA ligase (New England Biolabs; Cat.# M0317).

Heat shock transformation of recombinant plasmid constructs into *E. coli* DH5α. 5 µL of ligation reaction was added to 50 µL of competent DH5α cells which were then incubated on ice for 30 minutes. The cells were then heat-shocked in a 42°C water bath for 30 seconds and chilled on ice for 2 minutes. 1 mL of pre-warmed LB media was then added and the tubes were shake-incubated at 200 rpm at 37°C for 1 hour for recovery and outgrowth. 100 µL of transformants was then plated on LB + 20 µg/mL chloramphenicol plates and incubated overnight at 37°C.

Fluorescence agar plate imaging. Following transformation, colonies from the fresh plates were streaked to an LB + 20 µg/mL chloramphenicol plate (uninduced), as well as a LB + 20 µg/mL chloramphenicol plate with 200 µL of 0.5 mM IPTG (induced), and a LB + 20 µg/mL chloramphenicol plate with 150 µL of 1% glucose (repressed). These plates were incubated overnight at 37°C. The next day, plates were visualized using the Bio-Rad ChemiDocMP imaging system. Alexa 488 settings were then used to excite GFP.

Fluorescence quantification with plate reader and data processing. To quantify fluorescence using the BioTek Epoch 2 microplate spectrophotometer, 250 µL of LB with 20 µg/mL chloramphenicol media was added to the uninduced wells. For the induced and repressed wells, stock solutions of LB media with 20 µg/mL chloramphenicol and 0.5 mM

IPTG (induced) or 1% glucose (repressed) were created. A final volume of 250 µL was then added to the specific wells. Individual colonies were then inoculated to the wells. Each sample was prepared in triplicates and cells were left to grow on the microplate spectrophotometer overnight. OD_{600} reads and GFP reads, using an excitation wavelength of 485 nm and emission wavelengths of 528 nm, were collected. To adjust for cell density, the relative fluorescent units were normalized by dividing them by the OD_{600} values of each well. Exposure times and gain settings were kept consistent between all measured plates to ensure reliable and comparable results. Data was plotted with Prism®. Fold-change between induced, repressed, and uninduced samples was calculated by averaging the triplicates and dividing values for induced and repressed by values for uninduced samples. Statistical significance of fold-changes was calculated using 2-way ANOVA.

Whole plasmid sequencing. To prepare samples for whole plasmid sequencing via Plasmidsaurus, isolated plasmid constructs were diluted with nuclease free water to a final concentration of 30 ng/uL.

RESULTS

Three *brkA* promoter mapping amplicons were inserted in pSPPH21 and constructed pNSBR23_A1, pNSBR23_A2, and pNSBR23_A5. To identify possible promoter regions driving *brkA* transcription in pDO6935, the putative promoter sequences were amplified and inserted into GFP reporter plasmid pSPPH21. Following a visual analysis using SnapGene, a *lac* operon sequence was identified between *brkA* and the origin of replication, suggesting that the *lac* promoter could be driving *brkA* transcription. To investigate this possibility, the area upstream of *brkA* was divided into five regions using six primers (Figure S1). Primers 1 and 2 were used to amplify the region between *brkA* and up to, but not including, the *lac* operon, referred to here as A1 (amplicon 1). Primers 1 and 3 were used to amplify the region that included the *lac* operon, referred to as A2 (amplicon 2). Primers 1 and 4 were used to amplify the region between *brkA* and the origin of replication called A3 (amplicon 3). Primers 4 and 5 were used to amplify the region between the origin of replication and the *lac* operon, called A4 (amplicon 4). Lastly, primers 4 and 6 were used to amplify the region between the origin of replication up to and including the *lac* operon, called A5 (amplicon 5). Primers were designed to add XbaI and EcoRI restriction sites (Figure 1A, 1B). This was done to allow for amplicon insertion into pSB1C3 upstream of its promoterless GFP sequence. Two restriction sites were used to ensure that their orientation towards GFP in pSB1C3 would match that of *brkA* in pDO6935. Following the restriction sites design, we tested pSB1C3 as a negative control and learned that pSB1C3 expressed GFP in transformed colonies prior to amplicon insertion (Figure S2A, S2B). This prevented pSB1C3 use as a reporter plasmid. After testing the SIGEX reporter plasmid, pSPPH21, we discovered that it did not induce GFP expression in *E. coli*. This allowed for its use as a vector for our amplicons and as a negative control in subsequent steps of our experiment. As the XbaI restriction site would not result in a correct insertion alignment in pSPPH21, we used the blunt-ended NruI restriction site instead. Our vector was double-digested with EcoRI and NruI to allow for proper insertion. Our inserts, on the other hand, were only digested with EcoRI, as PCR amplification already yielded a blunt-ended amplicon. Five amplicons of expected lengths were amplified (Figure S3), and three amplicons (A1, A2, and A5) were successfully inserted into pSPPH21 to create pNSBR23_A1, pNSBR23_A2, and pNSBR23_A5 respectively.

***lac* operon sequence in pDO6935 may drive *brkA* expression in *E. coli*.** To test if the *lac* operon in pDO6935 was able to cause leaky expression of GFP in pSPPH21, we performed an induction assay. pNSBR23_A1, pNSBR23_A2, and pNSBR23_A5 were transformed into *E. coli* and grown on LB plates containing 20 µg/mL chloramphenicol and 0.5 mM IPTG (Figure 2A, 2B, 2C). Our positive control was a previously constructed pSPPH21 plasmid that contains an intact *lac* operon upstream of its GFP gene (8). Our negative control is the promoterless pSPPH21 vector that does not express GFP. Because amplicons 3 and 4 (A3 and A4) failed to insert, they were treated as negative controls for our IPTG induction assay (Figure 2C).

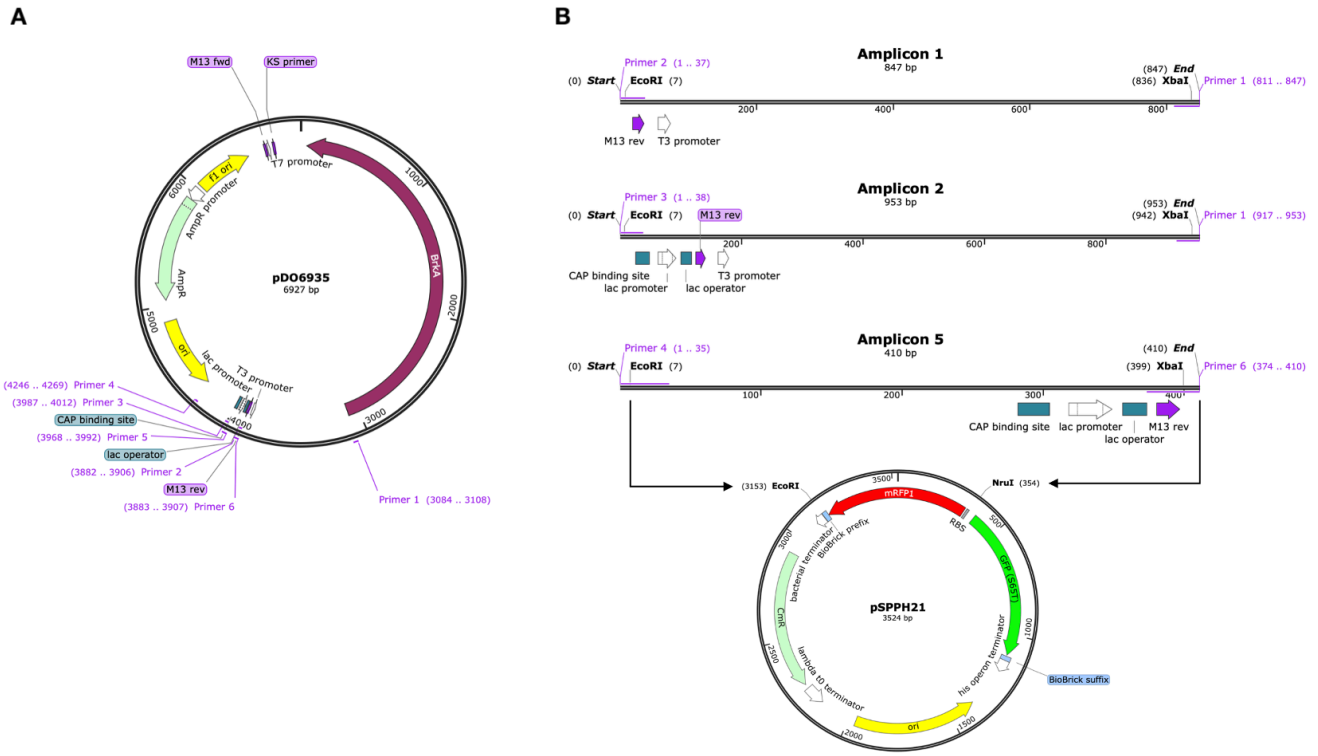

FIG. 1 Graphical Depiction of DNA plasmids and construct. (A) Primers were created to amplify select regions of pDO6935 within the space between the *brkA* gene and the plasmid origin of replication. (B) Amplicons were then ligated into pSPPH21 between its EcoRI and NruI sites after digestion. Amplicon 1 contained the region between *brkA* and the *lac* operon, Amplicon 2 was the same as Amplicon 1 with the addition of the *lac* operon, and Amplicon 5 was the region between the origin of replication up to and including the *lac* operon. All amplicons were inserted such that their orientation to the GFP of pSPPH21 matched their original orientation to *brkA* in pDO6935.

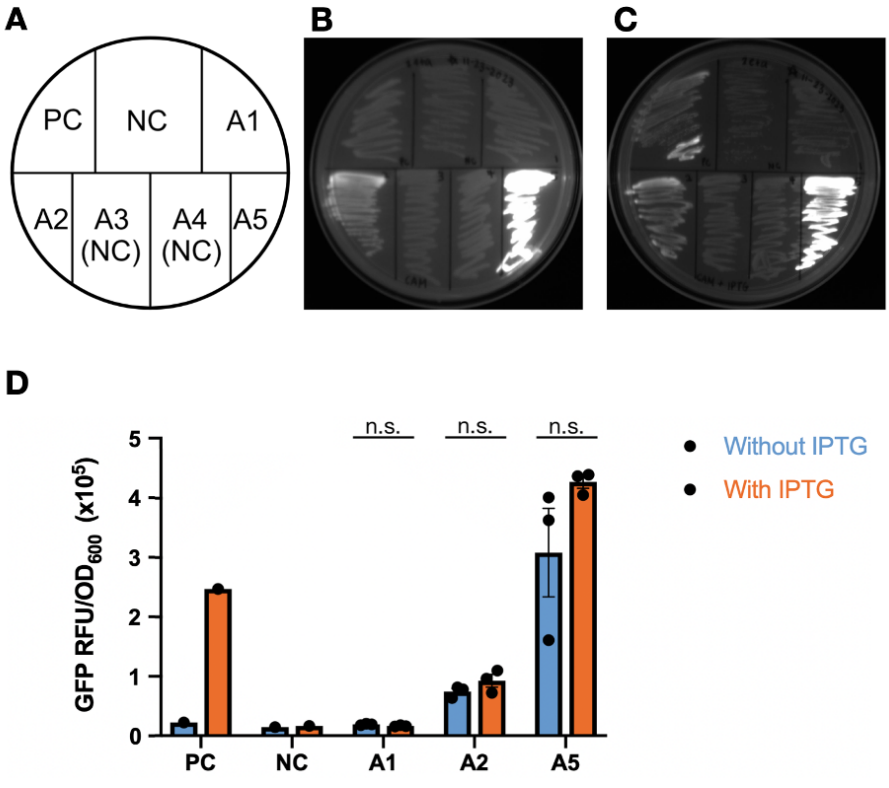

FIG. 2 IPTG does not significantly alter expression of GFP in *E. coli* transformed with pNSBR23_A1, pNSBR23_A2, or pNSBR23_A5. (A) Organization of plates streaked with *E. coli* transformed with pSPPH21 containing amplicons from pDO6935 used in panels B and C. PC = positive control, NC = negative control, A1-A5 = Amplicon 1 through Amplicon 5. (B-C) Representative GelDoc images showing GFP fluorescence of transformants at excitation wavelength of 488 nm when each transformants are plated on (B) LB containing 20 ng/µL chloramphenicol and (C) LB containing 20 ng/µL chloramphenicol and 0.5 mM IPTG. (D) Quantification of GFP RFU normalized to OD_{600} readings in 250 µL LB cultures containing 20 µg/mL chloramphenicol with and without 0.5 mM IPTG after 21 hours of incubation., n.s. = not significant. Significance calculated using two-way ANOVA.

Compared to the negative controls, pNSBR23_A2 and pNSBR23_A5 induced GFP expression. Both pNSBR23_A2 and pNSBR23_A5 contained the *lac* operon, suggesting that the *lac* promoter may have been responsible for driving *brkA* expression in pDO6935. The positive control contained a *lac* operon sequence and exhibited an inducible GFP expression through the addition of IPTG. However, none of the transformed cells seemed to qualitatively have an increase in GFP expression in the presence of IPTG (Figure 2C). Altogether, these results suggest that the *lac* operon drives *brkA* expression in pDO6935.

***lac* promoter expression levels in pNSBR23_A2 and pNSBR23_A5 were independent of IPTG induction in *E. coli*.** To quantitatively assess GFP expression levels in the presence of IPTG in *E. coli* transformed with pNSBR23_A1, pNSBR23_A2, and pNSBR23_A5, we ran a liquid IPTG induction assay. We cultured transformants in a clear-bottom 96 well plate with 250 µL of LB containing 20 µg/mL chloramphenicol. We included conditions with and without 0.5 mM IPTG. Using an overnight culture, we assessed the level of GFP fluorescence and normalized it to OD_{600} values for each sample. The negative control and pNSBR23_A1 did not express GFP regardless of IPTG induction as expected because the plasmid construct does not contain the lac operon (Figure 2D). In contrast, the positive control showed significant increase of GFP expression upon IPTG induction (Figure 2D). While pNSBR23_A2 and pNSBR23_A5 showed an increase in GFP expression in the presence of IPTG, the observed difference was not statistically significant (Figure 2D). Altogether, this suggests that *lac* operon in pDO6935 can cause expression of downstream genes regardless of IPTG presence.

The *lac* operon in pNSBR23_A2 and pNSBR23_A5 was inhibited by glucose in *E. coli*. To further test whether the transcription is modulated by the *lac* operon, we performed a glucose induction assay. We plated transformants on LB plates containing 20 µg/mL chloramphenicol and on LB plates with both 20 µg/mL chloramphenicol and 1% glucose (Figure 3A). We used the same positive control from our IPTG induction assay, pSPPH21

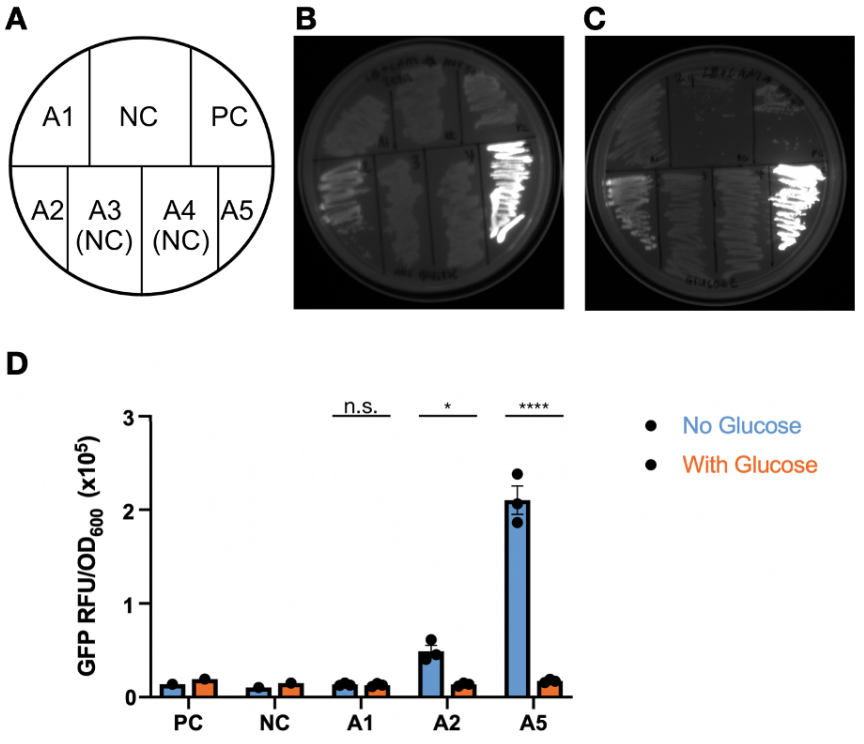

FIG. 3 Glucose significantly lowers expression of GFP in *E. coli* transformed with, pNSBR23_A2 and pNSBR23_A5. (A) Organization of plates colonized with *E. coli* transformed with pSPPH21 containing amplicons from pDO6935 used in panels B and C. PC = positive control, NC = negative control, A1-A5 = Amplicon 1 through Amplicon 5. (B-C) Representative GelDoc images showing GFP fluorescence of transformants at excitation wavelength of 488 nm when each transformants are plated on (B) LB containing 20 ng/µL chloramphenicol and (C) LB containing 20 ng/µL chloramphenicol and 1% glucose. (D) Quantification of GFP RFU normalized to OD_{600} readings in 250 µL LB cultures containing 20 µg/mL chloramphenicol with and without 1% glucose after 22 hours of incubation. n.s.=not significant, *p<0.05, ****p<0.0001. Significance calculated using two-way ANOVA.

plasmid with a *lac* promoter upstream GFP. Similarly, the negative control used is the promoterless pSPPH21 plasmid, which did not express GFP as anticipated. Transformants with pNSBR23_A3 and pNSBR23_A4 were used as additional negative controls due to the unsuccessful insertion. As expected, pNSBR23_A2 and pNSBR23_A5 transformants had a qualitatively lower expression of GFP when cultured on plates containing glucose. pNSBR23_A5 transformants had basal expression regardless of the presence of glucose

(Figure 3B, 3C). To obtain quantitative results, we cultured the transformants and controls in a clear-bottom 96 well plate with 250 µL of LB containing 20 µg/mL chloramphenicol and 1% glucose. Transformations with both pNSBR23_A2 and pNSBR23_A5, but not pNSBR23_A1, were found to have significantly reduced GFP expression in the presence of glucose (Figure 3D). This observation may suggest that *brkA* expression in *E. coli* transformed with pDO6935 may be decreased in glucose-rich environments. This further suggests that the *lac* operon sequence in pDO6935 can function as a SIGEX promoter of *brkA*.

DISCUSSION

Despite the many potential applications of pDO6935, its promoter has not yet been mapped. Because the *lac* promoter is known to exhibit basal levels of expression, it was investigated as a possible promoter driving expression of *brkA* in *E. coli*. To begin this research, various regions upstream of pDO6935 were amplified. These different amplicons were inserted into pSPPH21, a promoterless vector that contains GFP. Fluorescence was qualitatively and quantitatively assessed to determine whether the *lac* operon drives GFP expression. IPTG and glucose treatments were then conducted to support our hypothesis. Overall, findings demonstrate that the *lac* operon drives *brkA* expression in pDO6935 in *E. coli*.

Various factors resulted in poor ligation and transformation efficiencies. The vector used in our experiment, pSPPH21, contains NruI and EcoRI restriction sites upstream of the GFP region, where the amplicons were inserted (Figure 1B, S4-S8). When NruI and EcoRI cleave the plasmid, it leaves a blunt end and sticky end, respectively. The usage of blunt-ended and sticky-ended restriction sites in each amplicon ensured correct orientation of the inserts and was expected to prevent self-ligation of the vector. Despite these advantages, achieving successful ligation of the amplified regions into the pSPPH21 vector was challenging. Due to the incompatibility of these restriction enzymes, sequential digestion was performed as no available buffer allowed both enzymes to simultaneously exhibit >50% activity, according to manufacturer's instructions.

Blunt-ended ligation is also known to be 10-100x less efficient than sticky-end ligation due to the absence of single-stranded overhangs that facilitate complementary binding with the other corresponding end (9, 10). As a result of the lower efficiency observed in blunt-ended ligation, our culture yielded very minimal colonies, with the number of transformants per culture plate ranging anywhere from 0 to 6. Moreover, there was a low efficiency in insertion of our desired constructed plasmid into the vector. Initial sequencing results revealed that most of the transformants contained the same plasmid as the negative control, the vector itself. This suggested that the excised DNA in pSPPH21 (Figure 1B), once digested, likely re-ligated again to its original position.

To test whether the *brkA* of pDO6935 was being transcribed by leaky expression of the upstream *lac* operon, we divided the region between *brkA* and the origin of replication into three regions. The first region was between *brkA* and the *lac* operon, the second was the *lac* operon itself, and the third was between the *lac* operon and the origin of replication. We hoped to show that neither the regions upstream nor downstream of the *lac* operon were capable of inducing *brkA* expression unless the *lac* operon was included. To this extent, we created amplicons A1, A2, A4, and A5. Additionally, we also wanted to show that an amplicon containing the entire nucleotide sequence from *brkA* to the origin of replication would induce expression of a reporter gene, which we attempted to accomplish via amplicon A3.

Because of the low ligation and transformation efficiencies observed with all five amplicons, a laborious screening process was employed with repeated rounds of ligation, transformation, and plasmid extractions. Eventually, we were able to obtain three recombinant plasmids that served as our test constructs and we used the rest of the plasmids as negative controls. (Figure S9). While ligation and transformation concluded the experiment, higher efficiencies in both ligation and transformation are recommended to complete the mapping of the promoter regions and strengthen our findings.

lac operon drives expression in pDO6935. In this study, we have observed that the *lac* promoter upstream may be responsible for driving *brkA* expression. Analysis of the clones under the Bio-Rad ChemiDocMP Imaging system revealed that pNSBR23_A2 and

pNSBR23_A5 transformants exhibited a fluorescence signal (Figure 2B, 3B). pNSBR23_A2 contained the sequence upstream of *brkA*, encompassing the *lac* operon, while pNSBR23_A5 contained the entire *lac* operon and downstream base pairs up to the origin of replication (Figure 1B). Notably, a common factor between these two constructs was the *lac* operon, supporting our hypothesis that the leaky *lac* promoter was responsible for driving *brkA* expression (11). This is further supported by results observed from IPTG induction and glucose repression (Figure 2, 3).

Moreover, pNSBR23_A2 transformants exhibited decreased GFP expression compared to pNSBR23_A5 transformants (Figure 2, 3). This difference may be attributed to the location and proximity of the *lac* promoter to the GFP gene, which would align with previous research, where expression levels differed two to threefold based on chromosomal location, emphasizing the importance of the distance between the gene and promoter on promoter strength (12). The variation in GFP expression levels also corresponds with a previous study that employed a similar promoter trapping approach, revealing clones with varying levels of GFP expression based on distance between promoter and GFP (13).

IPTG induction has minimal effect on GFP expression. Our results demonstrated that induction with IPTG did not significantly increase GFP expression in plasmid constructs, as pNSBR23_A2 and pNSBR23_A5 transformants both exhibited high levels of fluorescence regardless of IPTG addition (Figure 2B, C). The minimal differences of IPTG induction for pNSBR23_A2 and pNSBR23_A5 compared to the positive control may be explained by the stoichiometric differences in the *lacI* gene and the *lac* operon. Our plasmid constructs did not contain the *lacI* gene, and since plasmids have the ability to replicate independently of the bacterial chromosome (14), copies of the plasmid *lac* operator were likely much higher than the levels of chromosomal *LacI*, thereby preventing repression of the gene. This possibility would align with previous research (15).

The positive control, on the other hand, included an insert that contained *lacI* and its promoter, as well as the *lac* promoter, and *lac* operator (8). The inclusion of *lacI* and its promoter resulted in the requirement for IPTG to be present before fluorescence was observed. Moreover, the extent that *lacI* is repressed has been shown to be determined largely by the position of the operator and kinetics of the repressor-operator complex formation (16). Because *lacI* is directly upstream of the *lac* operon in the positive control, we expected the operon to be strongly repressed by *lacI*, which was illustrated in our results (Figure 2D). Overall, this knowledge explained the differences in expression level observed in our plasmid constructs in comparison to the positive control.

Glucose represses GFP expression. Based on fluorescence data (Figure 3D), we observed that *E. coli* transformed with pNSBR23_A2 or pNSBR23_A5 exhibited significantly reduced GFP levels when incubated with 1% glucose. However, the positive control was unaffected by glucose repression due to the presence of *lacI*. Glucose repression is known to inhibit *lac* operon through catabolite repression (7). Specifically, since glucose is preferentially used as an energy source, lower concentrations of cyclic AMP (cAMP) are observed in the presence of abundant glucose (17). cAMP is necessary for the activation of the catabolite activator protein (CAP), which enhances the binding of RNA polymerase to the *lac* promoter (18). As such, repression of GFP expression by glucose in our plasmid constructs further supports our hypothesis that the *lac* operon may be driving *brkA* expression in pDO6935.

Limitations The aim of this project was to map out the promoter region responsible for *brkA* expression in pDO6935 by cloning various sequences to a SIGEX reporter plasmid. Due to various factors, minimal transformants were observed. One factor that led to this would be the use of sticky- and blunt-ended restriction sites, as no buffer was available that allowed for >50% activity of both restriction enzymes. We had initially planned to use pSB1C3 (with GFP) as our vector and designed primers with the appropriate sticky-end restriction sites. However, this vector demonstrated strong GFP fluorescence expression without insertion of any amplicons, making it unsuitable for our use. To mitigate these challenges, future work could involve designing new primers for pSPPH21 that utilize two sticky-ended restriction sites. We also faced the high likelihood that vector plasmids were re-ligating following digestion. A possible strategy to combat this in future experiments would be to run the digested plasmid solutions through an agarose gel, then excising and purifying the band

corresponding to the correctly digested plasmid. Other factors that may have affected our transformation efficiency included reagent contamination. Specifically, each digestion was followed by cleanup using the GeneJET PCR Purification Kit (Thermo Scientific™; Cat.# K0701), which was shared between teams. Poor transformation efficiency for our plasmid constructs could also have been associated with larger inserts being more difficult to ligate (19).

We also faced issues with our experimental setups. A major limitation of our study was the lack of statistical power, and increasing the biological replicates would strengthen the observed differences due to IPTG and glucose treatment. Additionally, instead of using qualitative analysis of fluorescence in agar plates, we could have used a tool that would provide a quantitative measure of fluorescence level normalized to the area occupied by the sample colonies, then ran statistical tests on the data. Finally, it is worth mentioning that a better positive control for our experiments studying the effect of glucose treatment would have been a transformant with a plasmid that caused a significantly decreased GFP expression in the presence of glucose.

Conclusions The reporter pSPPH21 constructed by Abrishamkar et al. was shown to be functional based on GFP expression when various promoter sequences from pDO6935 were inserted in a 5' to 3' direction towards GFP. The *lac* operon sequence in pNSBR23_A2 and pNSBR23_A5 induced GFP expression. Thus, we suggest that the *lac* operon may be driving *brkA* expression in pDO6935. Treatment with IPTG highlighted that the high copy number of *lac* operator present in our promoter regions likely stoichiometrically outnumbered the *lacI* repressor expressed by the host DH5α *E. coli* cells, leading to leaky expression regardless of IPTG induction. Glucose treatment demonstrated gene repression, likely because glucose is a known catabolite repressor, further supporting our argument that the *lac* operon drives expression of *brkA* in pDO6935.

Future Directions With the confirmed functionality of the pSPPH21 vector, researchers can pursue further promoter trapping experiments using this vector. Due to poor ligation and transformation efficiencies, optimizing the protocols is crucial to allow for successful cloning. Future investigation could consider refining the blunt-ended ligation efficiency by exploring longer incubation times and experimenting with different insert-to-vector ratios and reaction concentrations. Alternatively, ligation efficiency may be improved by designing new primers for the pSPPH21 vector, introducing sticky ends on both sides.

To further support the hypothesis that the *lac* promoter is responsible for *brkA* transcription, additional studies could insert a reporter gene such as GFP, downstream of the *lac* operon in pDO6935. If the reporter gene does express GFP, this would strengthen the finding that the *lac* promoter may be the driver of *brkA* expression.

While our study may have identified the promoter region of *brkA,* the Shine-Dalgarno (SD) sequence remains unknown (1). The SD sequence, typically characterized by the prevalence of AG-rich sequences at the center of the ribosome binding site, plays a crucial role in initiating translation (20). Translation initiation in prokaryotes often requires an interaction between the ribosome and SD sequence. However, many genes lack this sequence, and translation initiation takes place independent of the consensus SD sequence (21, 22). Further investigation could explore more into the translation initiation site of *brkA* within the pDO6935 plasmid by employing a similar approach. By using a reporter gene fusion with various potential translation initiation sites, researchers can identify the functional initiation site by assessing the expression of the reporter gene, a method used in previous studies (23).

ACKNOWLEDGEMENTS

We would like to thank Dr. David Oliver, Jade Muileboom, Herieth Ringo, and the rest of the MICB 471 teaching team for their continued support and guidance throughout the project. We would also like to acknowledge Abrishamkar et al. for providing us with pSPPH21 plasmid to use as a vector, which formed the basis of our project. Lastly, we would like to thank UBC's Department of Microbiology and Immunology for providing access to the laboratory resources and materials used in our research. We would also like to thank two anonymous reviewers for constructive feedback on this manuscript.

CONTRIBUTIONS

Nidhin Biju (NB), Sharisse Chan (SC), Dahyeon (Betty) Hong (BH), and Renz Po (RP) contributed to the conduction of the experiment. NB generated the figures, figure legends, and wrote sections of the introduction, results, and limitations. SC wrote sections of the abstract, methods, discussion, limitations and future directions. BH wrote sections of the introduction, discussion, and future directions. RP wrote sections of the methods, results, and limitations. All authors contributed to the writing of the abstract and editing of the manuscript.

REFERENCES

1. **Fernandez RC, Weiss AA.** 1994. Cloning and sequencing of a *Bordetella pertussis* serum resistance locus. *Infect Immun* **62**:4727–4738.
2. **Mattoo S, Cherry JD.** 2005. Molecular Pathogenesis, Epidemiology, and Clinical Manifestations of Respiratory Infections Due to *Bordetella pertussis* and Other Bordetella Subspecies. *Clin Microbiol Rev* **18**:326–382.
3. **Oliver DC, Huang G, Fernandez RC.** 2003. Identification of secretion determinants of the *Bordetella pertussis* BrkA autotransporter. *J bacteriol*, **185**:489–495.
4. **Weiss AA, Melton AR, Walker KE, Andraos-Selim C, Meidl JJ.** 1989. Use of the promoter fusion transposon Tn5 lac to identify mutations in *Bordetella pertussis* vir-regulated genes. *Infect Immun* **57**:2674–2682.
5. **Nielsen BL, Willis VC, Lin C-Y.** 2007. Western blot analysis to illustrate relative control levels of the lac and ara promoters in *Escherichia coli*. *Biochem Mol Biol Educ* **35**:133–137.
6. **Abrishamkar P, Dana HM, Oveisi S, Tabassi P.** 2022. Construction of SIGEX duo-directional reporter plasmid pSPPH21. *Undergrad J Exp Microbiol Immunol* **8**.
7. **Tyler B, Loomis WFJ, Magasanik B.** 1967. Transient repression of the *lac* operon. *J Bacteriol* **94**:2001–2011.
8. **Gawol D, Floyd R, Kohara K, Lee Y.** 2022. Duo-directionality of the substrate-induced expression screening vector pSPPH21 confirmed with a lac operon screen. *UJEMI+* **8**:1–12.
9. **Bercovich JA, Grinstein S, Zorzopulos J.** 1992. Effect of DNA concentration on recombinant plasmid recovery after blunt-end ligation. *BioTechniques* **12**:190, 192–193.
10. **Hung MC, Wensink PC.** 1984. Different restriction enzyme-generated sticky DNA ends can be joined in vitro. *Nucleic Acids Res* **12**:1863–1874.
11. **Jacob F, Monod J.** 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**:318–356.
12. **Sousa C, de Lorenzo V, Cebolla A.** 1997. Modulation of gene expression through chromosomal positioning in *Escherichia coli. Microbiology* **143**:2071–2078.
13. **Chen J, Haverty J, Deng L, Li G, Qiu P, Liu Z, Shi S.** 2013. Identification of a novel endogenous regulatory element in Chinese hamster ovary cells by promoter trap. *J Biotechnol* **167**:255–261.
14. **Jahn M, Vorpahl C, Hübschmann T, Harms H, Müller S.** 2016. Copy number variability of expression plasmids determined by cell sorting and Droplet Digital PCR. *Microb Cell Factories* **15**:211.
15. **Glascock CB, J. Weickert M.** 1998. Using chromosomal lacIQ1 to control expression of genes on high-copy-number plasmids in *Escherichia coli1*Published in conjunction with A Wisconsin Gathering Honoring Waclaw Szybalski on the occasion of his 75th year and 20years of Editorship-in-Chief of Gene, 10–11 August 1997, University of Wisconsin, Madison, WI, USA.1. *Gene* **223**:221–231.
16. **Lanzer M, Bujard H.** 1988. Promoters largely determine the efficiency of repressor action. *Proc Natl Acad Sci* **85**:8973–8977.
17. **Tagami H, Inada T, Kunimura T, Aiba H.** 1995. Glucose lowers CRP* levels resulting in repression of the *lac* operon in cells lacking cAMP. *Mol Microbiol* **17**:251–258.
18. **Fried MG, Crothers DM.** 1983. CAP and RNA polymerase interactions with the lac promoter: binding stoichiometry and long range effects. *Nucleic Acids Res* **11**:141–158.
19. **Gabor EM, Alkema WBL, Janssen DB.** 2004. Quantifying the accessibility of the metagenome by random expression cloning techniques. *Environ Microbiol* **6**:879–886.
20. **Shine J, Dalgarno L.** 1974. The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci U S A* **71**:1342–1346.
21. **Fargo DC, Zhang M, Gillham NW, Boynton JE.** 1998. Shine-Dalgarno-like sequences are not required for translation of chloroplast mRNAs in *Chlamydomonas reinhardtii* chloroplasts or in *Escherichia coli*. *Mol Gen Genet MGG* **257**:271–282.
22. **Skorski P, Leroy P, Fayet O, Dreyfus M, Hermann-Le Denmat S.** 2006. The highly efficient translation initiation region from the *Escherichia coli rpsA* gene lacks a shine-dalgarno element. *J Bacteriol* **188**:6277–6285.
23. **Saito K, Green R, Buskirk AR.** Translational initiation in *E. coli* occurs at the correct sites genome-wide in the absence of mRNA-rRNA base-pairing. *eLife* **9**:e55002.