Exploring a Putative Promoter Region in Mycobacteriophage JacoRen57

Ali Almail, Elizabeth Heeg , Byron Noordewier, Sara S. Tolsma
Northwestern College, Department of Biology, Orange City, IA 51041

Abstract

Phages are abundant particles that infect bacteria. For the SEA-PHAGES program, students discover phages and annotate their genomes. Throughout the annotation process, genes are identified based on bioinformatics evidence; however, little is known about mycobacteriophage promoters as they are not annotated. Promoters are necessary for gene expression, and in mycobacteriophages, a promoter typically precedes a series of genes that are expressed as a single transcript from which multiple proteins are translated. JacoRen57 is a singleton mycobacteriophage with a siphoviridae morphotype that possesses forward and reverse genes with gaps located at the transitions from forward to reverse genes. We hypothesized that these gaps contain promoters. We used BPROM and PePPER, prokaryotic promoter predictor software, which yielded matches to promoter consensus sequences in one of the gap regions. We cloned the putative promoter region into pLO86, a vector containing the mCherry reporter gene, to determine if the cloned region functions as a promoter by inducing mCherry expression in *Mycobacterium smegmatis*. The putative promoter region did not function as a promoter in vivo under standard *M. smegmatis* growth conditions.

Introduction

Mycobacteriophages are viruses that infect bacteria from the genus *Mycobacterium* [1]. JacoRen57 is a singleton soil mycobacteriophage that infects *Mycobacterium smegmatis* mc²155, which is a gram-positive acid-fast bacterium [2]. JacoRen57 was discovered by Tanner Rensink in a soil sample from Sioux Center, IA, USA in 2018 [2]. JacoRen57 exhibits a lytic cycle and a Siphoviridae morphotype with a non-contractile tail and double-stranded DNA genome. [2]. Its genome is 70,300 base pairs long composed of 33 forward genes followed by 16 reverse genes and 24 forward genes from the right to left arm of the genome [3]. The second transition from reverse to forward genes (40644-40974 bp) possesses a 331 bp region with minimal coding potential. We hypothesized that there would be a promoter for each of the forward and reverse genes that flank the region of interest. We used bioinformatics software BPROM and PePPER, which are online promoter predictor software, to assess the region of interest for promoter consensus sequences [4],[5]. These software utilize a database of prokaryotic promoter consensus sequences to identify if any are present in the inputted sequence [4],[5]. Both software yielded matches to putative promoter consensus sequences in the region of interest.

Prokaryotic promoters are sequences that are upstream of the transcription start site (+1) of a gene. A typical prokaryotic promoter contains -10 and -35 elements, each composed of 6 nucleotides that are recognized by a variety of sigma factors [6]. In *M. smegmatis*, the -10 promoter elements are like that of *Escherichia coli*; however, the -35 elements differ and can have sequential variability [7]. The TGN motif, which is an extended -10 element that is found in many *M. smegmatis* promoters allows for the binding of the **σ**70 factor [7]. Additionally, there are promoter consensus sequences of various sigma factors of *Mycobacterium tuberculosis*, which is in the same genus as *M. smegmatis*, that respond to cell stress signals [8].

Promoter sequences of mycobacteriophages have not been extensively examined. For the following experiments, we used reporter gene vectors pLO86 and pLO87 to assess promoter activity of the putative promoter region in JacoRen57 (40644-40974 bps) [3],[10]. These plasmids encode the mCherry gene, a reporter gene that codes for a pink fluorescent protein which is observable under visible light conditions [10]. Upstream of the mCherry gene is a ribosomal binding site, which is necessary for ribosome to bind and initiate translation, and a polylinker site for the insertion of putative promoter sequences [10]. These plasmids also possess a Kanamycin resistance gene for clonal selection [10]. If the cloned sequence functions as promoter in vivo then the mCherry gene will be transcribed when the sigma factors and RNA polymerase bind, resulting in a transcript that can be translated to produce mCherry. In this study we utilized reporter gene assays with pLO86 to assess the activity of the putative promoter region that lies in the transition from reverse to forward genes (40644-40974 bp) in JacoRen57 [3]. We demonstrate that the putative promoter region does not function as promoter in vivo in the phage host *M. smegmatis* under standard growth conditions.

Results

Objective 1: Amplify putative promoter region for cloning

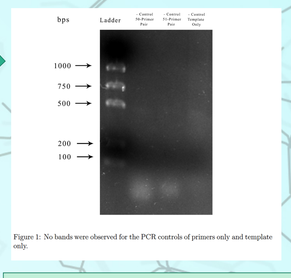

Figure 1: No bands were observed for the PCR controls of primers only and template only.

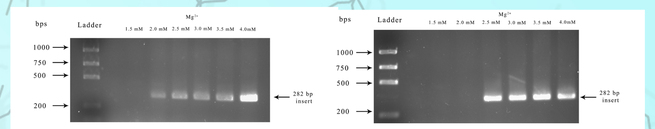

Figure 2: A 282 bp band resulted form PCR amplification of the putative promoter region using the 50-primer pair. We utilized a [Mg^{2+}] gradient from 1.5-4 mM to optimize the reaction, with [Mg^{2+}] of 2-4 mM yielding the expected 282 bp product.

Figure 3: A 282 bp band resulted form PCR amplification of the putative promoter region using the 51-primer pair. We utilized a [Mg^{2+}] gradient from 1.5-4 mM to optimize the reaction, with [Mg^{2+}] of 2.5-4 mM yielding the expected 282 bp product.

Putative promoter region amplified successfully.

Objective 2: Clone the 50 and 51 amplified inserts into pLO86 and transform into NEB5α.

Figure 4: Two NEB5α colonies were observed on the transformation plate from the 10-fold diluted 5:1 insert (50-primer pair amplified insert) to pLO86 ratio ligation. The dotted circles mark the colonies.

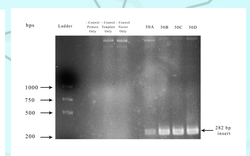

Figure 5: A 282 bp resulted form the PCR of the constructs using the 50-primer pair. 50A-50D are NEB5α colonies picked from the transformation plates of NEB5α to confirm the presence of the ligated insert in the vector using PCR.

Figure 6: NEB5α colonies were observed in the transformation plate from the 10-fold diluted 5:1 insert (51-primer pair amplified insert) to pLO86 ratio ligation.

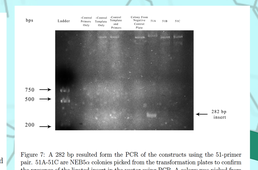

Figure 7: A 282 bp resulted form the PCR of the constructs using the 51-primer pair. 51A-51C are NEB5α colonies picked from the transformation plates to confirm the presence of the ligated insert in the vector using PCR. A colony was picked from the negative control plates in order to assess the transformed construct.

The 50 and 51 inserts were cloned successfully into pLO86.

Objective 3: Transform pLO86 and pLO87 into *M. smegmatis* as positive controls.

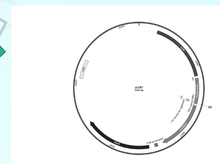

Figure 10: pLO87 vector map.

Figure 8: pLO86 vector map.

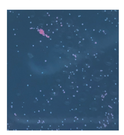

Figure 11: *M. smegmatis* colonies that are pink due to the expression of the reporter gene mCherry.

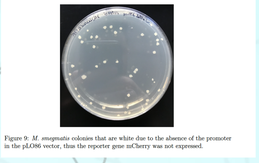

Figure 9: *M. smegmatis* colonies that are white due to the absence of the promoter in the pLO86 vector, thus the reporter gene mCherry was not expressed.

We successfully determined appropriate conditions to transform *M. smegmatis*

Objective 4: Transform pLO86 containing the 50 and 51 insert into *M. smegmatis*

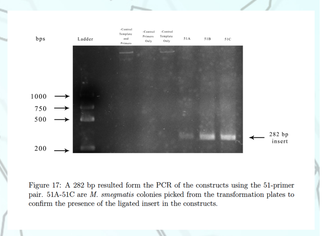

Figure 17: A 282 bp resulted form the PCR of the constructs using the 51-primer pair. 51A-51C are *M. smegmatis* colonies picked from the transformation plates to confirm the presence of the ligated insert in the constructs.

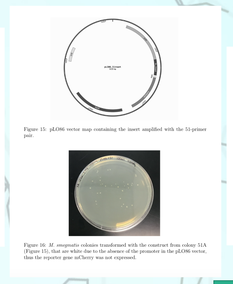

Figure 15: pLO86 vector map containing the insert amplified with the 51-primer pair.

Figure 16: *M. smegmatis* colonies transformed with the construct from colony 51A (Figure 15), that are white due to the absence of the promoter in the pLO86 vector, thus the reporter gene mCherry was not expressed.

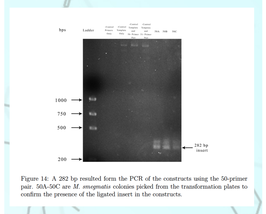

Figure 14: A 282 bp resulted form the PCR of the constructs using the 50-primer pair. 50A/50C are *M. smegmatis* colonies picked from the transformation plates to confirm the presence of the ligated insert in the constructs.

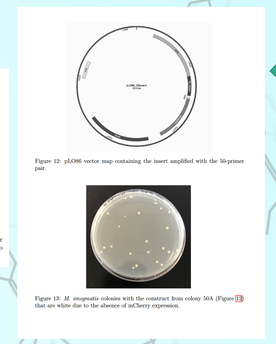

Figure 12: pLO86 vector map containing the insert amplified with the 50-primer pair.

Figure 13: *M. smegmatis* colonies with the construct from colony 50A (Figure 12) that are white due to the absence of mCherry expression.

We successfully transformed *M. smegmatis* with pLO86 containing the 50 and 51 insert. None of the colonies were pink in color.

Discussion

We successfully cloned the 50 insert and 51 insert into the pLO86 vector separately, as evidenced by the PCR using the 50 and 51-primer pairs which yielded a band at 282 bp, which is the predicted size of the insert, in only the transformed colonies and not the respective controls (Figures 5-7). For the construct with the 50 insert, which consists of the sequence preceding the gene with stop at 42969 bp from JacoRen57 in pLO86, *M. smegmatis* did not turn pink (Figure 13). This suggests that the cloned region is not a promoter in vivo for *M. smegmatis* under the standard growth conditions. For the 51 construct, which consists of the sequence proceeding the gene at stop 40254 bp from JacoRen57 in the pLO86 vector, the colonies did not turn pink in color (Figure 16). These results indicate that the cloned region does not function a promoter in vivo in *M. smegmatis* under standard growth conditions.

We hypothesize that this region might contain promoters. Since the putative promoter region is in the right hand of the genome, it might require phage transcription factors or sigma factors that are upstream of the region, which would then drive the expression of the mCherry gene in *M. smegmatis*. Furthermore, it may suggest that the spacing between the insert and +1 transcription start site exceeds the necessary limit; and therefore, the promoter does not function. However, the end of the putative promoter region is 21 bp from the ribosomal binding site (RBS), compared to the end of the hsp60 promoter which is 6 bp away from the RBS site in pLO87. Since the ribosomal site is fixed, we suspect that translation can occur if the gene is transcribed appropriately. These explanations would require further experimentation. One experiment would be sequencing to confirm proper insertion. A second experiment would be to conduct plaque assays with *M. smegmatis* transformed with either constructs to observe if certain phage transcription factors are required for the function of the promoter in vivo. If it is found that a phage transcription factor is required for the activity of the putative promoter region, then site directed mutagenesis can be utilized to elucidate the consensus sequence of the promoter.

Materials and Methods

PCR to Amplify Putative Promoter Region

Genomic DNA from JacoRen57 was isolated using the SEA-PHAGES DNA Extraction Protocol without DNase by Kristina Sevcik'19 [11]. We designed and purchased primers to amplify the gap region in JacoRen57 (40644-40974 bp), which yields a 282 bp product. We used the two primer pairs to amplify the gap region in both the forward and reverse directions to assay promoter activity in both directions. The primers we used are: 50 Forward: 5'-GCAAATGCGGCCGCCAGAGGCGTTCTGCAAGTA-3'; 50 Reverse: 5'-GCGGCGGTACCCAGAGGCGTTC TGCAAGTA-3'; 51 Forward: 5'-GCAAATGCGGCCGCCAGAGGCGTTCTGCAAGTA-3';51 Reverse: 5'-GCGGCGGTACCTCGCATGTCAGCACTGATTA-3'. We added NotI and KpnI restriction sites (underlined above) at the 5' ends of the primers for cloning purposes. We used GoTaq DNA Polymerase Kit from Promega (Madison, WI) to amplify the insert. For the PCR, we used an annealing temperature of 62°C and a [Mg2+] gradient to determine the optimal conditions for the amplification of the insert. We then purified the insert utilizing a 2% agarose gel and excised the band with a sterile scalpel under UV prep conditions using a BioRad Chemi-Doc (Hercules, California), and then purified it utilizing the Monarch DNA Gel Extraction Kit (Ipswich,MA).

Digestion of Insert and Vector

We digested both inserts (amplified product using both the 50 and 51-primer pairs) and the pLO86 vectors with NotI-HF and KpnI-HF. We assembled the reaction in 50 µL, with 40 units of each enzyme, approximately 800 ng of DNA, and 1x final CutSmart Buffer concentration. We placed the reaction in a 37°C water-bath for 3 hours. Then, we proceeded to purify the digestion products by running a 2% and 1% agarose gel for the insert and vector respectively. We excised the gel segments with a sterile scalpel under UV prep conditions using a BioRad Chemi-Doc (Hercules, CA). We extracted the DNA utilizing the Monarch DNA Gel Purification Kit (Ipswich, MA).

Ligation

We assembled the ligation reaction with insert to vector ratios of 3:1, 5:1, and 10:1. We used T4 DNA ligase and its respective buffer from NEB (Ipswich, MA) to carry out the ligation reaction for >16 hours. We diluted the reaction 10-fold in sterile nuclease-free water for the transformations.

Electroporation of NEB5α

We combined freshly thawed 25µL of electrocompetent *E. coli* cells with 2µL of 10-fold diluted ligation products in sterile nuclease-free water. We let the cells and DNA incubate on ice for 1 minute. We added 27µL of cells and DNA to a 1mm electroporation cuvette and electroporated using a BioRad preset protocol for E. coli and 1mm cuvettes (1.8 kV, 1 pulse) (Hercules, California). We resuspended the electroporated cells by adding 970µL of SOC outgrowth media to the cuvette and placed the liquid culture at 37µC for 1 hour for the cells to recover. We then plated the cell culture on Luria+1x Kanamycin (Millipore Sigma) plates and incubated them for 24 hrs. at 37µC.

Isolation of Cloned Constructs from NEB5a and *M. smegmatis*

We followed the Monarch Plasmid DNA Miniprep Kit Protocol (Ipswich, MA) to lyse, purify, and elute cloned constructs.

Transformation of M. smegmatis

We prepared *M. smegmatis* according to the instructions from BioRad (Hercules, California) [12]. We grew M. smegmatis in 1L of 7H9 media with 100mL ADC, 10mL 0.1M CaCl2, 2.5 mL of Tween 80, 20%, 1x Carbenicillin, 1x Cyclohexanone, 1x Kanamycin (Millipore Sigma) for 24 hrs. We then pelleted the cells and washed with 10% glycerol[12]. We stored 50µL aliquots of the cells at -80°C. For the transformation on ice, we added 2µL of DNA. After the incubation of the cells with DNA for a minute, we diluted the mixture with 150µL of sterile nuclease-free water. We added the mixture to a 2 mm cuvette from BioRad (Hercules, California) and electroporated according to the following conditions (1 pulse, 2.50 kV, 2.5_F, 500). After the pulse, we added 970µL of 7H9-ADC-Tween outgrowth medium, and incubated the cells for 2 hrs. at 37°C. We plated the culture on LB with 1x Carbenicillin, 1x Cyclohexanone, 1x Kanamycin (Millipore Sigma) and placed them at 37°C for 3-4 days until colonies were evident.

References

[1] Martha RJ Clokie, Andrew D Millard, Andrey V Letarov, and Shaun Heaphy. Phages in Nature. 1(1):31–45. [2] The Actinobacteriophage Database, JacoRen57. https://phagesdb.org/phages/JacoRen57/ [3] Mycobacterium phage JacoRen57, complete genome, MK279840.1. http://www.ncbi.nlm.nih.gov/nuccore/MK279840.1. [4] Victor Solovyev and Asaf A. Salamov. Automatic Annotation of MicrobialGenomes and Metagenomic Sequences. [5] Anne de Jong, Hilco Pietersma, Martijn Cordes, Oscar P. Kuipers, and Jan Kok. PePPER: a webserver for prediction of prokaryote promoter elements and regulons. 13(1):299. [6] M. A. Goldstein and R. H. Doi. Prokaryotic Promoters in Biotechnology. 1:105–128. [7] Murali D. Bashyam and Anil K. Tyagi. Identification and Analysis of Extended σ 10 Promoters from Mycobacteria. 180(9):2568–2573. [8] Riccardo Manganelli, Roberta Provedi, Sebastien Rodrigue, Jocelyn Beaucher, Luc Gaudreau, and Issar Smith. Sigma Factors and Global Gene Regulation in Mycobacterium tuberculosis. 186(4):895–902. [9] Matthew B. Avison. Measuring Gene Expression. The Basics. Taylor &Francis. [10] Lauren M. Oldfield and Graham F. Hatfull. Mutational Analysis of the Mycobacteriophage BPs promoter PR Reveals Context-Dependent Sequences for Mycobacterial Gene Expression. 196(20):3589–3597. [11] Marianne Poxleitner, et al. 9.1: Phage DNA Extraction. https://seaphagesphagediscoveryguide.helpdocsonline.com/9-1-protocol. [12] BioRad: Efficient Electro-transformation of Mycobacterium smegmatis.

Acknowledgements

I would like to thank Dr. Tolsma, Dr. Heeg, Dr. Noordewier for your wisdom and supervision. I would like to thank the Hatfull Lab at the University of Pittsburgh for the provision of the pLO86 and pLO87. Support for this project was provided by Northwestern College and Howard Hughes Medical Institute . I would also like to thank the directors of the NWC Honors Program Dr. Feenstra and Dr. Vonder Bruegge for the research opportunity.