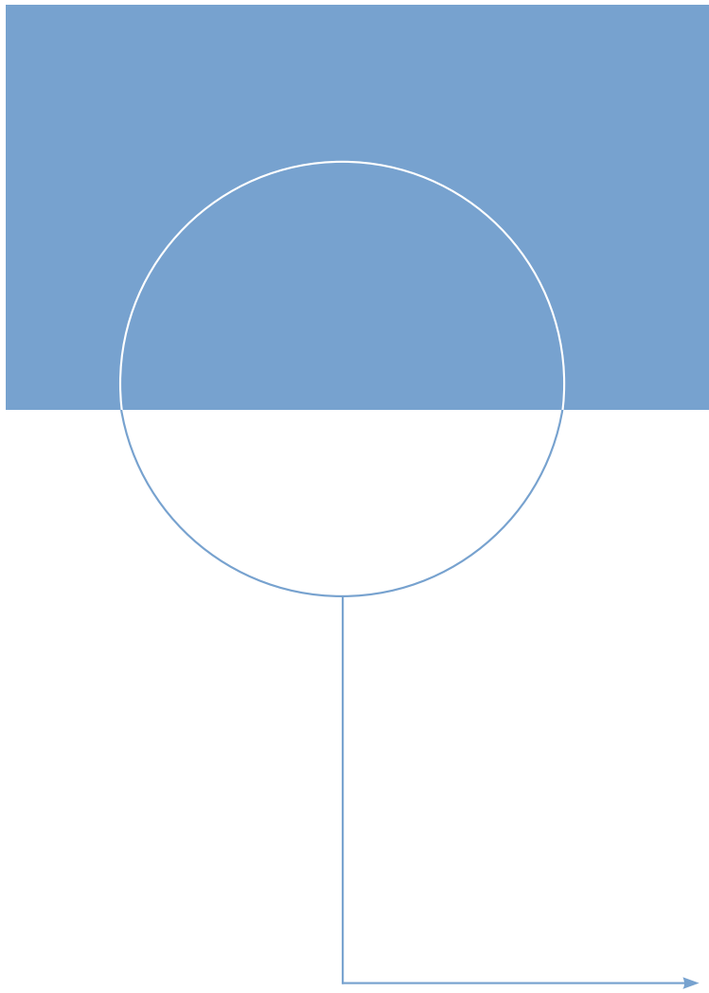

Master's thesis

2022

Arsal Javed

NTNU
Norwegian University of
Science and Technology
Faculty of Natural Sciences
Department of Biotechnology and Food Science

Arsal Javed

Investigating the binding of *Escherichia coli Sigma Factor 54* (σ^{54}) to specific DNA promoter sequences through *In vitro bulk* Transcription.

September 2022

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

NTNU

Norwegian University of
Science and Technology

Investigating the binding of *Escherichia coli Sigma Factor 54* (σ^{54}) to specific DNA promoter sequences through *In vitro bulk* Transcription.

Arsal Javed

Chemical Engineering and Biotechnology
Submission date: September 2022
Supervisor: Rahmi Lale
Co-supervisor: Essa Ahsan Khan

Norwegian University of Science and Technology
Department of Biotechnology and Food Science

NTNU

Norwegian University of Science and Technology

Investigating the binding of *Escherichia coli* Sigma Factor 54 (σ^{54}) to specific DNA promoter sequences through *In vitro bulk* Transcription.

Department of Biotechnology and Food Science
Trondheim
NTNU

TBT4901 – Biotechnology Master's Thesis

Supervisors:
Rahmi Lale
Essa Ahsan Khan

Arsal Javed

September 2022

Acknowledges

This project marks the end of my master's degree within Chemical Engineering and Biotechnology (MIKJ) at the Department of Biotechnology and Food Science at the Norwegian University of Science and Technology (NTNU). The practical aspect of this thesis was conducted at the Molecular Genetics Lab under the Microbial Biotechnology Scientific Research Division at the Department of Biotechnology and Food Science. During this time, I had the privilege of being a part of the research group Lale Lab under the supervision of Rahmi Lale. During this period, I have also had the privilege of learning several cutting-edge methodologies, which have contributed to the overall development of my knowledge within the field of Biotechnology with special focus on molecular and synthetic biology. The experience gained through this research project has definitely broaden my perspectives as a future scientist. It was a pleasure for me to work in such a competent environment full of friendly minded professionals willing to share their knowledge.

I would like to thank my supervisor Rahmi Lale and my co-supervisor Essa Ahsan Khan for their constant support, patience, and cooperation and for giving me the opportunity to be a part of the research group as a master student. Thank you for the constant support during the research period and for answering my inquiries regarding the research work. I would also like to thank my fellow researcher Maxime Fages-Lartaud for his help, support and guidance during lab work and fellow student and friend Dennis Raphael Winter for his support and friendship.

Last, but certainly not least, I would like to thank my family, friends and fellow research students for the moral support and comfort and for cheering me up during this period in my life.

Thank you all.

Norwegian University of Science and Technology (NTNU)
Faculty of Natural Sciences
Department of Biotechnology and Food Science
Trondheim, 9 September 2022
Arsal Javed

Abstract

The project aims to investigate the transcriptional regulation in *Escherichia coli* by σ54. Prokaryotic promoters are known to have UP elements and a core promoter region that is positioned upstream of the coding sequence (CDS). During the transcription process, the sigma factor recognizes a specific promoter sequence and aids in the specific binding of the core RNA polymerase (RNAP). There are many σ factors that belong to two major categories called σ54 and σ70. The present knowledge surrounding the binding of sigma (σ) factors to specific promoter sequence is filled with gaps that needs to be further researched and understood. Synthetic biology methods will be applied to bring together the *Escherichia coli* sigma 54 (σ54), Transcriptional activator protein and the core RNA polymerase for *In vitro* transcription. *Escherichia coli* sigma 54 (σ54) and Transcriptional activator protein will be expressed. The synthetically produced RNA transcript from the *in vitro* transcription will then be further sequenced using next generation sequencing, to confirm the binding of *Escherichia coli* sigma 54 (σ54) to specific DNA promoter sequences.

Through extensive lab work, Sigma Factor 54 Plasmid with T7 Promoter was cloned, and the colonies confirmed thorugh colony PCR. The cloned RpoN (Sigma Factor 54) plasmid was also successfully expressed with the T7 promoter system. PspF - Transcriptional activator Protein plasmid with T7 Promoter System was also cloned but no visible expression was shown. *E. coli* pO Promoter Sequence Plasmid with mCherry was successfully cloned and the cloning confirmed through colony PCR. The N200 Random Nucleotide DNA Library with Mango Aptamer was successfully cloned using golden gate cloning and the cloning was confirmed through colony PCR. The *in vivo* transcription results illustrated an unsuccessful reaction as both the induced sample and the negative control showed significant levels of similarity in mCherry fluorescence levels, hence indicating that the *E. coli* pO promoter sequence was not transcribed and the mCherry was not expressed in the pHH100-mCherry-S54pO plasmid. The IVT in bulk yielded a very high concentration for the induced sample with the inclusion of TOI-Biotin dye, giving the highest concentration of the RNA transcript. This result hence indicates a confirmed successful binding and expression. The synthetically produced RNA transcript from the *in vitro* transcription was then further sequenced using next generation sequencing, to confirm the binding of *Escherichia coli* sigma 54 (σ54) to specific DNA promoter sequences.

Sammendrag

Prosjektet tar sikte på å undersøke transkripsjonsreguleringen i *Escherichia coli* ved σ54. Prokaryote promotere er kjent for å ha UP-elementer og en kjernepromoterregion som er plassert oppstrøms for den kodende sekvensen (CDS). Under transkripsjonsprosessen gjenkjenner sigmafaktoren en spesifikk promotersekvens og hjelper til med den spesifikke bindingen av kjerne-RNA-polymerasen (RNAP). Det er mange σ-faktorer som tilhører to hovedkategorier kalt σ54 og σ70. Den nåværende kunnskapen rundt bindingen av sigma (σ)-faktorer til spesifikk promotersekvens er fylt med hull som må undersøkes og forstås videre. Syntetiske biologimetoder vil bli brukt for å bringe sammen *Escherichia coli* sigma 54 (σ54), transkripsjonsaktivatorprotein og kjerne - RNA-polymerasen for *in vitro* - transkripsjon. Escherichia coli sigma 54 (σ54) og transkripsjonsaktivatorprotein vil bli uttrykt. Det syntetisk produserte RNA-transkriptet fra *in vitro* - transkripsjonen vil deretter bli sekvensert videre ved hjelp av neste generasjons sekvensering, for å bekrefte bindingen av *Escherichia coli* sigma 54 (σ54) til spesifikke DNA-promotersekvenser.

Gjennom omfattende laboratoriearbeid ble Sigma Factor 54 Plasmid med T7 Promoter klonet, og koloniene bekreftet gjennom koloni-PCR. Det klonede RpoN (Sigma Factor 54) - plasmidet ble også vellykket uttrykt med T7-promotersystemet. PspF - transkripsjonsaktivator Proteinplasmid med T7 Promoter System ble også klonet, men ingen synlig ekspresjon ble vist. *E. coli* pO Promoter sekvens plasmid med mCherry ble vellykket klonet og kloningen bekreftet gjennom koloni-PCR. N200 Random Nucleotide DNA Library med Mango Aptamer ble vellykket klonet ved bruk av golden gate-kloning og kloningen ble bekreftet gjennom koloni-PCR. *In vivo* transkripsjonsresultatene illustrerte en mislykket reaksjon ettersom både den induserte prøven og den negative kontrollen viste signifikante nivåer av likhet i mCherry-fluorescensnivåer, og indikerte derfor at *E. coli* pO-promotersekvensen ikke ble transkribert og mCherry ikke ble uttrykt i pHH100 -mCherry-S54pO-plasmid. IVT i bulk ga en veldig høy konsentrasjon for den induserte prøven med inkludering av TOI-Biotin-fargestoff, noe som ga den høyeste konsentrasjonen av RNA-transkriptet. Dette resultatet indikerer derfor en bekreftet vellykket binding og uttrykk. Det syntetisk produserte RNA-transkriptet fra *in vitro* - transkripsjonen ble deretter sekvensert videre ved bruk av neste generasjons sekvensering, for å bekrefte bindingen av *Escherichia coli* sigma 54 (σ54) til spesifikke DNA-promotersekvenser.

Table of Contents

List of Abbreviations

Abbreviations:

E. coli	***Escherichia coli***
DNA	Deoxyribonucleic acid
SDS-PAGE	Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
LB	Lysogeny broth
Kan	Kanamycin
OD	Optical Density
SOC	Super Optimal Broth
ATP	Adenosine triphosphate
mRNA	Messenger RNA
RNAP	RNA Polymerase
RNA	Ribonucleic acid
CDS	Coding sequence
σ54	Sigma Factor 54
σ70	Sigma Factor 70
UTR	Untranslated Region
SD	Shine-Dalgarno
IHF	Integration Host Factor

Units of Measurement:

°C	**Degrees celsius**
rpm	Revolution per minute
µl	Microlitre
mg	Milligram
ml	Milliliter
M	Molar
g	Gram
nm	Nanometer
bp	Base Pair
ng	Nanogram
Pmol/µl	Picomole/Microliter
kDa	Kilo Daltons
%	Percentage
µg	Microgram
µg/ml	Microgram per millilitre
H	Hour
Min	Minute
ms	Millisecond
V	Voltage

1. Introduction

One of the most groundbreaking discoveries in the fields of molecular biology is that nucleic acids, DNA and RNA, carry the biological information. The knowledge of the chemical makeup of DNA has therefore offered a mechanistic basis for a variety of other phenomena, including cell proliferation and division in cancer. Nucleic acids are the molecules that encode genetic information, and the regulation of the genes is considered one of the most fundamental mechanisms in all forms of life *(Clark et al., 2018)*. Although, the list of sequenced genomes is constantly growing, our knowledge of the regulation of these genomes' genes at the sequence level is not keeping up with this rapid accumulation. We still don't know whether and how more than half of the genes are regulated, not even in *Escherichia coli*, the model organism that has been most thoroughly examined in molecular biology research (*Gama-Castro, 2016*).

Such knowledge is even more scarce in other model organisms. A genome-wide understanding of the transcriptional regulation of promoters is thus a problem that has not been solved in any area of life. Sigma Factors are in charge of identifying different *cis*-acting sequence elements in the promoter regions. These are part of a complex called the RNA polymerase (RNAp) holoenzyme. They plan the overall pattern of promoter-specific transcription's activation or repression. As a result, they are the primary regulatory components that regulate various kinds of promoters in response to intracellular or extracellular inputs *(Browning, 2016), (Iyer, 2020).*

Despite playing a crucial role, each bacterial sigma factor's knowledge on the *cis*-acting DNA sequence elements in the promoter regions is still far from clear. This knowledge gap restricts our ability to accurately predict and annotate promoter sequences, link them to their corresponding sigma factors, and, as a result, constricts our understanding of transcriptional control in bacteria as a whole. By addressing the promoter dependency of 21 sigma factors from three different bacterial species, *E. coli*, *Pseudomonas putida*, and *Bacillus subtilis*, we hope to increase our knowledge and understanding of the transcriptional control in bacteria within the aforementioned constraints. The chosen approaches entail a special fusion of high-throughput *in vitro* and *in vivo* screening efforts with a massive data creation and *in silico* data processing pipelines.

The overall project's primary goal is to decode *cis*-acting DNA sequences that are particular to the sigma factor through *in vivo* and *in vitro* experimental work in order to produce massive

data sets employing the three bacterial species. The second key goal is to create *in silico* data analysis pipelines that use statistics and machine learning to produce algorithms that can determine the number of matching transcripts and the precise *cis*-binding motifs for sigma factors. Establishing a cloud-based service with the intention of using it for more extensive scientific results, data, and tools would be the third and final key goal. An overview of the three goals is shown in the diagram below (figure 1).

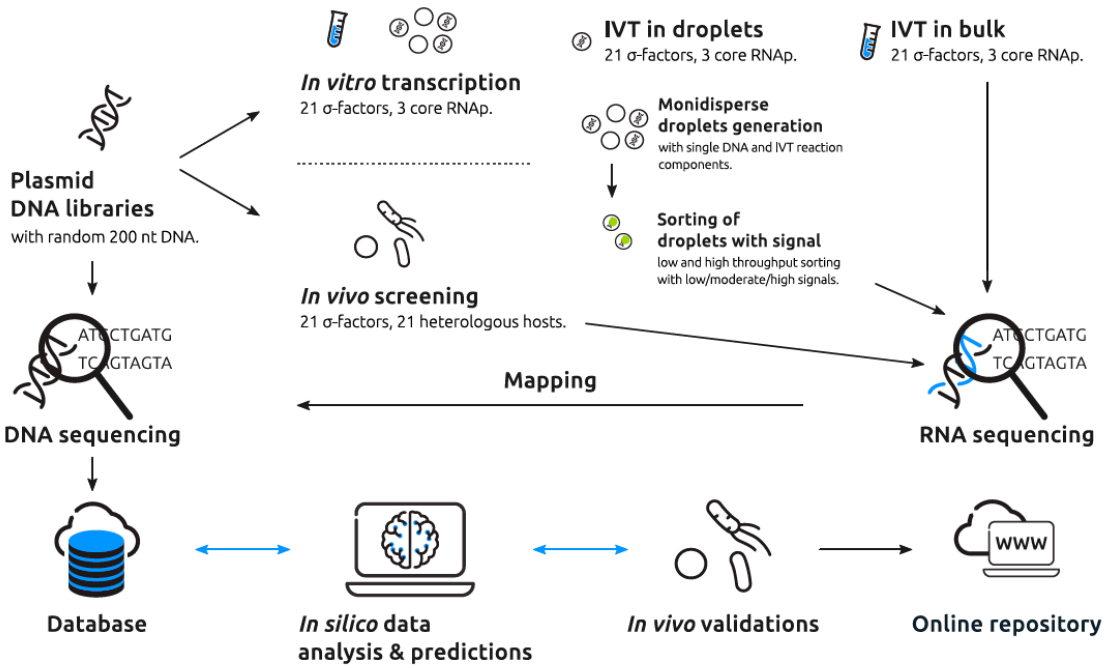

Figure 1: Summary of the experiments involved in the project. *Plasmid DNA libraries with random 200nt DNA will be used in In vivo and in vitro transcription experiments. The DNA libraries and the RNA transcripts generated will then be sequenced. The in-silico platforms will be used for data analysis (Lale, 2020).*

2. Theory

2.1. Bacterial Transcription Machinery

To benefit from the genetic information that is coded inside the DNA molecule, the expression of the genes is necessary. This procedure is achieved through the process of transcription. Although, the DNA molecule that contains the original copy of the genetic code is employed to store genetic data, it is not directly used as a source of instruction to power cells. As an alternative, RNA-based functioning copies of the genes are utilized, and transcription is therefore known as the process of information being transferred from DNA to RNA. The processing of the genetic data results in the formation of mRNA molecules that are also known as transcripts and carry genetic information that encode for a protein molecule. The process follows several steps that result in the successful transcription of a gene. The double stranded DNA is pulled apart exposing the sense and antisense strands separately (also known as the coding and template strands respectively). This allows for the new RNA strand to be synthesized by the RNA polymerase enzyme. The RNA polymerase binds to the DNA exactly at the starting point of a gene and unfolds the double helix DNA and uses it as a template to produce the RNA molecule. The RNA sequence that is produced is complementary to the antisense strand of the DNA molecule and this makes it identical to the sense strand of the DNA molecule as the synthesis process is conducted from the 5´to the 3´direction. The Figure 2 below illustrates the transcriptional machinery in prokaryotes *(Clark et al., 2018)*.

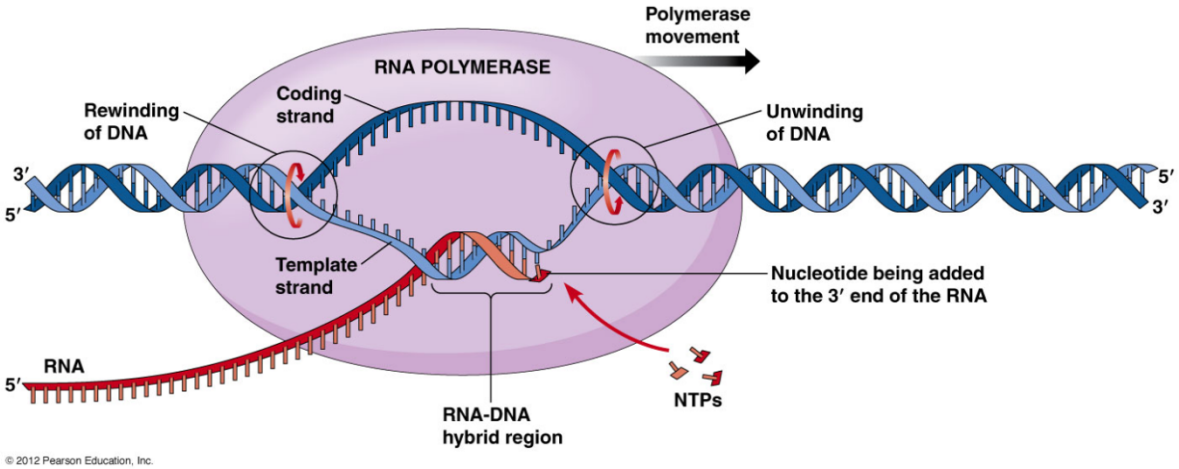

Figure 2: The Prokaryotic Transcriptional Machinery*. The illustration above depicts the different components that participate in the transcription process. The transcription process is completed in three subsets called the initiation, elongation, and the termination. The process of synthesis of RNA is initiated by copying the coding strand of DNA and during the elongation process proteins such as RNA polymerase II move in the direction from 5´to 3´direction. NTPs are used to add*

the nucleotides to the 3′end of the RNA, resulting in the production of an RNA strand. The illustration above depicts the RNA polymerase, NTPs, the coding strand, the template strand, and the RNA-DNA hybrid region (Karki, 2017).

The prokaryotic transcription is a complex process that varies in both eukaryotic and prokaryotic organisms. The difference can be seen in the details for binding, initiation, and the regulation of the transcriptional machinery. Each gene contains a region that regulates the transcription of the DNA. This regulatory region is called the promotor region and it is in this region the RNA polymerase binds and starts the process of transcription. This region is known as the upstream region as it's a stretch of DNA that is infront of the gene of interest at the 5′end. The gene of interest is composed of the 5′UTR region, mRNA region and the 3′UTR. The 5′UTR and the 3′UTR are known as untranslated regions as they are not translated during the process of translation that results in the synthesis of protein. Therefore, the initial base of the mRNA for a gene encoding for a protein is not directly identical with the first base of the whole sequence encoding for a protein. This can be visualized using the illustration below (Figure 3).

The transcriptional process initiates when the RNA polymerase binds to the promoter region of a specific gene. This bacterial RNA polymerase consists of two vital components called the core enzyme and the sigma subunit. The core enzyme is made up of α, α, β, β' and ω. These two components combined result in the formation of an enzyme called, the holoenzyme. The specific functions of each of the component is that the core enzyme is directly responsible for the synthesis of RNA, whereas the sigma subunit is responsible for recognizing the promoter. Specifically, the sigma subunit recognizes two specific sequences in the promoter called -10 sequence and the -35 sequence *(Clark et al., 2018).*

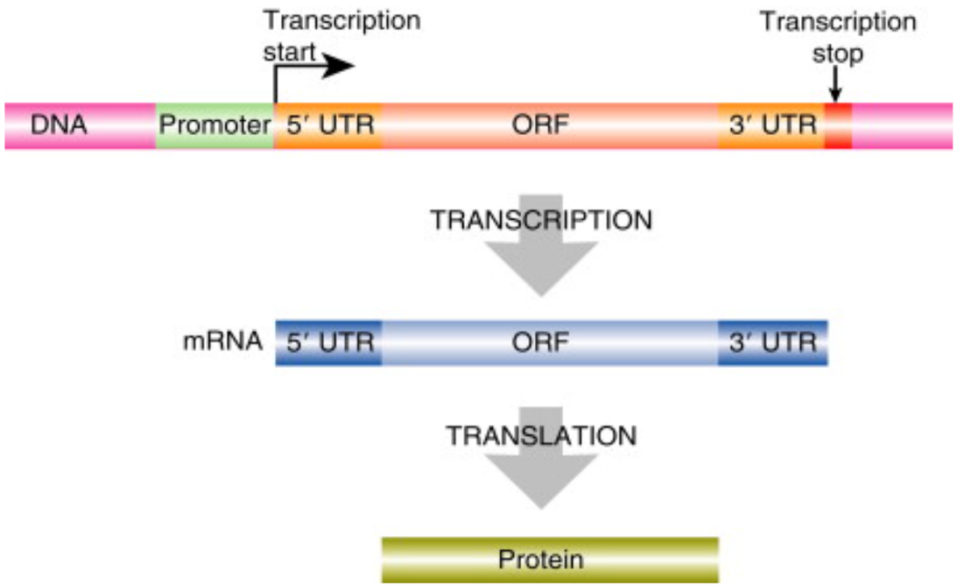

Figure 3: Upstream and Downstream regions of a typical Gene. *Upstream regions are regions positioned before the start point of mRNA transcription. Whereas Downstream regions are regions positioned after the start point of mRNA transcription. The transcriptional procedure is initiated at the end of the promoter sequence and stops at the end of the coding sequence. The resulting transcript contains more information than is required to produce a protein. Whereas the 5ÚTR and 3´UTR are regions in mRNA that are not used to make the final protein but contain important regulatory elements (Clark et al., 2018) (David P. Clark, 2016).*

After the formation of a new RNA strand, the sigma subunit fulfills its purpose and detaches from the DNA. This results in the core enzyme being left behind so that it can elongate the mRNA. As mentioned above, the bacterial core enzyme consists of five subunits, α, α, β, β' and ω. The β- and β'- subunits comprise the catalytic site of the enzyme, the α- subunit is required for the assembly and the recognition of promoters and the ω- subunit binds to the β'- subunit which stabilizes and aids its assembly into the core enzyme complex. As the end of the promoter sequence initiates the transcriptional process, there is also a special terminator sequence at the end of the coding sequence. This is called the terminator and is positioned in the template strand of DNA. The terminator consists of two complementary sequences that are separated by approximately 10 bases, and this results in the formation of a stem and loop or "hairpin" structure. Figure 4 below illustrates this structure. As a result, during the transcription procedure the stem and loop of the mRNA transcript causes the destabilization of the interaction between the mRNA with the RNA polymerase. This leads to instability in the complex, resulting in mRNA and DNA being released. Once the DNA and the RNA are separated from each other at the terminator, the RNA polymerase detaches and is released. There are two types of terminators, Rho-dependent terminators, and Rho-independent

terminators. The rho-dependent terminators require the aid of the Rho protein when separating the RNA polymerase from the DNA, whereas the Rho-independent terminators do not. In terms of *E. coli*, most terminators are Rho-independent *(Clark et al., 2018).*

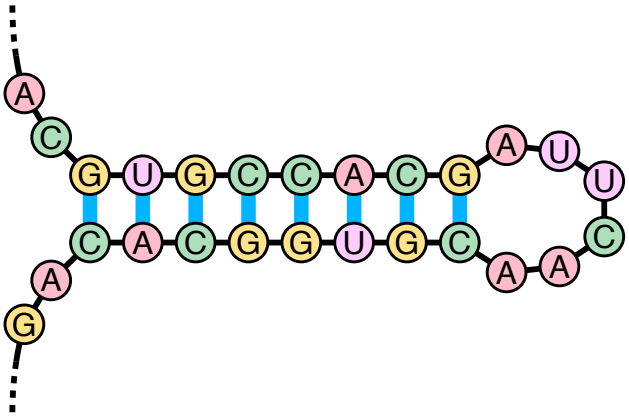

Figure 4: Stem and loop or "hairpin" structure of an RNA. *The stem and loop structure are formed when the complementary bases lead to the formation of the stem of the hairpin and the intervening bases leading to the formation of the loop (contributors, 2022).*

The transcriptional machinery often requires the need for activator proteins to regulate the gene transcription. Constitutive genes are continually expressed all the time. This means that they are always recognized by the sigma subunit of RNA polymerase and are therefore expressed under all types of conditions available. However, genes that are only required under specific conditions tend to have poor recognition sequences in the -10 and -35 regions of their promoter regions. Therefore, activator proteins are required in such cases as the promoter isn't recognized by the sigma subunit. The activator protein may stimulate the transcription of one or more genes, as a group of genes that are recognized by the same activator protein are expressed given that the conditions are similar *(Clark et al., 2018).*

3. Literature Review

3.1. Bacterial Promoters and consensus sequences

Microorganisms' diverse array of distinctive traits depend on the careful regulation of gene expression and promoters are at the heart of this outstanding control strategy. Promoters are DNA regions that facilitate the transcriptional regulation of gene expression and act in response to environmental cues through a complicated regulatory network. The interplay of cis-regulatory components and trans-acting variables results in this tightly coordinated regulation. Cis-regulatory elements are sections of DNA that are classified as non-coding and contain target-binding sites for trans-acting factors, such as transcription factor proteins, e.g., enhancers and promoters. As mentioned previously, the cellular machinery that recognizes promoters in conjunction with auxiliary factors is called RNA Polymerase. Promoters are composed of a number of components and are found upstream of coding sequences (CDS). This can be illustrated in figure 5. The cis-acting motifs are recognized by the RNA polymerase sigma subunit, which draws RNA polymerase and causes transcription to begin. Sigma factors in bacteria can be divided into two groups: Sigma factor 54 and Sigma factor 70 *(Zhang, 2019).* Alternative extra cytoplasmic function (ECF), a subset of Sigma 70, is another sizable group of sigma factors that waits to take action until it is necessary. We presently lack a complete understanding of the cis-acting DNA sequence components in bacterial promoters, despite the crucial part the sigma factor plays in the operation of RNA polymerase and transcription. Due to this knowledge gap, we have only a partial grasp of how the genome-wide regulatory interactions work, gene regulatory networks are only partially understood, and a number of significant biological phenomena are still unknown *(Sriram Kosuri, 2013).*

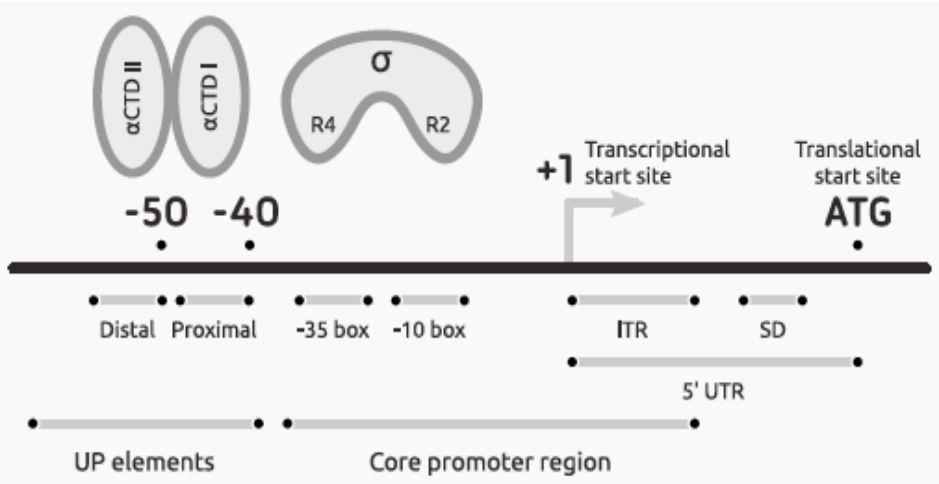

Figure 5: Schematic representation of bacterial promoters*. Sigma 70 subunit regions 2 (R2) and 4 (R4) engage with the -10 and -35 promoter motifs, respectively, as do RNA polymerase α-subunits, CTD I and II, with the proximal and distal subsites of the UP-element (UP). The acronyms stand for, initially transcribed region (ITR); Shine-Dalgarno sequence (SD); 5′ Untranslated Region (5′UTR) (Lale, 2020).*

3.1.1. 5′ Untranslated Region (5′UTR)

As explained above, the 5′UTR is known as the untranslated region as it is not translated during the process of translation. However, the region holds greater significance when it comes to gene expression as the elements within the 5′UTR of mRNA contribute to the gene expression of many genes and protein production *(Clark et al., 2018).* In terms of the DNA level, the DNA region corresponding to the 5′UTR is involved in transcript formation. This is as a result of the interaction between the promoter and the initially transcribed sequences (ITS), which in *Escherichia coli* cover the first 15 nucleotides *(Craig, 1998), (Goldman, 2009), (Hsu, 2006).* However, in terms of the mRNA level, the creation of secondary structures *(Kudla, 2009), (Osterman, 2013), (de Smit, 1990),* interactions with other molecules e.g., proteins *(Hajnsdorf, 2012)*, metabolites *(Mandal, 2004), (Coppins, 2007)* and small RNAs (*Repoila*, 2009), as well as ribosome binding and translation initiation, all affect transcript stability and translation at the mRNA level *(M., 2005).* The Shine-Dalgarno (SD) region and the 5′end of the coding sequence after the start codon *(Steitz, 1975)* are primarily involved in the translation initiation process *(Simone Balzer Le, 2020), (Laila Berg, 2009)*. Hence, In accordance with research, the regulatory process of transcription and translation relies heavily on the specific combination of traits that are found in the promoter, 5′ untranslated region (UTR), and 5′-end of the coding sequences (CDS). The influence of the 5' UTR sequence composition on transcript abundance and translation rates is highlighted by numerous research. According to two research articles *(Berg, 2009), (Lou, 2012),* point mutations within the 5' UTR not only impact translation rates but also affect transcript abundance, possibly by producing a more stable transcript. The

intersection of the promoter and 5' UTR, as well as the 5' UTR and CDS, is another area for gene expression optimization. According to another research article *(Mutalik, 2013),* these junctions have an impact on gene expression and can be used to optimize and control the amount of gene expression. Another study *(Mirzadeh, 2015)*, has also provided evidence for the significance of junction optimization for gene expression. The six nucleotides before the start codon were randomly assigned, then codons 2 and 3 were changed to any possible synonymous codons using degenerate primers. Consequently, there were multiple orders of magnitude variations in the amounts of gene expression when the nucleotides at the 5' UTR and CDS were altered *(Lisa Tietze, 2021).*

3.1.2. Artificial Promoters

One of the most essential parts of synthetic biological systems is promoters. By regulating the rate of transcription and turning on and off gene transcription in response to diverse stimuli, they have an impact on how the system behaves. A few well-characterized natural promoters have been used to stimulate gene expression *(Dahl, 2013), (Huo, 2018)*. There are, however, a finite number of natural promoters that can interact with the host's endogenous transcriptional machinery and interfere with the production of its genes. This implies that the selection and layout of the nearby genetic constructs affects their transcription, which may result in regulatory failures for the built biological systems. These issues can be resolved with synthetic (artificial) promoters. It is possible to quickly create novel promoters that do not occur in nature by joining synthetic fragments of promoter components *(Blazeck, 2013)*. These promoters have the potential to work more strongly than those produced naturally and to be independent of endogenous transcription control (Love, 2016)*, (Portela, 2017)*. The spacer sequence and the sigma factor binding sites are the sole elements seen in most synthetic promoters *(Jervis, 2019)*. The spacer sequence is often a 17 bp random segment sandwiched between the two binding sites, with the sigma factor binding sites typically being derivatives of the naturally determined -10 and -35 boxes. The synthetic core promoters can attain greater transcriptional activity and wider dynamic ranges as compared to natural promoters *(Yim, 2013), (Love, 2016), (Liyuan Jin, 2019)*. Since promoter sequences are functionally varied in multiple different organisms, it is highly useful to provide the host organism with artificial consensus promoters for expression of cloned genes *(Clark et al., 2018).* Highly efficient and regulatable promoters have a high demand in research as they can work in multiple organisms and when stimulated, they have a role in gene expression through the regulation of the transcriptional machinery

(Alisa Yurovsky, 2018), (Tanya Tschirhart, 2019). In terms of the project, this enables us to create 200nt long artificial promoter sequences with high-sequence variation that otherwise cannot be achieved by the use of natural promoters.

3.2. The Sigma Factors: σ54 and σ70

As mentioned previously, sigma factors are a vital part of the transcriptional machinery as they determine the promoter specificity. The majority, if not all, of the determinants for promoter recognition and the formation of open complexes are provided by sigma factors, but only in complex with the other members of the RNA polymerase (RNAP) *(Gruber, 2003)*. There are two different types of sigma factor families called the sigma factor 70 (σ70) and the sigma factor 54 (σ54) and most transcription in rapidly growing cells is mediated by the sigma factor 70 family. The majority of bacterial genomes also encode for a number of alternative sigma factors, which control particular groups of genes in response to various stresses and changes in development, and hence serve as the most fundamental method of bringing about significant transcriptional alterations. The majority of alternative sigma factors are also members of various subgroups of the sigma factor 70 family. The sigma factor 70 family, as a component of RNAP, can carry out open complex stabilization on their own. However, Sigma 70 class sigma factor regulation is primarily accomplished by altering the expression of sigma 70 and anti-sigma factors that bind to and inhibit them (*Wedel*, *1995*). On the other hand, members of the sigma factor 54, need the assistance of additional activators (activator proteins as mentioned previously) from the AAA+ ATPase family to unwind the DNA *(T.R.Hughes, 2011)*. Sigma factor 54 transcription has more control over gene expression and a larger dynamic range as a result of this activation requirement. As a result, rapid and accurate responses to environmental change are frequently produced via sigma factor 54 dependent gene expression *(Kazmierczak, 2005)*. Transcriptional initiation with sigma factor 54 is different compared to the transcriptional initiation by other sigma factors due to activation requirement. Sigma Factor 54 Transcriptional initiation occurs in four stages, assembly, activation, opening and elongation. Around 100 bp upstream of the gene's start site, the enhancer regions are where the AAA+ ATPase transcriptional activators interact as dimers. The active conformation, a hexameric ring capable of hydrolyzing ATP, will be formed by two sets of dimers aggregated at the enhancer region and a third dimer in solution. The transcriptional activator comes into close contact with sigma 54 when the Integration Host Factor (IHF) binds in the region between the enhancer sequence and bends the DNA severely. The hexameric transcriptional activators contacts sigma

54's N-terminus once the components have come together. This interaction requires the presence of a conserved GAFTGA motif close to the hexamer's central pore *(Lakshmi Tripathi, 2014)*. In addition, activation needs to occur at least once throughout an ATP hydrolysis cycle. The sigma 54 subunit of activated RNAP holoenzyme, which consists of core RNAP, can open DNA close to the promoter region of -12. The sigma factor 54 recognizes a -24 and -12 region within the consensus sequence. Core RNAP opens and starts elongating RNA from the DNA template at the +1 gene location. Core RNAP can transcribe RNA from opened DNA on its own, but it requires activated sigma 54 to open the DNA. Following DNA opening, sigma 54 eventually detaches from the central RNAP enzyme and elongation continues *(Wemmer, 2022)*. The figure 6 below illustrates a summary of this mechanism.

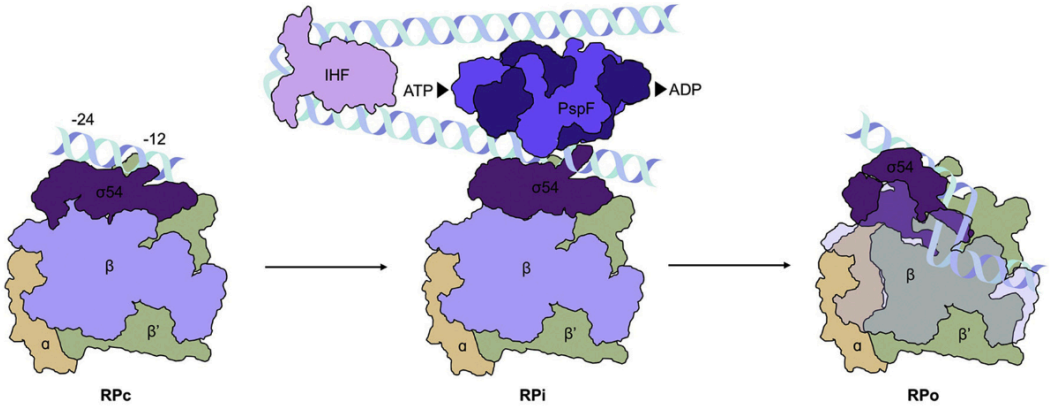

Figure 6: Transcriptional initiation with sigma factor 54*. Diagram illustrating the fundamental process by which the closed (RPc) and open (RPo) DNA complexes initiate transcription through the action of sigma 54. To see the open transcription bubble, theβ subunit of RPo is rendered translucent (Amy E.Danson, 2019).*

3.3. Bacterial RNA polymerase holoenzyme

The most important regulatory step in the expression of bacterial genes is transcription by RNA polymerase. The five protein subunits (as explained previously) that make up the core RNA polymerase are, two each of the alpha, beta, beta', and omega subunits and these can synthesize mRNA from a DNA template. Hence, the sigma factor, a sixth modular subunit, is necessary to melt the double-stranded DNA and produce a single-stranded DNA template for transcription. The complex is therefore referred to as the RNA polymerase holoenzyme when sigma factor attaches to the core RNA polymerase *(Wemmer, 2022)*.

The core enzyme must bind a sigma factor in order to start promoter-specific DNA transcription because this factor helps the polymerase go to particular promoters. For instance, *in vitro* testing of RNAP's transcriptional activity is simple when DNA and transcription factors

are present. To investigate the roles of the different subunits, RNAP can be simply recreated *in vitro* from its component subunits using either wild-type or mutant proteins *(Fujita, 1996)*. An overview of the RNAP structure and assembly is illustrated in the figure below (figure 7) *(Catherine Sutherland, 2018)*.

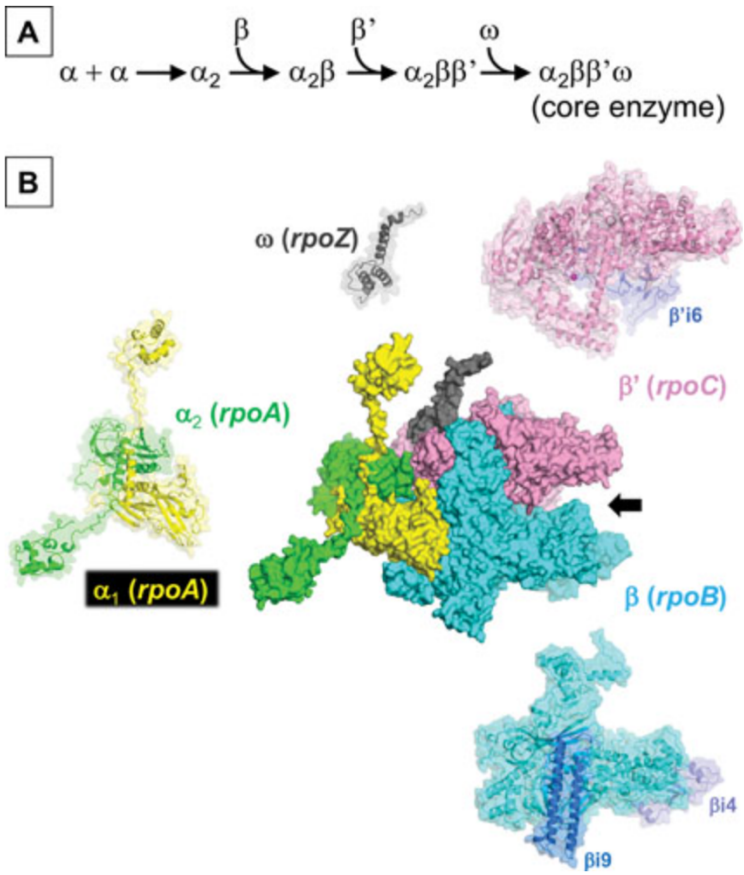

Figure 7:: Overview of RNAP assembly and structure. *The RNAP assembly is depicted in (A), and the molecular structure overview of the E. coli RNAP core enzyme is represented in (B). The black arrow represents the DNA binding region. The individual subunits are represented with partially transparent surface (Catherine Sutherland, 2018).*

3.4. Bacterial Transcriptional activator

The big AAA+ protein family, which includes the activators of sigma factor 54 - holoenzyme, uses ATP binding and hydrolysis to modify its substrates. Most of the central domains of sigma factor 54 activators, which contains ATP-binding and hydrolyzing determinants, conforms to the AAA core structure *(Chaney, 2001).* The RNA polymerase - sigma factor 54 complex attaches to the promoter DNA but is unable to proceed with transcription, in contrast to the RNA polymerase - sigma factor 70 holoenzyme, which is capable of transcription. The RNA polymerase - sigma factor 54 complex is changed from a "closed" (attached but stalled)

complex to a "open" complex that may actively transcribe when the activator protein PspF (a AAA+ family protein) binds to a bacterial enhancer-like region in the promoter and hydrolyzes ATP *(Jones, 2009)*. The expression of the psp operon depends on the 37 kDa constitutively active activator protein, PspF, which is encoded by the pspF gene. The family of enhancer binding proteins (EBP) includes PspF. It has a central catalytic and a DNA-binding domain, but unlike most family members, it lacks an N-terminal regulatory domain. In order to start transcription, σ54-RNAP must connect with an activator protein that catalyzes the isomerization of a closed to an open promoter complex. σ54-RNAP forms a closed complex with the promoter. For transcription to begin, the ATP bound in the central domain must be hydrolyzed. Once it has been created, the open complex can be kept going without the activator *(Jovanovic, 1999).*

3.5. Bacterial Host Organisms

Because of its capabilities for growth and expression, *E. coli* is a well-known and preferred choice. Under ideal circumstances, the bacterium may grow very quickly and is well renowned for possessing a distinctively well-known genomic system *(Jia, 2016).* A member of the *Enterobacteriaceae* family, *E. coli* is classified as a gram-negative prokaryote bacterium. *E. coli* BL21 is a manufacturing strain that is often utilized *(D., 2015)*, *(Hausjell, 2020)*. There are several advantages to using *E. coli* for laboratory investigations, including the fact that it has a smaller genome than most other bacteria and that it can grow quickly on affordable substrate media. This makes it possible to manipulate genetic information in controlled laboratory studies perfectly. Recombinant DNA technology, transcription, translation, molecular cloning, and heterologous protein production are examples of experiments utilizing genetic information manipulation *(D., 2015), (F., 1999).* There are several drawbacks to utilizing *E. coli* as an expression host organism in laboratory research, as well as some drawbacks to using *E. coli* as a bacterial host for the creation of recombinant proteins. Because recombinant proteins may be expressed at high quantities in *E. coli*, one drawback of employing this organism is the potential formation of highly aggregated proteins. The prokaryote organism's cytoplasm naturally produces inclusion bodies, which are collections of aggregated proteins *(Palmer, 2012).* As a result, when it comes to molecular biology research, various choices involving gram-positive or gram-negative bacteria are also considered.

4. Aim of Study

The project aims to investigate the transcriptional regulation in *Escherichia coli* by σ54. Prokaryotic promoters are known to have UP elements and a core promoter region that is positioned upstream of the coding sequence (CDS). During the transcription process, the sigma factor recognizes a specific promoter sequence and aids in the specific binding of the core RNA polymerase (RNAP). There are many σ factors that belong to two major categories called σ54 and σ70. The present knowledge surrounding the binding of sigma (σ) factors to specific promoter sequence is filled with gaps that needs to be further researched and understood. Synthetic biology methods will be applied to bring together the *Escherichia coli* sigma 54 (σ54), Transcriptional activator protein and the core RNA polymerase for *In vitro* transcription. Escherichia coli sigma 54 (σ54) and Transcriptional activator protein will be expressed. The synthetically produced RNA transcript from the *in vitro* transcription will then be further sequenced using next generation sequencing, to confirm the binding of *Escherichia coli* sigma 54 (σ54) to specific DNA promoter sequences.

5. Materials and Methods

5.1. Growth conditions, Bacterial Strains and Plasmid Vectors

E. coli cultures were grown with the right growth conditions. A detailed explanation of all the growth mediums and antibiotic solutions used for the optimal growth conditions during inoculation can be found in Appendix A and Appendix B.

5.1.1. E. coli Growth Conditions

The 2 *E. coli* strains, $DH_5\alpha$ and BL21, were grown in liquid LB medium. The liquid pre-cultures were grown in 15ml reagent tubes at a temperature of 37°C and rotation of 225rpm with shaking. Antibiotics such as, ampicillin and kanamycin, were added to the medium when required for selective purposes (depending on the plasmid being produced). The antibiotics were diluted from 50mg/ml to 50µg/ml. Appendix A and Appendix B gives an overview of various components involved in the growth mediums and the antibiotic stock solutions.

5.1.2. Bacterial Strains and Plasmid Vectors

The tables below summarise the bacterial strains and plasmid vectors used in the project. The plasmid maps used in this project thesis are shown in the figures below. Appendix D shows the corresponding primer sequences used in the amplification process for the different plasmid vectors.

Table 1: Bacterial Strains used and their purpose in the project thesis.

Bacterial Strain	Purpose
***E. coli* $DH_5\alpha$**	Cloning host strain
***E. coli* BL21**	Expression host strain

Table 2: Plasmid Vector Strains used and their purpose in the project thesis.

Plasmid Vector	Purpose
pHH100-mCherry-S54p0	Original *E. coli* Promoter Plasmid Sequence expressing mCherry. This was used for the *In vivo* experiment.
Lv3Δ-T7-PspF $_{(1-275)}$-AmpR	PspF activator protein gene expressing plasmid sequence with ampicillin resistant gene incorporated. Residues 1-275 is directly

	responsible for ATP hydrolysis and transcriptional activation *(Patricia Bordes, 2003)*.
pSEVA2311-N200-Mango	N200 DNA Promoter Library expressing Mango aptamer. This synthetic promoter was used for both In vitro bulk and In vivo transcription.
Lvt3t_T7_RpoN_AmpR	Plasmid expressing Sigma Factor 54 protein gene (RpoN) using the T7 promoter system and containing a 5ÚTR sequence. AmpR gene gives ampicillin resistance.

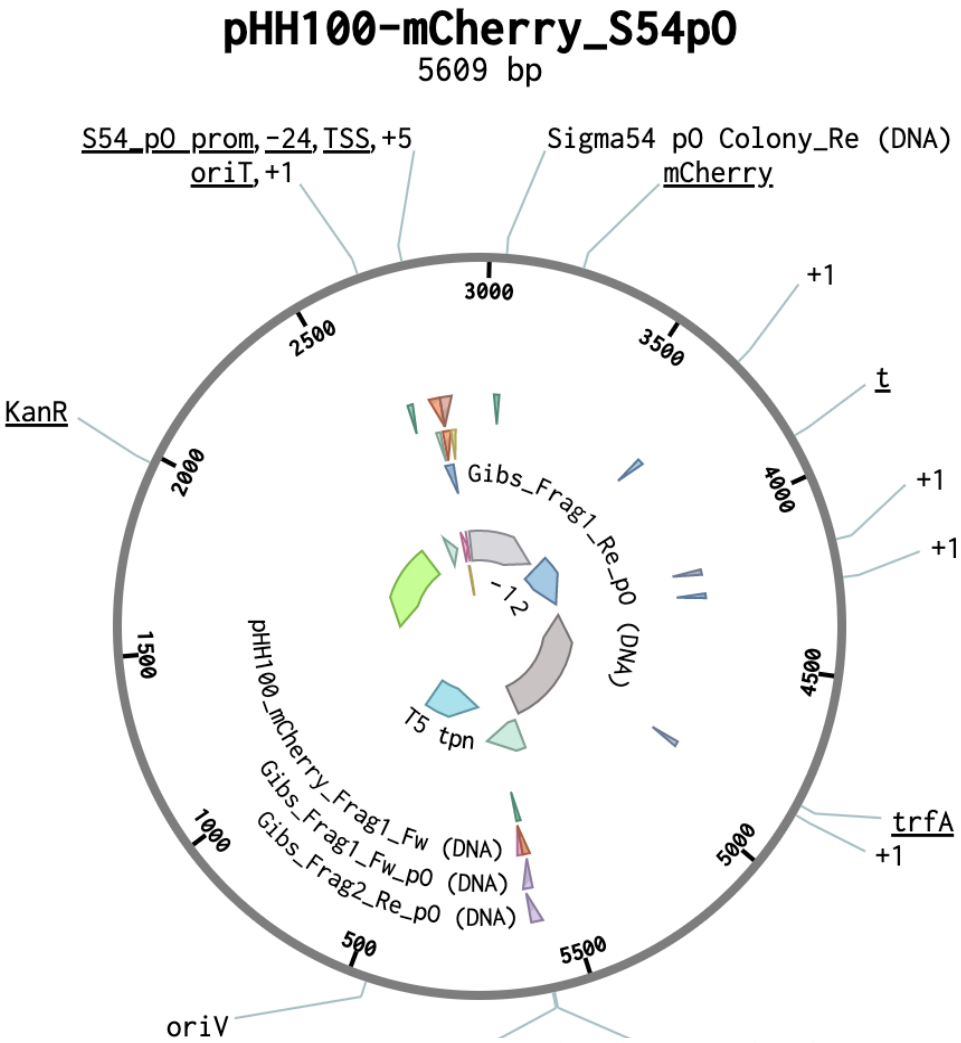

***Figure 8: E.coli pHH100-mCherry plasmid map**. Original E. coli Promoter Plasmid Sequence expressing mCherry. KanR gene gives kanamycin resistance. mCherry gene expresses the mCherry fluorescent protein.*

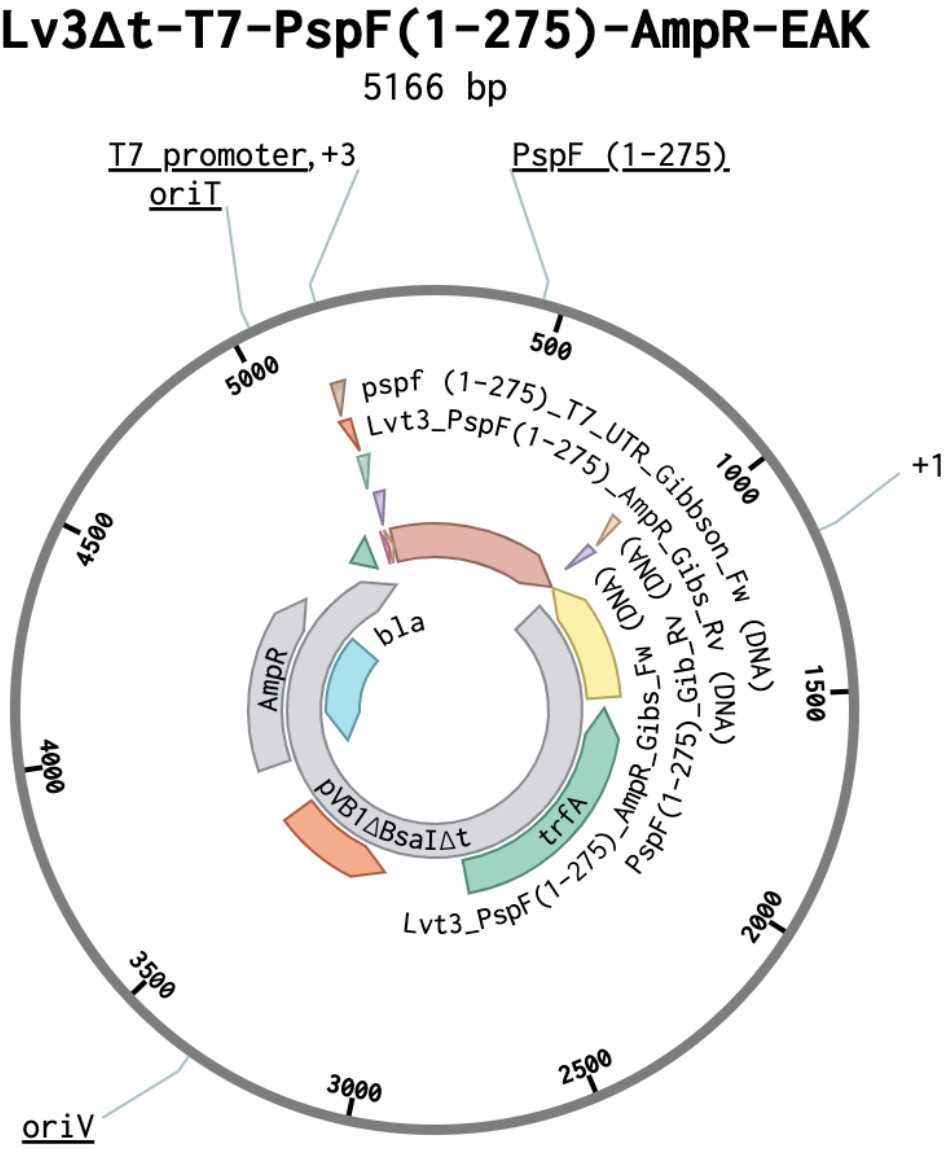

Figure 9: Lv3Δt – T7 – PspF(1-275) - AmpR plasmid map*. Plasmid expressing PspF (1-275) activator protein using the T7 promoter system and containing a 5ÚTR sequence. AmpR gene gives ampicillin resistance.*

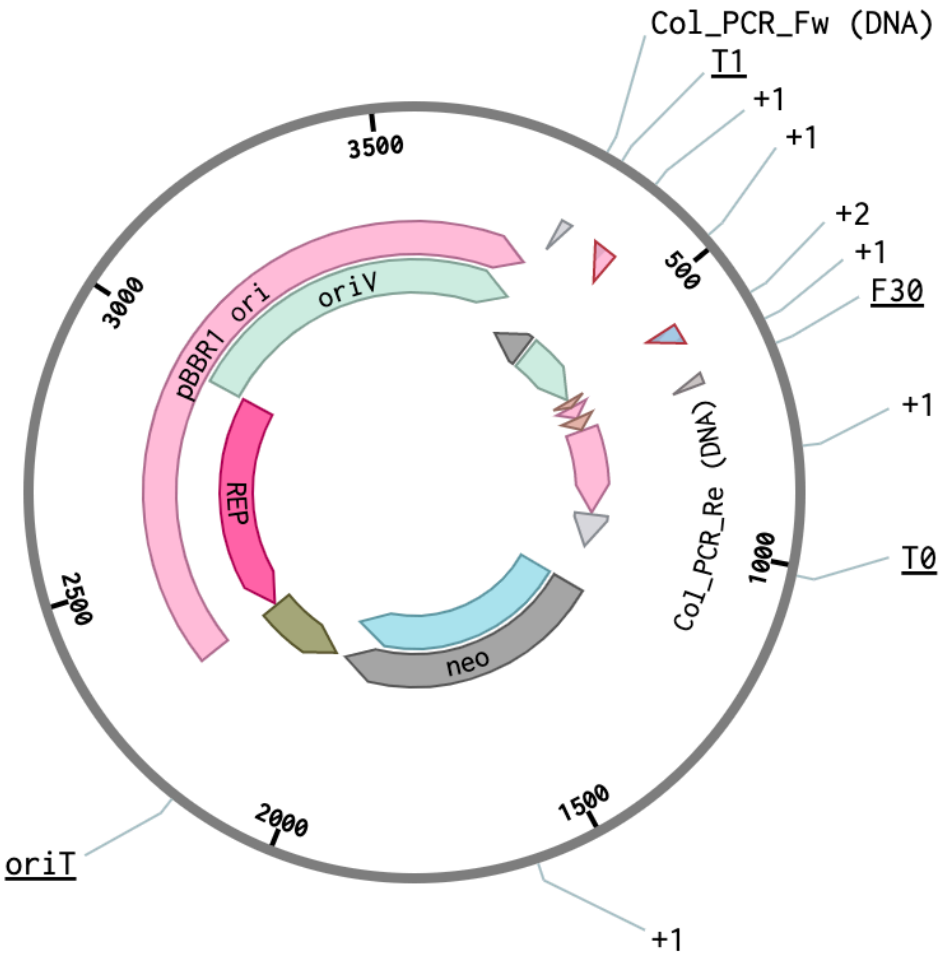

***Figure 10: pSEVA2311 plasmid map**. Plasmid expressing Mango aptamer fluorescent protein using the N200 synthetic DNA library as the promoter sequence.*

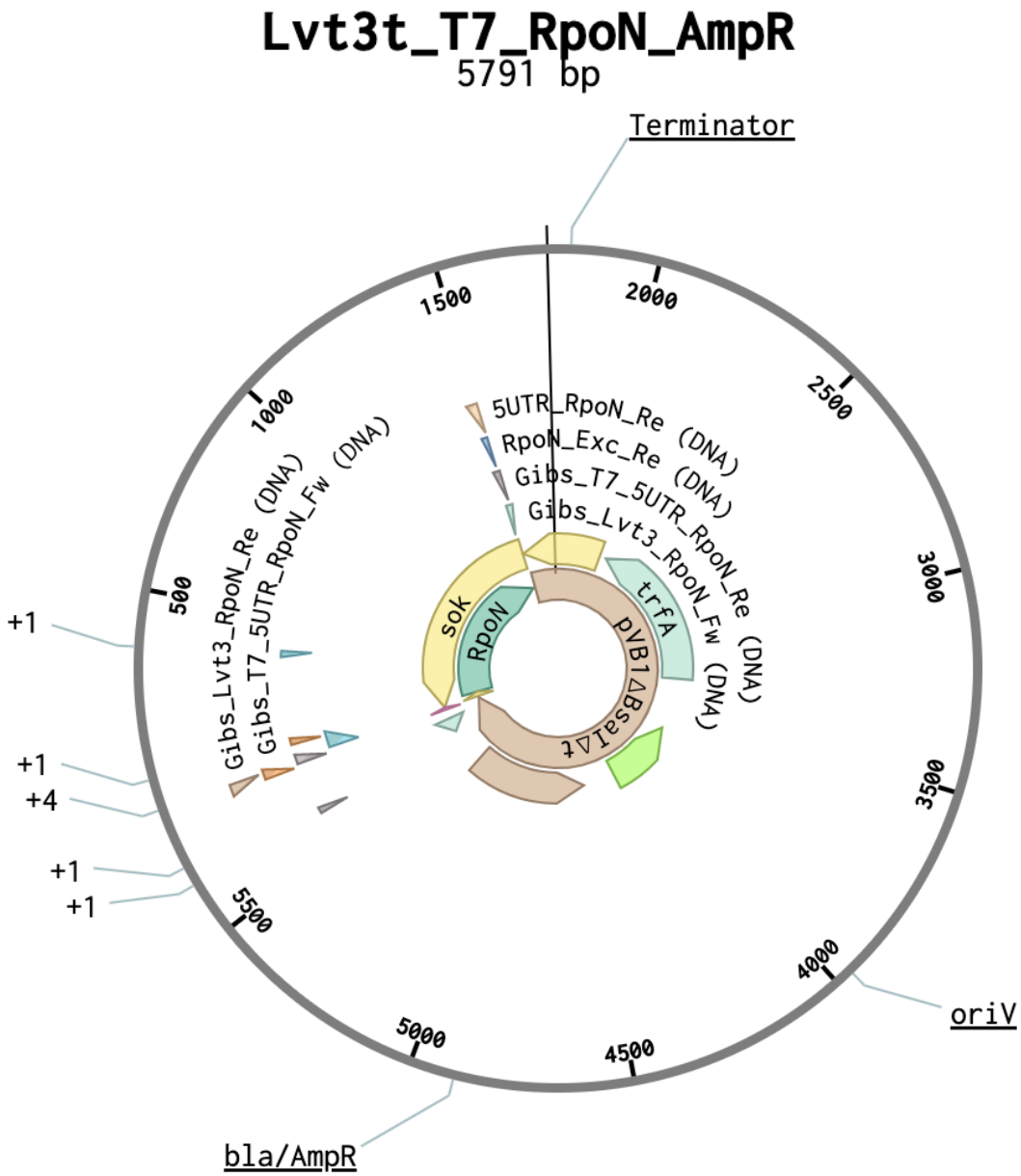

Figure 11: Lvt3t_T7_RpoN_ AmpR plasmid map*. Plasmid expressing Sigma Factor 54 protein gene (RpoN) using the T7 promoter system and containing a 5UTR sequence. AmpR gene gives ampicillin resistance.*

5.1.3. Preparation of E. coli competent cells

Appendix A, B, and C shows the mediums, antibiotics and buffers used in the preparation of *E. coli* cells.

A solution of *E. coli* culture was prepared in a small sterile Erlenmeyer flask with 10ml Psi medium and inoculated in a shaking incubator at 37°C overnight. The following day, 2ml of the overnight culture was transferred to a flask containing 200ml Psi medium. This was conducted under sterile conditions. The culture was inoculated in a shaking incubator at 37°C until the desired OD value was reached. To measure the OD at 600nm a Spectramax spectrophotometer was utilized. The expected OD at 600nm was 0.4 and it took 2.5 hours for the *E. coli* DH5α cells to reach this value. Once the culture reached an OD_{600} of 0.4, the flask along with the buffer solutions, TFB1 and TFB2, were transferred immediately to an ice for 15 min. The bacterial solution was then poured into a sterile centrifuge cup and centrifugated for 5min with 4000rpm at a temperature of 4°C. The supernatant was discarded, and the pellet was dissolved with 80ml TFB1. The sample was then left on ice for 5 minutes. The sample was then centrifugated for 5min at a speed of 4000rpm at a temperature of 4°C. The supernatant was poured off and the pellet was dissolved with 6ml of TFB2. The sample was dispensed into each Eppendorf tubes with a volume of 100µl. The tubes were left on ice until everything was aliquoted. The tubes were snap-freeze for 10 seconds with liquid nitrogen and stored in the -80 freezer.

5.1.4. Testing of E. coli competent cells

Appendix A, B, and C shows the mediums, antibiotics and buffers used in the testing of *E. coli* cells.

pLitmus28 plasmid containing an ampicillin resistance gene was miniprepped and stored at -20°C. The super competent cells were thawed on ice for 5-10 minutes. 1µl of the pLit28 plasmid was added to 100µl of the super competent cells and incubated on ice for 30 minutes. The cells were then heat-shocked in a 42°C water bath for approximately 35 seconds and later cooled on ice for 2 minutes. 1ml of the SOC medium was added and the tubes were then placed in a small Erlenmeyer flask. The flask was then placed in a shaking incubator for 1 hour at 37°C. A dilution series was made with LB medium. 900 µl of LB medium was added for each tube, 100µl from tube 1 was then transferred to the next tube and continued for all the tubes.

100µl of the sample was then spread onto plates and then incubated in a 37°C warming cabinet for overnight. The following day the bacteria colonies were counted and compared.

5.2. Cloning of Plasmid Constructs

5.2.1. PCR Amplification

Q5 PCR reactions were used to amplify sigma factor 54 (RpoN), transcriptional activator protein (PspF) from *E. coli* genomic DNA and to amplify the N200 random DNA library and the pO promoter sequence using primers designed beforehand. The Annealing temperatures were dependent on the primer information on the plasmid maps. The primer sequences used can be seen in Appendix D and the plasmid maps can be seen in sub-chapter 5.1.2. above. The reaction setup and the thermocycling conditions for the Q5 PCR reactions is illustrated in the table below.

Table 3: PCR Using Q5 High-Fidelity DNA Polymerase.

Component	Amount
5X Q5 Reaction Buffer	10 µl
10 mM dNTPs	1 µl
10 µM Forward Primer	2.5 µl
10 µM Reverse Primer	2.5 µl
Template DNA	1 µl
Q5 High-Fidelity DNA Polymerase	0.5 µl
Nuclease-Free Water	to 50 µl

The components in table 3 above were combined and setup in a PCR tube. The PCR tube was then run on a thermocycler.

Table 4: Thermocycler settings for Q5 PCR amplification

Step	Time	Temperature	Cycle
Initial Denaturation	98°C	30 econds	
Denaturation	98°C	10 seconds	
Annealing	50–72°C	20 seconds	Repeat step 2-4 30x
Extension	72°C	20 seconds	
Final Extension	72°C	2 minutes	
Hold	4°C	∞	

The table 4 above illustrates the Thermocycler settings for Q5 PCR amplification reaction. The Annealing temperature depends on the primer melting temperatures. The N200 DNA library amplification was done by 10-20 cycles, but the backbone and insert sequences was done with 30 cycles.

5.2.2. Purification of PCR Product

PCR purification was conducted on sigma factor 54 (RpoN), transcriptional activator protein (PspF), the N200 DNA promoter library and the pO promoter sequence after Q5 PCR amplification was conducted. The purpose of this is to remove any fragments between 100bp and 10kb, such as salts, buffers, enzymes, nucleotides, and primers. The protocol was adapted from the QIAquick® PCR Purification Kit, which is shown in Appendix E. 5 volumes of Buffer PB were added to 1 volume of the PCR reaction and mixed. A QIAquick column was placed in a provided 2ml collection tube. The sample was applied to the QIAquick column and centrifugated for 60 seconds to bind the DNA. The supernatant was discarded. 750 µl of Buffer PE was added to the QIAquick column and centrifugated for 60 seconds to wash the sample. The supernatant was discarded. The QIAquick column was then centrifugated again in the provided 2ml collection tube for 1min for the purpose of removing any residual wash buffer. The QIAquick column was then placed in a clean 1.5ml microcentrifuge tube. To elute the

DNA, 30 µl of Buffer EB was transferred to the center of the QIAquick membrane. This was centrifugated for 1min and the sample concertation was analyzed.

5.2.3. Gel electrophoresis

Agarose Gel electrophoresis was conducted to separate the DNA molecules by size and confirm their relative sizes and purity. The sample products produced from the PCR amplification and the PCR purification were run on the gel. Gel red was used for the agar gel and TAE buffer was used to fill gel electrophoresis chamber. The gel was made using 0.8% agarose cast into a casting tray and a comb placed to create wells. The gel solidified after 20-30 minutes and was placed in the gel chamber with a positive electrode on one side and a negative electrode on the other. 10µl of a DNA standard ladder was used on the gel. 1.7µl of Gel Loading Dye Purple (6X) was combined along with 3µl of the sample and 7µl of H_2O. 10µl of the prepared sample was applied to the gel. The gel was run at 90V for 45 minutes depending on the size of the product. After the gel was run completely, the gel was imaged by the geldoc system Image lab by Bio-Rad. The buffers and solutions used in the agarose gel electrophoresis and the DNA ladder are displayed in Appendix C and F.

5.2.4. Purification of Gel Fragment

After conducting a gel electrophoresis on the samples, the desired DNA product fragment was excised using a wide, clean, sharp scalpel. The purpose of this type of purification is to remove, isolate and later purify the desired DNA fragment based on size. The protocol was adapted from the E.Z.N.A.®Gel Extraction Kit, which is shown in Appendix E. The gel slice was weighted in a clean 1.5ml microcentrifuge tube. A density of 1g/ml was assumed, and volume of the gel was derived. 1 volume of XP2 Binding Buffer was added to the gel slice and incubated for 60°C for 7 minutes. The tube was vortexed in between the incubation time to to make sure the gel has completely melted in the 1.5ml microcentrifuge tube. A HiBind® DNA Mini Column was inserted in a 2 mL Collection Tube and 700 µL DNA/agarose solution from the Eppendorf tube was transferred. This was centrifuged for 1min at room temperature. The filtrate was discarded, and the collection tube was reused. This was repeated several times until all the sample was transferred to the column. 300µL of the XP2 Binding Buffer was later added and the sample centrifugated at maximum speed for 1 minute at room temperature. The filtrate was discarded, and the collection tube was reused. 700 µL of SPW Buffer was then added to the sample and then centrifugated at maximum speed for 1 minute at room temperature. The

filtrate was discarded, and the collection tube reused. The empty HiBind® DNA Mini Column was centrifugated for 2 minutes at maximum speed to dry the column matrix and then transferred to a clean 1.5ml microcentrifuge tube. 30µL of elution buffer was added to the column membrane and the sample was centrifugated for 1 minute after a rest period of 2 minutes. The resulted in a purified DNA gel fragment.

5.2.5. Nanodrop Analysis

Nanodrop ONE Thermo scientific spectrophotometer was used to measure the concentration of the different cloned plasmid constructs after cloning, PCR purification and Gel fragment purification, i.e., Sigma factor 54, Transcriptional activator Protein (PspF), N200 random DNA Library (Mango), pO Promoter Sequence and the In vitro RNA transcript.

5.2.6. DpnI Digestion

The PCR products that resulted from the PCR amplification cloning procedure were DpnI digested. The purpose of this method is to remove any template DNA that remains in the PCR product. This procedure cleaves the methylated site from the DNA template plasmids. The DpnI digestion mixture was made with the components and volumes displayed in table 5 below.

Table 5: Components and Volumes used for the DpnI digestion of the PCR product

Component	Amount
DNA Template (PCR Reaction Product)	18 µl
5X CutSmart Buffer	5 µl
H_2O	26 µl
DpnI	1 µl

1µl of the DpnI was added to a finished 49µl PCR product sample and mixed. The sample was incubated at 37°C for 1 hour and then purified using QIAquick® PCR Purification Kit.

5.2.7. Gibson Assembly

Gibson assembly was conducted to join the cloned linear insert to the respective plasmid vector backbone. This procedure enables fast and effective cloning of multiple DNA fragments into their respective vector backbones without requiring the use of restriction enzyme. This was done when cloning Sigma Factor 54, Transcriptional activator Protein (PspF), pO Promoter

Sequence and when incorporating the Mango aptamer to the pSEVA-2311 vector backbone. The table 6 below displays the sample components and their equivalent volume ratio for the reaction setup. The protocol was adapted from the NEBuilder HiFi DNA Assembly Reaction Protocol, which is shown in Appendix E.

Table 6: HiFi DNA Assembly components and their equivalent volume ratio.

DNA Molar Ratio	**Vector:insert = 1:3** **Backbone: 0.03 pmol** **Insert: 0.06 pmol**
Total Amount of Fragments	2.5 µl
NEBuilder HiFi DNA Assembly Master Mix	12.5 µl
Deionized H2O	10 µl
Total Volume	25 µl

The reaction mixture was incubated in a thermocycler at at 50°C for 60 minutes and later stored on ice for transformation preparation. Chemically competent *E. coli* $DH_5\alpha$ cells were thawed on ice. 2µl of the chilled assembled product mixture was added to the *E. coli* $DH_5\alpha$ cells and mixed gently. The cells were placed on ice for 30 minutes and later heat shocked at 42°C water bath for 30 seconds. The cells were then transferred to ice for 2 minutes and 950µl of room-temperature SOC media was added to the tube containing the cells. The cells were incubated at 37°C for 60 minutes with shaking. Depending on the plasmid being produced, selection plates containing LB agar with the relevant antibiotics, ampicillin or kanamycin, were warmed to 37°C. 100µl of the cells was transferred and spread onto the LB agar plates and incubated overnight at 37°C.

5.2.8. Golden Gate Cloning

Golden gate cloning is a cloning procedure that occurs through a one pot restriction and ligation. This procedure utilizes a type IIS restriction enzyme and a T4 DNA ligase. The N200 random DNA library was incorporated into the pSEVA – 2311 – Mango plasmid backbone using BsaI High Fidelity Restriction Enzyme and T4 DNA ligase enzyme. The table 7 below summarizes the components and their respective volumes that were used to setup the golden gate reaction.

Table 7: Components and amounts used for the Golden Gate cloning process of N200 random DNA library..

Component	Amount
T4 Ligation Buffer	2µL
Backbone	75ng
Insert	75ng/insert
T4 Ligase (500 U) (NEB)	1.25µL
BsaI HF v2 (15 U) (NEB)	0.75µL
H_2O	To 20µL

The reagent components were mixed in a PCR tube and run through a thermocycler. The conditions of the thermocycler are summarized in the table 8 below.

Table 8: Thermocycler settings used for the Golden gate cloning process of N200 random DNA Library.

Step	Temperature	Duration	Cycle
1	37°C	5 minutes	
2	16°C	5 minutes	Repeat step 1-2, 30x
3	65°C	10 minutes	
4	4°C	Hold	

To ensure an efficient transformation, the Golden gate assembled PCR product was used immediately after for transformation into *E. coli* $DH_5\alpha$ cells. 10µL of the golden gate assembly mix was used per transformation. To check for a successful transformation, clones were selected and confirmed using colony PCR.

5.2.9. Transformation

The different cloning plasmid vectors of Sigma Factor 54, Transcriptional activator protein, N200 random DNA library and pO promoter sequence were all transformed to the *E. coli* $DH_5\alpha$ to confirm the PCR products produced. For checking the expression, Sigma Factor 54 and Transcriptional activator protein were later transformed into *E. coli* BL21 cells and expressed using IPTG as an inducer. *E. coli* BL21 cells were also used to transform the pO promoter sequence for *In vivo* screening procedure. 100µL vials containing chemical competent cells were removed from the -80°C freezer and put on ice. The cells were defrosted for 5 minutes

on ice and 10µL of Gibson mix was added to the defrost cells. The sample was chilled on ice for 45 minutes after addition and heat-shocked at 42°C for 35 seconds. The sample was left to chill on ice for a further 5 minutes and then 1ml of LB medium was added to the sample. The sample was incubated in shaker at 37°C for 60 minutes. The sample was plated out on agar plates containing either Km50 (50µg/ml Kanamycin) or Amp50 (Ampicillin 50µg/ml), depending on the plasmid being produced. The sample was then incubated overnight at 37°C. The next day, a cell count was conducted to estimate the library size being produced.

5.2.10. Colony PCR

Colony PCR was used to screen colonies of *E. coli* $DH_5\alpha$ cells for the presence of the insert DNA. In a Colony PCR tube, 10µL of the mixture and 1 Colony from the pipette tip were mixed and amplified using Taq PCR. Multiple colonies were picked using a sterile pipette tip and used as a template DNA in the reaction master mix for Taq PCR. The PCR product was then analysed using agarose gel electrophoresis. Appendix D displays the primer sequences.

Table 9: Taq Colony PCR master mix components and respective volumes used for the determination of the length of the PCR product..

Components	Volume
10X TAQ buffer	9µL
10mM DNTPs	1.89µL
Reverse Primer	1.89µL
Forward Primer	1.89µL
Taq DNA polymerase Enzyme	0.45µL
Template DNA	Solid Colony
Water (dH_2O)	75.15µL

The table 9 above illustrates the components and their respective volumes used in the master mix for the colony PCR procedure.

Table 10: Thermocycler settings used for the Taq Colony PCR process.

Step	Temperature	Duration	Cycle
Destroying of cell walls	95°C	10 minutes	

Denaturation	95°C	10 seconds	
Annealing	45-68°C (Primer dependent)	30 seconds	Repeat step 2-4, 30x
Extension	68°C	1minute/kb	
Final Extension	68°C	2 minutes	
Hold	4°C	Hold	

The reagent components were mixed in a PCR tube and run through a thermocycler. The conditions of the thermocycler are summarized in the table 10 above. The annealing temperature is primer dependent and is therefore in the range 45-68°C.

5.2.11. Plasmid Preparation

Plasmid miniprep was conducted on sigma factor 54 (RpoN) for BL21 transformation and expression, transcriptional activator protein (PspF) for BL21 transformation, expression and *In vivo* screening, the N200 DNA promoter library for *In vitro* bulk transcription and the pO promoter sequence for *in vivo* screening. The purpose of this is to remove any small plasmid DNA from bacteria (e.g., *E. coli* $DH_5\alpha$ cells) while aiming to minimize contamination levels for the genomic DNA. The protocol was adapted from the QIAprep® Spin Miniprep Kit, which is shown in Appendix E. 5ml of the bacterial overnight culture was harvested through centrifugation at 8000rpm for 3min at room temperature. The bacterial cell pellet was resuspended in 250µl of Buffer P1 and transferred to a microcentrifuge tube. 250 µl of Buffer P2 was added and mixed until the solution became clear. 350 µl of Buffer N3 was later added and the sample was mixed thoroughly. The sample was then centrifugated for 10 minutes at 13,000 rpm using a table-top microcentrifuge. 800µl of the supernatant produced was applied to the QIAprep 2.0 spin column and the sample centrifuged for 60 seconds. The QIAprep 2.0 spin column was washed by adding 0.5ml of Buffer PB and centrifuged for 60 seconds. The QIAprep 2.0 spin column was washed again by adding 0.75ml of Buffer PE and centrifugated for 60 seconds. A collection tube was used for the QIAprep 2.0 spin column to remove residual wash buffer. This was done through centrifugation of the column for 1 minute. The flow-through produced were discarded in each step. A clean 1.5ml microcentrifuge tube was used to elute the DNA by adding 50 µl of Buffer EB (10 mM TrisCl, pH 8.5) to the QIAprep 2.0 spin column. After 1 minute, the sample was centrifugated for 1 min and the eluted DNA

collected and the concentration measured using Nanodrop ONE Thermo scientific spectrophotometer.

5.2.12. Plasmid Sequencing

Plasmid sequencing was conducted through Eurofins Genomics LightRun Tube. The plasmids that were isolated through plasmid miniprep were sent for sequence confirmation. The purpose is to determine whether the nucleotides are in the correct order (i.e., confirming the presence of the insert DNA) as expected in the genomic DNA after the cloning procedure. The samples were prepared by mixing 5µl of the DNA plasmid template with 5µl of the corresponding forward and reverse primers. The plasmid template DNA and the primers were prepared in an Eppendorf tube. The tube was labelled and sent for sequencing. The primers are listed in Appendix D.

5.2.13. Glycerol Stock

Glycerol stocks were made for plasmids that confirmed the presence of the nucleotides for the DNA insert. This was confirmed through plasmid sequencing of plasmids from the *E. coli* $DH_5\alpha$ transformant colonies that were produced during the Taq colony PCR. The stocks ensured the longevity and preservation of the cells for reusage in the future. The colonies showing presence of the plasmids were picked up from the plate and grown again in 4ml LB medium in 13ml reagent tubes. The tubes contained the appropriate volumes of the relevant antibiotic (kanamycin or ampicillin). The tubes were grown overnight at 37°C and 225rpm. The following day, 20ml of 50% glycerol was diluted with LB and this was diluted up to a volume of 50ml. 1ml of the diluted glycerol was then mixed with 1ml of the liquid culture and stored at -80°C.

5.3. **Expression of Plasmid Constructs**

5.3.1. Cell Growth and Protein Expression

The expression construct, pJF5401, contains the expression gene for Sigma Factor 54 (RpoN) and includes a heat inducible promoter. This expression construct was achieved from our collaborators in Sweden. The host *E. coli* strain is unknown. To express this plasmid construct, the *E. coli* strain was inoculated in 5ml of LB medium containing ampicillin with a concentration of 50µg/ml in 13ml culture tubes. The cells were incubated overnight at 37°C

with shaking at 225rpm in an incubator. The following day, 250µl of the overnight culture was inoculated in a 250ml flask with 25ml of LB medium containing ampicillin with a concentration of 50µg/ml. This was incubated at 30°C with shaking at 160rpm for 2.5 hours until the OD_{600} was in the interval 0.7-0.8. Once the expected OD_{600} was achieved, the induction was initiated by increasing the temperature to 42°C and the incubation was extended for another 5 hours. The negative control sample was grown without heat induced induction and thus remained at 30°C for a further 5 hours. The cells were harvested by centrifugation and washed once with M9 salts, and the pellet stored at -80°C for future use.

The expression construct, Lv3Δt – T7 – PspF (1-275) – AmpR**,** contains the expression gene for the Transcriptional Activator Protein (PspF) and includes a T7 promoter system. This expression construct was achieved through extensive research. The *E. coli* BL21 expression strain is unknown. To express this plasmid construct, the *E. coli* BL21 strain was inoculated in 5ml of LB medium containing ampicillin with a concentration of 50µg/ml in 13ml culture tubes. The cells were incubated overnight at 37°C with shaking at 225rpm in an incubator. The following day, 250µl of the overnight culture was inoculated in a 250ml flask with 25ml of LB medium containing ampicillin with a concentration of 50µg/ml. This was incubated at 30°C with shaking at 160rpm for 2.5 hours until the OD_{600} was in the interval 0.7-0.8. Once the expected OD_{600} was achieved, the induction was initiated by IPTG, and the incubation was extended for another 5 hours. The cells were harvested by centrifugation and the pellet stored at -80°C for future use.

5.3.2. Protein Isolation

The harvested cells for Sigma factor 54 and the Transcriptional activator protein (PspF) were run on SDS-PAGE gel to confirm the expression. SDS-PAGE is a protein separation method based on the migration of charged molecules to the opposite electrode when subjected to an electrical field. This helps in separating proteins of different molecular masses. The samples were retrieved from -80°C. The retrieved pellets were resuspended in 10ml lysis buffer. 260µg/ml of Lysozyme was added into the resuspended cells and separated into fractions. 3X Reducing Blue Protein Loading Dye was prepared by adding 1/10 volume of 30X Reducing Agent (DTT was used as a reducing agent) to 1 volume of 3X Blue Loading Dye. The sample was prepared by adding 1/2 volume of 3X Reducing Blue Protein Loading Dye. Therefore, 25µl of 3X Reducing Blue Protein Loading Dye and 25µl of the cell sample was added to a

tube to make 50µl sample. The sample was heated with a temperature of around 95–100°C for 5 minutes and then centrifugated. The sanple was then loaded onto the SDS gel in the SDS gel chamber. 1X SDS running buffer (BioRad, Tris/Glycin) was prepared. The SDS chamber was filled with the 1X SDS running buffer and 5µl of the sample and ladder were added to the SDS-PAGE gel. The gel was run for 120V, 400mA for 60 minutes. The SDS-PAGE was continuously running until all the sample was separated. The sample gel was stained with Bio-Rad Coomassie stain solution G250 once and then the gel was destained in water. The destaining procedure was conducted via a gel analyser machine (eStain® L1 Protein Staining System by GenScript) and then scanned through geldoc system Image lab by Bio-Rad.

5.4. *In Vivo* Transcription

In vivo transcription was conducted to determine the time the transcription process occurs in the cell using its own cellular components. pO promoter sequence was utilized and no inducer was used for this process. In accordance with the exact time the transcription process occurs using the *E. coli* pO promoter sequence, the experiment was expected to be reconducted using the artificial N200 random nucleotide DNA library. DmpR-His plasmid was used from *Pseudomonas Putida* and the bacteria was streaked out on LB agar plate and grown. From the growth, a colony was cultured and eventually the plasmid was isolated through miniprep. The isolated DmpR-His plasmid was then transformed into the *E. coli* BL21 strain. *E. coli* pO promoter sequence with mCherry was transformed into the *E. coli* BL21 strain and the cells were streaked out and grown on a plate. The positive colonies were confirmed using colony PCR and inoculated in 5ml LB medium with kanamycin present. The following day, the pre-culture was used to inoculate the *E. coli* strain into the main culture with 10ml of LB medium with kanamycin present. The culture was grown to OD_{600} 0.15 at 37°C and the incubation period took around 6-25 hours. The DmpR-His plasmid included an arabinose-inducible promoter P(BAD) promoter system and was induced using 2wt% arabinose. After induction, the sample was grown for a further 4 hours until an OD_{600} of 0.6-0.7 was reached. This is considered the mid to late log-phase. The OD_{600} was measured every hour to check the growth. Once the required OD_{600} value was reached, fluorescence was measured through TECAN 96-well microplate reader.

5.5. *In Vitro* Bulk Transcription

Four solution mixtures were made, called A, B, C and D for the *In vitro* bulk transcription reaction. Their components and their respective volumes are shown in tables 11-14 below.

Table 11: Mixture A components and their respective volumes. This mixture contains the Transcriptional Activator Protein DmpR-His.

Components of solution mixture A	Volume of Induced reaction	Volume of Control reaction
H_2O	2.77µl	4µl
Buffer	1µl	1µl
DmpR-His	1.23µl	Not Utilized

Table 12: Mixture B components and their respective volumes. This mixture contains the Core RNAP and the E. coli Sigma Factor 54.

Components of solution mixture B	Volume of Induced reaction	Volume of Control reaction
H_2O	3.43µl	10.63µl
Buffer	3µl	3µl
BSA (2mg/ml)	1.37µl	1.37µl
Core RNAP	4.7µl	Not utilized
***E. coli* Sigma Factor 54**	2.5µl	Not utilized

Table 13: Mixture C components and their respective volumes. This mixture contains the N200 Random Nucleotide Library.

Components of solution mixture C	Volume
H_2O	10.3µl
N200 Random Nucleotide Library	7.7µl
Buffer	12µl
ATP	30µl

Table 14: Mixture D components and their respective volumes. This mixture contains the RNAse inhibitor and the Nucleotide Mix. TO1 - Biotin was prepared by mixing 2µl of TO1 – Biotin with 33.45µl of H_2O.

Components of solution mixture D	Volume

H_2O	6.47µl
Buffer	4µl
RNAse inhibitor	0.27µl
Nucleotide Mix	5.76µl
TO1-Biotin	2.5µl
15µl KCl	1µl

Two *in vitro* bulk reactions were initiated with one without the enzyme included (Core RNAP and *P. putida* Sigma 54). Once all 4 solution mixtures were prepared, all mixtures were incubated at different time intervals at a constant temperature of 30°C. Mixture B was incubated for 20min, Mixture A and Mixture C were added and incubated for a further 30mins, Mixture C was added and then incubated for 60min. The reaction was inhibited with heparin and the sample was incubated for 10min. DNase was added twice, and the sample was incubated for 15min each time.

5.5.1. In Vitro NEB PCR & DNA cleanup Kit Protocol

The RNA transcript produced through *In Vitro* bulk transcription was cleaned using the Monarch® PCR & DNA Cleanup Kit (5 µg). The purpose of this method is to instantly purify concentrated DNA produced from enzymatic reactions. The protocol was adapted from the Monarch® PCR & DNA Cleanup Kit (5 µg), which is shown in Appendix E. All of the centrifugation steps were carried out at 13000rpm using a microcentrifuge. The RNA transcript sample were diluted using DNA cleanup binding buffer and mixed. A column was used with the collection tube and the sample was loaded onto the column. The sample was centrifuged for 1 minute and the flow-through was discarded. 200µl of DNA wash buffer was added and the sample was centrifuged for 1 minute. This step was repeated. The column was then transferred to a clean 1.5ml microfuge tube and 6µl of DNA elution buffer was added to the matrix. The tube was centrifuged for 1 minute and the RNA was eluted.

6. Results

6.1. Plasmid construct Creation

The results will be presented in this section for the Cloning of Plasmid Constructs, Expression of Plasmid Constructs, and the *in vivo* and *in vitro* bulk transcription reactions. The cloning results will be for RpoN, PspF, *E. coli* pO promoter sequence and the N200 random nucleotide library.

6.1.1. Sigma Factor 54 Plasmid with T7 Promoter

Q5 PCR reactions were used to amplify sigma factor 54 (RpoN) from *E. coli* genomic DNA. The expected band fragment was 1431bp. This was followed by the incorporation of 5´UTR sequence. This had an expected band fragment of 1479bp. Following this, T7 promoter sequence was incorporated into the template sequence. The expected band fragment for this was 1508bp. The template was then prepared for Gibbson assembly, and the expected band fragment was 1518bp. The backbone was prepared next for the Gibbson assembly, and the expected band fragment was 4313bp. The annealing temperatures were dependent on the primer information on the plasmid maps. The primer sequences used can be seen in Appendix D and the plasmid maps can be seen in sub-chapter 5.1.2. above. The PCR product was verified using gel electrophoresis.

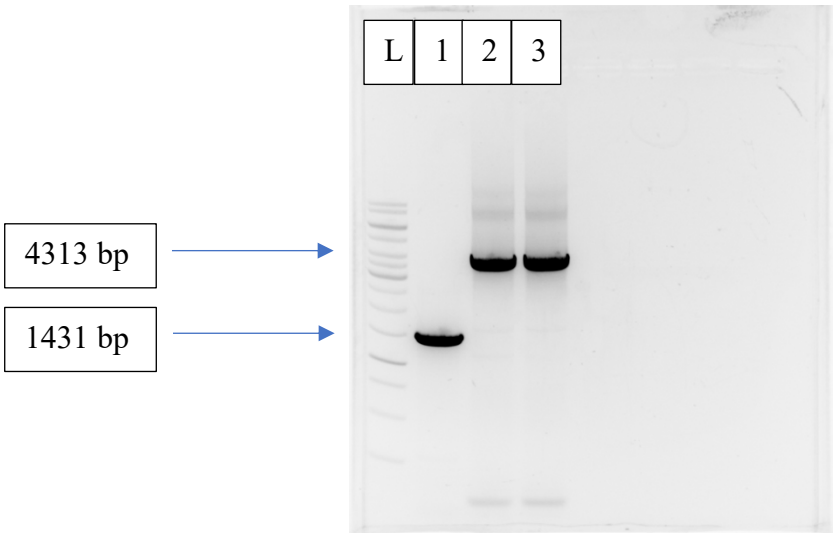

Figure 12: Agarose Gel Electrophoresis of Excision of RpoN and the Backbone for Gibbson Assembly from the E.coli Genome. *Well 1 contains the Excised RpoN sequence and Well 2&3 contains the Backbone sequence. The DNA Ladder can be seen in Appendix F.*

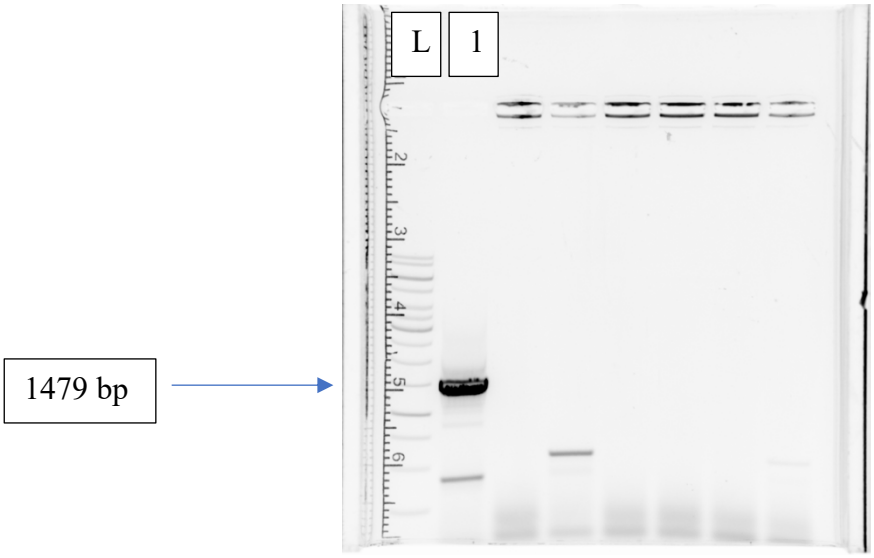

Figure 13: Agarose Gel electrophoresis of incorporation of 5ÚTR into E.coli RpoN*. The DNA Ladder can be seen in Appendix F.*

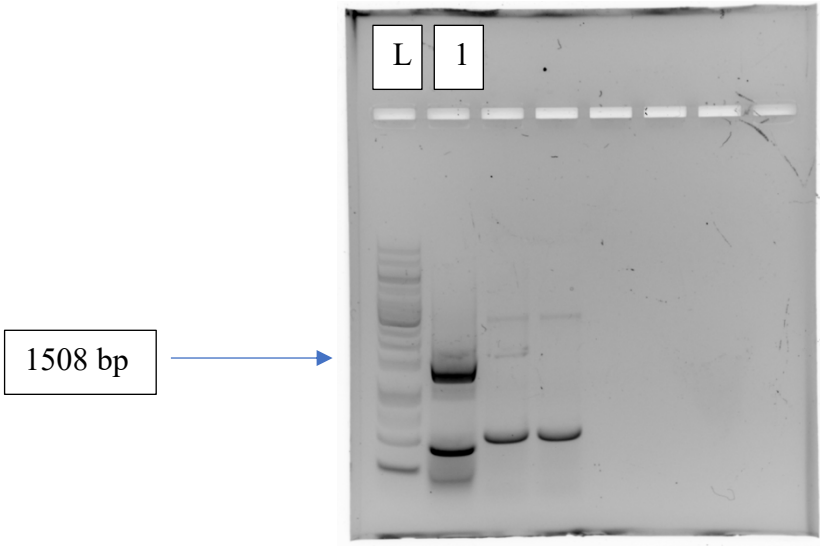

Figure 14: Agarose Gel electrophoresis of incorporation of T7 Promoter Sequence into E.coli RpoN – 5´UTR. *The DNA Ladder can be seen in Appendix F.*

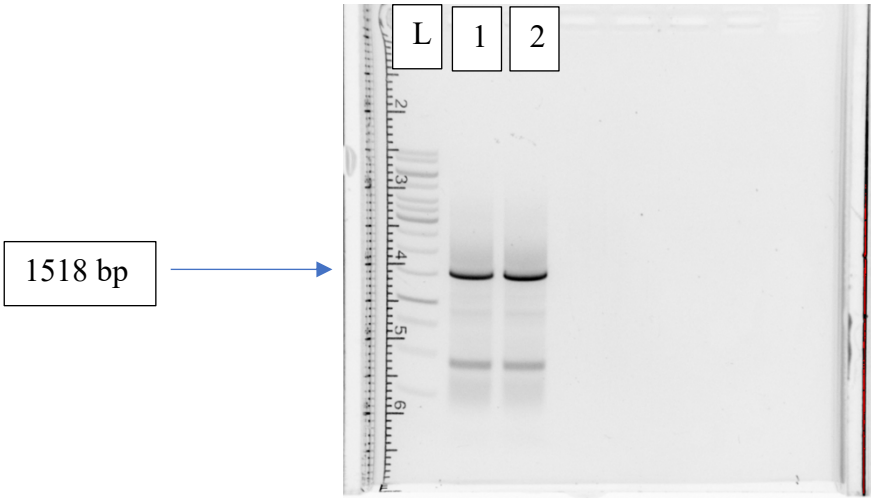

Figure 15: Agarose Gel electrophoresis of T7_5´UTR_RpoN template preparation for Gibson assembly. *The DNA Ladder can be seen in Appendix F*

Once the insert and the backbone for the sigma factor 54 plasmid with T7 promoter was cloned and transformed into $DH_5\alpha$ cells, it was grown, miniprepped and transformed in BL21 cells. Colony PCR was used to screen colonies of *E. coli* BL21 cells for the presence of the insert DNA. In a Colony PCR tube, 10µL of the mixture and 1 Colony from the pipette tip were mixed and amplified using Taq PCR. Multiple colonies were picked using a sterile pipette tip and used as a template DNA in the reaction master mix for Taq PCR. The PCR product was then analysed using agarose gel electrophoresis.

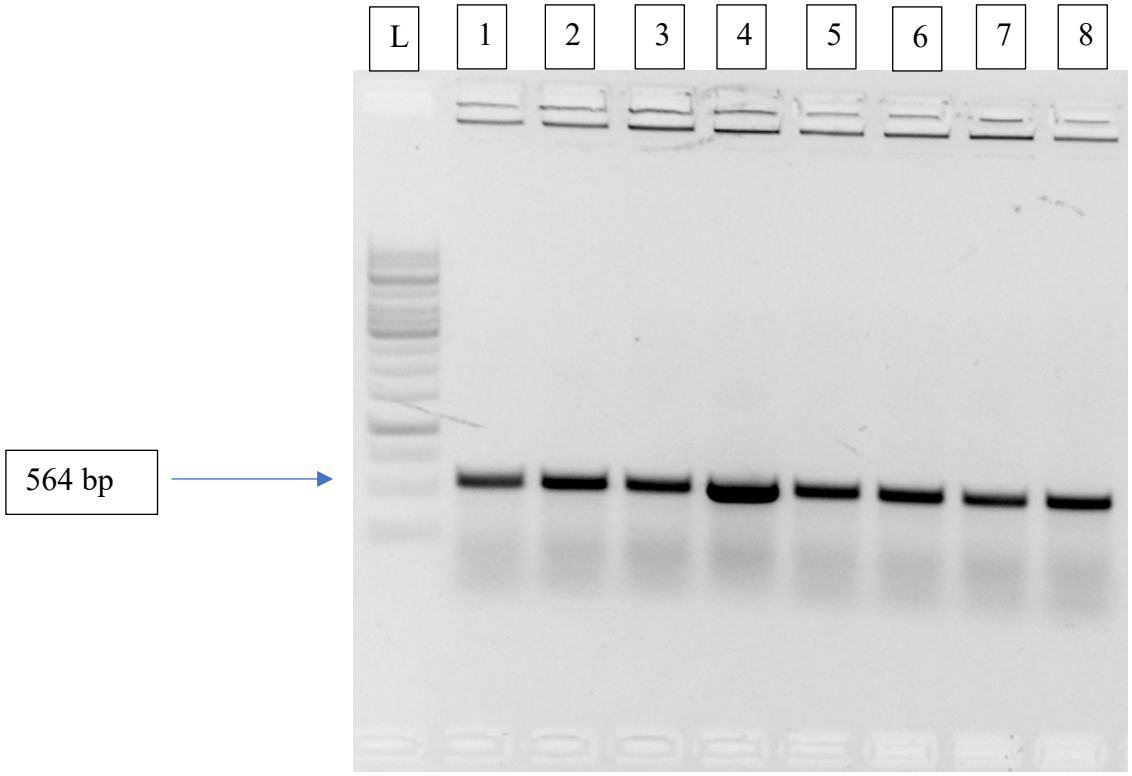

Figure 16: BL21 Colony PCR. *Multiple Colonies were tested for RpoN from BL21 variant. The DNA Ladder can be seen in Appendix F.*

6.1.2. PspF - Transcriptional activator Protein plasmid with T7 Promoter System

Q5 PCR reactions were used to amplify *PspF* – Transcriptional Activator Protein from *E. coli* genomic DNA. The expected band fragment was 828bp. This was followed by the incorporation of 5´UTR sequence. This had an expected band fragment of 854bp. Following this, T7 promoter sequence was incorporated into the template sequence. The expected band fragment for this was 873bp. The template was then prepared for Gibbson assembly, and the expected band fragment was 884bp. The backbone was prepared next for the Gibbson assembly, and the expected band fragment was 4319bp. The annealing temperatures were

dependent on the primer information on the plasmid maps. The primer sequences used can be seen in Appendix D and the plasmid maps can be seen in sub-chapter 5.1.2. above. The PCR product was verified using gel electrophoresis.

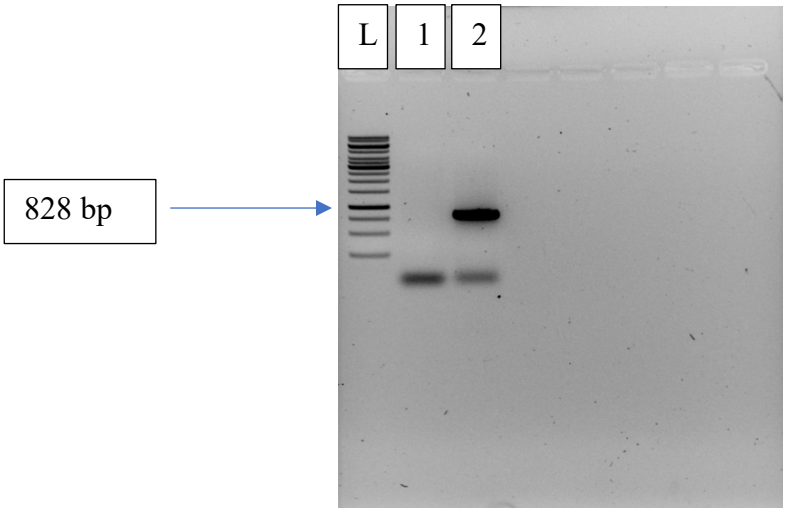

Figure 17: Agarose Gel Electrophoresis of Excision of PspF from the E.coli Genome. *Well 1 contains the Excised PspF sequence and Well 2 contains the PspF $_{(1-275)}$ sequence. The DNA Ladder can be seen in Appendix F.*

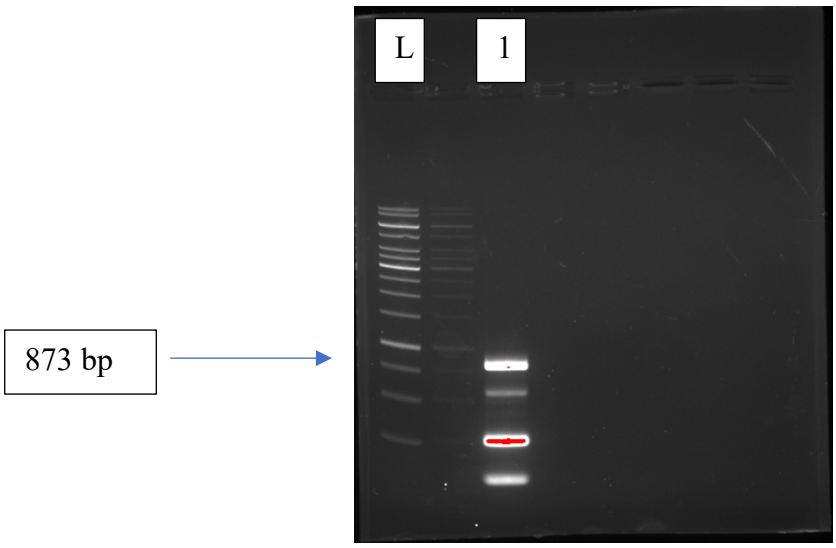

Figure 18: Agarose Gel electrophoresis of incorporation of T7 Promoter Sequence into PspF– 5´UTR. *The DNA Ladder can be seen in Appendix F.*

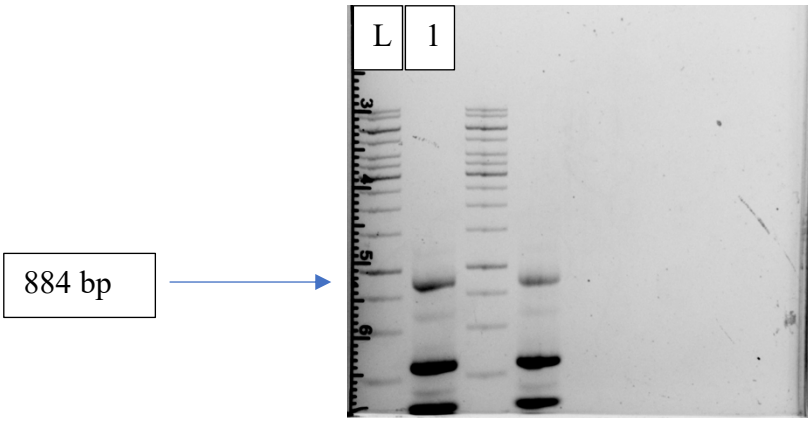

Figure 19: Agarose Gel electrophoresis of T7_5´UTR_PspF template preparation for Gibson assembly. *The DNA Ladder can be seen in Appendix F*

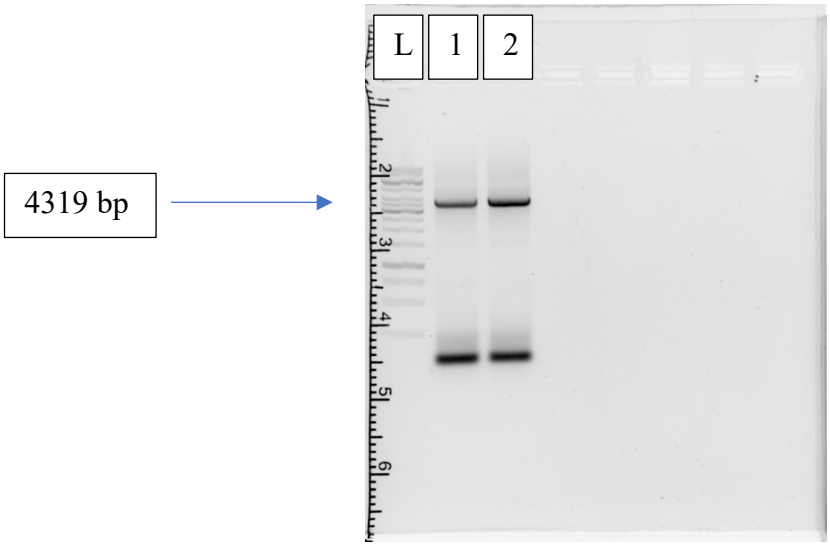

Figure 20: Agarose Gel Electrophoresis for the amplification of the Backbone (LvΔt – AmpR plasmid) for Gibbson Assembly. *The DNA Ladder can be seen in Appendix F.*

Gibson assembly was conducted to join the cloned linear T7 - 5´UTR - PspF linear template insert to the respective LvΔt – AmpR plasmid vector backbone. This resulted in the LvΔt – T7 - PspF $_{(1\text{-}275)}$ - AmpR plasmid. As a result of an unknown technical error, no visible expression of LvΔt – T7 - PspF $_{(1\text{-}275)}$ - AmpR was shown during SDS-PAGE. DmpR was therefore used as an alternative to PspF. This was received from our collaborators as a purified protein and as a plasmid and was used in the *in vitro* bulk transcription experiments. The plasmid was used in the *In vivo* transcription experiments and was induced by arabinose after transformation in the *E. coli* BL21 strains.

6.1.3. E. coli pO Promoter Sequence Plasmid with mCherry

pHH100 – mCherry – S54pO plasmid was constructed in several steps. *E. coli* Sigma 54 pO promoter insert was amplified from pHH100 – mCherry plasmid. The pHH100 – mCherry plasmid backbone was amplified into 2 fragments. One fragment gives an expected band fragment of 2793bp, and the other fragment gives an expected band fragment of 2766bp. The pO promoter sequence was then incorporated into the 2 backbone fragments with different band lengths. Immediately after, both fragments were prepared for Gibson assembly. Q5 PCR reactions were used to amplify the construction of pHH100 – mCherry – S54pO plasmid. The annealing temperatures were dependent on the primer information on the plasmid maps. The primer sequences used can be seen in Appendix D and the plasmid maps can be seen in sub-chapter 5.1.2. above. The PCR product was verified using gel electrophoresis.

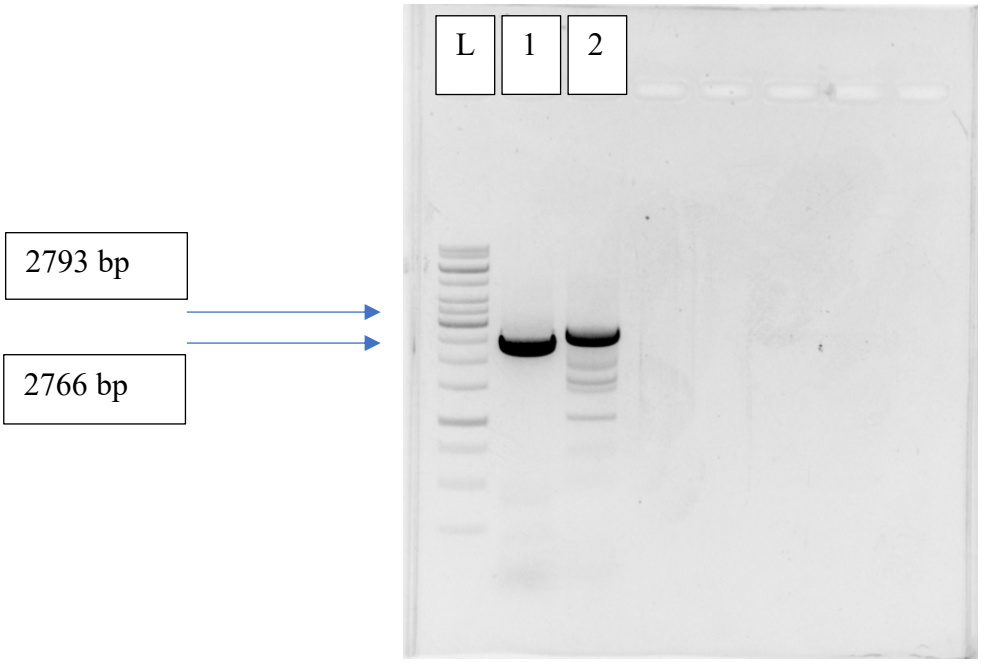

Figure 36: Agarose Gel Electrophoresis of Excision of Fragment 1 and Fragment 2 from the pHH100 – mCherry Plasmid. *Well 1 contains the Fragment 1 and Well 2 contains the Fragment 2 sequence. Fragment 1 gives an expected band fragment of 2766bp, and Fragment 2 gives an expected band fragment* ***of*** *2793bp. The DNA ladder can be seen in Appendix F.*

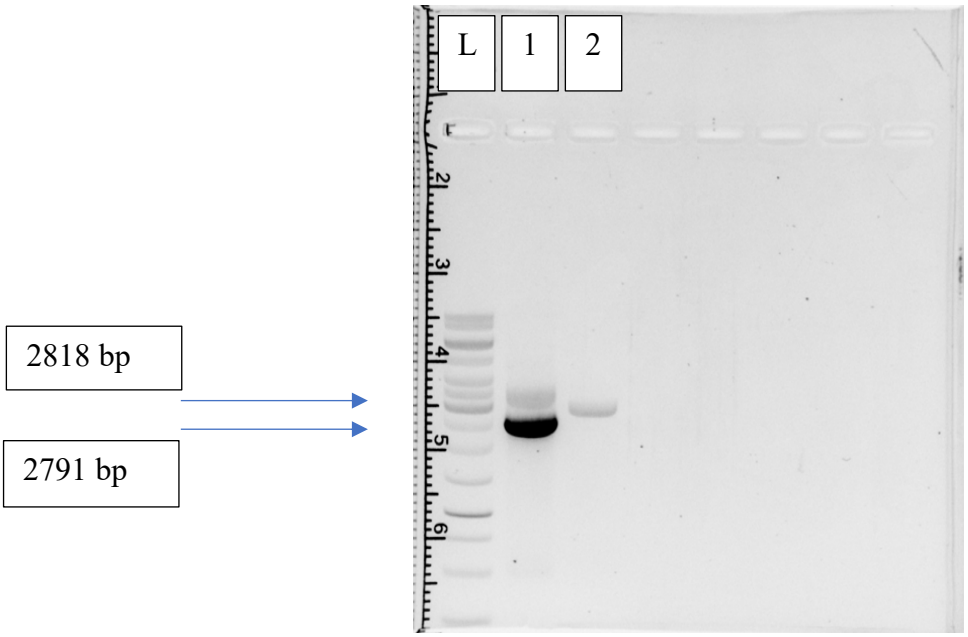

Figure 37: Agarose Gel Electrophoresis of incorporation of S54 pO Promoter to Fragment 1 and Fragment 2. *Well 1 contains the Fragment 1 and Well 2 contains the Fragment 2 sequence. Fragment 1 gives an expected band fragment of 2791bp, and Fragment 2 gives an expected band fragment of 2818bp. This gel was achieved after gel purification was conducted. The DNA Ladder can be seen in Appendix F.*

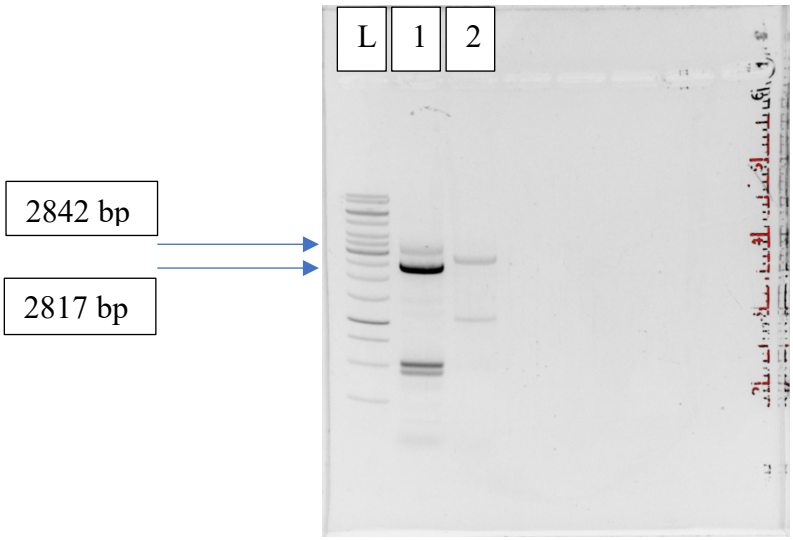

Figure 38: Agarose Gel Electrophoresis for preparation of Fragment 1 and Fragment 2 for Gibbson assembly. *Well 1 contains the Fragment 1 and Well 2 contains the Fragment 2 sequence. Fragment 1 gives an expected band fragment of 2817bp, and Fragment 2 gives an expected band fragment of 2842bp. The DNA Ladder can be seen in Appendix F.*

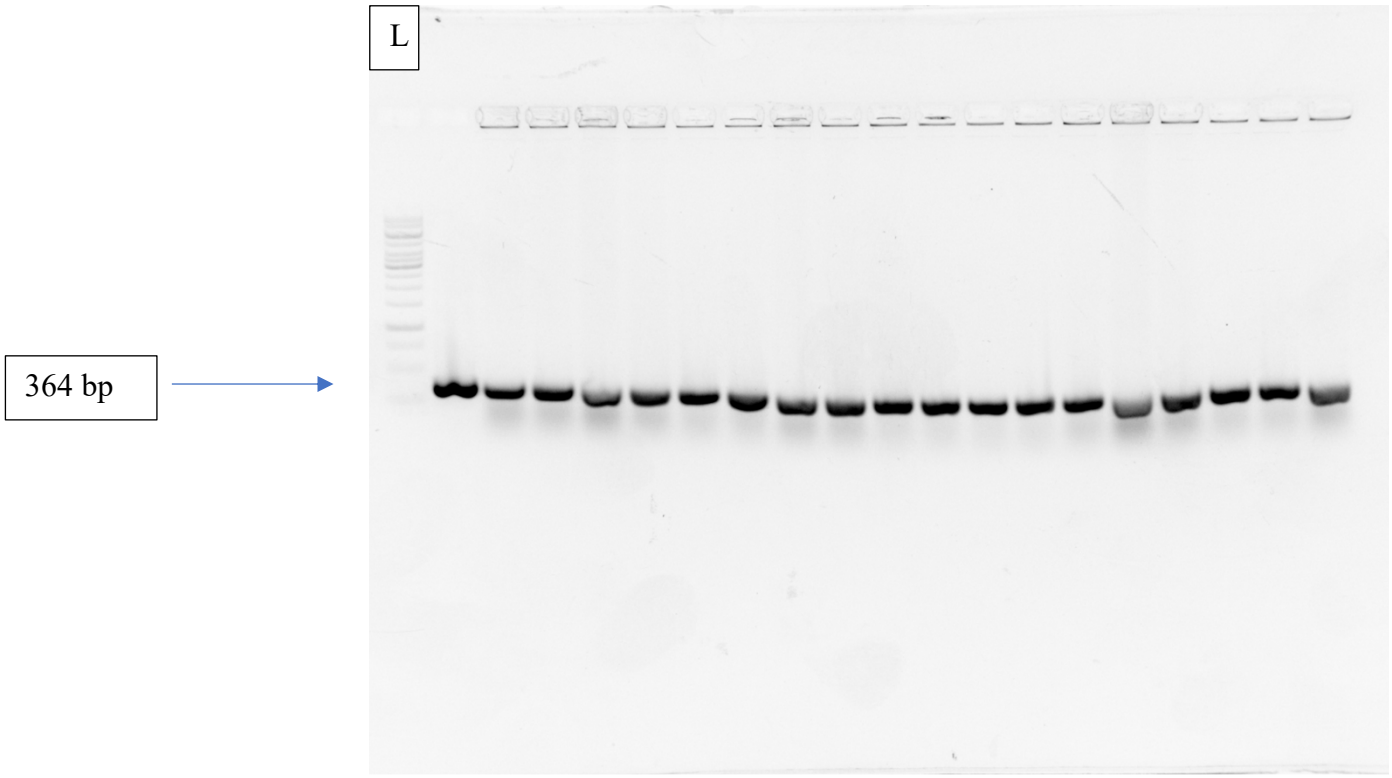

Figure 39: DH5α Colony PCR. *Multiple Colonies were tested for pHH100 – mCherry – S54pO plasmid with kanamycin resistance from DH5α variant. The PCR product has the band fragment size of 364bp. The DNA Ladder can be seen in Appendix F.*

Once the fragment 1 and fragment 2 were cloned and assembled through Gibbson assembly, the pHH100 – mCherry – S54pO plasmid was formed. The cloned construct was transformed into $DH_5\alpha$ cells and grown in agar plates. Colony PCR was used to screen colonies of *E. coli* $DH_5\alpha$ cells for the presence of the pHH100 – mCherry – S54pO. In a Colony PCR tube, 10µL of the mixture and 1 Colony from the pipette tip were mixed and amplified using Taq PCR. Multiple colonies were picked using a sterile pipette tip and used as a template DNA in the reaction master mix for Taq PCR. The PCR product was then analysed using agarose gel electrophoresis.

6.1.4. N200 Random Nucleotide DNA Library with Mango Aptamer

N200 Random nucleotide DNA library with Mango aptamer was constructed by first PCR amplifying the Mango - 3´UTR insert and the pSEVA backbone. The expected band fragment was 395bp and 3013bp respectively. The Mango-pSEVA template was confirmed through colony PCR with expected band fragment of 200bp. This Mango – pSEVA template was then prepared as a backbone for the incorporation of the N200 random nucleotide library through Golden gate cloning. The expected band fragment was 3397bp. The N200 Random nucleotide

library was amplified. The expected band fragment was 200bp. After Golden Gate cloning of the N200 Library insert with the Mango – pSEVA backbone, colony PCR was conducted to confirm the cloned plasmid. The expected band fragment for the insert was 200bp. Q5 PCR reactions were used to amplify the construction of pSEVA2311-N200-Mango plasmid. The annealing temperatures were dependent on the primer information on the plasmid maps. The primer sequences used can be seen in Appendix D and the plasmid maps can be seen in sub-chapter 5.1.2. above. The PCR product was verified using gel electrophoresis.

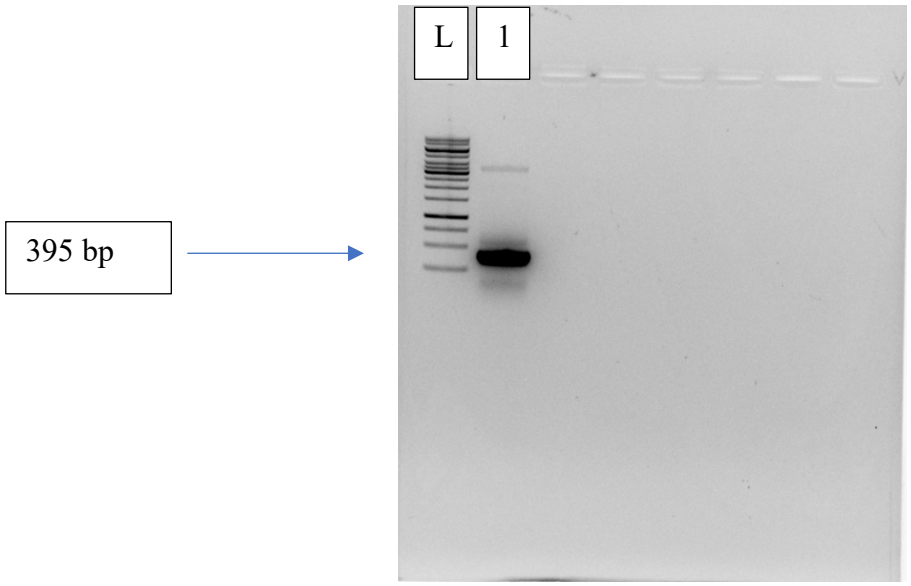

Figure 40: Agarose Gel Electrophoresis for amplification of Mango-3ÚTR Insert. *Well 1 gives an expected band fragment of 395bp. The DNA Ladder can be seen in Appendix F.*

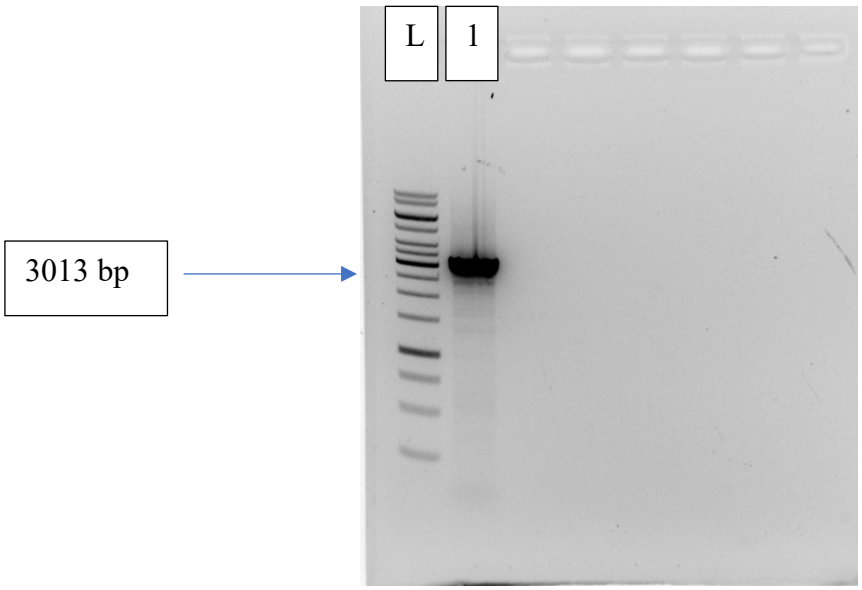

Figure 41: Agarose Gel Electrophoresis for amplification of pSEVA2311 backbone. *Well 1 gives an expected band fragment of 3013bp. The DNA Ladder can be seen in Appendix F.*

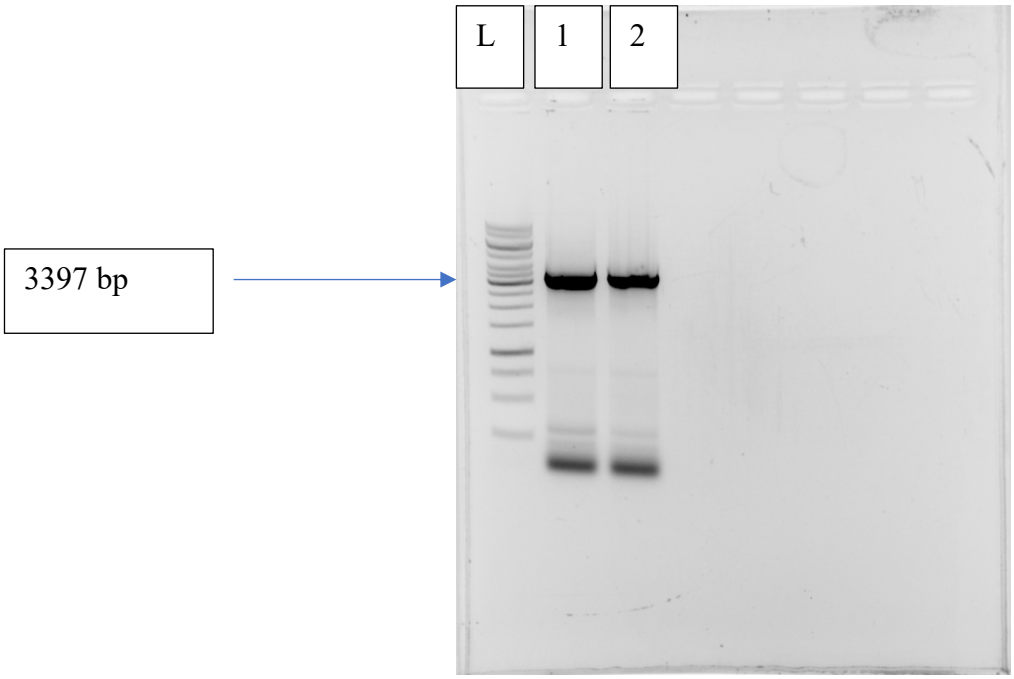

Figure 42: Preparing Golden Gate Backbone (Mango-pSEVA2311). *Well 1 & 2 gives an expected band fragment of 3397bp. The DNA Ladder can be seen in Appendix F.*

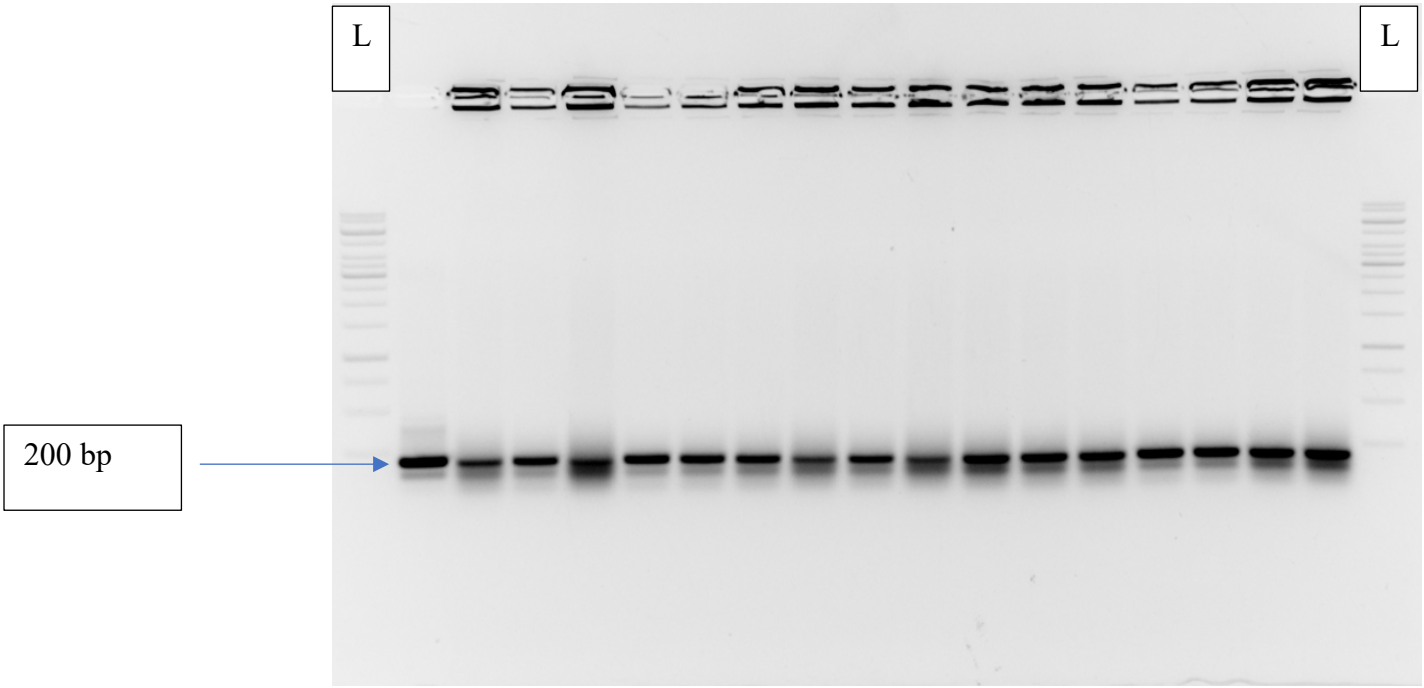

Figure 58: DH5α Colony PCR. *Multiple Colonies were tested for Mango-3ÚTR-pSEVA2311 backbone from DH5α variant. The PCR product has the band fragment size of 200bp. The DNA Ladder can be seen in Appendix F.*

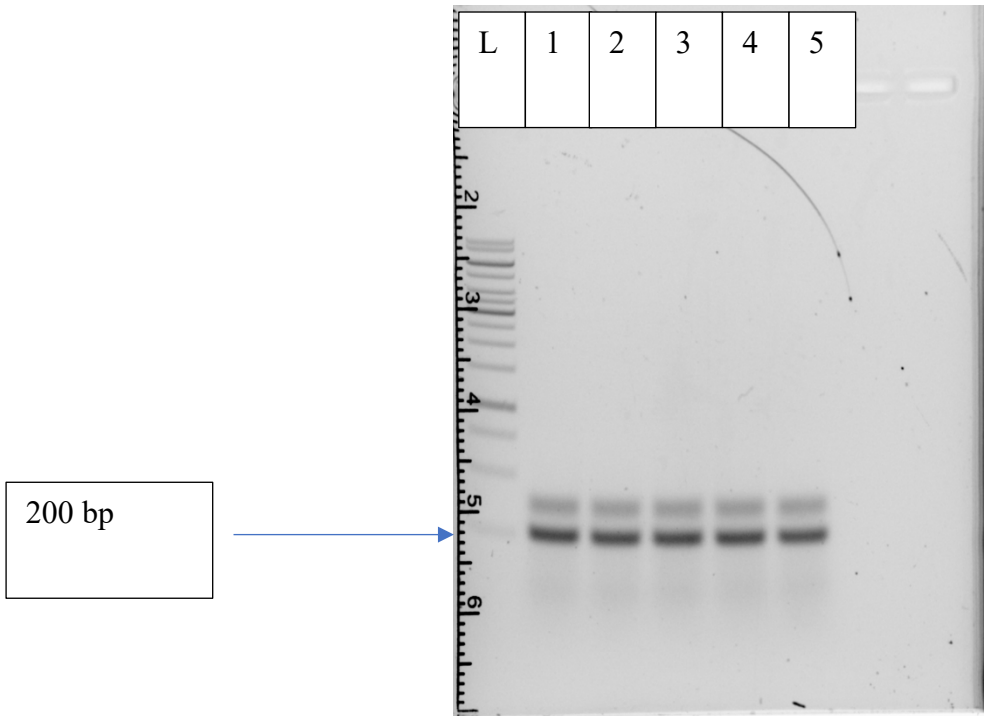

Figure 59: Amplification of N200 random nucleotide for Golden Gate Cloning. *Suffix and Prefix primers (from Lisa Tietze) were used to amplify the N200 Library. Primer sequences are shown in Appendix D. The DNA Ladder can be seen in Appendix F.*

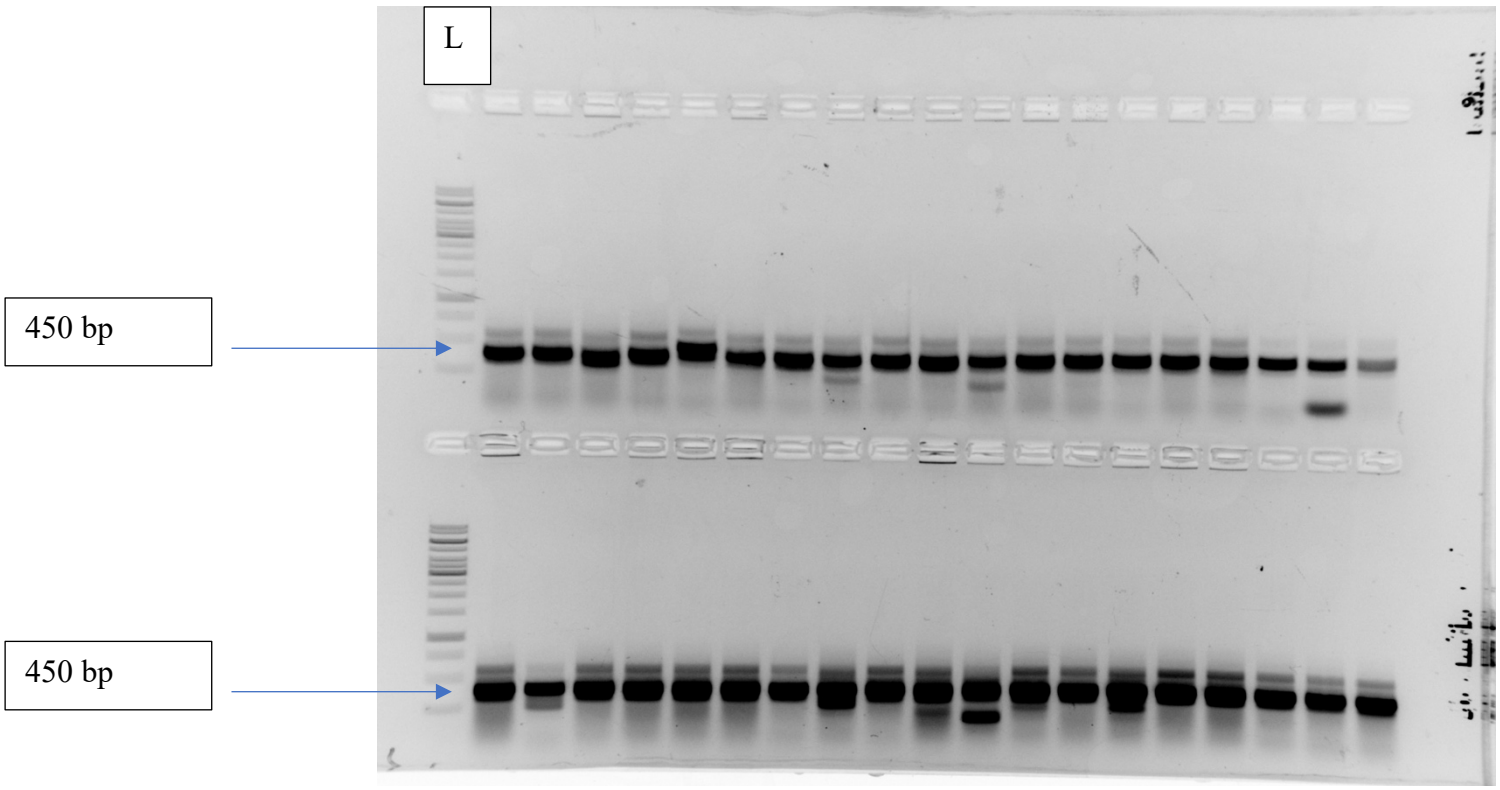

Figure 60: Colony PCR of N200 DNA Library (Mango). *After Golden gate cloning of N200 Random nucleotide Library with Mango-3ÚTR- pSEVA2311 backbone, colony PCR was conducted to confirm the N200 Random nucleotide Promoter insert. The expected fragment was 450bp. The DNA Ladder can be seen in Appendix F.*

6.2. Protein Expression

6.2.2. PspF - Transcriptional activator Protein

The expression construct, Lv3Δt – T7 – PspF (1-275) – AmpR**,** contains the expression gene for the Transcriptional Activator Protein (PspF) and includes a T7 promoter system. This expression construct was achieved through extensive research. The *E. coli* BL21 expression strain is unknown. To express this plasmid construct, the *E. coli* BL21 strain was inoculated in 5ml of LB medium containing ampicillin with a concentration of 50µg/ml in 13ml culture tubes. The cells were incubated overnight at 37°C with shaking at 225rpm in an incubator. The following day, 250µl of the overnight culture was inoculated in a 250ml flask with 25ml of LB medium containing ampicillin with a concentration of 50µg/ml. This was incubated at 30°C with shaking at 160rpm for 2.5 hours until the OD_{600} was in the interval 0.7-0.8. Once the expected OD_{600} was achieved, the induction was initiated by IPTG, and the incubation was extended for another 5 hours. The cells were harvested by centrifugation and the pellet stored at -80°C for future use. As a result of a technical error, no visible expression was shown during SDS-PAGE, DmpR was therefore used as an alternative to PspF. This was received from our collaborators as a plasmid and a purified protein and was used in the *in vitro* bulk transcription experiments. The plasmid was used in the *In vivo* transcription experiments and was induced by arabinose after transformation in the *E. coli* BL21 strains.

6.2.3. Sigma Factor 54

RpoN cells were received from our collaborators in Sweden during the early stages of this thesis and expression of these cells were conducted as follows. The expression construct received was called, pJF5401, and it contained the expression gene for Sigma Factor 54 (RpoN). This expression construct includes a heat inducible promoter. The host *E. coli* strain is unknown. To express this plasmid construct, the *E. coli* strain was inoculated in 5ml of LB medium containing ampicillin with a concentration of 50µg/ml in 13ml culture tubes. The cells were incubated overnight at 37°C with shaking at 225rpm in an incubator. The following day, 250µl of the overnight culture was inoculated in a 250ml flask with 25ml of LB medium containing ampicillin with a concentration of 50µg/ml. This was incubated at 30°C with shaking at 160rpm for 2.5 hours until the OD_{600} was in the interval 0.7-0.8. Once the expected OD_{600} was achieved, the induction was initiated by increasing the temperature to 42°C and the incubation was extended for another 5 hours. The negative control sample was grown without heat induced induction and thus remained at 30°C for a further 5 hours. The cells were

harvested by centrifugation and washed once with M9 salts, and the pellet run on SDS-PAGE gel to confirm the expression.

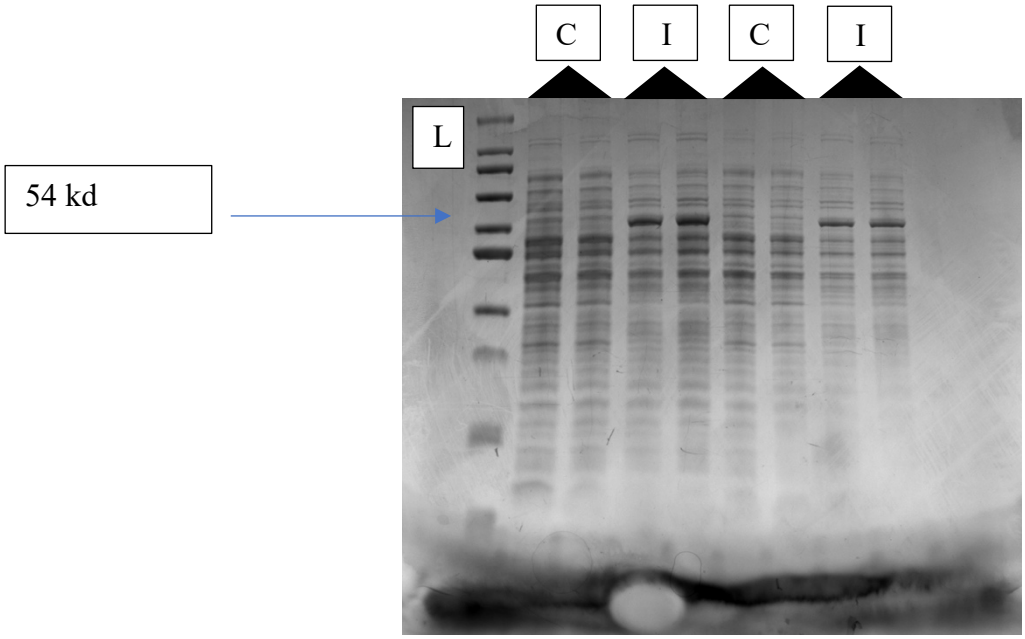

Figure 61: Expression of E.coli RpoN (Sigma Factor 54) with heat-inducible promoter. *The cells were harvested and washed. SDS-PAGE was conducted with two control samples and two induced samples. The "C" is the negative control, and "I" is the induced sample. The expected Protein fragment had a molecular mass of 54kDa. The Protein molecular weight standard can be seen in Appendix F.*

As a result of a technical error for an extended period, the *E. coli* strain from Sweden became unviable and therefore expired. This compelled the cloning of Lvt3t_T7_RpoN_ AmpR plasmid with a T7 promoter in *E. coli* BL21 expression strain. The same expression protocol was followed as for the expression construct, pJF5401, for the unknown *E. coli* strain, however, in this construct, IPTG was used to induce the T7 promoter system, and the sample incubated for a further 5 hours. The cells were harvested by centrifugation, washed once with M9 salts and the pellet run on SDS-PAGE gel to confirm the expression.

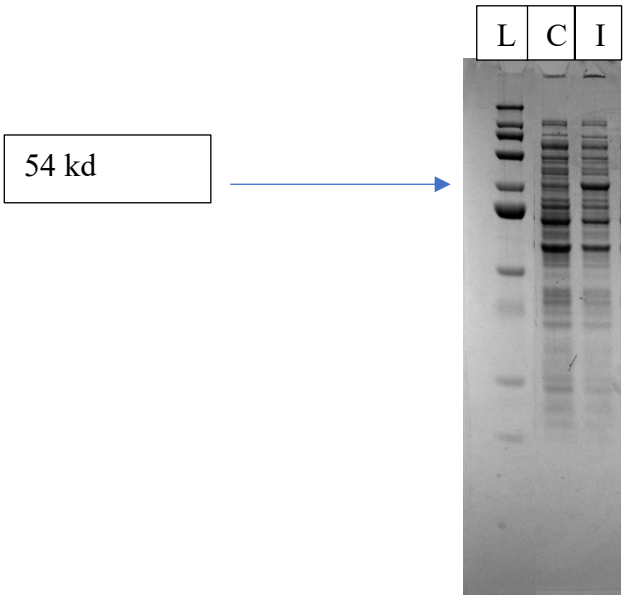

Figure 62: Expression of Cloned E.coli RpoN (Sigma Factor 54) with T7 promoter system. *The cells were harvested and washed. SDS-PAGE was conducted with one control sample and one induced sample. The "C" is the negative control, "I" is the induced sample and "L" is the MW Protein Ladder. The expected Protein fragment had a molecular mass of 54kDa. The Protein molecular weight standard can be seen in Appendix F.*

Since this sample was not purified through Fast Protein Liquid Chromatography (FPLC), it was therefore insoluble and unable to be utilized for the *In vitro* bulk transcription. As a result, an expressed and purified sigma factor 54 protein was received from our collaborators in Sweden again. This was then used in *In vitro* bulk transcription.

6.3. *In Vivo* Transcription

In vivo transcription was conducted to determine the time the transcription process occurs in the cell using its own cellular components. pO promoter sequence was utilized and no inducer was used for this process. In accordance with the exact time the transcription process occurs using the *E. coli* pO promoter sequence, the experiment was expected to be reconducted using the artificial N200 random nucleotide DNA library. DmpR-His plasmid was used from *Pseudomonas Putida* and the bacteria was streaked out on LB agar plate and grown. From the growth, a colony was cultured and eventually the plasmid was isolated through miniprep. The isolated DmpR-His plasmid was then transformed into the *E. coli* BL21 strain. *E. coli* pO promoter sequence with mCherry was transformed into the *E. coli* BL21 strain and the cells were streaked out and grown on a plate. The positive colonies were confirmed using colony

PCR and inoculated in 5ml LB medium with kanamycin present. The following day, the pre-culture was used to inoculate the *E. coli* strain into the main culture with 10ml of LB medium with kanamycin present. The culture was grown to OD_{600} 0.15 at 37°C and the incubation period took around 6-25 hours. The DmpR-His plasmid included an arabinose-inducible promoter P(BAD) promoter system and was induced using 2wt% arabinose. After induction, the sample was grown for a further 4 hours until an OD_{600} of 0.6-0.7 was reached. This is considered the mid to late log-phase. The OD_{600} was measured every hour to check the growth. Once the required OD_{600} value was reached, fluorescence was measured through TECAN 96-well microplate fluorescence reader.

Table 15: In vivo Transcription results showing OD_{600} and mCherry Fluorescence values for the Negative Control Sample (Not Induced with arabinose after 2.5hours).

Negative Control

Time (Hours)	OD_{600}	Fluorescence	*Cycle*
0	0.0538	11811	*14:30 PM*
2.5h	0.3716	10754	***Not Induced (17:00 PM)***
19h	0.6427	12215	*09:30 AM*
21h	0.6393	10429	11:30AM
23h	0.6674	9748	13:30PM
25h	0.3942	10767	15:30PM
45h	0.5781	13478	11:30 PM
47h	0.6045	10490	13:30 PM
66h	0.8148	10591	08:30 AM

Table 16: In vivo Transcription results showing OD_{600} and mCherry Fluorescence values for the Sample (Induced with arabinose after 2.5hours).

Induced Sample

Time	OD_{600}	Fluorescence	Cycle
0	0.0528	11717	*14:30 PM*
2.5h	0.3749	11073	***Induced (17:00 PM)***
19h	0.6385	11236	*09:30 AM*
21h	0.5965	10797	11:30AM

23h	0.6367	9652	13:30PM
25h	0.3804	10694	15:30PM
45h	0.5334	13654	11:30 PM
47h	0.5190	10864	13:30 PM
66h	0.6075	10509	08:30 AM

The graph depicts an unsuccessful In vivo transcription as both values show significant levels of similarity in mCherry fluorescence levels, hence indicating that the *E. coli* pO promoter sequence was not transcribed and the mCherry was not expressed in the pHH100-mcherry-S54pO plasmid. The reason behind this was later discovered to be a mutation in the pO promoter sequence. This was confirmed from the sequencing results that were received back later.

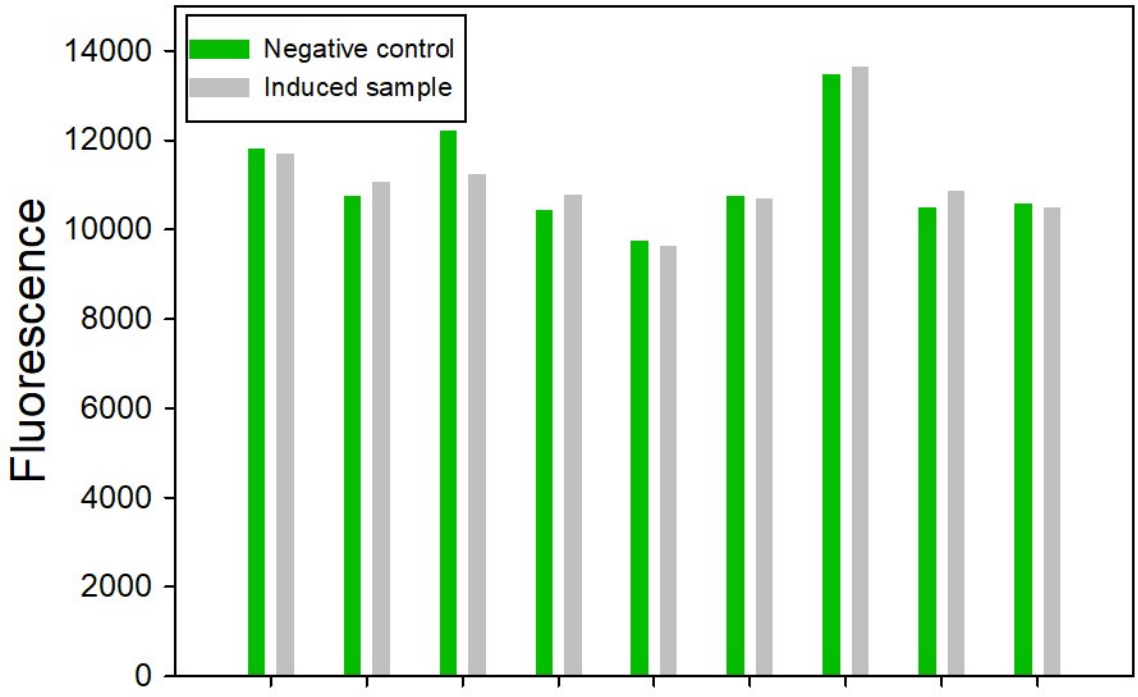

Figure 63: Graph Illustrating a comparison between the mCherry Fluorescence values over time for the Induced sample and the Non-induced Negative Control. *The graph depicts an unsuccessful In vivo transcription as both values show significant levels of similarity in mCherry fluorescence levels, hence indicating that the mCherry in the pHH100-mcherry-S54pO plasmid was not expressed.*

6.4. *In Vitro* Bulk Transcription

Since the Cloned *E. coli* RpoN (Sigma Factor 54) with T7 promoter system was not purified through Fast Protein Liquid Chromatography (FPLC), it was therefore insoluble and unable to be utilized for the *In vitro* bulk transcription. As a result, an expressed and purified sigma factor 54 protein was received from our collaborators in Sweden again. This was then used in *In vitro* bulk transcription. Three in vitro bulk transcriptions were conducted with different concentration values. The outcome for each experiment is illustrated in the tables below.

Table 17: In vitro transcription 1. *In vitro transcription by Sigma 54 E. coli using N200-Mango library.*

Sample Type	Concentration	Absorbance (A260/A280)	Purity (A260/A230)
Control	15.8	1.78	0.44
Induced	37.5	2.30	1.62

Table 18: In vitro Transcription 2. *In vitro transcription by Sigma 54 E. coli using N200-Mango library.*

Sample Type	Concentration	Absorbance (A260/A280)	Purity (A260/A230)
Control	51.2	2.23	0.79
Induced	20.9	1.82	0.75

Table 19: In vitro transcription 3. *In vitro transcription by Sigma 54 E. coli using N200-Mango library. Here the DNA batch deviates from the first two attempts. This is due to the inclusion of TOI-Biotin dye in the reaction.*

Sample Type	Concentration	Absorbance (A260/A280)	Purity (A260/A230)
Control	9.6	1.75	2.01
Induced	131.6	2.75	2.60

As we can see, *In vitro* Transcription with the inclusion of TOI-Biotin dye gives the highest concentration of the RNA transcript (Table 19). The following graph depicts the fluorescence levels of the *in vitro* transcription 3 for the induced sample and the control sample.

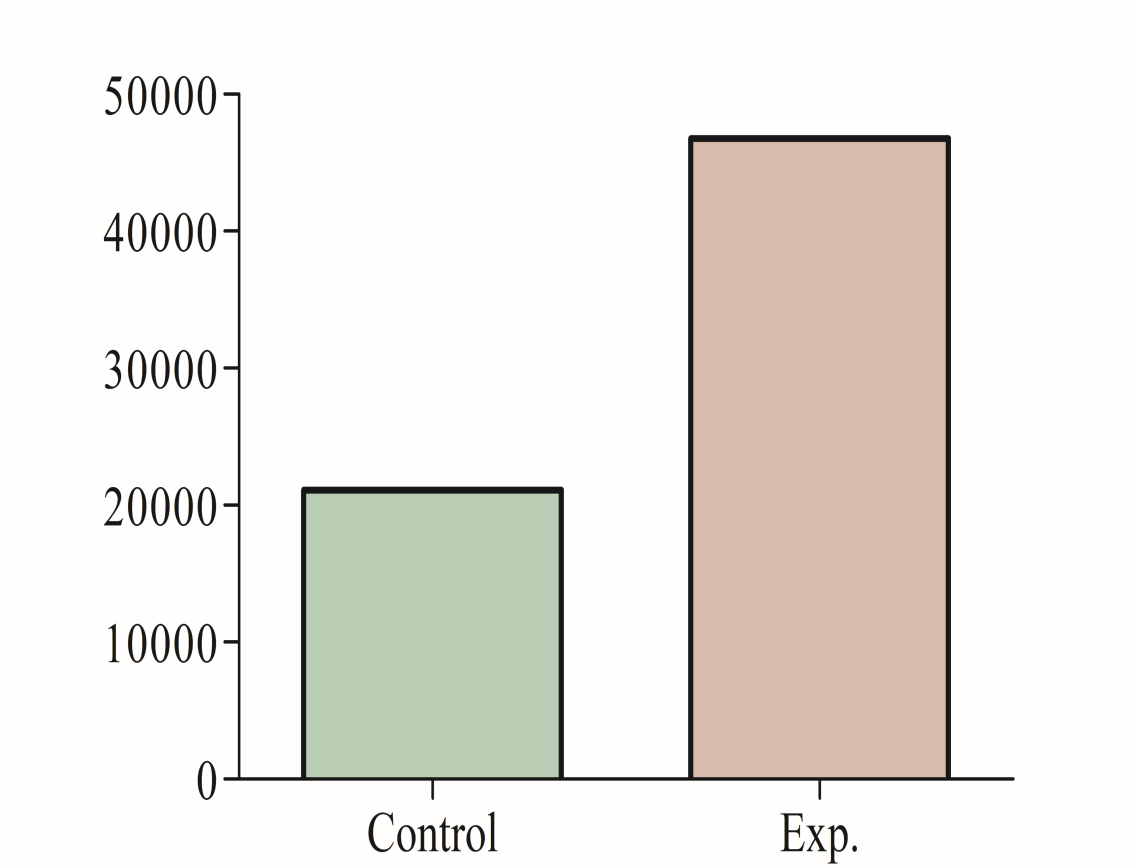

Figure 64: In vitro transcription by Sigma 54 E. coli using N200-Mango library. *The graph clearly depicts the fluorescence measurements of the induced sample and the control sample, hence showing a successful expression.*

7. Discussion

The overall project's primary goal is to decode the *cis*-acting DNA sequences that are particular to the sigma factor through *in vivo* and *in vitro* experimental work in order to produce massive data sets employing the three bacterial species. The second key goal is to create *in silico* data analysis pipelines that use statistics and machine learning to produce algorithms that can determine the number of matching transcripts and the precise *cis*-binding motifs for sigma factors. Establishing a cloud-based service with the intention of using it for more extensive scientific results, data, and tools would be the third and final key goal of the overall project. In relation to this thesis, the main aim was to investigate the transcriptional regulation in *Escherichia coli* by σ54. Synthetic biology methods were applied to bring together the *Escherichia coli* sigma 54 (σ54), Transcriptional activator protein and the core RNA polymerase for *In vitro* transcription.

The synthetic biology methods include, cloning of plasmid constructs, expression of the plasmid constructs, *in vivo* transcription, and *in vitro* bulk transcription. Q5 PCR reactions were used to amplify sigma factor 54 (RpoN), transcriptional activator protein (PspF) from *E. coli* genomic DNA and to amplify the N200 random DNA library and the pO promoter sequence using primers designed beforehand. PCR purification was conducted on sigma factor 54 (RpoN), transcriptional activator protein (PspF), the N200 DNA promoter library and the pO promoter sequence after Q5 PCR amplification. Agarose Gel electrophoresis was conducted to separate the DNA molecules by size and confirm their relative sizes and purity. The sample products produced from the PCR amplification and the PCR purification were run on the gel. After conducting a gel electrophoresis on the samples, the desired DNA product fragment was excised resulting in a purified DNA gel fragment. Nanodrop ONE Thermo scientific spectrophotometer was used to measure the concentration of the different cloned plasmid constructs after cloning, PCR purification and Gel fragment purification, i.e., Sigma factor 54, Transcriptional activator Protein (PspF), N200 random DNA Library (Mango), pO Promoter Sequence and the *In vitro* RNA transcript. The PCR products that resulted from the PCR amplification cloning procedure were DpnI digested. Gibson assembly was conducted to join the cloned linear insert to the respective plasmid vector backbone. This was done when cloning Sigma Factor 54, Transcriptional activator Protein (PspF), pO Promoter Sequence and when incorporating the Mango aptamer to the pSEVA-2311 vector backbone. The RNA aptamer, mango, has great significance in transcript level recognition and can be utilized to quantify the

transcript levels that are obtained as a result from expressing artificial promoters. The non-fluorescent aptamer binds with the non-fluorescent fluorogen resulting in the aptamer lighting up. In this case, the RNA aptamer mango binds with the TO1-biotin with high affinity to light up (fluoresces) brightly. Mango paired with the uorophore, Thizole Orange (TO)-biotin, also gives low background and has a high signal-to-noise ratio. As a result, Mango is a perfect indicator for successfull IVT reactions *(Trachman, 2019)*. These have therefore been utilised immensely when the purpose is to activate small-molecule fluorophores through *in vitro* attempts *(Trachman, 2019).* Golden gate cloning was used to incorporate the N200 random DNA library into the pSEVA – 2311 – Mango plasmid backbone using BsaI High Fidelity Restriction Enzyme and T4 DNA ligase enzyme. The different cloning plasmid vectors of Sigma Factor 54, Transcriptional activator protein, N200 random DNA library and pO promoter sequence were all transformed to the *E. coli* $DH_5\alpha$ cells to confirm the PCR products produced. For checking the expression, Sigma Factor 54 and Transcriptional activator protein were later transformed into *E. coli* BL21 cells and expressed using IPTG as an inducer. *E. coli* BL21 cells were also used to transform the pO promoter sequence for *In vivo* screening procedure. Colony PCR was used to screen colonies of *E. coli* $DH_5\alpha$ cells for the presence of the insert DNA. Plasmid miniprep was conducted on sigma factor 54 (RpoN) for BL21 transformation and expression, transcriptional activator protein (PspF) for BL21 transformation, expression and *In vivo* screening, the N200 DNA promoter library for *In vitro* bulk transcription and the pO promoter sequence for *in vivo* screening. Plasmid sequencing was conducted through Eurofins Genomics LightRun Tube. The plasmids that were isolated through plasmid miniprep were sent for sequence confirmation. This was to determine whether the nucleotides are in the correct order (i.e., confirming the presence of the insert DNA) as expected in the genomic DNA after the cloning procedure.

Escherichia coli sigma 54 (σ54) and Transcriptional activator protein were expressed. The expression construct, pJF5401, contained the expression gene for Sigma Factor 54 (RpoN) and included a heat inducible promoter. This expression construct was achieved from our collaborators in Sweden. The host *E. coli* strain is unknown. As a result of a technical error for an extended period, the *E. coli* strain from Sweden became unviable and therefore expired. This compelled the cloning of Lvt3t_T7_RpoN_ AmpR plasmid with a T7 promoter in *E. coli* BL21 expression strain. The same expression protocol was followed as for the expression construct, pJF5401, for the unknown *E. coli* strain, however, in this construct, IPTG was used

to induce the T7 promoter system. Since this sample was not purified through Fast Protein Liquid Chromatography (FPLC), it was therefore insoluble and unable to be utilized for the *In vitro* bulk transcription. As a result, an expressed and purified sigma factor 54 protein was received from our collaborators in Sweden again. This was then used in *In vitro* bulk transcription. The expression construct, Lv3Δt – T7 – PspF (1-275) – AmpR**,** contained the expression gene for the Transcriptional Activator Protein (PspF) and includes a T7 promoter system. This expression construct was achieved through extensive research. The *E. coli* BL21 expression strain is unknown. As a result of a technical error, no visible expression was shown during SDS-PAGE, DmpR was therefore used as an alternative to PspF. This was received from our collaborators as a plasmid and a purified protein and was used in the *in vitro* bulk transcription experiments. The plasmid was used in the *In vivo* transcription experiments and was induced by arabinose after transformation in the *E. coli* BL21 strains.

In vivo transcription was conducted to determine the time the transcription process occurs in the cell using its own cellular components. pO promoter sequence was utilized and no inducer was used for this process. In accordance with the exact time the transcription process occurs using the *E. coli* pO promoter sequence, the experiment was expected to be reconducted using the artificial N200 random nucleotide DNA library. In vivo screening has many knowledgeable outcomes, the differential expression profiles achieved through *in vivo* screening enables the determination of whether particular sequences are subjected to a specific *in vivo* regulation, e.g., anti-σ-factors, since such type of regulation would be absent In a IVT in bulk reaction *(Browning, 2016).*

The results from the *In vivo* transcription experiment depicted an unsuccessful outcome as both values show significant levels of similarity in mCherry fluorescence levels, hence indicating that the *E. coli* pO promoter sequence was not transcribed and the mCherry was not expressed in the pHH100-mCherry-S54pO plasmid. The reason behind this was later discovered to be a mutation in the pO promoter sequence. This was confirmed from the sequencing results that were received back later. An eventual follow up attempt would require the re - cloning of the pHH100-mcherry-S54pO plasmid for a potential successful result to be achieved. *In vitro* transcription in bulk was conducted for the purpose of RNA sequencing the generated mRNAs and map them to the DNA sequence data from the *in vivo* screening with random DNA, in order to determine the promoter sequences *(Lacour, 2003).* In addition, IVT in bulk will also serve the purpose of being a quality control experiment to assess the functionality of both the

RNAP and the Sigma Factor. Three *in vitro* bulk transcriptions were conducted with different concentration values. *In vitro* Transcription with the inclusion of TOI-Biotin dye gives the highest concentration of the RNA transcript, hence indicating a confirmed successful binding and expression. The synthetically produced RNA transcript from the *in vitro* transcription was then further sequenced using next generation sequencing, to confirm the binding of *Escherichia coli* sigma 54 (σ54) to specific DNA promoter sequences. One of the advantages from the *In vitro* in bulk and *in vivo* screening experiments is that it enables us to obtain functional (transcriptionally active) and non-functional (transcriptionally not active) DNA sequences. These functional and non-functional DNA sequences can potentially aid us in differentiating features and their role for the development of predictive algorithms.

Sigma factor and promoter association is largely obtained from genome-wide studies that has a very minor connection with experimental work that could prove promoter regulation by their corresponding sigma factors *(Rhodius, 2010).* This therefore means that promoter sequences are predicted, annotated, and utilized without having knowledge of whether they can be the subject of any type of regulation. This is therefore an unresolved problem that proposes a challenge in gene expression and synthetic biology research. In addition to the methods explained and utilized in this thesis work, there are several other methods that exist and can be used in finding a solution to the *cis*-acting sigma factor sequences in promoters. These methods include, chromatin immunoprecipitation (ChIP) combined with microarray (ChIP-chip), ChIP-sequencing, and systematic evolution of ligands by exponential enrichment (SELEX). However, with every method there's a disadvantage, in this case, the above-mentioned methods are only applicable for the determination of the binding sequences and not the quantification of the transcripts. These methods can also not be used in *in vivo* screening efforts in relation to the scale that is proposed for the overall research project. The other methods involve *In vitro* reactions that are based on creation of RNA transcripts using P^{32} radioactive labelled nucleotide A/UTP and using polyacrylamide gels (PAGE) to quantify the generated transcript. The disadvantage here is that the process is laborious, results in semi-quantitative information and can be initiated with a limited number of promoter sequences *(Lacour, 2003), (Dostálová, 2017)*. As a result, the above-mentioned methods do not provide a resolution to determine the sigma factor binding sequence or to quantify the resulting transcripts both *in vivo* and *in vitro.* One of the potential impacts of investigating the transcriptional regulation in *Escherichia coli* by σ54 is that it could provide us with opportunities for the construction of artificial devices. Not only does the correct association of sigma factors to promoter regions provide us with

knowledge of how bacteria can sense and respond to changes in the environment but also provide unique opportunities for the construction of artificial devices for the purpose of genetic control. An example of such an artificial device can be chimeric proteins that will be equipped with novel promoter-recognition characteristics aswell as the ability to function in multiple diverse organism *(Dostálová, 2017).* In addition, another benefit of this investigation is the expansion of the knowledge on bacterial virulence. Sigma-factors are known to play a significant role on bacterial virulence in human diseases and biofilm formation *(Iyer, 2020).* Since bacterial infections are gradually increasing in occurrence, the knowledge that will be generated through this investigation will be of very high value. As it can aid in the determination of sigma factor specific virulence gene regulation or by regulating genes that leads to the increase in shape and fitness of the bacterium during bacterial transmission and infection.

8. Conclusion

In conclusion for this study conducted in this thesis, the main aim of this master thesis was to investigate the transcriptional regulation in *Escherichia coli* by σ54 using synthetic biology methods. This was to bring together the *Escherichia coli* sigma 54 (σ54), Transcriptional activator protein and the core RNA polymerase for *In vitro* transcription and to therefore shed light upon the *cis*-acting DNA sequence elements that bind with different σ-factor in bacteria. Through extensive lab work, Sigma Factor 54 Plasmid with T7 Promoter was cloned, and the colonies confirmed thorugh colony PCR. The cloned RpoN (Sigma Factor 54) plasmid was also successfully expressed with the T7 promoter system. PspF - Transcriptional activator Protein plasmid with T7 Promoter System was also cloned but no visible expression was shown. *E. coli* pO Promoter Sequence Plasmid with mCherry was successfully cloned and the cloning confirmed through colony PCR. The N200 Random Nucleotide DNA Library with Mango Aptamer was successfully cloned using golden gate cloning and the cloning was confirmed through colony PCR. The *in vivo* transcription graph illustrated an unsuccessful reaction as both values show significant levels of similarity in mCherry fluorescence levels, hence indicating that the *E. coli* pO promoter sequence was not transcribed and the mCherry was not expressed in the pHH100-mcherry-S54pO plasmid. The IVT in bulk yielded a very high concentration for the induced sample with the inclusion of TOI-Biotin dye, giving the highest concentration of the RNA transcript. This result hence indicates a confirmed successful binding and expression. The synthetically produced RNA transcript from the *in vitro* transcription was then further sequenced using next generation sequencing, to confirm the binding of *Escherichia coli* sigma 54 (σ54) to specific DNA promoter sequences.

9. Future Work

In terms of future work, different applications can be utilized to study the Sigma factor specific promoter dependency in bacteria. Methods such as IVT through droplets (Microfluidics) can be utilized in comparison to IVT in bulk. Droplet assays enable the quantification for the concentration of the mRNA that is generated through IVT using the RNA aptamer mango. This also allows us to avoid the use of radioactive labels such as P^{32}. Microfluidics is one of the main applications for IVT in droplets and it is carried out in micro fluidic generated aqueous droplets. These microfluidic experiments are carried out in devices that are adapted to supporting the control of droplet size and the loading of biochemicals. IVT can be implemented by two ways. One way is by an off-chip followed by re-injection to a pico-injection chip for introduction of TO1-biotin. The other way is on-chip incubation on a chip that is designed to support the flow from the droplet generation, incubation, eventual pico-injection and sorting based on the mango signal.

References

1. ALISA YUROVSKY, M. R. A., JUSTIN GARDIN,YUPING CHEN,STEVE SKIENA,BRUCE FUTCHER 2018. Prokaryotic coding regions have little if any specific depletion of Shine-Dalgarno motifs. *PLOS ONE*.
2. AMY E.DANSON, M. J., MARTIN BUCK, XIAODONG ZHANG 2019. Mechanisms of σ54-Dependent Transcription Initiation and Regulation. *Journal of Molecular Biology*.
3. BERG, L., LALE, R., BAKKE, I., BURROUGHS, N., & VALLA, S. 2009. The expression of recombinant genes in Escherichia coli can be strongly stimulated at the transcript production level by mutating the DNA-region corresponding to the 5'-untranslated part of mRNA. *Microbial biotechnology*.
4. BLAZECK, J., & ALPER, H. S. 2013. Promoter engineering: recent advances in controlling transcription at the most fundamental level. *Biotechnology journal*.
5. BROWNING, D. F., & BUSBY, S. J. 2016. Local and global regulation of transcription initiation in bacteria. *Nature reviews. Microbiology*.
6. CATHERINE SUTHERLAND, K. S. M. 2018. An Introduction to the Structure and Function of the Catalytic Core Enzyme of Escherichia coli RNA Polymerase. *ASM Journals - EcoSal Plus*.
7. CHANEY, M., GRANDE, R., WIGNESHWERARAJ, S. R., CANNON, W., CASAZ, P., GALLEGOS, M. T., SCHUMACHER, J., JONES, S., ELDERKIN, S., DAGO, A. E., MORETT, E., & BUCK, M. 2001. Binding of transcriptional activators to sigma 54 in the presence of the transition state analog ADP-aluminum fluoride: insights into activator mechanochemical action. *Genes & development*.
8. CLARK, D. P., PAZDERNIK, N. & MCGEHEE, M. 2018. *Molecular Biology*, Elsevier Science Publishing Co Inc.
9. CONTRIBUTORS, W. 2022. *Stem-loop* [Online]. Wikipedia, The Free Encyclopedia. Available: https://en.wikipedia.org/w/index.php?title=Stem-loop&oldid=1070314016 [Accessed].
10. COPPINS, R. L., HALL, K. B., & GROISMAN, E. A. 2007. The intricate world of riboswitches. *Current opinion in microbiology*.
11. CRAIG, M. L., TSODIKOV, O. V., MCQUADE, K. L., SCHLAX, P. E., JR, CAPP, M. W., SAECKER, R. M., & RECORD, M. T., JR 1998. DNA footprints of the two kinetically significant intermediates in formation of an RNA polymerase-promoter open complex: evidence that interactions with start site and downstream DNA induce sequential conformational changes in polymerase and DNA. *Journal of molecular biology*.
12. D., B. Z. 2015. The unexhausted potential of E. coli. *eLife*.
13. DAHL, R. H., ZHANG, F., ALONSO-GUTIERREZ, J., BAIDOO, E., BATTH, T. S., REDDING-JOHANSON, A. M., PETZOLD, C. J., MUKHOPADHYAY, A., LEE, T. S., ADAMS, P. D., & KEASLING, J. D. 2013. Engineering dynamic pathway regulation using stress-response promoters. *Nature biotechnology*.
14. DAVID P. CLARK, N. J. P. 2016. DNA, RNA, and Protein. *Biotechnology (Second Edition)*.
15. DE SMIT, M. H., & VAN DUIN, J. 1990. Control of prokaryotic translational initiation by mRNA secondary structure. *Progress in nucleic acid research and molecular biology*.
16. DOSTÁLOVÁ, H., HOLÁTKO, J., BUSCHE, T. ET AL. 2017. Assignment of sigma factors of RNA polymerase to promoters in Corynebacterium glutamicum. *AMB Expr 7*.

17. F., B. 1999. Recombinant protein expression in Escherichia coli. . *Current opinion in biotechnology*.
18. FUJITA, N., & ISHIHAMA, A. 1996. Reconstitution of RNA polymerase. *Methods in enzymology*.
19. GAMA-CASTRO, S., SALGADO, H., SANTOS-ZAVALETA, A., LEDEZMA-TEJEIDA, D., MUÑIZ-RASCADO, L., GARCÍA-SOTELO, J. S., ALQUICIRA-HERNÁNDEZ, K., MARTÍNEZ-FLORES, I., PANNIER, L., CASTRO-MONDRAGÓN, J. A., MEDINA-RIVERA, A., SOLANO-LIRA, H., BONAVIDES-MARTÍNEZ, C., PÉREZ-RUEDA, E., ALQUICIRA-HERNÁNDEZ, S., PORRÓN-SOTELO, L., LÓPEZ-FUENTES, A., HERNÁNDEZ-KOUTOUCHEVA, A., DEL MORAL-CHÁVEZ, V., RINALDI, F., … COLLADO-VIDES, J. 2016. RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic acids research*.
20. GOLDMAN, S. R., EBRIGHT, R. H., & NICKELS, B. E. 2009. Direct detection of abortive RNA transcripts in vivo. *Science (New York, N.Y.)*.
21. GRUBER, T. M., & GROSS, C. A. 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual review of microbiology*.
22. HAJNSDORF, E., & BONI, I. V. 2012. Multiple activities of RNA-binding proteins S1 and Hfq. *Biochimie*.
23. HAUSJELL, J., KUTSCHA, R., GESSON, J. D., REINISCH, D., & SPADIUT, O. 2020. The Effects of Lactose Induction on a Plasmid-Free E. coli T7 Expression System. *Bioengineering (Basel, Switzerland)*.
24. HSU, L. M., COBB, I. M., OZMORE, J. R., KHOO, M., NAHM, G., XIA, L., BAO, Y., & AHN, C. 2006. Initial transcribed sequence mutations specifically affect promoter escape properties. *Biochemistry*.
25. HUO, Y. X., GUO, L., & MA, X. 2018. Biofuel production with a stress-resistant and growth phase-independent promoter: mechanism revealed by in vitro transcription assays. *Applied microbiology and biotechnology*.
26. IYER, S. C., CASAS-PASTOR, D., KRAUS, D., MANN, P., SCHIRNER, K., GLATTER, T., FRITZ, G., & RINGGAARD, S. 2020. Transcriptional regulation by σ factor phosphorylation in bacteria. *Nature microbiology*.
27. JERVIS, A. J., CARBONELL, P., TAYLOR, S., SUNG, R., DUNSTAN, M. S., ROBINSON, C. J., BREITLING, R., TAKANO, E., & SCRUTTON, N. S. 2019. SelProm: A Queryable and Predictive Expression Vector Selection Tool for Escherichia coli. *ACS synthetic biology*.
28. JIA, B., & JEON, C. O. 2016. High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives. *Open biology*.
29. JONES, S. 2009. Sigma 54 minds the gap. *Nat Rev Microbiol 7*.
30. JOVANOVIC, G., RAKONJAC, J., & MODEL, P. 1999. In vivo and in vitro activities of the Escherichia coli sigma54 transcription activator, PspF, and its DNA-binding mutant, PspFDeltaHTH. *Journal of molecular biology*.
31. KARKI, G. 2017. *Transcription in Prokaryotes* [Online]. Available: https://www.onlinebiologynotes.com/transcription-in-prokaryotes/ [Accessed].
32. KAZMIERCZAK, M. J., WIEDMANN, M., & BOOR, K. J. 2005. Alternative sigma factors and their roles in bacterial virulence. *Microbiology and molecular biology reviews : MMBR*.
33. KUDLA, G., MURRAY, A. W., TOLLERVEY, D., & PLOTKIN, J. B. 2009. Coding-sequence determinants of gene expression in Escherichia coli. *Science (New York, N.Y.)*.

34. LACOUR, S., KOLB, A., & LANDINI, P. 2003. Nucleotides from -16 to -12 determine specific promoter recognition by bacterial sigmaS-RNA polymerase. *The Journal of biological chemistry*.
35. LAILA BERG, R. L., INGRID BAKKE,NIGEL BURROUGHS,SVEIN VALLA 2009. The expression of recombinant genes in Escherichia coli can be strongly stimulated at the transcript production level by mutating the DNA-region corresponding to the 5′-untranslated part of mRNA. *Microbial Biotechnology*.
36. LAKSHMI TRIPATHI, Y. Z., ZHANGLIN LIN 2014. Bacterial sigma factors as targets for engineered or synthetic transcriptional control. *Front. Bioeng. Biotechnol.*
37. LALE, R. 2020. An interdisciplinary data-driven approach to resolve sigma-factor-speci c promoter dependency in bacteria (SiD).
38. LISA TIETZE, R. L. 2021. Importance of the 5′ regulatory region to bacterial synthetic biology applications. *Microbial Biotechnology*.
39. LIYUAN JIN, S. N., MENGLI XIA,XIAOYAN MA,YI-XIN HUO 2019. Context-dependency of synthetic minimal promoters in driving gene expression: a case study. *Microbial Biotechnology*.
40. LOU, C., STANTON, B., CHEN, Y. J., MUNSKY, B., & VOIGT, C. A. 2012. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nature biotechnology*.
41. LOVE, J. G. J. 2016. Synthetic promoter design for new microbial chassis. *Biochem Soc Trans*
42. M., K. 2005. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene*.
43. MANDAL, M., & BREAKER, R. R. 2004. Gene regulation by riboswitches. *Nature reviews. Molecular cell biology*.
44. MIRZADEH, K., MARTÍNEZ, V., TODDO, S., GUNTUR, S., HERRGÅRD, M. J., ELOFSSON, A., NØRHOLM, M. H., & DALEY, D. O. 2015. Enhanced Protein Production in Escherichia coli by Optimization of Cloning Scars at the Vector-Coding Sequence Junction. *ACS synthetic biology*.
45. MUTALIK, V. K., GUIMARAES, J. C., CAMBRAY, G., LAM, C., CHRISTOFFERSEN, M. J., MAI, Q. A., TRAN, A. B., PAULL, M., KEASLING, J. D., ARKIN, A. P., & ENDY, D. 2013. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nature methods*.
46. OSTERMAN, I. A., EVFRATOV, S. A., SERGIEV, P. V., & DONTSOVA, O. A. 2013. Comparison of mRNA features affecting translation initiation and reinitiation. *Nucleic acids research*.
47. PALMER, I., & WINGFIELD, P. T. 2012. Preparation and extraction of insoluble (inclusion-body) proteins from Escherichia coli. *Current Protocols in Protein Science, Chapter 6, Unit–6.3*.
48. PATRICIA BORDES, S. R. W., JORG SCHUMACHER, XIAODONG ZHANG, MATTHEW CHANEY, AND MARTIN BUCK 2003. The ATP hydrolyzing transcription activator phage shock protein F of Escherichia coli: Identifying a surface that binds 54.
49. PORTELA, R. M., VOGL, T., KNIELY, C., FISCHER, J. E., OLIVEIRA, R., & GLIEDER, A. 2017. Synthetic Core Promoters as Universal Parts for Fine-Tuning Expression in Different Yeast Species. *ACS synthetic biology*.
50. REPOILA, F., & DARFEUILLE, F. 2009. Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biology of the cell*.

51. RHODIUS, V. A., & MUTALIK, V. K. 2010. Predicting strength and function for promoters of the Escherichia coli alternative sigma factor, sigmaE. *Proceedings of the National Academy of Sciences of the United States of America*.
52. SIMONE BALZER LE, I. O., JON ANDREAS LORENTZEN, RAHMI LALE 2020. Dual UTR-A novel 5′ untranslated region design for synthetic biology applications. *Synthetic Biology*.
53. SRIRAM KOSURI, D. B. G., GUILLAUME CAMBRAY, VIVEK K. MUTALIK, YUAN GAO, ADAM P. ARKIN, DREW ENDY, AND GEORGE M. CHURCH 2013. Composability of regulatory sequences controlling transcription and translation in Escherichia coli. *PNAS*.
54. STEITZ, J. A., & JAKES, K. 1975. How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli. . *Proceedings of the National Academy of Sciences of the United States of America*.
55. T.R.HUGHES 2011. *A Handbook of Transcription Factors*, Springer.
56. TANYA TSCHIRHART, V. S., ERIN E. KELLY, ZACHARY SCHULTZHAUS, ERIN NEWRINGEISEN, JEFFREY S. ERICKSON, ZHENG WANG, WHITNEY GARCIA, EMALEIGH CURL, ROBERT G. EGBERT, ENOCH YEUNG, AND GARY J. VORA* 2019. Synthetic Biology Tools for the Fast-Growing Marine Bacterium Vibrio natriegens. *ACS Synth. Biol.*
57. TRACHMAN, R. J., 3RD, AUTOUR, A., JENG, S., ABDOLAHZADEH, A., ANDREONI, A., COJOCARU, R., GARIPOV, R., DOLGOSHEINA, E. V., KNUTSON, J. R., RYCKELYNCK, M., UNRAU, P. J., & FERRÉ-D'AMARÉ, A. R. 2019. Structure and functional reselection of the Mango-III fluorogenic RNA aptamer. *Nature chemical biology*.
58. WEDEL, A., & KUSTU, S. 1995. The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. *Genes & development*.
59. WEMMER, D. E. 2022. *Transcription Initiation by σ54* [Online]. Berkeley Lab. Available: http://www.cchem.berkeley.edu/wemmer/research/sigma54.html [Accessed].
60. YIM, S. S., AN, S. J., KANG, M., LEE, J., & JEONG, K. J. 2013. Isolation of fully synthetic promoters for high-level gene expression in Corynebacterium glutamicum. *Biotechnology and bioengineering*.
61. ZHANG, M., LI, F., MARQUEZ-LAGO, T. T., LEIER, A., FAN, C., KWOH, C. K., CHOU, K. C., SONG, J., & JIA, C. 2019. MULTiPly: a novel multi-layer predictor for discovering general and specific types of promoters. *Bioinformatics (Oxford, England)*.

Appendices

Appendix A: Growth Medium

Table A.1: LB Medium

5,0 g Tryptone
2,5 g Yeast extract
2,5 g NaCl
500 mL distilled water
Autoclave for 20 minutes at 120 °C

Table A.2: LB Agar

5,0 g Tryptone
2,5 g Yeast extract
2,5 g NaCl
500 mL distilled water
Autoclave for 20 minutes at 120 °C

Table A.3: Psi –medium

0,5 % Yeast Extract (5 g/l)
2 % Trypton (20g/l)
0,5 % MgSO4 (5g/l) (MgSO4 x 7 H2O – 10.24 g/l)
pH adjust to 7,6 with KOH (1M).
Autoclave

Table A.4: SOB medium

Per liter:
20g trypton
5g yeast extract
0,5 NaCl
10 ml 250 mM KCl: - **250 mM KCl, pH 7,0:** - **Per 100 ml:**

- **1,86g KCl**

Adjust pH to 7,0 with 5M NaOH (ca. 0,2 ml)

Autoclave

After autoclaving add: 5 ml 2M $MgCl_2$ (sterile)

- **2M $MgCl_2$:**
- **Per 100 ml:**
- **19g $MgCl_2$**
- **Autoclave**

Table A.5: SOC medium

Add 20ml 1M Glucose Solution (sterile) per liter to SOB medium.

- **1M Glucose Solution:**
- **Per 100 ml:**
- **18 g Glucose**
- **Sterile filter**

Aliquot and freeze down at -20°C.

Appendix B: Antibiotics Solutions

Table B.1: Kanamycin stock, 50 mg/mL

0,50 g Kanamycin
10 mL sterile distilled water
Filter through 0,22 µm syringe filter
Allocate into 1,5 mL Eppendorf tubes
Store at -20 °C

Table B.2: Ampicillin stock, 50 mg/mL

0,50 g Ampicillin
10 mL sterile distilled water
Filter through 0,22 µm syringe filter
Allocate into 1,5 mL Eppendorf tubes
Store at -20 °C

Appendix C: Buffer solutions

Table C.1: TFB1 Buffer – 200 ml

0,588g Potassium Acetate
2,42g Rubidium Chloride
0,294g Calcium Chloride
2,0g Manganese (II) chloride
30ml Glycerol
- **Fill up to 180 ml with ion-exchanged water, then measure the pH.** - **Adjust to 5,8 with diluted acetic acid (10%).** - **Fill up to 200 ml with ion-exchanged water.** - **Sterile filter**

Table C.2: TFB2 Buffer - 100 ml

0,21g MOPS
1,1g Calcium Chloride
0,121g Rubidium Chloride
15ml Glycerol
- **pH is adjusted to 6,5 with diluted NaOH.** - **Ion - exchanged water is added to 100 ml** - **Sterile filter**

Table C.3: Agarose solution

3.2 g Agarose (Cambrex)
400 mL 1x TAE solution
microwaved, 5 min
20 µL GelRed

Table C.4: M9 Salt

Dissolve 56,4g in 1 litre of distilled water
Autoclave for 15min at 121°C

Table C.5: Lysis Buffer

50mM Tris HCl 7.8

1mM DTT
1mM EDTA
10% Glycerol
Protease Inhibitor (Phenylmethanesulfonylfluoride, 23µg/ml

Appendix D: Primer Sequences

All the primer sequences that were used for PCR amplification in this thesis are listed in the tables below.

Table D.1: pHH100-mCherry plasmid (pO Promoter Plasmid) Primers used.

Primers	Sequence
pHH100_mCherry_Frag1_Fw	CTGCAAGCTACCTGCTTTCTCT
pHH100_mCherry_Frag1_Re	GAGCTCATGGCGCGCCTAG
pHH100_mCherry_Frag2_Fw	AACATGTACAATAATAATGGAG
pHH100_mCherry_Frag2_Re	TGGGCTTACATGGCGATAG
pHH100_Frag1_pO_Re	AATTTTTTATCTTTTTGATTTATAAGAGCTCATGGCGCGCCTAG
pHH100_Frag2_pO_Fw	GGCACGCAAATTGTATTAACAGTTCAACATGTACAATAATAATGGAG
pHH100_Frag2_pO_Re	TGGGCTTACATGGCGATAGCTAGACTGG
Gibs_Frag1_Fw_pO	CATGTAAGCCCACTGCAAGCTACCTGCTTTCTC
Gibs_Frag1_Re_pO	AATTTGCGTGCCAATTTTTTATCTTTTTGATTTATAAGAGCTCATGG
Gibs_Frag2_Fw_pO	AAGATAAAAAATTGGCACGCAAATTGTATTAACAGTTC
Gibs_Frag2_Re_pO	AGGTAGCTTGCAGTGGGCTTACATGGCGATAGCTAG

Sigma54 pO Colony_Fw	GAGCGGGTGTTCCTTCTTCAC
Sigma54 pO Colony_Re	CAGCGGACCACCTTTGGTAAC

Table D.2: Lv3Δt – T7 – PspF (1-275) - AmpR plasmid (PspF Transcriptional Activator Protein) Primers used.

Primers	**Sequence**
PspF_Exci_Fw	ATGGCAGAATACAAAGATAATTTAC
PspF (1-275) Re	CTACAGTGTTGGAAGCGAG
T7_PspF_Fw	TAATACGACTCACTATAGGATGGCAGAATAC
pspf_(1-275)_UTR_Fw	ATTGACGGGTCGTAAGGAGGTTTTTAATGGCAGAATACAAA
pspf (1-275)_T7_UTR_Fw	TAATACGACTCACTATAGGATTGACGGGTCGTAAGGAGG
pspf (1-275)_T7_UTR_Gibbson_Fw	CCTTTCTGCCTAATACGACTCACTATAGGATTGACG
Lvt3_PspF(1-275)_AmpR_Gibs_Fw	TCCAACACTGTAGCGAGCTAGAGTCAGCTTTATGCTTG
Lvt3_PspF(1-275)_AmpR_Gibs_Rv	GTGAGTCGTATTAGGCAGAAAGGGCCTCGTGAT
PspF-Gibs-Fw	GCCCTTTCTGCCTAATACGACTCACTATAGGATG

PspF(1-275)-Gibs-Re	GACTCTAGCTCGCTACAGTGTTGGAAGCGAGG
Col_PCR_Fw	GGGCAGGATAGGTGAAGTAGG
Col-PspF-Re	CGCTTCGTGACCAAACAGTTC

Table D.3: pSEVA2311 plasmid (N200 Random Nucleotide Promoter with Mango Aptamer) Primers used.

Primers	Sequence
Gibs_Mango_Fw	CCTTTAATTAAATGTAAATGTTGCCATGTGTATGTG
Gibs_Mango_Re	TCCAAGACTAGTACGCGCTACATCCGCTTTAG
Gibs_Frag1_Fw	GGATGTAGCGCGTACTAGTCTTGGACTCCTGTTG
Gibs_Frag1_Re	GCAACATTTACATTTAATTAAAGGCATCAAATAAAACGAAAG
Golden_gate_Re	GTACGCGTACGGTCTCAGGCATTAATTAAAGGCATCAAATAAAACG
Golden_gate_Fw	GTACGCGTACGGTCTCNATGTAAATGTTGCCATGTGTATG
Col_PCR_Fw	GTCCTACTCAGGAGAGCGTTC
Col_PCR_Re	GCCATGAATGATCCCGAAGG
Prefix_Forward	GAATTCGCGGCCGCTTCTAGAG
Suffix_Reverse	CTGCAGCGGCCGCTACTAGTA

Table D.4: lVT3T_T7_RpoN_AmpR plasmid (Sigma Factor 54 with T7 Promoter system) Primers used.

Primers	Sequence
RpoN_Exc_Fw	ATGAAGCAAGGTTTGCAACTCAG

RpoN_Exc_Re	TCAAACGAGTTGTTTACGCTGG
5UTR_RpoN_Fw	CCGTCAATACTTTGACGCTATCGATCAGCAGATTTAAGGAG GTATTTTATGAAGCAAGGTTTG
5UTR_RpoN_Re	TCAAACGAGTTGTTTACGCTGGTTTGACGGCGGAATG
T7_5UTR_RpoN_Fw	TAATACGACTCACTATAGGCCGTCAATACTTTG
Gibs_T7_5UTR_RpoN_Fw	CCTTTCTGCCTAATACGACTCACTATAGGCC
Gibs_T7_5UTR_RpoN_Re	CTCTAGCTCGTCAAACGAGTTGTTTACGCTG
Gibs_Lvt3_RpoN_Fw	ACTCGTTTGACGAGCTAGAGTCAGCTTTATGCTTG
Gibs_Lvt3_RpoN_Re	AGTCGTATTAGGCAGAAAGGGCCTCGTGATAC
Col_PCR_Fw	GGGCAGGATAGGTGAAGTAGG
Col_RpoN_Re	CAATGTAGTCACCGCTGGTGC

Appendix E: Original Protocols

QIAquick® PCR Purification Kit

1. Add 5 volumes Buffer PB to 1 volume of the PCR reaction and mix.
2. Place a QIAquick column in a provided 2 ml collection tube
3. To bind DNA, apply the sample to the QIAquick column and centrifuge for 30–60 s. Discard flow-through and place the QIAquick column back in the same tube.
4. To wash, add 750 µl Buffer PE to the QIAquick column and centrifuge for 30–60 s. Discard flow-through and place the QIAquick column back into the same tube.
5. Centrifuge the QIAquick column once more in the provided 2 ml collection tube for 1 min to remove residual wash buffer.
6. Place each QIAquick column in a clean 1.5 ml microcentrifuge tube.
7. To elute DNA, add 30 µl Buffer EB (10 mM Tris·Cl, pH 8.5) to the center of the QIAquick membrane and centrifuge the column for 1 min.

E.Z.N.A.®Gel Extraction Kit

1. Perform agarose gel/ethidium bromide electrophoresis to fractionate DNA fragments.
2. Carefully excise the DNA fragment of interest using a wide, clean, sharp scalpel. Minimize the size of the gel slice by removing extra agarose.
3. Assuming a density of 1 g/mL, determine the appropriate volume of the gel slice by weighing it in a clean 1.5 mL microcentrifuge tube.
4. Add 1 volume XP2 Binding Buffer.
5. Incubate at 60°C for 7 minutes or until the gel has completely melted. Vortex or shake the tube every 2-3 minutes.
6. Insert a HiBind® DNA Mini Column in a 2 mL Collection Tube.
7. Add no more than 700 µL DNA/agarose solution from Step 5 to the HiBind® DNA Mini Column.
8. Centrifuge at 10,000 x g for 1 minute at room temperature.
9. Discard the filtrate and reuse collection tube.
10. Repeat Steps 7-9 until all of the sample has been transferred to the column.
11. Add 300 µL XP2 Binding Buffer.
12. Centrifuge at maximum speed (≥13,000 x g) for 1 minute at room temperature.
13. Discard the filtrate and reuse collection tube.

14. Add 700 µL SPW Buffer.
15. Centrifuge at maximum speed for 1 minute at room temperature.
16. Discard the filtrate and reuse collection tube.
17. Centrifuge the empty HiBind® DNA Mini Column for 2 minutes at maximum speed to dry the column matrix.
18. Transfer the HiBind® DNA Mini Column to a clean 1.5 mL microcentrifuge tube.
19. Add 30µL Elution Buffer directly to the center of the column membrane.
20. Let sit at room temperature for 2 minutes.
21. Centrifuge at maximum speed for 1 minute.
22. Store DNA at -20°C.

NEBuilder HiFi DNA Assembly Transformation Protocol

1. Thaw chemically competent cells on ice.
2. Add 2 µl of the chilled assembled product to the competent cells.
3. Place the mixture on ice for 30 minutes.
4. Heat shock at 42°C for 30 seconds.
5. Transfer tubes to ice for 2 minutes.
6. Add 950 µl of room-temperature SOC media to the tube.
7. Incubate the tube at 37°C for 60 minutes. Shake vigorously (250 rpm).
8. Warm selection plates to 37°C.
9. Spread 100 µl of the cells onto the selection plates.
10. Incubate overnight at 37°C.

QIAprep® Spin Miniprep Kit

1. Pellet 1–5 ml bacterial overnight culture by centrifugation at >8000 rpm (6800 x g) for 3 min at room temperature (15–25°C).
2. Resuspend pelleted bacterial cells in 250 µl Buffer P1 and transfer to a microcentrifuge tube.
3. Add 250 µl Buffer P2 and mix.
4. Add 350 µl Buffer N3 and mix.
5. Centrifuge for 10 min at 13,000 rpm (~17,900 x g) in a table-top microcentrifuge.
6. Apply 800 µl supernatant from step 5 to the QIAprep 2.0 spin column. Centrifuge for 30–60s and discard the flow-through.

7. Wash the QIAprep 2.0 spin column by adding 0.5 ml Buffer PB. Centrifuge for 30–60s and discard the flow-through.
8. Wash the QIAprep 2.0 spin column by adding 0.75 ml Buffer PE. Centrifuge for 30–60s and discard the flow-through. Transfer the QIAprep 2.0 spin column to the collection tube.
9. Centrifuge for 1 min to remove residual wash buffer.
10. Place the QIAprep 2.0 column in a clean 1.5 ml microcentrifuge tube. To elute DNA, add 30 µl Buffer EB (10 mM TrisCl, pH 8.5) to the center of the QIAprep 2.0 spin column, let stand for 1 min, and centrifuge for 1 min.

Monarch® PCR & DNA Cleanup Kit (5 µg) (NEB #T1030)

1. Dilute sample with DNA Cleanup Binding Buffer according to the table below.

Table E.1: DNA Cleanup Binding Buffer Dilutions corresponding to the type of Sample

Sample Type	Ratio of Binding Buffer: Sample	Example
dsDNA > 2 kb (plasmids, gDNA)	2:1	200 µl:100 µl
dsDNA < 2 kb (some amplicons, fragments)	5:1	500 µl:100 µl
ssDNA > 200 nt**	7:1	700µl:100 µl

2. Insert column into collection tube and load sample onto column and close the cap. Spin for 1 minute, then discard flow-through.
3. Re-insert column into collection tube. Add 200 µl DNA Wash Buffer and spin for 1 minute.
4. Repeat wash (Step 3).
5. Transfer column to a clean 1.5 ml microfuge tube.
6. Add ≥ 6 µl of DNA Elution Buffer to the center of the matrix. Wait for 1 minute, then spin for 1 minute to elute DNA.

Appendix F: Protein & DNA Molecular Weight Size Markers

The molecular protein weight was estimated using the below molecular weight standard in SDS-PAGE.

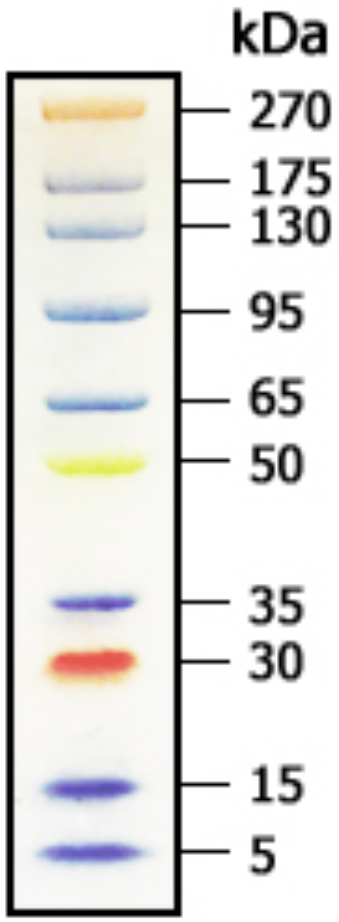

Figure F.1: GenScript Broad Multi Color Pre-Stained Protein Molecular Weight Standard.

The molecular DNA weight was estimated using the below molecular weight standard in Gel electrophoresis.

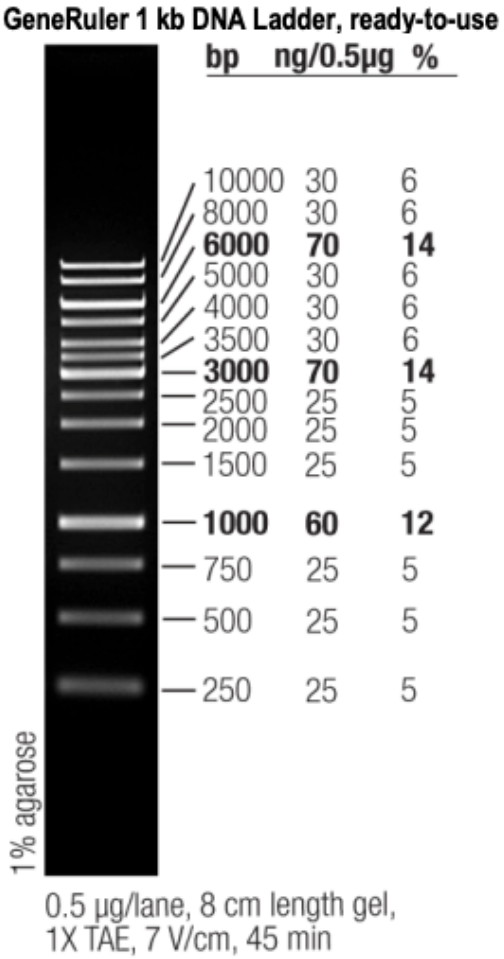

Figure F.2: GeneRuler 1 kb DNA Molecular Weight Standards Ladder.