



UNIVERSIDADE FEDERAL DO AMAZONAS – UFAM
PROGRAMA DE PÓS-GRADUAÇÃO MULTI-INSTITUCIONAL EM
BIOTECNOLOGIA - PPGBIOTEC

PROMOTORES QUE SÃO RECONHECIDOS POR DIFERENTES FATORES
SIGMA EM *Escherichia coli*.

Manaus - AM

2022

JOYCE BELENTANI DE SOUZA MACIEL

**PROMOTORES QUE SÃO RECONHECIDOS POR DIFERENTES FATORES
SIGMA EM *Escherichia coli*.**

Dissertação apresentada ao
Programa de Pós-Graduação em
Biotecnologia da Universidade
Federal do Amazonas como
requisito parcial para obtenção do
título de Mestre em Biotecnologia.

Orientador: Spartaco Astolfi Filho

Manaus - AM

2022

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

M152p	<p>Maciel, Joyce Belentani de Souza Promotores que são reconhecidos por diferentes fatores sigma em Escherichia coli / Joyce Belentani de Souza Maciel . 2022 58 f.: il. color; 31 cm.</p>
	<p>Orientador: Spartaco Astolfi Filho Dissertação (Mestrado em Biotecnologia) - Universidade Federal do Amazonas.</p>
	<p>1. Promotores. 2. Fatores sigma. 3. Proteínas recombinantes. 4. Fase estacionária. I. Astolfi Filho, Spartaco. II. Universidade Federal do Amazonas III. Título</p>

JOYCE BELENTANI DE SOUZA MACIEL

**PROMOTORES QUE SÃO RECONHECIDOS POR DIFERENTES FATORES
SIGMA EM *Escherichia coli*.**

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia da Universidade Federal do Amazonas como requisito parcial para obtenção do título de Mestre em biotecnologia.

Aprovado em: 27/05/2022

Banca examinadora

Spartaco Astolfi Filho

Adolfo José da Mota

Rudi Emerson de Lima Procópio

AGRADECIMENTOS

A Deus pela vida e oportunidade de chegar até aqui, a ele toda glória e louvor. A minha família, meu pai Wagner de Souza, meu avô Levy de Souza e minha irmã Jasminne Belentani, pelo apoio, incentivos e paciência nesse percurso.

Ao meu Orientador e chefe Spartaco Astolfi Filho por todo conhecimento compartilhado para realização deste trabalho e pela importância/influência que exerce na minha vida. Pela humildade, pela paciência, dedicação e pelo amor em me ensinar a Biotecnologia.

Aos meus parceiros Diego Moreira e Whendel Mesquita por todo conhecimento teórico e prático de laboratório, pela paciência, amizade e companheirismo em todos os momentos preciso, em especial a Enedina Nogueira, agradeço toda a ajuda e técnicas que ensinou com dedicação. Ao Joao Katchma e Maria Clara pela idéia e iniciativa do projeto na fase inicial. Agradeço aos amigos pelos força e incentivo nos momentos necessários, aos colegas do programa de biotecnologia. Ao CNPq pelo Fomento do projeto e pela bolsa concedida. E agradeço imensamente também a muitos que ajudaram e contribuíram diretamente e indiretamente.

RESUMO

Introdução:

Estratégias genéticas têm sido utilizadas por décadas, melhorando os bioprocessos bacterianos visando a simplificação e custo da produção de proteínas recombinantes. Tais estratégias incluem o desenho de vetores de expressão eficientes e a melhoria das linhagens bacteriana, neste contexto, os promotores são elementos de um vetor de expressão que determinam força e duração da transcrição, e consequentemente, o rendimento da proteína heteróloga a ser produzida. Os promotores bacterianos são reconhecidos pelos fatores sigmas, que são proteínas que se acoplam a RNA polimerase para realizar a transcrição. Os fatores sigmas se ligam aos promotores que respondem a diversas condições ambientais que a célula enfrenta, dentre eles os mais conhecidos são os fatores sigmas 70 e 38, que reconhecem promotores de fase exponencial e estacionária da célula.

Objetivo:

Design, clonagem e análise de promotores sintéticos reconhecidos pelos fatores sigmas σ^{70} e σ^{38} em *Escherichia coli*.

Material e Métodos:

A partir do conhecimento dos elementos consensuais dos promotores, reconhecidos pelos fatores sigma 70 e 38 de *Escherichia coli*, planejou-se um oligonucleotídeo onde mantinham-se as regiões consensuais mais importantes dos dois tipos de promotores e foram degeneradas as sequências de regiões adjacentes que poderiam influenciar na força e especificidade dos promotores. Um oligonucleotídeo parcialmente degenerado de 95 nucleotídeos foi sintetizado por síntese química, amplificado com um par de *primers* localizados em suas bordas. O amplicon resultante foi digerido com as enzimas de restrição *SpeI* e *NdeI*, purificado por eletroforese em gel de agarose 2%, eluído e clonado entre os sítios de *SpeI* e *NdeI* do plasmídeo pCDM, dessa forma utilizado como um vetor caça promotor que contém o gene da GFP como repórter. As colônias recombinantes foram escolhidas de acordo com a intensidade de produção da GFP, inicialmente a olho nu, posteriormente por fluorimetria para quantificar a produção. Os promotores sintéticos dos clones escolhidos nessa etapa foram sequenciados e analisados.

Resultados e Discussão:

Produziu-se em *E. coli*, uma biblioteca com 1.741.824 possibilidades de sequências construídas por síntese química. A partir de uma clonagem, foram obtidos cerca de 1200 clones recombinantes contendo diferentes promotores, em uma triagem inicial foram escolhidos 66 clones produtores de GFP, e após a confirmação de conterem promotores sintéticos foi feita a análise da expressão de GFP em meio líquido. Foram eleitos 16 clones: 5 que expressam baixos níveis de GFP, 5 níveis médios e 6 altos níveis; sendo alguns deles capazes de atuar na fase exponencial de crescimento, outros na fase estacionária e alguns em ambas as fases. Os novos promotores dos 16 clones escolhidos foram sequenciados e analisados para identificar as regiões consenso -10, -10 estendidas e -35. Analisou-se também as regiões não consensuais que foram degeneradas, bem como a intensidade de expressão de GFP. Alguns promotores foram modificados no espaçamento entre a posição -10 e a -35 e em um dos clones foimodificada a região -35. Os clones J8, J13 e J14 produziram altos níveis de GFP com mais de 900.000 unidades de fluorescência, e os clones J1, J2 e J3 foram os que produziram menos GFP de todos os clones analisados. Por essa estratégia foi possível obter promotores de diferentes forças e que se expressam em diferentes fases do crescimento celular, sendo que alguns apresentam níveis de expressão surpreendentemente altos com possíveis aplicações na expressão gênica para fins industriais.

Apoio: CNPq, FAPEAM e CAPES.

Palavras-chave: promotores; fatores sigma, proteínas recombinantes; fase estacionária

ABSTRACT

Introduction:

Genetic strategies have been used for decades, improving bacterial bioprocesses aimed at simplifying and costing the production of recombinant proteins. Such strategies include the design of efficient expression vectors and the improvement of bacterial strains, in this context, the promoters are elements of an expression vector that determine the strength and duration of the transcription, and consequently, the yield of the hertoiseprotein to be produced. Bacterial promoters are recognized by sigmas factors, which are proteins that engage rna polymerase to perform transcription. Sigma s factors are connected to the promoters that respond to various environmental conditions that the cell faces, among them the best known are the sigmas factors 70 and 38, which recognize exponential and stationary phase promoters of the cell.

Goal:

Design, cloning and analysis of synthetic promoters recognized by sigmas factors $\sigma 70$ and $\sigma 38$ in *Escherichia coli*.

Material and Methods:

From the knowledge of the consensual elements of the promoters, recognized by the sigma factors 70 and 38 of *Escherichia coli*, an oligonucleotide was planned where the most important consensual regions of the two types of promoters were maintained and the sequences of adjacent regions that could influence the strength and specificity of the promoters were degenerated. A partially degenerated oligonucleotide of 95 nucleotides was synthesized by chemical synthesis, amplified with a pair of primers located on its edges. The resulting amplicon was digested with the restriction enzymes *SpeI* and *NdeI*, purified by electrophoresis in 2% agarose gel, eluded and cloned between the *SpeI* and *NdeI* sites of the pCDM plasmid, thus used as a hunting vector that contains the GFP gene as a reporter. The recombinant colonies were chosen according to the production intensity of PFM, initially with the naked eye, later by fluorimetry to quantify the production. The synthetic promoters of the clones chosen at this stage were sequenced and analyzed.

Results and Discussion:

A library with 1,741,824 sequence possibilities constructed by chemical synthesis was produced in *E. coli*. From a cloning, about 1200 recombinant clones containing different promoters were obtained, at an initial screening 66 gfp-producing clones were chosen, and after confirmation of containing synthetic promoters, the expression of PFM was performed in liquid medium. 16 clones were elected: 5 that express low levels of PFM, 5 medium levels and 6 high levels; some of them are capable of acting in the exponential phase of growth, others in the stationary phase and some in both phases. The new promoters of the 16 clones chosen were sequenced and analyzed to identify the consensus regions – 10, -10 extended and -35. We also analyzed the non-consensual regions that were degenerated, as well as the intensity of PFM expression. Some promoters were modified in the spacing between position -10 and -35 and in one of the clones the region -35 was modified. Clones J8, J13 and J14 produced high levels of PFM with more than 900,000 fluorescence units, and clones J1, J2 and J3 produced less PFM of all clones analyzed. Through this strategy it was possible to obtain promoters of different forces and express themselves in different phases of cell growth, and some have surprisingly high levels of expression with possible applications in gene expression for industrial purposes.

Support: CNPq, FAPEAM and CAPES.

Keywords: promoters; sigma factors, recombinant proteins; stationary phase

LISTA DE TABELAS

Tabela 1: Fatores sigmas σ de <i>Escherichia coli</i>	19
Tabela 2: Regiões: consenso -35, espaçador e -10 dos fatores sigmas σ de <i>Escherichia coli</i>	25
Tabela 3: Meio de cultura Luria-Bertani	32
Tabela 4: Sequências dos oligos utilizados neste trabalho.	35
Tabela 5: Biblioteca de promotores reconhecidos por σ^{38}	40
Tabela 6: Sequência de nucleotídeos dos promotores J1 ao J16.....	49

LISTA DE FIGURAS

Figura 1: Imagem ilustrativa do reconhecimento e ancoragem da região 2 do fator σ com o elemento -10.....	15
Figura 2: Ilustração do ciclo de transcrição da RNA-polimerase bacteriana.....	17
Figura 3: Esquema do complexo enzimáticos realizando a transcrição.	18
Figura 4: Ilustração da interação do complexo enzimático holoenzima RNA polimerase com o DNA.....	21
Figura 5: ilustração dos domínios do fator σ^{70} e suas funções.....	22
Figura 6: Mapa plasmidial do vetor pCDM criado com o software SnapGene. 33	
Figura 7: desenho experimental de medidas de OD e GFP.....	38
Figura 8:Matriz de peso-posição plotada em gráfico pelo software MEME-Suite.	41
Figura 9: Perfil eletroforético do produto digestão enzimática dos promotores e dos plasmídeos	43
Figura 12:Fotografia das placas contendo os clones selecionados para analisar a fluorescência e confirmar se havia a presença do promotor.	45
Figura 13:gráficos de fluorescência em função dos clones nos tempos 16 e 24 horas.	50

LISTAS DE ABREVIACOES E SMBOLOS

α'	Alfa linha
α''	Alfa duas linhas
β	Beta
β'	Beta linha
ω	mega
σ	Sigma
σ^{24}	Sigma pertencente à famlia vinte e quatro
σ^{32}	Sigma pertencente à famlia trinte e dois
σ^{38}	Sigma pertencente à famlia trinta e oito
σ^{54}	Sigma pertencente à famlia cinquenta e quatro
σ^{70}	Sigma pertencente à famlia setenta
C-terminal	Carboxi-terminal
DNA	cido desoxirribonuclico
et al.	E colaboradores
GFP	Protena verde fluorescente
kDa	Quilodanton
mL	Mililitro
N-terminal	Amino-terminal
PCR	Reao em cadeia da polimerase
RNA	cido ribonucleico
rpm	Rotaes por minuto
μg	Micrograma
μL	Microlitro

SUMÁRIO

1. INTRODUÇÃO	12
2. REVISÃO BIBLIOGRÁFICA	14
2.1 Transcrição em procariotos.....	14
2.2 RNA polimerase em procariotos.....	16
2.3 Fatores sigmas.....	19
2.3.1 Fator sigma 70.....	21
2.3.2 Fator sigma 38.....	22
2.4 Promotores	23
2.4.1 Elementos -10 e -35	24
2.4.2 Elementos adicionais do promotor.....	25
2.5 Vetores de clonagem.....	27
2.5.1 Vetores	27
2.5.2 Plasmídeo caça-promotor.....	28
2.6 Expressão de proteínas heteróloga em <i>E. coli</i>	28
2.7 Construção de bibliotecas de promotores para produção de proteínas heterólogas	29
3. OBJETIVOS	31
3.1 Geral.....	31
3.2 Específicos	31
4. MATERIAIS E MÉTODOS	32
4.1 Desenho experimental.....	32
4.2 Hospedeira Bacteriana	32
4.3 Vetor caça promotor pCDM	33
4.4 Design dos oligonucleotídeos degenerados	33
4.4.1 Ferramentas de bioinformática para identificação das sequências consenso fatores sigmas σ^{70} e σ^{38}	33
4.4.2 Design da sequência promotora	34
4.5 Procedimentos de Clonagem Molecular e de Transformação Genética	34
4.5.1 Promotor	34
4.5.2 Preparo de células competentes	35
4.5.3 Vetor caça promotor pCDM	36
4.5.4 Digestão e purificação do vetor pCDM e dos diferentes promotores	36
4.5.5 Transformação Genética	37
4.6 Quantificação da expressão de GFP.....	37
4.7 Estatística	39
4.8 Sequenciamento	39
6 RESULTADOS E DISCUSSÃO	40
6.1 Design do oligonucleotídeo degenerado	40

6.2	Clonagem molecular dos promotores	42
6.3	Expressão da Proteína Verde Fluorescente - GFP	43
6.4	Sequenciamento	48
7	CONCLUSÃO	52
8	REFERENCIAS BIBLIOGRAFICAS	53

1. INTRODUÇÃO

Desde a descoberta da estrutura da molécula de DNA no início da década de 50, os progressos em biologia celular e molecular têm sido extraordinários. Vinte anos depois, descobertas provenientes de diferentes áreas (microbiologia, bioquímica, imunologia, genética microbiana e biologia molecular) apresentaram uma nova perspectiva para o estudo do DNA. O conjunto de técnicas denominado de Tecnologia do DNA Recombinante - TDR, possibilitou a manipulação do genoma, proporcionando isolar, caracterizar e clonar genes extraídos dos mais diversos organismos (ASTOLFI-FILHO et al, 1996).

O desenvolvimento dessa tecnologia foi possível graças à descoberta de enzimas chamadas endonucleases de restrição (capaz de reconhecer e clivar sequências específicas do DNA) e DNA ligase (capaz de unir dois pontos do DNA). Deste modo, foi possível isolar os fragmentos de ácidos nucleicos e inseri-los em vetores de clonagem, que em geral são plasmídeos ou genomas virais, passíveis de serem introduzidos no genoma de uma célula hospedeira compatível. Quando a célula mãe se divide em células filhas, duplica e expressa o fragmento de DNA inserido para produzir uma proteína recombinante (COHEN, 2013).

A TDR pode ser utilizada para diversos fins como: o estudo da evolução gênica, da expressão e regulação dos genes de organismos complexos (genes defeituosos, genes que codificam anticorpos, genes relacionados a apoptose, entre outros), produção de medicamentos (vacinas, antivirais, antibióticos, antifúngicos, hormônios, imunomoduladores, anticorpos, entre outros) e de enzimas de interesse industrial (amilases, lipases, celulases, proteases) produzidas em grande escala (KAMIONKA, 2011; PEREIRA, 2019; SANTOS-BENEIT e MOLLINEDO, 2000; WANG, 2016;).

Para obter uma produção de proteínas recombinantes em grandes quantidades, é necessário um sistema eficiente composto de célula hospedeira (*Escherichia coli* a célula procariótica mais utilizada) e de um vetor de expressão. Os vetores de clonagem compartilham de características essenciais: (1) uma região promotora para conduzir e regular a expressão do gene de interesse; (2) uma região de múltiplo sítios de clonagem para a inserção de fragmentos de DNA desejado; (3) um terminador de transcrição; (4) uma origem de replicação (ORI) o que lhes confere a capacidade de replicar independentemente do genoma hospedeiro; (5) possuem um marcador selecionável que normalmente é um gene de resistência a antibióticos.

A sequência promotora é o principal elemento que controla e regula expressão gênica. Os promotores são sequências de DNA localizados a montante do início da transcrição do gene.

Esta sequência dispõe de regiões específicas que atuam como sítios de reconhecimento para o acoplamento da RNA-polimerase (complexo enzimático responsável pela transcrição do DNA em RNA), (BROWNING e BUSBY, 2004). Em bactérias, o complexo enzimático da RNA polimerase (núcleo ou apoenzima) é constituído de cinco subunidades (α' , α'' , β , β' , ω) que é suficiente para alongamento e terminação da transcrição, mas é incapaz de reconhecer a sequência promotora (LEE e BURUKHOV, 2016; MURAKAMI, 2015).

Para o reconhecimento da região promotora é necessária a sexta subunidade denominada de fator sigma (σ), que se associa com o núcleo da RNA polimerase para formar a holoenzima. Os fatores σ foram classificados de acordo com a massa molecular: σ^{24} , σ^{28} , σ^{32} , σ^{38} , σ^{54} e σ^{70} , e a cada um deles foi designado um papel funcional no reconhecimento da sequência consenso dos distintos promotores, atuando como reguladores na expressão gênica em resposta à estímulos ambientais e condições desfavorável enfrentado pela célula.

A escolha cuidadosa da força dos promotores para produção de proteína heterólogas abundantemente é necessário para cada estratégia. Por esse motivo, este trabalho tem como objetivo fazer o design, clonagem e análises de promotores sintéticos reconhecidos pelos fatores sigmas 70 e 38 em *E. coli*, proporcionando promotores de forças variadas para diversas estratégias biotecnológicas, promotores que expressam níveis elevados de produto gênica nas fases distintas do ciclo celular (exponencial e estacionária), sem necessitar de indutores.

2. REVISÃO BIBLIOGRÁFICA

2.1 Transcrição em procariotos

O DNA é a molécula responsável por controlar as atividades celulares. As informações genéticas contidas nas sequências de DNA servem como uma “receita” coordenada para ajustar o bom funcionamento da célula, essa mensagem que os genes carregam precisam ser decifradas e traduzidas em proteínas, que é a molécula que atua nas reações metabólicas e estruturais da célula.

Para produzir uma proteína, a sequência de nucleotídeo contido no DNA é transcrita para RNA, moléculas estas que são constituídas por uma sequência de ribonucleotídeos formando uma fita simples que orientarão a síntese de proteínas. Todas as células expressam sua informação genética dessa maneira, um princípio fundamental que foi denominado dogma central da biologia molecular. Indiretamente o DNA controla as atividades da célula por meio da síntese de RNA, processo esse, conhecido por **transcrição** (MURAKAMI, 2015).

A transcrição utiliza um fragmento do DNA (gene) como molde, iniciando com a abertura e o desenrolar de uma pequena parte da fita de DNA, expondo as bases, que são utilizadas no pareamento de bases complementares dos ribonucleotídeos a serem incorporados. Esse processo é realizado por um complexo de subunidades enzimáticas chamada **RNA-polimerase** e geram três tipos principais de RNA (MURAKAMI, 2015; ZANG et al, 2012)

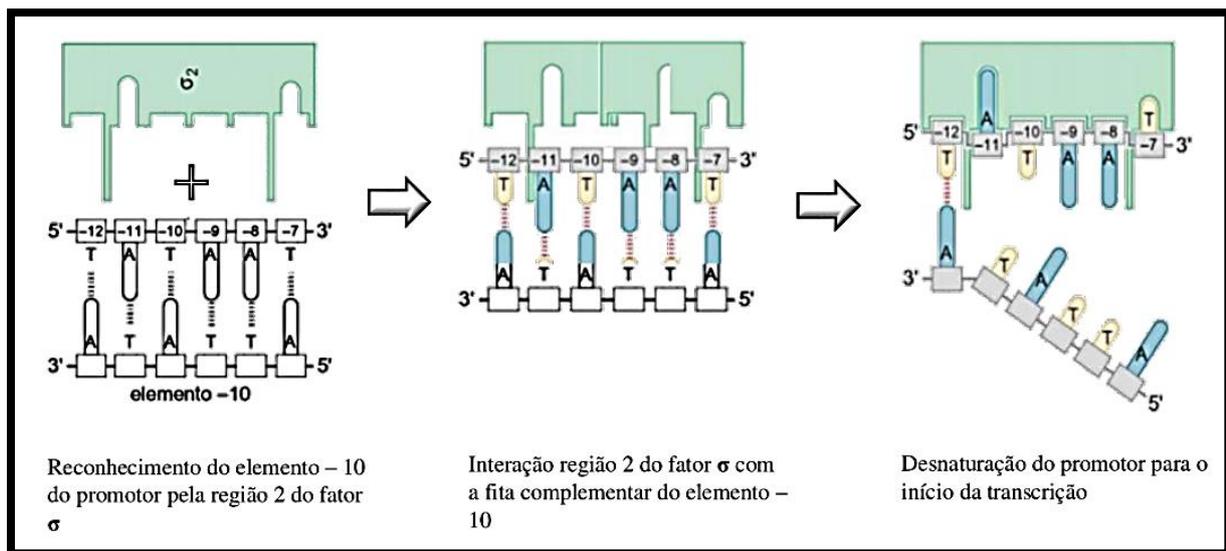
A transcrição de um gene, a RNA-polimerase realiza diversas tarefas que foram definidas em três etapas: **Início, alongamento e término**. A **iniciação** é uma etapa importante na expressão gênica, pois é nesse ponto que a frequência da transcrição é regulada. Nesse momento o núcleo da RNA-polimerase (apoenzima) se associa a uma subunidade adicional denominada fator sigma (σ), para formar a holoenzima RNA-polimerase e auxiliar no reconhecimento da sequência do DNA que indica o início da transcrição (LEE e BORUKHOV, 2016).

Quando ocorre a ligação entre holoenzima RNA polimerase e o promotor (**formando o complexo fechado**), a ligação inicial não altera a conformação da dupla-hélice do DNA (fig. 1), (MURAKAMI, 2002). Quando fator sigma faz os contatos específicos com as regiões do promotor, o DNA é desenrolado e um trecho de nucleotídeos é exposto (cerca de 15 nucleotídeos) formando a bolha de transcrição (**complexo aberto**).

Durante o complexo aberto, uma das fitas de DNA (5' → 3') é utilizada como molde para o pareamento de bases complementares com os ribonucleotídeos. No modelo do complexo aberto, o DNA se curva em torno da RNA polimerase formando dobras nos nucleotídeos -35 e -25 (ZANG et al, 2012).

A interação com o domínio 2 do fator σ ocorre no elemento -10, nos nucleotídeos -11 a -7 (fita não modelo do DNA), a bolha de transcrição inicia no nucleotídeo -11, então a abertura DNA se estende a jusante, passando o local de início da transcrição (+1) para completar a bolha da transcrição. O DNA modelo de fita simples é direcionada pelos domínios 2.4 e 3.0 do fator σ , para o sitio ativo da RNA polimerase através de um processo eletrostático para início da transcrição (MURAKAMI, 2015).

Figura 1: Imagem ilustrativa do reconhecimento e ancoragem da região 2 do fator σ com o elemento -10. Reconhecimento feito para abertura da bolha de transcrição (Fonte: Adaptado de Watson, 2015)



Os dez (ou mais) primeiros ribonucleotídeos incorporados são reconhecidos como um processo ineficiente, então a RNA polimerase se desprende da região promotora e do fator σ , liberando um transcrito em RNA curto e recomeça síntese (**Complexo de transcrição inicial**), (MURAKAMI, 2002). Em algumas ocasiões, o fator σ não é desagregado totalmente, porém as ligações feitas pelo fator σ com a RNA polimerase não atrapalha o acesso dos ácidos nucleicos. Comumente, o promotor permanece em um complexo de transcrição inicial que pode sofrer paradas abortivas e com isso não iniciar a transcrição (ZANG et al, 2012).

Quando de fato a RNA polimerase se dissocia do promotor, a transcrição é iniciada (**alongamento**). A RNA polimerase desempenha múltiplas tarefas, ao locomover-se sobre o DNA ela atua desenrolando a dupla-hélice a sua frente, ao mesmo tempo incorpora os ribonucleotídeos e dissocia a síntese de RNA. Na medida que avança os pares de base, ela recompõe a fita de DNA, pareamento corretamente e finaliza revisando a leitura. O alongamento da fita de RNA é realizado a uma velocidade de 50 nucleotídeos por segundo (MURAKAMI, 2015; ZANG et al, 2012).

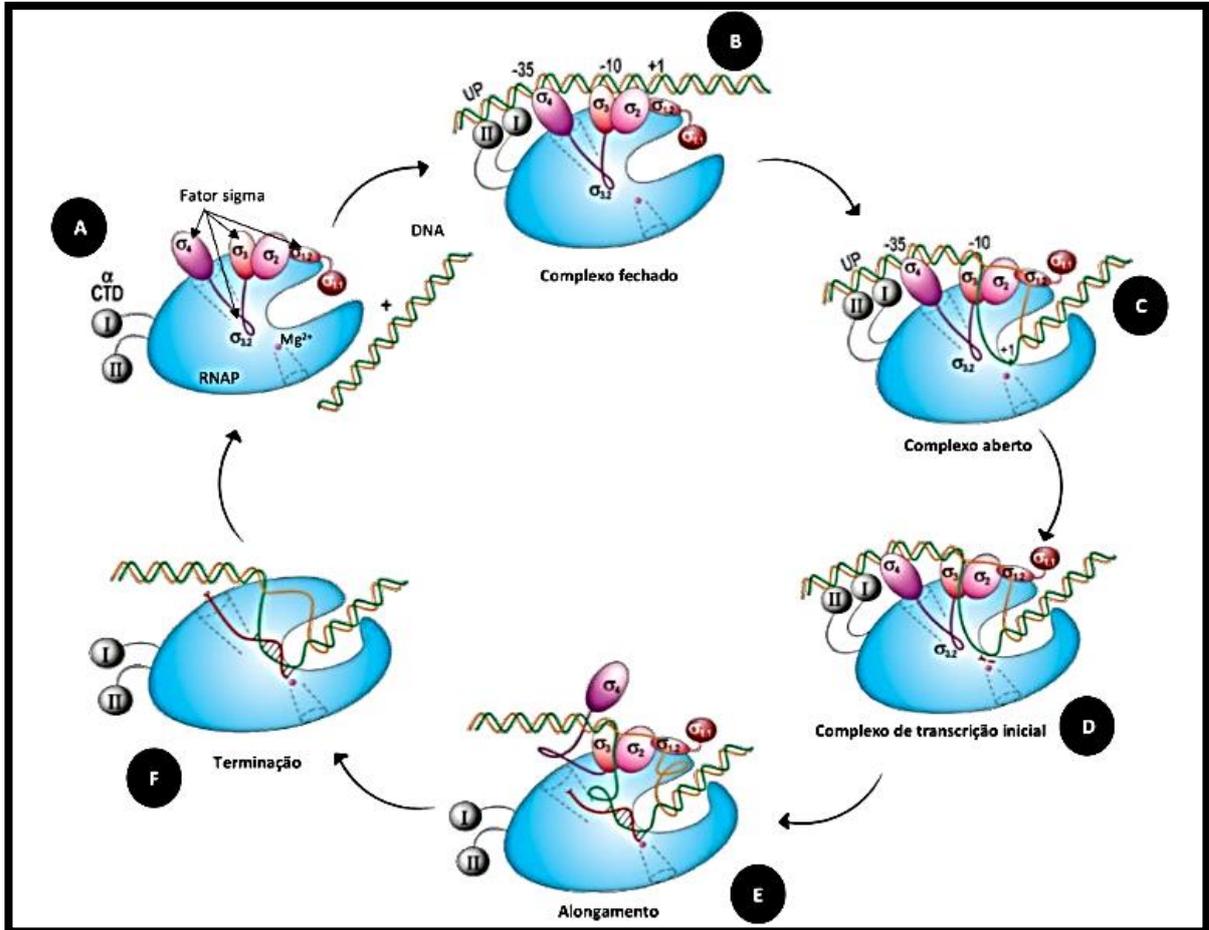
Para finalizar o alongamento, a RNA polimerase encontra uma sequência específica de DNA que consiste em pares de bases nucleotídicas A-T para sinalizar o **término da transcrição** (fig. 2). Para formar o grampo que ajuda no processo de desligamento da RNA com o sítio ativo, o sinal de término se apresenta de forma simétrica. Quando a RNA-polimerase se dissocia do RNA produzido e da fita molde de DNA, ela fica disponível para interagir com outro fator σ para dar início ao processo de transcrição novamente (LEE e BORUKHOV, 2016).

2.2 RNA polimerase em procariotos

A transcrição é o primeiro passo na expressão genética, sendo de extrema importância que essa etapa seja regulada. O complexo enzimático responsável por tal tarefa é a RNA polimerase, e ela é bem conservada entre os organismos vivos e as subunidades que compõem são muito semelhantes entre espécies, pois realizam a mesma função (BROWNING e BUSBY, 2004; EBRIGHT et al, 2000; LEE e BURUKHOV, 2016).

Em bactérias a RNA polimerase é composta por 5 subunidades, sendo elas: alfa (α' com $\cong 35$ kDa); ($\alpha'' \cong 10-35$ kDa), beta ($\beta \cong 150$ kDa), beta linha ($\beta' \cong 160$ kDa) e ômega ($\omega \cong 6$ kDa), juntas essas subunidades formam a apoenzima ou núcleo da RNA polimerase, possuindo um formato que se assemelha a uma garra de caranguejo. Essas garras são representadas principalmente pelas subunidades β e β' , entres elas se apresentam fissuras que acomodam os íons metálicos (Mg^{2+}), o DNA, o RNA sintetizado e os ribonucleotídeos a serem incorporados (EBRIGHT, 2000).

Figura 2: Ilustração do ciclo de transcrição da RNA-polimerase bacteriana. (A) a holoenzima RNA-polimerase é formada para localizar sequência de DNA do promotor; (B) A polimerase reconhece as regiões essenciais da sequência promotora, formando o complexo **fechado**; (C) a RNA polimerase expõe a fita molde no centro ativo da enzima formando o **complexo aberto**; (D) Com o DNA na posição em que a **transcrição inicial começa**; (E) Essa síntese de RNA inicial (iniciação abortiva) quando o transcrito curto é liberado a RNA polimerase se dissocia da região promotora, libera o fator σ e começa a síntese novamente (**alongamento**); e (F) o **término** ocorre quando a polimerase encontra o sinal de terminação e se dissocia do DNA-molde e libera o RNA recentemente transcrito (Fonte: adaptado de LEE e BORUKHOV, 2016).

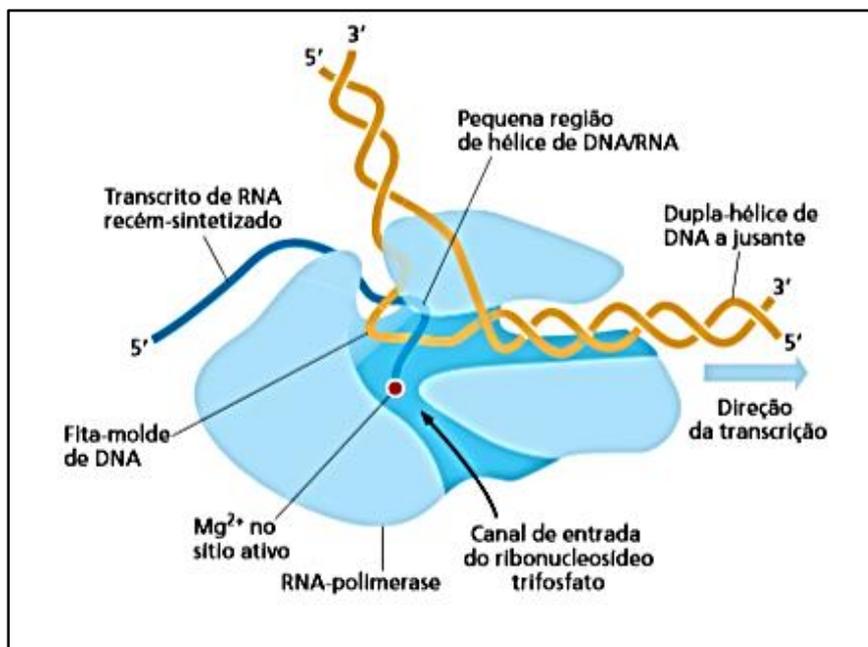


As subunidades α' e α'' situam-se próximo a fenda, apesar das sequências idênticas a localização dentro da RNA polimerase diferem: α' encontra-se próxima e α'' um pouco mais longe da fenda principal, e ambas interagem com as subunidades β e β' , respectivamente (LEE e BURUKHOV, 2016). As subunidades α são constituídas de dois domínios denominadas de N-terminal (α NTD) e C-terminal (α CTD) que estão ligados através de uma região de $\cong 20$ -aminoácidos (*linker*). Os domínios α NTD são responsáveis pela interação com as subunidades β e β' , e os domínios α CTD são responsáveis pela interação com uma região sequência promotora denominada de elemento UP, além de interagir com proteínas ativadoras e repressoras (EBRIGHT et al, 2000; MURAKAMI, 2015).

A subunidade ω possui um papel fundamental na síntese do RNA atuando como um chaperona molecular, localizado próximo ao sulco formado pelas subunidades β e β' na saída do RNA sintetizado para executar o enovelamento correto. (BROWNING e BUSBY, 2004). A subunidade maior β' consiste de uma massa central e dois lobos menores, e a subunidade β de duas massas centrais idênticas, a conexão entre essas enzimas formam: a fissura grande chamada de “canal principal” ou canal do centro ativo (onde ocorre a interação DNA-RNA polimerase e a incorporação dos ribonucleotídeos); o “canal secundário” por onde são introduzido os ribonucleotídeos; e o “canal terciário” ou canal de saída de RNA que é responsável por separar as moléculas de DNA-RNA e por interagir com os grampos finais no término da síntese (fig. 3), (LEE e BORUKHOV, 2016).

A conexão entre os canais é onde se acopla os íons metálicos (Mg^{2+}) formando o centro ativo. As subunidades β e β' juntas fazem a interação com o fator sigma σ e incorporam ribonucleotídeos. Quando se associa com o fator σ a RNA polimerase forma a holoenzima, que é capaz de reconhecer a sequência de DNA (promotor) que sinaliza o início da transcrição (MOREIRA et al, 2018; MURAKAMI, 2015).

Figura 3: **Esquema do complexo enzimáticos realizando a transcrição.** A RNA-polimerase está representada em azul-claro, atuando no desenrolar da molécula DNA representado em laranja, no seu sítio ativo indicado pelo Mg^{2+} na cor vermelha. Conforme avança, a polimerase adiciona os ribonucleotídeos no sítio de polimerização, usando uma fita de DNA exposta como molde e sintetizando o RNA mensageiro em azul escuro (Fonte: ALBERTS et al, 2017).



2.3 Fatores sigmas

O fator σ é componente essencial da RNA polimerase para a seletividade de um promotor. Os diferentes fatores σ presente nas células reconhecem as regiões promotoras de genes que são expressos em diferentes estados ou fases do metabolismo celular, e a cada um dos fatores foi designado um papel fundamental no reconhecimento da sequência consenso dos distintos promotores, atuando como reguladores na expressão gênica (BROWNING e BUSBY, 2004).

Em *E. coli* os fatores σ foram divididos em duas famílias de proteínas: σ^{70} e σ^{54} . A maioria dos fatores apresentam similaridade nos domínios pertencente à família σ^{70} . Os σ pertencente a família σ^{54} compreende proteínas semelhantes no funcionamento e distintos na estrutura, sua função baseia-se em assimilar o nitrogênio na célula (tabela 1), (CASES e LOREZO, 2005).

Tabela 1: Fatores sigmas σ de *Escherichia coli*

Fator σ	Gene	Função/Resposta
24	rpoE	Estresse térmico extra-citoplasmático
28	FliA ou rpoF	Mobilidade celular e patogenicidade ou síntese de flagelo
32	rpoH	Choque térmico extremo; renaturação de proteínas
38	rpoS	Resposta ao estresse e fase estacionária
54	rpoN	Assimilação ou limitação de nitrogênio (N ₂)
70	rpoD	Manutenção basal; crescimento celular

Fonte: Moreira et al, 2018

A família de proteínas σ^{70} foi dividida em quatro grupos filogenéticos embasados na estrutura e função. O grupo 1 possui os fatores sigmas que reconhecem os promotores de genes constitutivos, essenciais para o crescimento e manutenção da célula, o grupo 2 consiste em

reconhecimento de genes utilizados em fase de estresse ou fase estacionária, o grupo 3 responde a sinais específicos na esporulação, biossíntese de flagelos ou choque térmico e o grupo 4 reconhece genes com funções extra citoplasmáticas (ECF), (ÖSTERBERG E SHINGLER, 2011)

Fatores σ competem pelo sítio ativo formado por β e β' no núcleo da RNA polimerase. Após o σ se ligar na superfície da fissura principal, ele torna-se capaz de se ligar aos elementos do promotor (MURAKAMI, 2015). Alguns domínios participam da ligação com núcleo da RNA polimerase, porém esses contatos são relativamente fracos permitindo a competição pelo núcleo entre diferentes σ (VASSYLYEV et al, 2002).

Os membros da família σ^{70} possuem quatro regiões bem conservadas: σ^1 , σ^2 , σ^3 e σ^4 (unidos por *linkers*), mas apenas os domínios σ^2 e σ^4 são bem conservados em todos os membros, pois estão envolvidas no reconhecimento das regiões promotoras denominadas -10 e -35, respectivamente (HELMANN, 2001).

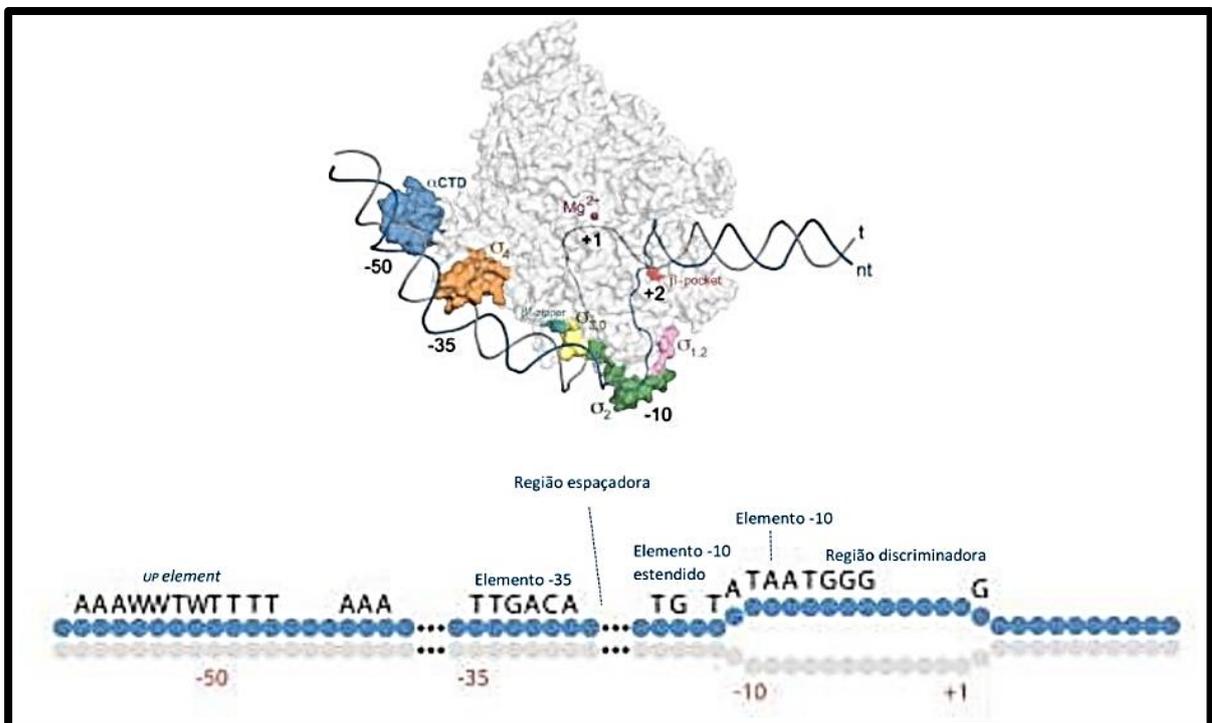
No domínio σ^1 é conservado entre os grupos 1 e 2, localizado próximo ao início da transcrição ele atua antagonicamente à ligação, impedindo o fator sigma de se ligar ao DNA quando estiver em solução sem a presença da RNA polimerase. Ao fazer a ligação com a RNA polimerase, o fator σ muda sua conformação para expor as superfícies de interação com o DNA. Esse domínio também muda sua conformação durante o processo de início da transcrição, mas não é essencial para os membros da família σ^{70} (FEKLISTOV et al, 2013; HELMANN, 2001).

Já o domínio σ^2 é bem conservado entre os sigmas, a região $\sigma^{2.1}$ atua na conexão do σ com o núcleo da RNA polimerase, o $\sigma^{2.3}$ participa da abertura da dupla fita de DNA e o $\sigma^{2.4}$ é responsável por reconhecer o elemento -10 da região promotora. O domínio σ^3 está dividido em duas regiões, o fator σ^3 é responsável por reconhecer o elemento -10 estendido, e a $\sigma^{3.2}$ está envolvida com o núcleo da RNA polimerase. O σ^4 participa do reconhecimento do elemento -35 e consiste em duas regiões, $\sigma^{4.1}$ e $\sigma^{4.2}$ (HELMANN, 2001; LONETTO et al, 2012).

O σ^3 e o σ^4 atravessam o canal principal ligado por um *linker*, localizando-se na saída do núcleo da RNA polimerase formando um *loop* desordenado no centro ativo para estabilizar o início da transcrição, e quando o RNA abortivo é sintetizado, o *linker* se desloca enfraquecendo a interação do fator σ com o núcleo da RNA polimerase, liberando o fator σ para continuar a transcrição (Fig. 4), (FEKLISTOV et al, 2013; HELMANN, 2001).

Os fatores σ são importantes para a regulação dos distintos genes dentro do genoma bacteriano, o reconhecimento das sequências promotoras é etapa mais importante para a regulação gênica, tarefas essas que foram atribuídas aos diferentes fatores σ . E cada domínio da sua estrutura tem uma função importante na atuação do reconhecimento do promotor e produção das proteínas necessárias.

Figura 4: **Ilustração da interação do complexo enzimático holoenzima RNA polimerase com o DNA.** A RNA polimerase núcleo está representada em branco, e os domínios do fator σ em colorido. O DNA está representado em fios: azul (fita não modelo) e cinza (fita molde). A sequência de nucleotídeos destacada representa o promotor com as regiões que são reconhecidas pelos domínios do fator σ (Fonte: Adaptado de FEKLISTOV, 2013).



2.3.1 Fator sigma 70

O σ^{70} faz parte do grupo 1 da família de proteínas σ^{70} , esses fatores transcrevem a maioria dos genes necessários para crescimento e manutenção celular (Fig 5), são codificados a partir dos genes RpoD. Estão presentes em maior quantidade na célula e possuem de 40 a 70 kDa de massa molecular, os domínios do σ que estão envolvidos no reconhecimento do promotor são altamente conservados (LONETTO et al, 2012).

Os fatores σ interagem com os elementos do promotor, sendo os hexâmeros -35 e -10 os mais conservados 5'-TTGaca-3' (-35) e 5'-TATAaT-3' (-10) (onde a maiúscula se refere a bases mais altamente conservadas), além dos elementos adicionais (ÖSTERBERG et al, 2011).

Por esse motivo, muitos promotores transcritos pela RNA polimerase associada a um fator σ^{70} também são reconhecidos pelo σ^{38} , dificultado o reconhecimento de promotores que permitem o reconhecimento seletivo da σ^{38} . As regiões consenso do promotor reconhecidos pelo σ^{70} são parcialmente iguais (com exceção de alguns nucleotídeos). Os σ^{38} são capazes de transcrever um promotor de forma eficiente sem a presença do elemento -35 consenso ou com a troca de base nucleotídicas T por uma C a jusante do elemento -10, mudanças essas, passíveis da redução do reconhecimento do σ^{70} .

Os fatores σ^{38} são expressos durante as condições de estresse (falta de nutrientes, pH, temperatura e molaridade não ótima, fase estacionária), e se acumula na célula à medida que cessa o crescimento celular. Então para capturar a RNA polimerase e expressar os genes necessários nessa fase celular, o σ^{38} utiliza de mecanismo específico pois sua afinidade com o núcleo da RNA polimerase é menor comparado com o fator σ^{70} , para se associar ao núcleo da RNA polimerase é produzido a proteína RSD (reguladora do sigma D), que se liga aos σ^{70} livres impedindo-os de se acoplar ao núcleo da RNA polimerase, permitindo então, que fatores alternativos se liguem (CASES e LOREZO, 2005).

Ao entrar na fase estacionária a produção de fatores σ^{38} começam a aumentar, porém, os fatores σ^{70} não diminuem. Então, uma estratégia adotada pela célula e produzir a proteína RSD atua no controle da σ^{70} , ela começa a ser transcrita na transição da fase exponencial para a estacionária, para o desligamento de genes reconhecidos pelo σ^{70} no início da fase estacionária para controlar o metabolismo da célula (PATIKOGLU et al, 2007; MITCHELL et al, 2007).

2.4 Promotores

A holoenzima RNA polimerase é capaz de reconhecer uma sequencias de promotores dentro do genoma, ao deslizar sobre o DNA ela reconhece uma sequência especial de nucleotídeos que indica o ponto de início para a síntese de RNA que é denominada de promotor. Os elementos do promotor definem a força de interação que a RNA polimerase irá realizar a transcrição.

A primeira base a ser transcrita é denominada como nucleotídeo +1, então o **promotor** vem a ser a sequência de DNA localizada a montante do início da transcrição, possuindo cerca de 90 pares bases. As diferenças entre essas sequências criam assinaturas conformacionais

únicas que determinam a força do promotor, definindo a quantidade de vezes que aquele gene será expresso, por exemplo, proteínas mais abundantes na célula tendem a ter promotores mais fortes do que proteínas raras (MURAKAMI, 2015; ÖSTERBERG et al, 2011).

Para alguns promotores as mudanças nas concentrações de temperatura, sal e soluto, bem como adições de fatores proteicos e ligantes, podem afetar a ligação com o fator σ . (FREKLISTOV, 2013). A modulação pela ação de proteínas ativadoras e repressoras atuam alterando a conformação do DNA promotor deixando atrativo ou obstruindo as regiões promotoras da RNA polimerase durante a fase de recrutamento (BROWNING e BUSBY, 2004)

Os promotores mostram variados níveis de diversidade entre eles, porém, eles possuem sequências consenso que são reconhecidas pelos fatores σ . Uma sequência consenso de nucleotídeos é derivada da comparação de várias sequências que se encontram alinhados numa determinada posição.

As sequências promotoras são assimétricas, isso significa que a RNA polimerase pode ligar-se em apenas uma orientação, sendo que a síntese de RNA ocorre apenas na direção 5' → 3', a orientação do promotor determina a fita que será utilizada como molde. A fita de DNA utilizada como molde para a síntese de RNA varia de gene para gene.

As sequências consenso foram chamadas de -35 e -10, estão situadas a 35 e a 10 pares de bases antecedentes a primeira base transcrita. Além da região consenso, alguns promotores contêm elementos adicionais para aumentar sua força, são eles: o elemento -10 estendido, a região discriminadora e a região *UP element* ou elemento UP (ZHANG et al, 2012).

2.4.1 Elementos -10 e -35

O **elemento -10** está localizado a montante do nucleotídeo +1 nas bases -7 a -12 com a sequência consenso 5'-TATAAT-3' que é reconhecida pelo domínio 2.4 do fator σ (tabela 2), promotores que contêm apenas os elementos -10 e o discriminador se mostraram capazes de realizar transcrição (BROWNING e BUSBY, 2004).

A RNA polimerase distorce a dupla fita de DNA na medida que procura um promotor e realiza sua conexão inicial no elemento -10, onde a um envolvimento com um intermediário da dupla fita estando desenovelada para facilitar a leitura das bases (FEKLISTOV et al, 2013). O elemento -10 direciona os passos da vinculação da RNA polimerase e as alterações

conformacionais na dupla fita de DNA, para que ocorra a bolha de transcrição em aproximadamente 15 pares de base (LEE e BURUKHOV, 2016).

O elemento **-35** está localizado a montante do -10, com a sequência consenso 5'-TTGACA-3' sendo o -35T, -34T e -33G os mais conservados. Ao interagir com o domínio 4 do fator σ o elemento -35 sofre uma curvatura de aproximadamente 36°, o elemento UP também interage com essa curvatura. A rotação do domínio 4 do σ é necessária para acomodar diferentes comprimentos da região espaçadora.

Tabela 2: Mostra as regiões: consenso -35, espaçador e -10 dos fatores sigmas σ de *E. coli*

Fator σ	Sequência consenso -35	Espaçador	Sequência consenso -10
24	GGAACTT	15 pares de base	GTCTAA
28	CTAAA	15 pares de base	GCCGATAA
32	CCCTTGAA	13 -15 pares de base	CCCGATNT
38	TTGACA	16-18 pares de base	TATACT
54	CTGGNA	6 pares de base	TTGCA
70	TTGACA	17 pares de base	TATAAT

Fonte: Moreira et al, 2018

2.4.2 Elementos adicionais do promotor

O elemento discriminador

Está localizado entre o início da transcrição e o elemento -10, possui de 6 a 8 pares de base sendo os nucleotídeos -6G, -5G e -4G consenso entre os promotores. Essa sequência é contatada pelo domínio 1.2 do σ , que interage na abertura da bolha de transcrição, a bolha de transcrição aumenta cerca de 13 pares de base em direção a jusante da base -11, essa sequência facilita a abertura e estabiliza o complexo aberto devido a presença das bases púricas (HAUGEN et al, 2008).

Quando há a presença de guanina a formação da bolha se mantém ordenada e empilhada, quando há a presença da citosina a composição aromática interage com o domínio

2 do σ . A presença dessas bases se faz uma característica importante na seletividade do promotor para iniciar a transcrição (LEE e BURUKHOV, 2016)

Elemento -10 estendido

Possui cerca de 4 pares de base, e pode estar presente em alguns promotores com o propósito de aumentar a força da transcrição e fortalecer o complexo aberto. Essa sequência é uma extensão do elemento -10 com as bases consenso TGn (-14 a -17), e faz contatos específicos com o domínio $\sigma^{3.0}$, sua presença se faz importante se o elemento -35 é ausente ou fraco e, também quando a região espaçadora é mais longa (LEE e BURUKHOV, 2016).

Ao fazer o reconhecimento dessa região, os domínios $\sigma^{2.4}$ e $\sigma^{3.0}$ que são carregados positivamente interagem com uma das fitas de DNA recrutando através de interações eletrostáticas nas bases -13 a -10, fazendo uma curvatura 90° para que a base +1 seja posicionada no sítio ativo dentro das fissuras do núcleo da RNA polimerase (LEE e BURUKHOV, 2016).

Região espaçadora

A **região espaçadora** está localizada entre os elementos -35 e -10, e não há uma sequência consenso para os promotores. Para o fator σ^{70} o comprimento ótimo é de 17 pares de bases, esses são os promotores que possuem alto nível de expressão, outros σ podem variar de 6 a 19 pares de bases. O comprimento da região espaçadora pode afetar a interação da região discriminadora com o domínio $\sigma^{1.1}$, isso pode afetar no início da transcrição.

A força do promotor pode ser afetada pela sequência da região espaçadora, influenciando no reconhecimento do elemento -10 ao fazer desvios na curvatura do promotor, porém, essas alterações por si só não podem alterar significativamente a força do promotor.

Dentro da região espaçadora as bases -24 a -18 foram identificadas interagindo com a subunidade β' e essa sequência foi denominada de elemento Z, essa região foi reconhecida aumentando a quantidade de transcrição de alguns promotores, atuando na formação dos complexos abertos quando as interações com os demais elementos são fracas (LEE e BURUKHOV, 2016).

Elemento UP

A região **elemento UP** é rica em A e T, e está presente em alguns promotores, localizada nas posições -40 a -70 (a montante do elemento -35), é considerada uma região do promotor. Os domínios α CTD das subunidades α do núcleo da RNA polimerase são responsáveis pelo reconhecimento dessa sequência, um fator importante para determinação à força do promotor (ROSS e ERNST, 2001)

A sequência elemento UP consiste em duas regiões, proximal e distal. A região proximal está localizada nas bases -45 a -37 e a distal nas bases -58 a -45, os elementos proximais, distais ou ambos que são contatados podem afetar o reconhecimento e atividade do promotor aumentando sua força (LEE e BURUKHOV, 2016).

Ao contatar a região elemento UP (quando tem) do promotor, a RNA polimerase praticamente dobra o DNA no elemento -10 e insere a dupla fita de DNA no sítio ativo do núcleo, em seguida ela abre o DNA para reajustar os contatos. A desnaturação do DNA ocorre nas base A:T da posição -11, a composição aromática dessas bases está envolvida com os estágios iniciais da desnaturação, o segundo nucleotídeo a ser contatado para a abertura da bolha de transcrição é a T:A da posição -7, essas duas posições são as mais conservadas do elemento -10. Outros nucleotídeos fazem contatos inespecíficos com o domínio 2 do σ (ZHANG et al, 2012).

Então, juntos, os elementos -10, -35, -10 estendido e o elemento UP fazem parte dos elementos do promotor que se ligam com o núcleo da RNA polimerase, cada promotor de maneira variada. Um promotor contendo todos esses elementos seria considerado perfeito, e no meio ambiente isso é irreal.

2.5 Vetores de clonagem

2.5.1 Vetores

Os vetores de clonagem são moléculas de DNA usadas para introduzir a informação genética (genes) desejada em uma célula hospedeira, e são capazes de amplificar aquele gene em centenas de cópias. Para isso, os vetores possuem características essenciais como: uma região promotora para regulação da expressão do gene de interesse; uma região de múltiplos sítios de clonagem para a inserção de gene desejado; um terminador de transcrição; uma origem

da replicação (ORI), que pode ser baixa ou alta cópia; e possuem um marcador para selecionar os recombinantes (STRYJEWSKA et al, 2013).

Atualmente, os tipos básicos de vetores usados na tecnologia do DNA recombinante apresentam características especiais que os tornam excelentes veículos de clonagem em diferentes situações. Cada vetor apresenta características diferentes para o implemento do tamanho do gene desejado e o tipo de proteína a ser expressa. (STRYJEWSKA et al, 2013; JONASSON, 2002)

2.5.2 Plasmídeo caça-promotor

Em *E.coli* os vetores mais utilizados são derivados dos plasmídeo. São moléculas de DNA fita dupla, circular e extracromossomal encontrado em espécies bacterianas. Todos contêm três características comuns: marca seletiva, origem de replicação e sítio múltiplo de clonagem.

Vetores utilizados em experimentos de clonagem e caracterização de promotores são chamadas de vetores de seleção de promotores (VSP), ou vetores caça-promotores. Esses vetores são destituídos de seus promotores originais, nesta região se faz presentes um ou mais sítios de clonagem para inserção da sequência de promotores, esses vetores carregam sequências gênicas de seleção, juntamente com um gene repórter. Quando a sequência promotora é inserida no vetor, ocorre a transcrição e conseqüentemente a expressão do gene repórter para seleção das colônias recombinantes, possibilitando a confirmação e quantificação dessa expressão (JONES, 2011; SCHOLZ et al, 2001).

2.6 Expressão de proteínas heteróloga em *E. coli*

A Tecnologia do DNA Recombinante possibilitou a manipulação do genoma, proporcionando isolar, caracterizar e clonar genes extraídos dos mais diversos organismos e a possibilidade de expressar em diferentes hospedeiros (ASTOLFI-FILHO et al, 1996). Essa técnica forneceu aos pesquisadores grandes quantidade de proteínas, que anteriormente eram impossíveis de obter por meios naturais.

Para que uma elevação na produção de proteínas heterólogas seja viável, é necessário pensar na fácil manipulação do hospedeiro, a estabilidade e rendimento da proteína escolhida.

A escolha do hospedeiro varia de acordo com suas propriedades e do uso final da proteína expressa. O hospedeiro mais utilizado é a bactéria gram-negativa *Escherichia coli*, por crescer em meio de cultura de baixo custo e de forma rápida, além de ter seu genoma sequenciado e suas linhagens bem caracterizadas. (JONASSON, 2002).

Técnicas que ampliam a expressão heteróloga de proteínas em *E. coli* tem sido estudada durante anos (KAUR et al, 2018). O objetivo é melhorar a expressão gênica e simplificar os processos de recuperação das proteínas, com o planejamento vetores de expressão mais eficientes e o melhoramento das cepas bacterianas (ex, deleção de genes de protease, expressão de chaperonas), (BERLEC ESTRUKELJ, 2013).

2.7 Construção de bibliotecas de promotores para produção de proteínas heterólogas

Tem sido muito útil a inclusão de novos componentes genéticos nas células, como ferramentas para melhorar e estudar os sistemas naturais. Algumas vias metabólicas elaboradas sinteticamente podem aumentar e otimizar o nível da expressão para determinadas aplicações (ALPER, 2005; DAVIS, 2011).

Técnicas experimentais e cálculos matemáticos foram utilizadas para identificar as etapas de alguma via metabólica. Quanto às técnicas experimentais, a construção de bibliotecas de promotores parece promissora, construindo promotores com diferentes forças de interação e isso faz com que altere a taxa de transcrição do gene e cada promotor produza diferentes quantidades de proteínas (DAVIS, 2011; SOLEM E JENSEN, 2002).

O ajuste da expressão se faz necessário para uma expressão equilibrada, dado que o desequilíbrio metabólico pode levar ao acúmulo de produtos gênicos, que possam ter efeitos citotóxico, esgotamentos necessários para o crescimento celular ou pode resultar em uma carga metabólica indevida sobre a célula, por isso a escolha cuidadosa da força do promotor para cada determinada proteína se faz importante para a expressão (DE MEY et al, 2007; SANTOS E STEPHANOPOULOS, 2008).

As bibliotecas de promotores é uma técnica utilizada para obtenção de diversos promotores com intensidades diferentes de expressão, na literatura, são descritas técnicas diferentes para a gerar bibliotecas de promotores para cada microrganismo ou grupo de organismos (DE MEY et al, 2007).

Uma das técnicas foi desenvolvida por Jensen & Hammer (1998), eles analisaram sequências promotoras, mantiveram as regiões consenso (-10 e -35) e os demais pares de base foram obtidos por mutações aleatórias utilizando oligonucleotídeos sintéticos degenerados. Como resultado, foi desenvolvida biblioteca de promotores para *Lactococcus lactis*, com 25%, 50%, 200% e 400% acima do nível normal de expressão do produto gênico. Deste modo, essa técnica começou a ser aplicada em outros organismos (JENSEN E HAMMER, 1998; RUD et al, 2006).

Outra técnica utilizada para gerar uma biblioteca de promotores foi desenvolvida por meio de PCR propenso a erros, essa técnica introduz mutações pontuais em promotores nativos bem caracterizados. Com um cofator não específico para *taq* polimerase, são geradas sequências com diferentes mutações aleatórias ocasionando uma grande diversidade de sequências promotoras, com forças de interações variadas, que podem ser usados para diversas aplicabilidades (GRUET et al, 2012; KAGIYA et al, 2005; SANTOS E STEPHANOPOULOS, 2008)

Em alguns casos promotores fortes apresentaram uma menor produção devido o aglomerado de proteínas na célula, podendo ocasionar morte celular. Em outros casos com promotores mais fracos na maior produção de proteína heteróloga, mostrando a importância de visar o equilíbrio da super expressão e o metabolismo celular (ALPER et al, 2005).

Este trabalho tem como objetivo gerar uma biblioteca de promotores que sejam reconhecidos tanto pelo fator sigma 70 como pelo fator sigma 38, capazes de programar a produção de proteínas em dois estágios de crescimento celular, exponencial e estacionária. Em produções de larga escala, essa estratégia nos proporciona um aumento na quantidade final de proteínas, sem haver aumentos de custos. Na biotecnologia esses promotores podem ser usados para os mais diversos fins (produção medicinal, farmacêutico, fitoterápico, entre outros).

3. OBJETIVOS

3.1 Geral

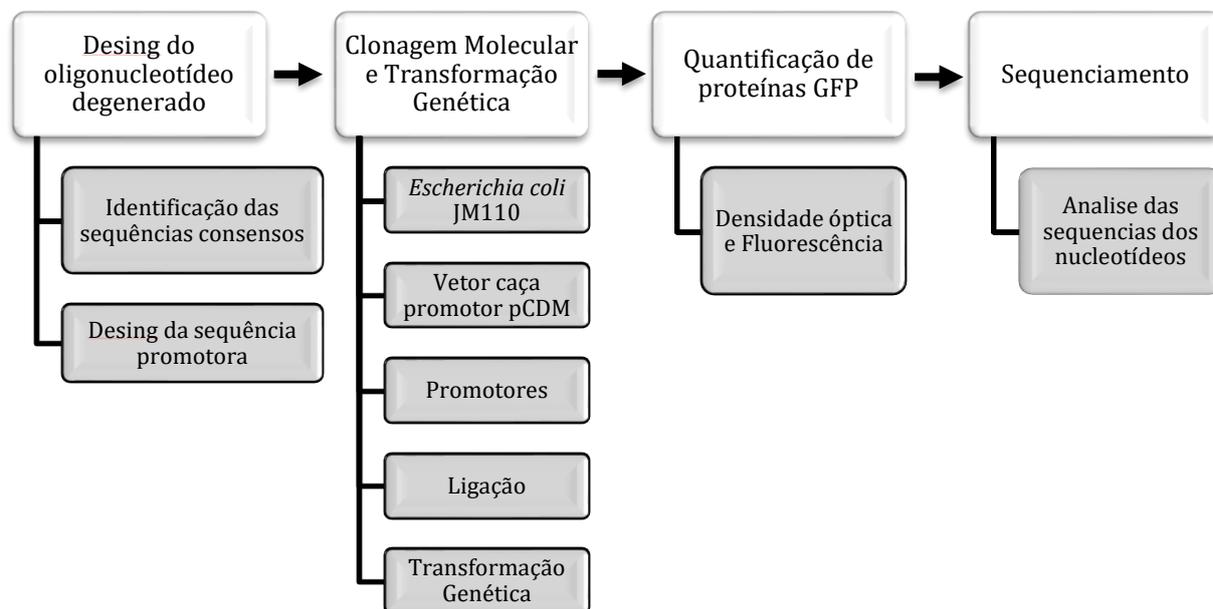
- Design, clonagem e análise de promotores sintéticos reconhecidos pelos fatores sigmas σ^{70} e σ^{38} em *Escherichia coli*.

3.2 Específicos

- Desenhar um oligonucleotídeo degenerado visando reconhecimento simultâneo pelos fatores sigmas σ^{70} e σ^{38} ;
- Amplificar por PCR convencional e clonar em *E. coli* os diferentes novos promotores sintéticos;
- Analisar a expressão da GFP sob o controle dos novos promotores sintéticos. através de ensaios de fluorescência.
- Determinar e analisar as sequências de nucleotídeos dos novos promotores.

4. MATERIAIS E MÉTODOS

4.1 Desenho experimental



4.2 Hospedeira Bacteriana

A linhagem hospedeira *E. coli* utilizada neste trabalho será a linhagem **JM110**: rpsL (Strr) thr leu thi-1 lacY galK galT ara tonA tsx dam dcm supE44 Δ (lac-proAB) [F' traD36 proAB lacI^q Z Δ M15]. A hospedeira JM110 apresenta resistência ao antibiótico estreptomicina e contém o gene *lacI^q* em um plasmídeo F' que codifica a proteína repressora *LacI* em alto nível capaz de bloquear a transcrição de promotores baseados no sistema *lac*. As células foram crescidas e trabalhadas em meio LB (tabela 3).

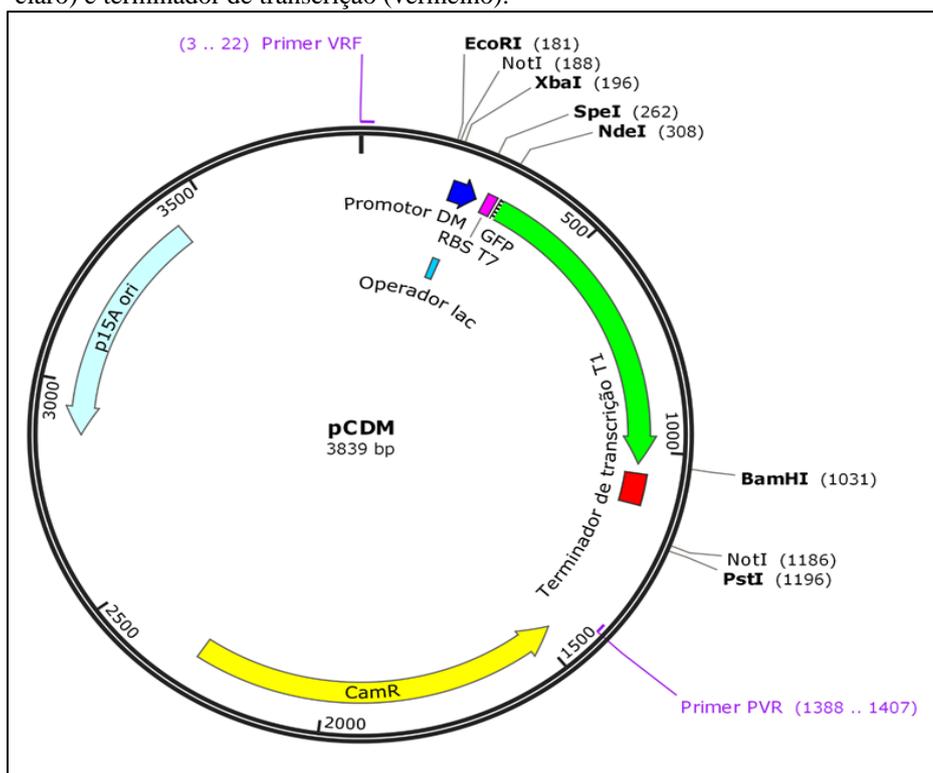
Tabela 3: Meio de cultura Luria-Bertani

Meio de cultura – Luria-Bertani	Composição
Líquido	Peptona 10g/L; Extrato de levedura 5g/L; Cloreto de sódio 10g/L; Água destilada 1000 mL

4.3 Vetor caça promotor pCDM

O vetor de expressão utilizado foi o pCDM (fig. 7). Este vetor possui uma origem de replicação de baixo número de cópias (10-12 cópias por célula), o gene de resistência a canamicina e o gene da proteína verde fluorescente (GFP) sob o controle do promotor bem regulado denominado de DM. Para ser usado como vetor caça-promotor o promotor DM é retirado por digestão com *EcoRI* e *SpeI* do plasmídeo pCDM sendo o restante utilizado para clonar e analisar outros promotores.

Figura 6: **Mapa plasmidial do vetor pCDM criado com o software SnapGene. O plasmídeo pCDM, contendo origem de replicação p15A ori (azul claro) e gene CamR (amarelo), promotor DM (azul escuro), gene da Proteína Verde Fluorescente (verde claro) e terminador de transcrição (vermelho).**



4.4 Design dos oligonucleotídeos degenerados

4.4.1 Ferramentas de bioinformática para identificação das sequências consenso fatores sigmas σ^{70} e σ^{38}

As sequências promotoras de diferentes genes foram obtidas através do *site National Center for Biotechnology Information* (NCBI) e pelo RegulonDB. Com o intuito de identificar as sequências dos elementos regulatórios reconhecidas pelos fatores σ^{70} e σ^{38} .

Os softwares SwissRegulon, MEME-Suite (do inglês, *Multiple Em for Motif Elicitation*) e *Regulatory Sequence Analysis Tools* (RSAT) tiveram como objetivo alinhar as sequências promotoras, a fim de inferir motivos padrões e consensuais presente nos diferentes fatores sigma, obtidos através de análises estatísticas da frequência dos pares de base presentes em cada posição no conjunto dos promotores, desenvolvendo uma matriz de peso-posição - PWM (do inglês, *Position-Weight Matrix*).

4.4.2 Design da sequência promotora

O design dos promotores foi embasado na análise da sequência consenso dos promotores reconhecidos pelos σ^{70} e σ^{38} . Após identificação dos elementos regulatórios, (como os hexâmeros -35 e -10, -10 estendido e o discriminador) que interagem com fatores σ , um oligonucleotídeo foi engenheirado sobrepondo em diferentes arquiteturas as sequências.

Após a escolha das sequências promotoras, foi determinado o promotor sintético, nele foi adicionado as extremidades uma região com múltiplos sítios de clonagem para ação das endonucleases de restrição *EcoRI*, *XbaI*, *BamHI*, *SpeI* e *PstI*, a síntese química do oligonucleotídeo foi realizada pela empresa Exxtend. O sítio de *EcoRI* e *SpeI* foi adicionado nas extremidades do promotor para clonar no plasmídeo pCDM, pois eles clivam a região onde está o promotor DM. O sítio de *BamHI* foi inserido nos novos promotores para clivar da região promotora até o final do gene da GFP, para liberar um fragmento de aproximadamente 750 pares de base na presença do promotor e com isso confirmar a clonagem de novos promotores.

4.5 Procedimentos de Clonagem Molecular e de Transformação Genética

4.5.1 Promotor

Os diferentes oligonucleotídeos fita simples (tabela 4) foram duplicados pelo método de PCR descrito por *Mullis et al*, 1980. Os *primers* Pm-1 e Pj-1 (tabela 4) foram utilizados para amplificação do promotor, a concentração foi de 20 pmol/ μ L para a reação de PCR, onde as condições foram ajustadas em termociclador BIO-RAD T100: temperatura do ciclo 96°C/ 1

min.; 25 ciclos (96 °C / 15seg, 50 °C / 15seg. 60 °C / 1,30 min. O resultado da reação de PCR foi analisado por eletroforese em gel de agarose 0,8%, corado com brometo de etídio e visualizado sob luz ultravioleta (UV) em fotodocumentador.

Tabela 4: Sequências dos oligos utilizados neste trabalho.

Nome	Sequência (5'-3')	Tamanho
<i>Uni-1</i>	TCTAGACTGAATTCTAATGCDDDDHHVVTTGWSWWWYCNYNNWW TGTGTCTATACTDDDTCDGGGATCCTACTAGTACTGCAG	85 bases
<i>Pm-1</i>	TCTAGACTGAATTCTAATGC	20 bases
<i>Pj-1</i>	CTGCAGTACTAGTAGGATCC	20 bases
<i>VF1</i>	TGCCACCTGACG TCTAAG AA	20 bases
<i>RCPI</i>	GAGAAAGTAGTGACAAGTGTT GGC C	25 bases

Onde, N= Qualquer base; H= Base A, C ou T; V= Base A, C ou G; y= Base T ou C; w= Base T ou A; D= Base T, G ou A; S= Base G ou C

5.5.2 Preparo de células competentes

As células foram reativadas por estrias cruzadas em meio sólido com antibiótico em placas de Petri, depois foi realizado um pré-inóculo de células competentes *E.coli* JM110 em 5 mL de meio LB com o antibiótico canamicina, e incubado em agitador de bancada (tipo *shaker*), com 160 rpm a 37°C durante a noite. Do pré-inóculo foi retirado 1 mL e foi inoculado em 200 mL de meio LB sem antibiótico. O inóculo foi incubado sob agitação de 150 rpm a 37°C e neste período foram retiradas alíquotas de 1 mL, em intervalos de 1 hora até atingir a densidade ótica de - O.D.600 = 0,6 nm.

Os tubos falcon utilizados foram resfriados no gelo, quando a O.D. foi alcançada foi distribuído 50 mL da cultura bacteriana para cada tubo. As células foram centrifugadas a 4.000 rpm por 10 minutos a 4°C. O sobrenadante foi descartado de imediato e ressuscitado gentilmente com 10 mL de glicerol 10% estéril gelado até homogeneizar todo o pellet. Esta metodologia foi repetida seguidamente por 4 vezes e as células foram reunidas em um único tubo falcon no final.

Logo após as lavagens e centrifugações, o pellet foi ressuscitado gentilmente em 1 mL de glicerol 10% gelado e estéril. As células eletrocompetentes foram separadas em alíquotas de 100 µL distribuídas em microtubos, que foram congeladas imediatamente e estocadas a -80°C.

5.5.3 Vetor caça promotor pCDM

O plasmídeo pCDM foi introduzido na célula eletrocompetente JM110 por eletroporação. A hospedeira foi deixada no gelo por 15 minutos, logo após, foi adicionado (2µL) do plasmídeo com as células (100µL) bacterianas e foram homogeneizados. E então, transferidos para a cubeta e inserido no aparelho “*Eletroporator 2510 Eppendorff*” onde foi aplicada uma corrente elétrica de 1900 mvolts.

Após o choque, foi adicionado 500 µL do meio LB líquido na cuveta foram homogeneizados e transferidos a um tubo *ependorff* de 2mL para a recuperação celular, logo após plaqueadas em alíquotas no meio LB sólido e incubadas a 37°C. Dos clones crescidos foi feito o pré-inóculo em 5mL durante a noite. Foi retirado 1mL e transferido para 50mL complementado com canamicina 50µg/mL e glicose a 1%.

A cultura foi incubada sob agitação de 150 rpm a 37°C no *shaker* durante 72 horas, em intervalos de 1 hora foram retiradas alíquotas de 1 mL, até obter a O.D. de 0,8. Após isso foi adicionado 2.5 mL de EDTA a 0.5M pH 7,5 e incubada novamente em agitação por 3 horas. Foi retirado alíquotas de 3 mL em *ependorfs* até formar *pellet* para extração plasmidial, a extração foi feita utilizando o kit da Cellco Fast-n-Easy Plasmid Mini-Prep, seguindo orientações do fabricante.

5.5.4 Digestão e purificação do vetor pCDM e dos diferentes promotores

Os fragmentos gerados da digestão dos plasmídeos e do promotor foram visualizados em gel de agarose 0,8% para análise. Para realizar a clonagem dos novos promotores aos vetores pCDM foi realizado uma reação com as enzimas de restrição *EcoRI* e *SpeI* em menor volume (digestão analítica) tanto para o vetor quanto para os promotores. O sistema de digestão do plasmídeo foi submetido a eletroforese em gel de agarose 0,8% e visualizado no transiluminador com marcador G Ladder 1Kb para confirmação do pCDM digerido.

Ao ser confirmado, foi retirado o fragmento do gel de agarose para purificação do plasmídeo com o Kit comercial ilustra GFX PCR DNA and Gel Band Purification, SIGMA. A purificação foi realizada de acordo com as especificações do fabricante.

Para visualizar a digestão analítica do promotor foi utilizado um gel de agarose de 2% utilizando o marcador molecular de DNA de 100 pb da Cellco. A purificação dos promotores dupla fita foi realizada cortando o fragmento digerido do gel de agarose e adicionando em um

ependorf com 200µL de Água ultrapura, foi feita a maceração do gel de agarose e incubados a 4°C por 16 horas.

Os fragmentos cortados e purificados (pCDM e promotores dupla fita) foram quantificados no Qubit fluorometric quantitation. Após, foi montado a reação de ligação de ambos: pCDM e diferentes promotores dupla fita com a T4 DNA ligase conforme instruções do fabricante. Usou-se a relação molar 2:1 entre plasmídeo para promotor, 5µL de tampão, 2µL de enzima para reação final de 50 µL. Foi utilizado 25 µL do vetor contendo 15ng e 15 µL do promotor contendo 1,05ng equivalente a $3,7 \times 10^9$ e $1,39 \times 10^{10}$ de moléculas respectivamente, foi elaborado a reação de ligação com a enzima T4 DNA ligase seguindo as recomendações do fabricante. A reação de ligação foi colocada em termociclador MJ Research Minicycler PTC-150 a 16 °C por 16 horas e após o período a reação foi purificada por precipitação.

5.5.5 Transformação Genética

A transformação foi feita por eletroporação, as células utilizadas foi a hospedeira *E. coli* JM110 eletrocompetentes. A hospedeira foi deixada no gelo por 15 minutos, logo após, foi adicionado (2µL) do plasmídeo ligado com o promotor e as células (100µL) bacterianas e foram homogeneizados. E então, transferidos para a cubeta e inserido no aparelho “*Eletroporator 2510 Eppendorff*” onde foi aplicado uma corrente elétrica de 1900 mvolts.

Após o choque, foi adicionado 500 µL do meio LB líquido na cuveta e foram homogeneizados e transferidos a *ependorf* de 2 mL para a recuperação celular. As amostras foram incubadas por 1 hora a 37°C com agitação de 150 rpm. Após esse tempo, alíquotas de 200µL de cultura bacteriana foram distribuídas em 3 placas de Petri contendo meio LB sólido suplementado com o antibiótico Canamicina a 50µg/mL.

5.6 Quantificação da expressão de GFP

Para quantificar a produção de GFP inicialmente os clones recombinantes crescidos com os diferentes promotores foram escolhidos em placa pela intensidade de expressão apresentada, a unidade de medida de fluorescência é simbolizada pela sigla UMF. Foi realizado a réplica *plate* dos clones, a partir da réplica *plate* foram selecionados clones com intensidades fracas, médias e fortes para eleger clones da biblioteca de promotores. Inicialmente foram escolhidos 30 clones, sendo 10 de intensidade fraca, 10 médias e 10 fortes. Foi feito o pré-

inoculo dos clones selecionados a partir dessa etapa, em 5 mL de meio LB, contendo antibiótico e incubados a 37°C no agitador durante a noite.

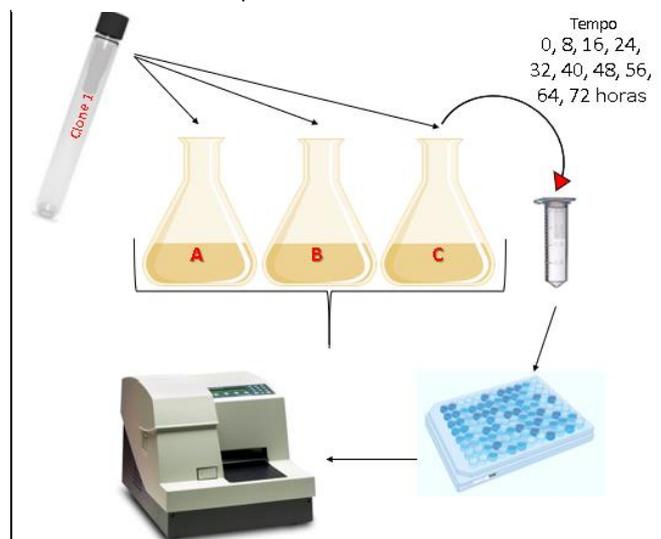
Então foram inoculados 1 mL da cultura em 50 mL de meio de cultura LB líquido suplementado com canamicina a 50µg/mL e incubado em *shaker* a 37°C, agitados a 150 rpm. Foram retiradas alíquotas de 1mL de 8 em 8 horas durante 72 horas para análise de O.D. e fluorescência.

Para medida de O.D. foram utilizadas 200µL de cultura em microplacas de 96 poços de fundo transparente. A densidade óptica foi medida como excitação de 595nm. Para medida de fluorescência a alíquota restante (800µL) foi centrifugado por 3 minutos, o sobrenadante foi descartado e adicionado o mesmo volume de tampão TN (0.15M NaCl + 0,10M Tris-HCL) homogeneizado a amostra e retirada a alíquota de 200uL para a leitura (fig. 8)

A leitura foi feita em microplaca preta com excitação de 395 nm e emissão a 509 nm em um leitor Hidex Plate CHAMELEON V. Como controle foi utilizado o plasmídeo p26 caracterizado para fase estacionária (WÖHLKE, et al 2002) e o promotor DM caracterizado para fase exponencial no vetor pCDM é patente do grupo de biologia molecular da UFAM.

O tempo de indução foi determinado visando que o design do promotor reconhecido pelo fator sigma 70 e 38, foi selecionado o tempo de intervalo de 8 por 72 horas, contendo 10 coletas incluindo o tempo 0. Após 8 horas de cultivo foi necessária a diluição das medidas de OD e fluorescência, sendo então diluídos 1:9, os resultados foram analisados no *software* IBM SPSS através do teste de normalidade, médias e desvio padrão.

Figura 7: Desenho experimental de medidas de OD e GFP. Nessa figura está apresentada a sequência de passos para realização da medida de densidade óptica e fluorescência dos clones.



5.7 Estatística

Os resultados da fluorescência foram analisados no teste de normalidade, e foi tirada a média e desvio padrão para elaborar os gráficos com as medidas de fluorescência e do crescimento (O.D) dos clones em função do tempo. No eixo primário, estão apresentadas as unidades de medida de fluorescência alcançada pelos clones apresentados em barras, no eixo secundários está a densidade ótica (O.D) exibida pelos clones que estão indicados nas abscissas.

5.8 Sequenciamento

Foram escolhidos 16 clones a partir da quantificação, sendo eles 5 com intensidades fracas, 5 médias e 6 fortes, e então foram crescidos em pré-inóculo em 5mL de meio LB durante a noite. Foi retirado 1mL e transferido para 50mL complementado com canamicina 50µg/mL e glicose a 1%. O inóculo foi incubado sob agitação de 150 rpm a 37°C no *shaker* e neste período foram retiradas alíquotas de 1 mL, em intervalos de 1 hora até atingir O.D. de 0,8.

Após isso foi adicionado 2.5 mL de EDTA a 0.5M pH 7,5 e incubado novamente em agitação por 3 horas. Foram então retiradas alíquotas de 3 mL, a partir das quais plasmídeos foram extraídos utilizando o kit da Cellco Fast-n-Easy Plasmid Mini-Prep, seguindo orientações do fabricante.

Foi visualizado em gel de agarose 0,8% a integridade dos plasmídeos e foi a quantificação no Qubit fluorometric. O sequenciamento dos promotores foi realizado a partir da amplificação de fragmento amplificado por PCR onde as condições foram ajustadas em termociclador BIO-RAD T100: termal cycle 96°C/ 1 min.; 25 ciclos (96 °C / 15seg, 50 °C / 15seg. 60 °C / 1,30 min. Para a reação de PCR a concentração dos plasmídeos foi de 10 ng, utilizando os *primers*: VF1 e RCP1 para amplificar 504 pares, onde inclui a região do promotor (tabela 4).

Após a PCR, para realizar as reações de sequenciamento, foram feitas purificações do DNA através do Kit comercial illustra GFX PCR DNA and Gel Band Purification e visualizou-se os amplicons em gel de agarose 0.8% corado com brometo de etídio em um transluminador e quantificou-se no Qubit fluorometric.

As reações de sequenciamento foram feitas utilizando o kit Sequencing-Big Dye Terminator Cycle Sequencing Ready ABI versão 3 – Applied Bioscience/Life Technologies, USA. Cada reação de sequenciamento foi ajustada para um volume final de 10 µL, com o DNA

amplificado, os oligonucleotídeos iniciadores VF1 e RCP, e tampão Big-Dye para 1X final e 2 μ L de Big-Dye. As condições de reação foram: 96°C / 1 min.; 25 ciclos (96°C / 15seg, 50 °C / 15seg. 60 °C / 1,30 min.) em sequenciador ABI, modelo 3500XL (Applied Biosystems).

6 RESULTADOS E DISCUSSÃO

6.1 Design do oligonucleotídeo degenerado

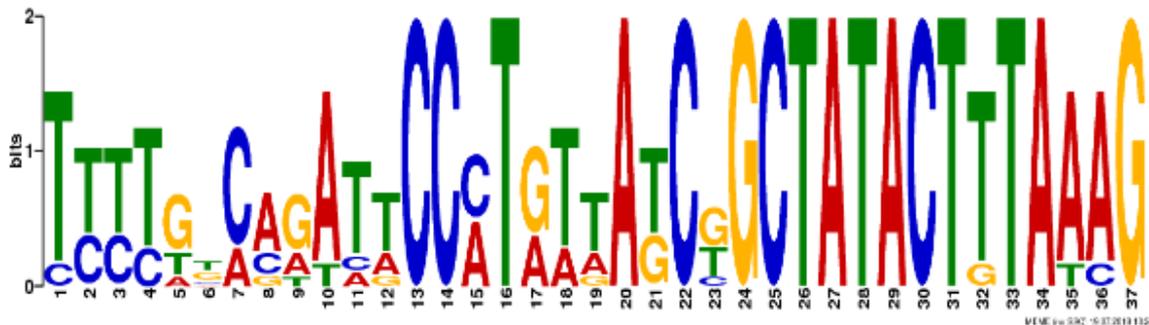
Para a obtenção de promotores funcionais tanto na fase exponencial quanto na fase estacionária do cultivo bacteriano, a estratégia foi construir uma biblioteca de promotores a partir de um oligonucleotídeo parcialmente degenerado ao qual na sua síntese química mantinha-se os nucleotídeos considerados mais importantes para a transcrição em cada uma das fases de crescimento e degenerava-se os menos importantes. Para tal utilizou-se o conhecimento já descrito na literatura das regiões consensuais dos dois tipos de promotores (sigma 70 e sigma 38) e também uma análise feita em nosso laboratório de 8 promotores que se expressam na fase estacionária, por: J.P.K. Albuquerque e M. C. T. Astolfi, descrita a seguir.

Foi realizada uma análise de consenso a partir dos 8 melhores promotores selecionados de uma biblioteca de promotores construída por Miksch, 2005, cuja sequências estão mostradas na tabela 5 a seguir. A análise de consenso foi feita utilizando-se o *software* MEME-Suite. O *software* gera algoritmos baseado na inferência de motivos, gerando uma matriz onde é designado um peso para cada base em uma determinada posição (figura 8).

Tabela 5: Biblioteca de promotores reconhecidos por σ 38. Promotores desenvolvido no trabalho de Miksch, 2005 com intenção de aumentar o nível de transcrição da proteína, em negrito a região consensual -10 reconhecida pelo sigma 38.

Promotor_1	TTTTTACAATCCCTGTTAGCGGCT TATACTTTAAAG
Promotor_2	TTCTGCCGATTCCCTGTTAGCTGCT TATACTTTAAAG
Promotor_3	TTCTGTCATATCCCTATTAGCTGCT TATACTTTAAAG
Promotor_4	TCCTGTCAGAATCCATGAAATCCGCT TATACTTTAAAG
Promotor_5	TTTTGCCAGATTCCCTGTGATCTGCT TATACTTTAAAG
Promotor_6	TCTCGGCAGATACCATATTATCGGCT TATACTGTATCG
Promotor_7	TCTCGGCAGATACCATATTATCGGCT TATACTTTAAAG
Promotor_8	CTTTAAAGATCGCCATGAAATCGGCT TATACTTTAAAG

Figura 8:Matriz de peso-posição plotada em gráfico pelo software MEME-Suite. Padrões de seqüências de DNA para cada posição nos promotores reconhecidos pelo sigma 38, tabela 5.



O desenho do promotor é:

**TCTAGACTGAATTCTAATGCDDDDHHVVVTTGWSWWWWYCNYNWWTGTGT
CTATACTDDDTCDGGGATCCTACTAGTACTGCAG**

Onde: N= Qualquer base; H= A, C ou T; V= A, C ou G; Y= T ou C; W= T ou A; D= T, G ou A; S= G ou C.

Descrevendo em mais detalhes, os nucleotídeos marcados com:

- Verde correspondem ao sítio da enzima de restrição *XbaI*;
- Marrom correspondem ao sítio da enzima de restrição *EcoRI*;
- Vermelho correspondem ao hexâmetro -35;
- Rosa correspondem aos nucleotídeos -26, -25 e -23 (encontrado nos promotores de fase estacionária);
- Roxo correspondem à região -10 estendida;
- Azul corresponde ao hexâmetro -10 (TATA Box) dos promotores de fase estacionária devido à alteração na penúltima base (C);
- Amarelo corresponde ao sítio da enzima de restrição *BamHI*;
- Verde claro corresponde ao sítio da enzima de restrição *SpeI*;
- Azul claro corresponde ao sítio da enzima de restrição *PstI*.

Manteve-se o distanciamento de 17 nucleotídeos entre os hexâmetros -35 e -10 que é considerado o ideal para os promotores reconhecidos pelo fator σ^{70} . Essa nova abordagem resultou na criação de uma biblioteca de promotores, com a expectativa de que diversos deles pudessem funcionar tanto na fase exponencial como na fase estacionária em *E. coli*, com diferentes forças. Pelas degenerações desenhadas no promotor a biblioteca deve conter 1.741.824 sequencias diferentes.

Os promotores de fase estacionária são reconhecidos pelo fator σ^{38} um fator alternativo ao σ^{70} que reconhece os genes constitutivos, porém as sequências consenso entre os fatores sigmas são muito semelhantes. Na posição -8 a especificidade do promotor de fase estacionaria é a citosina, a qual foi mantida junto com a constante da sequência consenso -10 -TATACT-, a região -10 estendida -TGTGTC-, essa região pode contribuir para o reconhecimento específico do σ^{38} (HENGGE-ARONIS, 2002; GAAL ET AL, 2001).

Duas posições o -23 e -26 foram randomizados para timina ou citosina, preferência do σ^{38} . Casualmente, o -35 não está presente nos promotores de fase estacionaria devido a presença da região -10 estendido, porém nesse design o intuito é ser reconhecido pelos dois fatores σ^{38} e σ^{70} , visando isso foi mantido o -35, a posição -30 e -32 foi randomizada para timina ou adenina.

Nas extremidades foram desenhados 5 locais de reconhecimento para endonucleases de restrição diferentes, sendo desenhada a estratégia de clonagem com as enzimas *EcoRI* e *SpeI* de acordo com o plasmídeo selecionado para clonagem do promotor, e *BamHI* para clivar o promotor até o terminador de transcrição do gene de GFP liberando um fragmento de aproximadamente 750 pares de base e certificar a presença do promotor. As demais bases foram randomizadas de acordo com os cálculos de frequência da ocorrência de nucleotídeos nos promotores selecionados (tabela 5) de fase estacionária relatados anteriormente.

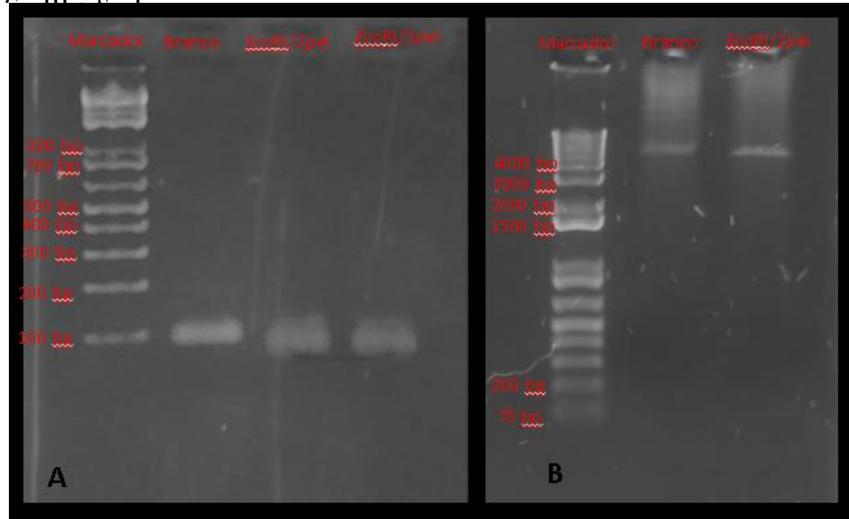
6.2 Clonagem molecular dos promotores

Para a clonagem molecular dos promotores utilizou-se o vetor de expressão pCDM do qual retirava-se a região do promotor DM por digestão com *EcoRI* e *SpeI*, que era substituído pelos promotores sintéticos obtidos por PCR a partir do oligonucleotídeo degenerado também digerido por *EcoRI* e *SpeI* em suas extremidades. A célula hospedeira para o processo de clonagem foi a *E. coli* JM110 eletrocompetente.

Os promotores de fita simples, contidos no oligonucleotídeo degenerado, foram amplificados por PCR usando os primers pJ1 e pM1. O produto da PCR e o plasmídeo pCDM foram digeridos com *EcoRI* e *SpeI*, e os fragmentos gerados foram purificados. A figura 9-A mostra uma fotografia do gel de agarose 2%, contendo marcador, o fragmento de PCR do promotor e duas amostras com o promotor digerido por *EcoRI* e *SpeI*. A figura 9-B mostra uma fotografia do gel de agarre 0,8% contendo o marcador, o plasmídeo intacto e o plasmídeo digerido com as enzimas *EcoRI* e *SpeI*. Os fragmentos purificados foram ligados usando a

enzima T4 DNA ligase por 16 horas e o produto da ligação foi concentrado por precipitação com etanol.

Figura 9: Perfil eletroforético do produto digestão enzimática dos promotores e dos plasmídeos. A- Gel de agarose 2% da esquerda para a direita, o marcador, controle (reação da PCR) e digerido com *EcoRI* e *SpeI*. B- Gel de agarose 0,8% da esquerda para a direita, o marcador, controle (plasmídeo pCDM) e digerido com

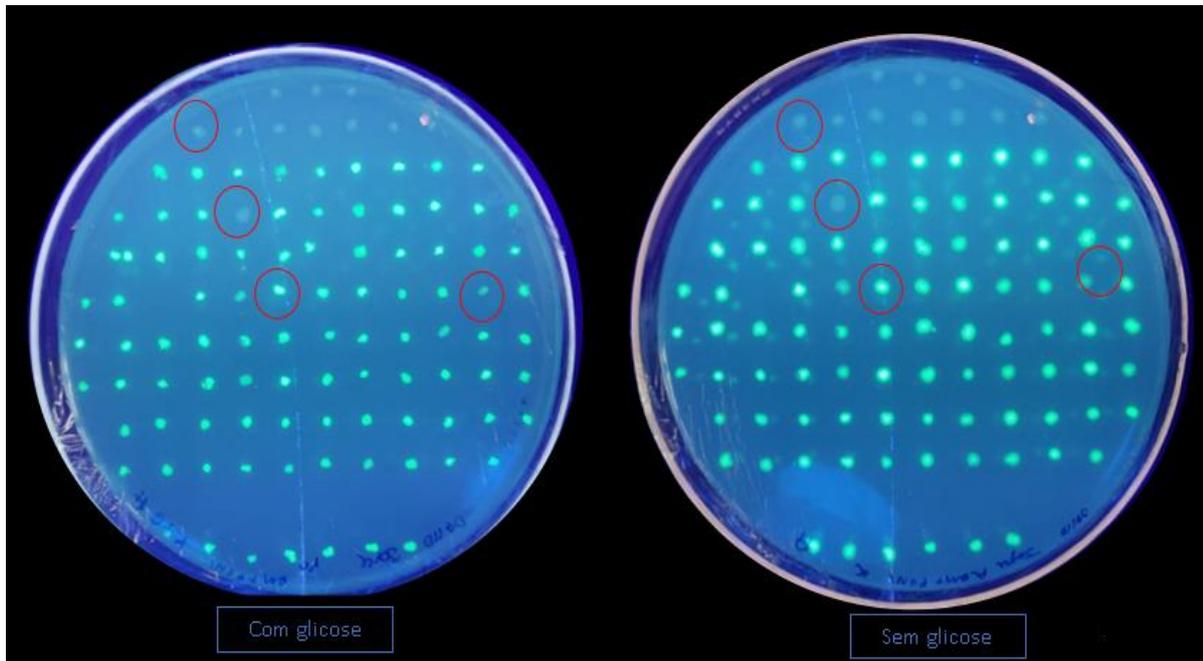


6.3 Expressão da Proteína Verde Fluorescente - GFP

O sistema de ligação purificado foi utilizado para transformar células eletrocompetentes de *E. coli* linhagem JM110, resultando em 1200 clones aproximadamente. Desses clones cerca de 400 foram analisados quanto expressão de GFP em réplicas *plate* em placas de meio LB contendo o antibiótico canamicina e glicose 1% ou apenas canamicina, uma das réplicas está mostrada na figura 10.

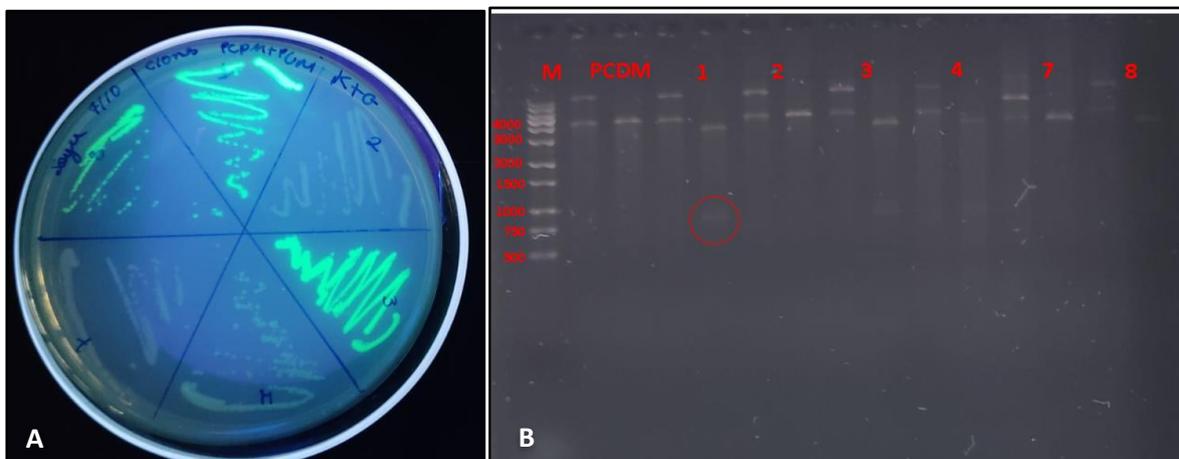
Foi possível, após 24 horas de crescimento, observar colônias exibindo intensidades diferentes de fluorescência, poucas colônias que não apresentaram fluorescência. Para confirmar a presença do promotor sintético foi feito um teste inicial com 6 clones: 2 sem fluorescência, 1 com fluorescência fraca, 3 com fluorescência média ou forte (figura 11-A).

Figura 10: Fotografia de placas de Petri contendo clones com diferentes intensidades de expressão. Esta imagem mostra réplica *plate* de alguns dos 1200 clones obtidos. Na placa da esquerda os clones foram cultivados na presença de canamicina e glicose enquanto na direita o cultivo foi feito apenas com canamicina



Os 6 clones foram cultivados em meio LB líquido para extração dos plasmídeos. Foi feita uma digestão com *Bam*HI desses plasmídeos (figura 11-B), para confirmar se foram ligados com os novos promotores, pois caso houvesse sido clonado o promotor sintético a digestão liberaria um fragmento de DNA de aproximadamente 750 pb.

Figura 11: Fotografia das placas de petri e do gel de agarose confirmando do promotor sintético UNI-1. A Placa de ágar contendo 6 clones com diferentes expressões de GFP, sem fluorescência, fluorescência fraca, média e forte.; B – Gel de agarose 8% contendo a digestão do pCDM com os 6 clones. No pCDM o controle mostra a topologia do plasmídeo não digerido, o controle digerido (segundo poço) lineariza o plasmídeo, assim como na amostra 2 e 7, onde não há a presença do pUNI, no clone 1 o controle mostra a topologia do plasmídeo e o digerido mostra um fragmento próximo do marcador de 700pb e uma queda no fragmento linearizado, assim como nos clones 3, 4 e 8, comprovando a presença do pUNI.



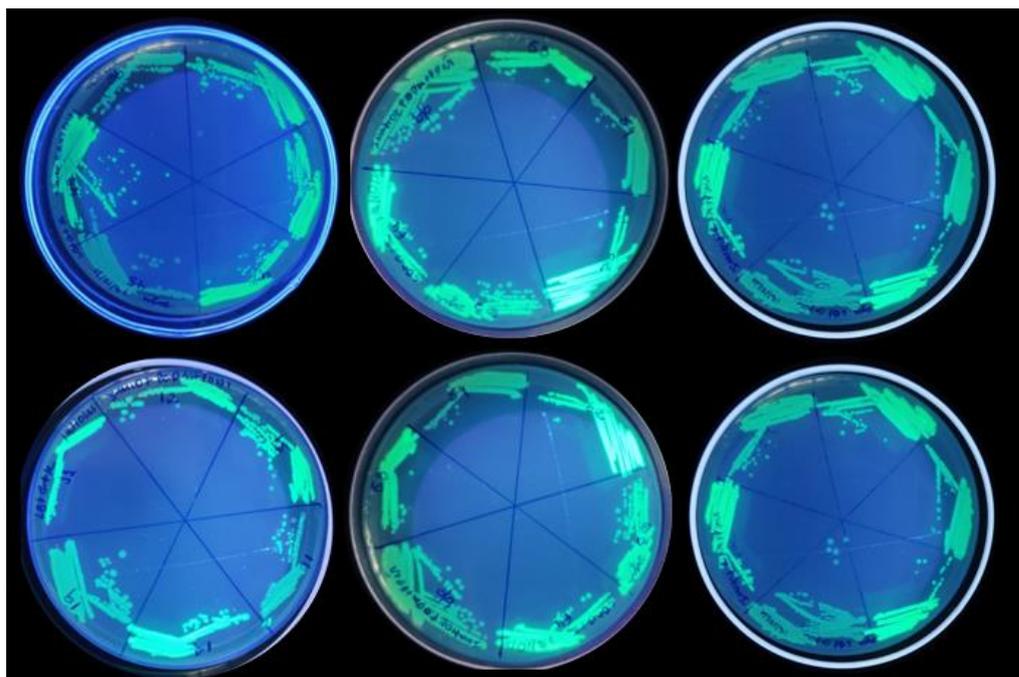
Os clones que não apresentavam fluorescência não possuíam os promotores sintéticos (figura 10 clones 2 e 7 juntamente com o pCDM que apenas linearizaram o plasmídeo), e os demais foram confirmados com a presença do promotor, pois no gel de agarose visualizou-se um fragmento de aproximadamente 750pb nos clones 1, 3, 4 e 8.

Após a confirmação da presença dos promotores nos clones, foram selecionados por análise visual 66 clones com diferentes intensidades de fluorescência para análise e digestão com *Bam*HI. Desses após plaqueamento por estrias cruzadas selecionou-se 30 com diferentes intensidades para estudo da expressão de GFP em meio líquido. A figura 12 mostra exemplos de clones crescidos em placas distribuídos por estrias cruzadas.

Para investigar a expressão dos promotores foi utilizado a GFP para análise em meio líquido e quantificação em espectrofluorímetro. O uso da GFP tem várias vantagens sobre as demais proteínas de sistema repórter, ela não é tóxica e não requer permeabilização de células bacterianas, a GFP não requer adição de substrato e a visualização da sua produção pode ser feita em tempo real (Miksch et al, 2005).

Os padrões iniciais de expressão dos promotores em meios líquidos foram analisados utilizando 30 clones, sendo eles 10 de intensidade fraca, 10 medias e 10 fortes. Os tempos de coleta das alíquotas dos meios líquidos de crescimento foram estabelecidos de 8 em 8 horas.

Figura 10: Fotografia das placas contendo os clones selecionados para analisar a fluorescência e confirmar se havia a presença do promotor. 30 clones foram selecionados para quantificação de GFP e foram estriados em placas contendo LB e ágar

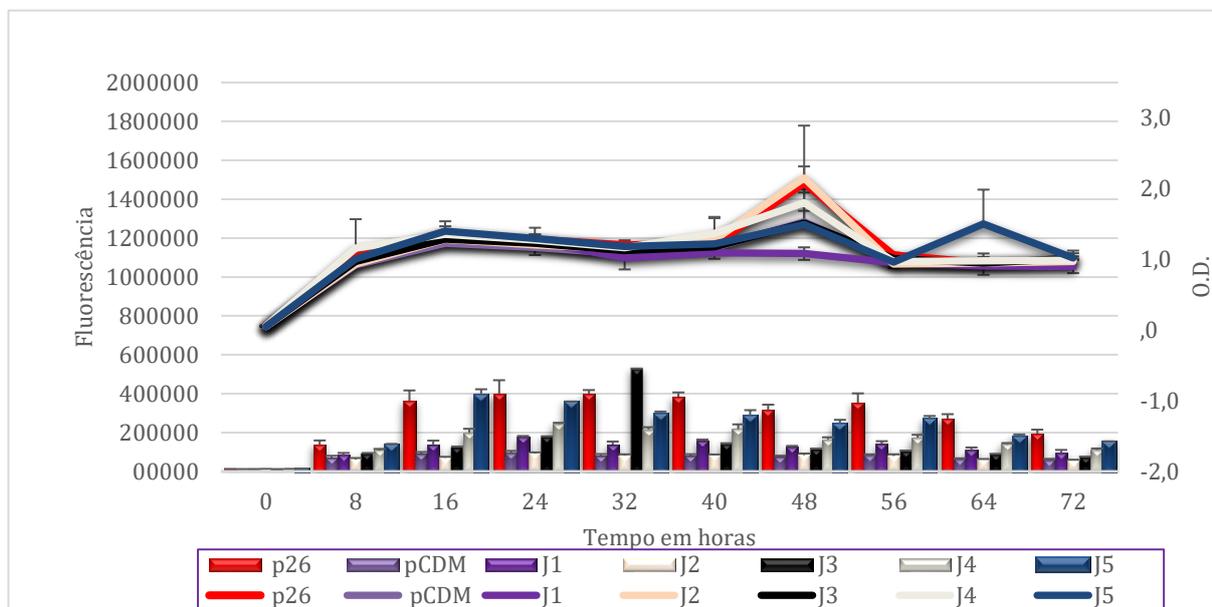


A expressão da GFP começou a intensificar a partir de 8 horas de cultivo, sendo que alguns clones alcançaram o seu ápice de expressão em tempos variados: 16, 24 e 32 horas. Comparando as atividades de expressão dos 30 clones selecionados, a gama de atividades de fluorescência (unidades de medida de fluorescência) variou de 60.000 até 912.000, alcançando seu ápice de produção variando de 16 até 32 horas. Destes 30 clones, foram escolhidos 16 clones sendo 5 de intensidade fraca, 5 de média e 6 de forte, para análise da expressão em meio líquido em triplicatas e determinação das sequencias de seus promotores sintéticos.

Os controles também foram analisados juntamente com os 16 clones escolhidos, o P26 caracterizado de fase estacionária, para analisarmos a expressão durante a fase estacionária, e o pCDM caracterizado de fase exponencial, para analisarmos a expressão na fase exponencial dos clones junto com os controles. Os resultados dessa análise estão mostrados nos gráficos 1, 2 e 3 a seguir.

No gráfico 1 pode-se observar alguns clones com expressão fraca, tendo o clone J3 ultrapassado os controles quando com 32 horas de cultivo, atingindo seu ápice de expressão com aproximadamente 500.000 unidades de GFP, e o clone J2 teve sua expressão muito semelhante ao pCDM.

Gráfico 1. Mostra a densidade óptica (OD) e unidade medida de fluorescência de cultivos dos clones p26 e pCDM (controle), clones J1 ao J5 em função do tempo determinados de 8 em 8 horas por 72 horas

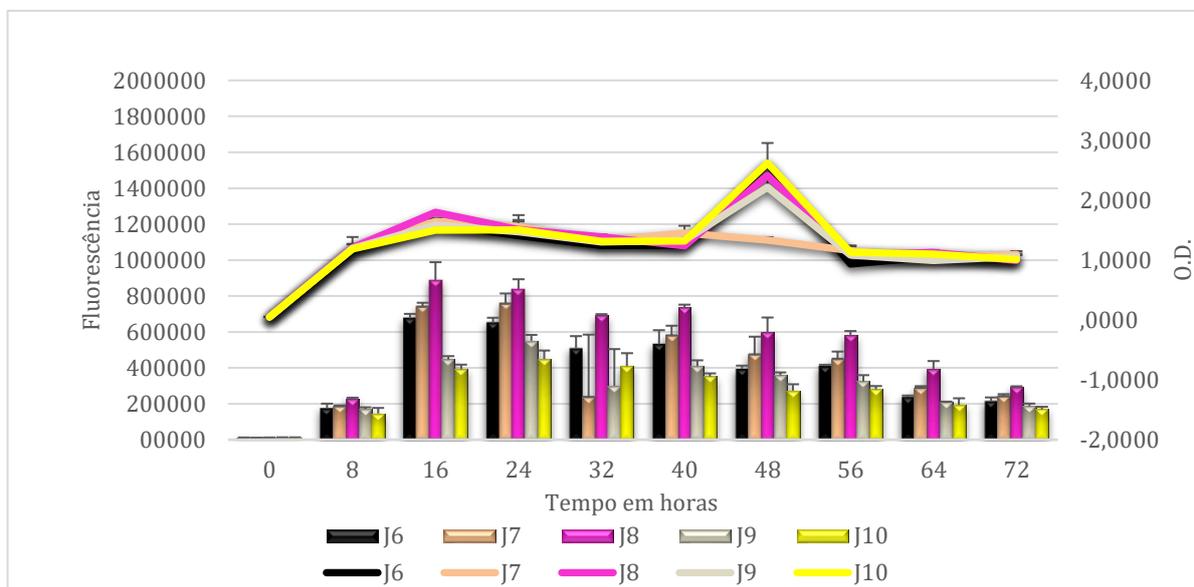


No crescimento, a maioria dos cultivos houve um aumento celular em absorvância (ODs) em 48 horas, foi observado muita produção de biofilmes, dificultando a homogeneização

das amostras. No gráfico analisado por WÖHLKE et al (2002) com o plasmídeo p26 (controle), quando o crescimento chega próximo das 48 horas também houve um aumento no número de células, porém o experimento encerrou com 48 horas.

No gráfico 2, foi possível observar intensidades médias e fortes de fluorescências dos clones, mantendo-se elevadas nos tempos 16, 24, 32 e 40 horas. Obtendo-se aproximadamente 150.000 a 880.000 UMF. O clone J8 apresentou a maior expressão do conjunto mostrado neste gráfico, em 16 horas atingiu seu ápice com 886.166 UMF. Os demais clones apresentaram seu ápice de expressão em tempos diferentes variando de 400.000 a 800.000 UMF.

Gráfico 2. Mostra a densidade óptica (OD) e medida de fluorescência dos cultivos dos clones J6 ao J10 em função do tempo

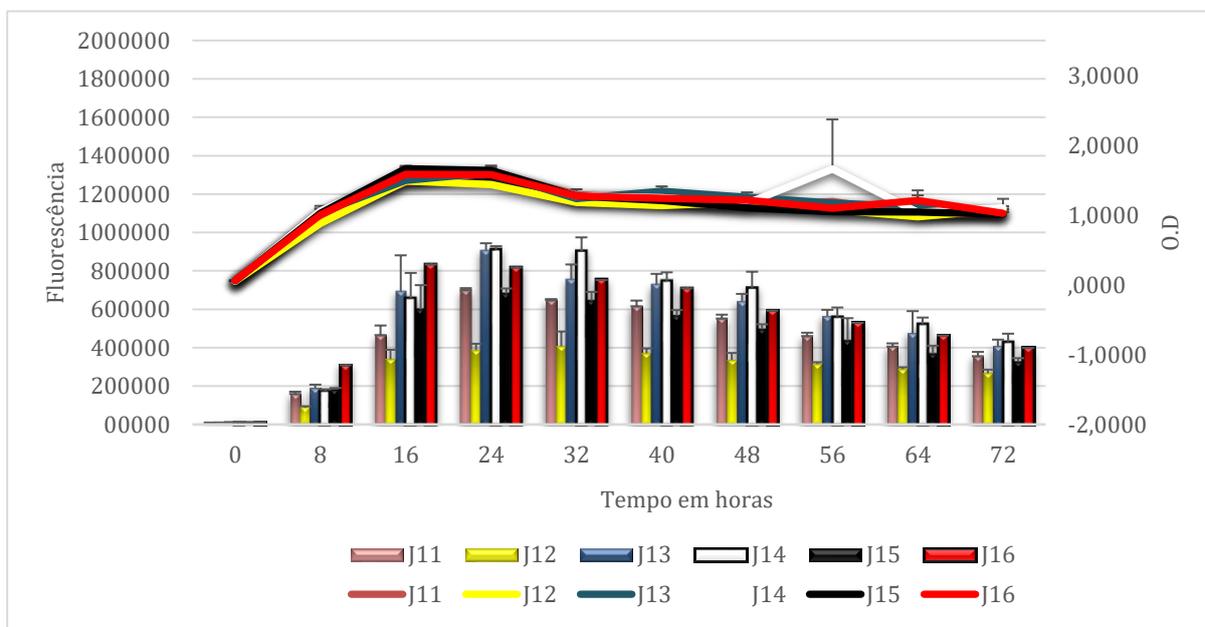


Em nenhum dos gráficos foi observado alteração no crescimento da célula em relação ao nível de expressão da GFP, comparando com o controle pCDM de fase exponencial que teve seu ápice de expressão em 24 horas com 96.063 UMF.

O gráfico 3 apresenta um ranking dos clones que apresentaram expressões elevadas de GFP, com seus ápices de 408.000 até 912.000 UMF em tempo diferentes, os clones J13, J14 e J16 possuem elevados níveis de expressão em relação aos demais, sendo o J14 o clone que mais expressou GFP. Em nenhum dos gráficos foi observado alteração no crescimento da célula em relação ao nível de expressão da GFP, comparando com o controle pCDM de fase exponencial que teve seu ápice de expressão em 24 horas com 96.063 UMF. O gráfico 3 apresenta um ranking dos clones que apresentaram expressões elevadas de GFP, com seus

ápices de 408.000 até 912.000 UMF em tempo diferentes, os clones J13, J14 e J16 possuem elevados níveis de expressão em relação aos demais, sendo o J14 o clone que mais expressou GFP

Gráfico 3. Mostra a densidade óptica (OD) e medida de fluorescência dos cultivos dos clones J11 ao J16 em função do tempo



Em geral o ranking de atividades promotores dos níveis de expressão determinados em culturas líquidas correspondia ao das colônias em placas de LB contendo ágar, porém ao contrário das líquidas não é possível análise quantitativa da fluorescência dos clones em placas.

6.4 Sequenciamento

De acordo com o desenho dos promotores sintéticos, eram possíveis de gerar 1.741.824 sequências diferentes. Para estudar a base molecular das diferentes forças dos promotores foram escolhidos 16 clones para sequenciamento com características de fase exponencial e estacionária, apresentando diferentes intensidades de expressão de GFP: fraca, média e forte. A tabela 6 apresenta os nomes dos promotores sequenciados, a sequência dos promotores, os dados de medida de fluorescência e o tempo em que foi apresentado a maior fluorescência.

Tabela 6: Sequência de nucleotídeos dos promotores. Nas extremidades estão os sítios das endonucleases de restrição nas cores rosa, azul e azul escuro. Em vermelho o elemento -10, em roxo o elemento -10 estendido e sublinhado o -35. UMF indica o pico de fluorescência de cada clone e o tempo em que foi produzido o pico de fluorescência.

PROMOTOR	SEQUÊNCIA DO PROMOTOR	UMF	TEMPO (HORAS)
	GAATTC TAATGCDDDDHHVVV TTG WSWWWYCN YNN WW TGTGTC TATACT DDDDTCDG GGATCC ACTAGTA		
J1	-GAATTC TAATGCTAGGTACA TTG GATAACCCCAATT TGTGTC TATACT GGATCAG GGATCC ACTAGTA	175620	24
J2	----GAATTC TAATGCATTATAACCTGTCTATTTCTTCAAT TGTGTC TATACT AAGTCTG GGATCC ACTAGTA	93830	24
J3	-----GAATTC TAATGCGTTGATAAA TTG TGAAATCTTCTAT TGTGTC TATACT GATT CAGGGATCC ACTAGTA	529080	32
J4	---GAATTC TAATGCTTTATTCAAT TTG TCTATTCCACCAATT TGTGTC TATACT TAATCTG GGATCC ACTAGTA	244170	24
J5	--GAATTC TAATGCGTATCTCCA TTG TCAATTTCCATTAAT TGTGTC TATACT TAATCAG GGATCC ACTAGTA	391943	16
J6	-GAATTC TAATGCATTTCCGAC TTG ACAATTTCCGCTCATT TGTGTC TATACT GGGYCT GGATCC ACTAGTA	675790	16
J7	-GAATTC TAATGCTAAACAGCG TTG ACATATTCTCCAAT TGTGTC TATACT AAGTCAG GGATCC ACTAGTA	757056	24
J8	-GAATTC TAATGCTAAGTTGGG TTG TGTTTTCCGTTTTT TGTGTC TATACT ATATCTG GGATCC ACTAGTA	886166	16
J9	GAATTC TAATGCGGTAACCGG TTG AGATTATCGTTGATT TGTGTC TATACT AATTCTG GGATCC ACTAGTA	550780	24
J10	-GAATTC TAATGCGTTAATCAG TTG ACAATTTCCGCACTT TGTGTC TATACT TTATCG GGATCC ACTAGTA	449260	24
J11	--GAATTC TAATGCATTAATAAC TTG TGTTATTCTTATAAT TGTGTC TATACT TTGTCTG GGATCC ACTAGTA	699336	24
J12	-GAATTC TAATGCTGAAAGCA TTG TCAATACCTTGCTAT TGTGTC TATACT GGATCTG GGATCC ACTAGTA	408273	32
J13	---GAATTC TAATGCTAGTAAAC TTG ACTATATCTTCTTAT TGTGTC TATACT ATATCTG GGATCC ACTAGTA	906463	24
J14	----GAATTC TAATGCAAATTGGC TTG TCTTATTCTTTAAT TGTGTC TATACT TTGTC KGGATCC ACTAGTA	912663	24
J15	--GAATTC TAATGCATAAACCC TTG TGATTTTCTATATT TGTGTC TATACT ATTT CAGGGATCC ACTAGTA	681573	24
J16	-GAATTC TAATGCTAAGTTGGG TTG TGTTTTCCGTTTTT TGTGTC TATACT ATATCTG GGATCC ACTAGTA	833463	16

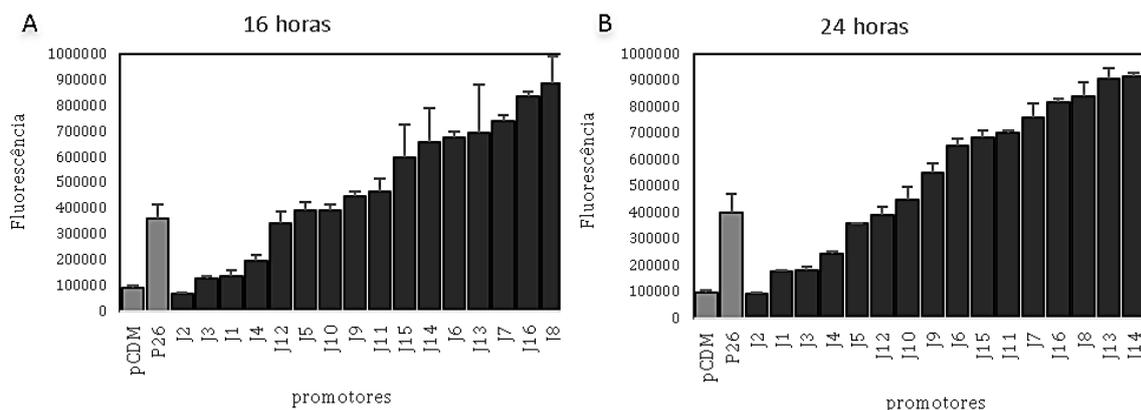
*UMF- unidade de medida de fluorescência

Os plasmídeos dos diferentes clones recombinantes apresentaram sequências idênticas nas regiões de reconhecimento das endonucleases de restrição como o esperado, *EcoRI*, *BamHI* e *SpeI*. A região - 10 e - 10 estendida também foram conservadas, o espaçamento de 17 bases

determinado mudou em alguns clones, e o consenso de nucleotídeos do **-35 (TTG)** se manteve em quase todos os promotores, pode-se notar que em termos de tamanho dos promotores poucos nucleotídeos foram acrescentados ou retirados. Os nucleotídeos que foram randomizados apresentaram sequencias diferentes suficientes para mudar a força do promotor e o tempo de melhor produção.

O J2 apresentou a expressão de GFP muito similar à do controle pCDM, ao analisar a sequência da J2 foi identificado a falta da sequência consenso -35 isso pode ter influenciado na força de expressão, resultando na expressão mais baixa desse clone. O J3 apresentou um espaçamento de 15 bases entre os elementos -10 e -35. Os promotores J8, J13, J14 e J16 apresentaram quantidades elevadas do nucleotídeo T na região espaçadora, e foram os promotores que apresentaram as maiores expressões de GFP nos experimentos.

Figura 11: Gráficos de fluorescência em função dos clones nos tempos 16 e 24 horas. Em ordem crescente, destacando os clones que expressaram pouco e muito GFP nos tempos de coleta de 16 e 24 horas.



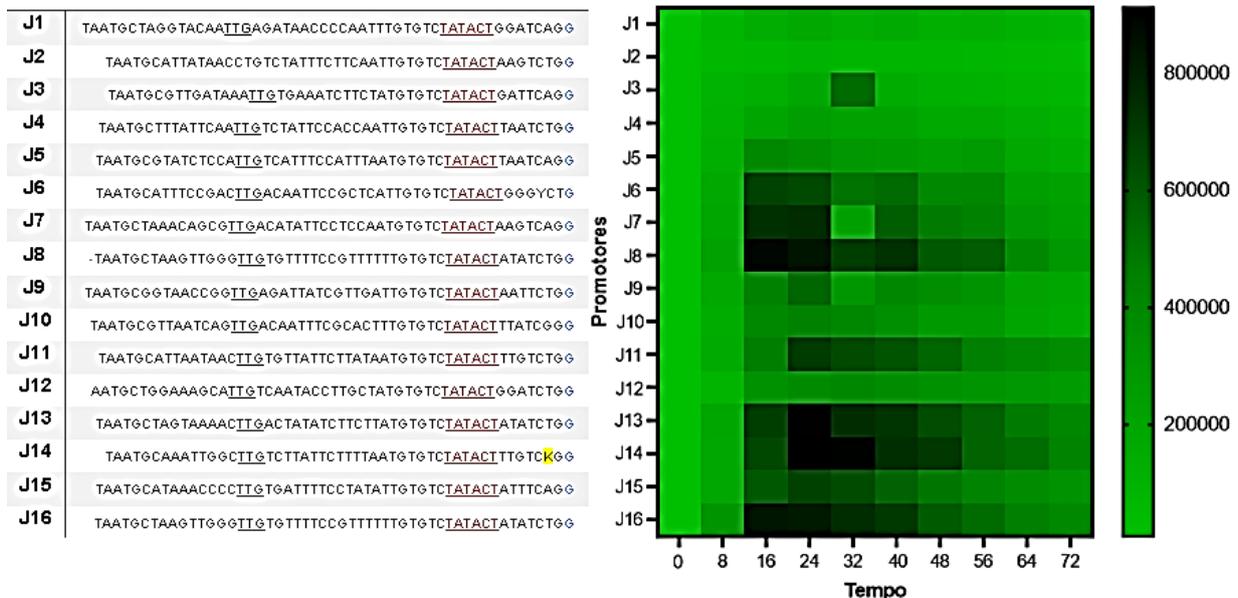
Na figura 13 os clones foram colocados em ordem crescente após os controles, para destacar a produção de fluorescência fraca, média e forte. Os clones J2, J3, J1 e J5 tiveram uma expressão fraca de fluorescência o J5 produziu mais em 16 horas, porém diminuiu em 24 horas, o J12 se manteve estável nesses dois tempos. O J12, J10, J9, J6 e J15 foram destacados como produção média, e mostram maior expressão com 24 horas. O J11, J7, J16, J8, J13 e J14 foram os clones que apresentaram expressões mais elevadas.

Na figura 14 está representada a sequência dos promotores J1 ao J16, contendo o elemento -10, o elemento -10 estendido, e sublinhado o -35. Ao lado está o gráfico de calor dos promotores em função do tempo. Do J1 ao J5 nota-se uma baixa produção de GFP com exceção de J3 que teve um pico em 32 horas, essa medida aparece em todas as réplicas desse clone. O clone J3 apresenta nível médio de expressão, mas com baixíssima expressão na fase

exponencial, apresenta na região -35 a sequência TTGTGA e apenas 15 nucleotídeos na região espaçadora.

Do J6 ao J8 e J11 houve uma expressão mais elevada, o J10 e J12 apresentaram um nível médio de expressão e do J13 e J16 houveram picos de expressão em 16, 24 e 32 horas.

Figura 14: Sequência de promotores e gráfico de calor dos promotores em função do tempo. A figura de calor apresentada ao lado do promotor, representando a escala de produção de fluorescência ao longo do tempo



Curiosamente dois clones J8 e J16 que apresentaram padrões de crescimento e de fluorescência muito parecidos em todas as análises, e com os resultados do sequenciamento mostraram ter a mesma sequência sintética clonada, ou seja, são os mesmos promotores. Embora a probabilidade desse evento ocorrer seja baixa esse é um evento possível.

No gráfico de calor pode-se observar que em dois clones: J7 e J9 há diminuição da fluorescência em 32 horas em relação à de 24 horas que volta a aumentar em 40 horas de cultivo, isso pode indicar uma atividade bimodal de funcionamento dos promotores, ou seja funcionaria mais intensamente na fase exponencial acumulando GFP, cessaria a atividade por um certo tempo e então voltaria a funcionar um uma fase estacionária mais tardia.

Durante as análises, foram escolhidos os clones crescidos inicialmente na placa de *petri* e nas medidas de fluorescência iniciais e foram selecionados clones com expressão: fraca do J1 ao J5, média do J6 ao J10 e forte do J11 ao J16. Após as análises finais, em triplicatas e o sequenciamento, pode-se observar alterações nas sequências dos promotores dos clones que mais produziam GFP.

Essa estratégia aqui apresentada, pode ser aplicada em outros microrganismos, podendo gerar promotores de diversas forças para se escolher a melhor estratégia para expressão de uma determinada proteína de interesse industrial, sem necessidade de se adicionar qualquer substância indutora.

Para estudos mais completos e precisos do funcionamento desses promotores de fase exponencial e estacionária, há necessidade de estudar o nível de expressão do mRNA da GFP, por exemplo por RT-qPCR. Uma outra alternativa, que tem sido muito utilizada recentemente, seria estudar a expressão de derivados instáveis de GFP, pois o fato de não acumularem fica mais fácil correlacionar nível de transcrição com o nível de GFP no interior das células. Pesquisas com mudança de temperatura, oxigenação e pH podem dar resultados também interessantes, além da clonagem com outra proteína para analisar a quantidade de expressão.

Dando continuidade a esse trabalho, um outro ponto que necessita ser estudado ainda seria verificar, especialmente nas células com promotores mais fortes, se GFP tem sido acumulada também na forma insolúvel com conformação 3D errada e, portanto, incapaz de fluorescer.

Conforme descrito nos trabalhos de Wölke, 2012, o promotor P26 programa a super expressão de GFP em *E. coli* enquanto os promotores descritos nesse trabalho como os J13, J14 e J16 programam mais que o dobro da expressão de GFP que o P26. Resultados de melhoria da expressão com os novos promotores mostram resultados análogos também em relação ao trabalho de Miskisch et al, 2005. Esses resultados indicam que estudos mais aprofundados sobre o funcionamento desses promotores, além de poder contribuir para um maior conhecimento dessa área da biologia molecular, poderão ser a base para a construção de vetores de expressão para *E. coli* ainda mais fortes que os utilizados atualmente, de o que seria grande utilidade para o setor industrial.

7 CONCLUSÃO

Com inspiração em pesquisas anteriores sobre fatores sigma 70 e 38 em *E. coli* e baseando-se em síntese química de desoxiligonucleotídeos com degeneração, concebeu-se uma estratégia para construção de uma biblioteca de promotores funcionais tanto na fase exponencial como na fase estacionária de crescimento da bactéria. A estratégia funcionou bem

e foi possível clonar diversos promotores com capacidades variadas de expressar o gene repórter da GFP, desde baixos níveis até super-expressão.

Estudos mais aprofundados do funcionamento destes promotores poderão agregar conhecimento a esta área da biologia molecular e os promotores fortes sintetizados e clonados nesse trabalho poderão ser de grande utilidade para expressão de altos níveis de proteínas recombinantes para fins industriais.

8 REFERENCIAS BIBLIOGRAFICAS

ALBERTS, B.; JOHNSON, A. & WALTER, P. **Biologia Molecular da Célula**. 6^a Ed., Artmed Editora, Porto Alegre, 2017.

ALPER, Hal et al. Tuning genetic control through promoter engineering. **Proceedings of the National Academy of Sciences**, v. 102, n. 36, p. 12678-12683, 2005.

ASTOLFI-FILHO, S. et al. **Noções básicas de tecnologia do DNA recombinante**. Universidade do Amazonas, 1996.

BERLEC, Aleš; ŠTRUKELJ, Borut. Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells. **Journal of industrial microbiology & biotechnology**, v. 40, n. 3-4, p. 257-274, 2013.

BROWNING, D. F.; BUSBY, S. J. W. The regulation of bacterial transcription initiation. **Nature Reviews Microbiology**, v. 2, n. 1, p. 57-65, 2004.

CANDEIAS, J. A. N. A engenharia genética. **Revista de saúde pública**, v. 25, n. 1, p. 3-10, 1991.

CASES, I., de LORENZO, V. Promoters in the environment: transcriptional regulation in its natural context. **Nature Reviews Microbiology**, 3(2), 105–118, 2005. doi:10.1038/nrmicro1084

CHANDRANGSU, Pete; HELMANN, John D. Sigma Factors in Gene Expression. **eLS**, 2014.

COHEN, S. N. DNA cloning: a personal view after 40 years. **Proceedings of the National Academy of Sciences**, v. 110, n. 39, p. 15521-15529, 2013.

DAVIS, Joseph H.; RUBIN, Adam J.; SAUER, Robert T. Design, construction and characterization of a set of insulated bacterial promoters. **Nucleic acids research**, v. 39, n. 3, p. 1131-1141, 2011.

DE MEY, Marjan et al. Construction and model-based analysis of a promoter library for *E. coli*: an indispensable tool for metabolic engineering. **BMC biotechnology**, v. 7, n. 1, p. 1-14, 2007.

EBRIGHT, R. H. RNA Polymerase: Structural Similarities Between Bacterial RNA Polymerase and Eukaryotic RNA Polymerase II. **Journal of Molecular Biology**, 304(5), 687–698, 2000. doi:10.1006/jmbi.2000.4309

FEKLISTOV, Andrey et al. Bacterial sigma factors: a historical, structural, and genomic perspective. **Annual review of microbiology**, v. 68, p. 357-376, 2014.

GRIFFITHS, A. J. F.; MILLER, J. H.; SUZUKI, D. T.; LEWONTIN, R. C.; GELBART, W. M.; WESSLER, S. R. **Introdução à genética**. 9ª edição. Rio de Janeiro, (RJ): Ed. Guanabara Koogan, 2009.

GRUET, Antoine; LONGHI, Sonia; BIGNON, Christophe. One-step generation of error-prone PCR libraries using Gateway® technology. **Microbial cell factories**, v. 11, n. 1, p. 14, 2012.

HAUGEN, S. P.; ROSS, W.; GOURSE, R. L. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. **Nature Reviews Microbiology**, v. 6, n. 7, p. 507-519, 2008.

HELMANN, J. D. Sigma Factors in Gene Expression. In: (Ed.). **eLS**: John Wiley & Sons, Ltd, 2001. ISBN 9780470015902.

HENGGE, Regine. Proteolysis of σ^S (RpoS) and the general stress response in *Escherichia coli*. **Research in microbiology**, v. 160, n. 9, p. 667-676, 2009.

HENRIQUE, Priscila Moraes et al. Molecular characterization of enterococci harboring genotype and phenotype incongruence related to glycopeptide resistance isolated in Brazilian hospitals. **Memórias do Instituto Oswaldo Cruz**, v. 103, n. 3, p. 301-305, 2008.

JENSEN, Peter Ruhdal; HAMMER, Karin. Artificial promoters for metabolic optimization. **Biotechnology and bioengineering**, v. 58, n. 2-3, p. 191-195, 1998.

JISHAGE, M. ISHIHAMA, A. A stationary phase protein in *Escherichia coli* with binding activity to the major σ subunit of RNA polymerase. **Proceedings of the National Academy of Sciences**, v. 95, n. 9, p. 4953-4958, 1998.

JONASSON, Per et al. Genetic design for facilitated production and recovery of recombinant proteins in *Escherichia coli*. **Biotechnology and applied biochemistry**, v. 35, n. 2, p. 91-105, 2002.

JONES, George H. Integrative, xylE-based promoter probe vectors for use in *Streptomyces*. **Plasmid**, v. 65, n. 3, p. 219-225, 2011.

KAGIYA, Go et al. Generation of a strong promoter for *Escherichia coli* from eukaryotic genome DNA. **Journal of biotechnology**, v. 115, n. 3, p. 239-248, 2005.

KAMIONKA, M. Engineering of therapeutic proteins production in *Escherichia coli*. *Current Pharmaceutical Biotechnology*, v. 12, n. 2, p. 268-274, 2011.

KAUR, Jashandeep; KUMAR, Arbind; KAUR, Jagdeep. Estratégias para otimização da expressão proteica heteróloga em *E. coli*: Bloqueios e reforços. **International Journal of Biological Macromolecules**, v. 106, p. 803-822, 2018.

LEE, J. BORUKHOV, S. Bacterial RNA polymerase-DNA interaction—the driving force of gene expression and the target for drug action. **Frontiers in molecular biosciences**, v. 3, p. 73, 2016.

LONETTO, M.; GRIBSKOV, M.; GROSS, C. A. The sigma 70 family: sequence conservation and evolutionary relationships. **Journal of Bacteriology**, v. 174, n. 12, p. 3843–3849, 1992.

MIKSCH, G., BETTENWORTH, F., FRIEHS, K. e FLASCHEL, E. A sequência a montante da sequência de consenso -10 modula a força e o tempo de indução dos promotores da fase estacionária em *Escherichia coli*. **Applied Microbiology and Biotechnology**, 69 (3), 312-320, 2005. doi: 10.1007 / s00253-005-0016-8

MITCHELL, Jennie E. et al. The *Escherichia coli* regulator of sigma 70 protein, Rsd, can up-regulate some stress-dependent promoters by sequestering sigma 70. **Journal of bacteriology**, v. 189, n. 9, p. 3489-3495, 2007.81

MOREIRA D. S., PINTO K. R., ASTOLFI M. C. T., ASTOLFI-FILHO S. A brief history of bacterial promoters: from sequence and regulatory mechanism to biotechnology application. **Scientia Amazonia**, v. 7, n.3, CB17-CB37, 2018.

MULLIS, Kary et al. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. In: **Cold Spring Harbor symposia on quantitative biology**. Cold Spring Harbor Laboratory Press, 1986. p. 263-273.

MURAKAMI, K. S. Structural biology of bacterial RNA polymerase. **Biomolecules**, v. 5, n. 2, p. 848-864, 2015.

MURAKAMI, K. S.; DARST, S. A. Bacterial RNA polymerases: The whole story. Current Opinion in **Structural Biology**, v. 13, n. 1, p. 31–39, 2003.

MURAKAMI, K. Structural Biology of Bacterial RNA Polymerase. **Biomolecules**, v. 5, n. 2, p. 848–864, 2015.

MURAKAMI, Katsuhiko S. et al. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. **Science**, v. 296, n. 5571, p. 1285-1290, 2002.

ÖSTERBERG, Sofia; PESO-SANTOS, Teresa del; SHINGLER, Victoria. Regulation of alternative sigma factor use. **Annual review of microbiology**, v. 65, p. 37-55, 2011.

PATIKOGLU, G. A., WESTBLADE, L. F., CAMPBELL, E. A., LAMOUR, V., LANE, W. J., DARST, S. A. Crystal Structure of the Escherichia coli Regulator of σ_{70} , Rsd, in Complex with σ_{70} Domain 4. **Journal of Molecular Biology**, 372(3), 649–659, 2007. doi:10.1016/j.jmb.2007.06.0

PEREIRA, W. E. S. Expressão heteróloga de uma protease coagulante produzida por *Thermomucor indicae-seudaticae* N31 em *Escherichia coli* e *Pichia pastoris*. 2019.

ROSS, W., ERNST, A., GOURSE, R. L. Fine structure of E. coli RNA polymerase-promoter interactions: α subunit binding to the UP element minor groove. **Genes & development**, v. 15, n. 5, p. 491-506, 2001.

ROSS, W., GOSINK, K., SALOMON, J., IGARASHI, K., ZOU, C., ISHIHAMA, A., GOURSE, R. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science*, 262(5138), 1407–1413, 1993. doi:10.1126/science.8248780

RUD, Ida et al. A synthetic promoter library for constitutive gene expression in *Lactobacillus plantarum*. **Microbiology**, v. 152, n. 4, p. 1011-1019, 2006.

SANTOS, Christine Nicole S.; STEPHANOPOULOS, Gregory. Combinatorial engineering of microbes for optimizing cellular phenotype. **Current opinion in chemical biology**, v. 12, n. 2, p. 168-176, 2008.

SANTOS-BENEIT, A. M., MOLLINEDO, F. Expression of genes involved in initiation, regulation, and execution of apoptosis in human neutrophils and during neutrophil differentiation of HL-60 cells. **Journal of leukocyte biology**, v. 67, n. 5, p. 712-724, 2000.

SCHOLZ, Holger C. et al. S-layer variation in *Bacillus stearothermophilus* PV72 is based on DNA rearrangements between the chromosome and the naturally occurring megaplasmids. **Journal of bacteriology**, v. 183, n. 5, p. 1672-1679, 2001.

SHETTY, Reshma P.; ENDY, Drew; KNIGHT, Thomas F. Engineering BioBrick vectors from BioBrick parts. **Journal of biological engineering**, v. 2, n. 1, p. 5, 2008.

SOLEM, Christian; JENSEN, Peter Ruhdal. Modulation of gene expression made easy. **Applied and Environmental Microbiology**, v. 68, n. 5, p. 2397-2403, 2002.

STRYJEWSKA, A., KIEPURA, K., LIBROWSKI, T., & LOCHYŃSKI, S. Biotechnology and genetic engineering in the new drug development. Part I. DNA technology and recombinant proteins. **Pharmacological Reports**, 65(5), 1075–1085, 2013. doi:10.1016/s1734-1140(13)71466-x

VASSYLYEV, Dmitry G. et al. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. **Nature**, v. 417, n. 6890, p. 712-719, 2002.

WANG, P., QIN, W., XU, J., YAN, Y., TIAN, J., WU, N., & YAO, B. Enhancing the soluble expression of an amylase in *Escherichia coli* by the mutations related to its domain interactions. **Protein expression and purification**, v. 120, p. 35-41, 2016.

WATSON, J. D., BAKER, T.A., BELL, S. P., LOSICK, G. R., LEVINE, M. **Biologia Molecular do Gene**. Editora Artmed, 7ª edição, Porto Alegre, 2015.

WÖHLKE, Jonathan Luiz et al. Construção de vetores de expressão para *Escherichia coli* baseados em promotores ativos na fase estacionária de crescimento. 2012.

ZHANG, G., CAMPBELL, E. A., MINAKHIN, L., RICHTER, C., SEVERINOV, K., & DARST, S. A. CRYSTAL Structure of *Thermus aquaticus* Core RNA Polymerase at 3.3 Å Resolution. **Cell**, 98(6), 811–824, 1999. doi:10.1016/s0092-8674(00)81515-9

ZHANG, Y., FENG, Y., CHATTERJEE, S., TUSKE, S., HO, M. X., ARNOLD, E., & EBRIGHT, R. H. Structural Basis of Transcription Initiation. **Science**, 338(6110), 1076–1080, 2012. doi:10.1126/science.1227786