

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

CONSTRUÇÃO DE NOVOS PROMOTORES FUNCIONAIS EM
Escherichia coli
POR MEIO DE SÍNTESE QUÍMICA DE DNA RANDOMIZADA

DIEGO DA SILVA MOREIRA

MANAUS

2016

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

DIEGO DA SILVA MOREIRA

CONSTRUÇÃO DE NOVOS PROMOTORES FUNCIONAIS EM
Escherichia coli
POR MEIO DE SÍNTESE QUÍMICA DE DNA RANDOMIZADA

Dissertação apresentada ao programa de Pós-graduação Multi-institucional em Biotecnologia da Universidade Federal do Amazonas, como requisito para a obtenção do título de Mestre em Biotecnologia, área de concentração Engenharia Genética.

Orientador: Prof. Dr. Spartaco Astolfi Filho

MANAUS

2016

Ficha Catalográfica

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

M838c Moreira, Diego da Silva
Construção de novos promotores funcionais em *Escherichia coli*
por meio de síntese química de DNA randomizada / Diego da Silva
Moreira. 2017
97 f.: il. color; 31 cm.

Orientador: Dr. Spartaco Astolfi Filho
Dissertação (Mestrado em Biotecnologia) - Universidade Federal
do Amazonas.

1. Síntese química de DNA. 2. Transcrição. 3. Regulação da
expressão gênica. 4. gene cat. 5. promotores. I. Astolfi Filho, Dr.
Spartaco II. Universidade Federal do Amazonas III. Título

“E sabemos que todas as coisas contribuem juntamente para o bem daqueles que amam a Deus, daqueles que são chamados segundo o seu propósito”.

Rm 8 :28

Agradecimentos

A Deus Pai Todo Poderoso no qual me mostra apenas uma ínfima parte do seu amor por mim através deste trabalho. Deus eu Te Amo e sei que teu amor por mim é Eterno.

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

Dedico também aos meus pais Olival Moreira e Dílce Monteiro, ao meu Tio Raimundo Moreira, meus irmãos Jardeson e Daiana, minha sobrinha Yanka e toda minha família por me entender nos momentos de estudo e por toda força e incentivo me concedidos até aqui.

A toda equipe Top do CAM, Drs. Enedina e Edson por todo o conhecimento teórico/prático de laboratório e pela

amizade, a Dona Elza pela simplicidade e pelas técnicas e macetes de Esterilização e também pelos cafezinhos.

A toda noíva da 8ª congregação, minha família, ao Pastor Edmar e aos meus líderes por todo o discernimento e sabedoria me concedido na minha vida espiritual. A todos pelos momentos de descontração, como os das farofas, filmes em casas, ... pela união. #SomosUm

A minha amiga Ruan Gama e Leíssandra Castelo, pela amizade profunda, pela imensa colaboração, pelos conselhos e momentos maravilhosos e pelo grau de liberdade intenso nas brincadeiras e demais colegas da pós-graduação.

Pro doído Arthur Ribeiro. Aos meus vizinhos pelos momentos de dominó, futebol de salão, risadas e brincadeiras e pelos farofões no meio da madrugada de vez em quando.

A FAPEAM pelo Formento do projeto e pela bolsa concedida.

Agradeço imensamente também a muitos que ajudaram e contribuíram diretamente e indiretamente.

SUMÁRIO

LISTA DE TABELAS	9
LISTA DE QUADROS.....	10
LISTA DE FIGURAS.....	11
LISTA DE ABREVIATURAS, SÍMBOLOS E UNIDADES.....	13
RESUMO.....	15
ABSTRACT	16
1 CONTEXTUALIZAÇÃO	17
2 REVISÃO BIBLIOGRÁFICA	19
2.1 O processo de transcrição	19
2.2 RNA polimerase em procariotos	21
2.3 Promotores procarióticos	23
2.3.1 Hexâmeros ou regiões -35 e -10.....	25
2.3.2 A região espaçadora entre as regiões -35 e -10	27
2.3.3 Sequência <i>Up elements</i>	28
2.4 Terminadores de transcrição	29
2.5 Shine-Dalgarno.....	31
2.6 Regulação da Expressão Gênica em procariotos baseado no operon <i>lac</i>	33
2.7 Expressão de genes heterólogos em hospedeiros bacterianos	37
2.8 Expressão de genes heterólogos em <i>E. coli</i> por meio do operon <i>lac</i>	39
2.9 Vetores	43
2.9.1 Vetores de seleção de promotores (VSP) ou Caça-Promotores.....	45
2.10 Construção de biblioteca de promotores por síntese química para produção de proteínas heterólogas	47
3 OBJETIVOS	50
3.1 Objetivo geral:.....	50
3.2 Objetivos específicos:.....	50
4 MATERIAL E MÉTODOS	51
4.1 Desenho Experimental.....	51
4.2 MATERIAL	51
4.2.1 Linhagem Bacteriana.....	51
4.2.2 Vetor	52

4.2.3 Oligonucleotídeos	52
4.2.4 Reagentes.....	53
4.2.5 Meios de cultura	53
4.2.6 Enzimas	53
4.2.7 Tampões e soluções.....	54
4.3 MÉTODOS	55
4.3.1 Construção dos promotores sintéticos.....	55
4.3.2 Preparação do vetor caça-promotor.....	58
4.3.3 Clonagem do promotor sintético.....	59
4.3.4 Clonagem no vetor caça-promotor pKK232-8	62
4.3.5 Análise dos clones recombinantes.....	63
4.3.6 Sequenciamento.....	64
4.3.7 Bioinformática	64
4.3.8 Escolha de outros clones para novas análises	64
4.3.9 Crescimento e indução em meio LB líquido	65
5 RESULTADOS E DISCUSSÃO	67
5.1 Construção e clonagem de promotores sintéticos	67
5.1.1 Construção da biblioteca de promotores sintéticos.....	67
5.1.2 Preenchimento dos oligonucleotídeos com a enzima <i>Klenow</i>	68
5.1.3 Digestão do vetor e dos promotores sintéticos com enzimas de restrição e purificação.....	69
5.1.4 Clonagem dos promotores no vetor pKK232-8	69
5.2 Análise dos clones recombinantes na ausência ou presença do indutor IPTG	70
5.3 Sequenciamento dos promotores dos clones recombinantes.....	72
5.4 Análise das sequências dos promotores sintéticos	73
5.5 Análise da resistência a cloranfenicol pelos clones com diferentes sequências promotoras.	78
5.5.1 Análise da resistência do cloranfenicol em meio sólido	78
5.5.2 Análise da resistência ao cloranfenicol em meio líquido.	81
6 CONCLUSÕES.....	86
BIBLIOGRAFIA DE REFERÊNCIA	87

LISTA DE TABELAS

Tabela 1: Algumas sequências nucleotídicas de promotores utilizados em vetores de expressão. 26

Tabela 2: Mutações que alteraram o espaçamento. 27

Tabela 3: Enzimas de restrição utilizadas neste trabalho. 53

LISTA DE QUADROS

Quadro 1: Oligonucleotídeos utilizados neste trabalho.	52
Quadro 2: Meios de cultura utilizados neste trabalho.	53
Quadro 3: Tampões e soluções utilizados neste trabalho.	54
Quadro 4: Sistemas e componentes para o preenchimento com a enzima Klenow.	57
Quadro 5: Sistemas das digestões do vetor e do oligonucleotídeo sintético.	60
Quadro 6: Clones diferentes em concentrações variadas de cloranfenicol com ou sem IPTG.	72
Quadro 7: Sequência Up element.	76
Quadro 8: Região -35.	76
Quadro 9: Região -10.	76
Quadro 10: Sequência depois da região -10 e antes do ponto de início de transcrição.	77

LISTA DE FIGURAS

Figura 1: Etapas da Transcrição.....	20
Figura 2: Esquema da holoenzima RNA polimerase no sítio inicial de transcrição	23
Figura 3: Frequência dos nucleotídeos presentes na região consenso e no espaçador de promotores de E. coli.....	24
Figura 4: Estrutura de um terminador intrínseco de transcrição	31
Figura 5: Sítios de ligação ao ribossomo	33
Figura 6: Desenho esquemático de um mRNA policistrônico.....	34
Figura 7: O operon lac é regulado pelo repressor LacI	354
Figura 8: Esquema dos sítios operadores no operon lac.....	365
Figura 9: Atividade de CAP/cAMP	376
Figura 10: Representação esquemática do plasmídeo caça-promotor pKK232-8 para obtenção de sequências promotoras.....	476
Figura 11: Fluxograma dos procedimentos experimentais realizados.	51
Figura 12: Desenho dos oligonucleotídeos sintéticos.....	56
Figura 13: Representação esquemática de anelamento dos oligonucleotídeos através da região espaçadora.....	56
Figura 14: Esquema da elongação dos oligonucleotídeos anelados.	57
Figura 15: Enlarmayers (inóculo) contendo 50 mL de meio LB líquido com cloranfenicol 250 µg/mL na presença e na ausência do indutor IPTG.....	665
Figura 16: Análise em gel de eletroforese do anelamento dos oligonucleotídeos randômicos.	67
Figura 17: Análise eletroforética do processo de obtenção dos promotores dupla fita.	68

Figura 18: Análise eletroforética do vetor e dos promotores sintéticos purificados..	69
Figura 19: Fotografia dos clones em meio de cultivo com cloranfenicol na ausência e presença do indutor sintético IPTG.	71
Figura 20: Desenho da sequência promotora sintética.	73
Figura 21: Figura das sequências promotoras sintéticas.	75
Figura 22: Fotografia dos clones recombinantes crescidos em placa de Petri contendo cloranfenicol (50 µg/mL e 500 µg/mL) com ou sem IPTG.	80
Figura 23: Gráfico que mostra o crescimento do clone 3.A na presença e ausência de IPTG.	82
Figura 24: Gráfico que mostra o crescimento do clone 4.A na presença e ausência de IPTG.	85

LISTA DE ABREVIATURAS, SÍMBOLOS E UNIDADES

α	Alfa
β	Beta
β'	Beta linha
σ	Sigma
σ^{70}	sigma pertencente à família setenta
ω	Ômega
(ρ)	Rho
ATP	Adenosina trifosfato
cAMP	monofosfato cíclico de adenosina
CAP	<i>Catabolite Activator Protein</i> ou proteína ativadora do catabolismo
CAT	Gene que codifica a proteína cloranfenicol acetil-transferase
Clf	Cloranfenicol
C-terminal	Carboxi-terminal
DNA	Ácido desoxirribonucléico
DNTp	Desoxirribonucleotídeos trifosfato
<i>E. coli</i>	<i>Escherichia coli</i>
et al.	e colaboradores
fMet	<i>N</i> -formil metionina
<i>Rv</i>	Reverse
IFs	Fatores iniciadores
IPTG	isopropil- β -D-tiogalactopiranosídeo
kb	Kilobases

kDa	Quilodalton
kV	Voltz
<i>lac</i>	Lactose
<i>Lac I</i>	Repressor <i>lac I</i>
mRNA	RNA mensageiro
mL	Militro
mM	Milimolar
N-terminal	Amino-terminal
ng	Nanograma
O.D.	Densidade óptica
pb	Pares de base
PCR	<i>Polimerase chain reaction</i> ou Reação em cadeia da polimerase
RNA	Ácido ribonucleico
RNApol	RNA polimerase
rpm	Rotações por minuto
SD	Shine-Dalgarno
tRNA	RNA transportador
µg	Micrograma
µL	Microlitro

RESUMO

O interesse por ferramentas de manipulação genética que aumentem a produção de proteínas heterólogas para fins biotecnológicos, especialmente proteínas terapêuticas e enzimas industriais, aumentou significativamente em nível mundial. Para a produção de proteínas heterólogas a escolha do sistema hospedeiro & vetor de expressão é fundamental. Para a construção de uma biblioteca de promotores reguláveis e funcionais em *E. coli*, dois desoxi-oligonucleotídeos com regiões degeneradas foram produzidos por meio de síntese química randômica de DNA. Para a obtenção dos promotores dupla-fita, as sequências do operador *Lac* foram aneladas e submetidas por tratamento com a enzima *Klenow*. Para a clonagem e análise da funcionalidade dos novos promotores, as bordas dos desoxi-oligonucleotídeos dupla-fita e do vetor foram digeridas com *Bam*HI e *Hind*III e após foram purificados. Os promotores sintéticos foram a seguir ligados ao vetor caça-promotores pKK232-8 de forma que ficassem em posição adequada para expressar o gene repórter **CAT**. Os plasmídeos recombinantes foram então introduzidos em *E. coli* DH5 α F'I^q e as células transformantes foram obtidas e classificadas de acordo com o tamanho da colônia em placas de Petri contendo diferentes concentrações de cloranfenicol na presença e ausência do indutor IPTG. Após as análises das colônias foram escolhidos 42 clones recombinantes, seus DNAs plasmidiais extraídos e as sequências dos seus promotores determinadas completamente. A seguir 6 plasmídeos recombinantes foram reintroduzidos em *E. coli* DH5 α F'I^q e os clones selecionados em cloranfenicol, com ou sem IPTG. Destes 6 clones, 4 foram escolhidos para analisar o crescimento em meio líquido com cloranfenicol na presença e ausência do indutor IPTG. Todos os clones recombinantes em *E. coli* demonstraram capacidade de crescer em meio líquido, sendo que 3 deles cresceram melhor com IPTG indicando que seus promotores estão sendo regulados. O outro clone não teve seu crescimento afetado pelo indutor, confirmado pela deficiência da sequência do operador *Lac*. Os resultados mostram que a metodologia utilizada para gerar e clonar novos promotores funciona e que os novos promotores têm potencialidade de serem utilizados para o desenvolvimento de novos, potentes e regulados vetores de expressão.

Palavras chaves: Síntese Química de DNA, Transcrição, Regulação da Expressão Gênica, CAT, Promotores.

ABSTRACT

The interest in genetic manipulation tools increase production of heterologous proteins for biotechnological purposes, especially therapeutic proteins, industrial enzymes and significantly increased worldwide. For the production of heterologous proteins choice of expression vector & host system is essential. For the construction of a library of adjustable and promoters functional in *E. coli*, two-deoxy oligonucleotides degenerate regions were produced by random chemical synthesis of DNA. To obtain the double-stranded promoters, *Lac* operator sequences were looped and submitted by treatment with *Klenow* enzyme. For cloning and analysis of the functionality of the novel promoters, the edges of the double-stranded deoxy-oligonucleotide and vector were digested with *Bam* HI and *Hind* III and following were purified. Synthetic promoters were then linked to the promoter vector pKK232-8 hunting so stay in proper position for expressing the CAT reporter gene. The recombinant plasmids were then introduced into *E. coli* DH5 α F'Iq and transformants cells were obtained and classified according to the size of the colony on Petri plates containing different concentrations of chloramphenicol in the presence and absence of the inducer IPTG. After analysis of the colonies were chosen 42 recombinant clones, their DNA extracted plamidiais and the sequences of the promoters completely certain. 6 The following recombinant plasmids were reintroduced into *E. coli* DH5 α F'Iq and clones selected on chloramphenicol, with or without IPTG. Of these six clones, 4 were selected to examine the growth in liquid medium with chloramphenicol in the presence and absence of the inducer IPTG. All recombinant clones in *E. coli* demonstrated ability to grow in liquid medium, 3 of them grew better with IPTG, indicating that their promoters are regulated. The other clone did not have their growth affected by the inductor, confirmed by deficiency following the *Lac* operator. The results show that the methodology used to generate and clone novel promoters running and new promoters have the potential of being used for the development of new, powerful and regulated expression vectors.

Key words: Synthesis DNA chemistry, Transcription, Gene Expression Regulation, CAT, Promoters.

1 CONTEXTUALIZAÇÃO

Após pouco mais de 40 anos do advento da Tecnologia do DNA Recombinante (Engenharia Genética) pode-se observar o grande impacto dessa metodologia tanto no desenvolvimento da área científica como biotecnológica (COHEN, 2013).

Na área científica essa tecnologia permitiu o isolamento, sequenciamento e estudo de genes de genomas complexos como o humano, como os oncogenes, anti-oncogenes, genes homeóticos, genes de hormônios, antígenos, anticorpos, genes responsáveis por doenças genéticas etc (WALSH, 2010; SANCHEZ-GARCIA, 2015).

Na área biotecnológica, além construção de seres transgênicos para uso direto, a engenharia genética possibilitou a produção de proteínas de interesse industrial em larga escala impactando tanto a produção de enzimas como: amilases, celulasas, lipases, proteases, etc (SACCO, 2013) como a produção de medicamentos biológicos como: imunomoduladores, hormônios, antígenos, vacinas e anticorpos monoclonais (INTERFARMA, 2012; BERLEC & STRUKELJ, 2013).

As primeiras proteínas recombinantes eram produzidas pela simples clonagem de um fragmento de DNA (gene) em um plasmídeo multicópia (vetor) e o DNA recombinante introduzido em uma célula hospedeira, normalmente uma bactéria ou uma levedura (HARTLEY, 2006). A síntese de proteínas recombinantes em altos níveis deve-se ao desenvolvimento de eficientes sistemas de expressão ou expressão/secreção compostos de vetores e suas respectivas hospedeiras (SORENSEN & MORTENSEN, 2005; ROSANO & CECCARELLI, 2014).

Um dos principais componentes de um vetor de expressão é o “cassete de expressão”, um conjunto seletivo de sequências que garantem a expressão de sequências codificadoras de proteínas de interesse. Um “cassete de expressão” é composto principalmente de um **promotor** que sinaliza o início da síntese do mRNA pela RNA polimerase e que na maioria dos vetores utilizados é passível de regulação, uma região de **múltiplos sítios de clonagem** (“polylinker”) para permitir a introdução da mensagem genética a ser expressa e um **terminador de transcrição** – para interromper a síntese do mRNA pela RNA polimerase (GUSAROV & NUDLER, 2001; HIJUM et al. 2009).

A região promotora (promotor) é a sequência responsável por indicar o ponto de início de transcrição do gene a ser expresso e pela regulação desse processo. Em seres procariontes o início da transcrição é a etapa mais relevante do processo de regulação da expressão gênica. Em um “cassete de expressão” é o principal elemento que garante a expressão do transgene em altos níveis e de forma regulada quando desejado (OVERTON, 2014).

Tanto por razões de propriedade industrial e especialmente para obter-se melhores vetores de expressão, a obtenção de novos promotores é a principal etapa. Esses elementos genéticos podem ser obtidos por clonagem diretamente de sequências de DNA de genomas naturais ou podem ser obtidos por síntese química (JULLESSON, 2015). Em ambos os casos pode-se valer de um precioso tipo de instrumento da engenharia genética para seleção e análise desses elementos - os vetores caça-promotores (ALI & MURRELL, 2009).

Sistemas de expressão heterólogos são muitas vezes utilizados para produzir rendimentos mais elevados de proteínas em relação ao hospedeiro natural. Embora haja uma variedade de sistemas de expressão procariontes e eucariontes, o hospedeiro mais comum e preferido para expressão, permanece a bactéria *Escherichia coli* (BANEYX & MUJACIC, 2004).

O presente trabalho refere-se à obtenção de novos e regulados promotores a partir de síntese química de DNA e clonagem e caracterização por meio de utilização de um vetor caça-promotor, para desenvolvimento de novos e eficientes vetores de expressão de transgenes em bactérias, especialmente em *Escherichia coli*.

2 REVISÃO BIBLIOGRÁFICA

2.1 O processo de transcrição

A informação genética contida na molécula de DNA (genes) possui relação com a síntese de proteínas necessárias ao metabolismo celular, na qual são expressos constitutivamente ou em certas condições, na forma de RNA mensageiro para posterior tradução gênica (SCHUMANN & FERREIRA, 2004; LEWIN, 2005). Esse processo de conversão de um molde de DNA para RNA é conhecido como transcrição gênica e possui a RNA polimerase como enzima responsável pela sua execução e, por conseguinte, é um dos processos mais fundamentais da vida (DRENNAN et al. 2012). Todos os organismos vivos codificam pelo menos uma RNA polimerase (RNAPol) sendo que a estrutura do núcleo e muitos elementos atuantes e reguladores são partilhados por RNAs polimerases bacterianas, eucarióticas e por archaea (RUFF et al. 2015).

O nível de transcrição gênica está relacionado com a frequência com que a RNA polimerase inicia o processo, bem como sua afinidade a sequência promotora (HIJUM et al. 2009; SAECKER et al. 2011).

A síntese ocorre primeiramente com uma associação não específica da RNA polimerase com a dupla fita de DNA (interação eletrostática) (FEKLISTOV, 2014). Logo após, ocorre a formação de um complexo mais estável de reconhecimento formado pelo núcleo da enzima RNA polimerase com a região promotora ou sequência conhecida como promotor (GESZVAIN & LANDICK, 2005). A ligação do núcleo da enzima com o fator sigma (σ) ocasiona o reconhecimento do promotor que dá início ao ciclo de transcrição. O núcleo da RNA polimerase interagindo com o fator sigma forma a holoenzima. Somente a holoenzima é capaz de reconhecer a sequência promotora e ocasionar as mudanças conformacionais no DNA necessárias para separar as duas fitas de DNA para formar uma bolha de transcrição e iniciar de fato a transcrição gênica (LEE et al. 2012).

Em todos os organismos, a transcrição por RNAs polimerases dependentes de DNA pode ser dividida em quatro etapas mecanicamente e estruturalmente distintas: reconhecimento da região promotora no DNA, iniciação, alongamento e terminação (Figura 1) (BORUKHOV & NUDLER, 2003; GESZVAIN & LANDICK, 2005; CAMPBELL et al. 2008; RUFF et al. 2015).

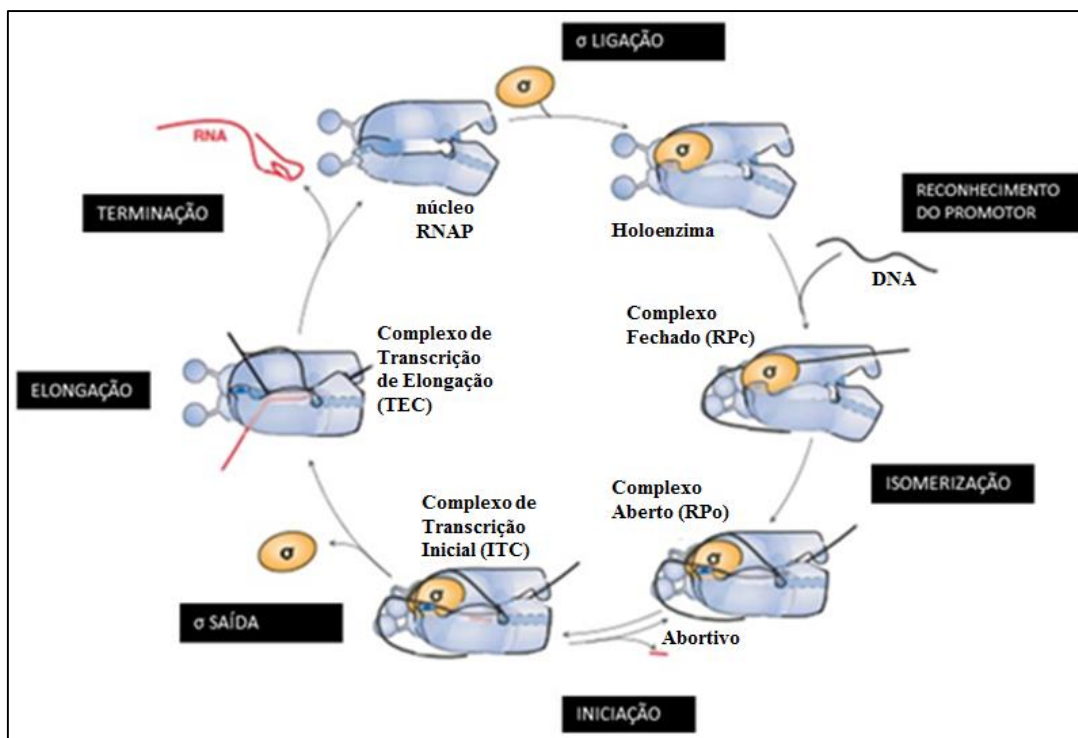


Figura 1: Etapas da Transcrição: A RNA polimerase holoenzima interage com a região promotora no DNA formando o complexo fechado (RP_c). Após o processo de reconhecimento do promotor ocorre uma isomerização em dois ou mais passos para formação do complexo aberto (RP_o), que permite que o molde de DNA encontre o sítio ativo da holoenzima. O complexo aberto é capaz de iniciar a transcrição, mas na maioria dos casos, o promotor mantém-se em um complexo de transcrição inicial (ITC) que sofre rodadas curtas de formação de transcrito e que é capaz de abortar a transcrição. Após a RNAPol deixar o promotor, é formado um complexo de alongamento da transcrição (TEC), onde a subunidade sigma (σ) se dissocia e o complexo segue de maneira processiva e estável. A transcrição termina quando a RNAPol atinge o sinal de término, fazendo com que a fita de DNA e RNA sejam lançadas e liberando a RNAPol para uma nova rodada de transcrição (Fonte: GESZVAIN E LANDICK, 2005).

A ligação específica da holoenzima ao promotor forma um complexo fechado inicial (RP_c) e põe em movimento uma série de grandes mudanças conformacionais em ambas as biomoléculas. Tais mudanças colocam o duplex de DNA no local do sítio ativo da RNA polimerase, que em seguida, abre o DNA para formar uma bolha de transcrição para a formação do complexo aberto (RP_o) (SAECKER et al. 2011; RUFF et al. 2015). Isto é devido a uma isomerização, pois supõe-se que deve resultar em um movimento da cadeia de DNA livre da bolha ao sítio ativo da enzima RNA polimerase de modo que a síntese de RNA possa começar (BROWNING & BUSBY, 2004). RP_o é capaz de iniciar a transcrição, mas na maioria dos casos, o promotor permanece em um complexo de

transcrição inicial (ITC) que pode sofrer paradas abortivas e com isso não iniciar a transcrição (GESZVAIN & LANDICK, 2005; SKANCKE et al. 2015).

O fator σ (sigma) não é necessário para a fase de alongação e é tipicamente liberado do núcleo da RNA polimerase, e apenas esta prossegue polimerizando o RNA (SAECKER et al. 2011). Ao alcançar o final da região a ser transcrita, sinais presentes na sequência do DNA ou do RNA levam ao término da transcrição, fazendo com que o complexo DNA-RNA polimerase e RNA seja desfeito. Tais sequências são chamadas de terminadores de transcrição e são reconhecidas pela RNA polimerase. Este processo pode ocorrer pela ação de fatores proteicos ou pelo simples desequilíbrio das forças que mantêm o híbrido DNA-RNA juntos (ROBERTS et al. 2008; HIJUM et al. 2009).

2.2 RNA polimerase em procariotos

Os procariotos utilizam dois conjuntos distintos de mecanismos para a ativação do início da transcrição, e estes estão focados no promotor ou na RNA polimerase (LEE et al. 2012). A RNA polimerase bacteriana é composta por pelo menos cinco subunidades constantes e pode ser encontrada em duas formas: a) a apoenzima ou núcleo da enzima que possui todas as subunidades exceto a subunidade (ou fator) σ ; e b) a holoenzima que possui as subunidades α , β , β' , ω , e σ , capaz de reconhecer promotores gênicos específicos e iniciar a síntese de RNA (BROWNING & BUSBY, 2004; CASES & DE LORENZO, 2005; FEKLISTOV, 2014).

Dentre as procarióticas, a RNA polimerase de *E. coli* é uma das mais estudadas e a melhor caracterizada (SAECKER et al. 2011). As duas subunidades alfa (α) possuem 40 kDa cada. A subunidade β (beta) possui em torno de 155 kDa e β' (beta linha) 160 kDa. A subunidade ω (ômega) possui 10 kDa e o fator sigma (σ) possui aproximadamente de 25-92 kDa. A subunidade α é formada por dois domínios de aminoácidos independentes unidos por um flexível ligante de 20 aminoácidos: o domínio N-terminal com resíduos de 8 a 235 aminoácidos e o C-terminal composto por resíduos de 249 a 329 aminoácidos. As subunidades β e β' possuem 1,342 e 1,407 aminoácidos. A subunidade ω apresenta 91 resíduos de aminoácidos (MURAKAMI & DARST, 2003; MARQUES, 2012).

Cada subunidade tem um papel específico na ligação ao DNA, na interação entre as subunidades da enzima e no processo catalítico. A parte N-terminal de α está envolvida na formação de dímeros e da ligação de β com β' , e o seu C-terminal está preso por meio de um ligante flexível ao seu N-terminal, é responsável pela interação com o *Up-element* presente a montante de alguns promotores que funciona como ativador transcricional (BROWNING & BUSBY, 2004; GESZVAIN & LANDICK, 2005). A subunidade β contém o sítio catalítico ativo da polimerase e juntamente com β' especifica a interação do σ (sigma) com o promotor e a fita de DNA, e também liga os ribonucleotídeos que irão ser polimerizados em RNA (MARQUES, 2012). O papel da subunidade ω (ômega) atua como uma “chaperona molecular” para o correto enovelamento da subunidade β' e auxilia na montagem estável da RNAPol (SCHUMANN & FERREIRA, 2004).

Como mencionado anteriormente, o fator σ está envolvido no reconhecimento dos promotores, sendo que cada um reconhece um conjunto distinto de genes, sendo que existem, pelos menos, seis tipos diferentes do fator σ em *E. coli*. Cada um desses fatores regula um subconjunto de genes em resposta a diferentes estímulos, como carência nutricional, aumento de temperatura, fase estacionária, estresse oxidativo, estresse osmótico, programa de desenvolvimento, etc (BROWNING & BUSBY, 2004; SAECKER et al. 2011).

Em contraste, todas as espécies de bactérias possuem pelo menos um fator σ primário. Em *E. coli*, o sigma 70 (σ^{70}) (codificado pelo gene *rpoD*), que ativa a transcrição principalmente dos genes constitutivos (*housekeeping*) ou transcrição dos genes de manutenção basal da bactéria é o principal, e por isso, é o fator sigma mais abundante na célula. Em *E. coli* o σ^{70} tem uma maior afinidade para o núcleo da RNA polimerase do que qualquer um dos seus sete fatores σ alternativos (Figura 2) (SACHDEVA et al. 2009; LEE et al. 2012).

Como parte da holoenzima funcional, σ^{70} sigma reconhece um promotor constituído por dois elementos bem conservados da sequência, os hexâmeros -35 e -10, que são separados por uma região espaçadora de 16-19 pb (RUFF et al. 2015). O fator σ^{70} pode ser dividido em quatro domínios proteicos peculiares unidos por ligantes flexíveis. Cada domínio de σ^{70} interage com ambos: RNA polimerase e DNA; os domínios 2 e 4 fazem contatos específicos com as regiões -10 e -35 respectivamente (CAMPBELL et al. 2008), ao passo que o domínio 3 interage com a região -10 estendida presente em alguns promotores. Embora os domínios 2, 3 e 4 são conhecidos por estarem envolvidos no

reconhecimento do promotor, o domínio 1 está envolvido na regulação do complexo aberto (Figura 2) (MURAKAMI & DARST, 2003; RODRIGUE et al. 2006; SACHDEVA et al. 2009).

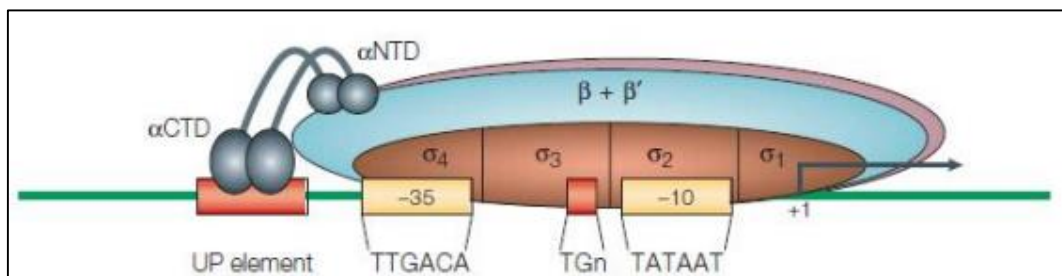


Figura 2: Esquema da holoenzima RNA polimerase no sítio inicial de transcrição. A fita verde sendo reconhecida pela RNAPol representa o DNA. As cinco subunidades da holoenzima estão descritas na figura 2, assim como os quatro domínios proteicos da subunidade σ . As sequências promotoras -10 (TATAAT) e -35 (TTGACA) são reconhecidos respectivamente pelos domínios 2 e 4 da subunidade σ (Fonte: BROWNING & BUSBY, 2004).

2.3 Promotores procarióticos

Os promotores controlam a transcrição de todos os genes e em procariotos existe um dado promotor para um conjunto de genes, sendo que os RNAs que são transcritos possuem informações variados de produtos gênicos que na maioria dos casos, estão relacionados metabolicamente (LI et al. 2015). Entretanto, embora cada conjunto de genes possua um promotor, nem todos são idênticos entre si, pois nas células bacterianas a RNA polimerase é confrontada com um conjunto com cerca de pelo menos 2.000 sequências promotoras diferentes (BROWNING & BUSBY, 2004).

Diversos trabalhos de codificação e análise de promotores de procariotos, através de programas de computador utilizando algoritmos (HAWLEY & MCCLURE, 1983; HARLEY & REYNOLDS, 1987; BURDEN et al. 2005; BEREZHNOY & SHCKORBATOV, 2005) reforçam um conceito existente de uma “sequência promotora consenso” que apresentam identidade entre algumas bases do promotor e que tais sequências são altamente conservadas (HARLEY E REYNOLDS, 1987; BEREZHNOY & SHCKORBATOV, 2005). Como consequência, houve a descoberta de muitas sequências promotoras dentro do genoma de *E. coli* reconhecidas pela holoenzima (RNAPol- σ^{70}) contendo sequências similares de DNA e que foi descoberta a importância

funcional destas sequências em muitas investigações (BROWNING & BUSBY, 2004). Dado que se estabeleceu duas regiões promotoras (PRIBNOW, 1975; BEREZHNOY & SHCKORBATOV, 2005) formadas respectivamente pelos hexâmeros – 35 (TTGACA) e -10 (TATAAT) e geralmente separadas entre si por 17 pares de bases, sendo a transcrição iniciada em uma purina localizada a 7 pares de bases após a região -10. Essas sequências tornaram-se conhecidas como duas regiões de identidade (-35 e -10) sendo localizadas nos procaríotos a aproximadamente 35 a 10 pares de bases a montante do início da transcrição. E em alguns casos, uma região de “*Up element*” que se localiza a montante da região -35 (HAWLEY & MCCLURE, 1983; BEREZHNOY & SHCKORBATOV, 2005; FEKLISTOV, 2014; RUFF et al. 2015).

Em média, promotores de bactérias contêm entre oito e nove bases iguais às do consenso, sendo raro ocorrer promotores com todas as sequências iguais às sequências consenso (nas 12 posições) (Figura 3). Ou seja, foi elucidado que existem variações nas sequências nucleotídicas de promotores e que estas variações afetam a ligação da RNA polimerase aos promotores, e ainda que existem regiões mais importantes do que outras para o funcionamento do promotor, como a região -10, que onde é altamente conservada e essencial para o reconhecimento pela holoenzima, independente da região -35 estar presente (HAWLEY & MCCLURE, 1983; BROWNING & BUSBY, 2004; BEREZHNOY & SHCKORBATOV, 2005; FEKLISTOV, 2014).

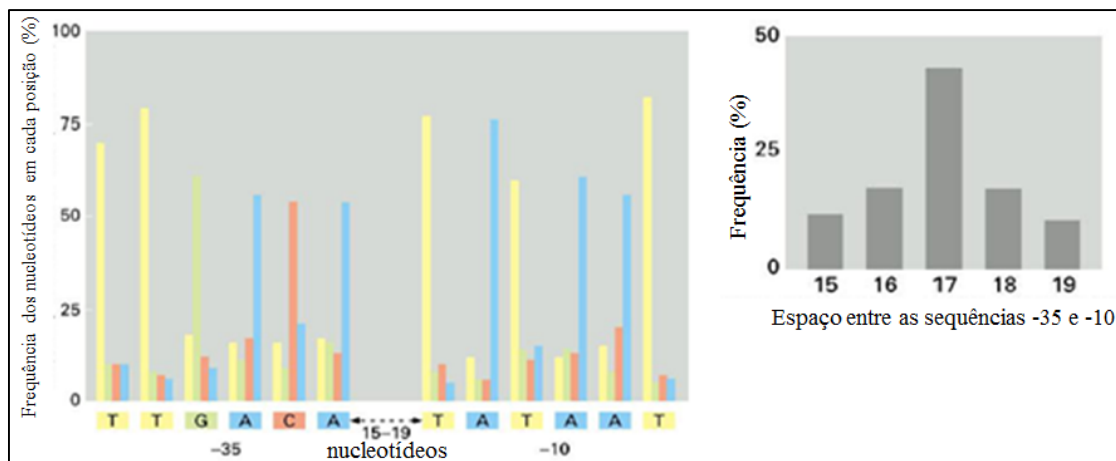


Figura 3: Frequência dos nucleotídeos presentes na região consenso e no espaçador de promotores de *E. coli*.

O verdadeiro significado biológico dessas sequências conservadas, confirma-se através dos relatos de mutações que diminuem a homologia de sua sequência com as sequências consenso, o inverso ocorrendo para mutações que aumentam a sua atividade ou melhoria da expressão/regulação gênica em hospedeiros (ANDRIANANTOANDRO et al. 2006; TIMOTHY, 2010; BLOUNT et al. 2012).

Sabe-se que as sequências promotoras que apresentam elementos mais parecidos com as da sequência consenso funcionam mais eficientemente, tendem a ser mais ativos e possuem uma maior atividade de ligação com a RNA polimerase do que promotores cujas sequências nucleotídicas difere da sequência consenso preferencial de sigma 70 (TTGACA – 17 pb –TATAAT) (SCHUMANN & FERREIRA, 2004). Isso implica que, promotores com maior afinidade pela RNA polimerase, resulta em uma maior frequência na transcrição dos genes. Ao contrário, promotores com baixa afinidade pela RNA polimerase ocasiona em uma menor transcrição de genes (PURVES et al. 2002). Assim, observa-se que a partir dessas sequências específicas de identidade presentes nos promotores, podem ser introduzidas mutações.

2.3.1 Hexâmeros ou regiões -35 e -10

As regiões de hexâmeros -35 e -10 desempenham um papel muito importante na seletividade da transcrição e na frequência da iniciação. O hexâmero ou região -10 desempenha um papel fundamental de modo a formar o complexo aberto, enquanto a região ou hexâmetro -35 serve como um alvo para ativadores de transcrição (CAMPBELL et al. 2008; RUFF et al. 2015). A região -35 possui a sequência de consenso 5' – TTGACA - 3' com o -35 = T; -34 = T; e -33 G sendo estas bases as mais altamente conservadas. Esta região também pode interagir com o núcleo da RNA polimerase para mediar o recrutamento de fatores de alongamento (RUFF et al. 2015). A região -10 apresenta uma região consenso (5' - TATAAT - 3') onde todos os pares de bases são ricos em A e T. As bases -11 = A; e -7 = T são as bases mais altamente conservadas do hexâmetro, sendo indispensáveis para a cinética de formação de iniciação do complexo aberto, caso esta região não seja rica em A e T (HEYDUK & HEYDUK 2014; RUFF et al. 2015).

Muitas mutações ocorrem dentro das regiões -35 e -10, em geral, muitas dessas mutações podem fugir da região consensual e conseqüentemente podem diminuir ou aumentar a frequência de iniciação.

Um promotor historicamente importante, caso à parte, é o promotor do operon *lac* de Jacob e Monod de 1961. O promotor *lac* não é tão semelhante a sequência de nucleotídeos presente na região -35 e -10 conhecida como consenso. Possui uma base na região -35 e duas na região -10 que difere da sequência consenso. E, a região espaçadora entre estas duas regiões é de 18 pares de bases (SCHUMANN & FERREIRA, 2004) (Tabela 1).

Tabela 1: Algumas sequências nucleotídicas de promotores utilizados em vetores de expressão.

Promotor	Região -35	Região espaçadora	Região -10
P_{lac}	TTtACA	18 pb	TATgtT
P_{lacUV5}	TTtACA	18 pb	TATAAT
P_{trp}	TTGACA	17 pb	TtaAcT
P_{tac}	TTGACA	16 pb	TATAAT
λP_L	TTGACA	17 pb	gATAcT
λP_R	TTGACt	17 pb	gATAAT
Consenso	TTGACA	17 pb	TATAAT

Os nucleotídeos presentes na sequência consenso são demonstrados em letras em negrito. Na sequência de promotores, os nucleotídeos idênticos à sequência consenso são mostrados em letras maiúsculas e os que não são idênticos, são mostrados em nucleotídeos minúsculos (Fonte: SCHUMANN & FERREIRA, 2004).

Entretanto, considerando que promotores que apresentam as sequências mais próximas às do consenso são mais fortes, construíram um outro promotor, o *placUV5* a partir do *plac* (tabela 1). O promotor *placUV5* difere em apenas uma base da região -35 consenso, e apresenta a região -10 idêntica à da região consenso. Com essa mudança, a atividade promotora aumentou cerca de 2,5 vezes (SCHUMANN & FERREIRA, 2004).

Além das regiões conservadas, a atividade promotora está também ligada com outras regiões do promotor bem como a região espaçadora.

2.3.2 A região espaçadora entre as regiões -35 e -10

A região espaçadora é um segmento de DNA que não apresenta nenhuma sequência de consenso de nucleotídeos e está localizada entre as regiões -35 e -10. Sendo que, há autores que afirmam que esta região é a mais importante que a sequência como um todo para a determinação do promotor, isso porque esse espaço seria o mais preferencial do fator sigma 70 (MITCHELL et al. 2003; KIRYU et al. 2004; SHIMADA et al. 2014). O comprimento do espaçador mais comum para promotores σ^{70} é de 17 pares de bases e mutações que alterem o tamanho do espaçador de 17 pares de bases existentes entre as sequências -35 e -10 também modificam o nível da transcrição dirigida pelo promotor mutado (KIRYU et al. 2004; SCHUMANN & FERREIRA, 2004; RUFF et al. 2015). O espaçamento entre as regiões consensuais apresenta grande influência na determinação da força e na eficácia do promotor (HOOK-BARNARD & HINTON, 2009). Em geral, o desvio a partir do comprimento do espaçador ótimo leva uma redução na atividade do promotor (MITCHELL et al. 2003).

Para *E. coli* o espaçamento usual entre as duas regiões -35 e -10 de promotores σ^{70} é de 17 pares de bases. Quando ocorre um aumento ou decréscimo do comprimento dessa região espaçadora, a força do promotor é diminuída (Tabela 2) (KIRYU et al. 2004).

Tabela 2: Mutações que alteraram o espaçamento.

Promotor	Mutação (pb)	Novo espaço	Fenótipo
ampC	+1	17	15 x >
lacp ^s	-1	17	10 x >
p22 ant	-1	16	Fraco
tyr T	-1	15	50 x <
lacP ⁺	+2	20	10 x <

Promotores com mutações do tipo inserção (+) ou deleção (-). Todos os fenótipos correspondem a mudança em nível de expressão *in vivo*, com exceção do lacp^s o qual foi caracterizado *in vitro*. O sinal de maior (>) ou menor (<) indica uma diminuição ou um aumento no nível da expressão (Fonte: HAWLEY & MCCLURE, 1983).

Casos em que a região espaçadora influencia na transcrição foi quando construíram o promotor pTAC, tendo a sequência consenso -35 igual à do promotor *trp* (pTRP) e a sequência consenso -10 igual ao do promotor pLACUV5 e com espaçamento

de 16 pb (DOMBROSKI et al. 1996). A partir disto, Brosius e Storella, 1985, construíram dois novos promotores se baseando a partir do pTAC, o pTRC com espaçamento de 17 pares de bases e o pTIC possuindo um espaçamento de 18 pares de bases. Após a construção, verificaram que o nível de expressão do gene CAT sob o controle dos dois novos promotores (pTRC e o pTIC), foi mais alta em comparação ao pTAC em até 90% maior.

Há outras interações menos definidas em comparação com a RNA polimerase (holoenzima) e a sequência espaçadora, dos quais podem aumentar a atividade transcricional em alguns promotores, como as sequências *Up elements*.

2.3.3 Sequências *Up elements*

Muitos promotores apresentam uma quarta região, o *Up element*, localizado a montante da região -35. O *Up element*, consiste em sequências ricas em A e T que interagem com o domínio C-terminal da subunidade α -CTD da RNA polimerase, aumentando a força do promotor (MENG et al. 2001; SCHUMANN & FERREIRA, 2004).

Pelo menos um terço dos promotores estáveis de RNA (rRNA e tRNA) e 4% de promotores de RNA mensageiros podem conter sequências *Up elements* (JI MA & HOWE, 2015). Há dois sub-sítios dentro do elemento UP: o proximal e o distal; sendo que qualquer um deles pode funcionar de forma independente, embora o sub-sítio proximal pode exercer maior efeito sobre a atividade do promotor (ESTREM et al. 1999; JI MA & HOWE, 2015).

Sequências consenso derivada previamente a partir de promotores de *E. coli* contêm regiões de identidade nos hexâmeros -10 e -35, mas não apresentam sequências conservadas a montante da região -35, o que sugere que os *Up elements* não são cruciais para a transcrição em todos os promotores. A razão em relação há funcionalidade destas sequências serem ricas em A e T não foram estabelecidas, e o número de sequências a montante da região -35 para interagir com α é insuficiente para derivar uma sequência consenso *Up element* bem definida (ESTREM et al. 1999; JI MA & HOWE, 2015). A melhor sequência *Up element* caracterizada está presente na *E. coli rrnB* promotor P1, em que os determinantes da sequência estão localizados entre as posições -40 e -60 em

relação ao local de início da transcrição, no qual aumenta a atividade do promotor em 70 vezes *in vivo* (MENG et al. 2001; SCHUMANN & FERREIRA, 2004).

Entretanto, desenvolveu-se uma seleção *in vitro* seguida por uma *in vivo* no âmbito de identificar sequências *Up elements* a partir de uma população aleatória de 31 sequências que conferem maiores aumentos no promotor P1 de *rrnB* que foram caracterizados por seus efeitos sobre a transcrição e interações com α *in vitro*, e derivado uma sequência consenso integral do *Up element* e dos seus sublocais (ESTREM et al. 1999; JI MA & HOWE, 2015). O grau de homologia com a sequência consenso deduzida (-59 NNAAA [A / T] [A / T] T [A / T] TTTTNNAAANN -38, onde N é qualquer nucleotídeo,) mostrou-se correlacionar com sequências naturais de *Up elements*, fusionados a jusante do promotor *lac* em procaríotos. A sequência *Up element* otimizada (*Up element* consenso) além de ser ricas em bases A e T, é cinco vezes mais eficaz do que o tipo selvagem no promotor P1 em *rrnB* (ESTREM et al. 1999; JI MA, 2015).

Estes resultados sugerem que as sequências *Up elements* a jusante dos promotores regulados, podem ser altamente ativas (MENG et al. 2001). Essa sequência a montante da região -35 pode ser considerada por alguns autores com uma sequência promotora propriamente dita, pois em contato com outros promotores como o *lacUV5*, estimula a transcrição. Para a produção de proteínas recombinantes, esta sequência em geral é muito útil para sistemas de expressão em alto nível, pois tem a capacidade de potencializar a transcrição de um gene quando ligado a um promotor heterólogo (SCHUMANN & FERREIRA, 2004).

2.4 Terminadores de transcrição

Um terminador de transcrição possui como papel principal permitir o término da síntese de transcrição (LEAVITT & ALPER, 2015). Um terminador de transcrição é colocado sempre depois da sequência que codifica para o gene alvo, que aumenta a estabilidade do mRNA, evitando que transcrições desnecessárias a partir de outros promotores presentes no próprio plasmídeo ou genoma ocorram (SORENSEN & MORTENSEN, 2005; CIAMPI, 2006). Assim, a “força” de um terminador pode ser avaliada com base na capacidade de leitura, o impacto sobre a meia vida do mRNA, a

estrutura do 3' UTR, e até mesmo sobre influência da atividade latente do promotor (LEAVITT & ALPER, 2015).

Em bactérias existem dois tipos de terminação: Rho-dependente e os independentes de fatores de Rho ou intrínseca (YACHIE et al. 2006; SINGH, 2014; ASSCHE et al. 2015).

Rho (ρ) é uma proteína essencial em *E. coli*, embora no seu genoma possua poucos terminadores dependentes dela, seu mecanismo é geral. Na bactéria *E. coli*, a terminação dependente de Rho está bem elucidada. (RICHARDSON, 1993; 2002; CIAMPI, 2006).

Terminadores Rho-dependentes diferem entre si, podendo ocorrer em sítios no qual a sequência de um terminador intrínseco seria incompatível, possuem uma organização mais complexa e alguns aspectos mecânicos que ainda não são totalmente compreendidos. Não possuem características comuns como sequência repetidas de base U, não podem ser definidos por uma sequência consenso, e por isso não são utilizados em sistemas de expressão de proteínas recombinantes (SCHUMANN & FERREIRA, 2004; CIAMPI, 2006, EPSHTEIN, 2010).

A terminação intrínseca, também conhecida como terminação Rho-independente, está presente em cerca da metade dos sítios de terminação das unidades de transcrição bacterianas. A terminação Rho-independente é formado por uma sequência palindrômica no DNA, que quando é transcrita no RNA fita simples causa um pareamento interno, gerando uma estrutura em forma de haste e alça (ERMOLAEVA et al. 2000; KINGSFORD, 2007). No entanto, o terminador possui uma sequência rica em T em seguida à sequência palindrômica rica em C e G, que no RNA será transcrita em U (KINGSFORD, 2007). A formação da haste e alça faz com que o pareamento DNA-RNA seja desfeito na bolha de transcrição e, como o pareamento A e U são ligações fracas ou instáveis, ocorre a dissociação do complexo de transcrição. Ambas as características são necessárias para a terminação (Figura 4) (SCHUMANN & FERREIRA, 2004; CIAMPI, 2006; YACHIE et al. 2006; SANTANGELO et al. 2009; MITRA, 2009). Terminadores intrínsecos apresentam características em comuns, mas variam em sequência, eficiência de rescisão e mecanismo (PETERS et al. 2011).

No genoma de *E. coli* existem aproximadamente 1100 sequências deste tipo. E esforços para o desenvolvimento de algoritmos e programas de análise de sequência

podem prever com relativa confiança a posição de todos os sítios terminadores intrínsecos em um genoma bacteriano (LESNIK, 2001; KINGSFORD, 2007; MITRA, 2009).

Uma pesquisa recente de 582 terminadores em *E. coli* por Chen e colaboradores, 2013 identificaram 39 terminadores fortes que reduzem a expressão de genes a junsante em mais de 50 vezes. Um conjunto de pontos em comum entre estes terminadores “fortes” levou a aprimorar parâmetros que podem informar tentativas futuras com a engenharia genética para a criação de terminadores sintéticos com forças desejáveis em *E. coli*, mostrando assim a importância de terminadores no controle de nível de transcrição (CHEN et al. 2013).

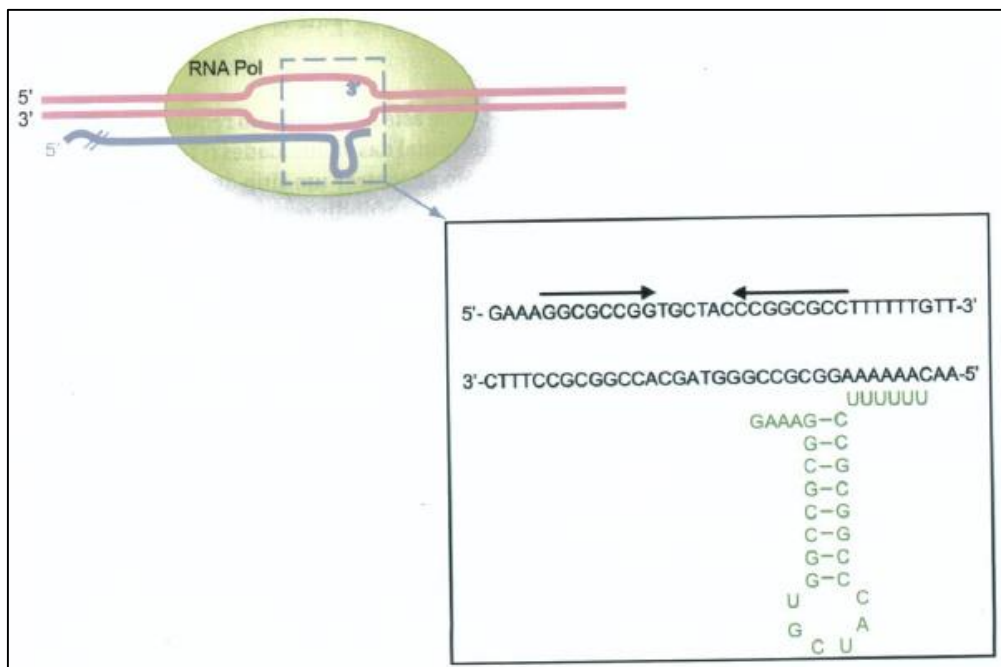


Figura 4: Estrutura de um terminador intrínseco de transcrição. O terminador intrínseco é composto por uma sequência palindrômica no DNA (setas), que é seguida por uma sequência rica em Timinas. Quando é transcrita em RNA fita simples, esta sequência forma uma estrutura de haste e alça, seguida de uma sequência pareada de Adeninas e Uracilas, o que faz com que o pareamento de DNA-RNA seja desfeito, ocorrendo assim a dissociação do complexo de transcrição (Fonte: MARQUES, 2012).

2.5 Shine-Dalgarno

Em bactérias, a biossíntese de proteínas é normalmente considerada o passo limitante da velocidade da tradução e um processo altamente regulado nos níveis de

expressão global de um gene (MA, 2002; SCHARFF, 2011). Procariotos usam um mecanismo relativamente simples da iniciação da tradução que envolvem o ribossomo, o aminoacilado e o iniciador formilado (fMet – tRNA_f^{Met}), o mRNA e um pequeno número de fatores de iniciação (IFs), como IF1, IF2 e IF3 para bactérias (GUALERZI, 2015).

As proteínas se iniciam, na maioria das vezes por um aminoácido metionina, mas no caso das bactérias é utilizado um aminoácido que apresenta uma modificação, A N-formil metionina (fMet) (JIN et al. 2006). Embora para os demais aminoácidos existam pelo menos dois códons (denominados de redundância do código), para o aminoácido metionina existe apenas um códon, AUG. Isso caracteriza um ponto de referência ao ribossomo, que pode reconhecer o códon no mRNA (JIN et al. 2006). Em eucariotos, o ribossomo reconhece o primeiro AUG do mRNA como o códon iniciador da proteína, mas em bactérias isto não é possível, porque o mRNA é policistrônico, e os ribossomos tem que reconhecer vários AUG iniciadores na mensagem (ALLEN & FRANK, 2007; MALYS & MCCARTHY, 2011).

Portanto, os códons iniciadores têm que ser identificados por mais algum sinal para a localização correta dos ribossomos. Este sinal está no sítio de ligação do ribossomo (ou RBS, do inglês *Ribosome Binding Site*) ou Shine-Dalgarno (SD) do mRNA, cerca de 5 a 10 nucleotídeos do a montante do códon iniciador e geralmente contém, a montante a partir de AUG, parte ou uma totalidade de uma sequência de várias purinas (UAAGGAGGU) (SHINE & DALGARNO, 1974; JIN et al. 2006). Em procariotos, a sequência de SD é complementar a uma sequência 3' rica em pirimidina (a anti-SD ou ASD região) do RNA 16S do ribossomo, embora sua sequência pode variar entre as espécies ou mesmo entre genes de um mesmo organismo (CHEN et al. 1994; GUALERZI, 2015).

Aparentemente, a sequência SD não é essencial para a ligação do mRNA aos ribossomos durante a tradução, embora a eficiência de iniciação é diminuída na sua ausência. Não é claro como é promovida iniciação em tais casos. Entretanto, a força de interação da SD, um espaço longo excessivo ou curto entre o SD e a escolha do códon de início pode acarretar um efeito direto sobre a produção de proteínas, assim como na eficiência da iniciação da tradução, em até 600 vezes, sendo AUG o mais eficaz do que os outros códons (MA, 2002; MUTALIK et al. 2013; MIRZADEH et al. 2015).

Muitos pesquisadores descobriram que existe um espaçamento ótimo entre o SD e o códon de início AUG em uma variedade de mRNAs, que variam de 5 a 13 nucleotídeos a partir do códon de início demonstrando confusão na determinação de uma região consenso de códons. Tais variações de diferentes sequências SDs pode ter um impacto significativo sobre a expressão de um gene (Figura 5) (PASOTTI & ZUCCA, 2014; OMOTAJÓ et al. 2015).

UCAGGGUGGUGAAUGUGAAACCAGU	<i>E. coli lacI</i>
GAUGGAGUGAAACGAUGGCGAUUGC	<i>E. coli araB</i>
UGUUUGAGGAUAAUAUGGCGAACGG	<i>C. crescentus cspC</i>
CUACUGAGGGUACGAUGGUCGAACU	<i>C. crescentus sdhB</i>
AGAAGGGGUCUAAAUGGCCGCUAA	<i>C. crescentus groEL</i>
AAGGGGGAGCAGACAUGAGUGAUGA	<i>B. subtilis katE</i>
UUCGAGGUGAGCUC AUGAAUUUUCU	<i>B. subtilis crp</i>
GGAGGGUUUACAUC AUGGCAAUACU	<i>B. subtilis galE</i>

Figura 5: Sítios de ligação ao ribossomo (sequência de Shine-Dalgarno). Algumas sequências de Shine-Dalgarno (em roxo) e sua relação ao códon iniciador (em laranja) (Fonte: MARQUES, 2012).

2.6 Regulação da Expressão Gênica em procariotos baseado no operon *lac*

Os procariotos vivem em muitos ambientes, onde os níveis de nutrientes, toxinas, acidez, temperatura, osmolaridade, umidade, temperatura, pH e outras condições não são propícios a sua sobrevivência e que podem mudar de maneira rápida e inesperadamente (BICH et al. 2015). Para tal sobrevivência, existe um monitoramento constante das condições intracelular e extracelular e ao mesmo tempo ocorrendo a adaptação da estrutura, fisiologia e comportamento celular frente a todas a estas alterações ambientais fazendo com que os níveis de proteína sejam ajustados de acordo com as necessidades celular (BICH et al. 2015; ASSCHE et al. 2015).

Apenas uma parte de genes estimados são transcritos em um dado momento. A bactéria para economizar na síntese proteica transcreve alguns produtos gênicos em larga escala, enquanto outros são necessários em menor escala (BOSDRiesz et al. 2015). O princípio que norteia a regulação gênica é o princípio da economia de energia (MARQUES, 2012).

A transcrição é o ponto-chave na expressão gênica em bactérias, sendo o principal ponto de regulação e os procariotos possuem um mecanismo relativamente simples para a regulação da expressão gênica. Muitos genes bacterianos que tem funções correlatas são aglomerados e estão sob o controle de um único promotor e geralmente são transcritos juntos. Esses genes em geral que são transcritos juntos é chamado de operon (HIJUM et al. 2009; GEROSA & SAUER, 2011).

O operon regula a expressão dos genes estruturais controlando a transcrição. Os mRNAs oriundos da transcrição são denominados policistrônicos (Figura 6).

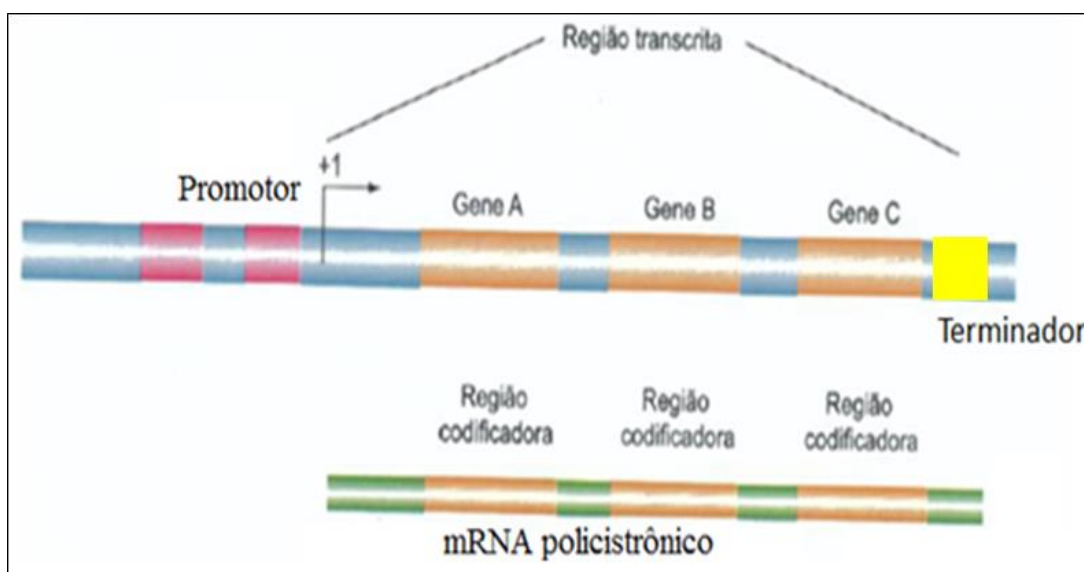


Figura 6: Desenho esquemático de um mRNA policistrônico. A figura mostra um óperon constituído pelos genes A, B e C. O promotor define o sítio de início de transcrição (+1), a partir do qual existirá um mRNA que compreende as regiões codificadoras (uma de cada gene) (Fonte: MARQUES, 2012).

Os primeiros estudos de expressão gênica em procariotos foram definidos através do metabolismo da lactose em *E. coli*. O operon *lac* foi o primeiro sistema de regulação gênica, descrito por François Jacob e Jacques Monod em 1961, tendo como sugestão a ação de um repressor controlando a transcrição em resposta a um indutor (lactose) (LEWIS, 2013). Devido a sua simplicidade, este sistema acabou se tornando o paradigma para o estudo da regulação de outros genes em bactérias para a biologia de sistemas atuais (BICH et al. 2015).

O operon *lac* consiste em uma região regulatória constituída como promotor (P), a partir de onde a RNA polimerase inicia a transcrição, e um operador (O), onde um

repressor se liga (PENUMETCHA et al. 2010). Essas regiões antecedem as três seqüências estruturais dos genes *lacZ*, *lacY* e *lacA*, respectivamente, que codificam as proteínas responsáveis para a utilização do açúcar lactose apenas quando a mesma está disponível. O gene *lacZ* codifica a enzima β -galactosidase, que cliva o dissacarídeo lactose em seus monômeros galactose e glicose, o gene *lacY* codifica uma permease, que realiza o transporte de lactose para dentro da célula, e o gene *lacA* codifica uma transacetilase, que acetila a lactose, não sendo essencial para o catabolismo da lactose (LEONARD et al. 2015). Na ausência de lactose, o operon *lac* é controlado por uma proteína tetramérica, o repressor *LacI*, uma proteína repressora codificada pelo gene *lacI* e que se encontra a montante do *lacZ* e regula a transcrição do operon *lac* impedindo a transcrição do operon através da repressão do operador (BICH et al. 2015).

Quando a glicose está presente no meio, ela desencadeia um processo chamado de repressão catabólica, ou seja, a lactose não é utilizada. Quando a concentração de glicose é baixa ou ausente, a lactose se liga na molécula repressora do *LacI*, que por sua vez muda de conformação de tal maneira que ele se desliga do operador *lacO* e permite que a RNA polimerase faça a transcrição dos três genes necessários para a utilização da lactose: *lacZ*, *lacY* e *lacA* (Figura 7) (BICH et al. 2015).

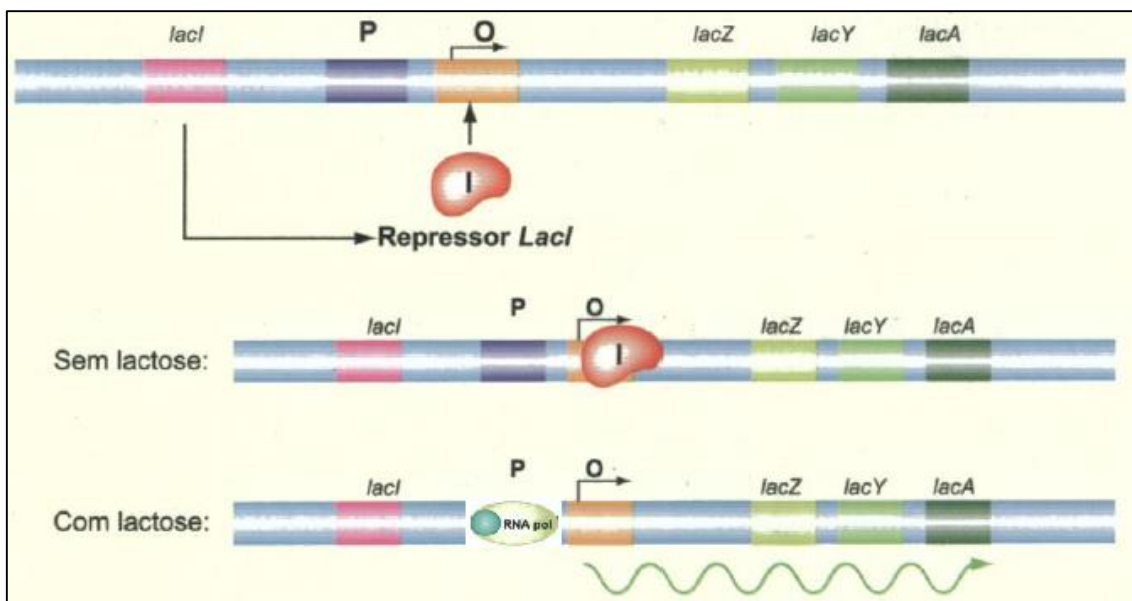


Figura 7: O operon *lac* é regulado pelo repressor *LacI*. O repressor *LacI* é codificado pelo gene *lacI*, que se liga ao operador impedindo que ocorra a transcrição pela RNA polimerase. Na ausência de lactose, o repressor *LacI* está ligado ao operador. Na presença de lactose, o repressor *LacI* desliga-se do operador e assim ocorre a transcrição (Fonte: MARQUES, 2012).

Atualmente sabe-se que existem três sítios operadores no operon *lac*, O1, O2 e O3. A molécula repressora pode se ligar a todos, porém com diferentes afinidades (Figura 8). O operador principal, O1, é o sítio que se sobrepõe o promotor, e é o mais importante e o mais forte. Porém, os sítios auxiliares O2 (localizado na posição +400) e O3 (localizado na posição -80), embora distantes e que são operadores mais fracos, também possuem um importante papel na repressão do operon *lac* (MARQUES, 2012; SEMSEY et al. 2013). A questão levantada é que o tetrâmero *LacI* se ligue a O1 e em outro sítio operador, ou seja, cada dímero se liga a um sítio, ocasionando a consequente formação de uma alça no DNA causando a repressão da transcrição. Este modelo de alça foi confirmado tanto *in vitro* e *in vivo* (CAMSUND et al. 2014). A ligação simultânea de O1 e O2 causa uma forte repressão, por causar a formação de uma alça mais favorável no DNA devido aos tetrâmeros de *LacI* ainda continuarem a serem produzidos. No entanto, quando O1 e O3 estão ligados, não é só o operon *lac* reprimido, mas também a produção de *LacI* é impedida (Figura 8) (PENUMETCHA et al. 2010).

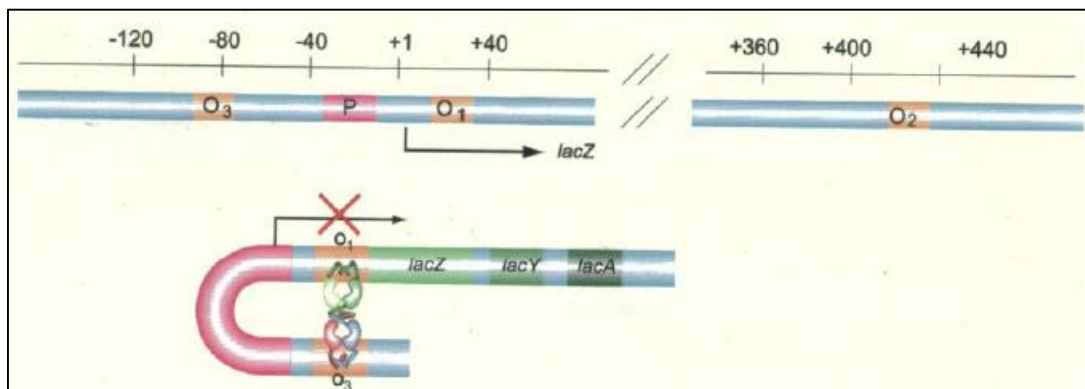


Figura 8: Esquema dos sítios operadores no operon *lac*. Existem três sítios operadores no operon *lac*, O1, O2 e O3, sendo que o repressor pode se ligar a todos, mas com afinidades diferentes. A ligação a mais de um sítio simultaneamente causa uma distorção da hélice do DNA que inibe a migração da RNA polimerase (Fonte: MARQUES, 2012).

Quando a bactéria *E. coli* é cultivada no meio contendo glicose e lactose, a expressão do operon *lac* permanece inativada, ou seja, enquanto houver glicose disponível, a bactéria irá consumi-la, e não utilizará a lactose devido a glicose ser a fonte de carbono principal, coerente com o fato de que o açúcar é mais energeticamente eficiente a ser metabolizado (LEWIS, 2013). Entretanto, somente quando houver o esgotamento da glicose é que o operon será ativado novamente e as enzimas para utilização da lactose

serão produzidas novamente. Desta forma exclui uma dispendiosa utilização de ambos, quando não é necessário (BICH et al. 2015).

Essa regulação do operon *lac*, é denominada de repressão catabólica, e é resultado dos baixos níveis de **cAMP** (monofosfato cíclico de adenosina), um pequeno sinalizador intracelular da ausência de glicose (BICH et al. 2015). Quando o nível de glicose no meio cai, os níveis de **cAMP** aumentam no citoplasma e os nucleotídeos cíclicos interagem a uma proteína chamada de **CAP** (do inglês *Catabolite Activator Protein*). A **CAP** ligada a **cAMP**, na forma de um dímero, é um ativador de transcrição, que posteriormente irá reconhecer uma sequência específica nas regiões regulatórias de vários genes, inclusive do operon *lac* (LEWIS, 2013). O sítio de ligação de **CAP/cAMP** se localiza a montante ao sítio de reconhecimento pela RNA polimerase, e a presença deste ativador ligado juntamente ao DNA favorece a interação da RNA polimerase com o promotor, ativando a transcrição e a aumentando em torno de 50 vezes a sua atividade (Figura 9) (MARQUES, 2012; BICH et al. 2015).

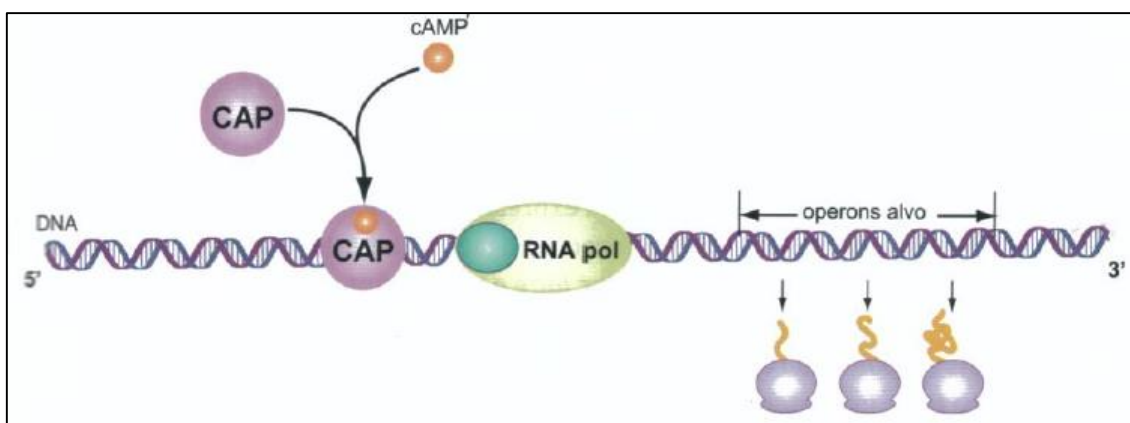


Figura 9: Atividade de CAP/cAMP. A proteína CAP ligada com a **cAMP** reconhecendo uma sequência específica na região regulatória do gene e consequentemente estimulando a ligação da RNA polimerase ao promotor, ocasionado a atividade da transcrição (Fonte: MARQUES, 2012).

2.7 Expressão de genes heterólogos em hospedeiros bacterianos

A seleção de um sistema de expressão para a produção de elevados níveis da proteína recombinante depende de muitos fatores, como: as condições de crescimento que são fatores cruciais para a obtenção de altos níveis de expressão, expressão intracelular e extracelular, modificações translacionais e a atividade biológica da proteína de interesse,

assim como questões regulatórias na produção de proteínas terapêuticas. A seleção de um sistema de expressão particular necessita de uma análise de custo em termos de processo, projeto e outras considerações econômicas (OVERTON, 2014).

A expressão de proteínas heterólogas funcionais em hospedeiros tem sido a grande questão da biotecnologia moderna (GUSTAFSSON et al. 2004; BERLEC & STRUKELJ, 2013). A síntese da proteína recombinante deve ser encarada como um processo estranho à célula, ou seja, não faz parte de seu conjunto metabólico natural e geralmente acaba causando problemas à célula hospedeira que reage de muitas maneiras antagônicas, comprometendo a obtenção da proteína de interesse. Por isso, deve-se considerar a escolha cuidadosa dos vetores compatíveis com a proteína que se quer expressar, normalmente, plasmídeos, manipulação dos hospedeiros, da regulação do promotor, a estabilidade e solubilidade da proteína obtida, assim como a velocidade e o rendimento da purificação (JONASSON et al. 2002; HUANG et al. 2012).

Após a clonagem num vetor de expressão, muitas das vezes a proteína não é expressa ou é expressa apenas em níveis baixos. E por isso muitos trabalhos têm sido desenvolvidos para melhorar a expressão de genes clonados, incluindo a otimização das condições de crescimento do hospedeiro e do desenvolvimento de novas estirpes bacterianas (BERLEC & STRUKELJ, 2013).

As bactérias são os organismos mais simples e os mais utilizados para a produção de proteínas recombinantes representando cerca de 30% de produtos biofarmacêuticos no mercado (FERRER-MIRALLES et al. 2009). Hospedeiros bacterianos são os mais vantajosos em relação a outros sistemas de expressão, tais como baixo custo do meio de crescimento e alta produtividade, e tempo de processos mais curtos, além de serem geneticamente bem caracterizados (PACHECO et al. 2012). A bactéria mais comumente utilizada para a produção de proteínas recombinantes é a *Escherichia coli*, que possui longa tradição de utilização segura nos laboratórios e indústria (HUANG et al. 2012).

Entretanto, os problemas mais comuns relatados em sistemas bacterianos incluem uma etapa de clonagem demorada, parâmetros de indução como: temperatura, concentração do indutor e tempo de indução, a expressão ineficiente e/ou a fraca solubilidade da proteína de interesse devido a deficiência nas modificações pós-traducionais, a falta de ligações dissulfeto e ausência de chaperonas necessárias para o correto enovelamento da proteína e a diferença da utilização de códons da proteína a ser

expressa e o hospedeiro pode atrapalhar na expressão (HART & TARENDEAU, 2006; PACHECO et al. 2012). Além disso, a alta expressão de genes/proteínas resulta na maioria das vezes uma carga metabólica excessiva na célula expressão (HART & TARENDEAU, 2006).

Variando vários parâmetros, tais como o hospedeiro bacteriano, temperatura de crescimento, promotores diferentes e regulados, e à manipulação de sequência de DNA com estratégias através do códon de otimização podem reduzir os problemas e aumentar a taxa de expressão da proteína recombinante (BAKKE et al. 2009). Uma ampla gama de hospedeiros geneticamente modificados está hoje atualmente disponível comercialmente para facilitar a expressão e a solubilidade da proteína-alvo, para facilitar a formação de ligações dissulfeto e hospedeiros livres de proteases endógenas (SAHDEV et al. 2008). Promotores tradicionalmente usados como o *lac*, *trc*, *araBAD* e pET oferecem maneiras de controlar o nível de expressão e as forças exercidas sobre o hospedeiro (RAMON & SMITH, 2011; PACHECO et al. 2012).

Um dos esforços no melhoramento dos hospedeiros para aplicação industrial, hoje em dia, é com base nas estratégias de melhoramento de estirpes naturais (HART & TARENDEAU, 2006; DERKX et al. 2014). Assim, muitos esforços recentes de engenharia genética, engenharia metabólica e biologia sintética têm se centrado no desenvolvimento de ferramentas no ajuste da expressão para facilitar a busca de um equilíbrio ótimo entre a viabilidade celular e o cassete de expressão (ANDRIANANTOANDRO et al. 2006; OVERTON, 2014).

2.8 Expressão de genes heterólogos em *E. coli* por meio do operon *lac*

Muitos microrganismos secretam quantidades limitadas de enzimas ou outras proteínas e peptídeos, e por isso a expressão heteróloga em outras células hospedeiras vem sendo utilizadas para potencializar a obtenção desses produtos de interesse biotecnológico através das técnicas da engenharia genética (YIN et al., 2007; DO CARMO, 2010).

A expressão do gene de proteínas heterólogas pode ser induzível ou não. Os sistemas de expressão induzível são mais utilizados, e em *E. coli*, um dos mais utilizados são baseados no operon *lac* (OVERTON, 2014). Neste sistema, o DNA de interesse

depende da presença de um indutor, seja sintético ou não, pois o gene heterólogo é clonado em fago ou plasmídeo do operon *lac* no lugar dos genes estruturais *lacZ*, *lacY* e *lacA* (SHOJAOOSADATI et al., 2008; RAVISHANKAR et al. 2015). Normalmente em laboratório, utiliza-se o isopropil- β -D-tiogalactopiranosídeo (IPTG), pois é um forte indutor químico análogo da lactose utilizado largamente em pesquisa básica na indução de promotores e representa atualmente uma das moléculas indutoras mais utilizadas nos mais variados sistemas de expressão (OVERTON, 2014; RAVISHANKAR et al. 2015). Na ausência do indutor como o próprio IPTG, o repressor *lacI* se liga ao operador situado logo após ao promotor. Quando o indutor IPTG está presente liga-se ao produto gênico do *lacI* (repressor) e causa a dissociação do operador ocasionando a transcrição dos genes (MARQUES 2012).

O IPTG é um indutor sintético que apresenta duas vantagens: não é dependente de *lac* permease, se difundindo pela membrana celular por difusão e não pode ser clivado pela β -galactosidase e assim não perde sua capacidade indutora, prevenindo o término da transcrição (SCHUMANN & FERREIRA, 2004; ROSANO & CECCARELLI, 2014). Porém, se por um lado ele apresenta efetividade na indução e não sendo consumido durante o processo (SHOJAOOSADATI et al. 2008), por outro lado o IPTG apresenta duas desvantagens para um processo industrial: é caro e tóxico, dificultando seu uso em larga escala. Porém o IPTG pode substituir a lactose em meios de indução em concentrações baixas para minimizar potenciais problemas de solubilidade de proteínas (HOEK & HOGEWEG, 2007; GOPAL & KUMAR, 2013).

Por ser tóxico, o uso de IPTG não é recomendado para a produção de proteínas com finalidades terapêuticas em saúde humana ou animal. Se utilizado nesses casos, procedimentos para sua remoção deverão ser incorporados ao processo e será necessário comprovar que está ausente no produto final (JANA & DEB, 2004; PETI & PAGE, 2007).

Muitas proteínas heterólogas são produzidas por microrganismos recombinantes, sendo a *Escherichia coli* o mais utilizado nestes eventos. Vários são os fatores que ocasionam a escolha de expressão em *E. coli*: a linhagem bacteriana utilizada, o número de cópias plasmidiais, a preferência do antibiótico apropriado para seleção de clones, **a força e a regulação do promotor, a eficiência dos terminadores de transcrição**, a estabilidade do RNA mensageiro, **a estrutura do sítio de ligação ao ribossomo**, e fáceis condições de cultivo, tais como tempo, temperatura, pH, características nutricionais e meio de cultura bem definidos (JANA & DEB, 2004; SORENSEN & MORTENSEN,

2005; RAMON & SMITH, 2011; HOCHKOEPLER, 2013). A *E. coli* é capaz de produzir proteínas em níveis tão altos quanto de 10 a 40% da massa celular total, além de ser bem caracterizada geneticamente, estando disponível um grande número de vetores de clonagem e expressão para serem utilizados e um grande número de linhagens mutantes, sendo seu metabolismo amplamente conhecido (OVERTON, 2014), apresenta rápido crescimento celular, permitindo a obtenção de altas densidades celulares, e consequentemente ocasionando a produção de proteína recombinante com baixo custo (TRIPATHI et al. 2009). Dos 211 produtos biofarmacêuticos que foram aprovados para uso regulamentar até o fim de 2011, 66 (31%) foram produzidos em *Escherichia coli* (BERLEC & STRUKELJ, 2013).

Entretanto, somente alguns genes podem ser expressos neste organismo. As principais desvantagens da *E. coli* como sistema de expressão inclui a incapacidade de fazer modificações pós-traducional encontradas em proteínas eucariotas, a não secreção da proteína para o meio de cultura e a limitada capacidade de formar pontes dissulfeto. Tais inconveniências proíbem a bactéria *E. coli* de serem hospedeiras para agentes terapêuticos complexos e importantes onde o correto enovelamento desempenha papéis cruciais nas suas atividades biológicas (HUANG et al. 2012; OVERTON, 2014).

No entanto, muitas proteínas terapêuticas recombinantes foram produzidas com êxito em *E. coli*, tais como: fatores sanguíneos, trombolíticos e anticoagulantes, interferons, interleucinas, enzimas, hormônio de crescimento, insulina, vacinas e anticorpos monoclonais (CHOI E LEE, 2004; HUANG et al. 2012; BERLEC & STRUKELJ, 2013).

Em *E. coli*, as proteínas recombinantes são expressas no citoplasma. No entanto, muitas proteínas são expressas em forma insolúvel, biologicamente inativa em corpos de inclusão, ocasionando sua recuperação complicada e de custo elevado. Muitas técnicas estão sendo desenvolvidas para que proteínas recombinantes em *E. coli* estejam em forma solúvel e/ou secretadas (SIVASHANMUGAM et al. 2009; BERLEC & STRUKELJ, 2013). A formação de corpos de inclusão continua sendo uma barreira para a expressão gênica no citosol. Porém, os corpos de inclusão apresentam algumas vantagens, no entanto, acabam sendo insignificantes, considerando a difícil tarefa de enovelar as proteínas agregadas, a incerteza de que a proteína enovelada detém a atividade biológica, e a redução do rendimento da proteína enovelada e purificada. Além disso, o custo do sistema é encarecido e o rendimento pode ser insatisfatório (OVERTON, 2014).

Existem várias maneiras de evitar a formação destes corpos de inclusão em *E. coli*, como a redução da velocidade de síntese proteica, que pode ser obtida usando baixas temperaturas para o crescimento bacteriano (SORENSEN & MORTENSEN, 2005). Isto resulta na síntese de duas classes de proteínas: proteases e chaperonas. Resumidamente, as proteases são responsáveis pela degradação de proteínas anormais, enquanto que as chaperonas facilitam o correto enovelamento das proteínas sintetizadas ou tentam redobrar as proteínas deformadas (MOGK et al. 2002; ARIIO DE MARCO, 2009; OVERTON, 2014).

Outra estratégia é fusionar a proteína alvo a um parceiro altamente solúvel para o periplasma, como uma proteína carreadora, já que esta pode mascarar as propriedades físicas do peptídeo, que permite a translocação das proteínas totalmente dobradas e, além disso, pode facilitar a purificação da proteína de fusão. A secreção para o periplasma é, por vezes, a única maneira de produzir uma proteína recombinante (ARIIO DE MARCO, 2009; BERLEC & STRUKELJ, 2013; HOCHKOEPLER, 2013; ROSANO & CECCARELLI, 2014).

A purificação é o último passo a ser realizado após a expressão e atualmente as construções dos sistemas de expressão são realizadas com a finalidade de facilitar este processo. Entre o grande número de sistemas existentes o mais utilizado são os que expressam as proteínas de fusão, onde o gene da proteína alvo está ligado a um gene de outra proteína que ficará retida na fase estacionária de uma coluna cromatográfica. Tendo vantagem adicional é que a proteína de fusão pode proteger a proteína de interesse do ataque de proteases (NEIVA, 2005; OVERTON, 2014). Apesar das vantagens proporcionais pela fusão, tem-se a desvantagem, em algumas construções, do produto desejado de apenas apresentar de uma pequena porção da proteína de fusão, ocasionando um baixo rendimento do peptídeo alvo. Entretanto, o rendimento pode ser aumentado pela produção em *tandem*, ou seja, uma proteína com múltiplas cópias do peptídeo alvo com tags metioninas ou histidinas intercalantes para permitir posterior clivagem química ou enzimática (BERLEC & STRUKELJ, 2013; ROSANO & CECCARELLI, 2014).

2.9 Vetores

São veículos de clonagem utilizados em engenharia genética para a introdução de DNA exógeno para dentro de um hospedeiro (SMILLIE et al. 2010; WAEGEMAN & SOETAERT, 2011). Com exceção dos vetores que se integram no genoma da bactéria, a grande maioria são moléculas com replicação independente e que apresentam marca de seleção, sítios únicos para enzimas de restrição que facilitam a sua ligação ao DNA exógeno (ASTOLFI-FILHO, 1985).

Os vetores mais simples e estritamente utilizados são conhecidos como plasmídeos devido sua fácil manipulação e a grande variabilidade de aplicações (BATISTA, 2008). Plasmídeos são moléculas circulares duplas de DNA existentes nesses microrganismos e fora dos seus cromossomos (YI-YWAN et al. 2011).

Essas moléculas atuam na transferência de material genético exógeno para o interior das células e a cada divisão da célula hospedeira, o DNA plasmidial é duplicado, ficando nas células-filhas cópias do mesmo (THOMAS, 2000; MOLBAK et al. 2003). São facilmente isolados e manipulados *in vitro* e ainda podem não causar danos as hospedeiras podendo carregar genes de resistência a antibióticos para seleção. Os antibióticos comumente utilizados na engenharia genética são ampicilina, kanamicina, tetraciclina e cloranfenicol (BATISTA, 2008; BERLEC & STRUKELJ, 2013).

Os vetores podem ser utilizados para muitas finalidades, incluindo a seleção de clones recombinantes, seleção de origem de replicação, expressão ou ainda expressão e secreção de uma dada proteína (YI-YWAN et al. 2011).

Um vetor de clonagem utilizado em *E. coli* deve possuir as seguintes propriedades: habilidade para se replicar na célula hospedeira como a presença de uma origem de replicação (ORI) que controla a taxa de replicação do plasmídeo e, portanto, o número de cópias (BERLEC & STRUKELJ, 2013), deve acomodar insertos de vários tamanhos sem danificar as funções de replicação, deve possuir a facilidade de inserção na célula hospedeira depois de manipulação *in vitro*, deve conter um ou mais genes marcadores de seleção para selecionar de maneira eficaz e eficiente somente as células que contêm o vetor e permitir uma manutenção estável dos plasmídeos (BERLEC & STRUKELJ, 2013), conter somente um sítio de múltipla clonagem (MCS) para

endonucleases de restrição, o que fornece várias alternativas para clonagem (BATISTA, 2008).

Os genes são clonados em plasmídeos após digestão com endonucleases de restrição, e previamente selecionados ao serem inseridos em bactérias, por meio da técnica de transformação (BATISTA, 2008). Estas células contendo o plasmídeo com o gene de interesse se multiplicam carregando este clone, em seguida podendo purificar este DNA plasmidial para diferentes fins (BATISTA, 2008).

Dentre as estratégias utilizadas para produção de clones contendo multicópias encontram-se, a seleção de transformantes em altas concentrações de antibiótico. Sob condições laboratoriais, tais plasmídeos multicópias são distribuídos aleatoriamente durante a divisão celular e na ausência de pressão seletiva são perdidos com baixa frequência, entretanto a maneira mais simples de resolver este problema é aproveitar os marcadores de resistência a antibióticos e com a concomitante adição do antibiótico ao meio de cultivo. Cria-se, assim, uma pressão seletiva para que as células retenham o plasmídeo (HAGG et al., 2004).

Um marco na história da clonagem molecular foi a construção do plasmídeo pBR322, por Bolívar e Rodrigues, no final da década de 70, dando origem a maioria dos plasmídeos hoje existentes (BATAUS, 1991). O plasmídeo pBR322 possui genes de resistência aos antibióticos ampicilina e tetraciclina contendo em cada um deles, sítios únicos de restrição para uso em experimentos de clonagem, origem de replicação que conferem um número de cópias relativamente baixo (15-20 cópias) e possui um tamanho pequeno (4.361 pb), razão pelo qual é considerado um bom vetor de clonagem para a produção em larga escala de agentes terapêuticos em *E. coli* (HUANG et al. 2012).

Os vetores pUC18 e pUC19 são vetores de *E. coli* de tamanho pequeno, com 2686 pares de bases de comprimento, apresentando-se em multicópias dentro da hospedeira devido a uma simples mutação pontual no gene *rep* (gene de origem de replicação). Esses dois vetores são idênticos, com exceção do múltiplo sítio de clonagem (MCS) que está presente nos dois vetores em orientação contrária. Os plasmídeos pUC18/19 contém: origem de replicação *rep* do plasmídeo pMB1, ambos com elevados número de cópias (500-700 exemplares); gene *bla*, codificando para beta-lactamase que confere resistência a ampicilina do plasmídeo pBR322. Além disso, contém o gene da beta-galactosidase que permite a seleção direta dos clones recombinantes de *E. coli* (HUANG et al. 2012).

Muitas hospedeiras que possuem vetores com essa tais configuração acima são capazes de crescer num meio contendo o antibiótico, enquanto que as células não transformadas acabam morrendo. Entretanto, existem ainda outros métodos de seleção úteis, os chamados vetores que envolvam a clonagem e caracterização de promotores que são denominados de caça-promotores.

2.9.1 Vetores de seleção de promotores (VSP) ou Caça-Promotores.

A construção *in vitro* de fusões de gene entre um promotor e um gene repórter revolucionou técnicas de engenharia genética que puderam analisar a expressão do gene. A tecnologia da fusão de genes é empregada utilizando diversos genes repórter em bactérias para monitorar a expressão gênica *in vivo*, especialmente em casos em que os produtos são difíceis de se obter em apenas um único experimento (ALI & MURRELL, 2009).

Em experimentos que envolvam a clonagem e caracterização de promotores usam-se os chamados vetores seleção de promotores (VSP), Caça promotores, ou sonda de promotores (*promoter-probes vectors*). Estes vetores caracterizam-se por portarem um ou mais genes marcadores e um gene repórter desprovido de seu promotor original justaposto a uma região de um ou mais sítios de clonagem (*polylinker*) para inserção de sequências promotoras randômicas de DNA (HENRIQUE et al. 2008). Uma vez realizando uma clonagem utilizando um vetor caça-promotor, podem ser identificados os plasmídeos recombinantes uma vez que tenham recebido a inserção de uma sequência promotora, pois quando a sequência promotora é inserida, ocorre a expressão do gene repórter ocorrendo a seleção de colônias recombinantes, onde o gene originalmente inativo tenha tido sua transcrição ativada, e possibilitando a quantificação de tal expressão (SCHOLZ et al. 2001; HENRIQUE et al. 2008).

Os vetores seleção de promotores não devem possuir sítios para as principais enzimas de restrição no gene repórter, na marca de seleção e na origem de replicação, pois estes sítios devem estar em uma região de múltiplos sítios de clonagem que antecede o gene indicador, na qual serão inseridas as sequências promotoras.

Jurgen Brosius, em 1984, construiu um vetor sonda, o plasmídeo pKK232-8 utilizando o gene CAT, como o gene repórter. Após transcrita, a enzima Cloranfenicol

acetiltransferase (CAT) é responsável pela inativação por acetilação do antibiótico cloranfenicol, tornando as células resistentes a tal droga (SCHLUNZEN et al., 2001; SORENSEN & MORTENSEN, 2005). Este vetor, o pKK232-8, possui múltiplos sítios de clonagem (*polylinker*) situado para inserção de DNA exógeno que, caso possua o promotor, irá proporcionar a expressão do gene repórter, conferindo resistência a cloranfenicol as células de *E. coli*. Possui ainda gene de resistência a ampicilina, origem de replicação do plasmídeo pBR322 (um moderado número de 15-20 cópias), gene CAT do plasmídeo pCM71 e outras regiões provenientes dos plasmídeos pKK3535 e pKK9-4. É usado para isolar, caracterizar promotores e avaliar a atividade de promotores quanto ao nível de força e regulação da expressão.

Através deste conceito, vários pesquisadores utilizaram como vetor sonda o plasmídeo pKK232-8 (Figura 10), sendo que os promotores só são capazes de ativar a transcrição do gene CAT sob seu controle quando suas bactérias receptoras eram crescidas em meio contendo cloranfenicol. Uma vez transferidas para o meio de cultura em placas, essas mesmas bactérias se mostravam capazes de expressar a resistência a cloranfenicol (FOURNIER et al. 1999; SCHOLZ et al. 2001; MRUK et al. 2007; HENRIQUE et al. 2008; KNIRSCHOVA et al. 2015).

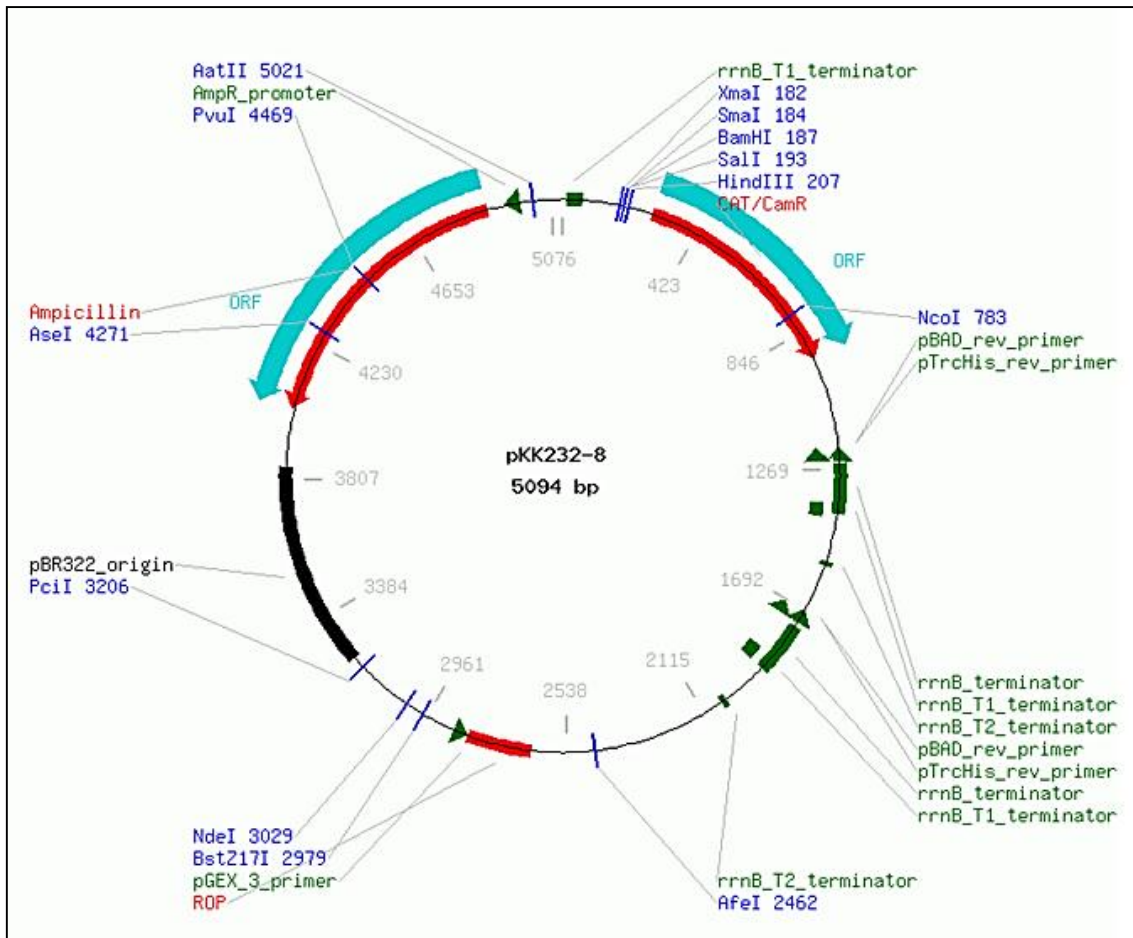


Figura 10: Representação esquemática do plasmídeo caça-promotor pKK232-8 para obtenção de sequências promotoras.

2.10 Construção de biblioteca de promotores por síntese química para produção de proteínas heterólogas

Um promotor gênico eficiente de *E. coli* adequado para a síntese de proteínas em alto nível deve apresentar várias características desejáveis. Entre elas, deve ser forte, resultando em um acúmulo de 10 a 30 % ou mais da proteína celular, ter um nível mínimo de atividade de expressão basal, ou seja, ser fortemente regulado, deve ser transferível com facilidade para outras cepas de *E. coli* para facilitar as tentativas de rendimento de proteína com um grande número de linhagens, e sua indução deve ser feita de forma simples, efetiva e eficaz em termos de custo (JANA & DEB, 2004; MIKSCH et al., 2005; BERLEC & STRUKELJ, 2013).

São vários os sistemas utilizados atualmente que utilizam promotores reguláveis, onde eles utilizam repressores: versões naturais ou modificadas do *LacI* promotor de *E. coli*, o promotor e repressor cI do fago λ , uma versão sensível a temperatura, sistemas com base na polimerase de RNA de T7 induzido por IPTG e um quarto denominado de sistema PET (DE3/T7) que consiste em um fago de RNA polimerase projetado a partir do promotor *lac* de *E. coli*, gene da polimerase de RNA do T7 e induzido por IPTG (ROSANO & CECCARELLI, 2014; OVERTON, 2014).

O promotor procariótico mais usado em pesquisas é sem dúvida o promotor *lac*, componente central do operon *lac*. O conhecimento acumulado sobre o funcionamento do sistema permitiu a utilização prolongada em vetores de expressão. Entretanto, o promotor *lac* é bastante fraco e, portanto, não é muito útil para a produção de proteínas recombinantes. Entretanto híbridos sintéticos, que utilizam outros promotores com associadas com as vantagens do promotor *lac* permitiu o aprimoramento e o aumento da expressão (ROSANO & CECCARELLI, 2014). Entretanto, o promotor gênico T7, derivado de bacteriófagos T7 e presente no sistema de expressão pET, é um dos mais utilizados atualmente para a produção de proteínas recombinantes. Sendo que em 2003, o sistema pET representou mais de 90% dos protocolos de preparação de proteínas e cujas estruturas foram depositadas no Protein Data Bank (PDB) (ROSANO & CECCARELLI, 2014).

Deve-se considerar que quanto mais próximo ao consenso ideal uma sequência regulatória estiver, mais forte será o promotor, mas deve-se ter em conta que promotores mais fracos são mais precisamente regulados quando necessário para otimizar a expressão de diferentes enzimas dentro da célula. A construção de promotores sintéticos tem sido cada vez mais utilizada em pesquisas atuais, no âmbito de ser obter um maior nível de expressão de determinados genes. Portanto o ajuste da força do promotor é de interesse sendo útil na fermentação industrial para produção de proteínas recombinantes (JENSEN & HAMMER, 2000; ALPER et al. 2006; RAMON & SMITH, 2011).

A busca de sequências promotoras não está voltada somente ao uso biotecnológico, mas pela compreensão em si do processo transcripcional, que permite otimizar tais utilizações para melhorar o rendimento, estabilidade, atividade específica, e solubilidade dos produtos, para isso utilizam a evolução dirigida para alcançar esses tipos de melhorias (RAMON & SMITH, 2011).

Existem atualmente duas técnicas estabelecidas para gerar essas bibliotecas de promotores. O primeiro método desenvolvido a partir de mutações aleatórias, Jensen e Hammer, 1998, construíram uma biblioteca de promotores sintéticos através da modulação das sequências espaçadoras entre as regiões -35 e -10 de promotores de bactérias, mantendo as regiões de consenso intacta. Com isso, construíram promotores artificiais utilizando oligonucleotídeos sintéticos degenerados com o intuito de aumentar o nível de expressão gênica em *Lactococcus lactis*. Como resultado, tiveram sucesso na expressão do gene da luciferase em 25, 50, 200 e 400% acima da expressão normal do gene e começaram a aplicar em outros organismos, como em *E. coli* (JENSEN & HAMMER, 2000; SANTOS & STEPHANOPOULOS, 2008).

Um outro método de criação de mutações é a construção de bibliotecas de promotores por meio de PCR de baixa fidelidade (*Error prone PCR*) (SANTOS & STEPHANOPOULOS, 2008; GRUET et al. 2012) o qual cria sequências promotoras com diferentes mutações aleatórias utilizando um cofator não específico para a enzima *Taq* polimerase. Tais mutações ocasionam uma grande diversidade de promotores que podem ser usados para fins biotecnológicos (KAGIYA et al. 2005). A partir do quanto ciclo de *Error prone PCR*, mutações do tipo transverso e transições acontecem ocasionando promotores construídos artificialmente. Kagiya et al. 2005, através desta técnica criaram promotores sintéticos e compararam com o promotor pTAc. Um dos promotores expressou mais fortemente o gene da luciferase e o de β -galactosidade.

Aplicações das bibliotecas de promotores são limitadas atualmente para a manipulação de um único gene ou operon sequencial. Uma desvantagem explícita da técnica é que ciclos de modificações são necessários para encontrar melhores níveis de expressão do gene (SANTOS & STEPHANOPOULOS, 2008). Os promotores sintéticos possuem um grande potencial para permitir a variação da transcrição simultânea de múltiplos genes de uma mesma via. Entretanto, bibliotecas de expressão podem ser criadas, simplesmente pela introdução de cada gene sob o controle de diferentes promotores e, em seguida, através do melhoramento das diferentes combinações que podem ser ajustadas.

3 OBJETIVOS

3.1 Objetivo geral:

- Produção, clonagem e caracterização de novas sequências gênicas promotoras de alta eficiência e funcionais em *Escherichia coli* com atividade regulável através de síntese química de DNA randomizada.

3.2 Objetivos específicos:

- Construção da biblioteca de promotores sintéticos que apresente diversos elementos UP (*Up elements*) com sequências aleatórias, regiões -35 com sequências aleatórias, regiões -10 aleatórias, região espaçadora idêntica ao operador *lac* e sequências nucleotídicas aleatórias no espaço entre a região -10 e o ponto de início de transcrição;
- Clonagem dos promotores no múltiplo sitio de clonagem do vetor caça-promotor pKK232-8 e selecionar clones recombinantes resistentes por meio de transformação genética a diferentes concentrações de cloranfenicol, sensíveis ou não à indução por IPTG;
- Determinar as sequências nucleotídicas dos promotores sintéticos regulados ou não por sequenciamento, analisar e comparar as suas sequências;

4 MATERIAL E MÉTODOS

4.1 Desenho Experimental

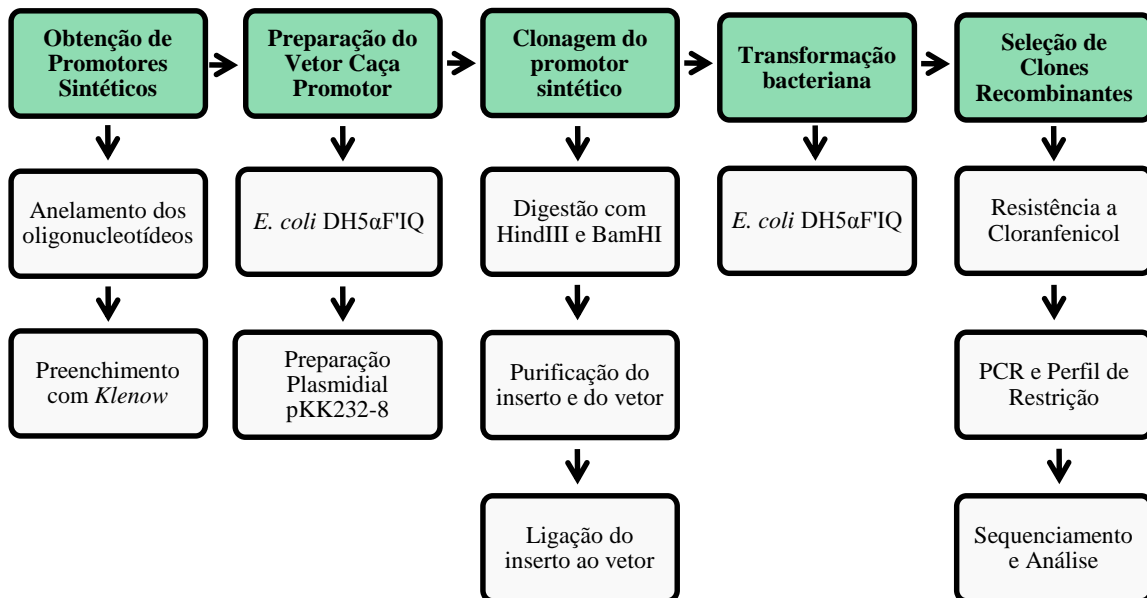


Figura 11: Fluxograma dos procedimentos experimentais realizados.

4.2 MATERIAL

4.2.1 Linhagem Bacteriana

A linhagem de *E. coli* utilizada neste trabalho foi a DH5 α F'IQ que apresenta o seguinte genótipo: F- ϕ 80lacZ Δ M15 Δ (lacZYA-argF) U169 recA1 endA1 hsdR17 (rk-, mk+) phoA supE44 λ - thi-1 gyrA96 relA1/F' proAB+ lacIqZ Δ M15 zcf: Tn5 [KmR].

A linhagem DH5 α F'IQ apresenta expressão para o produto do gene *lacIq*, conferindo a produção de repressor *LacI* em alto nível e apresenta resistência ao antibiótico kanamicina (JESSE & BLODGETT, 1988).

4.2.2 Vetor

O vetor utilizado neste trabalho foi o caça-promotor pKK232-8. O plasmídeo pKK232-8 apresenta grandes inovações de tamanha importância como: terminadores fortes de transcrição *rrnB* gene de *E coli* que impedem a transcrição do gene CAT por outros promotores existentes no vetor e códon de terminação da tradução nas três fases de leitura (a montante da região estrutural do gene CAT e a jusante dos sítios de clonagem). Isto é necessário para não ocorrer a formação de proteínas híbridas, favorecendo que a expressão do gene CAT esteja associada somente com a sequência promotora clonada (LEWIN et al. 2005).

4.2.3 Oligonucleotídeos

Os oligonucleotídeos DGN-1 e DGN-2 (*primers*) foram sintetizados quimicamente (*EXXTEND*). Os oligonucleotídeos utilizados neste trabalho estão listados no quadro 1, sendo que no DGN-1 foi adicionado o sítio de restrição para a enzima *Bam* HI e no DGN-2 foi adicionado um sítio de restrição para a enzima *Hind* III. O DGN-1 e DGN-2 representam os oligonucleotídeos com sequências correspondentes ao núcleo do promotor sintético. O *primer* utilizado no sequenciamento está descrito como Rv (*Reverse*).

Quadro 1: Oligonucleotídeos utilizados neste trabalho.

<i>Primers</i>	Sequência	MER*
DGN-1	5' ATCCCGGGATCC KHKMAAHWWHMSKVTGWSWTTGTGATCCGCTC ACAA 3'	49
DGN-2	5' ATCCCGGGAAGCTT YGKDRKSYAKYWTMTTGTGATCCGCTCACAA 3'	45
Rv	5' GGCATTTTCAGTCAGTTG 3'	17

**nucleotídeos*

V = A, C e G W = T e A R = G e A Y = C e T S = C e G

D = A, G e T H = A, C e T K = G e T M = A e C

4.2.4 Reagentes

Os reagentes e tampões utilizados para o preparo das soluções foram de grau analítico. A água utilizada neste trabalho foi destilada e deionizada (água ultrapura tipo 1).

As soluções foram esterilizadas a 120°C por 15 minutos ou, quando necessário, filtradas em membranas do tipo milipore com poro de 0,22µm.

4.2.5 Meios de cultura

Quadro 2: Meios de cultura utilizados neste trabalho.

MEIOS DE CULTURA	CONSIDERAÇÕES
Meio LB (Luria- Bertani) líquido	NaCl 1,0% Peptona 1,0% Extrato de levedura 0,5% pH 7,2
Meio LB – Ágar – sólido	Meio LB adicionado de: 1,5% de ágar bacteriológico pH 7,5
Meio SOFT	Meio LB adicionado de: 0,8 % (m/v) de ágar bacteriológico

4.2.6 Enzimas

Tabela 3: Enzimas de restrição utilizadas neste trabalho.

Enzima	Sítio de clivagem	Tampão	Temperatura de incubação	Procedência
<i>Bam</i> HI	G/GATCC CCTAG/G	NEB 3	2h /37°C	New England BioLabs® <i>Inc</i>
<i>Hind</i> III	A/AGCTT TTCGA/A	NEB 2	2h /37°C	New England BioLabs® <i>Inc</i>

Outras enzimas

Klenow DNA polimerase (New England BioLabs®*Inc*)

T4 DNA ligase (New England BioLabs®*Inc*)

Taq DNA polimerase (New England BioLabs®*Inc*)

4.2.7 Tampões e soluções

Quadro 3: Tampões e soluções utilizados neste trabalho.

SOLUÇÕES E TAMPÕES	COMPOSIÇÃO
Tampão TE	Tris-HCl - 20 mM; EDTA – 0,1 mM; pH 7,5
Tampão de corrida TEB 10x	Tris base - 0,89 M; Ácido Bórico - 0,89M; EDTA - 0,08M; pH 8.0- 8.4
Tampão de amostra TEB 5x	Glicerol 30% (p/v); Azul de bromofenol - 0,01% (p/v)
Tampão de lavagem	Tris-HCl (pH 7,5) 20 mM; EDTA 2 mM; NaCl 200 mM; Etanol absoluto 96 mL
Cloranfenicol	Solução estoque: 34 mg/mL Solução de trabalho: 25, 50, 100, 200, 500, 1000 e 2000 µg/mL (quando modificado será citado no teste).
Ampicilina	Solução estoque: 1g/mL Solução de trabalho: 100 µg/mL
Glicogênio	Solução estoque: 2 µg/µL
Kanamicina	Solucao estoque: 34 mg/mL Solução de trabalho: 50 µg/mL
Brometo de etídio	Solução estoque: 10 mg/mL Solução trabalho: 1 µg/mL
Etanol	70% (V/V)
MgCl₂	25 mM
dNTP	2,5 mM
Acetato de amônia	7,5M
NaCl	3,0 M
Marcador Ladder DNA GeneRuler™ 1 kb (Fermentas); Plus; Mix	250-10,000 pb; 75-20000 pb; 100-10000 pb.

Soluções do Kit GE HealthCare	<p>Solução I: Tris-HCl (pH 8) 25 mM EDTA 10 mM Glicose 10 mM</p> <p>Solução II: NaOH 0,2 M SDS 1% (p/v)</p> <p>Solução III: KAC 3M Ácido acético 2M pH 4,8-5,0</p>
--------------------------------------	---

4.3 MÉTODOS

4.3.1 Construção dos promotores sintéticos

4.3.1.1 Desenho e síntese dos promotores sintéticos

A construção dos promotores sintéticos foi realizada comparando-se as sequências de promotores fortes de *E. coli* descritas na literatura analisando as sequências na região -35 e -10, nos elementos a montante da região -35 (*Up elements*); na região a jusante da região -10, entre a região -10 e o início da transcrição (o discriminador) e o ponto de início da transcrição (+1). Foram observadas as bases conservadas (que se mantém igual e na mesma posição em grande parte das sequências promotoras quando tais sequências são alinhadas) e as não-conservadas. As que não são conservadas ou são menos conservadas, foram as bases escolhidas para apresentarem modificações aleatórias, a fim de proporcionar um maior número de sequências promotoras diferentes. Desta forma, praticamente se impediria de gerar sequências típicas (similares a consenso).

As bases mais conservadas foram mantidas durante a construção dos promotores sintéticos, exceto quatro bases da região -10 (TATA BOX) e três bases da região -35. Após a análise, os oligonucleotídeos foram desenhados conservando as regiões complementares entre si na região correspondente ao operador *lac* sem degenerações, localizada entre as regiões -35 e -10 contendo nas extremidades sítios para as enzimas de restrição *Bam*HI e *Hind*III (Figura 12). Portanto uma biblioteca de oligonucleotídeos com $20,736 \times 27 \times 16 \times 96 \times 2 \times 2 = 330.225.942.528$ diferentes regiões promotoras em potencial deveriam ser construídas.

Sítio da enzima *Bam* HI

Região de *Up element*

Região -35

Região espaçadora

Região -10

Região entre a região -10 e o ponto de início da transcrição

Ponto de início da transcrição

Sítio de *Hind* III

V = A, C ou G
W = T ou A
R = G ou A
Y = C ou T
S = C ou G
D = A, G ou T
H = A, C ou T
K = G ou T
M = A ou C

5` ATCCCGGGATCC**KHKMAAHWWHMSKV**TTG**WSW**TTGTGAGCGGATAACAA 3`
 5` ATCCCGGG**AAGCTTY**GKDRKSYAKYW**MT**TTGTGAGCGGATAACAA 3`

Figura 12: Desenho dos oligonucleotídeos sintéticos. As regiões em azul são as regiões espaçadoras complementares entre si que possibilitará o anelamento entre os dois oligonucleotídeos (DGN-1 e DGN-2). Nas regiões não-conservadas foram realizadas modificações aleatórias e as letras V, W, R, Y, S, D, H, K e M são os símbolos que indicam as degenerações e representam quais bases serão adicionadas durante a síntese química de oligonucleotídeos.

4.3.1.2 Anelamento dos promotores sintéticos

Para completar a dupla fita (Figura 13) (anelamento dos oligonucleotídeos através da região espaçadora) os oligonucleotídeos DGN-1 e DGN-2 foram diluídos a 1,47 µg/µL e 1,40 µg/µL, respectivamente com tampão Tris-HCl 2,5 mM pH 7,5. Utilizou-se 2 µL do DGN-1 (estoque de 22,0 nmol) e 2 µL do oligonucleotídeo DGN-2 (estoque de 19,4 nmol) separadamente e 2 µL de DGN-1 e DGN-2 (1 µL de cada). Nos três sistemas, completou-se o volume com tampão R (Tris-EDTA 2,5 M) para um sistema com volume final de 10 µL. O sistema foi aquecido por 5 minutos a 100°C seguido de anelamento em temperatura ambiente por 10 minutos. O resultado foi observado em gel de agarose 2%, corado com brometo de etídio e visualizado sob luz ultravioleta (UV).

5` ATCCCGGGATCC**KHKMAAHWWHMSKV**TTG**WSW**TTGTGAGCGGATAACAA 3`
 3` AACACTCGCCTATTGT**MT**WYK**AYS**KRDKGY**TT**CGAAGGGCCCTA 5`

Figura 13: Representação esquemática de anelamento dos oligonucleotídeos através da região espaçadora.

4.3.1.3 Preenchimento das fitas simples dos oligonucleotídeos

O preenchimento das regiões fita-simples do oligonucleotídeos anelados foi realizado utilizando a enzima *Klenow* DNA polimerase. Para a reação foi montado dois sistemas **A** e **B** como mostrado no quadro 4:

Quadro 4: Sistemas e componentes para o preenchimento com a enzima *Klenow*.

Componentes do sistema	Volume dos sistemas (µL)	
	A	B
Oligonucleotídeos anelados	5	10
Enzima <i>Klenow</i>	5	5
Tampão de Reação da enzima	2,5	2,5
dNTPs	7,5	7,5
Água deionizada filtrada	5	0
Volume Final	25	25



Figura 14: Esquema da elongação dos oligonucleotídeos anelados.

O resultado das reações foi observado em gel de agarose 2%, corado com brometo de etídio e visualizado sob luz ultravioleta (UV).

4.3.2 Preparação do vetor caça-promotor

4.3.2.1 Transformação genética de *Escherichia coli*

Utilizou-se células de *E. coli* competentes DH5 α F'IQ (Kit INVITROGEN MAX Efficiency®) quimio-competentes para transformação com plasmídeos pelo método do cloreto de cálcio (modificado de SAMBROOK et al. 2001). Para montar o sistema foi adicionado 2 μ L de DNA plasmidial (plasmídeo pKK 232-8) juntamente com 100 μ L de células, estocada a -80°C e homogeneizado suavemente com 21 mL de solução Tris-HCl 10mM pH 7,5 e CaCl₂ 80 mM para aumentar a eficiência da transformação. O sistema foi incubado durante 30 minutos no gelo e, em seguida, o sistema foi submetido a um choque térmico de 37°C por 5 minutos. Após esse choque térmico o sistema foi imediatamente retornado ao gelo e incubado por mais 5 minutos. Após a incubação no gelo, o sistema de transformação foi ressuspendido em 1 mL de meio LB e incubado por 1 hora a 37 °C. Em seguida, alíquotas foram plaqueadas em placas de Petri com meio LB-ágar contendo ampicilina (100 μ g/mL). Ao final do plaqueamento, as placas de Petri foram incubadas invertidas a 37°C durante a noite. Após 18 horas de incubação, as colônias transformantes foram selecionadas e removidas com palito pontiagudo de madeira estéril para 5 mL de meio de cultura LB-líquido. Os pré-inóculos foram incubados sob agitação constante durante a noite a 37°C, 13.000 rpm.

4.3.2.2 Extração dos plasmídeos dos clones recombinantes

O DNA plasmidial dos clones selecionados foram extraídos com o Kit de extração *Illustra plasmidPrep Mini Spin Kit* (GE Healthcare Bioscience) o qual possui soluções que atuará nas fases de lise celular, inibição de proteases, lavagem e eluição do DNA plasmidial. Os procedimentos foram realizados de acordo com as especificações do fabricante. Os DNAs plasmidiais extraídos foram analisados por eletroforese em gel de agarose 0,8% e visualizado sob luz ultravioleta (UV). Após foi mantido a 4°C até o seu uso novamente.

4.3.3 Clonagem do promotor sintético

4.3.3.1 Digestões do plasmídeo pKK232-8 e do promotor sintético

As digestões foram realizadas conforme instruções dos fabricantes das enzimas. O tempo de incubação, os tampões, as temperaturas de reação, as concentrações de DNA e enzima foram ajustados de acordo com o material a ser digerido e a eficácia da enzima.

Para a digestão do plasmídeo e do promotor sintético dupla fita foi realizada uma digestão dupla e sequencial. Foi utilizado inicialmente a endonuclease de restrição que requer a menor concentração de sal em seu tampão de acordo com o fabricante. Para isto foi montado um sistema primeiramente com a enzima de restrição *Hind* III tanto para o vetor pKK232-8 como para o oligonucleotídeo sintético para um volume final desejado de 50 μ L e incubado por 2 horas a 37°C. Após a conclusão foi ajustado a concentração de sal da mistura reacional utilizando uma pequena quantidade de solução salina concentrada para poder aproximar as condições de reação da segunda endonuclease de restrição *Bam* HI no sistema, que após a adição da enzima e dos reagentes foi ajustado para um volume final de 100 μ L, incubado por 2 horas por 37°C e submetido subsequentemente a 4°C para interromper a digestão. Todos os componentes do sistema e a reação em sequência está especificada no quadro 5.

Quadro 5: Sistemas das digestões do vetor e do oligonucleotídeo sintético.

Componentes dos Sistemas	Vetor	Oligonucleotídeo sintético
	<i>Hind</i> III	<i>Hind</i> III
H ₂ O deionizada	20 µL	20 µL
Tampão NEB 2	5 µL	5 µL
Tampão NEB 3	-	-
BSA 100X	-	-
Vetor pKK232-8	20 µL	-
Oligonucleotídeo	-	20 µL
Enzima	5 µL	5 µL
Volume Final	50 µL	50 µL
Digestão sequencial utilizando <i>Bam</i> HI		
Sistema <i>Hind</i> III		
H ₂ O deionizada	45 µL	45 µL
Tampão NEB 2	42 µL	42 µL
Tampão NEB 3	-	-
Tampão NEB 3	7,5 µL	7,5 µL
BSA 100X	0,5 µL	0,5 µL
Vetor pKK232-8	-	-
Oligonucleotídeo	-	-
Enzima	5 µL de <i>Bam</i> HI	5 µL de <i>Bam</i> HI
Volume Final	100 µL	100 µL

A sequência de corte das enzimas tanto no pKK232-8 e no oligonucleotídeo sintético geram extremidades adesivas. A enzima *Bam*HI e *Hind*III apresentam uma sequência de corte de 6 pares de bases. A complementaridade entre os pares de bases proporciona a ligação do fragmento da sequência promotora ao DNA plasmidial.

4.3.3.2 Purificação do vetor pKK 232-8 e dos oligonucleotídeos

Os dos sistemas de digestão (do vetor e do oligonucleotídeo) foram submetidos à eletroforese em gel de agarose 0,8% e 2,0% respectivamente e visualizado no transluminador. Aplicou-se os sistemas no gel e foi realizada uma eletroforese “longa” para que as bandas de DNA migrassem bastante e para que houvesse separação de possíveis moléculas contaminantes. Após a visualização, as bandas de interesse foram recortadas para purificação.

A purificação do vetor foi realizada através do *Kit comercial illustra GFX PCR DNA and Gel Band Purification* (GE Healthcare). Os procedimentos foram realizados de acordo com as especificações do fabricante. Os fragmentos foram quantificados utilizando o marcador GeneRuler™ 1 kb (Fermentas).

A purificação do oligonucleotídeo foi realizada por eletroforese, seguido de fragmentação do gel e difusão dos oligonucleotídeos seguido de precipitação. O fragmento que corresponde ao oligonucleotídeo foi submetido a 100 µL de tampão TE durante a noite. Após foi centrifugado a 10 minutos, retirado o sobrenadante, onde foi submetido a 1 µL de glicogênio, 2,5 volumes de álcool etílico 100 % e 0,1 volume de NaCl 3M, posto no freezer a 1 hora e centrifugado a 40 minutos por 12.000 rpm para o DNA ficar precipitando com o glicogênio, pois o NaCl interage com a água e o álcool, forçando o DNA a formar agregados. Após a precipitação, foi centrifugado novamente a 12.000 rpm, 10 minutos a 4°C para sedimentação do precipitado. Após isso, foi descartado o sobrenadante e adicionado 500 µL de álcool 70% para retirar o sal, sendo que o glicogênio se mantém na solução, colocado na estufa a 37°C e ressuspendido em tampão TE. Em seguida, os fragmentos foram analisados em gel de agarose (2,0%) e visualizado sob luz ultravioleta (UV).

4.3.3.3 Ligação dos promotores ao vetor pKK 232-8

O vetor e o inserto purificados foram submetidos a uma reação de ligação. Para a ligação dos promotores ao vetor caça-promotor, foi feita a relação inserto-vetor (3:1). O sistema de ligação foi feito com a T4 DNA ligase.

Foi utilizado com 15 µL de vetor, 2,35 µL de oligonucleotídeo, 1,0 µL de T4 DNA ligase, 2,5 µL de tampão de ligação DNA 10x com ATP e 4,0 µL de água deionizada. O sistema foi incubado a 16°C por 16-18 horas.

4.3.4 Clonagem no vetor caça-promotor pKK232-8

4.3.4.1 Preparo de células eletrocompetentes

Foi realizado um pré-inóculo de células competentes *E. coli* DH5 α F'IQ em 5 mL de meio LB com o antibiótico canamicina, e incubado em agitador de bancada (tipo shaker) de 160 rpm a 37°C durante a noite. Dessa cultura, 1 mL foi inoculado em 200 mL de meio LB isento de antibiótico. Esta cultura foi incubada sob agitação de 150 rpm a 37°C e neste período foram retiradas alíquotas de 1 mL, em intervalos de 1 hora até atingir O.D.₆₀₀ = 0,6 nm. Foi utilizado um branco contendo o mesmo meio para cultivo na ausência de crescimento bacteriano. Assim que a O.D. foi alcançada, foi pego tubos falcons pré-resfriados no gelo e distribuído 50 mL de cultura em cada um. As células foram centrifugadas a 4.000 rpm por 10 minutos a 4°C.

O sobrenadante foi descartado de imediato e ressuspendido gentilmente com 10 mL de glicerol 10% estéril gelado até soltar o pellet. As células foram novamente centrifugadas a 4.000 rpm por 10 minutos a 4°C e o sobrenadante descartado. Esta operação foi repetida seguidamente por 3 vezes e as células foram reunidas em um único tubo falcon no final. Logo após as lavagens e centrifugações, o pellet foi ressuspendido gentilmente em 1 mL de glicerol 10% gelado e estéril. As células eletrocompetentes foram separadas em alíquotas de 50 μ L distribuídas em microtubos, congeladas imediatamente em nitrogênio líquido e estocadas a -80°C.

4.3.4.2 Procedimento de eletroporação

Uma alíquota do sistema de ligação (1 μ L) foi misturado em 50 μ L de células *E. coli* DH5 α F'IQ eletrocompetentes que estavam a -80°C e transferidos para uma cuveta de eletroporação (0,2 cm). Foram feitos dois sistemas para a eletroporação: Apenas 1 μ L do vetor intacto + 50 μ L de células competentes e 1 μ L do vetor com inserto + 50 μ L de células competentes.

Foi aplicado um pulso elétrico de 1900 kV, e imediatamente após o pulso, adicionou-se 500 μ L de meio LB na cuveta revigorando as células transformadas. As

amostras foram incubadas a 37°C por 1 hora com agitação (160 rpm) para estabilização. Em seguida, alíquotas dos sistemas foram plaqueados em placas de Petri com meio LB-ágar contendo o antibiótico ampicilina (200 µg/mL). Ao final do plaqueamento as placas foram incubadas invertidas a 37°C durante a noite.

4.3.5 Análise dos clones recombinantes

4.3.5.1 Seleção dos clones transformantes

Para avaliar as diferentes taxas de expressão do gene de resistência ao cloranfenicol relacionada com diferentes sequências promotoras, os clones transformantes foram selecionados em placas de Petri com meio LB sólido contendo ampicilina 200 µg/mL e em placas de Petri contendo cloranfenicol 20 µg/mL, 50 µg/mL e 100 µg/mL, sendo usado como controles nas placas os vetores pKK232-8 intacto, pKG10 e pKG15 nas seguintes condições: Ampicilina + pKK232-8 intacto (Vetor) + células e Ampicilina + sistema de ligação (Vetor + Oligonucleotídeo) + célula. Ao final do plaqueamento as placas foram incubadas invertidas a 37°C durante a noite.

Tais clones recombinantes foram repicados em placas de Petri com meio LB-sólido contendo ampicilina 100 µg/mL e em placas contendo cloranfenicol nas concentrações de 20 µg/mL, 50 µg/mL, 100 µg/mL e 200 µg/mL na presença e ausência de indutor sintético IPTG (0,25 mM). Os clones crescidos em placas de Petri com cloranfenicol foram analisados e classificados quanto ao tamanho das colônias.

Foi utilizado para esse procedimento palitos estéreis, os quais foram imediatamente descartados e autoclavados após o uso. Ao final do plaqueamento as placas foram incubadas invertidas a 37°C durante a noite e depois um novo sistema de réplicaplate foi realizado.

4.3.5.2 Extração do DNA plasmidial dos clones recombinantes

Os clones recombinantes de *E. coli* foram inoculados durante a noite (aproximadamente 16 horas) à 37°C sob agitação de 150 rpm. O DNA plasmidial dos

clones selecionados foram extraídos com o Kit *Illustra plasmidPrep Mini Spin Kit* (GE Healthcare Bioscience), conforme o item **4.3.2.2**.

4.3.6 Sequenciamento

As reações de sequenciamento foram realizadas utilizando o kit Sequencing-Big Dye Terminator Cycle Sequencing Ready ABI versão 3 – Applied Bioscience/Life Technologies, USA, seguindo as recomendações do fabricante. Cada reação de sequenciamento foi ajustada para um volume final de 10 µL, utilizando cerca de 40 ng/µL dos fragmentos de interesse, 0,32 pmoles de oligonucleotídeos iniciadores, e tampão Big-Dye para 1X final e 2 µL 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 3130XL (Applied Biosystems). Seguido de precipitação por EDTA/Álcool absoluto.

4.3.7 Bioinformática

Os dados brutos provenientes do sequenciador de DNA foram submetidos e analisados pelo software Sequencing Analysis versão 5.4. e a seguir, para verificar qualidade das sequências nucleotídicas foi realizado procedimento via PHRED disponível no site da UnB (<http://www.biomol.unb.br/phph/index.html>).

As sequências foram alinhadas utilizando o programa BioEdit (disponível no site: www.mbio.ncsu.edu/bioedit/bioedit.html) para identificação das regiões com função promotora.

4.3.8 Escolha de outros clones para novas análises

Após obter centenas de colônias com diferentes perfis de expressão baseados na análise de resistência a cloranfenicol e em resposta a presença ou ausência de IPTG. Foram selecionados alguns clones e seus plasmídeos foram extraídos, sequenciados e analisados. Para prosseguir com as análises após o resultado do sequenciamento, novos

clones recombinantes juntamente com diferentes sequências promotoras foram submetidos a uma nova réplica-plate.

As sequências promotoras obtidas foram agrupadas de acordo com um perfil de crescimento máximo em cada grupo baseado na análise de resistência a cloranfenicol (20, 500, 1000 e 2000 $\mu\text{g/mL}$) e em resposta a presença ou ausência de IPTG (1 mM) em meio LB sólido. Após a análise das sequências promotoras, 6 plasmídeos foram escolhidos, reintroduzidos em clones *E. coli* por transformação e realizada uma réplica-plate para resistência em cloranfenicol em meio LB sólido (50, 500, 1000 e 2000 $\mu\text{g/mL}$) na ausência e presença de IPTG (1 mM) em meio LB sólido. O crescimento das colônias foi observado e os clones foram analisados também na presença de cloranfenicol em meio líquido.

4.3.9 Crescimento em meio líquido

Os 4 clones foram analisados em meio líquido, contendo o antibiótico cloranfenicol, na presença e ausência de IPTG. Cada clone contendo a sua sequência promotora foi inoculada primeiramente (pré-inóculo) em 5 mL de meio líquido durante a noite (16-18 horas) sob agitação de 150 rpm. Depois da determinação da absorbância do pré-inóculo e padronização, foi adicionado 500 μL de amostra em 50 mL de meio líquido com cloranfenicol a 250 $\mu\text{g/mL}$, na presença ou ausência de IPTG na concentração de 1 mM (Figura 15).



Figura 15: Enlermayers (inóculo) contendo 50 mL de meio LB líquido com cloranfenicol 250 µg/mL na presença e na ausência do indutor IPTG.

Para o início da observação do crescimento celular foi necessário medir o ponto zero. Para isso, de imediato após a inoculação de 500 µL de cultura (pré-inóculo) as garrafas foram agitadas manualmente e foi retirada 600 µL desse meio de cultura. Com essa primeira medição em absorvância a 600 nm, verificou-se que todas as amostras estavam com absorvâncias semelhantes e a partir desta verificação, as amostras foram inoculadas sob agitação de 150 rpm a 37°C, por um período de doze horas de incubação. Após três horas fez-se a primeira medição. Em seguida, as outras onze medições restantes foram feitas a cada uma hora de crescimento.

5 RESULTADOS E DISCUSSÃO

5.1 Construção e clonagem de promotores sintéticos

5.1.1 Construção da biblioteca de promotores sintéticos

Foi primeiramente realizado o anelamento dos oligonucleotídeos através da região espaçadora. A figura 16 mostra a obtenção do promotor fita dupla a partir do anelamento (conforme descrito no item 4.3.1.2) dos dois oligonucleotídeos DGN-1 (49 pb) e DGN-2 (45 pb). Como na canaleta **B** a banda de fluorescência mais intensa tem menor mobilidade eletroforética que as bandas das canaletas **A** e **C** pode-se inferir que o anelamento dos oligonucleotídeos ocorreu como esperado pois os dois oligonucleotídeos têm a sequências do operador lac perfeitamente complementares.



Figura 16: Análise em gel de eletroforese do anelamento dos oligonucleotídeos randômicos. A: oligonucleotídeo DGN-1 (49 mer); B: oligonucleotídeos DGN-1 e DGN-2 anelados e C: oligonucleotídeo DGN-2 (45 mer).

5.1.2 Preenchimento dos oligonucleotídeos com a enzima *Klenow*

A figura 17 mostra a análise eletroforética do processo utilizado para gerar promotores dupla fita com sequências parcialmente randômicas a partir dos oligonucleotídeos anelados tratados com a enzima *Klenow* (conforme descrito no item 4.3.1.3). Observou-se que nas canaletas 4 e 5 onde colocou-se os nucleotídeos anelados e tratados com *Klenow*, que as bandas com maior intensidade de fluorescência apresentaram também menor mobilidade eletroforética em comparação com os oligonucleotídeos apenas anelados (canaleta 3) e também em relação aos oligonucleotídeos não anelados e não tratados com *Klenow* (1 e 2). Havendo a mesma quantidade de oligonucleotídeos (DGN-1 e DGN-2) nas amostras e sabendo-se que DNA de fita dupla absorve mais brometo de etídio devido sua inserção na região espaçadora entre as bases nitrogenadas, pode-se concluir que houve a duplicação desejada das regiões fita simples.

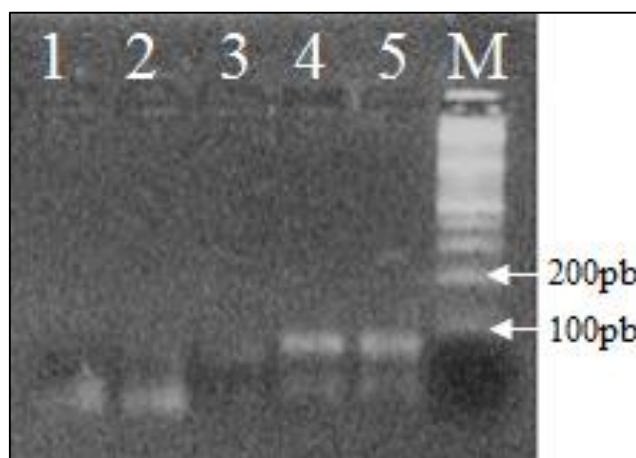


Figura 17: Análise eletroforética do processo de obtenção dos promotores dupla fita. M: Marcador GeneRuler™ 1 kb Mix (Fermentas); 1: oligonucleotídeo DGN-1; 2: oligonucleotídeo DGN-2. 3: oligonucleotídeos DGN-1 e DGN-2 somente anelados; 4 e 5: oligonucleotídeos DGN-1 e DGN-2 anelados e tratados com *Klenow*.

5.1.3 Digestão do vetor e dos promotores sintéticos com enzimas de restrição e purificação

O vetor caça-promotor pKK232-8 foi tratado com as enzimas de restrição *HindIII* e *BamHI* e purificado por eletroforese em gel de agarose 0,8%, de modo que os promotores randômicos (DGN-1 e DGN-2 tratados com *Klenow*) foram tratados com as mesmas enzimas e purificados por eletroforese em gel de agarose 2%. Os produtos das purificações estão mostrados na figura 18.

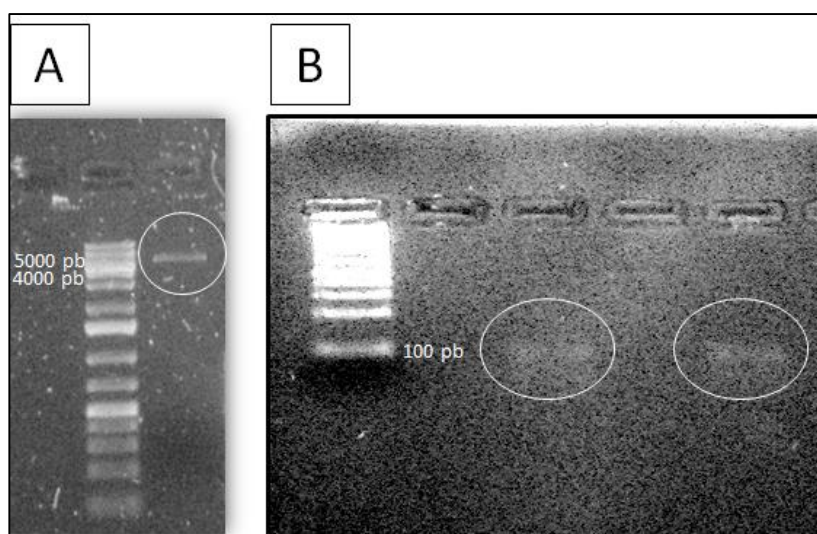


Figura 18: Análise eletroforética do vetor e dos promotores sintéticos purificados. No sistema A mostra o vetor purificado e no sistema B mostra os promotores randômicos purificados.

5.1.4 Clonagem dos promotores no vetor pKK232-8

Para a clonagem dos promotores no vetor pKK232-8, diferentes sequências promotoras (oligonucleotídeos dupla fita) com extremidades digeridas com *BamHI* e *HindIII* foram ligadas aos vetores pKK232-8 digeridas com as mesmas enzimas e a seguir células de *E. coli* da linhagem DH5 α F'IQ foram transformadas por eletroporação.

Após a eletroporação adicionou-se ao sistema 1 mL de meio LB e as células foram incubadas por 1 hora a 37°C e seguir foram semeadas em alíquotas de 50 e 100 μ L em placas de Petri contendo meio seletivo LB-ágar com ampicilina (200 μ g/mL). O resultado

foi a obtenção de transformantes com uma eficiência de transformação de aproximadamente $2,3 \times 10^5$ UFC*/ μg de vetor.

Para verificar se os clones transformantes continham plasmídeos recombinantes foi realizado uma análise por réplica-plate em meio de cultura LB-ágar contendo antibiótico ampicilina (100 $\mu\text{g}/\text{mL}$) e cloranfenicol (20, 50 e 100 $\mu\text{g}/\text{mL}$). Foi observado que cresceram 760 clones em ampicilina (controle), porém desses apenas 95 clones cresceram em todas as placas de Petri que continham o antibiótico cloranfenicol.

5.2 Análise dos clones recombinantes na ausência ou presença do indutor IPTG

Para uma melhor caracterização da atividade e regulação dos promotores sintéticos foram selecionados os 95 clones que cresceram em cloranfenicol. Estes clones foram submetidos a uma nova análise por réplica-plate em meios de cultura contendo ampicilina (100 $\mu\text{g}/\text{mL}$) e cloranfenicol (20, 50, 100 e 200 $\mu\text{g}/\text{mL}$) na presença ou ausência do indutor IPTG (0,25 mM).

A variedade de tamanho das colônias crescidas em placas de Petri indicou que o gene CAT estaria sob o controle de diferentes promotores e em alguns casos sendo induzido por IPTG. Diante destas análises foi possível também ter uma percepção semi qualitativa da expressão do gene CAT em função do tamanho das colônias (Figura 19). Dentre essas considerações pode-se observar clones que:

1. Não cresceram em meio contendo Clf sendo o indutor IPTG presente ou não (em azul);
2. Cresceram apenas em meio contendo Clf na ausência de IPTG (em preto);
3. Cresceram melhor quando o indutor estava presente (em amarelo);
4. Cresceram em Clf tanto na ausência quanto na presença do indutor e tiveram o mesmo tamanho de colônia (em vermelho);
5. Cresceram apenas em meio contendo Clf quando o indutor estava presente (em branco);

6. Cresceram tanto na presença quanto na ausência de IPTG, mas com o tamanho de colônia diferenciado (em laranja na figura 19). Nesse caso as colônias crescidas na presença de Clf sem IPTG são maiores que as crescidas com IPTG. A variação do tamanho das colônias cultivadas com e sem IPTG indica que o IPTG está interferindo na expressão do gene CAT.

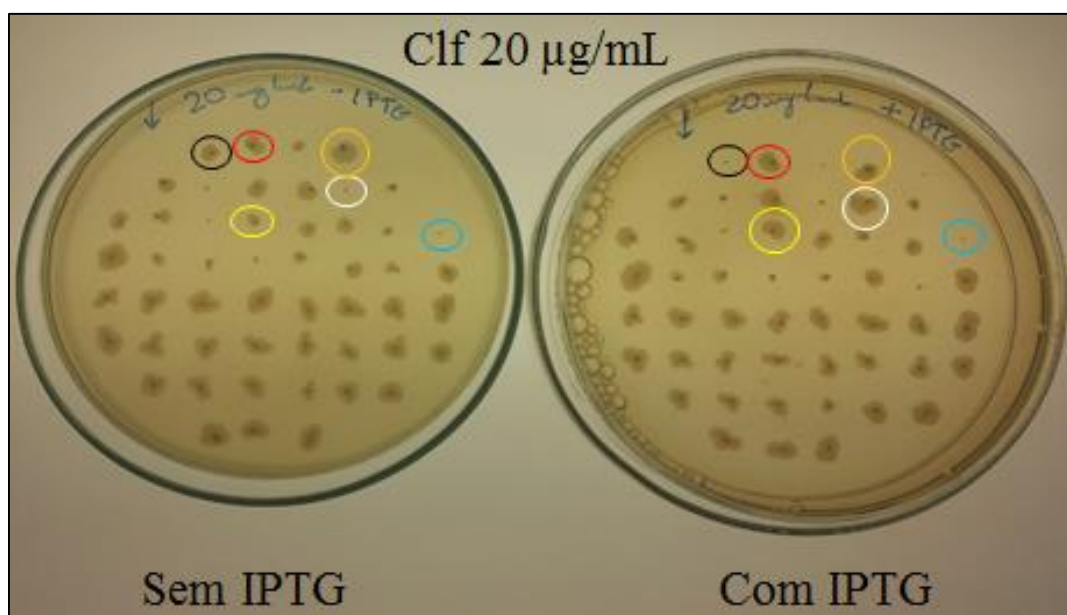


Figura 19: Fotografia dos clones em meio de cultivo com cloranfenicol na ausência e presença do indutor sintético IPTG. A figura mostra as diferentes características no crescimento dos clones com ou sem IPTG em cloranfenicol (círculos).

A partir das situações analisadas acima, foram realizadas outras réplicas-plate com novos clones em cultivos suplementados com cloranfenicol (20, 500, 1000 e 2000 µg/mL) na presença ou ausência do indutor IPTG (1 mM). A partir desta análise, os clones foram selecionados de acordo com suas maiores capacidades de crescer em cloranfenicol. Tais concentrações diferentes e os crescimentos dos clones estão mostrados no quadro 6. Os clones selecionados, foram então criopreservados em microtubos contendo glicerol e meio LB-líquido a -80 °C, e mantidos em placas de Petri com o antibiótico ampicilina e em meio soft-ágar em microtubo até o momento do uso.

Quadro 6: Clones diferentes em concentrações variadas de cloranfenicol com ou sem IPTG. O sinal de - (negativo), refere-se a clones que não cresceram e o sinal de positivo (+) se refere a clones que cresceram pouco (+), de maneira média (++) ou muito (+++).

Concentrações dos antibióticos	Identificações dos Clones	Sem IPTG	Com IPTG
20 µg/mL	1.A; 1.B; 1.C; 1.D; 1.E;	++	+
	1.F; 1.G;	++	-
	1.H; 1.I; 1.J; 1.K; 1.L;	+	-
500 µg/mL	2.A; 2.B; 2.C	-	+
	2.D; 2.E;	++	+++
	2.F; 2.J;	++	+
	2.G;	-	++
	2.H; 2.I,	+	+
	2.L; 2.K;	+	-
1000 µg/mL	3.A;	++	+++
	3.B; 3.C; 3.D; 3.E; 3.F; 3.G; 3.H; 3.I;	++	-
	3.J;	+	-
2000 µg/mL	4.A; 4.B; 4.C;	+	-
	4.D; 4.E; 4.F	+	+
	4.G; 4.H;	++	-
	4.I; 4.J; 4.K;	-	+

5.3 Sequenciamento

Foi observado primeiramente pela análise eletroforética que a maioria dos plasmídeos foram extraídos adequadamente mostrando um perfil típico de bandas. Alguns resultaram em pouco DNA ou não houve extração o que não permitiu que fossem utilizados no sequenciamento.

Visando confirmar a clonagem e determinar as sequências dos promotores construídos por síntese química, 42 dos plasmídeos extraídos foram então submetidos ao processo de sequenciamento automático de DNA pelo método de Sanger e Coulson utilizando o primer *Rv* (*reverso*) que se mostrou melhor para esse fim.

Após o sequenciamento foi avaliada a qualidade e o tamanho das sequências de cada fragmento pelo software *Electropherogram quality analysis* (<http://www.biomol.unb.br/phph/index.html>). Este software analisa as sequências geradas pelo sequenciador automático e dá a qualidade usando o programa PHRED.

Um total de 38 amostras resultaram em sequências de boa qualidade, tanto pela análise efetuada pelo programa do próprio sequenciador automático ABI (versão 3 – Applied Bioscience/Life Technologies) quanto pelo critério do programa PHRED. As sequências possuíam aproximadamente 900 pb e permitiram boa análise por bioinformática pois as sequências promotoras sintetizadas quimicamente deveriam possuir apenas 64 pares de bases.

5.4 Análise das sequências dos promotores sintéticos

Praticamente todas as sequências escolhidas para este trabalho possuem as regiões que constituem uma região promotora (a sequência “*Up element*”, a região -35, a região espaçadora de cerca de 17 pb, a região -10, a região entre a região -10 e o ponto de início da transcrição ou discriminador e o ponto de início da transcrição ou +1) (Figura 20).

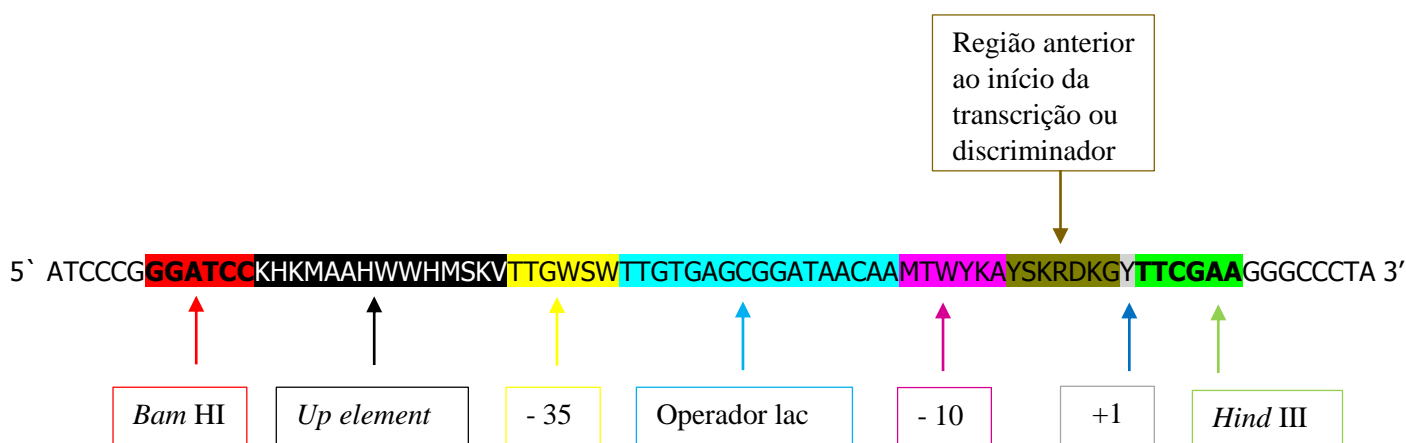


Figura 20: Desenho da sequência promotora sintética. Acima estão as principais sequências que compõem o promotor.

Para análise as sequências foram alinhadas manualmente tomando-se como referência a sequência do operador *lac* e os sítios de restrição das enzimas *HindIII* e

*Bam*HI que não deveriam variar no processo de construção dos novos promotores, ou seja, regiões constantes.

De todas as sequências, não houve clonagem do promotor sintético no múltiplo sítio de clonagem do vetor em clones que cresceram apenas em 20 µg/mL, esse fato não era esperado. Isso deve ter ocorrido pela baixa concentração do antibiótico (cloranfenicol) utilizado, ou pelo antibiótico utilizado estar parcialmente inativo.

A análise das sequências após o alinhamento revelou que diversas apresentavam a sequência do promotor sintético idênticas, por isso selecionou-se por alinhar e apresentar somente as sequências diferentes entre si, que estão apresentadas na figura 21.

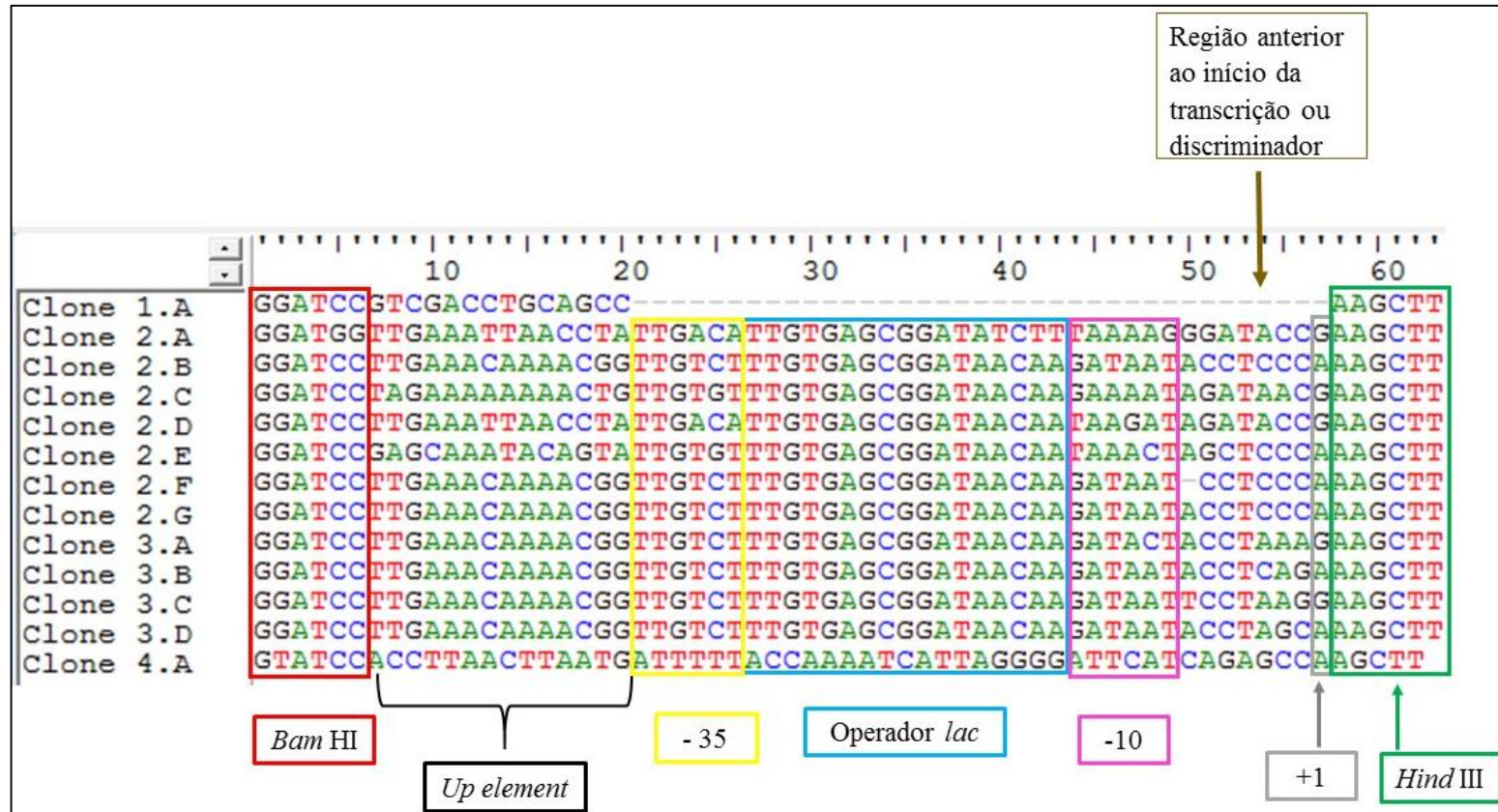


Figura 21: Figura das seqüências promotoras sintéticas. Na figura se observa o sítio da enzima de restrição de *Bam*HI em vermelho, a região de *Up element* em negro, a região -35 em amarelo, a região espaçadora está destacada em azul claro, a região -10 está destacada em rosa, a região do início do gene em cinza e a região marcada em verde exemplifica o sítio da enzima *Hind*III.

A seguir, os quadros abaixo mostram a sequência padrão, os diferentes tipos de sequências obtidas de diferentes regiões dos promotores sintéticos e o número de vezes que aparecem repetidas em todos os promotores sequenciados. Pode se observar que em algumas sequências obtidas, muitas bases nucleotídicas sofreram mutações (letras em vermelho) e algumas sequências não sofreram (letras em azul).

Quadro 7: Sequência *Up element*.

Sequência padrão	KHKM AA HWWH MSKV	Número de amostras obtidas
Sequências obtidas	TTGA AA CAAA ACGG	21
	CCTT AA CTTA ATGA	11
	TTG A AA TTAA CCTA	2
	TTG A AA CAAA ACGG	2
	TAGA AA AAAA ACTG	1
	GAGC AA ATAC AGTA	1

Quadro 8: Região -35.

Sequência padrão	TTGWSW	Número de amostras obtidas
Sequências obtidas	TTGTCT	21
	TTTTTA	11
	TTGTGT	2
	TTGACA	2
	TTGTCT	2

Quadro 9: Região -10.

Sequência padrão	MTWYKA	Número de amostras obtidas
Sequências obtidas	GAT AAT	22
	GAT ACT	1
	GAT TCA	11
	TAA AAG	1
	GAA AAT	1
	TAA GAT	1
	TAA ACT	1

Quadro 10: Sequência depois da região -10 e antes do ponto de início de transcrição.

Sequência padrão	YSKRDKG	Número de amostras obtidas
Sequências obtidas	ACCTCCC	16
	ACCTAAA	1
	ACCTCAG	1
	TCCTAAG	1
	ACCTAGC	1
	TCAGAGC	11
	GGATACC	1
	AGATAAC	1
	AGATACC	1
	AGCTCCC	1
	AGCTCCC	2
	CCTCCC	1

Tais variações entre as sequências promotoras foram analisadas e discutidas para dar continuidade no trabalho. Na região de “*Up element*” (Quadro 7) haviam 12 posições com bases degeneradas que podiam ser substituídos por duas ou três bases. Entretanto, obteve-se 6 sequências “*Up elements*” com degenerações diferentes das 20.736 possíveis. Na região -35 (Quadro 8) havia 3 posições com 3 bases degeneradas cada, dando um total de 27 possibilidades de sequências diferentes, neste aspecto foi possível obter apenas 5 sequências diferentes. Na região -10 (Quadro 9) havia 4 bases degeneradas que podiam ser substituídas por 2 bases cada, totalizando 16 possíveis combinações diferentes desta sequência, porém foram obtidas 7 possibilidades diferentes. Na região entre a própria região -10 e o ponto de início da transcrição ou discriminador (Quadro 10), obteve-se 12 diferentes possíveis sequências de 96 possibilidades ao todo, sendo que foram degeneradas 6 posições com 2 e 3 bases. E no ponto de início de transcrição obteve-se na sua maioria as bases A e G sendo que nas sequências dos clones em 2000 µg/mL detectou-se C, não previsto no desenho da sequência randômica devendo ser, portanto uma mutação.

Dos promotores sequenciados apenas dois apresentaram a sequência igual ao consenso TTGACA e curiosamente 21 promotores apresentaram a sequência TTGTCT. Nenhum promotor apresentou a sequência idêntica à região -10 consensual.

5.5 Análise da resistência a cloranfenicol pelos clones com diferentes sequências promotoras.

Com base nos resultados de resistência e indução por IPTG dos clones em meio LB-ágar (crescimento em cloranfenicol) e se fundamentando nas análises das sequências promotoras, foram escolhidos para novos testes de resistência 6 clones com suas respectivas sequências. As sequências escolhidas foram:

2.G

GGATCC TTGAAACAAAACGG TTGTCT TTGTGAGCGGATAACAA GATAAT ACCTCCCA AAGCTT

3.A

GGATCC TTGAAACAAAACGG TTGTCT TTGTGAGCGGATAACAA GATACT ACCTAAA GAAGCTT

3.B

GGATCC TTGAAACAAAACGG TTGTCT TTGTGAGCGGATAACAA GATAAT ACCTCAGAA AAGCTT

3.C

GGATCC TTGAAACAAAACGG TTGTCT TTGTGAGCGGATAACAA GATAAT TCCTAAGGA AAGCTT

3.D

GGATCC TTGAAACAAAACGG TTGTCT TTGTGAGCGGATAACAA GATAAT ACCTAGCA AAGCTT

4.A

GTATCC ACCTTA ACTTAATGATTTTTACCAAATCATTAGGGGATTCATCAGAGCC AAGCTT

5.5.1 Análise da resistência do cloranfenicol em meio sólido

Depois da escolha dos clones, os plasmídeos recombinantes com as 6 sequências promotoras foram reintroduzidos em *E. coli* DH5 α F[']IQ via transformação genética para resistência a cloranfenicol em meio LB-ágar (50 μ g/mL). Todos os 6 plasmídeos foram introduzidos na bactéria com uma eficiência de transformação da ordem de 10^5 / μ g de plasmídeo.

Posteriormente, prosseguiu-se com a análise em concentrações crescentes de antibiótico. Estes clones recombinantes foram repicados em concentrações de 50, 500, 1000 e 2000 µg/mL de cloranfenicol, com e sem IPTG (1 mM) para verificar se os promotores inseridos no pKK232-8 (clones recombinantes) seriam capazes de manter a capacidade de conferir igual resistência a cloranfenicol e serem induzidos por IPTG.

Na concentração de **50 µg/mL**, houve crescimento dos clones tanto na presença quanto na ausência do indutor IPTG com poucas diferenças no tamanho de colônia dos clones (Figura 22, sistema A).

Em **500 µg/mL** (concentração inicial), ocorreu um fato não esperado. Dos 6 clones plaqueados (sendo cada um dos seis plaqueados entre 4 e 6 vezes) todos os clones cresceram, porém com variações de tamanho das colônias do mesmo clone. Houve clones (a maioria) que cresceram com o mesmo tamanho nas placas que continham cloranfenicol com e sem IPTG (branco); clones que cresceram menos quando plaqueados em cloranfenicol sem IPTG (em negro) e clones que cresceram mais quando plaqueados em cloranfenicol com IPTG (vermelhos). Um maior crescimento com IPTG indica que o gene CAT estaria sendo induzido (Figura 22, sistema B).

Os clones em negro são os clones que cresceram melhor na presença de IPTG. Pode-se dizer então que o promotor deste clone pode ser classificado como mais forte e regulado, pois na presença do indutor eles conferem maior resistência contra o cloranfenicol. Os clones marcados em branco apresentaram o mesmo tamanho de colônia sem e com IPTG, isto significa que possivelmente a sequência do operador lac não estaria sendo modulada pelo IPTG.

Os clones nas concentrações de **1000 e 2000 µg/mL** tanto na presença e na ausência do indutor IPTG não apresentaram crescimento neste experimento. Isso pode ter sido consequência da utilização de um novo lote de cloranfenicol mais ativo.

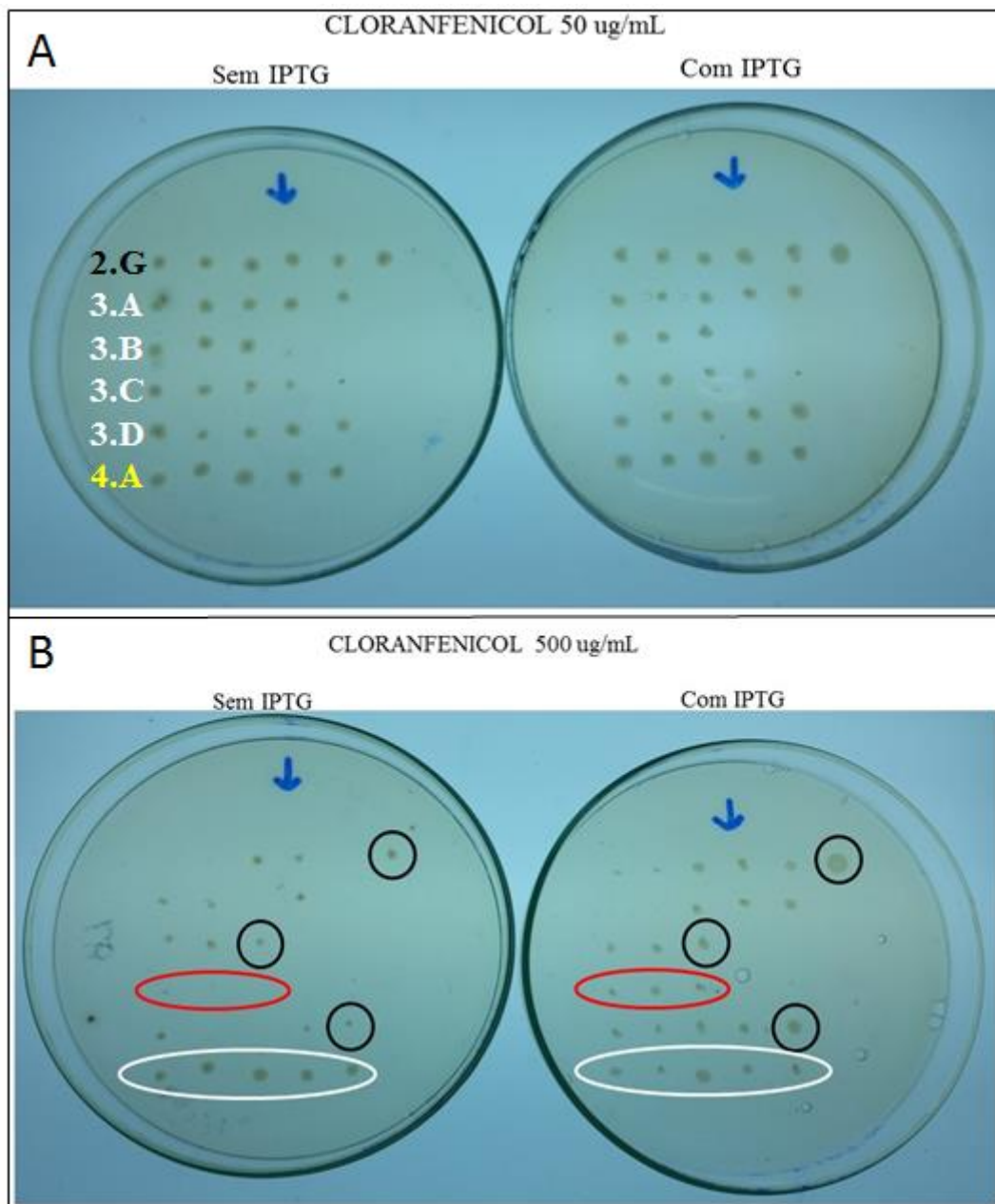


Figura 22: Fotografia dos clones recombinantes crescidos em placa de Petri contendo cloranfenicol (50 µg/mL e 500 µg/mL) com ou sem IPTG. No sistema A estão os clones que cresceram em 50 µg/mL e no sistema B os mesmos clones estão dispostos nas placas em concentração de 500 µg/mL. Os clones marcados em elipses vermelhas cresceram somente na presença do IPTG. Os clones marcados em preto cresceram melhor na presença de IPTG. Os clones em elipses brancas não mostraram grande diferença de crescimento de colônia e, portanto, cresceram sem e com IPTG.

5.5.2 Análise da resistência ao cloranfenicol em meio líquido.

Foi realizada uma análise em meio líquido quanto a resistência dos clones ao antibiótico cloranfenicol e a capacidade de os promotores serem modulados por IPTG.

Os gráficos a seguir mostram os valores de cada leitura de absorbância a 600 nm durante as doze horas de crescimento. O clone controle só com o vetor caça-promotor (pKK232-8) teve o resultado esperado, ou seja, não apresentou crescimento uma vez que esse crescimento só seria possível se no vetor tivesse clonado um promotor para a transcrição do gene de resistência a cloranfenicol (dados não mostrados).

Todos os seis clones que possuíam as suas respectivas sequências na análise em meio sólido foram primeiramente testados em meio líquido na concentração de 250 µg/mL de cloranfenicol na ausência e presença de IPTG (1 mM) em duplicata (dados não mostrados).

Após a análise das duplicatas, para aumentar a confiança experimental foram selecionados quatro clones com suas sequências promotoras e realizados experimentos em triplicatas. Após cada experimento em triplicata de cada clone, determinou-se a média e o desvio padrão. Tais clones possuíam as seguintes sequências promotoras selecionadas:

3.A

GGATCC TTGAAACAAAACGG TTGTCT TTGTGAGCGGATAACAA GATACT ACCTAAA GAAGCTT

3.B

GGATCC TTGAAACAAAACGG TTGTCT TTGTGAGCGGATAACAA GATAAT ACCTCAGAAAGCTT

3.D

GGATCC TTGAAACAAAACGG TTGTCT TTGTGAGCGGATAACAA GATAAT ACCTAGCAAAGCTT

4.A

GTATCC ACCTTAACTTAATGATTTTTACCAAAATCATTAGGGGATTCATCAGAGCC AAGCTT

Os clones **3.A**, **3.B** e **3.D** foram analisados e comparados quanto a sequência e ao crescimento, mostrando similaridade entre o crescimento quando cultivados com ou sem

IPTG em meio líquido, por esta razão apresentou-se apenas o gráfico do crescimento referente do clone **3.A** na figura 23.

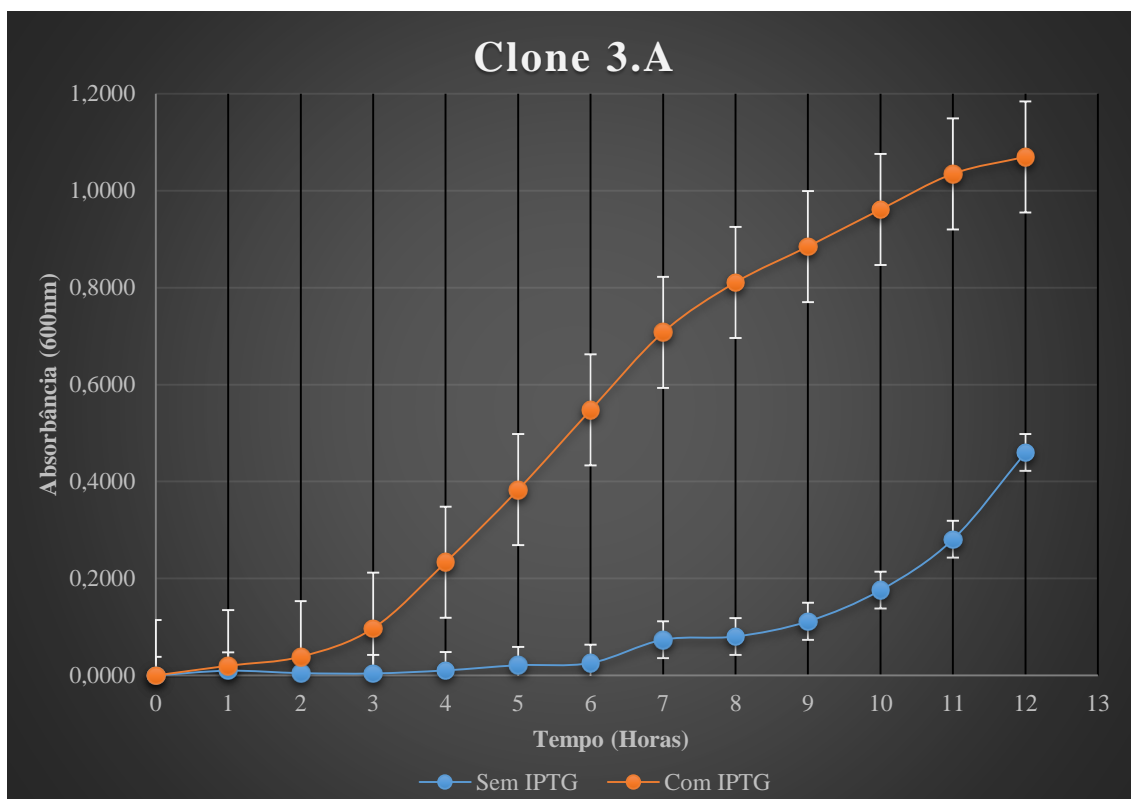


Figura 23: Gráfico que mostra o crescimento do clone 3.A na presença e ausência de IPTG. Em azul indica o crescimento na ausência de IPTG. Em laranja indica o crescimento na presença de IPTG.

Os clones **3.A**, **3.B** e **3.D** cresceram em todas as duas condições testadas (sem e com IPTG) com um perfil comparável durante as 6 horas iniciais de incubação. Todos os promotores clonados demonstraram ser fortes o suficiente para dar suporte ao crescimento da bactéria em meio líquido com **250 µg/mL** de cloranfenicol apresentando uma fase de latência com cerca de 3 horas na presença de IPTG e de cerca 9 horas na ausência de IPTG. Através dessas considerações, percebeu-se a diferença do crescimento com IPTG (em laranja) em relação ao controle sem IPTG (azul), tais resultados indicam que estes promotores clonados possuem força considerável e são regulados por IPTG.

Estes mesmos clones foram capazes de crescer em concentração de **500 µg/mL**, porém com fases de latência maior, porém somente após 22 horas de indução (dados não mostrados).

Foi observado que todos os clones **3.A**, **3.B** e **3.D** clones possuem a região de “*Up elements*”, a região -35 idênticas e a sequência do operador lac intacta e perfeita, sem nenhuma deleção ou inserção de bases. Essa estrutura deve estar permitindo que o repressor produzido pela hospedeira DH5 α F'IQ se ligue ao operador e controle a transcrição do gene CAT com eficiência.

Através destas metodologias utilizadas neste trabalho, nós não pudemos analisar se os potencializadores de transcrição (*Up element*) contribuíram para o aumento da expressão, para tal, experimentos adicionais independentes devem ser realizados. As sequências *Up element* apresentadas neste trabalho, presente em alguns promotores como nos clones que cresceram em 500 e 1000 μ g/mL mostraram uma rica composição de adeninas entre as posições -40 e -47, o que pode ter contribuído para potencializar a transcrição do gene CAT *in vivo* de forma análoga quando estes foram fusionados a outros promotores, tais como o *PlacUV5*, pois de acordo com Shawn, 1999 as sequências *Up element* proximais (posições -46 a -38; consenso 5'-AAAAAARNR-3') estimulam a transcrição em até 170 vezes (RAO, 1994; ESTREM, 1998; 1999; HAUGEN et al. 2008).

As regiões -35 e -10 destes clones neste estudo mostram-se diferentes das regiões -35 e -10 ditas consensuais (TGTACA e TATAAT), porém encontrou-se promotores que apresentam diversos nucleotídeos conservados, como na região -35, o que pode ter contribuído para a força desses promotores, assim como a afinidade pela subunidade σ^{70} . Ainda assim, algumas sequências nucleotídicas de promotores encontrados neste trabalho são idênticas as que são utilizados em vetores de expressão, como a região -10 do clone **3.A** que possui identidade com a região -10 do promotor λ PL (GATACT) e os clones **3.B** e **3.D** que possuem identidades com a região -10 do promotor λ PR (GATAAT) (SCHUMANN & FERREIRA, 2004). As região -10 dos clones **3.B** e **3.D** são idênticas diferindo apenas na quinta base do clone **3.A**.

Não foi possível, no desenho genético, introduzir a região operadora *Lac* completa entre as regiões conservadas -10 e -35 no desenho genético, pois o operador *Lac* principal O1 completo possui 21 pares de bases e sabendo que para *E. coli* o espaçamento usual entre as duas regiões -35 e -10 de promotores σ^{70} é de 17 pares de bases (KIRYU et al. 2004) optou-se em introduzir no desenho do promotor a sequência *core* (núcleo) do operador de 17pb, na esperança que o desenho de bases degeneradas, nas bordas do operador nas regiões -35 e -10, pudesse reconstituir em alguns casos o operador ativo.

O discriminador, região presente entre a região -10 e o ponto de início da transcrição (+1) possuem extrema importância para a expressão gênica *in vivo* e *in vitro* sendo fundamentais no envolvimento da regulação e estabilização do tempo de vida do complexo aberto (HAUGEN et al. 2008; RUFF et al. 2015). Tais sequências são reconhecidas pela sub-região $\sigma_{1.2}$ do sigma 70 em contato com a RNA polimerase. A maioria destas sequências variam de tamanho de 6 a 8 pares de bases e como o *Up element* a sequência de bases nesta região não é altamente conservada em promotores de *E. coli*. Estas sequências quanto mais ricas forem em C e G, resultam em tempos de vida curtos de complexos abertos ocorrendo a transcrição rapidamente sem sofrer muitas rodadas abortivas de transcritos curtos. (HAUGEN et al. 2008; RUFF et al. 2015). Neste trabalho, o discriminador mostrou alterações nos promotores a partir da quinta base nucleotídica e observou-se uma predominância de bases C e G nos discriminadores dos clones **3.B** e **3.D**. Estas sequências são idênticas nestes clones se diferenciando do clone **3.A**.

O ponto de início da transcrição é fundamental pois o afeta a cinética de iniciação de complexo fechado para o complexo aberto (RUFF et al. 2015). O conhecimento de cinética e seu mecanismo é necessário para saber como a sequência promotora funciona na presença de fatores regulatórios, ligantes, solutos/sais, inibidores, assim como antibióticos (RUFF et al. 2015). As bases mais comuns no ponto de início da transcrição são na ordem decrescente $A \geq G > T \gg C$, o que pode afetar a cinética de iniciação de complexo fechado para o complexo aberto (RUFF et al. 2015). Sendo que neste trabalho foi observado uma predominância para bases adeninas e guaninas como esperado pelo desenho experimental para todos os clones que cresceram nas concentrações de 500 $\mu\text{g/mL}$ e 1000 $\mu\text{g/mL}$, entretanto não foi possível detectar o sítio de início da transcrição dos clones que cresceram em 2000 $\mu\text{g/mL}$.

No caso dos promotores dos clones que cresceram em concentração de 2000 $\mu\text{g/mL}$ carecem de estudos mais aprofundados quanto a regulação, pois não responderam a indução por IPTG e em suas sequências não foi possível encontrar o operador lac conforme o desenho inicial dos promotores. Outro escolhido, o clone **4.A**, apresentou uma curva diferente dos outros clones analisados, como mostrado na figura 24. Este clone apresentou uma fase de latência menor que os outros clones com cerca de 2 horas (segundo ponto) crescendo até 12h. A sequência do promotor não apresenta o operador lac conforme desenhado e sintetizado, o que explica o fato do clone não ser afetado pela presença do IPTG (induzido). Não foi possível também identificar as regiões de *Up*

element, -10 e -35 em suas posições planejadas. Devido a essas profundas alterações na sequência desse promotor é possível que tenham ocorrido, durante o processo de clonagem, eventos de recombinação e mutação.

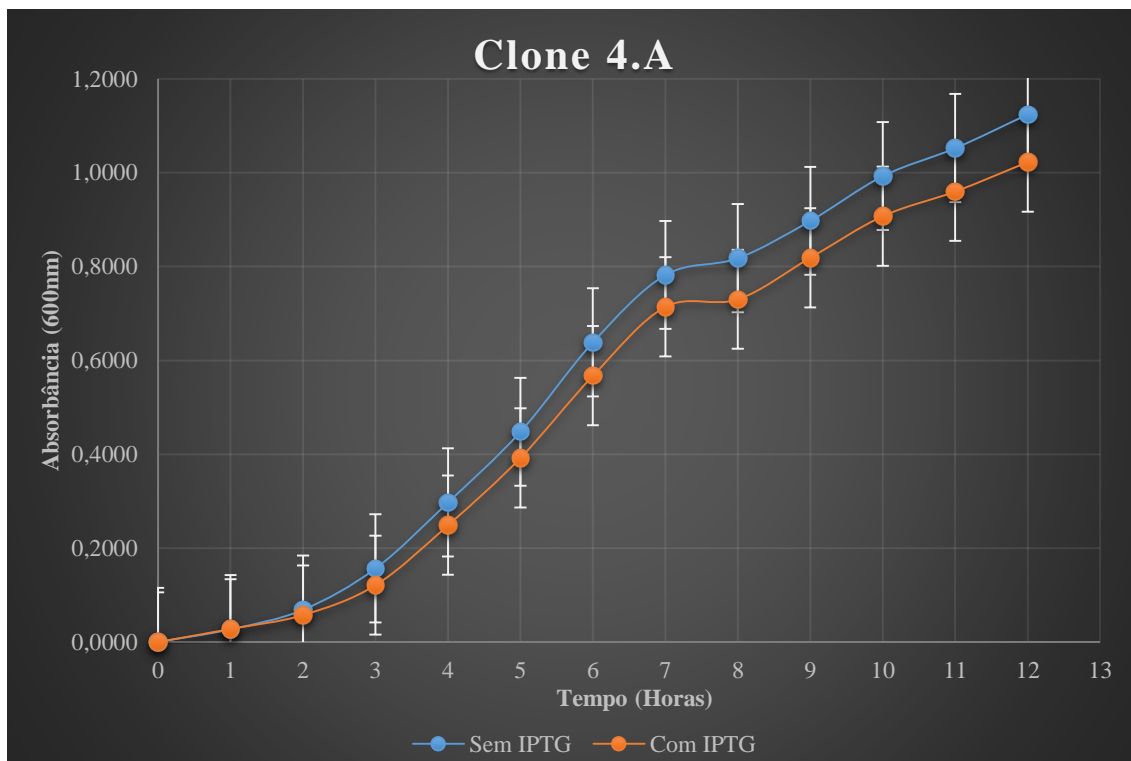


Figura 24: Gráfico que mostra o crescimento do clone 4.A na presença e ausência de IPTG. Em azul está indicado o crescimento na ausência de IPTG e em laranja na presença de IPTG.

Os resultados desse trabalho mostram que, conforme esperado, promotores regulados via operador *lac* puderam ser obtidos mesmo apresentando a sequência incompleta, diferente da sequência completa como descrita na literatura, como os clones de 3.A, 3.B e 3.D com a região operadora intacta situada entre a região -35 e -10 permitiu estericamente o acoplamento parcial da RNAPol no promotor na ausência do indutor sintético. Entretanto, na presença do indutor IPTG, este liga-se ao produto gênico do *lacI* (repressor) e este causa a dissociação do operador ocasionando a transcrição do gene CAT, ocasionando um modelo de regulação da expressão gênica (OEHLER et al. 1994).

6 CONCLUSÕES

Este trabalho reforçou que a biologia sintética e a biologia molecular podem juntas contribuir significativamente para a criação de bibliotecas de promotores visando a criação de novas sequências artificiais importantes, tanto contribuindo no enriquecimento do estudo de pesquisa científica e também como para aplicações biotecnológicas.

A estratégia que foi utilizada tem a potencialidade de gerar randomicamente muitas sequências promotoras diferentes e funcionais (330.225.942.528), sendo que algumas delas foram clonadas em *E. coli* utilizando o vetor caça-promotor pKK232-8 permitindo a expressão do gene CAT. De todas as sequências clonadas e analisadas, 44 promotores foram totalmente sequenciados e 6 deles foram analisados para estudar com mais detalhes se são regulados ou não por IPTG e como as células com os plasmídeos recombinantes tem a capacidade de crescer em diferentes concentrações de cloranfenicol.

Os promotores analisados em meio líquido com cloranfenicol na concentração de 250 µg/mL na ausência e presença de IPTG neste trabalho mostraram-se promissores para aplicações biotecnológicas, como por exemplo os clones: **3.A**, **3.B** e **3.D.**, no entanto, para uso destes promotores em vetores de expressão, haverá a necessidade de determinar diretamente a força de cada um e testá-los no que se refere a expressão de outros genes repórteres como por exemplo o da Proteína Fluorescente Verde (GFP). Além de estudar mais profundamente a sequência promotora do clone **4.A** que apresentou capacidade de resistência a cloranfenicol, porém não apresentou indução por IPTG.

Nenhum promotor clonado e sequenciado apresentou as duas regiões -35 e -10 idênticas as regiões consensuais (TGTACA e TATAAT), sendo que em apenas um promotor a sequência TTGACA foi encontrada. Dos clones analisados, os que apresentaram as regiões operadoras conforme desenhadas foram capazes de expressar o gene repórter de uma forma regulada, ou seja, induzida por IPTG.

Neste trabalho demonstrou-se um bom controle da expressão do gene CAT oferecidos por promotores sintéticos e é provável também que estes promotores possam ser funcionais em outros procariotos como *Lactococcus* (JENSEN & HAMMER, 1998) *Bacillus* (AMAYA et al. 2001), *Lactobacillus* (RUD et al. 2006), *Streptomyces* (SEGHEZZI et al. 2011) e *Corynebacterium* (YIM et al. 2013), que são algumas bactérias de grande interesse industrial.

BIBLIOGRAFIA DE REFERÊNCIA

ALI, H. & MURRELL, J. C. Development and validation of promoter-probe vectors for the study of methane monooxygenase gene expression in *Methylococcus capsulatus* Bath. *Microbiology*, 155: 761-771 March 2009, doi: 10.1099/mic.0.021816-0.

ALLEN, G. S & FRANK, J. Structural insights on the translation initiation complex ghosts of a universal initiation complex. *Mol Microbiol* **63**: 941-950. 2007.

ALPER, H.; FISCHER, C.; NEVOIGT, E.; STEPHANOPOULOS, G. Tuning genetic control through promoter engineering. *PNAS*, vol. 103. No. 8. **3007**. 2006.

AMAYA, E.; KHVOROVA, A.; PIGGOT, P. J. Analysis of Promoter Recognition In Vivo Directed by σ^F of *Bacillus subtilis* by Using Random-Sequence Oligonucleotides. *JOURNAL OF BACTERIOLOGY*, p. 3623–3630, June. 2001.

ANDRIANANTOANDRO, A.; BASU, S.; KARIG, D. K.; WEISS, R. Synthetic biology: new engineering rules for an emerging discipline. *Molecular Systems Biology*. 2006.

ARIO de MARCO. Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microbial Cell Factories*, **8**:26. 2009.

ASSCHE E. V.; VAN PUYVELDE, S.; VANDERLEYDEN J.; STEENACKERS H. P. RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front. Microbiol.* **6**:141. doi: 10.3389/fmicb.2015.00141. 2015.

ASTOLFI-FILHO, S.; AZEVEDO, J. L.; AZEVEDO, M. O. Genética de microrganismos em Biotecnologia e Engenharia genética. Piracicaba: Fealq, 1985.

BAKKE, I.; BERG, L.; AUNE, T. E.; BRAUTASET, T.; SLETTA, H.; TONDERVIK, A.; VALLA, S. Random mutagenesis of the PM promoter as a powerful strategy for improvement of recombinant-gene expression, *Applied Environmental Microbiology* **75** 2009.

BANEYX, F AND MUJACIC, M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology* **22**, 1399 – 1408. 2004. doi: 10.1038/nbt1029

BATAUS, L. A. M. Construção de um vetor de seleção de promotores (VSP) para *Escherichia coli*. Dissertação de Mestrado, Universidade de Brasília, Brasília, Distrito Federal. 87pp. 1991.

BATISTA, M. Construção de vetores para caracterização de genes de *Trypanosoma cruzi* em um sistema para clonagem em alta demanda. 124f. Dissertação (Mestrado em Biologia Celular e Molecular) - Universidade Federal do Paraná, Curitiba. 2008.

BEREZHNOY A. Y. & SHCKORBATOV Y. G. Dependence of the *E. coli* promoter strength and physical parameters upon the nucleotide sequence. *Journal of Zhejiang University SCIENCE*, **6B (11)**: 1063-1068. 2005.

BERLEC, A & STRUKELJ, B. Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells. *J Ind Microbiol Biotechnol* **40**:257–274. February 2013.

BICH, L; MOSSIO, M.; RUIZ-MIRAZO, K.; MORENO, A. Biological regulation: controlling the system from within. *Biol Philos*. SPRINGER. 2015. doi 10.1007/s10539-015-9497-8

BLOUNT, B. A.; WEENINK, T.; VASYLECHKO, S.; ELLIS, T. Rational Diversification of a Promoter Providing Fine-Tuned Expression and Orthogonal Regulation for Synthetic Biology. Rational Diversification of a Promoter Providing Fine-Tuned Expression and Orthogonal Regulation for Synthetic Biology. *PLoS ONE* **7**(3): e33279. doi:10.1371/journal.pone.0033279. 2012.

BOSDRIESZ, E.; MOLENAAR, D.; TEUSINK, B.; BRUGGEMAN, F. J. How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization *FEBS Journal* **282**:2029–2044. 2015.

BORUKHOV, S. & NUDLER, E. RNA polymerase holoenzyme: structure, function and biological implications. *Current Opinion in Microbiology*. **6**:93–100. 2003.

BROSIUS, J.; ERFIE, M.; STORELLA, JONH. Spacing of the -10 and -35 regions in the *tac* promoter. *The journal of the biological chemistry*. Vol. 260. No. 6. 1985.

BROWNING, D. F. & BUSBY, S. J. The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**:57–65. 2004.

BURDEN, S.; LIN, Y. X.; ZHANG, R. Improving promoter prediction for the NNPP2.2 algorithm: a case study using *Escherichia coli* DNA sequences. vol. 21 no. 5. 601–607. 2005.

CAMPBELL, E. A.; WESTBLADE, L. F.; DARST, S. A. Regulation of bacterial RNA polymerase σ factor activity: a structural perspective. *Current Opinion in Microbiology*, **11**:121–127. 2008.

CAMSUND, D.; HEIDORN, T LINDBLAD, P. Design and analysis of LacI-repressed promoters and DNA-looping in a cyanobacterium. *Journal of Biological Engineering*, **8**:4. 2014.

CASES I.; DE LORENZO V. Promoters in the environment: transcriptional regulation in its natural context. *Nat Rev Microbiol*. **3**(2):105-18, 2005.

CHEN, H.; BJERKNES, M.; KUMAR, R.; JAY, E. Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Research*, Vol. 22, No. 23 4953-4957. 1994.

CHEN, Y. J.; LIU, P.; NIELSEN, A. A. K.; BROPHY, J. A. N.; CLANCY, K.; PETERSON, T.; VOIGT, C. A. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nature methods*. Vol.10 NO.7. July, 2013.

CHOI, J. K. & LEE, S. Y. Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl Microbiol Biotechnol*. 64: 625–635. 2004.

CIAMPI, M. S. Rho-dependent terminators and transcription termination. *Microbiology* 152:2515-2528. 2006.

COHEN, S. N. DNA cloning: A personal view after 40 years. *PNAS*. vol. 110. no. 39 September 2013.

DERKX, P. M. F.; JANZEN, T.; SØRENSEN, K. I. CHRISTENSEN, J. E.; STUER-LAURIDSEN, B.; JOHANSEN, E.; The art of strain improvement of industrial lactic acid bacteria without the use of recombinant DNA technology. *Microbial Cell Factories*, **13**(Suppl 1):S5. 2014.

DO CARMO, E. J. Clonagem e expressão regulada do cDNA da glicoamilase de *Aspergillus awamori* em *Pichia Pastoris*. Instituto Nacional de Pesquisas da Amazônia. Programa de pós-graduação em genética, conservação e biologia evolutiva. Dissertação de Mestrado, 2010.

DOMBROSKI, A. J.; JOHNSON, B. D.; LONETTO, M.; GROSS, C. A. The sigma subunit of *Escherichia coli* RNA polymerase senses promoter spacing. *Proc. Natl. Acad. Sci. USA* Vol. 93, pp. 8858-8862, *Biochemistry*. August, 1996.

DRENNAN, A.; KRAEMER, M.; CAPP, M.; GRIES, T.; RUFF, E.; SHEPPAR, P.; WIGNESHWERARAJ, S.; ARTSIMOVITCH, I.; THOMAS RECORD JR, M. Key Roles of the Downstream Mobile Jaw of *Escherichia coli* RNA Polymerase in Transcription Initiation. **51**(47): 9447–9459. 2012.

EPSHTEIN, V.; DUTTA, D.; WADE, J.; NUDLER, E. An allosteric mechanism of Rho dependent transcription termination. *Nature*, 463:245-249, 2010.

ERMOLAEVA, M. D.; KHALAK, H. G.; WHITE, O.; SMITH, H. O.; SALZBERG, S. L. Prediction of transcription terminators in bacterial genomes. *J Mol Biol.*, 301:27-33, 2000.

ESTREM, S. T.; ROSS, W.; GAAL, T.; CHEN, Z. W. S.; NIU, W.; EBRIGHT, R. H.; GOURSE, R. L.; Bacterial promoter architecture: subsite structure of *Up elements* and interactions with the carboxy-terminal domain of the RNA polymerase α subunit. *GENES & DEVELOPMENT* 13:2134–2147 1999.

FEKLISTOV, A. RNA polymerase: in search of promoters. *Ann N Y Acad Sci*. 1293: doi:10.1111/nyas.12197. 2014.

FERRER-MIRALLES, N.; DOMINGO-ESPÍN, J.; CORCHERO, J. L.; VÁZQUEZ, E.; VILLAVERDE, A. Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories* 8:17. 2009. doi:10.1186/1475-2859-8-17.

FOURNIER, B.; GRAVEL, A.; HOOPER, D. C.; ROY, P. H. Strength and Regulation of the Different Promoters for Chromosomal β -Lactamases of *Klebsiella oxytoca*. *ANTIMICROBIAL AGENTS AND CHEMOTHERAPY*, p. 850–855. 1999.

GEROSA, L. & SAUER U. Regulation and control of metabolic fluxes in microbes. *Current Opinion in Biotechnology*. **22**:566–575. 2011.

GESZVAIN, K. AND LANDICK, R. The structure of bacterial RNA polymerase. In *The Bacterial Chromosome*, Higgins, N. P. (ed.) Washington, DC> American Society for Microbiology Press, pp. 283-296. 2005.

GOPAL, G. J. & KUMAR. A. Strategies for the Production of Recombinant Protein in *Escherichia coli*. *Protein J*, **32**:419–425. 2013.

GRUET, A.; LONGHI, S.; BIGNON, C. One-step generation of error-prone PCR libraries using Gateway® technology. *Microbial Cell Factories* **11**:14. 2012.

GUSAROV, I. & NUDLER E. Control of Intrinsic Transcription Termination by N and NusA: The Basic Mechanisms. *Cell*, Vol. 107, 437–449, 2001.

GUALERZI, C. O.; PON, C. L. Initiation of mRNA translation in bacteria: structural and dynamic aspects. *Cell. Mol. Life Sci.* 72:4341–4367. 2015. doi 10.1007/s00018-015-2010-3

GUSTAFSSON, C.; GOVINDARAJAN, S.; MINSHULL, J. Codon bias and heterologous protein Expression. *TRENDS in Biotechnology* Vol. 22 No.7. 2004.

HAGG, P.; DEPOHL, J. W.; ABDULKARIM, F. & ISAKSSON, L. A. A host/plasmid system that is not dependent on antibiotics and antibiotic resistance genes for stable plasmid maintenance in *E. coli*. *J. Biotechnology*. v. 111. p. 17-30. 2004.

HARLEY, D. K.; REYNOLDS, R. P. Analysis of *E. coli* promoter sequences. *Nucleic acids res.* 15: 2343-2361. 1987.

HART, D. J. & TARENDEAU, F. Combinatorial library approaches for improving soluble protein expression in *Escherichia coli*. *Acta Cryst.* **D62**, 19–26. 2006

HARTLEY, J. L. Cloning technologies for protein expression and purification. *Current Opinion in Biotechnology*. 17:359–366. 2006.

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*, AOP, 2008.

HAWLEY, D. K.; MCCLURE, W. R. Compilation and analysis of *E. coli* promoter DNA sequences. *Nucleic acids res.* 11: 2237-2255. 1983.

HENRIQUE, M. P.; PALAZZO, I. C. V.; ZANELLA, C. R.; DARINI, A. L. C. Molecular characterization of enterococci harboring genotype and phenotype incongruence related

to glycopeptide resistance isolated in Brazilian hospitals. *Mem Inst Oswaldo Cruz, Rio de Janeiro*, Vol. 103(3): 301-305, May 2008.

HEYDUK, E. & HEYDUK, T. Next Generation Sequencing-based Parallel Analysis of Melting Kinetics of 4096 Variants of a Bacterial Promoter. *Biochemistry*. **53(2)**: 282–292. doi:10.1021/bi401277w. 2014.

HIJUM, S. A. F. T. V.; MEDEMA, M. H.; KUIPERS, O. P. Mechanisms and Evolution of Control Logic in Prokaryotic Transcriptional Regulation. *Microbiology And Molecular Biology Reviews*. p. 481–509. vol. 73, No. 3. 2009. doi:10.1128/MMBR.00037-08.

HOCHKOEPLER, A. Expanding the landscape of recombinant protein production in *Escherichia coli*. *Biotechnol Lett*, **35**:1971–1981. 2013.

HOEK, M. V. & HOGEWEG, P. The Effect of Stochasticity on the Lac Operon: An Evolutionary Perspective. *PLoS Computational Biology*, vol. 3. Issue 6. e111. 2007.

HOOK-BARNARD, I. G.; HINTON, D. M. The promoter spacer influences transcription initiation via -70 region 1.1 of *Escherichia coli* RNA polymerase. *PNAS*. vol. 106. no. 3. 737–742. 2009

HUANG CHUNG-JR.; LIN, H.; YANG, X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol*, **39**:383–399. 2012.

INTERFARMA. Associação da Indústria Farmacêutica do Brasil. >www.interfarma.org.br. 2012.

JANA, S. & DEB, J. K. Strategies for efficient production of heterologous proteins in *Escherichia coli*. *Applied and Microbiological Biotechnology*, **67**: 289–298. 2004.

JENSEN, P. R.; HAMMER, K. The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl Environ Microbiol*. **64(1)**: 82–87. 1998

JENSEN, P. R.; HAMMER, K. Artificial promoters for metabolic optimization. *Biotechnol Bioeng*. **58**:191-195. 2000.

JESSE, J. & BLODGETT, K. Max Efficiency DH5 α F'IQ: A new M13 and phagemid host. *Focus* **10**, 69.

JI MA & HOWE, M. M. The Phage Mu Middle Promoter Pm Contains a Partial *Up element*. *Genes & Genomes & Genetics*. Vol. 5. 507. 2015.

JIN, H.; ZHAO, Q.; VALDIVIA, E. I. G.; ARDELL, D. H.; STENSTRÖM, M.; ISAKSSON, L. A.; Influences on gene expression in vivo by a Shine–Dalgarno sequence. *Molecular Microbiology*, **60** (2), 480–492. 2006.

JONASSON, P.; LIJEQVIST, S., NYGREN, P. A.; STHAL, S. Genetic design for facilitated production and recovery of recombinant proteins in *E. coli*. *Biotechnol. Appl. Biochem.* 35: 910105. 2002.

JULLESSON D.; DAVID, F.; PFLEGER, B.; NIELSEN, J. Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals, *Biotechnol Adv.* 2015.

JURGEN BROSIUS. Plasmid vectors for the selection of promoters. *Gene*, 27. 151-160. 1984.

KAGIYA, G.; OGAWA, R.; HATASHITA, M.; TAGAGI, K.; KODAKI, T.; HIROISHI, S.; YAMAMOTO, K. Generation of a strong promoter for *Escherichia coli* from eukaryotic genome DNA. *Journal of Biotechnology* 115, 2005.

KINGSFORD, C. L.; AYANBULE, K.; SALZBERG, S. L. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol.*, **8**:R22, 2007.

KIRYU, H.; OSHIMA, T; ASAI, K. Extracting relations between promoter sequences and their strengths from microarray data. *Bioinformatics*, **21**(7), 1062 – 1068. 2004.

KNIRSCHOVA, R.; NOVAKOVA, R.; MINGYAR, E.; BEKEOVA, C.; HOMEROVA, D.; KORMANEC, J. Utilization of a reporter system based on the blue pigment indigoidine biosynthetic gene *bpsA* for detection of promoter activity and deletion of genes in *Streptomyces*. *Journal of Microbiological Methods*. **113**:1–3. 2015.

LEAVITT, J. M. & ALPER S. H. Advances and current limitations in transcript-level control of gene expression. *Current Opinion in Biotechnology*, **34**:98–104. 2015

LEE, D. J.; MINCHIN, S. D.; BUSBY S. J. W. Activating Transcription in Bacteria. *Annu. Rev. Microbiol.* **66**:125–52. 2012.

LEONARD, S. R.; LACHER, D. W.; LAMPEL K. A. Acquisition of the *lac* operon by *Salmonella enterica*. *BMC Microbiology*, **15**:173. 2015.

LESNIK, E. A.; SAMPATH, R.; LEVENE, H. B.; HENDERSON, T. J.; MCNEIL, J. A.; ECKERA, D. J. Prediction of rho-independent transcriptional terminators in *Escherichia coli*. *Nucleic Acids Res.*, 29:3583–3594, 2001.

LEWIN, A.; MAYER, M.; CHUSAINOW, J.; JACOB, D.; APPEL, B. Viral promoters can initiate expression of toxin genes introduced into *Escherichia coli*. *BMC*, **5**:19. *Biotechnology*. 2005.

LEWIS, M. Allosterity and the *lac* Operon. *J. Mol. Biol.* **425**, 2309–2316. 2013.

LI, R.; XU, L.; SHI, H. Strategy of tuning gene expression ratio in prokaryotic cell from perspective of noise and correlation. *Journal of Theoretical Biology* **365**:377–389. 2015.

MALYS, N. & MCCARTHY, J. E. G. Translation initiation: variations in the mechanism can be anticipated. *Cell. Mol. Life Sci.* **68**:991–1003. 2011.

MARQUES, M. V. *Biologia molecular e genética bacteriana / Ribeirão Preto: Sociedade Brasileira de Genética.* 348. II. 2012.

MENG, W.; BELYAEVA, T.; SAVERY, N. J.; BUSBY, S. W. J.; ROSS, W. E.; GAAL, T.; GOURSE, R. L.; THOMAS, M. S. *Up element*-dependent transcription at the *Escherichia coli* *rrnB* P1 promoter: positional requirements and role of the RNA polymerase α subunit linker. *Nucleic Research.* vol. 29. no. 20. 2001.

MIKSCH, G.; BETTENWORTH, F.; FRIEHS, K.; FLASCHEL, E. The sequence upstream of the – 10 consensus sequence modulates the strength and induction time of stationary-phase promoters in *Escherichia coli*. *Applied and Microbiological Biotechnology*, 69: 312- 320. 2005.

MIRZADEH, K.; MARTÍNEZ, V.; TODDO, S.; GUNTUR, S.; HERRGARD, J. M.; ELOFSSON, A.; NØRHOLM, H. H. M.; DALEY*, O. D. Enhanced Protein Production in *Escherichia coli* by Optimization of Cloning Scars at the Vector–Coding Sequence Junction. *ACS Synth. Biol.*, 4, 959–965. 2015.

MITCHELL, E. J.; ZHENG, D; BUSBY, S. J. W.; MINCHIN, S. D. Identification and analysis of `extended -10' promoters in *Escherichia coli*. *Nucleic Acids Research.* Vol. 31, No. 16 4689±4695. 2003

MITRA, A.; ANGAMUTHU, K.; JAYASHREE, H.V.; NAGARAJA, V. Occurrence, divergence and evolution of intrinsic terminators across eubacteria. *Genomics*, **94**:110-116, 2009.

MOLBAK, L.; TETT, A.; USSERY, D. W.; WALL, K.; TURNER, S.; BAILEY, M.; FIELD, D. The Plasmid Genome Database. *Microbiology*, 149: 3043-3045, November 2003 doi: 10.1099/mic.0.C0123-0

MOGK, A.; MAYER, M.P.; DEUERLING, E. Mechanisms of protein folding: molecular chaperones and their application in biotechnology. *Chembiochem* 3, 807–814. 2002.

MRUK, I.; RAJESH, P.; BLUMENTHAL, R. M. Regulatory circuit based on autogenous activation-repression: roles of C-boxes and spacer sequences in control of the PvuII restriction-modification system. *Nucleic Acids Research*, Vol. 35, No. 20. **6935–6952**. 2007.

MURAKAMI, S. K. & DARST, S. A. Bacterial RNA polymerase: the wolo story. *Current Opinion In Structural Biology.* **13**:31-39. 2003

MUTALIK, V. K.; GUIMARAES, J. C.; CAMBRAY, G., LAM, C.; HRISTOFFERSEN, M. J.; MAI, Q. A.; TRAN, A. B.; PAULL, M.; KEASLING, J. D.; ARKIN, A. P.; ENDY, D. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* 10, 354–360. 2013.

NEIVA, M. Expressão em *Escherichia coli* de quitinase (E. C. 3.2.1.14) de *Chromobacterium violaceum*/Márcia Neiva – Manaus. Dissertação de Mestrado. 2005.

OEHLER, S.; EISMANN, R. E.; KRAMER, H.; MULLER-HILL, B. The three operators of the *lac* operon cooperate in repression, *The EMBO Journal* vol. 9 no. 4 pp. 973-979, 1990.

OEHLER, S.; AMOUYAL, M.; KOLKHOF, P.; WILCKEN-BERGMANN, B. V.; MULLER-HILL, B. Quality and position of the three *lac* operators of *E. coli* define efficiency of repression. *The EMBO Journal* vol. 13 no.14 pp. 3348-3355, 1994.

OMOTAJO, D.; TATE, T.; CHO, H.; CHOUDHARY, M. Distribution and diversity of ribosome binding sites in prokaryotic genomes. *BMC Genomics*, **16**:604. 2015.

OVERTON, T. W. Recombinant protein production in bacterial hosts. *Drug Discovery Today*. vol. 19. no. 5. May 2014.

PACHECO, B.; CROMBET, L.; LOPPNAU, P.; COSSAR, C. A screening strategy for heterologous protein expression in *Escherichia coli* with the highest return of investment. *Protein Expression and Purification* **81**. 33–41. 2012.

PASOTTI, L. & ZUCCA, S. Advances and Computational Tools towards Predictable Design in Biological Engineering. Hindawi Publishing Corporation. Article ID 369681, 16 pages. 2014.

PENUMETCHA, P.; LAU, K.; ZHU, X.; DAVIS, K.; ECKDAHL, T. T.; CAMPBELL T. T. Improving the Lac System for Synthetic Biology. *BIOS* **81(1)** 7 – 15. 2010.

PETERS, J. M.; VANGELOFF, D. A.; LANDICK, R. Bacterial Transcription Terminators: The RNA 3'-End Chronicles. *Journal of Molecular Biology*. *J. Mol. Biol.* **412**, 793–813. 2011.

PETI, W. & PAGE, R. Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expression and Purification*, **51**:1–10. 2007.

PRIBNOW, D. Bacteriophage T7 early promoters: nucleotide sequence of two RNA polymerase binding sites. *J. Mol. Biol.* **99**: 419-443. 1975.

PURVES, W.K.; SADAVA, D.; GORDON, H.O.; CRAIGH ELLER, H. Vida: A ciência da biologia. trad. Ana Paula Somer Vinagre et al. 6ed. Porto Alegre, Artmed, 2002.

RAMON, A. & SMITH, H. O. Single-step linker-based combinatorial assembly of promoter and gene cassettes for pathway engineering. *Biotechnol Lett*, **33**:549–555. 2011.

RAO, L.; ROSS, W.; APPLEMAN, J. A.; GAAL, T.; LEIRMO, S.; SCHLAX, P. J.; RECORD, M. T.; GOURSE, R. L. Factor independent activation of *rrnB* P1-an “extended” promoter with an upstream element that dramatically increases promoter strength. *J Mol Biol* **235**:1421-1435. 1994.

RAVISHANKAR, S.; AMBADY1, A.; RAMU, H.; MUDUGAL, N. V.; TUNDUGURU, R.; ANBARASU, A.; SHARMA, U. K.; SAMBANDAMURTHY, V. K.; RAMAIAH, S. An IPTG Inducible Conditional Expression System for Mycobacteria. PLOS ONE | DOI:10.1371/journal.pone.0134562. 2015.

RICHARDSON, J.P. Transcription termination. Crit. Rev. Biochem. Mol. Biol. 28: 1993.

RICHARDSON, J. P. Rho-dependent termination and ATPases in transcript termination. Biochem. Biophys. Acta, 1577:251-260, 2002.

RODRIGUE, S.; PROVVEDI, R.; JACQUES, P. E.; GAUDREAU, L.; MANGANELLI, R. The sigma factors of *Mycobacterium tuberculosis*. FEMS Microbiol. Rev., v. 30, p. 926 - 941, 2006.

ROBERTS, W. J.; SHANKAR, S.; FILTER, J. J. RNA Polymerase Elongation Factors. Annu Rev Microbiol. 62: 211.2008.

ROSANO, G. L. CICCARELLI, E. A. R. Recombinant protein expression in *Escherichia coli*. 1. 5. 172. April 2014.

RUD, I.; JENSEN, P. R.; NATERSTAD, K.; AXELSSON, L.; A synthetic promoter library for constitutive gene expression in *Lactobacillus plantarum*. Microbiology, 152, 1011–1019. 2006.

RUFF, E. F.; M. THOMAS RECORD, JR.; ARTSIMOVITCH, I. Initial Events in Bacterial Transcription Initiation. *Biomolecules*. 5, 1035-1062. 2015.

SACCO, Laís Postai. Isolamento de bactérias produtoras de enzimas de interesse em processos biotecnológicos. 2013. vii, 47 f. Dissertação (mestrado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Ciências Agrárias e Veterinárias, 2013. Disponível em: <<http://hdl.handle.net/11449/94874>>.

SACHDEVA, P., MISRA, R., TYIAGI, A.K., SINGH, Y. The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. FEBS J., v. 277, p. 605-626, 2009.

SAECKER, R. M.; THOMAS RECORD JR, M.; AND PIETER L. DEHASETH. Mechanism of Bacterial Transcription Initiation: RNA polymerase - Promoter Binding, Isomerization to Initiation- Competent Open Complexes, and Initiation of RNA Synthesis. J Mol Biol. 412(5): 754–77. 2011.

SAHDEV, S.; KHATTAR, S. K.; SAINI, K. S. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol Cell Biochem 307:249–264. 2008. doi 10.1007/s11010-007-9603-6

SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. Molecular Cloning. A Laboratory Manual. Two ed. Cold Spring Harbor: ColdSpring Harbor Laboratory Press, 2001.

SANCHEZ-GARCIA, L.; MARTÍN, L.; MANGUES, R.; FERRER-MIRALLES, N.; VÁSQUEZ, E.; VILLAVARDE, A. Recombinant pharmaceuticals from microbial cells: a update. *Microb Cell Fact.*15:33. 2016. DOI 10.1186/s12934-016-0437-3. 2015.

SANTANGELO, T.J.; CUBONOVÁ, L.; SKINNER, K.M.; REEVE, J.N. Archaeal intrinsic transcription termination in vivo. *J. Bacteriol.*, 191:7102-7108, 2009.

SANTOS, C. N. S. & STEPHANOPOULOS, G. Combinatorial engineering of microbes for optimizing cellular phenotype. *Current Opinion in Chemical Biology* **12**:168–176. 2008.

SEGHEZZI, N.; AMAR, P.; KOEBMANN, B.; & JENSEN, P. R.; MARIE-JOËLLE, V. The construction of a library of synthetic promoters revealed some specific features of strong *Streptomyces* promoters. *Appl Microbiol Biotechnol*, 90:615–623. 2011.

SCHUMANN, W. AND FERREIRA, L. C. S. Production of recombinant proteins in *Escherichia coli*. *Genet. Mol. Biol.*, vol.27, no.3, p.442-453. ISSN 1415-4757. 2004.

SCHOLZ, H. C.; RIEDMANN, E.; WITTE, A.; LUBITZ, W.; KUEN, B. S-Layer Variation in *Bacillus stearothermophilus* PV72 Is Based on DNA Rearrangements between the Chromosome and the Naturally Occurring Megaplasms. *J Bacteriol. Mar*; **183**(5): 1672–1679. 2001.

SEMSEY, S.; JAUFFRED, L.; CSISZOVSKI, Z.; ERDOSSY, J.; STÉGER, V.; HANSEN, S.; KRISHNA, S. The effect of LacI autoregulation on the performance of the lactose utilization system in *Escherichia coli*. *Nucleic Acids Research*, Vol. 41, No. 13 6381–6390. 2013.

SHIMADA, T.; YAMAZAKI, Y.; TANAKA, K.; ISHIHAMA, A. The Whole Set of Constitutive Promoters Recognized by RNA Polymerase RpoD Holoenzyme of *Escherichia coli*. *PLoS ONE* **9**(3): e90447. doi:10.1371/journal.pone.0090447. 2014.

SHAWN, T. E.; GAAL, T.; ROSS, W.; GOURSE, R. L. Identification of an *Up element* consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci. Biochemistry*. Vol. 95. pp. 9761-9766. 1999.

SHINE, J. & DALGARNO, L. The 3'-Terminal sequence of *E. coli* 16 S ribosomal RNA: Complementary to nonsense triplets and ribosome binding sites. *PNAS* 71, 1342-1346. 1974.

SHOJAOSADATI, S. A.; KOLAEI, S. M. V.; BABAEIPOUR, V. & FARNOUD, A. M. Recent advances in high cell density cultivation for production of recombinant protein. Review article. *Iranian Journal of Biotechnology*, v. 6. n. 2. p. 63-84. 2008.

SINGH, V. Recent advancements in synthetic biology: Current status and challenges. *Gene*, 535. 1–11. 2014.

SIVASHANMUGAM, A.; MURRAY, V.; CUI, C.; ZHANG, Y.; WANG, J.; LI, Q. Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Science*. Vol **18**:936—948. 2009.

SKANCKE, J.; BAR, N.; KUIPER, M.; HSU, M. L. Sequence-Dependent Promoter Escape Efficiency Is Strongly Influenced by Bias for the Pretranslocated State during Initial Transcription. *Biochemistry*. **54**, 4267–4275. 2015.

SMILLIE, C.; GARCILLÁN-BARCIA, M. P.; FRANCIÀ, V. M.; ROCHA, E. P. C.; CRUZ, F.; Mobility of Plasmids. *Microbiology and Molecular Biology Reviews*. p. 434–452. Vol. 74, No. 3. Sept. 2010. doi:10.1128/MMBR.00020-10.

SORENSEN, H.P.; MORTENSEN, K.K. Advanced genetic strategies for recombinant expression in *Escherichia coli*. *Journal of Biotechnology*, 115:113–128. 2005.

THOMAS, C. M. Paradigms of plasmid organization. *Molecular Microbiology* 37(3), 485-491. 2000.

TRIPATHI, N. K.; SHRIVASTVA, A; BISWAL, K. C.; RAO, P. V. L. Optimization of culture medium for production of recombinant dengue protein in *Escherichia coli*. Review methods. *Industrial Biotechnology* v. 5(3). p.179-183. 2009.

WAEGEMAN, H. & SOETAERT, W. Increasing recombinant protein production in *Escherichia coli* through metabolic and genetic engineering. *J Ind Microbiol Biotechnol*, **38**:1891–1910. 2011.

WALSH, G. Post-translational modifications of protein biopharmaceuticals. *Drug Discovery Today*. Volume 15, Numbers 17/18. September 2010.

YACHIE, N.; ARAKAWA, K.; TOMITA, M. On the interplay of gene positioning and the role of rho-independent terminators in *Escherichia coli*. *FEBS Lett*, 580:6909- 6914, 2006.

YI-YWAN, M. C.; HUI-RU S.; CHENG-TZER L.; SHIN-YI L. Properties and Construction of Plasmid pFW213, a Shuttle Vector with the Oral *Streptococcus* Origin of Replication. *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, p. 3967–3974. 2011.

YIM, S. S.; AN, S. J.; KANG, M.; LEE, J.; JEONG, K. J. Isolation of Fully Synthetic Promoters for High-Level Gene Expression in *Corynebacterium glutamicum*. *Biotechnology and Bioengineering*, Vol. 110, No. 11, November, 2013.