



Uma breve história de promotores bacterianos: da sequência e mecanismo regulatório à aplicação biotecnológica¹

Diego da Silva Moreira², Kerollen Runa Pinto³, Maria Clara Tavares Astolfi⁴ e Spartaco Astolfi-Filho⁵

Resumo

Nas células procarióticas, o principal nível da regulação da expressão gênica é o transcricional. Nesse processo, existe um *networking* onde componentes moleculares interagem para controlar a leitura de informações genéticas em resposta a uma gama de sinais ambientais e endógenos. Em nível do DNA, a sequência que determina o início da transcrição pela RNA polimerase é denominada de promotor. Esta revisão é dedicada ao: (I) início do processo de transcrição em *Escherichia coli*, (II) à arquitetura de sequências promotoras e como a RNA polimerase interage com tais sequências, (III) e à regulação desse processo. São abordados os mecanismos regulatórios de promotores que compõem cassetes de expressão de vetores amplamente utilizados em engenharia genética para produção de proteínas recombinantes, como os P_{Lac} , P_{Tac} , P_{araBAD} , P_L , P_R , P_{T7} . Embora a edição gênica, engenharia metabólica, terapia genética e a constante demanda por aprimoramento dos processos biotecnológicos industriais tenham estimulado grande avanço nessa área, o descobrimento de novos promotores naturais, ou design de promotores sintéticos, continua uma necessidade. Nesse sentido, a compreensão dos mecanismos já descritos por décadas de esforço científico pode abrir caminho e facilitar a decodificação e engenharia de novos componentes moleculares que atuam em nível transcricional, dando um passo adiante rumo às aplicações biotecnológicas.

Palavras-Chave: Transcrição, promotores bacterianos regulados, vetores de expressão, proteínas recombinantes de interesse industrial.

A brief history of bacterial promoters: from sequence and regulatory mechanism to biotechnology application. In prokaryotic cells, the main level of regulation of gene expression is the transcriptional. In this process, networking exists where molecular components interact to control the reading of genetic information in response to a range of environmental and endogenous signals. At the DNA level, the sequence that determines the start of transcription by RNA polymerase is called a promoter. This review is dedicated to: (I) initiation of the transcription process in *Escherichia coli*, (II) the architecture of promoter sequences and how RNA polymerase interacts with such sequences, (III) and the regulation of that process. We discuss the regulatory mechanisms of promoters that make up vector expression cassettes widely used in genetic engineering to produce recombinant proteins, such as P_{Lac} , P_{Tac} , P_{araBAD} , P_L , P_R , P_{T7} . Although genetics, metabolic engineering, gene therapy, and the constant demand for improvements in industrial biotechnology have spurred great strides in this area, the discovery of new natural promoters, or design of synthetic promoters, remains a necessity. In this sense, the understanding of the mechanisms already described by decades of scientific effort can open the way and facilitate the decoding and engineering of new molecular components that act at the transcriptional level, taking a step forward towards biotechnological applications

Keywords: Transcription, regulated bacterial promoters, expression vectors, recombinants proteins of industrial interest.

¹ Parte das Revisões Bibliográficas das Dissertações de Mestrado, dos dois primeiros autores, do Programa de Pós-Graduação em Biotecnologia na Universidade Federal do Amazonas

² Doutorando em Biotecnologia na Universidade Federal do Amazonas. Lab. de Tecnologias de DNA. Centro de Apoio Multidisciplinar. Universidade Federal do Amazonas. Manaus – AM. E-mail: diego.biotec02@gmail.com

³ Doutoranda em Biotecnologia na Universidade Federal do Amazonas. Lab. de Tecnologias de DNA. Centro de Apoio Multidisciplinar. Universidade Federal do Amazonas. Manaus – AM. E-mail: kerollen_runa@hotmail.com

⁴ Biotecnóloga pela Universidade Federal do Amazonas. Lab. de Tecnologias de DNA. Centro de Apoio Multidisciplinar. Universidade Federal do Amazonas. Manaus – AM. E-mail: mctastolfi@gmail.com

⁵ Professor Titular de Engenharia Genética. Departamento de Genética. Instituto de Ciências Biológicas Universidade Federal do Amazonas. 69077-000. Manaus – AM. E-mail: spartaco.biotec@gmail.com



1. Introdução

O avanço da Biologia Molecular e da Tecnologia do DNA Recombinante, a habilidade de editar e criar novos circuitos genéticos abriu caminho para uma gama de descobertas e uma revolução na biotecnologia; desde a compreensão da função de conjuntos gênicos, produção de biomoléculas em escala industrial à modelagem de complexas vias metabólicas (COHEN, 2013; XIAO et al., 2014; JAJESNIAK & WONG, 2015). Entretanto, embora tenha havido progresso nas quatro décadas desde o advento da Engenharia Genética, a alta produção de proteínas recombinantes em bactérias continua sendo um grande desafio (ITAKURA et al., 1977; BRAUN & LABAER, 2003).

Não é uma tarefa simples tornar realidade a superexpressão de genes no laboratório e menos simples ainda, é transferir a eficiência obtida no laboratório para a produção em escala industrial. Um dos principais desafios é o *design* preciso da sequência genética que codifica a função desejada na dinâmica do maquinário celular e a previsão do seu funcionamento. Nesse sentido, uma série de novas metodologias foram desenvolvidas para facilitar esse processo desde o nível de clonagem e expressão gênica, até a purificação e análise da proteína recombinante em questão (CHOU, 2007; OVERTON, 2014; PAPANEOPHYTOU & KONTOPIDIS, 2014; ROSANO & CECCARELLI, 2014).

As primeiras proteínas recombinantes eram produzidas pela simples clonagem de um fragmento de DNA (gene) ou cDNA (cópia do mRNA) em um plasmídeo multicópia (vetor) e esse DNA recombinante era introduzido em uma célula hospedeira, normalmente uma bactéria ou uma levedura (HARTLEY, 2006). Esses trabalhos iniciais foram importantes para consolidar a metodologia da clonagem, mas obviamente nem sempre garantiam grandes ganhos em relação ao nível de expressão. A síntese de proteínas recombinantes em altos níveis começou a se tornar realidade graças ao desenvolvimento de sistemas eficientes de expressão ou expressão/secreção compostos de vetores e suas respectivas hospedeiras (SØRENSEN et al., 2005; ROSANO & CECCARELLI, 2014).

Um dos principais componentes de um vetor de expressão é o “cassete de expressão”, um conjunto de sequências funcionais de DNA que

garante que genes heterólogos de interesse sejam expressos em um determinado hospedeiro. Um “cassete de expressão” é composto, principalmente por: 1) um promotor que sinaliza o início da síntese do mRNA pela RNA polimerase holoenzima (que na maioria dos vetores utilizados é passível de regulação); 2) um sítio de ligação do ribossomo (RBS ou *Ribosome Biding Site*) para sinalizar o início da tradução; 3) uma região de múltiplos sítios de clonagem (*polylinker*) para permitir a introdução das mensagens genéticas a serem expressas (também chamadas de *CoDing Sequence* ou CDS); 4) e um terminador de transcrição – para interromper as sínteses dos mRNAs pela RNA polimerase (GUSAROV & NUDLER, 2001; CRAMER, 2002; VAN 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 procariotos, o início da transcrição é a etapa mais relevante e estudada do processo de regulação da expressão gênica. Nesse sentido, a escolha do promotor ideal que irá compor o cassete de expressão é uma etapa fundamental do *design* genético com vistas a uma expressão de genes heterólogos com sucesso (HALL & COLLIS, 1995; MAKRIDES, 1996; BANEYX, 1999; SWARTZ, 2001; JANA & DEB, 2005).

Os vetores de expressão funcionais em *Escherichia coli*, para a produção de proteínas recombinantes em altos níveis, necessitam conter em seus cassetes de expressão promotores rigorosamente regulados, pois a constante superexpressão de genes heterólogos, com o acúmulo da respectiva proteína recombinante, pode ser prejudicial ou mesmo letal para as células hospedeiras, especialmente quando são tóxicas e inviabilizam o bioprocessamento como um todo (SCHUMANN & FERREIRA, 2004; SAIDA et al., 2006; MARSCHALL et al., 2016;).

Esses elementos genéticos que promovem a transcrição podem ser obtidos por clonagem diretamente de sequências de DNA de genomas naturais, podem ser fusionados para construção de promotores híbridos ou ainda, mais usualmente, elaborados com a ajuda de softwares que, além de possibilitarem a montagem dessas sequências, podem também realizar a análise e simulações computacionais para a melhor opção em cada caso



e subsequente solicitação de síntese química de DNA (JULLESSON et al., 2015).

Levando em consideração a importância dos promotores, a presente revisão de literatura tem por objetivo destacar os componentes que constituem a minuciosa regulação e expressão de informações genéticas em nível de transcrição, indicando os sistemas de expressão de proteínas recombinantes mais utilizados historicamente e descrevendo as vantagens e desvantagens do uso de cada um deles. Assim, almeja-se facilitar o *design* de novos cassetes de expressão funcionais tanto para uso em escala laboratorial como para a produção de proteínas recombinantes de interesse industrial em altos níveis.

2. Metodologia

Esta revisão de literatura levanta as principais informações sobre promotores procarióticos e suas principais aplicações em processos biotecnológicos. Foram consultadas as fontes bibliográficas disponíveis na internet como: *Scopus*, *Scirus*, *Pubmed*, *Chemical Abstract*, *SciELO*, *ScienceDirect* e portal de periódico da CAPES dentre outros sítios de fontes oficiais confiáveis. A análise crítica das informações levantadas foi com base na experiência adquirida pelos dois primeiros autores no curso de mestrado do Programa Multi-Institucional de Pós-Graduação em Biotecnologia (PPGBIOTEC) da Universidade Federal do Amazonas-(UFAM), bem como na experiência do grupo de Biotecnologia Molecular na construção e uso dos vetores de expressão. Foram consultados os principais artigos atuais porém valorizando também os artigos de referência histórica na área.

3. A transcrição gênica nos procariotos

A decodificação das informações genéticas contidas nas sequências de DNA dos genes inicia-se pelo processo de transcrição, ou síntese de ácidos ribonucleicos (RNAs), dentre os quais o mRNA carrega a mensagem transcrita do DNA para a síntese proteica (tradução da sequência de mRNA) (BAKSHI et al., 2015).

A transcrição nos procariotos não é um processo isolado, mas é fruto de uma orquestração no qual uma rede de eventos se dedica a síntese de diferentes tipos de RNA que acabam se encontrando na derradeira missão: “ler” o mRNA. Os principais componentes da transcrição do mRNA são: 1) uma enzima multimérica, a RNA

polimerase, que polimeriza o mRNA a partir da informação da fita de DNA molde; 2) um promotor, a sequência de DNA que indica o ponto de início da transcrição de um gene; 3) fator sigma (σ), peça central no processo de iniciação, responsável por direcionar a RNA polimerase ao reconhecimento dos elementos presentes nos promotores; 4) fatores de transcrição (TFs, do inglês *Transcription Factors*), que atuam na regulação como repressores ou ativadores deste processo dinâmico, controlando o momento que o mRNA será sintetizado de acordo com sinais físico-químicos ambientais ou celulares; 5) sequências operadoras, domínios nos promotores selecionados pela evolução, onde os reguladores (*i.e.* repressores ou ativadores) interagem e se ligam; 6) e, por fim, o terminador de transcrição, que sinaliza o fim da transcrição de um gene e desacopla a RNA polimerase (CRAMER, 2002; HAUGEN et al., 2008; VAN HIJUM et al., 2009; SAECKER et al., 2011; LEE et al., 2012; BROWNING & BUSBY, 2004; 2016).

O processo de transcrição nos procariotos é dividido em quatro etapas estruturalmente distintas: o reconhecimento da região promotora no DNA, a iniciação, o alongamento e o término do processo (BORUKHOV & NUDLER, 2003; GEZVAIN & LANDICK, 2004; MURAKAMI, 2015).

Para que o início da transcrição ocorra RNA polimerase, apoenzima que sintetiza RNA a partir de um molde de DNA na presença de ribonucleotídeos trifosfatados, deve primeiramente se associar a outra subunidade proteica, conhecida como fator sigma (σ). O complexo formado pela RNA polimerase e o fator sigma é chamado de RNA polimerase holoenzima. A RNA polimerase associada com um fator sigma é capaz de reconhecer um promotor e então iniciar o processo de transcrição (NUDLER, 2009; ROSS & GOURSE, 2009).

Ao alcançar o final da região a ser transcrita, domínios presentes na sequência do DNA e do RNA levam ao término da transcrição, fazendo com que o complexo DNA-RNA polimerase e RNA mensageiro sejam desfeitos. Tais domínios são chamados de terminadores de transcrição e são reconhecidas pela RNA polimerase. Este processo pode ocorrer pela ação de fatores proteicos que auxiliam na sinalização para a terminação da transcrição (como os Rho-dependentes, por exemplo) ou pelo simples



desequilíbrio das forças de pareamento que mantêm o híbrido DNA e RNA mensageiro juntos (ALBRECHTSEN et al., 1991; NUDLER & GOTTESMAN, 2002; RICHARDSON, 2002; CIAMPI, 2006; WASHBURN & GOTTESMAN, 2015).

3.1 A RNA polimerase de procariotos

A RNA polimerase bacteriana é composta por pelo menos cinco subunidades que compreendem a porção catalítica do complexo enzimático e pode ser encontrada em duas formas: a) a apoenzima ou o núcleo (cerne catalítico) da enzima, constituída pelas subunidades α , β , β' e ω a holoenzima que possui todas as subunidades, inclusive o fator σ , capaz de reconhecer promotores gênicos específicos e iniciar a síntese de RNA. Ao contrário das demais, a subunidade α está presente em duas cópias (MURAKAMI et al., 2002; MURAKAMI & DARST, 2003; LANE & DARST, 2010; BAE et al., 2015).

A RNA polimerase da bactéria Gram-negativa *Escherichia coli* é uma das mais estudadas e a melhor caracterizada dentre os procariotos. As duas subunidades alfa (α) possuem 40 kDa cada. A subunidade beta (β) possui em torno de 155 kDa e a beta linha (β') possui 160 kDa. A subunidade ômega (ω) possui 10 kDa e os fatores sigmas (σ) possuem massa molecular variando de 25 a 92 kDa. A subunidade α é formada por dois domínios de aminoácidos independentes e unida por um ligante flexível de 20 aminoácidos: o domínio N-terminal (9 kDa) com resíduos de 1 a 235 aminoácidos e o C-terminal (26 kDa) composto por resíduos de 250 a 329 aminoácidos. As subunidades β e β' possuem 1,342 e 1,407 aminoácidos, respectivamente e a subunidade ω (11 kDa) apresenta 91 resíduos de aminoácidos (BURGESS, 1969; ISHIHAMA, 2000; MURAKAMI et al., 2002; RUFF et al., 2015).

Cada subunidade tem um papel específico na ligação ao DNA, na interação entre as subunidades da enzima e no processo catalítico. O domínio N-terminal das subunidades α é responsável pela montagem da enzima, pois está envolvido na interação entre as duas subunidades α (formando um dímero) e na ligação de β com β' , já o C-terminal é responsável pela interação com sequências específicas de DNA, presentes a montante de alguns promotores, denominados elementos UP, e com fatores de transcrição

(AIYAR et al., 1998; BURGESS & ANTHONY, 2001; YAN & FONG, 2017).

A subunidade β contém o sítio catalítico ativo da polimerase e juntamente com β' especifica a interação do σ (sigma) com o promotor. Além disso, esta subunidade se liga aos ribonucleotídeos que irão ser polimerizados em RNA. A subunidade ω (ômega) atua como uma “chaperona molecular” para o correto enovelamento da subunidade β' e auxilia na montagem da RNA polimerase, estabilizando o complexo enzimático. As subunidades α também auxiliam neste processo ao conduzir o arranjo espacial das subunidades β e β' na estrutura da RNA polimerase. O fator σ consiste na subunidade “livre” da RNA polimerase e se associa ao seu cerne por meio de interações com β e β' (DARST, 2001; MINAKHIN et al., 2001; VASSYLYEV et al., 2002).

3.2 Fatores σ

A etapa central de identificação de um promotor e o início da transcrição pela RNA polimerase é dependente de uma subunidade específica: o fator sigma (σ). Esta subunidade tem como função o reconhecimento de promotores de um conjunto de genes que precisam ser expressos em determinado estado ou fase celular do metabolismo bacteriano (GRUBER & GROSS, 2003). Existem, pelos menos, seis tipos diferentes de fator σ em *E. coli*, que respondem a sinais ambientais ou celulares distintos. Somente quando ocorre a associação do cerne da RNA polimerase ao fator sigma é que a holoenzima se torna apta a reconhecer e se ligar de forma específica à região promotora alvo e iniciar de fato o processo de transcrição do gene adjacente (MURAKAMI & DARST, 2003; ÖSTERBERG et al., 2011; FEKLÍSTOV et al., 2014; RANGEL-CHAVEZ et al., 2017).

Com exceção da pequena família σ^{54} , todos os outros fatores sigmas compartilham características em comum (GUO et al., 2000). São proteínas que possuem até quatro domínios diferentes e estão envolvidos no reconhecimento dos promotores de genes em resposta a diferentes situações, como: carência nutricional; aumento de temperatura; fase estacionária; estresse oxidativo; estresse osmótico; transporte de íons; entre outros (Tabela 1). Assim, torna-se possível a um organismo apresentar uma resposta adequada a esses estímulos ambientais e sobreviver a certas condições adversas utilizando a mesma enzima



(RNA polimerase) e quando necessário ajustando somente o fator σ (GRUBER & GROSS, 2003; TRIPATHI et al., 2014).

Todas as bactérias têm um fator sigma primário que é suficiente para o crescimento em condições normais na presença dos nutrientes essenciais. A maioria dos fatores sigma bacterianos possuem homologia significativa com o sigma 70

($\sigma 70$) de *E. coli* e portanto, pertencem a família $\sigma 70$. Outros pertencem à família sigma 54 ($\sigma 54$), responsável pela expressão de genes envolvidos na utilização de nitrogênio (MERRICK, 1993; HELMANN, 2001; WIGNESHWERARAJ et al., 2008; GHOSH et al., 2010).

Tabela 1: Fatores sigmas (σ) de *Escherichia coli*.

Fator sigma (σ)	Gene	Função/Resposta	Seq. Consenso -35 Espaçador -10
24	rpoE	Estresse térmico extra-citoplasmático (ECF)	GGAACCTT- 15 pb- GTCTAA
28	FliA ou rpoF	Mobilidade celular e patogenicidade ou síntese de flagelo	CTAAA-15pb-GCCGATAA
32	rpoH	Choque térmico extremo; renaturação de proteínas	CCCTTGAA -13-15pb - CCCGATNT
38	rpoS	Resposta a estresse e fase estacionária	TTGACA- 16-18 pb- TATACT
54	rpoN	Assimilação ou limitação de nitrogênio (N ₂)	CTGGNA -6pb- TTGCA
70	rpoD	Manutenção basal; crescimento celular, sigma constitutivo	TTGACA - 16-18pb- TATAAT

Fonte: Adaptado de HELMANN, 2001.

Em *E. coli*, o $\sigma 70$ atua principalmente a transcrição dos genes constitutivos (*housekeeping*) ou dos genes de manutenção do metabolismo basal da bactéria. Por conta disso, possui afinidade maior pelo núcleo da RNA polimerase quando em comparação a fatores σ alternativos, sendo também o mais abundante na célula (LONETTO et al., 1992). À medida que os fatores σ alternativos se acumulam, passam a competir de forma dinâmica com o $\sigma 70$ pelo núcleo da RNA polimerase para iniciar a transcrição de promotores-alvo. Os promotores bacterianos reconhecidos pelo $\sigma 70$ apresentam sequências de DNA bem conservadas localizadas a -35 e -10 pares de bases localizados a montante do sítio em que se inicia a transcrição (PAGET & HELMANN, 2003; KAZMIERCZAK et al., 2005).

O fator $\sigma 70$ pode ser dividido em quatro domínios proteicos particulares. Cada domínio interage com a RNA polimerase e com o DNA, simultaneamente. Os domínios 4 e 2 ligam-se especificamente às regiões (ou hexâmeros) -35 e -

10, respectivamente. O domínio 3 interage com o elemento -10 estendido (TGn) presente em alguns promotores. Como visto, os domínios 2, 3 e 4 estão envolvidos no reconhecimento do promotor, entretanto a função do domínio 1 não é bem compreendida, ausente muitos fatores σ supõe-se que esteja envolvido na regulação do complexo aberto com a sequência de DNA, denominada de discriminador (Figura 1) (PAGET & HELMANN, 2003; SCHUMANN & FERREIRA, 2004; SAECKER et al., 2011).

3.3 Promotores procarióticos reconhecidos pelo fator $\sigma 70$

Os promotores procarióticos reconhecidos pelo fator $\sigma 70$ apresentam tamanho aproximado de 60~80 pares de bases. Estão localizados a montante do ponto de início de transcrição. Desta maneira, controlam os níveis celulares das proteínas codificadas por estes genes através de sua taxa de transcrição. Estes promotores possuem regiões específicas que atuam como sítios de

reconhecimento e ancoragem capazes de recrutar a RNA polimerase holoenzima (α 2, β , β' , ω e σ) e iniciar a síntese de RNA no sentido 5' para 3'. (BURGESS, 1969; ZHANG et al., 1999 CAMPBELL et al., 2001; LEE et al., 2012).

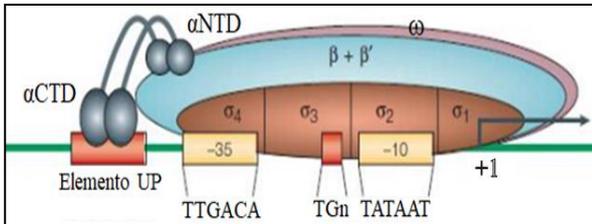


Figura 1: A holoenzima RNA polimerase em interação com o promotor gênico. Esquema do modelo do reconhecimento dos elementos que compõem a sequência promotora pela RNA polimerase holoenzima (α 2, β , β' , ω , e σ) baseado em estudos cristalográficos. A fita verde sendo reconhecida pela RNA polimerase holoenzima representa o DNA interagindo com os quatro domínios proteicos (σ 1, σ 2, σ 3 e σ 4) da subunidade σ 70. São mostradas as sequências de consenso para as regiões -35 (TTGACA), elemento -10 estendido (TGn) e região -10 (TATAAT) (Fonte: adaptado de BROWNING & BUSBY, 2004).

O passo principal na iniciação da transcrição é o reconhecimento do promotor pela RNA polimerase holoenzima. Os estudos cristalográficos levaram a geração de modelos que explicam como determinadas regiões da sequência de DNA, que compõem o promotor, são reconhecidos pela subunidade sigma. Estas regiões foram encontradas inicialmente através de análises computacionais comparativas das sequências de promotores, observando que, em certas regiões dos promotores, haviam sequências conservadas e por isso foram denominadas de consensuais. Desde essas análises iniciais considerava-se que, por serem conservadas, essas regiões deveriam ter papéis relevantes no início do processo de

transcrição (HAWLEY & MCCLURE, 1983; HARLEY & REYNOLDS, 1987; VASSYLYEV et al., 2002; BEREZHNOY & SHCKORBATOV, 2005; BURDEN et al., 2005).

Com a evolução dos procedimentos de sequenciamento de DNA, uma vasta quantidade de sequências de promotores de *E. coli* e de outros procarionotes foram determinadas, sendo estas reconhecidas pela RNA polimerase por meio do fator σ 70. Os dados de sequenciamento, juntamente com os gerados por outras estratégias permitiram a proposição do modelo apresentado na Figura 1 (BEREZHNOY & SHCKORBATOV, 2005; FEKLÍSTOV et al., 2014).

As regiões consensuais detectadas desde o início das análises eram dois hexâmeros, situados nas posições -35 (TTGACA) e -10 (TATAAT) em relação ao ponto de início da transcrição (+1). Geralmente, em promotores fortes foi observado o espaçamento de 17 nucleotídeos entre os hexâmeros consensuais -35 e -10. Posteriormente, foi confirmado por cristalografia de Raios-X que o espaçamento de 17 nucleotídeos era o ideal para o acoplamento da subunidade sigma σ 70 (HAWLEY & MCCLURE, 1983; HARLEY & REYNOLDS, 1987; ZHANG et al., 1999; SHIMADA et al., 2014; MURAKAMI, 2015;).

A região -35 (TTGACA) atua como sinal para que a RNA polimerase reconheça a sequência promotora, enquanto que a região -10 (TATAAT) atua na abertura das fitas de DNA e na conversão do complexo fechado para aberto. Estas sequências consensuais são consideradas as principais para o reconhecimento de um promotor e para determinar sua força, entretanto existem outras regiões de relevância que atuam no processo, como o elemento UP, o elemento -10 estendido, a sequência discriminadora e o sítio de início da transcrição ou +1 (Figura 2) (PRIBNOW, 1975; ENGSTROM & PFLEGER, 2017).

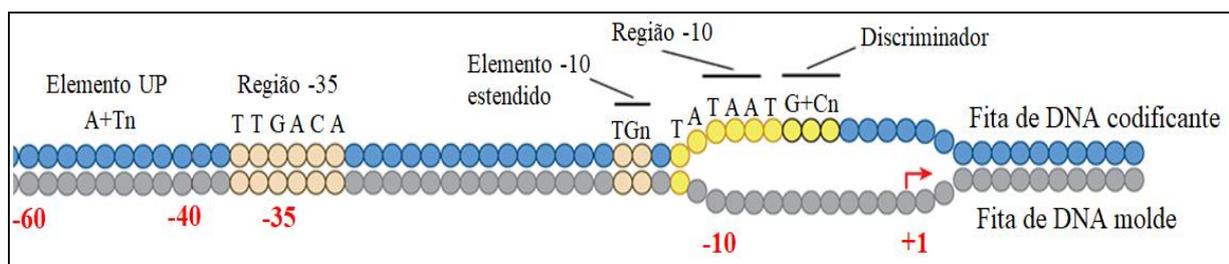


Figura 2. Promotor típico reconhecido pelo fator sigma 70. Está representado um promotor típico de fator sigma 70, indicando os principais elementos reconhecidos pela RNA polimerase (Fonte: Adaptado de FEKLÍSTOV et al. 2014).

O elemento -10 estendido, é um motivo de 3-6 nucleotídeos (5' TGn 3') localizado logo a montante da região -10 (5' TATAAT 3') pode afetar o nível de transcrição através de contatos específicos com domínio $\sigma 3$. Este elemento é importante em promotores com regiões -35 diferentes da sequência consenso 5' TTGACA 3' ou espaçadores mais longos (KEILTY & ROSENBERG, 1987; BARNE et al., 1997; HOOK-BARNARD; HINTON, 2007).

Os elementos UP são sequências situadas normalmente entre as posições -60 e -38 do ponto de início da transcrição (+1), são ricas em nucleotídeos A e T e apresentam como consenso a sequência

5'NNAAAWWTWTTTTNNNAAANN3' (W = A ou T e N = qualquer base). Atribui-se a função *enhancer* transcricional para esta sequência, ou seja, aumenta a transcrição em cerca de até 300 vezes *in vivo* (RAO et al., 1994; ESTREM et al., 1998; GOURSE et al., 2000; ROSS & GOURSE, 2005; RHODIUS et al., 2012).

O discriminador é uma região com cerca de 7-9 pb, e possui 5' GGG 3' como sequência consenso nas posições -4, -5 e -7, portanto está posicionado entre as regiões -10 e o sítio de início da transcrição (+1). Contribui para a duração do complexo de abertura do DNA pela RNA polimerase (HAUGEN et al., 2006; BARINOVA et al., 2008; RUFF et al., 2015).

O nucleotídeo correspondente ao sítio de início da transcrição e a sua localização relativa em relação à região -10 foi estudado em muitos promotores. Está localizado de 7-9 nucleotídeos a jusante da região -10, sendo encontrado nos promotores na seguinte ordem preferencial $A \geq G > T \gg C$ (HOOK-BARNARD & HINTON, 2007).

Esses elementos estão envolvidos diretamente na ligação (afinidade) da RNA polimerase holoenzima ao promotor, bem como no processo de abertura do complexo e início da transcrição. Assim, juntas, estas sequências determinam a força da expressão do promotor, tendo contribuição relativa que difere de promotor para promotor (DAVIS et al., 2017).

4. O uso de promotores de *Escherichia coli* em processos biotecnológicos

A regulação da expressão gênica em *E. coli*, e nos outros seres procariontos, ocorre principalmente em nível de transcrição, tendo como elemento gênico funcional o promotor.

Existem dois grandes tipos de promotores presentes nas células, os ditos constitutivos (ou não reguláveis) e os reguláveis, ou seja, os que são suscetíveis à regulação (ANTHONY et al., 2004; BLAZECK et al., 2012; BLAZECK & ALPER, 2013; DEHLI; et al., 2012; GILMAN & LOVE, 2016; HEISS et al., 2016).

Para a expressão de proteínas recombinantes em altos níveis em *E. coli* é necessário que o promotor do “cassete de expressão” seja regulado, pois a expressão em altos níveis da proteína “estranha” no interior da célula pode ser prejudicial e muitas vezes letal, principalmente se ela apresentar toxicidade à célula hospedeira (SAIDA et al., 2006; SAÏDA, 2007; MARSCHALL et al., 2017).

Um promotor eficiente de *E. coli* adequado para a síntese de proteínas em alto nível para aplicação industrial deve apresentar várias características desejáveis, entre elas: 1) deve ser forte, resultando em um acúmulo de 20 a 40% ou mais da proteína no citoplasma celular; 2) ter um nível mínimo de atividade de expressão basal, sendo rigorosamente regulado; 3) e sua indução deve ser feita de forma simples, efetiva e economicamente viável (JANA & DEB, 2005; BERLEC & ŠTRUKELJ, 2013).

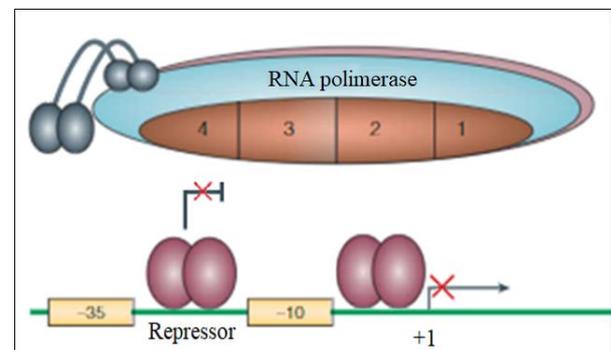


Figura 3: Repressão por impedimento estérico. A figura ilustra a ligação de um ou mais repressores (círculos ovais de cor rosa) aos elementos do núcleo do promotor bloqueia o reconhecimento do mesmo pela RNA polimerase holoenzima e o início da transcrição (Fonte: modificado de BROWNING & BUSBY, 2016).

A maioria dos sistemas que empregam promotores reguláveis nos vetores de expressão se baseia na repressão por impedimento estérico, onde reguladores de transcrição (repressores) ligam-se ao seu operador específico, posicionado sobreposto ou justaposto aos hexâmeros consensuais de reconhecimento (-35 e -10) da sequência promotora. Dessa maneira, com a ligação da

molécula repressora ao operador, haverá o impedimento do reconhecimento do promotor pelo fator sigma e pela RNA polimerase dificultando ou impedindo o início do processo transcricional do gene heterólogo (Figura 3) (ANDRIANANTANDRO et al., 2006; MINCHIN & BUSBY, 2009; BROWNING & BUSBY, 2016; ENGSTROM & PFLEGER, 2017).

Com o desenvolvimento da Biologia Molecular e Engenharia Genética, diversos promotores reguláveis se tornaram conhecidos, ou foram construídos, e utilizados nos vetores de expressão para linhagens de *E. coli*, capazes de expressar proteínas de procariotos e eucariotos. Não existe um sistema de expressão universal e a escolha de cada sistema (hospedeira e vetor de expressão) vai depender de cada proteína a ser expressa. Alguns sistemas que empregam promotores regulados em bioprocessos industriais estão descritos na tabela 2, assim como os seus indutores, nível de regulação e a quantidade de proteínas heterólogas produzidas na

hospedeira. Entre eles, destacam-se o promotor do operon *lac* e seus derivados, o promotor P_R e P_L do fago λ , o promotor P_{araBAD} do operon da arabinose e o promotor do fago P_{T7} (OVERTON, 2014; ROSANO & CECCARELLI, 2014).

4.1 Promotor *lac* e seus derivados

O promotor de transcrição mais estudado e utilizado em engenharia genética é, sem dúvida, o promotor *lac* do operon responsável pelo metabolismo da lactose (*lac*) em *E. coli*. Esse operon regulável é capaz de ser ativado/não ativado e reprimido/induzido. O operon *lac* é constituído por três genes estruturais designados por *lacZ*, *lacY* e *lacA*, que, respectivamente, codificam as proteínas β -galactosidase, permease e transacetilase, sob o controle do promotor *lac*, que atua juntamente com outros elementos regulatórios para captação e uso da lactose como substrato energético, quando há pouca oferta de glicose para a célula (JACOB & MONOD, 1961; COOPER & MAGASANIK, 1974; LEWIS, 2005; DEUTSCHER et al., 2006; ULLMANN, 2009).

Tabela 2: Sistemas de expressão comumente utilizados para a produção de proteínas recombinantes em *Escherichia coli*.

Sistema de expressão	Indutor	Regulação	Produção de proteínas em relação ao total	Referência
P_{Lac} ; P_{LacUV5}	IPTG; TMG; alolactose	Baixa	2-10%	Gronenborn, 1976.
P_{Tac} ; P_{Trc}	IPTG; TMG; alolactose	Moderada	15-32%	Brosius, et al. 1995;
P_{araBAD}	L-arabinose	Alta	10-30%	Guzman et al., 1995;
$\lambda P_L/cI857$ e $\lambda P_R/cI857$	28-42°C	Moderada/Alta	30-50%	Elvin, et al. 1990;
P_{T7} ; $P_{T7/lacUV5}$	IPTG	Moderada/Alta	40-75%	Studier, 1991;

O operon *lac* apresenta duas formas de regulação que atuam de forma complementar. O sistema de repressão/indução que envolve diretamente a região promotora/operadora onde se liga a proteína repressora *LacI* e o sistema de ativação que envolve uma sequência específica de DNA próxima ao promotor onde se liga a proteína CRP (*Cyclic-AMP Receptor Protein*), também denominada de CAP (Proteína Ativadora de Catabolismo) (BUSBY & EBRIGHT, 1999; LAWSON et al., 2004; KUHLMAN et al., 2007;

BICH et al., 2016; ENGSTROM & PFLEGER, 2017).

O sistema de repressão/indução funciona da seguinte forma: *E. coli* tem um gene constitutivo (não regulado) denominado de *lacI*, que sintetiza a proteína repressora *LacI*, constituída por 360 resíduos de aminoácidos, um produto difusível ativo que reconhece e se liga com afinidade ao locus operador *lac* (*lacO*) na ausência de lactose. Tal ligação impede que a RNA polimerase reconheça e se associe ao promotor *lac* (P_{Lac}),

diminuindo o nível de transcrição deste operon, pois o operador *lacO* está sobreposto ao promotor na posição +1. Na presença de lactose, essa é isomerizada no interior da célula, originando a alolactose, que se liga à proteína repressora LacI e altera sua conformação, fazendo com que essa perca a afinidade pelo operador *lacO*, liberando-o. Dessa forma, a RNA polimerase pode reconhecer o promotor (P_{Lac}) e iniciar o processo de transcrição dos genes estruturais do operon *lac*, ou seja, dos genes *lacZ*, *lacY* e *lacA*. Esse processo de

liberação da transcrição se chama indução, na qual a alolactose atua como molécula indutora. Normalmente, em laboratório, para a indução do operon *lac* utilizam-se análogos a alolactose, não metabolizáveis e mais eficientes denominados de isopropil- β -D-1-tiogalactopiranosídeo (IPTG) ou tiometil- β -D-galactosídeo (TMG) (JACOB & MONOD, 1961; CALOS, 1978; WILSON et al., 2007; ULLMANN, 2009; GARCIA et al., 2012; MARBACH & BETTENBROCK, 2012; FAUST et al., 2015).

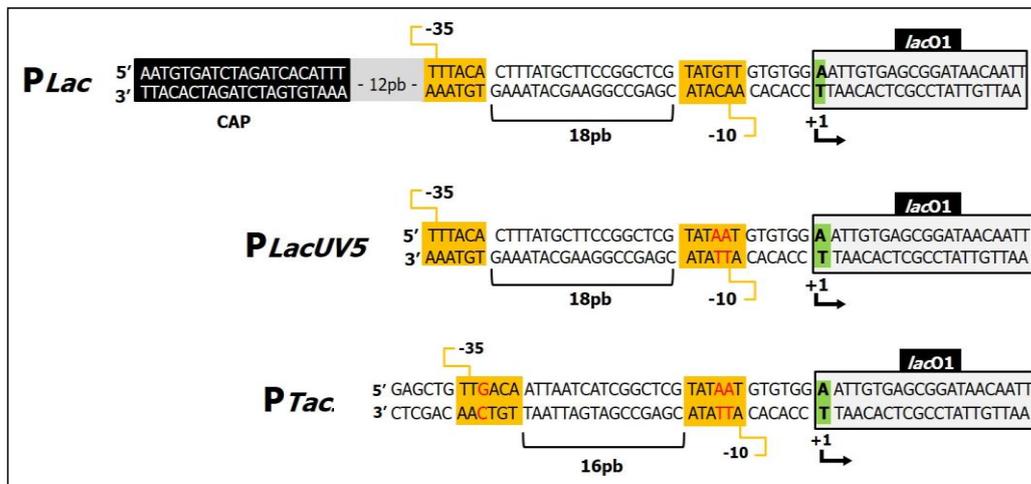


Figura 4: Sequências nucleotídicas dos promotores *lac*, *lacUV5* e *tac*. Os nucleotídeos marcados em vermelho nas regiões -35 e -10 mostram as diferenças de cada promotor (Fonte: adaptado de DE BOER et al., 1983).

O sistema de ativação/não ativação, sensível à glicose, envolve cAMP, a proteína CRP (ou CAP) e o locus de ligação de DNA do complexo CRP-cAMP. Se a lactose e a glicose estiverem presentes, a expressão do promotor *lac* não é totalmente induzida até que toda a glicose tenha sido utilizada pela célula. Quando a célula se encontra com pouca oferta ou ausência de glicose, é produzido cAMP (monofosfato de adenosina cíclica), um sinalizador intracelular. Este controle é conhecido como repressão catabólica. Na ausência de glicose, aumenta o nível de cAMP nas células de *E. coli*, e este liga-se a proteína CRP. O complexo cAMP-CRP liga-se em uma sequência específica de DNA que se situa a montante do promotor/operador, eventos necessários para a ativação completa do operon *lac*. O complexo cAMP-CRP interage com a RNA polimerase resultando na ativação do processo de transcrição dos genes estruturais do operon *lac* (WONG et al., 1997; BUSBY & EBRIGHT, 1999; LAWSON et al., 2004; LEWIS, 2005; BICH et al., 2016).

O sistema de ativação/não ativação, sensível à glicose, envolve cAMP, a proteína CRP (ou CAP) e o locus de ligação de DNA do complexo CRP-cAMP. Se a lactose e a glicose estiverem presentes, a expressão do promotor *lac* não é totalmente induzida até que toda a glicose tenha sido utilizada pela célula. Quando a célula se encontra com pouca oferta ou ausência de glicose, é produzido cAMP (monofosfato de adenosina cíclica), um sinalizador intracelular. Este controle é conhecido como repressão catabólica. Na ausência de glicose, aumenta o nível de cAMP nas células de *E. coli*, e este liga-se a proteína CRP. O complexo cAMP-CRP liga-se em uma sequência específica de DNA que se situa a montante do promotor/operador, eventos necessários para a ativação completa do operon *lac*. O complexo cAMP-CRP interage com a RNA polimerase resultando na ativação do processo de transcrição dos genes estruturais do operon *lac* (WONG et al., 1997; BUSBY & EBRIGHT, 1999; LAWSON et al., 2004; LEWIS, 2005; BICH et al., 2016).

Com o avanço das técnicas de engenharia genética, o promotor *lac* natural foi utilizado conjuntamente com a sequência estrutural do gene *lacZ* (como gene repórter) na composição dos vetores plasmidiais de clonagem, denominados pUC8, pUC9, pUC18 e pUC19. Esses vetores são amplamente utilizados para clonagem molecular de genes, pois possibilitam seleção direta dos clones recombinantes pelo seu múltiplo sítio de clonagem (*polylinker*) e também pelo seu alto número de cópias em *E. coli*. No que se refere à produção de proteínas recombinantes, só expressam proteínas na forma de fusão com β -galactosidase, com nível de expressão e regulação baixos (GRONENBORN, 1976). Esse foi um dos primeiros passos no uso de sequências promotoras no design de vetores moleculares na biotecnologia (RINGS, 1982;

YANISCH-PERRON et al., 1985; AMANN et al., 1988; STRIEDNER et al., 2003).

Foi desenvolvido um promotor mutante derivado do promotor *lac* denominado de P_{lacUV5} (Figura 4), esse promotor difere do promotor *lac* selvagem em 2 nucleotídeos na região conservada -10, tornando-a idêntica à sequência consenso TATAAT (a selvagem é TATGTT) o que o torna 2,5 vezes mais forte que o promotor *lac* nativo. Além disso, o P_{lacUV5} não contém o sítio de ligação da proteína CRP (CAP) e, portanto, não é passível de regulação por cAMP/glicose. Esse promotor integra cassetes de expressão quando se necessita expressar um gene em baixo/médio nível de uma forma bem regulada (SILVERSTONE et al., 1970; FULLER, 1982; YU & REZNIKOFF, 1984; LANZER & BUJARD, 1988; AIYAR et al., 1998; ROSS & GOURSE, 2005).

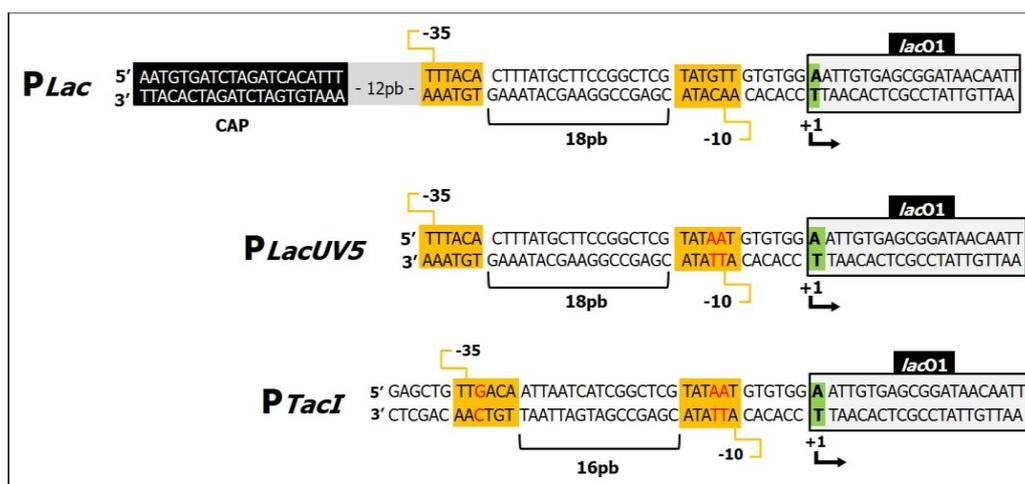


Figura 4: Sequências nucleotídicas dos promotores *lac*, *lacUV5* e *lac*. Os nucleotídeos marcados em vermelho nas regiões -35 e -10 mostram as diferenças de cada promotor (Fonte: adaptado de DE BOER et al., 1983).

Promotores híbridos que combinam a força de um promotor com as vantagens da regulação do promotor *lac* foram desenhados, construídos e utilizados para expressão de proteínas recombinantes, como os promotores sintéticos P_{Tac} e o P_{Trc} . O promotor *tac* híbrido formado por parte dos promotores *lac* e *trp* (do operon triptofano) contém a região consensual -35 (TTGACA) derivada do promotor *trp* e a região consensual -10 (TATAAT) derivada do promotor *lac*. A força do promotor foi aumentada em pelo menos 10 vezes em relação ao P_{lacUV5} (DE BOER et al., 1983). O promotor sintético *trc* é similar ao *tac* porém contém 1 nucleotídeo a mais na região espaçadora entre as regiões conservadas -35 e -10, ficando com 17 nucleotídeos, a distância considerada ideal para

o espaçamento (BROSIUS et al., 1985; AMANN et al., 1988).

Em condições adequadas de nível celular de moléculas repressoras para se ligar aos *loci* operadores, como por exemplo, clonando-se o gene *lacI^q* no próprio plasmídeo multicópia, esses promotores são bem regulados e de força considerada entre média e forte. Esses promotores entram na composição dos vetores da série pTTQs ou pTRCs, que expressam níveis médios/altos de proteínas heterólogas e com bom nível de regulação (DE BOER et al., 1983; STARK, 1987; AMANN et al., 1988).

4.2 Promotor *araBAD*

A série pBAD de plasmídeos permite a expressão das proteínas recombinantes controladas pelo promotor do operon da arabinose. A repressão do promotor *araBAD* é mais eficiente do que os promotores derivados do *lac*, reduzindo qualquer expressão basal indesejada. Os plasmídeos de

expressão baseados no promotor *araBAD* foram concebidos devido ao controle mais rigoroso da expressão, que pode ser graduada pelo indutor L-arabinose para a produção da proteína alvo. Um aumento linear da expressão do gene heterólogo é observado em relação ao aumento da concentração do indutor (GREENFIELD et al., 1978; GUZMAN et al., 1995; RENGBY et al., 2004).

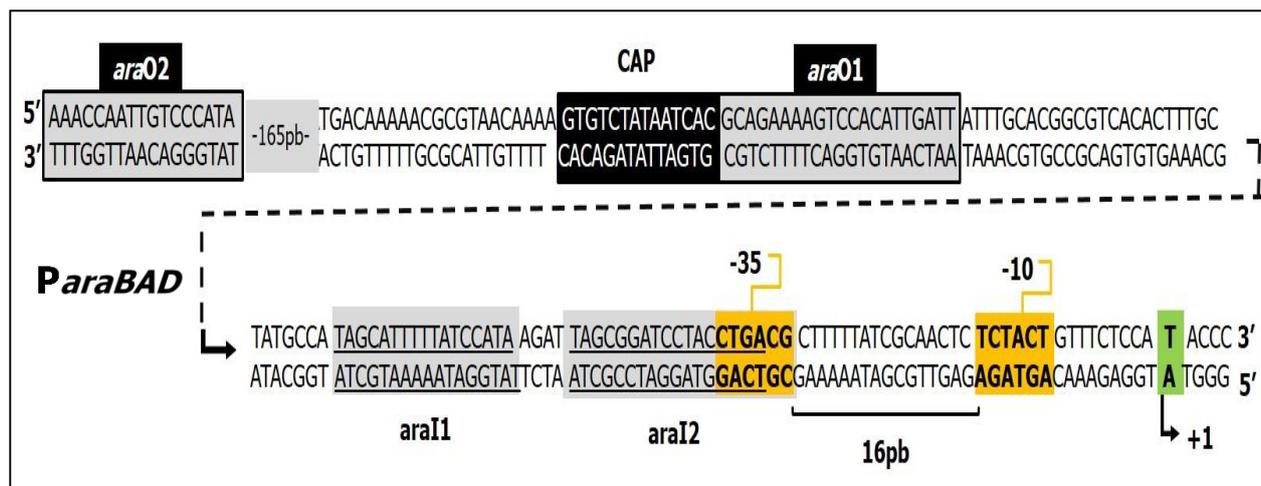


Figura 5: Sequência nucleotídica do promotor *araBAD* e seus elementos regulatórios (Fonte: adaptado de GREENFIELD et al., 1978).

A regulação do operon arabinose em *E. coli* é dirigida pelo produto do gene *araC*, que tem o duplo papel de repressor e ativador. O promotor C (P_C), adjacente ao promotor *araBAD* (Figura 5), transcreve o gene *araC* na direção oposta ao promotor *araBAD*. A proteína AraC forma um dímero e pode se ligar em três sítios no operon da arabinose: O2, I1 e I2. Na ausência de arabinose, AraC dimeriza-se e cada monômero faz contato com os sítios O2 (localizado dentro do gene *araC*) e I1 (a montante do promotor *araBAD*), fazendo uma curvatura (loop) de DNA com cerca de 210 pares de bases. Esta conformação 3D evita o reconhecimento do promotor pela RNA polimerase, bloqueando a transcrição dos genes responsáveis pela captação e uso da arabinose, assim como a ligação da CAP-cAMP (HIRSH & SCHLEIF, 1977; GREENFIELD et al., 1978; GUZMAN et al., 1995; SCHLEIF, 2000; 2010).

Na presença do indutor L-arabinose, a proteína AraC se liga aos sítios I1 e I2. Neste estado, permite o acesso da CAP-cAMP ao seu sítio de ligação, o que por sua vez, ajuda a recrutar

a RNA polimerase para os promotores P_{araBAD} e P_C e assim ativar a transcrição dos genes (SCHLEIF, 2000; 2010).

O promotor *araBAD* do operon arabinose de bactérias, regulado por L-arabinose e glicose, entra na composição dos vetores da série pBADs que expressam genes heterólogos com força média e relativamente bem regulados (GUZMAN et al., 1995; LIM et al., 2000; SAGMEISTER et al., 2014; MARSCHALL et al., 2017).

4.3 Promotores pR e pL do bacteriófago λ

As pesquisas sobre os ciclos de vida dos bacteriófagos mostraram que alguns desses vírus contém promotores muito fortes, que juntamente com outros fatores virais, propiciam eficiente multiplicação na bactéria hospedeira com depleção do metabolismo da célula bacteriana. Esses promotores têm sido muito utilizados para a expressão heteróloga de genes em altos níveis em *E. coli*.

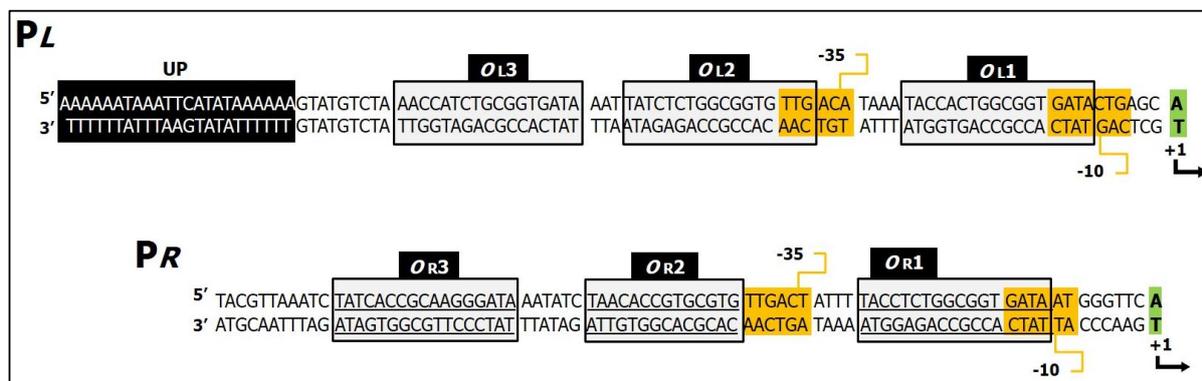


Figura 6: Sequências nucleotídicas dos promotores pL e pR e seus elementos regulatórios (Fonte: adaptado de LEWIS et al., 2016).

O bacteriófago λ tem dois promotores que atuam em sentido opostos denominados P_L (do inglês *promoter left* - promotor a esquerda) e P_R (do inglês *promoter right* - promotor a direita) que são essenciais ao ciclo de vida viral (LEWIS et al., 2016). Por apresentarem alta expressão, são amplamente utilizados em Engenharia Genética para expressão de genes heterólogos em *E. coli*. Os promotores P_L e P_R (figura 6) são regulados pelo produto do gene *cI*, o repressor CI. Durante as pesquisas com o fago λ , obteve-se um gene *cI* mutante (por mutagênese química) que produz um repressor termossensível denominado de *cI857*, que é ativo a 30°C e inativo a 37- 42°C. A expressão de genes heterólogos sob o controle desses promotores é reprimida a temperaturas inferiores a 33°C, ao passo que é induzida quando a temperatura é elevada a 37°C ou acima (normalmente 40°C) em consequência da mudança de conformação e inativação do repressor mutante *cI857* (ELVIN et al., 1990; MENART et al., 2003; VALDEZ-CRUZ et al., 2010).

Uma das vantagens desse sistema de indução térmica para aplicação industrial é o fato de ser vantajoso economicamente induzir a expressão do gene heterólogo através do aumento da temperatura em cerca de 10°C, ao invés de se adicionar ao sistema fermentativo uma substância química indutora. Por outro lado, para os promotores manterem-se reprimidos, as células hospedeiras de *E. coli* necessitam crescer abaixo de sua temperatura ótima, a 30°C, sendo uma limitação para o processo fermentativo industrial. Além disso, o aumento da temperatura para 40°C desencadeia a indução do sistema de estresse celular de choque térmico (HSR – *Heat Shock Response*) que é controlada a nível transcricional pelo fator sigma alternativo σ_{32} (VILLAVERDE

et al., 1993; MENART et al., 2003; VALDEZ-CRUZ et al., 2010).

A HSR inclui uma síntese rápida e seletiva de proteínas de choque térmico (*hsp* - *heat shock proteins*) logo após o aumento da temperatura. As proteínas HSP incluem as chaperonas moleculares e proteases envolvidas no enovelamento, degradação e regulação adequada em resposta ao HSR. Dessa forma, pode-se aumentar o rendimento da proteína heteróloga na conformação correta através das chaperonas moleculares. Entretanto, proteínas heterólogas pequenas poderiam ser reconhecidas como estranhas e degradadas pelas proteases. Além da síntese de HSP, a resposta fisiológica de *E. coli* após um choque térmico também inclui a diminuição temporária da taxa de crescimento e as alterações nas membranas celulares devido à modificação da proporção de lipídios e proteínas de membrana integral (YAMAMORI & YURA, 1980; GROSSMAN et al., 1984; MEJIA et al., 1999; GUIBERT et al., 2004; KELLY et al., 2005; LIMA et al., 2010; VALDEZ-CRUZ et al., 2010).

Os vetores de expressão baseados nesses promotores expressam, respectivamente, médios (P_R) e altos níveis (P_L) de proteínas heterólogas de forma bem regulada. Uma série de vetores baseado no promotor P_L e no repressor *cI857* denominadas de vetores pLEX estão disponíveis comercialmente (SIEGELE & HU, 1997; SØRENSEN et al., 2005). No Brasil, o primeiro fármaco produzido por engenharia genética/fermentação de *E. coli* foi a insulina humana, o vetor de expressão utilizado era pLMT8.5 de indução térmica baseado no promotor P_L do fago λ (ASTOLFI-FILHO et al., 2000).

4.4 Promotor do fago T7

Os bacteriófagos T7, diferentemente dos λ , ao infectarem as células de *E. coli* usam para transcrever seus genes uma RNA polimerase própria que reconhece especificamente seu promotor de apenas 18 pares de nucleotídeos (TAATACGACTCACTATAG), sendo que o último nucleotídeo (G) é o ponto de início da transcrição (figura 7) (OAKLEY & COLEMAN, 1977; TABOR & RICHARDSON, 1985; STUDIER, 1991; TABOR, 2001).

Studier e Moffatt descreveram em 1986 o primeiro sistema de expressão utilizando o sistema de transcrição (promotor e RNA polimerase) do fago T7 (STUDIER & MOFFATT, 1986). A partir desse trabalho foram construídos os primeiros vetores de expressão de uma série grande e amplamente utilizada, denominados de pETs. Desde o início, esse sistema mostrou-se altamente eficiente para expressão de proteínas heterólogas em *E. coli*, podendo expressar mais que 50% do total de proteínas celulares, revolucionando o campo da expressão heteróloga (TABOR & RICHARDSON, 1985; DUBENDORF & STUDIER, 1991; MATTHEY et al., 1999; GRAUMANN & PREMSTALLER, 2006).

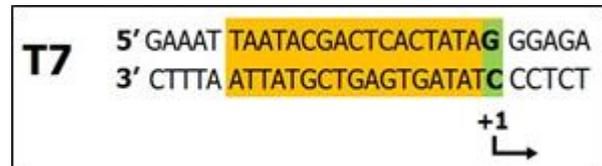


Figura 7: Sequência nucleotídica do promotor do fago T7 (Fonte: adaptado de OAKLEY & COLEMAN, 1977).

Como todo vetor de expressão, um vetor da série pET também tem seu cassete de expressão, só que nesse caso o promotor é o promotor T7 e o terminador de transcrição o próprio do T7, ambos específicos para a RNA polimerase do fago T7 (JENG et al., 1990; 1992; LYAKHOV et al., 1998; MAIRHOFER et al., 2015). Além do *polylinker* para clonagem da mensagem genética a ser expressa, esse tipo de vetor tem também a sequência codificadora do eficiente sítio de ligação ao ribossomo do gene da proteína 10 do próprio fago T7 (MOFFATT & STUDIER, 1987; ROSENBERG et al., 1987; OLINS et al., 1988; DERSCH et al., 1994).

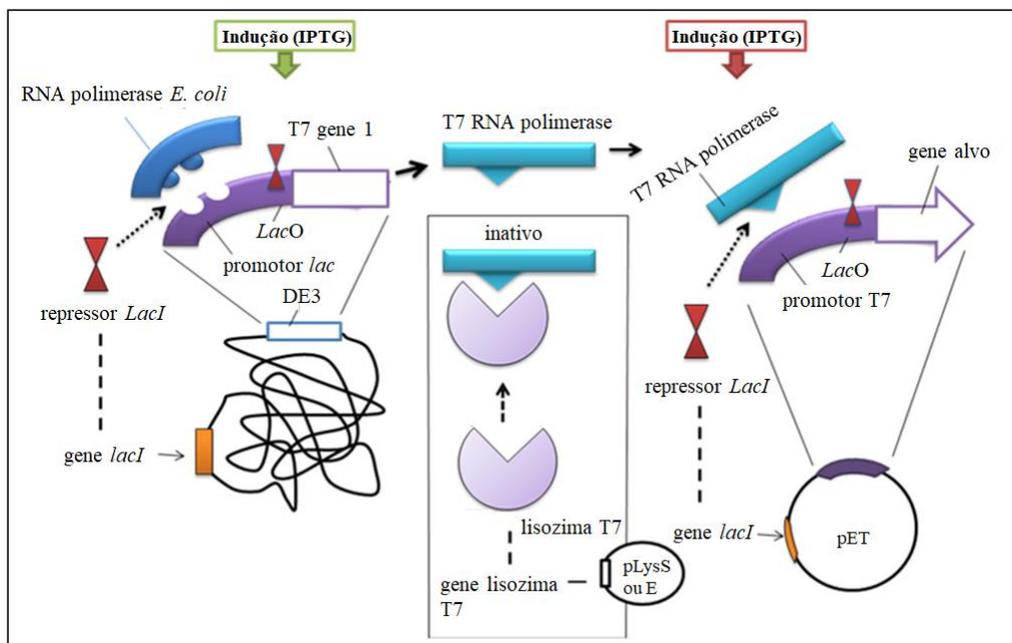


Figura 8: Expressão regulada de genes em *E. coli* por meio de vetores pET. Elementos de controle presentes no sistema da hospedeira BL21 e de vetores tipo pET que compõem o circuito genético para produção de proteínas heterólogas (Fonte: adaptado de DUBENDORF & STUDIER 1991).

Esses vetores podem ser propagados e amplificados em qualquer *E. coli* utilizada em

rotina para clonagem molecular de genes, mas para expressar a mensagem genética heteróloga, tem



que ser introduzido no genoma de *E. coli* que expresse o gene da T7 RNA polimerase (gene 1 do fago). Para tal e para evitar vazamento da expressão da T7 RNA polimerase, a região estrutural do seu gene foi introduzida em fago λ sob o controle do promotor *lacUV5*. Linhagens lisogênicas contendo esse fago recombinante são denominadas de DE3 e são utilizadas para expressar genes heterólogos clonados em vetores pETs com alta eficiência, utilizando IPTG como indutor (figura 8) (DAVANLOO et al., 1984; MERTENS et al., 1995; BORGEAUD & BLOKESCH, 2013).

Em conjunto, as estratégias de evitar vazamento e a elevada força de expressão do promotor T7 tornaram os vetores da série pET os mais utilizados para expressão de genes recombinantes em *E. coli*. Mais de 40 plasmídeos pETs diferentes, que expressam proteínas heterólogas em altos níveis e de maneira bem regulada, estão disponíveis comercialmente por mais de uma empresa (ZHANG & STUDIER, 1997; PIOLI et al., 1999; OVERTON, 2014).

Procurou-se nesse artigo de revisão descrever as sequências dos promotores funcionais em *E. coli*, como interação com a RNA polimerase e apresentar os promotores mais utilizados nos vetores de expressão para produção de proteínas heterólogas nessa bactéria.

5. Conclusão

Desde o início da Tecnologia do DNA Recombinante, tem sido constante a procura de novos promotores para expressão de mensagens genéticas heterólogas nos mais variados tipos de células, seja para expressão constitutiva ou regulada, para diferentes finalidades: superexpressão para fins industriais, terapia genética ou mesmo engenharia metabólica.

Nos primórdios os promotores eram clonados a partir de genes ou operons naturais e os mais fortes eram provenientes dos genomas de bacteriófagos. Hoje novos promotores ou derivados dos naturais são produzidos por síntese química de DNA e mesmo novas sequências têm sido geradas por síntese química randômica.

Para realizar engenharia metabólica com eficiência e velocidade para os mais variados propósitos é de grande relevância a disponibilidade de um conjunto grande e de fácil manipulação de

cassetes de expressão contendo promotores constitutivos e regulados das mais variadas formas.

Em síntese podemos dizer que, embora a pesquisa com promotores tenha evoluído muito nas últimas décadas, a busca por novos promotores funcionais, bem como o melhor entendimento de sistemas de regulação da expressão gênica para os mais variados tipos celulares e aplicações, continua sendo de extrema importância estratégica para as atividades contemporâneas de engenharia metabólica, edição gênica, terapia genética e mesmo a biotecnologia industrial.

Agradecimentos

Ao CNPq, CAPES e FAPEAM pela concessão de bolsas de mestrado (Diego S. Moreira e Kerollen Runa), pela bolsa de IC (Maria Clara T. Astolfi) e pela bolsa DT-C (Spartaco Astolfi-Filho). Aos Profs. Drs. Adolfo José da Mota e Carlos Gustavo Nunes da Silva pela revisão desse artigo.

Divulgação

Este artigo é inédito e não está sendo considerado para qualquer outra publicação. O(s) autor(es) e revisores não relataram qualquer conflito de interesse durante a sua avaliação. Logo, a revista Scientia Amazonia detém os direitos autorais, tem a aprovação e a permissão dos autores para divulgação, deste artigo, por meio eletrônico.

Referências

- A DARST, S. Bacterial RNA polymerase. **Current Opinion in Structural Biology**, v. 11, n. 2, p. 155–162, 2001.
- AIYAR, S. E.; GOURSE, R. L.; ROSS, W. Upstream A-tracts increase bacterial promoter activity through interactions with the RNA polymerase alpha subunit. **Proceedings of the National Academy of Sciences of the United States of America**, v. 95, n. 25, p. 14652–7, 1998.
- ALBRECHTSEN, B. et al. Transcriptional termination sequence at the end of the Escherichia coli ribosomal RNA G operon: Complex terminators and antitermination. **Nucleic Acids Research**, v. 19, n. 8, p. 1845–1852, 1991.
- AMANN, E.; OCHS, B.; ABEL, K. J. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli. **Gene**, v. 69, n. 2, p. 301–315, 1988.



- ANDRIANANTOANDRO, E. et al. Synthetic biology: New engineering rules for an emerging discipline. **Molecular Systems Biology**, v. 2, 2006.
- ANTHONY, L. C.; SUZUKI, H.; FILUTOWICZ, M. Tightly regulated vectors for the cloning and expression of toxic genes. **Journal of Microbiological Methods**, v. 58, n. 2, p. 243–250, 2004.
- ASTOLFI-FILHO, S.; LIMA, B. D.; THIEMANN, J. E.; SOUZA, H. R. T.; VILELA, L. Vector for expression of heterologous protein and methods for extracting recombinant protein and for purifying isolated recombinant insulin. 2000. Patente. Número do registro: US 6.068.993. United States Patent and Trademark Office, USA.
- BAE, B. et al. Structure of a bacterial RNA polymerase holoenzyme open promoter complex. **eLife**, v. 4, n. September 2015, 2015.
- BAKSHI, S.; CHOI, H.; WEISSHAAR, J. C. The spatial biology of transcription and translation in rapidly growing *Escherichia coli*. **Frontiers in Microbiology**, v. 6, 2015.
- BANEYX, F. Recombinant protein expression in *Escherichia coli*. **Curr Opin Biotechnol**, v. 10, n. 5, p. 411–421, 1999.
- BARINOVA, N. et al. Structural modules of RNA polymerase required for transcription from promoters containing downstream basal promoter element GCGA. **Journal of Biological Chemistry**, v. 283, n. 33, p. 22482–22489, 2008.
- BARNE, K. A. et al. Region 2.5 of the *Escherichia coli* RNA polymerase $\sigma 70$ subunit is responsible for the recognition of the “extended -10” motif at promoters. **EMBO Journal**, v. 16, n. 13, p. 4034–4040, 1997.
- 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. B**, v. 6, n. 11, p. 1063–8, 2005.
- BERLEC, A.; ŠTRUKELJ, B. Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells. **Journal of Industrial Microbiology and Biotechnology**, v. 40, n. 3-4, p. 257-274, 2013.
- BICH, L. et al. Biological regulation: controlling the system from within. **Biology and Philosophy**, v. 31, n. 2, p. 237–265, 2016.
- BLAZECK, J. et al. Controlling promoter strength and regulation in *Saccharomyces cerevisiae* using synthetic hybrid promoters. **Biotechnology and Bioengineering**, v. 109, n. 11, p. 2884–2895, 2012.
- BLAZECK, J.; ALPER, H. S. Promoter engineering: Recent advances in controlling transcription at the most fundamental level. **Biotechnology Journal**, v. 8, n. 1, p. 46-58, 2013.
- BORGEAUD, S.; BLOKESCH, M. Overexpression of the tcp Gene Cluster Using the T7 RNA Polymerase/Promoter System and Natural Transformation-Mediated Genetic Engineering of *Vibrio cholerae*. **PLoS ONE**, v. 8, n. 1, 2013.
- BORUKHOV, S.; NUDLER, E. RNA polymerase holoenzyme: Structure, function and biological implications. **Current Opinion in Microbiology**, v. 6, n. 2, p. 93-100, 2003.
- BRAUN, P.; LABAER, J. High throughput protein production for functional proteomics. **Trends in Biotechnology**, v. 21, n. 9, p. 383-388, 2003.
- BROSIUS, J.; ERFLE, M.; STORELLA, J. Spacing of the -10 and -35 regions in the tac promoter. Effect on its in vivo activity. **Journal of Biological Chemistry**, v. 260, n. 6, p. 3539–3541, 1985.
- BROWNING, D. F.; BUSBY, S. J. W. Local and global regulation of transcription initiation in bacteria. **Nature Reviews Microbiology**, v. 14, n. 10, p. 638–650, 2016.
- BROWNING, D. F. D. D. F.; BUSBY, S. J. W. S. The regulation of bacterial transcription initiation. **Nature reviews. Microbiology**, v. 2, n. 1, p. 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. **Bioinformatics (Oxford, England)**, v. 21, n. 5, p. 601–7, 2005.
- BURGESS, R. R. Separation and Characterization of the Subunits of Ribonucleic Acid Polymerase*. **THE JOURNAL OF BIOLOGICAL CHEMISTRY**, v. 244, n. 22, p. 6168–176, 1969.
- BURGESS, R. R.; ANTHONY, L. How sigma docks to RNA polymerase and what sigma does. **Current Opinion in Microbiology**, v. 4, n. 2, p. 126-131, 2001.
- BUSBY, S.; EBRIGHT, R. H. Transcription activation by catabolite activator protein (CAP). **Journal of Molecular Biology**, v. 4, n. 2, p. 126-131, 1999.
- CALOS, M. P. DNA sequence for a low-level



- promoter of the lac repressor gene and an "up" promoter mutation. **Nature**, v. 274, n. 5673, p. 762–765, 1978.
- CAMPBELL, E. A. et al. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. **Cell**, v. 104, n. 6, p. 901–912, 2001.
- CHOU, C. P. Engineering cell physiology to enhance recombinant protein production in *Escherichia coli*. **Applied Microbiology and Biotechnology**, v. 76, n. 3, p. 521–532, 2007.
- CIAMPI, M. S. Rho-dependent terminators and transcription termination. **Microbiology**, v. 152, n. 9, p. 2515–2528, 2006.
- COHEN, S. N. DNA cloning: a personal view after 40 years. **Proceedings of the National Academy of Sciences of the United States of America**, v. 110, n. 39, p. 15521–9, 2013.
- COOPER, T. G.; MAGASANIK, B. Transcription of the lac Operon of *Escherichia coli*. **J. Biol. Chem.**, v. 249, n. 20, p. 6556–6561, 1974.
- CRAMER, P. Multisubunit RNA polymerases. **Current Opinion in Structural Biology**, v. 12, n. 1, p. 89–97, 2002.
- DAVANLOO, P. et al. Cloning and expression of the gene for bacteriophage T7 RNA polymerase. **Proceedings of the National Academy of Sciences**, v. 81, n. 7, p. 2035–2039, 1984.
- DAVIS, M. C. et al. The essential activities of the bacterial sigma factor. **Canadian Journal of Microbiology**, v. 63, n. 2, p. 89–99, 2017.
- DE BOER, H. A.; COMSTOCK, L. J.; VASSER, M. The tac promoter: a functional hybrid derived from the trp and lac promoters. **Proceedings of the National Academy of Sciences**, v. 80, n. 1, p. 21–25, 1983.
- DEHLI, T.; SOLEM, C.; JENSEN, P. R. Tunable promoters in synthetic and systems biology. **Subcellular biochemistry**, v. 64, n. 3, p. 181–201, 2012.
- DERSCH, P.; FSIHI, H.; BREMER, E. Low-copy-number T7 vectors for selective gene expression and efficient protein overproduction in *Escherichia coli*. **FEMS microbiology letters**, v. 123, n. 1–2, p. 19–26, 1994.
- DEUTSCHER, J.; FRANCKE, C.; POSTMA, P. W. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. **Microbiology and Molecular Biology Reviews**, v. 70, n. 4, p. 939–1031, 2006.
- DUBENDORF, J. W.; STUDIER, F. W. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. **Journal of Molecular Biology**, v. 219, n. 1, p. 45–59, 1991.
- ELVIN, C. M. et al. Modified bacteriophage lambda promoter vectors for overproduction of proteins in *Escherichia coli*. **Gene**, v. 87, n. 1, p. 123–126, 1990.
- ENGSTROM, M. D.; PFLEGER, B. F. Transcription control engineering and applications in synthetic biology. **Synthetic and Systems Biotechnology**, 2017.
- ESTREM, S. T. et al. Identification of an UP element consensus sequence for bacterial promoters. **Proceedings of the National Academy of Sciences**, v. 95, n. 17, p. 9761–9766, 1998.
- FAUST, G.; STAND, A.; WEUSTER-BOTZ, D. IPTG can replace lactose in auto-induction media to enhance protein expression in batch-cultured *Escherichia coli*. **Engineering in Life Sciences**, v. 15, n. 8, p. 824–829, 2015.
- FEKLÍSTOV, A. et al. Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective. **Annual Review of Microbiology**, v. 68, n. 1, p. 357–376, 2014.
- FULLER, F. A family of cloning vectors containing the lacUV5 promoter. **Gene**, v. 19, n. 1, p. 43–54, 1982.
- GARCIA, H. G. et al. Operator sequence alters gene expression independently of transcription factor occupancy in bacteria. **Cell Reports**, v. 2, n. 1, p. 150–161, 2012.
- GEZVAIN, K. .; LANDICK, R. The structure of bacterial RNA polymerase. **The bacterial chromosome**, n. 40, p. 283–296, 2004.
- GHOSH, T.; BOSE, D.; ZHANG, X. Mechanisms for activating bacterial RNA polymerase. **FEMS Microbiology Reviews**, v. 34, n. 5, p. 611–627, 2010.
- GILMAN, J.; LOVE, J. Synthetic promoter design for new microbial chassis. **Biochemical Society Transactions**, v. 44, n. 3, p. 731–737, 2016.
- GOURSE, R. L.; ROSS, W.; GAAL, T. UPs and downs in bacterial transcription initiation: The role of the alpha subunit of RNA polymerase in promoter recognition. **Molecular Microbiology**, v. 37, n. 4, p. 687–695, 2000.



- GRAUMANN, K.; PREMSTALLER, A. Manufacturing of recombinant therapeutic proteins in microbial systems. **Biotechnology Journal**, v. 1, n. 2, p. 164 -186, 2006.
- GREENFIELD, L.; BOONE, T.; WILCOX, G. DNA sequence of the araBAD promoter in Escherichia coli B/r. **Proceedings of the National Academy of Sciences of the United States of America**, v. 75, n. 10, p. 4724–8, 1978.
- GRONENBORN, B. Overproduction of phage Lambda repressor under control of the lac promoter of Escherichia coli. **MGG Molecular & General Genetics**, v. 148, n. 3, p. 243–250, 1976.
- GROSSMAN, A. D.; ERICKSON, J. W.; GROSS, C. A. The htpR gene product of E. coli is a sigma factor for heat-shock promoters. **Cell**, v. 38, n. 2, p. 383-390, 1984.
- GRUBER, T. M.; GROSS, C. A. Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. **Annual Review of Microbiology**, v. 57, n. 1, p. 441–466, 2003.
- GUISBERT, E. et al. A chaperone network controls the heat shock response in E. coli. **Genes and Development**, v. 18, n. 22, p. 2812–2821, 2004.
- GUO, Y.; LEW, C. M.; GRALLA, J. D. Promoter opening by σ_{54} and σ_{70} RNA polymerases: σ Factor-directed alterations in the mechanism and tightness of control. **Genes and Development**, v. 14, n. 17, p. 2242–2255, 2000.
- GUSAROV, I.; NUDLER, E. Control of intrinsic transcription termination by N and NusA: The basic mechanisms. **Cell**, v. 107, n. 4, p. 437-449, 2001.
- GUZMAN, L. M. et al. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P(BAD) promoter. **Journal of Bacteriology**, v. 177, n. 14, p. 4121–4130, 1995.
- HALL, R. M.; COLLIS, C. M. Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. **Molecular Microbiology**, v. 15, n. 4, p. 593-600, 1995.
- HARLEY, C. B.; REYNOLDS, R. P. Analysis of E.Coli promoter sequences. **Nucleic Acids Research**, v. 15, n. 5, p. 2343–2361, 1987.
- HARTLEY, J. L. Cloning technologies for protein expression and purification. **Current Opinion in Biotechnology**, v. 17, n. 4, p. 359-366, 2006.
- HAUGEN, S. P. et al. rRNA Promoter Regulation by Nonoptimal Binding of σ Region 1.2: An Additional Recognition Element for RNA Polymerase. **Cell**, v. 125, n. 6, p. 1069–1082, 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**, v. 6, n. 7, p. 507-519, 2008.
- HAWLEY, D. K.; MCCLURE, W. R. Compilation and analysis of Escherichia coli promoter DNA sequences. **Nucleic acids research**, v. 11, n. 8, p. 2237–55, 1983.
- HEISS, S. et al. Evaluation of novel inducible promoter/repressor systems for recombinant protein expression in Lactobacillus plantarum. **Microbial Cell Factories**, v. 15, n. 1, 2016.
- HELMANN, J. D. Sigma Factors in Gene Expression. **eLS**, 2001.
- HIRSH, J.; SCHLEIF, R. The araC promoter: Transcription, mapping and interaction with the araBAD promoter. **Cell**, v. 11, n. 3, p. 545–550, 1977.
- HOOK-BARNARD, I. G.; HINTON, D. M. Transcription Initiation by Mix and Match Elements: Flexibility for Polymerase Binding to Bacterial Promoters. **Gene Regulation and Systems Biology**, v. 1, p. 117762500700100, 2007.
- ISHIHAMA, A. Functional modulation of Escherichia coli RNA polymerase. **Annual review of microbiology**, v. 54, p. 499–518, 2000.
- ITAKURA, K. et al. Expression in Escherichia coli of a Chemically Synthesized Gene for the Hormone Somatostatin. **Science**, v. 198, n. 4321, p. 1056–1063, 1977.
- JACOB, F.; MONOD, J. Genetic regulatory mechanisms in the synthesis of proteins. **Journal of Molecular Biology**, v. 3, n. 3, p. 318 –356, 1961.
- JALESNIAK, P.; SENG WONG, T. From genetic circuits to industrial-scale biomanufacturing: bacterial promoters as a cornerstone of biotechnology. **AIMS Bioengineering**, v. 2, n. 3, p. 277–296, 2015.
- JANA, S.; DEB, J. K. Strategies for efficient production of heterologous proteins in Escherichia coli. **Applied Microbiology and Biotechnology**, v. 67, n. 3, p. 289-298, 2005.
- JENG, S.; GARDNER, J.; GUMPORT, R. Transcription Termination in Vitro by Bacteriophage T7 RNA Polymerase. **Journal of Biological Chemistry**, v. 267, n. 27, p. 19306–



- 19312, 1992.
- JENG, S. T.; GARDNER, J. F.; GUMPORT, R. I. Transcription termination by bacteriophage T7 RNA polymerase at rho-independent terminators. **The Journal of biological chemistry**, v. 265, n. 7, p. 3823–3830, 1990.
- JULLESSON, D. et al. Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals. **Biotechnology Advances**, v. 33, n. 7, p. 1395-1402, 2015.
- KAZMIERCZAK, M. J.; WIEDMANN, M.; BOOR, K. J. Alternative sigma factors and their roles in bacterial virulence. **Microbiology and Molecular Biology Reviews**, v. 69, n. 4, p. 527–543, 2005.
- KEILTY, S.; ROSENBERG, M. Constitutive function of a positively regulated promoter reveals new sequences essential for activity. **Journal of Biological Chemistry**, v. 262, n. 13, p. 6389–6395, 1987.
- KELLY, A.; FILIPE, S.; GNADT, N.; BARBAS, A.; PISSARRA, P. Expression vectors and promoters for heterologous gene expression. 2005. Número de registro: US20050119462 A1. Data de depósito: 23/12/2002 Data de Concessão: 02/06/2005.
- KUHLMAN, T. et al. Combinatorial transcriptional control of the lactose operon of *Escherichia coli*. **Proceedings of the National Academy of Sciences**, v. 104, n. 14, p. 6043–6048, 2007.
- LANE, W. J.; DARST, S. A. Molecular Evolution of Multisubunit RNA Polymerases: Sequence Analysis. **Journal of Molecular Biology**, v. 395, n. 4, p. 671–685, 2010.
- LANZER, M.; BUJARD, H. Promoters largely determine the efficiency of repressor action. **Proceedings of the National Academy of Sciences of the United States of America**, v. 85, n. 23, p. 8973–7, 1988.
- LAWSON, C. L. et al. Catabolite activator protein: DNA binding and transcription activation. **Current Opinion in Structural Biology**, v. 14, n. 1, p. 10–20, 2004.
- LEE, D. J.; MINCHIN, S. D.; BUSBY, S. J. W. Activating Transcription in Bacteria. **Annual Review of Microbiology**, v. 66, n. 1, p. 125–152, 2012.
- LEWIS, D. E. A.; GUSSIN, G. N.; ADHYA, S. New Insights into the Phage Genetic Switch: Effects of Bacteriophage Lambda Operator Mutations on DNA Looping and Regulation of PR, PL, and PRM. **Journal of Molecular Biology**, v. 428, n. 22, p. 4438–4456, 2016.
- LEWIS, M. The lac repressor. **Comptes Rendus - Biologies**, v. 328, n. 6, p. 521–548, 2005.
- LIMA, B. D., SOUZA, H. R. T., THIEMANN, J. F., VILELA, L.; ASTOLFI-FILHO, S. Vetor para expressão de proteína heteróloga e métodos para extrair proteína recombinante e purificar insulina recombinante isolada. 2010. Número de registro: PI9810650-3. Data de depósito: 02/07/1998 Data de Concessão: 14/02/2010.
- LIM, H. K. et al. Production characteristics of interferon-alpha using an L-arabinose promoter system in a high-cell-density culture. **Applied Microbiology and Biotechnology**, v. 53, n. 2, p. 201–208, 2000.
- LONETTO, M.; GRIBSKOV, M.; GROSS, C. A. The sigma 70 family: sequence conservation and evolutionary relationships. **Journal of Bacteriology**, v. 174, n. 12, p. 3843–3849, 1992.
- LYAKHOV, D. L. et al. Pausing and termination by bacteriophage T7 RNA polymerase. **Journal of molecular biology**, v. 280, n. 2, p. 201–213, 1998.
- MAIRHOFER, J. et al. Preventing T7 RNA polymerase read-through transcription-A synthetic termination signal capable of improving bioprocess stability. **ACS Synthetic Biology**, v. 4, n. 3, p. 265–273, 2015.
- MAKRIDES, S. C. Strategies for achieving high-level expression of genes in *Escherichia coli*. **Microbiological reviews**, v. 60, n. 3, p. 512–538, 1996.
- MARBACH, A.; BETTENBROCK, K. Lac operon induction in *Escherichia coli*: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. **Journal of Biotechnology**, v. 157, n. 1, p. 82–88, 2012.
- MARSCHALL, L.; SAGMEISTER, P.; HERWIG, C. Tunable recombinant protein expression in *E. coli*: enabler for continuous processing. **Applied Microbiology and Biotechnology**, v. 100, n. 13, p. 5719–5728, 2016.
- MARSCHALL, L.; SAGMEISTER, P.; HERWIG, C. Tunable recombinant protein expression in *E. coli*: promoter systems and genetic constraints. **Applied Microbiology and Biotechnology**, v. 101, n. 2, p. 501–512, 2017.



- MATTHEY, B. et al. A new series of pET-derived vectors for high efficiency expression of Pseudomonas exotoxin-based fusion proteins. **Gene**, v. 229, n. 1–2, p. 145–153, 1999.
- MEJIA, R.; GOMEZ-EICHELMASS, M. C.; FERNANDEZ, M. S. Fatty acid profile of Escherichia coli During the Heat-shock response. **Biochemistry and Molecular Biology International**, v. 47, n. 5, p. 835–844, 1999.
- MENART, V. et al. Constitutive versus thermoinducible expression of heterologous proteins in Escherichia coli based on strong PR,PL promoters from phage lambda. **Biotechnology and Bioengineering**, v. 83, n. 2, p. 181–190, 2003.
- MERRICK, M. J. In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N). **Molecular microbiology**, v. 10, n. 5, p. 903–909, 1993.
- MERTENS, N.; REMAUT, E.; FIERS, W. Tight transcriptional control mechanism ensures stable high-level expression from T7 promoter-based expression plasmids. **Bio/Technology**, v. 13, n. 2, p. 175–179, 1995.
- MINAKHIN, L. et al. Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 3, p. 892–897, 2001.
- MINCHIN, S. D.; BUSBY, S. J. W. Analysis of mechanisms of activation and repression at bacterial promoters. **Methods**, v. 47, n. 1, p. 6–12, 2009.
- MOFFATT, B. A.; STUDIER, F. W. T7 lysozyme inhibits transcription by T7 RNA polymerase. **Cell**, v. 49, n. 2, p. 221–227, 1987.
- MURAKAMI, K. Structural Biology of Bacterial RNA Polymerase. **Biomolecules**, v. 5, n. 2, p. 848–864, 2015.
- MURAKAMI, K. S. et al. Structural basis of transcription initiation: An RNA polymerase holoenzyme-DNA complex. **Science**, v. 296, n. 5571, p. 1285–1290, 2002.
- MURAKAMI, K. S.; DARST, S. A. Bacterial RNA polymerases: The whole story. **Current Opinion in Structural Biology**, v. 13, n. 1, p. 31–39, 2003.
- MURAKAMI, K. S.; MASUDA, S.; DARST, S. A. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4?? resolution. **Science**, v. 296, n. 5571, p. 1280–1284, 2002.
- NUDLER, E. RNA Polymerase Active Center: The Molecular Engine of Transcription. **Annual Review of Biochemistry**, v. 78, n. 1, p. 335–361, 2009.
- NUDLER, E.; GOTTESMAN, M. E. Transcription termination and anti-termination in E. coli. **Genes to Cells**, v. 7, n. 8, p. 755–768, 2002.
- OAKLEY, J. L.; COLEMAN, J. E. Structure of a promoter for T7 RNA polymerase. **Proceedings of the National Academy of Sciences of the United States of America**, v. 74, n. 10, p. 4266–4270, 1977.
- OLINS, P. O. et al. The T7 phage gene 10 leader RNA, a ribosome-binding site that dramatically enhances the expression of foreign genes in Escherichia coli. **Gene**, v. 73, n. 1, p. 227–235, 1988.
- ÖSTERBERG, S.; PESO-SANTOS, T. DEL; SHINGLER, V. Regulation of Alternative Sigma Factor Use. **Annual Review of Microbiology**, v. 65, n. 1, p. 37–55, 2011.
- OVERTON, T. W. Recombinant protein production in bacterial hosts. **Drug Discovery Today**, v. 19, n. 5, p. 590–601, 2014.
- PAGET, M. S. B.; HELMANN, J. D. The $\sigma 70$ family of sigma factors. **Genome Biology**, v. 4, n. 1, p. 203, 2003.
- PAPANEOPHYTOU, C. P.; KONTOPIDIS, G. Statistical approaches to maximize recombinant protein expression in Escherichia coli: A general review. **Protein Expression and Purification**, v. 94, p. 22–32, 2014.
- PIOLI, D.; HOCKNEY, R. C.; KARA, B. V.; BUNDELL, K. R. T7 promoter-based expression system. 1999. Número de registro: WO1999005297 A1. Data de depósito: 21/07/1998 Data de Concessão: 4/02/1999.
- PRIBNOW, D. Bacteriophage T7 early promoters: Nucleotide sequences of two RNA polymerase binding sites. **Journal of Molecular Biology**, v. 99, n. 3, 1975.
- RANGEL-CHAVEZ, C.; GALAN-VASQUEZ, E.; MARTINEZ-ANTONIO, A. Consensus architecture of promoters and transcription units in Escherichia coli: design principles for synthetic biology. **Mol.**



BioSyst., v. 13, n. 4, p. 665–676, 2017.

RAO, L. et al. Factor independent activation of *rrnB* p1. An "Extended" promoter with an upstream element that dramatically increases promoter strength. **Journal of Molecular Biology**, v. 235, n. 5, p. 1421–1435, 1994.

RENGBY, O. et al. Assessment of production conditions for efficient use of *Escherichia coli* in high-yield heterologous recombinant selenoprotein synthesis. **Applied and Environmental Microbiology**, v. 70, n. 9, p. 5159–5167, 2004.

RHODIUS, V. A.; MUTALIK, V. K.; GROSS, C. A. Predicting the strength of UP-elements and full-length *E. coli* σ e promoters. **Nucleic Acids Research**, v. 40, n. 7, p. 2907–2924, 2012.

RICHARDSON, J. P. Rho-dependent termination and ATPases in transcript termination. **Biochimica et Biophysica Acta - Gene Structure and Expression**, v. 1577, n. 2, p. 251–260, 2002.

RINGS A. D. Method for microbial polypeptide expression. 1982. Número de registro: US4366246 A. Data de depósito: 05/11/1979 Data de Concessão: 28/12/1982.

ROSANO, G.; L., N L. G.; CECCARELLI, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. **Frontiers in Microbiology**, v. 5, n. APR, p. 1–17, 2014.

ROSENBERG, A. H. et al. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. **Gene**, v. 56, n. 1, p. 125–135, 1987.

ROSS, W.; GOURSE, R. L. Sequence-independent upstream DNA- CTD interactions strongly stimulate *Escherichia coli* RNA polymerase-*lacUV5* promoter association. **Proceedings of the National Academy of Sciences**, v. 102, n. 2, p. 291–296, 2005.

ROSS, W.; GOURSE, R. L. Analysis of RNA polymerase-promoter complex formation. **Methods**, v. 47, n. 1, p. 13–24, 2009.

RUFF, E. F.; THOMAS RECORD, M.; ARTSIMOVITCH, I. **Initial events in bacterial transcription initiation** *Biomolecules*, 2015.

SAECKER, R. M.; RECORD, M. T.; DEHASETH, P. L. Mechanism of bacterial transcription initiation: RNA polymerase - Promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. **Journal of Molecular Biology**, v. 412, n. 5, p. 754–771, 2011.

SAGMEISTER, P. et al. Tunable recombinant

protein expression with *E. coli* in a mixed-feed environment. **Applied Microbiology and Biotechnology**, v. 98, n. 7, p. 2937–2945, 2014.

SAIDA, F. et al. Expression of Highly Toxic Genes in *E. coli*: Special Strategies and Genetic Tools. **Current Protein and Peptide Science**, v. 7, n. 1, p. 47–56, 2006.

SAÏDA, F. Overview on the Expression of Toxic Gene Products in *Escherichia coli*. In: **Current Protocols in Protein Science**. [s.l.: s.n.]. p. 5.19.1-5.19.13.

SCHLEIF, R. Regulation of the L -arabinose operon of *Escherichia coli*. **Trends in Genetics**, v. 16, n. 12, p. 559–565, 2000.

SCHLEIF, R. AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. **FEMS Microbiology Reviews**, v. 34, n. 5, p. 779–796, 2010.

SCHUMANN, W.; FERREIRA, L. C. S. Production of recombinant proteins in *Escherichia coli*. **Genetics and Molecular Biology**, v. 27, n. 3, p. 442–453, 2004.

SHIMADA, T. et al. The whole set of constitutive promoters recognized by RNA polymerase RpoD holoenzyme of *Escherichia coli*. **PLoS ONE**, v. 9, n. 3, 2014.

SIEGELE, D. A.; HU, J. C. Gene expression from plasmids containing the *araBAD* promoter at subsaturating inducer concentrations represents mixed populations. **Proceedings of the National Academy of Sciences**, v. 94, n. 15, p. 8168–8172, 1997.

SILVERSTONE, A. E.; ARDITTI, R. R.; MAGASANIK, B. Catabolite-Insensitive Revertants of Lac Promoter Mutants. **Proceedings of the National Academy of Sciences of the United States of America**, v. 66, n. 3, p. 773–779, 1970.

SØRENSEN, H. P. et al. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. **Journal of Biotechnology**, v. 115, n. 2, p. 113–128, 2005.

STARK, M. J. Multicopy expression vectors carrying the *lac* repressor gene for regulated high-level expression of genes in *Escherichia coli*. **Gene**, v. 51, n. 2–3, p. 255–67, 1987.

STRIEDNER, G. et al. Tuning the Transcription Rate of Recombinant Protein in Strong *Escherichia coli* Expression Systems through Repressor Titration.



Biotechnology Progress, v. 19, n. 5, p. 1427–1432, 2003.

STUDIER, F. W. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. **Journal of Molecular Biology**, v. 219, n. 1, p. 37–44, 1991.

STUDIER, F. W.; MOFFATT, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. **Journal of Molecular Biology**, v. 189, n. 1, p. 113–30, 1986.

SWARTZ, J. R. Advances in Escherichia coli production of therapeutic proteins. **Current Opinion in Biotechnology**, v. 12, n. 2, p. 195–201, 2001.

TABOR, S. Expression Using the T7 RNA Polymerase/Promoter System. In: **Current Protocols in Molecular Biology**, 2001.

TABOR, S.; RICHARDSON, C. C. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes (T7 DNA polymerase/T7 gene 5 protein/proteolysis/13-lactamase/rifampicin). **Biochemistry**, v. 82, n. February, p. 1074–1078, 1985.

TRIPATHI, L.; ZHANG, Y.; LIN, Z. Bacterial Sigma Factors as Targets for Engineered or Synthetic Transcriptional Control. **Frontiers in Bioengineering and Biotechnology**, v. 2, 2014.

ULLMANN, A. Escherichia coli Lactose Operon. **Encyclopedia of Life Sciences**, 2009.

VALDEZ-CRUZ, N. A. et al. Production of recombinant proteins in E. coli by the heat inducible expression system based on the phage lambda pL and/or pR promoters. **Microbial Cell Factories**, v. 9, n. 18, 2010.

VAN HIJUM, S. A. F. T.; MEDEMA, M. H.; KUIPERS, O. P. Mechanisms and evolution of control logic in prokaryotic transcriptional regulation. **Microbiology and molecular biology reviews: MMBR**, v. 73, n. 3, p. 481–509, Table of Contents, 2009.

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

VILLAVERDE, A. et al. Fine regulation of cI857-controlled gene expression in continuous culture of

recombinant Escherichia coli by temperature. **Applied and Environmental Microbiology**, v. 59, n. 10, p. 3485–3487, 1993.

WASHBURN, R. S.; GOTTESMAN, M. E. Regulation of transcription elongation and termination. **Biomolecules**, v. 5, n. 2, p. 1063–1078, 2015.

WIGNESHWERARAJ, S. et al. Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor. **Molecular microbiology**, v. 68, n. 3, p. 538–46, 2008.

WILSON, C. J. et al. The lactose repressor system: Paradigms for regulation, allosteric behavior and protein folding. **Cellular and Molecular Life Sciences**, v. 64, n. 1, p. 3–16, 2007.

WONG, P.; GLADNEY, S.; KEASLING, J. D. Mathematical model of the lac operon: Inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose. **Biotechnology Progress**, v. 13, n. 2, p. 132–143, 1997.

XIAO, S.; SHILOACH, J.; BETENBAUGH, M. J. Engineering cells to improve protein expression. **Current Opinion in Structural Biology**, v. 26, n. 1, p. 32–38, 2014.

YAMAMORI, T.; YURA, T. Temperature-induced synthesis of specific proteins in Escherichia coli: Evidence for transcriptional control. **Journal of Bacteriology**, v. 142, n. 3, p. 843–851, 1980.

YAN, Q.; FONG, S. S. Study of in vitro transcriptional binding effects and noise using constitutive promoters combined with UP element sequences in Escherichia coli. **Journal of Biological Engineering**, v. 11, n. 1, 2017.

YANISCH-PERRON, C.; VIEIRA, J.; MESSING, J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. **Gene**, v. 33, n. 1, p. 103–119, 1985.

YU, X.-M.; REZNIKOFF, W. S. Deletion analysis of the CAP-cAMP binding site of the Escherichia coli lactose promoter. **Nucleic Acids Research**, v. 12, n. 13, p. 5449–5464, 1984.

ZHANG, G. et al. Crystal structure of thermus aquaticus core RNA polymerase at 3.3 Å resolution. **Cell**, v. 98, n. 6, p. 811–824, 1999.

ZHANG, X.; STUDIER, F. W. Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme. **Journal of Molecular Biology**, v. 269, n. 1, p. 10–27, 1997.