



Análise comparativa da literatura sobre expressão de proteínas recombinantes em *Escherichia coli* e *Pichia pastoris*

Comparative literature review of *Escherichia coli* and *Pichia pastoris* recombinant protein expression

Danielle Vitoria Santana Dias*, Cairo Alencar Zancanella‡,

Luís Felipe Minozzo Figueiredo†

RESUMO

Avanços na engenharia genética resultaram no desenvolvimento de sistemas de expressão bacterianos, como os que envolvem a bactéria *Escherichia coli*, que são capazes de produzir proteínas a partir dos genes clonados. Outro microrganismo amplamente utilizado para expressão de proteínas recombinantes é a levedura *Pichia pastoris*, cuja maior vantagem em relação aos sistemas bacterianos é a capacidade de realizar modificações pós traducionais. A produção de proteínas recombinantes é geralmente realizada utilizando uma abordagem de tentativa e erro. Porém, as interações entre as variáveis são perdidas, o que torna essa abordagem demorada. O objetivo deste trabalho foi levantar dados referentes à expressão de proteínas recombinantes em sistema de bactérias e leveduras comparando as condições de maior produtividade. As condições ótimas de expressão para *E. coli* foram definidas como: indução com IPTG 1 mM, temperatura e tempo de expressão de 30 °C e 12 h. Já os parâmetros de expressão para *P. pastoris* foram: crescimento a 38 °C, 250 rpm por 120 h; meios de cultivo YPD e BMGY com suplementos de antibióticos; indução por 120 h cuja concentração final de metanol seja 0,5 %.

Palavras-chave: Proteínas recombinantes, *Escherichia coli*, *Pichia pastoris*

ABSTRACT

Advances in genetic engineering have resulted in the development of bacterial expression systems, such as those involving the bacterium *Escherichia coli*, which are capable of producing proteins from cloned genes. Another microorganism widely used for expression of recombinant proteins is the yeast *Pichia pastoris*, whose greatest advantage over bacterial systems is the ability to perform post translational modifications. The production of recombinant proteins is usually performed using a trial-and-error approach. However, interactions between variables are lost, which makes this approach time-consuming. The objective of this study was to collect data regarding the expression of recombinant proteins in bacteria and yeast systems comparing the conditions of higher productivity. The optimal expression conditions for *E. coli* were defined as: induction with IPTG 1 mM, temperature and expression time of 30 °C and 12 h. The expression parameters for *P. pastoris* were: growth at 38 °C, 250 rpm per 120 h; YPD and BMGY culture media with antibiotic supplements; 120 h induction with a final methanol concentration of 0,5 %.

Keywords: Recombinant proteins, *Escherichia coli*, *Pichia pastoris*

1 INTRODUÇÃO

O fornecimento de proteínas que têm potencial uso clínico ou industrial é, muitas vezes, limitado por sua baixa disponibilidade natural. Porém, com os avanços modernos em genômica, proteômica e

* Engenharia de Bioprocessos e Biotecnologia, Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil; daniellevsd@gmail.com

‡ Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil; cairo.zancanella@gmail.com

† Universidade Tecnológica Federal do Paraná, Campus Toledo; luisfigueiredo@utfpr.edu.br



bioinformática, o número de proteínas produzidas por meio de técnicas recombinantes está aumentando exponencialmente, sendo uma área de pesquisa e desenvolvimento em constante crescimento (PAPANEOPHYTOU *et al.*, 2014).

Avanços na engenharia genética resultaram no desenvolvimento de sistemas de expressão bacterianos, como os que envolvem a bactéria *Escherichia coli* (*E. coli*), que são capazes de produzir proteínas a partir dos genes clonados. No entanto há dois desafios na produção de proteínas heterólogas nesse sistema, sendo eles a baixa expressão e o dobramento incorreto da proteína, expressa em agregados insolúveis chamados corpos de inclusão (KANE *et al.*, 1988). Outro microrganismo amplamente utilizado para expressão de proteínas recombinantes é a levedura *Pichia pastoris* (*P. pastoris*), cuja maior vantagem em relação aos sistemas bacterianos é a capacidade de realizar modificações pós traducionais, como N e O-glicosilações (YANG *et al.*, 2018).

A produção de proteínas recombinantes é geralmente realizada utilizando uma abordagem de tentativa e erro, com as diferentes variáveis de expressão sendo testadas independentemente umas das outras. Porém, dessa forma, as interações entre as variáveis são perdidas, o que torna essa abordagem demorada. Esse método tradicional foi sendo progressivamente substituído por abordagens fatoriais em cada etapa do processo, desde a expressão gênica até a purificação (PAPANEOPHYTOU *et al.*, 2014).

Devido à crise sanitária da COVID-19 não puderam ser realizadas atividades laboratoriais. Desta forma, o objetivo geral deste trabalho foi levantar dados referentes à expressão de proteínas recombinantes em sistema de bactérias e leveduras, comparando as condições de maior produtividade, e selecionar uma a ser desenvolvida no decorrer do projeto.

2 PROCEDIMENTOS OPERACIONAIS DA PESQUISA

Para o presente trabalho foi realizado um levantamento bibliográfico sobre a expressão de proteínas recombinantes em *E. coli* e *P. pastoris*, com enfoque comparativo nas condições de expressão. Isso foi feito utilizando os termos de busca “*Pichia pastoris recombinant protein expression*”, “*Escherichia coli recombinant protein expression*” além das palavras “*optimization*” e “*review*”, acopladas ou não aos termos supracitados. As plataformas utilizadas foram ScienceDirect, PubMed e SciFinder. Durante a seleção dos trabalhos deu-se preferência aos que descreveram as condições otimizadas de expressão detalhadamente, aos mais recentes, publicados a partir de 2014, e aos que apresentaram similaridades com o projeto. Por fim, levando em conta o rendimento, massa da proteína e outros fatores, foi selecionada uma metodologia a ser testada no decorrer do projeto, tanto com *P. pastoris* quanto com *E. coli*.

3 EXPRESSÃO DE PROTEÍNAS RECOMBINANTES EM *E. coli*

A *Escherichia coli* é uma das hospedeiras de expressão mais utilizadas para produção de proteínas recombinantes. Em comparação com outros sistemas de expressão, o sistema bacteriano apresenta diversas vantagens como baixo custo, alta produtividade, genes bem mapeados e conhecidos, meios de cultura simples e acessíveis, crescimento rápido, fácil indução com IPTG (isopropil β -D-1-tiogalactopiranosídeo) e uma purificação simples. Contudo este sistema também apresenta algumas desvantagens quanto à expressão insuficiente e baixa solubilidade da proteína de interesse. Com a pesquisa de fermentação recombinante, seria possível realizar uma produção econômica de uma proteína desejada, obtendo-se maior quantidade de proteína em determinado volume, em um período de tempo menor. Contudo existem fatores que influenciam nos níveis de expressão de uma proteína recombinante em *E. coli*., estes sendo primariamente centrados em torno do gene a ser expresso, o vetor de expressão utilizado, o hospedeiro de expressão e a condição de cultura (HIBIBI *et al.* 2015; MIYAMOTO, 2013; LAEMMLI, 1970).

Outrossim, por se tratar de um sistema barato e simples, porém extremamente promissor, a expressão de proteínas em *E. coli* está sob constantes estudos de otimizações de suas condições de expressão. Os fatores



abordados a seguir serão o meio de cultura, as condições de indução, a temperatura e o tempo de cultivo pós-indução e o rendimento apresentado pelo trabalho.

O estudo de Sadeghian-Rizi *et al.* (2019) se baseou na investigação de uma metodologia para aumentar a solubilidade da proteína recombinante, diminuindo o estresse celular e aumentando a produtividade. Já o objetivo de Montfort-Gardeazabal *et al.* (2021) foi expressar um peptídeo antimicrobiano utilizando proteínas de fusão a fim de obter uma proteína recombinante solúvel que contenha o peptídeo desejado. Os trabalhos apresentaram disparidades em suas condições de indução, incluindo a concentração de IPTG (1 mM no primeiro e 0,1 mM no segundo). Destarte, em ambos foram utilizados o meio LB suplementado com antibióticos, 25 °C e 16 h como temperatura e tempo de crescimento pós-indução. Sadeghian-Rizi *et al.* (2019) citam que com os parâmetros ótimos a expressão total foi aumentada em 5 vezes e Montfort-Gardeazabal *et al.* (2021) atingiram um rendimento de 10 mg/L com essas condições de expressão, em comparação com 6 mg/L obtidos em diferentes condições.

Da mesma forma que o trabalho de Montfort-Gardeazabal *et al.* (2021), Wu *et al.* (2021) utilizaram IPTG 0,1 mM como indutor, porém com a diferença do COCl₂ 0,5 mM. A temperatura de expressão desse trabalho foi próxima as dos supracitados, tendo sido definida em 24 °C, porém seu tempo de expressão foi de 12 h.

Socol *et al.* (2019) propôs a otimização dos parâmetros de temperatura, concentração do indutor e tempo de incubação para expressão da leptina. Suas condições de indução foram IPTG 0,01 mM a 37 °C e 180 rpm, e seus parâmetros de expressão foram temperatura e tempo pós-indução de 37 °C e 5 h, respectivamente, assemelhando-se ao trabalho de Marivate *et al.* (2021), cuja temperatura pós-indução foi a mesma citada anteriormente e o tempo de 2 h que, apesar de diferir do primeiro, foi diminuto da mesma maneira. Os trabalhos distinguiram-se também nos meios de cultura, Socol *et al.* tendo utilizado o meio LB suplementado com antibióticos e glucose 1 % e Marivate *et al.* o meio YT com antibióticos. Ambos, porém, afirmaram que os parâmetros se referiram aos maiores rendimentos de seus respectivos trabalhos.

Com o objetivo de expressar a proteína surfactante B dos pulmões de ratos de forma recombinante, utilizando uma proteína de fusão durante a expressão, Cao *et al.* (2021) induziram as células de *E. coli* com IPTG 1 mM, da mesma forma que Gao *et al.* (2020) e Maddi *et al.* (2021), cujos objetivos foram expressar uma proteína da subfamília LEA-3 originária da planta de chá *Camellia sinensis* (L.) O. Kuntze; e otimização de dez diferentes proteínas de fusão SAS-6 N-terminal, respectivamente. Os três trabalhos utilizaram a mesma temperatura de 30 °C pós-indução, diferindo apenas nos tempos de expressão que foram de 12, 10 e 4 h. Cao *et al.* concluíram que as condições de expressão permitiram alcançar a expressão solúvel da proteína desejada, da mesma forma que Maddi *et al.* Gao *et al.* citaram que a expressão heteróloga da proteína aumentou a tolerância da *E. coli* ao estresse pelo frio.

Diferentemente dos trabalhos anteriores, Dong *et al.* (2021) com o objetivo de expressar a proteína humana hATG7 e Larsen *et al.* (2008) que expressaram a lipase B da *Candida* em sistema bacteriano a fim de estudar sua glicosilação, utilizaram temperaturas mais baixas de expressão (20 e 16 °C, respectivamente), durante o mesmo tempo de 24 °C. Em ambos os trabalhos foi utilizado IPTG 0,1 mM para indução da expressão. Destarte, os dois afirmam terem atingido altos rendimentos durante as expressões.

Após analisar diversas metodologias de expressão de proteínas recombinantes utilizando a bactéria *E. coli*, foi possível descrever os parâmetros ótimos de expressão: meio de cultivo LB, indução com IPTG 1 mM, temperatura de 30 °C e tempo de expressão de 12 h.

4 EXPRESSÃO DE PROTEÍNAS RECOMBINANTES EM *P. pastoris*

A maior vantagem em utilizar o sistema de *Pichia pastoris*, em comparação com sistemas bacterianos, é que a levedura apresenta o potencial para realizar várias das modificações pós-traducionais tipicamente associadas a eucariotos superiores. Dentre estas, podem realizar o processamento de sequências de sinal, dobramento, formação de pontes dissulfeto, certos tipos de adição de lipídios e N e O-glicosilações. Para a



expressão de qualquer gene exógeno utilizando a *P. pastoris*, é necessário seguir três etapas básicas. A primeira é a inserção do gene em um vetor de expressão, seguida da introdução do vetor de expressão no genoma levedura, a última etapa é o exame de cepas de expressão potenciais para o produto do gene exógeno (CEREGHINO *et al.*, 2000). Outrossim, foi feito o levantamento bibliográfico comparativo das condições de indução e expressão de diferentes proteínas em *Pichia pastoris*.

Allonso *et al.* (2019) utilizaram o gene *DENV* do vírus da dengue com códons otimizados para o sistema de expressão em levedura. Para tal as células foram crescidas durante a noite a 30 °C e foram induzidas por 72 h com 0,7 % de metanol a cada 24 h. O trabalho de Duan *et al.* (2019), cujo objetivo foi a investigação dos efeitos de novos peptídeos sinais, foi conduzido de forma semelhante, diferindo apenas na concentração de metanol – 1 %. Em ambos os trabalhos foram utilizados os meios YPD e BMGY suplementados com antibióticos, além de apresentarem resultados semelhantes, com boa expressão das proteínas desejadas.

O trabalho de Li *et al.* (2020) teve como foco a otimização da expressão da proteína E2 do vírus da peste suína. As células foram mantidas a 28 °C e 250 rpm durante a indução, da mesma forma que no trabalho de Li *et al.* (2014), cujo objetivo foi a expressão da proteína de fusão M-IL-2. Os trabalhos diferiram na quantidade de metanol para indução celular, tendo o primeiro feito a indução com 150 µL a cada 24 h e o segundo adicionado metanol de forma a concentração final deste ser 0,5 % a cada 24 h. Os tempos de indução foram 96 e 120 h, respectivamente. Ao fim dos trabalhos Li *et al.* (2020) citaram que obtiveram uma alta concentração da proteína desejada, porém Li *et al.* (2014) afirmaram que a concentração de proteínas secretoras totais atingida foi de 1125 mg/L.

Com o objetivo de clonar e expressar o gene (BbCSN-1) da quitosanase de *Bauveria bassiana* em *P. pastoris*, Liu *et al.* (2020) utilizaram a mesma temperatura de indução que Li *et al.* (2014) e Li *et al.* (2020), assemelhando-se até mesmo nas condições de indução, onde adicionaram 0,5 % de metanol a cada 24 h. Seu tempo de indução, porém foi de 6 dias. Ao fim de seu trabalho a proteína alvo havia sido expressada.

Zhao *et al.* (2018), utilizaram a temperatura de 30 °C para indução e adicionou 0,5 % de metanol a cada 12 h, diferindo de L e Vasu *et al.* (2019) que, apesar de ter utilizado a mesma temperatura e concentração de metanol, fez a adição a cada 24 h. O tempo de indução do primeiro autor foi de 120 h, enquanto do segundo foi de 96 h. Apesar das discrepâncias, ambos os trabalhos afirmam que as proteínas foram expressas com sucesso.

Diferentemente dos trabalhos citados anteriormente, Lee *et al.* (2017) adicionaram, durante a indução, 0,4 % de glicerol, juntamente com 0,5 % de metanol, durante 48 h a 30 °C. Ele afirma que o método de autoindução é capaz de expressar proteínas de membrana em níveis comparáveis ao método tradicional de troca de meio. Semelhantemente a Lee *et al.* (2017), Akishev *et al.* (2021) também adicionaram uma substância além de metanol durante a indução, no caso dele a glucose que chega na concentração final de 0,8 %. Hopson *et al.* (2018) citam que induziu as células a 25 °C por 24 h, mas não detalha a concentração de metanol utilizada. Porém, ao fim de seu trabalho, sua proteína foi expressa com sucesso.

Após a análise bibliográfica os parâmetros ótimos de expressão para *P. pastoris* foram definidos como: 38 °C, 250 rpm e 120 h para condições de crescimento; os meios YPD e BMGY com suplementos de antibióticos, pois foram os mais utilizados na maioria das pesquisas; a indução utilizando 100 % de metanol (para concentração final ser de 0,5 %) a cada 24 h, com tempo total de indução de 120 h.

5 CONCLUSÃO

Os parâmetros ótimos definidos anteriormente serão utilizados de forma a avaliar e otimizar a expressão do fragmento de um anticorpo monoclonal recombinante, que é o foco principal do trabalho. Cada pesquisa trabalhou e expressou proteínas diferentes, com características, massa molecular e rendimentos distintos. Portanto, estes parâmetros foram escolhidos levando em consideração o alto rendimento apresentado e semelhanças entre o peso molecular da proteína que será utilizada no projeto.



AGRADECIMENTOS

Gostaríamos de registrar os agradecimentos a UTFPR-Toledo, ao CNPQ e a CAPES por disponibilizar os equipamentos e laboratórios, incentivar e auxiliar nas questões financeiras para continuidade da pesquisa.

REFERÊNCIAS

- AKISHEV, Z.; KIRIBAYEVA, A.; MUSSAKHMETOV, A.; BALTIN, K.; RAMANKULOV, Y.; KASSENOV, B. Constitutive expression of *Camelus bactrianus* prochymosin B in *Pichia pastoris*. **Heliyon**, v. 7, 2021.
- ALLONSO, D.; PEREIRA, I.B.; ALVES, A.M.B.; KURTENBACH, E.; BORGES-MOHANA, R. Expression of soluble, glycosylated and correctly folded dengue virus NS1 protein in *Pichia pastoris*. **Protein Expression and Purification**, v. 162, p. 7-17, 2019.
- CAO, W.; LIU, Q.; WANG, T.; ZHANG, Q.; CHENG, F.; TANG, Y.; MEI, C.; WEN, F.; WANG, W. Recombinant expression of the precursor of rat lung surfactant protein B in *Escherichia coli* and its antibacterial mechanism. **Protein Expression and Purification**, v. 179, 2021.
- CEREGHINO J.L.; CREGG J.M. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. **FEMS Microbiol Rev.** v. 24(1), p. 45-66, 2000.
- DONG, G.; ZHAO, X.; GUO, J.; MA, L.; ZHOU, H.; LIU, Q.; ZHAO, X.; WANG, C.; WU, K. Functional expression and purification of recombinant full-length human ATG7 protein with HIV-1 Tat peptide in *Escherichia coli*. **Protein Expression and Purification**, v. 182, 2021.
- DUAN, G.; DING, L.; WEI, D.; ZHOU, H.; CHU, J.; ZHANG, S.; QIAN, J. Screening endogenous signal peptides and protein folding factors to promote the secretory expression of heterologous proteins in *Pichia pastoris*. **Journal of Biotechnology**, v. 306, p. 193-202, 2019.
- GAO, T.; MO, Y.; HUANG, H.; YU, J.; WANG, Y.; WANG, W. Heterologous expression of *Camellia sinensis* late embryogenesis abundant protein gene 1 (CsLEA1) confers cold stress tolerance in *Escherichia coli* and yeast. **Horticultural Plant Journal**, v. 7, p. 89-96, 2021.
- HOPSON, S. E.; THOMPSON, M. J. Heterologous expression of the human polybromo-1 protein in the methylotrophic yeast *Pichia pastoris*. **Protein Expression and Purification**, v. 152, p. 23-30, 2018.
- KANE, J.F.; HARTLEY, D.L. Formation of recombinant protein inclusion bodies in *Escherichia coli*. **TIBTECH** v. 6, p. 95-101, 1988.
- L, S.; VASU, P. Cloning and expression of *in silico* modeled protein enriched with branched chain amino acids in *Pichia pastoris*. **Journal Pre-proofs**, 2019.
- LAEMMLI, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, v. 227, p. 680-685, 1970
- LARSEN, M.W.; BOENSCHUEER, U.T.; HULT, K. Expression of *Candida antarctica* lipase B in *Pichia pastoris* and various *Escherichia coli* systems. **Protein Expression and Purification**, v. 62, p. 90-97, 2008.
- LEE, J. Y.; CHEN, H.; LIU, A.; ALBA, B. M.; LIM, A. C. Auto-induction of *Pichia pastoris* AOX1 promoter for membrane protein expression. **Protein Expression and Purification**, v. 137, p. 7-12, 2017.



LI, D.; WU, J.; CHEN, J.; ZHANG, D.; ZHANG, Y.; QIAO, X.; YU, X.; ZHENG, Q.; HOU, J. Optimized expression of classical swine fever virus E2 protein via combined strategy in *Pichia pastoris*. **Protein Expression and Purification**, v. 167, 2020.

LI, L.; QIAN, D.; SHAO, G.; YAN, Z.; LI, R.; HUA, X.; SONG, X.; WANG, B. High-level expression purification and study of bioactivity of fusion protein M-IL-2 (88Arg, 125 Ala) in *Pichia pastoris*. **Protein Expression and Purification**, v. 101, p. 99-105, 2014.

LIU, Y.; LI, Y.; TONG, S.; YUAN, M.; WANG, X.; WANG, J.; FAN, Y. Expression of *Beauveria bassiana* chitosanase (BbCSN-1) in *Pichia pastoris* and enzymatic analysis of the recombinant protein. **Protein Expression and Purification**, v. 166, 2020.

MADDI, E. R.; NATESH, R. Optimization strategies for expression and purification of soluble N-terminal domain of human centriolar protein SAS-6 in *Escherichia coli*. **Protein Expression and Purification**, v. 183, 2021.

MARIVATE, A.; NJENGELE-TETYANA, Z.; FISH, M. Q.; MOSEBI, S. Recombinant expression, purification, and characterization of full-length human BST-2 from *Escherichia coli*. **Protein Expression and Purification**, v. 188, 2021.

MIYAMOTO, A.C. **Produção de proteínas recombinantes em *Escherichia coli***. Revista digital. 2013.

MONTFOR-GARDEAZABAL, J. M.; BALDERAS-RENTERIA, I.; CASILLAS-VEJA, N.G.; ZARATE, X. Expression and purification of the antimicrobial peptide Bin1b in *Escherichia coli* tagged with the fusion proteins CusF3H+ and SmbP. **Protein Expression and Purification**, v. 178, 2021.

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.

SADEGHIAN-RIZI T.; EBRAHIMI A.; MOAZZEN F.; YOUSEFIAN H.; JAHANIAN-NAJAFABADI, A. Improvement of solubility and yield of recombinant protein expression in *E. coli* using a two step system. **Res Pharm Sci**. p. 400-407, 2019.

SOCOL, C.T.; TRIF, M.; CRISTE, F. L.; MIERLITĂ, D.; SONEA, C. G.; RUSU, A. V. Optimization of expression and purification of recombinant protein in *E. coli* based on leptin model. **Rom biotechnol Lett**, v.25(4), p. 1810-1815, 2020.

WU, Z.; ZHANG, Z.; CAI, S.; ZHENG, R.; ZHENG, Y. High-level expression of nitrile hydratase from *Pantoea* sp. At-9b in *Escherichia coli*. **Process Biochemistry**, v. 101, p. 199-206, 2021.

YANG, Z.; ZHANG, Z. Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: a review. **Biotechnology Advances** v. 36, p. 182-195, 2018.

ZHAO, T.; LI, Z.; GUO, Z.; WANG, A.; LIU, Z.; ZHAO, Q.; LI, Y.; MCKENZIE, E. A.; DIAO, A. Functional recombinant human protein expression in *Pichia pastoris* to enable screening for Legumain small molecule inhibitors. **Protein Expression and Purification**, v. 150, p. 12-16, 2021.