UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA
Departamento de Engenharia Química

GABRIELA PANNUNZIO CARMIGNOTTO

**Estudo das condições de cultivo de *Escherichia coli* para a produção da
nuclease Cas9, utilizada em sistemas de edição de genoma CRISPR-Cas9**

**SÃO PAULO
2019**

UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA
Departamento de Engenharia Química

GABRIELA PANNUNZIO CARMIGNOTTO

**Estudo das condições de cultivo de *Escherichia coli* para a produção da
nuclease Cas9, utilizada em sistemas de edição de genoma CRISPR-Cas9**

**SÃO PAULO
2019**

GABRIELA PANNUNZIO CARMIGNOTTO

Estudo das condições de cultivo de *Escherichia coli* para a produção da nuclease Cas9, utilizada em sistemas de edição de genoma CRISPR-Cas9

Versão Corrigida

(Versão original encontra-se na unidade que aloja o Programa de Pós-graduação)

Tese apresentada ao Programa de Pós-Graduação em Engenharia Química da Escola Politécnica da Universidade de São Paulo para obtenção do título de Doutora em Ciências.

Área de concentração: Engenharia Química

Orientador: Prof. Dr. Adriano Rodrigues Azzoni

SÃO PAULO

2019

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catálogo-na-publicação

Carmignotto, Gabriela Pannunzio

Estudo das condições de cultivo de *Escherichia coli* para a produção da nuclease Cas9, utilizada em sistemas de edição de genoma CRISPR-Cas9 / G. P. Carmignotto -- versão corr. -- São Paulo, 2019.

137 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Química.

1.Biotecnologia 2.Bioprocessos 3.*Escherichia coli* 4.Proteínas recombinantes 5.Reatores bioquímicos I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Química II.t.

CARMIGNOTTO, G. P. **Estudo das condições de cultivo de *Escherichia coli* para a produção da nuclease Cas9, utilizada em sistemas de edição de genoma CRISPR-Cas9.** Versão corrigida. 2019. Tese (Doutorado no programa de engenharia química) – Departamento de Engenharia Química da Escola Politécnica, Universidade de São Paulo, São Paulo, 2019.

Aprovada em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

AGRADECIMENTOS

Agradeço ao professor Dr. Adriano Rodrigues Azzoni pela oportunidade, orientação, experiência e paciência no decorrer deste trabalho.

Aos meus pais, Ana e Atílio, e ao meu irmão Marco, pelo apoio e incentivo em todos os momentos da minha vida.

À minha prima Andrea, que me motivou e esteve sempre ao meu lado durante esta jornada. Às minhas amigas Thais, Tamires e Graziela, que me apoiaram em todos os momentos. À minha tia Carmem, que esteve sempre pronta a me ajudar.

Aos amigos e colegas de laboratório, em especial à Letícia, Isabella, Daniela, Carla, Ehsan, Orlinda, Andrea, Leme, Walter, Rafael da Gama, Rafael Risnik, Bianca, Georges, Davi, Sara, Fernando, Matheus, Guilherme, Karina, Luan, Juliano, Larissa e Suzanna, por toda ajuda e amizade. À Marina, que me apoiou nos primeiros anos de doutorado.

Ao professor Dr. Aldo Tonso por disponibilizar as instalações de seu laboratório.

A todos os professores e funcionários da Engenharia Química da Poli- USP que contribuíram para o desenvolvimento deste trabalho.

Agradeço todo o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Processos nº 444412/2014-0, nº 307739/2015-5 e 304125/2018-0) e pela concessão da bolsa de estudo. Agradeço também à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/PROEX), Código Financeiro 001.

RESUMO

CARMIGNOTTO, G. P. **Estudo das condições de cultivo de *Escherichia coli* para a produção da nuclease Cas9, utilizada em sistemas de edição de genoma CRISPR-Cas9**. 2019. Tese (Doutorado no programa de engenharia química) – Departamento de Engenharia Química da Escola Politécnica, Universidade de São Paulo, São Paulo, 2019.

A técnica CRISPR-Cas9 é uma ferramenta que vem sendo amplamente estudada na edição de genomas. Consiste em um complexo formado por uma nuclease chamada Cas9, que tem a função de cortar o DNA, e em um RNA guia (gRNA), que contém a sequência de reconhecimento do local de corte para a edição. A técnica CRISPR-Cas9 vem ganhando destaque para sua aplicação como ferramenta de biologia molecular e no tratamento de doenças, o que potencializa o uso da proteína Cas9 como biofármaco. Embora amplamente empregada, poucos são os estudos encontrados sobre a produção da proteína recombinante Cas9, de 158 kDa, derivada de *Streptococcus pyogenes*. Neste contexto, este trabalho teve como objetivo realizar um estudo sistemático da produção da proteína Cas9 contendo uma cauda de histidina em meio de cultura complexo e meio definido, empregando o vetor pET28b-Cas9His e duas cepas de *E. coli*: BL21(DE3) e BL21(DE3) Rosetta. Os resultados mostraram que a expressão da proteína Cas9 foi maior em *E. coli* BL21(DE3), enquanto que apenas traços da proteína Cas9 foram detectados na Rosetta, resultado do baixo nível de mRNA. O crescimento celular da *E. coli* BL21(DE3) e a expressão da proteína foram avaliados em meio complexo (LB) e meio definido (HDF) em frascos. A melhor condição de indução para a expressão da proteína Cas9 ocorreu em meio definido HDF, com adição de IPTG e temperatura de indução de 30°C, em um intervalo de tempo de 8h, obtendo-se a concentração final de 96,4 mg de Cas9/L de cultivo. Ao empregar biorreator operado em modo batelada com meio definido HDF, a concentração celular, a velocidade específica máxima de crescimento e o fator de conversão de substrato em biomassa obtidos foram 5,6 g/L, 0,53 h⁻¹ e 0,55, respectivamente. A concentração de proteína obtida em biorreator foi de 420,1 mg de Cas9/L de cultivo com 6h de indução. Este trabalho possibilitou produzir a proteína Cas9 de forma eficiente ao empregar-se a linhagem *E. coli* BL21(DE3) em biorreator usando meio definido HDF, com potencial para melhorias adicionais que poderiam facilitar a produção em larga-escala.

Palavras chave: CRISPR-Cas9; proteína Cas9; modo batelada; meio definido.

ABSTRACT

CARMIGNOTTO, G. P. **Study of the culture conditions of *Escherichia coli* for the production of Cas9 nuclease, used in CRISPR-Cas9 genome-editing systems**. 2019. Thesis (PhD in Chemical Engineering) – Department of Chemical Engineering of the Polytechnic School, University of Sao Paulo, Sao Paulo, 2019.

The CRISPR-Cas9 system is a tool that has been extensively studied in genome editing. It consists of a complex composed of a protein called Cas9, which has the function of cleaving the DNA at a specific site, and a guiding RNA (gRNA), which contains the cleavage target for editing. The CRISPR-Cas9 system has been gaining importance for its application as a molecular biology tool and in the treatment of diseases, which potentiates the use of the Cas9 protein as a biopharmaceutical. Although widely employed, few studies have been found on the production of the recombinant Cas9 protein (158 kDa) derived from *Streptococcus pyogenes*. In this context, this work aims to study the production of the Cas9 protein with a His-tag in complex and defined media using the pET28b-Cas9His vector and two strains of *E. coli*: BL21 (DE3) and BL21 (DE3) Rosetta. The results showed that Cas9 protein expression was higher in *E. coli* BL21 (DE3), whereas traces of the Cas9 protein were detected in Rosetta, as a result of the low level of mRNA. Cell growth of *E. coli* BL21 (DE3) and protein expression were evaluated in complex medium (LB) and defined medium (HDF) in shake-flasks. The best induction condition for Cas9 protein expression in shake-flasks occurred in HDF medium with addition of IPTG and induction temperature of 30°C in a time interval of 8h, reaching concentration of 96.4 mg Cas9/L culture. The cell concentration, the maximum specific growth rate and the substrate to cell yield factor obtained in bioreactor in batch mode was 5,6 g/L, 0,53 h⁻¹ e 0,55, respectively. The protein concentration obtained was 420.1 mg/L with 6 hours of induction time. This work showed the possibility of produce Cas9 protein efficiently using *E. coli* BL21(DE3) using defined HDF medium, with potential for further improvements that could facilitate large-scale production.

Keywords: CRISPR-Cas9; Cas9 protein; batch mode; defined medium.

LISTA DE FIGURAS

- Figura 2.1.** Sistema CRISPR-Cas9 nas bactérias. Fase1: armazenamento de sequências de DNA invasores. Fase 2: mecanismo do sistema CRISPR-Cas9 contra um DNA invasor. (Fonte: Adaptado de MALI *et al.*, 2013a)6
- Figura 2.2.** A: Representação do complexo Cas9-gRNA-DNA. B: Representação esquemática do complexo gRNA-DNA. C: Representação esquemática do sistema CRISPR-Cas9. (Fonte: Adaptado de NISHIMASU *et al.*, 2014 e HU *et al.*, 2016)7
- Figura 2.3.** Plasmídeo pRARE contendo os genes para os tRNAs correspondentes aos códons raros em *E. coli* (indicado em azul). Apresenta gene de resistência ao cloranfenicol. (Fonte: adaptado de NOVY *et al.*, 2001).18
- Figura 2.4.** Representação do metabolismo aeróbio e anaeróbio (adaptado de KONZ *et al.*, 1998).24
- Figura 3.1.** Vetor de expressão pET28b contendo o gene da proteína Cas9 (Fonte: <https://www.addgene.org/47327/>)30
- Figura 4.1.** Análise do vetor pET28b-Cas9His purificado de BL21(DE3) em eletroforese de gel de agarose 1%. Ladder- DNA Ladder 1kb RTU (Kasvi, Brasil); 1- Vetor pET28b-Cas9His não digerido; 2- Vetor pET28b-Cas9His digerido com a enzima NcoI; 3- Vetor pET28b-Cas9His digerido com as enzimas NcoI e NotI.46
- Figura 4.2.** Análise do vetor pET28b-Cas9His purificado de BL21(DE3) Rosetta em eletroforese de gel de agarose 1%. Ladder- DNA Ladder 1kb RTU (Kasvi, Brasil); 1- Vetor pET28b-Cas9His digerido com a enzima NcoI; 2- pRARE digerido com a enzima NotI; 3- Vetor pET28b-Cas9His e pRARE digeridos com a enzima NotI; 4- pRARE digerido com as enzimas NotI e NcoI; 5- Vetor pET28b-Cas9His e pRARE digeridos com as enzimas NotI e NcoI.46
- Figura 4.3.** Perfil eletroforético do cultivo da *E. coli* BL21(DE3) em meio LB (indução: 0,5 mM de IPTG à 25°C por 4h) para avaliar a expressão da proteína Cas9. Amostras dos *pellets* e das etapas de purificação da proteína empregando resina de Ni-agarose. Revelação por *Et al.* (a) Rompimento celular por 8 ciclos em ultrassom. (b) Rompimento celular por 16 ciclos em ultrassom. M- Marcador de massa molecular; 0h- *Pellet* com 0h de indução; Ins- Fração insolúvel após a sonicação; I- Amostra da injeção da coluna de purificação (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E- Eluição.47
- Figura 4.4.** Perfil eletroforético do cultivo da *E. coli* BL21(DE3) em meio LB (indução: 0,5 mM de IPTG à 25°C por 4h) para avaliar a expressão da proteína Cas9 e o

emprego de PMSF. Revelação por *Et al.*. M- marcador de massa molecular; Is- Injeção (fração solúvel após a sonicação) sem PMSF; Es- Eluição sem PMSF; Ic- Injeção (fração solúvel após a sonicação) com PMSF; Es- Eluição com PMSF.48

Figura 4.5. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta sem o plasmídeo pET28b-Cas9His em 100 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2, E3 e E4- Frações da etapa de eluição.51

Figura 4.6. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 100 mL de meio ZYP-5052 (indução: lactose à 27°C por 16h) e etapas da purificação para avaliar a expressão da proteína Cas9. Rompimento celular empregando 16 ciclos de sonicação e revelação por *Et al.*. M- marcador de massa molecular; S- sobrenadante extracelular; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2, E3 e E4- Frações da etapa de eluição.52

Figura 4.7. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 100 mL de meio ZYP-5052 (indução: lactose à 18°C por 24h) e etapas da purificação para avaliar a expressão da proteína Cas9. Rompimento celular empregando 16 ciclos de sonicação e revelação por *Et al.*. M- marcador de massa molecular; P- *pellet* após 24h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E - Eluição.53

Figura 4.8. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 100 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; R_{sem}- *pellet* Rosetta sem indução; R_{com}- *pellet* Rosetta após 4h de indução; BL_{sem}- *pellet* BL21(DE3) sem indução; BL_{com}- *pellet* BL21(DE3) após 4h de indução.53

Figura 4.9. Análise dos vetores pET28b-Cas9His e pRARE purificados de BL21(DE3) Rosetta em eletroforese de gel de agarose 1%. Ladder- DNA Ladder 1kb RTU (Kasvi, Brasil); 1- Vetor pRARE não linear; 2- Vetores pET28b-Cas9His e pRARE não lineares; 3- pRARE digerido com a enzima NotI; 4- Vetores pET28b-Cas9His e pRARE digeridos com a enzima NotI.....54

Figura 4.10. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 100 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) para avaliar a expressão da Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Colônia (1 e 2)- *pellet* após 4h de indução do cultivo com inóculo original de uma colônia de bactéria; Banco (1 e 2)- *pellet* após 4h de indução do cultivo com inóculo original do banco celular em -80°C.55

Figura 4.11. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 100 mL de meio LB (indução: 0,5 mM de IPTG à 18°C por 4 e 16h) para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada; Sem- *pellet* sem indução; 4h- *pellet* após 4h de indução; 16h- *pellet* após 16h de indução.56

Figura 4.12. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 850 mL de meio ZYP-5052 (indução: lactose à 18°C por 24h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada; Sem- *pellet* sem indução; Com- *pellet* após 24h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.57

Figura 4.13. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 500 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.58

Figura 4.14. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 500 mL de meio ZYP-5052 (indução: lactose à 18°C por 24h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Com- *pellet* após 24h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2, E3 e E4- Frações da etapa de eluição.58

Figura 4.15. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 500 mL de meio TB (indução: 0,5 mM de IPTG à 18°C por 20h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.59

Figura 4.16. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 250 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e etapas da purificação com 4 ciclos para rompimento celular. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.59

Figura 4.17. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) em 250 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa

molecular; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.60

Figura 4.18. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 250 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.61

Figura 4.19. *Western blot* do cultivo da *E. coli* BL21(DE3) e da BL21(DE3) Rosetta em 100 mL meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e frações da etapa de eluição da purificação para avaliar a expressão da proteína Cas9. M- marcador de massa molecular; R_{sem}- *pellet* da Rosetta sem indução; R_{com}- *pellet* da Rosetta com indução; E2_R e E3_R- frações 2 e 3 da etapa de eluição do cultivo com Rosetta; BL_{sem}- *pellet* da BL21(DE3) sem indução; BL_{com}- *pellet* da BL21(DE3) com indução; E2_{BL} e E3_{BL}- frações 2 e 3 da etapa de eluição do cultivo com BL21(DE3).61

Figura 4.20. Gel de agarose 1% para determinar a temperatura de anelamento dos *primers* 2 e 3 (Tabela 3.1) na amplificação do fragmento de 122 pb do gene da proteínas Cas9. 1- GeneRuler 1 kb Plus DNA Ladder (Thermo Fisher Scientific, EUA); 2- amostra diluída 10x de pET28b-Cas9His; 3- resultado da PCR convencional para amplificação do fragmento de 122 pb com temperatura de anelamento de 54°C.....63

Figura 4.21. Resultados da qRT-PCR do plasmídeo pET28b-Cas9His para determinar a curva padrão em número de cópias por C_T. (a) Curvas de amplificação das amostras de pET28b-Cas9His contendo número de cópias iniciais de: 6.10⁷, 6.10⁶, 6.10⁵, 6.10⁴, 6.10³, 6.10² cópias; (b) Curvas de *melting* (dissociação) para cada amostra; (c) Curva padrão da qRT-PCR (quantidade de pET28b-Cas9His x C_T).64

Figura 4.22. Resultados da qRT-PCR das amostras de *E. coli* BL21(DE3) com e sem o plasmídeo pET28b-Cas9His, com e sem indução. (a) Curvas de amplificação; (b) Curvas de *melting* (dissociação).65

Figura 4.23. Resultados da qRT-PCR das amostras de *E. coli* BL21(DE3) Rosetta com e sem o plasmídeo pET28b-Cas9His, com e sem indução. (a) Curvas de amplificação; (b) Curvas de *melting* (dissociação).66

Figura 4.24. Valores relativos de cópias de mRNA da proteína Cas9 nas amostras de *E. coli* BL21(DE3) e BL21(DE3) Rosetta com e sem o pET28b-Cas9His (+pET; -pET), induzidas ou não (-ind; +ind), normalizadas para o maior número de cópias

obtido (amostra da cepa BL21(DE3) contendo o pET28b-Cas9His após a indução). As barras de erro indicam o desvio padrão entre triplicatas.....67

Figura 4.25. Perfil do crescimento celular (●), concentração de glicose (▲), concentração de acetato (◆) e pH (■) do cultivo de *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em meio LB a 37°C e 250 rpm. Os experimentos foram realizados em triplicata e os valores de desvio padrão foram baixos (em torno de 0,03), não sendo possível visualizar as barras de erros no gráfico.....69

Figura 4.26. Perfil do crescimento celular (●), concentração de glicose (▲), concentração de acetato (◆) e pH (■) do cultivo de *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em meio HDF a 37°C e 250 rpm. Os experimentos foram realizados em triplicata e os valores de desvio padrão foram baixos (em torno de 0,03), não sendo possível visualizar as barras de erros no gráfico.....69

Figura 4.27. Determinação do parâmetro cinético velocidade específica máxima de crescimento ($\mu_{\text{máx}}$) da *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em (a) meio LB e (b) meio HDF a 37°C e 250 rpm.....70

Figura 4.28. Perfil eletroforético da expressão da Cas9 em *E. coli* BL21(DE3) nos meios LB e HDF (indução: 0,5 mM de IPTG à 25°C por 4h). Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada; LBsem- *pellet* sem indução; LBcom- *pellet* após 4h de indução; HDFsem- *pellet* sem indução; HDFcom- *pellet* após 4h de indução.72

Figura 4.29. Perfil eletroforético da expressão da Cas9 em *E. coli* BL21(DE3) em meio HDF em diferentes temperaturas de indução (indução: 0,5 mM de IPTG por 4h). As amostras contém *pellet* após 4h de indução nas temperaturas 18, 25, 30 e 37°C. Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada.....75

Figura 4.30. Perfil eletroforético da expressão da proteína Cas9 em *E. coli* BL21(DE3) utilizando meio HDF com indução à 30°C e 0,5 mM de IPTG para diferentes tempos de pós-indução (0h, 2h, 4h, 6h, 8h, 18h e 24h). Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada.....76

Figura 4.31. Perfil do crescimento celular (●), concentração de glicose (▲), concentração de acetato (◆) e pH (■) do cultivo de *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em meio HDF a 37°C em biorreator.....78

Figura 4.32. Determinação do parâmetro cinético velocidade específica máxima de crescimento ($\mu_{\text{máx}}$) da *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em meio HDF a 37°C em biorreator.....78

Figura 4.33. Variáveis temperatura (°C) (●), pH (●), rotação (rpm) (●) e oxigênio dissolvido (%) (●) ao longo de todo o cultivo em biorreator coletados pelo programa *LabVIEW*.79

Figura 4.34. Perfil do crescimento celular do cultivo de *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em meio HDF e biorreator na expressão da proteína Cas9 em dois instantes de indução. (●) Crescimento celular sem indução; (●) indução no início da fase exponencial de crescimento celular e (●) indução na metade da fase exponencial de crescimento celular. As setas indicam os dois instantes de indução que serão analisados para a expressão da proteína Cas9.81

Figura 4.35. Variáveis temperatura (°C) (●), pH (●), rotação (rpm) (●) e oxigênio dissolvido (%) (●) ao longo de todo o cultivo em biorreator na expressão da proteína Cas9 com indução no início da fase exponencial de crescimento celular coletados pelo programa *LabVIEW*.82

Figura 4.36. Variáveis temperatura (°C) (●), pH (●), rotação (rpm) (●) e oxigênio dissolvido (%) (●) ao longo de todo o cultivo em biorreator na expressão da proteína Cas9 com indução na metade da fase exponencial de crescimento celular coletados pelo programa *LabVIEW*.82

Figura 4.37. Perfil eletroforético da expressão da proteína Cas9 em *E. coli* BL21(DE3) utilizando meio HDF e biorreator com indução à 30°C e 0,5 mM de IPTG para diferentes instantes de indução. Revelação por *Et al.*. M- marcador de massa molecular; I_{sem}- amostra do instante de indução para o cultivo com indução no início da fase exponencial; I_{8h}- *pellet* após 8h da indução no início da fase exponencial; M_{sem}- amostra do instante de indução para o cultivo com indução na metade da fase exponencial; M_{4h}- *pellet* após 4h da indução na metade da fase exponencial; M_{6h}- *pellet* após 6h da indução na metade da fase exponencial; M_{8h}- *pellet* após 8h da indução na metade da fase exponencial.84

Figura 4.38. Perfil eletroforético da expressão da proteína Cas9 em *E. coli* BL21(DE3) utilizando meio HDF e biorreator (indução: 0,5 mM de IPTG à 30°C por 6h) e etapas da purificação da proteína Cas9 em resina de Ni-NTA-agarose. Revelação por *Et al.*. M- marcador de massa molecular; P_{sem}- *pellet* do instante de indução; P_{com}- *pellet* após 6h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3 - Frações da etapa de eluição.86

Figura 4.39. Gel de agarose 1% para verificar a atividade da proteína Cas9 produzida por *E. coli* BL21(DE3) em meio LB e meio HDF. O complexo contendo a proteína Cas9 purificada e o gRNA foi incubado com o plasmídeo pVAX1-GFP (3697 pb) linearizado e os fragmentos após o ensaio de digestão apresentaram 2960 pb e 737 pb.88

LISTA DE TABELAS

Tabela 2.1. Proteínas recombinantes produzidas por <i>E. coli</i> e aprovadas pelo FDA (Fonte: Adaptado de BERLEC e STRUKELJ, 2013; HUANG <i>et al.</i> , 2012; TRIPATHI, 2016).....	12
Tabela 2.2. Códonos raros para a <i>Escherichia coli</i> (Fonte: BURGESS-BROWN <i>et al.</i> , 2008).....	15
Tabela 3.1. Sequência e Temperatura de <i>melting</i> (T _m) dos <i>primers</i> para a síntese de cDNA e amplificação de fragmentos do gene da proteína Cas9.....	31
Tabela 3.2. Composição do meio HDF.	32
Tabela 3.3. Condições dos experimentos realizados para avaliar a expressão da proteína Cas9 em <i>E. coli</i> BL21(DE3) e BL21(DE3) Rosetta.	37
Tabela 4.1. Códonos raros para <i>E. coli</i> presentes na sequência do gene da proteína Cas9 (número de códonos na sequência: 1390).....	50
Tabela 4.2. Diferentes parâmetros avaliados para a expressão da proteína Cas9 utilizando <i>E. coli</i> BL21(DE3) Rosetta.	51
Tabela 4.3. Valores de concentração celular (g/L) e de massa de proteína Cas9 por litro de cultivo obtidos ao final do cultivo da expressão da proteína Cas9.	73
Tabela 4.4. Valores de concentração celular (g/L) e de massa de proteína Cas9 por litro de cultivo obtidos ao final do cultivo em diferentes temperaturas de indução.	75
Tabela 4.5. Valores de concentração celular (g/L) e de massa de proteína Cas9 por litro de cultivo obtidos para cada tempo de pós-indução.	76
Tabela 4.6. Valores de concentração celular (g/L) e de massa de proteína Cas9 por litro de cultivo obtidos ao final do cultivo da <i>E. coli</i> BL21(DE3) em biorreator para avaliar o efeito do instante de indução.....	84
Tabela 4.7. Recuperação e pureza relativa da proteína Cas9 nas etapas de purificação obtidas por densitometria.....	86

SUMÁRIO

1. INTRODUÇÃO.....	1
1.1. OBJETIVO	4
2. REVISÃO BIBLIOGRÁFICA	5
2.1. CRISPR-Cas9 e proteína Cas9.....	5
2.2. <i>E. coli</i> como sistema de expressão para proteínas recombinantes	12
2.2.1. <i>E. coli</i> BL21(DE3)	17
2.2.2. <i>E. coli</i> BL21(DE3) Rosetta.....	17
2.3. Condições de cultivo para produção de proteínas recombinantes.....	19
2.3.1. Modo de Operação Batelada	19
2.3.2. Meios de cultivo	20
2.3.3. Condições de indução	22
2.3.4. Oxigênio dissolvido.....	23
2.4. Extração da proteína recombinante: rompimento celular por ultrassom	25
2.5. Purificação por cromatografia de afinidade por íons metálicos imobilizados (IMAC).....	26
2.6. Produção da proteína Cas9: situação atual	27
3. MATERIAIS E MÉTODOS	29
3.1. Materiais.....	29
3.1.1. Microrganismos	29
3.1.2. Vetor de expressão.....	29
3.1.3. <i>Template</i> de DNA	30
3.1.4. <i>Primers</i> utilizados na qRT-PCR	31
3.1.5. Meios de cultivo	31
3.2. Métodos	32
3.2.1. Amplificação e purificação de plasmídeo.....	32
3.2.2. Competência das cepas <i>E. coli</i> BL21(DE3) e BL21(DE3) Rosetta	33
3.2.3. Transformação das cepas <i>E. coli</i> BL21(DE3) e BL21(DE3) Rosetta	33
3.2.4. Digestão dos plasmídeos com enzimas de restrição.....	33
3.2.5. Transcrição <i>in vitro</i> de RNA.....	34
3.2.6. Extração de RNA, tratamento com DNase e transcrição reversa	34
3.2.7. PCR (<i>Polimerase Chain Reaction</i>) convencional.....	35

3.2.8.	qRT-PCR (<i>Quantitative Reverse Transcription PCR</i>)	36
3.2.9.	Cultivos em frascos	36
3.2.9.1.	Preparo do inóculo	36
3.2.9.2.	Cinética de crescimento da <i>E. coli</i> em frascos	37
3.2.9.3.	Expressão da proteína Cas9	37
3.2.10.	Cultivo em biorreator da cepa <i>E. coli</i> BL21(DE3)	38
3.2.10.1.	Preparo do inóculo	39
3.2.10.2.	Cinética de crescimento da <i>E. coli</i> BL21(DE3) em biorreator.....	39
3.2.10.3.	Expressão da proteína Cas9 em biorreator	39
3.2.11.	Recuperação e purificação da proteína Cas9	39
3.3.	Métodos analíticos	40
3.3.1.	Determinação da concentração celular	40
3.3.2.	Dosagem de proteínas totais	41
3.3.3.	Dosagem de glicose e acetato.....	41
3.3.4.	Espectrometria de massa- MALDI-TOF MS e MS/MS.....	41
3.3.5.	Eletroforese em gel de agarose.....	41
3.3.6.	SDS-PAGE	42
3.3.7.	<i>Western blot</i>	42
3.3.8.	Densitometria em gel de poliacrilamida (SDS-PAGE)	43
3.3.9.	Atividade da proteína Cas9	43
3.3.10.	Cálculo dos parâmetros cinéticos.....	43
4.	RESULTADOS E DISCUSSÕES.....	45
4.1.	Análise do vetor de expressão pET28b-Cas9His.....	45
4.2.	Verificação da expressão da proteína Cas9.....	47
4.3.	Proteína Cas9 e expressão em <i>E. coli</i> BL21(DE3) Rosetta	49
4.4.	Avaliação da transcrição do gene da proteína Cas9 na BL21(DE3) e BL21(DE3) Rosetta por qRT-PCR	62
4.5.	Expressão da proteína Cas9 em <i>E. coli</i> BL21(DE3)	68
4.5.1.	Cinética de crescimento da <i>E. coli</i> BL21(DE3) em meio LB e meio HDF em frascos	68
4.5.2.	Expressão da proteína Cas9 em meio LB e meio HDF em frascos.....	72
4.5.3.	Avaliação das condições de indução na expressão da proteína Cas9 em frascos 73	
4.5.4.	Cinética de crescimento da <i>E. coli</i> BL21(DE3) em biorreator.....	77

4.5.5. Expressão da proteína Cas9 em biorreator	80
4.6. Purificação da proteína Cas9 por IMAC	85
4.7. Avaliação da atividade da proteína Cas9	87
5. CONCLUSÕES.....	89
6. REFERÊNCIAS BIBLIOGRÁFICAS	91
APÊNDICE A.....	106
APÊNDICE B.....	107
APÊNDICE C	114
APÊNDICE D	115
APÊNDICE E.....	118

1. INTRODUÇÃO

A técnica CRISPR-Cas9 (*Clustered Regularly-Interspaced Short Palindromic Repeats-Cas9*) vem sendo amplamente estudada como ferramenta para a edição de genoma (JINEK *et al.*, 2012; DOUDNA e CHARPENTIER, 2014; DASH *et al.*, 2019). Em 2012, a técnica ganhou destaque após ser utilizada com sucesso na edição de genoma de células humanas (MALI *et al.*, 2013b; WANG *et al.*, 2017). Em seguida, a técnica foi aplicada na edição de genoma *in vivo*, abrindo um extenso campo para sua aplicação no tratamento de doenças genéticas (GAGNON *et al.*, 2014; RAN *et al.*, 2015; YANG *et al.*, 2017). Estudos mais recentes mostraram que a ferramenta CRISPR-Cas9 foi capaz de corrigir o genoma de embrião humano com mutação para problemas cardíacos (MA *et al.*, 2017). Outro estudo demonstrou o emprego da técnica para eliminar genes de vírus presentes no genoma de porcos, um grande avanço no sentido de viabilizar o transplante de órgãos suínos em humanos (NIU *et al.*, 2017). Em 2016, um grupo da China iniciou o primeiro teste clínico usando a técnica CRISPR-Cas9 para editar células *ex vivo* do sistema imunológico de pacientes com câncer de pulmão (CYRANOSKI, 2016). Mais recentemente, Dash e colaboradores (2019) empregaram a técnica CRISPR-Cas9 combinada com a terapia anti-retroviral LASER ART (que reduz a taxa de replicação do vírus HIV) para eliminar o vírus HIV do genoma de camundongos, ampliando ainda mais as possibilidades de aplicação da técnica.

A edição de genoma via CRISPR-Cas9 consiste em inserir ou deletar genes com elevada precisão por meio de um complexo formado pela nuclease Cas9, capaz de clivar o DNA, e de um gRNA (RNA guia), responsável por guiar a Cas9 até o local específico da modificação (GAGNON *et al.*, 2014). Embora a técnica requeira o complexo formado por uma proteína e um RNA, existem três formas principais para o complexo ser entregue ao organismo de interesse (MOUNT *et al.*, 2017b; LI *et al.*, 2018). A primeira é por meio da entrega de vetores de expressão contendo os genes que codificam para a proteína Cas9 e para o gRNA. A segunda forma é entregando o mRNA da proteína Cas9 juntamente com o gRNA. Finalmente, uma terceira forma, mais utilizada, ocorre através da entrega do complexo proteína-gRNA (WANG *et al.*, 2017). As três formas de entrega possibilitam resultados promissores na técnica de edição de genoma (RAN *et al.*, 2015; MADDALO *et al.*, 2014; GAGNON *et al.*, 2014;

RAMAKRISHNA *et al.*, 2014; SUN *et al.*, 2015; CHU *et al.*, 2016; HWANG *et al.*, 2013; TIMIM *et al.*, 2018). Entretanto, a entrega do complexo proteína-gRNA apresenta algumas vantagens em relação à entrega dos vetores de expressão e do mRNA, pois possibilita maior controle das concentrações do complexo; diminuição dos efeitos *off-target* da técnica devido à degradação da proteína; e não apresenta problemas relativos a possível integração do vetor de expressão ou gene ao genoma das células alvo (BLENKE *et al.*, 2016; GAJ *et al.*, 2012; SUN *et al.*, 2015). Uma das desvantagens é a dificuldade em entregar o complexo formado pela proteína e o gRNA (LI *et al.*, 2018).

Com o grande interesse do emprego da técnica CRISPR-Cas9 no tratamento de inúmeras doenças somado às vantagens do uso do complexo proteína-gRNA, a proteína Cas9 tem surgido como um potencial biofármaco, a ser produzido como uma proteína recombinante. Além disso, atualmente a proteína Cas9 vem sendo amplamente empregada como ferramenta de biologia molecular. Embora o uso da proteína Cas9 seja crescente, poucos estudos relatam o seu processo de produção. Na literatura é possível encontrar algumas metodologias para a expressão da proteína Cas9 em *E. coli* em escala de laboratório, com a finalidade somente de produzir a proteína para sua aplicação em estudos da técnica CRISPR-Cas9. As metodologias encontradas são distintas e empregam diferentes vetores de expressão (pET-28a-c, pEC-K-MBP), diferentes cepas (*Rosetta* (DE3), *Rosetta* (DE3) pLysS, BL21, BL21 STAR (DE3)) e diferentes condições de cultivo (meio autoindutor a 18°C por 24h, meio LB com antibióticos específicos e indução com 0,5 ou 1,0 mM de IPTG, a 18, 20 ou 30°C e tempo pós-indução de 8h ou 16h) (GAGNON *et al.*, 2014; RAMAKRISHNA *et al.*, 2014; SUN *et al.*, 2015; ZURIS *et al.*, 2015). Além disso, dados de produtividade ou níveis de expressão são raros.

A proteína Cas9 contém 1390 aminoácidos e massa molecular de 158 kDa. É uma proteína derivada da bactéria *Streptococcus pyogenes* e não apresenta glicosilação, sendo possível sua expressão em *E. coli* (JINEK *et al.*, 2014; RAMAKRISHNA *et al.*, 2014). No entanto, a proteína Cas9 é uma proteína grande e apresenta códons raros para expressão em *E. coli*. Essas características podem limitar a expressão da proteína pela *E. coli*, fazendo com que condições não ideais para produção em larga escala sejam utilizadas, como baixas temperaturas de expressão (18°C) e longos períodos de indução (16 h) (CHEN e TEXADA, 2006;

GAGNON *et al.*, 2014; LIU *et al.*, 2015; Yan *et al.*, 2017). Neste sentido, para tornar a produção da proteína recombinante mais eficiente, algumas estratégias podem ser empregadas. A avaliação do sistema de expressão e a otimização das condições de cultivo são de grande importância para o desenvolvimento do processo de produção de proteínas recombinantes. Fatores como composição do meio de cultura (meios quimicamente definidos), cultivos em alta densidade celular, temperatura, pH, oxigênio dissolvido e as condições de indução devem ser avaliados para a obtenção de um processo não apenas viável, mas eficiente (HUANG *et al.*, 2012; BERLEC e STRUKELJ, 2013). O estudo sistemático permite ainda a obtenção de parâmetros cinéticos importantes ao entendimento de fenômenos fisiológicos e necessários para o aumento de escala do processo.

Tendo em vista a potencialidade da proteína Cas9, a escassez de trabalhos publicados na literatura que avaliem a sua produção e os desafios para a expressão da proteína Cas9, é relevante a realização de um estudo sistemático que avalie os efeitos das condições de cultivo sobre a expressão da proteína Cas9 em *E. coli* recombinante, levando ao desenvolvimento de um processo escalonável de produção desta proteína.

1.1.OBJETIVO

Este trabalho teve como objetivo estudar o efeito de diferentes condições de cultivo sobre a expressão e produção da proteína Cas9 recombinante contendo uma cauda de histidina, empregando o vetor pET28b-Cas9His e as bactérias *E. coli* BL21(DE3) e BL21(DE3) Rosetta como sistemas de expressão, em frascos e biorreator.

Para atingir o objetivo proposto, foram realizadas as seguintes etapas:

- 1) Estudo da expressão da proteína Cas9 em *E. coli* BL21(DE3) e BL21(DE3) Rosetta;
- 2) Cinéticas de crescimento celular da *E. coli* BL21(DE3) em meios de cultivo complexo (Luria-Bertani) e definido (HDF) em frascos;
- 3) Estudo da expressão da proteína Cas9 pela *E. coli* BL21(DE3) em meios de cultivo complexo (LB) e definido (HDF) em frascos;
- 4) Avaliação das condições de cultivo (temperatura de indução e tempo de pós-indução) na expressão da proteína Cas9 pela *E. coli* BL21(DE3) em meio definido HDF e frascos;
- 5) Cinética de crescimento celular da *E. coli* BL21(DE3) e expressão da proteína Cas9 em biorreator utilizando meio de cultivo definido HDF;
- 6) Etapa de purificação;
- 7) Ensaio de atividade da proteína Cas9.

2. REVISÃO BIBLIOGRÁFICA

2.1. CRISPR-Cas9 e proteína Cas9

A técnica CRISPR-Cas9 foi primeiramente publicada em 2012 como uma nova ferramenta a ser utilizada na edição de genoma (JINEK *et al.*, 2012; DOUDNA e CHARPENTIER, 2014). Essa técnica vem sendo amplamente estudada em laboratórios de todo o mundo com as mais distintas aplicações, desde ensaios *in vitro*, a nível celular, até ensaios *in vivo*, como, por exemplo, no tratamento de doenças em animais (DOUDNA e CHARPENTIER, 2014). A técnica consiste em utilizar um complexo formado pela proteína Cas9 guiada por um gRNA (RNA guia) capaz de cortar e modificar o DNA com alta eficiência e precisão (DOUDNA e CHARPENTIER, 2014; SINGHT *et al.*, 2017).

Essa técnica foi desenvolvida ao se estudar e verificar a existência de um sistema adaptativo imune nas bactérias e arqueobactérias contra vírus e plasmídeos. O sistema de defesa conta com uma região chamada CRISPR, que é a sigla para *Clustered Regularly-Interspaced Short Palindromic Repeats*, ou seja, repetições palindrômicas curtas agrupadas e regularmente interespaçadas. Essa região apresenta pequenas sequências que, quando transcritas, reconhecem o DNA de vírus e plasmídeos invasores, como se fossem memórias de um sistema imune. Além dessas sequências, o sistema de defesa requer genes que codificam proteínas Cas. Portanto, o mecanismo de defesa consiste em armazenar sequências de DNA invasores que serão transcritas em gRNA, que por sua vez guiarão a proteína Cas até o DNA invasor (Figura 2.1) (DOUDNA e CHARPENTIER, 2014; SINGHT *et al.*, 2017).

Phase1: immunization

Phase 2: immunity

Figura 2.1. Sistema CRISPR-Cas9 nas bactérias. Fase1: armazenamento de seqüências de DNA invasores. Fase 2: mecanismo do sistema CRISPR-Cas9 contra um DNA invasor. (Fonte: Adaptado de MALI *et al.*, 2013a)

Dentre as proteínas Cas estudadas, a proteína Cas9 foi identificada na bactéria *Streptococcus pyogenes* e atualmente é a mais conhecida, devido à simplicidade do sistema em que está inserida. Esse sistema é conhecido como CRISPR-Cas tipo II, ou CRISPR-Cas9, e é constituído de um gRNA (RNA guia) e da proteína Cas9 para o reconhecimento e clivagem (corte) do DNA (Figura 2.2)

(DOUDNA e CHARPENTIER, 2014; SINGHT *et al.*, 2017). A proteína Cas9 é uma nuclease com dois domínios de clivagem do DNA (HNH e RuvC), os quais clivam o DNA na cadeia complementar e não-complementar à sequência do gRNA (JINEK *et al.*, 2014). A proteína Cas9 é uma endonuclease monomérica que contém em torno de 1400 aminoácidos e massa molecular de aproximadamente 160 kDa (HWANG *et al.*, 2013; JINEK *et al.*, 2014; YAN *et al.*, 2017). O gRNA é o resultado da junção de dois RNAs: o crRNA (*short guide* CRISPR RNA), responsável por carregar a sequência específica de reconhecimento do DNA (aproximadamente 20 nucleotídeos); e o tracrRNA (*trans-activating* CRISPR RNA), que tem a função de formar o complexo com a Cas9 (DOUDNA e CHARPENTIER, 2014; SINGHT *et al.*, 2017). Os sistemas CRISPR-Cas tipo I e III utilizam complexos de proteínas Cas, sendo uma grande desvantagem para a aplicação desses sistemas como ferramenta na edição de genoma. (JINEK *et al.*, 2012; DOUDNA e CHARPENTIER, 2014; SINGHT *et al.*, 2017).

Figura 2.2. A: Representação do complexo Cas9-gRNA-DNA. B: Representação esquemática do complexo gRNA-DNA. C: Representação esquemática do sistema CRISPR-Cas9. (Fonte: Adaptado de NISHIMASU *et al.*, 2014 e HU *et al.*, 2016)

Uma das grandes vantagens em utilizar o sistema CRISPR-Cas9 para edição de genoma em diferentes organismos deve-se ao fato da técnica requerer apenas a proteína Cas9 e um gRNA, o qual pode ser alterado de acordo com a sequência alvo desejada (crRNA) (GAGNON *et al.*, 2014).

O emprego da técnica CRISPR-Cas9 vem sendo demonstrado por diversos autores em diferentes tipos de células e organismos. Em 2012, Mali e colaboradores (2013b) foram os pioneiros na aplicação do sistema CRISPR-Cas9 em células eucarióticas (células humanas), demonstrando o potencial da técnica para a edição de genoma. Li e colaboradores (2015) e Weninger e colaboradores (2016) avaliaram a técnica de edição de genoma da *E. coli* e da *P. pastoris*, respectivamente. Liang e colaboradores (2017), Cho e colaboradores (2013), Kim e colaboradores (2016) e Mount e colaboradores (2017a) estudaram a manipulação genética de células de mamíferos (HEK-293T, células-tronco pluripotentes induzidas, HeLa, Raw 264).

Outros autores utilizaram o sistema CRISPR-Cas9 *in vivo* para avaliar seu potencial terapêutico. Gagnon e colaboradores (2014) demonstraram ser possível utilizar a técnica para editar o genoma de *zebrafish*. Ran e colaboradores (2015) empregaram a técnica para reduzir os níveis de colesterol no sangue de camundongos. Kennedy e colaboradores (2014) e Zhu e colaboradores (2015) estudaram a técnica para o tratamento de doenças virais como o HPV (Human Papiloma Virus) e o HIV-1 (Human Immunodeficiency Virus), respectivamente. Os estudos foram realizados *in vitro* em células humanas. Yang e colaboradores (2017) apresentaram resultados promissores no tratamento da doença neurodegenerativa *Huntington* ao aplicar a técnica CRISPR-Cas9 em ratos para inibir a expressão do gene mutante.

Ainda visando à aplicação da técnica CRISPR-Cas9 para o tratamento de doenças, Ma e colaboradores (2017) empregaram a técnica CRISPR-Cas9 para corrigir o genoma de embrião humano com mutação para problemas cardíacos. Niu e colaboradores (2017) empregaram a técnica para eliminar genes de vírus presentes no genoma de porcos, um grande avanço para tentar viabilizar o transplante de órgãos suínos em humanos. Em 2016, um grupo da China iniciou o primeiro teste clínico usando a técnica CRISPR-Cas9 para editar *ex vivo* células do sistema imunológico de pacientes com câncer de pulmão (CYRANOSKI, 2016).

Em 2018, Amoasii e colaboradores (2018) utilizaram a técnica CRISPR-Cas9 em cães para editar uma mutação no gene da proteína distrofina, responsável pela integridade e funcionalidade dos músculos. A distrofia muscular de Duchenne está relacionada a esta mutação. No estudo, os componentes do sistema CRISPR-Cas9 foram entregues às células *in vivo* empregando vetor viral Adeno-Associado (*adeno-associated virus serotype 9- AAV9*) e cassete para a expressão do gRNA e da proteína Cas9. Estudos preliminares indicaram que a síntese de 15% de proteína distrofina sem mutação seria suficiente como terapia. O presente estudo que utilizou o sistema CRISPR-Cas9 apresentou um aumento de aproximadamente 80% dos níveis de distrofina normal. Este estudo mostrou um grande potencial da técnica no tratamento da doença Duchenne em animais de grande porte.

Um dos grandes desafios da aplicação da técnica CRISPR-Cas9 na edição de genoma é a forma de entrega do sistema. Os componentes do sistema podem ser entregues para as células através de plasmídeos que carregam os genes da proteína Cas9 e do gRNA (*gene delivery*), do mRNA (RNA mensageiro) da proteína Cas9 com o gRNA sintético, ou pelo *delivery* do complexo formado pela proteína Cas9 e pelo gRNA (MOUNT *et al.*, 2017b). Cada método apresenta vantagens e desvantagens e muitos pesquisadores vêm estudando as diferentes formas de *delivery* para alcançar melhores resultados na aplicação da técnica CRISPR-Cas9.

Ran e colaboradores (2015) aplicaram a técnica CRISPR-Cas9 utilizando o vetor viral Adeno-Associado (*adeno-associated virus vector- AAV*) para entregar o material genético contendo os genes que expressam a proteína Cas9 e o gRNA e obtiveram resultado positivo na edição de genoma ao reduzir os níveis de colesterol no sangue em camundongos. Maddalo e colaboradores (2014) também empregaram vetor adenoviral para utilizar a edição de genoma no estudo de tratamento de câncer em camundongos. Embora os resultados de *gene delivery* para a técnica CRISPR-Cas9 sejam promissores, a entrega de DNA esbarra na desvantagem da possibilidade deste se integrar ao genoma, o que pode causar sérios problemas genéticos ao paciente (BLENKE *et al.*, 2016). Além disso, Gaj e colaboradores (2012) mostraram que a expressão prolongada desse sistema apresenta mais um efeito não desejado da técnica CRISPR-Cas9, conhecido como corte *off-target*.

Outra forma de aplicar a técnica CRISPR-Cas9 é entregar o mRNA da proteína Cas9 com o gRNA. Embora o uso do mRNA para a expressão da proteína

Cas9 não apresente o risco de inserção no genoma do organismo alvo, seu emprego apresenta valores moderados na eficiência na edição, efeitos *off-target* e imunogenicidade (MOUNT *et al.*, 2017b). Hwang e colaboradores (2013) utilizaram a técnica para fazer *knockout* de genes em embriões de *zebrafish* e obtiveram sucesso na edição de aproximadamente 50% dos organismos. Chu e colaboradores (2016) aplicaram a técnica CRISPR-Cas9 para fazer *knockin* de genes na região *Rosa26* (região utilizada para expressão de genes constitutivos ou condicionais) de embriões de camundongos. Os autores empregaram mRNA da proteína Cas9 e gRNA e atingiram um valor de 10-20% de indivíduos contendo a mutação. O valor de indivíduos com a edição genética aumentou para 50% ao empregar a proteína Cas9 junto com o mRNA da Cas9.

A entrega da proteína Cas9 com o gRNA sintético vem ganhando destaque devido às suas vantagens em relação a entrega de genes e plasmídeos. Este método possibilita maior controle das concentrações intracelulares do complexo proteína-gRNA e apresenta um menor tempo para ação uma vez que sua expressão não é necessária (SUN *et al.*, 2015). Um desafio para a entrega da proteína está relacionado ao tamanho da proteína Cas9 (158 kDa) (LI *et al.*, 2018). Para entregar o complexo proteína-gRNA os sistemas de *delivery* que vêm sendo estudados são os lipídeos catiônicos, peptídeos membrano ativos (*Cell-penetrating peptides*- CPP), microinjeção e eletroporação (ZURIS *et al.*, 2015; GAGNON *et al.*, 2014; RAMAKRISHNA *et al.*, 2014; SUN *et al.*, 2015; KIM S *et al.*, 2016).

Liang e colaboradores (2015) avaliaram as 3 diferentes formas de empregar o sistema CRISPR-Cas9 (plasmídeo, mRNA e proteína) empregando duas diferentes formas de *delivery* (lipossomos e eletroporação) em diferentes células de mamíferos (HEK293FT, U2OS, Mouse ESCs, Human ESCs (H9), Human iPSCs, Jurkat, K562, A549, Human Keratinocytes (NHEK), Human Cord Blood Cells CD34+). Os resultados mostraram maior eficiência da técnica CRISPR-Cas9 ao empregar os complexos formados pela proteína Cas9 e o gRNA. Além disso, os efeitos de corte *off-target* ao utilizar a proteína foram reduzidos quando comparados aos da transfecção do plasmídeo, o que pode ser devido a expressão prolongada da proteína ao empregar o plasmídeo. Os resultados também apontaram melhor eficiência ao empregar eletroporação para a transfecção dos componentes da

técnica; entretanto, foi observado a necessidade de otimizar a etapa de transfecção para cada tipo de célula.

Ramakrishna e colaboradores (2014) conjugaram a proteína Cas9 com um peptídeo membrano ativo (*Cell-penetrating peptide- CPP*) para avaliar o *delivery* e a técnica CRISPR-Cas9 na edição do genoma de diferentes células humanas (células tronco embrionárias, fibroblastos dérmicos, HEK293T, células HeLa e células de carcinoma embrionário). Os resultados foram eficientes e apresentaram efeitos *off-target* reduzidos quando comparados à transfecção do plasmídeo.

Wu e colaboradores (2018) demonstraram ser possível realizar *knockout* e *knockin* em células B humanas utilizando eletroporação da proteína Cas9 complexada ao gRNA.

Kim e colaboradores (2018) conjugaram a proteína Cas9 com uma sequência para localização nuclear e um peptídeo membrano-ativo (protamina). O complexo proteína Cas9-conjugada e gRNA foram avaliados *in vitro* (células A549 e H23) e *in vivo* (camundongos) para inativar a mutação do gene KRAS, observada em câncer de pulmão. Os resultados foram promissores e mostraram a possibilidade futura de utilizar o complexo proteína-gRNA como terapêutico.

Em vista dos resultados promissores em relação ao emprego da técnica CRISPR-Cas9 utilizando a proteína Cas9 e o gRNA sintético *in vitro* e *in vivo*, a proteína Cas9 tem potencial para ser empregada como ferramenta de biologia molecular e, futuramente, como um biofármaco.

Embora exista anseio pelo desenvolvimento da técnica de edição de genoma CRISPR-Cas9, a aplicação da técnica esbarra em problemas éticos e de regulamentação. Atualmente, as discussões estão relacionadas à quando e como a técnica deve ser aplicada em humanos (BROKOWSKI e ADLI, 2019; MEMI *et al.*, 2018). A técnica deve ser utilizada para correção de genes (*gene therapy*)? E para modificações genéticas para melhoramento da fisiologia (*genetic enhancement*)? Quais células podem ser modificadas? Células somáticas e germinativas? Quais são os riscos da técnica envolvidos para cada aplicação (efeitos *off-target*, relação da informação genética com o fenótipo biológico, futuras gerações de células serão afetadas) (AMOASII *et al.*, 2018; BROKOWSKI e ADLI, 2019; MEMI *et al.*, 2018)? Enquanto a técnica vem sendo amplamente estudada para minimizar os potenciais riscos e maximizar o benefício da tecnologia, é fundamental que questões sobre

ética, regulamentação e segurança da técnica sejam discutidas pelas organizações nacionais e internacionais. Recentemente, um grupo de cientistas europeus fundou a *Association for Responsible Research and Innovation in Genome Editing* (ARRIGE) para fornecer orientações sobre o uso ético da edição do genoma (BROKOWSKI e ADLI, 2019). No Brasil, a Comissão Técnica Nacional de Biossegurança (CTNBio) e o Comitê de Ética em Pesquisa e pela Comissão Nacional de Ética em Pesquisa já discutem as questões sobre o uso da técnica CRISPR-Cas9.

2.2. *E. coli* como sistema de expressão para proteínas recombinantes

A bactéria *Escherichia coli* é um organismo amplamente utilizado como sistema de expressão para proteínas recombinantes (SINGHA *et al.*, 2017). Em 1982, a *E. coli* foi utilizada pela primeira vez para produzir uma proteína recombinante de uso terapêutico, a insulina (TRIPATHI, 2016). Desde então, aproximadamente 30% dos biofármacos aprovados são produzidos em *E. coli*. (BERLEC e STRUKELJ, 2013). Na Tabela 2.1 estão apresentados alguns exemplos de proteínas terapêuticas aprovadas pelo FDA (*Food and Drug Administration*).

Tabela 2.1. Proteínas recombinantes produzidas por *E. coli* e aprovadas pelo FDA (Fonte: Adaptado de BERLEC e STRUKELJ, 2013; HUANG *et al.*, 2012; TRIPATHI, 2016).

Nome	Indicação	Empresa	Ano de aprovação
Insulina	Diabetes	Eli Lilly	1982
Interferon alpha e beta	Esclerose múltipla, hepatite B e C	Merck Sharp & Dohme Limited, Bayer, Novartis	1986 e 1995
Glucagon	Hipoglicemia	Novo Nordisk, Boehringer Ingelheim	1998
PEG Interferon alpha 2b	Esclerose múltipla, hepatite B e C	Merck Sharp & Dohme Limited	2001
Hormônio do crescimento humano	Deficiência de hormônio do crescimento; Síndrome de Turner	Genentech, Eli Lilly, Pfizer, Novo Nordisk	2008
Certolizumab pegol (Cimzia)	Doença de Crohn	Union Chimique Belge	2008

A bactéria *E. coli* é preferencialmente escolhida no desenvolvimento de processos para produção de proteínas recombinantes devido ao amplo conhecimento já descrito sobre a sua biologia celular e molecular (DEMAIN e VAISHNAV, 2009; TRIPATHI, 2016; SINGHA *et al.*, 2017). Além disso, as ferramentas para sua manipulação genética são eficientes (BERLEC e STRUKELJ, 2013; BAESHEN *et al.*, 2015). A bactéria também apresenta rápido crescimento celular em meios de cultura simples e de relativo baixo custo. Os processos de cultivo são relativamente simples e de fácil manipulação (BAYNEYX, 1999; DEMAIN e VAISHNAV, 2009). Por fim, em alguns casos a expressão da proteína de interesse ocorre em elevados níveis poucas horas após a indução da bactéria (DEMAIN e VAISHNAV, 2009).

Embora a *E. coli* apresente diversas vantagens para sua aplicação como sistema de expressão de proteínas recombinantes, a bactéria também apresenta algumas desvantagens. Uma das grandes limitações do sistema deve-se ao fato da bactéria não ser capaz de realizar naturalmente modificações pós- traducionais para a síntese de proteínas mais complexas (AGARWAL *et al.*, 2010; BERLEC e STRUKELJ, 2013). Além disso, a presença de códons raros para a expressão da proteína, a instabilidade do mRNA, a exigência de um gene de seleção, a formação de corpos de inclusão e a baixa eficiência em secretar a proteína também são fatores negativos do sistema *E. coli* (AGARWAL *et al.*, 2010; BERLEC e STRUKELJ, 2013).

Algumas estratégias podem ser utilizadas para tornar eficiente o processo de expressão de proteína recombinante utilizando *E. coli*. De acordo com as características da proteína de interesse é possível utilizar cepas modificadas geneticamente ou diferentes vetores de expressão (plasmídeo) (DEMAIN e VAISHNAV, 2009; BERLEC e STRUKELJ, 2013).

Algumas proteínas recombinantes podem ser tóxicas ao seu hospedeiro. Essas proteínas tendem a reduzir a velocidade específica de crescimento celular ou causar a morte celular, sendo um problema para a obtenção de elevadas taxas de expressão (ROSANO e CACCARELLI, 2014; DUMON-SEIGNOVERT *et al.*, 2004).

Normalmente, a expressão de proteínas recombinantes é realizada empregando a *E. coli* BL21(DE3) e o sistema pET como vetor de expressão. A cepa BL21(DE3) possui o gene para a expressão da enzima T7 RNA polimerase,

necessária para iniciar a transcrição dos genes presentes no sistema pET (GOPAL e KUMAR, 2013; SORENSEN e MORTENSEN, 2005; FRANCIS e PAGE, 2010). Embora a expressão da T7 RNA polimerase esteja sob o controle do promotor *lac*, sua expressão pode ocorrer a níveis mínimos (*basal level expression*) mesmo sem que ocorra a indução (GIACALONE *Et al.*, 2006; ROSANO e CACCARELLI, 2014). Ainda que em baixas concentrações, se a proteína for tóxica para a célula, sua expressão pode acarretar em problemas na produção (SAIDA *et al.*, 2006). Como alternativas para minimizar o “escape” da expressão da proteína tóxica diferentes cepas e vetores de expressão são utilizados com a finalidade de se obter maior controle na expressão gênica (ROSANO e CACCARELLI, 2014; SAIDA *et al.*, 2006).

Uma forma de evitar a expressão indesejada das proteínas empregando promotores T7 é o uso da cepa *E. coli* BL21(DE3) contendo os plasmídeos pLysS ou pLysE (ROSANO e CACCARELLI, 2014; SAIDA *et al.*, 2006). Estes plasmídeos possuem o gene para a expressão constitutiva da lisozima, enzima capaz de inibir a T7 RNA polimerase. O plasmídeo pLysE é capaz de expressar maior quantidade de lisozima quando comparado ao plasmídeo pLysS, o que pode contribuir com uma redução na velocidade específica de crescimento da *E. coli* (ROSANO e CACCARELLI, 2014; SAIDA *et al.*, 2006). A expressão da proteína de interesse nesses sistemas continua sendo através de IPTG ou lactose.

A cepa *E. coli* BL21-AI também é utilizada para a expressão de proteínas tóxicas sob o controle de promotores T7. A cepa contém o gene que codifica para a T7 RNA polimerase sob o controle do promotor *araBAD*, sendo sua indução por arabinose. Este promotor apresenta forte regulação da expressão gênica (ROSANO e CACCARELLI, 2014; SAIDA *et al.*, 2006).

Por fim, para minimizar os efeitos da toxicidade da proteína após a indução e obter maiores taxas de produção da proteína recombinante, é necessário que a expressão da proteína seja eficiente e rápida antes que as células morram (SAIDA, 2007). Para isso também podem ser utilizados sistemas com promotores fortes e de elevado controle da expressão gênica (SAIDA *et al.*, 2006; SAIDA, 2007).

A necessidade de empregar cepas geneticamente modificadas também pode estar associada à presença de códons raros na sequência do gene da proteína de interesse para a sua expressão num determinado organismo (QUAX *et al.*, 2015). Isto é, algumas proteínas apresentam em sua sequência gênica códons que

possuem baixas concentrações de tRNAs correspondentes no sistema hospedeiro, com isso a expressão da proteína pode ser prejudicada (CHEN e TEXADA, 2006). A presença de códons raros é uma forma de regular a expressão do gene (ZHOU *et al.*, 2016). Erros de tradução, baixos níveis de expressão da proteína e síntese de proteínas truncadas podem estar associados à presença de códons raros (AGUIRRE-LÓPEZ *et al.*, 2017). Os códons raros para a *E. coli* estão apresentados na Tabela 2.2.

Tabela 2.2. Códons raros para a *Escherichia coli* (Fonte: BURGESS-BROWN *et al.*, 2008).

Aminoácidos	Códons raros
Arginina	AGG, AGA, CGG, CGA
Glicina	GGA, GGG
Isoleucina	AUA
Leucina	CUA, CUC
Serina	UCG, UCA, AGU, UCC
Prolina	CCC, CCU, CCA
Treonina	ACA

Agarwal e colaboradores (2010) utilizaram três cepas diferentes de *E. coli* para produzir α 1-antitripsina (inibidor de proteases) e demonstraram que a cepa BL21CP (CodonPlus-RIL) apresentou maiores valores de crescimento e expressão quando comparados às cepas DH5 α e TB1. A cepa BL21CP apresenta tRNAs em abundância para arginina, isoleucina e leucina, favorecendo o aumento da taxa de expressão. Tegel e colaboradores (2010) avaliaram a expressão de fragmentos de proteínas recombinantes em *E. coli* BL21(DE3) e BL21(DE3) Rosetta. Os resultados mostraram um aumento na expressão das proteínas ao empregar a cepa Rosetta, cepa que apresenta maior concentração de tRNAs para os códons AGG, AGA, AUA, CUA, CCC e GGA, códons raros para a *E. coli* mas comuns no genoma humano. Aguirre-López e colaboradores (2017) avaliaram o efeito da presença dos códons raros AGA e AGG para arginina na expressão em *E. coli* da proteína triose-fosfato isomerase de diferentes espécies. Foram utilizadas duas cepas, a BL21(DE3) pLysS e a BL21(DE3) CodonPlus-RIL, sendo que a segunda apresenta tRNAs necessários em abundância. Os resultados indicaram que a quantidade, a posição e

combinações dos códons raros de arginina apresentam efeito sobre a expressão da proteína em *E. coli*. Além disso, a expressão da proteína obtida na cepa CodonPlus foi maior.

Além do uso de diferentes cepas, a otimização dos códons na sequência gênica também pode tornar possível a expressão das proteínas que apresentam códons raros (BERLEC e STRUKELJ, 2013). Burgess-Brown e colaboradores (2008) avaliaram a expressão de diferentes genes da desidrogenase humana de cadeia curta em *E. coli* BL21(DE3) e BL21(DE3) Rosetta2 (cepa que possui maior concentração de tRNAs para códons raros). Além disso, os autores sintetizaram os genes com códons otimizados. Os resultados indicaram que embora as proteínas apresentem códons raros em sua sequência, a expressão pode ser obtida através da otimização de códons ou pelo uso de cepa com maior concentração de tRNAs para códons raros. Entretanto, foi observada maior expressão das proteínas ao empregar os genes com códons otimizados.

Outra estratégia para otimizar a expressão de proteínas recombinantes é a escolha do vetor de expressão (plasmídeo) (DEMAIN e VAISHNAV, 2009; BERLEC e STRUKELJ, 2013). Para garantir elevados níveis de transcrição do gene da proteína recombinante, é recomendado o emprego de promotores (iniciadores da síntese de mRNA) fortes e que apresentem o mínimo de expressão basal (a expressão deve ocorrer apenas quando induzida) (SINGHA *et al.*, 2017). Kim J e colaboradores (2016) avaliaram a expressão de esterases (enzimas catalisadora de hidrólise de éster) empregando os vetores pEQ30 e pET28a, respectivamente. Kim J e colaboradores (2016) demonstraram que a expressão da enzima aumentou ao utilizar o vetor pET28a. Essa diferença deve-se ao fato de que o vetor pET28a possui o promotor T7, promotor que apresenta maior eficiência de transcrição quando comparado ao promotor T5 encontrado no vetor pEQ30. Tegel e colaboradores (2011) avaliaram três diferentes promotores (T7, *trc* e *lacUV5*) na expressão das proteínas em *E. coli* BL21(DE3). Os resultados mostraram que ao empregar um promotor forte (T7), os níveis de transcrição e tradução foram altos, resultando em elevada expressão. O promotor mais fraco *lacUV5* apresentou os menores valores para a expressão da proteína. Schlegel e colaboradores (2013) e Sonnendecker e colaboradores (2017) avaliaram a expressão de proteínas no

periplasma e no meio extracelular, respectivamente, o que pode favorecer os processos de purificação da proteína.

Por fim, as condições de cultivo, como composição do meio de cultura, cultivos em alta densidade celular, temperatura, pH, oxigênio dissolvido e as condições de indução também podem ser otimizadas para melhorar a eficiência do processo de produção de proteínas (HUANG *et al.*, 2012; BERLEC e STRUKELJ, 2013). É importante ressaltar que cada proteína recombinante requer um estudo individual de todo processo de expressão para obter uma produção eficiente (BERLEC e STRUKELJ, 2013).

2.2.1. *E. coli* BL21(DE3)

Nos processos de produção de proteínas recombinantes, as cepas utilizadas devem apresentar baixos níveis de proteases, manter uma expressão estável e conter elementos genéticos necessários para a expressão. A *E. coli* BL21 é amplamente empregada para esse fim, pois apresenta a deleção dos genes *Lon* e *OmpT* que codificam para proteases, diminuindo assim a possibilidade de degradação da proteína recombinante durante a expressão e durante o processo de purificação (BERLEC e STRUKELJ, 2013). A cepa BL21(DE3) possui em seu genoma o gene que codifica para a enzima T7 RNA polimerase, sob o controle de um promotor *lac*. Essa enzima proporciona à bactéria a capacidade de transcrever genes sob o controle de um promotor T7 (do fago T7), muitas vezes utilizado nos vetores de expressão de proteínas recombinantes e reconhecido pela alta eficiência de expressão (GOPAL e KUMAR, 2013).

2.2.2. *E. coli* BL21(DE3) Rosetta

A *E. coli* BL21(DE3) Rosetta é uma cepa derivada da BL21(DE3). A cepa Rosetta foi modificada para aumentar a expressão de proteínas que contenham em seu gene códons raros para *E. coli* (GOPAL e KUMAR, 2013). A cepa apresenta um plasmídeo (pRARE) para aumentar a concentração de tRNAs dos códons raros para arginina (AGG, AGA), isoleucina (AUA), leucina (CUA), prolina (CCC) e glicina

(GGA). O pRARE está sob seleção do antibiótico cloranfenicol e possui 4694 pb (Figura 2.3).

Figura 2.3. Plasmídeo pRARE contendo os genes para os tRNAs correspondentes aos códons raros em *E. coli* (indicado em azul). Apresenta gene de resistência ao cloranfenicol. (Fonte: adaptado de NOVY *et al.*, 2001).

Feng e colaboradores (2015) avaliaram a expressão da proteína humana fator de crescimento de mechano (MGF) em *E. coli* BL21(DE3) e BL21(DE3) Rosetta. Esta proteína apresenta em seu gene códons raros para expressão em *E. coli*. Os resultados mostraram que a presença dos códons raros limitou a expressão da proteína pela BL21(DE3). Contudo, devido à presença de tRNAs extras para códons raros na cepa Rosetta, a expressão da proteína foi possível nesta cepa.

Tegel e colaboradores (2010) avaliaram a expressão de fragmentos de proteínas humanas recombinantes em *E. coli* BL21(DE3) e BL21(DE3) Rosetta. Eles demonstraram que a cepa Rosetta aumentou a expressão do número de proteínas quando comparada à cepa BL21(DE3), principalmente devido aos tRNAs extras presentes na Rosetta. Este resultado é de grande importância em projetos *high-throughput*, que dependem de elevadas taxas de expressão de diferentes proteínas. Neste estudo foi possível aumentar de 63% para 77% a produção das proteínas envolvidas ao substituir a cepa BL21(DE3) pela Rosetta.

2.3. Condições de cultivo para produção de proteínas recombinantes

O estudo das condições de cultivo é fator importante no desenvolvimento de um processo de produção de proteínas recombinantes em biorreatores. Os biorreatores podem ser operados de diferentes formas, podem ser conduzidos em batelada, batelada alimentada e contínuo, este último com ou sem perfusão de células. Além disso, a composição do meio de cultura e os parâmetros do cultivo, como temperatura, oxigênio dissolvido, pH e condições de indução, devem ser otimizados para cada processo (HUANG *et al.*, 2012; BERLEC e STRUKELJ, 2013).

Com a finalidade de obter elevados níveis de proteína recombinante, geralmente faz-se necessário o cultivo em altas densidades celulares (HUANG *Et al.*, 2012). Uma das grandes limitações nesses processos, em especial para a produção de proteínas por *E. coli*, é o elevado acúmulo de ácido acético durante os cultivos (ZAWADA *et al.*, 2003). O acetato é um metabólito que leva à inibição do crescimento celular e redução da expressão de proteínas (TURNER *et al.*, 1994). Excesso de glicose, baixas concentrações de oxigênio dissolvido e altas velocidades específicas de crescimento são fatores que contribuem para o aumento da concentração de acetato no meio (YEE e BLANCH, 1992). Uma das principais estratégias estudadas para reduzir esse problema é o uso da batelada alimentada, com controle principalmente da fonte de carbono (CHOI *et al.*, 2006).

A seguir, serão discutidas algumas das principais condições de cultivo estudadas neste trabalho, com vistas a otimizar a produção da proteína Cas9 em *E. coli*.

2.3.1. Modo de Operação Batelada

Cultivos em batelada são relativamente simples e possibilitam atingir altas densidades celulares em um curto espaço de tempo (HUANG *et al.*, 2012; BERLEC e STRUKELJ, 2013). Os cultivos em batelada normalmente são realizados em biorreatores de tanque agitado ou airlift (BERLEC e STRUKELJ, 2013).

Neste modo de operação todos os nutrientes necessários para o crescimento celular estão disponíveis logo no início do cultivo (JONASSON *et al.*, 2002). O volume permanece constante durante todo o cultivo. As células crescem em fase

exponencial, na qual a fonte de carbono e o oxigênio dissolvido são consumidos rapidamente e pode ocorrer produção de acetato, no caso de *E. coli*. Após o crescimento, as células entram no estado estacionário, no qual o consumo de oxigênio é reduzido e ocorre limitação de nutrientes (CHOU, 2007). A expressão de proteínas recombinantes é preferencialmente realizada na fase exponencial de crescimento celular devido à elevada atividade celular e metabólica desta fase (CHOU, 2007).

Contudo, a produtividade em cultivo batelada é menor quando comparada ao cultivo em batelada alimentada. Portanto, a batelada é utilizada principalmente na produção de proteínas terapêuticas de baixa dosagem ou na produção de medicamentos órfãos (para doenças raras) (HUANG *et al.*, 2012). Além disso, é empregado para produzir proteínas recombinantes para uso em diagnósticos ou em estudos de novas vacinas e terapêuticos (TRIPATHI, 2016).

2.3.2. Meios de cultivo

Os meios de cultivo para o crescimento de *E. coli* e expressão de proteínas requerem componentes essenciais como fonte de carbono e nitrogênio, aminoácidos, micronutrientes e íons metálicos (OVERTON, 2014). Os meios podem ser complexos (composição química desconhecida) ou definidos (com composição química bem estabelecida) (HUANG *et al.*, 2012; OVERTON, 2014). Uma das desvantagens do meio complexo deve-se ao fato deste possuir materiais de composição desconhecida (extrato de levedura, hidrolisados de proteína: triptonas e peptonas), apresentando variação de lote para lote, podendo gerar variabilidade em todo o processo de produção de proteínas, não sendo usado para a produção de proteínas terapêuticas de uso humano (OVERTON, 2014). Os meios complexos são de simples preparo e ricos em nutrientes (OVERTON, 2014).

Os meios definidos, por sua vez, facilitam o controle da produção, garantem a reprodutibilidade do processo e podem simplificar a purificação da proteína de interesse (HUANG *et al.*, 2012; OVERTON, 2014).

É importante ressaltar que a concentração de glicose no meio de cultivo deve ser controlada para minimizar a produção de acetato. O excesso de glicose faz com que o fluxo de carbono exceda sua capacidade e desvie o metabolismo para as vias

fermentativas, produzindo acetato (*overflow metabolismo*) (LARA *et al.*, 2008). Outras fontes de carbono como o glicerol, podem ser utilizadas para reduzir a formação de acetato no meio (OVERTON, 2014).

Chen e colaboradores (2007) demonstraram melhoria no processo de produção da enzima Heparinase I ao otimizar o meio de cultura M9. Os autores avaliaram diferentes fontes de carbono (glicerol, maltose, lactose e glicose) e verificaram que a glicose e o glicerol foram capazes de aumentar o crescimento celular e a produção da enzima. Além disso, o meio foi otimizado com a adição de cálcio e de fontes de nitrogênio presentes em extrato de levedura.

Meyer e colaboradores (1984) constataram que a formação de acetato em meio complexo contendo glicose foi maior que em meio definido. A quantidade de nutrientes disponíveis no meio complexo faz com que a demanda biossintética seja excedida e as vias fermentativas do metabolismo da *E. coli* sejam utilizadas, produzindo acetato (SUÁREZ *et al.*, 1998).

Para aumentar a produtividade de um processo geralmente são utilizados cultivos em alta densidade celular (HUANG *et al.*, 2012). O meio definido é normalmente empregado para este tipo de cultivo, pois possibilita o controle de nutrientes que favorecem ou inibem o crescimento celular (CARVALHO *et al.*, 2012; GNOTH *et al.*, 2008).

No trabalho aqui proposto, foi utilizado o meio definido denominado HDF (*High Density Fermentation*) para obter maiores densidades celulares de *E. coli*. A composição do meio foi adaptada de Seeger e colaboradores (1995) e está descrita no item Materiais e Métodos (Tabela 3.2). O meio contém como principais componentes a glicose, o fosfato de potássio monobásico, o fosfato de amônio dibásico, o sulfato de magnésio e solução traço de metais. Este meio foi utilizado por Carvalho e colaboradores (2012) e Silva e colaboradores (2012) com a finalidade de desenvolver e melhorar os processos de produção da Proteína A de superfície de *Streptococcus pneumoniae* e das hidrolases xilanase e endoglucanase, respectivamente. Caspeta e colaboradores (2013) também empregaram o meio HDF para estudar e otimizar a expressão de proteínas heterólogas utilizando como indutor a temperatura. Os resultados destes estudos mostraram que ao empregar o meio definido HDF e batelada alimentada altas concentrações celulares foram atingidas e, por consequência, a produção de proteínas também foi elevada. Além

disso, o meio definido possibilita o desenvolvimento de um processo escalonável, garantindo a reprodutibilidade do processo e facilitando as etapas de purificação do produto.

2.3.3. Condições de indução

As condições de indução são parâmetros importantes para obter níveis satisfatórios na expressão de proteínas recombinantes. A indução geralmente ocorre no meio da fase exponencial do crescimento celular. Nessa fase de crescimento as células estão com sua capacidade catabólica máxima e podem fornecer níveis suficientes de energia e metabólitos para a síntese das proteínas (CARNEIRO *et al.*, 2013). A indução no final da fase exponencial ou na fase estacionária também pode favorecer maior produção da proteína recombinante devido à alta densidade celular obtida nesses instantes do cultivo. Entretanto, o metabolismo da célula e a presença de componentes no meio, como proteases, podem reduzir os níveis de proteínas produzidas (GALLOWAY *et al.*, 2003; KAUR *et al.*, 2018).

San-Miguel e colaboradores (2013) verificaram que a melhor condição para a expressão da proteína ativa e solúvel progesterona 5 β -redutase foi com a indução da *E. coli* no início da fase exponencial em baixa temperatura. Olaofe e colaboradores (2010) e Hu e colaboradores (2015) também verificaram que a expressão das proteínas amidase termoestável (EC 3.5.1.4) e *Pfu* DNA polimerase, respectivamente, ocorreu em maiores níveis ao induzir a bactéria no início da fase exponencial de crescimento.

Galloway e colaboradores (2003) verificaram que a indução da *E. coli* no final da fase exponencial aumentou o nível de proteínas recombinantes de alto peso molecular solúveis. Collins e colaboradores (2013) constataram maior produtividade na produção da proteína ao realizar a indução no início da fase estacionária de crescimento celular, principalmente devido à alta densidade celular neste instante do cultivo.

A temperatura de indução é outro importante parâmetro que deve ser avaliado nos processos de produção de proteína (DONOVAN *et al.*, 1996). Em baixas temperaturas o metabolismo celular é reduzido, reduzindo assim as taxas de transcrição e de divisão celular, reduzindo a formação de corpos de inclusão

(aumentando, muitas vezes, a solubilidade da proteína recombinante produzida) (SAHDEV *et al.*, 2007; PAPANEOPHYTOUS e KONTOPIDIS, 2014). Outra vantagem do emprego de temperaturas baixas é a redução na atividade das proteases, reduzindo a degradação das proteínas de interesse. Geralmente a temperatura de crescimento para a bactéria é de 37°C e a indução entre 15 e 25°C (SAHDEV *et al.*, 2007). Ao reduzir a temperatura e a síntese de proteína pode ser necessário um maior tempo de indução para alcançar níveis satisfatórios de expressão (FRANCIS e PAGE, 2010). Entretanto, uma expressão prolongada pode resultar em degradação da proteína de interesse ao longo do tempo (APSANA *et al.*, 2016).

Volontè e colaboradores (2008) avaliaram a redução da temperatura de indução na expressão da enzima *cephalosporin-C-acylase* (GLA-utilizada na síntese de derivados da Cefalosporina) e constataram que a presença de corpos de inclusão à 37°C era maior que em 25°C. Portanto, a redução na temperatura possibilitou o aumento da quantidade de enzima solúvel e ativa.

Outro fator importante na indução da proteína é o indutor e sua concentração no meio de cultivo (DONOVAN *et al.*, 1996). Alguns indutores, como o IPTG, são de elevado custo para o processo e em altas concentrações podem ser tóxicos, inibindo o crescimento celular e/ou a expressão de proteínas recombinantes (LECINA *et al.*, 2013; DONOVAN *et al.*, 1996). Uma alternativa na indução de cepas que apresentam o *operon lac* é a utilização de diferentes indutores, por exemplo, a lactose (MENZELLA e al., 2003; GOMBERT e KILIKIAN, 1998).

Portanto, os níveis de expressão de proteínas recombinantes dependem de cada sistema utilizado, isto é, das características da proteína de interesse, da cepa utilizada e das condições de cultivo empregadas (OLOAFE *et al.*, 2010).

2.3.4. Oxigênio dissolvido

O oxigênio dissolvido é um parâmetro de grande importância nos processos de produção de proteínas recombinantes. Sua importância está relacionada ao metabolismo e a fisiologia celular da *E. coli*, pois o oxigênio está envolvido na cadeia de fornecimento de energia para o metabolismo celular, favorecendo elevados valores de velocidade específica de crescimento celular e baixa produção de acetato

(KONZ *et al.*, 1998). Em baixas concentrações de oxigênio dissolvido o metabolismo anaeróbico da *E. coli* ocorre e, por consequência, há produção de acetato (inibidor de crescimento celular) (Figura 2.4) (AKESSON *et al.*, 2001; KONZ *et al.*, 1998).

Phue e Shiloach (2005) avaliaram o metabolismo celular da *E. coli* BL21 em diferentes concentrações de oxigênio dissolvido em meio contendo glicose. Ao empregar as concentrações de oxigênio dissolvido de 30, 6 e 1% os autores obtiveram os valores finais de DO_{600nm} de 65, 45 e 40, respectivamente. Além disso, as concentrações de acetato foram de 10 e 4 g/L ao empregar oxigênio dissolvido a 1 e 6%, respectivamente, e zero ao utilizar 30% de oxigênio dissolvido. Os resultados mostraram que em baixas concentrações de oxigênio, o acúmulo de acetato foi maior e o crescimento celular reduzido.

Figura 2.4. Representação do metabolismo aeróbico e anaeróbico (adaptado de KONZ *et al.*, 1998).

As dificuldades em manter este parâmetro acima do valor crítico devem-se à baixa solubilidade do oxigênio em água e a alta demanda do nutriente em cultivos em alta densidade celular (KONZ *et al.*, 1998). Além disso, a homogeneização do meio em cultivos em alta densidade celular se torna mais difícil possibilitando a presença de gradientes, o que pode favorecer baixa demanda de oxigênio para algumas células (PHUE e SHILOACH, 2005). Formas de aeração e agitação do meio de cultivo são estudadas para garantir e melhorar a transferência de oxigênio.

2.4. Extração da proteína recombinante: rompimento celular por ultrassom

A maioria das proteínas recombinantes é produzida no meio intracelular (HARRISON, 1991). Deste modo, é necessário realizar uma etapa de rompimento celular para extrair a proteína de interesse do interior da célula. O rompimento celular pode ser realizado por diferentes métodos: mecânicos (ultrassom), físicos (choque osmótico), químicos (lise alcalina) ou biológicos (uso da proteína lisozima) (HARRISON, 1991). Os métodos são utilizados de acordo com a espécie de micro-organismo empregada (GECIOVA *et al.*, 2002). Em escala laboratorial, a extração da proteína de interesse do meio intracelular da *E. coli* ocorre principalmente por ultrassom (GECIOVA *et al.*, 2002).

O método de rompimento celular por ultrassom, também conhecido por sonicação, emprega ondas sonoras numa frequência capaz de romper a membrana celular dos micro-organismos em suspensão (GECIOVA *et al.*, 2002). Embora este método seja bastante utilizado, durante o procedimento ocorre aquecimento da solução, o que pode causar degradação da proteína de interesse (FELIU e VILLAYERDE, 1994; GECIOVA *et al.*, 2002). Portanto, para que o método seja eficiente, é necessário um controle da temperatura e os tempos de sonicação e volumes utilizados devem ser otimizados (PETERNEL, 2013). Outra desvantagem do método é devido ao limitado volume de células para a lise, não sendo viável para maiores escalas (JOSHI e JAIN, 2017).

Após a etapa de rompimento celular, a proteína de interesse é recuperada no sobrenadante por centrifugação e segue para a etapa de purificação.

2.5. Purificação por cromatografia de afinidade por íons metálicos imobilizados (IMAC)

A purificação de proteínas recombinantes por cromatografia de afinidade por íons metálicos imobilizados (IMAC) consiste em utilizar metais (Cu^{2+} , Ni^{2+} , Co^{2+} , Zn^{2+}) como ligantes imobilizados que apresentam afinidade por resíduos de aminoácidos acessíveis na molécula de interesse (PORATH, 1988). A purificação muitas vezes ocorre em apenas uma etapa, com elevados níveis de pureza e recuperação (ROSANO e CACCARELLI, 2014).

Com o avanço das tecnologias para a obtenção de proteínas recombinantes, proteínas podem ser sintetizadas com caudas ou “fusões” na estrutura primária (*tags*) que podem incorporar novas características para a proteína de interesse (TERPE, 2003). A cauda de histidina é amplamente utilizada na expressão e purificação de proteínas recombinantes (His-tag) (ROSANO e CACCARELLI, 2014). Normalmente são adicionadas seis histidinas à proteína, seja na porção N- ou C-terminal (TERPE, 2003). A principal característica adquirida com a adição do His-tag é a relativa facilidade de purificação da proteína de interesse utilizando a técnica de IMAC. Os íons metálicos (Cu^{2+} , Ni^{2+} , Co^{2+} , Zn^{2+}) apresentam afinidade por resíduos de histidina (ROSANO e CACCARELLI, 2014).

Vançan e colaboradores (2002) avaliaram a purificação de imunoglobulina G (IgG) a partir de plasma humano empregando IMAC com os íons Cu^{2+} , Ni^{2+} , Co^{2+} e Zn^{2+} imobilizados. Os resultados indicaram a possibilidade de purificação da molécula, principalmente em relação à albumina presente no plasma. Todorova-Balvay e colaboradores (2004) demonstraram que a IgG apresenta aminoácidos de histidina em sua estrutura, proporcionando afinidade pelos metais Cu^{2+} , Ni^{2+} , Co^{2+} e Zn^{2+} .

Tallet e colaboradores (2001) sintetizaram a glicoproteína de superfície do envelope HTLV-1 contendo uma cauda de histidina para determinar suas propriedades bioquímicas e sua estrutura. Os resultados mostraram a possibilidade de obter a proteína para os estudos seguintes com alta pureza e em uma única etapa de purificação ao empregar IMAC.

A proteína de interesse do atual trabalho, a Cas9, contém uma cauda de histidina no C-terminal e, portanto, passível de purificação pela técnica de IMAC.

2.6. Produção da proteína Cas9: situação atual

Como apresentado anteriormente, a técnica CRISPR-Cas9 pode ser realizada ao introduzir em células ou organismos o complexo formado pela proteína Cas9 e o gRNA (MOUNT *et al.*, 2017b). Atualmente, a produção da proteína Cas9 é descrita na literatura com a finalidade de obter a proteína para estudos em escala laboratorial. Entretanto, verifica-se uma diversidade de protocolos para sua produção, empregando-se variadas condições de cultivo e indução, sem preocupação com a produtividade.

Ramakrishna e colaboradores (2014) utilizaram a cepa *E. coli* BL21 e o vetor pET28a para expressar a proteína Cas9. A expressão ocorreu com 0,5 mM de IPTG a 30°C *overnight*. A proteína foi conjugada a um CPP (*cell-penetrating peptide*) e o complexo Cas9-CPP-gRNA foi entregue às células humanas.

Yan e colaboradores (2017), Liu e colaboradores (2015) e Kim e colaboradores (2018) empregaram a cepa *E. coli* BL21(DE3). Yan e colaboradores (2017) utilizaram o meio LB e o vetor pET29 com as condições de indução de 1 mM de IPTG a 18°C por 16h. A proteína foi marcada com Iodo com a finalidade de desenvolver um novo método de detecção. Liu e colaboradores (2015) utilizaram meio LB e o vetor pET28 para expressar a proteína utilizando 0,1 mM de IPTG a 18°C e 12h. Os autores descreveram protocolos para implementar a técnica de edição de genoma utilizando a entrega de complexos formado pela proteína em células embrionárias e primárias. Kim e colaboradores (2018) utilizaram o plasmídeo pET-Cas9-NLS-6xHis e meio LB, induzindo a expressão da proteína com 0,2 mM de IPTG a 18°C *overnight*. Os autores empregaram a proteína Cas9 fusionada com protamina de baixo peso molecular (LMWP) para avaliar o *delivery* e a eficiência do complexo Cas9-LMWP-gRNA em aplicações *in vivo* no tratamento de câncer de pulmão.

Zuris e colaboradores (2015) realizaram a expressão da proteína Cas9 empregando a cepa BL21(DE3) STAR e o vetor pEC-K-MBP em meio LB e 0,5 mM de IPTG a 20°C por 16h. O estudo teve como objetivo avaliar o emprego de lipídios catiônicos para entregar a proteína Cas9 complexada ao gRNA *in vitro* e *in vivo*. Rajagopalan e colaboradores (2018) empregaram o vetor pET-NLS-Cas9-6His e a *E. coli* BL21(DE3) STAR pLysS para a expressão da proteína Cas9. O cultivo foi

realizado em 2L de meio LB e a indução ocorreu com a adição de 0,5 mM de IPTG a 18°C por 16h. Os autores avaliaram a purificação da proteína Cas9 utilizando IMAC (cromatografia com metal imobilizado) e resina de troca catiônica.

De Witt e colaboradores (2017) utilizaram a *E. coli Rosetta 2* e o vetor pML-2CT. Os autores descreveram métodos para entregar a proteína Cas9 contendo o gRNA em culturas de células humanas. Gagnon e colaboradores (2014) empregaram a cepa *E. coli Rosetta* e o vetor pET28b para expressar a proteína em meio autoindutor ZYP-5052 a 18°C e 24h. A proteína Cas9 foi utilizada para a edição de genoma em *zebrafish*. Montagna e colaboradores (2018) utilizaram o mesmo vetor e condições de cultivo que Gagnon e colaboradores (2014). Os autores estudaram a entrega do sistema CRISPR-Cas9 utilizando vesículas formadas por glicoproteínas do envelope do vírus da estomatite vesicular.

Sun e colaboradores (2015) expressaram a proteína em *E. coli Rosetta* BL21(DE3) pLysS e vetor pET28a em meio LB e 0,5 mM IPTG a 20°C por 8h. O estudo avaliou o emprego de *DNA nanoclews* para o *delivery* do complexo proteína-gRNA. Liang e colaboradores, C. (2017) empregaram o vetor pET28a-Cas9-His e a *E. coli Rosetta* para expressar a proteína a 16°C. O estudo avaliou a edição de genoma em trigo empregando complexos Cas9-gRNA. Mout e colaboradores (2017) empregaram o vetor pET28b-Cas9 e a cepa Rosetta. A expressão da proteína ocorreu em meio 2xYT com 0,75 mM de IPTG a 18°C por 16h. Neste estudo, variações da proteína Cas9 foram utilizadas para avaliar a edição de genoma em células a partir do sistema CRISPR-Cas9 complexado a nanopartículas de ouro. Wang e colaboradores (2018) empregaram o vetor pHis-parallel1-NLS_{H2B}Cas9 e a *E. coli Rosetta* para expressar a proteína Cas9. Foram utilizados 1,5 L de meio LB e a indução ocorreu com a adição de 0,3 mM IPTG a 25°C por 5h a 160 rpm. Os autores avaliaram a proteína Cas9 modificada e o emprego de PEG na edição de genoma em fungos patogênicos.

Não foi possível encontrar na literatura dados cinéticos e de produtividade sobre a produção da proteína Cas9, apenas dados sobre as condições em que a expressão foi realizada.

3. MATERIAIS E MÉTODOS

3.1. Materiais

3.1.1. Microrganismos

Para avaliar a expressão da proteína Cas9 foram utilizadas as cepas *Escherichia coli* BL21(DE3) e BL21(DE3) Rosetta. A *E. coli* BL21(DE3) não apresenta os genes *Lon* e *OmpT* que codificam para proteases e possuem o gene que codifica para a enzima T7 RNA polimerase, sob o controle de um promotor *lac*. A cepa BL21(DE3) Rosetta é derivada da primeira e contém um plasmídeo (pRARE) resistente a cloranfenicol, fornecendo maiores quantidades de tRNAs para os códons AGG, AGA, AUA, CUA, CCC, GGA.

3.1.2. Vetor de expressão

O plasmídeo contendo a sequência da proteína Cas9 com cauda de histidina foi adquirido do banco de plasmídeos Addgene, chamado de pET28b-Cas9His e depositado pelo Dr. Alex Schier (*Addgene plasmid #47327*). O vetor pET28b possui 5368 pb (Figura 3.1). Contém sequência que codifica para cauda de histidina (His-tag) no C-terminal da proteína Cas9 e marcador de resistência a kanamicina. O gene de expressão da proteína Cas9 possui 4100 pb.

Os vetores do sistema pET (produzidos pela Novagen) são amplamente utilizados para clonagem e expressão de proteínas recombinantes em *E. coli*. Os vetores pET possuem o promotor T7, promotor forte que permite um elevado nível de transcrição na presença da T7 RNA polimerase (SORENSEN e MORTENSEN, 2005; FRANCIS e PAGE, 2010).

Figura 3.1. Vetor de expressão pET28b contendo o gene da proteína Cas9 (Fonte: <https://www.addgene.org/47327/>).

3.1.3. *Template* de DNA

Os *templates* de DNA com promotor T7 para a transcrição em gRNA foram sintetizados pela empresa IDT (Integrated DNA Technology, USA). O sgRNA possui sequência para reconhecimento da sequência da GFP (*Green fluorescent protein*) e o cgRNA foi desenhado como controle sem especificidade para nenhum gene. As sequências estão apresentadas a seguir.

- sgRNA

5'-GTTTTTTTTTAATACGACTCACTATAGGGCGAGGAGCTGTTACACCGTTTT
AGAGCTAGAAATAGCAAGTTAAATAAGGCTAGTCCGTTATCAACTTGAAAA
AGTGGCACCGAGTCGGTGCTTTTTTTT-3'

- cgRNA

5'-TTTTTTTAAATACGACTCACTATAGGGTAACCGTGCGGTCGTACGTTTT
AGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAA
AGTGGCACCGAGTCGGTGCTTTTTTTT-3'

Legenda para as sequências de RNA: *itálico: promotor T7*; sublinhado: sequência de reconhecimento da GFP e sequência sem reconhecimento para controle; normal: região para interação do gRNA com a proteína Cas9.

3.1.4. Primers utilizados na qRT-PCR

Foram desenhados três *primers* para a realização da qRT-PCR. Um *primer* foi utilizado para a reação de transcrição reversa do mRNA da proteína Cas9 gerando um fragmento de 1632 bases (item 1 da Tabela 3.1). Outros dois *primers* foram desenhados para a síntese e amplificação por qPCR de um fragmento de 122 pb (itens 2 e 3 da Tabela 3.1). Todos os *primers* foram adquiridos da empresa Exxtend (Brasil).

Tabela 3.1. Sequência e Temperatura de *melting* (T_m) dos *primers* para a síntese de cDNA e amplificação de fragmentos do gene da proteína Cas9.

Primer	Sequência (5'→3')	T_m (°C)
(1) Reverse_cDNA	CTG CTC TCC AGA CAG GAA TGC T	59,5
(2) Forward_qPCR	GCG AGT GAA CAC GGA GAT CAC C	61,4
(3) Reverse_qPCR	ACT TCT CAG GCA GTT GCT GTC T	57,7

3.1.5. Meios de cultivo

Os meios utilizados foram os meios complexos LB (Luria Bertani), TB (Terrific Broth) e ZYP-5052, e o meio definido HDF (*High Density Fermentation*). O meio LB foi preparado em água destilada e continha 10 g/L de triptona, 5 g/L de extrato de levedura e 10 g/L de cloreto de sódio. O meio TB foi preparado em água destilada e

continha 12 g/L de triptona, 24 g/L de extrato de levedura e 4 mL/L de glicerol. O meio ZYP-5052 foi preparado de acordo com Studier (2005).

O meio HDF foi preparado em ultrapura (Milli-Q, Millipore, EUA) de acordo com a Tabela 3.2. O pH do meio foi ajustado para 7,0 adicionando solução de 25% de NH_4OH (v/v).

Tabela 3.2. Composição do meio HDF.

Componente	Composição
Glicose	10,0 g/L
KH_2PO_4	13,3 g/L
$(\text{NH}_4)_2\text{HPO}_4$	4,0 g/L
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	1,2 g/L
Ácido cítrico	1,7 g/L
EDTA	14,1 mg/L
Citrato de ferro III	100,8 mg/L
Tiamina	4,5 mg/L
$\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$	2,5 g/L
$\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$	15,0 g/L
$\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$	1,5 g/L
H_3BO_3	3,0 g/L
$\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$	2,1 g/L
$\text{Zn}(\text{CH}_3\text{COO})_2 \cdot 2\text{H}_2\text{O}$	33,8 g/L

Os meios de cultivo foram esterilizados em autoclave a 121°C por 20 min e em seguida suplementados com $30 \mu\text{g} \cdot \text{mL}^{-1}$ de Kanamicina previamente filtrada em membrana de $0,22 \mu\text{m}$ (Millipore, EUA). Para o cultivo da cepa BL21(DE3) Rosetta foram adicionados $34 \mu\text{g} \cdot \text{mL}^{-1}$ de cloranfenicol.

3.2. Métodos

3.2.1. Amplificação e purificação de plasmídeo

O plasmídeo pET28b-Cas9His foi amplificado em *E. coli* DH5 α e o plasmídeo pVAX1-GFP, em *E. coli* DH10b. Colônias foram selecionadas e inoculadas em 3 mL

de meio LB contendo $30 \mu\text{g.mL}^{-1}$ de Kanamicina. As células foram incubadas por 16 h, sob agitação de 250 rpm e temperatura de 37°C . Em seguida, o plasmídeo foi purificado utilizando o *kit* QIAprep Spin Miniprep (Qiagen, Alemanha), segundo recomendações do fabricante. A concentração de DNA foi determinada utilizando espectrofotômetro DS-11 (DeNovix, EUA) a 260 nm.

3.2.2. Competência das cepas *E. coli* BL21(DE3) e BL21(DE3) Rosetta

Bactérias *E. coli* BL21(DE3) e BL21(DE3) Rosetta foram adicionadas a 5 mL de meio LB (contendo os respectivos antibióticos) na proporção 0,1 % (v/v). As células foram incubadas por 16 h, sob agitação de 250 rpm e temperatura de 37°C . Após este período 1 mL de solução foi inoculado em 50 mL de meio LB contendo antibiótico e incubado a 37°C e 250 rpm até atingir a absorbância (DO600) de 0,5. A cultura foi resfriada em água gelada sob agitação. Em seguida a solução foi centrifugada por 8 minutos a $10000g$ e 4°C . O *pellet* foi ressuspensionado em solução 0,1 M MgCl_2 gelado e centrifugado. O *pellet* foi ressuspensionado em 0,1 M CaCl_2 gelado e centrifugado novamente. O novo *pellet* foi ressuspensionado em 0,1 M CaCl_2 contendo glicerol e congelado a -80°C .

3.2.3. Transformação das cepas *E. coli* BL21(DE3) e BL21(DE3) Rosetta

As bactérias competentes foram transformadas por choque térmico com o vetor de expressão pET28b-Cas9His. Trinta nanogramas do vetor foram adicionados a 100 μL de células competentes descongeladas em gelo. A solução foi incubada por 1 h em gelo. Em seguida, a solução foi submetida a choque térmico a 42°C por 45 segundos, voltando ao banho de gelo por mais 2 minutos. Posteriormente foram adicionados 0,9 mL de meio SOC-Broth e a solução foi incubada por 1 h a 37°C . As suspensões foram plaqueadas em meio LB-ágar contendo antibiótico.

3.2.4. Digestão dos plasmídeos com enzimas de restrição

As reações de digestão dos plasmídeos pET28b-Cas9His e pVAX1-GFP foram realizadas conforme informações do fabricante das enzimas de restrição.

Para a linearização do plasmídeo pET28b-Cas9His foram utilizados a enzima NcoI e o tampão Tango (Thermo Fisher Scientific, EUA). Para a verificação da presença do gene da Cas9 no plasmídeo foi realizada uma *double digestion*, utilizando as enzimas NcoI e NotI e o tampão Orange (Thermo Fisher Scientific, EUA). As reações foram conduzidas à 37°C por 4 horas. As enzimas foram inativadas à 65°C por 20 minutos.

Na digestão do plasmídeo pVAX1-GFP a enzima XbaI e o tampão Tango (Thermo Fisher Scientific, EUA) foram utilizados. As reações foram conduzidas à 37°C por 4 horas e a enzima foi inativada à 65°C por 20 minutos. O plasmídeo linearizado foi purificado utilizando o *kit* ilustra GFX PCR DNA and Gel Band Purification (GE Healthcare, Reino Unido), segundo recomendações do fabricante.

3.2.5. Transcrição *in vitro* de RNA

A transcrição *in vitro* de RNA dos *templates* de DNA foram realizadas utilizando o *kit* GeneArt Precision gRNA Synthesis (ThermoFisherScientific, USA).

Primeiramente foi realizada uma PCR (*Polimerase Chain Reaction*) convencional para amplificação dos *templates* de DNA (item 3.1.3). A reação de amplificação foi realizada no termociclador Applied Biosystems Veriti Thermal Cycler (Thermo Fisher Scientific, EUA) nas condições: (1) desnaturação inicial a 98°C por 10 segundos, (2) desnaturação a 98°C por 5 segundos, anelamento a 55°C por 15 segundos, por 32 ciclos; (3) extensão final a 72°C por 1 minuto. Em seguida, foi realizada a reação de transcrição *in vitro*. Ao produto da PCR foram adicionados dNTPs e RNA polimerase. A solução foi incubada a 37°C por 2 horas. Após a reação, DNase I foi adicionada à solução para a remoção dos *templates* de DNA. O gRNA gerado foi purificado com a coluna e tampões de acordo com o kit e quantificado por absorbância com o equipamento NanoDrop ND-1000 (Thermo Fisher Scientific Inc, USA) a 260nm.

3.2.6. Extração de RNA, tratamento com DNase e transcrição reversa

As cepas BL21(DE3) e BL21(DE3) Rosetta (com e sem o pET28b-Cas9His) foram cultivadas em 100 mL de meio LB e induzidas com IPTG 0,5 mM a 30°C por 4h.

A extração do RNA foi realizada com o PureLink™ RNA Mini kit (Invitrogen, EUA), de acordo com as instruções do fabricante. A concentração de RNA total foi determinada com o equipamento NanoDrop ND-1000 (Thermo Fisher Scientific Inc, USA) a 260nm.

Para eliminar qualquer possível contaminação por DNA, as amostras de RNA foram tratadas com DNase. Amostras contendo a mesma quantidade de RNA total foram incubadas com DNase I (Promega™, EUA) a 37°C por 50 min. A enzima foi inativada pela adição de EGTA a 65°C por 10 min.

Em seguida, a síntese do cDNA da proteína Cas9 foi realizada empregando a enzima M-MLV *Reverse Transcriptase* (Invitrogen, EUA). Dez µL de solução de RNA tratado com DNase, 5 pmol/µL do primer (1) (Tabela 3.1), 10 mM de dNTPs e água RNase-free foram incubados a 65°C por 5 minutos. Em seguida, a solução foi submetida a um choque térmico em gelo e foram adicionados tampão, 0,1 M de DTT e 200 U de M-MLV, para um volume final de 20 µL. A reação de síntese do cDNA ocorreu a 37°C por 50 minutos. A enzima foi inativada a 70°C por 15 minutos. As amostras foram posteriormente analisadas por qRT-PCR.

3.2.7. PCR (*Polimerase Chain Reaction*) convencional

A PCR convencional foi realizada para amplificar o fragmento de 122 pb do gene da proteína Cas9 empregando os primers 2 e 3 descritos na Tabela 3.1. As reações ocorreram no volume de 25 µL contendo: 1 ng de DNA, 0,1 pmol/µL de primer (2), 0,1 pmol/µL de primer (3), 0,2 mM dNTPs (Sinapse Inc., Brasil), Taq buffer 1x (Sinapse Inc., Brasil), 1U Taq DNA Polimerase (Sinapse Inc., Brasil) e água ultrapura para completar o volume da reação.

A reação de amplificação foi realizada no termociclador Applied Biosystems Veriti Thermal Cycler (Thermo Fisher Scientific, EUA) nas condições: (1) desnaturação inicial a 94°C por 3 minutos, (2) desnaturação a 94°C por 1 minuto, anelamento a 54°C por 1 minuto, extensão a 72°C por 40 segundos, por 30 ciclos; (3) extensão final a 72°C por 7 minutos. O resultado da PCR foi avaliado em gel de agarose 1%.

3.2.8. qRT-PCR (*Quantitative Reverse Transcription PCR*)

Após a síntese do cDNA da proteína Cas9, as amostras foram quantificadas por qRT-PCR para avaliar a expressão do gene da proteína Cas9 pelas diferentes cepas. Foi utilizado o Kit SYBR™ Green PCR Master Mix (Thermo Fisher Scientific Inc, EUA) que contém dNTPS, MgCl₂, TaqDNA polimerase, Syber Green e tampão. Foram utilizados 3 µL de amostra e um volume final de reação de 15 µL. As reações foram realizadas em triplicatas empregando o termociclador StepOnePlus™ Real-Time PCR System (Applied Biosystems™- Thermo Fisher Scientific Inc, EUA), disponível no Laboratório de Imunologia Viral do Instituto Butantan . As condições dos ciclos de amplificação foram: 95°C por 10 minutos, 40 ciclos de 95°C por 30 segundos, 54°C for 20 segundos e 60°C por 20 segundos. Após a amplificação, foi verificada a presença de interações inespecíficas, através da análise das curvas de *melting* (70 a 95°C).

A quantidade de cópias para cada amostra foi determinada a partir de uma curva padrão (C_T x número de cópias iniciais) utilizando o pET28b-Cas9His (número de cópias iniciais: $6 \cdot 10^7$, $6 \cdot 10^6$, $6 \cdot 10^5$, $6 \cdot 10^4$, $6 \cdot 10^3$, $6 \cdot 10^2$). Os valores relativos das amostras foram normalizados para o maior número de cópias obtido (amostra da cepa BL21(DE3) contendo o pET28b-Cas9His após a indução).

3.2.9. Cultivos em frascos

Os ensaios de cultivo em incubador rotativo foram realizados em frascos de Erlenmeyer de 250, 500 e 1000 mL e incubados em incubador rotativo C24 com agitação orbital (New Brunswick Scientific, EUA).

3.2.9.1. Preparo do inóculo

Os inóculos foram preparados a partir de células preservadas em -80°C. Bactérias *E. coli* BL21(DE3) e BL21(DE3) Rosetta contendo o plasmídeo pET28b-Cas9His foram adicionadas em 50 mL de meio de cultivo (LB ou HDF contendo os respectivos antibióticos) na proporção 0,1 % (v/v). As células foram incubadas por 16 h, sob agitação de 250 rpm e temperatura de 37°C.

Para os ensaios que utilizaram uma colônia de *E. coli* como inóculo, foi realizado plaqueamento da bactéria em meio LB-ágar contendo antibiótico e a placa foi incubada overnight à 37°C.

3.2.9.2. Cinética de crescimento da *E. coli* em frascos

A cinética de crescimento foi realizada em frascos de 500 mL contendo 100 mL de meio de cultivo (LB ou HDF contendo antibiótico) inoculados de forma que a densidade óptica (600 nm) inicial fosse de 0,2. As condições do crescimento foram de 37°C e 250 rpm. Amostras de 2 mL foram coletadas de hora em hora para determinação de pH, absorbância (DO) e concentração de glicose e acetato. Os experimentos foram realizados em triplicata.

3.2.9.3. Expressão da proteína Cas9

A Tabela 3.3 apresenta um resumo das condições de cultivo dos experimentos realizados com as duas cepas (BL21(DE3) e BL21(DE3) Rosetta) em frascos.

Tabela 3.3. Condições dos experimentos realizados para avaliar a expressão da proteína Cas9 em *E. coli* BL21(DE3) e BL21(DE3) Rosetta.

BL21(DE3) Rosetta	Volume (mL)	Meio de cultura	Indutor	Temperatura de indução (°C)	Tempo pós- indução (h)	# Ciclos de sonicação	Inóculo	Cepa*	PMSF
1	100	ZYP	Lactose (autoindutor)	27	16	16	0,1 % (v/v)	2017	sem
2	100	ZYP	Lactose (autoindutor)	18	24	16	0,1 % (v/v)	2017	sem
3	100	LB	0,5 mM IPTG	30	4	-	0,1 % (v/v)	2017	sem
4	100	LB	0,5 mM IPTG	30	4	-	Colônia	2018	sem
5	100	LB	0,5 mM IPTG	30	4	-	0,1 % (v/v)	2018	sem
6	100	LB	0,5 mM IPTG	18	4, 16	-	0,1 % (v/v)	2018	sem
7	850	ZYP	Lactose (autoindutor)	18	24	8	Colônia	2018	sem
8	500	LB	0,5 mM IPTG	30	4	8	0,1 % (v/v)	2018	com
9	500	TB	0,5 mM IPTG	18	20	8	0,1 % (v/v)	2018	com
10	500	ZYP	Lactose (autoindutor)	18	24	8	0,1 % (v/v)	2018	com
11	250	LB	0,5 mM IPTG	30	4	4	0,1 % (v/v)	2018	sem
12	250	LB	0,5 mM IPTG	30	4	8	0,1 % (v/v)	2018	sem

BL21(DE3)	Volume (mL)	Meio de cultura	Indutor	Temperatura de indução (°C)	Tempo pós- indução (h)	# Ciclos de sonicação	Inóculo	Cepa*	PMSF
1	100	LB	0,5 mM IPTG	25	4	16	0,1 % (v/v)	2017	sem
2	100	LB	0,5 mM IPTG	25	4	8	0,1 % (v/v)	2017	sem
3	100	LB	0,5 mM IPTG	25	4	8	0,1 % (v/v)	2017	com
4	100	HDF	0,5 mM IPTG	25	4	8	0,1 % (v/v)	2017	sem
5	100	HDF	0,5 mM IPTG	18, 25, 30, 37	4	8	0,1 % (v/v)	2017	sem
6	100	HDF	0,5 mM IPTG	30	2, 4, 6, 8, 18, 24	8	0,1 % (v/v)	2017	sem

A expressão da proteína Cas9 foi avaliada em diferentes meios de cultivo, os meios complexos LB, TB e ZYP-5052 e o meio definido HDF. O crescimento da bactéria *E. coli* BL21(DE3) foi realizado conforme descrito no item 3.2.9.2. Ao obter absorbância (600 nm) no valor entre 0,6-0,8 para os meios LB e TB e entre 1,2-1,4 para o meio HDF, a temperatura foi alterada para a temperatura desejada e 0,5 mM de *isopropyl β-D-1-thiogalactopyranoside* (IPTG) foi adicionado ao meio de cultivo. Após o tempo de pós-indução, 90 mg de células (massa de células equivalente a um volume de 80 mL de cultivo com DO igual a 2,0) foram centrifugadas a 8000g por 10 minutos e 4°C e o *pellet* de células foi armazenado a -20°C para posterior recuperação, purificação e análise da proteína.

Para o meio ZYP-5052, que é um meio autoindutor (lactose), o inóculo foi realizado a partir de uma colônia de *E. coli* e o crescimento celular ocorreu por 12h à 37°C. Em seguida a temperatura foi reduzida para 18°C e o tempo de indução foi de 24h.

Para avaliar diferentes condições de indução, a temperatura de indução e o tempo de pós-indução foram variados. Este estudo foi realizado para a cepa BL21(DE3) e meio HDF. O crescimento da *E. coli* foi realizado conforme descrito no item 3.2.9.2. A indução ocorreu quando a absorbância (600 nm) do meio de cultivo foi de 1,2-1,4. Primeiro a indução foi realizada com 0,5 mM de IPTG em diferentes temperaturas (18, 25, 30 e 37°C). Em seguida, foi realizado um ensaio na temperatura de 30°C e amostras foram coletadas em diferentes tempos de pós-indução (0, 2, 4, 6, 8, 18 e 24 h) e avaliadas quanto à expressão da proteína Cas9. A expressão foi avaliada utilizando SDS-PAGE e densitometria (ImageJ).

3.2.10. Cultivo em biorreator da cepa *E. coli* BL21(DE3)

Os cultivos em biorreator (dorna de 4,0 L) (New Brunswick Scientific/Eppendorf – BioFlo III, EUA) foram realizados em modo batelada, com volume de meio de cultivo de 2,0 L, com controle de temperatura, pH, concentração de oxigênio dissolvido e agitação.

3.2.10.1. Preparo do inóculo

O inóculo foi preparado a partir de células preservadas em -80°C. Bactérias *E. coli* BL21(DE3) contendo o plasmídeo pET28b-Cas9His foram adicionadas em 100 mL de meio de cultivo (HDF contendo 30 µg.mL⁻¹ de kanamicina) na proporção 0,1 % (v/v). As células foram incubadas por 16 h, sob agitação de 250 rpm e temperatura de 37°C.

3.2.10.2. Cinética de crescimento da *E. coli* BL21(DE3) em biorreator

A cinética de crescimento foi realizada em 2,0 L de meio de cultivo definido HDF (contendo 30 µg.mL⁻¹ de kanamicina) inoculados de forma que a densidade óptica (600 nm) inicial fosse de 0,2. As condições de crescimento foram de 37°C, pH 7,0 e 30% de oxigênio dissolvido da saturação com o ar. A agitação foi ajustada de modo automático entre os valores de 200 e 800 rpm a fim de manter o valor desejado de oxigênio dissolvido. A vazão de ar foi mantida a 2,0 L/min e durante o cultivo foi necessário misturar oxigênio puro ao gás de entrada (0,4 L/min de oxigênio e 1,6 L/min de ar). O pH foi mantido constante adicionando solução de NH₄OH 25%. Amostras foram coletadas de hora em hora para determinar concentração celular, de glicose e acetato.

3.2.10.3. Expressão da proteína Cas9 em biorreator

A indução para a expressão da proteína Cas9 em biorreator foi realizada adicionando 0,5 mM de IPTG no início e na metade da fase exponencial de crescimento celular. O tempo pós-indução foi de 8h e a temperatura de indução de 30°C. Durante o ensaio, amostras foram coletadas e a expressão da proteína Cas9 foi avaliada utilizando SDS-PAGE e densitometria (programa ImageJ).

3.2.11. Recuperação e purificação da proteína Cas9

Após o crescimento das cepas *E. coli* BL21(DE3) e BL21(DE3) Rosetta contendo o plasmídeo pET-28b-CasHis e expressão da proteína Cas9 conforme

descrito no item 3.2.9.3, os *pellets* de células foram ressuspensos em 20mL de tampão de lavagem (20 mM Tris-HCl pH8,0; 30mM Imidazol; 500 mM de NaCl), seguido de rompimento celular por sonicator em 8 ciclos de 30 segundos cada, intercalados por 30 segundos, em banho de gelo. Em seguida, a solução foi centrifugada a 12000g por 20 minutos a 17°C. O sobrenadante seguiu para purificação da proteína Cas9 por cromatografia de afinidade empregando a resina Ni-NTA-agarose (Qiagen, Alemanha). O tampão de lavagem foi empregado como tampão de adsorção. A eluição ocorreu em tampão de lavagem contendo 500 mM de Imidazol. A proteína foi dialisada em tampão de diálise (50 mM NaH₂PO₄.H₂O pH 7,4, 1 M NaCl, 1 mM MgCl₂.6H₂O, 10% glicerol, 1 mM DTT) empregando ultrafiltração em membrana de celulose de 100 kDa de diâmetro de corte (Millipore, EUA). A concentração de proteína total foi determinada pelo método de Bradford e as amostras foram analisadas por SDS-PAGE e densitometria.

3.3. Métodos analíticos

3.3.1. Determinação da concentração celular

O crescimento celular foi monitorado através de medida de absorbância em espectrofotômetro (Quimis, Brasil) no comprimento de onda de 600 nm (DO600). Para determinar a concentração celular, foi obtida uma curva de massa seca (Apêndice A) e a equação obtida da regressão linear na fase exponencial do crescimento celular foi utilizada para conversão de DO em concentração celular (g.L⁻¹).

Para determinar a massa seca, alíquotas de 10 mL foram filtradas em membrana MF-Millipore 0,22 µm (Millipore, EUA) previamente pesadas. As membranas contendo biomassa foram secas em microondas por 15 min em 450 W de potência. Em seguida, a membrana com a biomassa foi novamente pesada e a massa de células secas foi determinada pela diferença entre a membrana com a biomassa e a membrana livre. A concentração celular em massa seca foi determinada dividindo a massa de células secas obtida pelo volume de filtrado.

3.3.2. Dosagem de proteínas totais

As proteínas totais e a proteína Cas9 foram quantificadas de acordo com o método proposto por Bradford (1976), utilizando albumina de soro bovino (BSA) como proteína de referência.

3.3.3. Dosagem de glicose e acetato

A quantificação de glicose foi realizada por método enzimático empregando o equipamento YSI 2700 (Yellow Spring Instrument, Marshall Scientific, EUA), disponível no Laboratório de Imunologia Viral do Instituto Butantan.

A dosagem de acetato foi realizada por cromatografia líquida de alta eficiência (HPLC) em equipamento Dionex® (Ultimate 3000, Thermo Fisher Scientific Inc, USA) utilizando coluna Aminex® HPX-87H (Biorad, EUA). A análise foi realizada no Laboratório de Bioprodutos do departamento de Microbiologia da Universidade de São Paulo.

3.3.4. Espectrometria de massa- MALDI-TOF MS e MS/MS

A espectrometria de massa foi realizada no Laboratório de Espectrometria de Massa Dempster do departamento de Engenharia Química da Poli-USP. Foi utilizado o equipamento espectrômetro de massa UltrafleXtreme MALDI-TOF (Bruker Daltonics). O resultado foi avaliado pelo *software* MASCOT.

3.3.5. Eletroforese em gel de agarose

Para confirmar a presença do plasmídeo pET-28b-Cas9His e do gene de interesse foi realizada eletroforese em gel de agarose 1% corado com SYBR safe (ThermoFisherScientific, EUA). O plasmídeo obtido foi digerido com as enzimas de restrição conforme item 3.2.4. A agarose foi dissolvida em tampão TAE (Tris-acetato-EDTA). A eletroforese ocorreu a 80 V.

3.3.6. SDS-PAGE

A proteína Cas9 foi analisada por eletroforese em gel de poliacrilamida, sob condições desnaturantes e não redutoras (SDS-PAGE). A eletroforese foi realizada em equipamento Mini Protean III (Biorad, EUA) com gel de poliacrilamida na concentração de 12%, como descrito em Laemmli (1970). As amostras das etapas de purificação foram preparadas com tampão contendo SDS e β -mercaptoetanol e desnaturadas por aquecimento a 100°C por 7 min. Para avaliar a expressão a partir de *pellet* celular, 0,7 mg de células (massa seca) foram tratadas com 20 μ L de tampão e aquecidas a 100°C por 10 min. Foram injetados 15 μ L de amostra no gel. Foi utilizado marcador de alto peso molecular *Precision Plus Protein standard* (Biorad, EUA). O gel foi submetido à voltagem de 150 V em cuba vertical. A coloração foi realizada por Coomassie blue.

3.3.7. Western blot

O ensaio imunoenzimático *Western blot* foi realizado de acordo com Towbin e colaboradores (1979). Primeiramente, foi realizada uma SDS-PAGE com gel 10% para separar as proteínas. Em seguida, foi realizada a transferência das proteínas para uma membrana de nitrocelulose 0,2 μ m (BioRad, EUA) por 1h30 à 0,1A. Após a transferência, a membrana foi incubada overnight à 4°C em tampão PBS (0,02 M) contendo 3% de leite em pó desnatado (Molico- Nestlé Brasil) e Tween-20 (0,05%) (etapa de bloqueio). Após o bloqueio, a membrana foi incubada por 1 h a temperatura ambiente no anticorpo primário anti-His-Tag (mouse monoclonal anti-His-Tag antibody) (#MA1-21315, Thermo Fisher Scientific Inc, USA) diluído 1:4000 em tampão PBS contendo 3% de leite em pó desnatado. Em seguida, a membrana foi lavada três vezes a cada 5 minutos com solução de lavagem (tampão PBS 0,02 M contendo 0,1% de Tween-20). Após a lavagem a membrana foi incubada por 1 h a temperatura ambiente no anticorpo secundário anti-IgG conjugado com peroxidase (anti-mouse IgG- peroxidase antibody produced in goat) (A4416, Sigma-Aldrich, USA) diluído 1:4000 em tampão PBS contendo 3% de leite em pó desnatado. Em seguida, foi realizada mais uma etapa de lavagem. As bandas foram reveladas por quimiluminescência empregando o kit Amersham ECL Select Western Blotting

Detection Reagent (GE Healthcare, USA). Foi utilizado o marcador Prestained Protein MW Marker (Thermo Scientific, USA). O ensaio de *Western blot* foi realizado no Laboratório de Desenvolvimento de Vacinas do Instituto de Ciências Biomédicas II da Universidade de São Paulo.

3.3.8. Densitometria em gel de poliacrilamida (SDS-PAGE)

Os valores da produção da proteína Cas9 por litro de cultivo (itens 3.2.9.3 e 3.2.10.3) foram estimados através de densitometria de gel de eletroforese utilizando o programa ImageJ (U. S. National Institutes of Health, USA) conforme metodologia descrita por Dagar e Khasa (2017) e Larentis e colaboradores (2011). O valor da área de cada banda de proteína foi determinado ao comparar com o valor obtido para a massa conhecida de proteína presente no padrão de eletroforese utilizado (*Precision Plus Protein standard* (Biorad, EUA)). Como base de cálculos foi utilizado o valor da área calculada pelo software ImageJ para a banda de 50 kDa presente no padrão de eletroforese. 10 µL de padrão foram aplicados ao gel, portanto, segundo fornecedor, 750 ng da proteína de massa 50 kDa. A densitometria foi determinada 3 vezes para cada banda.

3.3.9. Atividade da proteína Cas9

A reação para verificar a atividade da proteína ocorreu em um volume de 20 µL (em NEBuffer 3). Proteína Cas9 purificada (330 ng) foi pré-incubada com o gRNA (150 ng) por 10 minutos a 25°C. Em seguida, o plasmídeo pVAX1-GFP linear (150 ng) foi adicionado a solução. A reação foi finalizada com aquecimento. Após a digestão por 90 minutos a 37°C, o plasmídeo pVAX1-GFP foi analisado em gel de agarose 1%.

3.3.10. Cálculo dos parâmetros cinéticos

A velocidade específica de crescimento (μ) foi determinada utilizando os dados obtidos na fase exponencial dos ensaios de cinética de crescimento. O μ foi determinado através da Equação 3.1.

$$\mu = \frac{1}{X} \frac{dX}{dt} \quad (3.1)$$

Em que X é a concentração de células (g/L).

O fator de conversão de substrato em biomassa foi determinado pela Equação 3.2.

$$Y_{x/s} = \frac{X - X_0}{S_0 - S} \quad (3.2)$$

Em que X é a concentração de células (g/L), X_0 é a concentração inicial de células (g/L), S é a concentração de glicose final (g/L) e S_0 é a concentração inicial de glicose (g/L).

A produtividade do produto (P) foi determinada pela Equação 3.3.

$$P = \frac{P_f}{t_f} \quad (3.3)$$

Em que P_f é a concentração final de proteína obtida (mg/L) e t_f é o tempo total do cultivo (h).

4. RESULTADOS E DISCUSSÕES

4.1. Análise do vetor de expressão pET28b-Cas9His

O vetor pET28b contendo o gene de interesse para a expressão da proteína Cas9 foi adquirido junto ao banco de plasmídeos Addgene (EUA). O vetor recebido foi propagado em *E. coli* DH5 α (linhagem de clonagem) e purificado utilizando *kit* de minipreparação plasmidial. Em seguida, células da linhagem de expressão *E. coli* BL21(DE3) e BL21(DE3) Rosetta foram transformadas com o plasmídeo e o banco de células foi preparado. Para confirmar a presença do gene da proteína Cas9 no vetor pET28b-Cas9His, o plasmídeo foi novamente purificado e submetido ao ensaio de digestão utilizando enzimas de restrição. As enzimas de restrição utilizadas para a linearização e separação do gene de interesse foram a NcoI e a NotI. A verificação foi feita por eletroforese em gel de agarose 1% e os resultados obtidos estão apresentados nas Figuras 4.1 e 4.2.

A Figura 4.1 apresenta o resultado para o pET28b-Cas9His presente na cepa BL21(DE3). Foi possível notar uma banda de tamanho 9468 pb na amostra 2, referente ao plasmídeo linearizado, e duas bandas na amostra 3 de tamanhos 5368 pb e 4100 pb, que correspondem ao vetor pET28b e ao gene da proteína Cas9, respectivamente.

A Figura 4.2 apresenta os plasmídeos pET28b-Cas9His e pRARE (~4700 pb) constituintes da cepa BL21(DE3) Rosetta (amostra 3). Como a Rosetta apresenta um plasmídeo extra, foi necessário conhecer o resultado da digestão deste plasmídeo pelas enzimas NcoI e NotI. A amostra 4 indicou que o pRARE foi digerido pelas enzimas e produziu dois fragmentos. Em seguida, o pET28b-Cas9His e o pRARE purificados da cepa foram submetidos a digestão enzimática pelas enzimas e através da amostra 5 foi possível constatar a presença do vetor pET28b (5368 pb), do gene da proteína Cas9 (4100 pb) e dos dois fragmentos gerados pela digestão do pRARE. Os resultados confirmaram a presença do plasmídeo pET28b-Cas9His nas cepas BL21(DE3) e BL21(de3) Rosetta.

Figura 4.1. Análise do vetor pET28b-Cas9His purificado de BL21(DE3) em eletroforese de gel de agarose 1%. Ladder- DNA Ladder 1kb RTU (Kasvi, Brasil); 1- Vetor pET28b-Cas9His não digerido; 2- Vetor pET28b-Cas9His digerido com a enzima NcoI; 3- Vetor pET28b-Cas9His digerido com as enzimas NcoI e NotI.

Figura 4.2. Análise do vetor pET28b-Cas9His purificado de BL21(DE3) Rosetta em eletroforese de gel de agarose 1%. Ladder- DNA Ladder 1kb RTU (Kasvi, Brasil); 1- Vetor pET28b-Cas9His digerido com a enzima NcoI; 2- pRARE digerido com a enzima NotI; 3- Vetor pET28b-Cas9His e pRARE digeridos com a enzima NotI; 4- pRARE digerido com as enzimas NotI e NcoI; 5- Vetor pET28b-Cas9His e pRARE digeridos com as enzimas NotI e NcoI.

4.2. Verificação da expressão da proteína Cas9

Para verificar a expressão da proteína Cas9, a bactéria *E. coli* BL21(DE3) contendo o plasmídeo pET28b-Cas9His foi cultivada em meio LB contendo 30 $\mu\text{g.mL}^{-1}$ de kanamicina e induzida com 0,5 mM de IPTG a 25°C. Posteriormente, a cultura foi submetida a rompimento celular por sonicação e o lisado celular foi purificado por cromatografia de afinidade em coluna contendo resina de Ni-agarose. As amostras do processo foram analisadas por eletroforese (SDS-PAGE). Para a análise é importante saber que a proteína Cas9 é uma proteína monomérica e apresenta massa molecular de aproximadamente 160 kDa (JINEK *et al.*, 2014).

Na etapa de rompimento celular foram avaliados dois ciclos de rompimento: 8 ou 16 ciclos de 30 segundos cada, intervalados em 30 segundos, em banho de gelo. Os resultados da expressão e purificação da proteína Cas9 estão apresentados na Figura 4.3. Nota-se que ao empregar 16 ciclos para o rompimento celular não foi possível visualizar a banda da proteína na faixa de 150 kDa nas amostras do gel de eletroforese, enquanto que ao utilizar 8 ciclos a banda da proteína foi visível. Os resultados indicaram que o número elevado de ciclos (16) promoveu a degradação da proteína durante o rompimento celular, o que indica a suscetibilidade de degradação da proteína por ultrassom. A partir desse resultado, foi estabelecido que o rompimento celular fosse realizado com 8 ciclos de 30 segundos cada.

Figura 4.3. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) em meio LB (indução: 0,5 mM de IPTG à 25°C por 4h) para avaliar a expressão da proteína Cas9. Amostras dos *pellets* e das etapas de purificação da proteína empregando resina de Ni-agarose. Revelação por *Et al.*. (a) Rompimento celular por 8 ciclos em ultrassom. (b) Rompimento celular por 16 ciclos em ultrassom. M- Marcador de massa molecular; 0h- *Pellet* com 0h de indução; Ins- Fração insolúvel após a sonicação; I- Amostra da injeção da coluna de purificação (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E- Eluição.

Após a análise da etapa de rompimento celular por sonicator, a adição de PMSF ao tampão de lavagem (e ressuspensão do *pellet*) também foi avaliada. O PMSF (*Phenylmethylsulfonyl fluoride*) é um inibidor de proteases utilizado no preparo de lisados celulares. Para este estudo, o crescimento celular foi realizado em dois frascos de 500 mL contendo 100 mL de meio de cultivo LB em cada um. A etapa de indução ocorreu utilizando-se 0,5 mM de IPTG a 25°C por 4h. Após a indução, foram adicionados 1 mM de PMSF ao tampão de ressuspensão do *pellet* de um dos frascos. Seguiu-se com a purificação da proteína Cas9 e amostras da etapa de eluição para cada condição (com e sem adição de PMSF) foram avaliadas por SDS-PAGE e densitometria (para as bandas presentes na faixa de 150 kDa) (Figura 4.4; Apêndice B). Os resultados indicaram que na ausência de PMSF a amostra continha aproximadamente 2,3 µg de proteína, enquanto que com PMSF, 2,0 µg. Como os valores de proteína foram bastante semelhantes nas duas condições empregadas e a adição de qualquer componente ao processo pode ser fator crítico para o escalonamento do processo e para as etapas de purificação, decidiu-se não utilizar o PMSF para a recuperação da proteína Cas9.

Figura 4.4. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) em meio LB (indução: 0,5 mM de IPTG à 25°C por 4h) para avaliar a expressão da proteína Cas9 e o emprego de PMSF. Revelação por *Et al.*. M- marcador de massa molecular; Is- Injeção (fração solúvel após a sonicação) sem PMSF; Es- Eluição sem PMSF; Ic- Injeção (fração solúvel após a sonicação) com PMSF; Es- Eluição com PMSF.

Ao avaliar os resultados apresentados nas Figura 4.3 e Figura 4.4, as bandas visíveis na altura de aproximadamente 150 kDa nas amostras de lisado celular e da etapa de eluição da purificação por cromatografia de afinidade em Ni-agarose evidenciaram a expressão da proteína Cas9 na forma solúvel. Para confirmar que a proteína presente na etapa de eluição era realmente a nuclease Cas9, uma amostra desta etapa foi digerida empregando-se tripsina, e então avaliada por espectrometria de massa (MALDI-TOF). Os resultados confirmaram a presença da proteína Cas9, com massa molecular de 158 kDa (Apêndice C).

4.3. Proteína Cas9 e expressão em *E. coli* BL21(DE3) Rosetta

Após verificar que o plasmídeo pET28b-Cas9His apresentava o gene de interesse e constatar a expressão da proteína Cas9, iniciou-se o trabalho para otimização da expressão da proteína em *E. coli*. Alguns fatores podem ser levados em consideração para tentar aumentar o nível de expressão de proteínas recombinantes em *E. coli*, sendo eles o emprego de diferentes vetores de expressão (plasmídeo), sistemas de expressão (diferentes cepas) e condições de cultivo (BERLEC e STRUKELJ, 2013). Nesta etapa do trabalho foi avaliada a expressão da proteína Cas9 utilizando uma cepa modificada geneticamente, a BL21(DE3) Rosetta. A escolha da cepa a ser utilizada deve ser baseada nas características da proteína de interesse, neste caso, a Cas9 (TRIPATHI, 2016).

A proteína recombinante Cas9 é uma proteína considerada grande com ~1400 aminoácidos e massa molecular de aproximadamente 160 kDa (JINEK *et al.*, 2014). É uma endonuclease monomérica derivada da bactéria *Streptococcus pyogenes* (HWANG *et al.*, 2013). Além disso, a sequência do gene que codifica a proteína contém códons raros para sua expressão em *E. coli* (Apêndice D; Tabela 4.1). Levando em consideração que a presença de códons raros pode reduzir a expressão da proteína recombinante em *E. coli*, a cepa BL21(DE3) Rosetta foi escolhida como estratégia inicial para expressar a proteína Cas9. A bactéria BL21(DE3) Rosetta contém um plasmídeo extra que codifica para tRNAs raros (AGG, AGA, AUA, CUA, CCC, GGA), aumentando a concentração destes tRNAs favorecendo uma melhor tradução da proteína de interesse (FENG *et al.*, 2015).

Tabela 4.1. Códonos raros para *E. coli* presentes na sequência do gene da proteína Cas9 (número de códonos na sequência: 1390).
(<http://people.mbi.ucla.edu/sumchan/caltor.html>)

Aminoácidos	Códonos raros	Total de códonos na sequência da Cas9
Arginina	CGA	9
	CGG	20
	AGG	21
	AGA	15
Glicina	GGA	23
	GGG	15
Isoleucina	AUA	8
Leucina	CUA	0
Prolina	CCC	18
Treonina	ACG	8

Inicialmente foi realizado um experimento para verificar a presença de proteínas nativas da BL21(DE3) Rosetta que pudessem interferir na verificação da proteína Cas9. Para isso, a cepa BL21(DE3) Rosetta sem conter o pET28b-Cas9His foi cultivada em 100 mL de meio LB e induzida com 0,5 mM de IPTG à 30°C por 4h. Em seguida, o *pellet* celular foi rompido por 8 ciclos de sonicação e centrifugado. O sobrenadante foi aplicado na resina de Ni-agarose para verificar se alguma proteína além da Cas9 seria adsorvida pela resina. Os resultados apresentados na Figura 4.5 indicaram que a bactéria expressou proteína na banda de 150 kDa, mesma posição de verificação da proteína Cas9. Entretanto, após a etapa de purificação, foi possível verificar que a proteína deste tamanho não era a Cas9, uma vez que ela não foi retida (ausência de banda na faixa de 150 kDa nas frações de eluição). Além disso, notou-se que uma proteína de aproximadamente 25 kDa também foi adsorvida pela resina, porém esta banda não interfere na visualização da proteína Cas9 que apresenta banda na faixa de 150 kDa.

Figura 4.5. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta sem o plasmídeo pET28b-Cas9His em 100 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2, E3 e E4- Frações da etapa de eluição.

Após estabelecer o controle utilizando a cepa sem o plasmídeo pET28b-Cas9His, foram iniciados de fato os ensaios para avaliar a expressão da proteína Cas9 em BL21(DE3) Rosetta. A seguir estão apresentados os resultados dos experimentos realizados na busca por compreender e aumentar o nível de expressão da proteína Cas9 pela cepa Rosetta. Foram utilizados diferentes meios, volumes de cultivo, inóculos, indutores, temperatura de indução, tempo de pós-indução e condições de rompimento celular (Tabela 4.2).

Tabela 4.2. Diferentes parâmetros avaliados para a expressão da proteína Cas9 utilizando *E. coli* BL21(DE3) Rosetta.

Meios de cultivo	LB, ZYP-5052, TB
Volumes de cultivo	100, 250, 500, 850 mL
Indutores	IPTG e lactose
Inóculo	Colônia e banco de células (-80°C)
Temperatura de indução	18, 27, 30°C
Tempo pós-indução	4, 16, 20 e 24h
Rompimento celular (sonicação)	4 e 8 ciclos

Primeiramente, foi realizado um experimento utilizando condições semelhantes às utilizadas por Gagnon e colaboradores (2014), pois este empregou o mesmo plasmídeo e cepa para realizar a expressão da proteína Cas9. Foi utilizado o meio autoindutor ZYP-5052 contendo glicerol, glicose e lactose. A bactéria Rosetta transformada foi cultivada em dois frascos de 500 mL contendo 100 mL de meio ZYP-5052 por 12h a 37°C para crescimento celular. Em seguida, a temperatura de um dos frascos foi reduzida para 27°C e a indução ocorreu por 16h, e para o outro frasco a temperatura foi de 18°C por 24h. Estes experimentos ocorreram antes da verificação da condição de rompimento celular apresentada no item anterior, portanto foram utilizados 16 ciclos de sonicação. As Figura 4.6 e Figura 4.7 mostraram que não foi possível verificar a expressão da proteína Cas9, o que pode ser devido à degradação da mesma pelo sonicador, como constatado no item anterior (item 4.2).

Figura 4.6. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 100 mL de meio ZYP-5052 (indução: lactose à 27°C por 16h) e etapas da purificação para avaliar a expressão da proteína Cas9. Rompimento celular empregando 16 ciclos de sonicação e revelação por *Et al.*. M- marcador de massa molecular; S- sobrenadante extracelular; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2, E3 e E4- Frações da etapa de eluição.

Figura 4.7. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 100 mL de meio ZYP-5052 (indução: lactose à 18°C por 24h) e etapas da purificação para avaliar a expressão da proteína Cas9. Rompimento celular empregando 16 ciclos de sonicação e revelação por *Et al.*. M- marcador de massa molecular; P- *pellet* após 24h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E - Eluição.

Como descrito anteriormente (item 4.2), foi constatado que a linhagem BL21(DE3) foi capaz de expressar a proteína Cas9 em meio LB, decidiu-se assim realizar um experimento semelhante empregando a linhagem Rosetta. A cepa foi cultivada em 100 mL de meio LB e a indução ocorreu com 0,5 mM de IPTG a 30°C por 4h. A Figura 4.8 mostrou que nessa condição não foi possível visualizar uma expressão significativa da proteína Cas9 pela Rosetta, diferentemente do observado ao utilizar a cepa BL21(DE3) nas mesmas condições.

Figura 4.8. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 100 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; R_{sem}- *pellet* Rosetta sem indução; R_{com}- *pellet* Rosetta após 4h de indução; BL_{sem}- *pellet* BL21(DE3) sem indução; BL_{com}- *pellet* BL21(DE3) após 4h de indução.

Para verificar se a ausência da expressão da proteína estava relacionada à perda do plasmídeo pET28b-Cas9His pela *E. coli*, uma nova transformação da cepa Rosetta foi realizada. Para confirmar a presença dos plasmídeos na cepa BL21(DE3) Rosetta, os vetores pRARE e pET28b-Cas9His foram purificados e linearizados com a enzima de restrição NotI. A verificação foi feita por eletroforese em gel de agarose 1% e o resultado está apresentado na Figura 4.9. Foi verificado que os plasmídeos pRARE (~4700 pb) e o pET28b-Cas9His (9468 pb) estavam presentes na Rosetta.

Figura 4.9. Análise dos vetores pET28b-Cas9His e pRARE purificados de BL21(DE3) Rosetta em eletroforese de gel de agarose 1%. Ladder- DNA Ladder 1kb RTU (Kasvi, Brasil); 1- Vetor pRARE não linear; 2- Vetores pET28b-Cas9His e pRARE não lineares; 3- pRARE digerido com a enzima NotI; 4- Vetores pET28b-Cas9His e pRARE digeridos com a enzima NotI.

Em seguida, um novo ensaio foi realizado com as mesmas condições do anterior (100 mL de LB e indução com 0,5 mM de IPTG a 30°C por 4h), sendo que a Rosetta cultivada veio de um banco de células novas. Foram testados dois tipos de inóculo para o ensaio, um gerado pela colônia da placa de transformação da nova cepa e o outro a partir do banco novo de células (utilizando 0,1% v/v). Não foi possível visualizar a expressão da proteína Cas9 em nenhuma das condições avaliadas, indicando que o problema não estava associado à perda de plasmídeo pela bactéria (Figura 4.10).

Figura 4.10. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 100 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) para avaliar a expressão da Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Colônia (1 e 2)- *pellet* após 4h de indução do cultivo com inóculo original de uma colônia de bactéria; Banco (1 e 2)- *pellet* após 4h de indução do cultivo com inóculo original do banco celular em -80°C.

Sabe-se também que a temperatura de indução pode ser um fator importante para a expressão de proteínas recombinantes. Em baixas temperaturas o metabolismo celular é reduzido e contribui para que as taxas de transcrição e tradução também fiquem reduzidas, o que pode favorecer uma correta conformação da proteína e diminuir a formação de corpos de inclusão, aumentando sua solubilidade. Além disso, a atividade das proteases também fica reduzida. Entretanto, como o metabolismo fica mais lento, pode ser necessário um maior tempo de indução para uma significativa expressão da proteína (SAHDEV *et al.*, 2007).

Avaliando a possibilidade de a temperatura de indução ser um fator importante para a expressão da proteína Cas9 pela Rosetta, um novo experimento foi realizado em 100 mL de meio LB e a indução ocorreu com 0,5 mM de IPTG a 18°C por 16h. Mesmo assim, não foi possível visualizar a expressão da proteína Cas9 nessas condições (Figura 4.11).

Figura 4.11. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 100 mL de meio LB (indução: 0,5 mM de IPTG à 18°C por 4 e 16h) para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada; Sem- *pellet* sem indução; 4h- *pellet* após 4h de indução; 16h- *pellet* após 16h de indução.

Na possibilidade da expressão da proteína Cas9 estar sendo muito baixa e, portanto, de difícil visualização pelos métodos aqui utilizados, foi realizado um ensaio em maior volume de cultivo seguido de purificação da proteína Cas9. As condições utilizadas foram iguais às utilizadas pelo autor Gagnon e colaboradores (2014), já mencionado por utilizar o mesmo vetor e cepa. O inóculo foi realizado a partir de uma colônia de plaqueamento. Em seguida, o crescimento celular ocorreu em 850 mL de meio ZYP-5052 a 37°C por 12h e a indução a 18°C por 24h (meio autoindutor pela lactose). Avaliando os resultados apresentados na Figura 4.12, não foi possível visualizar a expressão da proteína Cas9, mesmo em frações das etapas de eluição, que geralmente apresentam maiores concentrações de proteína nas primeiras frações.

Figura 4.12. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 850 mL de meio ZYP-5052 (indução: lactose à 18°C por 24h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada; Sem- *pellet* sem indução; Com- *pellet* após 24h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.

Em seguida, foi realizado um novo experimento para verificar se o meio de cultivo poderia influenciar a expressão da proteína. Foram empregados 3 meios distintos, LB, ZYP-5052 e TB (Terrific Broth). Os cultivos foram realizados em 500 mL de meio. Para o meio LB, a condição de expressão foi mantida como 0,5 mM de IPTG a 30°C por 4h, condição na qual a BL21(DE3) consegue expressar significativamente a proteína Cas9. Para o meio ZYP-5052, a indução ocorreu a 18°C por 24h. Para o meio TB, a indução ocorreu com adição de 0,5 mM de IPTG a 18°C e 20h. O meio TB é mais rico em nutrientes quando comparado ao meio LB podendo favorecer concentrações celulares maiores. Para todas as condições, 1 mM de PMSF foram adicionados ao tampão de ressuspensão para garantir a inibição de proteases. As Figura 4.13, Figura 4.14 e Figura 4.15 apresentam os resultados dos experimentos realizados. Não foi possível visualizar a proteína Cas9 na etapa de eluição dos cultivos empregando meio LB e ZYP-5052; entretanto, para o meio TB, foi possível visualizar uma banda bem fraca na faixa de 150 kDa nas frações de eluição 2 e 3.

Figura 4.13. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 500 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.

Figura 4.14. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 500 mL de meio ZYP-5052 (indução: lactose à 18°C por 24h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Com- *pellet* após 24h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2, E3 e E4- Frações da etapa de eluição.

Figura 4.15. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 500 mL de meio TB (indução: 0,5 mM de IPTG à 18°C por 20h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.

Para verificar se a etapa de rompimento celular estava novamente prejudicando a visualização da expressão da proteína Cas9 pela Rosetta, foi realizado um cultivo em 250 mL de meio LB com indução 0,5 mM de IPTG a 30°C por 4h, mas com rompimento celular de 4 ciclos de 30 segundos cada. O resultado apresentado na Figura 4.16 mostrou que a não visualização da proteína Cas9 nas etapas de eluição não estava relacionada com a etapa de rompimento celular.

Figura 4.16. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 250 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e etapas da purificação com 4 ciclos para rompimento celular. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.

Embora a banda no cultivo com o meio TB tenha sido fracamente visualizada, após todo o estudo sistêmico para a expressão da proteína Cas9 realizado até aqui, constatou-se que a cepa BL21(DE3) Rosetta não foi uma boa alternativa para a expressão da proteína. Portanto, foi realizado um experimento para comparar a expressão da proteína Cas9 pelas cepas BL21(DE3) e BL21(DE3) Rosetta. Os ensaios foram realizados ao mesmo tempo em dois frascos de 1 L contendo 250 mL de meio LB. A indução ocorreu por adição de 0,5 mM de IPTG a 30°C por 4h. Pelos resultados apresentados pelas Figura 4.17 e Figura 4.18 e os anteriores, concluiu-se que a proteína Cas9 não foi obtida a níveis satisfatórios pela cepa Rosetta em nenhuma das condições empregadas e possíveis no laboratório, enquanto que com a cepa BL21(DE3) a produção da proteína Cas9 foi bastante expressiva e visualizada por SDS-PAGE e *Western blot* (Figura 4.19).

Figura 4.17. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) em 250 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.

Figura 4.18. Perfil eletroforético do cultivo da *E. coli* BL21(DE3) Rosetta em 250 mL de meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e etapas da purificação para avaliar a expressão da proteína Cas9. Revelação por *Et al.*. M- marcador de massa molecular; Sem- *pellet* sem indução; Com- *pellet* após 4h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3- Frações da etapa de eluição.

Figura 4.19. Western blot do cultivo da *E. coli* BL21(DE3) e da BL21(DE3) Rosetta em 100 mL meio LB (indução: 0,5 mM de IPTG à 30°C por 4h) e frações da etapa de eluição da purificação para avaliar a expressão da proteína Cas9. M- marcador de massa molecular; R_{sem}- *pellet* da Rosetta sem indução; R_{com}- *pellet* da Rosetta com indução; E2_R e E3_R- frações 2 e 3 da etapa de eluição do cultivo com Rosetta; BL_{sem}- *pellet* da BL21(DE3) sem indução; BL_{com}- *pellet* da BL21(DE3) com indução; E2_{BL} e E3_{BL}- frações 2 e 3 da etapa de eluição do cultivo com BL21(DE3).

É provável que a expressão da proteína Cas9 pela linhagem Rosetta não tenha sido significativa pois a Rosetta possui um plasmídeo extra (pRARE), que pode ter ocasionado uma maior carga metabólica para a célula. Apesar de o plasmídeo pRARE estar presente na cepa para auxiliar na tradução de algumas proteínas que apresentam códons raros em posições críticas para sua síntese, para

a proteína Cas9 essa característica acabou por reprimir sua expressão. Tegel e colaboradores (2010) e Feng e colaboradores (2015) observaram que para proteínas com altos valores de expressão na cepa BL21(DE3), a expressão foi semelhante ou reduzida na BL21(DE3) Rosetta, provavelmente devido ao metabolismo extra requerido pelo plasmídeo pRARE.

4.4. Avaliação da transcrição do gene da proteína Cas9 na BL21(DE3) e BL21(DE3) Rosetta por qRT-PCR

Para tentar identificar se o baixo nível de expressão da proteína Cas9 pela *E. coli* BL21(DE3) Rosetta estava relacionado à transcrição do gene da Cas9 (síntese de mRNA) ou na tradução da proteína, foi realizado o experimento qRT-PCR (*Quantitative Reverse Transcriptase PCR*). O qRT-PCR foi realizado para determinar a quantidade de mRNA da proteína Cas9 sintetizados pelas cepas BL21(DE3) e BL21(DE3) Rosetta. As cepas foram cultivadas em 100 mL de meio LB e induzidas com 0,5 mM de IPTG à 30°C por 4h. Cepas sem o plasmídeo de expressão também foram cultivadas nas mesmas condições para controle. Amostras sem indução também foram analisadas.

Inicialmente foi realizado uma PCR convencional empregando o plasmídeo pET28b-Cas9His para determinar a temperatura de anelamento dos *primers* 2 e 3 (Tabela 3.1) no gene da proteína Cas9. Esses primers foram utilizados na qRT-PCR para quantificar um fragmento de 122 pb do gene da proteína Cas9. A temperatura de anelamento de 54°C foi utilizada e o resultado da PCR está apresentado na Figura 4.20. É possível notar que a temperatura de anelamento de 54°C amplificou apenas o fragmento desejado, indicando ser essa uma temperatura adequada para ser utilizada na qRT-PCR empregando estes *primers*.

Figura 4.20. Gel de agarose 1% para determinar a temperatura de anelamento dos *primers* 2 e 3 (Tabela 3.1) na amplificação do fragmento de 122 pb do gene da proteínas Cas9. 1- GeneRuler 1 kb Plus DNA Ladder (Thermo Fisher Scientific, EUA); 2- amostra diluída 10x de pET28b-Cas9His; 3- resultado da PCR convencional para amplificação do fragmento de 122 pb com temperatura de anelamento de 54°C.

Em seguida, foi realizado a qRT-PCR. Os *pellets* obtidos dos cultivos das *E. coli* foram submetidos ao PureLink™ RNA Mini kit (Invitrogen, EUA) para a extração de RNA. Após a extração de RNA, a mesma massa de RNA total de cada amostra foi submetida ao tratamento com a enzima DNase para eliminar possíveis DNA contaminantes. Após o tratamento com a DNase, o cDNA foi sintetizado empregando a enzima M-MLV *Reverse Transcriptase* (Invitrogen, EUA). No Apêndice E estão apresentados os géis de agarose confirmando a eficiência do tratamento da DNase e a especificidade dos *primers* 2 e 3 (Tabela 3.1) para a amplificação do fragmento de 122 pb do gene da proteína Cas9 a partir do cDNA gerado.

Os cDNAs foram quantificados pelo ensaio de qRT-PCR empregando o kit SYBR™ Green PCR Master Mix (Thermo Fisher Scientific Inc, EUA). Após a amplificação, foi verificada a presença de interações inespecíficas, através da análise das curvas de *melting* (70 a 95°C). O número de cópias para cada amostra foi determinado a partir de uma curva padrão utilizando quantidades conhecidas de pET28b-Cas9His (Figura 4.21). Os resultados da qRT-PCR para as amostras de *E. coli* estão apresentados nas Figura 4.22 e Figura 4.23.

Figura 4.21. Resultados da qRT-PCR do plasmídeo pET28b-Cas9His para determinar a curva padrão em número de cópias por C_T . (a) Curvas de amplificação das amostras de pET28b-Cas9His contendo número de cópias iniciais de: $6 \cdot 10^7$, $6 \cdot 10^6$, $6 \cdot 10^5$, $6 \cdot 10^4$, $6 \cdot 10^3$, $6 \cdot 10^2$ cópias; (b) Curvas de *melting* (dissociação) para cada amostra; (c) Curva padrão da qRT-PCR (quantidade de pET28b-Cas9His x C_T).

Figura 4.22. Resultados da qRT-PCR das amostras de *E. coli* BL21(DE3) com e sem o plasmídeo pET28b-Cas9His, com e sem indução. (a) Curvas de amplificação; (b) Curvas de melting (dissociação).

Figura 4.23. Resultados da qRT-PCR das amostras de *E. coli* BL21(DE3) Rosetta com e sem o plasmídeo pET28b-Cas9His, com e sem indução. (a) Curvas de amplificação; (b) Curvas de melting (dissociação).

Após obter os valores de cópias para cada amostra utilizando a curva padrão, os resultados foram analisados por valores relativos normalizados para o maior número de cópias obtido (amostra da cepa BL21(DE3) contendo o pET28b-Cas9His após a indução). Os resultados estão apresentados na Figura 4.24 e indicaram baixa taxa de transcrição do mRNA da proteína Cas9 pela cepa Rosetta, o que pode justificar a baixa expressão da proteína Cas9 pela cepa. É possível que o tamanho da proteína Cas9 (158 kDa, proteína grande) e o stress metabólico gerado pelo plasmídeo pRARE extra, presente na cepa Rosetta, tenham afetado negativamente as taxas de transcrição do gene da Cas9 nesta cepa. Além disso, a baixa taxa de transcrição pode estar relacionada a um reduzido número de cópias do vetor na cepa (não avaliado neste trabalho). Søggaard e Nørholm (2016) avaliaram a expressão da proteína DasherGFP25 empregando a cepa BL21(DE3) contendo os plasmídeos pLysS e pLysSRARE2. Os resultados mostraram que a presença de cópias extras de tRNA no plasmídeo pLysSRARE2 proporcionaram efeito negativo na expressão da proteína. Os resultados obtidos por estes autores e os apresentados neste trabalho indicaram que a transcrição de tRNAs extras presentes nos plasmídeos pode ocasionar efeitos negativos na expressão de proteínas recombinantes.

Figura 4.24. Valores relativos de cópias de mRNA da proteína Cas9 nas amostras de *E. coli* BL21(DE3) e BL21(DE3) Rosetta com e sem o pET28b-Cas9His (+pET; -pET), induzidas ou não (-ind; +ind), normalizadas para o maior número de cópias obtido (amostra da cepa BL21(DE3) contendo o pET28b-Cas9His após a indução). As barras de erro indicam o desvio padrão entre triplicatas.

Como a BL21(DE3) apresentou resultado promissor para a expressão da proteína Cas9 com protocolo de cultivo relativamente simples, decidiu-se continuar os estudos com essa cepa para otimizar o processo de produção da proteína Cas9.

4.5. Expressão da proteína Cas9 em *E. coli* BL21(DE3)

Os efeitos das condições de cultivo sobre a expressão da proteína Cas9 em *E. coli* BL21(DE3) foram avaliados. Foram empregados o meio complexo LB e o meio definido HDF. A indução sob diferentes temperaturas de indução e tempos de pós-indução foram estudadas. Os ensaios foram realizados em frascos e em biorreator.

4.5.1. Cinética de crescimento da *E. coli* BL21(DE3) em meio LB e meio HDF em frascos

Para conhecer o comportamento da cinética de crescimento da *E. coli* BL21(DE3) contendo o plasmídeo pET28b-Cas9His foram realizados experimentos em meio LB e meio HDF contendo kanamicina à 37°C e 250 rpm em frascos. O meio LB é um meio complexo e rico em nutrientes que permite rápido crescimento celular. Além disso, embora seja um meio complexo, é possível obter reprodutibilidade dos experimentos durante a fase exponencial de crescimento celular (SEZONOV *et al.*, 2007). Já o meio HDF é um meio mineral definido empregado para obter altas densidades celulares, possibilitando maior produtividade ao processo de produção de proteínas. Entretanto, meios de alta densidade celular podem ter o crescimento celular e a produção de proteínas limitados pela transferência de oxigênio. (CHOI *et al.*, 2006; SEEGER *et al.*, 1995). As Figura 4.25 e Figura 4.26 apresentam os resultados obtidos para as cinéticas de crescimento em meio LB e meio HDF, respectivamente. Os experimentos foram realizados em triplicata e os valores de desvio padrão foram baixos (em torno de 0,03), não sendo possível visualizar as barras de erros nos gráficos. A partir dos dados obtidos, o parâmetro velocidade específica máxima de crescimento ($\mu_{\text{máx}}$) foi determinado (Figura 4.27). O $\mu_{\text{máx}}$ para a *E. coli* BL21(DE3) em meio LB foi de 0,68 h⁻¹ e 0,61 h⁻¹ em meio HDF. O fator de conversão de substrato em biomassa ($Y_{x/s}$) foi calculado para o meio HDF e o valor

obtido foi de 0,30 g de célula/g de glicose. O $Y_{x/s}$ não foi determinado para o meio LB pelo fato deste apresentar concentrações muito baixas de glicose.

Figura 4.25. Perfil do crescimento celular (●), concentração de glicose (▲), concentração de acetato (◆) e pH (■) do cultivo de *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em meio LB a 37°C e 250 rpm. Os experimentos foram realizados em triplicata e os valores de desvio padrão foram baixos (em torno de 0,03), não sendo possível visualizar as barras de erros no gráfico.

Figura 4.26. Perfil do crescimento celular (●), concentração de glicose (▲), concentração de acetato (◆) e pH (■) do cultivo de *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em meio HDF a 37°C e 250 rpm. Os experimentos foram realizados em triplicata e os valores de desvio padrão foram baixos (em torno de 0,03), não sendo possível visualizar as barras de erros no gráfico.

Figura 4.27. Determinação do parâmetro cinético velocidade específica máxima de crescimento ($\mu_{\text{máx}}$) da *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em (a) meio LB e (b) meio HDF a 37°C e 250 rpm.

A partir dos resultados obtidos foi possível notar que em meio LB a bactéria apresentou maior velocidade específica máxima de crescimento; entretanto, a fase exponencial de crescimento celular e a densidade celular foram maiores em meio HDF (3h e 1,2 g/L para o meio LB e 4h e 2,8 g/L para o meio HDF) ($\text{DO}_{600\text{nm}} = 1,98$ para o meio LB e $\text{DO}_{600\text{nm}} = 5,12$ para o meio HDF) (Figura 4.27).

O meio LB apresenta limitação de fonte de carbono (concentrações baixas de carboidratos) (Figura 4.25), forçando o metabolismo da célula a utilizar aminoácidos como fonte de carbono e limitando o crescimento celular (TAO *et al.*, 1999). Além

disso, foi possível notar que durante a fase exponencial de crescimento celular ocorreu a produção de acetato e redução do pH. Já na fase estacionária, o pH aumentou, provavelmente devido ao consumo do acetato presente no meio e ao catabolismo de aminoácidos, que resulta em produção de excesso de amônia (SEZONOV *et al.*, 2007).

Li e colaboradores (2014) avaliaram o potencial metabólico da *E. coli* BL21(DE3) em meio complexo (LB) e meio definido (contendo glicose). Os autores constataram que o meio complexo favoreceu um crescimento celular mais rápido (maior $\mu_{\text{máx}}$), pois devido às altas concentrações de nutrientes, a célula não precisa sintetizar muitos dos elementos já presentes no meio (ex: aminoácidos) e direciona seu metabolismo para sintetizar as moléculas necessárias para seu crescimento. Com isso, pode-se verificar velocidade específica de crescimento maior em meio complexo. Entretanto, devido à limitação de fontes de carbono, no meio complexo a concentração celular obtida foi inferior.

O meio HDF possui glicose como fonte de carbono e faz com que o crescimento celular seja inibido pela limitação da fonte de carbono ou pela síntese e acúmulo de subprodutos tóxicos ao crescimento celular, como o acetato. Na Figura 4.26 foi possível verificar que durante o crescimento celular em meio HDF houve redução do valor de pH, que está relacionado com a produção de ácido acético. Ao verificar o final da fase exponencial do crescimento celular, constatou-se que ainda existia glicose no meio (6,6 g/L); portanto, é possível que o crescimento celular tenha sido inibido pela presença de acetato (0,4 g/L). Nakano e colaboradores (1997) observaram que a presença de 0,5 g/L de acetato no meio foi capaz de inibir o crescimento da *E. coli* K12 em frascos.

Ferreira (2019) avaliou a cinética de crescimento celular da *E. coli* BL21(DE3) contendo plasmídeo pET28a para expressão da proteína β -glicosidase. A cinética foi realizada em frascos e meio HDF contendo 10 g/L de glicose. A velocidade específica máxima de crescimento obtida foi de 0,58 h⁻¹ e concentração celular máxima de 2,1 g/L. Nota-se que os valores de $\mu_{\text{máx}}$ e concentração celular obtidos nesse trabalho (0,61 h⁻¹ e 2,8 g/L, respectivamente) estão coerentes aos obtidos por Ferreira.

Beshay e colaboradores (2003) fizeram um estudo sistemático para otimizar o meio de cultura complexo (Terrific Broth) contendo fonte de carbono (lactose) para

produzir β -glucanase. Neste caso, ainda que o meio seja um meio complexo, ele foi suplementado com fonte de carbono para evitar inibição de crescimento celular por substrato. A concentração celular obtida para frascos foi de 3,5 g/L. Embora o valor obtido por Beshay e colaboradores tenha sido superior ao obtido neste estudo empregando o meio HDF (2,8 g/L), o meio HDF, por ser um meio definido contendo glicose, apresenta algumas vantagens para processos industriais, principalmente para a produção de biofármacos, pois facilita o controle do processo de produção e as etapas de purificação.

4.5.2. Expressão da proteína Cas9 em meio LB e meio HDF em frascos

A expressão da proteína Cas9 foi avaliada em relação aos meios LB e HDF. *E. coli* BL21(DE3) contendo o plasmídeo pET28b-Cas9His foram cultivadas em frascos de acordo com o item 3.2.9.3. A indução ocorreu com adição de 0,5 mM de IPTG a 25°C por 4 horas. A expressão da proteína Cas9 foi estimada através de densitometria de acordo com o item 3.3.8. A concentração celular foi determinada ao final do cultivo para o meio LB e o meio HDF, com e sem indução. Os resultados estão apresentados na Figura 4.28 e na Tabela 4.3.

Figura 4.28. Perfil eletroforético da expressão da Cas9 em *E. coli* BL21(DE3) nos meios LB e HDF (indução: 0,5 mM de IPTG à 25°C por 4h). Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada; LBsem- *pellet* sem indução; LBcom- *pellet* após 4h de indução; HDFsem- *pellet* sem indução; HDFcom- *pellet* após 4h de indução.

Tabela 4.3. Valores de concentração celular (g/L) e de massa de proteína Cas9 por litro de cultivo obtidos ao final do cultivo da expressão da proteína Cas9.

Meio de cultivo	Conc. Celular (g/L) Com indução	Conc. de Cas9 (mg Cas9/L cultivo)*	Produtividade Volumétrica (mg Cas9/L.h)	Produtividade Específica (mg Cas9/g CDW.h)
LB	1,02	13,6	1,81	1,8
HDF	2,64	35,0	4,67	1,8

* Determinada por densitometria.

Os resultados mostraram que a produção volumétrica de proteína em meio HDF (35,0 mg de Cas9.L⁻¹ de cultivo) foi aproximadamente 2,7 vezes maior que em meio LB (13,6 mg de Cas9.L⁻¹ de cultivo). Ao verificar a quantidade de proteína Cas9 produzida por célula constatou-se o mesmo valor para as duas condições (13,3 mg de proteína Cas9.g⁻¹ de célula), indicando que a diferença entre as composições dos meios não influenciou na fração de proteína Cas9 produzida pelas células. Portanto, a maior concentração de proteína alcançada ao empregar o meio HDF deve-se principalmente a maior concentração celular obtida ao final do cultivo (Tabela 4.3), como esperado de acordo com o item 4.5.1. É possível que a falta de carboidrato como fonte de carbono no meio LB tenha causado maior estresse metabólico para a célula, interferindo de forma negativa no crescimento celular (TAO *et al.*, 1999).

Com estes resultados foi possível constatar que o emprego de um meio que favorece alta densidade celular pode aumentar a produtividade do processo. Ao empregar o meio HDF, a densidade celular obtida foi maior e, por consequência, a produção da proteína Cas9. Babbal e colaboradores (2019) empregaram os meios LB e TB contendo glicerol para produzir a protease Ulp1 (*catalytic domain of ubiquitin-like protease 1*) em *E. coli*. Os autores verificaram que ao utilizar o meio TB, meio que favorece maiores concentrações celulares quando comparado ao meio LB, a produção aumentou de 142 mg/L para 195 mg/L.

4.5.3. Avaliação das condições de indução na expressão da proteína Cas9 em frascos

Nos processos de produção de proteínas recombinantes por *E. coli* as condições de indução devem ser avaliadas e otimizadas para alcançar maiores níveis de expressão da proteína. Como discutido anteriormente, a temperatura de

indução e o tempo de pós-indução são alguns dos parâmetros que podem favorecer uma melhor produção da proteína. Baixas temperaturas geralmente favorecem uma melhor síntese da proteína devido ao baixo metabolismo celular, assim como a redução da formação de corpos de inclusão. Entretanto, o crescimento celular também fica reduzido. Os tempos de pós-indução podem ser relevantes quando há degradação da proteína com o tempo. Além disso, devido ao baixo metabolismo celular em baixas temperaturas, maiores tempos de pós-indução podem ser necessários para atingir níveis satisfatórios de produção da proteína de interesse.

Em alguns trabalhos encontrados na literatura que descrevem metodologias para a produção da proteína Cas9 é possível encontrar uma variedade de valores de temperaturas e tempos de pós-indução para a expressão da proteína Cas9 (SUN *et al.*, 2015; ZURIS *et al.*, 2015; RAMAKRISHNA *et al.*, 2014). Portanto, os parâmetros temperatura de indução e tempo de pós-indução foram avaliados para otimizar a expressão da proteína Cas9 em *E. coli* BL21(DE3) empregando meio HDF. Os ensaios foram realizados conforme descrito no item 3.2.9.3.

Para determinar a melhor temperatura para a expressão da proteína Cas9 utilizando meio HDF e 0,5 mM de IPTG, foram avaliadas 4 diferentes temperaturas (18, 25, 30 e 37°C). Após determinar a temperatura ótima de expressão, o tempo de pós-indução foi avaliado. Os *pellets* celulares foram analisados em gel de eletroforese SDS-PAGE e a expressão da Cas9 foi estimada por densitometria (software ImageJ). Nas Figura 4.29 e Figura 4.30 e nas Tabela 4.4 e Tabela 4.5 estão apresentados os resultados dos experimentos realizados para as diferentes temperaturas de indução e tempos de pós-indução, respectivamente.

Figura 4.29. Perfil eletroforético da expressão da Cas9 em *E. coli* BL21(DE3) em meio HDF em diferentes temperaturas de indução (indução: 0,5 mM de IPTG por 4h). As amostras contém *pellet* após 4h de indução nas temperaturas 18, 25, 30 e 37°C. Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada.

Tabela 4.4. Valores de concentração celular (g/L) e de massa de proteína Cas9 por litro de cultivo obtidos ao final do cultivo em diferentes temperaturas de indução.

Temperatura (°C)	Conc. Celular (g/L) Com indução	Conc. de Cas9 (mg Cas9/L cultivo)*	Produtividade Volumétrica (mg Cas9/L.h)	Produtividade Específica (mg Cas9/g CDW.h)
18	1,76	12,1	1,61	0,91
25	2,63	29,6	3,94	1,50
30	3,02	45,1	6,01	1,99
37	3,58	30,8	4,10	1,15

* Determinada por densitometria.

Os resultados obtidos indicaram que com a temperatura de indução de 30°C os valores para a produtividade volumétrica e para a produtividade específica foram os maiores (Tabela 4.4). Portanto, foi determinado 30°C como melhor temperatura de expressão para a proteína Cas9.

Figura 4.30. Perfil eletroforético da expressão da proteína Cas9 em *E. coli* BL21(DE3) utilizando meio HDF com indução à 30°C e 0,5 mM de IPTG para diferentes tempos de pós-indução (0h, 2h, 4h, 6h, 8h, 18h e 24h). Revelação por *Et al.*. M- marcador de massa molecular; Cas9- proteína Cas9 purificada.

Tabela 4.5. Valores de concentração celular (g/L) e de massa de proteína Cas9 por litro de cultivo obtidos para cada tempo de pós-indução.

Tempo pós-indução (h)	Conc. Celular (g/L) Com indução	Conc. de Cas9 (mg Cas9/L cultivo)*	Produtividade Volumétrica (mg Cas9/L.h)	Produtividade Específica (mg Cas9/g CDW.h)
0	1,30	2,4	0,68	0,53
2	2,27	22,0	4,00	1,76
4	3,13	52,7	7,02	2,24
6	4,01	74,3	7,82	1,95
8	4,80	96,4	8,38	1,74
18	6,65	98,4	4,57	0,68
24	6,70	75,6	2,75	0,41

* Determinada por densitometria.

Após determinar que a temperatura ótima de expressão foi de 30°C, o tempo de pós-indução foi avaliado. Os resultados apresentados na Tabela 4.5 e na Figura 4.30 mostraram que com 0h de indução, ou seja, sem indução, ocorreu a síntese de proteína na banda de 150 kDa. É possível que a banda presente na faixa de 150 kDa esteja associada à síntese de proteína nativa da *E. coli* BL21(DE3) com massa molecular semelhante a da proteína Cas9 (158 kDa). Seria necessário realizar um experimento com a BL21(DE3) sem o pET28b-Cas9His para confirmar a presença de proteína nativa nessa banda. Outra possibilidade é que o vetor de expressão pET empregado apresenta “escape” da expressão da proteína de interesse. Embora a

expressão da T7 RNA polimerase necessária para iniciar a transcrição dos genes presentes no sistema pET esteja sob o controle do promotor *lac*, sua expressão pode ocorrer a níveis mínimos (*basal level expression*) mesmo sem que ocorra a indução (GIACALONE *Et al.*, 2006; ROSANO e CACCARELLI, 2014). Portanto, é possível que a proteína obtida seja a Cas9. Como a concentração de proteína neste caso foi de apenas 2,4 mg/L, este valor foi considerado baixo e sem influência na avaliação das condições de indução.

Foi observado que um tempo prolongado de cultivo (24h) resultou na degradação da proteína Cas9 presente no meio (Tabela 4.5). A degradação pode ter ocorrido devido à presença de proteases ou por uma possível instabilidade da proteína (TOMAZETTO *et al.*, 2007).

Por fim, os resultados indicaram que a concentração da proteína Cas9 foi maior com 8 e 18 horas de tempo de pós-indução. Levando em consideração que maiores tempos de cultivo acarretam maiores custos de produção e que a concentração de proteínas com 8 e 18h de pós-indução foram bastante semelhantes, determinou-se que o tempo de pós-indução para a produção da proteína Cas9 a 30°C e 0,5 mM de IPTG deve ser de 8 horas. Nestas condições, a concentração de proteína obtida foi de 96,4 mg/L em frascos.

4.5.4. Cinética de crescimento da *E. coli* BL21(DE3) em biorreator

Após determinar a melhor condição de indução para a expressão da proteína Cas9 em frascos e meio definido HDF (item 4.5.3), o desempenho do cultivo em maior escala nessas condições foi avaliado utilizando biorreator. Uma das principais vantagens ao utilizar biorreator é a manutenção do valor de pH constante e da condição de não limitação de oxigênio. Estes dois parâmetros podem ser de grande importância no crescimento celular e expressão de proteínas recombinantes, pois o oxigênio pode ser um limitante do crescimento celular e o pH pode causar estresse metabólico para as células.

Inicialmente, foi realizado um experimento para conhecer a cinética de crescimento da *E. coli* BL21(DE3) em meio HDF e biorreator. O experimento foi realizado conforme o item 3.2.10.2 e teve como finalidade determinar o instante para a indução e expressão da proteína Cas9 em biorreator. A Figura 4.31 apresenta o

resultado obtido da cinética de crescimento da *E. coli* BL21(DE3) em meio HDF. Neste ensaio, a concentração celular máxima foi de 5,6 g/L ($DO_{600nm} = 10,25$) e o valor da velocidade específica máxima de crescimento ($\mu_{m\acute{a}x}$) foi de $0,53\ h^{-1}$ (Figura 4.32). O valor do fator de conversão de substrato em biomassa ($Y_{x/s}$) foi de 0,55 g de célula/g de glicose.

Figura 4.31. Perfil do crescimento celular (●), concentração de glicose (▲), concentração de acetato (◆) e pH (■) do cultivo de *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em meio HDF a 37°C em biorreator.

Figura 4.32. Determinação do parâmetro cinético velocidade específica máxima de crescimento ($\mu_{m\acute{a}x}$) da *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em meio HDF a 37°C em biorreator.

Os valores das variáveis temperatura ($^{\circ}\text{C}$), pH, rotação (rpm) e oxigênio dissolvido (%) ao longo do cultivo foram coletados e armazenados pelo programa *LabVIEW* durante todo o cultivo em biorreator (Figura 4.33). No tempo de 7h foi observada uma perturbação devido à adição de antiespumante, substância tóxica à célula. No tempo próximo à 8h é possível notar um aumento no oxigênio dissolvido, indicando que a célula cessou a respiração (fim da glicose no meio).

Figura 4.33. Variáveis temperatura ($^{\circ}\text{C}$) (●), pH (●), rotação (rpm) (●) e oxigênio dissolvido (%) (●) ao longo de todo o cultivo em biorreator coletados pelo programa *LabVIEW*.

Ao avaliar as velocidades específicas de crescimento em frascos ($0,61\text{ h}^{-1}$) e em biorreator ($0,53\text{ h}^{-1}$) verificou-se que a velocidade de crescimento foi maior em frascos. Entretanto, em biorreator a fase exponencial foi mais longa e os valores de concentração celular e $Y_{x/s}$ foram superiores aos valores obtidos em frascos (4h, 2,8 g/L e 0,30 para frascos e 7h, 5,6 g/L e 0,55 para biorreator). É possível notar que a glicose foi o fator limitante do crescimento celular em biorreator e a produção de acetato foi baixa com valores na faixa de 0,1 g/L. Ademais, o controle de oxigênio dissolvido pode ter favorecido o crescimento celular e a conversão de glicose em célula, pois não foi um fator limitante ao processo.

Carvalho e colaboradores (2012) utilizaram meio HDF contendo 20 g/L de glicose para avaliar a expressão da Proteína A de superfície de *Streptococcus pneumoniae* em *E. coli* BL21(DE3). O $\mu_{\text{máx}}$ foi de $0,54\text{ h}^{-1}$ e a concentração celular obtida na batelada foi de 10,7 g/L. Silva e colaboradores (2013) obtiveram valor de

$\mu_{\text{máx}}$ em torno de $0,5 \text{ h}^{-1}$ e concentração celular de $10,4 \text{ g/L}$ no processo de expressão de hidrolases (xilanase e endoglucanase) empregando *E. coli* BL21(DE3) em meio HDF contendo 20 g/L de glicose em biorreator. Nota-se, portanto, que os valores obtidos para o $\mu_{\text{máx}}$ e para o $Y_{x/s}$ neste trabalho estão de acordo com os valores obtidos por estudos anteriores de crescimento celular de *E. coli* em meio HDF contendo glicose.

Além disso, os autores Beshay e colaboradores (2003), Lara e colaboradores (2008) e Lecina e colaboradores (2013) também obtiveram valores semelhantes de concentração celular e $Y_{x/s}$ para *E. coli* em diferentes meios. Beshay e colaboradores (2003) avaliaram a expressão da β -glucanase em *E. coli* BL21(DE3) usando biorreator e meio complexo (Terrific Broth) contendo lactose (7 g/L) otimizado e obtiveram concentração celular de $5,0 \text{ g/L}$. Lara e colaboradores (2008) obtiveram concentração celular de $14,7 \text{ g/L}$ ao cultivar *E. coli* K12 em biorreator e meio complexo contendo 30 g/L de glicose. Lecina e colaboradores (2013) obtiveram $4,5 \text{ g/L}$ de *E. coli* M15 ao empregar meio definido contendo 10 g/L de glicose. Com isso, pode-se concluir que o meio HDF contendo glicose apresenta potencial para ser utilizado no cultivo de *E. coli* para expressão de proteínas recombinantes quando comparado à outros meios.

Por fim, foram determinados dois instantes de indução para avaliar a expressão da proteína Cas9 em biorreator empregando meio HDF, sendo o primeiro no início da fase exponencial e o segundo na metade da fase exponencial de crescimento celular (Figura 4.31). Os resultados para a expressão da proteína Cas9 em biorreator estão apresentados no item a seguir.

4.5.5. Expressão da proteína Cas9 em biorreator

Após avaliar a cinética de crescimento celular da *E. coli* BL21(DE3) em meio HDF e biorreator, foram determinados dois instantes de indução para avaliar a expressão da proteína Cas9. Geralmente, a indução da expressão da proteína recombinante ocorre durante a fase exponencial de crescimento celular. Nesta fase as células estão com metabolismo ativo para sua duplicação e sintetizando proteínas (COLLINS *et al.*, 2013). Avaliar o impacto do instante de indução sobre a produção da proteína recombinante é de grande importância uma vez que a indução

pode ocasionar um efeito negativo no crescimento celular, e por consequência reduzir a produção celular e de proteínas (ZOU *et al.*, 2014). Portanto, foram realizadas duas bateladas em biorreator para avaliar o efeito do instante de indução sobre o crescimento celular da *E. coli* BL21(DE3) e da expressão da proteína Cas9. As induções ocorreram no início e na metade da fase exponencial de crescimento celular. A indução foi realizada com as condições determinadas para frascos conforme o item 4.5.3, utilizando 0,5 mM de IPTG à 30°C por 8h.

Na Figura 4.34 estão apresentados os resultados obtidos do crescimento celular do cultivo da *E. coli* BL21(DE3) durante a expressão da proteína Cas9 com dois instantes de indução diferentes. É possível notar que a indução teve um efeito negativo no crescimento celular principalmente quando a indução ocorreu no início da fase exponencial. Nesta condição, a concentração celular obtida ao final do cultivo foi de 2,7 g/L, o que pode ser devido ao desvio do metabolismo para a síntese da proteína recombinante, diminuindo o crescimento celular. Para o cultivo em que a indução ocorreu na metade da fase exponencial, o crescimento celular não foi afetado, chegando a uma concentração celular de 5,33 g/L. Por consequência, é preferível realizar a indução na metade da fase exponencial de crescimento, quando os valores de concentração celular são maiores.

Figura 4.34. Perfil do crescimento celular do cultivo de *E. coli* BL21(DE3) contendo vetor pET28b-Cas9His em meio HDF e biorreator na expressão da proteína Cas9 em dois instantes de indução. (●) Crescimento celular sem indução; (●) indução no início da fase exponencial de crescimento celular e (●) indução na metade da fase exponencial de crescimento celular. As setas indicam os dois instantes de indução que serão analisados para a expressão da proteína Cas9.

Os valores das variáveis temperatura ($^{\circ}\text{C}$), pH, rotação (rpm) e oxigênio dissolvido (%) ao longo dos cultivos em biorreator foram coletados e armazenados pelo programa *LabVIEW* (Figura 4.35 e Figura 4.36).

Figura 4.35. Variáveis temperatura ($^{\circ}\text{C}$) (●), pH (●), rotação (rpm) (●) e oxigênio dissolvido (%) (●) ao longo de todo o cultivo em biorreator na expressão da proteína Cas9 com indução no início da fase exponencial de crescimento celular coletados pelo programa *LabVIEW*.

Figura 4.36. Variáveis temperatura ($^{\circ}\text{C}$) (●), pH (●), rotação (rpm) (●) e oxigênio dissolvido (%) (●) ao longo de todo o cultivo em biorreator na expressão da proteína Cas9 com indução na metade da fase exponencial de crescimento celular coletados pelo programa *LabVIEW*.

Para avaliar a expressão da proteína Cas9, amostras contendo *pellet* celular foram coletadas e analisadas por SDS-PAGE. A concentração da proteína foi estimada por densitometria. Os resultados estão apresentados na Figura 4.37 e na Tabela 4.6. Para o tempo de pós-indução de 8h foram obtidos 130,6 mg Cas9/L de cultivo quando a indução ocorreu no início da fase exponencial de crescimento; e 320,2 mg Cas9/L de cultivo quando a indução foi na metade da fase exponencial. A produção de proteína Cas9 foi maior quando a indução ocorreu em maior densidade celular.

Avaliando o cultivo em que a indução ocorreu na metade da fase exponencial de crescimento celular, os maiores valores obtidos para a produtividade e concentração de proteína ocorreram com um tempo de pós-indução de 6h (Tabela 4.6). A concentração de proteína com 6h de pós-indução foi de 420,1 mg Cas9/L de cultivo, enquanto que para 8h de pós-indução foi de 320,2 mg Cas9/L de cultivo. Este resultado indicou que ocorreu degradação da proteína com o tempo devido à presença de proteases ou por uma possível instabilidade da proteína.

O cultivo em biorreator possibilitou um aumento no valor da concentração de proteína quando comparado aos valores obtidos em frascos, sendo que a produção aumentou de 96,4 mg Cas9/L de cultivo para 420,1 mg Cas9/L de cultivo. A melhor condição de indução em biorreator foi com adição de 0,5 mM de IPTG a 30°C na metade da fase exponencial de crescimento celular e com tempo de pós-indução de 6h.

Figura 4.37. Perfil eletroforético da expressão da proteína Cas9 em *E. coli* BL21(DE3) utilizando meio HDF e biorreator com indução à 30°C e 0,5 mM de IPTG para diferentes instantes de indução. Revelação por *Et al.*. M- marcador de massa molecular; I_{sem}- amostra do instante de indução para o cultivo com indução no início da fase exponencial; I_{8h}- *pellet* após 8h da indução no início da fase exponencial; M_{sem}- amostra do instante de indução para o cultivo com indução na metade da fase exponencial; M_{4h}- *pellet* após 4h da indução na metade da fase exponencial; M_{6h}- *pellet* após 6h da indução na metade da fase exponencial; M_{8h}- *pellet* após 8h da indução na metade da fase exponencial.

Tabela 4.6. Valores de concentração celular (g/L) e de massa de proteína Cas9 por litro de cultivo obtidos ao final do cultivo da *E. coli* BL21(DE3) em biorreator para avaliar o efeito do instante de indução.

Instante de indução	Tempo de pós-indução	Concentração celular (g/L)	Conc. de Cas9 (mg Cas9/L cultivo)*	Produtividade Volumétrica (mg Cas9/L.h)	Produtividade Específica (mg Cas9/g CDW.h)
Início da fase exponencial	Sem indução	0,51	2,0	0,7	1,31
	8h	2,27	130,6	11,9	5,23
Metade da fase exponencial	Sem indução	1,84	8,4	1,7	0,91
	4h	4,23	213,8	23,7	5,61
	6h	5,34	420,1	38,2	7,15
	8h	5,33	320,2	24,6	4,62

*Determinada por densitometria.

4.6. Purificação da proteína Cas9 por IMAC

Após avaliar e determinar a melhor condição para a expressão da proteína Cas9 em *E. coli* BL21(DE3) empregando meio HDF contendo glicose e biorreator, foi realizada uma breve avaliação da purificação da proteína Cas9 em 1mL de resina Ni-NTA-agarose (Qiagen, Alemanha). O volume de resina 1mL foi escolhido por ter sido utilizado por autores que já produziram a proteína Cas9 (GAGNON *et al.*, 2014; LIU *et al.*, 2015).

A proteína Cas9 estudada neste trabalho contém uma cauda de histidina na região C- terminal. A cauda de histidina muitas vezes é fusionada às proteínas de interesse para facilitar sua detecção e purificação (ROSANO e CACCARELLI, 2014). A adição de uma cauda His-tag à proteína favorece sua purificação empregando a técnica de IMAC (Cromatografia de afinidade por íons metálicos imobilizados). A técnica consiste em utilizar metais (Ni^{2+} , Co^{2+} , Zn^{2+}) como ligantes que apresentam afinidade por resíduos de histidina. A purificação muitas vezes ocorre em apenas uma etapa, com elevados níveis de pureza e recuperação (ARNAU *et al.*, 2006).

O *pellet* celular do cultivo da *E. coli* BL21(DE3) em biorreator com indução ao final da fase exponencial de crescimento (indução: 0,5 mM de IPTG à 30°C por 6h) foi ressuspendido em 20mL de tampão de lavagem (20 mM Tris-HCl pH8,0; 30mM Imidazol; 500 mM de NaCl), seguido de rompimento celular por sonicação em 8 ciclos de 30 segundos cada, intercalados por 30 segundos, em banho de gelo. Em seguida, a solução foi centrifugada e o sobrenadante seguiu para purificação da proteína Cas9 por cromatografia de afinidade empregando a resina Ni-NTA-agarose. O tampão de lavagem foi empregado como tampão de adsorção. A eluição ocorreu em tampão de lavagem contendo 500 mM de Imidazol. As amostras foram analisadas por SDS-PAGE e densitometria e os resultados estão apresentados na Tabela 4.7 e na Figura 4.38.

Figura 4.38. Perfil eletroforético da expressão da proteína Cas9 em *E. coli* BL21(DE3) utilizando meio HDF e biorreator (indução: 0,5 mM de IPTG à 30°C por 6h) e etapas da purificação da proteína Cas9 em resina de Ni-NTA-agarose. Revelação por *Et al.*. M- marcador de massa molecular; P_{sem}- *pellet* do instante de indução; P_{com}- *pellet* após 6h de indução; I- Injeção (fração solúvel após a sonicação); F- *Flowthrough*; L- Lavagem; E1, E2 e E3 - Frações da etapa de eluição.

Tabela 4.7. Recuperação e pureza relativa da proteína Cas9 nas etapas de purificação, calculadas por densitometria.

Etapa da purificação	Massa total de Cas9 (mg)	Recuperação (%)	Pureza relativa(%)
Injeção	2,8	---	---
<i>Flowthrough</i> + Lavagem	2,3	82	---
Eluição	0,8	28	96

Os resultados indicaram a presença de uma banda na faixa de 150 kDa nas primeiras frações da etapa de eluição (após a adição de imidazol), a qual corresponde à proteína Cas9 que possui massa molecular de 158 kDa. A recuperação da proteína Cas9 obtida na eluição foi de 28%. A pureza relativa da proteína Cas9 foi de 96%, determinada para a primeira fração de eluição. É provável que a perda de proteína Cas9 nas etapas de *flowthrough* e lavagem esteja relacionada com a saturação da resina, uma vez que a etapa de purificação não foi otimizada neste estudo. Além disso, é possível ainda que esteja ocorrendo impedimento estérico devido ao tamanho da proteína Cas9 (elevada massa molecular) e ao ligante utilizado (sem braço espaçador), o que pode reduzir a capacidade de adsorção da resina. Portanto, seria possível ainda otimizar o processo de purificação para obter maior recuperação de proteína Cas9 purificada.

4.7. Avaliação da atividade da proteína Cas9

Para verificar se a proteína Cas9 produzida estava biologicamente ativa, foi realizado um experimento *in vitro* baseado naqueles descritos por Sun e colaboradores (2015) e Yan e colaboradores (2017). O complexo contendo a proteína Cas9 e o gRNA foi incubado com o plasmídeo pVAX1-GFP e sua digestão pela proteína Cas9 foi avaliada.

Para formar o complexo do sistema CRISPR-Cas9, a proteína Cas9 produzida por *E. coli* BL21(DE3) em meio LB e também em meio HDF (item 3.2.9.3) foi purificada e pré-incubada com o gRNA. O gRNA foi sintetizado utilizando o kit GeneArt Precision gRNA Synthesis (Thermo Fisher Scientific, USA), que produz RNAs *in vitro* a partir de *templates* de DNA. Foram utilizados 2 *templates* de DNA para produzir 2 gRNAs, sendo que um deles continha uma sequência de reconhecimento para o gene da GFP (*Green Fluorescent Protein*) presente no pVAX1-GFP (sgRNA), e o outro uma sequência sem reconhecimento do plasmídeo, para ser utilizado como controle (cgRNA). O complexo formado foi adicionado a uma solução contendo o plasmídeo pVAX1-GFP (3697 pb), previamente linearizado com a enzima XbaI. Após a digestão, a atividade da proteína Cas9 foi confirmada através da análise do plasmídeo pVAX1-GFP em gel de agarose 1%. O resultado está apresentado na Figura 4.39.

Figura 4.39. Gel de agarose 1% para verificar a atividade da proteína Cas9 produzida por *E. coli* BL21(DE3) em meio LB e meio HDF. O complexo contendo a proteína Cas9 purificada e o gRNA foi incubado com o plasmídeo pVAX1-GFP (3697 pb) linearizado e os fragmentos após o ensaio de digestão apresentaram 2960 pb e 737 pb.

Os resultados mostraram que as proteínas Cas9 expressadas pela *E. coli* BL21(DE3) usando-se meio LB e também meio HDF apresentaram a atividade esperada. Ao analisar o gel de agarose foi possível verificar que a proteína Cas9 cortou o plasmídeo pVAX1-GFP apenas quando complexada ao sgRNA, que continha a sequência específica para o gene *GFP*.

5. CONCLUSÕES

Neste trabalho foram realizados experimentos em diferentes condições de cultivo para avaliar a expressão e a produção da proteína Cas9 em *E. coli*.

Após confirmar a presença do gene da proteína Cas9 nas cepas BL21(DE3) e BL21(DE3) Rosetta transformadas, foram realizados alguns estudos preliminares para constatar a expressão da proteína Cas9. Foi verificado que a etapa de rompimento celular é crítica ao processo, pois pode causar a degradação da proteína. A presença de PMSF para inibir a ação de proteases não apresentou efeito significativo. A proteína Cas9 foi confirmada por espectrometria de massa, apresentando massa molecular de 158 kDa.

Como primeira estratégia para avaliar a expressão da proteína Cas9 foi empregada a cepa *E. coli* BL21(DE3) Rosetta. Esta cepa contém um plasmídeo extra que expressa tRNAs podendo favorecer a expressão de proteínas que apresentem códons raros para expressão em *E. coli*. Os resultados indicaram que a baixa expressão da proteína Cas9 por esta cepa estava relacionada aos baixos níveis de transcrição do mRNA da proteína, que provavelmente ocorreu devido ao tamanho da proteína Cas9 e da carga metabólica do plasmídeo pRARE presente na Rosetta. Ao empregar a cepa BL21 (DE3), a expressão da proteína Cas9 foi significativa e, portanto, decisiva para continuar o processo de otimização para a expressão da proteína Cas9.

Após definir a cepa BL21(DE3) como sistema hospedeiro para a expressão da proteína Cas9, foram avaliadas a cinética de crescimento e a expressão da proteína em frascos em dois diferentes meios, meio complexo LB e meio definido HDF. A velocidade específica máxima de crescimento ($\mu_{\text{máx}}$) para o meio LB foi de $0,68\text{h}^{-1}$ e para o meio HDF, de $0,61\text{h}^{-1}$. A concentração celular ao final do cultivo em meio definido HDF foi de 2,8 g/L, enquanto que para o meio complexo LB o valor foi de 1,2 g/L. O fator de conversão de substrato em biomassa empregando meio HDF foi de 0,30. O valor da produção da proteína Cas9 em meio HDF foi de 35,0 mg de proteína Cas9/L de cultivo enquanto que em meio LB o valor foi de 13,6 mg de proteína Cas9/L de cultivo. A diferença foi devido à concentração celular final obtida nas duas condições, sendo maior em meio HDF. O ensaio de atividade para a proteína Cas9 expressada nestes dois meios foi positivo.

Para determinar a melhor condição de indução para a expressão da proteína Cas9 em meio HDF, os parâmetros temperatura de indução e tempo de pós-indução foram avaliados em ensaios utilizando frascos. Os resultados indicaram uma melhor expressão da proteína ao empregar uma temperatura durante a indução de 30°C com tempo de indução de 8h.

Para aumentar a escala de produção, a expressão da proteína em biorreator foi avaliada. Foi realizado um ensaio de crescimento celular da *E. coli* em biorreator utilizando meio HDF com a finalidade de conhecer a cinética de crescimento da *E. coli* para posterior avaliação da produção da proteína em biorreator. Os resultados mostraram que a velocidade específica máxima de crescimento obtida no biorreator foi de 0,53 h⁻¹. Embora este valor tenha sido inferior ao obtido em frascos, a fase exponencial de crescimento celular em biorreator foi mais longa. A concentração celular máxima obtida foi de 5,6 g/L e o fator de conversão de substrato em biomassa de 0,55, mostrando melhoria do processo ao empregar o biorreator. O oxigênio e o pH não foram fatores limitantes e de interferência ao crescimento celular, sendo, neste caso, a glicose o fator limitante.

A expressão da proteína Cas9 em biorreator foi avaliada para dois instantes de indução, um no início da fase exponencial de crescimento celular e o outro na metade. Os resultados mostraram que ao induzir a expressão da proteína com maior concentração celular a concentração de proteína aumentou de 130,6 mg de proteína Cas9/L de cultivo para 320,2 mg de proteína Cas9/L de cultivo. Ao realizar a indução com maior concentração celular, o pico de produção da proteína Cas9 foi com tempo de pós-indução de 6h, com concentração de proteína de 420,1 mg de Cas9/L de cultivo. Com isso, a melhor condição de indução para produzir a proteína Cas9 empregando *E. coli* BL21(DE3) em meio HDF e biorreator foi com adição de 0,5 mM de IPTG na metade da fase exponencial de crescimento celular a 30°C por 6h.

A etapa de purificação foi avaliada e uma otimização do processo é necessária para aumentar a recuperação da proteína de interesse.

Finalmente, este trabalho apresentou um estudo sistemático do processo de produção da proteína Cas9 empregando *E. coli* BL21(DE3) e meio definido HDF contendo glicose, contribuindo com informações importantes sobre a expressão da proteína Cas9 em *E. coli* e com potencial para ser empregado em larga escala.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- AGARWAL, S.; JHA, S.; SANYAL, I.; AMLA, D.V. expression. And purification of recombinant human α 1-proteinase inhibitor and its single amino acid substituted variants in *Escherichia coli* for enhanced stability and biological activity. **Journal of Biotechnology**, v. 147, p. 64-72, 2010.
- AGUIRRE-LÓPEZ, B.; CABRERA, N.; GÓMEZ-PUYOU, M.T.; PEREZ-MONTFORT, R.; GÓMEZ-PUYOU, A. The importance of arginine codons ADA and AGG for the expression in *E. coli* of triosephosphate isomerase from seven different species. **Biotechnology Reports**, v. 13, p. 42-48, 2017.
- AKESSON, M.; HAGANDE, P.; AXELSSON, J.P. Avoiding acetate accumulation in *Escherichia coli* cultures using feedback control of glucose feeding. **Biotechnology and Bioengineering**, v. 73, n. 3, 2001.
- AMOASII, L.; HILDYARD, J.C.W.; LI, H.; SANCHEZ-ORTIZ, E.; MIREAULT, A.; CABALLERO, D.; HARRON, R.; STATHOPOULOU, T.-R.; MASSEY, C.; SHELTON, J.M.; BASSEL-DUDY, R.; PIERCY, R.J.; OLSON, E.N. Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. **Science**, v. 362, p. 86-91, 2018.
- APSANA, R.; BALAMURUGAN, V.; VEEREGOWDA, B.M.; ABRAHAM, S.; RAJU, D.S.N.; RATHNAMMA, D.; BYREGOWDA, S.M.; RAHMAN, H.; SHAILA, M.S. expression. And characterization of immunodominant region of fusion protein of peste des petits ruminants virus in *E. coli*. **Small Ruminant Research**, v. 144, p. 75-82, 2016.
- ARNAU, J.; LAURITZEN, C.; PETERSEN, G.E.; PEDERSEN, J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. **Protein Expression and Purification**, v. 48, p. 1-13, 2006.
- BABBAL; ADIVITIYA; MOHANTY, S.; KHASA, Y.P. Biprocess optimization for the overproduction of catalytic domain of ubiquitin-like protease 1 (Ulp1) from *S. cerevisiae* in *E. coli* fed-batch culture. **Enzyme and Microbial Technology**, v. 120, p. 98-109, 2019.
- BAESHEN, M. N.; AL-HEJIN, A.M.; BORA, R.S.; AHMED, M.M.M.; RAMADAN, H.A.I.; SAINI, K.S.; BAESHEN, N.A.; REDWAN, E.M. Production of biopharmaceuticals in *E. coli*: current scenario and future perspectives. **Journal of Microbiology and Biotechnology**, v. 25, n. 7, p. 953-962, 2015.

- BANEYX, F. Recombinant protein expression in *Escherichia coli*. **Current Opinion in Biotechnology**, v. 10, n. 5, p. 411-421, 1999.
- BERLEC, A.; STRUKELJ, B. Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells. **Journal of Industrial Microbiology and Biotechnology**, v. 40, p. 257-274, 2013.
- BESHAY, U.; EL-ENSHASY, H.; ISMAIL, I.M.K.; MOAWAD, H.; WOJEIECHOWASKA, E.; ABD-EL-GHANY, S. β -Glucanase production from genetically modified recombinant *Escherichia coli*: effect of growth substrates and development of a culture medium in shake flasks and stirred tank bioreactor. **Process Biochemistry**, v. 39, p. 307-313, 2003.
- BLENKE, E.O.; EVERS, M.J.W.; MASTROBATTISTA, E.; van der OOST, J. CRISPR-Cas9 gene editing: Delivery aspects and therapeutic potential. **Journal of Controlled Release**, v. 244, p. 139-148, 2016.
- BRADFORD, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, v. 72, n. 1-2, p. 248-254, 1976.
- BROKOWSKI, C.; ADLI, M. CRISPR Ethics: Moral considerations for applications of a powerful tool. **Journal of Molecular Biology**, v. 431, p. 88-101, 2019.
- BURGESS-BROWN, N.A.; SHARMA, S.; SOBOTT, F.; LOENARZ, C.; OPPERMANN, U.; GILEADI, O. Codon optimization can improve expression of human genes in *Escherichia coli*: a multi-gene study. **Protein Expression and Purification**, v. 59, p. 94-102, 2008.
- CARNEIRO, S.; FERREIRA, E.C.; ROCHA, I. Metabolic responses to recombinant bioprocesses in *Escherichia coli*. **Journal of Biotechnology**, v. 164, p. 396-408, 2013.
- CARVALHO, R.J.; CABRERA-CRESPO, J.; TANIZAKI, M.M.; GONÇALVES, V.M. Development of production and purification processes of recombinant fragment of pneumococcal surface protein A in *Escherichia coli* using different carbon sources and chromatography sequences. **Applied Microbiology and Biotechnology**, v. 94, p. 683-694, 2012.
- CASPETA, L.; LARA, A.R.; PÉREZ, N.O.; FLORES, N.; BOLÍVAR, F.; RAMÍREZ, O.T. Enhancing thermo-induced recombinant protein production in *Escherichia coli* by temperature oscillations and post-induction nutrient feeding strategies. **Journal of Biotechnology**, v. 167, p. 47-55, 2013.

- CHEN, Y.; KING, X.-H.; YE, F.; KUANG, Y.; LUO, M. Production of MBP-HepA fusion protein in recombinant *Escherichia coli* by optimization of culture medium. **Biochemical Engineering Journal**, v. 34, p. 114-121, 2007.
- CHEN, D.; TEXADA, D.E. Low-usage codons and rare codons of *Escherichia coli*. **Gene Therapy and Molecular Biology**, v. 10, p. 1-10, 2006.
- CHO, S.W.; KIM, S.; KIM, J.M.; KIM, J.-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. **Nature Biotechnology**, v. 31, n. 3, p. 230-232, 2013.
- CHOI, J.H.; KEUM, K.C.; LEE, S.Y. Production of recombinant proteins by high cell density culture of *Escherichia coli*. **Chemical Engineering Science**, v. 61, p. 876-885, 2006.
- CHOU, C.P. Engineering cell physiology to enhance recombinant protein production in *Escherichia coli*. **Applied Microbiology and Biotechnology**, v. 76, p. 521-532, 2007.
- CHU, V.T.; WEBER, T.; GRAF, R.; SOMMERMAN, T.; PETSCH, K.; SACK, U.; VOLCHKOV, P.; RAJEWASKY, K.; KUHN, R. Efficient generation of Rosa 26 knock-in mice using CRISPR/Cas9 in C57BL/6 zygotes. **BMC Biotechnology**, v. 16, n. 4., 2016.
- COLLINS, T.; AZEVEDO-SILVA, J.; COSTA, A.; BRANCA, F.; MACHADO, R.; CASAL, M. Batch production of a silk-elastin-like protein in *E. coli* BL21(DE3): key parameters for optimization. **Microbial Cell Factories**, v. 12, n. 21, 2013.
- CYRANOSKI, D. CRISPR gene-editing tested in a person for the first time. **Nature**, v. 539, n. 479, 2016.
- DAGAR, V.K.; KHASA, Y.P. High-level expression. And efficient refolding of therapeutically important recombinant human Interleukin-3 (hIL-3) in *E. coli*. **Protein Expression and Purification**, v. 131, p. 51-59, 2017.
- DASH, P.K.; KAMINSKI, R.; BELLA, R.; SU, H.; MATHEWS, S.; AHOYI, T.M.; CHEN, C.; MANCUSO, P.; SARIYER, R.; FERRANTE, P.; DONADONI, M.; ROBINSON, J.A.; SILLMAN, B.; LIN, Z.; HILAIRE, J.R.; BANOUB, M.; ELANGO, M.; GAUTAM, N.; MOSLEY, R.L.; POLUEKTOVA, L.Y.; McMILLAN, J.; BADE, A.N.; GORANTLA, S.; SARIYER, I.K.; BURDO, T.H.; YOUNG, W.-B.; AMINI, S.; GORDO, J.; JACOBSON, J.M.; EDAGWA, B.; KHALILI, K.; GENDELMAN, H.E. Sequential LASER ART and CRISPR treatments

- eliminate HIV-1 in a subset of infected humanized mice. **Nature Communications**, v. 10, n. 2753, 2019.
- DEMAIN, A.L.; VAISHNAV, P. Production of recombinant proteins by microbes and higher organisms. **Biotechnology Advances**, v. 27, p. 297-306, 2009.
- DeWITT, M.A.; CORN, J.E.; CARROLL, D. Genome editing via delivery of Cas9 ribonucleoprotein. **Methods**, v.121-122, p. 9-15, 2017.
- DONOVAN, R.S.; ROBINSON, C.W.; GLICK, B.R. Review: Optimizing inducer and culture conditions for expression of foreign proteins under the control of the *lac* promoter. **Journal of Industrial Microbiology**, v. 16, p. 145-154, 1996.
- DOUDNA, J.A.; CHARPENTIER, E. The new frontier of genome engineering with CRISPR-Cas9. **Science**, v. 346, n. 6213, 2014.
- DUMON-SEIGNOVERT, L.; CARIOT, G.; VUILLARD, L. The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3) and C43(DE3). **Protein Expression and Purification**, v. 37, p. 203-206, 2004.
- FELIU, J.X.; VILLAVERDE, A. Na optimized ultrasonication protocol for bacterial cell disruption and recovery of β -galactosidase fusion proteins. **Biotechnology Techniques**, v. 8, n. 7, p. 509-514, 1994.
- FENG, J. WAN, R.; YI, Q.; HE, L.; YANG, L.; TANG, L. Examination of alternate codon bias solutions for expression and purification of recombinant mechano-growth factor in *Escherichia coli*. **Biotechnology and Applied Biochemistry**, v. 62, n. 5, p. 690-698, 2015.
- FERREIRA, R.G. Development and *in silico* evaluation of an expression platform based on *E. coli* for the production of a recombinant beta-glucosidase. 190 p. Tese (Doutorado em Engenharia Química)- Departamento de Engenharia Química, Escola Politécnica da Universidade de São Paulo, São Paulo, 2019.
- FRANCIS, D.M.; PAGE, R. Strategies to optimize protein expression in *E. coli*. **Current Protocols in Protein Science**, 5.24, 2010.
- GAGNON, J.A.; VALEN, E.; THYME, S.B.; HUANG, P.; AHKMETOVA, L.; PAULI, A.; MONTAGUE, T.G.; ZIMMERMAN, S.; RICHTER, C.; SCHIER, A.F. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. **Plos One**, v. 9, n. 5, 2014.

- GAJ, T.; GUO, J.; KATO, Y.; SIRK, S.J.; BARBAS iii, C.F. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. **Nature Methods**, v. 9, n. 8. P. 805-808, 2012.
- GALLOWAY, C.A.; SOWDEN, M.P.; SMITH, H.C. Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase. **Biotechniques**, v. 34, n. 3, p. 524-530, 2003
- GECIOVA, J.; BURY, D.; JELEN, P. Methods for disruption of microbial cells for potential use in dairy industry- a review. **International Dairy Journal**, v. 12, p. 541-553, 2002.
- GIACALONE, M.J.; GENTILE, A.M.; LOVITT, B.T.; BERKLEY, N.L.; GUNDERSON, C.W.; SURBER, M.W. Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. **BioTechniques**, v. 40, p. 355-364, 2006.
- GOMBERT, A.K.; KILIKIAN, B.V. Recombinant gene expression in *Escherichia coli* cultivation using lactose as inducer. **Journal of Biotechnology**, v. 60, p. 47-54, 1998.
- GOPAL, G.J.; KUMAR, A. Strategies for the production of recombinant protein in *Escherichia coli*. **The protein Journal**, v. 32, n. 6, p. 419-425, 2013.
- GNOTH, S.; JANZSCH, M.; SIMUTIS, R.; LUBBERT, A. Control of cultivation processes for recombinant protein production: a review. **Bioprocess and Biosystems Engineering**, v. 32, p. 21-39, 2008.
- HARRISON, T.L. Bacterial cell disruption: a key unit operation in the recovery of intracellular products. **Biotechnology Advances**, v. 9, n. 2, p. 217-240, 1991.
- HU, J.-H.; WANG, F.; LIU, C.-Z. Development of an efficient process intensification strategy for enhancing *Pfu* DNA polymerase production in recombinant *Escherichia coli*. **Bioprocess and Biosystems Engineering**, v. 38, p. 651-659, 2015.
- HU, J.H.; DAVIS, K.M.; LIU, D.R. Chemical biology approaches to genome editing: understanding, controlling, and delivering programmable nucleases. **Cell Chemical Biology**, v. 23, p. 57-73, 2016
- HUANG, C.-J.; LIN, H.; YANG, X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. **Journal of Industrial Microbiology and Biotechnology**, v. 39, p. 383-399, 2012.

- HWANG, W.Y.; FU, Y.; REYON, D.; MAEDER, M.L.; TSAI, S.Q.; SANDER, J.D.; PETERSON, R.T.; YEH, J.-R.J.; JOUNG, J.K. Efficient genome editing in zebrafish using a CRISPR-Cas system. **Nature Biotechnology**, v. 31, p. 227-229, 2013.
- JINEK, M.; CHYLINSKI, K.; FONFARA, I.; HAUER, M.; DOUDNA, J.A.; CHARPENTIER, E. A programmable Dual-RNA-Guided DNA endonuclease in adaptive bacterial immunity. **Science**, v. 337, n. 6096, p. 816-821, 2012.
- JINEK, M.; JIANG, F.; TAYLOR, D.W.; STERNBERG, S.H.; KAYA, E.; MA, E.; ANDERS, C.; HAUER, M.; ZHOU, K.; LIN, S.; KAPLAN, M.; IAVARONE, A.T.; CHARPENTIER, E.; NOGALES, E.; DOUDNA, J.A. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. **Science**, v. 343, n. 6176, 2014.
- JONASSON, P.; LILJEQVIST, S.; NYGREN, P.-A.; STAHL, S. Genetic design for facilitated production and recovery of recombinant proteins in *Escherichia coli*. **Biotechnology and Applied Biochemistry**, v. 35, p. 91-105, 2002.
- JOSHI, H.; JAIN, V. novel method to rapidly and efficiently lyse *Escherichia coli* for the isolation of recombinant protein. **Analytical Biochemistry**, v. 528, p. 1-6, 2017.
- KAUR, J.; KUMAR, A.; KAUR, J. Strategies for optimization of heterologous protein expression in *E. coli*: roadblocks and reinforcements. **International Journal of Biotechnology Macromolecules**, v. 106, p. 803-822, 2018.
- KENNEDY, E.M.; KORNEPATI, A.V.R.; GOLDSTEIN, M.; BOGERD, H.P.; POLING, B.C.; WHISNANT, A.W.; KASTAN, M.B.; CULLEN, B.R. Inactivation of the Human Papillomavirus E6 or E7 Gene in cervical carcinoma cells by using a bacterial CRISPR/Cas RNA-guided endonuclease. **Journal of Virology**, v. 88, n. 20, p. 11965-11972, 2014.
- KIM, S.-B.; LEE, W.; RYU, Y.-W. Cloning and characterization of thermostable esterase from *Archaeoglobus fulgidus*. **The Journal of Microbiology**, v. 46, n. 1, p. 100-107, 2008.
- KIM, S.; KIM, D.; CHO, S.W.; KIM, J.; KIM, J.-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. **Genome Research**, v. 24, p. 1012-1019, 2016.

- KIM, J.; KIM, S.I.; HONG, E.; RYU, Y. Strategies for increasing heterologous expression of a thermostable esterase from *Archaeoglobus fulgidus* in *Escherichia coli*. **Protein Expression and Purification**, v. 127, p. 98-104, 2016.
- KIM, S.M.; SHIN, S.C.; KIM, E.E.; KIM, S.-H.; PARK, K.; OH, S.J.; JANG, M. Simple *in vivo* gene editing via direct self-assembly of Cas9 ribonucleoprotein complexes for cancer treatment. **ACS Nano**, v. 12, p. 7750-7760, 2018.
- KONZ, J.O.; KING, J.; COONEY, C.L. Effects of oxygen on recombinant protein expression. **Biotechnology Progress**, v. 14, p. 393-409, 1998.
- LARA, A.R.; CASPETA, L.; GOSSET, G.; BOLÍVAR, F.; RAMÍREZ, O.T. Utility of na *Escherichia coli* strain engineering in the substrate uptake system for improved culture performance at high glucose and cell concentrations: an alternative to fed-batch cultures. **Biotechnology and Engineering**, v. 99, n. 4, 2008.
- LARENTIS, A.L.; ARGONDIZZO, A.P.C.; ESTEVES, G.S.; JESSOURON E.; GALLER, R.; MEDEIROS, M.A. Cloning and optimization of induction conditions for mature PsaA (pneumococcal surface adhesin A) expression in *Escherichia coli* and recombinant protein stability during long-term storage. **Protein Expression and Purification**, v. 78, p. 38-47, 2011.
- LECINA, M.; SARRÓ, E.; CASABLANCAS, A.; GÒDIA, F.; CAIRÓ, J.J. IPTG limitation avoids metabolic burden and acetic acid accumulation in induced fed-batch cultures of *Escherichia coli* M15 under glucose limiting conditions. **Biochemical Engineering Journal**, v. 70, p. 78-83, 2013.
- LI, Z.; NIMTZ, M.; RINAS, U. The metabolic potential of *Escherichia coli* BL21 in defined and rich medium. **Microbial Cell Factories**, v. 13, n. 45, 2014.
- LI, Y.; LIN, Z.; HUANG, C.; ZHANG, Y.; WANG, Z.; TANG, Y.; CHEN, T.; ZHAO, X. Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 mediated genome editing. **Metabolic Engineering**, v. 31, p. 13-21, 2015.
- LI, L.; HU, S.; CHEN, X. Non-viral delivery systems for CRISPR/Cas9-based genome editing: challenges and opportunities. **Biomaterials**, v. 171, p. 207-218, 2018.
- LIANG, X.; POTTER, J.; KUMAR, S.; ZOU, Y.; QUINTANILLA, R.; SRIDHARAN, M.; CARTE, J.; CHEN, W.; ROARK, N.; RANGANATHAN, S.; RAVINDER, N.; CHESNUT, J.D. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. **Journal of Biotechnology**, v. 208, p. 44-53, 2015.

- LIANG, Z.; CHEN, K.; LI, T.; ZHANG, Y.; WANG, Y.; ZHAO, Q.; LIU, J.; ZHANG, H.; LIU, C.; RAN, Y.; GAO, C. Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. **Nature Communications**, 8:14261, 2017.
- LIANG, X.; POTTER, J.; KUMAR, S.; RAVINDER, N.; CHESNUT, J.D. Enhanced CRISPR/Cas9-mediated precise genome editing by improves design and delivery of gRNA, Cas9 nuclease, and donor DNA. **Journal of Biotechnology**, v. 241, p. 136-146, 2017.
- LIU, J.; GAJ, T.; YANG, Y.; WANG, N.; SHUI, S.; KIM, S.; KANCHISWAMY, N.C.; KIM, J.S.; BARBAS, C.F. Efficient delivery of nuclease proteins for genome editing in human stem cells and primary cells. *Nature Protocols*, v. 10, p. 1842-1859, 2015.
- MA, H.; MARTI-GUTIERREZ, N.; PARK, S.-W.; WU, J.; LEE, Y.; SUZUKI, K.; KOSKI, A.; JI, D.; HAYAMA, T.; AHMED, R.; DARBY, H.; DYKEN C.V.; LI, Y.; KANG, E.; P. A.-R.; KIM, D.; KIM, S.-T.; GONG, J.; GU, Y.; XU, X.; BATTAGLIA, D.; KRIEG, S.A.; LEE, D.M.; WU, D.H.; WOLF, D.P.; KEITNER, S.B.; BELMONTE, J.C.I.; AMATO, P.; KIM, J.S.- KAUL, S.; MITALIPOV, S. Correction of a pathogenic gene mutation in human embryos. **Nature**, DOI: 10.1038/nature23305, 2017.
- MADDALO, D.; MANCHADO, E.; CAONCEPTION, C.P.; BONETTI, C.; VIDIGAL, J.A.; HAN, Y.-C.; OGRODOWSKI, P.; CRIPPA, A.; REKHTMAN, N.; STANCHINA E.; LOWE, S.W.; VENTURA, A. *In vivo* engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. **Nature**, v. 516, p. 423-502, 2014.
- MALI, P.; ESVELT, K.M.; CHURCH, G.M. Cas9 as a versatile tool for engineering biology. **Nature Methods**, v. 10, n. 10, p. 957-963, 2013a.
- MALI, P.; YANG, L.; ESVELT, K.M.; AACH, J.; GUELL, M.; DICARLO, J.E.; NORVILEE, J.E.; CHURCH, G.M. RNA-Guided human genome engineering via Cas9. **Science**, v. 339, 2013b.
- MEMI, F.; NTOKOU, A.; PAPANGELI, I. CRISPR/Cas9 gene-editing: research technologies, clinical applications and ethical considerations. **Seminars in Perinatology**, v. 42, p. 487-500, 2018.

- MENZELLA, H.G.; CECCARELLI, E.A.; GRAMAJO, H.C. Noval *Escherichia coli* strain allows efficient recombinant protein production using lactose as inducer. **Biotechnology and Bioengineering**, v. 82, n. 7, 2003.
- MEYER, H.-P.; LEINT, C.; FIECHTER, A. Acetate formation in continuous culture of *Escherichia coli* K12 D1 on defined and complex media. **Journal of Biotechnology**, v. 1, p. 355-358, 1984.
- MONTAGNA, C.; PATRIS, G.; CASINI, A.; MAULE, G.; FRANCESCHINI, G.M.; ZANELLA, I.; CONTI, L.; ARNOLDI, F.; BURRONE, O.R.; ZENTILIN, L.; ZACCHIGNA, S.; GIACCA, M.; CERESETO, A. VSV-G-Enveloped vesicles for traceless delivery of CRISPR-Cas9. **Molecular Therapy: Nucleic Acids**, v. 12, p. 453-462, 2018.
- MOUNT, R.; RAY, M.; TONGA, G.Y.; LEE, Y.-W.; TAY, T.; SASAKI, K.; ROTELLO V.M. Direct cytosolic delivery of CRISPR/Cas9-ribonucleoprotein for efficient gene editing. **ACS Nano**, v. 11, n. 3, p. 2452-2458, 2017a.
- MOUNT, R.; RAY, M.; LEE, Y.-W.; SCALETTI, F.; ROTELLO, V.M. *In vivo* delivery of CRISPR/Cas9 for therapeutic gene editing: progress and challenges. **Bioconjugate Chemistry**, v. 28, p. 880-884, 2017b.
- MOUT, R.; RAY, M.; TONGA, G.Y.; LEE, Y.-W.; TAY, T.; SASAKI, K.; ROTELLO, V.M. Direct cytosolic delivery of CRISPR-Cas9-ribonucleoprotein for efficient gene editing. **ACS Nano**, v. 11, p. 2452-2458, 2017.
- NAKANO, K.; RISCHKE, M.; SATO, S.; MARKL, H. Influence of acetic acid on the growth of *Escherichia coli* K12 during high-cell-density cultivation in a dialysis reactor. **Applied Microbiology and Biotechnology**, v. 48, p. 597-601, 1997.
- NIU, D.; WEI, H.-J.; LIN, L.; GEORGE, H.; WANG, T.; LEE, I.-H.; ZHAO, H.-Y.; WANG, Y.; KAN, Y.; SHROCK, E.; LESH, E.; WANG, G.; LUO, Y.; QING, Y.; JIAO, D.; ZHAO, H.; ZHOU, X.; WANG, S.; WEI, H.; GUELL, M.; CHURCH, G.M.; YANG, L. Inactivation of porcine endogenous retrovirus in pigs using CRISPR-Cas9. **Science**, DOI: <https://doi.org/10.1126/science.aan4187>, 2017.
- NISHIMASU, H.; RAN, F.A.; HSU, P.D.; KONERMANN, S.; SHAHATA, S.I.; DOHMAE, N.; ISHITANI, R.; ZHANG, F.; NUREKI, O. Crystal structure of Cas9 in complex with Guide RNA and Target DNA. **Cell**, v. 156, p. 935-949, 2014.

- NOVY, R.; DROTT, D.; YAEGER, K.; MIERENDORF, R. Overcoming the codons bias of *E. coli* for enhanced protein expression. **Innovations**, n. 12, p. 1-3, 2001.
- OLAOFE, O.A.; BURTON, S.G.; COWAN, D.A.; HARRISON, S.T.L. Improving the production of a thermostable amidase through optimizing IPTG induction in a highly dense culture of recombinant *Escherichia coli*. **Biochemical Engineering Journal**, v. 52, p. 19-24, 2010.
- OVERTON, T.W.; Recombinant protein production in bacterial hosts. **Drug Discovery Today**, v. 19, n. 5, 2014.
- 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.
- PETERNEL, S. Bacterial cell disruption: a crucial step in protein production. **New Biotechnology**, v. 30, n. 2, p. 250-254, 2013.
- PHUE, J.-N.; SHILOACH, J. Impact of dissolved oxygen concentration on acetate accumulation and physiology of *E. coli* BL21, evaluating transcription levels of key genes at different dissolved oxygen conditions. **Metabolic Engineering**, v. 7, p. 353-363, 2005.
- PORATH, J. IMAC- Immobilized metal ion affinity based chromatography. **TrAC Trends in Analytical Chemistry**, v. 7, n. 7, p. 254-259, 1988.
- QUAX, T.E.F.; CLAASSENS, N.J.; SOIL, D.; OOST, J.van der. Codon bias as a means to fine-tune gene expression. **Molecular Cell**, v. 59, n. 2, p. 149-161, 2015.
- RAJAGOPALAN, N.; KAGALE, S.; BHOWMIK, P.; SONG, H. A two-step method for obtaining highly pure Cas9 nuclease for genome editing, biophysical, and structural studies. **Methods and Protocols**, v. 1, n. 17, 2018.
- RAMAKRISHNA, S.; DAD, A.-B.K.; BELOOR, J.; GOPALAPPA, R.; LEE, S.-K.; KIM, H. Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. **Genome Research**, v. 21, p. 1020-1027, 2014.
- RAN, F.A.; CONG, L.; YAN, W.X.; SCOTT, D.A.; GOOTENBERG, J.S.; KRITZ, A.J.; ZETSCHE, B.; SHALEM, O.; WU, X.; MAKAROVA, K.S.; KOONIN, E.V.; SHARP, P.A.; ZHANG, F. *In vivo* genome editing using *Staphylococcus aureus* Cas9. **Nature**, v. 520, 2015.

- ROSANO, G.L.; CACCARELLI, E.A. Recombinant protein expression in *Escherichia coli*: advances and challenges. **Frontiers in Microbiology**, v. 5, n. 172, 2014.
- SAHDEV, S.; KHATTAR, S.K.; SAINI, K.S. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. **Molecular and Cellular Biochemistry**, v. 307, p. 249-264, 2008.
- SAIDA, F.; UZAN, M.; ODAERT, B.; BONTEMS, F. Expression of highly toxic genes in *E. coli*: Special strategies and genetic tools. **Current Protein and Peptide Science**, v. 7, p. 47-56, 2006.
- SAIDA, F. Overview on the expression of toxic gene products in *Escherichia coli*. **Current Protocols in Protein Science**, 50:5.19:5.19.1–5.19.13, 2007.
- SAN-MIGUEL, T.; PÉREZ-BERMÚDEZ, P.; GAVIDIA, I. Production of soluble eukaryotic recombinant proteins in *e. coli* is favored in early log-phase cultures induced at low temperature. **SpringerPlus**, v. 2, n. 89, 2013.
- SCHLEGEL, S.; RUJAS, E.; YTTERBERG, A.J.; ZUBAREV, R.A.; LUIRINK, J.; GIER, J.-W. Optimizing heterologous protein production in the periplasm of *E. coli* by regulating gene expression levels. **Microbial Cell Factories**, 12:24, 2013.
- SEEGER, A.; SCHNEPPE, B.; MCCARTHY, J.E.G.; DECKWER, W.-D.; RINAS, U. Comparison of temperature- and isopropyl- β -D-thiogalacto-pyranoside-induced synthesis of basic fibroblast growth factor in high-cell-density cultures of recombinant *Escherichia coli*. **Enzyme and Microbial Technology**, v. 17, p. 947-953, 1995.
- SEZONOV, G.; JOSELEAU-PETIT, D.; D'ARI R. *Escherichia coli* physiology in Luria-Bertani Broth. **Journal of Bacteriology**, v. 189, n. 23, p. 8746-8749, 2007.
- SILVA, M.R.; SEVERO, M.G.; DELABONA, P.S.; RULLER, R.; PRADELLA, J.G.C.; GONÇALVES, V.M.; FREITAS, S. High cell density co-culture for production of recombinant hydrolases. **Biochemical Engineering Journal**, v. 71, p. 38-46, 2013.
- SINGH, V.; BRADDICK, D.; DHAR, P.K. Exploring the potential of genome editing CRISPR-Cas9 technology. **Gene**, v. 599, p. 1-18, 2017.
- SINGHA T.K.; GULATI, P.; MOHANTY, A.; KHASA, Y.P.; KAPOOR, R.K.; KUMAR, S. Efficient genetic approaches for improvement of plasmid based expression

- of recombinant protein in *Escherichia coli*: A review. **Process Biochemistry**, v. 55, p. 17-31, 2017.
- SØGAARD, K.M.; NØRHOLM, M.H.H. Side effects of extra tRNA supplied in a typical bacterial protein production scenario. **Protein Science**, v. 25, p. 2102-2108, 2016.
- SONNENDECKER, C.; WEI, R.; KURZE, E.; WANG, J.; OESER, T.; ZIMMERMANN, W. Efficient extracellular recombinant production and purification of a *Bacillus cyclodextrin* glucanotransferase in *Escherichia coli*. **Microbial Cell Factories**, 16:87, 2017.
- SØRENSEN, H.P.; MORTENSEN, K.K. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. **Journal of Biotechnology**, v. 115, p. 113-128, 2005.
- STUDIER, F.W. Protein production by auto-induction in high-density shaking cultures. **Protein Expression and Purification**, v. 41, p. 207-234, 2005.
- SUÁREZ, D.C.; LIRIA, C.W.; KILIKIAN, B.V. Effect of yeast extract on *Escherichia coli* growth and acetic acid production. **World Journal of Microbiology & Biotechnology**, v. 14, p. 331-335, 1998.
- SUN, W.; JI, W.; HALL, J.M.; HU, Q.; WANG, C.; BEISEL, C.L.; GU, Z. Self-Assembled DNA nanoclews for the efficient delivery of CRISPR-Cas9 for genome editing. **Angewandte Zueschriften**, v. 127, p. 12197-12201, 2015.
- TALLET B.; ASTIER-GIN, T.; CASTROVIEJO, M.; SANTARELLI, X. One-step chromatographic purification procedure of a His-tag recombinant carboxyl half part of the HTLV-I surface envelope glycoprotein overexpressed in *Escherichia coli* as a secreted form. **Journal of Chromatography B**, v. 753, p. 17-22, 2001.
- TAO, H.; BAUSCH, C.; RICHMOND, C.; BLATTNER, F.R.; CONWAY, T. Functional genomics: expression anlysis of *Escherichia coli* growing on minimal and rich media. **Journal of Bacteriology**, v. 181, n. 20, p. 6425-6440, 1999.
- TEGEL, H.; TOURLE, S.; OTTOSSON, J.; PERSSON, A. Increased levels of recombinant human proteins with the *Escherichia coli* strain Rosetta(DE3). **Protein Expression and Purification**, v. 69, p. 156-167, 2010.

- TEGEL, H.; OTTOSSON, J.; HOBER, S. Enhancing the protein production levels in *Escherichia coli* with a strong promoter. **The FEBS Journal**, v. 278, p. 729-739, 2011.
- TERPE, K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. **Applied Microbiology and Biotechnology**, v. 60, p. 523-533, 2003.
- TIMIM, A.S.; MUSLIMOV, A.R.; LEPIK, K.V.; EPIFANOVSKAYA, O.S.; SHAKIROVA, A.I.; MOCK, U.; RIECKEN, K.; OKILOVA, M.V.; SERGEEV, V.S.; AFANASYEV, B.V.; FEHSE, B.; SUKHORUKOV, G.B. Efficient gene editing via non-viral delivery of CRISPR-Cas9 system using polymeric and hybrid microcarriers. **Nanomedicine: Nanotechnology, Biology, and Medicine**, v. 14, p. 97-108, 2018.
- TODOROVA-BALVAY, D.; PITIOT, O.; BOURHIM, M.; SRIKRISHNAN, T.; VIJAYALAKSHMI, M. Immobilized metal-ion affinity chromatography of human antibodies and their proteolytic fragments. **Journal of chromatography B**, v. 808, p. 57-62, 2004.
- TOMAZETTO, G.; MULINARI, F.; STANISÇUASKI, F.; SETTEMBRINI, B.; CARLINI, C.R.; AYUB, M.A.Z. Expression kinetics and plasmid stability of recombinant *E. coli* encoding urease-derived peptide with bioinsecticide activity. **Enzyme and Microbial Technology**, v. 41, p. 821-827, 2007.
- TOWBIN, H.; STAEBELIN, T.; GORDON, J. Electrophoretic transfer of proteins from poly-acrylamide gels to nitrocellulose sheets: procedure and some applications, **Proceedings of the National Academy of Sciences**, v. 76, p. 4350-4354, 1979.
- TRIPATHI, N.K. Production and purification of recombinant proteins from *escherichia coli*. **ChemBioEng Reviews**, v. 3, n. 3, p. 116-133, 2016.
- TURNER, C.; GREGORY, M.E.; TURNER, M.K. A study of the effect of specific growth rate and acetate on recombinant protein production of *Escherichia coli* JM107. **Biotechnology Letters**, v. 16, n. 9, p. 891-896, 1994.
- VANÇAN, S.; MIRANDA, E.A.; BUENO, S.M.A. IMAC of human IgG: studies with IDA-immobilized copper, nickel, zinc, and cobalt ions and different buffer systems. **Process Biochemistry**, v. 37, p. 573-579, 2002.
- VOLONTÈ, F.; MARINELLI, F.; GASTALDO, L.; SACCHI, S.; PILONE, M.S.; POLLEGIONI, L.; MOLLA, G. Optimization of glutaryl-7-aminocephalosporanic

- acid acylase expression in *E. coli*. **Protein Expression and Purification**, v. 61, p. 131-137, 2008.
- WANG, H.-X.; LI, M.; LEE, C.M.; CHAKRABORTY, S.; KIM, H.-W.; BAO, G.; LEONG, K.W. CRISPR/Cas9-Based genome editing for disease modeling and therapy: challenges and opportunities for nonviral delivery. **Chemical Reviews**, v. 117, p. 9874-9906, 2017.
- WANG, Q.; COBINE, P.A.; COLEMAN, J.J. Efficient genome editing in *Fusarium oxysporum* based on CRISPR/Cas9 ribonucleoprotein complexes. **Fungal Genetics and Biology**, v. 117, p. 21-29, 2018.
- WENINGER, A.; HATZL, A.-M.; SCHMID, C.; VOGL, T.; GLIEDER, A. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. **Journal of Biotechnology**, v. 235, p. 139-149, 2016.
- WU, C.-A.M.; ROTH, T.L.; BAGLAENKO, Y.; FERRI, D.M.; BRAUER, P.; ZUNIGA-PFLUCKER, J.C.; ROSBE, K.W.; WITHER, J.E.; MARSON, A.; ALLEN, C.D.C. Genetic engineering in primary human B cells with CRISPR-Cas9 ribonucleoproteins. **Journal of Immunological Methods**, v. 457, p. 33-40, 2018.
- YAN, Q.-L.; KONG, H.-T.; XIA, K.; ZHANG, Y.; ALADLBAHI, A.; SHI, J.-Y.; WANG, L.-H.; FAN, C.-H.; ZHAO, Y.; ZHU, Y. Expression and radiolabeling of Cas9 protein. **Nuclear Science and Techniques**, 28:11, (2017).
- YANG, S.; CHANG, R.; YANG, H.; ZHAO, T.; HONG, Y.; KONG, H.E.; SUN, X.; QIN, Z.; JIN, P.; LI, S.; LI, X.-J. CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington's disease. **The Journal of Clinical Investigation**, v. 127, n. 7, p. 2719-2724, 2017.
- YEE, L.; BLANCH, H.W. Recombinant protein expression in high cell density fed-batch cultures of *Escherichia coli*. **Nature Biotechnology**, v. 10, 1992.
- ZAWADA, J.; RICHTER, B.; HUANG, E.; LODES, E.; SHAH, A.; SWARTZ, J.R. High-density, defined media culture for the production of *Escherichia coli* cell extracts. In: **Fermentation Biotechnology**, October 7, p. 142-156, 2003.
- ZHOU, Z.; DANG, Y.; ZHOU, M.; LI, L.; YU, C.-H.; FU, J.; CHEN, S.; LIU, Y. Codon usage is an important determinant of gene expression levels largely through its effects on transcription. **PNAS**, v. 113, n. 41, E6117-E6125, 2016.

- ZHU, W.; LEI, R.; DUFF, Y.L.; LI, J.; GUI, F.; WAINBERG, M.A.; LIANG, C. The CRISPR/Cas9 system inactivates latent HIV-1 proviral DNA. **Retrovirology**, 12:22, 2015.
- ZOU, C.; DUAN, X.; WU, J. Enhanced extracellular production of recombinant *Bacillus deramificans* pullulanase in *Escherichia coli* through induction mode optimization and a glycine feeding strategy. **Bioresource Technology**, v. 172, p. 174-179, 2014.
- ZURIS, J.A.; THOMPSON, D.B.; SHU, Y.; GUILINGER, J.P.; BESSEN, J.L.; HU, J.H.; MAEDER, M.L.; JOUNG, J.K.; CHEN, Z.Y.; LIU, D.R. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing *in vitro* and *in vivo*. **Nature Biotechnology**, v. 33, n. 1, 2015.

APÊNDICE A

Curvas de calibração para conversão de absorbância (600nm) em massa seca (g/L) conforme item 3.3.1 para os meios LB e HDF.

Figura A.1. Curva de calibração para conversão de absorbância (600nm) em massa seca (g/L) para *E. coli* BL21(DE3) a 37°C e 250rpm em (a) meio LB e (b) meio HDF.

APÊNDICE B

Densitometria do gel de eletroforese SDS-PAGE para estimar o valor de produção da proteína Cas9 por litro de cultivo. Foi utilizado o software ImageJ. A seguir estão apresentados os valores obtidos para as bandas. Como base de cálculos foi utilizado o valor da área calculada pelo ImageJ para a banda de 50 kDa presente no padrão de eletroforese. 10 μ L de padrão foram aplicados ao gel, portanto, segundo fornecedor, 750 ng da proteína 50 kDa;

1- Valores de densitometria obtidos para os ensaios com e sem PMSF (item 4.2)

Tabela B.1. Valores de área obtidos para: (1) padrão correspondente a banda de 50 kDa (750 ng de proteína) do padrão de eletroforese, (2) sem PMSF, (3) com PMSF.

	Area1	Area2
1	4.612	4.864
2	13.520	17.036
3	10.609	14.054

(a)

(b)

Figura B.1. (a) Gel de SDS-PAGE analisado por densitometria utilizando programa ImageJ. M: marcador de massa molecular; 1: *pellet* sem indução sem PMSF; 2: amostra de eluição sem PMSF; 3: *pellet* sem indução com PMSF; 4: amostra de eluição com PMSF. (b) Imagem gerada pelo programa para determinar a área correspondente às bandas.

2- Valores de densitometria obtidos para os ensaios em meio LB e HDF (item 4.5.2).

Tabela B.2. Valores de área obtidos para: (1) padrão correspondente a banda de 50 kDa (750 ng de proteína) do padrão de eletroforese, (2) meio LB sem indução, (3) e (4) meio LB com indução em duplicata, (5) meio HDF sem indução, (6) e (7) meio HDF com indução em duplicata.

	Area1	Area2
1	5.778	5.000
2	1.009	956
3	6.071	5.586
4	6.353	5.912
5	1.256	1.086
6	6.721	6.375
7	7.420	7.051

(a)

(b)

Figura B.2. (a) Gel de SDS-PAGE analisado por densitometria utilizando programa ImageJ. M: marcador de massa molecular; Cas9: proteína Cas9 purificada; 1: *pellet* em meio LB sem indução; 2 e 3: *pellet* em meio LB com indução; 4: *pellet* em meio HDF sem indução; 5 e 6: *pellet* em meio HDF com indução. (b) Imagem gerada pelo programa para determinar a área correspondente às bandas.

3- Valores de densitometria obtidos para os ensaios em diferentes temperaturas (18, 25, 30 e 37°C) (item 4.5.3).

Tabela B.3. Valores de área obtidos para: (1) padrão correspondente a banda de 50 kDa (750 ng de proteína) do padrão de eletroforese, (2) e (6) 18°C, (3) e (7) 25°C, (4) e (8) 30°C, (5) e (9) 37°C.

	Area1	Area2
1	5.766	5.636
2	3.739	3.686
3	5.966	6.031
4	8.136	7.610
5	4.303	4.404
6	3.394	3.390
7	5.773	5.564
8	7.723	7.502
9	4.513	4.668

(a)

(b)

Figura B.3. (a) Gel de SDS-PAGE analisado por densitometria utilizando programa ImageJ. M: marcador de massa molecular; Cas9: proteína Cas9 purificada; 1 e 5: *pellet* com indução a 18°C; 2 e 6: *pellet* com indução a 25°C; 3 e 7: *pellet* com indução a 30°C; 4 e 8: *pellet* com indução a 37°C. (b) Imagem gerada pelo programa para determinar a área correspondente às bandas.

4- Valores de densitometria obtidos para os ensaios em diferentes tempos de pós-indução (0, 2, 4, 8, 18 e 24h) (item 4.5.3).

Tabela B.4. Valores de área obtidos para: (1) padrão correspondente a banda de 50 kDa (750 ng de proteína) do padrão de eletroforese, (2) 0h, (3) 2h, (4) 4h, (5) 6h, (6) 8h, (7) 18h, (8) 24h.

	Area1	Area2
1	5.873	5.615
2	964	997
3	5.174	4.922
4	8.573	8,944
5	9.394	9.853
6	10.196	10.677
7	7.671	7.707
8	5.745	5.981

(a)

(b)

Figura B.4. (a) Gel de SDS-PAGE analisado por densitometria utilizando programa ImageJ. M: marcador de massa molecular; Cas9: proteína Cas9 purificada; 1: *pellet* sem indução; 2: *pellet* após 2h de indução; 3: *pellet* após 4h de indução; 4: *pellet* após 6h de indução; 5: *pellet* após 8h de indução; 6: *pellet* após 18h de indução; 7: *pellet* após 24h de indução. (b) Imagem gerada pelo programa para determinar a área correspondente às bandas.

5- Valores de densitometria obtidos para os ensaios em biorreatores (item 4.5.5).

Tabela B.5. Valores de área obtidos para: (1) padrão correspondente a banda de 50 kDa (750 ng de proteína) do padrão de eletroforese, (2) indução meio fase exponencial, sem indução, (3) indução meio fase exponencial, 8h de indução, (4) indução final fase exponencial, sem indução, (5) indução final fase exponencial, 4h de indução, (6) indução final fase exponencial, 6h de indução, (7) indução final fase exponencial, 8h de indução.

	Area1	Area2
1	3.754	3.378
2	768	677
3	10.449	10.011
4	468	493
5	5.256	5.301
6	8.079	8.409
7	6.127	6.584

(a)

(b)

Figura B.5. (a) Gel de SDS-PAGE analisado por densitometria utilizando programa ImageJ. M: marcador de massa molecular; 1: *pellet* sem indução; 2: *pellet* após 8h de indução no meio da fase exponencial em biorreator; 3: *pellet* sem indução; 4: *pellet* após 4h de indução no final da fase exponencial em biorreator; 5: *pellet* após 6h de indução no final da fase exponencial em biorreator; 6: *pellet* após 8h de indução no final da fase exponencial em biorreator (b) Imagem gerada pelo programa para determinar a área correspondente às bandas.

6- Valores de densitometria obtidos para o ensaio de purificação (item 4.6).
Algumas amostras foram diluídas para uma melhor estimativa.

Tabela B.6. Valores de área obtidos para: (1) padrão correspondente a banda de 50 kDa (750 ng de proteína) do padrão de eletroforese, (2) Injeção, (3) Flowthrough, (4) Lavagem, (5) Eluição 1 (diluída 5x), (6) Eluição 2 (diluída 2x), (7) Eluição 3, (8) Eluição 4, (9) Eluição 5.

	Area1	Area2
1	3.754	3.378
2	4.404	3.660
3	1.488	1.417
4	1.825	1.712
5	3.682	3.446
6	3.887	3.634
7	575	444
8	63	42
9	0	0

Figura B.6. Gel de SDS-PAGE analisado por densitometria utilizando programa ImageJ. M: marcador de massa molecular; 1: *pellet* sem indução; 2: *pellet* com indução; 3: Injeção; 4: Flowthrough; 5: Lavagem; 6: Eluição 1; 7: Eluição 2; 8: Eluição 3; 9: Eluição 4; 10: Eluição 5.

7- Valores de densitometria obtidos para determinar a pureza no ensaio de purificação (item 4.6).

Tabela B.7. Valores de área obtidos para: (1) padrão correspondente a banda de 50 kDa (750 ng de proteína) do padrão de eletroforese, (2) banda da proteína Cas9, (3) (4) e (5) demais bandas presentes no gel.

	Area1	Area2
1	3.754	3.378
2	9.375	7.535
3	61	49
4	95	52
5	180	103

(a)

(b)

Figura B.7. (a) Gel de SDS-PAGE analisado por densitometria utilizando programa ImageJ. Fração 1 da etapa de eluição. (b) Imagem gerada pelo programa para determinar a área correspondente às bandas.

APÊNDICE C

Resultado da espectrometria de massa (MALDI-TOF MS e MSMS) de amostra de proteína digerida por tripsina.

[gi1489080018](#) Mass: 158519 Score: 108 Matches: 3(0) Sequences: 3(0)
type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes]

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
9	1022.5947	1021.5874	1021.4618	0.1256	0	45	6	1	U	K.VDDSFHR.L
12	1038.6900	1037.6827	1037.5618	0.1209	0	28	2.6e+002	2	U	K.HVAQILDSR.L
18	1148.7904	1147.7831	1147.6390	0.1442	0	35	61	1	U	R.IPYVGLAR

Figura C.1. Espectros de MS e MSMS da amostra de proteína digerida por tripsina e resultado da identificação da proteína utilizando o software MASCOT empregando os parâmetros: Taxonomia: Todas as entradas; Tolerância MS: 0,3Da; Tolerância MSMS: 0,6 Da; Base de dados: NCBI nr.

APÊNDICE D

Sequência de genes da proteína Cas9 (<https://www.addgene.org/47327/>).

ATG GAC AAG AAG TAC TCC ATT GGG CTC GAT ATC GGC ACA AAC AGC GTC GGC
TGG GCC GTC ATT ACG GAC GAG TAC AAG GTG CCG AGC AAA AAA TTC AAA GTT
CTG GGC AAT ACC GAT CGC CAC AGC AUA AAG AAG AAC CTC ATT GGC GCC CTC
CTG TTC GAC TCC GGG GAG ACG GCC GAA GCC ACG CGG CTC AAA AGA ACA
GCA CGG CGC AGA TAT ACC CGC AGA AAG AAT CGG ATC TGC TAC CTG CAG GAG
ATC TTT AGT AAT GAG ATG GCT AAG GTG GAT GAC TCT TTC TTC CAT AGG CTG
GAG GAG TCC TTT TTG GTG GAG GAG GAT AAA AAG CAC GAG CGC CAC CCA ATC
TTT GGC AAT ATC GTG GAC GAG GTG GCG TAC CAT GAA AAG TAC CCA
ACC AUA TAT CAT CTG AGG AAG AAG CTT GTA GAC AGT ACT GAT AAG GCT GAC
TTG CGG TTG ATC TAT CTC GCG CTG GCG CAT ATG ATC AAA TTT CGG GGA CAC
TTC CTC ATC GAG GGG GAC CTG AAC CCA GAC AAC AGC GAT GTC GAC AAA CTC
TTT ATC CAA CTG GTT CAG ACT TAC AAT CAG CTT TTC GAA GAG AAC CCG ATC
AAC GCA TCC GGAGTT GAC GCC AAA GCA ATC CTG AGC GCT AGG CTG TCC AAA
TCC CGG CGG CTC GAA AAC CTC ATC GCA CAG CTC CCT GGG GAG AAG AAG AAC
GGC CTG TTT GGT AAT CTT ATC GCC CTG TCA CTC GGG CTG ACC CCC AAC TTT
AAA TCT AAC TTC GAC CTG GCC GAA GAT GCC AAG CTT CAA CTG AGC AAA GAC
ACC TAC GAT GAT GAT CTC GAC AAT CTG CTG GCC CAG ATC GGC GAC CAG TAC
GCA GAC CTT TTT TTG GCG GCA AAG AAC CTG TCA GAC GCC ATT CTG CTG AGT
GAT ATT CTG CGA GTG AAC ACG GAG ATC ACC AAA GCT CCG CTG AGC GCT AGT
ATG ATC AAG CGC TAT GAT GAG CAC CAC CAA GAC TTG ACT TTG CTG AAG GCC
CTT GTC AGA CAG CAA CTG CCT GAG AAG TAC AAG GAA ATT TTC TTC GAT CAG
TCT AAA AAT GGC TAC GCC GGA TAC ATT GAC GGC GGA GCA AGC CAG GAG GAA
TTT TAC AAA TTT ATT AAG CCC ATC TTG GAA AAA ATG GAC GGC ACC GAG GAG
CTG CTG GTA AAG CTT AAC AGA GAA GAT CTG TTG CGC AAA CAG CGC ACT TTC
GAC AAT GGA AGC ATC CCC CAC CAG ATT CAC CTG GGC GAA CTG CAC GCT ATC
CTC AGG CGG CAA GAG GAT TTC TAC CCC TTT TTG AAA GAT AAC AGG GAA AAG
ATT GAG AAA ATC CTC ACA TTT CGG AUA CCC TAC TAT GTA GGC CCC CTC
GCC CGG GGA AAT TCC AGA TTC GCG TGG ATG ACT CGC AAA TCA GAA GAG ACC
ATC ACT CCC TGG AAC TTC GAG GAA GTC GTG GAT AAG GGG GCC TCT GCC CAG
TCC TTC ATC GAA AGG ATG ACT AAC TTT GAT AAA AAT CTG CCT AAC GAA AAG
GTG CTT CCT AAA CAC TCT CTG CTG TAC GAG TAC TTC ACA GTT TAT AAC GAG
CTC ACC AAG GTC AAA TAC GTC ACA GAA GGG ATG AGA AAG CCA GCA TTC CTG
TCT GGA GAG CAG AAG AAA GCT ATC GTG GAC CTC CTC TTC
AAG ACG AAC CGG AAA GTT ACC GTG AAA CAG CTC AAA GAA GAC TAT TTC AAA
AAG ATT GAA TGT TTC GAC TCT GTT GAA ATC AGC GGA GTG GAG GAT CGC TTC
AAC GCA TCC CTG GGA ACG TAT CAC GAT CTC CTG AAA ATC ATT AAA GAC AAG
GAC TTC CTG GAC AAT GAG GAG AAC GAG GAC ATT CTT GAG GAC ATT GTC CTC
ACC CTT ACG TTG TTT GAA GAT AGG GAG ATG ATT GAA GAA CGC TTG AAA ACT
TAC GCT CAT CTC TTC GAC GAC AAA GTC ATG AAA CAG CTC
AAG AGG CGC CGA TAT ACA GGA TGG GGG CGG CTG TCA AGA AAA CTG ATC

AAT GGG ATC CGA GAC AAG CAG AGT GGA AAG ACA ATC CTG GAT TTT CTT AAG
TCC GAT GGA TTT GCC AAC CGG AAC TTC ATG CAG TTG ATC CAT GAT GAC TCT
CTC ACC TTT AAG GAG GAC ATC CAG AAA GCA CAA GTT TCT GGC CAG GGG GAC
AGT CTT CAC GAG CAC ATC GCT AAT CTT GCA GGT AGC CCA GCT ATC AAA
AAG GGA AUA CTG CAG ACC GTT AAG GTC GTG GAT GAA CTC GTC AAA GTA
ATG GGA AGG CAT AAG CCC GAG AAT ATC GTT ATC GAG ATG GCC CGA GAG AAC
CAA ACT ACC CAG AAG GGA CAG AAG AAC AGT AGG GAA AGGATG AAG AGG ATT
GAA GAG GGT AUA AAA GAA CTG GGG TCC CAA ATC CTT AAG GAA CAC CCA GTT
GAA AAC ACC CAG CTT CAG AAT GAG AAG CTC TAC CTG TAC TAC CTG CAG AAC
GGC AGG GAC ATG TAC GTG GAT CAG GAA CTG GAC ATC AAT CGG CTC TCC GAC
TAC GAC GTG GAT CAT ATC GTG CCC CAG TCT TTT CTC AAA GAT GAT TCT ATT
GAT AAT AAA GTG TTG ACA AGA TCC GAT AAA AAT AGA GGG AAG AGT GAT AAC
GTC CCC TCA GAA GAA GTT GTC AAG AAA ATG AAA AAT TAT TGG CGG CAG CTG
CTG AAC GCC AAA CTG ATC ACA CAA CGG AAG TTC GAT AAT CTG ACT AAG GCT
GAA CGA GGT GGC CTG TCT GAG TTG GAT AAA GCC GGC TTC ATC AAA AGG CAG
CTT GTT GAG ACA CGC CAG ATC ACC AAG CAC GTG GCC CAA ATT CTC GAT TCA
CGC ATG AAC ACC AAG TAC GAT GAA AAT GAC AAA CTG ATT CGA GAG GTG AAA
GTT ATT ACT CTG AAG TCT AAG CTG GTC TCA GAT TTC AGA AAG GAC TTT CAG
TTT TAT AAG GTG AGA GAG ATC AAC AAT TAC CAC CAT GCG CAT GAT GCC TAC
CTG AAT GCA GTG GTA GGC ACT GCA CTT ATC AAA AAA TAT CCC AAG CTT GAA
TCT GAA TTT GTT TAC GGA GAC TAT AAA GTG TAC GAT GTT AGG AAA ATG ATC
GCA AAG TCT GAG CAG GAA AUA GGC AAG GCC ACC GCT AAG TAC TTC TTT TAC
AGC AAT ATT ATG AAT TTT TTC AAG ACC GAG ATT ACA CTG GCC AAT GGA GAG
ATT CGG AAG CGA CCA CTT ATC GAA ACA AAC GGA GAA ACA GGA GAA ATC GTG
TGG GAC AAG GGT AGG GAT TTC GCG ACA GTC CGG AAG GTC CTG TCC ATG CCG
CAG GTG AAC ATC GTT AAA AAG ACC GAA GTA CAG ACC GGA GGC TTC TCC AAG
GAA AGT ATC CTC CCG AAA AGG AAC AGC GAC AAG CTG ATC GCA CGC AAA AAA
GAT TGG GAC CCC AAG AAA TAC GGC GGA TTC GAT TCT CCT ACA GTC GCT TAC
AGT GTA CTG GTT GTG GCC AAA GTG GAG AAA GGG AAG TCT AAA AAA CTC AAA
AGC GTC AAG GAA CTG CTG GGC ATC ACA ATC ATG GAG CGA TCA AGC TTC GAA
AAA AAC CCC ATC GAC TTT CTC GAG GCG AAA GGA TAT AAA GAG GTC AAA AAA
GAC CTC ATC ATT AAG CTT CCC AAG TAC TCT CTC TTT GAG CTT GAA AAC
GGC CGG AAA CGA ATG CTC GCT AGT GCG GGC GAG CTG CAG AAA GGT AAC GAG
CTG GCA CTG CCC TCT AAA TAC GTT AAT TTC TTG TAT CTG GCC AGC CAC TAT
GAA AAG CTC AAA GGG TCT CCC GAA GAT AAT GAG CAG AAG CAG CTG TTC GTG
GAA CAA CAC AAA CAC TAC CTT GAT GAG ATC ATC GAG CAA AUA AGC GAA TTC
TCC AAA AGA GTG ATC CTC GCC GAC GCT AAC CTC GAT AAG GTG CTT TCT GCT
TAC AAT AAG CAC AGG GAT AAG CCC ATC AGG GAG CAG GCA GAA AAC ATT ATC
CAC TTG TTT ACT CTG ACC AAC TTG GGC GCG CCT GCA GCC TTC AAG TAC TTC
GAC ACC ACC AUA GAC AGA AAG CGG TAC ACC TCT ACA AAG GAG GTC CTG GAC
GCC ACA CTG ATT CAT CAG TCA ATT ACG GGG CTC TAT GAA ACA AGA ATC GAC
CTC TCT CAG CTC GGT GGA GAC AGC AGG GCT GAC CCC AAG AAG AAG AGG AAG
GTG GCG GCC GCA CTC GAG CAC CAC CAC CAC CAC

Tabela D.1. Códonos raros para *E. coli* presentes na sequência do gene da proteína Cas9 (número de códonos na sequência: 1390)
(<http://people.mbi.ucla.edu/sumchan/caltor.html>)

Aminoácidos	Códonos raros	Total de códonos na sequência da Cas9
Arginina	CGA	9
	CGG	20
	AGG	21
	AGA	15
Glicina	GGA	23
	GGG	15
Isoleucina	AUA	8
Leucina	CUA	0
Prolina	CCC	18
Treonina	ACG	8

APÊNDICE E

Géis de agarose 1% para avaliar a eficiência da etapa de tratamento com DNase das amostras de RNA após a extração (Figura D.1) e para verificar a especificidade dos *primers* 2 e 3 (Tabela 3.1) na amplificação do fragmento de 122 pb do gene da proteína Cas9 por PCR convencional (Figura D.2).

Figura E.1. Gel de agarose 1% para avaliar as amostras de RNA após o tratamento com DNase. 1- GeneRuler 1 kb Plus DNA Ladder (Thermo Fisher Scientific, EUA); 2- amostra da BL21(DE3) com pET sem indução; 3- amostra da BL21(DE3) com pET com indução; 4- amostra da BL21(DE3) sem pET sem indução; 5- amostra da BL21(DE3) sem pET com indução; 6- amostra da BL21(DE3) Rosetta com pET sem indução; 7- amostra da BL21(DE3) Rosetta com pET com indução; 8- amostra da BL21(DE3) Rosetta sem pET sem indução; 9- amostra da BL21(DE3) Rosetta sem pET com indução; 10-17: amostras da PCR das amostras de 2-9 com os *primers* 2 e 3 (Tabela 3.1) para amplificação do fragmento do gene da proteína Cas9 de 122 pb.

Figura E.2. Gel de agarose 1% para avaliar as amostras de cDNA após a reação de transcrição reversa. 1- GeneRuler 1 kb Plus DNA Ladder (Thermo Fisher Scientific, EUA); 2- amostra da BL21(DE3) com pET sem indução; 3- amostra da BL21(DE3) com pET com indução; 4- amostra da BL21(DE3) sem pET sem indução; 5- amostra da BL21(DE3) sem pET com indução; 6- amostra da BL21(DE3) Rosetta com pET sem indução; 7- amostra da BL21(DE3) Rosetta com pET com indução; 8- amostra da BL21(DE3) Rosetta sem pET sem indução; 9- amostra da BL21(DE3) Rosetta sem pET com indução; 10-17: amostras da PCR das amostras de 2-9 com os *primers* 2 e 3 (Tabela 3.1) para amplificação do fragmento do gene da proteína Cas9 de 122 pb.