UNIVERSIDADE DE SÃO PAULO
FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

**Expressão do fator estimulador de colônia de granulócito humano
recombinante (rhG-CSF) em sistema *Pichia pastoris* e otimização do
bioprocesso**

Bruna Samhan Archangelo

Ribeirão Preto

2017

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

Expressão do fator estimulador de colônia de granulócito humano
recombinante (rhG-CSF) em sistema *Pichia pastoris* e otimização do
bioprocesso

Dissertação de Mestrado apresentada ao
programa de Pós-Graduação em
Biociências Aplicadas à Farmácia para
obtenção do Título de Mestre em
Ciências.

Área de concentração: Biociências
Aplicadas à Farmácia

Orientada: Bruna Samhan Archangelo

Orientadora: Prof^a. Dr^a. Elisa Maria de
Sousa Russo

Ribeirão Preto – SP

2017

<p>ARCHANGELO, B.S.</p>	<p>Expressão do fator estimulador de colônia de granulócito humano recombinante (rhG-CSF) em sistema <i>Pichia pastoris</i> e otimização do bioprocesso.</p>		<p>MESTRADO FCFRPUSP</p> <p>2017</p>
------------------------------------	--	--	--

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.

Archangelo, Bruna Samhan

Expressão do fator estimulador de colônia de granulócito humano recombinante (rhG-CSF) em sistema *Pichia pastoris* e otimização do bioprocesso. Ribeirão Preto, 2017.

120p.: il.; 30cm

Dissertação de Mestrado apresentada à Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP – Área de concentração: Biociências Aplicadas à Farmácia.

Orientadora: Russo, Elisa Maria de Sousa

1. GCSF. 2. *Pichia pastoris*. 3. Proteína recombinante. 4. Bioprocesso

FOLHA DE APROVAÇÃO

Bruna Samhan Archangelo

Expressão do fator estimulador de colônia de granulócito humano recombinante (rhG-CSF) em sistema *Pichia pastoris* e otimização do bioprocesso

Dissertação de Mestrado apresentada ao programa de Pós-Graduação em Biociências Aplicadas à Farmácia para obtenção do Título de Mestre em Ciências.

Área de concentração: Biociências Aplicadas à Farmácia

Orientadora: Prof^a. Dr^a. Elisa Maria de Sousa Russo

Aprovado em: ____/____/____.

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Apoio financeiro:

Comissão de Aperfeiçoamento de Pessoal do Nível Superior (CAPES)

(CAPES – bolsa)

Dedicatória

Dedico este trabalho primeiramente a Deus e depois às pessoas mais importantes da minha vida: meu pai, minha mãe e meu irmão, pelo apoio incondicional em todas as etapas da minha vida.

Agradecimientos

Agradeço primeiramente a Deus, por ser minha força e meu auxílio em todos os momentos. Pela sabedoria e discernimento que me concede ao enfrentar desafios novos. Sem ele, nada seria possível;

À minha Mãe, Lamea Issa Samhan Archangelo, pela minha formação pessoal, por todo apoio recebido, por ser meu exemplo, pelo amor incondicional, pelas orações recebidas e por toda dedicação a mim em toda a vida;

Ao meu Pai, Luis Emilio Archangelo, que é o melhor homem que eu já conheci. Agradeço por toda a compreensão a mim dada, pelo carinho e amor incondicional que me dá, por ser meu exemplo, meu herói;

Ao meu Irmão, Victor Samhan Archangelo, que me faz ser uma pessoa melhor a cada dia, meu amigo e companheiro que mesmo longe está sempre tão perto;

Ao meu namorado, Rafael Pedezzi, minha melhor companhia. Agradeço por toda ajuda, companheirismo, amor, dedicação e cuidado. Sem ele, não teria chegado até aqui;

À Profa. Dra. Elisa Maria de Sousa Russo, pela dedicação, paciência, conhecimento e por ter me aberto portas para a realização de um sonho;

Aos meus amigos de laboratório, Aline Sanches Pereira, Carlos Eduardo Mendonça, Gabriel Antônio Nogueira, Gisele Lelis da Silva, Nathália Cristina Canicoba e Leandro Sobrinho Ávila por toda paciência, amizade, suporte e ajuda no laboratório;

Aos funcionários do Departamento de Análises Clínicas, Toxicológicas e Bromatológicas da Faculdade de Ciências Farmacêuticas de Ribeirão Preto (FCFRP)-USP, ao químico Luiz Fernando Fortunato Tucci por toda ajuda oferecida;

Ao Prof. Dr. Hamilton Cabral, pela colaboração com este trabalho e por sempre disponibilizar seu laboratório;

Ao Prof. Dr. Luis Alexandre de Pedro Freitas, da Faculdade de Ciências Farmacêuticas de Ribeirão Preto, pela colaboração com este trabalho e por toda ajuda;

À Profa. Dra. Cleni Mara Marzocchi Machado, pela convivência e, principalmente, pelos momentos compartilhados;

A toda a minha família, pelo amor, pelo companheirismo, por sempre torcerem muito por mim, por toda essa vida que caminhamos juntos;

Aos meus amigos, que me fizeram perceber que família pode ser escolhida;

Aos professores do Programa de Biociências Aplicadas à Farmácia da FCFRP-USP por auxiliarem meu aprendizado e crescimento acadêmico;

Aos coordenadores e funcionários do Programa de Biociências Aplicadas à Farmácia da FCFRP-USP pela disposição em ajudar sempre;

À Faculdade de Ciências Farmacêuticas de Ribeirão Preto-USP, pela oportunidade de realização do curso de Mestrado;

A Comissão de Aperfeiçoamento de Pessoal do Nível Superior (CAPES) pela concessão da bolsa de mestrado;

A todos aqueles que direta ou indiretamente contribuíram para a realização deste trabalho.

Epígrafe

“O Senhor é a minha luz e a minha salvação; a quem temerei? O Senhor é a força da minha vida; de quem me recearei?”

Salmos 27:1

“Que os vossos esforços desafiem as impossibilidades, lembrai-vos de que as grandes coisas do homem foram conquistadas do que parecia impossível.”

Charles Chaplin

RESUMO

ARCHANGELO, B. S. **Expressão do fator estimulador de colônia de granulócito humano recombinante (rhG-CSF) em sistema de *Pichia pastoris* e otimização do bioprocesso.** 2017. 120f. Dissertação (Mestrado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2017.

O fator estimulador de colônias de granulócito humano (hG-CSF) atua principalmente promovendo a maturação dos neutrófilos. Tem sido utilizado no tratamento de diversas patologias, sobretudo em neutropenias provocadas por eventos que suprimem a produção de células mieloides. No Brasil, este medicamento está inserido no Programa de Assistência Farmacêutica do SUS como medicamento excepcional, porém a um altíssimo custo para aquisição, pois ainda não é produzido no Brasil. Um estudo de viabilidade técnico-econômica da implantação da produção de rhG-CSF no Brasil demonstrou que esse medicamento pode ser produzido por custos até cerca de 90% inferiores aos valores pagos em compras governamentais. Este trabalho mostra a clonagem e expressão do rhG-CSF em sistema heterólogo de *P. pastori* e visou estabelecer os melhores parâmetros para a indução e produção da mesma em diferentes condições de cultivo, além de uma caracterização da linhagem produtora. A construção do plasmídeo de expressão foi feita com o vetor pPICZαA e confirmada por sequenciamento nucleotídico. Após a transformação em *P. pastoris*, algumas colônias foram testadas para avaliar as condições da expressão da proteína recombinante. A colônia que apresentou uma expressão melhor foi escolhida para a produção da proteína recombinante em uma escala maior. A detecção da proteína recombinante foi realizada por SDS-PAGE, *Western Blotting* e *Dot Blotting*. O planejamento Plackett-Burman foi aplicado para selecionar as variáveis significativas, que foram o pH e a adição de Sorbitol e aminoácidos no meio de cultura. Os dados obtidos passaram por análise estatística que permitiu esta seleção. O Delineamento Composto Central foi aplicado para uma análise em níveis maiores das variáveis significativas. Também foi feita uma caracterização de linhagem para que a colônia selecionada fosse caracterizada como uma linhagem produtora. Apesar das três variáveis terem sido significativas (pH, sorbitol e aminoácidos), apenas o aumento do valor de pH apresentou um efeito benéfico para a produção da rhG-CSF. Curvas de crescimento celular foram desenhadas, além do cálculo da velocidade máxima de crescimento e tempo de duplicação. Estes resultados mostram a produção da rhG-CSF em *P. pastoris* e estes são os primeiros passos para uma estratégia de tornar o Brasil um produtor deste biofármaco futuramente, além de contribuir para estudos de otimização de bioprocessos.

Palavras-chave: G-CSF, *Pichia pastoris*, *Screening design*, delineamento experimental, Plackett-Burman, Delineamento Composto Central.

ABSTRACT

ARCHANGELO, B. S. **Expression of human granulocyte colony stimulating factor (rhG-CSF) in *Pichia pastoris* and the optimization of the bioprocess.** 2017. 120f. Dissertation (Master). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2017.

The human granulocytic colony stimulating factor (hG-CSF) acts mainly by promoting the maturation of neutrophils. The hG-CSF has been used in the treatment of many diseases, in particular in neutropenic conditions. In Brazil, this medicine is inserted into the Pharmaceutical Assistance Program of the National Health Care System as an exceptional drug, but it has a high cost on purchasing because it is still not produced here. A study of technical and economic feasibility of the implementation of hG-CSF production in Brazil has shown that this drug can be produced by costs up to 90% lower than the amounts paid in government procurement. This work shows the cloning and expression of rhG-CSF in heterologous *P. pastoris* system and the best parameters for it, under different culture conditions, as well as a characterization of the producing line of recombinant protein. The expression plasmid was constructed with the pPICZ α A vector and confirmed by nucleotide sequencing. After transformation into *P. pastoris* some colonies were tested to evaluate the expression conditions of the recombinant protein. The colony that presented a higher expression level was chosen to produce the recombinant protein in a larger scale. Detection of the recombinant protein was performed by SDS-PAGE, Western Blotting and Dot Blotting. The steps of cloning and screening the transformants were successfully performed. Plackett-Burman planning was applied to select the significant variables, which were the pH and the addition of sorbitol and amino acids in the culture medium. The data obtained underwent statistical analysis that allowed this selection. The statistically significant variables were chosen and the Central Composite Design was applied for an analysis at higher levels of the significant variables. A lineage characterization was done so that the selected colony was characterized as a producer lineage. Although all three variables were significant (pH, sorbitol and amino acids) only the increase in pH had a beneficial effect on rhG-CSF production. Cell growth curves were made in addition to calculating the maximum growth time and doubling time. These results show the feasibility of rhG-CSF production in *P. pastoris* and this is the first step for the strategy to produce this biopharmaceutical in Brasil in a near future. In addition, the results contribute to the bioprocess optimization studies.

Keywords: G-CSF, *Pichia pastoris*, Screening Design, experimental design, Plackett-Burman, Central Composite Design.

LISTA DE FIGURAS

Capítulo I

Figura 1. Diagrama de fita da rhG-CSF. Estrutura secundária da proteína.....	5
Figura 2. Ilustração mostrando múltipla inserção gênica no genoma de <i>Pichia pastoris</i>	9
Figura 3. Via metabólica da oxidação do metanol em formaldeído na levedura <i>Pichia pastoris</i>	11
Figura 4. Esquema mostrando os passos principais dos métodos referentes ao capítulo 1.....	19
Figura 5. Mapa e região múltipla de clonagem do vetor pPICZαA.....	21
Figura 6. Mapa do vetor pJET1.2/blunt.....	23
Figura 7. Fotografia de um gel de agarose 1% corado com Gel Red mostrando a amplificação do fragmento correspondente à sequência da rhG-CSF a partir do pUC57_rhG-CSF, com reações em diferentes temperaturas de anelamento.....	32
Figura 8. Fotografia de um gel de agarose 1% resultado da PCR de colônia. Identificação de clones recombinantes em pJET1.2/blunt.....	33
Figura 9. Fotografia do E-Gel <i>SizeSelect Agarose</i> 2% (Invitrogen) em teste de restrição para identificação de clones recombinantes pPICZαA_rhG-CSF em DH10β.....	34
Figura 10. Caracterização do gene rhG-CSF: sequência completa do DNA codificante da proteína madura.....	35
Figura 11. Ensaio em menor escala da expressão da rhG-CSF em <i>Pichia pastoris</i>	36
Figura 12. Ensaio do perfil de expressão da rhG-CSF em <i>Pichia pastoris</i> . Linhagem não transformada.....	37

Figura 13. Análise da purificação da rhG-CSF por cromatografia de afinidade ao níquel. SDS-PAGE 12% corado com Comassie Blue. Frações da primeira corrida com aproximadamente 2 mL de sobrenadante.....	38
Figura 14. Análise da purificação da rhG-CSF por cromatografia de afinidade ao níquel. SDS-PAGE 12% corado com Comassie Blue. Frações da segunda, terceira, quarta e quinta corridas com aproximadamente 2 mL de sobrenadante cada.....	39
Figura 15. Perfil da cromatografia por afinidade ao Níquel para a purificação da rhG-CSF.....	40
Figura 16. Membrana de nitrocelulose mostrando um <i>Western Blotting</i> da proteína expressa em <i>P. pastoris</i> . As frações proteicas foram incubadas com o anticorpo primário anti-GCSF na diluição de 1:200.....	41
Figura 17. Membrana de nitrocelulose mostrando um <i>Dot Blotting</i> da proteína incubada com o anticorpo primário na diluição de 1:200.....	42

Capítulo II

Figura 1. Variáveis incluídas para o estudo de sua influência na produção de rhG-CSF usando-se o Delineamento Plackett-Burman.....	52
Figura 2. Esquema representando os passos do protocolo do Kit ELISA G-CSF Humano.....	55
Figura 3. Variáveis incluídas para o estudo de sua influência na produção de rG-CSF usando-se o Delineamento Composto Central (DCC).....	56
Figura 4. Esquema geral mostrando todos os passos do delineamento experimental....	60
Figura 5. Curva padrão Hu G-CSF pg/mL.....	67
Figura 6. Gráfico em colunas mostrando a diferença de concentração da rhG-CSF entre os 16 ensaios realizados.....	68
Figura 7. Efeitos estimados e coeficientes para proteína em pg/mL. Minitab.....	70

Figura 8. Gráfico de Pareto mostrando a relação entre a produção de proteína rhG-CSF e as variáveis testadas. A: Temperatura; B: pH; C: % metanol; D: Sorbitol; E: Aminoácidos; F: Tween 80.....	71
Figura 9. Gráfico da distribuição normal dos resíduos.....	72
Figura 10. Análise de superfície de resposta para as variáveis pH e aminoácidos.....	73
Figura 11. Curva padrão Hu G-CSF pg/mL para o DCC.....	77
Figura 12. Gráfico de colunas mostrando a diferença de concentração da rhG-CSF entre os 13 ensaios realizados.....	78
Figura 13. Efeitos estimados e coeficientes para proteína em pg/mL. Minitab.....	79
Figura 14. Análise de superfície de resposta para as variáveis pH e Sorbitol. Resultado do DCC.....	80
Figura 15. Perfil da cromatografia por afinidade ao Níquel para a purificação da rhG-CSF. Eluente B se refere ao Tampão de Lise com Imidazol (Tris 100 mM, NaCl 100 mM, NaH ₂ PO ₄ 17.4 mM, Imidazol 250 mM).....	82
Figura 16. Membrana de nitrocelulose mostrando um Dot Blotting da proteína incubada com o anticorpo primário na concentração de 1:200 mostrando em 1 : amostra Filgrastima comercial e 2 : amostra da rhG-CSF após os dez repiques para confirmar linhagem produtora.....	83
Figura 17. Curva de crescimento celular da colônia estoque e da colônia após os dez repiques (caracterização). O sinal + representa a caracterização e o sinal x representa o estoque.....	84
Figura 18. Regressão linear da amostra estoque e da amostra caracterização, mostrando as equações referentes a cada uma das retas.....	85

LISTA DE TABELAS

Capítulo I

Tabela 1. Sequência dos primers utilizados na clonagem do gene da rhG-CSF em vetor pPICZ α A.....	21
---	----

Capítulo II

Tabela 1. Valores utilizados no Delineamento Plackett-Burman.....	53
Tabela 2. Matriz de dados codificados para as variáveis estudadas.....	53
Tabela 3. Valores utilizados no Delineamento Composto Central (DCC).....	57
Tabela 4. Matriz de dados codificados para as variáveis estudadas.....	57
Tabela 5. Descrição das variáveis referentes a cada um dos ensaios mostrados na Figura 6.....	69
Tabela 6. Valores das variáveis referentes a cada um dos ensaios.....	78
Tabela 7. Tabela comparativa entre a amostra Estoque e a amostra Caracterização, mostrando a velocidade de crescimento específica, o tempo de duplicação e o rendimento final de cada.....	86

Sumário

RESUMO.....	i
ABSTRACT.....	ii
LISTA DE FIGURAS.....	iii
LISTA DE TABELAS.....	vi
Introdução Geral.....	1
Objetivos.....	6
Capítulo I – Clonagem e indução da expressão da rhG-CSF.....	i
RESUMO.....	i
ABSTRACT.....	ii
LISTA DE FIGURAS.....	iii
LISTA DE TABELAS.....	v
1. Introdução.....	8
1.1. Células progenitoras da linhagem mieloide.....	8
1.2. O Fator Estimulador de colônias de granulócito humano (hG-CSF).....	9
1.2.1. Caracterização molecular do hG-CSF.....	11
1.2.2. O hG-CSF recombinante.....	13
1.2.3. rhG-CSF e sua importância clínica.....	14
1.3. <i>Pichia pastoris</i>	15
1.3.1. Assimilação e formação de glicerol e metabolismo do metanol.....	17
1.3.2. Produção de proteínas heterólogas em <i>P. pastoris</i>	18
1.4. Brasil e os medicamentos biotecnológicos.....	19
2. Hipótese.....	23
3. Objetivos.....	24
3.1. Objetivo Geral.....	24
3.2. Objetivos Específicos/Metas.....	24
4. Materiais e métodos.....	25
4.1. Materiais.....	25
4.2. Métodos.....	26
4.2.1. Expressão recombinante do fator estimulador de colônia de granulócito humano..	26
4.2.1.1. Construção do plasmídeo de expressão.....	26
4.2.2. Sequenciamento nucleotídico da região codificadora da proteína recombinante de interesse	33
4.2.3. Transformação em <i>Pichia pastoris</i>	34

4.3.	Expressão em <i>Pichia pastoris</i> e purificação.....	36
4.4.	Deteção da proteína recombinante por <i>Dot Blotting</i> e <i>Western Blotting</i>	37
5.	Resultados e Discussão.....	39
5.1.	Clonagem do gene da rhG-CSF em <i>Pichia pastoris</i>	39
5.2.	Análise do sequenciamento.....	42
5.3.	Expressão de rhG-CSF em <i>P. pastoris</i>	43
5.4.	Purificação.....	44
5.5.	Deteção da proteína recombinante por <i>Western Blotting</i> e <i>Dot Blotting</i> com anticorpo anti-G-CSF.....	47
6.	Conclusão.....	50
Capítulo II – Delineamento experimental e caracterização de linhagem.....		i
RESUMO.....		i
ABSTRACT.....		ii
LISTA DE FIGURAS.....		iii
LISTA DE TABELAS.....		v
1.	Introdução.....	51
1.1.	Planejamento experimental.....	51
1.1.1.	Experimentos fatoriais.....	51
1.1.1.1.	Planejamento Plackett-Burman.....	52
1.1.1.2.	Delineamento Composto Central.....	54
2.	Hipótese.....	57
3.	Objetivos.....	58
3.1.	Objetivo Geral.....	58
3.2.	Objetivos Específicos.....	58
4.	Materiais e métodos.....	59
4.1.	Materiais.....	59
4.2.	Métodos.....	59
4.2.1.	Planejamento experimental para testar as diferentes condições físico-químicas para a indução da rhG-CSF.....	59
4.2.1.1.	Diferentes condições físico-químicas para a indução da rhG-CSF.....	59
4.2.1.2.	Delineamento Plackett-Burman (<i>Screening Design</i>).....	59
4.2.1.3.	Delineamento Composto Central (DCC).....	64
4.2.2.	Caracterização da linhagem.....	68
5.	Resultados e discussão.....	72

5.1. Planejamento experimental para testar as diferentes condições físico-químicas para a indução da rhG-CSF.....	72
5.1.1. Diferentes condições físico-químicas para a indução da rhG-CSF.....	72
5.1.2. Planejamento Plackett-Burman (<i>Screening Design</i>).....	74
5.1.3. Delineamento Composto Central (DCC).....	84
5.2. Caracterização de linhagem.....	89
6. Conclusão.....	96
CONSIDERAÇÕES FINAIS.....	97
REFERÊNCIAS BIBLIOGRÁFICAS.....	98

Introdução Geral

Todas as células sanguíneas se originam de células-tronco hematopoiéticas, que são células progenitoras multipotentes, capazes de se diferenciar em várias linhagens ou tipos celulares, e são auto-renováveis. Esse precursor origina dois tipos de progenitores celulares que são as células progenitoras da linhagem mieloide, que dará origem as linhagens que formarão hemácias e plaquetas e os granulócitos e monócitos, e as células progenitoras da linhagem linfoide são responsáveis pela formação dos linfócitos T e B (Zago, 2001; Kondo et al, 2003). Portanto, as células da linhagem mieloide darão origem aos granulócitos. Estes últimos são caracterizados por possuírem grânulos no citoplasma e entre eles estão os eosinófilos, basófilos e os neutrófilos (Chotinantakul, 2012).

Os neutrófilos constituem o maior subgrupo de fagócitos polimorfonucleados e perfazem de 50% a 75% do total de leucócitos circulantes, possuem um papel central nas respostas imunes e inflamatórias a microrganismos invasores e a células tumorais. Eles têm uma meia vida curta e, por isso, precisam ser constantemente repostos pelos progenitores hematopoiéticos indiferenciados na medula óssea (Abramson e Wheeler, 1993). A proliferação dos mesmos é principalmente controlada pelo fator estimulante de colônia de granulócitos (G-CSF), uma das moléculas estimuladoras do crescimento que são conhecidas como fatores estimulantes de colônias (CSF). Estes fatores constituem um grupo de proteínas que controlam a proliferação e a diferenciação das células hematopoiéticas (Clark and Kamen, 1987). O G-CSF é, portanto, central para defesas imunes baseadas em neutrófilos, através do seu papel regulatório no crescimento, diferenciação, sobrevivência e ativação dos neutrófilos e seus precursores (Callard, 1994). Este fator também está envolvido com o processo de segmentação nuclear dos

neutrófilos maduros (Babalola et al, 2004; Wittman et al, 2006). Além disso, modula as respostas inflamatórias, fazendo com que haja a redução da liberação das citocinas pró-inflamatórias por macrófagos e monócitos ativados (Barreda et al, 2004). A atividade biológica da proteína hG-CSF acontece pela interação a receptores específicos, que têm por função transmitir o sinal extracelular do hG-CSF para o citoplasma das células que expressam esses receptores específicos (Iwasaki et al, 1999).

Para auxiliar os estudos dessa citocina e também visando o uso dessa proteína como medicamento, G-CSF foi clonado e expressado como proteína recombinante e, atualmente, é um produto biológico estabelecido e comercialmente disponível, além de ser uma das proteínas recombinantes mais utilizadas para propósitos clínicos (Clark and Kamen, 1987; Hubel and Engert, 2003).

Atualmente, o Fator Estimulador de Colônias de Granulócitos humano recombinante (rhG-CSF) se encontra disponível comercialmente na forma não glicosilada, produzida em células de bactéria *E. coli* (Filgrastine e Granulokine), e na forma glicosilada, derivada da expressão em células de mamífero (Granocyte) (Codevilla et al, 2004).

No Brasil, muitos desses medicamentos se encontram inseridos no Programa de Assistência Farmacêutica do Sistema Único de Saúde e são considerados excepcionais (Encarnação, 2005; Ministério da Saúde, 2009). Esses produtos, respondem por uma parcela importante dos gastos do Governo Federal, sendo que um terço dos recursos aplicados na política de medicamentos é destinado a esses medicamentos (Delgado e Arcuri, 2013). A atual política do Ministério da Saúde é introduzir plataformas tecnológicas para reduzir gradativamente a dependência do Brasil por esses medicamentos importados e direcionar o país à auto-suficiência. No período de 2011 a

2012, o SUS gastou cerca de R\$ 2.487.355,20 com as importações deste medicamento (DataSUS, 2013)

Estudos de viabilidade técnico-econômica da implantação da produção de rhG-CSF no Brasil demonstraram que esse medicamento pode ser produzido por custos até cerca de 90% inferiores aos valores pagos em compras governamentais, provando que é importante desenvolver a tecnologia para produção deste medicamento no Brasil (Encarnação, 2005).

Em 2006, a Filgrastima, produzida pelo laboratório Amgen Inc, teve a sua primeira patente expirada. A partir da expiração das patentes sobre medicamentos recombinantes, novas oportunidades se abriram no sentido da produção de medicamentos conhecidos como “biossimilares”.

Desta forma, torna-se muito relevante o desenvolvimento de estudos que visem aprimorar métodos para o desenvolvimento e produção da proteína rhG-CSF nacionalmente, e assim gere informações necessárias para o caminho de independência da importação dos mesmos.

Pichia pastoris é uma levedura anaeróbia facultativa e um organismo metilotrófico, ou seja, capaz de metabolizar o metanol como única fonte de carbono. O sistema de expressão heteróloga de proteínas empregando essa levedura tem se mostrado muito eficiente, mesmo quando comparado com o sistema procariótico de expressão utilizando *Escherichia coli*. Este sistema apresenta vantagens como a secreção de proteínas para o meio extracelular, o processamento pós-traducional (por exemplo, glicosilação), correto enovelamento das proteínas sintetizadas e ausência de formação de corpos de inclusão. A possibilidade de integração dos vetores de expressão ao genoma da levedura aumenta a estabilidade das linhagens utilizadas na expressão heteróloga (Torres e Moraes, 2000).

Delineamento experimental, do inglês, *Design of Experiments*, é um instrumento empregado para avaliar a relação de causa e efeito de um determinado processo, através de vários testes nos quais são realizadas mudanças intencionais nas variáveis de entrada, que acaba por gerar alterações diferenciadas na variável de resposta deste mesmo processo. Os objetivos deste delineamento são: determinar quais são as variáveis mais influentes do processo na variável resposta; definir os níveis de ajuste dos fatores para que a resposta tenha um valor próximo do valor pretendido; determinar os níveis dos fatores para que a variabilidade no valor da resposta seja o menor possível; determinar os níveis de ajuste dos fatores controláveis para que os efeitos das variáveis incontroláveis sejam diminuídos (Montgomery, 2009; Bell et al, 2009).

Dentre as opções de delineamentos experimentais existentes os experimentos fatoriais são bastante triviais. Os experimentos fatoriais abrangem combinações entre dois ou mais fatores, sendo que cada combinação se refere a um tratamento.

Uma alternativa viável ao uso do arranjo fatorial completo, que permite analisar todos os possíveis ensaios e requer um número grande de experimentos, é o *Screening Design* ou delineamento de seleção. Este tipo de planejamento normalmente é realizado em estágios iniciais, nos quais é provável que certos fatores analisados possuam pouco ou nenhum efeito significativo sobre a variável resposta, podendo, assim, serem desconsiderados do planejamento.

O planejamento Plackett-Burman (PB) é classificado como delineamento fatorial fracionado de dois níveis, tendo disponíveis matrizes para qualquer n múltiplo de quatro fatores (12, 16, 20, 24...ensaios). O Delineamento Composto Central (DCC) é um dos mais populares para o ajuste de modelos de até segunda ordem (Myers e Montgomery, 1995) e servirá, nesse trabalho, para confirmação de significância dos parâmetros

resultantes do planejamento PB. O DCC é constituído por três partes: cúbica (fatorial), axial (α) e central.

Aperfeiçoar o bioprocesso tem se mostrado cada vez mais importante para a produção de biofármacos. Desta forma, realizar um planejamento experimental para testar as diferentes condições físico-químicas para a indução da rhG-CSF se torna bastante relevante, tanto para melhorar a produção da proteína específica quanto para auxiliar e contribuir com novos estudos nessa área. Torna-se necessário o estudo de novas metodologias que permitam um melhor rendimento final dessa proteína, estabelecendo-se protocolos que visem melhorar o processo de indução, expressão e, assim, a produção mais eficiente da mesma.

Objetivos

Objetivo geral

Produzir e caracterizar a proteína recombinante humana rhG-CSF em *Pichia pastoris* e estabelecer parâmetros de indução e expressão da mesma em diferentes condições de cultivo. Estabelecer parâmetros de indução e expressão da rhG-CSF, produzida em *Pichia pastoris*, em diferentes condições de cultivo. Pretende-se também realizar uma caracterização da linhagem produtora da proteína recombinante.

Objetivos específicos

- Construir o plasmídeo recombinante de expressão pPICZ α A::rhG-CSF;
- Transformar células de *Pichia pastoris* com o vetor de expressão pPICZ α A::rhG-CSF;
- Isolar transformantes produtores da proteína recombinante;
- Expressar a proteína recombinante;
- Selecionar os clones mais produtores;
- Purificar a proteína por cromatografia de afinidade, em coluna com resina contendo níquel;
- Detectar a proteína recombinante via *Western Blotting*;
- Expressar a proteína sob a indução de diferentes temperaturas, pH, concentrações de metanol e tempos de indução, além de modificações no meio de cultivo e adição de surfactantes;
- Aplicar a metodologia de Plackett-Burman e Delineamento Composto Central;
- Análise estatística dos dados;

- Avaliar a estabilidade genética da linhagem recombinante;
- Montar uma curva de crescimento celular em função do tempo;
- Calcular a velocidade máxima de crescimento e tempo de duplicação da linhagem recombinante.

Capítulo I – Clonagem e indução da expressão da rhG-CSF

RESUMO

ARCHANGELO, B. S. **Expressão do fator estimulador de colônia de granulócito humano recombinante (rhG-CSF) em sistema de *Pichia pastoris* e otimização do bioprocesso.** 2017. 120f. Dissertação (Mestrado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2017.

O fator estimulador de colônias de granulócito humano (hG-CSF) atua principalmente promovendo a maturação dos neutrófilos (granulócitos) e estimulando sua atividade fagocítica e quimiotática. O hG-CSF tem sido utilizado no tratamento de diversas patologias, sobretudo em neutropenias provocadas por eventos que suprimem a produção de células mieloides. O uso deste biofármaco é liberado pela agência reguladora dos Estados Unidos (FDA) desde 1991. No Brasil, este medicamento está inserido no Programa de Assistência Farmacêutica do SUS como medicamento excepcional, porém a um altíssimo custo para aquisição, pois ainda não é produzido no Brasil, precisando ser importado. Um estudo de viabilidade técnico-econômica da implantação da produção de rhG-CSF no Brasil demonstrou que esse medicamento pode ser produzido por custos até cerca de 90% inferiores aos valores pagos em compras governamentais. Este trabalho mostra a clonagem e expressão do rhG-CSF em sistema heterólogo de *P. pastoris* e estabelecer os melhores parâmetros para a indução e produção da mesma. A levedura *P. pastoris* possui características interessantes como sistema de expressão com possibilidade de modificação genética simplificada, capacidade de secretar proteínas e realizar modificações pós-traducionais, além de ser um organismo de fácil manipulação. A construção do plasmídeo de expressão foi feita com o vetor pPICZαA e confirmada por sequenciamento nucleotídico. Após a transformação em *P. pastoris*, algumas colônias foram testadas para avaliar as condições da expressão da proteína recombinante. A colônia que apresentou uma expressão melhor foi escolhida a produção da proteína recombinante em uma escala maior. A detecção da proteína recombinante foi realizada por SDS-PAGE, *Western Blotting* e *Dot Blotting*. Os passos de clonagem e *screening* dos transformantes foram realizados com sucesso. O SDS-PAGE da purificação mostrou a proteína alvo em diferentes bandas, mas o *Western Blotting* identificou apenas bandas de aproximadamente 50 kDa. Estes resultados mostram a facilidade da produção da rhG-CSF em *P. pastoris* e estes são os primeiros passos para uma estratégia de tornar o Brasil um produtor deste biofármaco futuramente, além de contribuir para estudos de otimização de bioprocessos.

Palavras-chave: G-CSF, *Pichia pastoris*, bioprocesso, fator estimulador de colônia de granulócito humano.

ABSTRACT

ARCHANGELO, B. S. **Expression of human granulocyte colony stimulating factor (rhG-CSF) in *Pichia pastoris* and the optimization of the bioprocess.** 2017. 120f. Dissertation (Master). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2017.

Progenitor cells of myeloid lineage are committed to the proliferation and differentiation of erythroid-megakaryocyte lineage and granulocytic-monocytic lineage. The human granulocytic colony stimulating factor (hG-CSF) acts mainly by promoting the maturation of neutrophils (granulocytes) and stimulating their phagocytic and chemotactic activity. The hG-CSF has been used in the treatment of many diseases, in particular in neutropenic conditions caused by treatments that suppress the production of myeloid cells. The use of this biopharmaceutical is approved by the Food and Drug Administration since 1991. In Brazil, this medicine is inserted into the Pharmaceutical Assistance Program of the National Health Care System as an exceptional drug, but it has a high cost on purchasing because it is still not produced in Brazil. A study of technical and economic feasibility of the implementation of hG-CSF production in Brazil has shown that this drug can be produced by costs up to 90% lower than the amounts paid in government procurement. This work shows the cloning and expression of rhG-CSF in heterologous *P. pastoris* system and the best parameters for it. *Pichia pastoris* expression system is quite attractive because it is easy to be genetically modified, its ability to secrete proteins and carry out post-translational modification as well as being a simple eukaryote. The expression plasmid was constructed with the pPICZαA vector and confirmed by nucleotide sequencing. After transformation into *P. pastoris* some colonies were tested to evaluate the expression conditions of the recombinant protein. The colony that presented a higher expression level was chosen to produce the recombinant protein in a larger scale. Detection of the recombinant protein was performed by SDS-PAGE, Western Blotting and Dot Blotting. The steps of cloning and screening the transformants were successfully performed. SDS-PAGE of the purification showed the target protein in different bands, but Western Blotting identified only bands of approximately 50 kDa. These results show the feasibility of rhG-CSF production in *P. pastoris* and this is the first step for the strategy to produce this biopharmaceutical in Brazil in a near future. In addition, the results contribute to the bioprocess optimization studies.

Keywords: G-CSF, *Pichia pastoris*, bioprocess, human granulocytic colony stimulating factor, biotechnology.

LISTA DE FIGURAS

Figura 1. Diagrama de fita da rhG-CSF. Estrutura secundária da proteína.....	5
Figura 2. Ilustração mostrando múltipla inserção gênica no genoma de <i>Pichia pastoris</i>	9
Figura 3. Via metabólica da oxidação do metanol em formaldeído na levedura <i>Pichia pastoris</i>	11
Figura 4. Esquema mostrando os passos principais dos métodos referentes ao capítulo 1.....	19
Figura 5. Mapa e região múltipla de clonagem do vetor pPICZ α A.....	21
Figura 6. Mapa do vetor pJET1.2/blunt.....	23
Figura 7. Fotografia de um gel de agarose 1% corado com Gel Red mostrando a amplificação do fragmento correspondente à sequência da rhG-CSF a partir do pUC57_rhG-CSF, com reações em diferentes temperaturas de anelamento.....	32
Figura 8. Fotografia de um gel de agarose 1% resultado da PCR de colônia. Identificação de clones recombinantes em pJET1;2/blunt.....	33
Figura 9. Fotografia do E-Gel <i>SizeSelect</i> Agarose 2% (Invitrogen) em teste de restrição para identificação de clones recombinantes pPICZ α A_rhG-CSF em DH10 β	34
Figura 10. Caracterização do gene rhG-CSF: sequência completa do DNA codificante da proteína madura.....	35
Figura 11. Ensaio em menor escala da expressão da rhG-CSF em <i>Pichia pastoris</i>	36
Figura 12. Ensaio do perfil de expressão da rhG-CSF em <i>Pichia pastoris</i> . Linhagem não transformada.....	37

- Figura 13.** Análise da purificação da rhG-CSF por cromatografia de afinidade ao níquel. SDS-PAGE 12% corado com Comassie Blue. Frações da primeira corrida com aproximadamente 2 mL de sobrenadante.....38
- Figura 14.** Análise da purificação da rhG-CSF por cromatografia de afinidade ao níquel. SDS-PAGE 12% corado com Comassie Blue. Frações da segunda, terceira, quarta e quinta corridas com aproximadamente 2 mL de sobrenadante cada.....39
- Figura 15.** Perfil da cromatografia por afinidade ao Níquel para a purificação da rhG-CSF.....40
- Figura 16.** Membrana de nitrocelulose mostrando um *Western Blotting* da proteína expressa em *P. pastoris*. As frações proteicas foram incubadas com o anticorpo primário anti-GCSF na diluição de 1:200.....41
- Figura 17.** Membrana de nitrocelulose mostrando um *Dot Blotting* da proteína incubada com o anticorpo primário na diluição de 1:200.....42

LISTA DE TABELAS

Tabela 1. Sequência dos primers utilizados na clonagem do gene da rhG-CSF em vetor pPICZ α A.....	21
---	----

Sumário Capítulo I

RESUMO	i
ABSTRACT	ii
LISTA DE FIGURAS.....	iii
LISTA DE TABELAS.....	vi
Introdução Geral	1
Objetivos	6
Capítulo I – Clonagem e indução da expressão da rhG-CSF.....	i
RESUMO	i
ABSTRACT	ii
LISTA DE FIGURAS.....	iii
LISTA DE TABELAS.....	v
1. Introdução	8
1.1. Células progenitoras da linhagem mieloide.....	8
1.2. O Fator Estimulador de colônias de granulócito humano (hG-CSF).....	9
1.2.1. Caracterização molecular do hG-CSF	11
1.2.2. O hG-CSF recombinante	12
1.2.3. rhG-CSF e sua importância clínica	14
1.3. <i>Pichia pastoris</i>	15
1.3.1. Assimilação e formação de glicerol e metabolismo do metanol	17
1.3.2. Produção de proteínas heterólogas em <i>P. pastoris</i>	18
1.4. Brasil e os medicamentos biotecnológicos	19
2. Hipótese	23
3. Objetivos	24
3.1. Objetivo Geral	24
3.2. Objetivos Específicos/Metas	24
4. Materiais e métodos	25
4.1. Materiais	25
4.2. Métodos	26
4.2.1. Expressão recombinante do fator estimulador de colônia de granulócito humano ..	26
4.2.1.1. Construção do plasmídeo de expressão	26
4.2.2. Sequenciamento nucleotídico da região codificadora da proteína recombinante de interesse	33
4.2.3. Transformação em <i>Pichia pastoris</i>	34

4.3.	Expressão em <i>Pichia pastoris</i> e purificação	36
4.4.	Detecção da proteína recombinante por <i>Dot Blotting</i> e <i>Western Blotting</i>	37
5.	Resultados e Discussão	39
5.1.	Clonagem do gene da rhG-CSF em <i>Pichia pastoris</i>	39
5.2.	Análise do sequenciamento	42
5.3.	Expressão de rhG-CSF em <i>P. pastoris</i>	43
5.4.	Purificação	44
5.5.	Detecção da proteína recombinante por <i>Western Blotting</i> e <i>Dot Blotting</i> com anticorpo anti-G-CSF	47
6.	Conclusão	50

1. Introdução

1.1. Células progenitoras da linhagem mieloide

Todas as células sanguíneas circulantes do adulto são geradas na medula óssea em um processo fisiológico chamado de hematopoese. No desenvolvimento do ser vivo este processo ocorre primeiramente em ilhotas sanguíneas do saco vitelínico e do mesênquima para-aórtico e posteriormente no baço e no fígado. A medula óssea assume essa função gradualmente. Após a puberdade ocorre principalmente nos ossos chatos, como no esterno, vértebras, ilíacos e nas costelas (Abbas, 2008). Quando a medula óssea é danificada ou quando há um aumento da demanda para a produção de células sanguíneas novas, o fígado e o baço são os órgãos alistados como locais de hematopoese extra medular (Chotinantakul, 2012).

Todas as células sanguíneas se originam de células-tronco hematopoiéticas, que são células progenitoras definidas como multipotentes, capazes de se diferenciar em várias linhagens ou tipos celulares, e que possuem características auto-renováveis. Esse precursor origina dois tipos de progenitores celulares que são as células progenitoras da linhagem mieloide e progenitoras da linhagem linfoide. As progenitoras mielóides darão origem a linhagens que formarão hemácias e plaquetas (eritroide-megacariocítica) e os granulócitos e monócitos (granulocítica-monocítica) enquanto que as células progenitoras da linhagem linfoide são responsáveis pela formação dos linfócitos T e B (Zago, 2001; Kondo et al, 2003).

As células da linhagem mieloide originam as Unidades Formadoras de Colônia (CFU). As células CFU, por serem mais diferenciadas, dão origem apenas a monócitos, macrófagos, células dendríticas, megacariócitos e aos granulócitos. Estes últimos são caracterizados por possuírem grânulos no citoplasma e entre eles estão os eosinófilos, basófilos e os neutrófilos (Chotinantakul, 2012).

As citocinas têm um papel extremamente importante no processo de diferenciação celular, pois estimulam o desenvolvimento e a proliferação de células precursoras. São glicoproteínas secretadas por células do sistema imune ou por outros tipos celulares em resposta a algum tipo de estímulo. Várias citocinas são chamadas de fatores de estimulação de colônia por possuírem a habilidade de estimular o crescimento e o desenvolvimento de muitas colônias leucocíticas ou eritroides (Alberts, 2011; Mendelson e Frenette, 2014).

1.2. O Fator Estimulador de colônias de granulócito humano (hG-CSF)

Granulócitos polimorfonucleares, macrófagos e células *Natural Killers* representam a primeira linha de defesa celular contra agentes infecciosos. Os neutrófilos constituem o maior subgrupo de fagócitos polimorfonucleados e perfazem de 50% a 75% do total de leucócitos circulantes, possuem um papel central nas respostas imunes e inflamatórias a microrganismos invasores e a células tumorais. Eles têm uma meia vida curta e, por isso, precisam ser constantemente repostos pelos progenitores hematopoiéticos indiferenciados na medula óssea (Abramson e Wheeler, 1993).

A proliferação dos neutrófilos é principalmente controlada pelo fator estimulante de colônia de granulócitos (G-CSF), uma das moléculas estimuladoras do crescimento que são conhecidas como fatores estimulantes de colônias (CSF). Estes fatores

constituem um grupo de proteínas que controlam a proliferação e a diferenciação das células hematopoiéticas (Clark and Kamen, 1987). O G-CSF é, portanto, central para defesas imunes baseadas em neutrófilos, através do seu papel regulatório no crescimento, diferenciação, sobrevivência e ativação dos neutrófilos e seus precursores (Callard, 1994).

O fator estimulante de granulócito humano (hG-CSF) é o responsável por estimular e regular a proliferação e a diferenciação das células hematopoiéticas granulocíticas (Clark and Kamen, 1987). Em uma escala menor, hG-CSF também atua em outros leucócitos, aumentando o número de células brancas agranulocíticas no sistema circulatório (linfócitos e monócitos) e nas células não hematopoiéticas, estimulando a proliferação e migração de células endoteliais vasculares (Sica et al, 1996; Barreda et al, 2004). Este fator também está envolvido com o processo de segmentação nuclear dos neutrófilos maduros (Babalola et al, 2004; Wittman et al, 2006). Além disso, modula as respostas inflamatórias, fazendo com que haja a redução da liberação das citocinas pró-inflamatórias por macrófagos e monócitos ativados (Barreda et al, 2004).

A atividade biológica da proteína hG-CSF acontece pela interação a receptores específicos, que têm por função transmitir o sinal extracelular do hG-CSF para o citoplasma das células que expressam esses receptores específicos, e que fazem parte da superfamília de receptores de citocina do tipo I. Esses receptores são glicoproteínas e sua estrutura é formada por um domínio transmembrânico e um domínio citoplasmático, o que possibilita a transdução de sinal intracelular. A interação do hG-CSF com o seu receptor fosforila as tirosinas quinases JAK1 e JAK2, o que ativa a via quinase Ras/MAP, que é responsável pela transdução do sinal (Iwasaki et al, 1999). Quando expressos em progenitores neutrofílicos e em neutrófilos maduros, o receptor de hG-

CSF tem a função de transmitir sinais de proliferação, diferenciação e sobrevivência, por outro lado, nos monócitos, levam a atenuação da liberação de citocinas pró-inflamatórias quando estes são ativados (Demetri e Griffin, 1991).

Para auxiliar os estudos dessa citocina e também visando o uso dessa proteína como medicamento, G-CSF foi clonado e expressado como proteína recombinante e, atualmente, é um produto biológico estabelecido e comercialmente disponível, além de ser uma das proteínas recombinantes mais utilizadas para propósitos clínicos, como tratamento de neutropenias causadas por doenças ou tratamentos imunossupressores (Clark and Kamen, 1987; Hubel and Engert, 2003).

1.2.1. Caracterização molecular do hG-CSF

O gene humano G-CSF está localizado no cromossomo 17 nas posições q21-22 e contém cinco éxons e quatro íntrons que abrangem uma área de aproximadamente 2,5 kb (Kanda et al, 1987). O comprimento total do mRNA do hG-CSF codifica uma proteína precursora de 204 aminoácidos, com um peptídeo sinal de 30 aminoácidos, duas pontes dissulfeto intercadeias preditas e nenhum sítio para N-glicosilação. A clivagem do peptídeo sinal produz uma proteína G-CSF madura de 174 aminoácidos com um peso molecular, um pouco maior do que o previsto, de 18,8 kDa, devido a uma O-glicosilação na posição Thr-133, que protege a molécula de agregações mas não parece influenciar diretamente na ligação com o receptor (Nagata et al, 1986; Souza et al, 1986). O ponto isoelétrico da proteína (pI) é 5,5 (Souza et al, 1986).

A estrutura tridimensional do hG-CSF foi determinada por cristalografia de raio-X. De acordo com Hill e colaboradores (1993), que produziram rhG-CSF, esta proteína é composta por quatro alfa-hélices dispostas de maneira antiparalela e com rotação para a esquerda, com dimensões de 45Å x 26Å x 26Å. As quatro α -hélices são denominadas

por A, B, C e D, sendo que a hélice A se forma entre os resíduos 11-39, a hélice B entre os resíduos 71-91, a C entre os resíduos 100-123 e a hélice D entre 143-172 (Hill et al, 1993).

Figura 1: Diagrama de fita da rhG-CSF. Estrutura secundária da proteína. A, B, C e D indicam as quatro α -hélices. Adaptado de Hill (1993).

A proteína hG-CSF contém um resíduo de cisteína livre em sua forma reduzida (SH) na posição 17 e duas pontes dissulfeto (Cys^{36, 42} e Cys^{64, 74}). Essas características são muito importantes para a estabilidade estrutural da proteína, dobramento correto da cadeia de aminoácidos e sua atividade biológica (Lu et al, 1989).

1.2.2. O hG-CSF recombinante

Os biofármacos são uma classe de proteínas terapêuticas produzidas por técnicas modernas de biotecnologia, especificamente por engenharia genética ou por tecnologia do hibridoma. São homólogos às proteínas humanas endógenas ou possuem um alto grau de semelhança com as mesmas. Portanto, para se dizer que uma proteína

recombinante terapêutica é um biofármaco, é necessário avaliar todo o processo utilizado para a sua obtenção, uma vez que estas são produzidas por métodos diferentes daqueles de extração direta de fontes biológicas nativas (Walsh, 2003; Wang, 1999).

Os Fatores Estimuladores de Colônia (CSF), quando *in vivo*, são produzidos em quantidades limitadas (Souetre et al, 1996). Portanto, a clonagem e expressão recombinante do gene hG-CSF facilitou grandemente o estudo de sua biologia molecular, dos seus efeitos no sistema hematopoiético e, assim, o seu uso clínico como medicamento biotecnológico (Nagata et al, 1986; Souza et al, 1986).

Em 1986, o gene hG-CSF foi clonado por dois grupos de pesquisa principais (Nagata et al, 1986; Souza et al, 1986). Em 1995, Larry Souza, clonou o gene G-CSF e produziu a sua proteína recombinante, permitindo com que o entendimento dos seus efeitos no sistema hematopoiético fosse mais bem estudado, além de proporcionar dados para o melhor entendimento da biologia molecular desse fator estimulador de colônia de granulócitos (Hammerling et al, 1995).

Babaeipour et al (2015) desenvolveram um processo para a produção e purificação da hG-CSF em *E. coli* que, de acordo com dados analisados, foi um método bastante eficiente. Algumas pesquisas mostram que a expressão da hG-CSF recombinante em *E. coli* é conseguida em um nível médio a alto, porém, o rendimento final da proteína purificada ainda é baixo, apesar dos bons resultados obtidos em trabalhos anteriores.

Alguns trabalhos foram realizados visando a expressão da rhG-CSF em *Pichia pastoris*. Maity et al (2015) produziram a proteína GCSF códon otimizada em sistema de *P. pastoris*, sob condições diversas, e tiveram um rendimento final de aproximadamente 18 mg/L, o que é considerado um rendimento relativamente baixo.

Apesar da O-glicosilação presente no fator hG-CSF endógeno não interferir diretamente na atividade biológica da proteína, ela tem importante função na prevenção de polimerização, proteção contra proteases e no aumento da termoestabilidade do fator em questão (Oh-eda et al, 1990).

Atualmente, o Fator Estimulador de Colônias de Granulócitos humano recombinante (rhG-CSF) se encontra disponível comercialmente na forma não glicosilada, produzida em células de bactéria *E. coli* (Filgrastine e Granulokine), e na forma glicosilada, derivada da expressão em células de mamífero (Granocyte) (Codevilla et al, 2004). Além destas, também há uma rhG-CSF produzida em bactéria que possui alguns aminoácidos substituídos, que, teoricamente, confere uma maior estabilidade e potência à mesma (Gomes et al, 2012).

O uso clínico da rhG-CSF foi aprovado pela agência reguladora dos Estados Unidos *U. S. Food & Drug Administration (FDA)*, em fevereiro de 1991 (U.S. Food and Drug Administration). No Brasil, o mesmo foi registrado pela Agência Nacional de Vigilância Sanitária em julho de 2002.

1.2.3. rhG-CSF e sua importância clínica

Após a sua liberação, o medicamento recombinante foi utilizado para tentar reduzir a incidência de infecções associadas à neutropenias desencadeadas por radioterapia, quimioterapia para o tratamento de tumores e uso de drogas que deixam o indivíduo imunossuprimido (Hartung et al, 1998). A droga também é utilizada para tratar casos de mielosupressão após a realização de transplantes autólogos ou alogênicos da medula óssea, além de neutropenias crônicas associadas à anemia aplástica, neutropenia idiopática, síndromes mielodisplásicas e neutropenia clínica (Hanada e Ono, 1990).

Mais recentemente, rhG-CSF vem sendo utilizado com sucesso para outras finalidades, assim como no reforço do sistema imune em pacientes com HIV, pneumonia, leucemia e neutropenia febril (Babalola et al, 2004; García-Carbonero et al, 2001). Alguns estudos mostram ainda que rhG-CSF pode ser empregado, com sucesso, no tratamento de infecções em pacientes não neutropênicos, como forma de prevenção ou para fortalecer o sistema imune (Wittman et al, 2006).

1.3. *Pichia pastoris*

Pichia pastoris é uma levedura anaeróbia facultativa que pertence à família *Saccharomycetales*, classe *Saccharomycetes*, filo *Ascomycota*, Reino *Fungi*. Sua morfologia é caracterizada por células esféricas ou ovais que podem variar de acordo com as condições de cultivo. Formam colônias não filamentosas com coloração creme, em meio de cultivo sólido (Kurtzman, 1987).

É um organismo metilotrófico, ou seja, capaz de metabolizar o metanol como única fonte de carbono. O primeiro passo na metabolização do metanol é sua oxidação a formaldeído e peróxido de hidrogênio utilizando oxigênio molecular pela enzima álcool oxidase. Esta enzima possui baixa afinidade por oxigênio molecular, que obriga o organismo a ter uma alta taxa de expressão da mesma como estratégia compensatória. O promotor utilizado para transcrever genes heterólogos é o que regula a expressão da enzima álcool oxidase 1 (AOX1) em *P. pastoris*, e é regulado pelo metanol. O gene AOX1 dá origem à maior parte das proteínas relacionadas ao metabolismo de metanol e o seu silenciamento gera organismos com crescimento lento em presença de metanol – fenótipo denominado Mut^s (Figura 2). O promotor do gene que codifica a álcool oxidase 2 (AOX2), exerce sua função normal na produção da enzima para a manutenção e sobrevivência da célula mesmo quando AOX1 está silenciado. Outra característica

importante é que *P. pastoris* não é uma levedura considerada forte fermentadora, portanto não atinge altos níveis de produção de etanol como metabólito intermediário. Isso evita que se possam atingir níveis tóxicos em culturas de alta densidade. O sistema de expressão heteróloga de proteínas empregando essa levedura tem se mostrado muito eficiente, mesmo quando comparado com o sistema procariótico de expressão utilizando *Escherichia coli*. Este sistema apresenta vantagens como a secreção de proteínas para o meio extracelular, o processamento pós-traducional (por exemplo, glicosilação), correto enovelamento das proteínas sintetizadas e ausência de formação de corpos de inclusão. A possibilidade de integração dos vetores de expressão ao genoma da levedura aumenta a estabilidade das linhagens utilizadas na expressão heteróloga (Torres e Morais, 2000).

Figura 2: Ilustração mostrando múltipla inserção gênica. Eventos multi-cópias podem ocorrer como inserção gênica no *aox1::ARG4* loci (EasySelect *Pichia* Expression Kit).

Embora os sistemas de expressão de leveduras, como a *P. pastoris*, possuam vias metabólicas de enovelamento como os organismos eucariotos, o padrão de glicosilação é diferente daqueles encontrados em humanos e em outros mamíferos. Alguns tipos de

carboidratos podem ser antigênicos e a *Food and Drug Administration* requer a completa caracterização de carboidratos dos produtos que possuem aplicações farmacêuticas, uma vez que glicosilações podem mudar a função e as características da proteína recombinante (Daly e Hearn, 2005).

1.3.1. Assimilação e formação de glicerol e metabolismo do metanol

Para muitas espécies de leveduras, incluindo a *Pichia pastoris*, o glicerol é utilizado como fonte de carbono em condições aeróbias. A via catabólica deste composto se caracteriza pela difusão passiva através da fosforilação da membrana plasmática, da ação da enzima glicerol quinase e, posteriormente, por uma oxidação causada pela enzima glicerol mitocondrial ubiquinona oxidoredutase fosfato (Poutou et al, 2005).

O metabolismo energético de leveduras metilotróficas, quando cultivadas com metanol como única fonte de carbono, começa dentro do peroxissomas (Figura 3). Dentro dos mesmos ocorre a oxidação de metanol a formaldeído pela ação da enzima álcool oxidase I e também a produção de peróxido de hidrogênio como um subproduto da reação. A oxidação acontece na presença de oxigênio molecular, que é oceptor final de elétrons ativo (Poutou et al, 2005; Xuan e Zhou, 2009).

cultura, enquanto que ao utilizar Tween 80 o rendimento foi mais de 100 vezes maior, 220 mg/L. Este grupo também utilizou pela primeira vez o Agilent Bioanalyzer para detectar glicoformas de proteínas, que se mostrou ser uma ferramenta eficiente para este tipo de análise. Bahrami et al (2009) também produziram a proteína rhG-CSF em *P. pastoris* utilizando Tween 80 e obtiveram um rendimento de aproximadamente 300 mg/L em batelada alimentada, enquanto que Apte-Deshpande et al (2009) obteve por volta de 220 mg/L nas mesmas condições.

Maity et al (2015) expressaram a rhG-CSF mediante diferentes condições e parâmetros, como diferentes temperaturas de indução, pH do meio, concentração de metanol e adição de surfactantes. Este estudo revelou um rendimento máximo de aproximadamente 16 mg/L da proteína rhG-CSF códon otimizada e expressa sob o controle do promotor AOXI em células de *Pichia pastoris*. Neste trabalho de Maity et al (2015) os melhores parâmetros para a obtenção da proteína foram pH 4, suplementação com mistura de aminoácidos, temperatura de 28°C e concentração de metanol a 1%. Não foi observada agregação ou degradação de proteínas.

Apesar dos trabalhos realizados com o fator rhG-CSF, produzido em sistema heterólogo de *Pichia pastoris*, ainda não foi alcançado um rendimento alto de produção e, portanto, muitos testes ainda precisam ser realizados buscando melhorar as condições de produção da mesma.

1.4. Brasil e os medicamentos biotecnológicos

Medicamentos obtidos por meio da biotecnologia estão sendo cada vez mais utilizados na clínica médica e, desta forma, cresce a cada ano o investimento mundial na produção e comercialização desse tipo de medicamento (Schmidt, 2004).

Em 2006, a Filgrastima, produzida pelo laboratório Amgem Inc, teve a sua primeira patente expirada. A partir da expiração das patentes sobre medicamentos recombinantes, novas oportunidades se abriram no sentido da produção de medicamentos conhecidos como “biossimilares”.

O programa Inova Saúde, que teve início em abril de 2013 e possui uma duração prevista até dezembro de 2017, é uma iniciativa do Ministério da Ciência, Tecnologia e Inovação e da FINEP, em cooperação com o Ministério da Saúde, o BNDES e o CNPq. Este programa foi criado com o objetivo de apoiar atividades de projetos de PD&I em farmoquímicos obtidos por processos de síntese química, biotecnológicos e extrativos, para desenvolvimento de medicamentos, novos ou genéricos, para atendimento às demandas da SUS (Sistema Único de Saúde), e também para o atendimento ao mercado nacional e internacional (FINEP – Inova Saúde). Este programa está inserido no Plano Inova Empresas e destina mais de três bilhões de reais para as atividades de inovação do Complexo da Saúde para que haja um aumento da independência tecnológica do Brasil em relação ao fornecimento de importantes insumos utilizados no campo da saúde humana (FINEP – Inova Saúde).

A possibilidade de desenvolvimento desses medicamentos é uma oportunidade de identificar um novo espaço de atuação, para empresas de pequeno porte e pouco internacionalizadas em países como o Brasil, desenvolvendo e estimulando o setor nacional de biotecnologia (Fardelone e Branchi, 2006).

No Brasil, muitos desses medicamentos se encontram inseridos no Programa de Assistência Farmacêutica do Sistema Único de Saúde e são considerados excepcionais, por possuírem um alto custo para o governo e serem administrados para doenças cujo tratamento deve ser continuado (Encarnação, 2005; Ministério da Saúde, 2009). É

importante ressaltar que este alto custo se deve ao fato de que o Brasil importa a maioria desses medicamentos. Esses produtos, portanto, respondem por uma parcela importante dos gastos do Governo Federal, sendo que um terço dos recursos aplicados na política de medicamentos é destinado a esses medicamentos (Delgado e Arcuri, 2013). A atual política do Ministério da Saúde é introduzir plataformas tecnológicas para reduzir gradativamente a dependência do Brasil por esses medicamentos importados e direcionar o país à auto-suficiência.

A rhG-CSF não é nacionalmente produzida e, desta forma, todo o medicamento comprado pelo governo é importado, o que gera gastos públicos relevantes com o mesmo. No período de 2011 a 2012, o SUS gastou cerca de R\$ 2.487.355,20 com as importações deste medicamento (DataSUS, 2013)

Estudos de viabilidade técnico-econômica da implantação da produção de rhG-CSF no Brasil demonstraram que esse medicamento pode ser produzido por custos até cerca de 90% inferiores aos valores pagos em compras governamentais, provando que é importante desenvolver a tecnologia para produção deste medicamento no Brasil (Encarnação, 2005).

O medicamento recombinante rhG-CSF é de extrema importância para o tratamento de doenças associadas à neutropenia e é muito utilizado no Brasil, principalmente pelo Sistema de Saúde Pública, que investe valores altíssimos na importação desse medicamento. Desta forma, torna-se muito relevante o desenvolvimento de estudos que visem aprimorar métodos para o desenvolvimento e produção da proteína rhG-CSF nacionalmente, e assim gere informações necessárias para o caminho de independência da importação dos mesmos.

Essa proteína tem sido bastante estudada, porém, só foi produzida na sua forma insolúvel, quando em *E. coli*, necessitando de técnicas posteriores de desnaturação e

reenovelamento para tê-la na sua forma ativa, o que pode levar a uma diminuição de seu rendimento final e aumento nos custos das etapas *downstream*. Também existem estudos da produção desta proteína em *Pichia pastoris*, porém ainda não se tem um protocolo eficiente que leve a um alto rendimento da mesma. A rhG-CSF também é produzida em células de mamíferos, o que implica em alto custo e um aumento da complexidade do processo. Portanto, a expressão heteróloga em *Pichia pastoris* é interessante no sentido de ser muito mais barata e facilmente manipulada, em relação às células de mamíferos, e apresentar vantagens como o processamento pós-traducional e correto enovelamento das proteínas sintetizadas, facilitando todo o processo *downstream*.

Portanto, torna-se interessante o uso do sistema de expressão de *Pichia pastoris* para a produção da proteína recombinante humana G-CSF.

2. Hipótese

O sistema de *Pichia pastoris* pode ser usado para produzir a proteína recombinante hG-CSF ativa em altos níveis em um bioprocesso otimizado.

3. Objetivos

3.1. Objetivo Geral

Produzir e caracterizar a proteína recombinante humana rhG-CSF em *Pichia pastoris* e estabelecer parâmetros de indução e expressão da mesma em diferentes condições de cultivo.

3.2. Objetivos Específicos/Metas

- Construção do plasmídeo recombinante de expressão pPICZ α A::rhG-CSF;
- Transformar células de *Pichia pastoris* com o vetor de expressão pPICZ α A::rhG-CSF;
- Isolar transformantes produtores da proteína recombinante;
- Expressar a proteína recombinante;
- Selecionar os clones mais produtores;
- Purificar a proteína por cromatografia de afinidade, em coluna com resina contendo níquel;
- Detectar a proteína recombinante via *Western Blotting*.

4. Materiais e métodos

4.1. Materiais

O estoque primário das linhagens modificadas geneticamente de *Pichia pastoris* KM71H, assim como o vetor de expressão pPICZαA, foram fornecidos pelo grupo de pesquisa da Profa. Eliane Candiani Arantes, do Departamento de Física e Química da faculdade de Ciências Farmacêuticas de Ribeirão Preto – USP. As células de *P. pastoris* foram mantidas em YPD e Glicerol 20% em freezer -80 °C. Os reagentes utilizados foram de grau analítico.

4.2. Métodos

Figura 4: Esquema mostrando os passos principais dos métodos referentes ao capítulo 1.

4.2.1. Expressão recombinante do fator estimulador de colônia de granulócito humano

4.2.1.1. Construção do plasmídeo de expressão

A sequência do gene endógeno utilizado, que codifica uma G-CSF, foi obtida do banco de dados do *National Center for Biotechnology Information* (NCBI), depositada sob o número de acesso NP_757373.1. A Fase Aberta de Leitura (ORF) da G-CSF

humana foi sintetizada pela *GeneScript Corporation* (Piscataway, NJ, EUA), e clonada ao vetor pUC57. Este gene possui um peptídeo sinal para secreção, porém foi sintetizado sem o mesmo, para a utilização do fator α do vetor de expressão (pPICZ α A), que está inserido no genoma de *Pichia pastoris*. Este vetor foi escolhido porque possibilita a ligação do gene de interesse sob a regulação do promotor AOX1 presente no vetor, desta forma, quando o vetor é integrado no genoma de uma levedura com o gene AOX1 cromossomal não funcional – cepa KM71H, por exemplo – a mesma produz grande quantidade da proteína recombinante dirigido pelo AOX1 do vetor, desde que possua metanol como única fonte de carbono.

Para a construção do plasmídeo, o gene da G-CSF foi amplificado por uma Reação em Cadeia da Polimerase (PCR) utilizando os primers específicos pPICZ α A_rhG-CSF Fw e pPICZ α A_rhG-CSF Rv (Tabela 1) que incluem sítios para as enzimas *XhoI* e *XbaI*, respectivamente, que possibilita a clonagem *in frame* no plasmídeo de expressão final pPICZ α A. Estas enzimas foram escolhidas para a clonagem para evitar a tradução do sítio múltiplo de clonagem na porção N-terminal da proteína. Porém, de acordo com o mapa do pPICZ α A (Figura 5), isso resultaria em uma proteína sem o sinal de clivagem para a protease *Kex-II*, que atua na clivagem do fator- α no retículo endoplasmático da levedura. Isto nos levou a adicionar a sequência que codifica os aminoácidos de reconhecimento (lisina e arginina) no primer *forward*. Aos primers foram adicionados três nucleotídeos na extremidade 5' de cada primer, que serviram como facilitadores para a ancoragem das enzimas de restrição da digestão do produto de PCR.

Figura 5: Mapa e região múltipla de clonagem do vetor pPICZ α -A. Destaque em retângulos pretos para os sítios de restrição das enzimas *XhoI* e *XbaI*, utilizadas na clonagem. Destaque em vermelho para o sinal de clivagem da enzima Kex2, perdido na restrição. (Adaptado de: EasySelectTMPichia Expression Kit, Invitrogen).

Primers rhG-CSF	
Primer Forward	5'- AAACCTCGAGAAAAGAATGACCCCGCTGGGGCC – 3'
Primer Reverse	5'- TTTGTGCGACCGTTGTGCCAGGTGACG – 3'

Tabela 1: Sequência dos primers utilizados na clonagem do gene da rhG-CSF em vetor pPICZ α A.

A PCR se deu por 2,5 μL de Tampão 100 mM Tris-HCl pH 8,5 e 500 mM KCl (Ludwig), 0,75 μL de MgCl_2 50 mM (LUDWIG), 1 μL de dNTPs 10 mM (LUDWIG), 1 μL de Primer Fw 10 pmol/ μL , 1 μL de Primer Rv 10 pmol/ μL , 0,5 μL de Taq DNA Polimerase 5 U/uL (LUDWIG), 17,25 μL de água e 1 μL de DNA pPUC57_rhG-CSF (8 ng/ μL), totalizando um volume final de 25 μL . No ciclo utilizado, em termociclador (BIORAD), foi testado um gradiente de temperaturas de anelamento dos primers, que foi de 50 °C à 58 °C. Este ciclo foi 95 °C por 5 minutos, 92 °C por 45 segundos, 50-58 °C por 40 segundos, 72 °C por 45 segundos, repetidos 35 vezes, e 72 °C por 10 minutos. Os produtos de PCR amplificados foram analisados em gel de agarose 0,8 % e purificados do gel utilizando o kit *Wizard SV Gel and PCR Clean-Up System* (PROMEGA).

Tendo a fase aberta de leitura do gene sido amplificada, foi realizada uma subclonagem utilizando o *CloneJET PCR Cloning Kit* (THERMO SCIENTIFIC), que é um sistema de seleção positiva avançado para clonagem de produtos de PCR gerados por DNA polimerases termoestáveis. Com este kit, fragmentos de DNA com extremidades coesivas ou cegas, podem ser clonados com sucesso. Neste processo foi utilizado o vetor de clonagem pJET1.2/blunt (Figura 6), que contém um gene letal, o qual é desligado quando um fragmento de DNA é ligado ao sítio de clonagem. Como resultado, apenas células com plasmídeos recombinantes são capazes de se propagar. O vetor também contém um amplo sítio múltiplo de clonagem, o T7 promotor para transcrição *in vitro* e primers para sequenciamento nucleotídico.

foi semeada em placas de Petri contendo meio LB (Luria Bertani) Ágar e Ampicilina 100 mg/mL (antibiótico para o qual o vetor confere resistência), as quais foram incubadas overnight na estufa 37 °C. Foram realizadas PCR de 10 colônias para análise da presença e orientação do inserto de DNA. A reação para análise da presença do inserto foi composta por 2 µL de *10X Taq Buffer*, 2 µL de dNTP 2mM, 1,2 µL de MgCl₂ 25 mM, 0,4 µL de *primer pJET1.2 Forward* 10 mM, 0,4 µL de *primer pJET1.2 Reverse* 10 mM, 0,1 µL de *Taq DNA Polimerase* 5 U/ µL e 11,9 µL de água e 2 µL de DNA das células, totalizando um volume de 20 µL. O ciclo utilizado foi 95 °C por 3 minutos, 94 °C por 30 segundos, 60 °C por 30 segundos, 72 °C por 1 minuto por 25 vezes. Após a verificação da presença dos insertos em gel de agarose 1%, foi realizado o mesmo protocolo de PCR para verificar se o inserto de interesse (rhG-CSF) estava na orientação correta. Para tanto, a única mudança foi a utilização do primer rhG-CSF Rv juntamente com o primer pJET1.2. Fw

Após a seleção de uma colônia positiva foi feita a extração do DNA plasmidial do pJET1.2._rhG-CSF. Concomitantemente, também foi realizada a extração plasmidial do vetor de expressão pPICZαA (INVITROGEN) armazenado em células DH10β. Para as extrações dos plasmídeos foi utilizado o *QIAprep Spin Miniprep Kit* (QIAGEN) de acordo com as instruções da fabricante.

Os plasmídios isolados (pJET1.2_rhG-CSF e o vetor expressão pPICZαA) foram digeridos com as enzimas *XhoI* e *XbaI* (THERMO SCIENTIFIC) para posterior clonagem do gene rhG-CSF em vetor pPICZαA. A reação de restrição para o vetor pPICZαA foi realizada com 24,3 µL de água, 4 µL de *Buffer Tango* 10x concentrado (THERMO SCIENTIFIC), 10,7 µL de DNA (~1500 ng) e 1 µL de *XbaI* (10 U/ µL), totalizando um volume de 40 µL. Esta reação foi incubada em banho-maria a 37°C por 3 horas. Após este tempo, foram adicionados na reação 3 µL de água, 6 µL de Buffer

Tango 10x (concentração final 2x) e 1 μL de *XhoI* (10 U/ μL), atingindo volume final de 50 μL . A reação permaneceu em banho-maria *overnight*. A reação de restrição para o isolamento do fragmento de interesse do vetor pJET1.2_rhG-CSF foi realizada com 22,1 μL de água, 4 μL de *Buffer Tango* 10X, 12,4 μL de DNA (1500 ng) e 1,5 μL de *XbaI*. Esta reação permaneceu no banho-maria a 37 °C por 3 horas. Após as três horas, foram adicionados à mistura 2,5 μL de água, 6 μL de *Buffer Tango* 10X e 1,5 μL de *XhoI*, totalizando um volume de 50 μL . A reação permaneceu a 37 °C *overnight*. Após a restrição, as enzimas foram inativadas a 80 °C por 20 minutos. As duas amostras foram analisadas em gel de agarose 0,8 % e as seções do gel referentes ao inserto e ao vetor foram recortadas e purificadas utilizando o kit *Wizard SV Gel and PCR Clean-Up System* (PROMEGA) de acordo com as instruções do fabricante.

Para a ligação do vetor pPICZ α A e do inserto rhG-CSF, que foi elaborada respeitando a proporção de 5:1 entre moléculas de inserto e plasmídeo, foram estabelecidas as seguintes condições: 12,47 μL de água, 4 μL de *Buffer Ligase* 5X (INVITROGEN), 1,13 μL de inserto rhG-CSF (30 ng/ μL), 1,4 μL de vetor pPICZ α A (70 ng/ μL) e 1 μL de T4 DNA Ligase (INVITROGEN), tendo um volume final de 20 μL . A reação foi mantida a 16 °C *overnight* no termociclador (BIORAD). Essa reação (20 μL) foi empregada na transformação de células competentes DH10 β (200 μL) por choque térmico, 90 segundos a 42 °C e 2 minutos no gelo (referência de choque térmico). Após o choque, foram adicionados 800 μL de LB caldo no tubo de células competentes e este permaneceu por 1 hora no banho a 37 °C. A mistura de reação de transformação foi utilizada para semear placas de Petri contendo meio LB Low Salt Ágar e o Zeocina 100 $\mu\text{g/mL}$ (antibiótico para o qual o vetor confere resistência), as quais foram incubadas *overnight* em estufa a 37 °C.

Das colônias de células transformadas que cresceram na placa, três foram escolhidas para o teste de restrição com as enzimas *XbaI* e *XhoI*, com o intuito de confirmar a presença do gene alvo integrado no vetor de expressão. Foram preparados três pré-inóculos (um de cada colônia) com LB *Low Salt* e zeocina, que cresceram a 37 °C por aproximadamente 16 horas. A partir dos pré-inóculos foi feita a extração de DNA plasmidial utilizando o *QIAprep Spin Miniprep Kit* (QIAGEN). O DNA plasmidial extraído foi usado para o teste de restrição, que seguiu exatamente o mesmo protocolo das reações de restrições descritas anteriormente, para a confirmação de colônias positivas para pPICZαA_rhG-CSF (colônias 1 e 3). Para as colônias positivas em DH10β, foram feitos estoques em glicerol 10 % e armazenados à -80 °C.

4.2.2. Sequenciamento nucleotídico da região codificadora da proteína recombinante de interesse

O sequenciamento nucleotídico foi realizado no laboratório do Prof. Dr. Sergio Akira, da Faculdade de Ciências Farmacêuticas de Ribeirão Preto. O plasmídeo recombinante, denominado rhG-CSF, foi sequenciado via método de Sanger, método da terminação de cadeia utilizando dideoxynucleotídeos marcados com fluorocromos (BigDye Terminator v3.1 Cycle Sequencing, Applied Biosystems, EUA). As reações de amplificação ocorreram de acordo com o protocolo recomendado pelo fabricante e o sequenciamento foi realizado em sequenciador automático 3500 Genetic Analyzer 8-capillary (Applied Biosystems, EUA). Para o sequenciamento foram utilizados os primers AOX 5' (5'-GACTGGTTCCAATTGACAAGC-3') e AOX 3' (5'-GCAAATGGCATTCTGACATCC-3'). A reação foi realizada com aproximadamente 400 ng de DNA, tampão de sequenciamento (3 µL), BigDye Terminator 3.1 (2 µL), 10 µM de primer e água para um volume final de 20 µL. A sequência obtida foi analisada

com o auxílio do programa Bioedit Sequence Alignment Editor vc7.2.5 (Hall, 2013) e da ferramenta BLAST (Basic Local Alignment Search Tool, NCBI).

4.2.3. Transformação em *Pichia pastoris*

Com o objetivo de se obter células de *P. pastoris* recombinantes para o plasmídeo contendo o gene da rhG-CSF foram produzidas células competentes da linhagem KM71H, segundo o manual “EasySelect™ PichiaExpression kit” (Invitrogen). A metodologia é baseada no crescimento prévio das células até que elas atinjam a fase logarítmica de crescimento celular. Em seguida, há a transferência das células para meio nutriente YEPD (Extrato de levedura 1%, Peptona 2%, Dextrose 2% e Ágar 2%) adicionado de HEPES (ácido N- (2-hidroxietil) piperazina-N`-2- etanossulfônico), um tampão que mantém o pH fisiológico e DTT (Ditiotreitol), que tem a função de evitar a formação de pontes de sulfeto intracelular. Por fim, lavagens nas células foram feitas com água gelada para a transferência das mesmas para uma solução de Sorbitol 1M.

A transformação das células competentes foi realizada com o plasmídeo recombinante pPICZαA_rhG-CSF. Cerca de 1 µg do plasmídeo purificado foi linearizado com a enzima de restrição de sítio único *PmeI* (New England Biolab). A reação de linearização permaneceu em banho maria a 37 °C por 16 horas. Após este período, a enzima foi inativada por 20 minutos a 65 °C e o DNA foi precipitado com etanol.

Aproximadamente 1 µg do vetor linearizado foi adicionado em 40 µL de células competentes. As células foram incubadas em gelo por 5 minutos e submetidas a pulsos elétricos utilizando o dispositivo *Gene Pulser* (Bio-Rad) em cubeta de 0,2 cm, com os parâmetros: 1,5 Kv, 25 µF, 200 Ω, seguindo as instruções do fabricante (Bio-Rad). Foi adicionado 1 mL de sorbitol 1M e incubou-se a 30 °C por 2 horas sem agitação. As

células transformadas foram selecionadas em meio de cultivo YEPDS (Extrato de levedura 1%, Peptona 2%, Dextrose 2%, Ágar 2% e Sorbitol 1M) contendo o antibiótico zeocina em diferentes concentrações (100-500 µg/mL), identificando, assim, recombinantes multicópias do gene de interesse.

Das colônias crescidas nas placas, 11 foram testadas para avaliar as condições de expressão da proteína recombinante. O experimento contou com dois controles negativos: um na ausência de levedura, para comprovar a experimentação em condições assépticas e outro com a inoculação da levedura não transformada, para posterior comparação do padrão de bandas em SDS-PAGE da expressão.

Para a expressão da proteína recombinante as células transformadas foram mantidas em meio de cultivo complexo tamponado contendo glicerol BMGY (1% de Extrato de Levedura, 2% Peptona, 100mM de Tampão Fosfato pH 6,0, 1,34% de YNB, 4×10^{-5} % Biotina), suplementado com 1 % de glicerol como única fonte de carbono por 48 horas a 30 °C em tubos tipo falcon contendo 10 mL do meio em cada. Posteriormente, foi realizada a indução em meio complexo tamponado contendo metanol BMMY (1% de Extrato de Levedura, 2% Peptona, 100 mM de Tampão Fosfato pH 6,0, 1,34% de YNB, 4×10^{-5} % Biotina), contendo metanol como única fonte de carbono, na concentração final de 0,75 %, com 7 mL do meio em cada tubo. Todos os dias alíquotas de metanol absoluto 0,75 % foram adicionadas no meio BMMY. Foram realizadas coletas diárias (0, 24, 48, 72, 96, 120 e 144 horas) de 100 µL que, posteriormente, foram analisadas em SDS-PAGE 15 %. Para o gel, foi aplicado um volume de 20 µL de cada amostra analisada. As leveduras transformadas que mostraram maior produção das proteínas recombinantes foram selecionadas para o processo de indução em maior volume.

4.3. Expressão em *Pichia pastoris* e purificação

Após os testes de expressão em pequena escala e identificação de uma colônia produtora da proteína recombinante, esta foi utilizada para a produção da proteína recombinante em maior escala. Esta colônia foi colocada em 5 mL do meio BMGY por 24 horas, 30 °C e 250 rpm. Após este período, o volume total foi transferido para 250 mL de BMGY, em Erlenmeyer de 1 litro, e foi incubada por mais 24 horas nas mesmas condições. A DO 600 nm foi medida e a cultura foi centrifugada em Centrifuge 5810 R Eppendorf a 3000 x g, 5 minutos e 20 °C. O sobrenadante foi descartado e as células foram ressuspensas em 50 mL do meio BMMY, em Erlenmeyer de 500 mL, para a indução da expressão proteica por até 144 horas, a 30 °C e 250 rpm. Por sete dias, metanol foi adicionado ao meio a cada 24 horas para uma concentração final de 0,75% possibilitando a indução da expressão da proteína recombinante. Ao fim do processo, o meio foi centrifugado por 5 minutos, 3000 x g e a 20 °C para a separação da fração celular e o sobrenadante foi coletado e filtrado em membrana de 0,45 µm (Millipore). As amostras foram analisadas em SDS-PAGE 12% com a aplicação de cerca de 20 µL de cada amostra no gel.

Uma purificação por cromatografia de afinidade foi feita e conduzida em coluna de resina carregada com níquel HisTrap™ HP 1 mL (GE Healthcare), na qual as proteínas ficam aderidas pela cauda de histidina por afinidade química. Cinco corridas de purificação foram realizadas com cerca de 2 mL de amostra do sobrenadante contendo a rhG-CSF, para um volume final maior de amostra purificada. O equipamento utilizado foi um FPLC AKTA (GE Healthcare). Para coletar as proteínas, a coluna foi lavada com solução tampão (Tris 100 mM, NaCl 100 mM, NaH₂PO₄ 17,4 mM) com concentrações crescentes de imidazol, o qual compete com as histidinas pelo níquel, liberando, assim, a proteína recombinante na forma purificada. As amostras

foram analisadas em SDS-PAGE 12 % para confirmar as frações que apresentaram a proteína purificada. As frações proteicas foram armazenadas em tubo falcon a 4 °C.

4.4. Detecção da proteína recombinante por *Dot Blotting* e *Western Blotting*

Para detectar a proteína no sobrenadante produzido, foi realizado um *Dot Blotting* utilizando o anticorpo primário G-CSF (K-15) sc-49679 Goat Polyclonal IgG (Santa Cruz Biotechnology) e o secundário Mouse Anti-Goat IgG-AP sc-2355 AP Conjugate (Santa Cruz Biotechnology). Para tanto, pegou-se uma membrana de nitrocelulose 0,45 µm (BioRad) e marcou-se com um lápis a região que seria feito o blot. Pipetou-se 10 µL de Filgrastine rHu G-CSF 300 µg/mL (Blau) à esquerda da membrana, para atuar como controle positivo. Do lado direito da membrana pipetou-se 10 µL do sobrenadante contendo a rhG-CSF proveniente da indução em larga escala. Após a aplicação das amostras, esperou-se até a membrana secar e foi feito um bloqueio dos sítios não específicos com 5 % BSA/TBS-T (0,05% Tween 20 em TBS 1x e 0,1% BSA) por aproximadamente 1 hora. Após o bloqueio, lavou-se a membrana 3 vezes com BSA/TBS-T. A membrana foi incubada com o anticorpo primário dissolvido em BSA-TBS-T por 30 minutos a temperatura ambiente, em uma diluição de 1:200. Lavou-se a membrana por três vezes de aproximadamente 5 minutos com TBS-T. Então, a membrana foi incubada com o anticorpo secundário, em uma diluição de 1:5000, por 30 minutos a temperatura ambiente. Por último, lavou-se a membrana três vezes com TBS-T, uma vez TBS e foi feita a revelação utilizando o AP Conjugate Substrate Kit (BioRad).

Também foi feito um *Western Blotting*, para identificar qual banda representava a proteína de interesse no sobrenadante da indução realizada. Os mesmos anticorpos

primários e secundários foram utilizados neste protocolo. Separou-se em gel SDS-PAGE 15 % as seguintes amostras: *Precision Plus Protein Standards Dual Color* (Bio-Rad), Filgrastine (Blau), rhG-CSF concentrada em concentrador a vácuo RVC 2-18 CDplus (Christ), rhG-CSF não concentrada. Após a corrida do gel nas condições: 150 mA, 110 V por aproximadamente 2 horas, o mesmo foi colocado no tampão de transferência (0,25% Tris e 0,4% Glicina) sem metanol por 15 minutos, juntamente com as esponjas e os papéis de filtro. Posteriormente, misturou-se o metanol com o tampão de transferência e a membrana de nitrocelulose ficou imersa por 5 minutos. Montou-se o “sanduíche” com o gel e a membrana para fazer a transferência. A transferência foi feita usando o *Mini Trans-Blot Eletrophoretic transfer Cell System* (Bio-Rad) nos seguintes parâmetros: 150 mA por 2 horas. Após a transferência a membrana foi corada com solução Ponceau por 5 minutos e as bandas de interesse foram marcadas. Descorou-se o gel com água Mili-Q. A membrana foi incubada em solução de bloqueio por 1 hora sob agitação suave e temperatura ambiente. Lavou-se a membrana com TBS 1x por quatro vezes de 5 minutos sob agitação suave. Depois disto, a membrana foi incubada com o anticorpo primário anti-G-CSF (Santa Cruz Biotechnology) *overnight*, na diluição de 1:200. Lavou-se a membrana novamente com TBS 1x, por quatro vezes. Incubou-se a membrana, em seguida, com o anticorpo secundário Mouse Anti-Goat IgG AP (Santa Cruz Biotechnology), na diluição de 1:5000 por 1 hora e 30 minutos e depois lavou-se a membrana novamente. A membrana foi revelada utilizando o AP Conjugate Substrate Kit (Bio-Rad) e a imagem da membrana após a revelação foi capturada em fotodocumentador ChemiDocTM MP Imaging System (Bio-Rad).

5. Resultados e Discussão

5.1. Clonagem do gene da rhG-CSF em *Pichia pastoris*

A amplificação do gene da rhG-CSF, com *primers* específicos, está indicada na figura 7. Observa-se uma banda, com o tamanho esperado de aproximadamente 500 pb além de banda de massa molecular maior, provavelmente devido a uma amplificação inespecífica. As temperaturas testadas para o anelamento dos *primers* estão indicadas na legenda e a imagem indica que a amplificação ocorreu em todas estas temperaturas testadas. Desta forma, todas as reações de amplificação foram reunidas e a banda correspondente ao gene, de aproximadamente 500 pb, foi purificada do gel para que a subclonagem pudesse seguir sem problemas posteriores.

Figura 7: Fotografia de um gel de agarose 1% corado com Gel Red mostrando a amplificação do fragmento correspondente à sequência da rhG-CSF a partir do pUC57_rhG-CSF com 500 pb, com reações em diferentes temperaturas de anelamento. **M:** marcador de peso molecular GeneRuler 1Kb (Fermentas); **1:** controle negativo; **2-8:** amostras amplificadas a 50°C; 50.5°C; 51.5°C; 53.1°C; 54.9°C; 56.4°C; 57.4°C e 58°C.

A sequência amplificada foi subclonada no vetor pJET1.2./blunt, através da reação de *blunt* e da ligação do gene ao vetor. Realizou-se uma transformação em células competentes de *E. coli*, linhagem DH10β. A confirmação da clonagem e da

orientação do inserto nos clones foi realizada por PCR das colônias geradas. Os resultados das reações podem ser observados a seguir.

Figura 8: Fotografia de um gel de agarose 1%, resultado da PCR de colônia. Identificação dos clones recombinantes, clonagem em pJET1.2. **M:** marcador de peso molecular GeneRuler 1 Kb (Fermentas); **1:** controle negativo; **2,3,4,6,7,10,11:** colônias com inserto não inserido ou em orientação incorreta; **5,8,9:** colônias recombinantes. Eletroforese em gel de Agarose 1% corado com GelRed.

Os *primer forward* utilizado para a reação de identificação da orientação se hibridiza no vetor de clonagem pJET1.2 na posição 310 pb do vetor e o *primer reverse* hibridiza no final do inserto rhG-CSF. Assim sendo, as amostras recombinantes apresentam tamanho aproximado do *amplicom* igual a 610 pb.

A partir da subclonagem, foi realizada a clonagem do gene de interesse no vetor pPICZαA (Invitrogen). A ORF da rhG-CSF foi retirada do vetor pJET1.2 pela ação das enzimas de restrição *Xho*-I e *Xba*-I. O vetor também foi linearizado com as mesmas enzimas, possibilitando a clonagem *in frame* no plasmídeo de expressão. Após a reação de ligação da rhG-CSF com o pPICZαA, foi formada a construção pPICZαA_rhG-CSF. O gene da proteína foi produzido em fusão com a sequência His_(6x)-Tag na sua extremidade C-terminal.

O plasmídeo pPICZ α A_rhG-CSF foi utilizado para transformar células de *Escherichia coli* DH10 β e as colônias foram analisadas por um teste de restrição com as enzimas XbaI e XhoI (Figura 9).

Figura 9: Fotografia do E-Gel *SizeSelect Agarose 2%* (Invitrogen) em teste de restrição para identificação de clones recombinantes pPICZ α A_rhG-CSF em DH10 β . **M:** E-Gel *96 High range DNA Marker* 100 ng/lane (Invitrogen). **1 e 2:** colônias recombinantes, confirmadas pela presença da banda referente ao inserto. **Setas pretas:** indicam os pesos moleculares referentes às bandas do marcador. **Seta azul:** indica banda referente ao vetor restringido, de aproximadamente 3600 pb. **Seta vermelha:** indica banda referente ao inserto, de aproximadamente 600 pb.

Como podemos observar na figura, a banda mais baixa de aproximadamente 600 pb corresponde ao tamanho do inserto, enquanto a banda mais alta corresponde ao tamanho do vetor restringido, com tamanho de aproximadamente 3600 pb. Desta forma, pôde-se confirmar as colônias recombinantes.

Considerando que as amostras recombinantes foram encontradas, seguiu-se para a confirmação da clonagem em correta fase aberta de leitura por sequenciamento das extremidades de clonagem.

5.2. Análise do sequenciamento

Após a construção do plasmídeo recombinante, o mesmo foi sequenciado via método de Sanger. A sequência de nucleotídeos e a sequência de aminoácidos predita a partir do resultado do sequenciamento podem ser verificadas na figura 10.

Figura 10: Caracterização do gene da rhG-CSF: sequência completa do DNA codificante da proteína madura. As setas vermelhas indicam as regiões dos *primers* contendo os sítios das enzimas de restrição. A caixa azul indica a cauda de histidina. O triângulo roxo indica o início da sequência da proteína madura e o asterisco mostra o *stop* códon.

Com a análise do sequenciamento pudemos confirmar a correta clonagem do gene da rhG-CSF no plasmídeo de expressão pPICZαA. Nenhuma mutação foi inserida na sequência e não houve saída de frame, possibilitando a continuação dos experimentos.

5.3. Expressão de rhG-CSF em *P. pastoris*

Após a caracterização do plasmídeo pPICZαA_rhG-CSF, este foi linearizado e usado na transformação de células de *Pichia pastoris*, linhagem KM71H. Com o objetivo de identificar transformantes produtores da proteína recombinante, colônias positivas que cresceram em meio contendo concentração de 500 µg/mL do antibiótico zeocina, foram utilizadas para a seleção dos transformantes multicópias.

Para os ensaios de expressão da rhG-CSF, alíquotas foram retiradas a cada 24 horas de indução, as quais podem ser visualizadas da figura 11.

Figura 11: Ensaio em menor escala da expressão da rhG-CSF em *P. pastoris*. SDS-PAGE 15% mostrando em **M**: marcador de peso molecular Precision Plus Protein Standards Dual Color (Bio-Rad); **1-9**: amostra de diferentes colônias após 144 horas de indução. **Setas pretas** indicam as possíveis bandas referentes à proteína recombinante rhG-CSF, de aproximadamente 25 e 50 kDa.

A análise da expressão em menor escala mostrou que é possível observar possíveis bandas correspondentes à rhG-CSF após 24 horas de indução. Foi identificado um aumento na intensidade de uma banda de, aproximadamente, 25 kDa referente a proteína de interesse. Análises em *sílico* mostraram que o tamanho esperado da proteína

recombinante seria por volta deste valor. O mesmo aumento de intensidade foi observado para bandas de aproximadamente 50 kDa, podendo ser proveniente de uma forma mais glicosilada da proteína ou até mesmo resultado de formação de dímeros entre as moléculas. Pode-se comparar também o padrão de bandas entre o perfil de expressão de uma cultura recombinante para rhG-CSF de 0 a 144 horas com o perfil de expressão de uma cultura de *Pichia pastoris* não transformada (Figura 12). Nota-se que há uma diferença significativa entre as culturas, sustentando a ideia de que as bandas indicadas são referentes à proteína recombinante em questão.

Figura 12: Ensaio do perfil de expressão de *P.pastoris*. Linhagem não transformada. SDS-PAGE 15% mostrando em **M**: marcador de peso molecular Precision Plus Protein Standards Dual Color (Bio-Rad); **1**: amostra de 0h de indução; **2**: amostra após 24h de indução; **3**: amostra após 48h de indução; **4**: amostra após 72h de indução; **5**: amostra após 96h de indução; **6**: amostra após 120h de indução; **7**: amostra após 144h de indução; **8**: amostra após 144h de indução de *P. pastoris* não transformada. Setas vermelhas indicam bandas de 25 kDa e 50 kDa, referente às bandas das amostras 1-8.

5.4. Purificação

A purificação da rhG-CSF foi realizada por cromatografia de afinidade em coluna HisTrapTM HP 1 mL (GE Healthcare) contendo Níquel no equipamento FPLC Akta (GE Healthcare). A figura 10 mostra as frações de purificação frente à

concentrações crescentes de imidazol. Vê-se que a proteína começou a ser eluída em concentrações pequenas de imidazol, porém o pico referente à proteína foi identificado nas amostras 8 e 9 (Figura 13). As frações de lavagem aparecem praticamente sem bandas aparentes e o volume morto de coluna (*void*) aparece um pouco mais carregado, podendo conter ainda alguma quantidade da proteína recombinante que não se ligou na coluna. A figura 14 mostra as frações referentes aos picos de absorbância de outras cromatografias realizadas nas mesmas condições.

Figura 13: Análise da purificação da rhG-CSF por cromatografia de afinidade ao níquel. SDS-PAGE 12% corado com Comassie Blue. Frações da primeira corrida, com aproximadamente 2 mL de sobrenadante, mostrando em **M**: marcador de peso molecular Precision Plus Protein Standards Dual Color (Bio-Rad); **1**: volume morto de coluna; **2**: lavagem; **3**: lavagem; **4**: lavagem; **5**: amostra eluída em 5% de imidazol; **6**: amostra eluída com 10% de imidazol; **7**: amostra eluída com 15% de imidazol; **8**: amostra eluída com 20% de imidazol; **9**: amostra eluída com 25% de imidazol. “Seta pretas” indicam a proteína possivelmente glicosilada, com tamanho de aproximadamente 50 kDa e a proteína não glicosilada, de aproximadamente 25 kDa.

Figura 14: Análise da purificação da rhG-CSF por cromatografia de afinidade ao níquel. SDS-PAGE 12% corado com Comassie Blue. Frações da segunda, terceira, quarta e quinta corridas, como aproximadamente 2 mL de sobrenadante cada, mostrando em **M**: marcador de peso molecular Precision Plus Protein Standards Dual Color (Bio-Rad); **1**: amostra anterior ao pico da segunda corrida; **2**: amostra do pico da segunda corrida; **3**: amostra do pico da terceira corrida; **4**: amostra anterior ao pico da terceira corrida; **5**: amostra anterior ao pico da quarta corrida; **6**: amostra do pico da quarta corrida; **7**: amostra anterior ao pico da quinta corrida; **8**: amostra do pico da quinta corrida. “Setas pretas” indicam a proteína possivelmente glicosilada, com tamanho de aproximadamente 50 kDa e a proteína de aproximadamente 25 kDa.

A proteína purificada aparece com massa molecular de aproximadamente 25 kDa e também de 50 kDa, o dobro da prevista. Este resultado indica que a proteína foi, provavelmente glicosilada pela *P. pastoris*, o que é esperado devido a presença de sítios de O-glicosilação na sequência de aminoácidos da rhG-CSF. Também pode ter ocorrido a formação de dímeros de proteína devido à presença de pontes dissulfeto intercadeias e um resíduo de cisteína livre presente na posição 17.

Pode-se observar que houve um padrão para todas as corridas realizadas. Todos os picos foram obtidos em uma concentração de imizadol de aproximadamente 20 %, como mostra a figura 13. Provavelmente, todas as bandas aparentes são referentes à rhG-CSF, sendo as que estão acima de 25 kDa são as formas glicosiladas ou diméricas.

O gráfico abaixo mostra o perfil da cromatografia de afinidade para a purificação da proteína rhG-CSF.

Figura 15: Perfil da cromatografia por afinidade ao Níquel para a purificação da rhG-CSF. Eluente B se refere ao Tampão de Lise com Imidazol (Tris 100 mM, NaCl 100 mM, NaH_2PO_4 17.4 mM, Imidazol 250 mM).

5.5. Detecção da proteína recombinante por *Western Blotting* e *Dot Blotting* com anticorpo anti-G-CSF

Para detectar a expressão da proteína recombinante no sobrenadante da indução em *P. pastoris*, amostras da expressão no tempo de 96h tiveram suas frações proteicas separadas por SDS-PAGE 12% e depois transferidas para uma membrana de nitrocelulose (Bio-Rad).

O anticorpo primário G-CSF (K-15) sc-49679 Goat Polyclonal IgG (Santa Cruz Biotechnology) foi eficiente para o reconhecimento da proteína recombinante em uma diluição de 1:200. O *Western Blotting* realizado com estes anticorpos específicos mostrou que o anticorpo foi capaz de reconhecer a proteína na amostra utilizada e também foi específico. Além de reconhecer a proteína recombinante, foi eficiente na detecção do medicamento recombinante Filgrastine rHu G-CSF 300 $\mu\text{g/mL}$ (Blau) que

foi utilizada como um controle positivo (Figura 16). As banda marcadas possuem um tamanho de aproximadamente 50 kDa, de acordo com o marcador de peso molecular. Isto indica que a proteína produzida pode estar glicosilada ou que haja a formação de dímeros, já que o tamanho esperado da proteína recombinante é de aproximadamente 20 kDa. As bandas de menor tamanho não foram marcadas pelo anticorpo, possivelmente devido à baixa quantidade das mesmas.

Figura 16: Membrana de nitrocelulose mostrando um *Western Blotting* da proteína expressa em *Pichia pastoris*. As frações proteicas foram incubadas com o anticorpo primário anti-G-CSF na diluição de 1:200. **M:** marcador de peso molecular Precision Plus Protein Standards Dual Color (Bio-Rad); **1:** amostra da rhG-CSF após 96h de indução; **2:** amostra do sobrenadante contendo a rhG-CSF após 96h de indução; **3:** amostra do sobrenadante contendo a rhG-CSF após 96h de indução concentrada; **4:** Filgrastima comercial. A **seta preta** indica as bandas referentes a proteína recombinante e a **seta azul** indica a banda referente à Filgrastima comercial.

A detecção da proteína recombinante no sobrenadante de *P. pastoris* também foi feita por *Dot Blotting* com amostras de sobrenadante após 96 horas de indução. O anticorpo primário G-CSF (K-15) sc-49679 Goat Polyclonal IgG (Santa Cruz Biotechnology) foi eficiente contra a proteína recombinante em uma diluição de 1:200 (Figura 17).

O círculo marcado referente a rhG-CSF apareceu bem mais claro do que o círculo referente à Filgrastima comercial utilizada, muito provavelmente, devido à baixa concentração da amostra utilizada.

Figura 17: Membrana de nitrocelulose mostrando um Dot Blotting da proteína incubada com o anticorpo primário na concentração de 1:200 mostrando em **1**: amostra Filgrastima comercial e **2**: amostra da rhG-CSF após 96h de indução da expressão.

6. Conclusão

Podemos concluir neste capítulo que é possível a produção da rhG-CSF em sistema de expressão *Pichia pastoris*. Estes resultados mostram os primeiros passos da produção da rhG-CSF em *P. pastoris* e isto pode ser muito importante para a estratégia de tornar o Brasil um produtor deste biofármaco futuramente.

Capítulo II – Delineamento experimental e caracterização de linhagem.

RESUMO

ARCHANGELO, B. S. **Expressão do fator estimulador de colônia de granulócito humano recombinante (rhG-CSF) em sistema de *Pichia pastoris* e otimização do bioprocesso.** 2017. 120f. Dissertação (Mestrado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2017.

Delineamento experimental é um instrumento empregado para avaliar a relação de causa e efeito de um determinado processo, através de vários testes nos quais são realizadas mudanças intencionais nas variáveis de entrada, que acaba por gerar alterações diferenciadas na variável de resposta deste mesmo processo. Dentre as opções de delineamentos experimentais, os experimentos fatoriais são bastante triviais. Uma alternativa viável ao experimento fatorial completo, que requer um número de ensaios muito grande, é um delineamento de seleção ou *Screening Design* que se baseia em delineamentos prévios fundamentados em frações fatoriais, dos fatores e/ou níveis, que realmente levam a respostas experimentais ótimas. Dentro destes está o planejamento Plackett-Burman. Este é classificado como fatorial fracionado de dois níveis, tendo disponíveis matrizes para qualquer n múltiplo de quatro fatores. Possibilita uma simetria que permite que os efeitos principais de cada fator sejam individualmente determinados. Este tipo de planejamento é bastante útil na seleção de fatores. Planejamentos baseados em delineamentos compostos também são interessantes. O Delineamento Composto Central (DCC) é um dos mais populares para o ajuste de modelos de até segunda ordem e é constituído por três partes: cúbica (fatorial), axial (alpha) e central. Os pontos axiais contribuem na estimativa dos termos quadráticos do modelo de segunda ordem. Os responsáveis pela estimativa dos efeitos de interação são os pontos cúbicos, que são combinações geralmente utilizadas em ajustes de modelos de primeira ordem com ou sem interações duplas entre os fatores. O ponto central possibilita a estimativa do erro puro, quando ele é repetido, além de ser responsável pela estimativa dos efeitos quadráticos. Esta parte do trabalho visou estabelecer melhores parâmetros de indução e expressão da rhG-CSF, produzida em *Pichia pastoris*, em diferentes condições de cultivo, além de uma caracterização da linhagem produtora da proteína recombinante. O planejamento Plackett-Burman foi aplicado para selecionar as variáveis significantes, que foram o pH e a adição de Sorbitol e aminoácidos no meio de cultura. Os dados obtidos passaram por análise estatística que permitiu esta seleção. Tendo estas sido selecionadas como estatisticamente significativas, o Delineamento Composto Central foi aplicado para uma análise em níveis maiores das variáveis significativas. Também foi feita uma caracterização de linhagem para que a colônia selecionada fosse caracterizada como uma linhagem produtora. Apesar das três variáveis terem sido significativas (pH, sorbitol e aminoácidos), apenas o aumento do valor de pH apresentou um efeito benéfico para a produção da rhG-CSF. Curvas de crescimento celular foram desenhadas, além do cálculo da velocidade máxima de crescimento e tempo de duplicação.

Palavras-chave: Screening design, delineamento experimental, Plackett-Burman, Delineamento Composto Central.

ABSTRACT

ARCHANGELO, B. S. **Expression of human granulocyte colony stimulating factor (rhG-CSF) in *Pichia pastoris* and the optimization of the bioprocess.** 2017. 120f. Dissertation (Master). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2017.

Experimental design is an instrument used to evaluate the cause-effect relationship of a given process through several tests in which intentional changes are made to the input variables. These intentional changes result in differential alterations in the response variable of this process. Among the options of experimental designs, the factorial experiments are quite trivial. These experiments include two or more factors and each combination is considered a treatment. A feasible alternative to the complete factorial experiment, which requires a very large number of tests, is a selection design or Screening Design which is based on previous delineations reasoned on factorial fractions, factors and/or levels, which actually lead to optimal experimental responses. The Plackett-Burman planning is classified as a fractional factorial of two levels with matrices available for any n multiple of four factors. It gives a symmetry that allows the main effects of each factor to be individually determined, assuming that the effects of interactions are negligible without affecting the quality of the results. This type of planning is very useful in selecting factors. Another interesting planning is the composite design. The Central Composite Design (DCC) is one of the most popular for the adjustment of models up to second order and it is constituted of three parts: cubic (factorial), axial (alpha) and central. The axial points contribute to the estimation of the quadratic terms of the second order model, thus not contributing to the estimation of the effects of interactions. Those responsible for this are the cubic points, which are combinations commonly used in adjustment of first order models with or without double interactions between the factors. The central point allows the estimation of the pure error, when it is repeated, besides being responsible for the estimation of the quadratic effects. This part of the project aimed to establish best parameters of induction and expression of rhG-CSF in *Pichia pastoris*, under different culture conditions, as well as a characterization of the producing line of the recombinant protein. Plackett-Burman planning was applied to select the significant variables, which were the pH and the addition of sorbitol and amino acids in the culture medium. The data obtained underwent statistical analysis that allowed this selection. The statistically significant variables were chosen and the Central Composite Design was applied for an analysis at higher levels of the significant variables. A lineage characterization was done so that the selected colony was characterized as a producer lineage. Although all three variables were significant (pH, sorbitol and amino acids) only the increase in pH had a beneficial effect on rhG-CSF production. Cell growth curves were made in addition to calculating the maximum growth time and doubling time.

Key words: Screening Design, experimental design, Plackett-Burman, Central Composite Design (DCC).

LISTA DE FIGURAS

Figura 1. Variáveis incluídas para o estudo de sua influência na produção de rhG-CSF usando-se o Delineamento Plackett-Burman.....	52
Figura 2. Esquema representando os passos do protocolo do Kit ELISA G-CSF Humano.....	55
Figura 3. Variáveis incluídas para o estudo de sua influência na produção de rG-CSF usando-se o Delineamento Composto Central (DCC).....	56
Figura 4. Esquema geral mostrando todos os passos do delineamento experimental....	60
Figura 5. Curva padrão Hu G-CSF pg/mL.....	67
Figura 6. Gráfico em colunas mostrando a diferença de concentração da rhG-CSF entre os 16 ensaios realizados.....	68
Figura 7. Efeitos estimados e coeficientes para proteína em pg/mL. Minitab.....	70
Figura 8. Gráfico de Pareto mostrando a relação entre a produção de proteína rhG-CSF e as variáveis testadas. A: Temperatura; B: pH; C: % metanol; D: Sorbitol; E: Aminoácidos; F: Tween 80.....	71
Figura 9. Gráfico da distribuição normal dos resíduos.....	72
Figura 10. Análise de superfície de resposta para as variáveis pH e aminoácidos.....	73
Figura 11. Curva padrão Hu G-CSF pg/mL para o DCC.....	77
Figura 12. Gráfico de colunas mostrando a diferença de concentração da rhG-CSF entre os 13 ensaios realizados.....	78
Figura 13. Efeitos estimados e coeficientes para proteína em pg/mL. Minitab.....	79
Figura 14. Análise de superfície de resposta para as variáveis pH e Sorbitol. Resultado do DCC.....	80

- Figura 15.** Perfil da cromatografia por afinidade ao Níquel para a purificação da rhG-CSF. Eluente B se refere ao Tampão de Lise com Imidazol (Tris 100 mM, NaCl 100 mM, NaH_2PO_4 17.4 mM, Imidazol 250 mM).....82
- Figura 16.** Membrana de nitrocelulose mostrando um Dot Blotting da proteína incubada com o anticorpo primário na concentração de 1:200 mostrando em **1**: amostra Filgrastima comercial e **2**: amostra da rhG-CSF após os dez repiques para confirmar linhagem produtora.....83
- Figura 17.** Curva de crescimento celular da colônia estoque e da colônia após os dez repiques (caracterização). O sinal + representa a caracterização e o sinal x representa o estoque.....84
- Figura 18.** Regressão linear da amostra estoque e da amostra caracterização, mostrando as equações referentes a cada uma das retas.....85

LISTA DE TABELAS

Tabela 1. Valores utilizados no Delineamento Plackett-Burman.....	53
Tabela 2. Matriz de dados codificados para as variáveis estudadas.....	53
Tabela 3. Valores utilizados no Delineamento Composto Central (DCC).....	57
Tabela 4. Matriz de dados codificados para as variáveis estudadas.....	57
Tabela 5. Descrição das variáveis referentes a cada um dos ensaios mostrados na Figura 6.....	69
Tabela 6. Valores das variáveis referentes a cada um dos ensaios.....	78
Tabela 7. Tabela comparativa entre a amostra Estoque e a amostra Caracterização, mostrando a velocidade de crescimento específica, o tempo de duplicação e o rendimento final de cada.....	86

Sumário Capítulo II

Capítulo II – Delineamento experimental e caracterização de linhagem.....	i
Resumo.....	i
Abstract	ii
LISTA DE FIGURAS.....	iii
LISTA DE TABELAS.....	v
1. Introdução.....	51
1.1. Planejamento experimental.....	51
1.1.1. Experimentos fatoriais	51
1.1.1.1. Planejamento Plackett-Burman	52
1.1.1.2. Delineamento Composto Central.....	54
2. Objetivos	58
2.1. Objetivo Geral	58
2.2. Objetivos Específicos.....	58
3. Materiais e métodos	59
3.1. Materiais	59
3.2. Métodos	59
3.2.1. Planejamento experimental para testar as diferentes condições físico-químicas para a indução da rhG-CSF	59
3.2.1.1. Diferentes condições físico-químicas para a indução da rhG-CSF	59
3.2.1.2. Delineamento Plackett-Burman (<i>Screening Design</i>)	59
3.2.1.3. Delineamento Composto Central (DCC).....	64
3.2.2. Caracterização da linhagem	68
4. Resultados e discussão.....	72
4.1. Planejamento experimental para testar as diferentes condições físico-químicas para a indução da rhG-CSF	72
4.1.1. Diferentes condições físico-químicas para a indução da rhG-CSF	72
4.1.2. Planejamento Plackett-Burman (<i>Screening Design</i>)	74
4.1.3. Delineamento Composto Central (DCC).....	84
4.2. Caracterização de linhagem	89
5. Conclusão	96
CONSIDERAÇÕES FINAIS.....	97
REFERÊNCIAS BIBLIOGRÁFICAS	98

1. Introdução

1.1. Planejamento experimental

Delineamento experimental, do inglês, *Design of Experiments*, é um instrumento empregado para avaliar a relação de causa e efeito de um determinado processo, através de vários testes nos quais são realizadas mudanças intencionais nas variáveis de entrada, que acaba por gerar alterações diferenciadas na variável de resposta deste mesmo processo. Os objetivos deste delineamento são: determinar quais são as variáveis mais influentes do processo na variável resposta; definir os níveis de ajuste dos fatores para que a resposta tenha um valor próximo do valor pretendido; determinar os níveis dos fatores para que a variabilidade no valor da resposta seja o menor possível; determinar os níveis de ajuste dos fatores controláveis para que os efeitos das variáveis incontroláveis sejam diminuídos (Montgomery, 2009; Bell et al, 2009).

1.1.1. Experimentos fatoriais

Dentre as opções de delineamentos experimentais existentes os experimentos fatoriais são bastante triviais. Os experimentos fatoriais abrangem combinações entre dois ou mais fatores, sendo que cada combinação se refere a um tratamento.

O arranjo fatorial completo permite analisar todos os possíveis ensaios, uma vez que o número de ensaios é igual ao número de níveis elevado ao número de fatores, representado por N^k , onde N é o número de níveis de cada fator e k é o número de fatores em questão. À medida que se aumenta o número dos fatores neste tipo de esquema fatorial completo, aumenta-se rapidamente o número de ensaios, o que acaba por impossibilitar a utilização deste tipo de arranjo para 6 ou mais fatores com 2 níveis

(64 ensaios) e para 4 ou mais fatores com 3 níveis (81 ensaios) (Box, Hunter e Hunter, 2005; Montgomery, 2009; Rodrigues e Iemma, 2009).

Desta forma, uma alternativa viável é o uso de um *Screening Design* ou delineamento de seleção. Segundo Rodrigues e Iemma (2009), uma solução para o problema do rápido aumento no número de ensaios em esquemas fatoriais completos é a utilização de um *Screening Designs*, através de delineamentos prévios fundamentados em frações fatoriais, dos fatores e/ou níveis, que realmente levam a respostas experimentais ótimas. Este tipo de planejamento normalmente é realizado em estágios iniciais, nos quais é provável que certos fatores analisados possuam pouco ou nenhum efeito significativo sobre a variável resposta, podendo, assim, serem desconsiderados do planejamento. Dentre os delineamentos de seleção estão os fatoriais fracionados e planejamento Plackett-Burman.

1.1.1.1. Planejamento Plackett-Burman

Nos processos multivariáveis, numerosos fatores potencialmente influentes estão envolvidos, portanto, é essencial determinar quais são os mais fundamentais. É necessário submeter o processo a uma triagem inicial, *Screening Design*, antes da otimização. A metodologia de Plackett-Burman é uma das ferramentas empregadas para essa triagem inicial, proporcionando um pequeno número de ensaios e tornando possível a determinação da influência de vários fatores, evitando os esquemas fatoriais completos que acarretam em tempo e custos inviáveis (Plackett e Burman, 1946; Box, Hunter e Hunter, 2005).

O planejamento Plackett-Burman (PB) é classificado como delineamento fatorial fracionado de dois níveis, tendo disponíveis matrizes para qualquer n múltiplo de quatro

fatores (12, 16, 20, 24...ensaios). Nestes delineamentos, todas as colunas são ortogonais e cada uma delas contém o mesmo número de sinais positivos, para os níveis superiores, e sinais negativos, para os níveis inferiores. Esta ortogonalidade possibilita uma simetria que permite que os efeitos principais de cada fator sejam individualmente determinados, admitindo-se que os efeitos de interação sejam desprezíveis e, desta forma, certamente dispensáveis os cálculos dessas interações sem prejudicar a qualidade dos resultados. Este tipo de planejamento é bastante útil na seleção de fatores (Box, Hunter e Hunter, 2005; Barros Neto et al, 2007).

O delineamento PB assume que não haja interações entre os diferentes componentes de média na série de variáveis, considerando satisfatório para a triagem uma abordagem linear. É uma concepção fatorial fracionária e o efeito principal pode ser calculado pela diferença entre a média das medições feitas no nível alto (1) do fator e a média de medições no nível baixo (-1) (Box, Hunter e Hunter, 2005).

Segundo Rodrigues e Iemma (2209) é mais interessante dar início às análises de triagem com um fatorial fracionado ou Plackett-Burman, analisando os efeitos principais das variáveis e depois realizar outros experimentos, diminuindo o número de variáveis e alterando as faixas de estudo em função do impacto que elas tiveram sobre as respostas.

A utilização de um planejamento fatorial de um grupo limitado de variáveis é vantajosa em relação ao método mais comum de estudo de um único parâmetro por ensaio, já que com esta última abordagem não se consegue verificar e quantificar efeitos sinérgicos e antagônicos entre os fatores em questão. Isto pode acarretar falhas na localização das condições ótimas. Adicionalmente, o planejamento fatorial torna

possível tirar proveito do conhecimento prático sobre o processo durante a análise final da superfície de resposta (Box, Hunter e Hunter, 2005).

Barros Neto et al (2007) mencionam que a desvantagem na utilização dos planejamentos PB é a complexidade das relações entre os contrastes calculados e os efeitos de um fatorial completo. Isto requer conhecimentos matemáticos específicos, como Álgebra e Estatística-Matemática, matrizes de Hadamard e funções de Legrange, segundo o enfoque Plackett-Burman.

Uma recomendação relevante para quando se usa este tipo de metodologia é escolher um delineamento PB com no mínimo quatro ensaios a mais do que o número de fatores, assim os três graus de liberdade restantes podem ser usados para estimar os valores calculados para os efeitos principais. Outra recomendação imprescindível é a realização de pelo menos três repetições na condição do ponto central, para que haja graus de liberdade no cálculo do Erro Padrão devido à ausência de repetição dos ensaios. Desta forma, é possível estimar o erro puro e, assim, avaliar a reprodutibilidade do processo; além de verificar a existência de curvatura quando os valores das respostas forem aproximadamente iguais (Rodrigues e Iemma, 2009).

1.1.1.2. Delineamento Composto Central

Muitas tentativas têm sido feitas a partir de delineamentos de tratamentos mais complexos, que utilizam todas as combinações dos fatores em estudo, para tentar reduzir o número de pontos experimentais através de metodologias como a de confundimento e de repetição fracionada. Além destas, foi criada a técnica que utiliza os delineamentos compostos, que reduzem o número de pontos experimentais. Este tipo de delineamento foi desenvolvido, inicialmente, por Box e Wilson (1951) com o objetivo

de estudar funções polinomiais de resposta na indústria, onde o erro experimental é normalmente pequeno e as condições do experimento são mais facilmente controladas. Desta forma, é comum repetir apenas um tratamento relativo ao ponto central.

Dentre os delineamentos padrões, o Delineamento Composto Central é considerado um delineamento ótimo (Mateus et al, 2001).

O Delineamento Composto Central (DCC) é um dos mais populares para o ajuste de modelos de até segunda ordem (Myers e Montgomery, 1995). O mesmo é constituído por três partes: cúbica (fatorial), axial (α) e central. Em geral, um DCC para k fatores, cujos níveis codificados serão combinados, é dado pelo conjunto constituído pelos seguintes tratamentos com níveis codificados: $(1,1,...,1)$; $(-1,1,...,-1)$; ...; $(-1,-1,...,-1)$ referentes à parte cúbica, totalizando 2^k combinações; $(-\alpha, 0,..., 0)$; $(\alpha, 0,..., 0)$; $(0, -\alpha,..., 0)$; $(0, \alpha,..., 0)$; ...; $(0, 0,..., -\alpha)$; $(0, 0,..., \alpha)$ referentes à parte axial, totalizando $2k$ combinações e $(0, 0,..., 0)$ referente ao ponto central. No total, são $k^2 + 2k + 1$ combinações entre os níveis dos k fatores estudados (Mendonça, 2012).

Os pontos axiais são situados nos eixos do sistema de coordenadas com distância $\pm \alpha$ da origem. São esses pontos que contribuem, em maior parte, na estimativa dos termos quadráticos do modelo de segunda ordem, não contribuindo, assim, para a estimativa dos efeitos das interações. Os responsáveis por isto são os pontos cúbicos, que são combinações geralmente utilizadas em ajustes de modelos de primeira ordem com ou sem interações duplas entre os fatores. O ponto central possibilita a estimativa do erro puro, quando ele é repetido, além de ser responsável pela estimativa dos efeitos quadráticos (Myers e Montgomery, 1995).

A natureza sequencial deste tipo de delineamento é claramente observada. Os pontos fatoriais contribuem para a estimativa dos termos lineares e interações entre os

fatores e os pontos centrais informam sobre a existência de curvatura no sistema. Caso seja encontrada uma curvatura, a adição dos pontos axiais permite uma estimativa dos termos quadráticos.

Em um DCC, qualquer valor pode ser adotado para α e para o número de repetições do ponto central. Porém, um DCC com dois fatores apresenta propriedades especiais quando α é igual a 1,4142 e o número de pontos centrais é igual a cinco. Neste caso, o planejamento assume propriedades de rotacionalidade e ortogonalidade, o que facilita o tratamento estatístico dos resultados (Diamond, 1989; Perázio, 2010).

Aperfeiçoar o bioprocesso tem se mostrado cada vez mais importante para a produção de biofármacos. Desta forma, realizar um planejamento experimental para testar as diferentes condições físico-químicas para a indução da rhG-CSF se torna bastante relevante, tanto para melhorar a produção da proteína específica quanto para auxiliar e contribuir com novos estudos nessa área.

Torna-se necessário o estudo de novas metodologias que permitam um melhor rendimento final dessa proteína, estabelecendo-se protocolos que visem melhorar o processo de indução, expressão e, assim, a produção mais eficiente da mesma.

2. Hipótese

É possível se determinar uma linhagem produtora de proteína heteróloga em *Pichia pastoris*. Também é possível melhorar a produção da mesma com modificações no meio de cultura e em condições físico-químicas durante a fase indução da expressão.

3. Objetivos

3.1.Objetivo Geral

Estabelecer parâmetros de indução e expressão da rhG-CSF, produzida em *Pichia pastoris*, em diferentes condições de cultivo. Pretende-se também realizar uma caracterização da linhagem produtora da proteína recombinante.

3.2.Objetivos Específicos

- Expressar a proteína sob a indução de diferentes temperaturas, pH, concentrações de metanol e tempos de indução, além de modificações no meio de cultivo e adição de surfactantes;
- Aplicar a metodologia de Plackett-Burman e Delineamento Composto Central;
- Análise estatística dos dados;
- Avaliar a estabilidade genética da linhagem recombinante;
- Montar uma curva de crescimento celular em função do tempo;
- Calcular a velocidade máxima de crescimento e tempo de duplicação da linhagem recombinante.

4. Materiais e métodos

4.1. Materiais

O estoque primário das linhagens de *Pichia pastoris* KM71H, assim como o vetor de expressão pPICZ α A, foram fornecidos pelo grupo de pesquisa da Profa. Eliane Candiani Arantes, do Departamento de Física e Química da faculdade de Ciências Farmacêuticas de Ribeirão Preto – USP. As células de *P. pastoris* foram mantidas em YPD e Glicerol 20% em freezer -80 °C. Os reagentes utilizados foram de grau analítico.

4.2. Métodos

4.2.1. Planejamento experimental para testar as diferentes condições físico-químicas para a indução da rhG-CSF

4.2.1.1. Diferentes condições físico-químicas para a indução da rhG-CSF

As condições físico-químicas que serão analisadas nos ensaios serão a temperatura durante o período de indução, o pH do meio de cultivo e a concentração de metanol adicionado no meio de cultivo. Também será analisado a adição de Sorbitol, aminoácidos e Tween 80 no meio de cultura durante a fase de indução da proteína recombinante.

4.2.1.2. Delineamento Plackett-Burman (*Screening Design*)

A estratégia utilizada inicialmente para analisar as respostas das diferentes condições físico-químicas durante a indução da rhG-CSF foi um delineamento Plackett-Burman (PB). Neste delineamento foram avaliados estatisticamente os efeitos de seis

variáveis, visando à seleção dos fatores mais importantes a serem considerados em estudos posteriores de modelagem e otimização (Figura 1).

Figura 1: Variáveis incluídas para o estudo de sua influência na produção de rhG-CSF usando-se o Delineamento Plackett-Burman.

Seguindo as recomendações da escolha de um delineamento Plackett-Burman com um número mínimo de quatro ensaios a mais do que o número de variáveis a serem analisadas, propostas por Rodrigues e Iemma (2009), foi realizado um estudo adotando um PB de 12 ensaios, acrescidos de 4 pontos centrais (pontos executados com réplicas e dará uma estimativa interna do erro puro e de pontos axiais, que irão determinar os termos quadráticos), totalizando 16 ensaios para a avaliação de 6 fatores. Os fatores testados foram: temperatura, pH e concentração de metanol durante o período de indução e adição de sorbitol, aminoácidos e tween 80 no meio de cultura. Os níveis testados para cada fator estão representados na Tabela 1.

Tabela 1 – Valores utilizados no Delineamento Plackett-Burman (PB)

Variáveis	Código	Intervalo de Variação	Nível		
			-1	0	+1
Temperatura	X1	10	25	30	35
pH	X2	1	5	6	7
Metanol	X3	0,25	0,5	0,75	1
Sorbitol	X4	0,5	0	0,5	1
Aminoácidos	X5	10	0	10	20
Tween	X6	80	80	160	240

Este delineamento PB é uma metodologia de *Screening Design* para a seleção de variáveis em que se têm vários fatores e pouco conhecimento das condições otimizadas, permitindo analisar simultaneamente os efeitos destas variáveis. Desta forma, reduz-se drasticamente o número de ensaios que deveriam ser realizados em um fatorial completo, sendo imprescindível a realização de no mínimo três repetições do ponto central.

A matriz de dados utilizados na montagem dos ensaios está apresentada na Tabela 2. Cada linha da mesma representa um ensaio realizado, mostrando também os valores referentes a temperatura, pH, metanol, Sorbitol, aminoácidos e Tween que foram utilizados em cada um dos ensaios.

Tabela 2 – Matriz de dados codificados para as variáveis estudadas

A	B	C	D	E	F	TEMPER	pH	METANOL	SORB	AMINOAC	TWEEN
-1	1	1	1	-1	1	25	7	1,00	1,0	0	240
-1	1	-1	-1	-1	1	25	7	0,50	0,0	0	240
-1	-1	-1	-1	-1	-1	25	5	0,50	0,0	0	80
0	0	0	0	0	0	30	6	0,75	0,5	10	160
0	0	0	0	0	0	30	6	0,75	0,5	10	160
1	1	1	-1	1	1	35	7	1,00	0,0	20	240
0	0	0	0	0	0	30	6	0,75	0,5	10	160
1	-1	1	-1	-1	-1	35	5	1,00	0,0	0	80
-1	-1	1	1	1	-1	25	5	1,00	1,0	20	80
1	-1	1	1	-1	1	35	5	1,00	1,0	0	240
-1	-1	-1	1	1	1	25	5	0,50	1,0	20	240
0	0	0	0	0	0	30	6	0,75	0,5	10	160
1	1	-1	1	-1	-1	35	7	0,50	1,0	0	80
1	-1	-1	-1	1	1	35	5	0,50	0,0	20	240
1	1	-1	1	1	-1	35	7	0,50	1,0	20	80
-1	1	1	-1	1	-1	25	7	1,00	0,0	20	80

Todas as variáveis foram testadas durante a indução da rhG-CSF, que ocorreu da seguinte forma. As células foram induzidas em meio de cultivo complexo tamponado BMGY (1% de Extrato de Levedura, 2% Peptona, 100mM de Tampão Fosfato pH 6,0, 1,34% de YNB, $4 \times 10^{-5}\%$ Biotina), suplementado com 1 % de glicerol como única fonte de carbono, por 48 horas a 30 °C em tubos tipo falcon 50 mL Corning *Mini Bioreactor* contendo 10 mL do meio em cada. Posteriormente foi realizada a indução em meio complexo tamponado BMMY (1% de Extrato de Levedura, 2% Peptona, 100mM de Tampão Fosfato pH 6,0, 1,34% de YNB, $4 \times 10^{-5}\%$ Biotina), contendo metanol como única fonte de carbono, com 7 mL do meio em cada tubo. As variáveis testadas só foram modificadas durante a fase de produção da proteína recombinante pelas células, que ocorre durante a indução da expressão em meio complexo contendo metanol BMMY. Portanto, para todos os ensaios, a fase de crescimento em BMGY se manteve a mesma, como descrito anteriormente. Todos os dias alíquotas de metanol absoluto foram adicionados no meio BMMY, que variaram de acordo com o proposto em cada ensaio (Tabela 2). Foram realizadas coletas diárias (0, 24, 48, 72, 96, 120 e 144 horas) de 100 µL que, posteriormente, foram analisadas em SDS-PAGE 12 %.

Ao fim do processo, o meio foi centrifugado por 8 minutos, 3200 x g e a 20 °C para a separação da fração celular e o sobrenadante foi coletado e filtrado em membrana de 0,45 µm (Millipore).

Para quantificar as proteínas produzidas em cada um dos ensaios, utilizou-se o Kit ELISA G-CSF Humano (Novex *by Life Technologies*), segundo recomendações do fabricante (Figura 2). Este kit é utilizado para determinação quantitativa *in vitro* da G-CSF humana natural ou recombinante em soro humano, plasma, solução tampão ou meio de cultura celular. Esta metodologia foi aplicada, pois nos permite avaliar

10% de significância, realizado no software Minitab 17 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com) Minitab.

4.2.1.3. Delineamento Composto Central (DCC)

A partir da análise inicial dos efeitos principais pelo Plackett-Burman, outro modelo foi escolhido e aplicado para uma análise mais profunda das variáveis significantes apontadas.

Neste segundo ensaio, a estratégia utilizada para analisar as respostas das diferentes condições físico-químicas durante a indução da rhG-CSF, foi um Delineamento Composto Central (DCC). Neste delineamento foram avaliados os efeitos de duas variáveis, que foram estatisticamente significantes de acordo com o ensaio de Plackett-Burman realizado anteriormente.

Figura 3: Variáveis incluídas para o estudo de sua influência na produção de rhG-CSF usando-se o Delineamento Composto Central (DCC).

Seguindo as recomendações da escolha de um Delineamento Composto Central (DCC), foi realizado um estudo adotando um DCC de 8 ensaios, acrescidos de 5 pontos

centrais, totalizando 13 ensaios para a avaliação de dois fatores. Os fatores testados foram: pH e adição de Sorbitol no meio de cultura. Os níveis testados para cada fator estão representados na tabela 3.

Tabela 3: Valores utilizados no Delineamento Composto Central (DCC).

Variáveis	Código	Intervalo de Variação	Nível				
			$-\alpha$	-1	0	+1	$+\alpha$
pH	Y1	1	4,59	5	6	7	7,41
Sorbitol	Y2	0,75	0	0,25	1	1,75	2

Este delineamento permite analisar simultaneamente os efeitos de variáveis em vários níveis ($-\alpha$, -1, 0, 1, α). Com ela, reduz-se o número de ensaios que deveriam ser realizados em um fatorial completo, sendo imprescindível a realização de algumas repetições do ponto central e um valor de α definido, preferencialmente para 1,4142.

A matriz de dados utilizados na montagem dos ensaios está representada na Tabela 4.

Tabela 4 – Matriz de dados codificados para as variáveis estudadas.

A	B	Sorbitol	pH
0	0	1	6
-1	1	0,25	7
0	0	1	6
0	0	1	6
0	0	1	6
1	1	1,75	7
-1	-1	0,25	5
0	1,41	1	7,41
1	-1	1,75	5
1,41	0	2	6
0	0	1	6
0	-1,41	1	4,59
-1,41	0	0	6

As variáveis testadas anteriormente e que não apresentaram significância estatística no experimento anterior, como temperatura e concentração de metanol, foram fixadas e mantidas em todos os ensaios do DCC. A temperatura foi mantida a 30°C e a concentração final de metanol foi de 0,75%. A adição de Tween no meio de cultura não apresentou significância estatística no *screening* realizado e não foi realizada. A adição de aminoácidos teve um efeito significativo negativo para a produção da rhG-CSF e por isso não foi realizada.

A avaliação das variáveis testadas durante a indução da rhG-CSF ocorreu como descrito a seguir. A expressão da proteína recombinante foi realizada em meio de cultivo complexo tamponado BMGY (1% de Extrato de Levedura, 2% Peptona, 100mM de Tampão Fosfato pH 6,0, 1,34% de YNB, $4 \times 10^{-5}\%$ Biotina), suplementado com 1 % de glicerol como única fonte de carbono, por 48 horas a 30 °C em tubos tipo falcon Corning 50 mL *Mini Bioreactor* contendo 10 mL do meio em cada. Posteriormente foi realizada a indução em meio complexo tamponado BMMY (1% de Extrato de Levedura, 2% Peptona, 100mM de Tampão Fosfato pH 6,0, 1.34% de YNB, $4 \times 10^{-5}\%$ Biotina), contendo metanol como única fonte de carbono, com 7 mL do meio em cada tubo. As variáveis testadas só foram modificadas durante a fase de produção da proteína recombinante pelas células, que ocorre durante a indução da expressão em meio complexo contendo metanol BMMY. Portanto, para todos os ensaios, a fase de crescimento em BMGY se manteve a mesma, como descrito anteriormente. Todos os dias alíquotas de metanol absoluto foram adicionadas no meio BMMY, que variaram de acordo com o proposto em cada ensaio (Tabela 4). Foram realizadas coletas diárias (0, 24, 48, 72, 96, 120 e 144 horas) de 100 µL que, posteriormente, foram analisadas em SDS-PAGE 15 %.

Ao fim do processo, o meio foi centrifugado por 8 minutos, 3200 x *g* e a 20 °C para a separação da fração celular e o sobrenadante foi coletado e filtrado em membrana de 0,45 µm (Millipore).

Para quantificar as proteínas produzidas em cada um dos ensaios, utilizou-se o Kit ELISA G-CSF Humano (Novex *by Life Technologies*), segundo recomendação do fabricante (Figura 2). Esta metodologia foi aplicada, pois nos permite avaliar quantitativamente a produção da proteína recombinante em questão e se mostrou bastante sensível e eficiente na detecção da mesma.

Após a realização das quantificações da proteína recombinante produzida nas diferentes condições, as análises estatísticas foram realizadas pelo Professor Doutor Luis Alexandre Pedro de Freitas, da Faculdade de Ciências Farmacêuticas da USP, campus Ribeirão Preto. A metodologia de superfície de resposta foi utilizada para análise dos dados. Os coeficientes foram devidamente testados pelo teste t de Student a 5% de significância, realizado no software Minitab 17 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com) Minitab.

Abaixo está representado um esquema geral, mostrando os passos sequenciais de todo o delineamento experimental realizado:

Figura 4: Esquema geral mostrando todos os passos do delineamento experimental.

4.2.2. Caracterização da linhagem

Uma linhagem produtora de proteína recombinante deve ser estável geneticamente. As células recombinantes de *Pichia pastoris* produzindo rhG-CSF foram avaliadas por dez gerações (dez repiques consecutivos). Assim, espera-se uma maior

estabilidade genética para a linhagem gerada. Os repiques foram feitos a cada 24 horas, durante dez dias.

Após os dez repiques sucessivos, foi feita a indução da proteína rhG-CSF a partir de uma colônia da décima placa (denominada de caracterização) e também uma nova indução da colônia estoque (denominada de estoque). Esta colônia foi colocada em 5 mL do meio BMGY por 24 horas, 30 °C e 250 rpm em *shaker*. Após este período, o volume total foi transferido para 250 mL de BMGY, em Erlenmeyer de 1 litro por 24 horas adicionais nas mesmas condições. Mediu-se a DO a 600 nm encontrando-se resultados ao redor de 6. O meio foi centrifugado em Centrifuge 5810 R Eppendorf a 3000 x g, 5 minutos e 20 °C. O sobrenadante foi descartado e as células foram ressuspendidas em 50 mL do meio BMMY, em Erlenmeyer de 500 mL para a indução da expressão proteica, onde foram incubadas por 144 horas, 30 °C e 250 rpm em *shaker*. Metanol em uma concentração final de 0,75% foi adicionado no meio, para indução da proteína recombinante, a cada 24 horas, durante sete dias. Ao fim do processo, o meio foi centrifugado por 5 minutos, 3000 x g e a 20 °C para a separação da fração celular e o sobrenadante foi coletado e filtrado em membrana de 0,45 µm (Millipore).

Uma purificação por cromatografia de afinidade foi feita e conduzida em coluna de resina carregada com níquel HisTrapTM HP 1 mL (GE Healthcare), na qual as proteínas ficam aderidas pela cauda de histidina por afinidade química. O equipamento utilizado foi um FPLC AKTA (GE Healthcare). Para coletar as proteínas, a coluna foi lavada com solução tampão (Tris 100 mM, NaCl 100 mM, NaH₂PO₄ 17,4 mM) de concentrações crescentes de imidazol, o qual compete com as histidinas pelo níquel, liberando a proteína recombinante na forma purificada. As amostras foram analisadas

em SDS-PAGE 12 % para confirmar as frações que apresentaram a proteína purificada. As proteínas foram armazenadas em tubo falcon a 4 °C.

Todas as amostras foram concentradas em cinco vezes em *SpeddVac*.

Com as amostras purificadas foi realizado o teste de *Dot Blotting* para avaliar a linhagem produtora, ou seja, para confirmar se mesmo após os dez repiques a linhagem ainda estava produzindo a rhG-CSF. O anticorpo primário utilizado foi o G-CSF (K-15) sc-49679 Goat Polyclonal IgG (Santa Cruz Biotechnology) e o secundário foi o Mouse Anti-Goat IgG-AP sc-2355 AP Conjugate (Santa Cruz Biotechnology). Para o *Dot Blotting* pegou-se uma membrana de nitrocelulose 0,45 µm (BioRad) e marcou-se com um lápis a região que seria feito o blot. Pipetou-se 10 µL de Filgrastine rHu G-CSF 300 µg/mL (Blau) no canto superior esquerdo da membrana, para atuar como controle positivo. Pipetou-se também 10 µL do sobrenadante contendo a rhG-CSF proveniente da indução em escala maior referente à décima placa da caracterização no canto inferior direito da membrana. Após a aplicação das amostras, esperou-se até a membrana secar e foi feito um bloqueio dos sítios não específicos com 5 % BSA/TBS-T (0,05% Tween 20 em TBS 1x e 0,1% BSA) por aproximadamente 1 hora. Após o bloqueio, lavou-se a membrana 3 vezes com BSA/TBS-T. A membrana foi incubada com o anticorpo primário dissolvido em BSA-TBS-T por 30 minutos a temperatura ambiente, em uma diluição de 1:200. Lavou-se a membrana por três vezes de aproximadamente 5 minutos com TBS-T. Em sequência, a membrana foi incubada com o anticorpo secundário em uma diluição de 1:5000, por 30 minutos a temperatura ambiente. Por último, lavou-se a membrana três vezes com TBS-T, uma vez TBS e foi feita a revelação utilizando o AP Conjugate Substrate Kit (BioRad).

A caracterização da linhagem produtora da proteína recombinante, em relação ao crescimento, foi dada por meio de uma curva de crescimento celular em função do tempo. A taxa de crescimento celular é proporcional ao aumento de massa de células em função do tempo. Sendo assim, foram coletadas amostras em diferentes tempos de crescimento. Estas serão mensuradas via espectrofotometria. O cálculo do crescimento celular foi realizado a partir da equação a seguir:

$$\mu = \frac{1}{X} \cdot \frac{dX}{dt}$$

Onde: μ - velocidade específica (h^{-1})

X – concentração de células (cél/mL)

t – tempo (h)

Na fase exponencial de crescimento da célula pode-se considerar que:

$$\frac{1}{X} \cdot \frac{dX}{dt} = cte = \mu_{m\acute{a}x}$$

Integrando esta equação da concentração de células inicial X_0 a uma concentração X em um intervalo de tempo t, tem-se que,

$$\ln X = \ln X_0 + \mu_{m\acute{a}x} \cdot t$$

Nesta expressão, o valor de $\mu_{m\acute{a}x}$ é o coeficiente angular do gráfico de $\ln X$ em função do tempo. Uma vez determinado o valor de $\mu_{m\acute{a}x}$, é possível determinar o tempo de duplicação (t_d) na fase de crescimento exponencial através da equação abaixo:

$$t_d = \frac{\ln 2}{\mu_{m\acute{a}x}}$$

5. Resultados e discussão

5.1. Planejamento experimental para testar as diferentes condições físico-químicas para a indução da rhG-CSF

5.1.1. Diferentes condições físico-químicas para a indução da rhG-CSF

A temperatura ideal de crescimento para a levedura *Pichia pastoris* é de 30 °C, mas sabe-se que a variação de temperatura influencia na cinética de produção de proteínas. Essa influência pode ocorrer de duas maneiras: pode-se ocorrer um aumento da cinética de produção com o aumento da temperatura ou diminuição dessa taxa com o decréscimo de temperatura. Porém, o aumento da cinética de produção de proteínas pode saturar o retículo endoplasmático e provocar deficiência na etapa de secreção proteica. Neste caso, a diminuição de temperatura pode ser vantajosa, pois a produção proteica passará a ocorrer em uma taxa na qual o retículo endoplasmático não é estressado e a secreção proteica ocorre normalmente. Mudanças significativas no rendimento de proteínas heterólogas têm sido reportadas, em relação à temperatura de indução, possivelmente devido ao enovelamento correto em baixas temperaturas, facilitando a secreção (Jafari et al, 2011; Jahic et al, 2003). Estudos mostraram que a alteração da temperatura na indução da rG-CSF não teve nenhum efeito, enquanto que outros estudos já mostraram que a eficiência na produção da hG-CSF em *P. pastoris* aumentou quando a indução ocorreu sob temperaturas mais baixas (Maity et al, 2015). Portanto, foram testadas as temperaturas de 25, 30 e 35 °C.

A variação de pH influencia diretamente a atividade das enzimas relacionadas com o metabolismo celular de um microrganismo unicelular. Portanto, esta variação pode beneficiar ou prejudicar a produção de enzimas pela célula. A estabilidade de proteínas secretadas na cultura extracelular de *Pichia pastoris* é dependente da ação de

proteases e da sua habilidade em manter o seu enovelamento no ambiente. O pH exerce uma influência considerável nesses dois aspectos. Cada proteína possui suas propriedades particulares e, portanto, estabilidades diferentes (Maity et al, 2015). A levedura *P. pastoris* tolera uma ampla faixa de pH, de 3-9, porém a maioria das proteínas recombinantes são produzidas na faixa de pH 5-6. Foram testados os seguintes pHs durante a indução da proteína recombinante: 5, 6 e 7. Todos os valores de pH foram indicados pelo tampão fosfato utilizado durante a indução da proteína.

Genes que são expressos sob o controle do promotor AOX1 sofrem enorme influência dos níveis de concentração de metanol. Teoricamente, quando maior a concentração de metanol, maior será a produção de proteína recombinante, porém, estudos mostram que enquanto uma baixa concentração de metanol pode não ser suficiente para desencadear o processo de transcrição, uma alta concentração do mesmo pode levar a um acúmulo de formaldeído tóxico, causando lise celular e liberando proteases para o meio (Macauley-Patrick et al, 2005). Considerando que o metanol é a molécula indutora neste sistema de expressão, deve haver uma concentração ideal para que a indução seja máxima, sem que haja citotoxicidade. Desta forma, foram testadas concentrações de 0,5%, 0,75% e 1% de metanol no meio de cultura BMMY após 48 horas de crescimento e indução.

Dados mostram que a adição de Tween pode aumentar a liberação extracelular de GCSF em *P. pastoris* devido a um aumento na permeabilidade da membrana por conta de uma variedade de conteúdos hidroxilas (Apte-Deshpande et al, 2009). Um estudo mostrou que o uso de Tween não afetou da produção da proteína rhG-CSF, na linhagem SMD1168 de *P. pastoris* (Maity et al, 2015). Neste experimento foram testadas as concentrações de 80, 160 e 240 mg/L de Tween 80 no meio de cultura BMMY.

A maioria dos estudos na área de produção de proteínas recombinantes em *P. pastoris* utiliza um meio complexo para o crescimento e para indução da expressão proteica. Apesar de estes meios serem excelentes por fornecerem o ambiente químico correto requerido para o crescimento celular da *Pichia*, eles possuem limitações quando consideramos larga escala e fermentação industrial. Para se produzir uma grande quantidade de proteínas recombinantes e também para estudos estruturais e de caracterização, é necessário o uso de um meio definido, assim como um entendimento básico do desenvolvimento do projeto, com o intuito de poder manipular o ambiente físico-químico para se maximizar o crescimento e a produção de proteínas (Macauley-Patrick et al, 2005).

Para aumentar a produção de proteínas heterólogas, bem como reduzir o tempo de indução, uma abordagem comum é o uso de diferentes fontes de carbono além do metanol (Files et al, 2001). O sorbitol é uma fonte de carbono interessante interessante que pode ser adicionado juntamente com o metanol para a produção de proteínas recombinantes em *Pichia pastoris* (Arnau et al, 2010; Wang et al, 2010).

A partir disto, testamos a suplementação do meio BMMY com Sorbitol, nas concentrações de 0%, 0,5% e 1%, e também com aminoácidos em concentrações de 0, 10 e 20 mg/100mL.

5.1.2. Planejamento Plackett-Burman (*Screening Design*)

Para o início dos testes do *Screening Design* foi feito, primeiramente, a indução da expressão da rhG-CSF como descrito na metodologia. Após 144h de indução, as amostras foram centrifugadas e o sobrenadante foi utilizado. Para identificar e

quantificar as proteínas produzidas em cada um dos ensaios foi utilizado o Kit ELISA G-CSF Humano (Novex *by Life Technologies*).

A quantificação específica da rhG-CSF foi eficiente utilizando o Kit ELISA G-CSF Humano (Novex *by Life Technologies*). Amostras de G-CSF humana com concentrações conhecidas foram ensaiadas como padrão do ensaio, em duplicata. Este padrão abrange uma faixa de 0 a 2500 pg/mL de G-CSF humana. Desta forma, foi feita uma curva com os dados referentes às amostras padrão Hu G-CSF do ensaio (Figura 5), que serviu como parâmetro para a quantificação das amostras a serem analisadas contendo a hG-CSF recombinante. Esta curva corresponde à relação gráfica entre os valores de absorbância e os de concentração e, com base nesta análise gráfica, é possível verificar a linearidade da reação e calcular um fator de conversão de valores de absorbância em concentração.

Figura 5: Curva padrão Hu G-CSF pg/mL.

De acordo com o gráfico obtido, podemos dizer que os valores são confiáveis, já que o valor de R^2 foi igual a 0,9995, mostrando que os dados são adequados para a análise.

Os resultados da curva padrão e a equação da reta foram utilizados para quantificar as 16 amostras provenientes dos ensaios de expressão em que foram aplicados o Delineamento de Plackett-Burman. Essas quantificações foram realizadas em duplicata. Foi feito um gráfico de colunas para mostrar as diferenças dos valores de quantificação (pg/mL) entre os ensaios realizados, como podemos ver na Figura 6. A tabela 5 nos mostra os valores de cada variável referente a cada um dos ensaios mostrados no eixo X do gráfico da Figura 6.

Figura 6: Gráfico em colunas mostrando a diferença de concentração da rhG-CSF entre os 16 ensaios realizados.

Tabela 5: Descrição das variáveis referentes a cada um dos ensaios mostrados na Figura 6.

1	pH 7.0, 1% metanol, 1% sorbitol, 0 aminoácidos, 240 mg/L Tween 80
2	pH 7.0, 0,5% metanol, 0% sorbitol, 0 aminoácidos, 240 mg/L Tween 80
3	pH 5.0, 0,5% metanol, 0% sorbitol, 0 aminoácidos, 80 mg/L Tween
4	pH 5.0, 1% metanol, 1% sorbitol, 20 mg/100 mL aminoácidos, 80 mg/L Tween 80
5	pH 5.0, 0,5% metanol, 1% sorbitol, 20 mg/100 mL aminoácidos, 240 mg/L Tween 80
6	pH 7.0, 1% metanol, 0 sorbitol, 20 mg/100 mL aminoácidos, 80 mg/L Tween 80
7	pH 6.0, 0,75% metanol, 0,5% sorbitol, 10 mg/100 mL aminoácidos, 160 mg/L Tween 80
8	pH 6.0, 0,75% metanol, 0,5% sorbitol, 10 mg/100 mL aminoácidos, 160 mg/L Tween 80
9	pH 6.0, 0,75% metanol, 0,5% sorbitol, 10 mg/100 mL aminoácidos, 160 mg/L Tween 80
10	pH 6.0, 0,75% metanol, 0,5% sorbitol, 10 mg/100 mL aminoácidos, 160 mg/L Tween 80
11	pH 7.0, 1% metanol, 0 sorbitol, 20 mg/100 mL aminoácidos, 240 mg/L Tween 80
12	pH 5.0, 1% metanol, 0 sorbitol, 0 aminoácidos, 80 mg/L Tween 80
13	pH 5.0, 1% metanol, 1% sorbitol, 0 aminoácidos, 240 mg/L Tween 80
14	pH 7.0, 0,5% metanol, 1% sorbitol, 0 aminoácidos, 80 mg/L Tween 80
15	pH 5.0, 0,5% metanol, 0% sorbitol, 20 mg/100 mL aminoácidos, 240 mg/L Tween 80
16	pH 7.0, 0,5% metanol, 1% sorbitol, 20 mg/100 mL aminoácidos, 80 mg/L Tween 80

A Figura 6 mostra que houve bastante diferença na concentração da proteína recombinante obtida nos diferentes ensaios, sendo que nas condições do ensaio 2 foi obtida a maior produção de rhG-CSF, com 218,32 pg/mL, e o ensaio 5 possivelmente foi o que resultou em menor produção, já que seu valor foi negativo, mostrando que a quantidade de proteína produzida estava abaixo do limite de detecção do kit.

Os valores exatos da concentração de rhG-CSF obtidos em cada um dos ensaios passaram por uma análise estatística para avaliar quais destes resultados tiveram uma diferença estatisticamente significativa e indicar quais das variáveis testadas foram realmente importantes. As análises estatísticas foram feitas em colaboração com o Prof. Dr. Luis Alexandre Pedro de Freitas, da Faculdade de Ciências Farmacêuticas de Ribeirão Preto da Universidade de São Paulo. Os programas utilizados na análise estatística foram o Minitab 17 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com) Minitab.

Com os dados obtidos após a análise estatística (Figura 7) foi feito um gráfico de Pareto dos efeitos padronizados (Figura 8), que é um tipo especial de gráfico de barras onde os valores representados graficamente são organizados do maior para o menor. Neste gráfico está representada a relação das variáveis estudadas (temperatura, pH, metanol e adição de Sorbitol, aminoácidos e Tween 80) com a produção da proteína recombinante rhG-CSF em pg/mL. Também mostra quais variáveis foram significativas para um valor de α de 0,1. Neste caso, como mostra a linha vertical presente no gráfico, foram as variáveis pH, sorbitol e aminoácidos.

Term	Effect	Coef	SE	Coef	T	P	
Constant		67,27	11,90	5,65	0,000		
A	15,04	7,52	11,90	0,63	0,545		
B	47,70	23,85	11,90	2,00	0,080	*	
C	-43,83	-21,92	11,90	-1,84	0,103		
D	-44,26	-22,13	11,90	-1,86	0,100	*	
E	-67,90	-33,95	11,90	-2,85	0,021	*	
F	-25,78	-12,89	11,90	-1,08	0,310		
Ct Pt		-38,46	23,80	-1,62	0,145		

Figura 7: Efeitos estimados e coeficientes para proteína em pg/mL. Minitab.

Figura 8: Gráfico de Pareto mostrando a relação entre a produção de proteína rhG-CSF e as variáveis testadas. A: Temperatura; B: pH; C: % metanol; D: Sorbitol; E: Aminoácidos; F: Tween 80.

A figura 7 nos mostra os efeitos estimados e coeficientes da produção de proteína em pg/mL. Os valores de P mostrados na tabela nos permitem identificar quais das variáveis tiveram um efeito significativo no *Screening Design*, ou seja, aqueles valores menores ou iguais a 0,1 (referente ao valor de α 0,1) indicam as variáveis com efeito significativo. Os asteriscos mostram as variáveis estatisticamente significantes. Podemos ver também os valores dos efeitos, que indicam se as variáveis tiveram um efeito positivo ou negativo no experimento. Os resultados para os efeitos de parâmetros individuais comprovaram que, para as variáveis significantes, que foram pH, aminoácidos e sorbitol, apenas o aumento do pH teve um efeito positivo, enquanto a adição de aminoácidos e sorbitol no meio de cultivo apresentou um efeito negativo, sendo assim necessário diminuir o nível para aumentar a produção enzimática.

Para avaliar a confiabilidade dos dados também foi feito um gráfico da distribuição normal de resíduos, que é um conjunto de técnicas utilizadas para investigar a adequabilidade de um modelo de regressão com base nos resíduos (Gráfico 4). O resíduo é dado pela diferença entre a variável resposta observada e a variável resposta estimada. O gráfico de probabilidade normal dos resíduos deve seguir uma linha reta. Este determina se os dados obedecem a uma distribuição hipotética, baseado no exame visual dos dados. Se os pontos caem em torno de uma reta a distribuição é adequada. Como podemos ver, temos uma boa distribuição dos pontos sobre a reta, o que indica a normalidade dos dados (Figura 9). Com isto, podemos dizer que os dados obtidos são considerados adequados para a análise.

Figura 9: Gráfico da distribuição normal dos resíduos.

Para ajudar a entender e a otimizar melhor a resposta, foi feita uma análise de superfície de resposta, que é um conjunto de técnicas avançadas de planejamento de experimentos. Normalmente, esta metodologia é utilizada para refinar modelos após a determinação dos fatores importantes por meio de experimentos fatoriais, especialmente

quando há suspeita de curvatura na superfície de resposta. Uma equação de superfície de resposta permite a adição dos termos quadráticos que permitem modelar curvatura na resposta, tornando-as úteis para entender ou mapear uma região de uma superfície de resposta, modelar a forma como certas mudanças nas variáveis afetam uma resposta de interesse, determinar os níveis das variáveis que otimizam uma resposta e selecionar as condições operacionais que atendem às especificações.

Os gráficos gerados a partir da superfície de resposta estão apresentados a seguir (Figura 10). As figuras nos mostram a variável resposta, que no caso é a produção da proteína rhG-SCF em pg/mL, sendo graficamente representada em função dos fatores identificados como importantes, ou significativos, no processo (pH, adição de sorbitol e aminoácidos no meio de cultura). A partir da função que define uma superfície de resposta é possível encontrar qual condição de operação que leva ao ponto ótimo da variável resposta. Neste caso, qual a melhor condição para uma maior produção da proteína recombinante rhG-CSF.

Figura 10: Análise de superfície de resposta para as variáveis pH e aminoácidos.

Como podemos ver na figura 10, que mostra a relação da adição de aminoácidos no meio de cultura e a variação do pH na produção da proteína recombinante, o ponto ótimo (indicado na figura pelo vermelho intenso) é apresentado onde o valor de pH é maior e adição de aminoácidos é nula. Desta forma, podemos dizer que para a produção de rhG-CSF o aumento do pH até o valor de 7,4 foi positivo, enquanto que a adição de aminoácidos no meio de cultivo teve um efeito negativo.

A relação da adição de sorbitol no meio de cultura e a variação de pH na produção da proteína recombinante se comportaram da mesma forma. Assim como no gráfico anterior, o ponto ótimo é indicado onde o valor de pH é maior e adição de sorbitol é mínima, ou igual a zero. Desta forma, a adição do sorbitol no meio de cultura durante a indução da rhG-CSF também apresentou um efeito negativo para a produção da mesma. Desta forma, para este experimento, a única variável que teve um efeito positivo no aumento da produção da rhG-CSF foi o pH, enquanto que as outras variáveis (adição de sorbitol e aminoácidos) apresentaram um efeito negativo, porém significativos.

Apesar de outros autores, como Arnau et al (2010), Wang et al (2010) e Maity et al (2015), observarem um efeito positivo da adição de sorbitol no meio de cultivo como outra fonte de carbono para a produção de proteínas recombinantes em *Pichia pastoris*, o mesmo não foi observado neste experimento. A fonte de carbono alternativa não apresentou um efeito positivo no processo, possivelmente devido ao fato de que algum outro efeito secundário dessa adição de sorbitol pode ter causado algum estresse para as células produtoras da proteína recombinante. O mesmo aconteceu com a adição de aminoácidos. Outros autores, como Maity et al (2015) haviam descrito que esta adição teria sido benéfica para produção da G-CSF em *P. pastoris*, aumentando a sua produção em até duas vezes. Porém, para esta proteína específica estes dois fatores tiveram um

influência negativa na sua produção. Muito provavelmente estas diferenças ocorreram devido ao fato de as linhagens celulares e vetores utilizados serem diferentes, no trabalho de Maity et al (2015) as linhagens utilizadas são GS115 e SMD1168 his4pep4 (Mut+) e os vetores pGAPZαA e pPICZαB. Em relação ao pH, apesar de a maioria das proteínas recombinantes de *P. pastoris* serem produzidas na faixa de pH de 5-6, um valor de pH mais alto, próximo de 7-8, foi mais eficiente para o aumento na produção da rhG-CSF. Esta variação de pH influencia diretamente a atividade das enzimas relacionadas com o metabolismo celular de um microrganismo unicelular.

O aumento do valor de pH beneficiou a produção da proteína pela célula. Possivelmente isto ocorreu, pois estes valores mais altos de pH conferiram uma maior estabilidade às proteínas recombinantes secretadas na cultura extracelular, diminuindo a ação de proteases e permitindo que a proteína mantivesse sua integridade no ambiente. Este resultado também vai de encontro ao trabalho de Maity et al (2015), que observou um aumento da produção de G-CSF nos pHs de 4 e 5 devido ao fato de que os íons citrato utilizados para a preparação do tampão em pHs mais baixos serviu como uma fonte adicional de carbono para o acúmulo de biomassa e assim da formação do produto. Portanto, apesar de ser a mesma proteína produzida, ela se comportou de maneiras diferentes em relação às variáveis, confirmando o fato de que cada proteína possui suas propriedades particulares e estabilidades diferentes que são estabelecidas de acordo com o processo de indução envolvido.

É importante lembrar que, apesar de as outras variáveis não terem sido estatisticamente significantes, qualquer variação mínima das mesmas pode interferir positiva ou negativamente quando se trata de cultivo celular. Mesmo quando não há uma diferença estatística, a variação de um fator pode ser muito significativa para o bioprocesso. Quando se trata de cultivo de células, as mudanças físico-químicas ou

outras condições podem mudar completamente a produção ou o rendimento de uma proteína recombinante.

Tendo feito a análise inicial dos efeitos principais pelo Plackett-Burman, outro modelo foi escolhido e aplicado para uma análise mais profunda das variáveis significantes apontadas. Devido ao fato de duas das variáveis significativas terem um efeito negativo na produção da rhG-CSF (adição de aminoácidos e sorbitol), apenas uma delas foi escolhida para entrar no modelo como um confirmatório do seu efeito negativo. Desta forma, foram analisados o pH e a adição de sorbitol no meio de cultura.

5.1.3. Delineamento Composto Central (DCC)

Para o início do Delineamento Composto Central (DCC), uma nova indução da rhG-CSF foi realizada, a partir da mesma colônia estoque utilizada nos ensaios do Plackett-Burman, como descrito na metodologia. Após 144 horas de indução as amostras foram centrifugadas e o sobrenadante foi utilizado. Para identificar e quantificar as proteínas produzidas em cada um dos ensaios foi utilizado o Kit ELISA G-CSF Humano (Novex *by Life Technologies*).

A quantificação específica da rhG-CSF pelo kit de ELISA utilizado foi novamente eficiente. Para este ensaio, foi feito uma curva padrão com amostras de G-CSF humana com concentrações conhecidas, em duplicata. Este padrão cobre uma faixa de que vai de 0 a 2500 pg/mL de G-CSF humana. Desta forma, foi feita uma curva com os dados referentes às amostras padrão Hu G-CSF do ensaio (Figura 11), que serviu como base para a quantificação das amostras a serem analisadas contendo a hG-CSF recombinante. Esta curva corresponde à relação gráfica entre os valores de absorbância

e os de concentração e, com base nesta análise gráfica, é possível verificar a linearidade da reação e calcular um fator de conversão de valores de absorbância em concentração.

Figura 11: Curva padrão Hu G-CSF pg/mL para o DCC.

De acordo com o gráfico obtido, podemos dizer que os valores são confiáveis, já que o valor de R^2 foi igual a 0,9996, mostrando que os dados são adequados para a análise.

Com os dados da curva padrão e a equação da reta, as 13 amostras foram quantificadas, todas em duplicata. Um gráfico de colunas foi feito para mostrar as diferenças dos valores de quantificação (pg/mL) entre os ensaios realizados, como podemos ver na Figura 12. A tabela 6 nos mostra os valores de cada variável referente a cada um dos ensaios mostrados no eixo X do gráfico da Figura 12.

Figura 12: Gráfico de colunas mostrando a diferença de concentração da rhG-CSF entre os 13 ensaios realizados.

Tabela 6: Valores das variáveis referentes a cada um dos ensaios.

1	pH 6.0; 1% Sorbitol
2	pH 7.0; 0,25% Sorbitol
3	pH 6.0; 1% Sorbitol
4	pH 6.0; 1% Sorbitol
5	pH 6.0; 1% Sorbitol
6	pH 7.0; 1,75% Sorbitol
7	pH 5.0; 0,25% Sorbitol
8	pH 7.41; 1% Sorbitol
9	pH 5.0; 1,75% Sorbitol
10	pH 6.0; 2% Sorbitol
11	pH 6.0; 1% Sorbitol
12	pH 4.59; 1% Sorbitol
13	pH 6.0; 0% Sorbitol

Neste experimento também houve bastante diferença na concentração da proteína recombinante entre os ensaios, sendo o ensaio 13 o que mais produziu a rhG-CS, com 119,3 pg/mL, e o ensaio 11 o que menos produziu, com valor nulo, mostrando que possivelmente a quantidade de proteína produzida estava abaixo do limite de detecção do kit utilizado.

Da mesma forma que os ensaios anteriores, estes dados obtidos (Figura 13) passaram por uma análise estatística para avaliar quais destes resultados apresentaram diferenças significativas estatisticamente. As análises estatísticas foram feitas em colaboração com o Prof. Dr. Luis Alexandre Pedro de Freitas, da Faculdade de Ciências Farmacêuticas de Ribeirão Preto da Universidade de São Paulo. Os programas utilizados na análise estatística foram o Minitab 17 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com) Minitab.

Term	Coef	SE Coef	T	P
Constant	34,501	12,574	2,744	0,029
A	-24,153	9,941	-2,430	0,045
B	2,392	9,941	0,241	0,817
A*A	16,606	10,660	1,558	0,163
B*B	-8,787	10,660	-0,824	0,437
A*B	8,464	14,058	0,602	0,566

Figura 13: Efeitos estimados e coeficientes para proteína em pg/mL. Minitab.

Para ajudar a entender e a otimizar melhor as respostas, foi feita uma análise de superfície de resposta. O gráfico gerado a partir desta metodologia está representado a seguir (Figura 14). O gráfico nos mostra a variável resposta, a produção de rhG-CSF em pg/mL, sendo representada em função dos fatores analisados pelo DCC (pH e Sorbitol). A partir da função que define uma superfície de resposta é possível encontrar qual condição de operação leva ao ponto ótimo da variável resposta. Neste caso, para uma maior produção da proteína recombinante em estudo, o aumento do valor de pH se mostrou bastante importante, pelo menos até o valor estudado que foi de 7,41. Confirmando o resultado do Plackett-Burman, a adição de Sorbitol no meio de cultura apresenta um efeito negativo na produção da rhG-CSF, ou seja, quanto mais Sorbitol era adicionado, menor era a produção da proteína recombinante.

Figura 14: Análise de superfície de resposta para as variáveis pH e Sorbitol. Resultado do DCC.

Como podemos ver no gráfico, o ponto ótimo (indicado pela cor vermelha intensa) é localizado onde o valor de pH é maior e a adição de Sorbitol é menor.

Os resultados do Delineamento Composto Central foram eficientes para analisar em níveis maiores os efeitos do pH do meio e da adição de Sorbitol no meio de cultura, mostrando que um aumento no valor de pH ainda é positivo para a produção da rhG-CSF nestas condições e que a adição de Sorbitol realmente apresenta um efeito negativo para a produção da mesma. Estes dados confirmam o resultado da análise de Plackett-Burman.

É importante lembrar que, apesar de os experimentos ocorrerem em condições controladas, não é possível ter total controle sobre todas as variáveis envolvidas no processo. Muitas variáveis com seus níveis requerem muitas unidades experimentais, devido à complexidade do processo que, no caso, é fisiológica e implica que os resultados experimentais sejam objeto de variabilidade inerente. Isto ocorre, pois quando se trabalha com cultivo de células para produção de proteínas é impossível o

controle das variáveis inerentes ao processo. A maior dificuldade de experimentos deste tipo está em controlar as influências de vários fatores externos ao experimento, como, por exemplo, variações nos lotes dos reagentes utilizados no processo ou até mesmo mutações inseridas no DNA das células, entre outros.

Apesar das dificuldades que se impõem neste tipo de análise, torna-se cada vez mais importante o uso de metodologias que visem otimizar a produção do alvo de interesse.

5.2. Caracterização de linhagem

Após a indução da expressão da rhG-CSF referente à primeira placa (primeiro repique) e da placa após os dez repiques, foi feita uma purificação das amostras, conduzida em coluna de resina carregada com níquel HisTrapTM HP 1 mL (GE Healthcare), na qual as proteínas ficam aderidas pela cauda de histidina por afinidade química. O equipamento utilizado na purificação foi um FPLC AKTA (GE Healthcare). O gráfico abaixo mostra o perfil da cromatografia de afinidade para a purificação da proteína rhG-CSF proveniente das células que haviam sofrido dez repiques. O pico foi dado em uma concentração de imizadol de aproximadamente 20 %, como mostra a figura abaixo.

Figura 15: Perfil da cromatografia por afinidade ao Níquel para a purificação da rhG-CSF. Eluente B se refere ao Tampão de Lise com Imidazol (Tris 100 mM, NaCl 100 mM, NaH_2PO_4 17.4 mM, Imidazol 250 mM).

A detecção da proteína recombinante nas amostras purificadas do sobrenadante de *P. pastoris* após os dez repiques (décima placa) foi feita por *Dot Blotting*. As amostras utilizadas foram da expressão após 144 horas. O anticorpo primário G-CSF (K-15) sc-49679 Goat Polyclonal IgG (Santa Cruz Biotechnology) foi eficiente contra a proteína recombinante em uma diluição de 1:200 (Figura 16).

Figura 16: Membrana de nitrocelulose mostrando um Dot Blotting da proteína incubada com o anticorpo primário na concentração de 1:200 mostrando em **1**: amostra Filgrastima comercial e **2**: amostra da rhG-CSF após os dez repiques para confirmar linhagem produtora.

O círculo marcado referente à rhG-CSF apareceu bem mais claro do que o círculo referente à Filgrastima comercial utilizada, muito provavelmente, devido à baixa concentração da amostra utilizada. Porém, a amostra foi marcada, comprovando a presença da proteína recombinante no sobrenadante da indução. Outro ponto que comprova a presença da rhG-CSF são os dados da purificação, ou seja, a presença do pico referente à saída das proteínas na cromatografia, considerando que, em teoria, apenas a proteína recombinante possui uma cauda de histidina capaz de se ligar à coluna utilizada. Desta forma, podemos considerar a colônia selecionada como uma linhagem produtora, que mesmo após vários repiques as células não perderam a capacidade de produzir a proteína recombinante, conferindo uma maior estabilidade genética para a linhagem gerada.

A partir dos valores mensurados via espectrofotometria das amostras de cultura coletadas durante a fase de crescimento da levedura, foram feitos gráficos mostrando a curva de crescimento (Figura 17). Isto foi feito tanto para a colônia estoque quanto para a colônia testada após os dez repiques. A partir do crescimento da levedura em meio

BMGY foi possível verificar no período de 24 horas o crescimento celular. Os gráficos estão apresentados a seguir.

Figura 17: Curva de crescimento celular da colônia estoque e da colônia após os dez repiques (caracterização). O sinal + representa a caracterização e o sinal x representa o estoque.

Pode-se observar que as colônias testadas tiveram um crescimento bastante parecido, como mostram as curvas de crescimento no gráfico. Com os pontos obtidos, pode-se observar a fase de crescimento inicial (fase lag), a fase log, ou fase de crescimento exponencial, e possivelmente o início da fase estacionária, sinalizada pelo último ponto obtido. A fase estacionária não pode ser bem definida pela falta de pontos coletados.

A partir dos dados obtidos do crescimento celular, foi feita uma regressão linear, com o objetivo de verificar a existência de uma relação funcional entre a variável dependente (absorbância) com uma variável independente (tempo), ou seja, consiste na obtenção de uma equação que tenta explicar a variação da absorbância pela variação dos níveis de tempo.

A regressão linear foi feita para transformar a fase log de crescimento em uma reta e, assim, obter-se a equação do gráfico (Figura 18). O coeficiente angular da reta gerada pela regressão linear reflete a velocidade de crescimento das células. Desta forma, a partir desta equação foi possível encontrar a taxa de crescimento específico. Considerando que os dados tomados para a confecção da regressão linear são da fase log de crescimento, o valor de velocidade encontrado reflete a velocidade máxima de crescimento ($\mu_{m\acute{a}x}$).

Figura 18: Regressão linear da amostra estoque e da amostra caracterização, mostrando as equações referentes a cada uma das retas.

Desta forma, de acordo com o gráfico acima e tendo a equação da reta, a velocidade máxima de crescimento do Estoque é 0,2133 ($\mu_{m\acute{a}x} = 0,2133$) e da Caracterização é 0,2582 ($\mu_{m\acute{a}x} = 0,2582$) para estes pontos coletados na fase log de crescimento. Pode-se observar que as velocidades de crescimento das amostras são bastante próximas, o que indica que não houve uma mudança muito grande na linhagem

mesmo após os dez repiques, em relação à curva de crescimento celular e também em relação à velocidade máxima de crescimento.

Uma vez determinado o valor de $\mu_{m\acute{a}x}$, foi possível determinar o tempo de duplicação (t_d) na fase de crescimento exponencial através da equação abaixo:

$$t_d = \frac{\ln 2}{\mu_{m\acute{a}x}}$$

Para o estoque o tempo de duplicação foi de 2,68 h e para a caracterização foi de 3,24. Estes dados mostram que não houve também diferença grande entre a colônia estoque e colônia após os dez repiques sucessivos.

	$\mu_{m\acute{a}x}$	Tempo de duplicação (t_d)	Rendimento (mg/L)
Estoque	0,2133	2,684535943	18
Caracterização	0,2582	3,249635164	83

Tabela 7: Tabela comparativa entre a amostra Estoque e a amostra Caracterização, mostrando a velocidade de crescimento específica, o tempo de duplicação e o rendimento final de cada.

A tabela acima nos mostra os resultados da velocidade de crescimento máxima, o tempo de duplicação e também o rendimento, em relação à produção da proteína rhG-CSF, das colônias testadas. O rendimento da caracterização foi maior do que o do estoque, possivelmente devido a alguma variável durante o processo de indução e expressão da proteína.

O maior rendimento obtido neste trabalho foi de 83 mg/L de rhG-CSF em Erlenmeyer agitado. Apte-Deshpande et al (2015) obtiveram cerca de 2 mg/L de G-CSF em condições normais e cerca de dez vezes mais, 200 mg/L, com a adição de surfactantes no meio de cultivo, também em escala de Erlenmeyer. Neste caso, a adição

de surfactantes aumentou muito consideravelmente a produção de proteína recombinante. Bahrami et al (2009) conseguiu um rendimento de aproximadamente 18 mg/L de G-CSF em Erlenmeyer com a adição de Tween 80 no meio de cultivo. Nestes dois casos, a adição de algum tipo de detergente foi bastante benéfica para a produção de G-CSF, enquanto que para proteína recombinante produzida neste trabalho eles não apresentaram efeito significativo. Em 2015, Maity et al, conseguiram um rendimento de 4 mg/L em condições normais, enquanto que após algumas mudanças em condições nutricionais e operacionais a produção aumentou de duas a cinco vezes, tendo um rendimento máximo de 0,17 mg/L/h. Todos os trabalhos citados acima produziram a rhG-CSF em sistema de expressão de *Pichia pastoris*. As diferenças observadas em relação às condições de cultivo e ao rendimento final da proteína podem ser explicadas pelo fato de que, apesar de ser a mesma proteína e o mesmo sistema de expressão, as linhagens e os vetores de expressão são bastante diferentes. Isto influencia diretamente na maneira que a célula se comporta e na sua produtividade. Desta forma, podemos dizer que neste trabalho obtivemos um bom rendimento final da rhG-CSF para um escala pequena, em Erlenmeyer, quando comparado com outros trabalhos na literatura.

6. Conclusão

O protocolo de expressão e indução em *Pichia pastoris* é muito bem estabelecido. Apesar disto, verificamos que o aumento de pH proporcionou um aumento pontual na produtividade da rhG-CSF, de acordo com as análises estatísticas.

De acordo com o os resultados apresentados, conclui-se que foi possível estabelecer um protocolo otimizado de expressão de uma linhagem celular geneticamente estável.

CONSIDERAÇÕES FINAIS

Neste trabalho conseguimos um rendimento final da rhG-CSF de aproximadamente 83 mg/L. Este valor de rendimento é consideravelmente bom quando comparado com outros resultados obtidos da literatura. Desta forma, torna-se interessante fazer um escalonamento desta produção para se tentar obter parâmetros necessários para uma maior produção da proteína recombinante.

Outra análise importante a ser feita é análise de glicosilação das proteínas, já que para um possível biofármaco, é extremamente importante se conhecer a cadeia de carboidratos presente em sua estrutura.

Da mesma maneira, a realização de testes *in vivo* e *in vitro* se mostra bastante interessante no intuito de observar como se dá a resposta de alguns tipos de células e organismos em relação à proteína recombinante produzida.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABBAS, A.; LICHTMANN, A.; PILLAI, S. **Imunologia Celular e Molecular**. Editora Elsevier, 6ª edição, 2008.
- ABRAMSON, J. S.; WHEELER, J. G. **The neutrophil**. Oxford: IRL Press at Oxford University Press; 1993.
- ALBERTS, B.; JOHNSON, A.; LEWIS, J.; et al. **Biologia Molecular da célula**, 5. Ed. Porto Alegre: Artmed, 2011.
- APTE-DESHPANDE, A.; SOMANI, S.; MANDAL, G.; SOORAPANENI, S.; PADMANABHAN, S. Over expression and analysis of O-glycosylated recombinant human granulocyte colony stimulating factor in *Pichia pastoris* using Agilent 2100 Bioanalyzer. **Journal of Biotechnology**, V. 143, P. 44–50, 2009.
- BABAEIPOUR, V.; KHANCHEZAR, S.; MOFID, M. R.; ABBAS, M. P. H. Efficient process development of recombinant human granulocyte colony-stimulating factor (rh-GCSF) production in *Escherichia coli*. **Iranian Biomedical Journal**, v. 2, p. 102-110, 2015.
- BABALOLA, C.; NIGHTINGALE, C.; NICOLAU, D. Adjunctive efficacy of granulocyte colony stimulating factor on treatment of *Pseudomonas aeruginosa* pneumonia in neutropenic and non-neutropenic hosts. **Journal of antimicrobial Chemotherapy**, v. 53, n. 6, p. 1098-100, 2004.
- BAHRAMI, A.; SHOJAOSADATI, S. A.; KHALILZADEH, R.; MASOUMIAN, M. R. Prevention of human granulocyte colony-stimulating factor protein aggregation in recombinant *Pichia pastoris* fed-batch fermentation using additives. **Biotechnology and Applied Biochemistry**, v. 52, p. 141-149, 2008.
- BARREDA, D. R.; HANINGTON, P. C.; BELOSEVIC, M. Regulation of myeloid development and function by colony stimulating factors. **Developmental & Comparative Immunology**, v. 28, n. 5, p. 509-554, 2004.
- BARROS NETO, B.; SCARMINIO, I. S.; BRUNS, R. E. **Como fazer experimentos: pesquisa e desenvolvimentos na ciência e na indústria**. 3ª ed. Campinas, SP: Editora UNICAMP, 2007. 480 p.
- BAYNES, B. M.; TROUT, B. L. Rational design of solution additives for the prevention of protein aggregation. **Biophysical Journal**, v. 87, p. 1631-1639, 2004.
- BELL, G. H.; LEDOLTER, J.; SWERSEY, A. A Plackett-Burman Experiment to Increase Supermarket Sales of National Magazine. **Interfaces**. V. 39, n. 2, p. 145-158, 2009.
- BOX, G. E. P.; HUNTER, J. S.; HUNTER, W. G. **Statistics for experimenters: design, innovation and discovery**. 2nd ed. New Jersey: John Wiley & Sons, Inc., 2005.

- BOX, G. E. P.; WILSON, K. B. On the experimental attainment of optimum conditions. **Journal of the Royal Statistical Society**, v. 13, p. 1-45, 1951.
- BRASIL. Agência Nacional de Vigilância Sanitária. Resolução RE nº 1316, de 23 de julho de 2002. Diário Oficial da União, 2002; jul 25.
- BRASIL. Ministério da Saúde. Portaria SAS nº 125 de 19 de abril de 2001. Altera a Tabela Descritiva do SAI/SUS estabelecida pela Portaria GM/MS nº 1230, de 14 de outubro de 1999 no que diz respeito ao Grupo 36. Diário Oficial da União, 2001; 20 abr.
- CALLARD, R. E. **The cytokine FactsBook**. San Diego: Academic Press, p. 265, 1994.
- CHOTINANTAKUL, K; LEEANANSAKSIRI, W. Hematopoietic stem cell development, niches, and signaling pathways. **Bone Marrow Research**, v. 2012, 16 pages, 2012.
- CLARK, S. C.; KAMEN, R. The human hematopoietic colony-stimulating factors. **Science**, 236: 1229-1237, 1987.
- CODEVILLA, C. F. et al. Biological potency evaluation and characterization of rhG-CSF in pharmaceutical products. **Rev. Bras. Hematol. Hemoter**, v. 26, n.2, p. 104-108, 2004.
- DALY, R.; HEARN, M. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. **Journal of Molecular Recognition**, v. 18, n. 2, p. 119-138, 2005.
- DATA SUS [internet]. [cited 2013 oct 1]; Available from: www.datasus.gov.br
- DELGADO, I. G.; ARCURI, R. B. Um pacto pela saúde e pelo desenvolvimento. **Valor econômico**, 26 fev. 2013. Disponível em: <<http://clippingmp.planejamento.gov.br/cadastros/noticias/2013/2/26/um-pacto-pela-saude-e-pelo-desenvolvimento>>. Acesso em: 19 agos. 2013.
- DEMETRI, G. D.; GRIFFI, J. D. **Granulocyte colony-stimulating factor and its receptor**. Blood, 1991.
- ENCARNAÇÃO, A.; ALMEIDA, C.; BRUM, P.; KARELE, S. Projeto de uma planta para a produção de G-CSF recombinante. Rio de Janeiro (Brasil): **Universidade Federal do Rio de Janeiro**, 2005.
- FARDELONE, L. C.; BRANCHI, B. A. **O setor de biofármacos e as oportunidades para o Brasil**. Revista da FAE, p. 29-38, 2006.
- GARCÍA-CARBONERO, R.; MAYORDOMO, J. I.; TORNAMIRA, M. V.; et al. Granulocyte colony-stimulating factor in the treatment of high-risk febrile neutropenia: a multicenter randomized trial. **Journal of the National Cancer Institute**, v. 93, n. 1, p. 31-38, 2001.
- GOMES, F. R.; MALUENDA, A. C.; TÁPIAS, J. O.; OLIVEIRA, F. L.; et al. Expression of recombinant human mutant granulocyte colony stimulating factors (Nartograstim) in *Escherichia coli*. **World Journal of Microbiology and Biotechnology**, v. 28, n. 7, p. 2593-2600, 2012.

- HAMMERLING, U.; KROON, R.; SJODIN, L. In vitro bioassay with bioassay with enhanced sensitivity for human granulocyte colony-stimulating factor. **Journal of pharmaceutical and biomedical analysis**, v. 13, n.1, p. 9-20, 1995.
- HANADA, J.; ONO, I. Disappearance of neutrophil ascilation in a child with cyclic neutropenia after treatment with recombinant human granulocyte colony-stimulating factor. **European Journal of Hematology**, v. 45, p. 181-182, 1990.
- HARTUNG, T.; AULOCK, S.; VON, W.A. Role of granulocyte colony-stimulating factor in infection and inflammation. **Medical Microbiology and Immunology**, v. 187, n. 2, p. 61-69, 1998.
- HENRY, C. J.; BUSS, M. S.; LOTHROP, C. D. Veterinary uses of recombinant human granulocyte colony-stimulating factor. **Compendium on continuing education for the practicing veterinarian**; v. 20, p. 728-738, 1998.
- HEYLAND, J.; BLANK, L. M.; SCHMID, A. Quantification of metabolic limitations during recombinant protein production in *Escherichia coli*. **J Biotechnol**, v. 155, n. 2, p. 84-178, 2011.
- HILL, C. P.; OSSLUND, T. D.; EISENBERG, D. The structure of granulocyte-colony-stimulating factor and its relationship to other growth factors. **Biochemistry**, v. 90, p. 5167-5171, 1993.
- HUBEL, K.; ENGERT, A. Clinical applications of granulocyte colony-stimulating factor: an update and summary. **AnHematol**. V. 87, p. 207-213, 2003.
- IWASAKI, H.; SHIMODA, K.; OKAMURA, S.; OTSUKA, T.; NAGAFUJI, K.; HARADA, N. et al. Production of soluble granulocyte colony-stimulating factor receptors from myelomonocytic cells. **The Journal of Immunology**, v. 163, n. 12, p. 6907, 1999.
- JAFARI, R.; SUNDSTROM, B. E.; HOLM, P. Optimization of production of the anti keratin 8 single chain Fv TS-218 in *Pichia pastoris* using design of experiments. **Microbial Cell Factories**, 10, 34, 2011.
- JAHC, M.; WALLBERG, F.; BOLLOCK, M.; GARCIA, P.; & Enfors, S. O. Temperature limited fed-batch technique for control of proteolysis in *Pichia pastoris* bioreactor cultures. **Microbial Cell Factories**, 2, 6, 2003.
- JEVSEVAR, S.; GABERC-POREKAR, V.; FONDA, I.; PODOBNIK, B.; GRDADOLNIK, J.; MENART, V. Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. **Biotechnology progress**, v. 17, p. 687-692, 2005.
- KANDA, N.; FUKUSHIGE, S.; MUROTSU, T.; et al. Human gene coding for granulocyte-colony stimulating factor is assigned to the q21-q22 region of chromosome 17. **Somat Cell Mol Genet**, v. 13, n. 6, p. 679-684, 1987.
- KONDO, M.; et al. Biology of hematopoietic stem cells and progenitors: implications for clinical application. **Annual Review of Immunology**, v. 21, p. 759-806, 2003.

- KURTZMAN, C. P. Impact of nucleic acid comparison on the classification of fungi. **Proc. Indian Acad.** v. 97, p. 185-201, 1987.
- LASNIK, M. A.; POREKAR, V.G.; STALE, A. Human granulocyte colony stimulating factor (hG-CSF) expressed by methylotrophic yeast *Pichia pastoris*. Pflügers Archives of European Journal of Physiology, v. 442, p r184-r186, 2001.
- LU, H. S.; BOONE, T. C.; SOUZA, L. M.; LAI, P. H. Disulfide and secondary structure of recombinant human granulocyte colony-stimulating factor. **Archives of Biochemistry and Biophysics**, v. 268, n.1, p. 81-92, 1989.
- MACAULEY-PATRICK, S.; FAZENDA, M. L.; MCNEIL, B.; HARVEY, L.M. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*, V. 22, P. 249–270, 2005.
- MAITY, N.; THAWANI, A.; SHARMA, A.; GAUTAM, A.; MISHRA, S.; SAHAI, V. Expression and Control of Codon-Optimized Granulocyte Colony-Stimulating Factor in *Pichia pastoris*. **Applied Biochemistry and Biotechnology**, v. 178, n. 1, p 159-172, 2015.
- MATEUS, N.B. O delineamento compost central e sua viabilidade de uso em algumas áreas de pesquisa. Dissertação (mestrado) – Escola Superior de Agricultura “Luiz Queiroz”, Universidade de São Paulo, Piracicaba, v. 1, 105p., 2001.
- MENDELSON, A.; FRENETTE, P. S. Hematopoietic stem cell niche maintenance during homeostasis and regeneration. **Nature Medicine**, v. 20; n. 8, p. 833-846, 2014.
- MINISTÉRIO DA SAÚDE [homepage on the Internet]. Disponível em: <http://portal.saude.gov.br/saude>. [2009 out 20].
- MONTGOMERY, D. C.; MYERS, R. H.; ANDERSON-COOK, C. M. **Response surface methodology: process and product optimization using designed experiments**. Third Edition. John Wiley & Sons, Inc., Hoboken, New Jersey, 2009.
- MYERS, R. H.; MONTGOMERY, D. C. Response surface methodology: process and product optimization using designed experiments. **New York: John Wiley & Sons**, 700 p., 1995.
- NAGATA, S.; TSUCHIYA, M.; ASANO, S.; et al. Molecular cloning and expression of cDNA for human granulocyte colony-stimulating factor. **Nature**; v. 5, n. 319, p. 415-418, 1986.
- OH-EDA, M.; HESEGAWA, M.; HATTORI, K.; KUBONIWA, H.; KOJIMA, T.; ORITA, T.; et al. O-linked sugar chain of human granulocyte colony-stimulation factor protects it against polymerization and denaturation allowing it to retain its biological activity. **J. Biol. Chem.** v. 265, p. 432-435, 1990.
- ORAWAN, K.; SUNUTCHA, S.; Strategies for production of active eukaryotic proteins in bacterial expression system. **Asian Pac J Trop Biomed**.v.2, n. 2, p. 159-162, 2012.

- PERÁZIO, B. A. A. Efeitos dos pontos axiais e centrais sobre a eficiência do delineamento composto central. Dissertação (mestrado) – Universidade Federal de Viçosa, Viçosa, MG, 2010.
- PLACKETT, R. L.; BURMAN, J. P. The design of optimum multifactorial experiments. **Biometrika**, v. 33, n. 4, p. 305-325, 1946.
- POUTOU, R. et al. Influencia de la fuente de carbon sobre la expresión de proteínas AOX 1-reguladas em *Pichia pastoris*. ISSN: 1794-2470, VOL.3 No. 3:1-120, enero – junio, 2005.
- RODRIGUES, M. I.; IEMMA, A. F. **Planejamento de experimentos e otimização de processos**. 2^a. ed. Campinas, SP: Cárita Editora, 2009.
- RUSSO-CARBOLANTE, E. M. S.; PICANÇO-CASTRO, V. SWIECH, K.; COVAS, D. T. Produção estável e em larga escala de FVII humano em linhagem celular humana Sk-Hep-1. 2011, Brasil.
- RUSSO, E. M. S. Large Scale and stable production of human FVII in the human cell line SK-Hep1. 2013, Estados Unidos. Patente: Privilégio de Inovação. Número do registro: 13/981,446, título “Large Scale and stable production of human FVII in the human cell line SK-Hep1”, Instituição de registro: United States Patent and Trademark Office. Depósito: 10/04/2013.
- SHADLE, P. J.; ALLEN, J. I.; GEIER, M. D.; et al. Detection of endogenous macrophage colony-stimulating factor (M-CSF) in human blood. **ExpHematol**, v. 17, n. 2, p. 154-159, 1989.
- SCHMIDT, F. Recombinant expression system in the pharmaceutical industry. **Applied microbiology and biotechnology**, v. 65, n. 4, p. 3636-3672, 2004.
- SHIRAKI, K.; KUDOU, M.; FUJIWARA, S.; IMANAKA, T.; TAKAGI, M. Biophysical effect of amino acids on the prevention of protein aggregation. **J. Biochem**, v. 132, p. 591-595, 2002.
- SICA, S.; RUTELLA, S.; DI MARIO, A.; SALUTARI, P. RUMI, C. ORTU LA BARBERA, E.; ETUK, B.; MENICHELLA, G.; D’ONOFRIO, G.; LEONE, G. rhG-CSF in healf donors: mobilization of peripheral hemopoietic progenitors and effect on peripheral blood leukocytes. **J Hemat-other**, v. 5, p. 391-397, 1996.
- SOUETRE, E.; QING, W.; PENELAUD, P. F. Economic analysis of the use of recombinant human granulocyte colony stimulating factor in autologous bone marrow transplantation. **European Journal of Cancer**; v. 32, n. 7, p. 1162-1165, 1996.
- SOUZA, L. Production of pluripotent granulocyte colony-stimulating factor. 1989. Patentee US 4, 810, 643.

- SOUZA, L. M.; BOONE, T. C.; GABRILOVE, J. et al. Recombinant human granulocyte colony-stimulating factor: effects on normal and leukemic myeloid cells. **Science**, v. 232, n. 4746, p. 61-65, 1986.
- TORRES, F. A. G.; MORAES, L. M. P. **Proteínas recombinantes produzidas em leveduras**. Biotecnologia Ciência e Desenvolvimentos, Brasil, v. 12, p. 20-22, 2000.
- U.S Food and Drug Administration [homepage on the internet].Disponível em: <http://www.accessdata.fds.gov>. [2015 mai 15].
- VALLEJO, L. F.; RINAS, U. Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. **Microbial Cell Factories**, v. 3, p. 1-11, 2004.
- VANZ, A. L. S.; RENARD, G.; PALMA, M. S.; et al. Human granulocyte colony stimulating factor (hG-CSF): cloning, overexpression, purification and characterization. **Microbial Cell Factories**, v.7, n. 13, 2008.
- VASINA, J.; BANEYX, F. Expression of aggregation-prone recombinant proteins at low temperature: a comparative study of the *Escherichia coli* cspA and tac promoter systems. **Protein Expression and Purification**, v. 9, p. 211-218, 1997.
- WALSH, G. Biopharmaceutical: biochemistry and biotechnology. Hoboken, N.J.: John Wiley & Sons, 2nd edition. New York: Wiley, 2003.
- WANG, W.; Instability, stabilization and formulation of liquid protein pharmaceuticals. **International Journal of Pharmaceutics**. V. 185, n. 2, p. 129-188, 1999.
- WITTMAN, B; HORAN, J.; LYMAN, G. H. Prophylactic colony-stimulating factor in children receiving myelosuppressive chemotherapy: a meta-analysis of randomized controlled trials. **Cancer Treatment Reviews**, v. 32, n. 4, p. 289-303, 2006.
- XUAN, Y. et al. An upstream activation sequence controls the expression. Of AOX1 gene in *Pichia pastoris*. **FEMS Yeast Res**, v. 8, n. 9, p. 1271-1282, 2009.
- ZAGO, M.; FALCÃO, R.; PASQUINI, R. **Hematologia: fundamentos e prática**. São Paulo: Atheneu, 2001.