



Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Biologia Celular
Programa de Pós-Graduação em Tecnologias Química e Biológica

Produção de Fator de Crescimento Epidermal Humano em *Komagataella phaffii*

Ana Laura Alfonso Pérez

Orientador: Prof. Fernando Araripe Gonçalves Torres

Brasília-DF, fevereiro de 2018

Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Biologia Celular
Programa de Pós-Graduação em Tecnologias Química e Biológica

Produção do fator de crescimento epidermal humano (hEGF) em *Komagataella phaffii*

Ana Laura Alfonso Pérez

Orientador: Prof. Fernando Araripe Gonçalves Torres

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Tecnologias Química e Biológica da Universidade de Brasília como requisito parcial para obtenção do título de mestre.

Brasília-DF, fevereiro de 2018

Trabalho realizado no Laboratório de Biologia Molecular, Departamento de Biologia Celular, Instituto de Ciências Biológicas, Universidade de Brasília, sob a orientação do professor Fernando Araripe Gonçalves Torres e com apoio financeiro do PEC-PG CNPq.

Banca Examinadora

Professor Dr. Fernando Araripe Gonçalves Torres
Universidade de Brasília
Presidente da Banca

Professor Dr. Joao Ricardo de Almeida
Embrapa Agroenergia
Membro titular

Professora Dra. Viviane Castelo Branco Reis
Universidade de Brasília
Membro titular

Professora Dra. Lídia Maria Pepe de Moraes
Universidade de Brasília
Membro suplente

*"Caminhando em linha reta não se
pode chegar muito longe"*

Antoine de Saint-Exupéry

Aos meus pais, que são meu paradigma.

Aos meus irmãos, ao meu esposo.

AGRADECIMENTOS

À minha família pelo apoio durante todo este tempo, apesar da distância sempre estiveram comigo.

À minha mãe Mary Tere e ao meu pai Mauro por ser meu exemplo de vida, minha meta. Por acreditar em mim. Por serem os responsáveis de ser quem sou. Por me ensinar que cada desafio me faz crescer e ser melhor pessoa. Obrigada por tudo!

Aos meus irmãos: a Alejandra, pela amizade, pelo amor, por ser tudo que preciso e estar sempre para mim. Por sem falar, saber o que preciso. Ao Mauri, pelo carinho, e a pesar da distância, me deu um dos melhores presentes da minha vida: as minhas sobrinhas!

À minha família do Brasil, ao meu irmão Dani por estar sempre para mim, ainda quando não preciso. Por ser minha família aqui. A Sabrina, por cuidar dele e por fazer de meus finais de semana mais divertidos.

Ao meu esposo Alejandro, por me ajudar a cumprir este sonho. Porque decidimos começar esta nova etapa juntos e tem sido meu apoio sempre. Porque não importa se é difícil, é o nosso sonho. À Belkys, pelo carinho e pelos conselhos sempre bons, pelo amor incondicional.
À Adri, pelas alegrias.

Ao professor Fernando Araripe, por me permitir entrar no programa sem me conhecer. Por acreditar em mim. Por todos os ensinamentos e pela preocupação sempre. Por estar sempre para mim.

Aos professores que me ajudaram em todo momento. À Vivis, pelo apoio e sempre estar presente para tirar todas as dúvidas, e me ensinar a escutar as palavras mais baixas.

À Janice, pela ajuda no AKTA e pela preocupação sempre, pelo carinho.

À professora Lídia, pelos conselhos sempre úteis.

Aos meus amigos no Lab 2 “Rainha” por compartilhar tantos momentos de alegrias e trabalho, tantas horas na copa falando de exatamente qualquer coisa: Tiago, Chris, Vanessa, Marciano, Francisco, Carol, Roberta, Ieda, Dani, Igor, Diego. A Myrna pela ajuda sempre, pelos conselhos.

À Lu, pela ajuda na correção da dissertação, pelas dúvidas a qualquer hora. Por ser uma amiga sempre presente. Por estar sempre olhando com os olhos do coração.

À Mari, que mais que uma companheira de laboratório, foi uma amiga, uma irmã. Obrigada pelos ratos de trabalhos, as festas, tudo!!! Saudades Mari.

Ao professor Osmindo, pela ajuda nos experimentos de purificação no AKTA, e sobre tudo pela paciência! Obrigada por acreditar que no final todo ia dar certo. À Ana pela ajuda com as células e a Carlos pela ajuda no espectrômetro de massas. Tudo acaba!

Às minhas amigas “mijischurritis”: Gaby e Wendy por estar sempre para mim. Por tantos anos juntas, por tantas horas divertidas. Amigas como vocês fazem falta no Brasil.

Ao Amauri pela ajuda nas correções e pela disposição sempre a me ajudar, por ser a família que precisamos.

À Eve pelo carinho, pelas risadas e a ajuda sempre.

Aos membros da banca examinadora, pela disponibilidade e paciência para a correção do meu trabalho.

Ao PEC-PG/CNPq, pelo apoio financeiro.

A todos que de alguma forma colaboraram para a realização deste trabalho.

ÍNDICE

LISTA DE FIGURAS	i
LISTA DE TABELAS.....	iii
LISTA DE ABREVIATURAS E SÍMBOLOS	iv
RESUMO.....	vi
ABSTRACT	vii
RESUMEN	viii
1. INTRODUÇÃO	1
1.1. <i>Komagataella phaffii</i> e expressão heteróloga.....	1
1.1.1. Vetores de expressão.....	3
1.1.2. Promotores	4
1.1.3. Marcas de seleção	5
1.1.4. Linhagens hospedeiras. Cepas mais utilizadas	6
1.1.5. Secreção de proteínas.....	8
1.1.6. Outros aspectos de interesse	9
1.2. <i>K. phaffii</i> e a produção de biofármacos	9
1.2.1. O fator de crescimento epidermal humano (hEGF) e sua produção em leveduras....	11
1.2.1.1. O hEGF	11
1.2.1.2. Produção do hEGF	12
2. OBJETIVO.....	16
2.1. Objetivo Geral	16
2.2. Objetivos específicos	16
3. ESTRATÉGIA	16
4. MATERIAIS E MÉTODOS	17
4.1. Materiais	17
4.1.1. Células.....	17

4.1.2.	Meios de cultura.....	17
4.1.3.	Soluções	18
4.1.4.	Kits.....	23
4.1.5.	Oligonucleotídeos (<i>primers</i>)	23
4.1.6.	Marcadores de massa molecular para DNA.....	24
4.1.7.	Marcadores de massa molecular para proteínas.....	24
4.1.8.	Enzimas de restrição	24
4.1.9.	Vetores	24
4.2.	Métodos	25
4.2.1.	Cultivo de micro-organismos.....	25
4.2.2.	Amplificação por PCR.....	25
4.2.3.	Preparação de células bacterianas competentes	25
4.2.4.	Manipulação do DNA	25
4.2.5.	Construção de plasmídeos.....	26
4.2.6.	Preparação de DNA plasmidial.....	26
4.2.7.	Tranformação de leveduras por eletroporação.....	27
4.2.8.	PCR de colônia	28
4.2.9.	Avaliação do crescimento em meio líquido.....	28
4.2.10.	Crescimento e expressão de proteínas recombinantes em frasco.....	28
4.2.11.	Concentração do sobrenadante.....	29
4.2.12.	Precipitação das proteínas com TCA	29
4.2.13.	Quantificação de proteínas pelo Método de Bradford	29
4.2.14.	Análise eletroforética em gel de poliacrilamida desnaturante	29
4.2.15.	Western Blot.....	30
4.2.15.1.	Transferência para a membrana	30
4.2.15.2.	Bloqueio da membrana.....	30
4.2.15.3.	Anticorpos e revelação	30

4.2.16.	Purificação do hEGF	30
4.2.16.1.	Cromatografia de Exclusão Molecular (FPLC)	31
4.2.16.2.	Purificação da fração do EGF/FPLC por HPLC-fase reversa.....	31
4.2.16.3.	Purificação da fração do EGF/HPLC por HILIC	31
4.2.17.	Espectrometria de massa MALDI-TOF/TOF	32
4.2.18.	Degradação de Edman e sequenciamento automático.....	32
4.2.19.	Análise de dados.....	32
5.	RESULTADOS	33
5.1.	Construção dos cassetes de expressão para hEGF contendo diferentes peptídeos sinais 33	
5.2.	Construção dos vetores pK α FEGF-ld, pKSUC2EGF-ld e pKPHO1EGF-ld	34
5.3.	Transformação de <i>K. phaffii</i> M12-K	35
5.4.	Confirmação da presença do gene do hEGF	37
5.5.	Avaliação do crescimento dos clones transformantes	38
5.6.	Crescimento e expressão dos clones positivos em frasco.....	39
5.7.	Identificação do hEGF.....	40
5.7.1.	Avaliação das proteínas presentes no sobrenadante	40
5.7.2.	Confirmação da presença do EGF	40
5.8.	Purificação e identificação da proteína recombinante	41
5.8.1.	Purificação por Cromatografia de Exclusão Molecular (FPLC).....	42
5.8.1.1.	SDS-PAGE das frações da cromatografia de exclusão molecular	45
5.8.1.2.	Western blot das frações da cromatografia de exclusão molecular	46
5.8.2.	HPLC em fase reversa	47
5.8.3.	Espectrometria de massas	49
5.9.	Melhoramento do processo de purificação do hEGF na amostra clone 1 ^{αF}	52
5.9.1.	HILIC.....	52
5.9.2.	Espectrometria de massas	53

5.9.3. Sequenciamento automático de Edman	54
6. DISCUSSÃO.....	56
7. CONCLUSÕES E PERSPECTIVAS.....	60
8. REFERÊNCIAS	61
ANEXOS	69
Anexo 1. Mapa físico do vetor pK-LeudEGF	69
Anexo 2. Perfil cromatográfico em FPLC do EGF padrão (Sigma, EUA).	70
Anexo 3. Perfil cromatográfico em RP-HPLC do EGF padrão (Sigma, EUA).	71
Anexo 4. Espectro de massas do EGF padrão (Sigma, EUA).	72

LISTA DE FIGURAS

Figura 1. Representação esquemática para a expressão de um gene de interesse em <i>K. phaffii</i> .	3
Figura 2. Sítios de clivagem no fator α .	8
Figura 3. Sequência de aminoácidos do hEGF.	11
Figura 4. Estratégia de trabalho para a produção de hEGF.	16
Figura 5. Amplificação por PCR dos cassetes α FEGF, SUC2EGF e PHO1EGF.	33
Figura 6. Mapas físicos dos vetores de expressão utilizados para a construção dos plasmídeos.	35
Figura 7. Análise da clonagem, perfil de restrição dos vetores construídos.	35
Figura 8. Linearização dos vetores pK α FEGF-ld, pKSUC2EGF-ld e pKPHO1EGF-ld com a enzima SacI.	36
Figura 9. Resultado da transformação de <i>K. phaffii</i> M12-K com os vetores pK α FEGF-ld (A), pKSUC2EGF-ld (B) e com pKPHO1EGF-ld (C).	37
Figura 10. PCR de colônia de 6 clones de cada sistema.	38
Figura 11. Cinética de crescimento dos clones transformados com os vetores pK α FEGF-ld, pKSUC2EGF-ld e pKPHO1EGF-ld.	38
Figura 12. SDS-PAGE 15% de clones selecionados.	40
Figura 13. Análise da produção de EGF por Western Blot em clones selecionados.	41
Figura 14. Perfil cromatográfico em FPLC da amostra clone1 ^{αF} .	43
Figura 15. Perfil cromatográfico em FPLC da amostra clone23 ^{SUC2} .	43
Figura 16. Perfil cromatográfico em FPLC da amostra clone1 ^{PHO1} .	44
Figura 17. Perfil cromatográfico em FPLC da amostra M12-K.	44
Figura 18. SDS-PAGE das frações coletadas em FPLC. SDS-PAGE 15%.	45
Figura 19. Western blot das frações coletadas em FPLC.	46
Figura 20. Perfil cromatográfico em RP-HPLC da amostra clone 1 ^{αF} .	47

Figura 21. Perfil cromatográfico em RP-HPLC da amostra clone 23 ^{SUC2}	48
Figura 22. Perfil cromatográfico em RP-HPLC da amostra clone 1 ^{PHO1}	48
Figura 23. Perfil cromatográfico em RP-HPLC da amostra M12-K.....	49
Figura 24. Espectro de massa da amostra clone 1 ^{αF}	50
Figura 25. Espectro de massa da amostra clone 23 ^{SUC2}	51
Figura 26. Perfil cromatográfico em HILIC da amostra clone 1 ^{αF}	53
Figura 27. Espectro de massa da amostra clone 1 ^{αF}	54
Figura 28. Resultado do alinhamento da sequência de aminoácidos obtida pelo método de degradação de Edman.....	55

LISTA DE TABELAS

Tabela 1. Genótipos e fenótipos das principais linhagens de <i>Komagataella sp.</i>	7
Tabela 2. <i>Primers</i> utilizados. Os sítios de restrição estão sublinhados	23
Tabela 3. Enzimas de restrição utilizadas.....	24
Tabela 4. Vetores utilizados	24
Tabela 5. Velocidade específica máxima de crescimento ($\mu_{\text{máx}}$) dos clones avaliados.....	39
Tabela 6. OD _{600nm} dos clones selecionados após 24 h de crescimento.	39
Tabela 7. OD _{600nm} e fatores de concentração dos clones selecionados após 24 h.	42

LISTA DE ABREVIATURAS E SÍMBOLOS

ACN	acetonitrila
ADP	adenosina difosfato
α -F	fator α
APS	persulfato de amônio
ATP	adenosina trifosfato
BSA	albumina sérica bovina
DNA	ácido desoxirribonucleico
EDTA	ácido etileno-diamino-tetra-acético
EGF	<i>epidermal growth factor</i> ou fator de crescimento epidermal
FDA	<i>federal Drug Administration</i>
HILIC	cromatografia líquida com interação hidrofílica
MALDI TOF/TOF	ionização e dessorção a laser assistida por matriz-tempo de voo/tempo de voo
MF- α	<i>mating factor α</i>
OD ₆₀₀	densidade celular óptica medida a 600 nm
PCR	reação em cadeia da polimerase
RNase A	ribonuclease A
RP-HPLC	cromatografia líquida de alta eficiência em fase reversa
SCP	<i>single cell protein</i>
SDS	dodecilsulfato de sódio
TAE	tampão tris-acetato-EDTA
TCA	ácido tricloroacético
TEMED	N,N,N',N'-Tetrametiletenodiamina

TFA	ácido trifluoroacético
VMA	adesão vitreomacular sintomática
YNB	<i>yeast nitrogen base</i>

RESUMO

O sistema de expressão baseado na utilização da levedura *Komagataella phaffii* tem sido utilizado com sucesso na produção de uma grande variedade de proteínas heterólogas. Esta levedura reúne características como fácil manipulação molecular, crescimento celular rápido, capacidade de realizar modificações pós-traducionais e secreção eficiente de proteínas, além de atingir altas densidades celulares com grande produção da proteína heteróloga. Uma das proteínas de grande interesse na indústria biofarmacêutica e cosmética é o fator de crescimento epidermal humano (hEGF). O objetivo desse estudo foi desenvolver um sistema para a produção de hEGF na levedura *K. phaffii*. Para isso foram utilizados os peptídeos sinais dos genes αF , *SUC2* e *PHO1*. Três vetores foram construídos, cada um deles contendo um dos peptídeos sinal. Foram construídos os vetores de expressão sob o controle do promotor do *PGK1* e utilizados para transformação de *K. phaffii* M12-K, uma linhagem mutante para o gene *KEX1*. Os clones recombinantes foram confirmados por PCR de colônia e foram crescidos em meio mínimo para avaliar a cinética de crescimento. Os clones positivos foram selecionados e utilizados na expressão em frasco empregando meio complexo. A presença do hEGF foi avaliada por SDS-PAGE e confirmada por Western blot na presença de um anticorpo específico. Com isso, um clone de cada sistema foi escolhido para a purificação do hEGF por cromatografia de exclusão molecular e RP-HPLC. A partir desses experimentos foi confirmada a presença do hEGF e o polipeptídeo foi parcialmente purificado. Finalmente, foi escolhido o clone 1 ^{αF} para otimizar o processo de purificação, adicionando uma etapa em HILIC. O hEGF foi satisfatoriamente purificado. O tamanho correto e a sequência do N-terminal igual ao EGF presente em humanos foram confirmados por espectrometria de massas e sequenciamento automático. Os resultados obtidos mostram que o hEGF foi produzido com sucesso em *K. phaffii* com a sequência primária correta.

ABSTRACT

The expression system based on the utilization of the yeast *Komagataella phaffii* has been used successfully in the production of a large variety of heterologous proteins. This yeast combines several important features such as easy molecular manipulation, rapid cell growth, ability to perform post-translational modifications and efficient secretion of proteins, in addition to large heterologous protein production at high cell densities. One of the proteins of great interest in the biopharmaceutical and cosmetic industry is the human epidermal growth factor (hEGF). The aim of this study was to develop a system for hEGF production in *K. phaffii*. For this, the 3 different signal peptides sequences from genes αF , *SUC2* and *PHO1* were tested. Three vectors were constructed, each containing one of the signal peptides. Expression vectors were constructed under the control of the *PGK1* promoter and used for transformation of *K. phaffii* M12-K, a strain mutant for the *KEX1* gene. Recombinant clones were confirmed by colony PCR and grown in minimal medium to evaluate growth kinetics. Positive clones were selected and used in flask expression using complex media. The presence of hEGF was assessed by SDS-PAGE and confirmed by western blot with a specific antibody. One clone from each system was chosen for the production and purification of hEGF by molecular-exchange chromatography and RP-HPLC. From these experiments the presence of hEGF was confirmed and the polypeptide was partially purified. Finally, clone 1 ^{αF} was chosen to optimize the purification process by adding one step in HILIC. The hEGF was successfully purified. Correct size and N-terminal sequence equal to EGF present in humans were confirmed by mass spectrometry and Edman sequencing. Together, our results show that hEGF was successfully produced in *K. phaffii* with the proper primary structure.

RESUMEN

El sistema de expresión basado en la utilización de la levadura *Komagataella phaffii* ha sido utilizado con éxito en la producción de una gran variedad de proteínas heterólogas. Esta levadura reúne características como fácil manipulación molecular, rápido crecimiento celular, capacidad de realizar modificaciones pos-traduccionales y eficiente secreción de proteínas, además de llegar hasta altas densidades celulares con elevada producción heteróloga. Una de las proteínas de gran interés en la industria biofarmacéutica y cosmética es el factor de crecimiento epidérmico humano (hEGF). El objetivo de este estudio fue desarrollar un sistema para la producción de hEGF en la levadura *K. phaffii*. Para eso fueron utilizadas las señales de los genes αF , *SUC2* e *PHO1*. Fueron construidos tres vectores, cada uno de ellos conteniendo una de las secuencias señales. Los vectores fueron construidos sobre el control del promotor del gen *PGK1* y fueron utilizados para la transformación de la cepa M12-K de *K. phaffii*, que es mutante para el gen *KEX1*. Los clones recombinantes fueron confirmados mediante PCR de colonias y fueron crecidos en medio mínimo para analizar la cinética de crecimiento. Los clones positivos fueron seleccionados y utilizados para expresión en frasco utilizando medio complejo. La presencia de hEGF fue analizada por SDS-PAGE e confirmada por western blot en presencia de un anticuerpo específico. A partir de ahí, una muestra de cada sistema fue escogida para la purificación de hEGF utilizando cromatografía de exclusión molecular y RP-HPLC. A partir de esos experimentos fue confirmada la presencia de hEGF y el polipeptídeo fue parcialmente purificado. Finalmente, fue escogido el clon 1^{αF} para optimizar el proceso de purificación, adicionando una cromatografía en HILIC. El hEGF fue satisfactoriamente purificado. El tamaño correcto y la secuencia del N-terminal igual al EGF presente en humanos fueron confirmados por espectrometría de masas y secuenciamiento automático. Los resultados obtenidos muestran que fue producido hEGF en *K. phaffii*, con la secuencia primaria correcta.

1. INTRODUÇÃO

1.1. *Komagataella phaffii* e expressão heteróloga

O uso das leveduras como sistema de expressão vem sendo realizado com sucesso crescente ao longo dos últimos 20 anos (Damasceno *et al.*, 2012). Esses organismos são uma eficiente plataforma para a expressão de proteínas recombinantes e representam uma das alternativas mais utilizadas para a produção de proteínas em larga escala (Spohner *et al.*, 2015). *Komagataella phaffii* (inicialmente classificada como *Pichia pastoris*) (Kurtzman, 2005) é uma das leveduras mais empregadas para expressão heteróloga (Kurtzman, 2009). Foi introduzida pela empresa *Phillips Petroleum* há mais de quatro décadas na produção comercial de “*single cell protein* (SCP)” como suplemento animal (Cereghino *et al.*, 2000; Spohner *et al.*, 2015). A mais notável característica fisiológica dessa levedura é o fato de ser metilotrófica, ou seja, capaz de crescer em meio contendo metanol como única fonte de carbono (Bertheleme *et al.*, 2015). Pertence ao grupo dos ascomicetos, é homotática e haploide a menos que seja submetida a condições de limitação por nitrogênio (Celik *et al.*, 2012).

Não existe uma solução universal para a produção de proteínas recombinantes e ainda não é possível prever qual sistema de expressão é o melhor para produzir uma proteína com alto rendimento e capacidade funcional (Bill, 2014). Há uma série de vantagens que fazem de *K. phaffii* uma ótima escolha como sistema de expressão de proteínas recombinantes (Spohner *et al.*, 2015; Weinacker *et al.*, 2013).

Dentre suas principais características destaca-se a fácil manipulação genética: as técnicas moleculares desenvolvidas para *Saccharomyces cerevisiae*, um dos sistemas experimentais mais bem caracterizados, são facilmente aplicáveis em *K. phaffii* (Cereghino *et al.*, 2000; Spohner *et al.*, 2015). Recentemente, foi descrito o uso do sistema CRISPR/cas9 em *K. phaffii* como ferramenta para realizar modificações genéticas específicas: ruptura gênica e integração de cassetes de DNA homólogo (Weninger *et al.*, 2016). Essa levedura tem a capacidade de crescer em meio contendo metanol, o que inibe até certo ponto o crescimento de contaminantes. É um sistema barato para o estabelecimento e manutenção em escala industrial. *K. phaffii* tem rápido crescimento celular, podendo chegar até altas densidades celulares (200 g/L) durante uma fermentação em batelada alimentada com limitação de glicose (Heyland *et al.*, 2010). A secreção de poucas proteínas endógenas é outra importante

vantagem, pois facilita a purificação de proteínas que sejam exportadas ao meio extracelular. Por último, a levedura tem a habilidade de realizar modificações pós-traducionais como clivagem proteolítica, glicosilação e formação de ligações dissulfeto, além da ausência do padrão de hiperglicosilação presente em *S. cerevisiae* (Weinacker *et al.*, 2013).

Todas estas vantagens fazem com que atualmente *K. phaffii* constitua um versátil sistema de expressão e a plataforma escolhida para a produção de várias proteínas de interesse na indústria alimentícia e na saúde (He *et al.*, 2016; Spohner *et al.*, 2015; Weinacker *et al.*, 2013). O sistema de expressão baseado no uso da levedura *K. phaffii* vem sendo usado com sucesso na produção de proteínas bacterianas, fúngicas, de plantas, animais e humanos (Weinacker *et al.*, 2013). Mais de 5000 proteínas já foram produzidas em *K. phaffii* e mais de 70 encontram-se no mercado ou na etapa final de desenvolvimento. Exemplos de proteínas heterólogas produzidas por *K. phaffii* vão desde enzimas bacterianas até biofármacos e anticorpos monoclonais. Biofármacos como Kalbitor® (ecallantide: proteína recombinante inibidora de calicreína) e Insugen® (insulina humana recombinante), usados para tratar a angioedema e a diabetes, respectivamente, são produzidos em *K. phaffii* e foram aprovados pela *Federal Drug Administration* (FDA, EUA) (Bernstein *et al.*, 2012). Algumas das principais áreas de investigação e aplicação de proteínas expressas em *K. phaffii* são, diabetes, alergias, câncer, papiloma vírus humano e aterosclerose (Vogl *et al.*, 2013; Weinacker *et al.*, 2013)

A expressão de qualquer gene de interesse em *K. phaffii* consta de três etapas básicas: uma primeira etapa onde é inserido o gene de interesse em um vetor de expressão; uma segunda etapa para integração do vetor de expressão dentro do genoma da levedura; e uma terceira para a seleção dos clones que potencialmente estiverem expressando o gene de interesse (Cereghino *et al.*, 2000; Li *et al.*, 2007) (Figura 1).

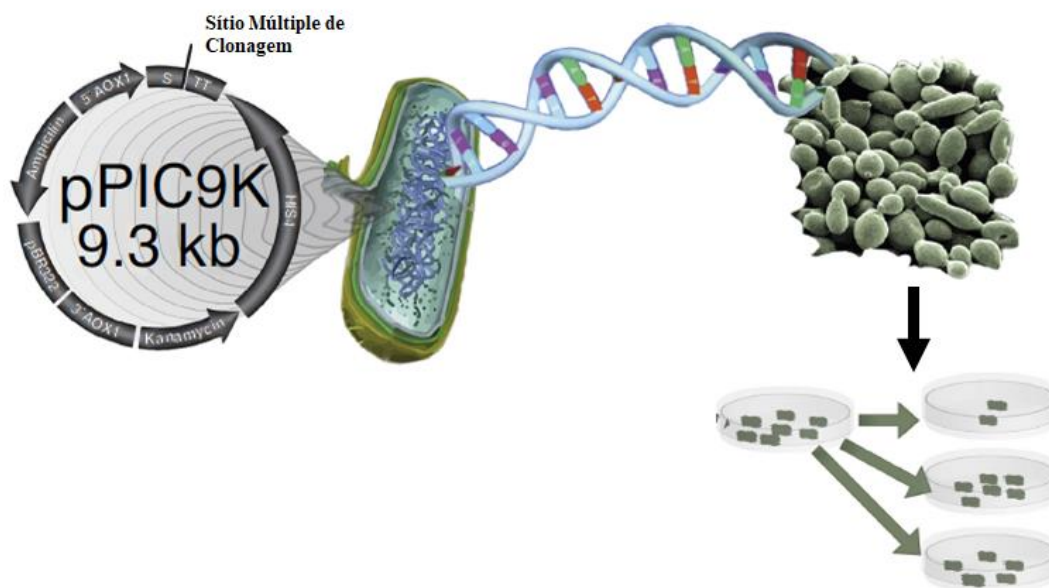


Figura 1. Representação esquemática para a expressão de um gene de interesse em *K. phaffii*. Expressão por integração no genoma. Uma vez confirmada a sequência correta, o vetor de expressão contendo o gene de interesse é transformado em células de *K. phaffii*. Posteriormente, são selecionadas as colônias recombinantes que expressam a proteína desejada. Adaptado de (Byrne, 2015).

Quando são desenhadas as estratégias para a clonagem e expressão de proteínas heterólogas em *K. phaffii* alguns pontos importantes devem ser considerados. Por exemplo: a seleção de uma combinação promotor-terminador, a presença de marcadores de seleção e a aplicação dos sistemas tanto para expressão intracelular como para secreção (com o uso de um adequado peptídeo sinal). A seleção do vetor de expressão e da linhagem apropriados são os mais importantes pré-requisitos para a expressão de proteínas recombinantes com sucesso (Ahmad et al., 2014).

1.1.1. Vetores de expressão

Os vetores de expressão para *K. phaffii* são geralmente plasmídeos integrativos. São desenhados para conter elementos genéticos para manipulação e propagação em *Escherichia coli*; sequências regulatórias (promotores), induzíveis ou constitutivos, que são reconhecidas pela célula hospedeira; sequências terminadoras de transcrição de levedura (*AOX1tt*), marcas de seleção auxotrófica ou dominante e sítios múltiplos de restrição para clonagem (Cereghino et al., 2000; Damasceno et al., 2012; Logez et al., 2012). Para a secreção da proteína recombinante ao meio extracelular, os vetores de expressão contêm um sinal de secreção, uma sequência que codifica um peptídeo sinalizador que direciona a exportação da proteína heteróloga para o meio extracelular, como é o caso do fator alfa (α -MF) de *S. cerevisiae*

(Wang *et al.*, 2016). Alguns vetores ainda podem conter, opcionalmente, sequências “tag” (*HIS6* ou *c-Myc*), que facilitam a identificação e purificação da proteína heteróloga (Cereghino *et al.*, 2000).

1.1.2. Promotores

Para a expressão de genes heterólogos, diferentes promotores endógenos, fortes ou fracos, podem ser utilizados. Os promotores utilizados nos vetores de expressão podem ser induzíveis ou constitutivos (Spohner *et al.*, 2015). Entre os promotores induzíveis, o mais usado é o promotor do gene *AOXI* (P_{AOXI}). O gene *AOXI* codifica a primeira enzima na via de utilização de metanol. Este promotor é firmemente controlado em nível transcricional por um mecanismo de repressão/desrepressão mediado pela presença do agente indutor, o metanol. Ao mesmo tempo, é fortemente reprimido na presença de glicose ou glicerol, fontes de carbono preferenciais (Ahmad *et al.*, 2014; Cereghino *et al.*, 2000; Tschopp *et al.*, 1987; Vogl *et al.*, 2013).

Apesar de ser o promotor mais estudado e, como consequência, mais utilizado, o P_{AOXI} nem sempre é a melhor opção. Seu emprego com produtos utilizados para a indústria biofarmacêutica é desaconselhado, uma vez que altas doses de metanol podem ter efeitos tóxicos. Além disso, uma vez adicionado à cultura, é difícil reduzir sua dose. Seu uso em biorreatores de larga escala também pode ser problemático, já que o metanol é altamente inflamável e o armazenamento de grandes quantidades desse combustível é perigoso (Prielhofer *et al.*, 2013; Vogl *et al.*, 2013; Wang *et al.*, 2017)

Outros promotores induzíveis utilizados são os dos genes que codificam a enzima formaldeído desidrogenase (P_{FLDI}) e uma proteína de membrana do peroxissomo (P_{PEX8}) que, além do metanol, podem ser induzidos com metilamina e ácido oleico respectivamente. Embora o P_{PEX8} proporcione uma alternativa ao sistema induzível por metanol, não é muito usado devido ao seu baixo nível de expressão (Vogl *et al.*, 2013). Outros três novos promotores induzíveis de *K. phaffii* são aqueles que regulam a expressão dos genes da álcool desidrogenase (P_{ADHI}), glicerol quinase (P_{GUT}) e enolase (P_{ENO1}) (Ahmad *et al.*, 2014; Liu *et al.*, 1995; Shen *et al.*, 1998; Stadlmayr *et al.*, 2010). Recentemente, foram descobertos cinco genes induzíveis por ramnose e, dentre eles, P_{LRA3} e P_{LRA4} são ativados na presença de ramnose e reprimidos na presença de glicose (Liu *et al.*, 2016)

Dentre os promotores constitutivos, destacam-se aqueles que transcrevem genes da via glicolítica, como o P_{GAP} (gliceraldeído 3-fosfato desidrogenase) e o P_{PGKI} (fosfoglicerato quinase) (Arruda *et al.*, 2016; de Almeida *et al.*, 2005; Waterham *et al.*, 1997). Estes promotores permitem a expressão de genes heterólogos em meios contendo glicose, glicerol e outras fontes de carbono como substrato. O P_{TEFI} (fator de alongação da tradução) é também um promotor constitutivo que apresenta uma força comparável ao P_{GAP} (Ahn *et al.*, 2007; Logez *et al.*, 2012; Vogl *et al.*, 2013). No nosso grupo de pesquisa foi isolado o promotor do gene *PGKI* que codifica para a enzima glicolítica 3-fosfoglicerato quinase, a qual transfere um grupo fosforil do 1,3-bifosfoglicerato para ADP, produzindo ATP e 3-fosfoglicerato (de Almeida *et al.*, 2005). Foi observada uma alta eficiência na produção da enzima α -amilase de *Bacillus subtilis* o que demonstrou a possibilidade de usar o P_{PGKI} como um promotor constitutivo (Arruda *et al.*, 2016). A grande vantagem de utilizar P_{PGKI} para controlar a expressão do gene heterólogo é que não é necessário realizar mudanças no meio de cultura para passar de uma fase de crescimento para uma fase de expressão.

Quando o correto dobramento da proteína ou o processamento na via secretória são as etapas limitantes para obter elevadas quantidades do produto ativo, um promotor que apresente uma forma menor pode ser desejável (Mattanovich *et al.*, 2004). Por isso é importante a identificação de promotores com diferentes propriedades regulatórias e níveis de expressão. Alguns destes promotores são o P_{PET9} , o P_{YPT1} e o P_{GMP1} (Sears *et al.*, 1998; Stadlmayr *et al.*, 2010).

1.1.3. Marcas de seleção

Independentemente do tipo de vetor de expressão utilizado, este deve possuir uma marca de seleção que permita identificar as células que receberam o DNA exógeno. Existem poucas marcas disponíveis para a manipulação genética em *K. phaffii*. Existem as marcas de seleção limitadas a genes de vias biossintéticas como *HIS1*, *HIS2*, *HIS4*, *ARG4*, *URA3*, *URA5*, *ADE1*, *LEU2*, dentre outras (Ahmad *et al.*, 2014; Cereghino *et al.*, 2000). Na utilização deste tipo de marca, primeiramente é preciso criar uma linhagem auxotrófica por meio da inativação de um gene de via biossintética. Nesse caso, o vetor de transformação a ser utilizado possui uma cópia funcional do gene alterado, assim permitindo o crescimento dos transformantes sem a necessidade de suplementação com a biomolécula produzida pela via metabólica em questão (Cereghino *et al.*, 2000; Piva *et al.*, 2017).

As marcas auxotróficas defectivas consistem em genes de vias biossintéticas que possuem seu promotor reduzido. Tal característica faz com que a taxa de transcrição do gene seja baixa e, para compensar isso a levedura precisa integrar um alto número de cópias da marca para poder crescer em meio sem suplementação. Um exemplo é a marca *leu2-d*, alelo truncado do gene *LEU2*. O alelo *leu2-d* possui apenas 29 pb do promotor original e seu uso como marca de seleção, em combinação com o uso de meio complexo, propiciou um aumento da produção de ácido artemisínico em 8,1 vezes em *S. cerevisiae* (Ro *et al.*, 2008). Esses resultados demonstraram que a marca defectiva permite a levedura manter o plasmídeo com alta estabilidade em meio não seletivo. No nosso laboratório foi desenvolvido um sistema genético baseado em uma marca auxotrófica defectiva, o *leu2-d*. Esse sistema pode ser aplicado para aumentar a produção de proteínas heterólogas em *K. phaffii*, permitindo a seleção de clones multicópias após a transformação (Betancur *et al.*, 2017).

O outro tipo de marca de seleção é a presença de genes que conferem resistência a drogas, por exemplo, *Sh ble*, *kanR* e *blasticidina S desaminase*, que conferem resistência a zeocina, G418 e blasticidina, respectivamente (Spohner *et al.*, 2015). As marcas de seleção dominantes de resistência a drogas geralmente codificam para proteínas que protegem a célula dos efeitos de um antibiótico, por exemplo, a fosfotransferase produzida pelo gene *kan* (Delneri *et al.*, 2000), ou para enzimas que permitam o metabolismo de um novo composto. Dessa forma, essas marcas aumentam a carga de proteínas a serem produzidas pela célula e isso pode sobrecarregar o seu metabolismo. A resistência a antibióticos também apresenta problemas no caso de leveduras naturalmente resistentes, que podem originar falsos positivos nas transformações (Kanda *et al.*, 2014). Como outro exemplo de marcas de seleção dominantes, no nosso laboratório foi desenvolvido um sistema para o uso da marca *amdS* (acetamidase), permitindo a reciclagem da marca através da contrasseleção com fluoroacetamida (Piva *et al.*, 2018)

1.1.4. Linhagens hospedeiras. Cepas mais utilizadas

A escolha de uma cepa específica de *K. phaffii* é determinada pela aplicação desejada. Há várias cepas desta levedura com uma grande variedade de genótipos que estão disponíveis atualmente. (Cereghino *et al.*, 2000; Logez *et al.*, 2012). Todas as cepas modificadas de *K. phaffii* são derivadas da linhagem nativa NRRL-Y 11430 (Northern Regional Research Laboratories) (Cereghino *et al.*, 2000).

Mutantes auxotróficos e linhagens deficientes em proteases encontram-se disponíveis para serem usadas como hospedeiras dos vetores de expressão. A tabela 1 mostra os genótipos e fenótipos das principais linhagens hospedeiras.

Tabela 1. Genótipos e fenótipos das principais linhagens de *Komagataella sp.*

Linhagem	Genótipo	Fenótipo	Referência
X-33	<i>Δhis4::HIS4</i>	Prototrófica	(Higgins <i>et al.</i> , 1998)
GS115	<i>his4</i>	His ⁻	(Cregg <i>et al.</i> , 1985)
KM71	<i>Δaox::SARG4 his4 arg4</i>	Mut ^S , His ⁻	(Tschopp <i>et al.</i> , 1987)
MC100-3	<i>Δaox::SARG4 Δaox2::Phis4 his4 arg4</i>	Mut ⁻ , His ⁻	(Cregg <i>et al.</i> , 1989)
SMD1168	<i>Δpep4::URA3 his4 ura3</i>	His ⁻ , deficiente em proteinase A	(Gleeson <i>et al.</i> , 1998)
SMD1165	<i>pbr1 his4</i>	His ⁻ , deficiente em proteinase B	(Gleeson <i>et al.</i> , 1998)
SMD1163	<i>pep4 pbr1 his4</i>	His ⁻ , deficiente em proteinases A e B	(Gleeson <i>et al.</i> , 1998)
JC220	<i>ade1</i>	Ade ⁻	(Cregg <i>et al.</i> , 1998)
JC254	<i>ura3</i>	Ura ⁻	(Cregg <i>et al.</i> , 1998)
GS190	<i>arg4</i>	Arg ⁻	(Cregg <i>et al.</i> , 1998)
M12	<i>leu2</i>	Leu ⁻	(Betancur <i>et al.</i> , 2017)

O uso de genes de vias biossintéticas como marcas de seleção requer que as linhagens hospedeiras sejam auxotróficas para o nutriente em questão, devido à ausência de uma cópia funcional desse gene no seu genoma. Desta forma, a menos que a levedura seja transformada com um alelo funcional do gene para recuperar a prototrofia, uma linhagem auxotrófica só consegue crescer em meio contendo o fator de crescimento apropriado (Cereghino *et al.*, 2000). As linhagens auxotróficas podem ser obtidas por mutações pontuais no gene biossintético ou pela deleção completa ou de uma parte desse gene (ruptura gênica). A ruptura gênica é mais estável, pois evita reversões que podem ocorrer no caso das mutações pontuais.

No nosso laboratório foi construída a linhagem M12-K derivada da linhagem M12 (Betancur *et al.*, 2017) que apresenta uma deleção no gene *KEX1*.

1.1.5. Secreção de proteínas

Os vetores de expressão podem conter também um sinal de secreção (peptídeo-sinal) que direciona a proteína heteróloga para a via secretória. A expressão de proteínas recombinantes em *K. phaffii* pode ser intracelular ou extracelular. A vantagem desta última é que essa levedura secreta poucas proteínas endógenas e seu meio de crescimento não contém proteínas adicionais. As proteínas exógenas podem representar até 80 % do total das proteínas extracelulares, o que pode atuar como primeiro passo no processo de purificação (Ahmad *et al.*, 2014; Larsen *et al.*, 2013; Wang *et al.*, 2016).

A sequência sinal mais utilizada é o fator α de acasalamento ou “*alpha mating factor*” (α -MF) de *S. cerevisiae* (Ahmad *et al.*, 2014; Cereghino *et al.*, 2000). Na via de secreção, o peptídeo sinal do α -MF é clivado por uma peptidase-sinal presente na membrana do retículo endoplasmático rugoso. Após a proteína ganhar o lúmen do retículo endoplasmático, a protease Kex2 da levedura reconhece o sítio de clivagem Lys-Arg e libera a porção pró do fator- α . A presença de sequência Glu-Ala próxima ao sítio de clivagem da Kex2 pode aumentar a atividade desta enzima, e ser posteriormente removida da proteína madura pela ação de uma diaminopeptidase, produto do gene *STE13* (Figura 2) (Ahmad *et al.*, 2014; Amorin, 2008; Cereghino *et al.*, 2000).

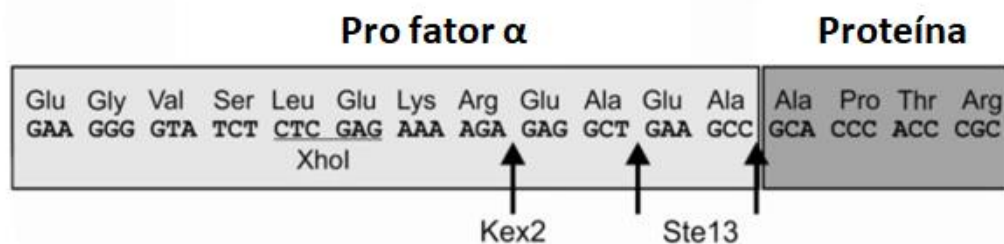


Figura 2. Sítios de clivagem no fator α . Sítios de reconhecimento das principais proteases envolvidas no processamento do peptídeo sinal

Apesar de ser o sinal de secreção mais utilizado e que geralmente apresenta os melhores resultados, há alguns casos em que o α -MF de *S. cerevisiae* não resolve o problema de interesse (Damasceno *et al.*, 2012). Então, outros sinais de secreção são utilizados como, por exemplo, o sinal de secreção da fosfatase ácida (*PHO1*) de *K. phaffii* ou da invertase (*SUC2*) de *S. cerevisiae* (Cereghino *et al.*, 2000; Liang *et al.*, 2013). Em último caso, quando os sinais de secreção disponíveis comercialmente falham, a sequência sinal nativa da proteína

heteróloga ou outra associada a altos níveis secretórios também podem ser utilizadas (Yu *et al.*, 2015).

As diferentes sequências sinais diferem amplamente na eficiência de secreção das proteínas recombinantes. Por isso, é necessário identificar diferentes sinais de secreção para encontrar a ótima para cada proteína (Massahi *et al.*, 2015). Os diferentes peptídeos sinais também podem ser utilizados em conjunto para comparar a expressão de uma determinada proteína a fim de escolher qual é o melhor sinal de secreção para obter os melhores resultados e assim ter o melhor sistema.

1.1.6. Outros aspectos de interesse

Existem vários aspectos que podem ser otimizados para a expressão de proteínas recombinantes em *K. phaffii*. A otimização de códons típicos de leveduras no gene de interesse resulta em um aumento muito grande nos níveis de expressão das proteínas. Também algumas sequências influem nos níveis de expressão, por exemplo, a sequência de iniciação da tradução, adaptação no conteúdo de GC e a adaptação no ponto isoelétrico da proteína desejada (Spohner *et al.*, 2015).

K. phaffii demonstrou ser uma eficiente plataforma para a produção de proteína recombinante em larga escala (Cereghino *et al.*, 2000), sendo que a fermentação pode ser facilmente aplicada para atender a demandas maiores, e parâmetros que influenciam a produtividade e a atividade da proteína, tais como pH, aeração e taxa de alimentação da fonte de carbono, podem ser controlados (Macauley-Patrick *et al.*, 2005). E por preferir um crescimento aeróbico, essa levedura não é considerada uma forte fermentadora como *S. cerevisiae*, assim, os produtos de fermentação como o etanol e o ácido acético não se acumulam rapidamente levando, portanto, mais tempo para atingir níveis tóxicos à célula (Celik *et al.*, 2012).

Estas características junto aos meios de crescimento de baixo custo fazem de *K. phaffii* um sistema de expressão altamente eficiente e com uma ótima relação custo-benefício.

1.2. *K. phaffii* e a produção de biofármacos

A produção de proteínas recombinantes tornou-se um negócio multimilionário, sendo que 25 % dos produtos farmacêuticos comercializados são biofármacos. O objetivo da

indústria farmacêutica é atingir processos de produção competitivos, ou seja, processos altamente produtivos, eficientes e econômicos (Weinacker *et al.*, 2013).

K. phaffii é uma plataforma ideal para a produção de proteínas de interesse biofarmacêutico, já que oferece as vantagens da eficiência da fácil manipulação das células procariotas (*E. coli*) e também a qualidade das células eucariotas (Yu *et al.*, 2015). Para a produção de biofármacos, a utilização de leveduras permite ter por um lado um crescimento robusto em meios de cultura simples (em biorreatores de grande escala) com modificações genéticas facilmente possíveis de realizar e com a introdução de modificações pós-traducionais desejáveis (Vogl *et al.*, 2013).

Uma das aplicações na indústria biofarmacêutica é a produção de subunidades recombinantes de vacinas (Wang *et al.*, 2016). Uma proteína de grande interesse na atualidade é a insulina. Este biofármaco é produzido como proteína recombinante utilizando dois processos diferentes. Um deles é a produção do precursor da insulina na forma de corpos de inclusão, usando *E. coli*, e o outro envolve a utilização do sistema de expressão em leveduras, em *K. phaffii*. O sistema de expressão com *K. phaffii* utiliza o α -MF como sinal de secreção, para obter o precursor da insulina solúvel no sobrenadante do cultivo. Usa a linhagem X-33 e o precursor da insulina com os códons otimizados para expressão em *K. phaffii* (Weinacker *et al.*, 2013).

K. phaffii também tem sido utilizada na produção de anticorpos terapêuticos, enzimas para aplicações biotecnológicas e citocinas (Bertheleme *et al.*, 2015). O primeiro polipeptídeo terapêutico expressado em *K. phaffii*, o ácido “ecallantide”, foi aprovado pela FDA no ano 2009 e é comercializado pela Dyax Corporation (Cambridge, MA) para o tratamento do angioedema hereditário (Bernstein *et al.*, 2012; Lunn *et al.*, 2011). O “ecallantide” é um inibidor da calicreína plasmática. A calicreína é uma glicoproteína, que faz parte da família das proteases serínicas presentes nos tecidos e nos fluidos corporais dos mamíferos. As calicreínas plasmáticas atuam ao nível da cascata da coagulação sanguínea e na fibrinólise. Apesar da sua importância as calicreínas apenas estão presentes nos órgãos e fluidos biológicos em baixas concentrações (Bos *et al.*, 2003; Raspi, 1996).

O segundo produto aprovado pela FDA foi a ocriplasmina, para o tratamento da adesão vitreomacular sintomática (VMA). A ocriplasmina é uma forma truncada da plasmina, uma serina-protease humana. Consiste só no domínio catalítico da plasmina. A ocriplasmina

tem várias vantagens sobre a plasmina: é consideravelmente mais estável, tem menor tamanho o que resulta em melhor penetração nos tecidos. É a primeira droga aprovada para o tratamento da VMA (Khan *et al.*, 2016; Stefanini *et al.*, 2014).

1.2.1. O fator de crescimento epidermal humano (hEGF) e sua produção em leveduras

1.2.1.1.O hEGF

Um dos compostos de grande interesse na indústria biofarmacêutica e cosmética é o fator de crescimento epidermal humano ou “*human epidermal growth factor*” (hEGF). O hEGF é um polipeptídeo de aproximadamente 6,0 kDa (6222,01 Da), formado por 53 resíduos de aminoácidos e três ligações dissulfeto (Figura 3) (Cohen *et al.*, 1975).

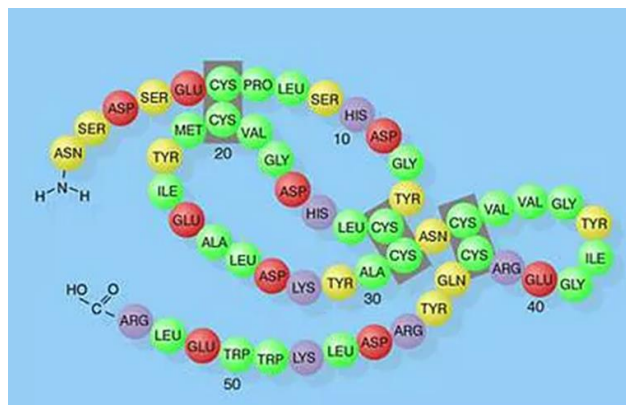


Figura 3. Sequência de aminoácidos do hEGF. Figura adaptada de (Cohen, 1987) com modificações.

O EGF inicialmente foi isolado das glândulas submandibulares de camundongos adultos e, quando aplicado em camundongos recém-nascidos, foram observadas algumas mudanças anatômicas como abertura precoce das pálpebras e erupção prematura dos dentes (Cohen, 1962). Posteriormente, foi detectado na urina humana um polipeptídeo com algumas das propriedades biológicas e imunológicas do EGF de camundongo, o qual foi identificado como fator de crescimento epidermal humano (hEGF) (Cohen *et al.*, 1975; Mu *et al.*, 2016).

O gene do hEGF está localizado no *locus* 4q25 (Gene/Locus MIM number 131530). Tem aproximadamente 110 kb e 24 éxons. Codifica para a síntese de uma grande pré-proteína de 1207 resíduos de aminoácidos que, após processamento proteolítico, resulta no polipeptídeo maduro com 53 resíduos de aminoácidos (Carpenter *et al.*, 1979).

O hEGF se liga ao seu receptor (EGFR). Este receptor é ativado quando seu ligante se liga ao domínio extracelular, propiciando assim a formação de dímeros com outro EGFR

ativado. O EGFR ativado é responsável por estimular inúmeras vias de sinalização intracelulares. Há duas vias principais de sinalização citoplasmática decorrentes da ativação do EGFR: RAS e PI3 quinase/Akt (Capdevila *et al.*, 2009).

A sinalização do EGFR tem impacto em muitos aspectos na biologia tumoral. Alterações no EGFR são frequentes em diversos tipos de câncer e podem promover a sobrevivência celular, proliferação, invasão, reparo e angiogênese (Hynes *et al.*, 2005). A partir dessa sinalização então o hEGF estimula a proliferação de uma grande variedade de células, incluindo precursores neurais *in vitro* (Kilpatrick *et al.*, 1995) e células tumorais. O sistema formado pelo EGF e seu receptor tem sido utilizado para o tratamento do câncer (Crombet Ramos *et al.*, 2015).

O hEGF pode inibir a secreção de ácido gástrico no estômago, aumentar a proliferação e queratinização dos tecidos epiteliais e acelerar a cicatrização de feridas (He *et al.*, 2016). Já foi demonstrado que o tratamento tópico de ratos com uma pomada de hEGF recombinante estimula a cicatrização de feridas devido ao incremento da taxa de proliferação epidermal e de miofibroblastos e à acumulação de colágeno na derme (Kwon *et al.*, 2006). Também foi demonstrado o efeito cicatrizante do hEGF em pacientes com úlcera do pé diabético (Pai-Dhungat *et al.*, 2015). Essas características fazem com que as aplicações terapêuticas deste fator de crescimento sejam amplas: cicatrização de feridas e queimaduras, cura da abrasão da superfície da córnea e tratamento de úlceras gástricas (He *et al.*, 2016; Tang *et al.*, 2016).

1.2.1.2. Produção do hEGF

Versões sintéticas do cDNA do hEGF têm sido expressas em células hospedeiras como *E. coli* (Oka *et al.*, 1985) e *S. cerevisiae* (Coppella *et al.*, 1989). Embora as leveduras sejam uma atraente plataforma para a produção de hEGF na sua forma solúvel, a ação de uma carboxipeptidase (Kex1) resulta na remoção da arginina C-terminal. Em *S. cerevisiae*, essa peptidase é codificada pelo gene *KEX1*. Em *K. phaffii* esse gene também já foi deletado para produzir endostatina humana e murina contendo a lisina C-terminal; porém, não foram feitos testes com o peptídeo completo e o peptídeo truncado para avaliar se realmente esse aminoácido é essencial para a atividade biológica da molécula (Boehm *et al.*, 1999).

Ainda, outro estudo demonstrou que as versões de hEGF com 52 ou 53 resíduos de aminoácidos possuem a mesma atividade e a mesma afinidade pelo receptor em fibroblastos (Hollenberg *et al.*, 1980). Contudo, para a utilização de hEGF na indústria biofarmacêutica, é

desejável a integridade da molécula e, portanto, a deleção do gene *KEX1* é uma estratégia promissória para diminuir o problema da degradação de proteínas com aminoácidos básicos na região C-terminal.

No nosso laboratório foi construída uma linhagem de *K. phaffii* com uma deleção no gene *KEX1* para ser usada como plataforma para a produção de hEGF. Testes iniciais de expressão mostraram que o hEGF havia conservado o resíduo de arginina C-terminal, o que pode ser atribuído à deleção de *KEX1*. Todavia a molécula apresentava resíduos Glu-Ala no N-terminal, o que foi atribuído a um processamento incompleto por parte da diaminopeptidase codificada pelo gene *STE13*, que atua no processamento do sinal de secreção α -MF. Esses resultados mostram que, para que se possa produzir hEGF em *K. phaffii*, além da deleção de *KEX1*, é necessário incluir um sinal de secreção que possa ser eficientemente processado, a fim de que o N-terminal correto da molécula seja obtido.

K. phaffii foi o sistema heterólogo escolhido para a expressão do hEGF devido a diversas vantagens dessa plataforma de produção (Cereghino *et al.*, 2000). Essa levedura apresenta crescimento rápido em meios de cultura simples e sua manipulação genética é fácil e rápida. Altos níveis de produção proteica já foram relatados para ela. Além disso, esse é um sistema eucariótico capaz de realizar modificações pós-traducionais e de secretar as proteínas para o meio de cultura (Higgins, 2001; Looser *et al.*, 2015). O correto processamento pós-traducional de proteínas é uma característica essencial para a produção de hEGF, já que a presença de ligações dissulfeto (o hEGF apresenta três ligações dissulfeto) é essencial para sua atividade biológica (Alewood *et al.*, 2005; Chen *et al.*, 2017). Quando comparado a outros sistemas eucarióticos de expressão, *K. phaffii* é o sistema mais barato e de escalonamento mais simples, o que o torna muito atrativo industrialmente.

Devido às suas potencialidades terapêuticas, o EGF já foi obtido de forma recombinante a partir da sua secreção para o meio de cultura na levedura *S. cerevisiae* (Brake *et al.*, 1984). Além disso, foi produzido no periplasma de *E. coli* (Abdull Razis *et al.*, 2008). Sua expressão extracelular também foi documentada em *Bacillus brevis* (Yamagata *et al.*, 1989) e em plantas de *Nicotiana benthamiana* (Thomas *et al.*, 2014).

Em *K. phaffii*, já foi obtido o EGF murino (Clare *et al.*, 1991) e porcino (Lee *et al.*, 2006), em ambos os casos utilizando o promotor do gene *AOX1*. Existem alguns estudos que apresentam a expressão do EGF humano em *K. phaffii* onde foi utilizado o promotor do gene

AOXI (Eissazadeh et al., 2017; Khan et al., 2017). O uso de indutores químicos, como o metanol, indutor do P_{AOXI} , apresenta limitações. Especificamente, altas doses podem ter efeitos tóxicos, podem ser bastante caros e, uma vez adicionados à cultura, é difícil reduzir a dose (o composto químico teria que ser retirado do meio). Em fermentações em grande escala essas questões devem ser levadas em consideração. Além disso, o metanol é inflamável. Na indústria biofarmacêutica não é desejado o uso do metanol na expressão de proteínas (Prielhofer et al., 2013; Vogl et al., 2013). Por isso, a expressão do EGF utilizando um promotor que não use metanol pode ser uma estratégia promissora.

Em *S. cerevisiae*, o gene *KEX1* já foi caracterizado e o produto desse gene, Kex1, está relacionado com o processamento de proteínas secretadas ao meio de cultura (Cooper et al., 1989). Kex1 é uma glicoproteína associada à membrana e tem atividade carboxipeptidásica com especificidade para aminoácidos básicos no C-terminal de proteínas e peptídeos. Em *K. phaffii*, o gene precursor da carboxipeptidase *KEX1* foi identificado e relatado (GenBank número de acesso AF095574). *Aspergillus oryzae* também apresenta um gene similar ao *KEX1* denominado *KEXA*, que apresenta atividade de carboxipeptidase do tipo serina (Morita et al., 2012). A comparação da estrutura primária da proteína Kex1 em *K. phaffii* e *S. cerevisiae* revelou apenas 36% de identidade e 43,7% de similaridade (Boehm et al., 1999). Mesmo assim, os resíduos envolvidos na tríade catalítica em *S. cerevisiae* (Ser198, Asp406 e His470) são conservados (Shilton et al., 1997). Já foi demonstrado que Kex1 é responsável pelo processamento incompleto da proteína glargina insulina, um análogo da insulina. Kex1 foi a responsável pela remoção da arginina C-terminal da cadeia β desta proteína (Sreenivas et al., 2016).

Com a ideia de produzir o fator de crescimento epidermal humano íntegro (53 aminoácidos), foi utilizada neste trabalho a linhagem M12-K construída no nosso laboratório, na qual foi deletado o gene *KEX1* para evitar a degradação de proteínas heterólogas pela proteinase codificada por esse gene. A linhagem M12-K é derivada da linhagem M12, que é auxotrófica para leucina, e é utilizada com vetores que apresentam a marca *leu2-d* (Betancur et al., 2017). As marcas auxotróficas defectivas consistem em genes de vias biossintéticas que possuem seu promotor truncado, o que faz com que a taxa de transcrição do gene seja baixa. Esse é o caso da marca *leu2-d* e, para compensar isto, a levedura tem que integrar um alto número de cópias da marca para poder crescer em meio mínimo (Betancur et al., 2017). Esse processo garante também integrar um elevado número de cópias do gene de interesse. Isto

representa uma maneira de obter clones que contenham múltiplas cópias de um determinado vetor de expressão.

2. OBJETIVO

2.1. Objetivo Geral

Desenvolver um sistema para a produção de hEGF na levedura *K. phaffii*.

2.2. Objetivos específicos

- Amplificar o gene hEGF contendo os peptídeos sinais dos genes *SUC2*, *PHO1* e αF .
- Clonar os fragmentos amplificados no vetor pKGFP-Id.
- Transformar a levedura *K. phaffii* e selecionar clones que expressem hEGF.
- Produzir o hEGF em frasco e purificar a proteína recombinante.
- Analisar o hEGF por espectrometria de massa.

3. ESTRATÉGIA

A estratégia utilizada para produção de hEGF em *K. phaffii* está descrita na figura 4.

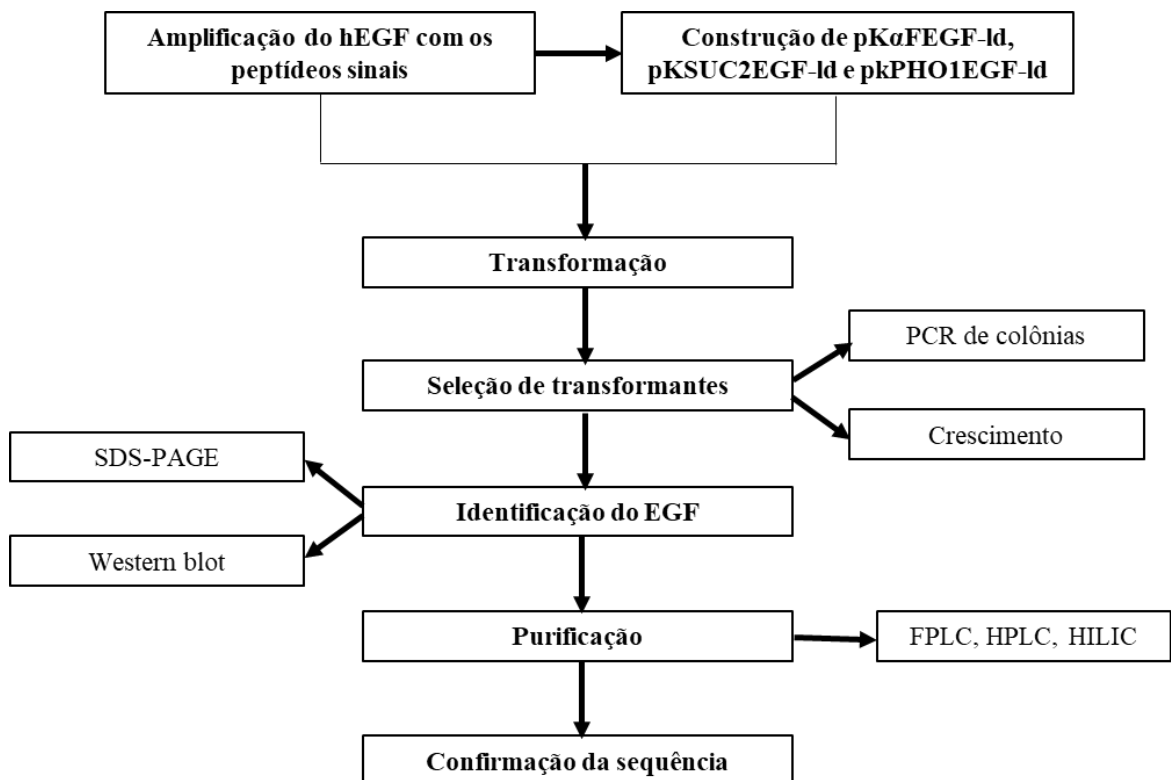


Figura 4. Estratégia de trabalho para a produção de hEGF. Estratégia para a produção de hEGF com os os peptídeos sinais αF , *SUC2* e *PHO1* em *K. phaffii*.

4. MATERIAIS E MÉTODOS

4.1. Materiais

4.1.1. Células

Báctérias

- *Escherichia coli* XL10-gold (Agilent Technologies, EUA): *Tet^r Δ(mcrA)183Δ(mcrCB-hsdSMR- mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F'proAB lacI^qZΔM15 Tn10 (Tet^r) Amy Cam^r]*

Leveduras

- *K. phaffii* X-33: *his4::HIS4* (derivada da linhagem GS115) (Invitrogen, EUA)
- *K. phaffii* M12: *leu2* (derivada da linhagem X-33) (Betancur *et al.*, 2017)
- *K. phaffii* M12-K: *leu2 kex1* (derivada da linhagem M12, construída no nosso laboratório)

4.1.2. Meios de cultura

Meio Luria-Bertani – LB

Extrato de levedura	0,5% (p/v)
Peptona	1% (p/v)
NaCl	1% (p/v)

O pH foi ajustado para 7,2. Para o meio sólido, foi adicionado ágar 2% (p/v)

SOB

Triptona	20 g/L
Extrato de levedura	5 g/L
NaCl	0,6 g/L
KCl	0,5 g/L
MgCl ₂	10 mM
MgSO ₄	10 mM

Foram feitas soluções de MgCl₂ e MgSO₄ na concentração de 1 M e filtradas. Depois do meio ser autoclavado foram adicionadas as soluções para a concentração final de 10 mM.

YPD

Extrato de levedura	1% (p/v)
Peptona	2% (p/v)
Glicose	2% (p/v)

Para o meio sólido, foi adicionado ágar 2% (p/v)

MD (Meio Mínimo com Dextrose)

YNB com sulfato de amônio	1,34% (p/v)
Biotina	4×10^{-5} % (p/v)
Glicose	2% (p/v)

Para o meio sólido, foi adicionado ágar 2% (p/v)

BMGY (Meio Complexo Tamponado com Glicerol)

Extrato de Levedura	1% (p/v)
Peptona	2% (p/v)
Fosfato de Potássio pH 6,0	100 mM
YNB com sulfato de amônio	1,34% (p/v)
Biotina	4×10^{-5} % (p/v)
Glicerol	1% (v/v)

Para os meios mínimos e os meios complexos tamponados, foram feitas soluções estoque de todos os reagentes e misturadas assepticamente no momento de preparar as placas. As soluções dos aminoácidos com uma concentração de 1% (p/v), a solução de glicose 10X, a solução YP 2X [extrato de levedura 2% (p/v) e peptona 4% (p/v)], a solução de fosfato de potássio 1 M pH 6,0 e a solução de ágar 2% (p/v) foram autoclavadas separadamente. A solução de biotina 500X e a solução YNB com sulfato de amônio 10X foram esterilizadas por filtração (utilizando um filtro de 0,22 μ m).

4.1.3. Soluções

Soluções estoque para meios de cultura

- Solução YNB (Yeast Nitrogen Base) com sulfato de amônio 10X

Yeast Nitrogen Base w/o ammonium sulfate (Difco)	3,4% (p/v)
Sulfato de Amônio (NH ₄) ₂ SO ₄	10% (p/v)

Esterilizada por filtração

- Solução de glicose 10X

Glicose	20% (p/v)
---------	-----------

- Solução de Biotina 500X

Biotina	0,02% (p/v)
---------	-------------

Esterilizada por filtração

Antibióticos

- Solução de ampicilina (1000X)

Ampicilina	100 mg/mL
------------	-----------

Dissolvida em água milli Q e esterilizada por filtração

- Solução de kanamicina (1000X)

Kanamicina	50 mg/mL
------------	----------

Dissolvida em água milli Q e esterilizada por filtração

- Solução de G418 (100X)

G418	50 mg/mL
------	----------

Dissolvida em água milli Q e esterilizada por filtração

Soluções para células bacterianas competentes

- Tampão de transformação I

RbCl	12 g/L
MnCl ₂ ·4H ₂ O	9,9 g/L
Acetato de potássio	0,03 M
CaCl ₂ ·2H ₂ O	1,5 g/L
Glicerol	150 g/L

Uma solução estoque de acetato de potássio 1 M foi preparada e seu pH ajustado para 7,5 com ácido acético glacial. A solução foi esterilizada por filtração. O pH do tampão de transformação foi ajustado para 5,8 com ácido acético 0,2 M e a solução foi esterilizada por filtração.

- Tampão de transformação II

MOPS	0,02 M
RbCl	1,2 g/L
CaCl ₂	11 g/L
Glicerol	150 g/L

Uma solução estoque de MOPS 1 M foi preparada e seu pH ajustado para 6,8 com NaOH. A solução foi esterilizada por filtração. O pH do tampão de transformação foi ajustado para 6,8 com NaOH e a solução foi esterilizada por filtração.

Soluções para preparação de DNA plasmidial (miniprep)

- Solução I

Tris-HCl pH 8,0	25 mM
EDTA pH 8,0	10 mM

- Solução II

NaOH	0,2 M
SDS	1% (p/v)

Solução preparada no momento de uso.

- Solução III

Acetato de sódio	3 M
Ácido acético	2 M

O pH foi ajustado para 5,0.

Soluções para eletroforese em gel de agarose

- Brometo de Etídio

EtBr	10 mg/mL
------	----------

- Tampão Tris-acetato EDTA (TAE) 50X

Tris	2 M
Ácido acético	1 M
EDTA pH 8,0	0,05 M

- Tampão de amostra

TAE 20X	50% (p/v)
Glicerol	30% (v/v)
Azul de bromofenol	0,25 (p/v)

Soluções para eletroforese em gel de poliacrilamida

- Acrilamida:bisacrilamida (29:1)

Acrilamida	29% (p/v)
Bisacrilamida	1% (p/v)

- Tris-HCl 1,5 M pH 8,8
- Tris-HCl 1 M pH 6,8
- SDS 10% (p/v)
- Persulfato de amônio 10% (p/v)

Armazenado a -20°C.

- Tampão de amostra desnaturante 5X

Tris-HCl pH 6,8	60 mM
SDS	2% (p/v)
β -Mercaptoetanol	14,4 mM
Glicerol	25% (p/v)
Azul de bromofenol	0,1% (p/v)

- Tampão de corrida Tris-Glicina 5X

Tris	16,7 g/L
Glicina	104,5 g/L
SDS	0,5% (p/v)

Revelação de proteínas por coloração do gel com Azul de Coomassie

- Solução corante com Coomassie

Coomassie brilliant blue G-250	0,25% (p/v)
Metanol	30% (v/v)
Ácido acético glacial	7% (p/v)
• Solução descorante com Coomassie	
Metanol	30% (v/v)
Ácido acético glacial	7% (p/v)

Revelação das proteínas por coloração do gel com Nitrato de Prata

• Solução de fixação	
Acetona	50% (v/v)
TCA	50% (v/v)
Formaldeído	37% (v/v)
• Solução de pré-tratamento I	
Acetona	50% (v/v)
• Solução de pré-tratamento II	
Tiosulfato de sódio	10% (v/v)
• Solução de impregnação	
Nitrato de prata	20% (v/v)
Formaldeído	37% (v/v)
• Solução de revelação	
Carbonato de sódio	2% (p/v)
Tiosulfato de sódio	10% (v/v)
Formaldeído	37% (v/v)
• Solução de parada	
Ácido acético	1% (v/v)

Todas as soluções foram feitas em água destilada em um volume final de 60 mL.

Soluções para transferência e Western blot

- Solução salina de Tris com Tween 20 (TBST)

Tris pH 7,5	20 mM
NaCl	150 mM
Tween 20	0,1% (v/v)
• Solução de transferência	
Tris pH 7,5	25 mM
Glicina	190 mM
Metanol	20% (v/v)
• Solução de Bloqueio	
BSA	3% (p/v)

A solução de bloqueio foi diluída em TBST

4.1.4. Kits

- Qiagen Plasmid Maxi kit (Qiagen): purificação de plasmídeos em grande escala.
- In-Fusion HD Cloning Kit (Clontech): clonagem de fragmentos por recombinação homóloga *in vitro*.
- Wizard SV Gel and PCR Clean-Up System (Promega): purificação de fragmentos de DNA a partir de gel de agarose e purificação de produtos de PCR.
- Optiblot ECL Detect Kit (23 pg-187 ng) ab133406 (abcam): detecção para análise por Western blot.

4.1.5. Oligonucleotídeos (*primers*)

A tabela 2 apresenta os *primers* utilizados neste trabalho com seus respectivos sítios de restrição sublinhados.

Tabela 2. *Primers* utilizados. Os sítios de restrição estão sublinhados

<i>Primer</i>	Sequência (5' - 3')	Sítio de Restrição
SUC2EGF-F	AATTGATTACGAAAGGATCCAAACGATGCTTTTG CAAGCCTTCCTTTTCCCTTTTGGCTGGTTTTGCAGC TAAGATATCTGCAAACCTCCGACTCCGAATGTCC	BamHI
PHO1EGF-F	AATTGATTACGAAAGGATCCAAACGATGTTCTCT CCAATTTTGTCTTGGAAATTATTTTAGCTTTGGC TACTTTGCAATCTGTCTTCGCTAACTCCGACTCCG AATGTCC	BamHI

αFEGF-F	GAGGAGGGTGTCTCTCTCGAGAAAAGAACTCC GACTCCGAATGTCC	XhoI
EGFAOX1-R	TTATCTCAATTCCCACCACTGCGGCCGCGTCGAC CAT	NotI

4.1.6. Marcadores de massa molecular para DNA

- 2-log DNA ladder, 100-10.000 pb (New England Biolabs).
- O' GeneRuler 1 kb Plus DNA Ladder 75-20.000 pb (Thermo Scientific).

4.1.7. Marcadores de massa molecular para proteínas

- SeeBlue Pre-Stained Protein Standard, 3-198 kDa (Life Technologies)

4.1.8. Enzimas de restrição

A tabela 3 apresenta as enzimas de restrição usadas neste trabalho, indicando o sítio de clivagem, assim como a temperatura de incubação e o tampão recomendado para cada enzima.

Tabela 3. Enzimas de restrição utilizadas

Enzima	Sítio de clivagem	Tampão	Temperatura de incubação
BamHI	G ⁺ GATCC	NEBuffer 3.1	37 °C
NotI	GC ⁺ GGCCGC	NEBuffer 3.1	37 °C
XhoI	C ⁺ TCGAG	NEBuffer 3.1	37 °C
BglII	A ⁺ GATCT	NEBuffer 3.1	37 °C
SacI	GAGCT ⁺ C	NEBuffer 1.1	37 °C

4.1.9. Vetores

A tabela 4 exhibe todos os vetores usados neste trabalho e apresenta uma breve descrição de cada um deles.

Tabela 4. Vetores utilizados

Vetor	Descrição
pKGFp-Id	Vetor construído em nosso laboratório contendo o gene <i>LEU2</i> como marca de seleção (Betancur <i>et al.</i> , 2017).
pk-LeudEGF	Vetor para amplificação do hEGF
pKαFChym-Id	Vetor contendo o α F como peptídeo sinal (construído neste trabalho)

pKαFEGF-Id	Vetor para a expressão de hEGF contendo o peptídeo sinal α F (construído neste trabalho)
pKSUC2EGF-Id	Vetor para a expressão de hEGF contendo o peptídeo sinal <i>SUC2</i> (construído neste trabalho)
pKPHO1EGF-Id	Vetor para a expressão de hEGF contendo o peptídeo sinal <i>PHO1</i> (construído neste trabalho)

4.2. Métodos

4.2.1. Cultivo de micro-organismos

As bactérias foram cultivadas a 37 °C e as leveduras a 28 °C. Para ambos casos, uma velocidade de agitação de 250 rpm foi usada quando o crescimento foi em meio líquido.

4.2.2. Amplificação por PCR

Os *primers* utilizados nesse trabalho estão listados na Tabela 2. As amplificações foram realizadas utilizando-se Phusion High-Fidelity DNA Polymerase (Thermo Scientific, EUA), de acordo com as instruções do fabricante.

4.2.3. Preparação de células bacterianas competentes

Células de *E. coli* XL10-gold foram semeadas em meio LB ágar a partir de um estoque armazenado a -80°C e incubadas a 37°C durante a noite. Uma colônia isolada foi inoculada em 10 mL de meio SOB em um frasco Erlenmeyer de 125 mL e incubada a 37°C sob agitação a 250 rpm por 16 h. Dois mililitros desse pré-inóculo foram inoculados em 100 mL de meio SOB em um frasco Erlenmeyer de 500 mL e incubados a 37°C sob agitação a 250 rpm por aproximadamente 1 h, até atingir uma OD₆₀₀ de 0,3. A cultura foi resfriada por 15 min em banho de água-gelo e, posteriormente, as células foram coletadas a 3.000 x g por 5 min a 4 °C. O precipitado foi ressuspensionado em 32 mL de tampão de transformação I e incubado no gelo por 15 min. Em seguida, as células foram concentradas por centrifugação sob as mesmas condições e ressuspensionadas em 4 mL de tampão de transformação II. Finalmente, foram feitas alíquotas de 100 μ L em tubos de 1,5 mL e armazenadas a -80°C.

4.2.4. Manipulação do DNA

As digestões de DNA foram realizadas segundo as instruções do fabricante das enzimas. A quantidade de enzima foi ajustada à quantidade de DNA a ser digerido considerando que o volume da enzima não devia ser maior do que 10% do volume total de reação. A tabela 3 apresenta os tampões e temperaturas recomendadas para cada enzima. Todas as reações foram incubadas por 2 horas.

Para retirar sais e enzimas após a digestão, o DNA foi precipitado adicionando acetato de sódio para uma concentração final de 0,3 M e 2,5 volumes de etanol 100% gelado. Após incubação a -20 °C por 16 h, o sistema foi coletado por centrifugação a 10.000 x g por 10 min. O sobrenadante foi descartado e o precipitado lavado duas vezes com 500 µL de etanol 70% por centrifugação a 10.000 x g por 2 min. O sobrenadante foi descartado e, uma vez seco, o precipitado foi ressuspenso em 10 µL de água milli Q.

A análise de fragmentos de DNA foi feita utilizando a técnica de eletroforese em gel de agarose adaptada de Sambrook (1989) (Sambrook *et al.*, 1989). O gel preparado na concentração de 1% (p/v) de agarose em tampão TAE 1X continha brometo de etídeo na concentração final de 0,5 µg/mL. As amostras foram aplicadas no gel e submetidas à corrente elétrica (70 V) para permitir a migração e separação dos fragmentos de DNA. A visualização das bandas de DNA foi feita com a exposição do gel à luz ultravioleta.

Os fragmentos de DNA eluídos de gel de agarose foram purificados com o kit Wizard SV Gel and PCR Clean-Up System (Promega, EUA) seguindo as orientações especificadas pelo fabricante.

4.2.5. Construção de plasmídeos

A ligação do fragmento pKαFChym-IId com o vetor pKGFP-IId digerido com as enzimas de restrição BglII e NotI foi feita utilizando a enzima T4 DNA ligase (usb) com o tampão fornecido pelo fabricante. O sistema foi incubado a 16°C durante 16 h.

O sistema de ligação foi utilizado para transformar células bacterianas competentes pelo método de choque térmico. Uma alíquota de células competentes armazenada a -80 °C foi descongelada no gelo (de 2-5 minutos) e em seguida foram adicionados 10 µL do sistema de ligação. As células foram incubadas no gelo por 30 min e posteriormente foram submetidas a choque térmico a 42 °C por 90 s. Foram adicionados 900 µL de meio LB ao sistema, que foi incubado a 37 °C por 1 h. As células foram semeadas em meio LB ágar contendo o antibiótico adequado.

A construção dos plasmídeos contendo sequências dos diferentes peptídeos sinais (αF, *SUC2* e *PHO1*) e o EGF foi feita utilizando o kit In-Fusion® HD Cloning Kit (Clontech, EUA), seguindo as indicações do fabricante.

4.2.6. Preparação de DNA plasmidial

Para extrair os plasmídeos das células bacterianas em pequena escala (miniprep) foi utilizado o protocolo descrito por (Sambrook *et al.*, 1989) com adaptações. Cinco mililitros de meio LB contendo o antibiótico adequado foram inoculados com uma colônia transformante e incubados a 37 °C por 16-18 h sob agitação. Três mililitros da cultura foram coletados por centrifugação a 10.000 x g por 2 min e o precipitado de células foi ressuscitado em 200 µL de solução I. Em seguida, foram adicionados 360 µL de solução II recém-preparada; a mistura foi homogeneizada por inversão e incubada à temperatura ambiente por 5 min. Posteriormente, foram adicionados 300 µL de solução III gelada, a mistura homogeneizada e incubada no gelo por 5 min. Após esse tempo a amostra foi concentrada por centrifugação a 10.000 x g por 5 min. O sobrenadante foi transferido para outro tubo e foram adicionados 750 µL de isopropanol. A mistura foi homogeneizada por inversão e submetida a centrifugação a 10.000 x g por 5 min. O sobrenadante foi descartado e o precipitado ressuscitado em 200 µL de solução I. Em seguida, foram adicionados 110 µL de acetato de amônio 7,5 M e a mistura homogeneizada vigorosamente seguido de centrifugação a 10.000 x g por 10 min. O sobrenadante foi transferido para outro tubo e foram adicionados 750 µL de etanol 100% gelado. Após centrifugação a 10.000 x g por 5 min, o sobrenadante foi descartado e o precipitado lavado com 500 µL de etanol 70% seguido de uma nova centrifugação a 10.000 x g por 2 min. O sobrenadante foi descartado e o excesso de etanol foi seco à temperatura ambiente. O precipitado foi ressuscitado em 50 µL de água milli Q contendo RNase A (0,1 mg/mL) e incubado a 37 °C por 15 min. O DNA foi armazenado a -20 °C até seu uso.

Para a obtenção dos plasmídeos em maior quantidade foi utilizado o kit Qiagen Plasmid Maxi (Qiagen), seguindo as orientações do fabricante.

4.2.7. Transformação de leveduras por eletroporação

A transformação com vetores integrativos foi feita por eletroporação segundo o protocolo descrito no EasySelect Pichia Expression Kit (Invitrogen, EUA) com algumas modificações. Uma colônia fresca de levedura foi inoculada em 5 mL de YPD. Após 12 h de crescimento a 28 °C sob agitação de 250 rpm, 500 µL desse pré-inóculo foram usados para inocular mais 5 mL de meio. Após 8 h de incubação, o volume necessário dessa nova cultura foi inoculado em 100 mL de meio YPD para atingir uma OD₆₀₀ de 1,3 após 16 h de crescimento. As células foram coletadas e lavadas 3 vezes com água estéril gelada (100 mL na primeira lavagem e 50 mL nas outras duas) e uma vez com sorbitol 1 M gelado (3 mL) centrifugando-se a 1.500 x g por 5 min a 4 °C. Posteriormente, as células foram

ressuspendidas em 300 μ L de sorbitol 1 M gelado e 80 μ L dessas células competentes foram misturadas com 5-10 μ g de DNA linearizado e transferidas a uma cubeta de eletroporação de 2 mm (Bio-Rad). A cubeta foi incubada no gelo por 5 min antes de aplicar um pulso elétrico sob as seguintes condições: 1500 V, 400 Ω , e 25 μ F. Imediatamente após o pulso, 1 mL de sorbitol 1 M gelado foi adicionado. As células foram semeadas em meio mínimo para seleção de transformantes.

4.2.8. PCR de colônia

A PCR de colônia é realizada a partir de colônias isoladas em cada transformação e tem por objetivo detectar quais são os clones positivos (que inseriram o fragmento em estudo). Com a ponta da pipeta ou com um palito, a colônia é riscada o mais precisamente possível e diluída em 5 μ L de água. O controle negativo contém a reação de PCR sem DNA. A reação de PCR de colônia contém os “Primers Forward” e “Reverse”, DNA polimerase (Phusion), Tampão HF 1X, dNTPs a 2 mM e água, perfazendo o volume de 15 μ l. Os resultados são analisados em gel de agarose a 1%.

4.2.9. Avaliação do crescimento em meio líquido

Uma colônia nova da levedura foi inoculada em 500 μ L de MD em placa *deep-well* e incubada a 30 °C e 250 rpm por 24 h. Um volume de 100 μ L de MD foi inoculado com o volume necessário do pré-inóculo para começar a cultura com uma OD₆₀₀ de 0,1 em placa de 96 poços. A análise de crescimento foi realizada com o espectrofotômetro Eon High Performance Microplate Spectrophotometer (Biotek, EUA) incubando a 30 °C sob agitação de 300 rpm por 72 h, coletando dados de OD₆₀₀ a cada 30 minutos.

Para plotar as curvas de crescimento foram usados os valores de logaritmo natural da OD₆₀₀ em função do tempo. A velocidade máxima de crescimento foi calculada como a inclinação da parte linear dessas curvas que corresponde à fase exponencial de crescimento.

4.2.10. Crescimento e expressão de proteínas recombinantes em frasco

Uma colônia da levedura foi inoculada em 5 mL de meio BMGY e incubada a 28 °C por 16 h. O volume apropriado dessa cultura foi inoculado em 100 mL de meio BMGY para começar a expressão com uma OD₆₀₀ de 0,05. A nova cultura foi incubada a 28 °C durante 24 h. Posteriormente as amostras foram centrifugadas a 3000 x g por 5 min e o sobrenadante foi recuperado.

4.2.11. Concentração do sobrenadante

Uma vez obtido o sobrenadante dos clones crescidos em meio complexo, este foi concentrado para os experimentos de purificação da proteína de interesse. O sobrenadante foi concentrado utilizando o sistema Amicon ultrafiltration setup com uma membrada de 1 kDa. As amostras foram concentradas entre 9 e 12 vezes.

4.2.12. Precipitação das proteínas com TCA

As proteínas presentes no sobrenadante foram precipitadas com TCA, para isso 250 μ L de TCA 100% foram adicionados a 1 mL do sobrenadante e a mistura foi incubada a -20 °C durante a noite. Depois foi centrifugada a 14000 x g por 15 min. O precipitado foi lavado duas vezes com acetona 100 %. Finalmente, o precipitado foi ressuspenso em 25 μ L de tampão de amostra 1X.

4.2.13. Quantificação de proteínas pelo Método de Bradford

A determinação da concentração de proteínas presentes nas amostras foi feita pelo Método de Bradford (Bradford, 1976). Foram adicionados 10 μ L de amostra (sobrenadante concentrado) ou água (branco) e 200 μ L do reagente de Bradford numa placa de 96 poços. As reações foram incubadas a temperatura ambiente por 5 minutos e foi medida a OD_{595nm}. Cada amostra foi aplicada por duplicata. Para o cálculo das concentrações foram interpolados os resultados obtidos com uma curva padrão previamente realizada.

4.2.14. Análise eletroforética em gel de poliacrilamida desnaturante

Para separação eletroforética de proteínas foi utilizado o sistema SDS-PAGE na concentração de 15%. Foi utilizado um sistema de cubas verticais de eletroforese (BioRad, EUA). A corrida eletroforética foi conduzida em tampão 1X com uma voltagem inicial de 60 V e depois de 100 V. As amostras a serem analisadas foram diluídas com tampão de amostra 5X.

Para visualização das proteínas após eletroforese, os géis foram corados com Coomassie Brilliant Blue ou com Nitrato de Prata, dependendo do caso. Para o caso da coloração com Coomassie, o gel foi incubado na solução corante por 16 h sob agitação, seguido de incubação na solução descorante, sob agitação, até o aparecimento das bandas.

Para o caso da coloração com Nitrato de Prata foi seguido o protocolo descrito por Nesterenko (Nesterenko *et al.*, 1994).

A estimativa de massa molecular foi feita por comparação com marcador de massa molecular (Life Technologies) e com o EGF padrão (Sigma) aplicados no gel.

4.2.15. Western Blot

Para a conformação da presença de EGF foi feito um ensaio de Western Blot, segundo o protocolo da BioRad, EUA (General Protocol for Western Blotting), com algumas modificações.

4.2.15.1. Transferência para a membrana

Uma vez concluída a corrida eletroforética, o gel foi incubado em solução de transferência por 15 minutos. Para a transferência, foi utilizada uma membrana de PVDF. A membrana foi ativada 40 segundos em metanol, seguido de 5 minutos em água e 10 minutos em solução de transferência. A transferência foi feita a 15 V, 150 mA, durante 30 minutos. Foi utilizado o Trans-Blot® SD Semi-Dry Transfer Cell (BioRad).

4.2.15.2. Bloqueio da membrana

Depois da transferência, a membrana foi incubada 1 h em solução de bloqueio a temperatura ambiente.

4.2.15.3. Anticorpos e revelação

A membrana foi incubada com o anticorpo primário (anti-EGF antibody ab9695, abcam) durante a noite a 4°C. Após a incubação com o anticorpo primário, foram feitas três lavagens com TBST durante 5 minutos cada uma. A membrana foi incubada 1 h com o anticorpo secundário (Goat anti-rabbit IgG H&L HRP ab205718, abcam). Depois, a membrana foi lavada três vezes com TBST. Os anticorpos foram diluídos em TBST. A revelação foi feita com o Optiblot ECL Detect Kit (23pg-187ng) ab133406 (abcam, Reino Unido), conforme as especificações do fabricante. A detecção por quimioluminescência foi feita no equipamento Amersham Imager 600 (GE Life Sciences, EUA).

4.2.16. Purificação do hEGF

O processo de purificação final do EGF foi feito em três etapas. Uma primeira onde foi feita uma Cromatografia de Exclusão Molecular, a segunda em HPLC-fase reversa e uma terceira em HILIC.

4.2.16.1. Cromatografia de Exclusão Molecular (FPLC)

A cromatografia de exclusão molecular foi feita utilizando-se a coluna SuperdexTM 30 PG (AKTA Pure, GE Healthcare). O sistema e a coluna foram previamente equilibrados com o tampão fosfato de potássio 50 mM, pH 6 a um fluxo de 1 mL/min. O tampão foi filtrado e degaseificado antes de ser aplicado no sistema e na coluna. Em cada corrida cromatográfica foram injetados 1 mL de sobrenadante de cultura concentrado em filtro Amicon de 1 kDa e filtrado em membrana de 0,22 µm. A eluição foi realizada a um fluxo contínuo de 1 mL/min. A presença de proteínas foi monitorada pela absorvância a 216 nm e 280 nm. Frações de 1 mL foram coletadas automaticamente. Este procedimento repetiu-se várias vezes. As frações de interesse foram secadas em sistema de SpeedVac e armazenadas a -20 °C para o próximo passo de purificação. Antes, 500 µL de cada fração foram separados e precipitados com TCA para análise por gel de poliacrilamida desnaturante e por Western blot.

4.2.16.2. Purificação da fração do EGF/FPLC por HPLC-fase reversa

As frações da cromatografia de exclusão molecular em que a proteína desejada foi visualizada foram ressuspensas em água milli Q e utilizadas em cromatografia em fase reversa. Das frações obtidas, 200 µl de amostra foram injetados no sistema em fase reversa (Shimadzu Co. Kioto, Japan série LC 20A), em coluna analítica C₁₈ (Kinetex Phenomenex 300Å, 250 x 4.6 mm, EUA). As fases móveis utilizadas foram TFA 0.12% (v/v) em água deionizada (solução A) e TFA 0.12% (v/v) em acetonitrila (solução B). A eluição foi feita em um gradiente linear de 0-50 % de B em 50 minutos. A detecção foi feita a 216 nm e a fração correspondente ao EGF (36,5 minutos de retenção para a primeira purificação e 32,5 minutos de tempo de retenção para a segunda) foi coletada manualmente em microtubos de 1,5ml, secas em sistema de SpeedVac e armazenadas a -20 °C para posterior análise.

4.2.16.3. Purificação da fração do EGF/HPLC por HILIC

A fração correspondente ao hEFG purificado cromatografia por HPLC-fase reversa, foi submetido a uma recromatografia em fase normal com coluna HILIC (Phenomenex 300Å, 150 x 4.6 mm). As fases móveis utilizadas foram TFA 0.12% (v/v) em água deionizada 10% (solução A) e TFA 0.12% (v/v) em água deionizada (solução B). A eluição foi feita em um

gradiente linear de 0-20% de B em 25 minutos. A detecção foi feita a 216 nm e a fração correspondente ao EGF (18 minutos de tempo de retenção) foi coletada manualmente em tubos tipo eppendorf de 1,5ml, seca em sistema de SpeedVac e armazenadas a -20 °C para posterior análise.

4.2.17. Espectrometria de massa MALDI-TOF/TOF

As frações cromatográficas de interesse foram analisadas em espectrômetro de massa do tipo MALDI-TOF/TOF (*matrix-assisted laser desorption/ionization -time of flight/time of flight*) modelo Autoflex II (Bruker Daltonics, Alemanha). As frações foram ressuspendidas em acetonitrila 50% contendo TFA 0,1% e 0,5 µL de cada fração cromatográfica aplicada em placa de aço inoxidável juntamente com matriz ácido sinapínico (SA 20 µg/µL) ressuspendida em acetonitrila 50% contendo 0,1% de TFA (v/v). O procedimento foi realizado no modo linear positivo, sendo a faixa m/z analisada a partir de 4000 a 20.000 Da.

4.2.18. Degradação de Edman e sequenciamento automático

Após a determinação da massa molecular, a proteína foi sequenciada automaticamente pelo método de degradação de Edman. A determinação da estrutura primária por meio do sequenciamento automático do N-terminal de peptídeos foi realizada em um sequenciador de aminoácidos PPSQ-23 (Shimadzu, Japão) conforme instruções do fabricante.

4.2.19. Análise de dados

As análises estatísticas e os gráficos foram feitos com o software GraphPad Prism 5. Foi aplicado ANOVA seguido do teste de Tukey para comparação entre as amostras. As sequências foram analisadas e alinhadas utilizando o programa CLUSTAL OMEGA (CLUSTAL OMEGA 1.2.4 *multiple sequence alignment*).

5. RESULTADOS

5.1. Construção dos cassetes de expressão para hEGF contendo diferentes peptídeos sinais

Com o objetivo de produzir o hEGF mantendo sua integridade foram testados três peptídeos sinais diferentes que são amplamente utilizados para expressão heteróloga em *K. phaffii*. A fim de expressar o cDNA codificador de hEGF em *K. phaffii*, três construções gênicas foram desenhadas. Foram utilizados os peptídeos sinais: o fator α de *S. cerevisiae*, do qual o final do fator foi otimizado para *K. phaffii* (α F) (foram eliminadas as sequências GLU-ALA no final do sinal de secreção e feita otimização de códon para *K. phaffii*), o sinal de secreção da invertase de *S. cerevisiae* (*SUC2*) e o sinal de secreção da fosfatase ácida (*PHO1*) de *K. phaffii*. O cDNA do hEGF foi amplificado por PCR a partir do vetor pK-LeudEGF (Anexo 1), utilizando os *primers* α FEGF-F e EGFAOX1-R; SUC2EGF-F e EGFAOX1-R; PHO1EGF-F e EGFAOX1-R, respectivamente (Tabela 2).

Na figura 5 pode-se observar, para cada um dos sistemas, uma banda do tamanho correspondente ao fragmento do cDNA do hEGF com o respectivo peptídeo sinal (258 pb para o *SUC2*, 203 pb para o α F, e 267 pb para o *PHO1*). Os fragmentos foram obtidos com pares de *primers* que contêm os peptídeos sinais e que contêm 15 pb homólogas ao vetor de expressão para realizar a clonagem por recombinação homóloga.

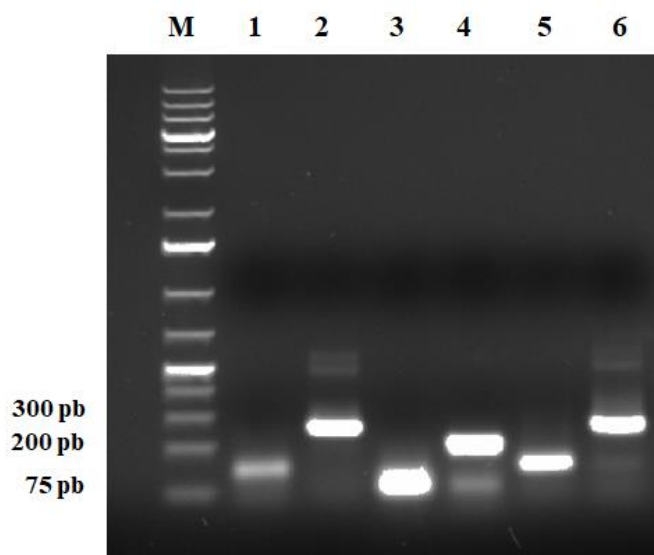


Figura 5. Amplificação por PCR dos cassetes α FEGF, SUC2EGF e PHO1EGF. Eletroforese em gel de agarose 1% corado com brometo de etídeo. M: O'GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). 1: SUC2EGF (controle negativo), 2: SUC2EGF (258 pb), 3: α FEGF (controle negativo), 4: α FEGF (203 pb), 5: PHO1EGF (controle negativo), 6: PHO1EGF (267 pb)

Os produtos de PCR foram purificados.

5.2. Construção dos vetores pK α FEGF-Id, pKSUC2EGF-Id e pKPHO1EGF-Id

Depois de purificados, os fragmentos foram clonados por recombinação homóloga no vetor também purificado. Para os sistemas com os peptídeos sinais *SUC2* e *PHO1* foi utilizado como molde o vetor pKGFP-Id. Este vetor foi digerido com as enzimas BamHI e NotI e clonado com os fragmentos SUC2EGF ou PHO1EGF. O vetor pKGFP-Id (Betancur *et al.*, 2017) empregado neste trabalho utiliza uma versão de 618 pb do *P_{PGK1}* de *K. Phaffii* que é capaz de controlar a expressão de proteínas heterólogas e de promover a integração do vetor no *locus PGK1* no genoma da levedura (Arruda *et al.*, 2016; de Almeida *et al.*, 2005).

Com o objetivo de construir o vetor com o α F, foi construído um vetor intermediário contendo o α -MF fusionado com o gene da quimosina bovina construído em nosso laboratório. Este vetor intermediário (pK α FChym-Id) foi obtido da digestão do vetor pKGFP-Id com as enzimas de restrição BglIII e NotI e a posterior ligação deste fragmento purificado pK α FChym. Este sistema de ligação foi utilizado para transformar *E. coli* XL-10 Gold. O vetor resultante pK α FChym-Id foi digerido com as enzimas XhoI e NotI e clonado com o fragmento purificado α FEGF. Os vetores utilizados estão esquematizados na figura 6.

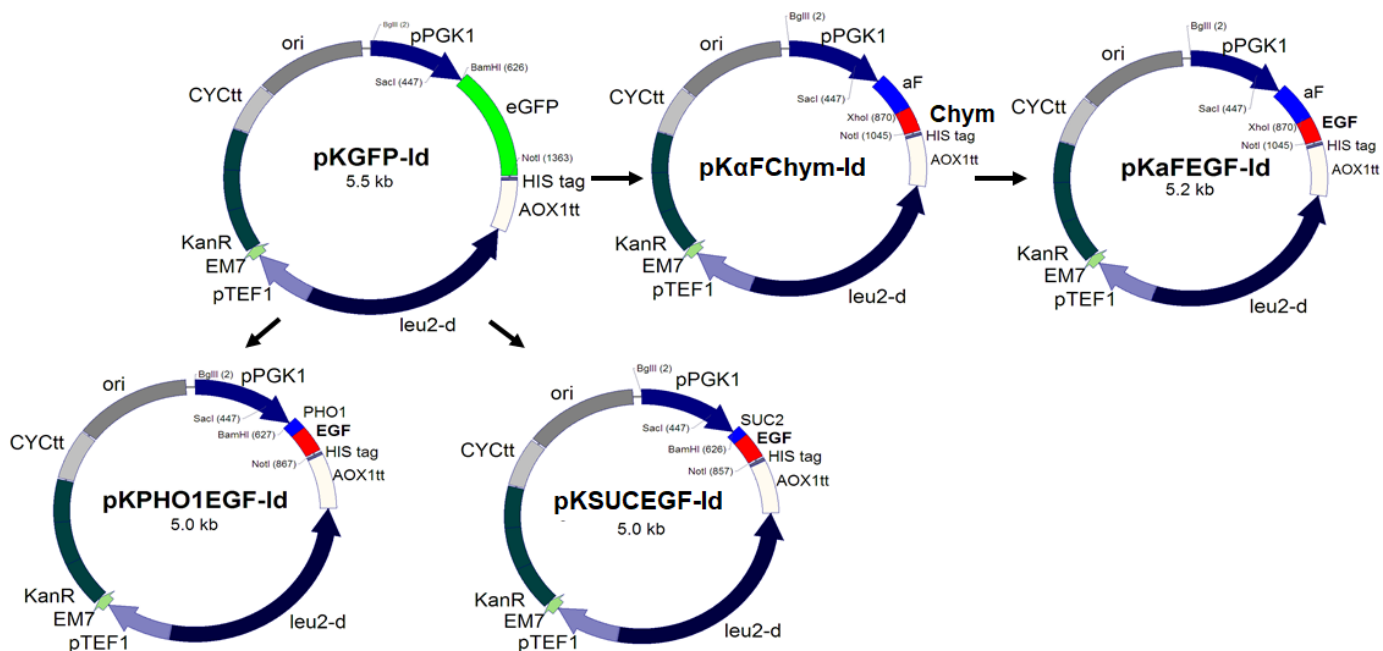


Figura 6. Mapas físicos dos vetores de expressão utilizados para a construção dos plasmídeos. São mostradas as enzimas de restrição e os sítios utilizados para as digestões.

Os produtos da clonagem foram utilizados para transformar *E. coli* XL-10 Gold. Após a seleção, foi feita a extração plasmidial de 6 clones transformantes de cada sistema e o DNA foi digerido com BglIII e NotI (para os sistemas SUC2EGF e PHO1EGF) ou com XhoI e NotI (sistema α FEGF), para confirmar a inserção do EGF com os respectivos peptídeos sinais no vetor de expressão. Se o fragmento tivesse sido clonado, a digestão com as respectivas enzimas devia liberar um fragmento de 855 pb para o sistema SUC2EGF, 865 pb para o sistema PHO1EGF e 175 pb para o sistema α FEGF. A figura 7 mostra o resultado da digestão para os clones selecionados.

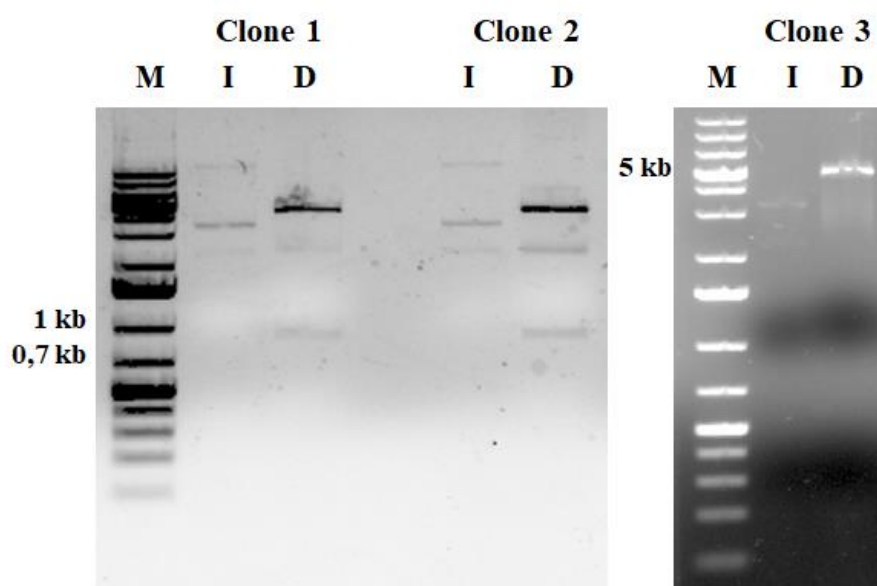


Figura 7. Análise da clonagem, perfil de restrição dos vetores construídos. Eletroforese em gel de agarose 1% corado com brometo de etídeo apresentando o DNA plasmidial intacto (I) e digerido com BglIII e NotI (D) para 2 clones transformantes (clone1: pKEGFPHO1EGF-ld e clone 2: pKEGF SUC2EGF-ld) ou digerido com XhoI e NotI (D) para um clone transformante (clone 3: pK α FEGF-ld). M: O'GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific, EUA).

Como esperado, o fragmento menor (175 pb) não foi observado devido à sua baixa concentração. Estes clones foram selecionados para fazer extração plasmidial em grande escala (midiprep), a fim de obter quantidade maior de DNA. Os vetores obtidos foram chamados pK α FEGF-ld, pKSUC2EGF-ld e pKPHO1EGF-ld, respectivamente.

5.3. Transformação de *K. phaffii* M12-K

Como apresentado na figura 8, os vetores pK α FEGF-ld, pKSUC2EGF-ld e pKPHO1EGF-ld foram linearizados com a enzima SacI, previamente à transformação.

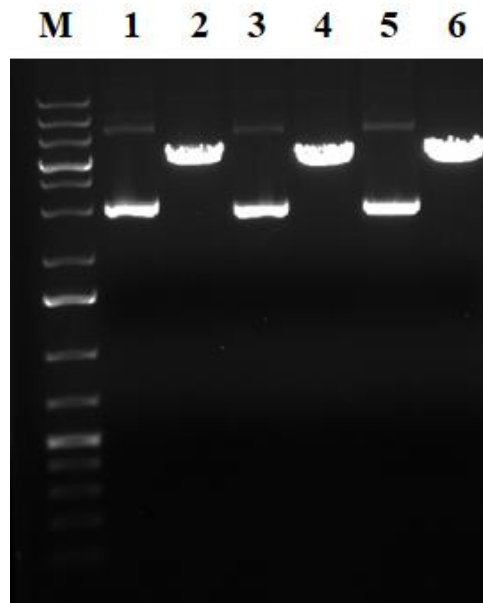


Figura 8. Linearização dos vetores pK α FEGF-ld, pKSUC2EGF-ld e pKPHO1EGF-ld com a enzima SacI. Eletroforese em gel de agarose 1% corado com brometo de etídeo apresentando o DNA plasmidial intacto (1,3,5) e digerido com SacI (2,4,6) para 1 clone transformante dos sistemas pK α FEGF-ld, pKSUC2EGF-ld e pKPHO1EGF-ld respectivamente. M: O'GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific, EUA).

Com a ideia de produzir o hEGF íntegro, foi deletado o gene *KEX1* da linhagem M12 (Betancur *et al.*, 2017) para evitar a degradação do peptídeo pela proteinase codificada por esse gene. A linhagem M12-K de *K. phaffii* foi transformada por eletroporação com cada um dos vetores de expressão pK α FEGF-ld, pKSUC2EGF-ld e pKPHO1EGF-ld.

Quando a levedura foi transformada com os diferentes vetores, começaram a aparecer colônias nas placas após quatro dias de incubação. Na placa controle, onde foi plaqueada a levedura submetida a eletroporação sem DNA, não houve crescimento. Foram selecionadas 30 colônias de cada sistema e isoladas para uma nova placa de MD (Figura 9). Foi semeada a levedura sem transformar para ser usada como controle.

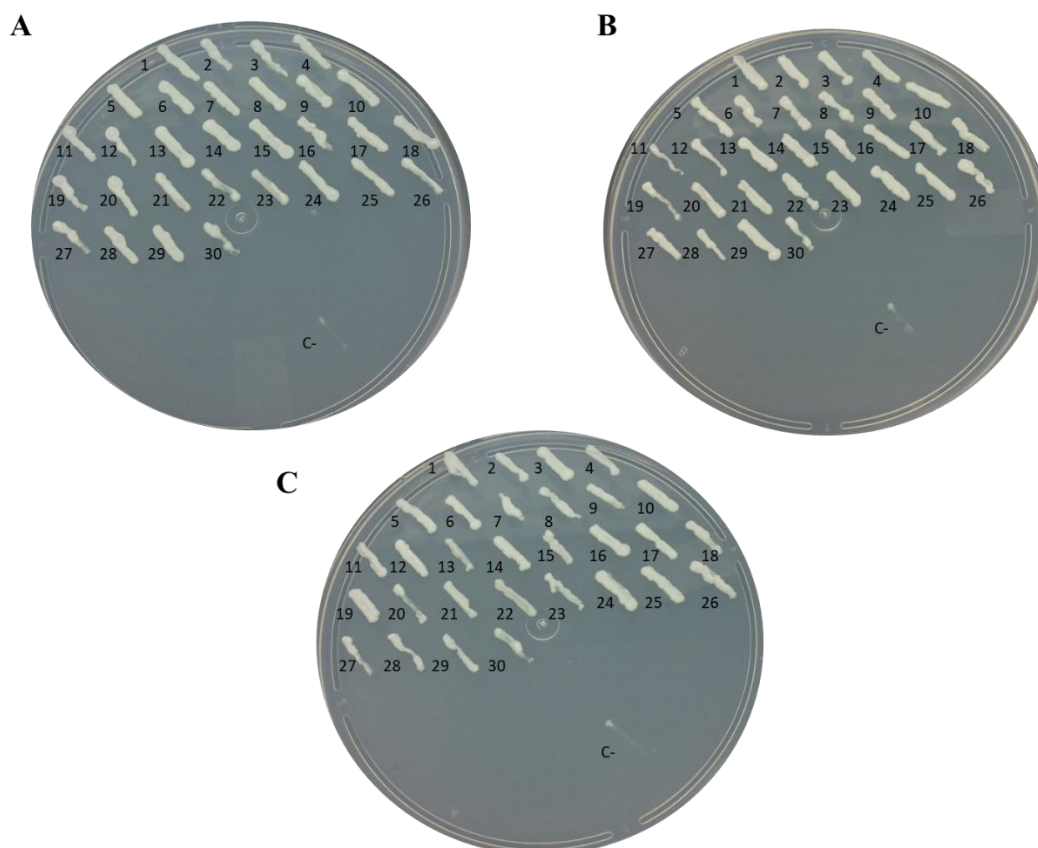


Figura 9. Resultado da transformação de *K. phaffii* M12-K com os vetores pKaFEGF-ld (A), pKSUC2EGF-ld (B) e com pKPHO1EGF-ld (C). Colônias selecionadas em placas com meio MD após 4 dias de incubação a 28 °C. Controle negativo (c-): *K. phaffii* M12-K sem transformar.

5.4. Confirmação da presença do gene do hEGF

A partir dos 30 clones plaqueados após a transformação, foram selecionados 6 clones de cada sistema para avaliar a presença do gene do hEGF com seu respectivo peptídeo sinal por PCR de colônia. As colônias transformantes foram isoladas e foi feita uma PCR de colônia para testar quais dos transformantes eram realmente recombinantes, utilizando os *primers* α FEGF-F e EGFAOX1-R; SUC2EGF-F e EGFAOX1-R; PHO1EGF-F e EGFAOX1-R, respectivamente. Foram usados como controles negativos reações de PCR com os respectivos *primers* para cada sistema e sem a colônia. Se a integração tivera ocorrido satisfatoriamente deveríamos esperar fragmentos de 203, 258 e 267 pb para os sistemas com o α F, *SUC2* e *PHO1*, respectivamente. A figura 10 mostra o resultado da eletroforese onde pode-se observar que todas as colônias transformantes selecionadas são recombinantes.

Os clones recombinantes foram selecionados e armazenados para caracterização posterior.

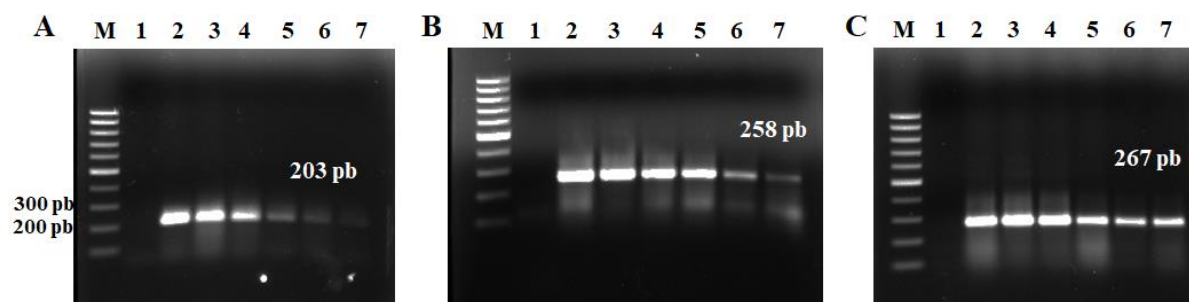


Figura 10. PCR de colônia de 6 clones de cada sistema. Eletroforese em gel de agarose 1% corado com brometo de etídeo. A: Sistema pKaFEGF-ld, B: Sistema pKSUC2EGF-ld e C: pKPHO1EGF-ld. M: marcador 100 pb (New England Biolabs, EUA), 1: Controle negativo, 2-7: colônias de cada sistema.

5.5. Avaliação do crescimento dos clones transformantes

Para avaliar se havia diferença na velocidade de crescimento foram selecionados 5 clones de cada sistema e foi feita uma cinética de crescimento em MD. O crescimento foi comparado com o da levedura X-33. As análises foram realizadas em triplicata. A figura 11 mostra as curvas de crescimento com valores de $\ln OD_{600nm}$ em função do tempo, em horas.

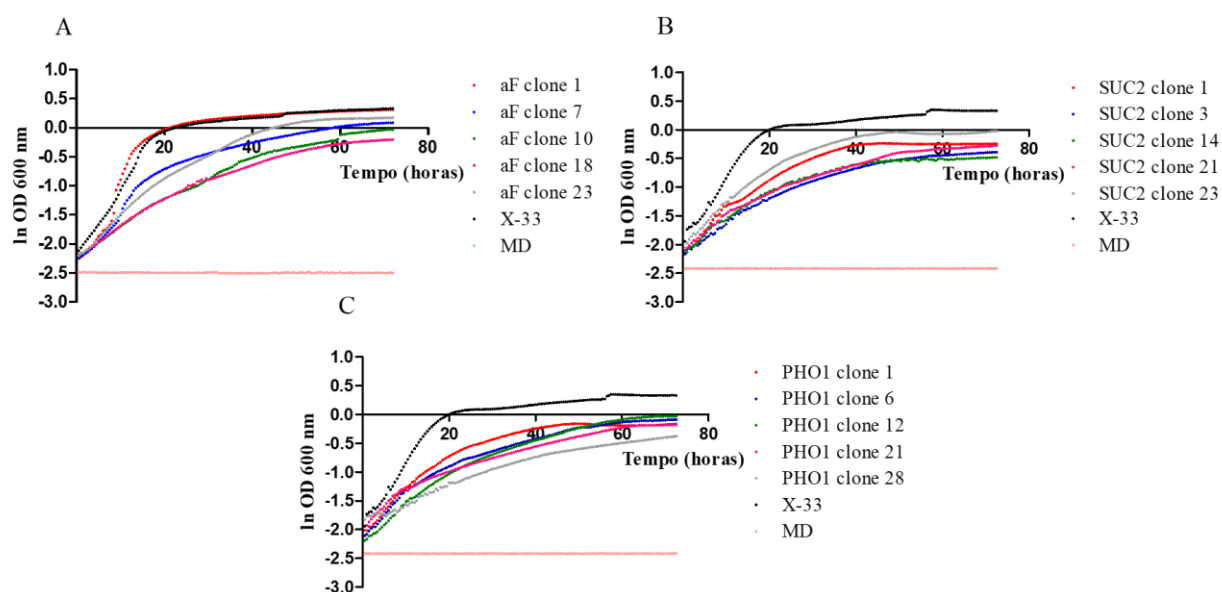


Figura 11. Cinética de crescimento dos clones transformados com os vetores pKaFEGF-ld, pKSUC2EGF-ld e pKPHO1EGF-ld. Células crescidas em meio mínimo (MD) a 30°C durante 72 h. OD_{600} medida a cada 30 min. Todas as culturas começaram com uma OD_{600nm} de 0,1. X-33: linhagem prototrófica de *K. phaffii*. A: Sistema pKaFEGF-ld, B: Sistema pKSUC2EGF-ld e C: Sistema pKPHO1EGF-ld.

Como observado na figura 11, houve clones que apresentaram um perfil de crescimento parecido ao de X-33 e outros que cresceram muito mais lentamente. A tabela 5

apresenta a velocidade específica máxima de crescimento ($\mu_{\text{máx}}$) para cada um dos clones avaliados.

Tabela 5. Velocidade específica máxima de crescimento ($\mu_{\text{máx}}$) dos clones avaliados.

Sistema	Clone	$\mu_{\text{máx}}$ (h^{-1})
αF	1	$0,1421 \pm 0,007395$ ns
	7	$0,08950 \pm 0,003217$
	10	$0,05148 \pm 0,0006114$
	18	$0,05009 \pm 0,0007032$
	23	$0,06841 \pm 0,001070$
	X-33	$0,1301 \pm 0,001750$
SUC2	1	$0,06256 \pm 0,001713$
	3	$0,04911 \pm 0,0008656$
	14	$0,05305 \pm 0,0009771$
	21	$0,04833 \pm 0,001223$
	23	$0,06606 \pm 0,001036$
	X-33	$0,1113 \pm 0,001861$
PHO1	1	$0,04905 \pm 0,0008365$
	6	$0,02915 \pm 0,0006097$
	12	$0,03560 \pm 0,0007036$
	21	$0,02646 \pm 0,0004889$
	28	$0,02923 \pm 0,0004225$
	X-33	$0,1113 \pm 0,001861$

ns: não significativo em comparação à linhagem selvagem X-33. Análise de variância (ANOVA) seguida do teste de Tukey.

A partir deste experimento, foi selecionado um clone de cada sistema para os próximos ensaios (clone 1 ^{α F}, clone 23^{SUC2} e clone 1^{PHO1}).

5.6. Crescimento e expressão dos clones positivos em frasco

Os clones 1 ^{α F}, 23^{SUC2} e 1^{PHO1} foram selecionados para crescimento e expressão em frasco. Foi colocada para crescimento também a linhagem M12-K, como controle negativo. Os clones foram crescidos em frascos de 1L, em um volume final de 100 mL. O ensaio foi realizado em meio complexo e, para isso, os clones foram crescidos em meio BMGY durante 24 h e o sobrenadante foi coletado. A tabela 6 mostra os valores de OD_{600nm} obtidos no final das 24 h de crescimento. As proteínas do sobrenadante foram precipitadas com TCA e analisadas em gel de poliacrilamida em condições desnaturantes e por Western blot.

Tabela 6. OD_{600nm} dos clones selecionados após 24 h de crescimento.

Clone	OD600nm
clone 1^{αF}	28
clone 23^{SUC2}	29,5
clone 1^{PHO1}	20,5

M12-K	26,5
-------	------

5.7. Identificação do hEGF

5.7.1. Avaliação das proteínas presentes no sobrenadante

Para confirmar a presença de uma banda do tamanho do hEGF (~6 kDa), foi feita uma eletroforese em gel de poliacrilamida em condições desnaturantes (SDS-PAGE). Foram aplicados em cada poço do gel 10 µg de proteínas totais, quantificadas pelo método de Bradford. A figura 12 mostra o resultado do gel.

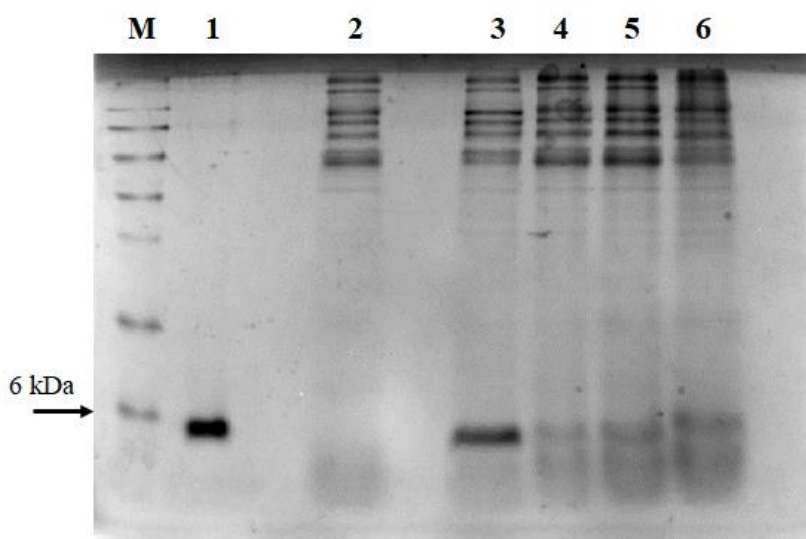


Figura 12. SDS-PAGE 15% de clones selecionados. Análise feita por SDS-PAGE 15%, corado com Coomassie Brilliant Blue. Sobrenadante de clones crescidos em meio BMGY durante 24 h a 28 °C e 250 rpm. M: Marcador SeeBlue Pre-Stained Protein Standard (Life Technologies, EUA). 1: EGF padrão (4µL) (Sigma, EUA), 2: M12-K, 3: clone 1^{aF}, 4: clone 23^{SUC2} e 5: clone 1^{PHO1}. A banda correspondente ao EGF é a de ~6 kDa.

Foi observada uma banda de ~6 kDa em todas as construções, que corresponde ao tamanho do hEGF maduro, assim como foi confirmada a referida banda no poço onde foi aplicado o EGF padrão (Sigma, EUA). No entanto, a banda de ~6 kDa não foi observada no controle correspondente a M12-K, indicando que essa banda corresponde a uma proteína que não está presente no sobrenadante bruto de *K. phaffii*, a qual, muito provavelmente, seria a nossa proteína de interesse.

5.7.2. Confirmação da presença do EGF

Com o objetivo de confirmar se a banda de ~6 kDa presente na eletroforese SDS-PAGE se trata especificamente de nossa proteína de interesse, foi realizado um Western blot utilizando um anticorpo específico para hEGF. Um mililitro de sobrenadante dos clones

selecionados foi precipitado e aplicado no gel SDS-PAGE. Depois de finalizado, o gel foi transferido à membrana de PVDF e, posteriormente, incubado com os anticorpos primário (anti-EGF) e secundário (Goat anti-rabbit IgG H&L HRP), até finalizar o protocolo com a revelação da membrana. A figura 13 mostra o resultado final do processo de revelação da membrana.

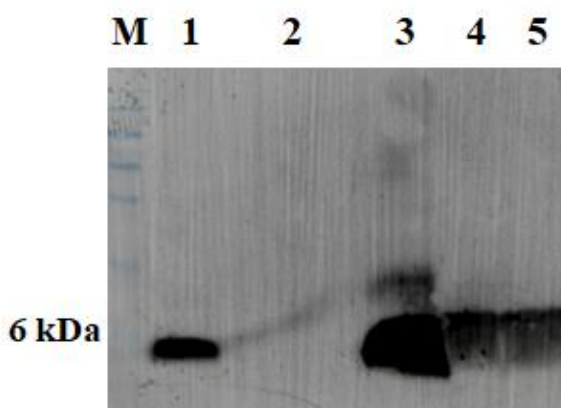


Figura 13. Análise da produção de EGF por Western Blot em clones selecionados. Foi aplicado 1 mL de cada sobrenadante precipitado com TCA e ressuspendido em 25 μ L. M: Marcador SeeBlue Pre-Stained Protein Standard (Life Technologies, EUA), 1: EGF padrão (2 μ L) (Sigma, EUA), 2: M12-K, 3: clone 1 ^{α F}, 4: clone 23^{SUC2} e 5: clone 1^{PHO1}. Anticorpos: anti-EGF antibody ab9695 e Goat anti-rabbit IgG H&L HRP ab205718 (abcam, Reino Unido).

O Western blot realizado revelou o aparecimento de uma banda de ~6 kDa nos clones com os três peptídeos sinais α F, *SUC2* e *PHO1*. Não foi verificado o aparecimento de nenhuma banda na amostra correspondente ao controle negativo M12-K. Foram aplicados no gel 2 μ L do EGF padrão (Sigma, EUA) como controle positivo e confirmada a presença da banda correspondente ao EGF.

5.8. Purificação e identificação da proteína recombinante

Uma vez confirmada a presença do hEGF nas amostras, iniciou-se o processo de purificação. Inicialmente foram utilizados dois passos para a purificação: uma cromatografia de exclusão molecular e um HPLC em fase reversa.

As amostras de sobrenadante de cada um dos clones em estudo foram concentradas em sistema Amicon ultrafiltration setup, com uma membrana de 1 kDa, entre 9-12 vezes (Tabela 7).

Tabela 7. OD_{600nm} e fatores de concentração dos clones selecionados após 24 h.

Clone	OD _{600nm}	Concentração Amicon 1 kDa
clone 1^{aF}	28	12X
clone 23^{SUC2}	29,5	10X
clone 1^{PHO1}	20,5	9,2X
M12-K	26,5	12X

Um mililitro de sobrenadante concentrado foi submetido a uma cromatografia de exclusão molecular Superdex 30 PG. As frações onde foi confirmada a presença do hEGF foram reunidas e aplicadas em um HPLC em fase reversa. A fração do HPLC foi analisada por espectrometria de massas para confirmar o tamanho do hEGF.

5.8.1. Purificação por Cromatografia de Exclusão Molecular (FPLC)

Inicialmente foi feita uma purificação parcial utilizando-se cromatografia de exclusão molecular em uma coluna Superdex 30 PG. Para cada clone foram aplicados 1 mL de sobrenadante concentrado no sistema Amicon 1 kDa. Para a identificação do tempo de eluição do EGF dentro da corrida cromatográfica, foi feita uma primeira cromatografia aplicando o EGF padrão (Sigma, EUA) (Anexo 2). Para cada um dos clones (clone 1^{aF}, clone 23^{SUC2} e clone 1^{PHO1}), foram feitas as cromatografias seguindo as mesmas condições do EGF padrão. As figuras 14 a 17 mostram os cromatogramas obtidos para cada amostra. A partir do cromatograma obtido para o EGF padrão, foram selecionadas e coletadas as frações da 28 a 35 para confirmar a presença do hEGF. Para o caso do sistema com o *PHO1* como peptídeo sinal foram selecionadas as frações 30 a 37. As frações coletadas foram analisadas em gel de poliacrilamida desnaturante e em Western blot.

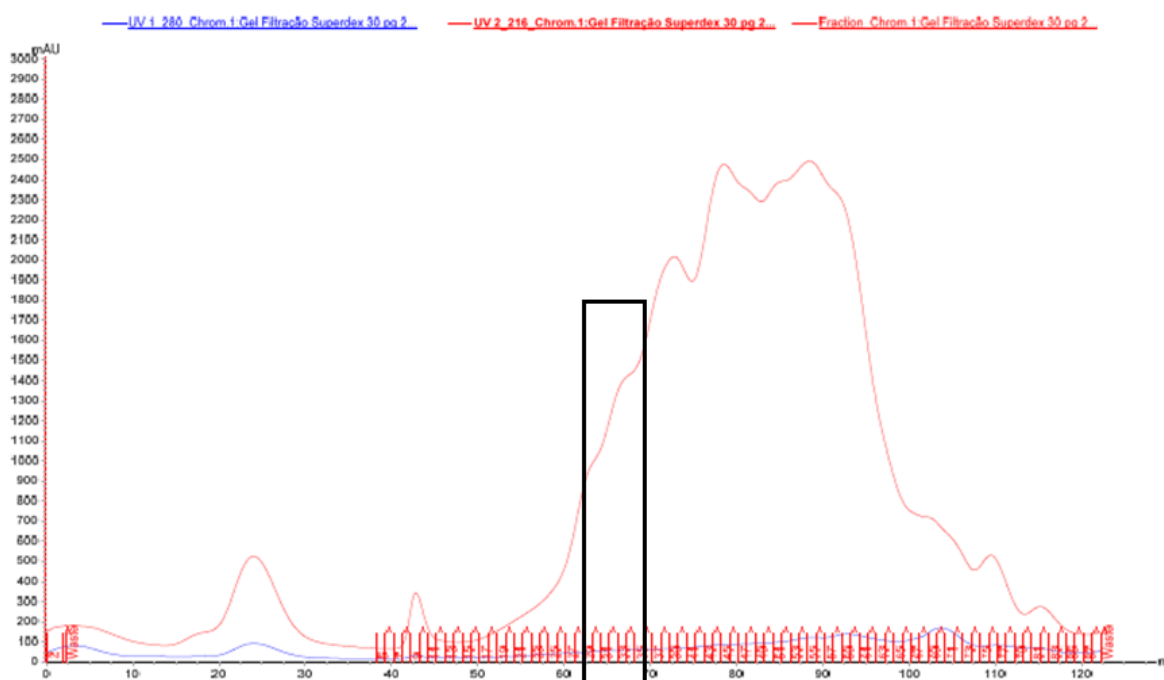


Figura 14. Perfil cromatográfico em FPLC da amostra clone1^{af}. Cromatografia Superdex 30 PG correspondente a 1 mL de sobrenadante concentrado em Amicon 1 kDa. Tampão Fostato de Potássio 50 mM, pH 6. Em vermelho, $\lambda = 216$ nm e em azul $\lambda = 280$ nm. O retângulo representa as frações 28 a 35.

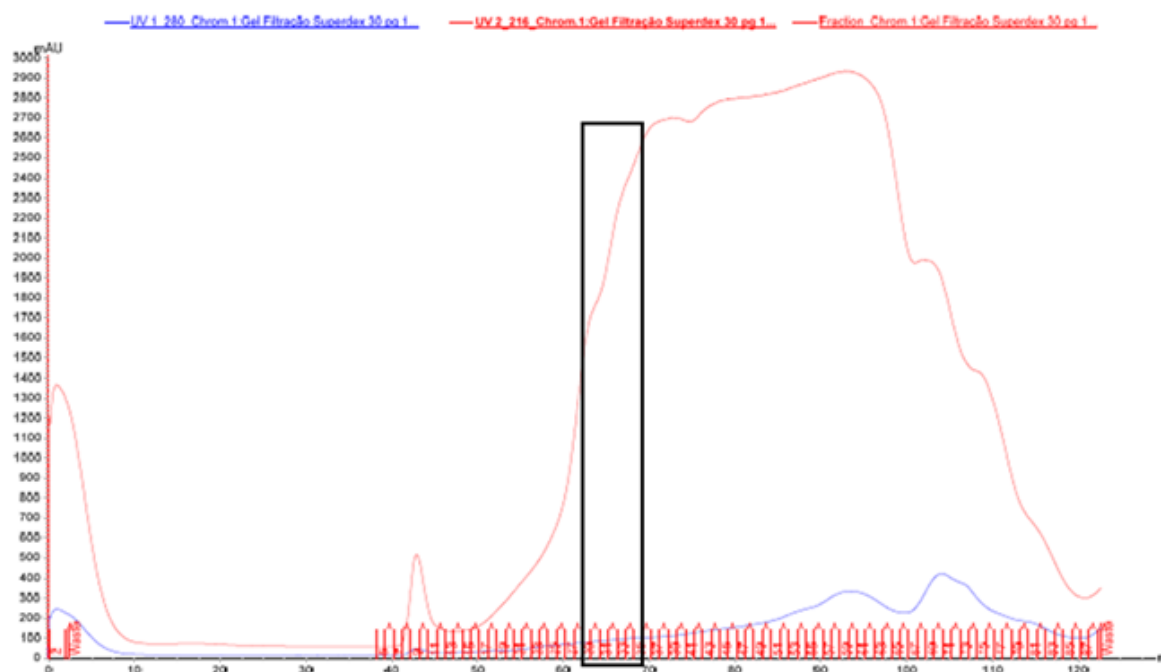


Figura 15. Perfil cromatográfico em FPLC da amostra clone23^{SUC2}. Cromatografia Superdex 30 PG correspondente a 1 mL de sobrenadante concentrado em Amicon 1 kDa. Tampão Fostato de Potássio 50 mM, pH 6. Em vermelho $\lambda = 216$ nm e em azul $\lambda = 280$ nm. O retângulo representa as frações 28 a 35.

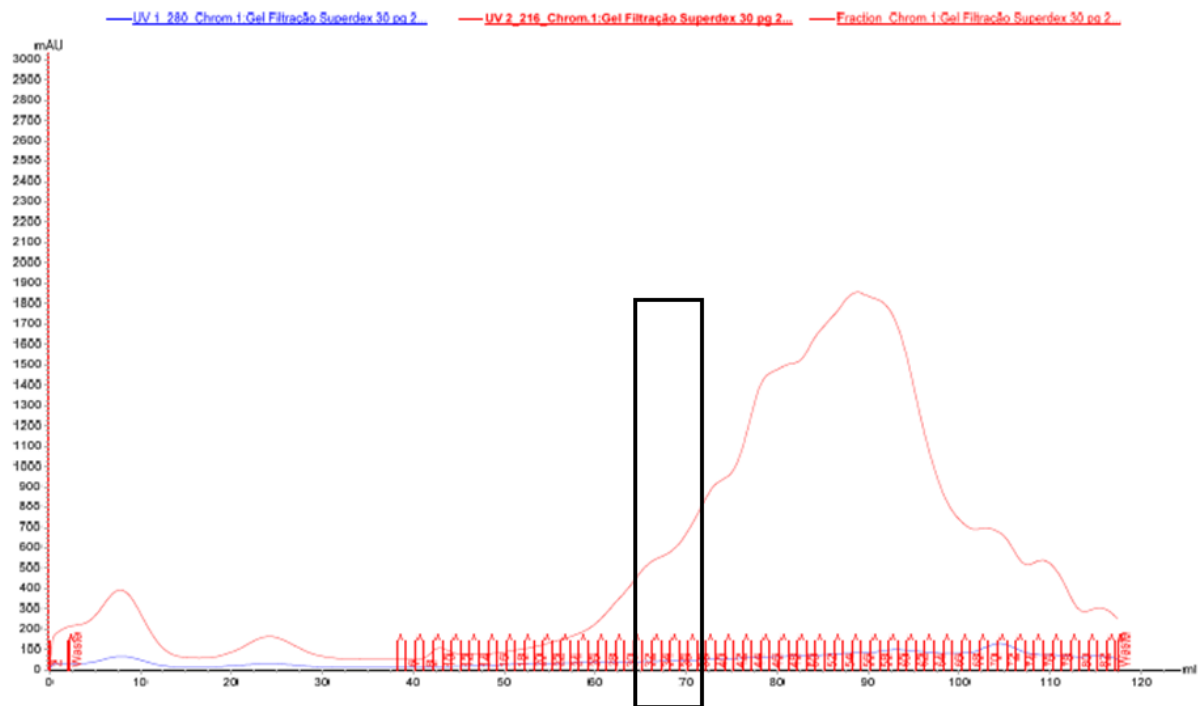


Figura 16. Perfil cromatográfico em FPLC da amostra clone1^{PHO1}. Cromatografia Superdex 30 PG correspondente a 1 mL de sobrenadante concentrado em Amicon 1 kDa. Tampão Fostato de Potássio 50 mM, pH 6. Em vermelho, $\lambda = 216$ nm e em azul $\lambda = 280$ nm. O retângulo representa as frações 30 a 37.

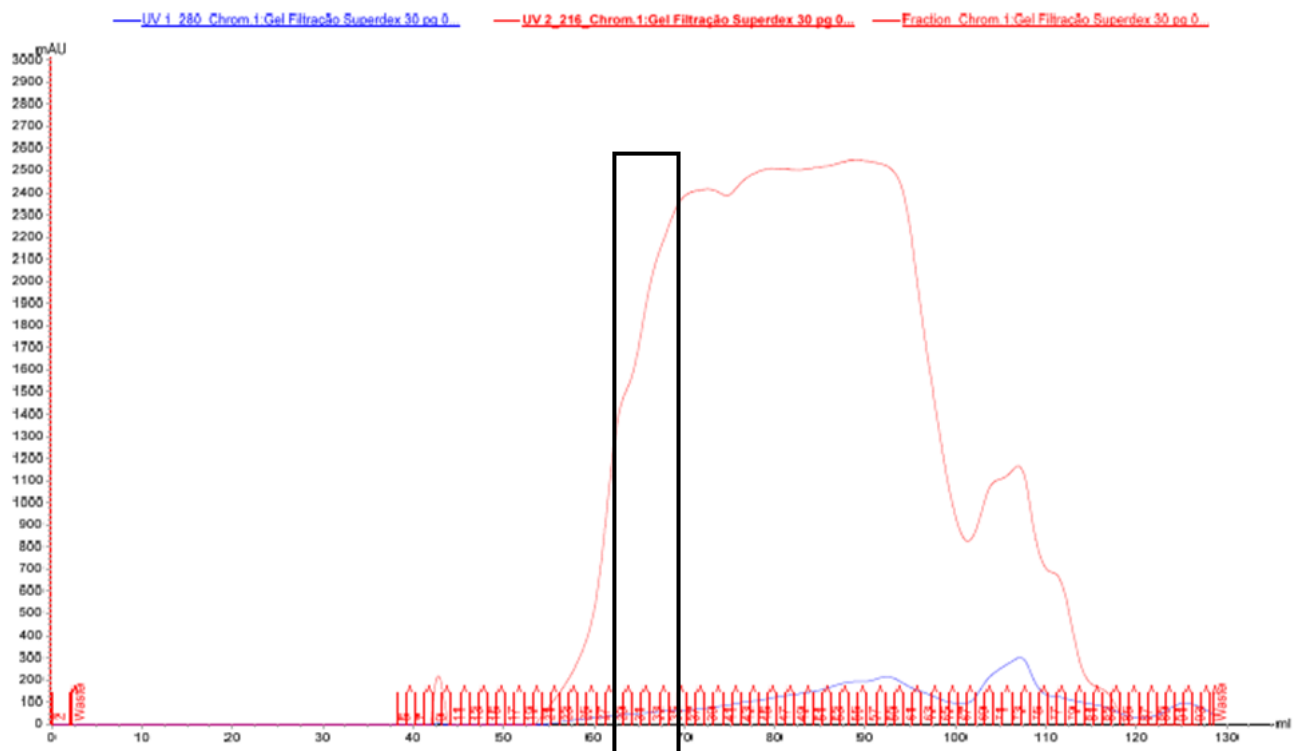


Figura 17. Perfil cromatográfico em FPLC da amostra M12-K. Cromatografia Superdex 30 PG correspondente a 1 mL de sobrenadante concentrado em Amicon 1 kDa. Tampão Fostato de Potássio

50 mM, pH 6. Em vermelho, $\lambda = 216$ nm e em azul $\lambda = 280$ nm. O retângulo representa as frações 28 a 35.

5.8.1.1. SDS-PAGE das frações da cromatografia de exclusão molecular

Uma vez coletadas as frações de interesse foi feito um gel de poliacrilamida para avaliar a presença do hEGF e de outras proteínas ou componentes nesta fração. De cada 1 mL de fração coletada, 500 μ L foram precipitados com TCA e ressuspensos em 20 μ L de tampão de amostra 1X para serem aplicados em um gel (SDS-PAGE). Este mesmo procedimento foi feito para cada um dos clones. Como a amostra poderia ter sido diluída durante a cromatografia, após a corrida do gel a coloração foi feita com nitrato de prata. A figura 18 mostra os resultados obtidos para os clones avaliados.

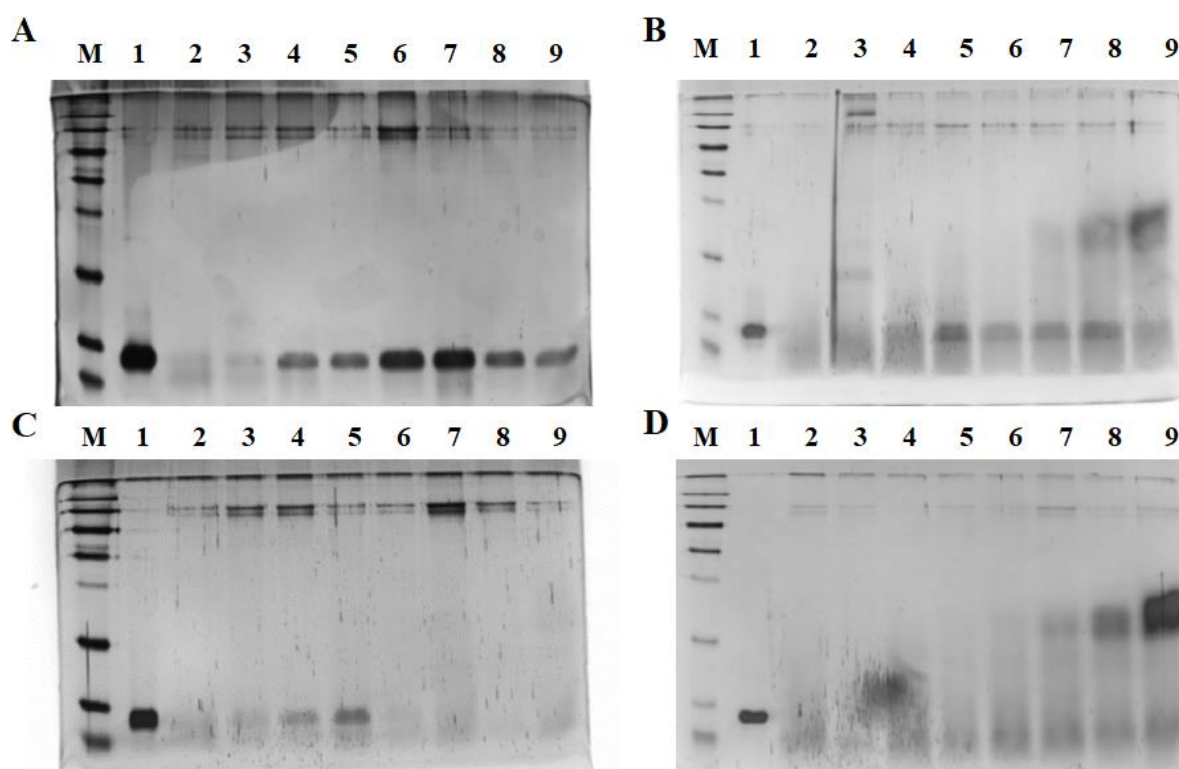


Figura 18. SDS-PAGE das frações coletadas em FPLC. SDS-PAGE 15%. M: Marcador de peso molecular, SeeBlue Pre-Stained Protein Standard (Life Technologies, EUA), 1: EGF padrão (2 μ L) (Sigma, EUA), 2-9: frações 28 a 35 (para A, B e D) e 2-9: frações 30-37 (para C). A: clone 1^{off}, B: clone 23^{SUC2}, C: clone 1^{PHO1} e D: M12-K. Coloração com nitrato de prata.

A partir dos géis obtidos das frações da cromatografia de exclusão molecular foi confirmada a presença de uma banda do tamanho do hEGF, e que é do mesmo tamanho que o EGF padrão também aplicado nos géis como controle positivo. A banda foi confirmada em cada um dos sistemas com os diferentes peptídeos sinais e não foi observada no gel

correspondente a M12-K. A partir desse resultado, foram selecionadas as frações 30 a 33, para confirmar a presença do hEGF por Western blot.

5.8.1.2. Western blot das frações da cromatografia de exclusão molecular

Para confirmar que a banda coletada das frações da cromatografia de exclusão molecular era o hEGF foi feito um Western blot com as frações 30 a 33 de cada clone. De cada uma das frações, 500 μ L foram precipitados e ressuspensos em 20 μ L de tampão de amostra. A figura 19 mostra o resultado do Western blot para cada uma das frações das amostras aplicadas.

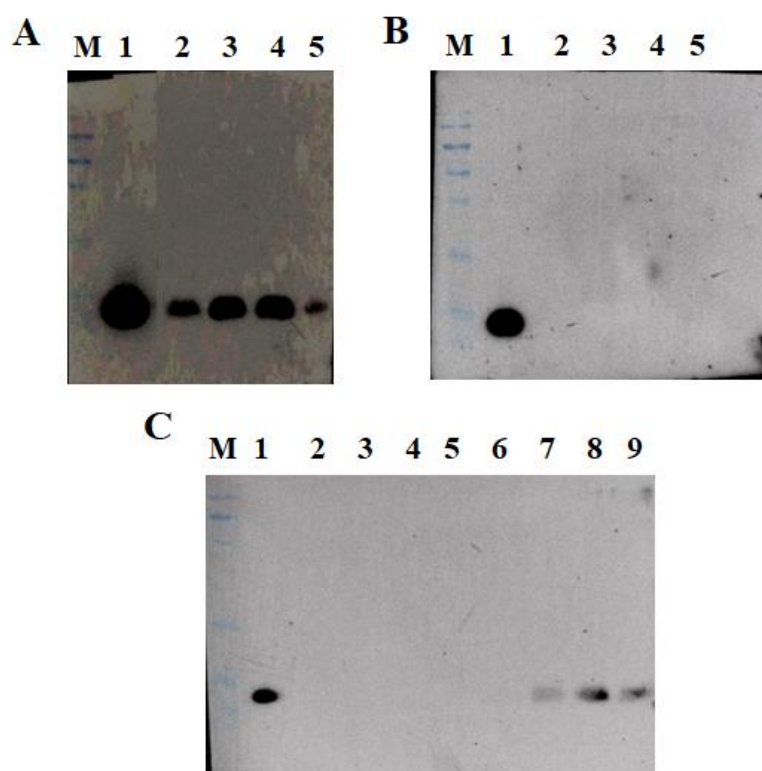


Figura 19. Western blot das frações coletadas em FPLC. M: Marcador SeeBlue Pre-Stained Protein Standard (Life Technologies, EUA), 1: EGF padrão (2 μ L) (Sigma, EUA), 2-5: frações 30 a 33 (para A, B e C) e 6-9: frações 30-33 (C). A: clone 1 ^{α F}, B: clone 1^{PHO1}, C: M12-K (poços 2-5) e clone 23^{SUC2} (poços 6-9).

A partir dos resultados do Western blot foi possível confirmar a presença do hEGF para os sistemas com os peptídeos sinais α F e *SUC2*. Para o caso do sistema com o peptídeo sinal *PHO1* não foi possível identificar a presença do hEGF.

Isto pode ser devido à baixa concentração presente nas frações, o que fez com que a sensibilidade do método não fosse capaz de detectar o hEGF, já que no gel de poliacrilamida as bandas correspondentes às frações estavam muito fracas. Para o caso do controle M12-K,

como esperado, não foi observada nenhuma banda, indicando a ausência do hEGF. As frações 30 a 33 de cada clone foram secas em sistema de SpeedVac para o próximo passo.

5.8.2. HPLC em fase reversa

Uma vez que foi confirmada por Western blot a presença de EGF nas frações 30 a 33 da cromatografia de exclusão molecular para dois dos sistemas, estas frações de todos os sistemas foram submetidas a um ensaio em RP-HPLC. Assim como no primeiro passo de purificação, uma alíquota de EGF padrão foi aplicada em RP-HPLC (Anexo 3). Conhecendo-se seu tempo de retenção, foram passadas as amostras correspondentes aos diferentes peptídeos sinais. Os perfis cromatográficos podem ser observados nas figuras 20 a 23.

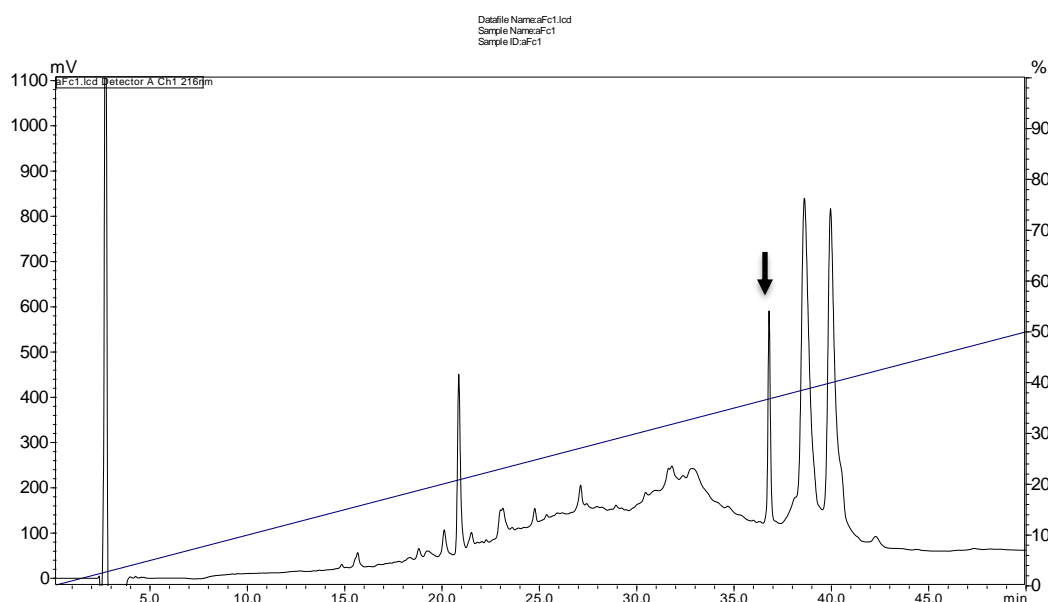


Figura 20. Perfil cromatográfico em RP-HPLC da amostra clone 1^aF. Cromatografia realizada em coluna de fase reversa C18 (Kinetex Phenomenex 300Å, 250 x 4.6 mm). A fração de interesse está indicada com uma seta. O experimento foi monitorado a 216 nm, com um fluxo de 1 mL/min. Amostra das frações 30 a 33 coletadas da cromatografia de exclusão molecular. A linha azul representa o gradiente crescente de acetonitrila. A seta indica o pico de interesse.

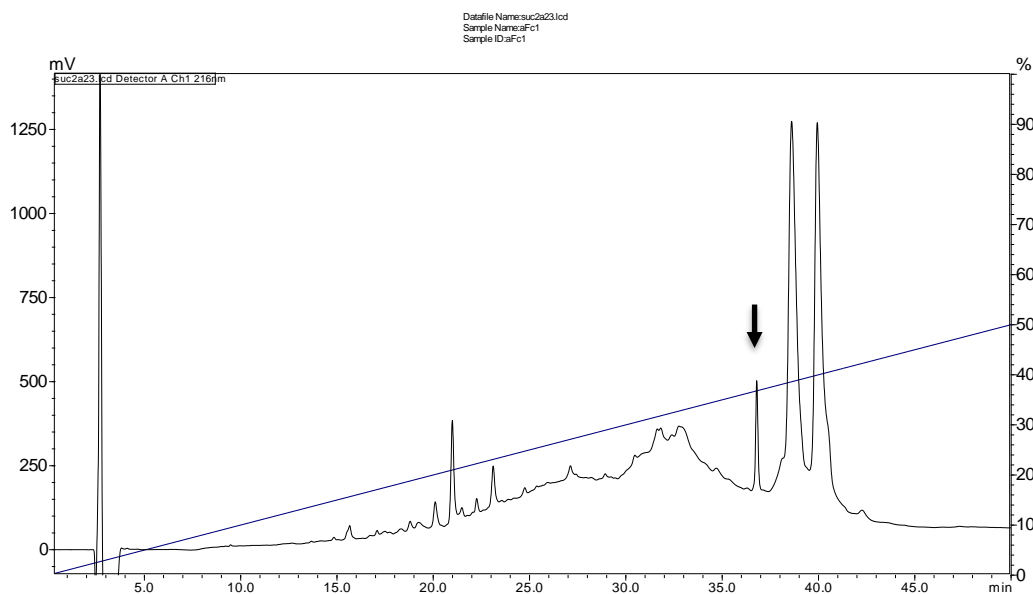


Figura 21. Perfil cromatográfico em RP-HPLC da amostra clone 23^{SUC2}. Cromatografia realizada em coluna de fase reversa C18 (Kinetex Phenomenex 300Å, 250 x 4.6 mm). A fração de interesse está indicada com uma seta. O experimento foi monitorado a 216 nm, com um fluxo de 1 mL/min. Amostra das frações 30 a 33 coletadas da cromatografia de exclusão molecular. A linha azul representa o gradiente crescente de acetonitrila. A seta indica o pico de interesse.

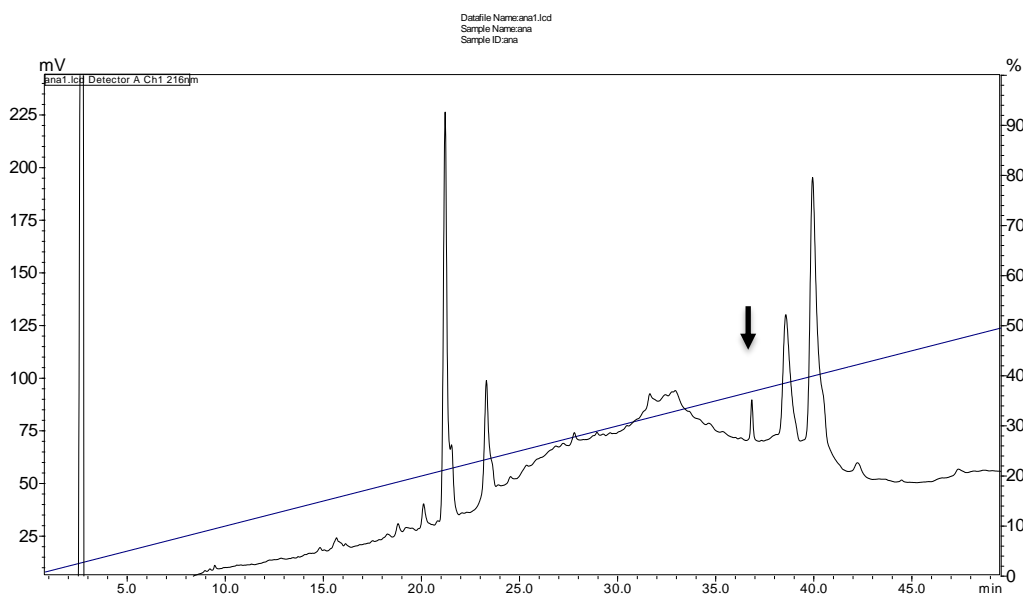


Figura 22. Perfil cromatográfico em RP-HPLC da amostra clone 1^{PHO1}. Cromatografia realizada em coluna de fase reversa C18 (Kinetex Phenomenex 300Å, 250 x 4.6 mm). A fração de interesse está indicada com uma seta. O experimento foi monitorado a 216 nm, com um fluxo de 1 mL/min. Amostra das frações 30 a 33 coletadas da cromatografia de exclusão molecular. A linha azul representa o gradiente crescente de acetonitrila. A seta indica o pico de interesse.

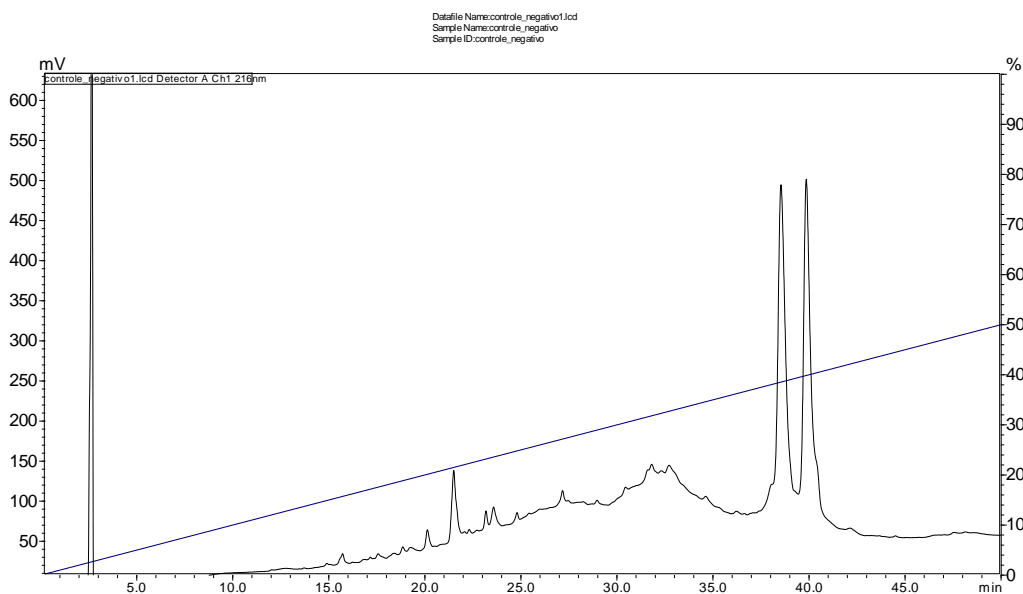


Figura 23. Perfil cromatográfico em RP-HPLC da amostra M12-K. Cromatografia realizada em coluna de fase reversa C18 (Kinetex Phenomenex 300Å, 250 x 4.6 mm). O experimento foi monitorado a 216 nm, com um fluxo de 1 mL/min. Amostra das frações 30 a 33 coletadas da cromatografia de exclusão molecular. A linha azul representa o gradiente crescente de acetonitrila.

A partir dos resultados obtidos nos cromatogramas, é possível identificar a fração que corresponde ao hEGF. Além disso, as amostras aplicadas na cromatografia ainda não estão puras, já que apresentam outros sinais, que podem indicar outros componentes presentes nas amostras. Como esperado, a fração correspondente à amostra com o peptídeo sinal *PHO1* tem um intensidade muito baixa, já que no gel de poliácridamida apresentou umas bandas bem fracas e não exibiu nenhum sinal no Western blot. A partir deste ponto, o sistema *PHO1* foi descartado para as análises posteriores.

5.8.3. Espectrometria de massas

A fim de se determinar o grau de pureza e a massa molecular do hEGF obtida após a etapa de cromatografia em C18 (HPLC em fase reversa), a fração coletada da cromatografia foi analisada em espectrômetro de massa MALDI-TOF/TOF modelo Autoflex II (Bruker Daltonics, Alemanha). Semelhantemente aos casos anteriores, foi feita espectrometria de massas para o EGF padrão (Anexo 4), para confirmar a massa molecular. As amostras correspondentes ao clone 1^{αF} e clone 23^{SUC2} foram submetidas à análise e as figuras 24 e 25 mostram os espectros de massas obtidos, respectivamente.

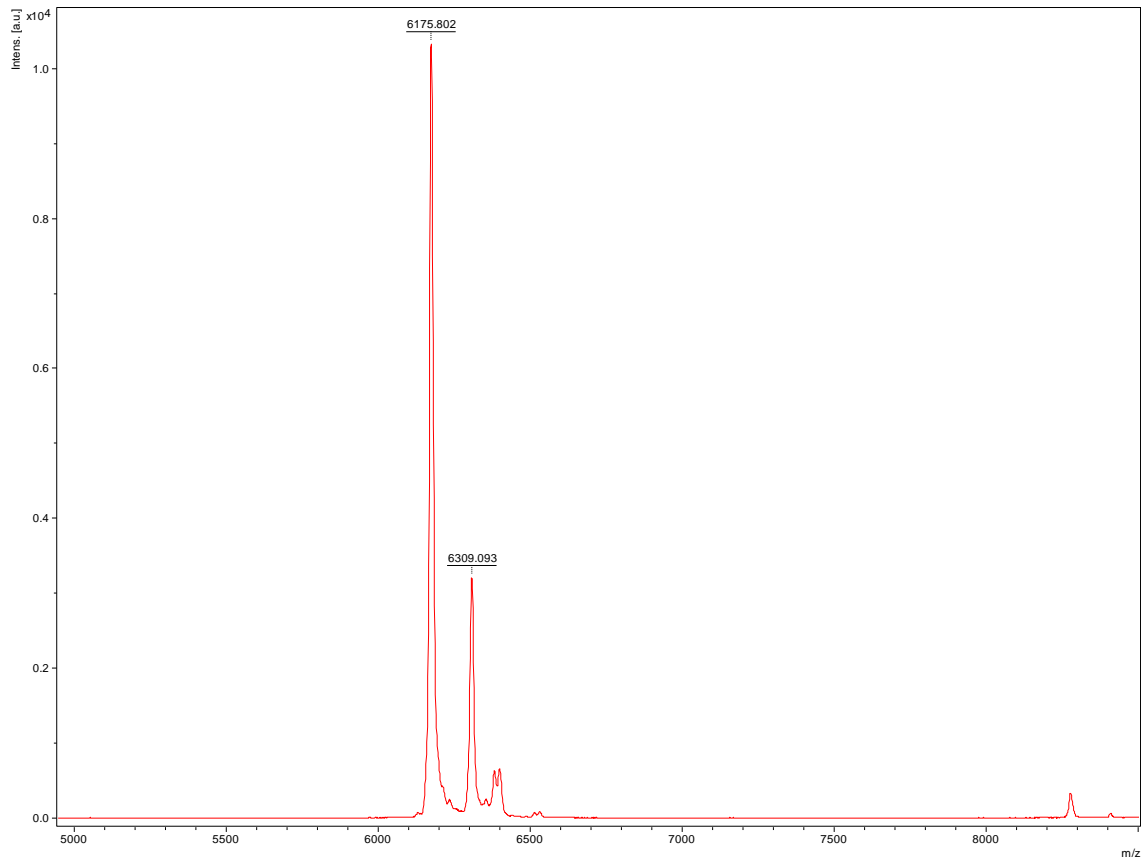


Figura 24. Espectro de massa da amostra clone 1^{aF}. O íon de maior intensidade, representado pela massa 6175,802 Da, é a proteína de interesse. Experimento realizado em espectrômetro de massa MALDI-TOF/TOF modelo Autoflex II (Bruker Daltonics, Alemanha).

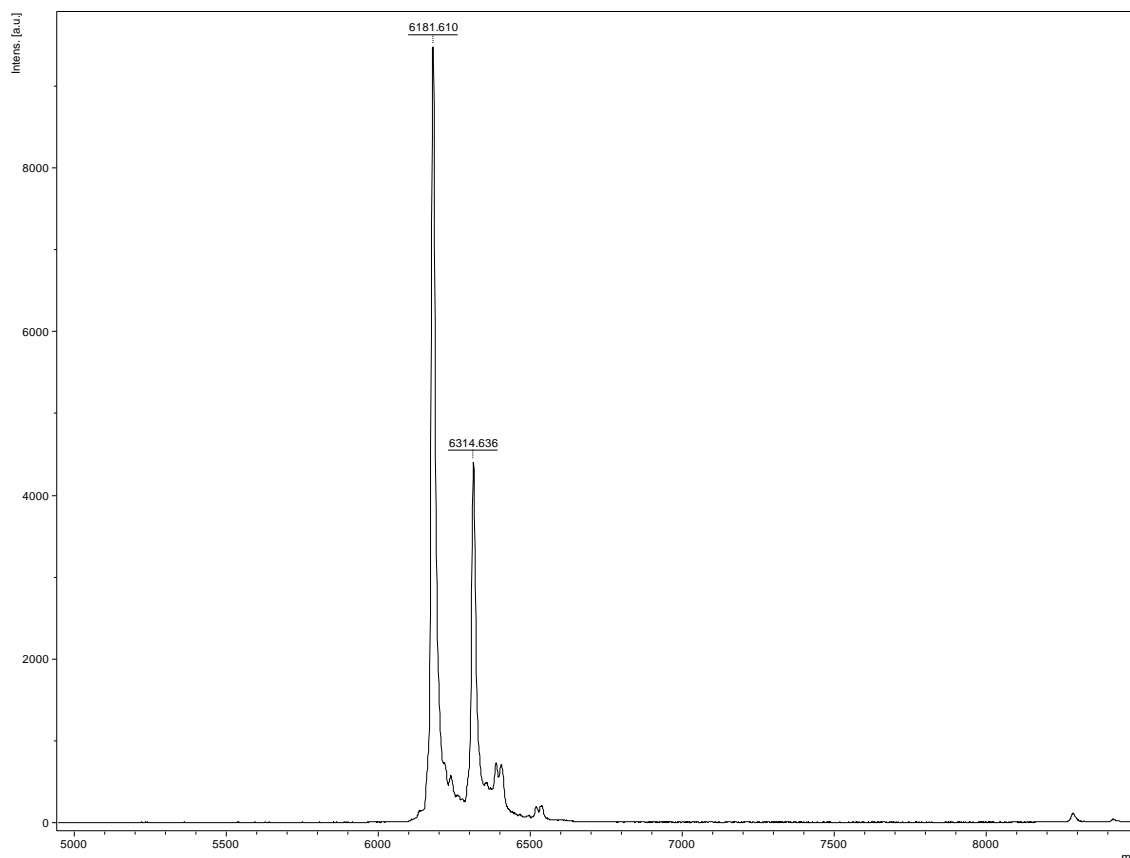


Figura 25. Espectro de massa da amostra clone 23^{SUC2}. O íon de maior intensidade, representado pela massa 6175,802 Da, é a proteína de interesse. Experimento realizado em espectrômetro de massa MALDI-TOF/TOF modelo Autoflex II (Bruker Daltonics, Alemanha).

Pelos resultados obtidos é possível afirmar que no pico observado nos cromatogramas da RP-HPLC não há apenas a proteína de interesse, já que foram obtidos dois picos no espectro de massas. De acordo com a análise de massas, um dos picos corresponde ao hEGF e o outro pode ser um contaminante presente na amostra. Pode ser outra proteína presente no sobrenadante de cultura ou algum contaminante do meio de cultura.

A massa obtida para o EGF padrão foi de **6307,491 g/mol** e, para o caso dos clones 1^{aF} e clone 23^{SUC2}, foi de **6175,802 g/mol** e **6181,610 g/mol** respectivamente.

O EGF padrão (Sigma, EUA) utilizado como controle neste estudo é produzido em *E. coli*. Por isso, ele apresenta na estrutura primária uma metionina adicional no N-terminal, apresentando então 54 aminoácidos. Para que a massa do hEGF de nossos clones esteja correta, a diferença de massas entre o EGF padrão e eles deve ser igual à massa de um resíduo de metionina (131 g/mol). A diferença entre o EGF padrão e as amostras clone^{aF} 1 e clone 23^{SUC2} foi de 131,67 g/mol e 126 g/mol, respectivamente, indicando que o tamanho do hEGF produzido em *K. phaffii* está compatível com o esperado. Essa diferença entre os dois clones

deve-se a um erro na calibração do equipamento. Então, o hEGF secretado por *K. phaffi* apresenta o mesmo tamanho que o EGF humano maduro. No entanto, até esta etapa, o hEGF não pode ser considerado puro.

Uma vez alcançados esses resultados, o passo seguinte foi tentar melhorar o processo de purificação com o objetivo de obter uma fração que, no espectrômetro de massas, defina apenas o hEGF. Para os experimentos seguintes, decidiu-se trabalhar com o sistema α F, já que era o que vinha dando os melhores resultados, com a amostra clone 1 ^{α F}.

5.9. Melhoramento do processo de purificação do hEGF na amostra clone 1 ^{α F}.

A amostra clone 1 ^{α F} foi escolhida para dar continuidade e melhorar o processo de purificação do hEGF. O clone foi crescido em frascos de 1 L, utilizando três frascos contendo 100 mL de cultura, a fim de obter suficiente material para os experimentos de purificação e posteriormente de atividade biológica. Após 24 horas de crescimento, o sobrenadante foi coletado e concentrado utilizando o sistema Amicon ultrafiltration setup utilizando uma membrana de 1 kDa. A amostra foi concentrada 12 vezes. Como aparentemente o clone 1 ^{α F} apresentou uma melhor produção, as frações tanto da cromatografia de exclusão molecular como do HPLC em fase reversa foram submetidas a nova cromatografia para uma melhor pureza.

5.9.1. HILIC

Tendo em consideração que, quando foi analisada a fração da cromatografia em HPLC em fase reversa por espectrometria de massas, foram obtidas duas frações, indicando a presença de mais de uma proteína ou contaminante na amostra, a fração do HPLC em fase reversa foi submetida a uma cromatografia em HILIC, para tentar separar o hEGF do outro componente. A figura 26 mostra o cromatograma obtido.

Satisfatoriamente, a amostra aplicada produziu duas frações no cromatograma, indicando que, efetivamente, apenas a cromatografia em HPLC em fase reversa não era suficiente para separar o hEGF.

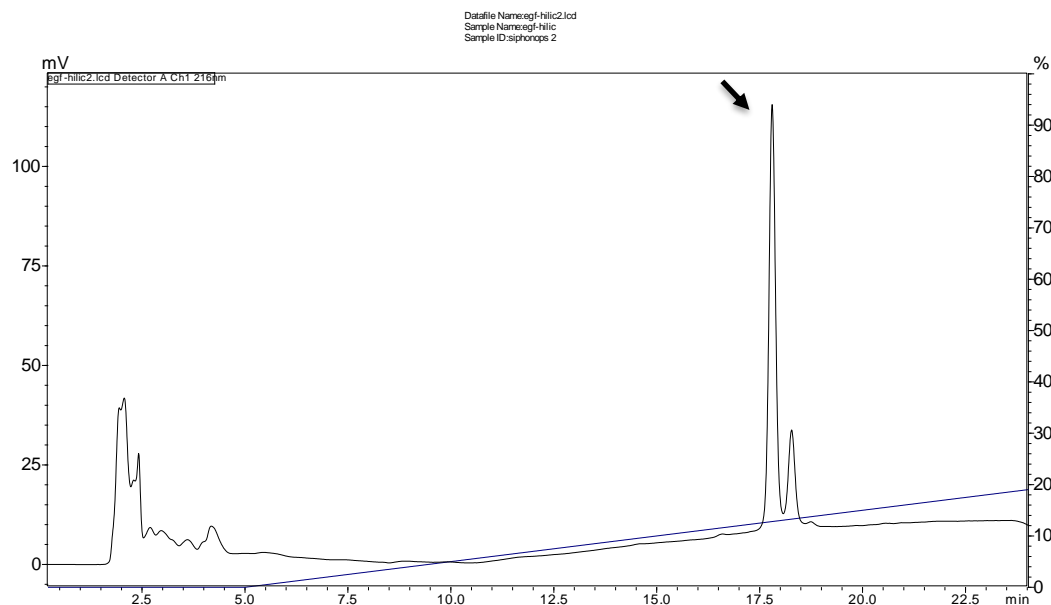


Figura 26. Perfil cromatográfico em HILIC da amostra clone 1^{af}. Cromatografia realizada em fase normal com coluna HILIC (Phenomenex 300Å, 150 x 4.6 mm). O experimento foi monitorado a 216 nm, com um fluxo de 1 mL/min. Amostra do pico coletado da cromatografia em HPLC-fase reversa. A linha azul representa o gradiente crescente de acetonitrila. A seta indica o pico de interesse.

A fração foi coletada manualmente e guardada para posterior análise.

5.9.2. Espectrometria de massas

Para confirmar que efetivamente o hEGF havia sido purificado, a fração de maior intensidade obtida na cromatografia em HILIC foi analisada por espectrometria de massas. Na figura 27 pode-se observar que a fração obtida corresponde ao EGF humano e que a massa monoisotópica obtida no espectro, 6216,269 Da, corresponde à massa monoisotópica do EGF humano.

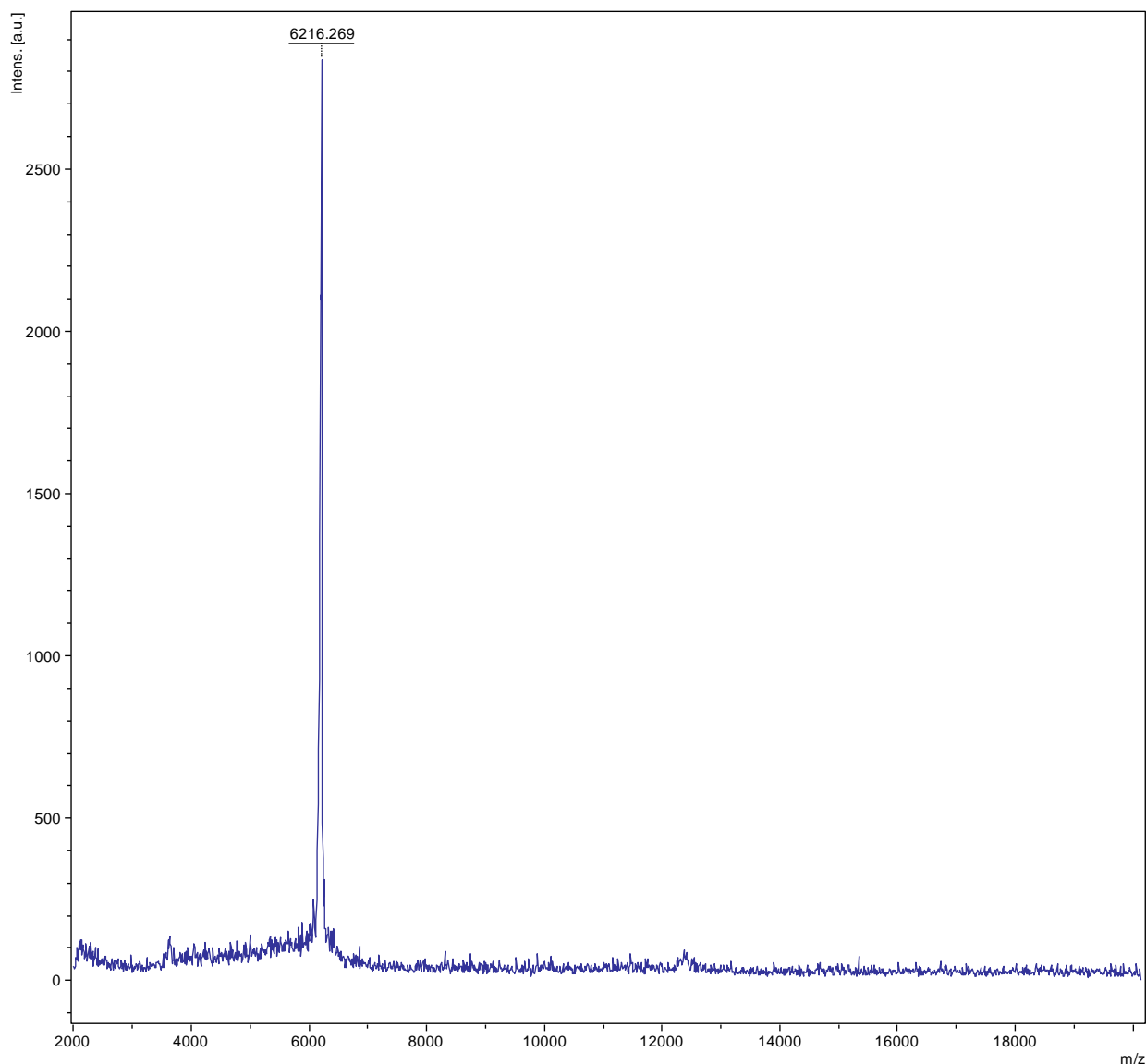


Figura 27. Espectro de massa da amostra clone 1^{aF}. O íon de maior intensidade, representado pela massa monoisotópica 6216,269 Da, é o hEGF. Experimento realizado em espectrômetro de massa MALDI-TOF/TOF modelo Autoflex II (Bruker Daltonics, Alemanha).

Como pode ser visualizado no resultado obtido, este método cromatográfico em HILIC utilizado como etapa final na purificação do hEGF foi efetivo, uma vez que o polipeptídeo recombinante foi detectado numa única fração, que foi confirmada depois no espectro de massas. Com a adição de um terceiro passo no processo de purificação, a cromatografia em HILIC, foi possível a separação do EGF humano com um elevado grau de pureza.

5.9.3. Sequenciamento automático de Edman

Para confirmar a sequência correta do N-terminal do hEGF obtido após a cromatografia em HILIC a amostra foi submetida a um sequenciamento automático pelo

método de degradação de Edman. A sequência de aminoácidos do hEGF produzido em *K. phaffii* foi alinhada (CLUSTAL O (1.2.4) *multiple sequence alignment*) com a sequência do EGF humano P01133 EGF_HUMAN (www.uniprot.org) e com o EGF padrão (Sigma, Alemanha, sequência obtida com o fabricante) (Figura 28). O sequenciamento chegou até o resíduo 29, o que representa 55 % do total de aminoácidos.

P01133 EGF humano	1	-NSDSECLSHDGYCLHDGVCMYIEALDKYACNCVVGYIGERCQYRDLKHWELR	53
hEGF (α F clone 1)	1	-NSDSECLSHDGYCLHDGVCMYIEALDKY-----	29
EGF padrão (Sigma)	1	MNSDSECLSHDGYCLHDGVCMYIEALDKYACNCVVGYIGERCQYRDLKHWELR	54

Figura 28. Resultado do alinhamento da sequência de aminoácidos obtida pelo método de degradação de Edman. São comparadas as sequências do EGF humano, EGF padrão e hEGF produzido em *K. phaffii*. (CLUSTAL O (1.2.4) *multiple sequence alignment*)

A partir dos resultados obtidos foi confirmada a estrutura primária até o aminoácido 29 do hEGF produzido em *K. phaffii*. Esta estrutura é a mesma presente no EGF humano confirmando que o hEGF secretado utilizando o α F foi corretamente processado no N-terminal.

6. DISCUSSÃO

As leveduras vêm se estabelecendo como importantes ferramentas biotecnológicas para a produção de proteínas heterólogas de aplicação tecnológica, farmacêutica e industrial. Dentre as leveduras, *K. phaffii* se destaca como importante sistema de expressão heteróloga (Fraczek *et al.*, 2018; Piva *et al.*, 2017; Weninger *et al.*, 2016). Essa levedura já foi utilizada com sucesso para a expressão de uma grande variedade de proteínas de interesse industrial (Cereghino *et al.*, 2000; Higgins, 2001). Neste trabalho foi demonstrada a produção do hEGF em *K. phaffii*, baseada na utilização de um vetor sob o controle do promotor do gene *PGK1*, que codifica a 3-fosfoglicerato quinase, e permite a expressão constitutiva de proteínas heterólogas (de Almeida *et al.*, 2005). Para a expressão do hEGF em *K. phaffii* foram escolhidos três peptídeos sinais já utilizados para a expressão heteróloga tanto em *K. phaffii* como em *S. cerevisiae*. A primeira etapa realizada foi a construção dos vetores contendo o hEGF e os peptídeos sinais: o fator α (α F), o *SUC2* e o *PHO1*.

A seleção dos transformantes foi feita em meio mínimo já que, com a presença da marca *leu2-d*, a levedura deveria recuperar a prototrofia. As três construções foram transformadas na linhagem M12-K e as colônias transformantes foram obtidas em MD. A partir da obtenção de colônias transformantes, foram feitos os testes para saber se tais colônias eram recombinantes. A partir de PCR de colônia, foi identificada a presença do gene do hEGF com os respectivos peptídeos sinais, indicando que as sequências de interesse estavam presentes nas colônias. Todavia, a presença do gene não necessariamente implica na expressão do mesmo, já que a proteína expressada precisa ser corretamente processada e secretada ao meio extracelular.

A secreção das proteínas desejadas para o meio de cultura foi analisada nos experimentos de SDS-PAGE e Western blot no sobrenadante de cultura. Com o SDS-PAGE, foi possível demonstrar a presença de uma banda do tamanho do hEGF (aproximadamente 6 kDa), presente nas amostras com os três peptídeos sinais. O experimento de Western blot mostrou que a secreção da proteína utilizando três peptídeos sinais, α F, *SUC2* e *PHO1*, para o ambiente extracelular ocorreu e que a proteína secretada foi eficientemente reconhecida por um anticorpo específico pelo hEGF.

Com o objetivo de se buscar uma caracterização físico-química e estrutural mais apurada do hEGF produzido, tentou-se sua purificação. No início, os três sistemas foram

escolhidos para dar continuidade ao trabalho e, finalmente, somente o sistema baseado no αF foi empregado. A escolha do hEGF com este peptídeo sinal foi devida ao fato deste ter propiciado melhores níveis de secreção como observado nos experimentos de SDS-PAGE e western blot.

Inicialmente, a proteína produzida a partir dos três sistemas de secreção foi parcialmente purificada por cromatografia de exclusão molecular e por HPLC em fase reversa. A partir dessas duas etapas cromatográficas, não se conseguiu a purificação completa da proteína. O sistema de HPLC em fase reversa, normalmente, mostra-se eficiente na separação de compostos proteicos de baixa massa molecular. No entanto, nos resultados apresentados naquele estágio do trabalho, as etapas de purificação utilizadas não foram suficientes para que se obtivesse o material com o grau de pureza necessário para a realização dos ensaios biológicos.

A partir dos resultados da purificação parcial do hEGF, foi escolhido o clone com o αF e otimizado o processo de purificação com a adição de uma cromatografia em HILIC. Os resultados apresentados demonstram que o hEGF obtido estava com um grau de pureza suficientemente adequado para a realização de uma caracterização mais detalhada e ensaios de atividade biológica.

Embora não tenham sido feitos estudos de cinética de produção nem avaliados os parâmetros fermentativos, no presente trabalho, os níveis de expressão de hEGF em frasco foram bons. Isso é evidenciado pelas bandas visualizadas nos géis de poliacrilamida em condições desnaturantes. Ainda, quando aplicada a mesma quantidade de proteínas totais no gel, foi possível comparar que a expressão tinha sido mais eficiente no sistema baseado no αF como peptídeo sinal. Embora seja um resultado bastante preliminar, pois as culturas foram feitas somente durante 24 horas, é possível especular que poderia haver um aumento da produtividade em tempos maiores de cultivo ou em condições fermentativas otimizadas.

Existem poucos estudos completos de expressão de hEGF em *K. phaffii*, e eles são utilizando o P_{AOXI} e empregando o α -MF como sinal de secreção. Em um estudo publicado no ano 2014, foi expresso em *K. phaffii* o hEGF derivado da linhagem celular Huh-7 (Khan *et al.*, 2014). Só que o estudo só chegou até confirmar a presença da banda de aproximadamente 6 kDa no gel de poliacrilamida, mas não foram feitos estudos posteriores para saber se o hEGF produzido tinha sido corretamente processado e tinha o tamanho correto. Em outro estudo

publicado por esses mesmo autores o hEGF foi parcialmente purificado e foi observado que uma fração do polipeptídeo produzido estava sendo glicosilada (Khan *et al.*, 2017). O hEGF não é glicosilado nem apresenta potenciais sítios de glicosilação, então esse resultado pode ser devido a um incorreto processamento que adicionou ou eliminou resíduos e fez com que se formassem potenciais sítios de glicosilação, que foram reconhecidos posteriormente. Tanto a fração glicosilada obtida como a não glicosilada tiveram atividade biológica medida pelo ensaio de MTT (Khan *et al.*, 2017).

Outro estudo onde foi produzido hEGF utilizou a linhagem GS115 e também um vetor sob o controle do PAOX1. Para a detecção do hEGF foi feito um ELISA e foi confirmada a presença do polipeptídeo. Nesse estudo foram comparadas algumas condições e a secreção obteve os melhores resultados em 60 horas (2,48 µg/mL) (Eissazadeh *et al.*, 2017). Mas o meio empregado foi BMMY que utiliza metanol como agente indutor e que seu uso não é desejável na indústria biofarmacêutica.

No nosso estudo a identidade das bandas obtidas em SDS-PAGE foi confirmada como sendo a proteína hEGF por western blot. Em alguns estudos relacionados com a expressão do EGF, não foi possível detectar a presença da proteína por western blot. Em um trabalho publicado por Khan *et al.*, se descreve a expressão em *K. phaffii* do hEGF derivado da linhagem humana *Huh-7* e o sinal não foi detectado por western blot. Possivelmente, não houve detecção devido à baixa eficiência na expressão e secreção da proteína (Khan *et al.*, 2014).

A verificação da presença do hEGF em SDS-PAGE e a confirmação por western blot no sobrenadante de cultura é indicativo de que a proteína foi processada corretamente. De fato, o dobramento proteico e formação de ligações dissulfeto são fatores limitantes na via de secreção de proteínas (Hohenblum *et al.*, 2003; Kowalski *et al.*, 1998). As proteínas são dobradas no retículo endoplasmático, onde existe um sistema de controle de qualidade que assegura que somente proteínas corretamente dobradas serão secretadas para o meio extracelular (Agaphonov *et al.*, 2002). Proteínas mal dobradas são rapidamente degradadas pela via de degradação proteica associada ao retículo endoplasmático ou por proteases vacuolares (Agaphonov *et al.*, 2002; Hohenblum *et al.*, 2003).

Ainda assim, às vezes o processamento não é totalmente correto e as proteínas secretadas ao meio extracelular não apresentam alguns resíduos do N- ou C-terminal. Sabe-se

também que o processamento correto no N-terminal da proteína é crucial na obtenção da mesma com a estrutura nativa (Li *et al.*, 2012). Em particular a escolha do sinal de secreção pode ter uma grande influência, uma vez que esta sequência precisa ser removida após a proteína ter sido direcionada para o meio extracelular.

Já foi reportado que proteínas que apresentam lisina ou arginina no C-terminal são reconhecidas pela protease Kex1 e este último aminoácido é clivado (Boehm *et al.*, 1999; Sreenivas *et al.*, 2016). Por isso, o desenvolvimento no nosso laboratório da linhagem M12-K é uma alternativa para expressar proteínas que apresentem esses resíduos no seu C-terminal, o que garante o correto processamento dessa extremidade da proteína. O hEGF apresenta arginina no seu C-terminal.

Para garantir que, além do dobramento correto, o hEGF apresentasse o tamanho e a estrutura primária corretos, foram feitos estudos de espectrometria de massas e sequenciamento pelo método de degradação de Edman. Esses estudos confirmaram o tamanho correto da proteína produzida e a estrutura do N-terminal até o aminoácido 29. Para garantir que o hEGF produzido em *K. phaffii* seja igual ao EGF produzido em humanos será sequenciado o polipeptídeo completo.

Para o escalonamento da produção, a utilização de fermentadores é a metodologia mais empregada para a obtenção de altos níveis de proteínas recombinantes em *K. phaffii*. Por isso, para se produzir maiores quantidades de hEGF para a realização de testes de atividade biológica *in vivo* e para a futura produção industrial dessa proteína, a expressão dos clones recombinantes em fermentador é a opção mais promissora. Uma vez padronizados e ajustados os parâmetros fermentativos, a eficiência na produção de hEGF pode aumentar muito.

Para confirmar que o hEGF produzido apresenta as mesmas propriedades que o EGF humano é necessário o desenvolvimento de testes de atividade biológica, para saber se o hEGF produzido é capaz de estimular o crescimento e proliferação de uma linhagem de células dependentes de hEGF. A linhagem MCF-10A (ATCC® CRL-10317™) é uma linhagem epitelial não-tumorigênica e dependente de EGF para seu crescimento (Bessette *et al.*, 2015). Por tanto, a bioatividade de nosso hEGF pode ser avaliada utilizando-o como suplemento para o crescimento desta linhagem celular.

7. CONCLUSÕES E PERSPECTIVAS

Neste trabalho o hEGF foi produzido em *K. phaffii* utilizando três diferentes peptídeos sinais (α F, *SUC2* e *PHO1*).

O hEGF produzido foi purificado em três etapas cromatográficas. O processo pode ser otimizado e escalonado para uma caracterização mais completa do polipeptídeo produzido.

De acordo com os dados de espectrometria de massa, a proteína recombinante possui a massa molecular esperada sugerindo que não houve processamento do C-terminal. Sendo assim, o hEGF produzido apresenta o mesmo tamanho que o EGF produzido em humanos.

A sequência do N-terminal do hEGF foi confirmada pelo método de degradação de Edman, indicando que o sinal de secreção utilizado foi eficientemente clivado.

Com o objetivo de avaliar a bioatividade do hEGF purificado serão feitos testes de atividade biológica utilizando a linhagem celular MCF-10A.

8. REFERÊNCIAS

Abdull Razis AF, Ismail EN, Hambali Z, Abdullah MN, Ali AM, Mohd Lila MA. Expression of recombinant human epidermal growth factor in *Escherichia coli* and characterization of its biological activity. *Applied biochemistry and biotechnology*. 2008;144(3):249-61.

Agaphonov MO, Romanova NV, Trushkina PM, Smirnov VN, Ter-Avanesyan MD. Aggregation and retention of human urokinase type plasminogen activator in the yeast endoplasmic reticulum. *BMC molecular biology*. 2002;3:15.

Ahmad M, Hirz M, Pichler H, Schwab H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Applied microbiology and biotechnology*. 2014;98(12):5301-17.

Ahn J, Hong J, Lee H, Park M, Lee E, Kim C, et al. Translation elongation factor 1-alpha gene from *Pichia pastoris*: molecular cloning, sequence, and use of its promoter. *Applied microbiology and biotechnology*. 2007;74(3):601-8.

Alewood D, Nielsen K, Alewood PF, Craik DJ, Andrews P, Nerrie M, et al. The role of disulfide bonds in the structure and function of murine epidermal growth factor (mEGF). *Growth factors*. 2005;23(2):97-110.

Amorin J. Produção de quimosina B de *Bos taurus* em *Pichia pastoris* 2008.

Arruda A, Reis VC, Batista VD, Daher BS, Piva LC, De Marco JL, et al. A constitutive expression system for *Pichia pastoris* based on the PGK1 promoter. *Biotechnology letters*. 2016;38(3):509-17.

Bernstein JA, Moellman JJ. Progress in the emergency management of hereditary angioedema: focus on new treatment options in the United States. *Postgraduate medicine*. 2012;124(3):91-100.

Bertheleme N, Singh S, Dowell S, Byrne B. Heterologous expression of G-protein-coupled receptors in yeast. *Methods in enzymology*. 2015;556:141-64.

Bessette DC, Tilch E, Seidens T, Quinn MC, Wiegmanns AP, Shi W, et al. Using the MCF10A/MCF10CA1a Breast Cancer Progression Cell Line Model to Investigate the Effect of Active, Mutant Forms of EGFR in Breast Cancer Development and Treatment Using Gefitinib. *PloS one*. 2015;10(5):e0125232.

Betancur MO, Reis VCB, Nicola AM, De Marco JL, de Moraes LMP, Torres FAG. Multicopy plasmid integration in *Komagataella phaffii* mediated by a defective auxotrophic marker. *Microbial cell factories*. 2017;16(1):99.

Bill RM. Playing catch-up with *Escherichia coli*: using yeast to increase success rates in recombinant protein production experiments. *Frontiers in microbiology*. 2014;5:85.

Boehm T, Pirie-Shepherd S, Trinh LB, Shiloach J, Folkman J. Disruption of the *KEX1* gene in *Pichia pastoris* allows expression of full-length murine and human endostatin. *Yeast*. 1999;15(7):563-72.

Bos IG, de Bruin EC, Karuntu YA, Modderman PW, Eldering E, Hack CE. Recombinant human C1-inhibitor produced in *Pichia pastoris* has the same inhibitory capacity as plasma C1-inhibitor. *Biochimica et biophysica acta*. 2003;1648(1-2):75-83.

Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry*. 1976;72:248-54.

Brake AJ, Merryweather JP, Coit DG, Heberlein UA, Masiarz FR, Mullenbach GT, et al. Alpha-factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*. 1984;81(15):4642-6.

Byrne B. *Pichia pastoris* as an expression host for membrane protein structural biology. *Current opinion in structural biology*. 2015;32:9-17.

Capdevila J, Elez E, Macarulla T, Ramos FJ, Ruiz-Echarri M, Tabernero J. Anti-epidermal growth factor receptor monoclonal antibodies in cancer treatment. *Cancer treatment reviews*. 2009;35(4):354-63.

Carpenter G, Cohen S. Epidermal growth factor. *Annual review of biochemistry*. 1979;48:193-216.

Celik E, Calik P. Production of recombinant proteins by yeast cells. *Biotechnology advances*. 2012;30(5):1108-18.

Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS microbiology reviews*. 2000;24(1):45-66.

Chen J, Li H, Chen J. Human epidermal growth factor coupled to different structural classes of cell penetrating peptides: A comparative study. *International journal of biological macromolecules*. 2017;105(Pt 1):336-45.

Clare JJ, Romanos MA, Rayment FB, Rowedder JE, Smith MA, Payne MM, et al. Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies. *Gene*. 1991;105(2):205-12.

Cohen S. Isolation of a mouse submaxillary gland protein accelerating incisor eruption and eyelid opening in the new-born animal. *The Journal of biological chemistry*. 1962;237:1555-62.

Cohen S, Carpenter G. Human epidermal growth factor: isolation and chemical and biological properties. *Proceedings of the National Academy of Sciences of the United States of America*. 1975;72(4):1317-21.

Cooper A, Bussey H. Characterization of the yeast *KEX1* gene product: a carboxypeptidase involved in processing secreted precursor proteins. *Molecular and cellular biology*. 1989;9(6):2706-14.

Coppella SJ, Dhurjati P. alpha-Factor directed expression of the human epidermal growth factor in *Saccharomyces cerevisiae*. *Biotechnology and bioengineering*. 1989;33(8):976-83.

Cregg JM, Barringer KJ, Hessler AY, Madden KR. *Pichia pastoris* as a host system for transformations. *Molecular and cellular biology*. 1985;5(12):3376-85.

Cregg JM, Madden KR, Barringer KJ, Thill GP, Stillman CA. Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*. *Molecular and cellular biology*. 1989;9(3):1316-23.

Cregg JM, Shen S, Johnson M, Waterham HR. Classical genetic manipulation. *Methods in molecular biology*. 1998;103:17-26.

Crombet Ramos T, Rodriguez PC, Neningen Vinageras E, Garcia Verdecia B, Lage Davila A. CIMAvax EGF (EGF-P64K) vaccine for the treatment of non-small-cell lung cancer. *Expert review of vaccines*. 2015;14(10):1303-11.

Damasceno LM, Huang CJ, Batt CA. Protein secretion in *Pichia pastoris* and advances in protein production. *Applied microbiology and biotechnology*. 2012;93(1):31-9.

de Almeida JR, de Moraes LM, Torres FA. Molecular characterization of the 3-phosphoglycerate kinase gene (PGK1) from the methylotrophic yeast *Pichia pastoris*. *Yeast*. 2005;22(9):725-37.

Delneri D, Tomlin GC, Wixon JL, Hutter A, Sefton M, Louis EJ, et al. Exploring redundancy in the yeast genome: an improved strategy for use of the cre-loxP system. *Gene*. 2000;252(1-2):127-35.

Eissazadeh S, Moeini H, Dezfouli MG, Heidary S, Nelofer R, Abdullah MP. Production of recombinant human epidermal growth factor in *Pichia pastoris*. *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]*. 2017;48(2):286-93.

Fraczek MG, Naseeb S, Delneri D. History of Genome Editing in Yeast. *Yeast*. 2018.

Gleeson MA, White CE, Meininger DP, Komives EA. Generation of protease-deficient strains and their use in heterologous protein expression. *Methods in molecular biology*. 1998;103:81-94.

He Y, Schmidt MA, Erwin C, Guo J, Sun R, Pendarvis K, et al. Transgenic Soybean Production of Bioactive Human Epidermal Growth Factor (EGF). *PloS one*. 2016;11(6):e0157034.

Heyland J, Fu J, Blank LM, Schmid A. Quantitative physiology of *Pichia pastoris* during glucose-limited high-cell density fed-batch cultivation for recombinant protein production. *Biotechnology and bioengineering*. 2010;107(2):357-68.

Higgins DR, Cregg JM. Introduction to *Pichia pastoris*. *Methods in molecular biology*. 1998;103:1-15.

Higgins DR. Overview of protein expression in *Pichia pastoris*. *Current protocols in protein science*. 2001;Chapter 5:Unit5 7.

Hohenblum H, Borth N, Mattanovich D. Assessing viability and cell-associated product of recombinant protein producing *Pichia pastoris* with flow cytometry. *Journal of biotechnology*. 2003;102(3):281-90.

Hollenberg MD, Gregory H. Epidermal growth factor-urogastrone: biological activity and receptor binding of derivatives. *Molecular pharmacology*. 1980;17(3):314-20.

Hynes NE, Lane HA. ERBB receptors and cancer: the complexity of targeted inhibitors. *Nature reviews Cancer*. 2005;5(5):341-54.

Kanda K, Ishida T, Hirota R, Ono S, Motomura K, Ikeda T, et al. Application of a phosphite dehydrogenase gene as a novel dominant selection marker for yeasts. *Journal of biotechnology*. 2014;182-183:68-73.

Khan MA, Khan F, Ahmad N, Khan MI, Zafar AU, Husnain T. Expression line approach to recombinant human epidermal growth factor into the yeast, *Pichia pastoris* from Huh-7 cell line. *Molecular biology reports*. 2014;41(3):1445-51.

Khan MA, Haller JA. Ocriplasmin for Treatment of Vitreomacular Traction: An Update. *Ophthalmology and therapy*. 2016.

Khan MA, Ahmed N, Khan MI, Zafar AU, Tahir S, Anjum MS, et al. Bioactivity studies of *Huh-7* cells derived human epidermal growth factor expressed in *Pichia pastoris*. *Bioscience, biotechnology, and biochemistry*. 2017;81(6):1114-9.

Kilpatrick TJ, Bartlett PF. Cloned multipotential precursors from the mouse cerebrum require FGF-2, whereas glial restricted precursors are stimulated with either FGF-2 or EGF.

The Journal of neuroscience : the official journal of the Society for Neuroscience. 1995;15(5 Pt 1):3653-61.

Kowalski JM, Parekh RN, Mao J, Wittrup KD. Protein folding stability can determine the efficiency of escape from endoplasmic reticulum quality control. The Journal of biological chemistry. 1998;273(31):19453-8.

Kurtzman CP. Description of *Komagataella phaffii* sp. nov. and the transfer of *Pichia pseudopastoris* to the methylotrophic yeast genus *Komagataella*. International journal of systematic and evolutionary microbiology. 2005;55(Pt 2):973-6.

Kurtzman CP. Biotechnological strains of *Komagataella (Pichia) pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis. Journal of industrial microbiology & biotechnology. 2009;36(11):1435-8.

Kwon YB, Kim HW, Roh DH, Yoon SY, Baek RM, Kim JY, et al. Topical application of epidermal growth factor accelerates wound healing by myofibroblast proliferation and collagen synthesis in rat. Journal of veterinary science. 2006;7(2):105-9.

Larsen S, Weaver J, de Sa Campos K, Bulahan R, Nguyen J, Grove H, et al. Mutant strains of *Pichia pastoris* with enhanced secretion of recombinant proteins. Biotechnology letters. 2013;35(11):1925-35.

Lee DN, Kuo TY, Chen MC, Tang TY, Liu FH, Weng CF. Expression of porcine epidermal growth factor in *Pichia pastoris* and its biology activity in early-weaned piglets. Life sciences. 2006;78(6):649-54.

Li P, Anumanthan A, Gao XG, Ilangovan K, Suzara VV, Duzgunes N, et al. Expression of recombinant proteins in *Pichia pastoris*. Applied biochemistry and biotechnology. 2007;142(2):105-24.

Li X, Beeson WT, Phillips CM, Marletta MA, Cate JH. Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. Structure. 2012;20(6):1051-61.

Liang S, Li C, Ye Y, Lin Y. Endogenous signal peptides efficiently mediate the secretion of recombinant proteins in *Pichia pastoris*. Biotechnology letters. 2013;35(1):97-105.

Liu B, Zhang Y, Zhang X, Yan C, Zhang Y, Xu X, et al. Discovery of a rhamnose utilization pathway and rhamnose-inducible promoters in *Pichia pastoris*. Scientific reports. 2016;6:27352.

Liu H, Tan X, Russell KA, Veenhuis M, Cregg JM. *PER3*, a gene required for peroxisome biogenesis in *Pichia pastoris*, encodes a peroxisomal membrane protein involved in protein import. The Journal of biological chemistry. 1995;270(18):10940-51.

- Logez C, Alkhalfioui F, Byrne B, Wagner R. Preparation of *Pichia pastoris* expression plasmids. *Methods in molecular biology*. 2012;866:25-40.
- Looser V, Bruhlmann B, Bumbak F, Stenger C, Costa M, Camattari A, et al. Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. *Biotechnology advances*. 2015;33(6 Pt 2):1177-93.
- Lunn M, Banta E. Ecallantide for the treatment of hereditary angioedema in adults. *Clinical Medicine Insights Cardiology*. 2011;5:49-54.
- Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*. 2005;22(4):249-70.
- Massahi A, Calik P. In-silico determination of *Pichia pastoris* signal peptides for extracellular recombinant protein production. *Journal of theoretical biology*. 2015;364:179-88.
- Mattanovich D, Gasser B, Hohenblum H, Sauer M. Stress in recombinant protein producing yeasts. *Journal of biotechnology*. 2004;113(1-3):121-35.
- Morita H, Tomita S, Maeda H, Okamoto A, Yamagata Y, Kusumoto K, et al. Serine-type carboxypeptidase KexA of *Aspergillus oryzae* has broader substrate specificity than *Saccharomyces cerevisiae* Kex1 and is required for normal hyphal growth and conidiation. *Applied and environmental microbiology*. 2012;78(22):8154-7.
- Mu W, Wang Q, Bourland WA, Jiang C, Yuan D, Pan X, et al. Epidermal growth factor-induced stimulation of proliferation and gene expression changes in the hypotrichous ciliate, *Stylonychia lemnae*. *Gene*. 2016;592(1):186-92.
- Nesterenko MV, Tilley M, Upton SJ. A simple modification of Blum's silver stain method allows for 30 minute detection of proteins in polyacrylamide gels. *Journal of biochemical and biophysical methods*. 1994;28(3):239-42.
- Oka T, Sakamoto S, Miyoshi K, Fuwa T, Yoda K, Yamasaki M, et al. Synthesis and secretion of human epidermal growth factor by *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. 1985;82(21):7212-6.
- Pai-Dhungat JV, Parikh F. Diabetic Foot and Heberprot-P. *The Journal of the Association of Physicians of India*. 2015;63(3):142-3.
- Piva LC, Bentacur MO, Reis VCB, De Marco JL, Moraes LMP, Torres FAG. Molecular strategies to increase the levels of heterologous transcripts in *Komagataella phaffii* for protein production. *Bioengineered*. 2017;8(5):441-5.
- Piva LC, De Marco JL, de Moraes LMP, Reis VCB, Torres FAG. Acetamidase as a dominant recyclable marker for *Komagataella phaffii* strain engineering. *Applied microbiology and biotechnology*. 2018;102(6):2753-61.

Prielhofer R, Maurer M, Klein J, Wenger J, Kiziak C, Gasser B, et al. Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*. *Microbial cell factories*. 2013;12:5.

Raspi G. Kallikrein and kallikrein-like proteinases: purification and determination by chromatographic and electrophoretic methods. *Journal of chromatography B, Biomedical applications*. 1996;684(1-2):265-87.

Ro DK, Ouellet M, Paradise EM, Burd H, Eng D, Paddon CJ, et al. Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC biotechnology*. 2008;8:83.

Sambrook J, Fritsch EF, Maniatis T. *Molecular cloning: a laboratory manual*, 2o ed. New York. 1989.

Sears IB, O'Connor J, Rossanese OW, Glick BS. A versatile set of vectors for constitutive and regulated gene expression in *Pichia pastoris*. *Yeast*. 1998;14(8):783-90.

Shen S, Sulter G, Jeffries TW, Cregg JM. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene*. 1998;216(1):93-102.

Shilton BH, Thomas DY, Cygler M. Crystal structure of Kex1deltap, a prohormone-processing carboxypeptidase from *Saccharomyces cerevisiae*. *Biochemistry*. 1997;36(29):9002-12.

Spohner SC, Muller H, Quitmann H, Czermak P. Expression of enzymes for the usage in food and feed industry with *Pichia pastoris*. *Journal of biotechnology*. 2015;202:118-34.

Sreenivas S, Krishnaiah SM, Shyam Mohan AH, Mallikarjun N, Govindappa N, Chatterjee A, et al. Disruption of *KEX1* gene reduces the proteolytic degradation of secreted two-chain Insulin glargine in *Pichia pastoris*. *Protein expression and purification*. 2016;118:1-9.

Stadlmayr G, Mecklenbrauker A, Rothmuller M, Maurer M, Sauer M, Mattanovich D, et al. Identification and characterisation of novel *Pichia pastoris* promoters for heterologous protein production. *Journal of biotechnology*. 2010;150(4):519-29.

Stefanini FR, Maia M, Falabella P, Pfister M, Niemeyer M, Kashani AH, et al. Profile of ocriplasmin and its potential in the treatment of vitreomacular adhesion. *Clinical ophthalmology*. 2014;8:847-56.

Tang X, Liu H, Yang S, Li Z, Zhong J, Fang R. Epidermal Growth Factor and Intestinal Barrier Function. *Mediators of inflammation*. 2016;2016:1927348.

Thomas DR, Walmsley AM. Improved expression of recombinant plant-made hEGF. *Plant cell reports*. 2014;33(11):1801-14.

Tschopp JF, Brust PF, Cregg JM, Stillman CA, Gingeras TR. Expression of the *lacZ* gene from two methanol-regulated promoters in *Pichia pastoris*. *Nucleic acids research*. 1987;15(9):3859-76.

Vogl T, Glieder A. Regulation of *Pichia pastoris* promoters and its consequences for protein production. *New biotechnology*. 2013;30(4):385-404.

Vogl T, Hartner FS, Glieder A. New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris*. *Current opinion in biotechnology*. 2013;24(6):1094-101.

Wang J, Wang X, Shi L, Qi F, Zhang P, Zhang Y, et al. Methanol-Independent Protein Expression by *AOXI* Promoter with trans-Acting Elements Engineering and Glucose-Glycerol-Shift Induction in *Pichia pastoris*. *Scientific reports*. 2017;7:41850.

Wang M, Jiang S, Wang Y. Recent advances in the production of recombinant subunit vaccines in *Pichia pastoris*. *Bioengineered*. 2016;7(3):155-65.

Waterham HR, Digan ME, Koutz PJ, Lair SV, Cregg JM. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene*. 1997;186(1):37-44.

Weinacker D, Rabert C, Zepeda AB, Figueroa CA, Pessoa A, Farias JG. Applications of recombinant *Pichia pastoris* in the healthcare industry. *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]*. 2013;44(4):1043-8.

Weninger A, Hatzl AM, Schmid C, Vogl T, Glieder A. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. *Journal of biotechnology*. 2016;235:139-49.

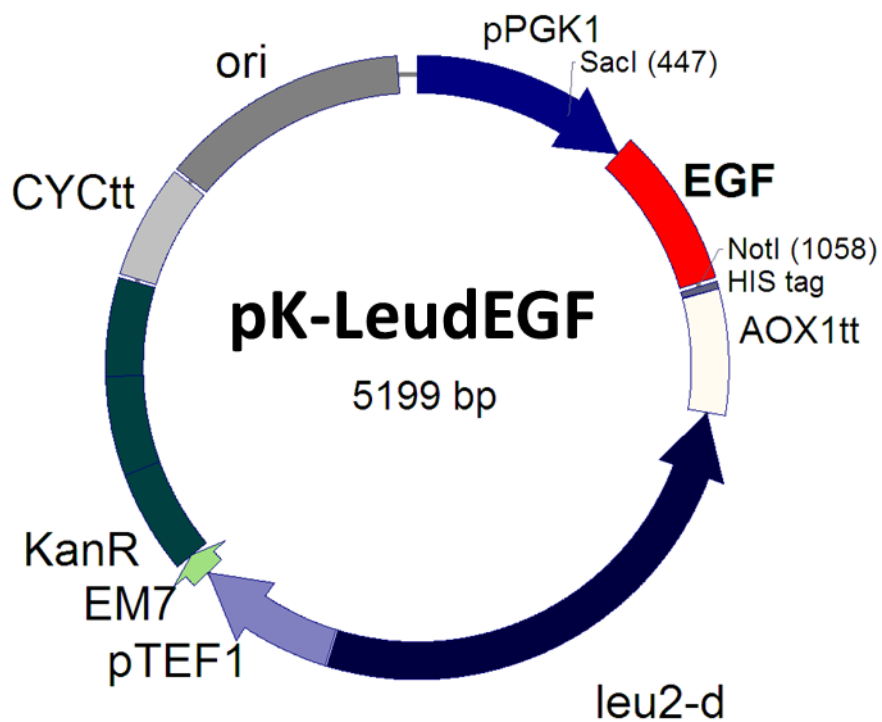
Yamagata H, Nakahama K, Suzuki Y, Kakinuma A, Tsukagoshi N, Udaka S. Use of *Bacillus brevis* for efficient synthesis and secretion of human epidermal growth factor. *Proceedings of the National Academy of Sciences of the United States of America*. 1989;86(10):3589-93.

Yu P, Zhu Q, Chen K, Lv X. Improving the secretory production of the heterologous protein in *Pichia pastoris* by focusing on protein folding. *Applied biochemistry and biotechnology*. 2015;175(1):535-48.

ANEXOS

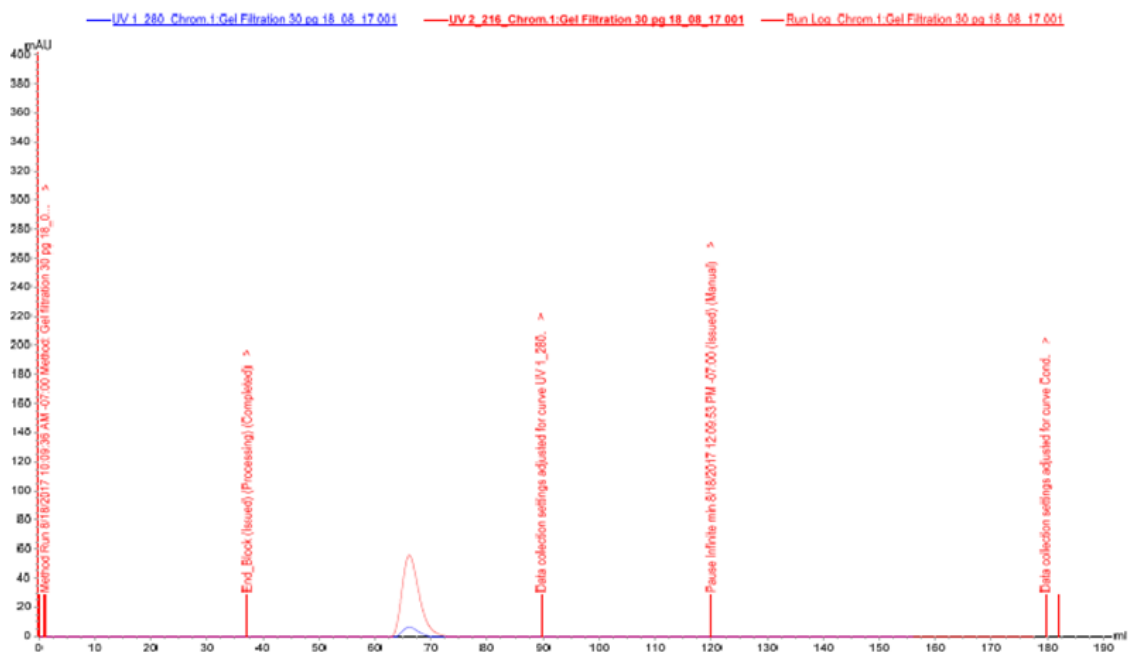
Anexo 1. Mapa físico do vetor pK-LeudEGF .

Vetor utilizado para expressão do cDNA do hEGF em *K. phaffii*.



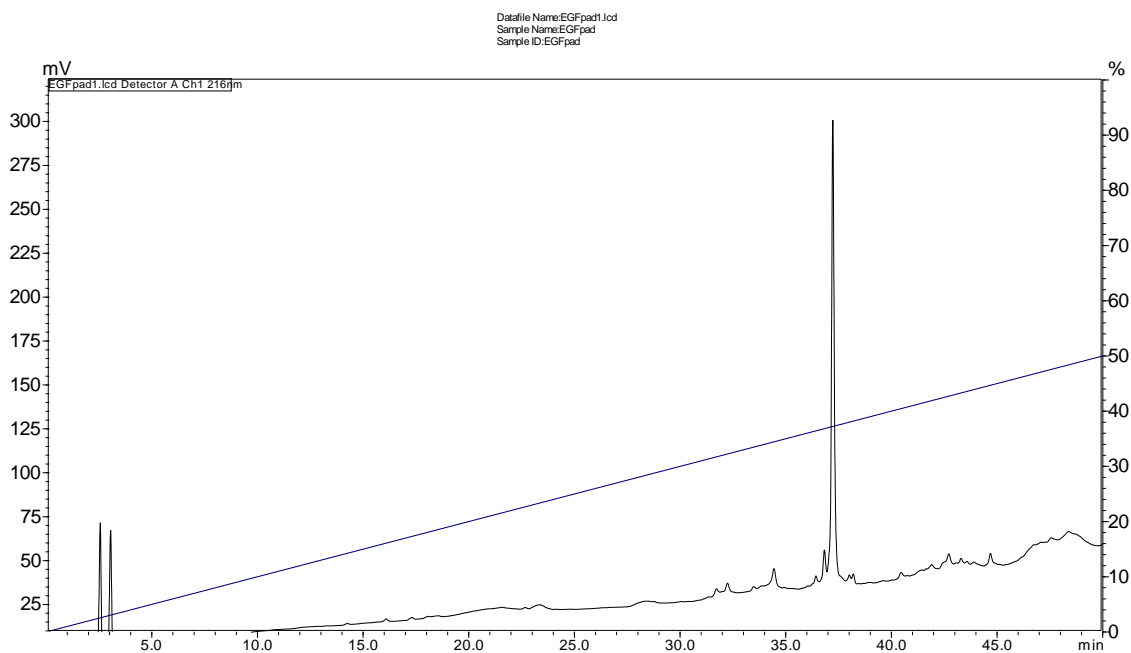
Anexo 2. Perfil cromatográfico em FPLC do EGF padrão (Sigma, EUA).

Cromatografia Superdex 30 PG correspondente a 1 mL de sobrenadante concentrado em Amicon 1 kDa. Tampão Fostato de Potássio 50 mM, pH 6. Em vermelho $\lambda = 216$ nm e em azul $\lambda = 280$ nm



Anexo 3. Perfil cromatográfico em RP-HPLC do EGF padrão (Sigma, EUA).

Cromatografia realizada em coluna de fase reversa C18 (Kinetex Phenomenex 300Å, 250 x 4.6 mm). A fração de interesse está indicada com uma seta. O experimento foi monitorado a 216 nm, com um fluxo de 1 mL/min. Amostra das frações 30 a 33 coletadas da cromatografia de exclusão molecular. A linha azul representa o gradiente crescente de acetonitrila



Anexo 4. Espectro de massas do EGF padrão (Sigma, EUA).

O íon de maior intensidade, representado pela massa 6175,802 Da, é a proteína de interesse. Experimento realizado em espectrômetro de massa MALDI-TOF/TOF modelo Autoflex II (Bruker Daltonics, Alemanha).

