UNIVERSIDADE DE SÃO PAULO INSTITUTO DE QUÍMICA Programa de Pós-Graduação em Ciências Biológicas (Bioquímica)

MARCELA DE OLIVEIRA VITARELLI

Humanização específica do sistema de glicosilação de

Pichia pastoris pela técnica CRISPR/Cas9

visando a expressão de glicoproteínas humanas

Versão corrigida da Dissertação defendida

São Paulo

Data do depósito na SPG:

28/09/2016

MARCELA DE OLIVEIRA VITARELLI

Humanização específica do sistema de glicosilação de *Pichia pastoris* pela técnica CRISPR/Cas9 visando a expressão de glicoproteínas humanas

Dissertação apresentada ao Instituto de Química da Universidade de São Paulo para obtenção do título de Mestre em Ciências (Bioquímica)

Orientador: Prof. Dr. Paulo Lee Ho

São Paulo

2016

Ficha Catalográfica Elaborada pela Divisão de Biblioteca e Documentação do Conjunto das Químicas da USP.

Vitarelli, Marcela de Oliveira

V837h Humanização específica do sistema de glicosilação de Pichia pastoris pela técnica CRISPR/Cas9 visando a expressão de glicoproteínas humanas / Marcela de Oliveira Vitarelli. -- São Paulo, 2016.

101p.

Dissertação (mestrado) – Instituto de Química da Universidade de São Paulo. Departamento de Bioquímica. Orientador : Ho, Paulo Lee

1. Plasmídeo : Engenharia genética 2. Leveduras I. T. II. Ho, Paulo Lee, orientador.

574.88 CDD

À minha amada família, por toda a dedicação e incentivo. Obrigada por sempre acreditarem nos meus sonhos.

AGRADECIMENTOS

Agradeço ao meu orientador, Dr. Paulo Lee Ho, pelos desafios apresentados e acreditar em mim ao longo desses anos. Obrigada pela confiança, pela oportunidade e pelos ensinamentos, eles me fizeram crescer não só como profissional, mas também como pessoa.

Agradeço à Dra. Eliane Namie Miyaji pela disponibilidade e por todo o conhecimento compartilhado, eles foram essenciais ao desenvolvimento do projeto.

Agradeço à Dra. Maria Lenor Sarno de Oliveira pelo incentivo, disponibilidade e pelas valiosas contribuições ao trabalho.

Agradeço aos pesquisadores do Laboratório de Biotecnologia Molecular I, Dr. Enéas de Carvalho, Dra. Josefa da Silva, Dra. Alessandra Schanoski e Dra. Cláudia Sossai, por todo o suporte e aprendizado.

Agradeço à Dra. Juliana Branco Novo, por me acompanhar nessa trajetória, compartilhando todo o seu conhecimento.

Agradeço aos meus queridos amigos do Laboratório de Biotecnologia Molecular I, Rafaella, Júlia, Stefanni, Fabiana, Tasson e Rubia por todos os conselhos, conhecimentos divididos, risadas, desabafos, cafés, chocolates e happy hours. Obrigada pela amizade!

Agradeço às técnicas do laboratório, Vera, Aline e Nídia, por tornar meu trabalho mais simples e por me ajudarem sempre que possível.

Agradeço aos integrantes da República Só Sucesso, Gabriella, Luiza e Pedro, pela amizade, pela convivência, pelas contribuições ao projeto, pelas inúmeras risadas, almoços, jantares e cervejas. Obrigada pela companhia diária!

Agradeço à minha mãe pela motivação, pelo apoio, pela dedicação e carinho de sempre!

Agradeço à minha irmã por estar sempre ao meu lado mesmo distante, por me incentivar e acreditar sempre em mim!

Agradeço ao meu pai pela motivação e por sempre acreditar que não existem caminhos errados na ciência.

Agradeço ao Felipe por toda amizade, amor, carinho e compreensão. Obrigada por estar ao meu lado e me motivar sempre!

Agradeço ao apoio financeiro da FAPESP, CAPES, CNPq e Fundação Butantan, pelo auxílio à pesquisa, possibilitando o desenvolvimento do trabalho.

E a todos que contribuíram direta ou indiretamente para a realização deste trabalho, muito obrigada!

Errata:

Agradeço ao apoio financeiro pelo processo nº 2014/22200-4, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). "As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade dos autores e não necessariamente refletem a visão da FAPESP".

" A ideia ainda brilhava, irreal como uma bolha de sabão, e Lyra não ousou examinála muito de perto para que não estourasse. Mas estava familiarizada com ideias daquele tipo, e assim, ela a deixou brilhar... "

Philip Pullman

RESUMO

Vitarelli, M.O. Humanização específica do sistema de glicosilação de *Pichia pastoris* pela técnica CRISPR/Cas9 visando a expressão de glicoproteínas humanas. 2016. 101p. Dissertação – Programa de Pós-Graduação em Ciências Biológicas (Bioquímica). Instituto de Química, Universidade de São Paulo, São Paulo.

A produção de proteínas terapêuticas recombinantes compreende moléculas complexas e de alto valor agregado, incluindo a enzima glucocerebrosidase (GCase). Sua deficiência resulta na Doença de Gaucher, passível de tratamento por meio da terapia de reposição enzimática. A forma ativa da GCase recombinante usada na terapia apresenta resíduos terminais de manose expostos no seu perfil de glicosilação. Perfil este que espera-se ser reproduzido por meio da construção de uma linhagem de Pichia pastoris com um padrão de glicosilação humanizado, por meio da deleção de dois genes envolvidos no sistema de glicosilação da levedura: alg3 e och1, responsáveis pela posterior hiper-manosilação característica desse organismo. Assim, a expressão da GCase será usada como modelo no desenvolvimento desta linhagem de Pichia pastoris que permita a expressão de glicoproteínas com um perfil humanizado específico de glicosilação. Além da produção da linhagem mutante pela técnica de CRISPR/Cas9, propomos a construção de duas linhagens controle: uma expressando a proteína GCase para análise do seu padrão selvagem de glicosilação em P. pastoris e outra expressando a proteína Cas9 de Streptoccocus pyogenes (SpCas9). A linhagem P. pastoris/GCase foi construída testando-se duas sequências sinal de secreção diferentes: fosfatase alcalina (PHO1) e albumina humana (Alb). Resultados de western blot mostraram a GCase no lisado celular e baixos níveis de proteína secretada no sobrenadante de cultura, sendo mais expresso na linhagem contendo a sequência PHO1. A linhagem P. pastoris/SpCas9 foi construída e a enzima SpCas9 foi detectada via western blot no lisado celular

após indução com metanol. Para a produção da linhagem com padrão de glicosilação humanizado propôs-se a deleção dos genes *alg3* e *och1* e a inserção, pela via de reparo por recombinação homóloga (HDR), de marcas de resistência aos antibióticos higromicina ou canamicina. Para tal, propusemos a construção de dois vetores finais de expressão do sistema CRISPR/Cas9 em P. pastoris, cada um contendo a enzima SpCas9 e os RNAs guia (gRNAs) para deleção do gene alg3 ou och1, e também a construção de dois fragmentos para HDR contendo o gene de resistência ao antibiótico flanqueado por regiões de 1Kb de homologia com a região de deleção do gene *alg3* ou *och1*. A construção dos vetores e fragmentos para HDR foram inicialmente feitas por meio de técnicas de clonagem clássica. No entanto, apesar de inúmeras tentativas, resultados de PCR e sequenciamento mostraram o insucesso das construções. Partiu-se então para a técnica de Gibson Assembly[®], através da qual os dois fragmentos para HDR foram construídos. Porém, os vetores de expressão contendo SpCas9 e os gRNAs ainda apresentam dificuldades na sua construção. Esforços ainda estão sendo feitos para a construção dos vetores e consequente tentativa de estabelecimento das linhagens mutantes. O sucesso no estabelecimento de um sistema de expressão de proteínas heterólogas com este padrão de glicosilação humano específico permitirá a obtenção e possível comercialização da GCase em sua forma terapêutica. Além disso, permitirá possíveis edições genômicas futuras para um padrão de maior complexidade de glicosilação humanizado, criando uma plataforma nacional para produção de outras glicoproteínas terapêuticas de interesse biotecnológico.

Palavras-chave: Pichia pastoris. Glucocerebrosidase. CRISPR-Cas9. Edição gênica.

ABSTRACT

Vitarelli, M.O. Specific humanization of *Pichia pastoris* glycosylation system with the CRISPR/Cas9 technique aiming the expression of human glycoproteins. 2016. 101p. Master Thesis – Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, São Paulo.

The production of therapeutic recombinant protein comprises complex and high valued molecules, including the glucocerebrosidase enzyme (GCase). Its deficiency results in Gaucher Disease, susceptible of treatment by enzymatic replacement therapy. The active form of recombinant GCase employed in therapy presents exposed terminal mannose residues in its glycosylation pattern. We hope to reproduce such pattern by constructing a Pichia pastoris strain with a specific human glycosylation pattern through the deletion of two genes involved in yeast glycosylation system, alg3 and och1, responsible for the final hyper-mannosylation characteristic of this organism. Therefore, the expression of GCase will be a case model for the development of the recombinant Pichia pastoris strain that could allow the expression of glycoproteins with a specific humanized glycosylation profile. Despite the establishment of the mutant strain using the CRISPR/Cas9 technique, we propose the construction of two control strains: one expressing the GCase protein for analysis of its wild type glycosylation pattern and another one expressing the Cas9 protein from Streptoccocus pyogenes (SpCas9). The P. pastoris/GCase strain was constructed testing two different secretion signal sequences: alkaline fosfatase (PHO1) and human albumin (Alb). Western blot results have shown GCase in cell lysate and in low expression levels in culture supernatant, being more expressed in the strain containing the PHO1 signal sequence. P. pastoris/SpCas9 strain was constructed and SpCas9 enzyme was detected via western blot in cell lysate after the induction with methanol. To produce the strain with the

humanized glycosylation pattern, the deletion of *alg3* and *och1* genes was proposed along with the insertion, by homology directed repair pathway (HDR), of hygromycin and kanamycin antibiotics resistance marks. In order to do so, we have proposed the construction of two final expression vectors of the CRISPR/Cas9 system in P. pastoris, each one containing SpCas9 enzyme and the guide RNAs (gRNAs) for deletion of *alg3* or *och1*, and also the construction of two fragments for HDR containing the antibiotics resistance gene flanked by 1Kb regions of homology with the deleted regions of alg3 or och1. Vectors and HDR fragments constructions were initially performed using classic cloning techniques. However, despite numerous tries, PCR and sequencing results have shown the failure of the constructions. Then, we moved on to the Gibson Assembly[®] technique, through which the two HDR fragments were built. Still, the expression vectors containing SpCas9 and the gRNAs presented difficulties in its assembly. Efforts continue to be made to successfully construct the remaining vectors and to establish the mutant lineage. Success in the establishment of a heterologous protein expression system with specific human glycosylation pattern will allow the obtainment and possible commercialization of the therapeutic form of GCase. Furthermore, it will also allow possible future genomic editing to a high complexity human glycosylation pattern, creating a national platform for the production of other therapeutic glycoproteins of biotechnological interest.

Keywords: Pichia pastoris. Glucocerebrosidase. CRISPR-Cas9. Genome editing.

LISTA DE ILUSTRAÇÕES E TABELAS

Figura 1 – Esquema representativo das principais vias de N-glicosilação em humanos e em P. pastoris	21
FIGURA 2 –MECANISMOS DE REPARO DA QUEBRA DA DUPLA FITA DO DNA	25
FIGURA 3 – MODELO REPRESENTATIVO DE CLIVAGEM DA DUPLA FITA DE DNA PELO SISTEMA CRISPR-CAS9	26
FIGURA 4 – ESQUEMA DO AUTO PROCESSAMENTO DAS RIBOZIMAS PARA A PRODUÇÃO DE RNAS GUIA MADUROS	28
Figura 5 – Processamento da glucocerebrosidase para obtenção da Cerezyme®	31
FIGURA 6 – DESENHO DO GENE SINTÉTICO GRNA_ALG3	49
FIGURA 7 - DESENHO DO GENE SINTÉTICO GRNA_OCH1	50
FIGURA 8 - DESENHO DO GENE SINTÉTICO HDR	53
FIGURA 9 – ESQUEMA DOS FRAGMENTOS A SEREM UTILIZADAS NO EVENTO DE RECOMBINAÇÃO HOMÓLOGA	53
FIGURA 10 - CONSTRUÇÃO DOS VETORES PGEM-TEASY/GCASEALB E PGEM-TEASY/GCASEPHO1	62
FIGURA 11 - CONSTRUÇÃO DOS VETORES PPIC3.5/GCASEALB E PHIL-S1/GCASEPHO1	63
FIGURA 12 - ANÁLISE DA EXPRESSÃO DE GCASEALB E GCASEPHO1 NO LISADO CELULAR DE <i>P. PASTORIS</i>	65
FIGURA 13 - ANÁLISE DA EXPRESSÃO DE GCASEALB E GCASEPHO1 NO SOBRENADANTE DE CULTURA DE <i>P. PASTORIS</i>	66
FIGURA 14 - CONSTRUÇÃO DO VETOR PPGEM-TEASY/SPCAS9CONT	67
FIGURA 15 - CONSTRUÇÃO DO VETOR PPICHOLI-1/SPCAS9CONT	68
FIGURA 16 - ANÁLISE DA EXPRESSÃO DE SPCAS9 NO LISADO CELULAR DE <i>P. PASTORIS</i>	69
FIGURA 17 - CONSTRUÇÃO DO VETOR PPGEM-TEASY/SPCAS9	70
FIGURA 18 - CONSTRUÇÃO DOS VETORES PMA-RQ/GRNA_ALG3/SPCAS9 E PMA-RQ/GRNA_OCH1/SPCAS9	71
FIGURA 19 - CONSTRUÇÃO DO VETOR PPIC6A/GRNA_OCH1/SPCAS9	71
FIGURA 20 - CONSTRUÇÃO DOS VETORES PGEM-TEASY/E3, PGEM-TEASY/D3, PGEM-TEASY/E1 E PGEM-TEASY/D1	72
FIGURA 21 - CONSTRUÇÃO DOS VETORES PGEM-TEASY/KAN E PGEM-TEASY/HYGRO	73
FIGURA 22 - CONSTRUÇÃO DOS VETORES PMK-RQ-BB/HDR/KAN E PMK-RQ-BB/HDR/HYGRO	74
FIGURA 23 - CONSTRUÇÃO DO VETOR PGEM-TEASY/E1/KAN/D1 – GIBSON ASSEMBLY	75
FIGURA 24 - CONSTRUÇÃO DO VETOR PMK-RQ-BB/HDR/HYGRO – GIBSON ASSEMBLY	76
FIGURA 25 - CONSTRUÇÃO DO VETOR PGEM-TEASY/E3/HYGRO/D3 - GIBSON ASSEMBLY	77
FIGURA 26 – ESQUEMA DE EDIÇÃO GÊNICA PROPOSTO PELO SISTEMA CRISPR-CAS9	82

TABELA 1 - LISTA DE ORGANISMOS UTILIZADOS	36
TABELA 2 - LISTA DE VETORES UTILIZADOS E CONSTRUÍDOS	38
TABELA 3 – DESENHO DA SEQUÊNCIA GUIA E NUCLEOTÍDEOS INICIAIS HHR	48

LISTA DE ABREVIATURAS E SIGLAS

Alb – Sequência sinal albumina humana	HDR – Reparo por Recombinação
Amp – Ampicilina	Homóloga
Blas – Blasticidina	HDV – Hepatitis Delta Virus
Cas9 – CRISPR-associated protein 9	HHR – Ribozyme Hammerhead
Cas9cont - CRISPR-associated protein 9	Hygro - Higromicina
controle	IgG – Imunoglubulina G
CHO – Células Chinese hamster ovary	Kan - Canamicina
CRISPR - Clustered Regularly Interspaced	Kb - KiloBases
Short Palindromic Repeats	kDa – KiloDalton
crRNA - CRISPR targeting RNA	Man – Manose
D1 – Braço de homologia direito <i>och1</i>	NHEJ – Reparo por ligação de pontas não
D3 – Braço de homologia direito <i>alg3</i>	homólogas
DO - Densidade Óptica	PAM - Protospacer Adjacent Motif
DSB – Double strand break	PARS – Sequência de Replicação Autônoma
E1 – Braço de homologia esquerdo <i>och1</i>	de Pichia pastoris
E3 – Braço de homologia esquerdo <i>alg3</i>	pb - Pares de Bases
ER – Retículo Endoplasmático	PCR - Polymerase Chain Reaction
FDA – Food and Drug Administration	PHO1 – Sequência sinal fosfatase alcalina
GCase – Glucocerebrosidase	PFU – Unidades Formadoras de Colônia
GD1/2/3 – Doença de Gaucher tipo 1, 2 ou 3	pol II – RNA polimerase II
Gen – Geneticina	SDS-PAGE - Sodium Dodecyl Sulfate
gRNA – guide RNA	Polyacrylamide Gel Electrophoresis

Sp - *Streptococcus pyogenes*

Zeo - Zeocina

TBS - Tris Buffered Saline

TRE – Terapia de Reposição Enzimática

U – Unidade

SUMÁRIO

1.	INT	ſRODUÇÃO	17
	1.1.	PRODUÇÃO DE PROTEÍNAS TERAPÊUTICAS EM LEVEDURAS	17
	1.2.	GLICOSILAÇÃO EM MAMÍFEROS E EM LEVEDURAS	19
	1.3.	SISTEMA CRISPR-CAS9	23
	1.4.	DOENÇA DE GAUCHER	29
	1.5.	TERAPIA DE REPOSIÇÃO ENZIMÁTICA PARA A DOENÇA DE GAUCHER	31
2.	OB	JETIVOS	35
3.	MA	TERIAIS E MÉTODOS	36
	3.1.	ORGANISMOS E LINHAGENS UTILIZADOS	36
	3.2.	MEIOS DE CULTURA	36
	3.3.	SOLUÇÕES	37
	3.4.	VETORES	38
	3.5.	OBTENÇÃO DE LINHAGEM CONTROLE P. PASTORIS EXPRESSANDO A PROTEÍNA GCASE	39
	3.5	.1. Construção dos vetores pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1	39
	3.5	.2. Desenho de oligonucleotídeos específicos para clonagem de GCaseAlb e GCasePHO1	40
	3.5	.3. Reação de PCR para amplificação dos genes GCaseAlb e GCasePHO1	40
	3.5	.4. Subclonagem no vetor pGEM-Teasy	41
	3.5	.5. Transformação de bactérias quimiocompetentes	41
	3.5	.6. PCR de colônia	42
	3.5	.7. Mini-preparações plasmidiais: pGEM-Teasy/GCaseAlb e pGEM-Teasy/GCasePHO1	42
	3.5	.8. Sequenciamento de DNA: pGEM-Teasy/GCaseAlb e pGEM-Teasy/GCasePHO1	42
	3.5	.9. Digestões enzimáticas	42
	3.5	.10. Reação de ligação: pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1	43
	3.5	.11. Precipitação de DNA com etanol	43
	3.5	.12. Transformação de bactérias eletrocompetentes	43
	3.5	.13. Seleção de clones recombinantes	44
	3.5	.14. Transformação de P. pastoris com vetores pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1.	44
	3.5	.15. Expressão de GCase em P. pastoris	44
	3.5	.16. Ensaio por western blotting	45
	3.6.	OBTENÇÃO DE LINHAGEM CONTROLE DE <i>P. PASTORIS</i> EXPRESSANDO A PROTEÍNA <i>SP</i> CAS9	
	(SPCA	AS9CONT)	45
	3.6	.1. Desenho de oligonucleotídeos específicos para clonagem de SpCas9cont	45
	3.6	.2. Construção do vetor pGEM-Teasy/SpCas9cont	46
	3.6	.3. Construção do vetor pPICHOLI-1/SpCas9cont	46
	3.6	.4. Transformação de P. pastoris com vetor pPICHOLI-1/SpCas9cont	47
	3.6	.5. Expressão de SpCas9 em P. pastoris pPICHOLI-1/SpCas9cont	47
	3.7.	OBTENÇÃO DE LINHAGEM <i>P. PASTORIS-SP</i> CAS9 COM DELEÇÃO DOS GENES <i>ALG3</i> E <i>OCH1</i> E	
	INTEC	RAÇÃO DAS MARCAS DE SELEÇÃO PELO REPARO POR RECOMBINAÇÃO HOMÓLOGA	48
	3.7	.1. Desenho dos RNAs guia (gRNAs) flanqueados pelas ribozimas HHR e HDV	48
	3.7	.2. Desenho dos genes sintéticos contendo sítio de clonagem para SpCas9 e os gRNAs	
	flai	nqueados por ribozimas	49
	3.7	.3. Construção do vetor pGEM-Teasy/SpCas9cont	50

	3.7.4.	Construção dos vetores pMA-RQ/gRNA_alg3/SpCas9 e pMA-RQ/gRNA_och1/SpCas9	50
	3.7.5.	Construção do vetor pPIC6A/gRNA_och1/SpCas9 – Via Clássica	51
	3.7.6.	Desenho dos braços de homologia	52
	3.7.7.	Desenho dos fragmentos para recombinação homóloga	52
	3.7.8.	Construção dos vetores pGEM-Teasy/E3, pGEM-Teasy/D3, pGEM-Teasy/E1 e pGEM-	
	Teasy/	D1	54
	<i>3.7.9</i> .	Construção dos vetores pGEM-Teasy/Kan e pGEM-Teasy/Hygro	55
	3.7.10.	Construção dos vetores pMK-RQ-Bb/HDR/Kan e pMK-RQ-Bb/HDR/Hygro	56
	3.7.11.	Construção do vetor pMK-RQ-Bb/HDR/Kan/E1	56
	3.7.12.	Construção do vetor pGEM-Teasy/E1/Kan/D1 – Gibson Assembly	57
	3.7.13.	Construção do vetor pGEM-Teasy/E3/Hygro/D3 – Gibson Assembly	59
	3.7.14.	Construção dos vetores pPIC6A/gRNA_och1/SpCas9 e pPICHOLI-1/gRNA_alg3/SpCas	s9 –
	Gibson	Assembly	60
1	DESII	TADOS	67
ч.	$\frac{11}{10}$	TADOS	02 62
	4.1. OB	Construção dos vatoras pPIC3 5/GCasa Alb a pHIL \$1/GCasa PHO1	02 62
	4.1.1.	Expressão da CCasa Alb a CCasa PHO1 am P. pastoris	02 61
	4.1.2.	Expression de OCUSERIO e OCUSERITOT em 1. pusioris	04
	4.2. OB	TENÇAO DE LINHAGEM CONTROLE DE I . <i>FASTORIS</i> EXPRESSANDO A PROTEINA CAS5	66
	(CAS9CO	Construção do votor pPICHOLI 1/SpCosQcont	00
	4.2.1.	Evenyassão da Cospeant am P. pastoris	00
	4.2.2.	Expression de Casaconi em 1 : pasionis	00
	4.J. UD	TENÇÃO DE LINHAGEM I . PASTORIS-CASO COM DELEÇÃO DOS GENES ALGO E OCHT E	60
		AO DAS MARCAS DE SELEÇÃO PELO REPARO POR RECOMBINAÇÃO HOMOLOGA	09 60
	4.5.1.	Construção dos vetores pMA-RQ/gRIVA_dig5/SpCas9 e pMA-RQ/gRIVA_ocn1/SpCas9	09
	4.3.2.	Construção dos vateros pCEM Taggy/E2, pCEM Taggy/D2, pCEM Taggy/E1 a pCEM	/ 1
	4.3.3. Toggau/	Construção dos vetores pGEM-Teasy/E5, pGEM-Teasy/D5, pGEM-Teasy/ET e pGEM-	72
	1 easy/1	Construção dos vistorios pMV PO Ph/UDP/Kan a pMV PO Ph/UDP/Uvero	/ 2
	4.3.4.	Construção dos vetores pMR-RQ-B0/HDR/Kan e pMR-RQ-B0/HDR/Hygro	7 4
	4.5.5.	Construção do vetor pGEM-Teasy/E1/Kan/D1 – Gloson Assembly	/ 4
	4.5.0.	Construção do velor pGEM-1easy/E5/Hygro/D5 – Gloson Assembly	/5
	4.3.7. Cibror	Construção dos vetores priCoA/gRNA_ocn1/spCas9 priChOLI-1/gRNA_dig5/spCas9 -	- 77
	Gloson	I Assembly	//
5.	. DISCU	SSÃO	78
6.	. CONC	LUSÕES	84
7.	REFE	RÊNCIAS BIBLIOGRÁFICAS	85
	NEVO 1		01
A	INEAU I -	LISTA DE L'KIMERS	91
A	NEXO 2 -	SEQUÊNCIAS gRNA E RIBOZIMAS	93
A	NEXO 3 -	OUTRAS TÉCNICAS DE EDIÇÃO GÊNICA	94
A	NEXO 4 -	SÚMULA CURRICULAR	97

1. Introdução

1.1. Produção de Proteínas Terapêuticas em Leveduras

As proteínas são macromoléculas essenciais que possuem diversas funções nos organismos vivos. Elas participam da composição de membranas biológicas, possuem funções regulatórias, de sinalização, transporte e replicação celular e são responsáveis pela catálise de reações metabólicas, desempenhando um papel de muita importância no organismo humano. Portanto, qualquer irregularidade na síntese proteica que cause a formação de um produto essencial defeituoso ou não funcional pode acarretar no desenvolvimento de inúmeras doenças que podem ser corrigidas pela simples administração clínica da proteína (Manning *et al.*, 1989). Devido à sua possível utilização no tratamento de diversas doenças, a produção de proteínas terapêuticas recombinantes compreende uma das maiores atividades produtivas das indústrias na área de medicina molecular atual, com estimativa de faturamento de U\$169 bilhões no ano de 2014 (Goodman, 2009; Kim *et al.*, 2014).

A obtenção de proteínas eucarióticas a partir da purificação direta das fontes naturais onde são encontradas é muitas vezes dispendiosa e laboriosa, além de apresentar alto risco de contaminação e baixo rendimento (Ma *et al.*, 2003). Com a introdução da engenharia genética nos anos 70 tornou-se possível a expressão heteróloga de proteínas em microrganismos, um marco para a expansão da indústria biotecnológica (Cohen *et al.*, 1973). Em 1980 foi licenciado para uso clínico pelo Food and Drug Administration (FDA) uma insulina recombinante obtida a partir de *Escherichia coli*, sendo a primeira proteína recombinante a entrar no mercado (Fda, 1982). Desde então este setor não parou de crescer, outras drogas recombinantes foram lançadas e novos sistemas de expressão foram desenvolvidos, possibilitando a introdução de linhagens de

microrganismos adaptados à produção proteica e a incorporação de leveduras e sistemas eucarióticos como "fábricas biológicas" (Ferrer-Miralles *et al.*, 2009).

A maioria das proteínas recombinantes de valor terapêutico, incluindo anticorpos, apresentam estruturas complexas e, por isso, são comumente produzidas em células de mamíferos. A complexidade dessas estruturas se deve às modificações pós-traducionais que são essenciais para a estrutura e função proteica, sendo a mais comum delas a glicosilação. Mas a baixa produção volumétrica (concentração do produto/tempo de cultura), a heterogeneidade do produto, o elevado custo de produção e o tempo gasto na produção de linhagens celulares estáveis são alguns dos inconvenientes desse modelo de expressão (Choi *et al.*, 2003).

A expressão heteróloga de proteínas em *E. coli* ainda é muito utilizada, uma vez que existem ferramentas bem desenvolvidas para sua manipulação genética e apresentam elevada densidade celular e rápido crescimento em cultura, exigindo baixo investimento (Martínez *et al.*, 2012). Porém, a produção de proteínas eucarióticas em sistemas procarióticos geralmente resulta na formação de corpos de inclusão e baixa taxa de rendimento específico (Panda, 2003). Além disso, a ineficiência da expressão de alguns códons eucarióticos e a ausência de modificações póstraducionais fazem com que esse sistema de expressão não seja o mais eficiente para a produção de proteínas glicosiladas, que compreendem a maior parte das proteínas terapêuticas (Bill, 2014).

As leveduras são organismos eucarióticos unicelulares que apresentam muitas vantagens na produção de proteínas complexas de interesse biofarmacêutico. Elas possuem um custo de produção moderado, são capazes de crescer em meio mínimo (ao contrário do meio complexo utilizado para células de mamíferos), podem ser cultivadas em biorreatores, o que possibilita a produção em larga escala, são menos sensíveis ao estresse no processo de produção devido à presença da parede celular e apresentam alto rendimento do produto no processo fermentativo, que é de apenas alguns dias (Redden *et al.*, 2014). Mesmo sendo um sistema eucariótico, a glicosilação das proteínas em levedura é diferente daquela realizada por células de mamíferos, o que a princípio pode ser um empecilho para a produção de proteínas terapêuticas, uma vez que seu padrão de glicosilação é imunogênico aos seres humanos. Mas, a partir de técnicas de engenharia genética, é possível criar linhagens de leveduras com o mesmo padrão de glicosilação humano, possibilitando a produção de proteínas humanas funcionais em escala industrial (Vervecken *et al.*, 2004; Abe *et al.*, 2009).

A primeira insulina recombinante humana produzida em *Saccharomyces cerevisiae* em 1987 impulsionou a utilização desse sistema de expressão, que hoje corresponde a 15% do mercado de proteínas terapêuticas (Goodman, 2009; Berlec and Strukelj, 2013). Em 2009 foi licenciado pelo FDA o primeiro biofármaco recombinante produzido na levedura *Pichia pastoris*, o Inibidor de Calicreína Kalbitor[®], produzido pela Shire. Em 2012, o medicamento Jetrea[®], utilizado para o tratamento de adesões vitreomaculares e produzido pela empresa Thrombogenics, também foi licenciado pelo FDA. Vários outros medicamentos estão na fase de testes clínicos (Walsh, 2010). A combinação de novas técnicas de engenharia genética com a facilidade de cultivo e produção em leveduras tem contribuído cada vez mais para a adoção desse sistema de expressão na produção de proteínas recombinantes.

1.2. Glicosilação em Mamíferos e em Leveduras

A glicosilação é a mais comum das modificações pós-traducionais existentes em organismos eucarióticos e possui papel fundamental no dobramento, estabilidade, atividade biológica, direcionamento celular e imunogenicidade das proteínas (Bailon and Won, 2009). A glicosilação ocorre no retículo endoplasmático (ER) e no complexo de Golgi e consiste na ligação covalente de carboidratos a um polipeptídio, geralmente catalisado por glicosiltransferases que utilizam

nucleotídeos ativados na forma NTP-açúcar como doador de substrato (Varki *et al.*, 2009). Resumidamente, os oligossacarídeos podem ser adicionados a resíduos de asparagina (Asn), formando os N-glicanos, ou a resíduos de treonina ou serina, formando O-glicanos. A Nglicosilação é a mais frequente, caracterizada por um padrão heterogêneo de formas estruturais, agrupadas em três tipos principais: alta manose, complexo e híbrido (Skropeta, 2009).

A N-glicosilação de proteínas em mamíferos e leveduras é iniciada no ER até a formação de um intermediário comum e, a partir da sua transferência para o Golgi, a via de N-glicosilação desses dois organismos começa a se diferenciar (Figura 1a). No ER, um complexo de oligosacariltransferases transfere o precursor Glc₃Man₉GlcNAc₂ para o resíduo de Asn presente na cadeia polipeptídica crescente que vai sofrer diferentes reações até originar o intermediário comum Man₈GlcNAc₂ (Man8) que é então transferido para o Golgi. Em leveduras, mais resíduos de manose são adicionados à cadeia, resultando em estruturas hiper-manosiladas, enquanto que em mamíferos, Man8 é processada até Man₅GlcNAc₂ e a cadeia pode ser estendida para estruturas híbridas ou complexas com resíduos de N-acetil-glicosamina, fucose, galactose e ácido siálico (Vervecken *et al.*, 2004; Varki *et al.*, 2009; Loos and Steinkellner, 2012).



Figura 1 – (a) Esquema representativo das principais vias de N-glicosilação em humanos (esquerda) e em leveduras (direita). Um intermediário comum (Man8GlcNAc2) é formado no retículo endoplasmático (ER) e, após seu transporte para o Golgi, as vias se diferem. Em humanos, são formadas estruturas híbridas ou complexas com resíduos de GlcNAc, galactose e ácido siálico. Em leveduras, a via resulta em estruturas com perfil de glicosilação GlcNAc, N-acetil-glucosamina, Gal, Galactose; NANA, ácido Nhiper-manosilados. Man, Manose; acetilneuramínico, Fuc, fucose; Mns, Manosidase; N-acetil-glicosaminiltransferase; GnT, GalT, Galactosiltransferase; ST, Sialiltransferase; MnT, Manosiltransferase. (b) Mutantes sintéticos de P. pastoris com deleção nos genes alg3 e och1 e inserção de enzimas da via humana podem ser usados para recriar o sistema de Nglicosilação humano, levando à formação de estruturas complexas em levedura. (c) Estrutura da Cerezyme®, glucocerebrosidase (GCase) comercial recombinante com resíduos expostos de manose. Os asteriscos em vermelho mostram as formas de glicosilação da GCase, semelhantes à da enzima comercial, que se pretende obter ao longo do desenvolvimento desse projeto. Ressalta-se que a fucose ligada a N-acetil-glucosamina em (c) pode ou não estar presente neste core, como mostrado em (Tekoah et al., 2013). (d) GCase de interesse, que se deseja a princípio ser obtida pelo Instituto Butantan. Adaptado de Genzyme (Corporation, 2001); (Wildt and Gerngross, 2005).

A estratégia utilizada para humanizar o sistema de glicosilação de leveduras envolve a eliminação das reações de glicosilação responsáveis pela hiper-manosilação, seguida da introdução das enzimas que perfazem a via de N-glicosilação humana (Figura 1b). Em leveduras, a primeira transferência de um resíduo de manose que ocorre no lúmen do retículo endoplasmático à cadeia é catalisada por uma manosiltransferase codificada pelo gene alg 3, essa enzima garante o substrato para adições subsequentes de manose (Davidson et al., 2004). Já a adição lateral de manoses ao núcleo Man8 em leveduras, inicia-se com a ação da enzima α -1,6manosiltransferase, codificada pelo gene och 1 (Wildt and Gerngross, 2005). Portanto, a deleção desses dois genes impede a formação de estruturas com alto grau de manosilação, originando um estrutura do tipo Man₅GlcNAc₂ (Figura 1b, indicado por *), semelhante à presente na via de glicosilação humana (Bobrowicz et al., 2004). Em seguida, a adição de uma α-1,2-manosidase para gerar o intermediário Man₃GlcNAc₂ (Figura 1b, indicado por **) se faz necessária para a produção de glicoproteínas híbridas e complexas, pois é ele que vai servir de substrato para as enzimas subsequentes: N-acetilglucosaminiltransferase I, N-acetilglucosaminiltransferase II, galactosiltransferase e sialiltransferase, juntamente com a adição dos seus respectivos NTPaçúcares e seus transportadores, dependendo do padrão de glicosilação da proteína de interesse (Hamilton et al., 2006). É importante destacar que basicamente a estrutura Man₃GlcNAc₂ é aquela presente predominantemente na enzima glucocerebosidase comercial (*Cerezyme*[®]) (Figura 1c) (Veja itens 1.4 e 1.5 abaixo). Na verdade, a enzima terapêutica apresenta uma mistura de formas glicosiladas, sendo a forma mais abundante, aquela com 3 resíduos de manose (Figura 1c). No entanto, as formas ativas que melhor se ligam-se aos receptores de manose nos macrófagos podem ter de 3-5 resíduos de manose, como as formas indicadas com * e ** na Figura 1b (Van Patten et al., 2007; Tekoah et al., 2013). Assim, a dupla deleção dos genes alg 3 e och 1 seria suficiente para a produção da glucocerebrosidase com 5 manoses terminais e a inserção e expressão da enzima α -1,2-manosidase a seguir, resultaria em uma produção da glucocerebrosidase com 3 manoses terminais, muito semelhante ao produto comercial *Cerezyme*[®]. É possível que somente a deleção dos genes *alg 3* e *och 1* já seja suficiente para a produção da glucocerebrosidase clinicamente funcional, sem a necessidade da expressão heteróloga da enzima α -1,2-manosidase.

1.3. Sistema CRISPR-Cas9

A tecnologia de edição de genomas foi revolucionada nos últimos anos com o desenvolvimento do sistema CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) Cas9, que permite de forma bem simples, a inserção, deleção, substituição e regulação gênica (Ver outras técnicas no Anexo 3) (Mali *et al.*, 2013; Hsu *et al.*, 2014). CRISPR-Cas codifica um sistema de imunidade adaptativa que defende procariotos contra infecções virais e outros plasmídios. A imunidade é mediada por nucleases Cas (CRISPR-associated) que usam RNAs guias, conhecidos como crRNAs (CRISPR *targeting* RNA), para direcionar o sítio de clivagem no genoma do ácido nucleico invasor (Charpentier and Marraffini, 2014; Gersbach and Perez-Pinera, 2014). No sistema CRISPR-Cas do tipo II, o mais simples dos três tipos conhecidos, a atividade de clivagem do DNA é realizada por uma única enzima, Cas9, guiada por um duplex de RNA, que permite que as células bacterianas reconheçam determinadas sequências de um genoma viral e se protejam por meio do reconhecimento e clivagem dessas sequências no caso de uma nova infecção (Cong *et al.*, 2013; Esvelt *et al.*, 2014).

Devido a sua alta especificidade e por requerer apenas uma pequena sequência de RNA para direcionar a clivagem via Cas9:crRNA, surgiu a hipótese de que seria possível programar a clivagem de Cas9 para um alvo específico (Jinek *et al.*, 2012). De fato, a expressão de um crRNA quimérico, também denominado RNA guia (gRNA), contendo uma sequência inicial de 20 pares

de base idêntica à sequência alvo, permitiu a reconstituição funcional da unidade Cas9-crRNA e possibilitou a edição de genomas de diversos organismos eucarióticos (Sampson and Weiss, 2013). O primeiro trabalho utilizando CRISPR-Cas9 realizado em leveduras, utilizando o organismo *S. cerevisiae* como modelo foi descrito em 2013 (Dicarlo *et al.*, 2013) e o único trabalho descrevendo esse sistema em *P. pastoris* foi publicado recentemente por (Weninger *et al.*, 2016).

A clivagem pela Cas9 pode ativar dois mecanismos de reparo da dupla fita de DNA: ligação de pontas não homólogas (NHEJ) e recombinação homóloga (HDR) (Hou *et al.*, 2013). No reparo por NHEJ, as pontas do DNA são processadas por uma maquinaria de reparo endógena e são então religadas, o que pode gerar mutações pontuais aleatórias no sítio da junção. Caso essas mutações ocorram numa região codificadora do gene, pode resultar em um *frameshift* ou na criação de um *stop codon* prematuro, ocasionando o *knock-out* do gene. De forma alternativa e concomitantemente à clivagem do DNA, é possível introduzir uma construção de DNA flanqueada pelas sequências adjacentes no reparo por HDR, permitindo uma edição gênica precisa e de alta fidelidade (Ran *et al.*, 2013) (Figura 2).



Specific change introduced to the genomic DNA

Figura 2 – (a) Mecanismos de reparo da quebra da dupla fita do DNA (DSB): reparo por ligação de pontas não homólogas (NHEJ), que gera mutações pontuais aleatórias no sítio da junção, ou reparo por recombinação homóloga (HDR), que permite a inserção de um fragmento flanqueado por sequências adjacentes ao local da quebra (HDR repair template), como mostrado na Figura (b). (Addgene CRISPR guide, 2014 e (Ran *et al.*, 2013), respectivamente)

Para que ocorra o reconhecimento e a clivagem é necessário que haja complementariedade de bases entre a sequência alvo e o crRNA, assim como a presença de uma sequência PAM (*Protospacer Adjacent Motif*) na extremidade 3' da sequência alvo (Figura 3) (Pyzocha *et al.*, 2014). O motivo PAM é um componente essencial do sistema CRISPR, pois permite o direcionamento correto da clivagem e ainda atua como um sistema de auto reconhecimento,

impedindo que o próprio DNA bacteriano seja reconhecido como alvo (Mali *et al.*, 2013). Os requerimentos para o sítio PAM variam de acordo com os diferentes organismos procariotos, sendo o mais comum dos sistemas engenheirados baseados na sequência PAM de *Streptococcus pyogenes*, que é constituída da sequência NGG, sendo N qualquer nucleotídeo (Ran *et al.*, 2013).



Figura 3 – Modelo representativo de clivagem da dupla fita de DNA pelo sistema CRISPR-Cas9. A enzima Cas9 (em amarelo) interage com a sequência alvo no genoma (em azul) com a ajuda do RNA guia apresentando 20 pares de bases, complementar à sequência alvo. A sequência alvo é reconhecida e direcionada para clivagem por Cas9 por meio da sequência PAM (NGG) (Pyzocha *et al.*, 2014).

Uma sequência correta do gRNA é essencial para o reconhecimento e para a clivagem do DNA, fazendo com que qualquer adição de bases ou outra modificação nas regiões 5' ou 3' impeça que essa construção guie a quebra do DNA pela enzima Cas9 (Ryan and Cate, 2014). Promotores de RNA polimerase II (pol II) normalmente não são utilizados para expressão do sistema CRISPR-Cas9, uma vez que passam por um extenso processamento e modificação das extremidades (Jacobs *et al.*, 2014). Esse problema pode ser contornado pela adição de ribozimas que sofrem auto-clivagem flanqueando o gRNA: a ribozima Hammerhead (HHR) na região 5' do gRNA e a ribozima Hepatitis Delta Virus (HDV) na região 3'. Após sofrerem auto clivagem, o gRNA maduro é liberado sem a adição de nenhum nucleotídeo ou modificação em sua estrutura, possibilitando o seu reconhecimento pela enzima Cas9 para formação do complexo Cas9-gRNA (Figura 4) (Gao and Zhao, 2014).

O processo de edição gênica envolve então: a escolha da sequência alvo que se deseja alterar, o desenho dos gRNAs e a introdução da enzima Cas9, dos RNAs guias e do fragmento para inserção por recombinação homóloga, caso seja de interesse.



Figura 4 – Esquema do auto processamento da fita de RNA para a produção de RNAs guia (gRNA) maduros. Estrutura do pré-gRNA contendo a ribozima Hammerhead (HHR) na região 5' (em verde), seguida da sequência guia específica contendo 20pb do gRNA (em azul) e da sequência do esqueleto do gRNA (em vermelho), e finalmente a ribozima Hepatitis Delta Virus (HDV) na região 3' (em laranja). As regiões de hairpin da HHr são denominadas H1, H2 e H3, e P1, P2, P3 e P4 para as mesmas regiões da HDV. A região 5' da sequência guia é complementar à região 3' da HHR. O pré-gRNA sofre auto processamento liberando o gRNA maduro. Adaptado de (Gao and Zhao, 2014).

1.4. Doença de Gaucher

A doença de Gaucher é uma desordem metabólica autossômica recessiva, sendo a principal doença de depósito lisossômica dentre outras 50 conhecidas e a primeira a ser descrita, em 1882, por Philippe Gaucher (Bennett and Mohan, 2013; Ferraz *et al.*, 2014). A doença é causada por um erro genético do metabolismo que torna defeituosa a enzima glucocerebrosidase (GCase), uma enzima associada à membrana lisossomal, responsável pela degradação do glicolipídeo glucocerebrosídeo em glicose e ceramida (Aerts *et al.*, 1985). Os defeitos na função da GCase resultam de mais de 350 mutações diferentes, incluindo mutações pontuais, inserções, deleções e rearranjos (Grabowski *et al.*, 2001) no gene GBA1, que possui 7,6Kb, 11 exons e está localizado na região 21q do cromossomo 1 (Rosenbloom and Weinreb, 2013).

A sequência nucleotídica completa do cDNA da GCase humana foi descrita por (Sorge *et al.*, 1985; Tsuji *et al.*, 1986), em trabalhos simultaneamente desenvolvidos. A proteína GCase madura contém 497 aminoácidos e 5 potenciais sítios de N-glicosilação, sendo 4 deles normalmente ocupados (Grace *et al.*, 1994). A GCase nativa, na sua forma ativa, encontra-se na forma monomérica, com massa molecular variando entre 59 e 69 kDa, de acordo com o padrão de glicosilação (Bergmann and Grabowski, 1989). Cadeias de carboidratos do tipo rico em manose são inicialmente adicionadas a resíduos de Asn durante a síntese da proteína no retículo endoplasmático. Estas cadeias são modificadas e unidades oligossacarídicas são constantemente adicionadas e removidas durante o processamento da proteína no complexo de Golgi. Uma vez nos lisossomos, a GCase sofre a ação de exoglicosidases que formam cadeias oligossacarídicas apenas do tipo complexo, não terminadas em manose, resultando na GCase madura de aproximadamente 66-69 kDa (Fabrega *et al.*, 2000).

Quando a atividade da GCase é deficiente, os glucocerebrosídeos acumulam-se em macrófagos, especialmente no figado, baço e medula óssea, mas também podem ser encontrados

nos pulmões, intestino, rim, coração e, em casos mais raros, no tecido cerebral (Connock *et al.*, 2006). O acúmulo progressivo de glucocerebrosídeos nos macrófagos, que nessa condição são também conhecidos como células de Gaucher, compromete o funcionamento normal dos órgãos, e, além disso, o acúmulo em excesso de substrato pode levar à ruptura dessas células e ao aumento da produção de citocinas inflamatórias e quimiocinas, que por sua vez podem causar fibrose e outros danos teciduais (Grabowski, 2012).

A redução parcial ou a perda total da atividade enzimática resulta em uma ampla gama de sintomas que caracterizam as três maiores formas da doença de Gaucher, que são classificadas de acordo com a presença ou ausência de danos neurológicos (Germain, 2004).

A doença de Gaucher tipo 1 (GD1) ocorre principalmente em adultos e corresponde a cerca de 95% dos casos da doença. A GD1 não apresenta envolvimento do sistema nervoso central e possui uma variedade de manifestações clínicas, sendo as mais comuns: aumento exagerado do figado e baço, anemia, trombocitopenia e leucopenia, podendo haver lesões ósseas progressivas e variáveis (Martins *et al.*, 2009; Bennett and Mohan, 2013).

Já as doenças dos tipos 2 e 3 (GD2 e GD3, respectivamente) são variantes neuropáticas. A GD2 é a forma mais rara e grave da doença, caracteriza-se por deterioração neurológica progressiva e rápida desde os primeiros meses de vida, e morte até os 2 anos de idade. A GD3 também é uma forma rara da doença e é caracterizada como uma neuropatia crônica. Apresenta um quadro visceral semelhante à GD1 e envolvimento neurológico variável, com pacientes chegando à idade adulta (Futerman *et al.*, 2004).

1.5. Terapia de Reposição Enzimática para a Doença de Gaucher

O tratamento mais comum e melhor caracterizado para a doença de Gaucher é a terapia de reposição enzimática (TRE), no qual a GCase defeituosa é suplementada com a enzima ativa (Kacher *et al.*, 2008). O conceito de TRE para o tratamento de doenças de depósito lisossômico foi primeiramente sugerido por (De Duve, 1964), e, a fim de explorar essa nova possibilidade, a enzima GCase proveniente da placenta humana foi parcialmente purificada por Pentchev *et al.* 1973, e administrada sem sucesso em pacientes por Brady *et al.* 1974. Em 1978, descobriu-se que a GCase humana não entrava nos macrófagos de forma eficiente devido à ausência de resíduos expostos de manose em sua estrutura para se ligar aos receptores de manose presentes na superfície dos macrófagos, justificando o fracasso da terapia inicial (Stahl *et al.*, 1978). Em uma tentativa de expor os resíduos de manose, foi feita a remoção sequencial de resíduos de ácido siálico, galactose e N-acetil-glicosamina (Figura 5), o que se revelou eficiente para aumentar em 50 vezes a ligação da GCase modificada aos macrófagos (Furbish *et al.*, 1981; Brady and Furbish, 1982).



Figura 5 – Modificação enzimática da glucocerebrosidase (G) para criar a Cerezyme®. Man, Manose; GlucNAc, N-acetilglucosamina; Gal, Galactose; NeuNAc, Ácido N-acetilneuramínico. Adaptado de Genzyme Corporation, 2001.

Por meio da modificação dos açúcares terminais, três produtos de TRE foram desenvolvidos e estão em uso atualmente para a doença de Gaucher: Imiglucerase, Taliglucerase-alfa e Velaglucerase alfa (Tekoah *et al.*, 2013).

A Imiglucerase ou *Cerezyme*® (*Genzyme Corporation*) é uma GCase produzida em células de mamíferos da linhagem CHO (*Chinese-Hamster-Ovary*) modificadas pós-produção pela ação de três exoglicosidases: neuraminidase, β -galactosidase e β -N-acetilglucosaminidase. A GCase recombinante difere da placentária em um aminoácido na posição 495, onde uma arginina foi substituída por uma histidina, para resultar em uma enzima funcional de maior estabilidade, sendo aprovada pelo FDA em 1994 e tornando-se o medicamento de referência mundial. O padrão de glicosilação predominante pode ser visto na Figura 1c (Corporation, 2001; Bennett and Mohan, 2013; Tekoah *et al.*, 2013).

A empresa *Shire Human Genetic Therapies* produziu a GCase gene-ativada Velaglucerase alfa, produzida em fibroblastos humanos que utiliza um inibidor da manosidase I para obter o perfil de glicosilação desejado. A inibição da maturação normal do glicano produz predominantemente oligossacarídeos com alto grau de manose. Foi aprovada em 2010 pelo FDA e possui sequência de aminoácidos idêntica à da GCase humana (Brumshtein *et al.*, 2010; Bennett and Mohan, 2013; Tekoah *et al.*, 2013)

Já a Taliglucerase alfa (*Protalix Biotherapeutics*), aprovada em 2012 pelo FDA, é produzida utilizando um sistema de expressão de proteínas recombinantes em plantas *ProCellEx*®. Essa plataforma utiliza células de cenoura geneticamente modificadas, direcionando exoglicosidases para o vacúolo, onde os resíduos terminais dos glicanos são removidos (Bennett and Mohan, 2013; Tekoah *et al.*, 2013). Apresenta sequência de aminoácidos semelhante à da Imiglucerase, porém apresenta 2 aminoácidos adicionais na extremidade N-terminal e 7 na extremidade C-

terminal, correspondentes ao linker para fusão do peptídeo sinal e à sequência de direcionamento ao vacúolo, respectivamente (Shalltiel *et al.*, 2007).

De acordo com (Van Patten *et al.*, 2007), a diferença nos tamanho das cadeias de manose presentes nas enzimas comerciais, que variam de 2 a 9 resíduos, não influência na captação da GCase nem na atividade da enzima recombinante pelos macrófagos.

O tratamento com a TRE consiste em infusões intravenosas permanentes da enzima GCase, geralmente durante 1 a 2 horas, a cada 14 dias, em regime ambulatorial e sem necessidade de internação. A extensão da resposta clínica à terapia pode variar de acordo com o regime usado, a gravidade e a taxa de progressão da doença. Por isso, as doses devem ser individualizadas e os doentes necessitam de avaliação periódica (Barranger and O'rourke, 2001). A maioria dos pacientes inicia o tratamento com doses que variam entre 30 e 60 unidades da enzima por quilograma de massa corpórea (U/kg). Devido ao alto custo do medicamento por paciente, custo de cada dose - Cerezyme® U\$432,978 (Novo *et al.*, 2012), a determinação da menor dose efetiva, inicial e de manutenção no tratamento, é imprescindível.

Segundo o Ministério da Saúde, existem no Brasil atualmente, cerca de 610 pacientes de Gaucher sob tratamento da enzima GCase, sendo 170 apenas no Estado de São Paulo. Os pacientes estão cadastrados no Componente Especializado da Assistência Farmacêutica e fazem uso do medicamento *Cerezyme*®, importado pelo Ministério da Saúde.

Por tratar-se de um medicamento extremamente caro e por ser considerado essencial à melhoria na qualidade de vida dos portadores sintomáticos, o *Cerezyme*® foi incluído na lista de Medicamentos Excepcionais do SUS. Assim, os pacientes têm direito por lei ao tratamento gratuito. O Ministério da Saúde disponibiliza o medicamento aos estados, por meio de aquisição centralizada, seguindo a Portaria GM nº 2.577/06. Segundo dados de 2010, o consumo médio anual é de 12.602 frascos, que são importados ao custo unitário de U\$555,00, o que representa

um gasto de aproximadamente U\$84 milhões ao ano apenas com a compra do medicamento (Novo *et al.*, 2012).

Apesar da eficiência comprovada no tratamento dos doentes de Gaucher, o alto custo da terapia, que é subsidiada pelo sistema público no Brasil, destaca a importância econômica deste tratamento, ainda que a doença seja de baixa incidência.

2. Objetivos

Esse projeto de mestrado visa o estabelecimento de uma linhagem de *Pichia pastoris* com um padrão específico de glicosilação humanizado por meio da técnica de edição de genomas CRISPR-Cas9 e, futuramente, caso seja bem sucedido, a obtenção de uma nova enzima glucocerebrosidase (GCase) recombinante purificada, com resíduos de manose expostos, apresentando potencial para comercialização. O projeto possui os seguintes objetivos específicos:

- Obtenção de linhagem selvagem de *P. pastoris* expressando a proteína GCase como controle (*P. pastoris*/GCaseAlb e *P. pastoris*/GCasePHO1);
- Obtenção de linhagem *P. pastoris* expressando a proteína Cas9 de *Streptoccocus* pyogenes também como controle (*P. pastoris/Sp*Cas9);
- Obtenção de linhagem *P. pastoris* com deleção dos genes *alg3* e *och1* e integração das marcas de resistência aos antibióticos canamicina e higromicina pelo reparo por recombinação homóloga (*P. pastoris/Δalg3/Δoch1/*Hygro⁺/Kan⁺).
3. Materiais e Métodos

3.1. Organismos e linhagens utilizados

Tabela	1 -	Lista	de	organismos	utilizados
--------	-----	-------	----	------------	------------

Linhagem	Genótipo	Empresa
<i>E. coli</i> DH5α	F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK-mK+), λ-	Invitrogen
<i>E. coli</i> DH10B	F– endA1 deoR+ recA1 galE15 galK16 nupG rpsL Δ(lac)X74 φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) StrR λ–	Invitrogen
<i>E. coli</i> TOP10F'	F' {lacIq, Tn10(TetR)} mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG	Life Technologies
P. pastoris GS115	his4	Life Technologies

3.2. Meios de cultura

- Meio LB (Luria-Bertani): triptona 1% (m/v), extrato de levedura 0,5% (m/v) e NaCl 1% (m/v); para meio sólido adicionar 1,5% (m/v) de ágar.
 - LB/Amp/IPTG/Xgal: Meio LB acrescido de ampicilina 100µg/mL, IPTG 0,1M e Xgal 40µg/mL.
 - LB/Amp: Meio LB acrescido de ampicilina 100µg/mL
 - LB/Kan: Meio LB acrescido de canamicina 50µg/mL
 - LB/Hygro: Meio LB acrescido de higromicina 100µg/mL
 - LB/Zeo: Meio LB acrescido de zeocina 25µg/mL
- Meio LB low salt (Luria-Bertani): triptona 1% (m/v), extrato de levedura 0,5% (m/v) e NaCl 0,5% (m/v); para meio sólido adicionar 1,5% (m/v) de ágar.
 - LB/Blas: Meio LB low salt acrescido de blasticidina 75µg/mL

- Meio SOC (Super Optimal broth with Catabolite repression) : triptona 1% (m/v), extrato de levedura 0,5% (m/v), NaCl 5M 0,2% (v/v) e KCl 1M 0,25% (v/v).
- Meio MD (Minimal Dextrose) e MDH (Minimal Dextrose + Histidine): yeast nitrogen base with ammonium sulfate without amino acids (YNB) 1,34% (v/v), biotina 4x10⁻⁵% (v/v) e dextrose 2% (v/v); para MDH adicionar histidina 4x10⁻³% (v/v). Para meio sólido adicionar 1,5% (m/v) de ágar.
- Meio BMMY (Buffered Methanol-complex Medium): extrato de levedura 1% (m/v), peptona 2% (m/v), 100mM fosfato de potássio pH6,0, YNB 1,34% (v/v), biotina 4x10⁻⁵ % (v/v), metanol 0,5% (v/v).
- Meio YPD (Yeast Extract Peptone Dextrose): extrato de levedura 1% (m/v), peptona 2% (m/v), dextrose 2% (v/v); para meio sólido adicionar 2% (m/v) de ágar.
 - YPD/Zeo: Meio YPD acrescido de zeocina 100 µg/mL
 - YPD/Blas: Meio YPD acrescido de blasticidina 300 µg/mL
 - YPD/Gen Meio YPD acrescido de geneticina 250 µg/mL

3.3. Soluções

- Soluções para SDS-PAGE:
 - Tampão de amostra para SDS-PAGE 5X: 250 mM Tris-HCl pH 6,8, SDS 10% (m/v), azul de bromofenol 0,5% (m/v), glicerol 50% (v/v) e 1M β-mercaptoetanol (v/v).
 - Tampão Tris-glicina 5X: Tris base 1,5% (m/v), glicina 9,4% (m/v) e SDS 0,5% (m/v).
 - Solução corante de SDS-PAGE: Comassie brillant blue 0,25% (m/v), etanol 40% (v/v) e ácido acético 10% (v/v).
 - Solução descorante de SDS-PAGE: etanol 30% (v/v) e ácido acético 10% (v/v).
- Soluções para western blot:

- Solução TBS-T 0,1%: 100 mM de Tris-HCl pH 8,0, 150 mM de NaCl, 0,1% de Tween 20 (v/v).
- Tampão de transferência para western blot: tampão Tris-glicina 5X, 20% de etanol.
- Solução de bloqueio: solução TBS-T 0,1% e 5% de leite em pó desnatado.
- Soluções para gel de agarose:
 - Tampão TAE10X: 40 mM Tris Base, 1 mM EDTA e 20 mM ácido acético.

3.4. Vetores

Tabela 2 - Lista de vetores utilizados e construídos

Vetor	Organismo	Marca de seleção em	Marca de seleção em P nastoris	Sequência sinal de	Promotor	Empresa	Sucesso na construção	Tipo de clonagem
pGEM-Teasy	E. coli	Amp/IPTG/ Xgal	- -	-	-	Promega	-	-
pGEM-Teasy/ GCaseAlb	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	~	Clássica
pGEM-Teasy/ GCasePHO1	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	~	Clássica
pGEM-Teasy/ SpCas9cont	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	~	Clássica
pGEM-Teasy/ SpCas9	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	~	Clássica
pGEM-Teasy/ Kan	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	~	Clássica
pGEM-Teasy/ Hygro	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	~	Clássica
pGEM-Teasy/ E1	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	~	Clássica
pGEM-Teasy/ D1	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	~	Clássica
pGEM-Teasy/ E3	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	~	Clássica
pGEM-Teasy/ D3	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	~	Clássica
pGEM-Teasy/ E1_Kan_D1	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	\checkmark	Gibson Assembly
pGEM-Teasy/ E3_Hygro_D3	E. coli	Amp/IPTG/ Xgal	-	-	-	Presente trabalho	\checkmark	Gibson Assembly
pH7WG2.0	E. coli	Hygro	-	-	-	Novo <i>et al.</i> , 2012	-	-
pENTR/D-TOPO	E. coli	Kan	-	-	-	Novo <i>et al.</i> , 2012	-	-
pPIC3.5	E. coli / P. pastoris	Amp	His	-	AOX1	Life Technologies	-	-
pPIC3.5/GCaseAl b	E. coli / P. pastoris	Amp	His	Alb	AOX1	Presente trabalho	~	Clássica
pHIL-S1	E. coli / P. pastoris	Amp	His	PHO1	AOX1	Life Technologies	-	-
pHIL-	E. coli / P.	Amp	His	PHO1	AOX1	Presente	✓	Clássica

S1/GCasePHO1	pastoris					trabalho		
pPICHOLI-1	E. coli / P. pastoris	Zeo	Zeo	-	AOX1	MobiTech	-	-
pPICHOLI-1/ SpCas9cont	Ē. coli / P. pastoris	Zeo	Zeo	-	AOX1	Presente trabalho	\checkmark	Clássica
pPICHOLI-1/ gRNA_ <i>alg3/</i> <i>Sp</i> Cas9	E. coli / P. pastoris	Zeo	Zeo	-	TEF1	Presente trabalho	×	Gibson Assembly
pPIC6A	E. coli / P. pastoris	Blas	Blas	-	AOX1	Life Technologies	-	-
pPIC6A / gRNA_och1/ SpCas9	E. coli / P. pastoris	Blas	Blas	-	TEF1	Presente trabalho	×	Clássica/ Gibson Assembly
pMA- RQ/gRNA <i>alg3</i>	E. coli	Amp	-	-	-	GeneArt	-	-
pMA-RQ/ gRNA <i>alg3/</i> <i>Sp</i> Cas9	E. coli	Amp	-	-	-	Presente trabalho	~	Clássica
pMA- RQ/gRNA och1	E. coli	Amp	-	-	-	GeneArt	-	-
pMA-RQ/ gRNA_och1/ SpCas9	E. coli	Amp	-	-	-	Presente trabalho	×	Clássica
pMK-RQ- Bb/HDR	E. coli	Kan	-	-	-	GeneArt	-	-
pMK-RQ- Bb/HDR/Kan	E. coli	Kan	-	-	-	Presente trabalho	~	Clássica
pMK-RQ- Bb/HDR/ Hygro	E. coli	Kan	-	-	-	Presente trabalho	1	Clássica/ Gibson Assembly
p <i>Sp</i> Cas9(BB)-2A- GFP (PX458)	E.coli	Amp	-	-	-	Addgene	-	-

3.5. Obtenção de linhagem controle P. pastoris expressando a proteína GCase

3.5.1. Construção dos vetores pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1

Duas estratégias para clonagem da proteína GCase humana por *P. pastoris* foram usadas para garantir maior expressão e secreção da mesma no sobrenadante de cultura de *P. pastoris*. O vetor pHIL-S1 possui a sequência sinal da fosfatase alcalina (PHO1), e, para tentar obter uma maior eficiência na secreção, uma sequência sinal de albumina humana (Alb) (GenBank: M12523.1) foi adicionada à extremidade 5' do gene que codifica a proteína GCase no vetor pPIC3.5. Essa estratégia foi adotada devido à alta secreção de albumina (>1g/L) em linhagens controle de *P. pastoris* contendo o gene da soro albumina humana clonada com a sua sequência sinal, de acordo com informações obtidas no manual do fabricante. A GCase contendo a sequência sinal de

albumina humana será referida neste trabalho como GCaseAlb e a contendo a sequência sinal PHO1, como GCasePHO1.

3.5.2. Desenho de oligonucleotídeos específicos para clonagem de GCaseAlb e GCasePHO1

Oligonucleotídeos específicos foram desenhados para amplificar o gene da GCase humana madura (GenBank: M16328.1) do plasmídeo pH7WG2.0 (Novo *et al.*, 2012) através de reação da polimerase em cadeia (PCR).

Os oligonucleotídeos 1 e 2 (ver Anexo 1) foram desenhados para ligação da GCaseAlb no plasmídeo pPIC3.5 e contêm a sequência sinal Alb e apresenta a sequência consenso de leveduras, inserida junto ao códon de iniciação da tradução ATG: (A/Y)A(A/T)A<u>ATG</u>, semelhante à sequência Kosak de mamíferos (Romanos *et al.*, 1992). Os oligonucleotídeos 3 e 4 (ver Anexo 1) foram desenhados para ligação da GCase no plasmídeo pHIL-S1, e, para inserção em fase com o sítio de clivagem da sequência PHO1, dois aminoácidos extras foram adicionados, inserindo um códon que codifica para o aminoácido glicina.

3.5.3. Reação de PCR para amplificação dos genes GCaseAlb e GCasePHO1

Para amplificação da GCaseAlb e GCasePHO1 por PCR, utilizou-se cerca de 100ng do plasmídeo pH7WG2.0/GCase, 2μL da enzima *PFU* (produzida pelo próprio laboratório), 1nM dNTP, 1,5pmol do primer forward, 1,5pmol do primer reverse, 8mM de MgSO₄ e tampão *PFU* 10X.

Para amplificar o fragmento GCaseAlb a reação foi realizada com temperatura de desnaturação de 94°C durante 5 min e 30 ciclos de: 94°C por 1 min/ 60°C por 45 seg/ 72°C por 7 min, finalizando com 72°C por 7min.

Para amplificar o fragmento GCasePHO1 a reação foi realizada com temperatura de desnaturação de 94°C durante 5 min e 30 ciclos de: 94°C por 1 min/ 55°C por 45 seg/ 72°C por 7 min, finalizando com 72°C por 7min.

Os produtos das reações de PCR foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. As bandas dos fragmentos nos tamanhos aproximados de 1.566 pb para GCaseAalb e de 1.494 pb para GCasePHO1 foram extraídas do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante.

3.5.4. Subclonagem no vetor pGEM-Teasy

Os produtos de PCR purificados foram submetidos a uma reação de adenilação realizada a 72°C por 30 min (aproximadamente 75ng do DNA, 0,2µL da enzima *Taq DNA polimerase* (Fermentas), 1,5mM MgCl2, 0,2mM dATP e tampão Taq buffer KCl 10X) e posterior reação de ligação no vetor de subclonagem pGEM-Teasy (Promega), conforme instruções do fabricante, para construção dos vetores pGEM-Teasy/GCaseAlb e pGEM-Teasy/GCasePHO1.

3.5.5. Transformação de bactérias quimiocompetentes

Cerca de 3 μ L do produto da reação de ligação foi usado para transformar bactérias *E. coli* DH5 α quimiocompetentes. A mistura foi incubada em gelo por 30 min, aquecida a 42°C por 2 min e novamente incubada em gelo por 5 min. Após a adição de 350 μ L de meio LB, a mistura foi incubada a 37°C por 1 hora e 30 min, sob agitação, e plaqueada em meio sólido LB/Amp-IPTG-Xgal. As placas foram incubadas a 37°C por 16h.

3.5.6. PCR de colônia

Colônias resistentes à ampicilina foram submetidas a reações PCR para os genes GCaseAlb e GCasePHO1, conforme descrito no item 3.5.3, para confirmação da inserção dos genes.

3.5.7. Mini-preparações plasmidiais: pGEM-Teasy/GCaseAlb e pGEM-Teasy/GCasePHO1

Clones positivos foram inoculados em meio LB/Amp (100 µg/mL) e incubados a 37°C por 16h, sob agitação. A cultura foi então utilizada para realização de mini-preparações plasmidiais com o kit QiAprep spin miniprep kit (Qiagen), conforme instruções do fabricante.

3.5.8. Sequenciamento de DNA: pGEM-Teasy/GCaseAlb e pGEM-Teasy/GCasePHO1

Para analisar a sequência completa da GCaseAlb e GCasePHO1 foram desenhados oligonucleotídeos internos (ver Anexo 1). As reações foram realizadas pelo método de dideoxinucleotídeos (dNTPs) (Sanger *et al.*, 1977), usando o sequenciador automático ABI 3500 – *Genetic Analyser* (Applied Biosystems do Brasil) e o kit *Big Dye*, de acordo com as instruções do fabricante.

3.5.9. Digestões enzimáticas

Os plasmídeos pPIC3.5 e pGEM-Teasy/GCaseAlb, foram digeridos com as enzimas de restrição *EcoRI* e *NotI* (New England Biolabs), de acordo com as instruções do fabricante. Já os plasmídeos pHIL-S1 e pGEM-Teasy/GCasePHO1 foram digeridos com as enzimas de restrição *SmaI* e *XhoI* (New England Biolabs) também de acordo com as instruções do fabricante.

Os produtos das reações de digestão foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. Os fragmentos foram extraídos do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante.

3.5.10. Reação de ligação: pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1

Para a reação de ligação foi utilizada uma razão molar vetor:inserto de 1:2. A reação foi realizada utilizando a enzima *T4 DNA ligase* (Fermentas) conforme instruções do fabricante.

3.5.11. Precipitação de DNA com etanol

Todo o volume da reação de ligação pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1 foi precipitado com etanol: 1:10 v/v de acetato de sódio 3M pH5,2 foi adicionado e o restante do volume completado com H₂O milli-Q totalizando 100µL. 3 volumes de reação de etanol 100% gelado foram adicionados e a mistura foi incubada a -20°C por 30 min. Após esse período a mistura foi centrifugada a 13.000 rpm, 4°C, 15 min . O sobrenadante foi descartado e o pellet lavado com etanol 70% gelado, seguido de uma nova etapa de centrifugação a 13.000 rpm, 4°C, 10 min. Seguiu-se mais uma etapa de lavagem com etanol 70% e centrifugação e o pellet foi seco à temperatura ambiente.

3.5.12. Transformação de bactérias eletrocompetentes

O DNA precipitado com etanol referente às ligações pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1 foram ressuspendidos em alíquotas de *E. coli* DH10B eletrocompetente, que em seguida foram eletroporadas a 2500V em cubetas geladas de 0,2 cm. As bactérias eletroporadas foram ressuspedidas em 1mL de meio SOC e incubadas a 37°C por 1h, sob agitação e plaqueadas em meio sólido LB/Amp. As placas foram incubadas a 37°C por 16h.

3.5.13. Seleção de clones recombinantes

A seleção de clones positivos foi realizada por meio de PCR de colônia de clones resistentes a ampicilina (item 3.5.6), por sequenciamento (item 3.5.8) e digestão enzimática (item 3.5.9).

3.5.14. Transformação de *P. pastoris* com vetores pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1

Células de *P. pastoris* GS115 eletrocompetentes foram preparadas conforme descrito no manual Pichia Expression Kit (Life Techonologies).

Os plasmídeos pPIC3.5, pPIC3.5/GCaseAlb, pHIL-S1 e pHIL-S1/GCasePHO1 foram linearizados com a enzima de restrição *SacI F.D.* (Thermo Scientific), conforme instruções do fabricante. As reações de digestão foram precipitadas com etanol conforme descrito no item 3.5.11. O DNA precipitado foi adicionado a uma alíquota de *P. pastoris* GS115 eletrocompetente e a mistura foi eletroporada a 1500V, 200W, 25µF, em cubeta gelada de 0,2 cm. As células eletroporadas foram ressuspedidas em 1mL de sorbitol 1M e incubadas a 30°C por 1h, sob agitação e plaqueadas em meio MD sólido. As placas foram incubadas a 30°C por aproximadamente 3 dias.

3.5.15. Expressão de GCase em P. pastoris

A expressão de GCaseAlb e GCasePHO1 em *P. pastoris* foi realizada conforme descrito no manual do fabricante: Pichia Expression Kit (Life Technologies). Após 72h de indução por metanol, a cultura foi centrifugada a 14.000rpm, 4°C, por 3min, e o pellet e o sobrenadante foram separados e mantidos no gelo. O pellet da cultura foi lisado conforme descrito no manual do fabricante utilizando 1µL de *Protease Inhibitor Cocktail* (Sigma) e 10 mL do sobrenadante de cultura foram precipitados com acetona: 4 volumes de acetona gelada foram adicionado à amostra de sobrenadante de cultura, que foi agitada em vortex e incubada a -20°C por 1h. A mistura foi centrifugada a 13.000 x g por 10 min, o sobrenadante descartado e o pellet de proteínas foi deixado secar à temperatura ambiente.

3.5.16. Ensaio por western blotting

As amostras do lisado de extrato celular, do sobrenadante de cultura e do sobrenadante de cultura precipitado com acetona foram separadas por eletroforese em gel de SDS-poliacrilamida 12% para o precipitado e o sobrenadante de cultura e 10% para o sobrenadante de cultura precipitado com acetona, sob condições desnaturantes e transferidos para membranas de nitrocelulose. As membranas foram bloqueadas por 12 horas a 4°C em solução de bloqueio e incubadas por 90 min a temperatura ambiente com anticorpo policional anti-GCase produzido por (Novo *et al.*, 2012). As membranas foram lavadas 3 vezes com TBST 1X por 10 min e então incubadas com anti-IgG de camundongo gerado em cabra conjugado à peroxidase (Sigma) diluída em solução de bloqueio por 1h a temperatura ambiente. Após 3 lavagens em TBST 1X, as membranas foram reveladas com reagente de quimioluminescência ECL[®] (GE Healthcare).

3.6. Obtenção de linhagem controle de *P. pastoris* expressando a proteína *Sp*Cas9 (*Sp*Cas9cont)

3.6.1. Desenho de oligonucleotídeos específicos para clonagem de SpCas9cont

Oligonucleotídeos específicos foram desenhados para amplificar o gene da *Sp*Cas9 contendo a sequência de localização nuclear SV40 do plasmídeo PX458, gentilmente cedido pelo Dr. Sergio Verjovski Almeida (ver Anexo 1).

3.6.2. Construção do vetor pGEM-Teasy/SpCas9cont

Para amplificação da Cas9cont por PCR foi utilizada a enzima *Phusion High-Fidelity Polymerase* (Thermo Fisher) de acordo com o manual do fabricante, a reação foi realizada com temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 60°C por 30 seg/ 72°C por 3 min, finalizando com 72°C por 5 min.

O produto da reação de PCR foi submetido à eletroforese em gel de agarose 1% em tampão TAE 1X. A banda do fragmento no tamanho aproximado de 5 Kb para *Sp*Cas9cont foi extraída do gel e purificada com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante.

Seguiu-se então a subclonagem no vetor pGEM-Teasy (ver item 3.5.4), a transformação do produto de ligação em bactérias *E. coli* DH10B conforme procedimento descrito no item 3.5.12, PCR de colônia conforme descrito acima, mini preparações plasmidiais dos clones positivos (ver item 3.5.7), sequenciamento de DNA utilizando os primers 12 a 24 (ver Anexo 1) conforme descrito no item 3.5.8 e digestão enzimática utilizando as enzimas *SalI* e *NotI* (New England Biolabs), conforme instruções do fabricante.

3.6.3. Construção do vetor pPICHOLI-1/SpCas9cont

Os plasmídeos pPICHOLI-1 e pGEM-Teasy/*Sp*Cas9cont foram digeridos com as enzimas de restrição *SalI* e *NotI* (New England Biolabs), de acordo com as instruções do fabricante. Os produtos das reações de digestão foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. Os fragmentos foram extraídos do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante. Para a reação de ligação foi utilizada uma razão molar vetor:inserto de 1:3. A reação foi realizada utilizando a enzima *T4 DNA ligase* 400U (New England Biolabs) conforme instruções do fabricante. Conforme descritos

nos itens 3.5.11 e 3.5.12, o produto de ligação foi precipitado com etanol e utilizado para transformar bactérias *E. coli* TOP10F' eletrocompetentes. As bactérias eletroporadas foram plaqueadas em meio sólido LB/Zeo. A seleção de clones positivos foi realizada por meio da realização de PCR de colônia de clones resistentes a zeocina, seguida por sequenciamento (item 3.5.8) e digestão enzimática utilizando as enzimas *SalI* e *NotI* (New England Biolabs), conforme instruções do fabricante.

3.6.4. Transformação de *P. pastoris* com vetor pPICHOLI-1/SpCas9cont

Células de *P. pastoris* GS115 eletrocompetentes foram preparadas conforme descrito no manual do plasmídeo pPICHOLI-1 (MobiTec). 500ng da construção pPICHOLI-1/*Sp*Cas9cont foram adicionados a uma alíquota de *P. pastoris* GS115 eletrocompetente. Em seguida, a mistura foi eletroporada a 1500V, 200W, 25µF, em cubeta gelada de 0,2 cm. As células eletroporadas foram ressuspedidas em 1mL de sorbitol 1M e incubadas a 30°C por 1h, sob agitação e plaqueadas em meio sólido YPD/Zeo. As placas foram incubadas a 30°C por aproximadamente 3 dias.

3.6.5. Expressão de SpCas9 em P. pastoris pPICHOLI-1/SpCas9cont

Somente o pellet da cultura foi lisado conforme descrito no manual do fabricante (MobiTec) utilizando 5mM de *Phenylmethanesulfonyl fluoride* (PMSF) (Sigma) e mantido a - 20°C. As amostras do lisado de extrato celular foram separadas por eletroforese em gel de SDS-poliacrilamida 8% sob condições desnaturantes e transferidos para membrana de nitrocelulose. A membrana foi bloqueada por 12 horas a 4°C em solução de bloqueio e incubada por 90 min a temperatura ambiente com anticorpo monoclonal de camundongo anti-Cas9 (#MAC133, clone A79, Millipore). A membrana foi lavadas 3 vezes com TBST 1X por 10 min e então incubada com anti-IgG de camundongo gerado em cabra conjugado à peroxidase (Sigma) diluído em

solução de bloqueio, por 1h a temperatura ambiente. Após 3 lavagens em TBST 1X, as membranas foram reveladas com reagente de quimioluminescência ECL[®] (GE Healthcare).

3.7. Obtenção de linhagem *P. pastoris-Sp*Cas9 com deleção dos genes *alg3* e *och1* e integração das marcas de seleção pelo reparo por recombinação homóloga

3.7.1. Desenho dos RNAs guia (gRNAs) flanqueados pelas ribozimas HHR e HDV

Quatro gRNAs foram desenhados, um para a região 5' e um para a região 3' do gene *alg3*, e da mesma forma para o gene *och1*. As sequências dos genes *alg3* (GenBank: AY653304.1) e *och1* (NCBI Reference Sequence: XM_002492801.1) foram utilizadas para o desenho dos gRNAs utilizando o software Cas9 Design (<u>http://cas9.cbi.pku.edu.cn/</u>).

Os gRNAs foram sintetizados flanqueados por duas ribozimas, conforme descrito por Gao e Zaho, 2014: Hammerhead ribozyme (HHR) e hepatitis delta vírus ribozyme (HDV) da seguinte maneira: HHR-gRNA-HDV. Os 6 primeiros nucleotídeos da HHR devem ser complementares aos 6 primeiros nucleotídeos da sequência alvo do gRNA, conforme apresentado na Tabela 3.

As sequências do esqueleto do gRNA quimérico descrita por Ran *et al.* 2013 e das ribozimas estão descritas no anexo 2.

Tabela 3 – Desenho da sequência guia e nucleotídeos iniciais HHR

Alvo	Sequência guia (20pb) 5' – 3'	6 nucleotídeos iniciais HHR 5' – 3'
alg3 - 5'region	TGGTTCTATCGTTTCCTCTT	GAACCA
alg3 - 3'region	GATTGATCATGACCTGATTA	TCAATC
och1 - 5'region	GCTCACGGTTGGACTCATAA	GTGAGC
och1 - 3'region	ACTGCCGATTATCAGCTTCT	GGCAGT

3.7.2. Desenho dos genes sintéticos contendo sítio de clonagem para *Sp*Cas9 e os gRNAs flanqueados por ribozimas

Dois conjuntos de genes foram sintetizados de forma a ligar os fragmentos contendo todo o aparato de CRISPR-Cas9 para deleção dos genes *alg3* e *och1* nos vetores de expressão de *P*. *pastoris*: pPICHOLI-1 e pPIC6A respectivamente.

O primeiro deles foi desenhado para se ligar ao plasmídeo pPICHOLI-1 deletado da região promotora AOX1 e contém a seguinte sequência, conforme mostrado na Figura 6: sítios de restrição, seguido do promotor de *P. pastoris TEF1 (TEF1p)*, sítio de clonagem para *Sp*Cas9, terminador de *P. pastoris CYC1 (CYC1t)*, sítios de restrição, seguido do *TEF1p*, ribozima HHR, gRNA para quebra da região 5' do gene *alg3*, ribozima HDV e *CYC1t*, sítios de restrição e um novo módulo contendo *TEF1p*, ribozima HHR, gRNA para quebra da região 3' do gene *alg3*, ribozima HDV, *CYC1t* e sítio de restrição.



Figura 6 – Desenho do gene sintético gRNA_alg3, contendo um sítio para clonagem do gene Cas9 nos sítios de restrição *Hind III* e *NotI*, além do promotor *TEF1* (*TEF1p*) e do terminador *CYC1* (*CYC1t*), seguido por fragmentos contendo os gRNAs para as regiões 5' e 3' do gene *alg3* flanqueadas pelas ribozimas Hammerhead (HHR) e Hepatitis Delta Virus (HDV). Diversos sítios de restrição foram incluídos na construção, sendo os sítios para *HpaI* e *SmaI* utilizados para ligação ao vetor pPICHOLI-1.

O segundo foi desenhado de forma a se ligar no plasmídeo pPIC6A também deletado da região promotora AOX1 e contém a mesma estrutura do gene anterior, porém apresenta gRNAs para quebra das extremidades 5' e 3' do gene *och1*, além de apresentar uma sequência de replicação autônoma de *P. pastoris* (PARS) que já está presente no vetor pPICHOLI-1, mas

ausente no vetor pPIC6A. Essa sequência permite a replicação da sequência desejada sem integrá-la ao genoma da levedura. O segundo gene sintético está esquematizado na Figura 7.



Figura 7 - Desenho do gene sintético gRNA_och1, contendo um sítio para clonagem do gene Cas9 nos sítios de restrição *Hind III e NotI*, além do promotor *TEF1 (TEF1p)* e do terminador *CYC1 (CYC1t)*, seguido por fragmentos contendo os gRNAs para as regiões 5' e 3' do gene *och1* flanqueadas pelas ribozimas Hammerhead (HHR) e Hepatitis Delta Virus (HDV). Além disso, uma sequência de replicação autônoma de *P. pastoris* (PARS) foi incluída. Diversos sítios de restrição foram incluídos na construção, sendo os sítios para *HpaI* e *BamHI* utilizados para ligação ao vetor pPIC6A.

Os dois genes foram sintetizados pela empresa GenArtTM e vieram inseridos no vetor pMA-RQ/gRNA *alg3* e pMA-RQ/gRNA *och1*.

3.7.3. Construção do vetor pGEM-Teasy/SpCas9cont

Oligonucleotídeos específicos foram desenhados para amplificar o gene da *Sp*Cas9 contendo a sequência de localização nuclear SV40 do plasmídeo PX458, ver Anexo 1. E seguiuse o mesmo procedimento do item 3.6.2 para construção do vetor pGEM-Teasy/*Sp*Cas9.

3.7.4. Construção dos vetores pMA-RQ/gRNA_*alg3/Sp*Cas9 e pMA-RQ/gRNA *och1/Sp*Cas9

Os plasmídeos pMA-RQ/gRNA_*alg3*, pMA-RQ/gRNA_*och1* e pGEM-Teasy/*Sp*Cas9, foram digeridos sequencialmente com a enzima de restrição *Hind III* (Thermo Scientific), seguido da enzima de restrição *Not I* (New England Biolabs), de acordo com as instruções do fabricante. Os produtos das reações de digestão foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. Os fragmentos foram extraídos do gel e purificadas com o kit

QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante. Para a reação de ligação foi utilizada uma razão molar vetor:inserto de 1:5. A reação foi realizada utilizando a enzima *T4 DNA ligase* 400U (New England Biolabs) conforme instruções do fabricante. Conforme descritos no item 3.5.12, 3μ L dos produtos de ligação foram utilizados para transformar bactérias *E. coli* TOP10F' eletrocompetentes. As bactérias eletroporadas foram plaqueadas em meio sólido LB/Amp. A seleção de clones positivos foi realizada por meio da realização de PCR de colônia de clones resistentes a ampicilina, seguida por sequenciamento (item 3.5.8) e digestão enzimática com as enzimas de restrição *Hind III* (Thermo Scientific) e *Not I* (New England Biolabs).

3.7.5. Construção do vetor pPIC6A/gRNA och1/SpCas9 – Via Clássica

Para deleção da região promotora *AOX1* do vetor pPIC6A este foi digerido com as enzimas de restrição *BamHI* e *AleI* (Thermo Fischer), de acordo com as instruções do fabricante.

A construção pMA-RQ/gRNA_*och1/Sp*Cas9 foi digerida sequencialmente com as enzimas *BamHI* e *HpaI* (New England Biolabs), de acordo com as instruções do fabricante.

Os produtos das reações de digestão foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. As bandas dos fragmentos no tamanho aproximado de 2 Kb para o vetor pPIC6A e de 8Kb para o fragmento contendo gRNA_*och1/Sp*Cas9 foram extraídas do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante.

Para a reação de ligação foram utilizadas diferentes razões molares vetor:inserto, de 1:3, 1:5 e 3:1. As reações foram realizadas utilizando a enzima *T4 DNA ligase 400U* (New England Biolabs) conforme instruções do fabricante. Conforme descrito no item 3.5.12, cerca de 3 - 5μL dos produtos de ligação foram utilizados para transformar bactérias *E. coli* TOP10F' eletrocompetentes. As bactérias eletroporadas foram plaqueadas em meio sólido LB/Amp. A seleção de clones positivos foi realizada por meio da realização de PCR de colônia para o gene de *Sp*Cas9 de clones resistentes a blasticidina, conforme metodologia descrita no item 3.6.2.

3.7.6. Desenho dos braços de homologia

Para que ocorra o evento de recombinação homóloga durante o reparo da quebra de fita dupla originado pelo sistema CRISPR-Cas9, o gene que se deseja inserir deve estar flanqueado por regiões de homologia ao genoma nas regiões 5' e 3' à quebra da dupla fita (ver Figura 2).

Portanto, foram analisadas as sequências dos cromossomos 3 (NC_012965.1) e 4 (NC_012966.1) de *P. pastoris* GS115 que contém respectivamente os genes *och1* e *alg3*, e após simulação da deleção dos genes de acordo com o sítio de quebra de Cas9 e com os desenhos dos gRNAs realizados (ver item 3.7.1) foram identificados fragmentos do tamanho de 1 Kb a montante e a jusante à região de quebra da dupla fita para os dois genes.

A região de homologia 5' anterior à região de deleção do gene *alg3* será referida como esquerda de *alg3* (E3), assim como a região de homologia 3' posterior à região de deleção de *alg3* será referida como direita de *alg3* (D3). Da mesma forma, A região de homologia 5' anterior à região de deleção do gene *och1* será referida como esquerda de *och1* (E1), assim como a região de homologia 3' posterior à região de deleção de *och1* será referida como direita de *och1* (D1).

3.7.7. Desenho dos fragmentos para recombinação homóloga

Um conjunto de genes foi sintetizado de forma a unir os fragmentos para recombinação homóloga durante o reparo da quebra de fita dupla originado pelo sistema CRISPR-Cas9, contendo a seguinte sequência (Figura 8): sítios de restrição *SnaBI, Bam HI* e *Hind III*, seguido do promotor de *P. pastoris GAP1 (GAPp)*, sítio de clonagem para as marcas de resistência aos antibióticos Canamicina (Kan) ou Higromicina (Hygro) contendo enzimas de restrição *Not I* e *Xho I*, o terminador *TEF51A (TEF51At)* e os sítios de restrição das enzimas *Pac I, Cla I* e *Pme I*.



Figura 8 - Desenho do gene sintético HDR, contendo um sítio para clonagem das marcas de resistência aos antibióticos kanamicina ou hygromicina, contendo os sítios de restrição *NotI* e *XhoI*, além do promotor *GAP1* (*GAP1p*) e do terminador *TEF51A* (*TEF51At*). Diversos sítios de restrição foram incluídos na construção, sendo os sítios de *SnaBI* e *BamHI* utilizados para ligação aos braços de homologia esquerdo e os sítios de *ClaI* e *PmeI* para ligação dos braços de homologia direito (ver Figura 9).

Os sítios de restrição *SnaBI* e *Bam HI* foram desenhados para posterior ligação do braço esquerdo de homologia, assim como os sítios *Cla I* e *Pme I* foram desenhados para posterior ligação do braço direito de homologia. O gene foi sintetizado pela empresa GenArtTM e veio inserido no vetor pMK-RQ-Bb: pMK-RQ-Bb/HDR

Um esquema com as construções finais dos fragmentos desejados para recombinação homóloga está apresentado na Figura 9:



Figura 9 – Esquema final das construções a serem utilizadas no evento de recombinação homóloga (HDR). Fragmento contendo braços de homologia do tamanho de 1Kb na região de deleção do gene *och1* para integração da marca de resistência à kanamicina (Kan) no reparo por HDR. Fragmento contendo braços de homologia do tamanho de 1Kb na região de deleção do gene *alg3* para integração da marca de resistência à hygromicina (Hygro) no reparo por HDR.

3.7.8. Construção dos vetores pGEM-Teasy/E3, pGEM-Teasy/D3, pGEM-Teasy/E1 e pGEM-Teasy/D1

Para amplificar as regiões de homologia E3, D3, E1 e D1 do genoma de *P. pastoris* GS115, foi realizada a extração do DNA conforme metodologia descrita por (Tomita *et al.*, 2002). Oligonucleotídeos específicos foram desenhados para amplificar os braços de homologia E3, D3, E1 e D1 do genoma de *P. pastoris* GS115, através de reação de PCR (ver Anexo 1). Para as reações de PCR foram utilizados 5µL de uma diluição 1:100 do total de DNA extraído de *P. pastoris* GS115 e a enzima *Phusion High-Fidelity Polymerase* (Thermo Fisher) de acordo com o manual do fabricante.

Para amplificar os fragmentos E3 e D3, as reações foram realizadas a uma temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 60°C por 30 seg/ 72°C por 30 seg, finalizadas com 72°C por 5 min.

Para amplificar o fragmento E1, a reação foi realizada a uma temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 64°C por 30 seg/ 72°C por 30 seg, finalizada com 72°C por 5 min.

Para amplificar o fragmento D1, a reação foi realizada a uma temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 54°C por 30 seg/ 72°C por 30 seg, finalizada com 72°C por 5 min.

Os produtos das reação de PCR foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. As bandas dos fragmentos no tamanho aproximado de 1 Kb para cada braço de homologia foram extraídas do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante. Seguiu-se então a subclonagem no vetor pGEM-Teasy (ver item 3.5.4), a transformação de 5µL de cada produto de ligação em bactérias *E. coli* TOP10F' eletrocompetentes conforme procedimento descrito no item 3.5.12, PCR de colônia

conforme descrito acima, mini-preparações plasmidiais dos clones positivos (ver item 3.5.7) e sequenciamento de DNA utilizando os primers M13 e T7 (ver Anexo 1) conforme descrito no item 3.5.8.

Ainda para confirmação da clonagem dos braços de homologia no plasmídeo pGEM-Teasy, estes foram digeridos com a enzima de restrição *EcoRI* (Thermo Fisher) para liberação do inserto, de acordo com as instruções do fabricante. Os produtos das reações de digestão foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X.

3.7.9. Construção dos vetores pGEM-Teasy/Kan e pGEM-Teasy/Hygro

Oligonucleotídeos específicos foram desenhados para amplificar as marcas de resistência a Kan e Hygro dos vetores pENTR/D-TOPO ou pH7WG2.0 respectivamente (ver Anexo 1). Para as reações de PCR foi utilizada a enzima *Phusion High-Fidelity Polymerase* (Thermo Fisher) de acordo com o manual do fabricante.

Para amplificar o fragmento Kan, a reação foi realizada a uma temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 58°C por 30 seg/ 72°C por 30 seg, finalizada com 72°C por 5 min.

Para amplificar o fragmento Hygro, a reação foi realizada a uma temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 45°C por 30 seg/ 72°C por 1min, finalizada com 72°C por 5 min.

Os produtos das reações de PCR foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. As bandas dos fragmentos no tamanho aproximado de 810pb para Kan e de 1789pb para Hygro foram extraídas do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante. Os produtos de PCR das marcas de resistência aos antibióticos purificados foram submetidos a uma reação de ligação no vetor de subclonagem pGEM-Teasy conforme descrito no item 3.5.4. A transformação de 5 μ L de cada produto de ligação em bactérias *E. coli* TOP10F' eletrocompetentes foi realizada conforme procedimento descrito no item 3.5.12, PCR de colônia conforme descrito acima, mini preparações plasmidiais dos clones positivos (ver item 3.5.7) e sequenciamento de DNA utilizando os primers 25 e 26 (ver Anexo 1) conforme descrito no item 3.5.8.

3.7.10. Construção dos vetores pMK-RQ-Bb/HDR/Kan e pMK-RQ-Bb/HDR/Hygro

Os plasmídeos pMK-RQ-Bb/HDR, pGEM-Teasy/Kan e pGEM-Teasy/Hygro foram digeridos com as enzimas de restrição *Not I* e *Xho I* (New England Biolabs), de acordo com as instruções do fabricante. Os produtos das reações de digestão foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. Os fragmentos foram extraídos do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante.

Para a reação de ligação foi utilizada uma razão molar vetor:inserto de 1:3. As reações foi realizada utilizando a enzima *T4 DNA ligase* 400U (New England Biolabs) conforme instruções do fabricante. Conforme descrito no item 3.5.12, 3µL dos produtos de ligação foram utilizados para transformar bactérias *E. coli* TOP10F' eletrocompetentes. As bactérias eletroporadas foram plaqueadas em meio sólido LB/Kan. A seleção de clones positivos foi realizada por meio da realização de PCR de colônia para os genes Kan ou Hygro de clones resistentes a canamicina, conforme metodologia descrita no item 3.7.9.

3.7.11. Construção do vetor pMK-RQ-Bb/HDR/Kan/E1

As construções pMK-RQ-Bb/HDR/Kan e pGEM-Teasy/E1 foram digeridas sequencialmente com as enzimas *SnaBI* e *BamHI* (New England Biolabs), de acordo com as

instruções do fabricante. Os produtos das reações de digestão foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. As bandas dos fragmentos no tamanho aproximado de 4 Kb para o vetor pMK-RQ-Bb/HDR/Kan e de 1 Kb para o fragmento E1 foram extraídas do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante.

Para a reação de ligação foi utilizada uma razão molar vetor:inserto de 1:3. A reação foi realizada utilizando a enzima *T4 DNA ligase* 400U (New England Biolabs) conforme instruções do fabricante. Conforme descritos no item 3.5.12, $3 \mu L$ do produto de ligação foram utilizados para transformar bactérias *E. coli* TOP10F' eletrocompetentes. As bactérias eletroporadas foram plaqueadas em meio sólido LB/Kan. A seleção de clones positivos foi realizada por meio da realização de PCR de colônia para o gene E1 de clones resistentes a canamicina, conforme metodologia descrita no item 3.7.9.

3.7.12. Construção do vetor pGEM-Teasy/E1/Kan/D1 – Gibson Assembly

Após sucessivos insucessos nas tentativas de construção utilizando a via clássica de clonagem, por meio de enzimas de restrição, tentou-se obter as construções por meio de reações baseadas na técnica de Gibson Assembly, utilizando o kit NEBuilder® HiFi DNA Assembly Cloning Kit (New England Biolabs) que contém, além do mix para a reação de Assembly, bactérias *E. coli* DH5α quimiocompetentes.

Para tal, oligonucleotídeos específicos, contendo 25pb de homologia com a sequência anterior ou posterior na sequência foram desenhados (ver Anexo 1) para amplificação dos produtos por meio de reações de PCR utilizando a enzima *Phusion High-Fidelity Polymerase* (Thermo Fisher), de acordo com o manual do fabricante.

Para amplificar o fragmento E1 do vetor pGEM-Teasy/E1 (3.7.8) com os overlaps, a reação foi realizada com temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 67°C por 30 seg/ 72°C por 30 seg, finalizando com 72°C por 5 min.

Para amplificar o fragmento D1 do vetor pGEM-Teasy/D1 (3.7.8) com os overlaps, a reação foi realizada com temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 70°C por 30 seg/ 72°C por 30 seg, finalizando com 72°C por 5 min.

Para amplificar o fragmento HDR/Kan presente na construção pMK-RQ-Bb/HDR/Kan (3.7.11) com os overlaps, a reação foi realizada a uma temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 64°C por 30 seg/ 72°C por 1min, finalizada com 72°C por 5 min.

Os produtos das reações de PCR foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. As bandas dos fragmentos no tamanho aproximado de 1.900pb para HDR/Kan e de 1.041pb para E1 e D1 foram extraídas do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante.

A reação de Assembly e a transformação foram realizadas conforme manual do fabricante, utilizando-se 0,05pmol de cada fragmento e do vetor pGEM-Teasy. As bactérias quimiocompetentes transformadas foram plaqueadas em meio sólido LB/Amp/IPTG/Xgal. A seleção de clones positivos foi realizada por meio da realização de PCR de colônia para o fragmento de recombinação homóloga de clones resistentes a ampicilina, sequenciamento (item 3.5.8) e digestão enzimática com *EcoRI* (Thermo Scientific), de acordo com as instruções do fabricante.

3.7.13. Construção do vetor pGEM-Teasy/E3/Hygro/D3 – Gibson Assembly

Primeiramente realizou-se uma reação de Assembly para construção do vetor pMK-RQ-Bb/HDR/Hygro. Da mesma forma que no item anterior, oligonucleotídeos específicos foram desenhados (ver Anexo 1) para a reação de Assembly e foi realizada a reação de PCR utilizando a enzima *Phusion High-Fidelity Polymerase* (Thermo Fisher) de acordo com o manual do fabricante:

Para amplificar o fragmento Hygro do vetor pGEM-Teasy/Hygro (item 3.7.9) com os overlaps, a reação foi realizada a uma temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 72°C por 90 seg, finalizada com 72°C por 5 min.

O vetor pMK-RQ-Bb/HDR (item 3.7.7) foi digerido com as enzimas de restrição *XhoI* e *NotI* (New England Biolabs), conforme instruções do fabricante.

O produto das reações de PCR e de digestão foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. As bandas dos fragmentos no tamanho aproximado de 3.340pb para pMK-RQ-Bb/HDR e de 1.800pb para Hygro foram extraídas do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante.

A reação de Assembly e a transformação foram realizadas conforme manual do fabricante, utilizando-se 0,05pmol do vetor e 0,1pmol do inserto. As bactérias quimiocompetentes transformadas foram plaqueadas em meio sólido LB/Kan. A seleção de clones positivos foi realizada por meio da realização de PCR de colônia para o fragmento de recombinação homóloga de clones resistentes a canamicina, sequenciamento (item 3.5.8) e digestão enzimática com *XhoI* e *NotI* (New England Biolabs), de acordo com as instruções do fabricante.

Oligonucleotídeos específicos foram desenhados (ver Anexo 1) para a segunda reação de Assembly e foi realizada a reação de PCR utilizando a enzima *Phusion High-Fidelity Polymerase* (Thermo Fisher) de acordo com o manual do fabricante:

Para amplificar o fragmento E3 do vetor pGEM-Teasy/E3 (item 3.7.8) com os overlaps foi utilizada a mesma reação descrita para amplificar o fragmento D1 no item anterior.

Para amplificar o fragmento D3 do vetor pGEM-Teasy/D3 (item 3.7.8) com os overlaps foi utilizada a mesma reação descrita para amplificar o fragmento E1 no item anterior.

Para amplificar o fragmento HDR/Hygro presente na construção pMK-RQ-Bb/HDR/Hygro (item 3.7.12) com os overlaps, a reação foi realizada a uma temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 67°C por 30 seg/ 72°C por 90 seg, finalizada com 72°C por 5 min.

Novamente foi realizada a reação de Assembly contendo 0,05pmol de cada fragmento e do vetor pGEM-Teasy.

3.7.14. Construção dos vetores pPIC6A/gRNA_*och1/Sp*Cas9 e pPICHOLI-1/gRNA *alg3/Sp*Cas9 – Gibson Assembly

Primeiramente tentou-se amplificar por PCR os fragmentos gRNA alg3 e gRNA *och1/Sp*Cas9 dos vetores pMA-RQ/gRNA alg3 (item 3.7.2) e pMA-RQ/gRNA och1/SpCas9 (item 3.7.4), respectivamente, utilizando oligonucleotídeos específicos que possuíam 25pb de overlap com os vetores de interesse. Após o teste de várias condições, não foi possível padronizar a reação de PCR. Portanto, amplificamos por PCR o esqueleto dos vetores pPICHOLI-1 e pPIC6A sem a região promotora AOX1, e com 25pb de overlap com a sequência dos gRNAs, utilizando a enzima Phusion High-Fidelity Polymerase (Thermo Fisher) de acordo com o manual do fabricante:

Para amplificar o fragmento pPICHOLI-1∆AOX1 do vetor pPICHOLI-1 com os overlaps, a reação foi realizada a uma temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 60°C por 30 seg/ 72°C por 1min, finalizada com 72°C por 5 min.

Para amplificar o fragmento pPIC6A-1∆AOX1 do vetor pPIC6A com os overlaps, a reação foi realizada a uma temperatura de desnaturação de 98°C durante 30 seg e 35 ciclos de: 98°C por 5 seg/ 55°C por 30 seg/ 72°C por 1min, finalizada com 72°C por 5 min.

O vetor pMA-RQ/gRNA_*alg3* foi digerido com a enzima de restrição *SfiI* (New England Biolabs), conforme instruções do fabricante. O vetor pMA-RQ/gRNA_*och1/Sp*Cas9 foi digerido sequencialmente com as enzimas *BamHI* (Invitrogen) e *HpaI* (New England Biolabs), conforme instruções do fabricante.

O produto das reações de PCR e de digestão foram submetidos à eletroforese em gel de agarose 1% em tampão TAE 1X. As bandas dos fragmentos no tamanho aproximado de 2.300pb para pPICHOLI-1, 2Kb para pPIC6A, 2.700pb para gRNA_*alg3* e 8Kb para gRNA_*och1/Sp*Cas9 foram extraídas do gel e purificadas com o kit QiAquick Gel Extraction Kit (Quiagen), conforme instruções do fabricante.

A reação de Assembly e a transformação foram realizadas conforme manual do fabricante, utilizando-se 0,05pmol do vetor e 0,1pmol do inserto para cada reação. As bactérias quimiocompetentes transformadas foram plaqueadas em meio sólido LB/Zeo para pPICHOLI-1/gRNA_*alg3/Sp*Cas9 ou LB/Blas para pPIC6A/gRNA_*och1/Sp*Cas9. A seleção de clones positivos foi realizada por meio da realização de PCR de colônia para o gRNA ou para Cas9.

4. Resultados

4.1. Obtenção de linhagem controle *P. pastoris* expressando a proteína GCase

4.1.1. Construção dos vetores pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1

Para a construção da linhagem de *P. pastoris* GS115 expressando as proteínas GCaseAlb e GCasePHO1, os genes GCaseAlb e GCase foram amplificados do vetor pH7WG2.0 por PCR e bandas nos tamanhos aproximados de 1.566 pb para GCaseAlb e de 1.494 pb para GCasePHO1 foram observadas (Figura 10a). Bactérias *E.coli* DH5 α foram transformadas com os plasmídeos pGEM-Teasy/GCaseAlb e pGEM-Teasy/GCasePHO1, construídos a partir dos fragmentos amplificados por PCR, e clones positivos crescidos em placa LB/Amp/IPTG/Xgal foram selecionados. A presença de GCaseAlb e GCasePHO1 foi confirmada por PCR de colônia (Figuras 10b e 10c, respectivamente), e após a realização de mini-preparações plasmidiais as construções foram analisadas por ensaios de digestão enzimática (Figuras 10d e 10e) e sequenciamento.



Figura 10 - Construção dos vetores pGEM-Teasy/GCaseAlb e pGEM-Teasy/GCasePHO1. (a) PCR para amplificação dos fragmentos GCaseAlb e GCase do vetor pH7WG2.0: 1- 1Kb plus, 2- GCaseAlb e 3- GCase. (b) PCR de colônia da transformação pGEM-Teasy/GCaseAlb para confirmação da presença de GCaseAlb: 1- 1 Kb plus, 2- Clone 1, 3- Clone 2 e 4- Clone 3; (c) PCR de colônia da transformação pGEM-Teasy/GCasePHO1 para confirmação da presença de GCasePHO1: 1- 1Kb plus, 2- Clone 1, 3- Clone 2 e 4- Clone 3; (c) PCR de colônia da transformação pGEM-Teasy/GCasePHO1 para confirmação da presença de GCasePHO1: 1- 1Kb plus, 2- Clone 1, 3- Clone 2 e 4- Clone3; (d) Digestão do vetor pGEM-Teasy/GCaseAlb com as enzimas de restrição *EcoRI* e *NotI*: 1- 1 Kb plus, 2- pGEM-Teasy/GCaseAlb *EcoRI/NotI*; (e) Digestão do vetor pGEM-Teasy/GCasePHO1 com as enzimas de restrição *SmaI* e *XhoI*: 1- 1 Kb plus, 2- pGEM-Teasy/GCasePHO1, 3- pGEM-Teasy/GCasePHO1 *SmaI*, 4- pGEM-Teasy/GCasePHO1 *XhoI* e 5- pGEM-Teasy/GCasePHO1 *XhoI* e 5- pGEM-Teasy/GCasePHO1 *SmaI*, 4- pGEM-Teasy/GCasePHO1 *XhoI* e 5- pGEM-Teasy/GCasePHO1 *SmaI*, 4- pGEM-Teasy/GCasePHO1 *XhoI* e 5- pGEM-Teasy/GCasePHO1 *XhoI* e 5- pGEM-Teasy/GCasePHO1 *SmaI*, 4- pGEM-Teasy/GCasePHO1 *SmaI*, 4- pGEM-Teasy/GCasePHO1 *XhoI* e 5- pGEM-Teasy/GCasePHO1 *SmaI*, 4- pGEM-Teasy/GCasePHO1 *XhoI* e 5- pGEM-Teasy/GCasePHO1 *SmaI*, 4- pGEM-Teasy/GCasePHO1 *XhoI* e 5- pGEM-Teasy/GCasePHO1 *SmaI*, 4- p

Em seguida os vetores foram digeridos e os fragmentos correspondentes à GCaseAlb e GCasePHO1 foram ligados nos vetores de expressão pPIC3.5 e pHIL-S1. Bactérias *E.coli* DH10B foram transformadas, clones positivos resistentes a ampicilina foram selecionados e, a partir do PCR de colônia, foi possível observar a amplificação dos fragmentos nos tamanhos esperados de 1.494 pb para GCasePHO1 e de 1.566 pb para GCAseAlb, confirmando a inserção dos genes (Figuras 11a e 11b). As construções foram confirmadas após a realização de minipreparações plasmidiais por ensaios de digestão enzimática (Figuras 11c e 11d) e sequenciamento.



Figura 11 - Construção dos vetores pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1. (a) PCR de colônia da transformação pHIL-S1/GCasePHO1 para confirmação da presença de GCase: 1- 1 Kb plus, 2- Clone 1, 3- Clone 2 e 4- Clone 3; (b) PCR de colônia da transformação pPIC3.5/GCaseAlb para confirmação da presença de GCaseAlb: 1- 1 Kb plus, 2- pGEM-Teasy/GCaseAlb e 18 clones foram analisados nas canaletas de 3 a 20, respectivamente; (c) Digestão pPIC3.5/GCaseAlb com enzimas de restrição *EcoRI* e *NotI*: 1- 1 Kb plus, 2- pPIC3.5/GCaseAlb com enzimas de restrição *EcoRI* e *NotI*: 1- 1 Kb plus, 2- pPIC3.5/GCaseAlb, 3- pPIC3.5/GCaseAlb *EcoRI*, 4- pPIC3.5/GCaseAlb *NotI* e 5- pPIC3.5/GCaseAlb *EcoRI/NotI*; (d) Digestão do vetor pHIL-S1/GCasePHO1 com as enzimas de restrição *SmaI* e *XhoI*: 1- 1 Kb plus, 2- pHIL-S1/GCasePHO1, 3- pHIL-S1/GCasePHO1 *SmaI*, 4- pHIL-S1/GCasePHO1 *XhoI* e 5- pHIL-S1/GCasePHO1 *SmaI/XhoI*.

4.1.2. Expressão de GCaseAlb e GCasePHO1 em *P. pastoris*

Após transformação de *P. pastoris* com as construções pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1, foi observado o crescimento em placas MD, que não possui o aminoácido histidina no meio de cultivo, resultando em colônias individualizadas.

Seis clones foram escolhidos para o procedimento de expressão, além de um controle negativo contendo apenas o vetor de expressão vazio. A expressão de GCaseAlb e GCasePHO1 foi realizada conforme descrito no item 3.5.15. Os lisados celulares e os sobrenadantes de cultura foram submetidos à análise da expressão por western-blotting. Como controle positivo foi utilizada a proteína GCase expressa em células COS7 (aproximadamente 66 kDa) ou a proteína purificada de *E.coli* (aproximadamente 56 kDa) em uréia 8M, utilizada na imunização dos animais para produção do soro por Novo *et al.*, 2012.

O western blot com soro anti-GCase mostrou fortes bandas para o lisado celular das linhagens transformadas tanto com o vetor pPIC3.5/GCaseAlb (Figura 12a, bandas fracas de aproximadamente 66 kDa e bandas mais forte de aproximadamente 45 kDa) quanto para o vetor pHIL-S1/GCasePHO1 (Figura 12b, bandas aproximadamente de 66 e 50 kDa). A presença de várias bandas se deve, muito provavelmente, à degradação da GCase por proteases. A expressão de GCase não foi detectada nos clones expressando apenas os vetores vazios.



Figura 12 - Análise da expressão de GCaseAlb e GCasePHO1 no lisado celular de *P. pastoris* após 72h de indução com metanol por western blot utilizando soro anti-GCase. (a) Expressão de GCaseAlb após transformação com o vetor pPIC3.5/GCaseAlb: 1- LMW, 2- COS7/GCase, 3- pPIC3.5 vazio, 4- Clone 1, 5- Clone 2, 6- Clone 3, 7- Clone 4, 8- Clone 5, 9- Clone 6; (b) Expressão de GCasePHO1 após transformação com o vetor pHIL-S1/GCasePHO1: 1- LMW, 2- GCase purificada de *E. coli*, 3- pHIL-S1 vazio, 4- Clone 1, 5- Clone 2, 6- Clone 3, 7- Clone 4, 8- Clone 6

Um western blot foi realizado com soro anti-GCase utilizando amostras em condições desnaturantes dos sobrenadantes de cultura dos seis clones selecionados mas nenhuma banda foi detectada. Portanto, o sobrenadante de cultura foi precipitado com acetona, conforme metodologia descrita no item 1.1.14, para clones selecionados com alta expressão de GCase segundo o resultado obtido para o western blot do lisado celular. Um novo western-blot foi realizado com soro anti-GCase e bandas fracas com massa molecular maior que 97 kDa para a proteína GCase foram observadas, principalmente no clone transformado com o vetor pHIL-S1/GCase (Figura 13, bandas indicadas pelas setas vermelhas). Não foram detectadas bandas para os vetores de expressão vazios. As bandas fracas podem ser explicadas pela baixa expressão da proteína no sobrenadante de cultura.



Figura 13 - Análise da expressão de GCasealb e GCasePHO1 no sobrenadante de cultura de *P. pastoris* precipitado com acetona após 72h de indução com metanol por western blot utilizando soro anti-GCase. 1- LMW, 2- GCase purificada de *E. coli*, 3- pPIC3.5 vazio, 4- Clone 3 pPIC3.5/GCaseAlb, 5- pHLI-S1 vazio, 6- Clone 3 pHIL-S1/GCasePHO1.

4.2. Obtenção de linhagem controle de *P. pastoris* expressando a proteína Cas9 (Cas9cont)

4.2.1. Construção do vetor pPICHOLI-1/SpCas9cont

Para a construção da linhagem controle de *P. pastoris* GS115 expressando a proteína *Sp*Cas9 (*Sp*Cas9cont), o gene da Cas9 foi amplificado do vetor PX458 por PCR e uma banda no tamanho esperado de 5Kb, correspondente a Cas9cont foi observada (Figura 14a). Bactérias *E.coli* TOP 10F' foram transformadas com o plasmídeo pGEM-Teasy/*Sp*Cas9cont, construído a partir do fragmento amplificado por PCR, e clones positivos crescidos em placa LB/Amp/IPTG/Xgal foram selecionados. A presença de *Sp*Cas9cont foi confirmada por PCR de colônia (Figura 14b), e após a realização de mini-preparações plasmidiais as construções foram analisadas por ensaios de digestão enzimática (Figura 14c) e sequenciamento.



Figura 14 - Construção do vetor pPGEM-Teasy/SpCas9cont. (a) PCR para amplificação do fragmento *Sp*Cas9cont do vetor PX458: 1- 1Kb plus, 2- *Sp*Cas9cont; (b) PCR de colônia da transformação pGEM-Teasy/*Sp*Cas9cont para confirmação da presença de *Sp*Cas9cont: 1- 1 Kb plus, 2- Clone 1; (c) Digestão do vetor pGEM-Teasy/*Sp*Cas9cont com as enzimas de restrição *SalI* e *NotI*: 1- 1 Kb plus, 2- pGEM-Teasy/*Sp*Cas9cont, 3- pGEM-Teasy/*Sp*Cas9cont *SalI*, 4- pGEM-Teasy/*Sp*Cas9cont *NotI* e 5- pGEM-Teasy/*Sp*Cas9cont *SalI*/*NotI*.

Em seguida o vetor foi digerido e o fragmento correspondente à *Sp*Cas9cont foi ligado ao vetor de expressão pPICHOLI-1. Bactérias *E.coli* TOP 10F' eletrocompetentes foram transformadas, clones positivos resistentes a zeocina foram selecionados e, a partir do PCR de colônia, foi possível observar a amplificação do fragmento no tamanho esperado de 5 Kb para *Sp*Cas9, confirmando a inserção do gene (Figura 15a). A construção foi confirmada após a realização de mini-preparações plasmidiais por ensaios de digestão enzimática (Figura 15b) e sequenciamento.



Figura 15 - Construção do vetor pPICHOLI-1/SpCas9cont. (a) PCR de colônia da transformação pPICHOLI-1/SpCas9cont para confirmação da presença de SpCas9cont: 1- 1 Kb plus e 12 clones foram analisados nas canaletas de 2 a 13, respectivamente; (b) Digestão dos clones 5 e 9 com enzimas de restrição SalI e NotI: 1- 1 Kb plus, 2-Clone 5 pPICHOLI-1/SpCas9cont, 3- Clone 5 pPICHOLI-1/SpCas9cont Sal I, 4- Clone 5 pPICHOLI-1/SpCas9cont NotI, 5- Clone 5 pPICHOLI-1/SpCas9cont SalI/NotI, 6- Clone 9 pPICHOLI-1/SpCas9cont, 7- Clone 9 pPICHOLI-1/SpCas9cont SalI, 8- Clone 9 pPICHOLI-1/SpCas9cont NotI e 9- Clone 9 pPICHOLI-1/SpCas9cont SalI/NotI.

4.2.2. Expressão de Cas9cont em *P. pastoris*

Após transformação de *P. pastoris* com a construção pPICHOLI-1/*Sp*Cas9cont, foi observado o crescimento de colônias resistentes à zeocina, resultando em algumas colônias individualizadas. Três clones foram escolhidos para o procedimento de expressão, além de um clone contendo apenas o vetor de expressão vazio. A expressão de *Sp*Cas9 foi realizada conforme descrito no item 3.6.5. O lisado celular foi submetido à análise da expressão por western-blotting.

O western blot realizado com anticorpo monoclonal anti-Cas9, se mostrou específico para os clones positivos induzidos por metanol, indicado pela presença de bandas no tamanho aproximado de 160kDa (Figura 16). As bandas mais altas provavelmente são devido à formas de *Sp*Cas9 que sofreram modificação pós-traducional, como glicolisação por exemplo. A expressão de *Sp*Cas9 não foi detectada nos clones expressando apenas o vetor vazio nem nos clones positivos não induzidos.



Figura 16 - Análise da expressão de SpCas9 no lisado celular de *P. pastoris* após 72h de indução com metanol por western blot utilizando anticorpo monoclonal anti-Cas9. 1- pPICHOLI-1 vazio não induzido, 2- pPICHOLI-1 vazio induzido, 3- Clone 1 pPICHOLI-1/*Sp*Cas9cont não induzido, 4- Clone 1 pPICHOLI-1/*Sp*Cas9cont induzido, 5- Clone 2 pPICHOLI-1/*Sp*Cas9cont não induzido, 6- Clone 2 pPICHOLI-1/*Sp*Cas9cont induzido, 7- Clone 3 pPICHOLI-1/*Sp*Cas9cont não induzido, 8- Clone 3 pPICHOLI-1/*Sp*Cas9cont induzido.

4.3. Obtenção de linhagem *P. pastoris*-Cas9 com deleção dos genes alg3 e och1 e integração das marcas de seleção pelo reparo por recombinação homóloga

4.3.1. Construção dos vetores pMA-RQ/gRNA_alg3/SpCas9 e pMA-RQ/gRNA och1/SpCas9

Para a construção da linhagem de *P. pastoris* GS115 mutante, o gene da *Sp*Cas9 foi amplificado do vetor PX458 por PCR e uma banda no tamanho esperado de 5Kb, correspondente a *Sp*Cas9 foi observada (Figura 17a). Bactérias *E.coli* TOP 10F' foram transformadas com o plasmídeo pGEM-Teasy/*Sp*Cas9, construído a partir do fragmento amplificado por PCR, e clones positivos crescidos em placa LB/Amp/IPTG/Xgal foram selecionados. A presença de *Sp*Cas9 foi confirmada por PCR de colônia (Figura 17b), e após a realização de mini-preparações plasmidiais as construções foram analisadas por ensaios de digestão enzimática (Figura 17c) e sequenciamento.



Figura 17 - Construção do vetor pPGEM-Teasy/SpCas9. (a) PCR para amplificação do fragmento *Sp*Cas9 do vetor pX458: 1- 1Kb plus, 2- *Sp*Cas9; (b) PCR de colônia da transformação pGEM-Teasy/*Sp*Cas9 para confirmação da presença de Cas9: 1- 1 Kb plus, 2- Clone 1, 3- Clone 2, 4- Clone 3; (c) Digestão do vetor pGEM-Teasy/*Sp*Cas9 com as enzimas de restrição *Hind III* e *NotI*: 1- 1 Kb plus, 2- pGEM-Teasy/*Sp*Cas9 *HindIII/NotI*.

Em seguida o vetor foi digerido e o fragmento correspondente à *Sp*Cas9 foi ligado nos vetores de expressão pMA-RQ/gRNA_*alg3* ou pMA-RQ/gRNA_*och1*. Bactérias *E.coli* TOP 10F' foram transformadas, clones positivos resistentes a ampicilina foram selecionados e, a partir do PCR de colônia, foi possível observar a amplificação do fragmento no tamanho esperado de 5 Kb para *Sp*Cas9, confirmando a inserção do gene apenas no vetor pMA-RQ/gRNA_*och1* (Figura 18b). Nenhuma colônia foi positiva para a ligação de Cas9 ao vetor pMA-RQ/gRNA_*alg3* (Figura 18a). A construção pMA-RQ/gRNA_*och1/Sp*Cas9 foi confirmada após a realização de mini-preparações plasmidiais por ensaios de digestão enzimática (Figura 18c) e sequenciamento.



Figura 18 - Construção dos vetores pMA-RQ/gRNA_alg3/SpCas9 e pMA-RQ/gRNA_och1/SpCas9. (a) PCR de colônia da transformação pMA-RQ/gRNA_*alg3/Sp*Cas9 para confirmação da presença de *Sp*Cas9: 1- 1 Kb plus, 2-pGEM-Teasy/*Sp*Cas9 e 7 clones foram analisados nas canaletas de 3 a 9, respectivamente; (b) PCR de colônia da transformação pMA-RQ/gRNA_*och1/Sp*Cas9 para confirmação da presença de *Sp*Cas9: 1- 1 Kb plus, 2- pGEM-Teasy/*Sp*Cas9 e 12 clones foram analisados nas canaletas de 3 a 14, respectivamente ; (c) Digestão do clones 12 pMA-RQ/gRNA_*och1/Sp*Cas9 com enzimas de restrição *XhoI* e *NdeI*: 1- 1 Kb plus, 2- pMA-RQ/gRNA_*och1/Sp*Cas9 *XhoI/NdeI*.

4.3.2. Construção do vetor pPIC6A/gRNA_och1/SpCas9

Os vetores pPIC6A e pMA-RQ/gRNA_*och1/Sp*Cas9 foram digeridos e bactérias *E.coli* TOP 10F' foram transformadas com o produto da ligação dos fragmentos digeridos e purificados. Clones resistentes crescidos em placas LB/Blas foram selecionados e a presença de *Sp*Cas9 foi analisada por PCR de colônia (Figura 19). Não foi verificada a presença nenhuma banda com o tamanho estimado de 5Kb referente à *Sp*Cas9 em nenhum dos clones selecionados.



Figura 19 - Construção do vetor pPIC6A/gRNA_och1/SpCas9. PCR de colônia da transformação pPIC6A/gRNA_*och1/Sp*Cas9 para confirmação da presença de *Sp*Cas9: 1-1 Kb plus, 2- pGEM-Teasy/*Sp*Cas9 e 13 clones foram analisados nas canaletas de 3 a 15, respectivamente
4.3.3. Construção dos vetores pGEM-Teasy/E3, pGEM-Teasy/D3, pGEM-Teasy/E1 e pGEM-Teasy/D1

Os genes referentes aos braços de homologia E3, D3, E1 e D1 foram amplificados do genoma de *P. pastoris* por PCR e bandas nos tamanhos esperados de 1Kb, correspondente aos braços foram observadas (Figura 20a). Bactérias *E.coli* TOP 10F' foram transformadas com os plasmídeos pGEM-Teasy/E3, pGEM-Teasy/D3, pGEM-Teasy/E1 e pGEM-Teasy/D1, construídos a partir dos fragmentos amplificados por PCR, e clones positivos crescidos em placa LB/Amp/IPTG/Xgal foram selecionados. As construções foram analisadas por ensaios de digestão enzimática (Figura 20b) e sequenciamento.



Figura 20 - Construção dos vetores pGEM-Teasy/E3, pGEM-Teasy/D3, pGEM-Teasy/E1 e pGEM-Teasy/D1. (a) PCR para amplificação dos braços de homologia do DNA genômico de *P. pastoris*: 1- 1Kb plus, 2- E3, 3- D3, 4- E1, 5- D1; (b) Digestão dos vetores pGEM-Teasy/E3 e pGEM-Teasy/D3 com a enzima de restrição *EcoRI*: 1- 1 Kb plus, 2- pGEM-Teasy/E3, 3- pGEM-Teasy/E3 *EcoRI*, 4- pGEM-Teasy/D3 e 5- pGEM-Teasy/D3 *EcoRI*; (c) Digestão dos vetores pGEM-Teasy/D1 com a enzima de restrição *EcoRI*: 1- 1 Kb plus, 2- pGEM-Teasy/E1 e pGEM-Teasy/D1 com a enzima de restrição *EcoRI*: 1- 1 Kb plus, 2- pGEM-Teasy/E1, 3- pGEM-Teasy/E1 e pGEM-Teasy/D1 e 5- pGEM-Teasy/D1 *EcoRI*.

4.3.4. Construção dos vetores pMK-RQ-Bb/HDR/Kan e pMK-RQ-Bb/HDR/Hygro

Os genes das marcas de resistência aos antibióticos canamicina e higromicina foram amplificadas dos vetores pENTR/D-TOPO e pH7WG2.0, respectivamente, por PCR e bandas nos

tamanhos esperados de aproximadamente 810pb para Kan (Figura 21a) e de 1789pb para Hygro (Figura 21d) foram observadas. Bactérias *E.coli* TOP 10F' foram transformadas com os plasmídeos pGEM-Teasy/Kan e pGEM-Teasy/Hygro, construídos a partir dos fragmentos amplificados por PCR, e clones positivos crescidos em placa LB/Amp/IPTG/Xgal foram selecionados. A presença de Kan e Hygro foi confirmada por PCR de colônia (Figuras 21b e 21e, respectivamente), e após a realização de mini-preparações plasmidiais as construções foram analisadas por ensaios de digestão enzimática (Figuras 21c e 21f) e sequenciamento.



Figura 21 - Construção dos vetores pGEM-Teasy/Kan e pGEM-Teasy/Hygro. (a) PCR para amplificação de Kan do vetor pENTR-D/TOPO: 1- 1Kb plus e 2- Kan; (b) PCR de colônia da transformação pGEM-Teasy/Kan para confirmação da presença de Kan: 1- 1 Kb plus, 2- Clone 1 e 3- Clone 2; (c) Digestão do vetor pGEM-Teasy/Kan com as enzimas de restrição *XhoI* e *NotI*: 1- 1 Kb plus e 2- pGEM-Teasy/Kan *XhoI/NotI*; (d) PCR para amplificação de Hygro do vetor pH7WG2.0: 1- 1 Kb plus e 2- Hygro; (e) PCR de colônia da transformação pGEM-Teasy/Hygro para confirmação da presença de Hygro: 1- 1 Kb plus, 2- Clone 1, 3- Clone 2 e 4- Clone 3; (f) Digestão do vetor pGEM-Teasy/Hygro com as enzimas de restrição *XhoI* e *NotI*: 1- 1 Kb plus, 2- pGEM-Teasy/Hygro, 3-pGEM-Teasy/Hygro *XhoI* e 4- pGEM-Teasy/Hygro *XhoI/NotI*;

Em seguida os vetores foram digeridos e os fragmentos correspondentes à Kan e à Hygro foram ligados no vetor de expressão pMK-RQ-Bb/HDR. Bactérias *E.coli* TOP 10F' foram transformadas, clones positivos resistentes a canamicina foram selecionados e, a partir do PCR de colônia, foi possível observar a amplificação dos fragmentos nos tamanhos esperados de 810pb para Kan e de fragmentos de aproximadamente 1.786pb para Hygro (Figura 22a). Apenas a construção pMK-RQ-Bb/HDR/Kan foi confirmada após a realização de mini-preparações plasmidiais por ensaios de digestão enzimática (Figura 22b) e sequenciamento. Nenhuma colônia foi de fato positiva para a ligação de Hygro ao vetor pMK-RQ-Bb/HDR.

Posteriormente, tentou-se ainda a ligação do fragmento E1 ao vetor pMK-RQ-Bb/HDR/Kan, mas sem sucesso.



Figura 22 - Construção dos vetores pMK-RQ-Bb/HDR/Kan e pMK-RQ-Bb/HDR/Hygro. (a) PCR de colônia das transformações para confirmação da presença de Kan ou Hygro: 1- 1 Kb plus, 2- pGEM-Teasy/Kan, 3- Clone 1 pMK-RQ-Bb/HDR/Kan, 4- Clone 2 pMK-RQ-Bb/HDR/Kan, 5- Clone 3 pMK-RQ-Bb/HDR/Kan, 6- pGEM-Teasy/Hygro, 7- Clone 1 pMK-RQ-Bb-sínteseIII/Hygro, 8- Clone 2 pMK-RQ-Bb/sínteseIII/Hygro, 9- Clone 3 pMK-RQ-Bb/HDR/Hygro; (b) Digestão de pMK-RQ-Bb/HDR/Kan com enzimas de restrição *XhoI* e *NotI*: 1- 1 Kb plus, 2- pMK-RQ-Bb/HDR/Kan, 3- pMK-RQ-Bb/HDR/Kan *XhoI*, 4- pMK-RQ-Bb/HDR/Kan *NotI*, 5- pMK-RQ-Bb/HDR/Kan *XhoI*/NotI.

4.3.5. Construção do vetor pGEM-Teasy/E1/Kan/D1 – Gibson Assembly

Devido ao grande número de insucessos nas etapas de clonagem clássica, por meio do uso de enzimas de restrição, decidiu-se utilizar a técnica de clonagem por Gibson Assembly. O vetor pGEM-Teasy, os fragmentos E1 e D1 e a marca de resistência a Kan, ligada ao promotor e ao terminador (HDR/Kan), foram amplificados por PCR e bandas nos tamanhos esperados de

aproximadamente 3Kb para o vetor pGEM-Teasy, 1Kb para os braços de homologia E1 e D1 e de 1.900pb para HDR/Kan foram identificados (Figura 23a) e purificados. Bactérias *E.coli* DH5α foram transformadas com a reação de assembly (ver item 3.7.12), e clones positivos crescidos em placa LB/Amp/IPTG/Xgal foram selecionados e reações de mini-preparações plasmidiais foram realizadas (Figura 23b). O clone 3 (canaleta 4, Figura 23b) apresentou tamanho aproximado de 6Kb e a construção foi confirmada por digestão enzimática (Figura 23c) e sequenciamento.



Figura 23 - Construção do vetor pGEM-Teasy/E1/Kan/D1 – Gibson Assembly. (a) Amplificação por PCR dos fragmentos para Assembly: 1- 1Kb plus, 2- pGEM-Teasy, 3- vazio, 4- E1, 5- vazio, 6- HDR/Kan, 7- vazio, 8- D1; (b) Mini-preparações plasmidiais de colônias brancas: 1- 1 Kb plus, 2 - 14 clones foram analisados nas canaletas de 2 a 15, respectivamente; (c) Digestão enzimática do Clone 3 com *EcoRI*: 1 – 1Kb plus, 2- Clone 3 pGEM-Teasy/E1/Kan/D1 *EcoRI*.

4.3.6. Construção do vetor pGEM-Teasy/E3/Hygro/D3 – Gibson Assembly

Inicialmente foi construído o vetor pMK-RQ-Bb/HDR/Hygro. O vetor pMK-RQ-Bb/HDR foi digerido com as enzimas XhoI e NotI (Figura 24b) e a marca de resistência a Hygro foi amplificada por PCR e uma banda no tamanho esperado de aproximadamente 1.800pb foi detectada (Figura 24a) e purificada. Bactérias *E.coli* DH5α foram transformadas com a reação de assembly (ver item 3.7.13), e clones positivos crescidos em placa LB/Amp/IPTG/Xgal foram selecionados. A presença de Hygro foi confirmada por PCR de colônia e um clone se mostrou positivo (Figura 24c). Foi então realizada uma reação de mini-preparação plasmidial (Figura 24d) e seguiu-se para a construção do vetor pGEM-Teasy/E3_Hygro_D3.



Figura 24 - Construção do vetor pMK-RQ-Bb/HDR/Hygro – Gibson Assembly. (a) Amplificação por PCR do fragmento Hygro para Assembly: 1- 1 Kb plus, 2- Hygro; (b) Digestão enzimática do vetor pMK-RQ-Bb/HDR com as enzimas *XhoI* e *NotI*: 1- 1Kb plus, 2- pMK-RQ-Bb/HDR *XhoI/NotI*; (c) PCR de colônia das transformações para confirmação da presença de Hygro: 1- 1 Kb plus, 2- pGEM-Teasy/Hygro, 3- 13 clones foram analisados nas canaletas de 3 a 15, respectivamente; (d) Mini-preparação plasmidial do Clone 4.

Os fragmentos E3 e D3 e a marca de resistência a Hygro (HDR/Hygro), foram amplificados por PCR e bandas nos tamanhos esperados de aproximadamente 1Kb para os braços de homologia E3 e D3 (Figura 25b) e de 2.800pb para HDR/Hygro foram identificados (Figura 25a) e purificados. Bactérias *E.coli* DH5α foram transformadas com a reação de assembly (ver item 3.7.14), e clones positivos crescidos em placa LB/Amp/IPTG/Xgal foram selecionados. A presença do fragmento E3/HDR/Hygro/D3 foi confirmada por PCR de colônia e dois clones se mostraram positivos apresentando tamanho aproximado de 4.800pb (Figura 25c). Minipreparações plasmidiais foram realizadas e a construção foi confirmada por digestão enzimática (Figura 25d) e sequenciamento.



Figura 25 - Construção do vetor pGEM-Teasy/E3/Hygro/D3 - Gibson Assembly. (a) Amplificação por PCR do fragmento HDR/Hygro para Assembly: 1- 1 Kb plus, 2- HDR/Hygro; (b) Amplificação por PCR dos fragmentos E3 e D3 para Assembly: 1- 1 Kb plus, 2- E3, 3- vazio, 4- D3; (c) PCR de colônia das transformações para confirmação da presença de Hygro: 1- 1 Kb plus, 2- 19 clones foram analisados nas canaletas de 2 a 20, respectivamente; (d) Digestão enzimática dos clones 3 e 16 com *SnaBI* e *Pme1*: 1- 1Kb plus, 2- Clone 3 pGEM-Teasy/E3/HygroD3, 3 – Clone 3 pGEM-Teasy/E3/HygroD3 *SnaBI*, 4- Clone 3 pGEM-Teasy/E3/HygroD3 *PmeI*, 5- Clone 3 pGEM-Teasy/E3/HygroD3 *SnaBI*, 8- Clone 16 pGEM-Teasy/E3/HygroD3 *SnaBI*, 9- Clone 16 pGEM-Teasy/E3/HygroD3 *SnaBI/PmeI*.

4.3.7. Construção dos vetores pPIC6A/gRNA_och1/SpCas9 pPICHOLI-

1/gRNA_alg3/SpCas9 – Gibson Assembly

Tentou-se ainda a construção dos vetores para expressão em *P. pastoris* contendo o sistema CRISPR-Cas9 por meio da técnica de Assembly. Porém, apesar de diversas tentativas,

nas poucas colônias que cresceram para ambas as transformações, nenhuma se mostrou positiva.

5. Discussão

Proteínas recombinantes terapêuticas representam uma grande parte da produção de biofármacos no cenário mundial, abrangendo cerca de 300 produtos disponíveis segundo (Laukens *et al.*, 2015). Por apresentarem grande interesse clínico e comercial, o estudo de novas formas de produção em larga escala desses produtos está em constante evolução.

A maioria das proteínas humanas terapêuticas são naturalmente glicosiladas, sendo essa modificação pós-traducional de extrema importância tanto para a conformação, quanto para atividade biológica proteica. A expressão de proteínas em cultura de células de mamíferos apresenta a grande vantagem de possuir um padrão de glicosilação muito semelhante ao humano, porém é um método extremamente dispendioso. Uma alternativa é a expressão de proteínas em leveduras, que podem ser cultivadas em biorreatores apresentando amplo rendimento e baixo custo.

As espécies Saccharomyces cerevisiae e Pichia pastoris são amplamente utilizadas para a produção de proteínas heterólogas, sendo aceitas como GRAS (Generally Recognized as Safe) pelo FDA. Porém, seu padrão de glicosilação de proteínas é diferente do humano, apresentando estruturas hiper-manosiladas que podem ser imunogênicas quando administradas terapeuticamente. A levedura P. pastoris possui algumas vantagens para a expressão de proteínas recombinantes em escala industrial, como a capacidade de secreção proteica, maior rendimento, hipermanosilação menos pronunciada e a possibilidade de manipulação do seu padrão de glicosilação para seguir o padrão complexo humano até o final da via, incluindo a adição de resíduos de ácido siálico, já descrito na literatura (Kavšček *et al.*, 2015).

Visando criar uma linhagem de *P. pastoris* geneticamente modificada, com um padrão de glicosilação humano customizado para produção de glicoproteínas terapêuticas pelo Instituto Butantan, propôs-se a construção do primeiro passo da via de glicosilação humana

(Man₅GlcNAc₂) em *P. pastoris* a partir da técnica de edição de genomas CRISPR-Cas9, para a expressão da proteína terapêutica glucocerebrosidase (GCase) como modelo.

O sistema CRISPR-Cas9 revolucionou o campo de engenharia genética nos últimos anos. Originário do sistema imune de bactérias, ele foi adaptado para a edição de genomas de diferentes organismos, permitindo, de uma maneira simples, a deleção e inserção de genes por uma única enzima, *Sp*Cas9, que realiza uma quebra da dupla fita do DNA alvo, guiada por um pequeno RNA guia quimérico específico (gRNA).

Para obtenção do padrão de glicosilação humanizado, propusemos a deleçãodos genes *alg3* e *och1* de *P. pastoris*, responsáveis pela hiper-manosilação característica desse organismo, como descrito por (Wildt and Gerngross, 2005). A deleção de ambos os genes permite a predominância de estruturas no padrão de glicosilação desejado: Man₅GlcNAc₂.

A GCase humana é uma enzima associada à membrana lisossomal responsável pela degradação do glicolipídeo glucocerebrosídeo. A atividade deficiente da enzima resulta na doença de Gaucher, a mais comum das desordens de depósito lisossomal (Jmoudiak and Futerman, 2005). A forma mais freqüente e menos grave, conhecida como Tipo I, vem sendo tratada com sucesso através da terapia de reposição enzimática (TRE) (Ilan *et al.*, 2009). O tratamento foi aprovado em 1991, usando GCase purificada de placenta humana (Aglucerase, *Ceredase*®), substituída em 1994 pela forma recombinante da enzima, produzida em células CHO (Imiglucerase, *Cerezyme*®). Embora muito eficaz, o tratamento com TRE é extremamente custoso, motivando a busca por terapias alternativas. Atualmente, a enzima velaglucerase alfa (GCase gene-ativada), produzida pela *Shire*, e a taliglucerase alfa (GCase recombinante obtida de células de planta), produzida pela *Protalix* também estão sendo comercializadas.

O Instituto Butantan domina a tecnologia de produção de proteínas recombinantes em sistemas de biorreatores, podendo ser adaptado para o cultivo de *P. pastoris*. A missão pública do

Instituto é atender às necessidades do sistema público de saúde, produzindo soros, biofármacos e vacinas a um custo reduzido para o Sistema Único de Saúde (SUS). Portanto, a proteína GCase humana poderia ser produzida em escala industrial pelo Instituto, a um custo menor em relação aos medicamentos importados, e atender às necessidades dos pacientes brasileiros. Além disso, a partir da criação de uma plataforma de expressão em *P. pastoris* com um sistema de glicosilação humanizado, diversas outras glicoproteínas terapêuticas poderão ser produzidas e comercializadas, fortalecendo a produção nacional e biofármacos e biossimilares.

As enzimas GCase já comercializadas para TRE apresentam resíduos expostos de manose em sua estrutura para se ligar aos receptores de manose presentes na superfície dos macrófagos. O número de resíduos de manose da enzima não influencia sua captação pelos macrófagos nem sua atividade, desde que possua entre 2 a 9 resíduos em sua estrutura (Van Patten *et al.*, 2007). A linhagem de *P. pastoris* com padrão de glicosilação humanizado que se deseja construir apresentará predominantemente 5 resíduos de manose, devendo expressar, portanto, uma forma terapêutica funcional da enzima.

Primeiramente duas linhagens controle foram estabelecidas: uma expressando a proteína GCase com o padrão de glicosilação selvagem de *P. pastoris* e outra expressando apenas a proteína *Sp*Cas9 para avaliação da expressão dessa endonuclease na levedura.

Para a construção da linhagem controle expressando GCase dois vetores de expressão em *P. pastoris* foram construídos com sucesso: pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1, e a avaliação da expressão de GCase por western blot mostrou uma fraca expressão da proteína no sobrenadante de cultura. Como foi utilizada a sequência do gene da GCase humana, a baixa expressão pode ser devido às diferenças na utilização de códons pela levedura *P. pastoris* como descrito por (Sinclair and Choy, 2002). Novos experimentos serão refeitos para avaliação da

expressão de GCase com padrão de glicosilação selvagem de *P. pastoris* utilizando uma sequência códon-otimizada para este organismo.

Também foi construído um vetor para controle de expressão da proteína *Sp*Cas9: pPICHOLI-1/*Sp*Cas9cont. Verificou-se que *Sp*Cas9 é expressa no lisado celular da cultura em clones induzidos com metanol. Aparentemente a proteína não demonstrou ser tóxica para o microrganismo utilizado, mas, para evitar possíveis clivagem em regiões não desejadas na linhagem humanizada, a expressão de *Sp*Cas9 será mediada pela sequência de replicação autônoma de *P. pastoris* (PARS) que permitirá a expressão temporária da endonuclease e dos gRNAs, sem integrá-los ao genoma.

Para a obtenção da linhagem de *P. pastoris* humanizada foi necessária a síntese de diversas construções: um vetor para expressão de *Sp*Cas9 contendo os gRNAs para deleção do gene *alg3* (pPICHOLI-1/gRNA_*alg3/Sp*Cas9), um vetor também para expressão de *Sp*Cas9, porém contendo gRNAs para deleção do gene *och1*(pPIC6A/gRNA_*och1/Sp*Cas9), e dois fragmentos para utilização no reparo por recombinação homóloga (HDR), contendo regiões de 1Kb de homologia à região da dupla quebra e uma marca de resistência ao antibiótico canamicina ou higromicina a ser introduzida no genoma da levedura (E1/Kan/D1 e E3/Hygro/D3) para facilitar a seleção dos clones com os genes deletados e reparados por HDR (Figura 26).



Figura 26 – Esquema de edição gênica proposto pelo sistema CRISPR-Cas9. (a) Os genes *alg3* (em verde) e *och1* (em azul) presentes no genoma de *P. pastoris* serão excisados pelo sistema CRISPR-Cas9 (em destaque) por meio das construções pPICHOLI-1/gRNA_*alg3/Sp*Cas9 e pPIC6A/gRNA_*och1/Sp*Cas9, cada uma contendo gRNAs para as regiões 5'e 3'do gene de interesse que irão guiar a quebra da dupla fita de DNA pela enzima *Sp*Cas9 (indicado pelas setas em vermelho). (b) Deleção dos genes *alg3* e *och1*. (c) Juntamente com o complexo CRISPR-Cas9 serão inseridos templates de reparo contendo genes de resistência aos antibióticos higromicina (em laranja) e canamicina (em vermelho), flanqueados por regiões apresentando 1Kb de homologia a cada região anterior e posterior à quebra da dupla fita: E3 e D3 respectivamente para as regiões esquerda de *alg3* e direita de *alg3*, assim como E1 e D1 para as regiões esquerda de *och1* e direita de *och1*. (d) Inserção dos genes de resistência pelo mecanismo de HDR

Primeiramente, foram feitas diversas tentativas para a construção dos fragmentos de recombinação utilizando-se a ligação dos fragmentos por meio de enzimas de restrição, porém esse processo se mostrou muito pouco eficiente e não foi obtido êxito nas construções. Os fragmentos para HDR foram construídos com sucesso uma vez que se passou a utilizar a técnica de clonagem de Gibson Assembly.

A construção dos vetores contendo o sistema CRISPR-Cas9, pPICHOLI-1/gRNA_*alg3/Sp*Cas9 e pPIC6A/gRNA_*och1/Sp*Cas9, não foram bem sucedidas independente da técnica de clonagem empregada até o momento. Diversas tentativas foram realizadas pela técnica de clonagem clássica, mas ainda poucas tentativas foram feitas pelo método de Assembly.

A baixa eficiência da técnica de clonagem clássica que resultou em um grande número de insucessos nos eventos de clonagem, se deve possivelmente aos erros que podem ocorrer nos inúmeros processos que antecedem o evento de transformação, como a reação de polimerização em cadeia do gene, digestões enzimáticas e etapas de purificação e ligação. Além disso, o fragmento de *Sp*Cas9 é consideravelmente grande, podendo dificultar as reações de ligação, e ainda pode ser tóxico para *E. coli*, o que também explicaria a dificuldade na obtenção de clones positivos para as construções contendo *Sp*Cas9.

O resultado do projeto é até o momento promissor, uma vez que mais estudos de engenharia genética de leveduras utilizando o sistema CRISPR-Cas9 estão sendo publicados. Esforços continuarão sendo feitos para que se obtenha êxito nas construções e que possamos testar o sistema CRISPR-*Sp*Cas9 proposto em *P. pastoris*, visando a expressão de uma GCase recombinante com potencial terapêutico.

6. Conclusões

- Os vetores pPIC3.5/GCaseAlb e pHIL-S1/GCasePHO1 foram construídos com sucesso;
- A proteína GCase foi expressa em *Pichia pastoris*, apresentando um padrão de glicosilação selvagem (massa molecular de aproximadamente 66 kDa quando comparado aos 56kDa da proteína não glicosilada). Apresentou baixos níveis de secreção no sobrenadante de cultura, sendo a sequência sinal de secreção PHO1 um pouco mais eficiente que a da albumina humana (Alb). A códon-otimização da sequência humana para *P. pastoris* pode se mostrar mais eficaz na expressão da proteína no sobrenadante (Sinclair and Choy, 2002);
- O vetor pPICHOLI-1/SpCas9cont foi construído com sucesso;
- A proteína SpCas9 foi expressa em P. pastoris, não aparentando ser tóxica para a levedura;
- A técnica de Gibson Assembly se mostrou mais eficiente na construção dos vetores contendo os fragmentos para recombinação homóloga quando comparado à técnica de clonagem clássica;
- Não foi obtido êxito na construção dos vetores contendo o sistema CRISPR-Cas9 para expressão em *P. pastoris*: pPICHOLI-1/gRNA_*alg3/Sp*Cas9 e pPIC6A/gRNA_*och1/Sp*Cas9. Esforços continuarão sendo feitos para alcançar este objetivo baseado no trabalho publicado recentemente por (Weninger *et al.*, 2016), que mostrou pela primeira vez a eficiência do sistema CRISPR-Cas9 em *P. pastoris*.

7. Referências Bibliográficas

ABE, H. et al. Development of valuable yeast strains using a novel mutagenesis technique for the effective production of therapeutic glycoproteins. **Glycobiology**, v. 19, n. 4, p. 428-36, Apr 2009.

AERTS, J. M. et al. The occurrence of two immunologically distinguishable betaglucocerebrosidases in human spleen. **Eur J Biochem**, v. 150, n. 3, p. 565-74, Aug 1985.

BAILON, P.; WON, C. Y. PEG-modified biopharmaceuticals. **Expert Opin Drug Deliv,** v. 6, n. 1, p. 1-16, Jan 2009.

BARRANGER, J. A.; O'ROURKE, E. Lessons learned from the development of enzyme therapy for Gaucher disease. J Inherit Metab Dis, v. 24 Suppl 2, p. 89-96; discussion 87-8, 2001.

BENNETT, L. L.; MOHAN, D. Gaucher disease and its treatment options. **Ann Pharmacother**, v. 47, n. 9, p. 1182-93, Sep 2013.

BERGMANN, J. E.; GRABOWSKI, G. A. Posttranslational processing of human lysosomal acid beta-glucosidase: a continuum of defects in Gaucher disease type 1 and type 2 fibroblasts. **Am J Hum Genet**, v. 44, n. 5, p. 741-50, May 1989.

BERLEC, A.; STRUKELJ, B. Current state and recent advances in biopharmaceutical production in Escherichia coli, yeasts and mammalian cells. **J Ind Microbiol Biotechnol,** v. 40, n. 3-4, p. 257-74, Apr 2013.

BILL, R. M. Playing catch-up with Escherichia coli: using yeast to increase success rates in recombinant protein production experiments. **Front Microbiol**, v. 5, p. 85, 2014.

BOBROWICZ, P. et al. Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast Pichia pastoris: production of complex humanized glycoproteins with terminal galactose. **Glycobiology**, v. 14, n. 9, p. 757-66, Sep 2004.

BRADY, R. O.; FURBISH, F. S. Enzyme replacement therapy: specific targeting of exogenous enzymes to storage cells.: In Membranes and transport 1982.

BRUMSHTEIN, B. et al. Characterization of gene-activated human acid-beta-glucosidase: crystal structure, glycan composition, and internalization into macrophages. **Glycobiology**, v. 20, n. 1, p. 24-32, Jan 2010.

CHARPENTIER, E.; MARRAFFINI, L. A. Harnessing CRISPR-Cas9 immunity for genetic engineering. **Curr Opin Microbiol**, v. 19, p. 114-9, Jun 2014.

CHOI, B. K. et al. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast Pichia pastoris. **Proc Natl Acad Sci U S A,** v. 100, n. 9, p. 5022-7, Apr 2003.

COHEN, S. N. et al. Construction of biologically functional bacterial plasmids in vitro. **Proc** Natl Acad Sci U S A, v. 70, n. 11, p. 3240-4, Nov 1973.

CONG, L. et al. Multiplex genome engineering using CRISPR/Cas systems. Science, v. 339, n. 6121, p. 819-23, Feb 2013.

CONNOCK, M. et al. The clinical effectiveness and cost-effectiveness of enzyme replacement therapy for Gaucher's disease: a systematic review. **Health Technol Assess**, v. 10, n. 24, p. iii-iv, ix-136, Jul 2006.

CORPORATION, G. Monografia do produto: Cerezyme[®] Imiglucerase para injeção. Terapia de reposição enzimática dirigida aos macrófagos para a doença de Gaucher 2001.

DAVIDSON, R. C. et al. Functional analysis of the ALG3 gene encoding the Dol-P-Man: Man5GlcNAc2-PP-Dol mannosyltransferase enzyme of P. pastoris. **Glycobiology**, v. 14, n. 5, p. 399-407, May 2004.

DE DUVE, C. From cytases to lysosomes.: Fed Proc 1964.

DICARLO, J. E. et al. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. **Nucleic Acids Res,** v. 41, n. 7, p. 4336-43, Apr 2013.

ESVELT, K. M. et al. Concerning RNA-guided gene drives for the alteration of wild populations. **Elife**, p. e03401, Jul 2014.

FABREGA, S. et al. Human glucocerebrosidase: heterologous expression of active site mutants in murine null cells. **Glycobiology**, v. 10, n. 11, p. 1217-24, Nov 2000.

FDA. Human insulin receives FDA approval: FDA Drug Bull 1982.

FERRAZ, M. J. et al. Gaucher disease and Fabry disease: new markers and insights in pathophysiology for two distinct glycosphingolipidoses. **Biochim Biophys Acta**, v. 1841, n. 5, p. 811-25, May 2014.

FERRER-MIRALLES, N. et al. Microbial factories for recombinant pharmaceuticals. **Microb Cell Fact,** v. 8, p. 17, 2009.

FURBISH, F. S. et al. Lysosomal hydrolases as glycoproteins.: Biochim. Biophys 1981.

FUTERMAN, A. H. et al. New directions in the treatment of Gaucher disease. **Trends Pharmacol Sci**, v. 25, n. 3, p. 147-51, Mar 2004.

GAO, Y.; ZHAO, Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. **J Integr Plant Biol**, v. 56, n. 4, p. 343-9, Apr 2014.

GERMAIN, D. P. Gaucher's disease: a paradigm for interventional genetics. **Clin Genet,** v. 65, n. 2, p. 77-86, Feb 2004.

GERSBACH, C. A.; PEREZ-PINERA, P. Activating human genes with zinc finger proteins, transcription activator-like effectors and CRISPR/Cas9 for gene therapy and regenerative medicine. **Expert Opin Ther Targets,** v. 18, n. 8, p. 835-9, Aug 2014.

GOODMAN, M. Market watch: sales of biologics to show robust growth through to 2013.: Nat Rev Drug Discov 2009.

GRABOWSKI, G. A. Gaucher disease and other storage disorders. Hematology Am Soc Hematol Educ Program, v. 2012, p. 13-8, 2012.

GRABOWSKI, G. A.; PETSKO, G. A.; KOLODNY, E. **Gaucher disease**. NY: The Online Metabolic and Molecular Bases of Inherited Disease 2001.

GRACE, M. E. et al. Analysis of human acid beta-glucosidase by site-directed mutagenesis and heterologous expression. **J Biol Chem**, v. 269, n. 3, p. 2283-91, Jan 1994.

HAMILTON, S. R. et al. Humanization of yeast to produce complex terminally sialylated glycoproteins. **Science**, v. 313, n. 5792, p. 1441-3, Sep 2006.

HOU, Z. et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. **Proc Natl Acad Sci U S A,** v. 110, n. 39, p. 15644-9, Sep 2013.

HSU, P. D.; LANDER, E. S.; ZHANG, F. Development and applications of CRISPR-Cas9 for genome engineering. **Cell**, v. 157, n. 6, p. 1262-78, Jun 2014.

ILAN, Y.; ELSTEIN, D.; ZIMRAN, A. Glucocerebroside: an evolutionary advantage for patients with Gaucher disease and a new immunomodulatory agent. **Immunol Cell Biol,** v. 87, n. 7, p. 514-24, Oct 2009.

JACOBS, J. Z. et al. Implementation of the CRISPR-Cas9 system in fission yeast. Nat Commun, v. 5, p. 5344, 2014.

JINEK, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. **Science**, v. 337, n. 6096, p. 816-21, Aug 2012.

JMOUDIAK, M.; FUTERMAN, A. H. Gaucher disease: pathological mechanisms and modern management. **Br J Haematol**, v. 129, n. 2, p. 178-88, Apr 2005.

KACHER, Y. et al. Acid beta-glucosidase: insights from structural analysis and relevance to Gaucher disease therapy. **Biol Chem**, v. 389, n. 11, p. 1361-9, Nov 2008.

KAVŠČEK, M. et al. Yeast as a cell factory: current state and perspectives. Microb Cell Fact, v. 14, p. 94, 2015.

KIM, H.; YOO, S. J.; KANG, H. A. Yeast synthetic biology for the production of recombinant therapeutic proteins. **FEMS Yeast Res**, Aug 2014.

LAUKENS, B.; DE VISSCHER, C.; CALLEWAERT, N. Engineering yeast for producing human glycoproteins: where are we now? **Future Microbiol**, v. 10, n. 1, p. 21-34, 2015.

LOOS, A.; STEINKELLNER, H. IgG-Fc glycoengineering in non-mammalian expression hosts. Arch Biochem Biophys, v. 526, n. 2, p. 167-73, Oct 2012.

MA, J. K.; DRAKE, P. M.; CHRISTOU, P. The production of recombinant pharmaceutical proteins in plants. **Nat Rev Genet,** v. 4, n. 10, p. 794-805, Oct 2003.

MALI, P.; ESVELT, K. M.; CHURCH, G. M. Cas9 as a versatile tool for engineering biology. **Nat Methods**, v. 10, n. 10, p. 957-63, Oct 2013.

MANNING, M. C.; PATEL, K.; BORCHARDT, R. T. Stability of protein pharmaceuticals. **Pharm Res**, v. 6, n. 11, p. 903-18, Nov 1989.

MARTINS, A. M. et al. Recommendations on diagnosis, treatment, and monitoring for Gaucher disease. **J Pediatr,** v. 155, n. 4 Suppl, p. S10-8, Oct 2009.

MARTÍNEZ, J. L. et al. Pharmaceutical protein production by yeast: towards production of human blood proteins by microbial fermentation. **Curr Opin Biotechnol**, v. 23, n. 6, p. 965-71, Dec 2012.

NOVO, J. B. et al. Generation of a Chinese hamster ovary cell line producing recombinant human glucocerebrosidase. **J Biomed Biotechnol**, v. 2012, p. 875383, 2012.

PANDA, A. K. Bioprocessing of therapeutic proteins from the inclusion bodies of Escherichia coli. Adv Biochem Eng Biotechnol, v. 85, p. 43-93, 2003.

PYZOCHA, N. K. et al. RNA-guided genome editing of mammalian cells. **Methods Mol Biol**, v. 1114, p. 269-77, 2014.

RAN, F. A. et al. Genome engineering using the CRISPR-Cas9 system. **Nat Protoc,** v. 8, n. 11, p. 2281-308, Nov 2013.

REDDEN, H.; MORSE, N.; ALPER, H. S. The synthetic biology toolbox for tuning gene expression in yeast. **FEMS Yeast Res**, Jul 2014.

ROMANOS, M. A.; SCORER, C. A.; CLARE, J. J. Foreign gene expression in yeast: a review. **Yeast,** v. 8, n. 6, p. 423-88, Jun 1992.

ROSENBLOOM, B. E.; WEINREB, N. J. Gaucher disease: a comprehensive review. **Crit Rev Oncog**, v. 18, n. 3, p. 163-75, 2013.

RYAN, O. W.; CATE, J. H. Multiplex engineering of industrial yeast genomes using CRISPRm. **Methods Enzymol,** v. 546, p. 473-89, 2014.

SAMPSON, T. R.; WEISS, D. S. Exploiting CRISPR/Cas systems for biotechnology.: Bioessays 2013.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proc Natl Acad Sci U S A**, v. 74, n. 12, p. 5463-7, Dec 1977.

SHALLTIEL, Y.; BARTFELD, D.; HASMUELI, S. Production of glucocerebrosidase with terminal mannose glycans for enzyme replacement therapy of Gaucher's disease using a plant cell system.: Plant Biotechnol J 2007.

SINCLAIR, G.; CHOY, F. Y. Synonymous codon usage bias and the expression of human glucocerebrosidase in the methylotrophic yeast, Pichia pastoris. **Protein Expr Purif,** v. 26, n. 1, p. 96-105, Oct 2002.

SKROPETA, D. The effect of individual N-glycans on enzyme activity. **Bioorg Med Chem,** v. 17, n. 7, p. 2645-53, Apr 2009.

SORGE, J. et al. Molecular cloning and nucleotide sequence of human glucocerebrosidase cDNA. **Proc Natl Acad Sci U S A,** v. 82, n. 21, p. 7289-93, Nov 1985.

STAHL, P. D. et al. Evidence for receptor-mediated binding of glycoproteins, glycoconjugates, and lysosomal glycosidases by alveolar macrophages. **Proc Natl Acad Sci U S A,** v. 75, n. 3, p. 1399-403, Mar 1978.

TEKOAH, Y. et al. Glycosylation and functionality of recombinant β -glucocerebrosidase from various production systems. **Biosci Rep**, v. 33, n. 5, 2013.

TOMITA, E. Y. et al. Isolation of genomic DNA from Pichia pastoris without hydrolases.: Biotecnologia Aplicada 2002.

TSUJI, S. et al. Nucleotide sequence of cDNA containing the complete coding sequence for human lysosomal glucocerebrosidase.: J Biol Chem 1986.

VAN PATTEN, S. M. et al. Effect of mannose chain length on targeting of glucocerebrosidase for enzyme replacement therapy of Gaucher disease. **Glycobiology**, v. 17, n. 5, p. 467-78, May 2007.

VARKI, A. et al. Essentials of Glycobiology. 2nd edition. 2009.

VERVECKEN, W. et al. In vivo synthesis of mammalian-like, hybrid-type N-glycans in Pichia pastoris. **Appl Environ Microbiol**, v. 70, n. 5, p. 2639-46, May 2004.

WALSH, G. Biopharmaceutical benchmarks: Nat Biotechnol 2010.

WENINGER, A. et al. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast Pichia pastoris. **J Biotechnol**, Mar 2016.

WILDT, S.; GERNGROSS, T. U. The humanization of N-glycosylation pathways in yeast. **Nat Rev Microbiol**, v. 3, n. 2, p. 119-28, Feb 2005.

ANEXO 1 – LISTA DE PRIMERS

Função	#	Forward (F) ou Reverse (R)	Primer 5' -> 3'	Tm (°C)	pb
Amplificação GCaseAlb	1	F	GAATTCAATAATGAAGTGGGTAACCTTTATTTC CCTTCTTTTTCTCTTTAGCTCGGCTTATTCCGC CCGCCCCTGCATCCCT	71	83
	2	R	GCGGCCGCTCACTGGCGATGCCACAGGTA	72	30
Amplificação GCasePHO1	3	F	CTCGAGGTGCCCGCCCCTGCATCCCTAAAAGCT TCGGCTACAGCTCGGTGGTGTGTGTCTGCAATG CCACATACTGT	74,6	78
	4	R	CCCGGGTCACTGGCGATGCCACAGGTA	69,2	27
	5	F	GCCCGCCCTGCATCCCTAAAA	64,6	23
	6	R	TGTGCCCGTGTGATTAGC	55,5	19
Sequenciamento GCase	7	F	CACCGAGCCCTGCAGTTG	59,3	18
	8	F	TCCAGCCAAAGCCACCCTA	59	20
	9	R	TCACTGGCGATGCCACAGGTA	61	22
Amplificação SpCaço	10	F	GTCGACATGGACTATAAGGACCAC	56,7	24
Ampinicação Speaso	11	R	GCGGCCGCTTAGAATTCCTTGTACAG	61,6	26
Amplificação de SpCas9	12	F	AAGCTTATGGACTATAAGGACCACG	56,2	25
Amplificação Kon	13	F	GCGGCCGCATGAGCCATATTCAACGG	66	26
Amplificação Kan	14	R	CTCGAGTTAGAAAAACTCATCGAGC	55,5	25
Amplificação do Unaro	15	F	GCGGCCGCTCAGCTTGCATGCCGGTCGAT	72,5	29
Amplificação de Hygro	16	R	CTCGAGATCATACATGAGAATTAAGG	53	26
	17	F	TACGTACGCCAGTACCGGGTAGGCTA	63,9	26
Amplificação E3	18	R	GGATCCAGGAAACGATAGAACCAG	57	24
	19	F	ATCGATTTATGGTTTGACATCGGAGG	57	26
Amplificação D3	20	R	GTTTAAACTTGCCCAAGAGACGGAGC	60	26
	21	F	TACGTAGCCGTAACACGGCTCTTCA	61,4	25
Amplificação E1	22	R	GGATCCTGAGTCCAACCGTGAGCC	63,3	24
	23	F	ATCGATTCTCGGCCGGTGCAGGTA	64,1	24
Amplificação D1	24	R	GTTTAAACAATAGACATAAATTAAAAAGACAC	51,4	32
Sequenciamento Kan	25	F	TGATGCGCTGGCAGTGTT	58	18
	26	R	AACACTGCCAGCGCATCA	58	18
Sequenciamento Hygro	27	F	ACTGTGCCACCAAGCGTA	57,1	18
	28	F	GCGATCCTGCAAGCTCCG	59,3	18
	29	F	ATTGACCGATTCCTTGCG	53	18
	30	F	CTTCTCGACAGACGTCGC	55,5	18
	31	R	TACGCTTGGTGGCACAGT	57,1	18
Sequenciamento Cas9	32	F	GTGGACGACAGCTTCTTC	53,1	18
	33	F	CAAGAGCAGACGGCTGGA	57,5	18
	34	F	CAGAGCAAGAACGGCTAC	53,2	18
	35	F	GCAGATTCGCCTGGATGA	55	18

	36	F	CGGCGTGGAAGATCGGTT	58	18
	37	F	CACGACGACAGCCTGACC	58,5	18
	38	F	CAGGAACTGGACATCAAC	50,3	18
	39	F	GTCCAAGCTGGTGTCCGA	57,4	18
	40	F	AGACCGAGGTGCAGACAG	56,8	18
	41	F	ACTTCCTGTACCTGGCCA	55,8	18
	42	F	GATCGACCTGTCTCAGCT	53,7	18
	43	F	CGACCACATGAAGCAGCA	55,7	18
	44	R	GAAGAAGCTGTCGTCCAC	53,1	18
Sequenciamento gRNAs	45	F	CAGTGACCTCCATTGATA	48,2	18
	46	R	CTAGACTTCAGGTTGTCT	47,8	18
	47	F	GATCCCCCACACCATAGCTTCAA	61,2	25
	48	R	ATGTGTGAGAATGGACCTGTGGATG	59,6	25
	49	F	CATCCACAGGTCCATTCTCACACATAACCGATC	72.9	50
Gibson Assembly		•	GACCGGTCTCGAGCCCA	,.	
s9	50	R	ATCGATAGCTTGCAAATTAAAG	67,2	55
	51	F	ACGTAGTTTAAACCCCACACACCATAGCTTCAA AATGTTTCTACTC	65,1	46
	52	R	TTGAAGCTATGGTGTGTGGGGGGATCCGATAAGC TGGGGGAACATTCGCGA	72	50
	53	F	CTGATGAATATCTTGTGGTAGGGGTTTG	59,6	28
	54	R	ATGTGTGAGAATGGACCTGTGGATG	59,6	25
Cibeen Assembly	55	F	CATCCACAGGTCCATTCTCACACATAACCGATC GACCGGTCTCGAGCCCA	72,9	50
	56	R	CCTTATAGTCCATAAGCTTGATTAATTGTCAAC ACCGCCCCTTAGATTAG	66,3	50
pPICHOLI-	57	F	ATTAATCAAGCTTATGGACTATAAGGACCACGA CGGAG	64	38
1/gRNA_alg3/SpCas9	58	R	ACGTGGCGGCCGCTTAGAATTCCTTGTACAGCT CGTCC	70,5	38
	59	F	CAAGGAATTCTAAGCGGCCGCCACGTCCGACGG CGGCC	74,8	38
	60	R	ACCCCTACCACAAGATATTCATCAGGGGAGCTT GCAAATTAAAGCCTTCGAGCGTCCCAAAACCTT CTCAAGC	72,7	73
Gibson Assembly pGEM-Teasy/ E1/Kan/ D1	61	F	CCATGGCGGCCGCGGGAATTCGATTGCCGTAAC ACGGCTCTTCACCTGTA	74,5	50
	62	R	TGGATCCTACGTATACTTTTCTAAACAGCTGTT GTTGCATTAGG	64,4	44
	63	F	GTTTAGAAAAGTATACGTAGGATCCAAGCTTTT TTTGTAG	60,7	40
	64	R	GCACCGGCCGAGAGTTTAAACATCGATTTAATT AACCTGC	65,9	40
	65	F	TCGATGTTTAAACTCTCGGCCGGTGCAGGTAGT GGAAA	67,2	38
	66	R	GCAGGCGGCCGCGAATTCACTAGTGATAAATAG ACATAAATTAAAAAGACACAGCCAGAATATTAA ACCAGAC	69,3	73
	67	F	TATCACTAGTGAATTCGCGGCCGCCTGC	65,5	28
	68	R	AATCGAATTCCCGCGGCCGCCATGG	66,1	25
Gibson Assembly	69	F	CCATGGCGGCCGCGGGAATTCGATTCGCCAGTA CCGGGTAGGCTAATGAG	75,3	50

pGEM-Teasy/ E3/Hygro/ D3	70	R	TGGATCCTACGTAAGGAAACGATAGAACCAGAG ACTAGGC	66,9	40
	71	F	TCTATCGTTTCCTTACGTAGGATCCAAGCTTTT TTTGTAGAAATG	63,4	45
	72	R	GCATGCAAGCTGAGCGGCCGCATAGTTGTTCAA TTGAT	67,2	38
	73	F	ACTATGCGGCCGCTCAGCTTGCATGCCGGTCGA TCTAG	71,6	38
	74	R	TGCGGCCCTCGAGATCATACATGAGAATTAAGG GAGTCACGTTATGAC	69,1	48
	75	F	TCTCATGTATGATCTCGAGGGCCGCACCGGTTA ACATC	68,3	38
	76	R	TGTCAAACCATAAGTTTAAACATCGATTTAATT AACCTGCAGGACTCGAAC	64,6	51
	77	F	TCGATGTTTAAACTTATGGTTTGACATCGGAGG AGGTTG	63,9	39
	78	R	GCAGGCGGCCGCGAATTCACTAGTGATATTGCC CAAGAGACGGAGCTCTTAGA	72,9	53
	79	F	TATCACTAGTGAATTCGCGGCCGCCTGC	65,5	28
	80	R	AATCGAATTCCCGCGGCCGCCATGG	66,1	25
Gibson Assembly 2 pPIC6/ pMA- RQ/gRNA_och1/SpCas9	81	F	TCGCGAATGTTCCCCCAGCTTATCGGATCCCCC ACACACCATAG	71,8	44
	82	R	TGGGCTCGAGACCGGTCGATCGGTTATGTGTGA GAATGGACCTGTG	72,2	46
Gibson Assembly 2 pPICHOLI/pMA-RQ gRNA_alg3/SpCas9	83	F	TTGCAAGCTCCCGGGCTGGGCCTCACTGATGAA TATCTTGTGGTAGGG	72,5	48
	84	R	CGGTCGATCGGTTAACATGCGGCCTATGTGTGA GAATGGACCTGTG	71,3	46
T7	85	F	TAATACGACTCACTATAGGG	47	20
M13	86	R	CAGGAAACAGCTATGAC	47	17

ANEXO 2 - SEQUÊNCIAS GRNA E RIBOZIMAS

Duas ribozimas serão usadas para flanquear os gRNAs, conforme esquema mostrado na Figura 3. Suas sequências estão apresentadas abaixo, conforme descrito por Gao e Zaho, 2014.

• Sequência ribozima **HHR**:

5' NNNNNCTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTC 3'

Os 6 primeiros nucleotídeos da HHR devem ser complementares aos 6 primeiros nucleotídeos da sequência alvo do gRNA, conforme apresentado na Tabela 3.

• Sequência ribozima HDV:

5'GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCAT GGCGAATGGGAC 3'

ANEXO 3 – OUTRAS TÉCNICAS DE EDIÇÃO GÊNICA

As técnicas de engenharia genética estão em constante evolução, sendo aquelas que visam modificações sítio-específicas no DNA as que apresentam os maiores progressos na área de edição de genomas.

Antes da era CRIPSR/Cas, a edição de sequências gênicas era realizadas por meio da fusão de motivos proteicos contendo um domínio de reconhecimento a uma sequência específica do DNA com um domínio efetor, que catalisa a mudança estrutural ou funcional do gene alvo (Perez-Pinera *et al.*, 2012). O motivo de reconhecimento proteico é baseado na estrutura de proteínas que se ligam ao DNA, como proteínas do tipo *Zing Fingers* (ZF) ou *Transcription Activator-Like Effectors* (TALEs), que quando fusionadas ao domínio de nuclease da enzima *FokI* formam as *Zing Fingers Nucleases* (ZFN) e as *Transcription Activator-Like Effectors Nucleases* (TALEN) e permitem a quebra da dupla fita do DNA de maneira sítio-específica (Tan *et al.*, 2016).

1. ZF

O motivo dedo de zinco Cys₂-Hys₂ é o mais comum dos domínios de ligação ao DNA conhecidos em eucariotos e se baseia na ligação de 30 resíduos de aminoácidos em conformação $\beta\beta\alpha$, na qual a α -hélice se projeta para dentro do sulco maior do DNA e reconhece uma sequência de 3-4 nucleotídeos. Esse domínio foi manipulado de forma a gerar sequências de aminoácidos específicas que reconhecem quase todas as 64 combinações de 3 nucleotídeos existentes. Estes domínios foram então engenheirados de forma que seja possível ligar um domínio ZF a outro, possibilitando sua ligação a uma sequência específica de DNA (Hu *et al.*, 2015). Ver Figura 27a.



Figura 27 – **(a) Domínio** *Zing Finger:* estrutura tridimensional do motivo dedo de zinco Cys₂-Hys₂ com aproximadamente 30 aminoácidos e os domínios de ligação ao DNA em evidência, reconhecendo 3 nucleotídeos; **(b) Domínio** *Transcription Activator-Like Effector:* estrutura tridimensional do motivo TALE com 34 aminoácidos que reconhece apenas um par de base no DNA. (Perez-Pinera *et al.*, 2012).

2. TALE

TALES são pequenas proteínas que se ligam ao DNA produzidas pelo patógeno *Xanthomonas* de plantas para regular a expressão gênica do hospedeiro. As TALEs reconhecem sequências específicas do DNA por meio de domínios de ligação contendo entre 33-35 resíduos de aminoácidos, sendo que cada domínio reconhece um único nucleotídeo. A especificidade da ligação ao DNA é determinada por dois aminoácidos hipervariáveis nas posições 12 e 13 denominados *repeat-variable diresidues* (RVDs). Da mesma forma que nas ZFs, esses aminoácidos foram manipulados para gerar domínios TALEs que reconheçam nucleotídeos específicos e que possam sem ligados, permitindo o reconhecimento de sequências específicas de DNA (Ma and Liu, 2015). Ver Figura 27b.

3. ZFN e TALEN

Os motivos ZF e TALE modificados são fusionados à uma endonuclease para que ocorra a quebra da dupla fita (DSB) do DNA no locus de interesse. A endonuclease *FokI* foi modificada para aumentar sua atividade e especificidade e é o domínio catalítico mais utilizado nas técnicas de ZFN e TALEN (Figuras 38a e 38b, respectivamente).

A nuclease atua como um dímero, sendo necessário duas sequências de proteínas de ligação ao DNA para sequências adjacentes de um mesmo alvo separadas por um região espaçadora, na qual o domínio catalítico da nucleasse possa se dimerizar e clivar o DNA no sítio de interesse (Ma and Liu, 2015).

Ambas as técnicas foram amplamente utilizadas no universo da edição gênica e apresentam suas vantagens e desvantagens. Apesar de reconhecer um número maior de pares de base por motivo, a técnica ZFN possui eficiência variável por depender de um motivo específico Cys₂-Hys₂, e frequentemente requer otimização das sequências utilizadas, necessitando de um longo período de trabalho e um alto investimento. Já a técnica TALEN possui uma eficiência maior e um design mais simples que a ZFN, mais ainda sim é necessário customizar uma ampla sequência proteica para cada alvo do DNA (Perkel, 2014). Apesar de serem técnicas bem estabelecidas e cumprirem a função de modificação sítio específica do DNA ainda saem em desvantagem quando comparadas com técnica de edição de genomas CRISPR-Cas9, que apresenta um design mais simples, mais barato, menos trabalhoso e mais eficaz (Figura 38c).



Figura 28 – (a) *Zing Finger Nuclease:* esquema dos domínios de ligação ao DNA das proteína ZF ligados entre si e fusionados à nuclease *Fok I*, sendo que cada domínio reconhece 3 nucleotídeos na sequência de DNA. Uma ZFN dimérica reconhece e quebra o DNA contendo sítio adjacentes ao local da DSB; (b) Domínio Transcription Activator-Like Effector: esquemas dos domínios de ligação ao DNA das proteínas TALEs ligados entre si e reconhecendo 1 nucleotídeo cada (sendo a especificidade determinada pelos aminoácidos denominados *repeatvariable diresidues* (RVDs)) fusionados à nuclease *Fok I*. Da mesma forma que nas ZFN, apresenta sequências adjacentes que vão permitir a quebra da dupla fita pela forma dimérica da enzima; (c) CRISPR-Cas9: Esquema da técnica de edição de genomas CRISPR-Cas9, contendo a endonuclease Cas9, a sequência de RNA-guia (sgRNA) que irá hibridizar com a sequência alvo, seguida de uma sequência PAM (protospacer adjacent motif). (Adaptado de Hu *et al.*, 2015).

4. Referêncas bibliográficas

HU, J. H. et al. Chemical Biology Approaches to Genome Editing: Understanding, Controlling, and Delivering Programmable Nucleases. **Cell Chem Biol**, n. 23, p. 57-73, Jan 2016.

MA, D.; LIU, F. Genome Editing and Its Applications in Model Organisms. Genom Proteom Bioinf, n. 13, p.336-344, Dec 2015.

PEREZ-PINERA, P. et al. Advances in targeted genome editing. Curr Op Chem Biol, n. 13, p. 268-277, Jul 2012.

PERKEL, J. Editorial: The power and possibilities of genome engineering. CRISPR-Cas9: Engineering a Revolution in Gene Editing. **Science AAAS**, p. 4-6, Dec 2014.

TAN, W. et al. Gene targeting, genome editing: from Dolly to editors. **Transgenic Res**, n. 25, p. 273-287, Jan 2016.

ANEXO 4 - SÚMULA CURRICULAR

1 - Dados pessoais:

Nome: Marcela de Oliveira Vitarelli

Local de data de nascimento: Belo Horizonte, 07/12/1989

2 - Educação:

- Colégio Santa Dorotéia, Belo Horizonte/MG, 2007. Ensino Médio
- 2. Universidade Federal de Viçosa (UFV), Viçosa/MG, 2014. Bacharel em Bioquímica

a. Université d'Evry-Val-d'Essonne, Evry, França, 2013.
Master 1, Sciences du Génome et des Organismes – 16 meses, Bolsista CAPES, Ciência Sem Fronteiras

b. University of West Florida, Pensacola, EUA, 2011.
 BSc Chemistry – 6 meses, Bolsista UFV

 Universidade de São Paulo (USP), São Paulo/SP, 2016 (em andamento). Mestrado em Ciências Biológicas (Bioquímica)

3 - Formação complementar:

- 1. Bolsista Programa de Aperfeiçoamento de Ensino (PAE). Universidade de São Paulo, São Paulo/SP, 2015.
- 2. XVII Reunião Cinentífica Anual. Instituto Butantan, São Paulo/SP, 2015.
- 3. 7º Encontro em Genética e Melhoramento da UFV e 2º Simpósio Internacional em Genética e Melhoramento. Universidade Federal de Viçosa, Viçosa/MG. 2010.
- 4. Curso de Cultivo Celular. Universidade Federal de Viçosa, Viçosa/MG, 2010.
- 5. Curso de Biossegurança. Universidade Federal de Viçosa, Viçosa/MG, 2010.
- Curso de Introdução de técnica de ELISA. Universidade Federal de Viçosa, Viçosa/MG, 2010.

- II Semana Acadêmica de Bioquímica da UFV. Universidade Federal de Viçosa, Viçosa/MG, 2010.
- 8. Workshop Writing Techniques. Real English Center, Viçosa/MG, 2010.
- Seminário Técnicas de Pipetagem Thermo Scientific. Universidade Federal de Viçosa, Viçosa/MG, 2010.
- VI Encontro em Genética e Melhoramento da UFV; I TECGEM; I Simpósio Internacional de Tecnologia em Genética e Melhoramento e I Encontro da APL – Biotec. Universidade Federal de Viçosa, Viçosa/MG, 2009.

4 – Ocupação:

Bolsista de Mestrado, FAPESP, Jan/2014 a Dez/2016.

5 – Resumos publicados em anais de Congressos:

2016 – 12th EWGGD Zaragoza, The European Working Group on Gaucher Disease. Analysis of Human Glucocerebrosidase Expressed in *Pichia pastoris*.

2016 – 45^a Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular. Cas9 Expression in *Pichia pastoris* GS115.

2012 –III Semana de Enfermagem da UFV. Novos campos e possibilidades: a atuação da enfermagem na obtenção de antígenos visando à produção de um kit diagnóstico para dengue.

2011 – 40^a Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular. Expression and Purification of the Splicing-Related Kinase SRPK2.