

RENATO LIMA SENRA

**DESENVOLVIMENTO DE UMA PLATAFORMA DE PRODUÇÃO DE
GLICOPROTEÍNAS POR EXPRESSÃO HETERÓLOGA DE SIALILTRANSFERASE E
INTERFERON-BETA EM *Leishmania tarentolae***

Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Bioquímica Aplicada, para obtenção
do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2018

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

S478d
2018 Senra, Renato Lima, 1992-
Desenvolvimento de uma plataforma de produção de glicoproteínas por expressão heteróloga de sialiltransferase e interferon-beta em *Leishmania tarentolae* / Renato Lima Senra. – Viçosa, MG, 2018.
xiii, 56 f. : il. (algumas color.) ; 29 cm.

Orientador: Tiago Antônio de Oliveira Mendes.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Referências bibliográficas: f. 46-56.

1. Glicoproteínas. 2. *Leishmania tarentolae*. 3. Interferon.
4. Sialiltransferase. I. Universidade Federal de Viçosa.
Departamento de Bioquímica e Biologia Molecular. Programa de Pós-Graduação em Bioquímica Aplicada. II. Título.

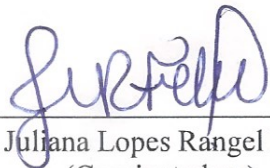
CDD 22. ed. 572.6

RENATO LIMA SENRA

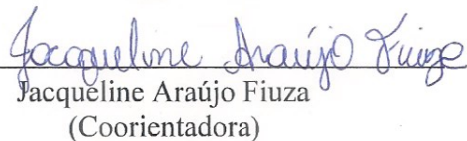
**DESENVOLVIMENTO DE UMA PLATAFORMA DE PRODUÇÃO DE
GLICOPROTEÍNAS POR EXPRESSÃO HETERÓLOGA DE SIALILTRANSFERASE E
INTERFERON-BETA EM *Leishmania tarentolae***

Dissertação apresentada à
Universidade Federal de Viçosa, como
parte das exigências do Programa de
Pós-Graduação em Bioquímica
Aplicada, para obtenção do título de
Magister Scientiae.


APROVADA: 23 de fevereiro de 2018.



Juliana Lopes Rangel Fietto
(Coorientadora)



Jacqueline Araújo Fiuza
(Coorientadora)



Eduardo de Almeida Marques da Silva



Tiago Antônio de Oliveira Mendes
(Orientador)

”O nitrogênio em nosso DNA, o cálcio em nossos dentes, o ferro em nosso sangue, o carbono em nossas tortas de maçã... Foram feitos no interior de estrelas em colapso, agora mortas há muito tempo. Nós somos poeira das estrelas.” (Carl Sagan, Cosmos, 1980)

AGRADECIMENTOS

A Deus, por todas as bênçãos e oportunidades concedidas, por ser meu guia, essência e força.

Aos meus pais que são tudo na minha vida, pelo amor de sempre. Ao meu pai José Renato por todo apoio, tantos sacrifícios e que sempre foi o meu modelo de ser humano, sinônimo de honestidade e integridade. À minha mãe Marli, mulher guerreira, pela dedicação, carinho, e tantos conselhos que nunca me deixaram desistir. Este trabalho leva o nome de vocês também!

Aos meus irmãos Roberto e Rayana, pelo carinho, incentivo, companheirismo e torcida de sempre.

Aos meus avós, tios, primos e toda minha família que são a minha base e raiz que me sustentaram nessa jornada.

Aos meus padrinhos Teitei e Deia e toda família, que sempre fizeram parte da minha vida, agradeço por estarem sempre presente, me apoiando e contribuindo para a construção de tudo isso.

À Universidade Federal de Viçosa (UFV), pela oportunidade de me graduar em um lugar de excelência, que transforma as pessoas e que tem me oferecido cada vez mais experiências enriquecedoras e um ensino de qualidade.

Ao professor e Orientador Dr. Tiago Antônio de Oliveria Mendes pela confiança em me tornar responsável de guiar este projeto. Por toda orientação, paciência, conselhos, disponibilidade e por proporcionar-me um enorme crescimento pessoal e profissional.

Às coorientadoras Prof. Dr^a. Andréa de Oliveira Barros Ribon, Prof. Dr^a Juliana Lopes Rangel Fietto e Dr^a. Jacqueline Araújo Fiuza por toda colaboração, contribuição científica, disponibilidade para tirar minhas dúvidas e agregar excelência a este trabalho.

À Professora Dr^a Maria Cristina Baracat-Pereira, por todos os conselhos, apoio, confiança, amizade e os ensinamentos que foram concedidos durante a minha graduação e que também me auxiliaram durante o mestrado.

Ao Higor, meu amigo, pela parceria de aulas, laboratório e de vida. Por todos os conselhos, por estar sempre presente me apoiando, incentivando, e me ajudando desde o começo quando eu ainda não sabia nada de Biologia Molecular.

Ao Laboratório de Biotecnologia Molecular (LBM) onde fui recebido da melhor forma possível. Uma verdadeira relação de amor e ódio onde na maioria da vezes se tornou a minha segunda casa (ou talvez a primeira). Agradeço a todos os membros que direta ou indiretamente me ajudaram e que sempre estavam dispostos à transmitir o conhecimento. Agradeço pela estrutura que tornou possível a realização desse projeto, pelo cafézinho diário, pelas brocas e pela ótima convivência.

Um agradecimento especial à Amanda, Paty, Gêssica, Gilza e Silvana que me aturaram no desespero diário da rotina de experimentos e que estavam sempre dispostas à ajudar. Com certeza vocês também foram minhas coorientadoras. Muito obrigado pela amizade, conselhos e por tornar tudo isso muito mais leve e prazeroso.

Ao Laboratório de Bioquímica Celular e Bioprodutos por toda colaboração e estrutura onde foi possível realizar os experimentos de cultivo. Agradeço a Tatiana e a Lethícia por todo apoio, paciência, disponibilidade e ensinamentos no cultivo de *L.tarentolae*.

Ao Laboratório de Proteômica e Bioquímica de Proteínas onde iniciei meus aprendizados em bioquímica e que até hoje posso contar com a disponibilidade e colaboração.

À Yaankha e todos alunos de graduação e Iniciação Científica, pela paciência nas minhas tentativas de passar conhecimento e por toda contribuição para este trabalho.

A todos os laboratórios do Departamento de Bioquímica e Biologia Molecular que tive oportunidade de conhecer e trabalhar, pela receptividade, colaboração e todo conhecimento adquirido.

Aos funcionários do Departamento de Bioquímica e Biologia Molecular, obrigado pela trabalho e disponibilidade.

Aos meus amigos que foram fundamentais durante todos dias proporcionando momentos inesquecíveis. Pela amizade, companheirismo, paciência, festas e risadas. Vocês são demais!

À CAPES pelo apoio financeiro e incentivo na realização deste trabalho.

Aos membros da banca pela disponibilidade em avaliarem este trabalho.

A todos que ajudaram, direta ou indiretamente na realização desse trabalho, obrigado.

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS	vii
LISTA DE FIGURAS E TABELAS	x
RESUMO.....	xii
ABSTRACT	xiii
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	2
2.1. Produção Heteróloga de Proteínas Terapêuticas.....	2
2.2. Proteínas Glicosiladas.....	3
2.3. Interferon-Beta.....	5
2.4. Sistemas de Expressão de Proteínas	7
2.5. <i>Leishmania tarentolae</i> como hospedeiro de expressão	8
2.6. Engenharia de proteínas.....	11
3. OBJETIVOS.....	13
3.1. Objetivos Específicos	13
4. MATERIAL E MÉTODOS.....	14
4.1. Técnicas Básicas de Bioquímica e Biologia Molecular.....	14
4.2. Organismos	14
4.3. Genômica	14
4.4. Manutenção e cultura de <i>L. tarentolae</i>	14
4.5. Extração de metabólitos.....	15
4.6. Identificação e quantificação de metabólitos por ¹ H-RMN	15
4.7. Síntese dos genes otimizados e desenho de primers.....	16
4.8. Amplificação dos genes sintéticos do Interferon- β e Sialil-transferase.....	17
4.9. Clonagem dos plasmídeos pGEM T e pGEM T-Easy	17
4.10. Cultivo de <i>E. coli</i>	18
4.11. Preparo de <i>E. coli</i> DH5 α ultra competentes e transformação por choque térmico	19
4.12. Digestão	20
4.13. Clonagem e características dos plasmídeos pLEXY	20
4.14. Estratégias de clonagem.....	21
4.15. Transfecção e seleção clonal de <i>L. tarentolae</i>	23

4.16.	PCR diagnóstico de recombinação dos clones de <i>L.tarentolae</i>	24
4.17.	Análise de transcritos por RT-PCR.....	24
4.18.	Análise da expressão dos clones selecionados por SDS-PAGE 14%.....	25
4.19.	Análise da expressão dos clones selecionados por Dot Blotting	25
5.	RESULTADOS E DISCUSSÃO	26
5.1.	Genômica	26
5.2.	Metabolômica	28
5.3.	Amplificação por PCR dos genes sintéticos otimizados IFN- β e HST	31
5.4.	Digestão dos plasmídeos pGEM-Teasy e pGEM-T clonados	33
5.5.	Digestão dos plasmídeos pLEXY-NEO2 e pLEXY-egfp-SAT2 com enzimas de restrição	33
5.6.	Clonagem dos genes nos plasmídeos digeridos pLEXY	35
5.7.	PCR diagnóstico para verificação de recombinação homóloga.....	38
5.8.	Análise da presença de transcritos por RT-PCR.....	40
5.9.	Avaliação da expressão proteica de IFN- β por SDS-PAGE e Dotblotting.....	42
6.	CONCLUSÃO	43
8.	BIBLIOGRAFIA	46

LISTA DE ABREVIATURAS E SIGLAS

μF	Microfarad
μL	Microlitro
1D-1H	Espectro de uma dimensão com sinal de hidrogênio
Alfa-Gal	Galactose-alfa-1,3-galactose
ANVISA	Agência Nacional de Vigilância Sanitária/Brasil
Asn	Aminoácido asparagina
BHI	Brain heart infusion
BHK	Células de rim de filhotes de hamster (<i>do inglês, baby kidney hamster</i>)
BSA	Albumina de soro bovino
CaCl₂	Cloreto de cálcio
CHO	Células de ovário de hamster chinês (<i>do inglês, Chinese hamster ovary</i>)
CMP	Citidina monofosfato
D.O.	Densidade Óptica
D₂O	Óxido de deutério
DAB	3'-Diaminobenzidine
DMSO	Dimetilsulfóxido
DNA	Ácido desoxirribonucleico
<i>E.coli</i>	<i>Escherichia coli</i>
EPO	Eritropoietina humana
EUA	Estados Unidos da América
FW	Forward
g	Gramma
Glc	Glicose
GlcNAc	N-acetilglicosamina
GNE	UDP-N-acetilglicosamina 2-epimerase
HEPES	Ácido 2--etanossulfônico
HST	Sialiltransferase Humana
IFN- β	Interferon-β
K	Kelvin
KCl	Cloreto de potássio

kDa	Quilodalton (unidade de massa atômica)
L	Litro
<i>L.tarentolae</i>	<i>Leishmania tarentolae</i>
LB	Meio de cultivo Luria Bertani
Man	Manose
ManNAc	N-acetilmanosamina
mg	Miligrama
MHz	Mega-Hertz
mL	Mililitro
mm	Milímetros
MnCl₂	Cloreto de Manganês
NaCl	Cloreto de Sódio
NANP	N-acetilneuraminato-9-fosfatase
NANS	N-Acetilneuraminato Sintase
CMAS	N-acetilneuraminato citidiltransferase
NB1	Nível de Biossegurança 1
Neu5Ac	Ácido N-Acetilneuramínico
Neu5Gc	Ácido N-Glicolilneuramínico
nm	Nanômetro
°C	Graus Celsius
pb	Pares de base (<i>do inglês base pair</i>)
PBS	Tampão fosfato salino
PCR	Reação em cadeia da polimerase
PDB	Protein Data Bank
RMN	Ressonância magnética nuclear
RNA	Ácido ribonucleico
RNA_m	RNA mensageiro
RT-PCR	Reverse transcription polymerase chain reaction
RV	Reverse
<i>S.cerevisae</i>	<i>Saccharomyces cerevisae</i>
SDS-PAGE	Eletroforese em gel de Poliacrilamida na presença de Dodecil Sulfato de Sódio
SOB	Meio Super Optimal Broth

ssu	Subunidade ribossomal menor – 18S (do inglês, small subunit)
SUS	Sistema Único de Saúde
TBS-T	Tampão Tris Salino com Tween
TCA	Ácido tricloroacético
TSP	Trimetilsilil-2,2,3,3-tetradeteropropionato de sódio
UDP-GlicNAc	Uridina difosfato N-acetilglicosamina
V	Volts
WT	Wild Type

LISTA DE FIGURAS E TABELAS

Figura 1. Biossíntese de N-glicanos	4
Figura 2. Estrutura do interferon- β não glicosilado (A) e glicosilado com o carboidrato em evidência na cor azul (B).	6
Figura 3. Arquitetura do sistema LEXSY constitutivo..	9
Figura 4. Comparação do padrão de N-glicosilação nos diferentes sistemas de expressão de proteína recombinante.....	10
Figura 5. Mapa dos vetores de clonagem pGEM-T (A) e pGEM-Teasy (B).	18
Figura 6. Mapa do vetor de expressão constitutiva integrativa pLESXY-neo2 (A) e PLESXY-egfp-SAT2 (B).	21
Figura 7. Esquema da estrutura N-glicanos.....	27
Figura 8. Espectro obtido por RMN em 1D-1H (A) e Esquema de reações envolvidas no metabolismo de biossíntese do ácido siálico (B).	29
Figura 9. Eletroforese em gel de agarose 1% com os produtos da amplificação dos genes IFN- β e HST.	32
Figura 10. Gel de agarose 1% contendo a amplificação dos genes clonados em pGEM-T e pGEM-Teasy..	32
Figura 11. Gel de agarose 0,8% contendo os produtos de digestão do vetor pGEM-Teasy (A) e pGEM-T (B).	33
Figura 12. Gel de agarose 0,8% contendo os produtos de digestão do vetor PLESXY-NEO 2.	34
Figura 13. Gel de agarose 0,8% contendo os produtos de digestão do vetor PLESXY-egfp-SAT2.....	35
Figura 14. Desenho esquemático dos clones construídos após a reação de ligação.....	35
Figura 15. PCR de verificação de integração do gene IFN- β no plasmídeo pLESXY-NEO2 (A) e do gene HST no plasmídeo pLESXY-egfp-SAT2 (B).	36
Figura 16. Eletroforese em gel de agarose 0,8% com o resultado da digestão da construção pLESXY-NEO2+IFN- β , extraída dos clones recombinantes, com a enzima de restrição SwaI.	37
Figura 17. Eletroforese em gel de agarose 0,8% com o resultado da digestão da construção pLESXY-egfp-SAT2+HST, extraída dos clones recombinantes, com a enzima de restrição SwaI.	38

Figura 18. Eletroforese em gel de agarose 1% com o produto reação de PCR diagnóstico..	39
Figura 19. Eletroforese em gel de agarose 0,8% com os produtos da extração de RNA total das culturas de <i>L.tarentolae</i> transfectadas.....	41
Figura 20. Eletroforese em gel de agarose 1% com os produto da amplificação da reação de PCR utilizando os primers específicos para o gene IFN- β (A) e HST (B) a partir do molde de cDNA obtido pelo RT-PCR usando o RNA celular total das células selvagens e dos clones.	41
Figura 21. SDS-PAGE 14% com as amostras de sobrenadante de culturas clonais (IFN- β) e selvagens (WT).....	42
Figura 22. Dot Blotting.	43
Tabela 1. Conjunto de iniciadores utilizados nos experimentos de PCR.	17
Tabela 2. Diferentes estratégias abordadas para a construção dos clones em plasmídeos pGEM-T e pLEXSY.....	22
Tabela 3. Enzimas envolvidas no processo de biossíntese de N-glicanos.....	27
Tabela 4: Metabólitos detectados por RMN e que estão envolvidos no metabolismo de N-glicosilação... ..	29
Tabela 5: Enzimas envolvidas no processo de biossíntese do ácido siálico.	30

RESUMO

SENRA, Renato Lima, M.Sc., Universidade Federal de Viçosa, fevereiro de 2018. **Desenvolvimento de uma plataforma de produção de glicoproteínas por expressão heteróloga de sialiltransferase e interferon-beta em *Leishmania tarentolae*.** Orientador: Tiago Antônio de Oliveira Mendes. Coorientadoras: Andréa de Oliveira Barros Ribon, Juliana Lopes Rangel Fietto e Jacqueline Araújo Fiuza.

A produção de glicoproteínas terapêuticas possui alto custo devido à necessidade de utilização de cultura de células de mamíferos que dependem de meios de cultivo complexos e caros, além de possuírem baixo rendimento de produção. *Leishmania tarentolae*, um protozoário não patogênico para mamíferos, tem sido sugerido como um sistema alternativo para expressão heteróloga de glicoproteínas devido à existência de métodos eficientes de expressão heteróloga, ser facilmente adaptado para produção em alta escala de baixo custo. Além disso, este protozoário apresenta modificações pós-traducionais ausentes em bactérias e leveduras, organismos mais utilizados para produção industrial de proteínas recombinantes. Entre elas, a proteína interferon- β . Esta é uma molécula glicosilada utilizada no tratamento de doenças neurodegenerativas inflamatórias e autoimunes, como a esclerose múltipla e a artrite reumatoide. Embora o perfil de glicosilação de proteínas expressas em *L. tarentolae* seja semelhante ao de mamíferos, existem algumas diferenças no perfil de carboidratos presentes nas terminações das cadeias glicídicas devido à ausência de enzimas de biossíntese e incorporação de ácido siálico. A presente proposta de projeto, portanto, teve como objetivo o desenvolvimento de uma plataforma de expressão de interferon- β utilizando linhagens de *L. tarentolae* selvagem e geneticamente otimizadas para síntese de glicoproteínas com potencial incorporação de ácido siálico na extremidade da cadeia glicídica, proporcionando uma ação mais eficaz das proteínas e reduzindo os custos de produção. As abordagens utilizadas, que visaram entender mais sobre as vias metabólicas encontradas nestes organismos utilizando dados de genômica e metabolômica, indicaram uma alternativa de tornar o sistema de expressão em *L. tarentolae* mais otimizado pela expressão heteróloga de sialiltransferase, enzima ausente nestes organismos. As linhagens geneticamente modificadas foram obtidas por transfecção com genes otimizados para o gênero *Leishmania*. A integração das sequências heterólogas de sialiltransferase e IFN- β foi confirmada por PCR e pela presença de transcritos obtidos por RT-PCR. Ensaios de eletroforese em gel de poliacríamida e marcação de proteínas com anticorpos específicos (anti-his e anti-IFN- β) também confirmaram produção heteróloga de proteína nas linhagens selecionadas.

ABSTRACT

SENRA, Renato Lima, M.Sc., Universidade Federal de Viçosa, February, 2018. **Development of glycoprotein's production platform by heterologous expression of sialyltransferase and interferon-beta in *Leishmania tarentolae*.** Adviser: Tiago Antônio de Oliveira Mendes. Co-advisers: Andréa de Oliveira Barros Ribon, Juliana Lopes Rangel Fietto and Jacqueline Araújo Fiuza.

The production of therapeutic glycoproteins is expensive due to need to use mammalian cell culture that depend of complex and expensive culture media and have low yield of production. *Leishmania tarentolae*, a non-pathogenic trypanosomal protozoan for mammals, has been suggested as an alternative system for heterologous expression of glycoproteins due to the existence of efficient methods for heterologous expression, it is easily adapted for high-scale production using culture media with low cost. In addition, this protozoan presents post-translational modifications absent in bacteria and yeasts, organisms most used for the industrial production of recombinant proteins. Among them, the interferon- β protein is a glycosylated molecule used in the treatment of inflammatory and autoimmune neurodegenerative diseases, such as multiple sclerosis and rheumatoid arthritis. Although the glycosylation profile of proteins expressed in *L. tarentolae* is similar to mammals, there are some differences in the carbohydrate profile present in the endings of the glycidic chains due to the absence of biosynthesis enzymes and the incorporation of sialic acid. The present project proposal aims to develop a interferon- β expression platform using wild type and genetically optimized *L. tarentolae* strains for glycoprotein synthesis with potential sialic acid incorporation at the end of the glycine chain providing a more effective action of proteins with reduced costs. The approaches used to understand more about the metabolic pathways found in these organisms using genomic and metabolomic data indicated an alternative to make the *L.tarentolae* expression system more optimized by the heterologous expression of sialyltransferase, an enzyme absent in these organisms. The genetically modified strains were obtained by transfection with genes optimized for the genus *Leishmania*. Integration of the heterologous sialyltransferase and IFN- β sequences was confirmed by PCR and the presence of transcripts obtained by RT-PCR. Polyacrylamide gel electrophoresis and labeling of proteins with specific antibodies (anti-his and anti-IFN- β) also confirmed heterologous protein production in the selected lineages.

1. INTRODUÇÃO

Embora glicoproteínas e proteínas oligoméricas apresentem elevado potencial biotecnológico na terapêutica, indústria de alimentos e outras, sua produção em alta escala é comprometida. Devido à sua complexidade bioquímica, a melhor forma de produção é pela obtenção de proteínas recombinantes em organismos geneticamente modificados. Os sistemas mais favoráveis para cultivo em escala industrial, com alto rendimento do produto e curto período de produção, são baseados em bactérias e leveduras. Bactérias não são capazes de realizar glicosilação. Apesar de possuírem parcialmente a maquinaria necessária, leveduras produzem perfil de glicosilação diferente de eucariotos superiores e glicoproteínas que necessitem de um padrão de glicosilação específico de mamíferos produzidos neste organismo falham em atividade biológica e estabilidade. Além disso, ambos os sistemas podem produzir incorreto dobramento de proteínas que prejudiquem sua atividade. Células de mamíferos produzem proteínas com correto perfil de glicosilação e dobramento, porém ambas possuem dificuldade de cultivo em grande escala, longo tempo de cultivo necessário e baixa produtividade de proteínas. Dentro deste contexto, uma boa opção seria a utilização de *Leishmania tarentolae* para expressão heteróloga de glicoproteínas e proteínas oligoméricas. Este eucarioto é apatogênico para humanos e a manipulação genética deste organismo necessita de laboratório de segurança NB1 igual a bactérias e leveduras, é passível de cultivo em larga escala com alto rendimento de produção de proteínas. Além disso, o processo de glicosilação e formação de pontes dissulfeto em *L. tarentolae* são muito semelhantes aos processos que ocorrem em mamíferos e existe vetor comercial para manipulação genética como ocorre para bactérias e leveduras. Apesar destas vantagens, o perfil de glicosilação é semelhante, mas não idêntico ao de mamífero devido à ausência de enzimas da via de biossíntese de ácido siálico e existem poucos estudos sobre otimização de produção heteróloga de proteínas neste organismo.

Atualmente existem técnicas computacionais que guiam a otimização de produção do biocomposto de interesse, além de protocolos e vetores que possibilitam engenharia genética de *Leishmania*. O interferon- β foi selecionado como um primeiro alvo com validada atividade terapêutica para avaliar o potencial das novas linhagens a serem utilizadas para expressão heteróloga de proteínas. Essa proteína é utilizada para o tratamento de doenças que causam impacto econômico tanto para sistemas públicos e privados de saúde devido à alta demanda e a seu alto custo de produção. Como depende de

glicosilação com padrão específico de mamíferos para estabilidade e aumento da atividade da proteína, sua correta síntese utilizando uma plataforma de expressão mais otimizada proporcionaria uma ação mais eficaz nos tratamentos que utilizam esse medicamento, aumentando o bem-estar de pacientes por diminuir efeitos adversos além de reduzir custos do produto. A produção com padrão correto de glicosilação diminui a quantidade de vezes que a proteína precisa ser administrada ao paciente por injeção e aumenta a estabilidade e o tempo de meia-vida desta proteína no corpo humano.

Por meio da otimização direcionada por modelos computacionais com a adição dos genes de mamíferos importantes para a biossíntese e incorporação de ácido siálico e posterior validação experimental utilizando metodologias de engenharia genética, as linhagens de *L. tarentolae* geradas poderão ser utilizadas para produzir centenas de outras glicoproteínas em larga escala com aplicações nas mais diferentes áreas industriais como farmacêutica, de alimentos, de biocombustíveis, entre outras.

2. REVISÃO DE LITERATURA

2.1. Produção Heteróloga de Proteínas Terapêuticas

A produção industrial de proteínas é uma área em expansão que atinge diversos segmentos de mercado com produtos proteicos com potencial para serem utilizados na produção de biocombustíveis mais sustentáveis, no uso na indústria alimentícia, de papel e celulose e na terapêutica, representada por diversos fármacos (Rabert et al., 2013; Walsh, 2014). Muitos dos medicamentos atuais, como anticorpos, antibióticos e hormônios são compostos por proteínas, e observa-se o uso crescente desses produtos terapêuticos no combate de diversas doenças (Vega 2016). Esse mercado tem gerado uma receita de mais de 140 bilhões de dólares (Walsh 2006, Lalonde and Durocher 2017), com mais de 240 proteínas aprovadas para uso terapêutico nos Estados Unidos e União Europeia e 500 em fase de triagem clínica (Russel, Wang and Rathore, 2007; Walsh, 2014)(Usmani, Bedi et al. 2017).

No passado, a produção de proteínas para suprir a necessidade das indústrias era realizada primariamente por processos laboriosos, caros e de baixo rendimento pela extração do produto de sua fonte natural como tecidos humanos, animais e vegetais (Gräslund et al., 2008; Nehete et al., 2013). O desenvolvimento da tecnologia do DNA recombinante e o progresso feito na otimização de organismos para produção heteróloga de

proteínas oferecem uma variedade de alternativas para produção de proteína com alto rendimento para uso em processos industriais com novas ou melhores propriedades (Gräslund et al., 2008).

2.2. Proteínas Glicosiladas

Mais de um terço dos produtos proteicos terapêuticos são glicoproteínas (Albrecht, 2014) e estima-se que a taxa de crescimento anual deste mercado seja de 26% (Gerngross, 2004; Mistry, Wraight and Cox, 1996; Walsh, 2006, 2014). Glicosilação é a mais prevalente e a mais estruturalmente diversa modificação pós-traducional encontrada em proteínas, presente em aproximadamente metade das proteínas de eucariotos. A ocorrência de glicosilação em proteínas é crítica para vários processos biológicos incluindo adesão celular, interação proteína-proteína e funções importantes em receptores e transdução de sinais (Dwek, 1998; Dwek et al., 2002; Albrecht, 2014).

Glicosilação também é considerada uma das mais complexas modificações pós-traducionais porque envolve um grande número de reações enzimáticas altamente ordenadas (Spiro, 2002; Wieland, 1988). De acordo com o átomo dos resíduos de aminoácidos ao qual os glicídios são ligados, pode-se classificar o processo de adição de carboidratos a proteínas em *O*-glicosilação e *N*-glicosilação. Em *O*-glicosilação, monossacarídeos são ligados a hidroxilas preferencialmente dos aminoácidos serina e treonina e o processo pode ocorrer em diferentes compartimentos como retículo endoplasmático, complexo de golgi, citosol e núcleo (Zauner et al., 2012). A *N*-glicosilação é extensivamente mais frequente que *O*-glicosilação (Gerngross, 2004). Nesse tipo de glicosilação, os carboidratos são ligados a grupos aminos dos resíduos de asparagina com processo inicialmente ocorrendo no retículo endoplasmático e finalizando no complexo de golgi (Figura 1) (Kukuruzinska and Lennon, 1998).

Do ponto de vista de utilização terapêutica de proteínas, glicosilação é importante tanto para atividade biológica das mesmas como para alteração das propriedades farmacocinéticas de produtos de interesse médico. Por esse motivo, muitas proteínas glicosiladas são produzidas por sistemas de expressão eucarióticos, que possuem o padrão de glicosilação semelhante ao de humano, como células imortalizadas de rim de hamster neonato (BHK-21), linhagens de células de camundongos como mieloma e hibridoma (NS0 e Sp2/0) e células de ovário de hamster chinês (CHO), que são responsáveis por

cerca de 50% da produção das proteínas terapêuticas (Yoo, Chintalacharuvu et al. 2002, Demain and Vaishnav 2009, Ghaderi, Zhang et al. 2012).

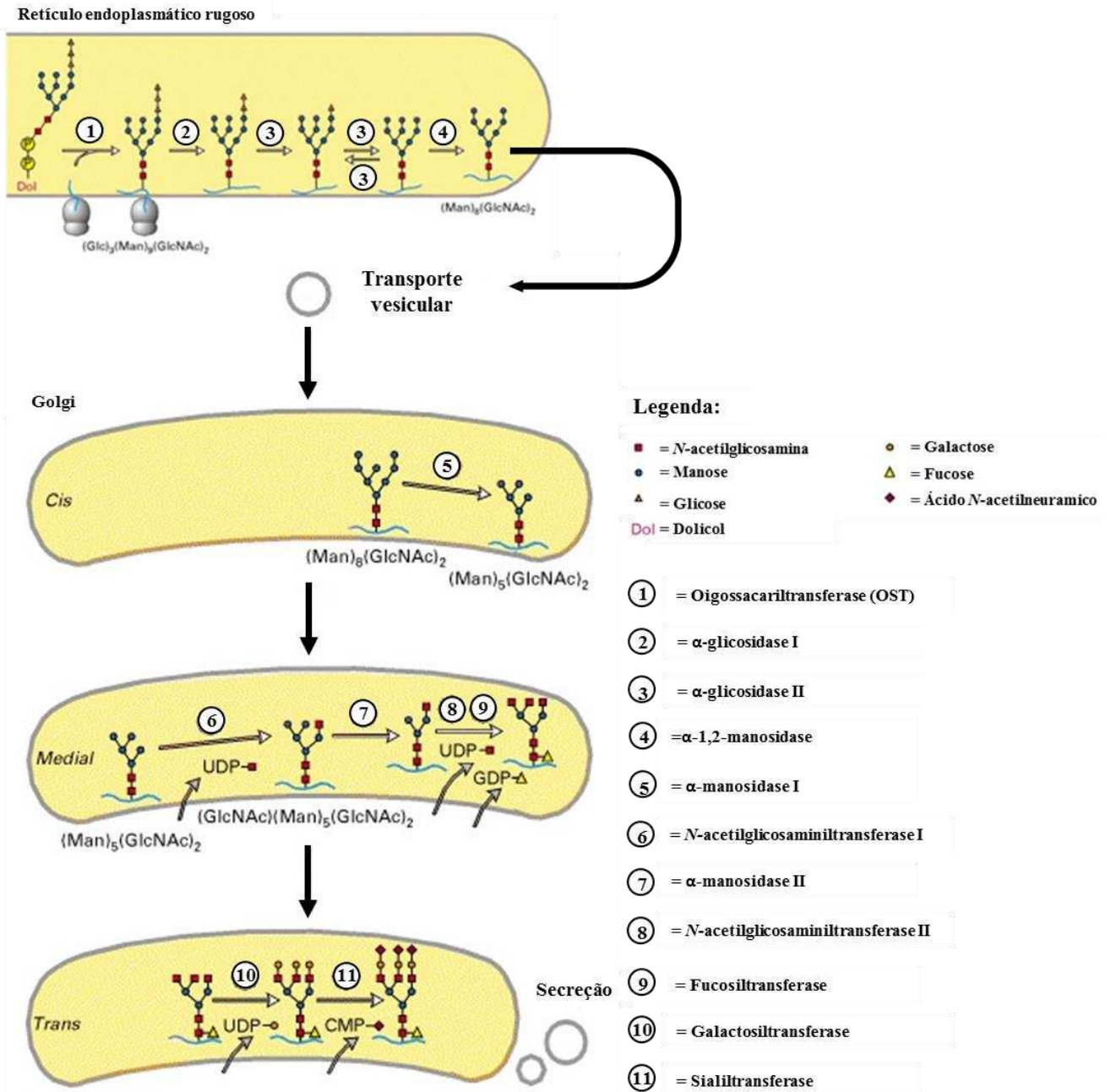


Figura 1. Biossíntese de N-glicanos: O precursor $\text{Glc}_3\text{Man}_9(\text{GlcNAc})_2$ é transferido do transportador de dolicol para uma proteína recém traduzida assim que um resíduo de aminoácido de asparagina (em uma sequência consenso de Asn-X-Ser/Thr) passa para o lúmen do retículo endoplasmático rugoso (1). Posteriormente, reações, retiram resíduos de glicose e manose (1-5). A re-adição de um único resíduo de glicose (3) faz parte do processo de controle do ER. Após estas reações, a proteína glicosada é transportada por vesícula até o Golgi. Após a remoção de três resíduos de manose na região cis, a proteína se move para região medial e remove-se mais dois resíduos de manose (5-7), são adicionados então três resíduos de N-acetilglicosamina e um de fucose (6-9). O processamento é completado na região trans com a adição de três resíduos de galactose (10) e por fim um resíduo de ácido N-acetilneurâmico (ácido siálico) a cada um dos

resíduos de galactose (11). Figura adaptada, fonte: Lodish H, Berk A, Zipursky SL, et al. Molecular Cell Biology. 4th edition.

A capacidade da eritropoietina humana (EPO), hormônio que regula na medula óssea a produção de eritrócitos, de controlar os níveis de eritrócitos e ser utilizada para o tratamento de anemias ser eliminada quando é removido o sítio de *N*-glicosilação (Takeuchi et al., 1989) é um exemplo da importância de glicosilação na atividade biológica de moléculas. Além disso, imunoglobulinas como IgG humanas tem recebido destaque no mercado terapêutico nos últimos anos, onde mais de 50% dos novos compostos terapêuticos de natureza glicoproteica aprovados nos Estados Unidos e Europa são anticorpos monoclonais (Walsh, 2006, 2014). Estas imunoglobulinas também possuem atividade dependente de glicosilação no qual se destaca um conservado sítio de *N*-glicosilação localizado no domínio C γ 2 (Abès and Teillaud, 2010; Glockshuber, Schmidt and Plückthun, 1992; Hristodorov, Fischer and Linden, 2013).

2.3. Interferon-beta

Os Interferons (IFNs) são proteínas glicosiladas da família das citocinas produzidas como resposta do organismo de muitos vertebrados a diversos agentes como vírus, bactérias, parasitas e células oncogênicas. Das três principais classes de interferon, o Interferon-beta (IFN- β) se inclui na do tipo I e é conhecido por exercer diversas funções como antiviral, antisséptica, antioncogênica, antiproliferativa e imunomoduladora (Dissing-Olesen, Thaysen-Andersen et al. 2008).

O uso terapêutico dessa proteína inclui a forma glicosilada ou IFN- β 1a (Rebif® ou Avonex®) produzida por células de ovário de hamster chinês (CHO) e a forma não glicosilada IFN- β 1b (Betaferon®) produzida por bactérias (Gerngross 2004, Han, Koo et al. 2009). Ambas as formas variantes do interferon-beta recombinante humano ((hr)IFN- β) são aprovadas e utilizadas para tratamento de doenças neurodegenerativas inflamatórias e auto imunes como esclerose múltipla (Goldenberg, 2012) e artrite reumatoide (Holten, van, Plater-Zyberk and Tak, 2002). Entretanto, a forma não glicosilada demanda uma quantidade muito maior de proteína, quando comparada com a forma glicosilada para alcançar o mesmo efeito e eficácia (Bertolotto et al., 2001). Observou-se uma redução significativa de 10 vezes na resposta de IFN- β glicosilado nativo quando submetido a diferentes glicosidases (Dissing-Olesen, Thaysen-Andersen et al. 2008, Han, Koo et al.

2009). Estudos comprovam que frações de IFN- β sem a sialização na extremidade da glicosilação tem sua atividade reduzida quando comparada com a sua forma inicial, sendo relacionado com esse efeito a dependência desse carboidrato na interação com os receptores de IFN- β (Mastrangeli, Rossi et al. 2015).

A proteína IFN- β 1 possui em torno de 18 kDa sem a glicosilação e aproximadamente 25 kDa quando glicosilada (Han, Koo et al. 2009). É composta por 166 resíduos de aminoácidos na sua forma madura e contém um único sítio de glicosilação na posição 80 em um resíduo de asparagina (Dissing-Olesen, Thaysen-Andersen et al. 2008) e uma ponte dissulfeto entre a cisteínas nas posições 31(Cys31) e 141 (Cys141) (Lennon, Cherry et al. 1983).

Quanto à melhoria de propriedades farmacocinéticas, a produção de interferon- β glicosilado proporciona o aumento da meia vida da proteína comparada com a isoforma não glicosilada devido ao aumento do volume hidrodinâmico (Figura 2) e consequente diminuição da eliminação renal (Gerngross, 2004). Além disso, foi observado um aumento na termoestabilidade quando comparadas na mesma concentração, com a forma glicosilada desnaturando em cerca de 5°C a mais que a forma não glicosilada (Runkel, Meier et al. 1998). Esse aumento da estabilidade pode melhorar o rendimento na produção e reduzir a frequência na qual a proteína deve ser injetada no paciente, reduzindo custos e efeitos adversos que são comumente observados.

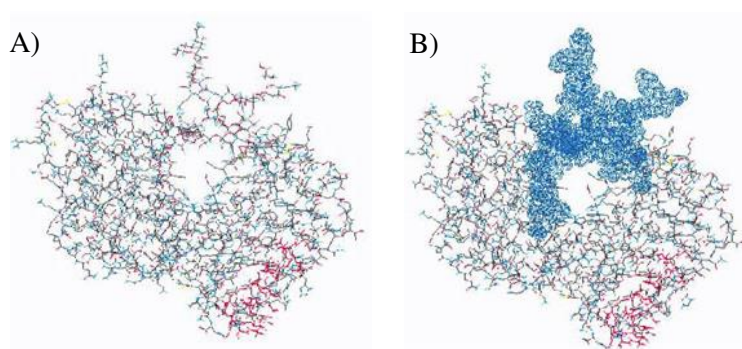


Figura 2. Estrutura do interferon- β não glicosilado (A) e glicosilado com o complexo N-glicano em evidência na cor azul (B). A ilustração demonstra o impacto que o glicano exerce na molécula aumentando seu volume hidrodinâmico (Gerngross 2004). Fonte: <http://www.dkfz.de/spec/glycosciences.de/modeling/glyprot/php/main.php>.

O INF- β glicosilado está na lista dos principais fármacos vendidos no mundo, com uma receita de aproximadamente 3 bilhões de dólares e tendo uma taxa de crescimento de 3% em 2013 (EvaluatePharma, 2014). O produto está incluso na lista de Componente

Especializado da Assistência Farmacêutica (CEAF) do SUS e apresentam um custo que varia de aproximadamente R\$ 4.000,00 a R\$ 8.000,00 para compras públicas e podendo chegar a quase R\$ 10.000 para o consumidor no estado de Minas Gerais (ANVISA, 2017).

2.4. Sistemas de Expressão de Proteínas

Cada método ou sistema de expressão apresenta características, vantagens e limitações diferentes e na escolha de qual utilizar deve-se levar em consideração a natureza da proteína que se deseja expressar e a qualidade que se espera obter da molécula produzida. Um bom organismo com potencial utilização como sistema de expressão de proteínas recombinantes deve ser eficiente e produzir a proteína de interesse em grandes quantidades, ter custo reduzido e possibilidade de cultivo em grandes quantidades em equipamentos de escala industrial como fermentadores, além de permitir facilidade e rapidez de purificação do produto final (Gerngross, 2004).

Sistemas de expressão em bactérias, representada principalmente por *Escherichia coli* (Figura 4), são os mais utilizados por serem facilmente manipulados, apresentarem rápida expressão de proteínas, cultivo barato e viabilidade de cultivo em alta escala (Gräslund et al., 2008). *E. coli* tem sido utilizada para produzir mais de 88% de todas as cadeias proteicas descritas no Protein Data Bank (PDB), o que demonstra a importância, a facilidade e a versatilidade desse organismo para produção de proteínas recombinantes (Fernandez and Vega 2013). Porém, esse sistema falha na realização de modificações pós-traducionais como pontes dissulfetos e glicosilação por ausência de maquinaria apropriada, além da possibilidade de gerar endotoxinas (Demain and Vaishnav 2009). Além disso, proteínas hidrofóbicas ou com alta massa molecular são frequentemente expressas em corpúsculos de inclusão, o que inviabiliza o processo de purificação em condições não desnaturantes (Mus-Veteau, 2002; Palomares, Estrada-Mondaca and Ramírez, 2004; Yoon, Kim and Kim, 2010). Por isso, esse meio de expressão tem sido escolhido para produção de proteínas não muito complexas como insulina e hormônio de crescimento humano (Dingermann 2008). Apesar da existência de proteínas que são naturalmente glicosiladas, como IFN- α , β e γ , interleucinas e Tumor Necrosis Factor- α (TNF- α) e que são produzidas por *E.coli* para fins terapêuticos, a eficiência desses medicamentos é muito reduzida quando comparada com a forma glicosilada produzida por outros sistemas (Dingermann 2008).

Leveduras como *Saccharomyces cerevisiae* e *Pichia pastoris* fazem parte do segundo sistema de expressão mais utilizado por também serem facilmente manipuláveis, permitirem produção em escala industrial e ainda apresentarem maquinaria para diversas modificações pós-traducionais (Gerngross, 2004; Rabert et al., 2013). Devido às leveduras possuírem um padrão de glicosilação diferente de mamíferos e outros eucariotos superiores, a maior dificuldade de utilização deste sistema está relacionada a produção de proteínas que necessitam de um padrão específico de *N* e *O*-glicosilação para possuírem atividade biológica (Berger, Kaup and Blanchard, 2012; Gerngross, 2004). Tanto sistemas de bactérias como de levedura também podem gerar um dobramento incorreto das proteína que podem comprometer sua atividade (Palomares, Estrada-Mondaca and Ramírez, 2004).

Esses problemas de padrão de glicosilação e dobramento de proteínas podem ser resolvidos utilizando células de inseto ou células de mamífero, porém estes dois sistemas de expressão falham no escalonamento industrial, seja por dificuldade de cultivo de suspensão de célula em fermentadores ou devido ao tempo necessário para multiplicação e expressão de proteínas que contribuem para elevar os custos de produção (Dalton and Barton, 2014; Douris et al., 2006). Além disso, por mais semelhanças que células de outros mamíferos apresentam com células humanas, os humanos não possuem maquinaria de biossíntese de um glicano terminal chamado galactose-alfa-1,3-galactose (alfa-Gal) e do principal ácido siálico de mamíferos, *N*-glicolilneuramínico (Neu5Gc). Estas diferenças no padrão de glicosilação podem causar efeitos imunogênicos quando proteínas são produzidas por sistemas baseados em células de mamíferos não humanas (Ghaderi, Zhang et al. 2012).

2.5. *Leishmania tarentolae* como hospedeiro de expressão

Recentemente, *L. tarentolae* tem sido sugerida como um sistema para expressão heteróloga de proteínas. *L. tarentolae* é um protozoário tripanossomatídeo parasito de lagartos *Tarentola mauritanica* do mesmo gênero de espécies de *Leishmania* que causam leishmaniose em mamíferos incluindo humanos. O genoma de *L. tarentolae* foi sequenciado (Raymond et al., 2012) e, por ser eucarioto, este organismo possui maquinaria para diversas modificações de proteínas realizadas após o processo de tradução, inclusive com padrões de modificações pós-traducionais mais similares aos de células de mamíferos se comparado com leveduras (Breitling et al., 2002; Raymond et al., 2012). Organismos do gênero *Leishmania* são conhecidos por possuírem muitas glicoproteínas, que podem chegar

a mais de 10% do total de proteínas (Ferguson 1997, Basile and Peticca 2009). Essa característica se deve provavelmente ao estilo de vida parasitário dos tripanossomatídeos, já que a maioria dos representantes dessa família parasitam mamíferos e apresentam padrão de glicoproteínas semelhantes (Breitling, Klingner et al. 2002).

L. tarentolae também exibe rápida multiplicação celular em cultura, os meios de cultura não exigem suplementações complexas como exigidas por células de mamífero e de insetos, seu cultivo pode ser facilmente escalonado para fermentadores industriais com volumes de 100 litros, com densidade de $0,65-1 \times 10^9$ células/mL (Meehan, Lundberg et al. 2000) e esta espécie apresenta um tempo de duplicação celular de 6-8 horas (Simpson, Frech and Maslov, 1996). Diferente das espécie causadoras de leishmaniose, *L. tarentolae* é não patogênica para mamíferos (Raymond et al., 2012) e exige estrutura laboratorial para cultivo com nível de segurança igualmente aplicado para bactérias e leveduras não patogênicas que são usuais sistemas de expressão heteróloga (Basile and Peticca, 2009; Breton et al., 2005).

O sistema LEXSY (Leishmania Expression System), comercializado pela empresa Jena® e desenvolvido por Breitling em 2002, disponibiliza vetores de expressão para *L. tarentolae* e possui modelos constitutíveis e induzíveis para expressão intracelular ou secretada. A expressão de forma constitutiva acontece por meio da integração do gene de interesse ao locus da subunidade 18S do DNA ribossômico (ssu) e é controlada pela RNA polimerase I endógena (Figura 3) (JenaBioscience, 2013).

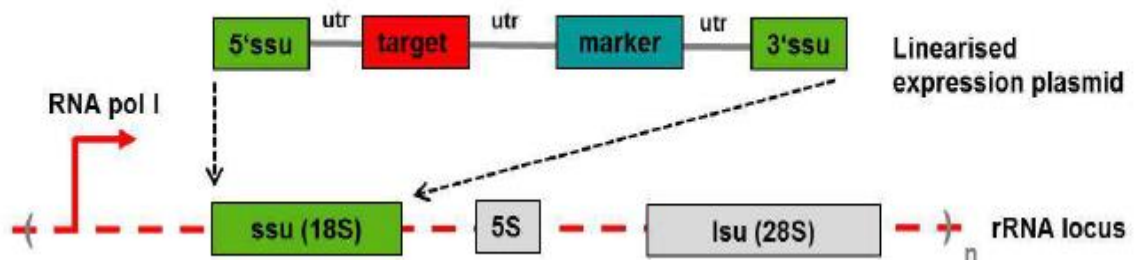


Figura 3. Arquitetura do sistema LEXSY constitutivo. Os genes de interesse são inseridos ao sítio alvo do genoma. Após a linearização com enzima de restrição, o plasmídeo é introduzido no hospedeiro de expressão (Cepa P10) por eletroporação e é integrado por recombinação homóloga em uma das cópias dos genes 18S rRNA (ssu). Para seleção de clones, ainda é disponibilizado marcadores alternativos de resistência aos antibióticos. Fonte: Manual LEXSYcon2 Expression Kit - Jena Bioscience, 2015).

A principal vantagem de expressão heteróloga de proteínas em *L. tarentolae* é sua capacidade para gerar dobramento e modificações pós-traducionais de proteína como formação de pontes de dissulfeto e glicosilação semelhante a mamíferos (Basile and Peticca, 2009; Hemayatkar et al., 2010; Raymond et al., 2012). Entre proteínas que já foram expressas neste sistema e necessitam de modificações pós-traducionais para serem ativas estão eritropoietina (Breitling et al., 2002), proteína ativadora de plasminogênio tecidual (Hemayatkar et al., 2010) e fator VII da coagulação (Mirzaahmadi et al., 2011). Embora o perfil de glicosilação de proteínas expressas em *L. tarentolae* seja semelhante ao processo que ocorre em mamíferos, existem algumas diferenças no perfil de carboidratos presentes nas terminações das cadeias glicídicas devido à ausência de enzimas de biossíntese e incorporação de ácido siálico nos tripanossomatídeos (Ghoshal and Mandal, 2011; Ghoshal et al., 2010). Espera-se que significantes avanços no desenvolvimento de novas linhagens e vetores, melhoria de técnicas de manipulação genética, melhor entendimento da biologia de espécies de *Leishmania* contribuam para expansão das vantagens do uso deste sistema para expressão heteróloga em escala industrial (Basile and Peticca, 2009).

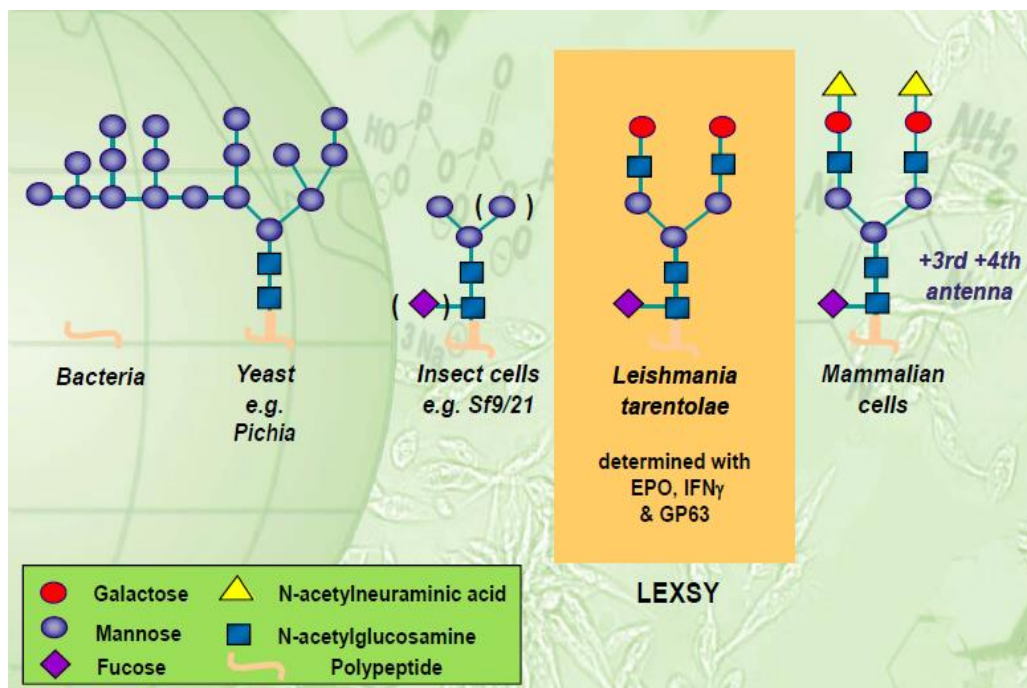


Figura 4. Comparação do padrão de N-glicosilação nos diferentes sistemas de expressão de proteína recombinante. Em evidência no quadro alaranjado o perfil de glicosilação encontrado em *L. tarentolae*, semelhante ao de células de mamíferos com exceção da adição de um ácido N-acetilneuramínico (ácido siálico). (Fonte: Manual LEXSYcon2 Expression Kit - Jena Bioscience, 2015).

2.6. Engenharia de proteínas

Nas últimas décadas houve expansão de métodos que possibilitam modificar as propriedades das proteínas, permitindo acúmulo de conhecimento sobre mecanismo funcional e relação estrutura-atividade destas macromoléculas (Baker, 2011). Esses métodos também podem ser aplicados para modificar proteínas de forma a obter novas ou melhoradas capacidades, em que o planejamento racional de modificação de proteínas é denominado engenharia de proteínas. Dentro do contexto de expressão de macromoléculas de interesse biotecnológico em alta escala, tanto proteínas recombinantes obtidas como produto desejado como proteínas que pertencem a vias metabólicas importantes para obtenção do produto final podem ser otimizadas. Um exemplo de modificação é a adição de peptídeos sinais à região amino terminal de proteínas para direcionar para vias de secreção ou envio para outro compartimento subcelular. Proteínas secretadas são mais facilmente purificadas pela menor complexidade dos meios mínimos de cultura geralmente utilizados comparado ao ambiente intracelular e, além disso, protegem as proteínas recombinantes da degradação por proteases intracelulares (Yoon, Kim and Kim, 2010). Quimeras de proteínas contendo dois ou mais sítios catalíticos de diferentes enzimas que catalisam reações subsequentes em uma mesma via bioquímica podem ser desenhadas para garantir expressão de ambas as enzimas e restringir as reações ao mesmo compartimento subcelular, assegurando que produtos da primeira reação sejam substratos da reação subsequente (Jermutus et al., 2001; Sun, Geng and Shamoo, 2006). Alteração de sequências de amino ácidos em regiões não essenciais para função da proteína podem aumentar a estabilidade da estrutura contribuindo para a produção de proteínas com maior tempo de meia vida (Kumar, Tsai and Nussinov, 2000; Teng, Srivastava and Wang, 2010).

Diversas técnicas têm sido desenvolvidas para alteração da sequência primária de proteínas e consequentemente alteração de sua estrutura e função como mutação sítio dirigida, mutações aleatórias, síntese química de genes codificadores de proteínas, adição de modificações pós-traducionais, entre outras (Bornscheuer and Kazlauskas, 2011; Kazlauskas and Bornscheuer, 2009). Mutações sítio-dirigidas apresentam vantagens sobre mutação aleatória por permitir maior controle das alterações, inclusive impedindo modificações de aminoácidos essenciais para função da proteína (Muhl and Filloux, 2014). Por outro lado, sem o conhecimento do mecanismo da atividade de uma proteína, mutações aleatórias podem gerar grande quantidade de dados em curto espaço de tempo que permite elaboração de um planejamento mais racional (Chusacultanachai and Yuthavong, 2004). A

síntese química de genes codificadores de proteínas contendo as alterações ou montagens necessárias permite produção de proteínas com alta taxa de modificações em menor espaço de tempo (Hughes, Miklos and Ellington, 2011; Xiong et al., 2008). Combinações de diferentes métodos também permitem a elaboração de modificações mais complexas e efetivas (Zawaira et al., 2012).

3. OBJETIVOS

3.1. Objetivo Geral

Desenvolver uma plataforma de expressão da proteína terapêutica interferon- β utilizando linhagens de *L. tarentolae* selvagem e geneticamente otimizadas para síntese de glicoproteínas com incorporação de ácido siálico.

3.2. Objetivos Específicos

- Integrar dados de genômica e metabolômica de promastigotas de *L. tarentolae* para identificar alvos enzimáticos e metabólitos para otimização da produção de glicoproteínas com incorporação de ácido siálico nas terminações das cadeias glicídicas;
- Desenhar e sintetizar os genes interferon- β e sialiltransferase humano com códons otimizados para *L. tarentolae*;
- Expressar o gene sialiltransferase humano de linhagens de *L. tarentolae* para permitir produção de proteínas glicosiladas com incorporação de ácido siálico;
- Expressar interferon- β em *L. tarentolae* selvagem e geneticamente modificada para incorporação de ácido siálico.

4. MATERIAL E MÉTODOS

4.1. Técnicas Básicas de Bioquímica e Biologia Molecular

As técnicas básicas de bioquímica e biologia molecular não descritas detalhadamente nesta seção, tais como preparo de géis de agarose e de poliacrilamida (SDS-PAGE), eletroforese de DNA e de proteínas, extração de DNA plasmidial (Miniprep), extração de RNA total e preparo de tampões, foram realizadas de acordo com os protocolos propostos por Sambrook e colaboradores (Sambrook et al., 1989)

Todas as dosagens e quantificações dos materiais genéticos (DNA genômico, DNA plasmidial e RNA) e proteicos utilizados neste projeto foram realizadas utilizando o equipamento Qubit 3.0 Fluorometer (Thermo Fisher Scientific).

4.2. Organismos

Durante este projeto foi cultivada a estirpe de *E. coli* da linhagem DH5 α para as técnicas de multiplicação de plasmídeos (antes ou após as clonagens) e a linhagem de *L. tarentolae*: cepa P10 (Código: LT-101, Jena Bioscience) utilizada para expressão constitutiva integrativa.

4.3. Genômica

As sequências proteicas envolvidas nas reações metabólicas de biossíntese de *N*-glicanos foram recuperadas do KEGG (Kanehisa, 2002). Foram selecionadas as sequências correspondentes aos genes de *Homo sapiens* (código “HSA”) e então contrastadas com as proteínas codificadas no genoma de *L. tarentolae* recuperadas do banco de dados TritypDB versão 35 (Aslett et al., 2010). Para isso utilizou-se duas abordagens: a primeira procurou pela anotação das proteínas envolvidas no metabolismo de glicosilação no genoma do tripanossomatídeo e a segunda foi realizada com a ferramenta Blastp comparando as proteínas de *H. sapiens* contra o proteoma predito de *L. tarentolae*. Ambas as abordagens confirmaram a função da glicosilação pela validação da presença de domínios funcionais específicos pelo banco de dados de famílias de proteínas Pfam (<http://pfam.xfam.org/>).

4.4. Manutenção e cultura de *L. tarentolae*

Formas promastigotas de *L. tarentolae* da cepa P10 foram descongeladas e mantidas em meio BHI (Brain Heart Infusion), preparado de acordo com as

recomendações do manual do kit LEXSY (37g/L). O meio de cultivo foi inicialmente esterilizado por 15 minutos a 121°C e 1 atm de pressão e no momento da inoculação foi suplementado com hemina (10 mg/mL), antibióticos Penicilina (10000u/mL) e Streptomina (10000µg/mL) para evitar contaminações bacterianas. As células foram mantidas a 26°C ao abrigo de luz, em volumes de 5-10 mL com as garrafas posicionadas horizontalmente sem agitação (cultura estática) sendo checadas diariamente no microscópio para verificar sua viabilidade e presença de contaminantes no meio de cultivo. Para mensuração da quantidade de células, a contagem foi realizada em câmara de Neubauer diluindo alíquotas das células em cultivo em 1:40 ou mais em solução de formalina 4%.

4.5. Extração de metabólitos

Extração de metabólitos foi realizada pelo método de metanol/clorofórmio/água conforme TENG et al., 2009. Amostras contendo 5×10^8 parasitos foram centrifugadas a 2000 x g por 10 minutos a 4° C, o meio de cultura foi descartado e os parasitos foram lavados duas vezes com PBS. Ao sedimento contendo as células foi adicionado 750 µL de solução 2:1 de metanol e clorofórmio e levado ao vortex (VELP Scientifica, Usmate Velate, Itália) por 5 minutos a 3000 rpm. Seguiu-se com adição de 100 µL de água ultrapura gelada, e levado ao vortex por 5 minutos a 3000 rpm. As amostras foram colocadas em banho de gelo por 15 minutos seguida por adição de 750 µL de água ultrapura gelada. Após serem novamente levadas ao vortex por 5 minutos a 3000 rpm, as amostras foram centrifugadas a 2000 x g por 10 minutos a 4°C. A fase superior de metanol/água foi coletada e a fase inferior de clorofórmio foi reextraída por adição de 750 µL de solução 2:1 metanol e água. As amostras foram levadas ao vortex por 5 minutos a 3000 rpm seguida por incubação em banho de gelo por 5 minutos e centrifugação a 2000 x g por 10 minutos a 4 °C. A fase superior foi recolhida e combinada com a fase superior previamente recuperada. As amostras foram secas em centrífuga evaporadora e estocadas a -80 °C até utilização.

4.6. Identificação e quantificação de metabólitos por 1H-RMN

Inicialmente, amostras com metabólitos extraídos pelo método de metanol/clorofórmio/água foram ressuspensas em 500 µL de óxido de deutério (D₂O) (Sigma-Aldrich, Saint Louis, EUA) contendo uma concentração final de 0,58 mmol.L⁻¹ de

trimetilsilil-2,2,3,3-tetradeuteropropionato de sódio (TSP) como referência interna de deslocamento químico e padrão interno para quantificação de metabólitos (Teng et al., 2009). Os espectros de RMN foram obtidos em equipamento Bruker DRX400 (Bruker, Billerica, EUA) de 9,4 T (400,13 MHz para frequência do hidrogênio), temperatura de 303 K e equipado com uma sonda de detecção inversa de 5 mm. Para o processamento foi utilizado o alargamento de linha de 0,3 Hz, anterior à transformada de Fourier. O processamento de imagens e adaptação dos espectros foi realizado pelo software ACD/Labs Spectrus Processor 2017.2 (Advanced Chemistry Development, Inc., Toronto, On, Canada, www.acdlabs.com). As fases e linhas de bases foram corrigidas e o espectro foi calibrado pelo sinal do TSP utilizando o programa NMR suite 8.31 (Chenomx).

Para identificação e quantificação dos metabólitos presente em cada amostra, inicialmente os espectros de RMN obtidos foram carregados no programa rNMR versão 1.1.9 (Lewis; Schommer; Markley, 2009) implementada em plataforma R versão 3.4.1 (Team, 2013). O ruído das leituras foi eliminado utilizando um valor de corte como 2 vezes o desvio padrão da linha de base, considerando como sinal apenas valores acima desta linha de corte. A lista de deslocamento químico e intensidade de sinal foi gerada e utilizada como arquivo de entrada para a ferramenta de busca NMR_based Search implementada no banco de dados de metabólitos Madison Metabolomics Consortium Database (CUI et al., 2008). Este programa permite identificação dos metabólitos presentes nas amostras através da comparação de deslocamentos químicos presentes na leitura da amostra com dados de deslocamento químico de compostos de referência depositados no banco de dados. Os parâmetros para análise de misturas foram ajustados para um mínimo de 80% do número de picos esperados para um determinado composto para assegurar presença do composto na mistura, desvio do deslocamento esperado para sinais provenientes do hidrogênio de 0,05 ppm, tipo de espectro 1D-1H e força do campo do RMN de 400,13 MHz. A concentração relativa dos metabólitos em relação ao TSP foi multiplicada por 0,58 para obtenção da concentração absoluta em mM dos compostos.

4.7. Síntese dos genes otimizados e desenho de *primers*

A sequências dos genes de interesse foram projetados com os códons otimizados para serem expressos em *L. tarentolae* utilizando a ferramenta Codon Optimization Tool da Integrated DNA Technologies (IDT)(<https://www.idtdna.com/CodonOpt>) utilizando

como espécie de referência *Leishmania donovani*. Os *primers* específicos para cada gene foram desenhados adicionando os sítios de restrição estratégicos para a ligação do gene de interesse nos plasmídeos de expressão utilizando o software OligoExplorer 1.2 (GeneLink). As sequências resultantes dos genes e dos primers foram encaminhadas à empresa Integrated DNA Technologies (IDT) para síntese química.

4.8. Amplificação dos genes sintéticos do Interferon- β e Sialil-transferase

Após ensaios para verificação das melhores condições experimentais, as amplificações das sequências gênicas de interesse foram realizadas utilizando os pares de iniciadores previamente adquiridos como demonstrado na tabela 1. A amplificação foi realizada em termociclador primeiramente a 95°C por 5 minutos para desnaturação inicial e alternando em um total de 35 ciclos a temperatura de 95°C por 1 minuto para desnaturação, 50° C (para primers IFN- β) e 55° C (para primers HST) por 40 segundos para anelamento e 72°C por 1 minutos para extensão, seguido por um passo de extensão final a 72°C por 7 min.

Tabela 1. Conjunto de iniciadores utilizados nos experimentos de PCR.

Nome do iniciador (enzima para clonagem)	Sequência (5'-3')	Orientação
IFN- β FW (<i>XbaI</i>)	TCTAGAAATCAATTACAAGCAGCTC	Direto
IFN- β RV (<i>KpnI</i>)	GGTACCGTTTTGGAAATTCCG	Reverso
HST FW (<i>BglII</i>)	AGATCTATGGGGCTGCTGG	Direto
HST RV (<i>NheI</i>)	GCTAGCTCAATGGTGATGGTG	Reverso

As amplificações tinham como objetivo obter o produto de PCR para posterior clonagem em sistema pGEM-T para isso foi utilizado a enzima DNA polimerase GoTaq® (Promega) que adiciona por meio de uma reação denominada *A-tailing procedure* uma base nitrogenada adenina (A) na extremidade 3' do gene amplificado. Os amplicons foram purificados com o kit Wizard® SV Gel and PCR Clean-Up System (Promega).

4.9. Clonagem nos plasmídeos pGEM T e pGEM T-Easy

O Sistema pGEM-T e pGEM-Teasy (Figura 6) de clonagem (Promega) é constituído de plasmídeos adquiridos na forma linear e que possuem em suas extremidades

3' uma deoxitimidina (T) que permite a inserção direta dos amplicons adenilados obtidos pela reação de PCR. A inserção da sequência de interesse ocorre mediante a incubação do plasmídeo e do produto adenilado na presença da enzima T4 DNA ligase. Após a reação de circularização com a inserção do gene o plasmídeo recombinante foi introduzido em bactérias competentes *E. coli* previamente preparadas para a propagação plasmidial.

A escolha de diferentes versões do vetor “pGEM-T” foi decorrente da necessidade de obter o sítio de restrição *NotI* para o gene da HST que posteriormente seria inserido no vetor de expressão PLEXY-egfp-SAT2. O plasmídeo pGEM-Teasy possui dois sítios para *NotI* o que impossibilitaria a retirada do mesmo com apenas um sítio de restrição para essa enzima, por isso a escolha pelo pGEM-T para a clonagem desse gene. Para a confirmação do posicionamento de inserção foi utilizado enzimas de restrição diferentes daquelas utilizadas para clonagem, como a *BsrGI* que possui sítio de clivagem no gene da HST.

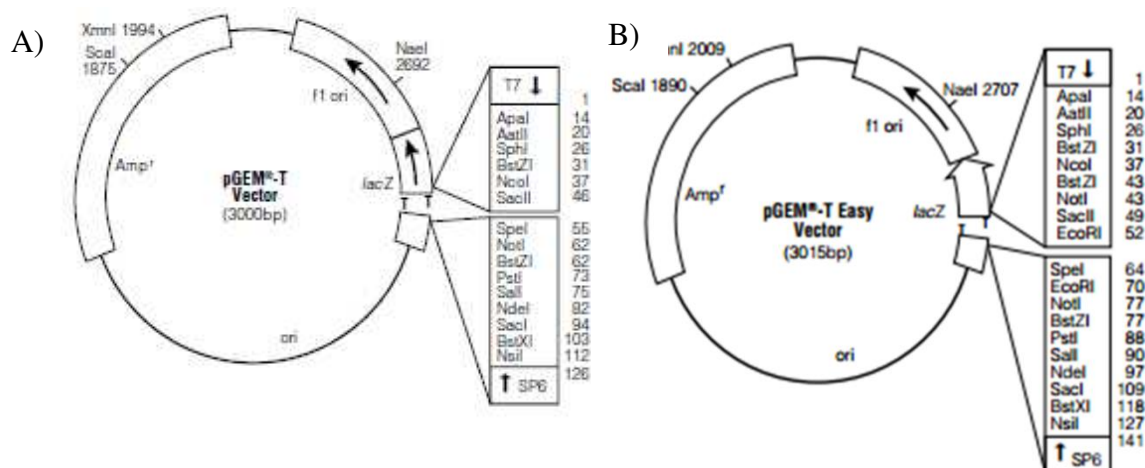


Figura 5. Mapa dos vetores de clonagem pGEM-T (A) e pGEM-Teasy (B) com os sítios de restrição disponíveis e a localização onde existe a deoxitimidina (T) livre para inserção da sequência de interesse.

4.10. Cultivo de *E. coli*

A linhagem de *E. coli* DH5α foram cultivadas em meio LuriaBertani (LB) líquido estéril, constituído de 1% de triptona (p/v); 0,5% de extrato de levedura (p/v); 1% de NaCl (p/v). Foi adicionado como método de seleção das células transformadas tanto com os vetores de clonagem (pGEM-T e pGEM-Teasy), quanto com os vetores de expressão (PLEXY-NEO2 e PLEXY-EFGP-SAT2), o antibiótico ampicilina (concentração final de 100 µg/mL) após a esterilização do meio e antes da realização dos inóculos. Para a

seleção dos clones em placas de petri após as transformações foi utilizado o meio LB sólido que contém todos os reagentes do meio LB líquido com a adição de 1,5% (p/v) de ágar.

4.11. **Preparo de *E. coli* DH5 α competentes e transformação por choque térmico**

O preparo de células competentes, foi realizado com pré-inóculo de 5mL em meio LB líquido sem adição de antibióticos e com alíquotas de 50 μ L de *E. coli* DH5 α estocadas em - 80°C. Após incubação overnight sob agitação a 180 rpm a 37°C, o pré-inóculo foi expandido para 250mL em meio SOB (2% de triptona (p/v), 0,5% de extrato de levedura (p/v), 0,05% de cloreto de sódio (p/v), 2,5 mmol.L⁻¹ de cloreto de potássio e 20 mmol.L⁻¹ de cloreto de magnésio) e incubado a 18°C sob agitação de 150 rpm até atingir D.O 600 nm entre 0,5 de 0,6. Atingida a D.O, a cultura foi mantida no gelo por 10 minutos, centrifugada a 4000 rpm por 10 minutos e 4°C. O sobrenadante foi descartado, e as células sedimentadas foram gentilmente ressuspensas em 30 mL de solução TB (15 mmol.L⁻¹ de CaCl₂, 250 mmol.L⁻¹ de KCl, 10 mmol.L⁻¹ de HEPES, 55 mmol.L⁻¹ de MnCl₂ em pH 6,7) gelada e filtrado, incubadas em banho de gelo por 10 minutos. Decorrido este tempo repetiu-se o passo anterior e ressuspendeu as células em 8mL de TB gelado e 7% de DMSO e incubadas em banho de gelo por 10 minutos. Posteriormente o volume total foi distribuído em alíquotas de 100 μ L, congelados primeiramente em nitrogênio líquido e estocados em ultrafreezer a - 80°C.

As transformações bacterianas foram realizadas com as alíquotas de 100 μ L de células ultracompetentes previamente preparadas. Após a retirada do ultrafreezer foram adicionados os DNAs plasmidiais transformantes (mínimo de 50ng), as células mantidas em banho de gelo por 30 minutos e depois submetidas a choque térmico em banho-maria a 42 °C por 45 segundos. Após esta etapa, retornamos as células para o banho de gelo por mais 2 minutos, depois suspensas em 1 mL de meio LB ou SOB (preferencialmente) sem antibiótico e incubadas de 1 a 2 horas sob agitação a 180 rpm a 37°C (no caso dos vetores exclusivamente bacterianos) ou 30°C (temperatura recomendada para manutenção dos plasmídeos LEXSY em *E. coli*). Após a recuperação das culturas, as mesmas foram centrifugadas a 5000 rpm por 5 minutos, e o sobrenadante descartado até o restante de 100 μ L que foi usado para ressuspender o pellet. Esse volume então foi utilizado para plaqueamento com auxílio de uma alça de Drigalski em LB-ágar, com o antibiótico de

seleção ampicilina. As placas foram mantidas a 37° por 16 horas (para vetores bacterianos) ou 30°C por 40 horas (para vetores LEXSY).

4.12. Digestão

As reações e condições de digestão, como tampão, unidades de enzima, e tempo de digestão foram adaptadas para cada enzima utilizada. As enzimas *XbaI*, *KpnI*, *BglII* e *NheI* (Códigos: ER0681, ER0521, ER0081, ER0971) foram adquiridas pela empresa Thermofisher/Fermentas e a enzima *NotI* (R6431) adquirida pela empresa Promega. Os protocolos de digestão foram desenvolvidos de acordo com as recomendações deste fabricante. Após as reações de digestão os DNAs foram aplicados em eletroforese em gel de agarose 0,8% para verificação da eficiência do processo e posterior purificação das bandas de interesse com o kit Wizard® SV Gel and PCR Clean-Up System (Promega).

4.13. Clonagem e características dos plasmídeos pLEXSY

Os vetores de expressão integrativos constitutíveis do Sistema LEXSY permitem que por meio dos sítios de restrição disponíveis e como o gene é inserido, as proteínas recombinantes de interesse sejam expressas e mantidas no citoplasma da célula hospedeira ou endereçadas para secreção extracelular. O que define a secreção é a presença da sequência da fosfatase ácida secretada de *L. Mexicana* (LMSAP), utilizada como peptídeo sinal, disponível nos plasmídeos.

Dos plasmídeos utilizados neste trabalho, apenas pLEXSY-neo2 (figura 6.A)) apresenta essa configuração. O plasmídeo pLEXSY-egp-SAT2 (figura 6.B) foi utilizado apenas retirando a sequência codificante “enhanced green fluorescent protein” (egfp) para a inserção do gene e expressão citoplasmática. Cada um foi digerido por enzimas de restrição para linearização e liberação dos sítios de clonagem de interesse. Cada construção utilizou os sítios de restrição específicos para cada gene (Tabela 1).

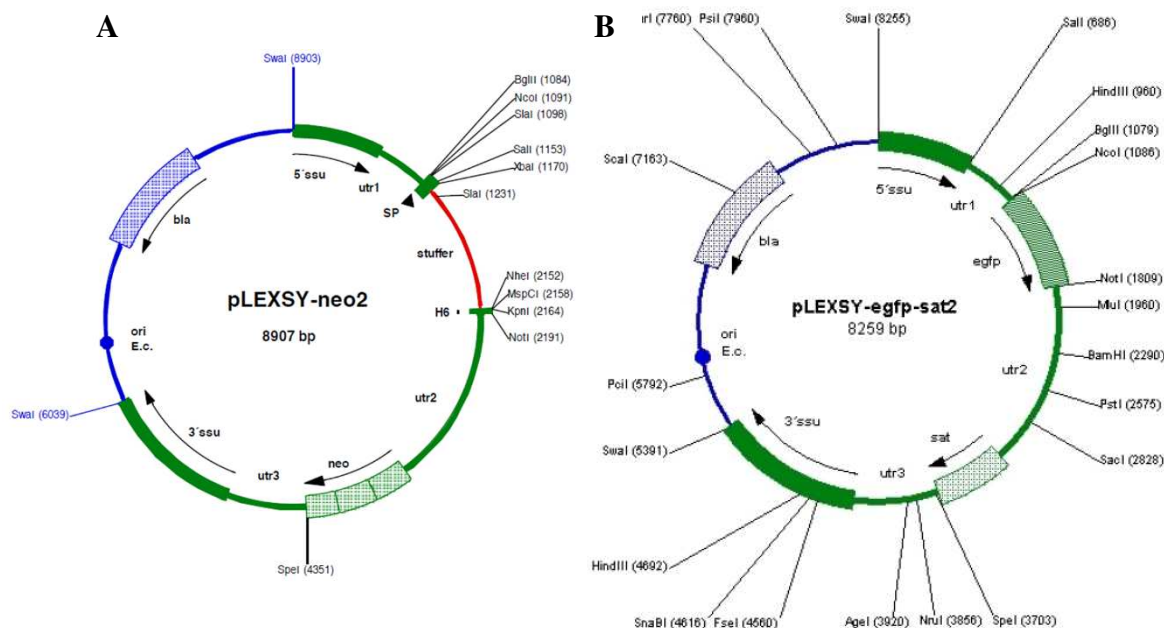


Figura 6. Mapa do vetor de expressão constitutiva integrativa pLEXY-neo2 (A) e pLEXY-egfp-SAT2 (B). O fragmento destacado em vermelho em (A) e o *egfp* em (B) correspondem a sequência do *stuffer* liberado durante as digestões para inserção do gene de interesse. As regiões em azul destaca a sequência que possui a origem de replicação bacteriana e marca de seleção para ampicilina que deve antes da transfecção deve ser retirada pela enzima de restrição *SwaI*. Após a linearização do vetor realizada pela enzima *SwaI*, as regiões de recombinação homóloga 5' *ssu* e 3' *ssu*, são utilizadas para integração no genoma do hospedeiro (cepa P10). Os sítios *utr* são regiões transcritas e não traduzidas de *L. tarentolae*, adicionadas com a finalidade de fornecer os sinais de *trans-splicing* apropriados para as modificações pós-transcricionais dos mRNAs da proteína alvo e sinais de início de tradução típicos de tripanossomatídeos.

4.14. Estratégias de clonagem

Primeiramente elaboramos as estratégias de clonagem baseadas no sistema de clonagem pGEM-Teasy e de expressão integrativa constitutiva LEXSY utilizando os vetores pLEXY-NEO2 e pLEXY-SAT2. Foi determinado um vetor de expressão diferente para cada gene, cada um com uma marca de seleção por antibiótico, para que posteriormente os dois genes clonados separadamente pudessem ser expressos e selecionados simultaneamente no mesmo meio de cultivo. Para isso os conjuntos de iniciadores (primers) foram planejados para que o gene otimizado do IFN- β possuísse os sítios de restrição para *XbaI* e *KpnI*, mantivesse o peptídeo sinal e adquirisse a sequência His-tag do vetor de expressão. E para que gene da sialtransferase humana (HST) possuísse os sítios de restrição para *BglIII* e *NheI*, não mantivesse o peptídeo sinal para

secreção e não adquirisse a sequência His-tag (já que a gene otimizado ja possuía essa sequência).

Após a síntese dos genes verificamos a disponibilidade de uso de apenas um dos vetores de expressão planejados, o pLEXSY-NEO2 planejado para clonagem do gene IFN- β , necessitando usar um outro vetor, o pLEXSY-egfp-SAT2 para clonagem do gene HST. Esse vetor não possui o sítio de clonagem *NheI* previamente inserido no primer do gene HST, o que nos fez também alterar a abordagem do vetor de clonagem pGEM-Teasy para o vetor pGEM-T. Os detalhes das abordagens de clonagem estão especificados no quadro 1.

Tabela 2. Diferentes estratégias abordadas para a construção dos clones em plasmídeos pGEM-T e pLEXSY, especificadas de acordo com o nome, plasmídeo utilizado, sítios de restrição utilizados e objetivo.

Nome da Construção	Plasmídeo utilizado	Sítios de restrição	Objetivo
IFNβ + pGEM-Teasy	pGEM-Teasy (Código: A1360)	<i>XbaI</i> e <i>KpnI</i>	Multiplicação de plasmídeo
HST + pGEM-T	pGEM-T (Código: A3600)	<i>BglIII</i> e <i>NheI</i>	Multiplicação de plasmídeo e obtenção do sítio de restrição NotI
IFNβ + PLEXSY NEO	pLEXSY-NEO2 (Código: EGE-233)	<i>XbaI</i> e <i>KpnI</i>	Expressão constitutiva secretória da proteína madura (Mantendo o peptídeo sinal e His-tag*)
HST + PLEXSY efpSAT	pLEXSY-EGFP-SAT2 (Código: EGE-235)	<i>BglIII</i> e <i>NotI</i>	Expressão constitutiva com deleção do peptídeo sinal

*HisTag: cauda de hexa-histidina na região carboxi-terminal da proteína alvo.

Para confirmar a inserção dos genes de interesse nos vetores de expressão foi realizado um PCR utilizando iniciadores específicos dos plasmídeos pLEXSY, o *primer forward* P1442 (5'- CCGACTGCAACAAGGTGTAG-3' – Código PM110) e *reverse* A264 (5'-CATCTATAGAGAAGTACACGTAAAAG-3'- Código PM101) incluídos no kit LEXSY. O primer direto P1442 anela na região 5' de inserção (cerca de 80 pb a montante do ATG) e o iniciador inverso A264 anela na região 3'do inserto (cerca de 80 pb a jusante do códon de parada).

4.15. **Transfecção e seleção clonal de *L. tarentolae***

Para a Transfecção as células da linhagem P10 com densidade de 6×10^7 células/mL foram sedimentadas por centrifugação a $3000 \times g$ por 5 minutos e suspensas no meio remanescente na concentração de 10^8 células/mL. As células foram transferidas para cubetas de eletroporação de 2 mm (Bio-Rad) e mantidas por 10 minutos em gelo juntamente com os DNAs plasmidiais transformantes (na concentração de 5 a 10 μg) em um volume total de 400 μL . A eletroporação foi efetuada no aparelho Gene Pulser (Bio-Rad) seguindo seguinte programa: voltagem de 450V, capacitância de 450 μF e pulsos de 5 a 6 milissegundos. Imediatamente após o pulso elétrico as cubetas voltaram para o gelo por mais 10 minutos e decorrido este tempo as células foram transferidas para 5 mL de BHI suplementado com hemina, penicilina-estreptomicina e incubadas nas condições de cultivo por 48 horas para recuperação das células. Após esse tempo foi adicionado o antibiótico de seleção e analisado o seu tempo de ação com o auxílio de uma cultura de células como controle negativo transformadas sem material plasmidial, apenas com água. A seleção clonal foi realizada em placas de cultivo BHI-ágar preparadas com o meio BHI na concentração 2X, além da adição de soro fetal bovino inativado, tampão HEPES 1M pH 7.4, hemina e antibióticos (incluindo o de seleção). As placas foram incubadas em condições de cultivo por no mínimo 48 horas até ser possível visualizar colônias. Essas colônias foram expandidas em meio BHI líquido (com todas as suplementações já citadas) gradativamente em média de 1 a 2 dias: primeiramente em 200 μL , depois em 500 μL , 1 mL, 5 mL e finalmente 10 mL. O preparo de novos estoques para congelamento, tanto das cepas transformadas como das não transformadas, e o plaqueamento em meio sólido era realizado sempre que as células já estavam recuperadas, e com uma densidade de 8×10^7 células/mL. Para o congelamento, em 9 mL das culturas adicionou-se 3 mL de glicerol 80% estéril, misturado gentilmente por inversão e as células foram distribuídas em alíquotas de 2 mL por criotubo, que foram mantidos 10 minutos a temperatura ambiente, 1 hora incubados em gelo, por 16 horas a -20°C e depois transferidos para ultrafreezer a -80°C . Para reativação das células de *L. tarentolae* estocadas a -80°C , o conteúdo total de um criotubo era inoculado em 10 mL de BHI (suplementado com hemina, contendo antibióticos apropriados, e 10% de soro fetal bovino) e as células recuperadas por sucessivas passagens a cada 2 ou 3 dias, em diluições de 1:10 ou 1:20, dependendo da densidade celular.

4.16. PCR diagnóstico de recombinação dos clones de *L. tarentolae*

Com o objetivo de se avaliar se houve a inserção do cassete de expressão no locus do cromossomo da célula hospedeira, e também verificar a presença dos genes do IFN- β e HST foi realizado um ensaio de PCR. Para a reação utilizou-se o material genômico das culturas de *L.tarentolae* após a transfecção, extraído com o kit Wizard® Genomic DNA Purification-Promega. A partir da fita molde de DNA genômico extraído das células selecionadas, foi realizado um PCR com um iniciador direto específico fornecido pela empresa Jena Bioscience, o forward F3001 (5'-GATCTGGTTGATTCTGCCAGTAG-3'-Código: PM-105). O oligonucleotídeo forward F3001 hibridiza com a região 5'ssu presente somente no cromossomo da célula hospedeira, enquanto que para oligonucleotídeos reversos foram utilizados os iniciadores específicos de cada gene clonado. Como controle negativo da reação foi utilizado o material genômico extraído de células selvagens.

4.17. Análise de transcritos por RT-PCR

Para analisar a presença de transcritos de RNAm correspondentes aos genes clonados foi realizado um ensaio de RT-PCR. Nesse experimento utilizaram-se os clones positivos confirmados pelo PCR diagnóstico de recombinação que foram submetidos à extração de RNA celular total. A extração de RNA foi realizada após uma lavagem das células em suspensão com PBS 1X e posterior adição de 1mL de Trizol. Após o tempo de 5 minutos de incubação no gelo foi adicionado 200 μ L de clorofórmio para a separação das fases, e essa solução deixada no gelo por 15 minutos. Centrifugou-se por 12000g a 4°C e 15 minutos. Transferiu-se a fase aquosa adquirida na preparação das amostras para um tubo limpo para posterior adição de 500 μ L de isopropanol gelado (misturado por inversão). Após 10 minutos de incubação no gelo, as amostras foram centrifugadas por 10 minutos. Descartou-se o sobrenadante e lavou-se o RNA sedimentado com 1mL de etanol 70%. Após novamente serem submetidas a centrifugação o sobrenadante foi descartado e utilizou-se 20 μ L de água livre de RNase para ressuspensão do pellet.

Após a extração do RNA celular total o material foi quantificado e em seguida o submetido a digestão com uma enzima DNase I (Invitrogen, Cat# 18068-015) seguindo as orientações do fabricante. O produto da digestão com a DNase foi utilizado para a síntese de uma molécula de cDNA numa reação de RT-PCR. Após a síntese do cDNA o produto

da reação foi utilizado em outra reação de PCR utilizando-se a enzima Taq DNA Polimerase (Cellco) e os primers e condições de reação de cada gene descrita no item 4.8.

4.18. Análise por SDS-PAGE 14% da expressão dos clones selecionados

Após a confirmação da presença dos transcritos, 10mL de cultura foram separados para análise das proteínas expressas. As células coletadas foram centrifugadas a 5000g por 5 minutos e as frações de sedimentos celulares e sobrenadante separadas em tubos distintos. O preparo das amostras foi realizado segundo manual LEXSY. Para precipitação das proteínas secretadas, ao sobrenadante foi adicionado uma solução de TCA 50% (concentração final de 10%). O volume final foi mantido em gelo por 30 minutos e posteriormente centrifugado a 15000 x g, por 15 minutos e 4°C. Após a centrifugação, descartou-se o sobrenadante e ressuspende-se o pellet em acetona 80% gelada. Foi realizada outra centrifugação (com as mesmas configurações da anterior) e o pellet ressuspendido com PBS. Os sedimentos de células previamente separados do meio de cultivo foram suspensos diretamente em tampão de amostra de proteínas. Para análise dos extratos os materiais (tanto do sobrenadante quanto do pellet) foram fervidos após a adição de tampão de amostra em alíquotas de 20 a 30 µL e foram aplicadas em SDS-PAGE 14%.

4.19. Análise por Dot Blotting da expressão dos clones selecionados

Como um ensaio preparatório para técnica de *Western Blotting* foi utilizado a técnica de Dot Blotting. Nestes ensaios não há necessidade de aplicação das amostras em SDS-PAGE e posterior transferência das proteínas para membrana de nitrocelulose. As amostras foram aplicadas diretamente sobre a membrana por pipetagem. Após a secagem das amostras, foi realizado o bloqueio da membrana com uma solução de BSA (Albumina de soro bovino) 3% em TBS-T 1X incubada por 1 hora. Após o bloqueio as membranas foram lavadas com solução de TBS-T 1X por três vezes e incubadas por 1 hora com primário monoclonal de camundongo, contra cauda de histidina, diluído em 1:3000 (GE Healthcare; código: 27-4710-01) ou com o anticorpo primário policlonal de coelho contra IFN-β diluído também em 1:3000 (Sigma, código: PRS4243). As membranas foram novamente lavadas por três vezes com solução de TBS-T 1X e posteriormente incubadas por 1 hora com anticorpo secundário contra camundongo ou coelho conjugado à peroxidase (Sigma; código: A4416), em diluição de 1:10000. Após mais três lavagens a atividade peroxidásica foi detectada com o uso de solução de revelação contendo 10 mL de

Tris 50 mM pH 7,5, 10 mg/mL de DAB e 10 µL peróxido de hidrogênio 30%. Todos anticorpos foram diluídos em tampão TBS-T 1X.

5. RESULTADOS E DISCUSSÃO

5.1. Genômica

Como o processo de glicosilação em *L. tarentolae* ainda apresenta algumas dúvidas (Breitling et al., 2002), optamos por utilizar a metodologia de genômica comparativa como uma tentativa de elucidar pontos ainda não discutidos no processo de biossíntese de glicoproteínas nesses organismos.

Após as análises utilizando as abordagens de busca pela anotação das proteínas envolvidas no metabolismo de glicosilação e com a ferramenta Blastp, indentificou-se um grupo de proteínas presente no metabolismo de *N*-glicanos no proteoma predito de *L. tarentolae* correspondente às mesmas encontradas em humanos, com excessão da sialiltransferase (Tabela 3). Essa enzima tem a função de catalisar a adição de ácido siálico na extremidade da cadeia glicídica, sendo a última enzima do processo de formação de proteínas *N*-glicosiladas. A função de algumas dessas proteínas no processo de *N*-glicosilação foi também predita pelo reconhecimento de domínios proteicos específicos após a submissão das sequências no banco de dados de família de proteínas, Pfam. Na tabela 3 estão demonstradas as enzimas utilizadas para a análise, bem como os códigos das proteínas correspondentes ao KEGG e ao banco de dados TrytriDB, o valor de e-value e o nome da família validados pelo Pfam e na figura 7 o esquema estrutural de *N*-glicanos e as enzimas envolvidas no processo.

Nossas análises corroboram as observações e ensaios realizados por Breitling em 2002: por técnicas de sequenciamento por eletroforese em gel de poliacrilamida, foi demonstrado que *L. tarentolae* de fato possui ativa a via de *N*-glicosilação capaz de produzir o complexo biantenário de eucariotos superiores presente em glicoproteínas, apenas faltando a sialização. Também foi demonstrada uma importante característica da expressão proteica nesses organismos, que é a homogeneidade da estrutura, muito importante para proteínas terapêuticas.

Tabela 3. Enzimas envolvidas no processo de biossíntese de *N*-glicanos, códigos relacionados ao bancos de dados e e-value da validação de domínio funcional específico.

Número da Enzima	Nome	Código Kegg	Código TrytripDB	Família	E-value Pfam	
1	Glicosiltransferase	K00717	LtaP02.0260	Glyco_transf_10	3,0E-09	
2, 3 4, 5, 6 e 7	Acetil-glicosamina-transferase	K00737				
		K00744				
		K13748	LtaP34.3920	PIG-Y	6,1E-14	
		K00736	LtaP36.1700	GPI2	7,9E-67	
		K00778	LtaP36.4890	PIG-P	3,60E-27	
8	Galactosil-transferase	K00738				
		K00726				
			LtaP02.0130	-	-	
			LtaP02.0140	-	-	
			LtaP02.0150	-	-	
			LtaP02.0160	-	-	
9	Sialiltransferase	K00778				
				LtaP07.1260	-	-
				LtaP14.1470	-	-
				LtaP29.1150	-	-

Legenda:

1. Glicoproteína 6-alfa-L-fucosiltransferase
2. Beta-1,4-manosil-glicoproteína-beta-1,4-N-acetilglicosaminiltransferase
3. Alfa-1,3(6)-manosilglicoproteína-beta-1,6-N-acetil-glicosaminiltransferase
4. Alfa-1,3-manosilglicoproteína-beta-1,4-N-acetilglicosaminiltransferase C
5. Alfa-1,6-manosil-glicoproteína beta-1,2-N-acetilglicosaminiltransferase
6. Alfa-1,3-manosil-glicoproteína beta-1,4-N-acetilglicosaminiltransferase A/B)
7. Alfa-1,3-manosil-glicoproteína beta-1,2-N-acetilglicosaminiltransferase
8. Beta-1,4-galactosiltransferase 1
9. Beta-galactose-alfa-2,6-sialiltransferase (sialiltransferase 1)

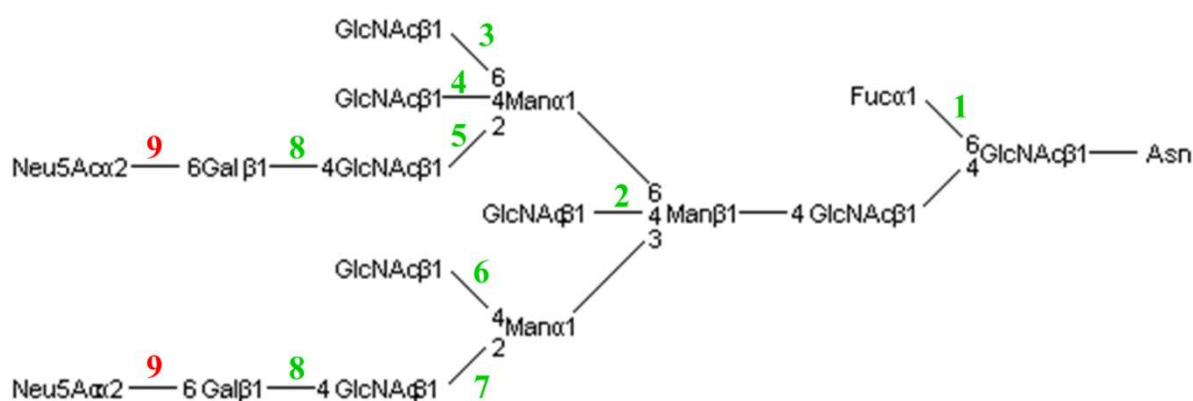


Figura 7. Esquema da estrutura *N*-glicanos. Os números em verde correspondem as enzimas presentes em *L. tarentolae*. Em vermelho está representado a enzima sialiltransferase, não identificada no genoma predito de *L. tarentolae*. 1: Fucosiltransferase, 2, 3, 4, 5, 6 e 7: acetilglicosaminatransferases, 8: galactosiltransferase, 9:sialiltransferase. Figura adaptada. (Fonte: Kegg- N-Glycan biosynthesis - Reference pathway).

5.2. Metabolômica

Existem diferentes metodologias utilizadas para análises de metabólitos, entre elas a espectrometria de massas, métodos cromatográficos e ressonância magnética nuclear (RMN). A técnica de RMN se destaca das demais, por diversas vantagens como a não necessidade de introduzir mudanças estruturais nas moléculas a serem analisadas e especificidade na separação de compostos com massas semelhantes (Pontes, João Guilherme et al. 2016).

A abordagem de metabolômica por RMN teve como o intuito a detecção de metabólitos presentes em *L. tarentolae* e relacioná-los com as potenciais rotas metabólicas associadas com *N*-glicosilação. Após a análise realizada na abordagem de genômica, onde avaliaram os genes que codificam as enzimas presentes na via de biossíntese de *N*-glicanos em humanos e se as mesmas estavam presentes em *L. tarentolae*, observamos a ausência da enzima sialitransferase que adiciona uma molécula de Neu5Ac (ácido siálico) à extremidade da cadeia glicídica das glicoproteínas. Nosso foco, então, foi identificar nos metabólitos detectados por RMN, aqueles que participam da biossíntese do ácido siálico (Figura 8 e Tabela 4). Após a análise, foi possível encontrar todos os substratos precursores da molécula de Neu5Ac, não sendo possível identificar somente a sua forma associada a uma molécula de CMP (citidina monofosfato). Uma molécula de CTP (citidina trifosfato) é utilizada por meio de uma hidrólise e perda de um grupo fosfato para ativação do ácido siálico no núcleo, sendo formado então o CMP-Neu5Ac que está apto para ser transportado para o complexo de Golgi. A ausência da molécula de CMP-Neu5Ac nos permite inferir que *L. tarentolae* também não possui a expressão da enzima que catalisa a ativação do ácido siálico, a *N*-acetilneuraminato citidiltransferase (CMAS) (Figura 8B). Outra proteína importante nesse metabolismo é o transportador específico de CMP-Neu5Ac, que se encontra acoplado na região medial-trans intermembrana do Golgi, e é responsável por transferir o ácido siálico ativo sintetizado no núcleo para a síntese das glicoproteínas sializadas no interior do Golgi (van Karnebeek, Bonafe et al. 2016). Seguindo a metodologia semelhante à abordada na análise genômica, na tabela 5 estão representadas as enzimas envolvidas no metabolismo do ácido siálico e aquelas que foram possíveis de se encontrar anotadas no proteoma predito de *L. tarentolae*, e também validadas pelo domínio funcional específico.

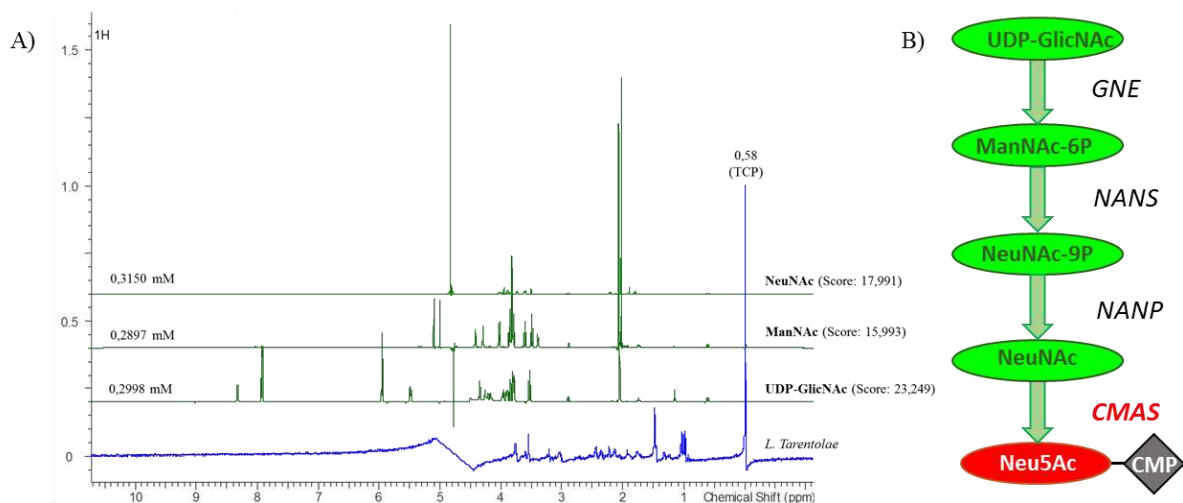


Figura 8. Espectro obtido por RMN em 1D-1H (A) e Esquema de reações envolvidas no metabolismo de biosíntese do ácido siálico (B). Em (A) estão alinhados todos os espectros padrões de RMN (H) dos metabólitos encontrados na via de biosíntese de ácido siálico (em verde) com o espectro obtido experimentalmente de células de *L.tarentolae* em cultivo (azul). Acima de cada espectro está a concentração encontrada para cada substrato (mensurada por meio da adição do calibrante TCP) e número de score calculado pelos matches encontrados por similaridade dos picos. Em (B) estão representados na cor verde todos os substratos da via de biosíntese de ácido siálico identificados pelo experimento de RMN e em vermelho o metabólito não encontrado, CMP-Neu5Ac. (GNE = UDP-N-acetilglicosamina 2-epimerase/N-acetilmanosamina quinase; NANS= N-acetilneuraminato sintase; NANP = N-acetilneuraminato-9-fosfatase; CMAS =N-acetilneuraminato citidiltransferase).

Tabela 4: Metabólitos detectados por RMN e que estão envolvidos no metabolismo de *N*-glicosilação. Na tabela abaixo se encontra os nomes, o código de identificação (EXPNMR_ID) do experimento padrão encontrado no Madison-Qingdao Metabolomics Consortium Database (MMCD), a concentração absoluta e o score para cada substrato identificado.

Nome do metabólito	EXPNMR_ID (MMCD)	Concentração absoluta (mM)	Score
Uridina 5'-difosfo-N-Acetilglicosamina	expnmr_00188	0,29981	23,249
N-acetil-D-manosamina	expnmr_00103	0,2897	15,993
N-acetilneuraminico-tipo ácido IV S	expnmr_00102	0,3150	17,991

Tabela 5: Enzimas envolvidas no processo de biossíntese do ácido siálico, códigos relacionados aos bancos de dados e e-value obtido pelo Blast realizado banco TrytripDB e da validação de domínio funcional específico no banco de dados Pfam.

Abreviação	Nome	Código Kegg	Código TrytripDB	E-value TrytripDB	Família	E-value Pfam
GNE	UDP-N-acetilglicosamina 2-epimerase/ N-acetilmanosamina quinase	K12409	LtaP33.2750	-	UDPGP	1E-23
NANS	N-acetilneuraminato sintase	K05304	-	-	-	-
NANP	N-acetilneuraminato-9-fosfatase	K01097	-	-	-	-
CMAS	N-acetilneuraminato citidiltransferase	K21749	-	-	-	-
CMP-SAT	CMP-ácido siálico transportador	K15272	LtaP24.0350	9E-24	Nuc_sug_transp	6,6E-51

Realizando o Blastp no banco de dados de tripanossmatídeos encontramos a proteína específica transportadora de CMP-ácido siálico. Essa informação é discutível pois não há evidências que *L. tarentolae* possua a enzima que catalisa a ativação do ácido siálico e portanto, não exista nesses organismos a molécula utilizada por esse transportador. Há então a possibilidade, de devido à anotação de genoma encontrada no banco de dados TrytriDB não ser validado, e estar disponível somente na forma automática, a proteína encontrada pela ferramenta Blastp possua domínios muitos semelhantes ao transportador de ácido siálico de outro organismo modelo, mas a sua função não é de transportar especificamente essa molécula. De fato, como demonstrado por Riemersma em 2015, o transportador em questão realiza outras funções independentes do ácido siálico, como a *O*-manosilação de α -dístroglicanos, que são receptores associados à matriz extracelular e ao citoesqueleto das células (Riemersma, Sandrock et al. 2015).

Já a presença do metabólito Neu5Ac pode ser explicada por essa molécula ser utilizada em outros processos metabólicos como a de biossíntese de glicolípideos. Lipídeos sializados são importantes para o reconhecimento de epitopos em células de mamíferos, por exemplo, muito utilizadas para o mecanismo de parasitismo (Barreto-Bergter, Eliana et al. 2010).

Como o substrato CMP-Neu5Ac atualmente é produzido industrialmente e comercializado, a sua inserção nas células de *L. tarentolae* por meio de adição ao meio de cultivo é uma alternativa à expressão heteróloga da enzima a *N*-acetilneuraminato citidiltransferase. Com essa abordagem associada à expressão heteróloga da enzima

sialiltransferase espera-se obter uma linhagem qualitativamente capaz de produzir glicoproteínas com o padrão de glicosilação semelhantes ao de humanos.

5.3. Amplificação por PCR dos genes sintéticos otimizados IFN- β e HST

As sequências proteicas da sialiltransferase e Interferon- β foram obtidas no Kegg (Kanehisa, 2002) e as sequências gênicas sintéticas foram geradas para estas proteínas com os códons otimizados para o gênero *Leishmania*. Por meio da característica básica do código genético ser degenerado, é possível alterar os códons ou nucleotídeos de uma sequência gênica e ainda manter a mesma sequência proteica traduzida. Nesse sentido, a técnica de otimização de códons tem sido muito utilizada para alterar a sequência gênica, retirando códons raros que podem tornar a tradução ineficiente e gerando códons mais abundantes no organismo e conseqüentemente mais RNAt (RNAs transportadores) aumentando a eficiência da produção proteica. Estudos já demonstraram que a otimização de códons pode dobrar a concentração proteica expressa no meio de cultivo, portanto barateando a produção, já que aumenta o rendimento mantendo-se as mesmas condições (Kianmehr, Mahrooz et al. 2016)

Para o desenho das sequências também foi planejada a adição de um peptídeo sinal (adquirida no vetor de expressão) que permite que a proteína sintetizada seja direcionada para o meio extracelular e adição de cauda de histidina (seja por síntese química ou adquirida no vetor expressão) que permite que a proteína seja identificada por anticorpos específicos. Ambas as estratégias são utilizadas para facilitar e otimizar a purificação das proteínas expressas (Wiese, Ilg et al. 1995, Klatt and Konthur 2012).

Os genes de interesse foram amplificados a partir dos genes sintéticos utilizando os pares de *primers* desenhados para cada um como demonstrado na tabela 1. Os amplicons obtidos através de PCR resultaram em bandas únicas e definidas, de 496 pb para o IFN- β e 1068 pb para HST observadas após a eletroforese (Figura 9). O material purificado dessas bandas foi utilizado nas subseqüentes reações de clonagem.

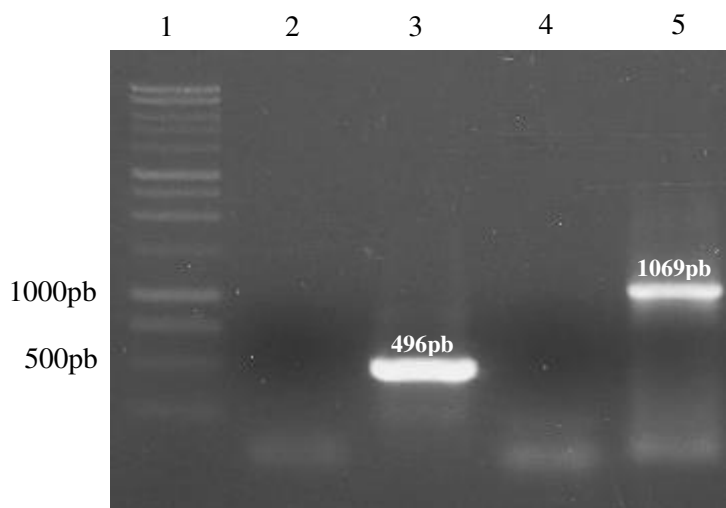


Figura 9. Eletroforese em gel de agarose 1% com os produtos da amplificação dos genes IFN- β e HST. Linha 1: padrão 1Kb DNA ladder (Promega). Linha 2: controle negativo da amplificação do gene IFN- β . Linha 3: Amplificação do gene IFN- β demonstrado na banda de 496pb. Linha 4: controle negativo da amplificação do gene HST. Linha5: Amplificação do gene HST demonstrado na banda de 1069pb.

Após a etapa de amplificação dos genes de interesse, o amplicon purificado do gel correspondente ao gene IFN- β foi submetido a reação de ligação com o plasmídeo linear pGEM-Teasy e o gene HST com pGEM-T. Os produtos de reação de clonagem foram utilizados para transformar células *E. coli* DH5 α . As colônias crescidas em meio sólido em placa de LB-ágar foram inoculadas em meio LB-líquido e posteriormente foi realizado a extração de plasmídeos para verificação das colônias positivas por PCR (Figura 10).

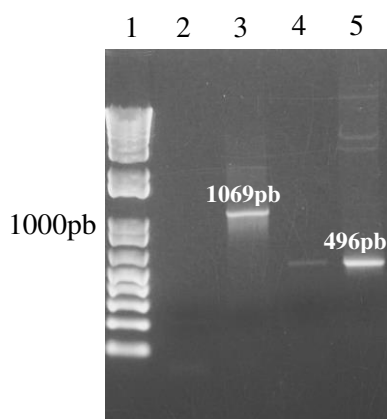


Figura 10. Gel de agarose 1% contendo a amplificação dos genes clonados em pGEM-T e pGEM-Teasy. Os plasmídeos recuperados após a clonagem foram utilizados para confirmação das colônias positivas. Linha 1: padrão 1Kb Plus DNA ladder (Invitrogen). Linha 2: controle negativo da reação de amplificação com os primers HST. Linha 3: produto da amplificação do gene HST com 1069pb. Linha 4: controle negativo da reação de amplificação com os primers IFN- β . Linha 5: produto da amplificação do gene IFN- β com 496pb.

5.4. Digestão dos plasmídeos pGEM-Teasy e pGEM-T clonados

A confirmação de amplificação dos genes de interesse com os plasmídeos clonados permitiu prosseguir com a digestão dos plasmídeos pGEM-Teasy e pGEM-T com as enzimas de restrição correspondentes a cada estratégia de clonagem desenhada (Figura 11), detalhadas na Tabela 2.

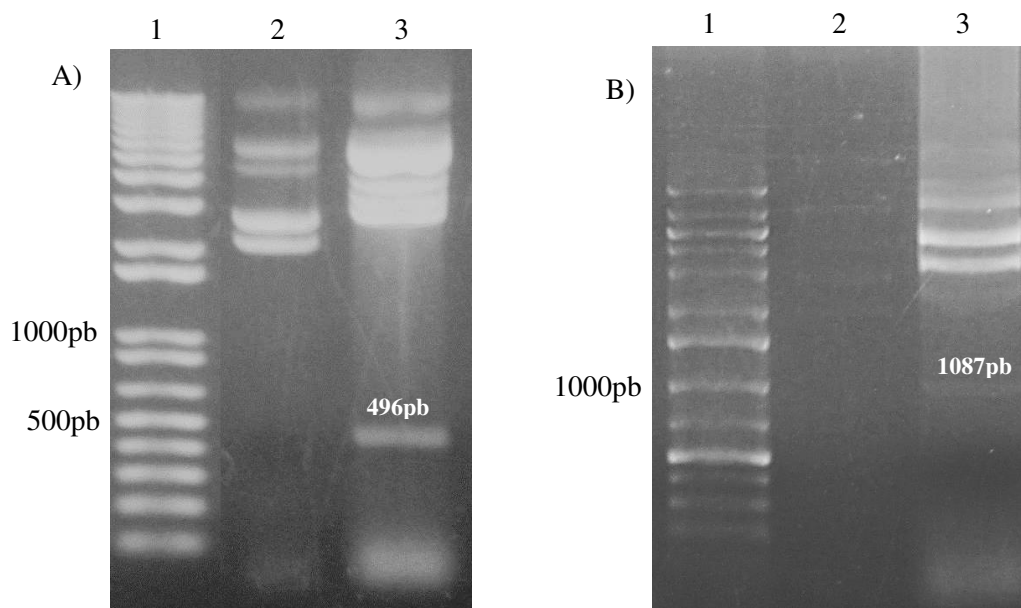


Figura 11. Gel de agarose 0,8% contendo os produtos de digestão do vetor pGEM-Teasy (A). Linha 1: padrão 1Kb Plus DNA ladder (Invitrogen). Linha 2: Controle negativo da reação de digestão. Linha 3: Digestão do DNA plasmidial por *XbaI* e *KpnI*, gerando o fragmento de 496pb. **Gel de agarose 0,8% contendo os produtos de digestão do vetor pGEM-T (B).** Linha 1: padrão 1Kb Plus DNA ladder (Cellco). Linha 2: Controle negativo da reação de digestão. Linha 3: Digestão do DNA plasmidial por *BglIII* e *NotI* gerando o fragmento de 1087pb.

5.5. Digestão dos plasmídeos pLEXY-NEO2 e pLEXY-egfp-SAT2 com enzimas de restrição

Com a obtenção dos fragmentos purificados decorrentes da digestão dos plasmídeos pGEM-T e pGEM-Teasy o próximo passo foi digerir os plasmídeos LEXSY com as enzimas de restrição correspondentes a cada estratégia de clonagem desenhada, detalhadas na Tabela 1. Como esses vetores possuem entre os sítios de clonagem uma sequência stuffer (1000 pb) ou egfp (731 pb), liberada quando a dupla digestão ocorre corretamente, permite confirmar a eficácia do processo, visto que se observar esses fragmentos por

eletroforese. As digestões duplas realizadas com os plasmídeos demonstraram em gel de agarose dois fragmentos, como esperado (Figura 12 e 13).

Os fragmentos maiores, de aproximadamente 7900 pb para o pLEXSY-NEO2 e de aproximadamente 7500 pb para o pLEXSY-egfp-SAT2 correspondem aos vetores digeridos e sem a sequência stuffer ou egfp (figuras 12 e 13). Os vetores digeridos foram extraídos dos géis e purificados para posterior uso nas reações de clonagem. Os fragmentos menores correspondentes às sequências stuffer e egfp foram descartados. Como controle negativo, utilizou-se uma amostra do vetor não digerido.

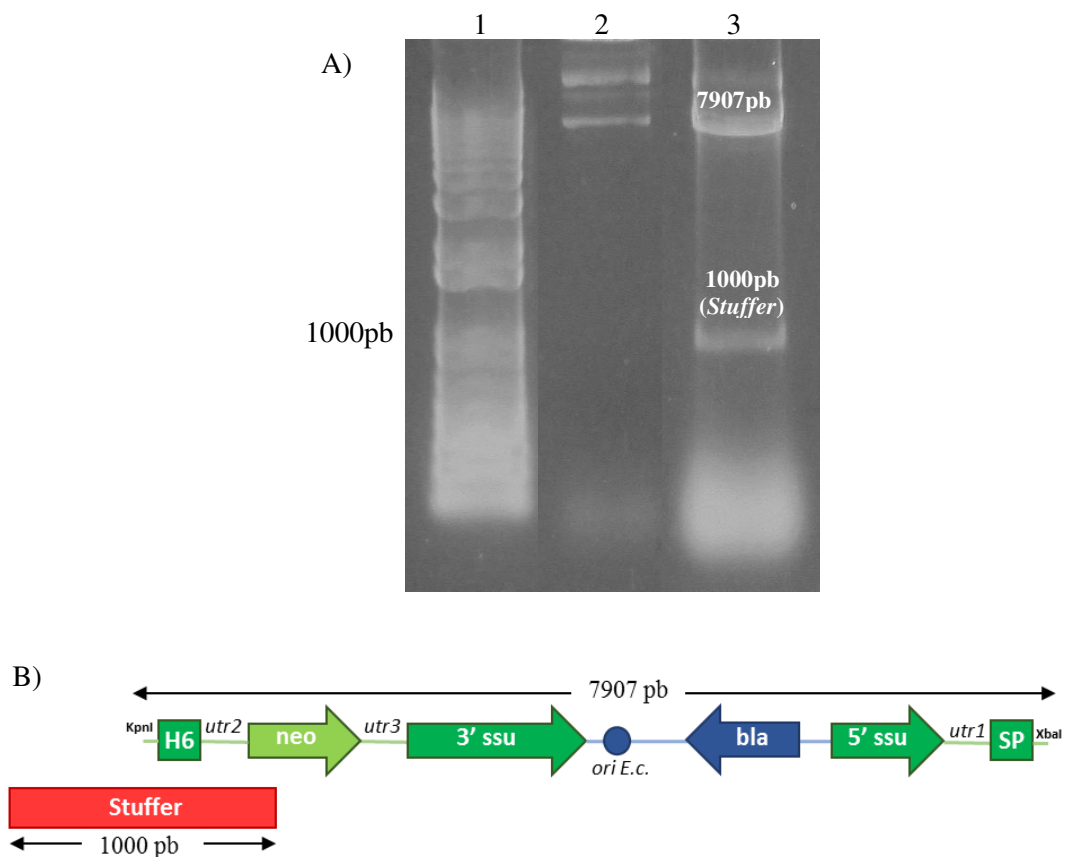


Figura 12. Gel de agarose 0,8% contendo os produtos de digestão do vetor PLEXSY-NEO 2 (A) e esquema do vetor digerido (B) Digestão do pLEXSY-NEO2 por XbaI e KpnI (para construção do Clone pLEXSY-NEO2 + IFN- β). Na linha 3, o fragmento inferior de aproximadamente 1000 pb corresponde à região stuffer do plasmídeo, liberada após a dupla digestão e o fragmento superior linearizado de aproximadamente 7907 pb foi purificado do gel para ser utilizado na reação de clonagem. Linha 1: padrão de pares de bases 1 Kb DNA ladder (Invitrogen).

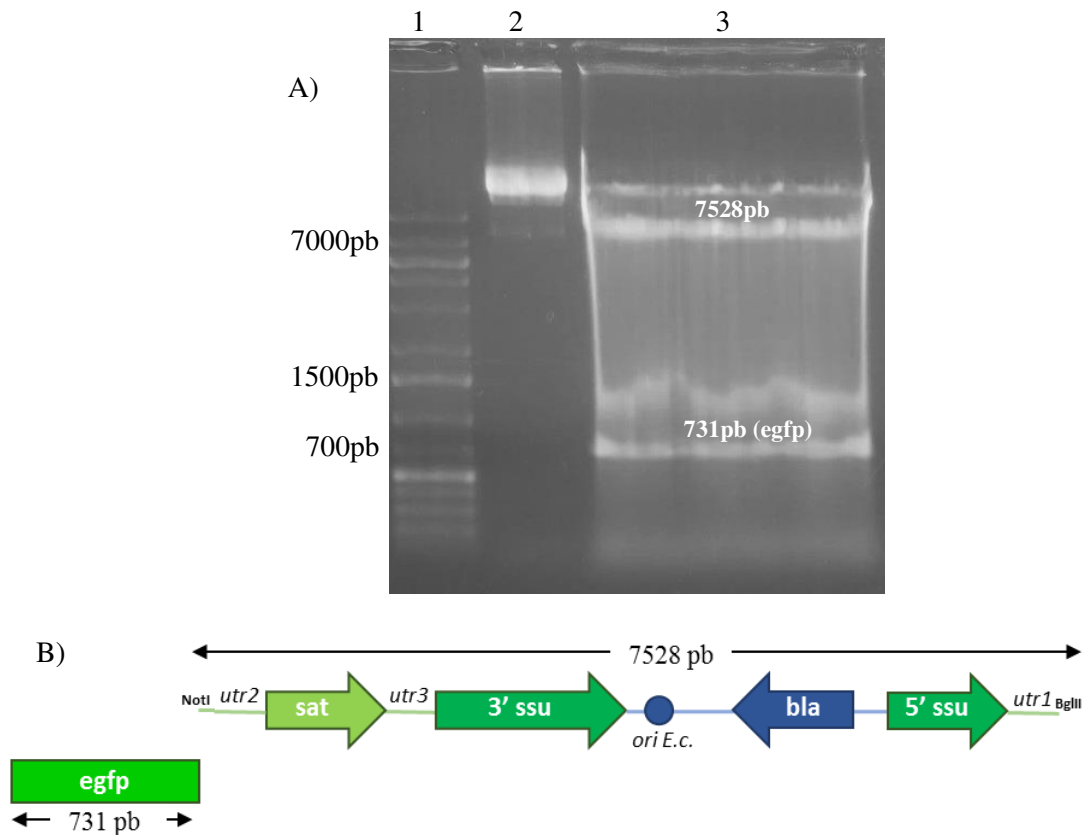


Figura 13. Gel de agarose 0,8% contendo os produtos de digestão do vetor PLEXSY-egfp-SAT2 (A) e esquema do vetor digerido (B). Digestão do pLEXSY-egfp-SAT2 por BglII e NotI (para construção do Clone pLEXSY-egfp-SAT2 + HST). Na linha 3, o fragmento inferior de aproximadamente 731 pb corresponde à região egfp do plasmídeo, liberada após a dupla digestão e o fragmento superior linearizado de aproximadamente 7528 pb também foi purificado do gel para ser utilizado na reação de clonagem. Linha 1: padrão de pares de bases 1 Kb DNA ladder (Cellco).

5.6. Clonagem dos genes nos plasmídeos digeridos pLEXSY

Os genes purificados decorrentes da digestão dos plasmídeos pGEM-T e os produtos lineares obtidos na digestão dos plasmídeos pLEXSY, foram submetidos reação de clonagem, utilizando a enzima T4 ligase (metodologia citada no item 4.9) para inserção dos genes nos vetores (Figuras 14).

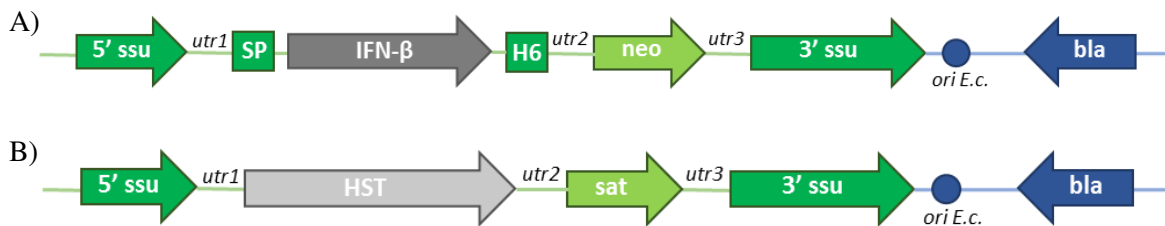


Figura 14. Desenho esquemático dos clones construídos após a reação de ligação. O gene IFN- β inserido no vetor pLEXSY-NEO2 (A) e o gene HST inserido no vetor pLEXSY-egfp-SAT2 (B).

Após a reação de ligação os vetores foram utilizados como amostras para um PCR de verificação de integração dos genes utilizando os *primers* direto P1442 e reverso A264 (Figura 15), como mencionado no item 4.14 da metodologia.

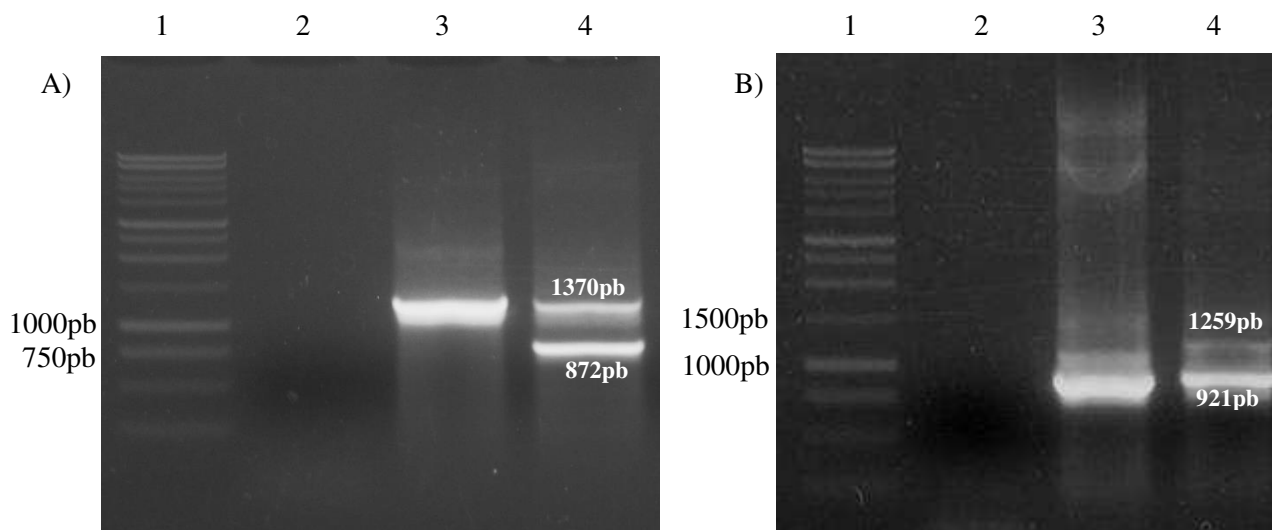


Figura 15. PCR de verificação de integração do gene IFN- β no plasmídeo pLEXSY-NEO2 (A). Linha 1: padrão 1Kb DNA ladder (Promega). Linha 2: controle negativo onde o DNA foi substituído por água. Linha 3: controle negativo tendo como amostra um plasmídeo vazio (com *stuffer*). Linha 4: amplificação do vetor clonado demonstrada pela banda de 872 pb. **PCR de verificação de integração do gene HST no plasmídeo pLEXSY-egp-SAT2 (B).** Linha 1: padrão 1Kb DNA ladder (Promega) .Linha 2: controle negativo onde o DNA foi substituído por água. Linha 3: controle negativo tendo como amostra um plasmídeo vazio (com *egfp*). Linha 4: amplificação do vetor clonado demonstrada pela banda de 1259 pb

Posterior a confirmação da clonagem os vetores foram inseridos por transformação em *E.coli* (DH5 α) para multiplicação plasmidial e posteriormente extraídos. Os testes para verificação da eficácia da clonagem foram realizados utilizando PCR com os *primers* específicos de cada gene e tendo com DNA molde os plasmídeos extraídos dos clones, digestão para observar a liberação do fragmento clonado e sequenciamento. Tanto o clone pLEXSY-*egfp*-SAT2+HST, quanto pLEXSY-NEO2+IFN- β foram positivos para o teste de PCR e digestão, porém apenas o clone contendo o gene do IFN- β pôde ser sequenciado. Após a extração plasmidial dos clones selecionados como positivos, os mesmos foram submetidos à reação de linearização pela digestão com a enzima de restrição *SwaI*. Após a eletroforese em gel de agarose 0,8% observou-se a formação de duas bandas no gel. A de aproximadamente 3000pb, correspondente a origem de replicação de *E. coli* e do gene de resistência a ampicilina e a outra banda de aproximadamente 6043 para o vetor clonado pLEXSY-NEO2+IFN β (Figura 16) ou 5395pb para o vetor clonado pLEXSY-*egfp*-

SAT2+HST (Figura 17), representando fragmento linearizado do cassete de expressão que será integrado por recombinação homóloga no hospedeiro.

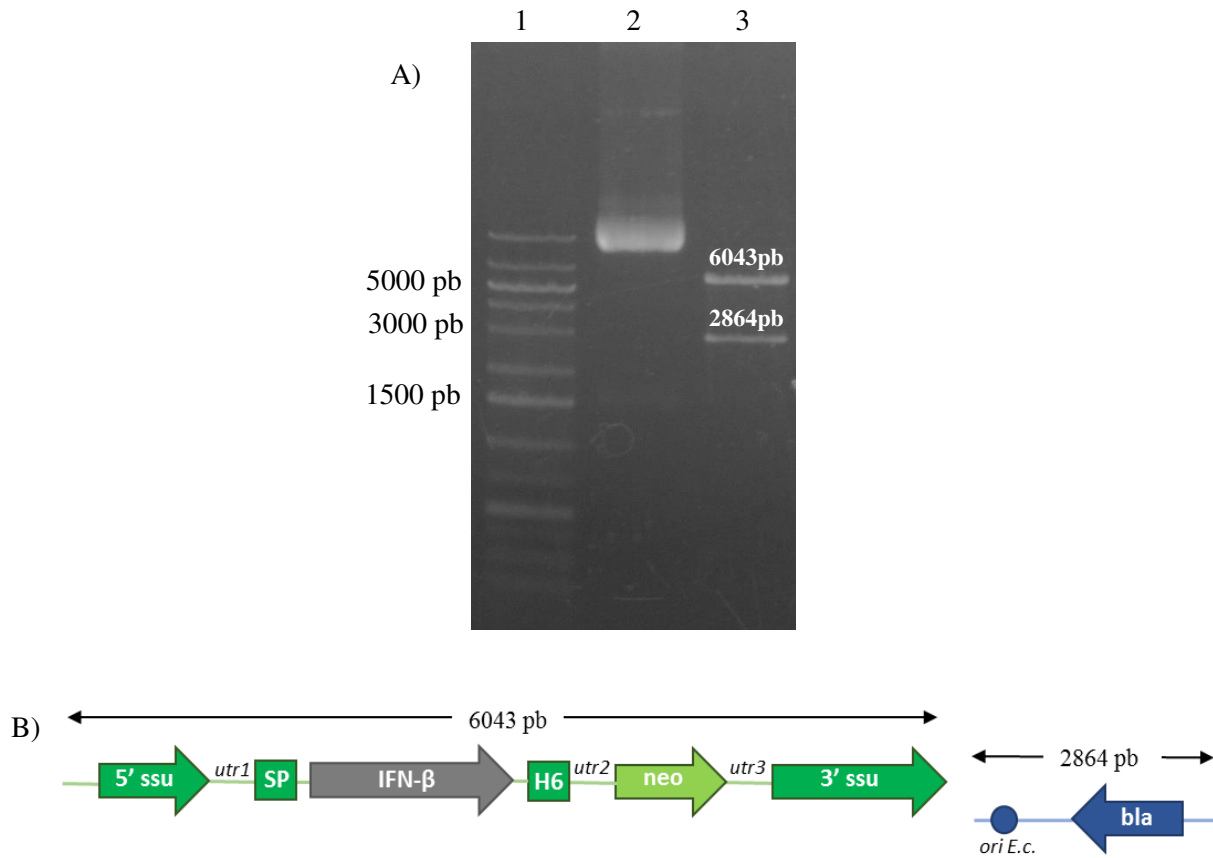


Figura 16. Eletroforese em gel de agarose 0,8% com o resultado da digestão da construção pLEXY-NEO2+IFN-β, extraída dos clones recombinantes, com a enzima de restrição SwaI. Linha 1: padrão 1Kb plus DNA ladder (Cellco). Linha 2: Controle negativo. Linha 3: fragmentos oriundos da digestão da construção extraída de um clone recombinante com a enzima SwaI.

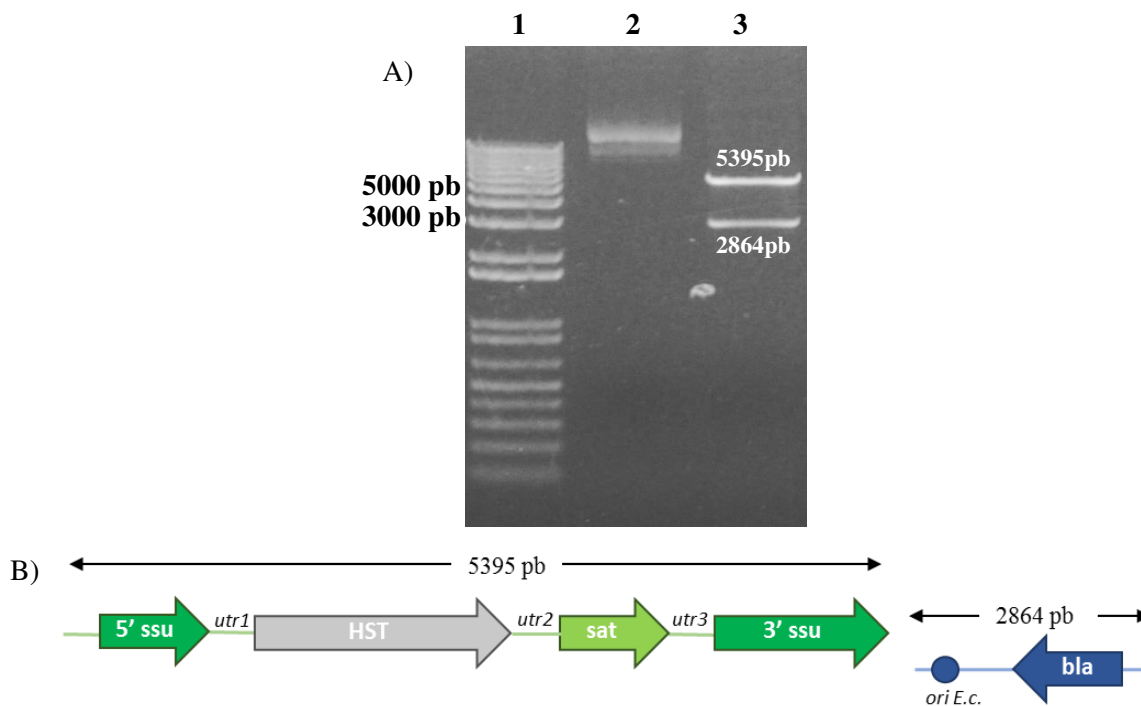


Figura 17. Eletroforese em gel de agarose 0,8% com o resultado da digestão da construção pLEXY-egfp-SAT2+HST, extraída dos clones recombinantes, com a enzima de restrição *Swa*I. Linha 1: padrão 1Kb plus DNA ladder (Invitrogen). Linha 2: Controle negativo. Linha 3: fragmentos oriundos da digestão da construção extraída de um clone recombinante com a enzima *Swa*I.

Os fragmentos linearizados contendo os genes de interesse foram transfectados em *L.tarentolae* (5 – 10 µg) seguindo as recomendações do protocolo fornecido pela Jena Bioscience. Como a eletroporação com DNAs, as células depois de recuperadas foram plaqueadas em meio sólido após este processo. Após dois dias de incubação já foi observado o surgimento de colônias isoladas (duas), que foram inoculadas em meio BHI para expansão das células.

5.7. PCR diagnóstico para verificação de recombinação homóloga

Neste ensaio foram utilizadas duas reações de PCR que serviram para constatar a inserção do cassete de expressão do gene IFN-β no locus gênico das células transfectadas com as construções linearizadas.

Existem algumas vantagens para se utilizar a expressão gênica de forma integrativa constitutiva. No caso da inserção gênica no locus ssu de tripanossomatídeos, podemos citar a presença de muitas cópias de genes do 18S rDNA (em torno de 20 a 40) permitindo a integração de sequências exógenas e ainda mantendo a função e expressão dos genes de

origem. Além disso, genes ribossomais geralmente são muito expressos, portanto há a possibilidade de uma alta taxa de expressão do gene recombinado (Inga, De Doncker et al. 1998, Misslitz, Mottram et al. 2000, Soysa, Tran et al. 2015).

Em uma reação utilizaram-se os pares de oligonucleotídeos específicos *forward* F3001 e o *reverse* específico para o gene IFN- β e em outra utilizou-se o par de *primers* específicos para o gene. Como um dos controles negativos foi utilizado água no lugar de DNA e como outro controle foi utilizado o material genômico extraído de células de *L. tarentolae* selvagens da cepa P10. O oligonucleotídeo *forward* F3001 é específico para a região 5'ssu presente apenas no cromossomo do hospedeiro, e espera-se para validar a inserção observar a formação de um fragmento de aproximadamente 1847pb que corresponde a essa região até a região 3' do gene inserido. Foram observadas as bandas de interesse, tanto a que comprova a inserção quanto a da presença do gene (Figura 18).

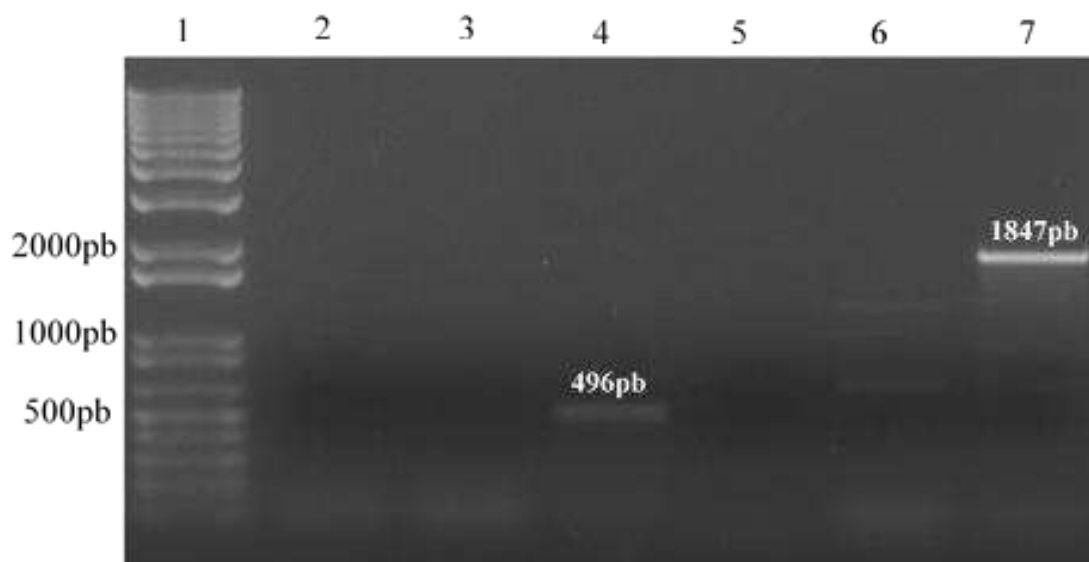


Figura 18. Eletroforese em gel de agarose 1% com o produto reação de PCR diagnóstico. Linha 1: padrão 1Kb plus DNA ladder (Invitrogen). Linha 2 a 4 estão os produtos da reação de PCR utilizando os primers específicos do gene IFN- β . Linha 2: controle negativo onde o DNA foi substituído por água. Linha 3: controle negativo com produto da reação com o material extraído da *Leishmania* selvagem. Linha 4: Amplificação do gene IFN- β em torno de 496 pb Linhas 5, a 7 representam os produtos das reações com o primer de recombinação na região 5'ssu (*forward* F3001) e com o reverso do gene IFN- β . Linha 5: controle negativo onde o DNA foi substituído por água. Linha 6: controle negativo com produto da reação com o material extraído de *Leishmania* selvagem. Linha 7: Amplificação da região onde corresponde a região 5'ssu cromossomal da *Leishmania* e a região 3' do gene IFN- β , em torno de 1847pb.

5.8. Análise da presença de transcritos por RT-PCR

Nesse ensaio foi possível detectar que os clones selecionados até o momento foram integrados ao genoma nos genes de interesse e produziam de fato as moléculas de RNA mensageiro correspondentes as proteínas IFN- β e HST clonadas. No entanto, ainda não podemos afirmar que estejam sendo de fato sintetizadas as moléculas de RNAm maduras ou que esteja acontecendo a tradução eficientemente. Esses organismos possuem a característica de seus RNAs mensageiros serem produzidos como transcritos precursores únicos ou policistrônicos. Esses precursores então são processados pós-transcrição por meio de reações *trans-splicing* e poliadenilação nas regiões intergênicas (Figura 19) (De Gaudenzi, Noe et al. 2011). Estudos demonstram também que os níveis de RNAm não necessariamente estão relacionados com os níveis de proteínas e que há um rígido controle em eventos transcripcionais e pós transcripcionais nesses parasitos (Martínez-Calvillo S et al. 2010).

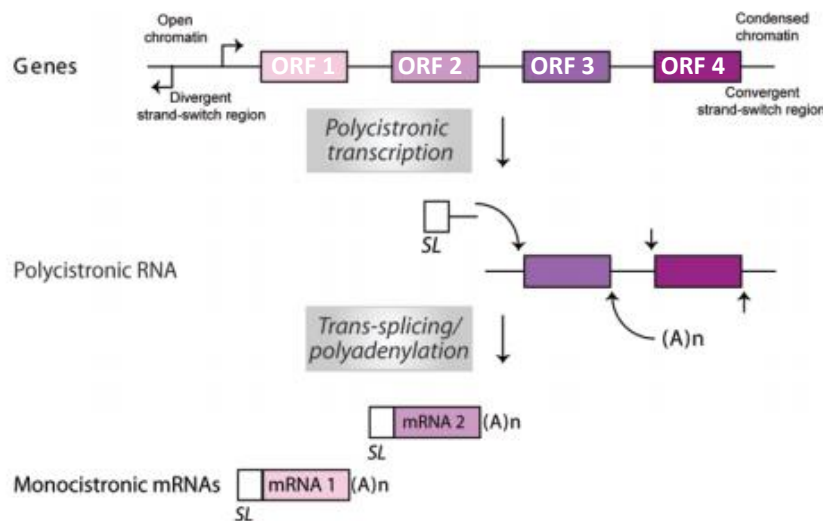


Figura 19. Esquema geral do mecanismo de transcrição em tripanossomatídeos. Os genes que podem possuir funções bem distintas são organizados em grupos policistrônicos (ORFs) mantidas na mesma direção de transcrição, que será realizada pela RNA Polimerase II. A esses genes então são adicionados os RNA Splice Leader na região 5' terminal (trans-splicing) e posteriormente acontece a poliadenilação na região 3' terminal para gerar os mRNAs maduros.(De Gaudenzi, Noe et al. 2011).

Após a extração de RNA total (Figura 20) e síntese do cDNA em uma reação de RT-PCR, os *primers* específicos para as sequências nucleotídicas referentes às proteínas desejadas foram utilizados para o PCR diagnóstico (Figura 21). O teste foi realizado após o cultivo das células recombinantes e selvagens para posterior extração do RNA e digestão com a enzima DNase para ter certeza de que todo o DNA das amostras de RNA celular

total teria sido degradado. Os resultados mostraram a presença de transcritos em todos os clones positivos testados, inclusive para a transfecção dupla. Como controle negativo ambas as reações tiveram em comum o RNA de células selvagens ou o material genético substituído por água.

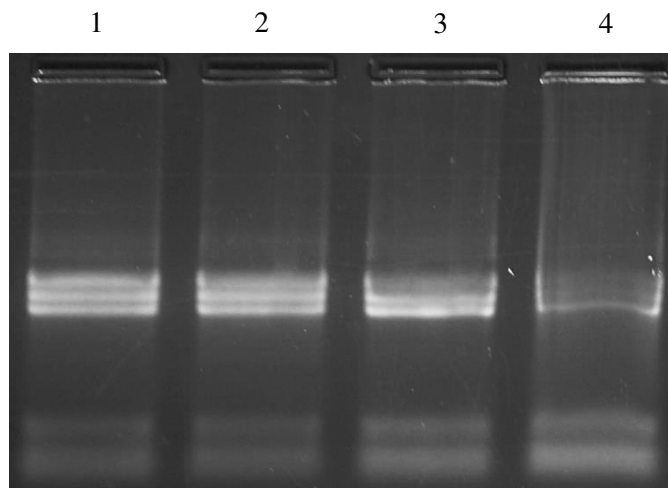


Figura 20. Eletroforese em gel de agarose 0,8% com os produtos da extração de RNA total das culturas de *L.tarentolae* transfectadas. Linha 1: controle negativo com o material extraído da Leishmania selvagem. Linha 2: material extraído das culturas clonadas com o gene IFN- β . Linha3: material extraído das culturas clonadas com o gene HST. Linha 4: material extraído das culturas clonadas com os dois genes, IFN- β e HST.

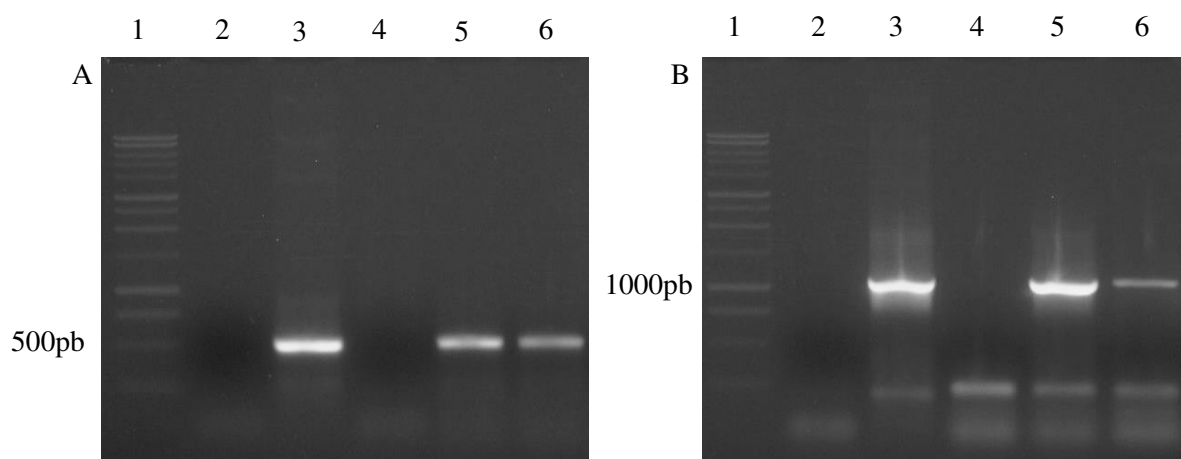


Figura 21. Eletroforese em gel de agarose 1% com os produto da amplificação da reação de PCR utilizando os primers específicos para o gene IFN- β (A) e HST (B) a partir do molde de cDNA obtido pelo RT-PCR usando o RNA total das células selvagens e dos clones. Linhas 1: padrão 1Kb DNA ladder (Promega). Linhas 2: controle negativo onde o DNA foi substituído por água. Linhas 3: controle positivo. Linhas 4: amplificação obtida tendo com molde o cDNA das culturas de *L. tarentolae* selvagens. Linha 5: amplificação obtida tendo com molde o cDNA das culturas de *L.tarentolae* clonadas com o gene IFN- β (A) e HST (B). Linhas 6: amplificação obtida tendo com molde o cDNA das culturas de *L. tarentolae* clonadas com os dois genes (IFN- β e HST).

5.9. Avaliação da expressão proteica de IFN- β por SDS-PAGE e Dotblotting

Para verificar a expressão proteica de IFN- β foram aplicadas em um gel de poliacrilamida 14% (figura 22) as amostras proteicas obtidas pela precipitação com TCA do meio de cultivo das culturas de *L. tarentolae* selvagens e dos transformadas (20 μ g de proteínas totais). Foi adicionado tampão de amostra às amostras preparadas como descrito no tópico 4.18 e então submetidas à fervura por 5 min antes da aplicação.

A figura 22 mostra duas bandas características (entre 15 e 25kDa) do controle positivo e que são compatíveis ao tamanho molecular encontrados nas diferentes isoformas expressas da proteína IFN- β (forma glicosilada e não glicosilada) como descrito em diferentes trabalhos já publicados ((Runkel, Meier et al. 1998, Dissing-Olesen, Thaysen-Andersen et al. 2008)). Não foram observadas as bandas equivalente na linha onde foram aplicadas as amostras das células selvagens (WT).

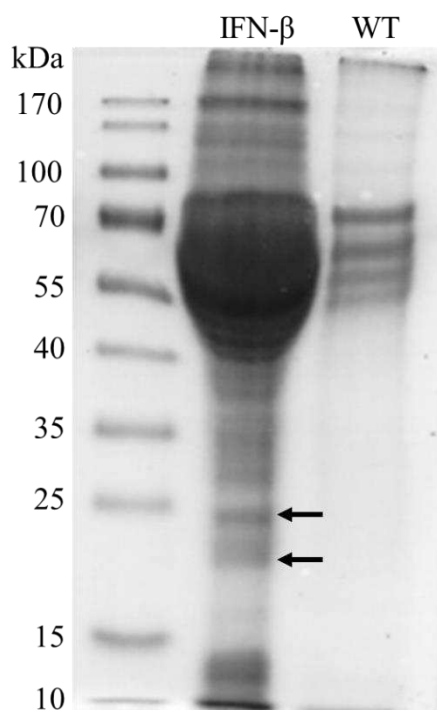


Figura 22. SDS-PAGE 14% com as amostras de sobrenadante de culturas clonais (IFN- β) e selvagens (WT). Padrão molecular: PageRuler Prestained Protein Ladder (ThermoScientific).

Para o ensaio de Dotblotting as mesmas amostras utilizadas no SDS-PAGE foram aplicadas sobre uma membrana de nitrocelulose e submetidas aos procedimentos descritos no item 4.19 da metodologia. Após incubação com os anticorpos (inicialmente os

primários anti-His e anti-IFN- β e depois os secundários conjugados à peroxidase), a revelação da membrana com DAB e peróxido de hidrogênio identificou precipitação nos dois tratamentos, sendo indicada a presença da proteína IFN- β e/ou proteínas contendo cauda de histidina nos extratos aplicados (figura 23).

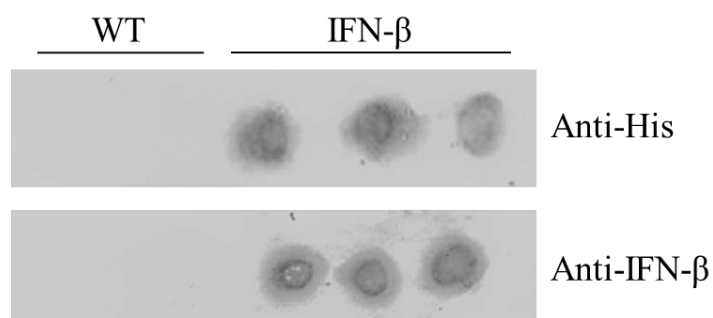


Figura 23. Dot Blotting. O ensaio teve como amostras as proteínas precipitadas do meio de cultivo dos clones selecionados que possuíam a presença do transcrito do gene IFN- β . Foram aplicados em torno de 40 μ g de proteína em cada ponto. Para o controle negativo foram utilizadas as proteínas precipitadas do meio de cultivo de células selvagens. Observou-se marcação do precipitado indicando interação dos anticorpos tanto anti-His (diluição 1:3000) quanto anti-IFN- β (diluição 1:3000) com as amostras obtidas. Não foram observados precipitados nos controles negativos.

Todos os ensaios descritos acima foram realizados também com os extratos proteicos obtidos com o *pellet* das culturas, porém observou-se resultados semelhante entre os tratamentos e controles negativos. Como há uma grande variedade de proteínas encontradas nesse tipo de extrato comparada com a do meio extracelular, limitaria ainda mais a especificidade dos métodos como *Dotblotting* e como espera-se que a proteína de interesse seja expressa e secretada pela presença do peptídeo sinal, decidimos não apresentar os resultados com esses extratos nesse trabalho, demonstrando apenas os resultados obtidos com os extratos proteicos oriundos do meio de cultivo.

6. CONCLUSÃO

O principal intuito deste trabalho se baseia no desenvolvimento de uma plataforma de expressão mais acessível de glicoproteínas com propriedades bioquímicas e farmacocinéticas mais semelhantes ao organismo de origem utilizando *L. tarentolae*. Esse organismo tem sido muito estudado como alternativa a outros sistemas de expressão, se tornando um meio mais fácil, barato e relativamente mais eficiente na produção de

glicoproteínas terapêuticas. Como apresentado na seção de revisão de literatura (item 2.5), apesar de todas as vantagens que a produção de proteínas recombinantes utilizando *L. tarentolae* como hospedeiro apresenta, o perfil de glicosilação destes organismos não é idêntico ao de mamíferos.

As abordagens apresentadas neste projeto visaram entender mais sobre as vias metabólicas encontradas nestes organismos que pudessem indicar uma alternativa de tornar o sistema de expressão em *L. tarentolae* mais otimizado. Como já se sabe, o sistema LEXSY de expressão é característico por possuir toda uma metodologia já estabelecida com plasmídeos planejados e desenvolvidos para a produção de proteínas heterólogas com o padrão de glicosilação semelhante ao de mamíferos com alto rendimento. Porém, esse método falha na adição do último glicano da cadeia glicídica, característico das proteínas *N-glicosiladas*, o ácido siálico. Somente essa diferença pode acarretar na diminuição da atividade de fármacos, com também gerar efeitos colaterais e adversos ao tratamento.

Por meio das metodologias de genômica e metabolômica apresentadas foi possível desenvolver uma primeira estratégia como tentativa de resolver esse problema por meio da expressão heteróloga da enzima sialiltransferase, e tendo a proteína interferon- β como alvo para validação do potencial das novas linhagens otimizadas. O uso do IFN- β neste trabalho é destacado tanto pela sua importância como fármaco, pois é utilizado no tratamento de doenças graves, quanto pelo impacto econômico gerado em sistemas públicos e privados de saúde devido à alta demanda e alto custo de produção. Também não são encontradas publicações que citam a expressão dessa proteína em *L. tarentolae*.

Todas as etapas de clonagem, bem como os ensaios confirmatórios apresentaram resultados satisfatórios que demonstram a eficiência do método. Ainda não temos a confirmação de expressão das proteínas por metodologias mais confiáveis como *WesternBlot* juntamente com os testes de atividades das proteínas expressas. Mas, os indícios apresentados pelo sucesso nas técnicas de clonagem, pela presença dos transcritos confirmada no RT-PCT, e nas técnicas de SDS-PAGE e Dotblotting, fornecem ainda mais embasamento para a nossa proposta.

7. PERSPECTIVAS

- Usar dados de simulação de fluxo associados a metodologias de cultivo para analisar o impacto da adição do substrato CMP-Neu5Ac e da enzima sialiltransferase ao metabolismo endógeno de *L. tarentolae* selvagens.
- Otimizar e escalar a expressão heteróloga de proteínas recombinantes dependente de glicosilação em linhagens otimizadas de *L. tarentolae*.
- Analisar a expressão proteica diferencial de linhagens otimizadas de *L.tarentolae* utilizando ferramentas proteômicas, como eletroforese bidimensional.
- Analisar o perfil e presença de glicosilação encontrados nas proteínas expressas pelas linhagens de *L.tarentolae* otimizadas por espectrometria de massas.
- Testar a atividade de interferon- β produzido em *L. tarentolae* em reduzir o processo inflamatório em cultura de células de mamíferos.
- Realizar teste de atividade comparativa de proteínas recombinantes expressas em linhagens geneticamente otimizadas de *L. tarentolae* e expressas em sistemas de *E. coli* e *S. cerevisiae* e células de mamíferos CHO (*Chinese hamster ovary*).

8. BIBLIOGRAFIA

ABÈS, R.; TEILLAUD, J. L. Impact of glycosylation on effector functions of therapeutic. IgGPharmaceuticals, 2010.

AGREN, R. et al. The RAVEN Toolbox and Its Use for Generating a Genome-scale Metabolic Model for *Penicillium chrysogenum*. PLoS Computational Biology, v. 9, 2013.

ALBRECHT, S. (2014). Therapeutic proteins: facing the challenges of glycobiology. Journal of Health Policy & Outreach Research. 1. 12-17

AMERICAN TYPE CULTURE COLLECTION. MTT Cell Proliferation Assay Instruction Guide. Components, v. 6597, p. 1–6, 2011.

ANVISA, 2017. Disponível em : <http://portal.anvisa.gov.br/listas-de-precos>. Acesso em 24/09/2017.

ASLETT, M. et al. TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic acids research, v. 38, p. D457–62, 2010.

BAKER, M. Protein engineering: navigating between chance and reason. Nature methods, v. 8, p. 623–626, 2011.

BARRETO-BERGTER, E. et al. Structures of Glycolipids Found in Trypanosomatids: Contribution to Parasite Functions, 2010.

BASILE, G.; PETICCA, M. Recombinant Protein Expression in *Leishmania tarentolae*. Molecular Biotechnology, v. 43, p. 273–278, 2009.

BERGER, M.; KAUP, M.; BLANCHARD, V. Protein glycosylation and its impact on biotechnology. Advances in Biochemical Engineering/Biotechnology, 2012.

BONDT, A. et al. IgG Fab glycosylation analysis using a new mass spectrometric high-throughput profiling method reveals pregnancy-associated changes. Molecular & cellular proteomics : MCP, v. 31, p. 1–30, 2014.

BORNSCHEUER, U.; KAZLAUSKAS, R. J. Survey of protein engineering strategies. Current Protocols in Protein Science, v. 1, 2011.

BREITLING, R., S. KLINGNER, N. CALLEWAERT, R. PIETRUCHA, A. GEYER, G. EHRLICH, R. HARTUNG, A. MULLER, R. CONTRERAS, S. M. BEVERLEY AND K. ALEXANDROV (2002). "Non-pathogenic trypanosomatid protozoa as a platform for protein research and production." *Protein Expr Purif* 25(2): 209-218.

BRETON, M. et al. Live nonpathogenic parasitic vector as a candidate vaccine against visceral leishmaniasis. *Infection and Immunity*, v. 73, p. 6372–6382, 2005.

CHUSACULTANACHAI, S.; YUTHAVONG, Y. Random mutagenesis strategies for construction of large and diverse clone libraries of mutated DNA fragments. *Methods in molecular biology* (Clifton, N.J.), v. 270, p. 319–334, 2004.

CUI, Q. et al. Metabolite identification via the Madison Metabolomics Consortium Database. *Nature biotechnology*, 2008.

DALTON, A. C.; BARTON, W. A. Over-expression of secreted proteins from mammalian cell lines. *Protein science : a publication of the Protein Society*, v. 23, p. 517–25, 2014.

DANGORIA, N. S. et al. HLA-B27 misfolding is associated with aberrant intermolecular disulfide bond formation (dimerization) in the endoplasmic reticulum. *Journal of Biological Chemistry*, v. 277, p. 23459–23468, 2002.

DE GAUDENZI, J. G., G. NOE, V. A. CAMPO, A. C. FRASCH AND A. CASSOLA (2011). "Gene expression regulation in trypanosomatids." *Essays Biochem* 51: 31-46.

DEMAIN, A. L. AND P. VAISHNAV (2009). "Production of recombinant proteins by microbes and higher organisms." *Biotechnol Adv* 27(3): 297-306.

DINGERMANN, T. (2008). "Recombinant therapeutic proteins: production platforms and challenges." *Biotechnol J* 3(1): 90-97.

DISSING-OLESEN, L., M. THAYSEN-ANDERSEN, M. MELDGAARD, P. HOJRUP AND B. FINSEN (2008). "The function of the human interferon-beta 1a glycan determined in vivo." *J Pharmacol Exp Ther* 326(1): 338-347.

DOURIS, V. et al. Stably Transformed Insect Cell Lines: Tools for Expression of Secreted and Membrane-anchored Proteins and High-throughput Screening Platforms for Drug and Insecticide Discovery. *Advances in Virus Research*, 2006.

DWEK, R. A. Biological importance of glycosylation. *Developments in biological standardization*, v. 96, p. 43–47, 1998.

DWEK, R. A. et al. Targeting glycosylation as a therapeutic approach. *Nature reviews. Drug discovery*, v. 1, p. 65–75, 2002.

EDGAR, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, v. 32, p. 1792–1797, 2004.

EVALUATEPHARMA: World preview 2014, outlook to 2020: Returning to growth

FERGUSON, M. A. (1997). "The surface glycoconjugates of trypanosomatid parasites." *Philos Trans R Soc Lond B Biol Sci* 352(1359): 1295-1302.

FERNANDEZ, F. J. AND M. C. VEGA (2013). "Technologies to keep an eye on: alternative hosts for protein production in structural biology." *Curr Opin Struct Biol* 23(3): 365-373.

FINN, R. D.; CLEMENTS, J.; EDDY, S. R. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Research*, v. 39, 2011.

FORTES-DIAS, C. L. et al. Insights on the structure of native CNF, an endogenous phospholipase A2 inhibitor from *Crotalus durissus terrificus*, the South American rattlesnake. *Biochimica et biophysica acta*, v. 1844, n. 9, p. 1569–1579, set. 2014.

GERNGROSS, T. U. Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. *Nature biotechnology*, v. 22, p. 1409–1414, 2004.

GHADERI, D., M. ZHANG, N. HURTADO-ZIOLA AND A. VARKI (2012). "Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation." *Biotechnol Genet Eng Rev* 28: 147-175.

GHOSHAL, A. et al. Sialic acids in different *Leishmania* sp., its correlation with nitric oxide resistance and host responses. *Glycobiology*, v. 20, p. 553–566, 2010.

GHOSHAL, A.; MANDAL, C. A Perspective on the Emergence of Sialic Acids as Potent Determinants Affecting *Leishmania* Biology. *Molecular Biology International*, 2011.

GLOCKSHUBER, R.; SCHMIDT, T.; PLÜCKTHUN, A. The disulfide bonds in antibody variable domains: effects on stability, folding in vitro, and functional expression in *Escherichia coli*. *Biochemistry*, v. 31, p. 1270–1279, 1992.

GOLDENBERG, M. M. Multiple sclerosis review. [s.l: s.n.]. v. 37p. 175–84.

GRÄSLUND, S. et al. Protein production and purification. *Nature methods*, v. 5, p. 135–146, 2008.

GUO, Z. Y.; JIA, X. Y.; FENG, Y. M. Replacement of the interchain disulfide bridge-forming amino acids A7 and B7 by glutamate impairs the structure and activity of insulin. *Biological Chemistry*, v. 385, p. 1171–1175, 2004.

HAN, Y. K., T. Y. KOO AND G. M. LEE (2009). "Enhanced interferon-beta production by CHO cells through elevated osmolality and reduced culture temperature." *Biotechnol Prog* 25(5): 1440-1447.

HEMAYATKAR, M. et al. Increased expression of recombinant human tissue plasminogen activator in *Leishmania tarentolae*. *Biotechnology Journal*, v. 5, p. 1198–1206, 2010.

HOLTEN, J. VAN; PLATER-ZYBERK, C.; TAK, P. P. Interferon-beta for treatment of rheumatoid arthritis? *Arthritis research*, v. 4, p. 346–352, 2002.

HRISTODOROV, D.; FISCHER, R.; LINDEN, L. With or without sugar? (A)glycosylation of therapeutic antibodies. *Molecular Biotechnology*, 2013.

HUA, Q. X. et al. The folding nucleus of the insulin superfamily: A flexible peptide model foreshadows the native state. *Journal of Biological Chemistry*, v. 281, p. 28131–28142, 2006.

HUGHES, R. A.; MIKLOS, A. E.; ELLINGTON, A. D. Gene synthesis: Methods and applications. *Methods in Enzymology*, v. 498, p. 277–309, 2011.

INGA, R., S. DE DONCKER, J. GOMEZ, M. LOPEZ, R. GARCIA, D. LE RAY, J. AREVALO AND J. C. DUJARDIN (1998). "Relation between variation in copy number of ribosomal RNA encoding genes and size of harbouring chromosomes in *Leishmania* of subgenus *Viannia*." *Mol Biochem Parasitol* 92(2): 219-228.

J. SAMBROOK, E. F. FRITSCH, AND T. MANIATIS, *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor laboratory press. 1989

JERMUTUS, L. et al. Structure-based chimeric enzymes as an alternative to directed enzyme evolution: Phytase as a test case. *Journal of Biotechnology*, v. 85, p. 15–24, 2001.

KANEHISA, M. The KEGG database. *Novartis Foundation symposium*, v. 247, p. 91–101; discussion 101–103, 119–128, 244–252, 2002.

KAUFMAN, R. J. Post-translational modifications required for coagulation factor secretion and function. *Thromb Haemost*, v.79, n.6, p.1068-79, Jun. 1998.

KAZLAUSKAS, R. J.; BORNSCHEUER, U. T. Finding better protein engineering strategies. *Nature chemical biology*, v. 5, p. 526–529, 2009.

KIANMEHR, A., A. MAHROOZ, M. OLADNABI, Y. SAFDARI, J. ANSARI, K. VEISI, M. EVAZALIPOUR, H. SHAHBAZMOHAMMADI AND E. OMIDINIA (2016). "Purification and Characterization of Recombinant Darbepoetin Alfa from *Leishmania tarentolae*." *Mol Biotechnol* 58(8-9): 566-572.

KLATT, S. AND Z. KONTHUR (2012). "Secretory signal peptide modification for optimized antibody-fragment expression-secretion in *Leishmania tarentolae*." *Microb Cell Fact* 11: 97.

KUKURUZINSKA, M. A.; LENNON, K. Protein N-glycosylation: molecular genetics and functional significance. *Critical reviews in oral biology and medicine: an official publication of the American Association of Oral Biologists*, v. 9, p. 415–448, 1998.

KUMAR, S.; TSAI, C. J.; NUSSINOV, R. Factors enhancing protein thermostability. *Protein engineering*, v. 13, p. 179–191, 2000.

LALONDE, M. E. AND Y. DUROCHER (2017). "Therapeutic glycoprotein production in mammalian cells." *J Biotechnol* 251: 128-140.

LEE, J. M.; GIANCHANDANI, E. P.; PAPIN, J. A. Flux balance analysis in the era of metabolomics. *Briefings in Bioinformatics*, 2006.

LENNON, D. R., J. D. CHERRY, A. MORGENSTEIN, J. G. CHAMPION AND Y. J. BRYSON (1983). "Longitudinal study of influenza B symptomatology and interferon production in children and college students." *Pediatr Infect Dis* 2(3): 212-215.

LIU, H.; MAY, K. Disulfide bond structures of IgG molecules: Structural variations, chemical modifications and possible impacts to stability and biological function. *mAbs*, 2012.

LODISH H, BERK A, ZIPURSKY SL, ET AL. *Molecular Cell Biology*. 4th edition. New York: W.H. Freeman; 2000. Section 17.7, Protein Glycosylation in the ER and Golgi Complex. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK21744/>

MARTÍNEZ-CALVILLO S, VIZUET-DE-RUEDA JC, FLORENCIO-MARTÍNEZ LE, MANNING-CELA RG, FIGUEROA-ANGULO EE. Gene expression in trypanosomatid parasites. *J Biomed Biotechnol*. 2010; Feb 11.

MASTRANGELI, R., M. ROSSI, M. MASCIA, W. PALINSKY, A. DATOLA, M. TERLIZZESE AND H. BIERAU (2015). "In vitro biological characterization of IFN-beta-1a major glycoforms." *Glycobiology* 25(1): 21-29.

MCAULEY, A. et al. Contributions of a disulfide bond to the structure, stability, and dimerization of human IgG1 antibody C(H)3 domain. *Protein Science : A Publication of the Protein Society*, v. 17, n. 1, p. 95–106, 20 jan. 2008.

MEEHAN, H. A., R. A. LUNDBERG AND G. J. CONNELL (2000). "A trypanosomatid protein specifically interacts with a mammalian iron-responsive element." *Parasitol Res* 86(2): 109-114.

MENZEL, N. et al. MAP-kinase regulated cytosolic phospholipase A2 activity is essential for production of infectious Hepatitis C virus particles. *PLoS Pathogens*, v. 8, p. 21, 2012.

MINISTÉRIO DA SAÚDE, Portal da Saúde, SUS. Avanços no tratamento do SUS beneficiam pacientes hemofílicos. 2015. (<http://portalsaude.saude.gov.br/index.php/cidadao/principal/agencia-saude/17480-avancos-no-tratamento-do-sus-beneficiam-pacientes>. Acesso em 24/09/2017).

- MIRZAAHMADI, S. et al. Expression of recombinant human coagulation factor VII by the lizard *Leishmania* expression system. *Journal of Biomedicine and Biotechnology*, v. 2011, 2011.
- MISSLITZ, A., J. C. MOTTRAM, P. OVERATH AND T. AEBISCHER (2000). "Targeted integration into a rRNA locus results in uniform and high level expression of transgenes in *Leishmania* amastigotes." *Mol Biochem Parasitol* 107(2): 251-261.
- MISTRY, P. K.; WRAIGHT, E. P.; COX, T. M. Therapeutic delivery of proteins to macrophages: Implications for treatment of Gaucher's disease. *Lancet*, v. 348, p. 1555–1559, 1996.
- MUHL, D.; FILLOUX, A. Site-directed mutagenesis and gene deletion using reverse genetics. *Methods in molecular biology* (Clifton, N.J.), v. 1149, p. 521–39, 2014.
- MUS-VETEAU, I. Heterologous expression and purification systems for structural proteomics of mammalian membrane proteins. *Comparative and Functional Genomics*, 2002.
- NEHETE, J. Y. et al. Natural proteins: Sources, isolation, characterization and applications. *Pharmacognosy reviews*, v. 7, p. 107–116, 2013.
- NIWA, J. I. et al. Disulfide bond mediates aggregation, toxicity, and ubiquitylation of familial amyotrophic lateral sclerosis-linked mutant SOD1. *Journal of Biological Chemistry*, v. 282, p. 28087–28095, 2007.
- PALOMARES, L. A.; ESTRADA-MONDACA, S.; RAMÍREZ, O. T. Production of recombinant proteins: challenges and solutions. *Methods in molecular biology* (Clifton, N.J.), v. 267, p. 15–52, 2004.
- PERALES, J. et al. Molecular structure and mechanism of action of the crotoxin inhibitor from *Crotalus durissus terrificus* serum. *European journal of biochemistry / FEBS*, v. 227, p. 19–26, 1995.
- PONTES, JOÃO GUILHERME M. et al. NMR-based metabolomics strategies: plants, animals and humans, 2016.
- RABERT, C. et al. Recombinants proteins for industrial uses: Utilization of *Pichia pastoris* expression system. *Brazilian Journal of Microbiology*, 2013.

RABINER, L. R. A tutorial on hidden Markov models and selected applications in speech recognition Proceedings of the IEEE, 1989. Disponível em: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=18626@ieeejrns.

RAYMOND, F. et al. Genome sequencing of the lizard parasite *Leishmania tarentolae* reveals loss of genes associated to the intracellular stage of human pathogenic species. *Nucleic Acids Research*, v. 40, p. 1131–1147, 2012.

RIEMERSMA, M., J. SANDROCK, T. J. BOLTJE, C. BULL, T. HEISE, A. ASHIKOV, G. J. ADEMA, H. VAN BOKHOVEN AND D. J. LEFEBER (2015). "Disease mutations in CMP-sialic acid transporter SLC35A1 result in abnormal alpha-dystroglycan O-mannosylation, independent from sialic acid." *Hum Mol Genet* 24(8): 2241-2246.

RUNKEL, L., W. MEIER, R. B. PEPINSKY, M. KARPUSAS, A. WHITTY, K. KIMBALL, M. BRICKELMAIER, C. MULDOWNNEY, W. JONES AND S. E. GOELZ (1998). "Structural and functional differences between glycosylated and non-glycosylated forms of human interferon-beta (IFN-beta)." *Pharm Res* 15(4): 641-649.

RUSSEL, E.; WANG, A.; RATHORE, A. Process Scale Bioseparations for the Biopharmaceutical Industry. In: *Process Scale Bioseparations for the Biopharmaceutical Industry*. [s.l: s.n.]. p. 1–58.

SANTOS, R. M. M. DOS et al. Inhibition of crotoxin binding to synaptosomes by a receptor-like protein from *Crotalus durissus terrificus* (the South American rattlesnake). *Biochimica et Biophysica Acta - Biomembranes*, v. 1717, p. 27–33, 2005.

SCHELLENBERGER, J. et al. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature protocols*, v. 6, p. 1290–1307, 2011.

SCHMIDT, Cleber A. et al. Avaliação da atividade e caracterização de eritropoietina humana recombinante em produtos farmacêuticos. *Arq Bras Endocrinol Metab*, São Paulo, v. 47, n. 2, p. 183-189, Apr. 2003.

SEGRÈ, D.; VITKUP, D.; CHURCH, G. M. Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences of the United States of America*, v. 99, p. 15112–15117, 2002.

SELMAN, M. H. J. et al. Immunoglobulin G glycopeptide profiling by matrix-assisted laser desorption ionization Fourier transform ion cyclotron resonance mass spectrometry. *Analytical chemistry*, v. 82, p. 1073–1081, 2010.

SEO, Y. J. et al. Validation of a HeLa Mx2/Luc reporter cell line for the quantification of human type I interferons. *Pharmacology*, v. 84, p. 135–144, 2009.

SIMPSON, L.; FRECH, G. C.; MASLOV, D. A. [10] RNA editing in trypanosomatid mitochondria. In: GIUSEPPE M. ATTARDI, A. C. B. T.-M. IN E. (Ed.). *Mitochondrial Biogenesis and Genetics Part B*. [s.l.] Academic Press, 1996. v. Volume 264p. 99–121.

SMALL, G. W.; MCLEOD, H. L.; RICHARDS, K. L. Analysis of innate and acquired resistance to anti-CD20 antibodies in malignant and nonmalignant B cells. *PeerJ*, v. 1, p. e31, 2013.

SOYSA, R., K. D. TRAN, B. ULLMAN AND P. A. YATES (2015). "Integrating ribosomal promoter vectors that offer a choice of constitutive expression profiles in *Leishmania donovani*." *Mol Biochem Parasitol* 204(2): 89-92.

SPIRO, R. G. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, v. 12, p. 43R–56R, 2002.

SUN, S.; GENG, L.; SHAMOO, Y. Structure and enzymatic properties of a chimeric bacteriophage RB69 DNA polymerase and single-stranded DNA binding protein with increased processivity. *Proteins: Structure, Function and Genetics*, v. 65, p. 231–238, 2006.

TAKEUCHI, M. et al. Relationship between sugar chain structure and biological activity of recombinant human erythropoietin produced in Chinese hamster ovary cells. *Proceedings of the National Academy of Sciences of the United States of America*, v. 86, p. 7819–7822, 1989.

TANABE, M.; KANEHISA, M. Using the KEGG database resource. *Current Protocols in Bioinformatics*, 2012.

TENG, S.; SRIVASTAVA, A. K.; WANG, L. Sequence feature-based prediction of protein stability changes upon amino acid substitutions. *BMC genomics*, v. 11 Suppl 2, p. S5, 2010.

USMANI, S. S., G. BEDI, J. S. SAMUEL, S. SINGH, S. KALRA, P. KUMAR, A. A. AHUJA, M. SHARMA, A. GAUTAM AND G. P. S. RAGHAVA (2017). "THPdb: Database of FDA-approved peptide and protein therapeutics." *PLoS One* 12(7): e0181748.

VAN KARNEBEEK, C. D., L. BONAFE, X. Y. WEN, M. TARAILO-GRAOVAC, S. BALZANO, B. ROYER-BERTRAND, A. ASHIKOV, L. GARAVELLI, I. MAMMI, L. TUROLLA, C. BREEN, D. DONNAI, V. CORMIER, D. HERON, G. NISHIMURA, S. UCHIKAWA, B. CAMPOS-XAVIER, A. ROSSI, T. HENNET, K. BRAND-ARZAMENDI, J. ROZMUS, K. HARSHMAN, B. J. STEVENSON, E. GIRARDI, G. SUPERTI-FURGA, T. DEWAN, A. COLLINGRIDGE, J. HALPARIN, C. J. ROSS, M. I. VAN ALLEN, A. ROSSI, U. F. ENGELKE, L. A. KLUIJTMANS, E. VAN DER HEEFT, H. RENKEMA, A. DE BROUWER, K. HUIJBEN, F. ZIJLSTRA, T. HEISSE, T. BOLTJE, W. W. WASSERMAN, C. RIVOLTA, S. UNGER, D. J. LEFEBER, R. A. WEVERS AND A. SUPERTI-FURGA (2016). "NANS-mediated synthesis of sialic acid is required for brain and skeletal development." *Nat Genet* 48(7): 777-784.

VEGA, M. C. (2016). "Advanced Technologies for Protein Complex Production and Characterization." *Anticancer Res* 36(8): 4375.

WALSH, G. (2006). "Biopharmaceutical benchmarks 2006." *Nat Biotechnol* 24(7): 769-776.

WALSH, G. Biopharmaceutical benchmarks 2014. *Nat Biotech*, v. 32, n. 10, p. 992–1000, out. 2014.

WEDEMEYER, W. J. et al. Disulfide bonds and protein folding. *Biochemistry*, v. 39, p. 4207–4216, 2000.

WIELAND, F. Structure and biosynthesis of prokaryotic glycoproteins. *Biochimie*, v. 70, p. 1493–1504, 1988.

WIESE, M., T. ILG, F. LOTTSPEICH AND P. OVERATH (1995). "Ser/Thr-rich repetitive motifs as targets for phosphoglycan modifications in *Leishmania mexicana* secreted acid phosphatase." *EMBO J* 14(6): 1067-1074.

XIONG, A. S. et al. Chemical gene synthesis: Strategies, softwares, error corrections, and applications. *FEMS Microbiology Reviews*, 2008.

YOON, S. H.; KIM, S. K.; KIM, J. F. Secretory production of recombinant proteins in *Escherichia coli*. *Recent patents on biotechnology*, v. 4, p. 23–29, 2010.

ZAUNER, G. et al. Protein O-glycosylation analysis. *Biological Chemistry*, 2012.

ZAWAIRA, A. et al. A discussion of molecular biology methods for protein engineering. *Molecular Biotechnology*, 2012.

ZHANG, L.; CHOU, C. P.; MOO-YOUNG, M. Disulfide bond formation and its impact on the biological activity and stability of recombinant therapeutic proteins produced by *Escherichia coli* expression system. *Biotechnology Advances*, 2011.