

UNIVERSIDADE FEDERAL DO PAMPA

JÉSSICA SILVA TAPIA

**CLONAGEM E EXPRESSÃO DA TRANSPOZASE *mos1* DE *Drosophila simulans*
E O DESENHO *IN SILICO* DE UM VETOR PARA EXPRESSÃO HETERÓLOGA DE
PROTEÍNAS**

**São Gabriel
2017**

JÉSSICA SILVA TAPIA

**CLONAGEM E EXPRESSÃO DA TRANSPOSASE *mos1* DE *Drosophila simulans*
E O DESENHO *IN SILICO* DE UM VETOR PARA EXPRESSÃO HETERÓLOGA DE
PROTEÍNAS**

Dissertação apresentada à Comissão de Pós-Graduação em Ciências Biológicas, Universidade Federal do Pampa – UNIPAMPA, Campus São Gabriel, como requisito necessário à obtenção do grau de Mestre em Ciências Biológicas.

Orientador: Dr. Paulo Marcos Pinto

Co-orientador: Dr. Gabriel da Luz Wallau

**São Gabriel
2017**

Ficha catalográfica elaborada automaticamente com os dados fornecidos pelo(a) autor(a) através do Módulo de Biblioteca do Sistema GURI (Gestão Unificada de Recursos Institucionais) .

T172 Tapia, Jéssica Silva
CLONAGEM E EXPRESSÃO DATRANSPOSASE mos1 DE *Drosophila simulans* E O DESENHO IN SILICO DE UM VETOR PARA EXPRESSÃO HETERÓLOGA DE PROTEÍNAS / Jéssica Silva Tapia.
66 p.

Dissertação (Mestrado)-- Universidade Federal do Pampa, MESTRADO EM CIÊNCIAS BIOLÓGICAS, 2017.
"Orientação: Paulo Marcos Pinto".

1. Clonagem e expressão. 2. Transposons. 3. *Drosophila simulans*. 4. Vetor de expressão. 5. Expressão heteróloga de proteínas. I. Título.

JÉSSICA SILVA TAPIA

**CLONAGEM E EXPRESSÃO DE UMA TRANSPOSASE DE *Drosophila simulans*
E O DESENHO RACIONAL DE UM VETOR PARA EXPRESSÃO HETERÓLOGA
DE PROTEÍNAS**

Dissertação apresentada à Comissão de Pós-Graduação em Ciências Biológicas, Universidade Federal do Pampa – UNIPAMPA, Campus São Gabriel, como requisito necessário à obtenção do grau de Mestre em Ciências Biológicas.

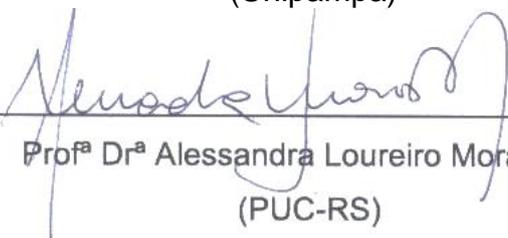
Dissertação defendida e aprovada em 16 de março de 2017.

Banca examinadora:

Prof. Dr. Paulo Marcos Pinto

Orientador

(Unipampa)



Prof.ª Dr.ª Alessandra Loureiro Morassutti
(PUC-RS)

Prof. Dr. Andrés Delgado Cañedo

(Unipampa)

São Gabriel

2017

DEDICATÓRIA

Dedico este trabalho aos meus pais Sandro Luís de Lima Tapia e Angela Maria Vieira da Silva Tapia, por todo o esforço que fizeram para que chegasse até aqui, por todo amor e pela educação me deram. Esta conquista é de vocês.

AGRADECIMENTO

Á Deus agradeço pelo dom da vida e pela saúde para poder trabalhar e estudar. Por me guiar e iluminar em todas as horas de dificuldade.

Aos meus pais, Sandro e Angela, por todo suporte, pois sem vocês não estaria aqui, pelo amor, paciência e por acreditarem em mim, mesmo nos momentos em que eu mesma não acreditava. Por serem meus pilares e meu modelo a seguir. Obrigada por me levantarem a cada tombo, por enxugar as minhas lágrimas e comemorarem comigo a cada vitória. Obrigada pelas horas ao telefone, conversas, por me deixarem livre para fazer minhas próprias escolhas e tomar minhas decisões e por estarem comigo em cada uma delas. Obrigada pela educação que me deram, e por me ensinarem a sempre escolher o certo e o justo, estes valores para mim estão acima de tudo. Todo o esforço foi por vocês, eu os amo mais do que tudo!

Aos meus irmãos, Rafael e Gabriel por serem meus primeiros amigos na vida, por serem meu orgulho, por cuidarem de mim mesmo longe, por me apoiarem, e por estarem comigo em cada etapa da minha vida. Que eu carregue em mim a pureza e a bondade de vocês, pois são os seres humanos mais lindos que já conheci em toda minha vida.

Ao meu amor Diorgenes, por todas as horas de paciência, por me incentivar a não desistir, por me fazer sentir segura nos momentos de medo e por acreditar no meu potencial. Obrigada por aguentar as crises de choro, por todo companheirismo e por tornar essa fase mais fácil. Acredito que Deus nos manda anjos nos momentos mais difíceis da vida, obrigada por ser o meu. Obrigada por dividir comigo sua família e me permitir fazer parte dela, agradeço por todo cuidado, carinho, atenção e preocupação que todos eles tiveram comigo, são uma família para mim. Amo vocês!

Ao meu orientador Paulo Marcos Pinto, primeiramente pela oportunidade e por ter me recebido em seu laboratório nesses dois anos, pelos ensinamentos, pela confiança e por fazer parte desta minha fase de amadurecimento pessoal e profissional. Obrigada!

Aos professores, Gabriel da Luz Wallau, Juliano Tomazzoni Boldo, pelo apoio direto a este trabalho, por todas as horas prestadas a me ajudar, por estarem dispostos a tirar dúvidas e a me incentivar a não desistir.

Aos professores Andrés Delgado Cañedo e Analía Del Valle Garneró e aos

seus respectivos laboratórios pelas inúmeras ajudas prestadas, reagente emprestados, conselhos e por todas as dúvidas sanadas, muito obrigada por sempre estarem dispostos a me ajudar e por serem mais que mestres, amigos que eu fiz nesta etapa. Tenho admiração profunda por vocês.

Aos amigos que fiz nesses 6 anos nesta nova cidade, Gabriela, Thiago, Jaciara, Cleusa, Bernardo, Rodrigo Kapucha, Fabiane, obrigada por me alegrarem, por serem mais que amigos, por serem irmãos. Quero levá-los por toda a vida, e serei eternamente grata pelo apoio nessa caminhada.

A minhas amigas de profissão, Thays e Natasha, por serem companheiras, dividirem comigo suas preocupações e alegrias, começamos juntas e terminaremos esta etapa juntas. Obrigada por tudo, vocês sempre terão um espaço especial no meu coração.

Aos colegas de laboratório, por todos os dias de laboratório, congressos e seminários, obrigada pelos ensinamentos e por me acolherem na vida de vocês. Em especial a Thalita, Laís, Darlene e Filipe pela ajuda direta neste trabalho, por dividirem seus conhecimentos e pela ajuda no laboratório. Obrigada, aprendi muito com vocês e agradeço por tudo.

Enfim a todos que contribuíram para este caminho ser trilhado, dentro e fora de universidade. Muito obrigada!

EPÍGRAFE

"Não vos preocupeis, pois, com o dia de amanhã: o dia de amanhã terá as suas preocupações próprias. A cada dia basta o seu cuidado."

Mateus 6: 34

RESUMO

A importância da expressão de proteínas de forma heteróloga tem crescido surpreendentemente, tendo em vista a progressiva necessidade de sua produção para aplicações na indústria farmacêutica, para uso terapêutico, ou uso comerciais, e estudo de estrutura e função. No entanto, a expressão das proteínas pode ser limitada por diversos fatores, podendo apresentar diversas dificuldades como a formação de corpos de inclusão ou proteínas com conformação incorreta, por exemplo. Estudos em hospedeiro para expressão de proteínas heterólogas tem sido desenvolvido, a fim de otimizar fatores como temperatura e concentração de indutor, com o intuito de melhorar a qualidade das proteínas obtidas. Neste estudo, inicialmente objetivou-se otimizar a expressão da transposase do elemento *mariner mos1*, em diferentes cepas de *Escherichia coli* BL21(DE3), *E. coli* BL21 (DE3) pLysS, *E. coli* BL21-CodonPlus (DE3) -RP; *E. coli* BL21-CodonPlus (DE3) -RIL; *E. coli* BL21(DE3) pT-GROE, usando para subclonagem o vetor plasmidial de expressão pGEX-4T1. Analisando o melhor sistema de expressão para este elemento em particular, comparando as temperaturas de 30°C e 37°C para expressão e utilizando concentrações de 0,1mM; 0,5mM e 1,0mM de IPTG. Mesmo obtendo sucesso na expressão, com todas as condições de IPTG e em ambas as temperaturas testadas, o sistema de expressão não foi eficaz para obter proteínas de forma solúvel. Frente a este problema procuramos desenvolver um vetor de expressão de proteínas heterólogas que afim de tornar este processo mais rápido e mais eficiente, utilizando apenas um evento de clonagem e de transformação da bactéria de expressão. O vetor para expressão foi desenhado *in silico* a partir de um promotor Lac responsivo a IPTG, seguido por um sinal de exportação celular (peptídeo-sinal), um fragmento *Strep-tagII* para purificação e um sítio para trombina, permitindo a clivagem da proteína de interesse dos *tags* fusionados. Foram inseridos dois sítios para endonucleases de restrição (BamHI e XbaI) que permitem a clivagem do gene clonado no constructo. Desta forma demonstramos que todas as cepas foram eficazes na expressão da transposase *mos1*, sendo a temperatura mais adequada de 37°C em todas as concentrações de IPTG, já que não houve nenhuma diferença quanto a mudança deste fator. Apresentamos um novo vetor de expressão de proteínas heterólogas, com a finalidade

de facilitar a obtenção de proteínas heterólogas purificadas, tais problemas encontrados durante a tentativa de expressão da transposase.

ABSTRACT

The importance of heterologous protein expression has surprisingly grown in view of the progressive need for its production for applications in the pharmaceutical industry, for therapeutic use, or commercial use, and study of structure and function. However, the expression of the proteins may be limited by several factors, and may present several difficulties such as the formation of inclusion bodies or proteins with incorrect conformation, for example. Host studies for expression of heterologous proteins have been developed in order to optimize factors such as temperature and concentration of inducer in order to improve the quality of proteins obtained. In this study, we initially aimed to optimize transposase expression of the mariner element *mos1* in different strains of *Escherichia coli* BL21 (DE3), *E. coli* BL21 (DE3) pLysS, *E. coli* BL21-CodonPlus (DE3) -RP; *E. coli* BL21-CodonPlus (DE3) -RIL; *E. coli* BL21 (DE3) pT-GROE, using for subcloning the pGEX-4T1 expression plasmid vector. Analyzing the best expression system for this particular element, comparing the temperatures of 30°C and 37°C for expression and using concentrations of 0.1 mM; 0.5mM and 1.0mM IPTG. Even with expression success, with all IPTG conditions and at both temperatures tested, the expression system was not effective in obtaining proteins in a soluble form. In view of this problem we have tried to develop an expression vector of heterologous proteins in order to make this process faster and more efficient, using only one event of cloning and transformation of the expression bacterium. The vector for expression was drawn in silico from an IPTG responsive Lac promoter, followed by a cell export signal (peptide-signal), a Strep-tagII fragment for purification and a site for thrombin, allowing cleavage of the protein from Interest of merged tags. Two sites for restriction endonucleases (BamHI and XbaI) were inserted which allow the cleavage of the cloned gene in the construct. In this way, we demonstrated that all strains were effective in the expression of the transposase *mos1*, with the most suitable temperature being 37 ° C in all concentrations of IPTG, since there was no difference in the change of this factor. We present a new expression vector of heterologous proteins, in order to facilitate the obtaining of purified heterologous proteins, such problems encountered during the attempt to express the transposase.

LISTA DE FIGURAS

- Figura 1 - *Escherichia coli*.** Microscopia eletrônica de varredura da bactéria em meio de cultura.....21
- Figura 2 - Enzima transposase do elemento transponível *mos1***.....27
- Figura 3 - Mapa do vetor de expressão pGEX-4T-1, utilizado para expressar a transposase *mos1*.**.....32
- Figura 4 - Gel de poliacrilamida SDS PAGE 12% da indução das cepas de *Escherichia coli*.** Em (A) *E. coli* BL21 (DE3) pT-GROE, em (B) *E.coli* BL21 (DE3) PlysS, e em (D) *E. coli* BL21 (DE3) RIL onde as amostras foram induzidas com 1mM, 0,5mM e 0,1mM de IPTG respectivamente a 30°C e 37°C. Em (C) *E. coli* BL21 (DE3) RP, induzida a 37°C e 30°C.37
- Figura 5 - Gel de poliacrilamida SDS PAGE 12% da cepa *E. coli* BL21 (DE3) pT-GROE.** PM – Peso molecular. Na figura (A) as amostras foram induzidas a temperatura de 30°C e em (B) 37°C, intercalando sobrenadante e pellet induzidos por IPTG 1,0mM, 0,5mM e 0,1mM.38
- Figura 6- Gel de poliacrilamida SDS PAGE 12%; cepa *E. coli* BL21-CodonPlus (DE3) -RP.**PM: peso molecular. Na figura (a) as amostras foram induzidas a temperatura de 30°C e em (b) 37°C, intercalando sobrenadante e pellet induzidos por IPTG 1,0mM, 0,5mM e 0,1mM.39
- Figura 7 - Gel de poliacrilamida SDS PAGE 12% *E. coli* BL21-CodonPlus (DE3) – RIL.** PM – Peso molecular. Na figura (a) as amostras foram induzidas a temperatura de 30°C e em (b) 37°C, intercalando sobrenadante e pellet induzidos por IPTG 1,0mM, 0,5mM e 0,1mM.39
- Figura 8 - Gel de poliacrilamida SDS PAGE 12%; cepa *E. coli* BL21(DE3) PlysS.** PM - Peso molecular. Na figura (A) as amostras foram induzidas a temperatura de 30°C e em (B) 37°C, intercalando sobrenadante e pellet induzidos por IPTG 1,0mM, 0,5mM e 0,140

LISTA DE FIGURAS – MANUSCRITO

Figure 1 - Schematic map of the construct, and corresponding sequence. The figure depicts the major features present in construct. In (A) you demonstrate the sequences of the construct with their identification below the underline, and in (B) a schematic with the fragments of the construct represented by colors and the respective value of base pairs in the upper part of the figure 52

Figure 2 - Schematic representation of the expression vector. Schematic representation of the expression vector with the restriction sites for the endonucleases (XbaI, EcoRI, BamHI and HindIII) selection gene, origin of replication derived from plasmid pMB1 (rep pMB1), sites of thrombin and Strep-tagII 53

Figure 3 - Electrophoresis 1% (w / v) agarose of construct. Channel 1 corresponds to the molecular weight marker described in detail on the right side of the image; the channels 1 through 2 correspond to the amplification of the expression construct with the corresponding primers, totaling 852 bp 54

Figure 4 - Electrophoresis 1% (w / v) agarose of pUC19 linearized. The channeled 1 of the gel corresponds to the molecular weight marker described in detail on the right side of the image; the next channel corresponds to the extracted pUC19 vector cleaved with the BamHI restriction endonuclease, for linearization. 55

Figure 5 - Expression vector – Expression vector (pUC19 + insert) clivated with the BamHI endonuclease and precipitated with PEG13%. The vector used to receive the insert was pUC19 with 2686 base pairs, which will total after the addition of the insert 3538pb 56

LISTA DE TABELAS

Tabela 1 - Propriedades da Ordem TIR - Propriedades estruturais e moleculares das nove superfamílias de transposon de DNA pertencentes à classe II, subclasse I.....28

LISTA DE QUADROS

Quadro 1 – Cepas utilizadas e seus respectivos genótipos.....33

Quadro 2 - Análise da expressão das cepas de *E.coli*.....41

SUMÁRIO

1	INTRODUÇÃO GERAL	18
1.1	Técnica do DNA recombinante	18
1.2	Clonagem gênica.	18
1.3	Expressão heteróloga de proteínas	19
1.4	<i>Escherichia coli</i>	21
1.5	Vetores gênicos plasmidiais	23
1.6	Elementos de transposição	24
1.7	Superfamília Tc1 / <i>mariner</i> – Mos 1	26
1.8	Gênero <i>Drosophila</i>	27
2	JUSTIFICATIVA	29
	CAPÍTULO I	30
3	OBJETIVO	31
3.1	Objetivo geral	31
3.2	Objetivos específicos	31
4	METODOLOGIA	32
4.1	Clonagem do elemento <i>mos1</i>	32
4.2	Transformação bacteriana	33
4.3	Indução por IPTG	33
4.4	Análise de expressão	34
4.5	Análise de solubilização.....	34
5	RESULTADOS E DISCUSSÃO	35
5.1	Expressão.....	35
5.2	Solubilidade.....	38
	CAPÍTULO II	42
	Manuscrito.....	42
	O artigo intitulado: “Design and construction of <i>Escherichia coli</i> expression vector: addressing to for soluble proteins” está disposto na forma na qual será submetido para o periódico Protein Expression and Purification (ISSN: 1046-5928).....	42
	INTRODUCTION	44
	MATERIALS AND METHODS	45
	<i>Designing the expression construct</i>	45
	<i>Preparation of the plasmid</i>	47
	<i>Expression vector synthesis</i>	48

<i>Transformation of Escherichia coli</i>	48
RESULTS AND DISCUSSION	49
<i>Designing the expression construct</i>	49
<i>Synthesis of the expression vector</i>	49
<i>Cloning</i>	50
FINAL REMARKS	50
6 CONCLUSÃO GERAL	57

1 INTRODUÇÃO GERAL

1.1 Técnica do DNA recombinante

Em 1953, os pesquisadores Watson e Crick determinaram a estrutura molecular do ácido desoxirribonucleico (DNA), molécula responsável pelo armazenamento da informação genética, que consiste em duas cadeias longas nucleotídicas compostas por 4 subunidades nucleotídicas que se enrolam ao redor de um eixo comum formando uma dupla hélice (Alberts et al. 2010). Tais estudos, aliados a outras descobertas “construíram” o Dogma Central da Biologia Molecular, postulado por Francis Crick em 1958, o qual sumariza que a informação genética contida no DNA das células e dos vírus é preservada, transcrita em RNA e traduzida em proteínas. (Reys et al. 2011).

Até a década de 70, o DNA era o componente celular mais difícil de ser analisado, geralmente sendo estudado de forma geral e indireta, somente com o desenvolvimento de novas tecnologias foi possível que a molécula de DNA fosse estudada com mais propriedade, permitindo o isolamento e a purificação de regiões específicas, e a obtenção desta molécula em grande quantidade, além de determinar sua sequência (Nascimento et al. 2003). A tecnologia do DNA recombinante, começou a ser desenvolvida na década de 70, quando os bioquímicos Berg, Boyer e Cohen demonstraram que seria possível a transferência de genes de um organismo para outro. Desde então é conhecida como um conjunto de técnicas para localizar, isolar, alterar e estudar segmentos de DNA. O termo recombinante conferido a técnica vem do uso frequente destas técnicas para combinar o material genético de fontes distintas (Alberts et al. 2010).

A tecnologia do DNA recombinante tem diversas finalidades, dentre elas o estudo de mecanismos de replicação e expressão de genes, é usada também para determinar a sequência dos mesmos e, conseqüentemente, auxiliar estudos avançados que visam analisar a estrutura e função das proteínas, além da produção de proteínas recombinantes em sistemas utilizando um procaríoto ou um eucarioto como hospedeiro (Nascimento et al. 2003). Portanto, sua aplicação biotecnológica e industrial tem um amplo potencial.

1.2 Clonagem gênica

A técnica central da metodologia do DNA recombinante é a clonagem

molecular, a qual consiste na propagação e isolamento de moléculas de DNA idênticas. As novas tecnologias desenvolvidas na Microbiologia, Bioquímica, Imunologia e Genética Microbiana permitiram a clonagem gênica (Zaha et. al,2003).

A clonagem molecular, consiste na replicação de moléculas de DNA idênticas e baseia-se na propagação natural de células ou indivíduos geneticamente idênticos ao inicial (Brown, 2003). O experimento de clonagem gênica baseia-se em ao menos dois estágios importantes: inicialmente o fragmento de DNA de interesse é ligado a uma molécula carreadora, para formar o DNA recombinante (Brown, 2003). Na etapa seguinte esta molécula é introduzida dentro de células compatíveis, denominadas hospedeiras, num processo conhecido como transformação. Se as manipulações forem executadas corretamente, o gene será expresso e a proteína será sintetizada pela célula bacteriana (Brown, 2003).

A célula que recebeu o DNA modificado (recombinante) é chamada de célula transformada, passando por muitos ciclos de divisão, desta forma a célula carreadora também se multiplicará, gerando assim diversas cópias do DNA de interesse (Brown, 2003). As células de cada colônia são idênticas entre si, daí então a origem da palavra clonagem, que surgiu da genética bacteriana que considera a colônia de bactéria como um clone (Nascimento et al. 2003).

1.3 Expressão heteróloga de proteínas

A expressão de proteínas heterólogas é uma importante ferramenta biotecnológica, tendo em vista a crescente demanda por produzir proteínas em quantidade suficientes para estudo de sua função e estrutura, bem como aplicações na indústria farmacêutica. (Schmidt, 2004).

A escolha do hospedeiro que cederá o maquinário para a síntese da proteína, é crucial, e este define a tecnologia necessária para o projeto, a variedade das ferramentas moleculares, equipamento, reagentes, entre outros (Demain & Vaishnav 2009). Vários hospedeiros são possíveis para a expressão de proteínas heterólogas, tais como fungos, algas unicelulares, bactérias entre outros. Todos possuem vantagens e desvantagens quanto ao uso, e sua escolha está atrelada ao tipo de proteína de interesse (Adrio & Demain 2010).

Dentre os hospedeiros mais utilizados se encontra a bactéria bacilar *Escherichia coli* devido ao vasto conhecimento sobre suas características e as vantagens que apresentam em relação aos demais para este tipo de processo (Terpe,

2006). A *E. coli* é mais vantajosa em comparação com outros organismos devido ao seu ciclo de crescimento extremamente rápido. E em condições ambientais ideais seu tempo de duplicação é de aproximadamente 20 minutos (Sezonov et al. 2007), o meio de cultivo utilizado para crescer esta bactéria tem um baixo custo. Transformar com DNA exógeno usando um método de transformação simples, além de permitir a produção de proteínas heterólogas em grande quantidade, considerando que podem ser obtidas altas densidades celulares (Demain & Vaishnav 2009).

Os passos principais para a produção de uma proteína recombinante são simples no âmbito teórico, uma vez que compreendem a escolha do gene de interesse, a escolha do sistema de expressão, clonagem e expressão e, em seguida, a proteína está pronta para purificação e caracterização (Miyamoto 2013). No entanto a expressão heteróloga de proteínas pode ser limitada por inúmeros fatores (Miroux & Walker 1996; apud Dumon-seignovert et al. 2004) já que a expressão da proteína no hospedeiro acaba por ser considerada um processo estranho a célula, ou seja, algo não natural, metabolicamente dizendo, causando em muitas das vezes problemas a célula hospedeira pois reage dificultando a obtenção da proteína de interesse (Vaz 2008).

Desta forma, sabe-se que para produzir uma proteína recombinante, é necessário analisar diversos aspectos, como escolher cuidadosamente o vetor, o hospedeiro e as condições do crescimento como fatores essenciais. A escolha do hospedeiro deve levar em conta sua facilidade de manutenção, além das propriedades da proteína, do seu destino final, sua solubilidade e estabilidade, assim como rendimento da produção (Jonasson et al. 2002; Williamson et al. 2000). Outro fator a ser analisado é a possibilidade da formação de corpos de inclusão insolúveis, causados pelo excesso de proteínas produzidas (Kim et al. 2016), e reverter esta formação acaba tornando o processo mais complexo e dispendioso (Jacobson et al. 2014; apud Gabrielczyk et al. 2017).

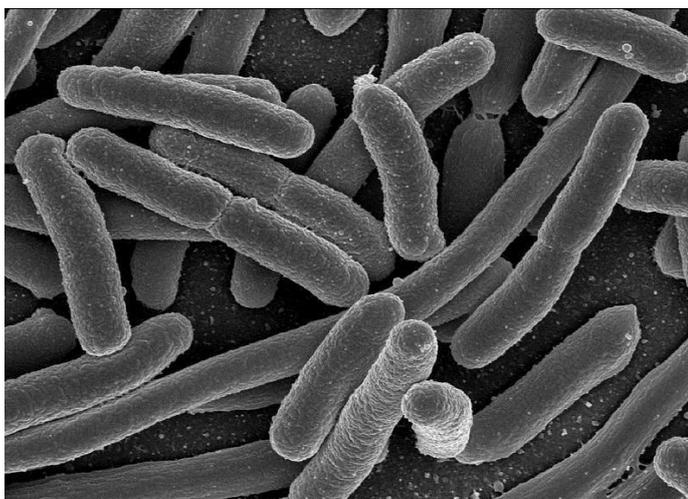
Portanto, é crucial a escolha do sistema de expressão de proteínas, pois esse processo envolve diversos fatores importantes devido ao dobramento e conformação correta da proteína, além do local correto para sua expressão (Mergulhão et al. 2005). A proteína pode ser expressa em três localizações: no citoplasma, no periplasma e extracelularmente. Em um dos modelos mais conhecidos, a *E. coli*, a proteína é frequentemente armazenada em grandes quantidades no citoplasma bacteriano, onde a proteína é expressa porém de forma agregada, porém existem na literatura casos

de outros locais de endereçamento, como ao espaço periplasmático (entre a membrana celular e parede bacteriana) um ambiente oxidativo que permite a conformação natural da proteína (Makrides 1996), e mais raramente ainda, as moléculas são secretadas para o meio extracelular, onde a proteína é enviada ao meio de cultura (Sorensen & Mortensen 2005).

1.4 *Escherichia coli*

A bactéria *E. coli* é um organismo bem estudado, e sabe-se muito sobre seu metabolismo, regulação e fisiologia. Descrita pela primeira vez em 1886 por Theodor Escherich, um pediatra alemão, que a identificou em fezes de animais saudáveis e sua primeira denominação foi "commune bactéria coli", devido ao local em que era encontrada. Esta bactéria, representada na figura 1, tem forma de bacilo, é filamentosa, gram-negativa e não esporulada, pertencente à família Enterobacteriaceae, anaeróbia facultativa e fermentadora de açúcares (Reed 2005).

Figura 1 - *Escherichia coli*. Microscopia eletrônica de varredura da bactéria em meio de cultura.



Fonte: Rocky Mountain Laboratories

Encontrada normalmente na microbiota entérica (intestino) de aves e mamíferos, incluindo os seres humanos. Tem crescimento ótimo a 37°C, e pelo fato de não esporular tem sua esterilização de forma simples;(Reed 2005).

O genoma da *E. coli* é composto por um cromossomo circular único composto por 4.639.221 pares de base de DNA agrupados em 4.288 genes. Destes, 70% correspondem a genes monocistrônicos, e 6% policistrônicos e apenas cerca de 30% das ORFs (*Open Reading Frame*) sequenciadas têm funções desconhecidas (Madigan et al. 2016). Mesmo com esta estrutura simples, desempenha um

metabolismo complexo para manter seu crescimento e divisão celular. (Madigan et al. 2016).

Até meados dos anos 70, a *E. coli* não tinha sido descoberta com a importância industrial, o potencial desta bactéria foi demonstrado após o advento da tecnologia do DNA recombinante, que desenvolveu seus estudos em grande parte em torno deste modelo (Lima, 2004). Além disto até hoje são os organismos estudados com mais frequência e os mais utilizados como hospedeiros devido a riqueza de informações sobre seu metabolismo.

A *E. coli* tem vantagens frente aos demais sistemas de expressão, pois possui um ciclo de crescimento rápido em condições ambientais ideais, seu tempo de duplicação é de aproximadamente 20 minutos (Sezonov et al. 2007), permite o uso de substratos de cultivo de baixo custo, se comparado aos demais desta categoria, possui a capacidade de transformação com DNA exógeno usando um método de transformação simples, além de possibilitar a produção de proteínas heterólogas em grande quantidade, tendo em vista que altas densidades celulares podem ser alcançadas (Jong Hyun Choi et al. 2006).

As cepas desenvolvidas para utilização em laboratório são, em sua grande maioria, derivadas de uma única estirpe a *E. coli* K-12, isolada em 1922 (Tatum & Lederberg 1947) Uma das cepas mais utilizadas é a DH5 α , desenvolvida por D. Hanaham como uma estirpe de clonagem com múltiplas mutações que permitem transformações de alta eficiência (Durfee et al. 2008). O genótipo da cepa DH5 α é descrito da seguinte forma: F⁻ Φ 80lacZ Δ M15 Δ (lacZYA-argF) U169 recA1 endA1 hsdR17 (rK⁻, mK⁺) phoA supE44 λ ⁻ thi-1 gyrA96 relA1. Estas características correspondem às características distintas que tornam a cepa DH5 α excelente em gerar procedimentos de clonagem. A mutação endA1 possibilita uma menor degradação do DNA recombinante pela endonuclease que garante taxas de transferência de plasmídeo mais elevadas, e pôr fim a mutação recA1 acaba por reduzir a recombinação homóloga estabilizando a molécula recombinante.

As cepas de expressão BL21 e BL21 (DE3) foram desenvolvidas por William Studier e Barbara Moffatt, que têm como característica principal a deleção da protease OmpT e Lon, características comuns da linhagem B da qual ela pertence (Jeong et al. 2015), com genótipo *fhuA2 [lon] ompT gal [dcm] Δ hsdS*. A cepa de *E. coli* BL21 (DE3) é derivada de BL21, abriga um pró-fago DE3 do bacteriófago λ , que transporta o gene da RNA polimerase do bacteriófago sob o controle do promotor *lacUV5* (Jeong et al.

2015), Este promotor pode ser induzido por adição de *isopropil-β-D-1-thiogalactopyranoside* (IPTG) um análogo de lactose não hidrossolúvel, este mecanismo fornece o controle da expressão de sequencias sob seu controle (Long et al. 2015). Este sistema visa por fim reduzir problemas na regulação do promotor com repressão catabólica (Terpe 2006). Seu genótipo é F - ompT HSD_{SB} (r_B - m_B) gal DCM (DE3).

A linhagem BL21 (DE3) pLysS é resistente a cloranfenicol e possuem o gene para a lisozima T7. Já as cepas BL21-CodonPlus-RP (DE3) e BL21-CodonPlus-RIL (DE3), como o nome sugere, são também derivados de BL21 (DE3). Estas estirpes fornecem cópias adicionais de genes raros de RNAt, para superar a diferença no uso de códons entre diferentes organismos. As estirpes RIL carregam genes para os RNAt de Arg, Ile e Leu e RP para os de Arg e Pro, estas também resistentes a cloranfenicol. Por fim a cepa de *E. coli* BL21 (DE3) pT-GROE, que possui o plasmídeo pT-GroE que possui tem um seletor de resistência a cloranfenicol e expressa chaperonas GroES e GroEL sob o controle do promotor T7. Estas chaperonas estão diretamente relacionadas com o correto dobramento das proteínas e sua solubilidade (Herrmann et al. 2004) Carvalho & Meneghini 2008)

1.5 Vetores gênicos plasmidiais

Os vetores gênicos de expressão, são vetores de clonagem que possuem todos os elementos genéticos para permitir a expressão de proteínas recombinantes, os mais utilizados são os plasmídeos, que são pequenas moléculas circulares de DNA de fita dupla, derivadas de plasmídeos maiores, que ocorrem naturalmente em células bacterianas (Alberts et al. 2010).

Os plasmídeos de expressão bacteriano são diversos e com o propósito de obter a proteína heteróloga na sua conformação correta é necessário respeitar o sinal de tradução dos genes procarióticos (sinal Shine-Delgarno), além de respeitar a fase de leitura correta (ficar em fase com o ATG inicial) (Jana & Deb 2005). Os plasmídeos de expressão possuem alguns elementos essenciais, os mais utilizados resultam de múltiplas combinações de origens de replicação, marcadores de seleção, e um gene que codifica um produto que distingue a célula transformada da célula não transformada, promotores, sítios de clonagem múltipla, onde o inserto é incorporado ao vetor de clonagem e sítios de fusão de proteínas (Davison 2002).

A origem de replicação controla o número de copias desejado, no entanto, é

importante levar em conta que um elevado número de cópias pode significar maior quantidade de proteínas, porém estas cópias podem acabar por sobrecarregar o metabolismo bacteriano e diminuir o seu crescimento e produzindo plasmídeos instáveis. (Rosano & Ceccarelli 2014). O promotor é a região onde se inicia a transcrição, portanto é onde se controla a expressão do gene clonado. Os promotores que são usados em vetores de expressão são em grande parte induzíveis, ou seja, a síntese proteica se inicia apenas quando necessário, a partir da indução de um indutor. (Madigan et al. 2016).

Os principais plasmídeos de expressão utilizam os promotores Lac, Tac e T7. O promotor Lac é semelhante ao utilizado pelas cepas selvagem de *E. coli* e é reprimido pela proteína lacI na ausência de indutor e induzido em sua presença. O promotor Tac contém a região operadora do Operon Trp e a região promotora do Operon Lac (Jonasson et al. 2002). Assim, ambos promotores são induzidos por lactose ou IPTG e utilizam a RNA polimerase bacteriana para transcrever o gene de interesse. O promotor T7, utiliza a RNA polimerase do fago T7, que é mais ativa do que a RNA polimerase bacteriana, que resulta em níveis de expressão proteica bem maior (Stano & Patel 2004).

Para seleção das cópias recombinantes um marcador de seleção é adicionado ao plasmídeo. Na *E. coli*, os genes de resistência a antibiótico são geralmente utilizados para esse propósito, que é habitualmente representada pela ampicilina o qual tem a resistência conferida pelo gene *bla*, do qual o produto é uma enzima que inativa o anel β -lactâmico dos antibióticos β -lactâmicos e permite que as colônias transformadas cresçam em meios de cultivo com a presença de antibióticos. (Korpimäki et al. 2003)

1.6 Elementos de transposição

Transposons são elementos genéticos móveis, também conhecidos como “jumping genes” amplamente distribuídos no genoma de plantas, bactérias, invertebrados e vertebrados, capazes de se mover dentro do genoma (Kidwell & Lisch 2001). Tais elementos constituem uma parte significativa do genoma de eucariotos, sendo maior que 70% em algumas plantas e anfíbios, e 45% em humanos (Venner et al. 2010). A grande maioria dos genomas investigados até hoje apresentaram elementos transponíveis (Biémont & Vieira 2006).

Estes genes saltadores foram inicialmente identificados no genoma do milho, onde aparentemente, controlavam a coloração das sementes, pela geneticista Barbara McClintock, na década de 40 em Nova York. Inicialmente a comunidade científica foi cética com esta descoberta, no entanto, ao longo das próximas décadas, se evidenciou a existência dos transposons, contudo somente na década de 80 os elementos controladores, foram chamados definitivamente de elementos transponíveis (Kazazian 1998).

Os transposons desempenham um papel importante na evolução dos genes da maioria das espécies (Bire et al. 2013). Devido sua grande diversidade, estudos foram realizados, a fim de compreender qual o real papel, ou seja, a real função destes elementos na evolução do genoma. Por possuírem a capacidade de se inserirem ou deletarem em locais variados do genoma, acabam por causar algumas alterações, contribuindo para modificações genéticas, e em muitas vezes interferindo na regulação da expressão genica (Böhne et al. 2008). Outra importante característica destes elementos é a capacidade de se moverem tanto por transferência vertical, passando de um ancestral a um descendente ou por transferência horizontal (HTs), entre as espécies (Filée et al. 2015).

Devido à grande diversidade de elementos transponíveis, e as proteínas codificadas por eles, a partir da década de 80 percebeu-se a necessidade de desenvolver um sistema de classificação para os mesmos. A primeira classificação foi proposta por Finnegan em 1989, essa classificação discriminava os transposons de acordo com seu mecanismo de transposição, dividindo-os em classe I (retrotransposons) e classe II (transposons). Atualmente, com o sequenciamento de novos organismos, novas características foram adicionadas a classificação de Finnegan, por Wicker e colaboradores em 2007. Com estas os TEs podem ser agrupados de forma hierárquica, em classes, subclasses, ordens, superfamílias, famílias e subfamílias (Wallau 2010).

Os elementos transponíveis de classe I, são conhecidos como retrotransposons, por possuírem a estrutura semelhante aos retrovírus, estes TEs tem o mecanismo de mobilidade baseado em um intermediário de RNA, conhecido como “copia-cola”, estes não codificam a enzima transposase, dependendo assim da transcriptase reversa e a integrase para converter a molécula de RNA em DNA. O RNAm é inicialmente transcrito no hospedeiro e então revertido para DNA pela transcriptase reversa, e finalmente a integrase insere o fragmento de DNA em outra

região do genoma, diferente daquela de origem, gerando uma cópia do fragmento. Esta classe possui uma subdivisão, que subdivide pela presença de repetições terminais longas denominadas LTRs, ou a ausência dessas repetições terminais os não-LTRs.

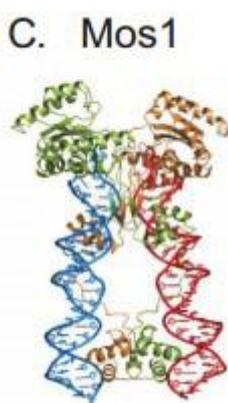
Os elementos de classe II, possuem a transposição baseada em um intermediário de DNA, Transposons de DNA são segmentos nucleotídicas que apresentam repetições terminais invertidas nas suas extremidades (TIRs) (Bire et al. 2013), onde a transposase retira o elemento de uma região do genoma e insere em outra, em um movimento de “corta e cola”. Apesar das diferenças nos mecanismos de transposição, a integrase de alguns elementos de RNA e a transposase de alguns elementos de DNA provavelmente têm uma origem comum (Capy et al. 1992). Os elementos desta classe se subdividem em duas subclasses baseadas na geração de cortes de DNA de cadeia simples ou dupla durante o processo de transposição. A primeira subclasse, engloba os TEs que realizam o processo “corta e cola” que são flanqueadas por repetições terminais invertidos (TIRs), composta por nove superfamílias, que se distinguem pelo motivo de sequência das suas repetições terminais e o comprimento das duplicações de locais de destino (TSDs). A segunda subclasse requer em seu processo de transposição replicação e não a quebra de DNA de cadeia dupla (Kapitonov & Jurka 2001; Pritham et al. 2007).

1.7 Superfamília Tc1 / *mariner*– Mos 1

A superfamília TC1 pertence aos elementos de classe II, subclasse I, demonstrada na tabela 1, que compreende os elementos que realizam o processo de transposição “corta e cola” (Dawson & Finnegan 2003), da ordem TIR, caracterizada pelas repetições terminais invertidas (terminal inverted repeats - TIRs), esta superfamília se destaca por ser uma das mais difundidas na natureza (Yuan & Wessler 2011) presentes em alguns fungos, peixes, plantas e mamíferos (Arkhipova & Meselson 2005) este elementos possuem entre 1300 e 2400 pb de comprimento codificam uma transposase com cerca de 340 aminoácidos (tabela 1), que é flanqueada por repetições invertidas simples (TIRs) com cerca de 30 pb (Tellier et al. 2014).

Nesta família encontra-se o elemento *mos1*, um elemento autônomo isolado em *Drosophila mauritiana* (Crénés et al. 2010). Possui 1.286 pb de extensão, com repetições terminais invertidas de 28 pb (TIRs) e contém um único quadro de leitura aberto, que codifica a transposase *mos1* de 345 aminoácidos (Tnp) (Figura 2) responsável por promover todos os passos de transposição (Auge et al. 2005). Este elemento se mostrou ser ativo em organismos como zebra-fish e em *Aedes aegypti*,

Figura 2 - Enzima transposase do elemento transponível *mos1*



Fonte: Hickman Dyda, 2015

mas é ativa in vitro, confirmando que não depende de fatores do hospedeiro (Nowacki et al. 2009) é capaz de causar modificações genéticas em células eucarióticas e em outros organismos (Mamoun et al. 2000). A transposase é o único fator responsável para que se realize a transposição no elemento *mos1* (Tosi & Beverley 2000). Em geral a maioria dos transposons da superfamília mariner são independentes na sua transposição, ou seja, não depende de fatores do hospedeiro. Como todos os passos da transposição são catalisados pela transposase, o elemento *mos1* não apresenta restrições em se transpor em eucariotos ou procariotos (Picardeau 2010).

1.7.1 Gênero *Drosophila*

O gênero *Drosophila* é constituído por aproximadamente 1149 espécies, subdivididas em subgêneros (BÄCHLI, 2008) pertencente à família *Drosophilidae*, e da superfamília *Ephydroidea* (THOMPSON, 2005), uma das mais diversas e amplamente distribuídas entre as mais de 150 famílias da ordem Diptera (YEATES E WIEGMANN, 2005). Já o subgênero *Drosophila* é composto atualmente por 721 espécies subdivididas em 43 grupos. Esse subgênero constitui o maior representante da família *Drosophilidae* na região neotropical, onde é representado por 25 grupos de espécies. Muitas espécies do gênero *Drosophila* vêm sendo utilizadas, há mais de 100 anos, como um excelente

modelo para estudos genéticos e evolutivos, devido sua disponibilidade de genoma de espécies deste gênero (CLARK *et al.*, 2007), vários genomas populacionais (<http://www.dpgp.org/>), e um número razoável de espécies com relação filogenética esclarecida estas informações tornaram estes organismos ainda mais atrativos para responder às questões sobre evolução de TEs e possibilita um amplo campo de estudos visando esclarecer os mecanismos gerais de evolução genômica.

Tabela 1 - Propriedades da Ordem TIR - Propriedades estruturais e moleculares das nove superfamílias de transposon de DNA pertencentes à classe II, subclasse I

Subclasse I / Ordem TIR								
Superfamília	IS	Ocorrência	Tamanho (kb)	TIRs (bp)	TSDs (bp)	Proteínas codificadas	DBD	Núcleo Catalítico
<i>Tc1/mariner</i>	IS630	Eucariontes	1.2-5.0	17-1100	2 (TA)	Tnp	HTH	(D,D,E/D)Tnp
hAT	Nd	Eucariontes	2.5-5.0	5-27	8	Tnp	BEDZnF	(D,D,E) Tnp
<i>Mutator</i>	IS1016	Eucariontes	1.3-7.4	0 – muitos kb	9-11	Tnp	WRKY/GCM1 ZnF	(D,D,E) Tnp
<i>Merlin</i>	Nd	Animais e Eubactérias	1.4-3.5	21-462	8-9	Tnp	Nd	(D,D,E) Tnp
<i>Transib</i>	Nd	Metazoos e Fungi	3-4	9-60	5	Tnp	Nd	(D,D,E) Tnp
P	nd	Plantas e Metazoos	3-11	12-150	8	Tnp	THAPZnF	Nd
<i>piggyBac</i>	IS1380	Eucariontes	2.3-6.3	12-19	4 (TTAA)	Tnp	nd	Nd
<i>PIF/Harbinger</i>	IS5	Eucariontes	2.3-5.	15-270	3 (CWG ou TWA)	Tnp+Myb-likeprotein	Myb/SANT	(D,D,E) Tnp
CACTA	nd	Plantas, Metazoos e Fungi	4.5-1.5	10-54	2-3	TnpA+TnpD	nd	nd

Fonte: Adaptado de Sinzelle et al. 2009

2 JUSTIFICATIVA

Os elementos transponíveis estão presentes em grande maioria dos genomas investigados até hoje (Biémont & Vieira 2006). Estes elementos podem influenciar no processo evolutivo do seu hospedeiro (Feschotte & Pritham 2007), e apesar de serem bem conhecidos, ainda há uma busca por informações sobre suas estruturas moleculares e seus mecanismos de transposição. Para uma melhor compreensão destes aspectos uma análise completa da enzima ativa desses elementos (transposase), por exemplo, é fundamental. A transposase é uma enzima responsável por catalisar o movimento do transposon para outra parte do genoma por um mecanismo de corte e cola ou um mecanismo de transposição replicativo, sendo assim um dos elementos mais importantes para a elucidação dos processos de transposição destes elementos. Assim fundamenta-se a necessidade de expressão da transposase do elemento *mos1* pertencente a uma das famílias de transposons mais difundidos e capaz de causar modificações em células eucarióticas e em outros organismos (Mamoun et al. 2000). No entanto todo novo projeto que necessite uma proteína purificada, se depara com a necessidade de como obtê-la. O processo de obtenção destas proteínas heterólogas podem ser dificultadas por inúmeros fatores, tendo em vista que além do processo necessitar diversas clonagens subsequentes, que acabam por aumentar o custo e tempo dos experimentos, a proteína obtida pode não ser solúvel ou ativa, podem haver a formação de corpos de inclusão e até mesmo a não obtenção da proteína de interesse. Desta forma, justifica-se desenvolver um vetor como um novo método para produzir uma proteína solúvel, evitando a formação de agregados insolúveis, tais como corpos de inclusão em bactérias, e em quantidade suficiente que permita seu estudo e/ou sua aplicação industrial.

CAPÍTULO I

EXPRESSÃO DA TRANSPOSASE DE *mos1* EM DIFERENTES CEPAS DE *Escherichia coli* BL21 (DE3)

3 OBJETIVO

3.1 Objetivo geral

Expressar a transposase do elemento *mos1* e determinar as condições ideais para expressão e purificação da transposase de *Drosophila simulans*.

3.2 Objetivos específicos

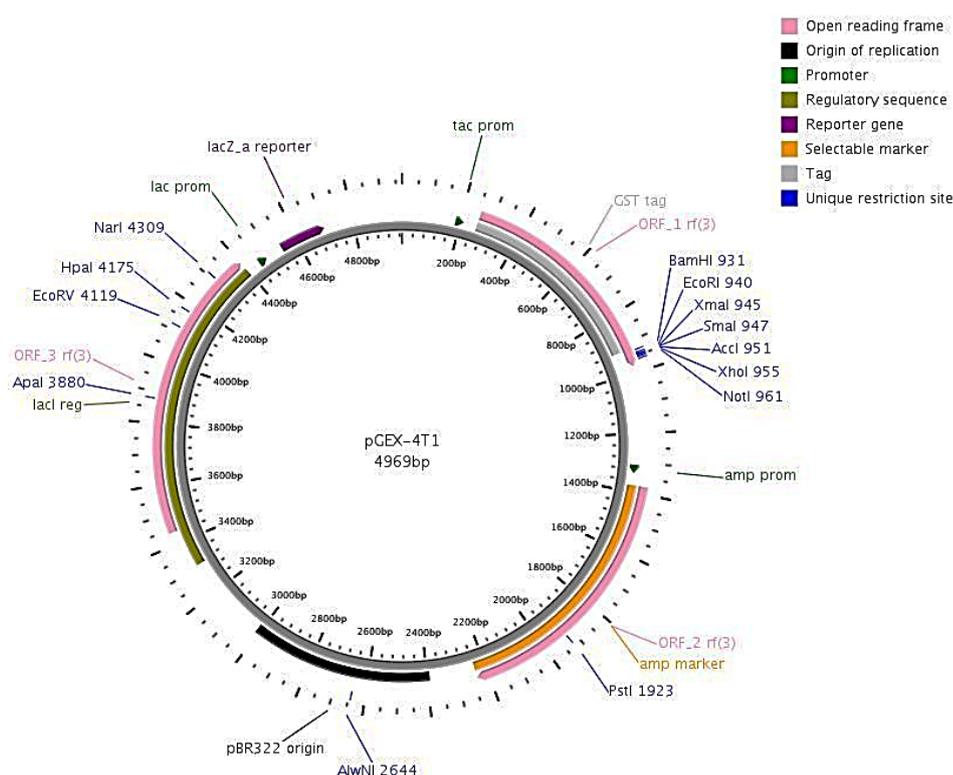
- Clonagem da sequência de DNA da transposase no vetor pGEX 4T1;
- Transformar eficientemente o vetor nas diferentes linhagens de *E. coli* BL21 (DE3);
- Testar, diferentes temperaturas de indução e diferentes concentrações de IPTG afim de otimizar as condições de expressão;
- Expressar de forma solúvel a transposase do elemento transponível *mos1*.

4 METODOLOGIA

4.1 Clonagem do elemento *mos1*

A amplificação do gene da transposase de *mos1* foi realizado através da técnica de Reação de Cadeia de Polimerase (PCR) em termociclador SureCycler 8800 (Agilent), utilizando primer direto PGEX 4T1 - Mos Fw 5' CTG GTT CCG CGT GGA TCC CCG ATG TCG AGT TTC GTG CCG AA 3' e como primer reverso PGEX 4T1 - Mos Rv 5' ACG CGC GAG GCA GAT CGT CAT TAT TCT AAG TAT TTG CCG TCG C 3'. Afim de anexar ao inserto sequencias complementares ao plasmídeo pGEX 4T1, com o propósito de cloná-lo através da técnica PCR circular descrita por Quan e Tian em 2009. Para a realização da clonagem, o plasmídeo pGEX 4T1 foi linearizado com enzima de restrição *XhoI*, seguindo protocolo de uso (*Invitrogen*). Foram utilizados 1 μ L (10 μ g/ μ l) do plasmídeo, 1 μ L de *mos1*, 1 μ L *XhoI*, 2 μ L de 10X Buffer K (Tris-HCl 200 mM, pH 8,5, MgCl₂ 100 mM, Dithiothreitol 10 mM, KCl 1 000 mM) e 10 μ l de água estéril. Na técnica de PCR circular utilizou-se o inserto e o plasmídeo pGEX4T1 (figura 3) clivado, as condições da PCR foram 35 ciclos, iniciando com desnaturação a 98°C durante 10 segundos, uma etapa de anelamento a 56°C por 10 segundos e finalizando com um passo de extensão a 72°C por 3 minutos.

Figura 3 - Mapa do vetor de expressão pGEX-4T-1, utilizado para expressar a transposase *mos1*.



4.2 Transformação bacteriana

A transformação bacteriana ocorreu através do método de transformação química seguindo protocolo padrão (Sambrook & Russel, 2000). As bactérias foram submetidas a cultivo *overnight*, a 37°C com agitação de 250RPM em 3ml de meio de Cultura Luria-Bertani [1% (p/v) NaCl, 0.5% (p/v) extrato de levedura e 1% (p/v) Triptona], as células foram então centrifugadas e acrescidas de 1ml de MgCl₂ e posteriormente novamente centrifugadas e adicionadas a 1ml de CaCl₂, nesta etapa as células foram deixadas em banho de gelo (0°C) durante 30 (trinta) minutos. Com as células devidamente preparadas o plasmídeo PGEX 4T1 foi adicionado (1µl de plasmídeo na concentração 10mg/µl), submetidas a processo de choque térmico, com as etapas de 20 minutos a 0°C, 2 min a 42°C e 5min a 0°C. As células foram adicionadas em 600µl de meio de cultura LB e incubadas a 37°C por 1 hora. A linhagem transformada foi cultivada em 25ml de Luria-Bertani [1% (p/v) NaCl, 0.5% (p/v) extrato de levedura e 1% (p/v) Triptona], com 100µg/ml de ampicilina e 170µg/ml de cloranfenicol e incubou-se a 37°C em período *overnight* para seleção das transformantes. As eficiências de transformação foram expressas como o número de transformantes por micrograma de DNA (unidade formadora de colónias, CFU). Esta etapa foi repetida para todas as linhagens presentes no Quadro 1.

Quadro 1 – Cepas utilizadas e seus respectivos genótipos

CEPA DE <i>Escherichia coli</i>	GENÓTIPO
<i>E. coli</i> BL21 (DE3) pLysS	B F ⁻ <i>ompT gal dcm lon hsdS_B (r_B⁻m_B⁻)</i> λ (DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>]) [<i>malB⁺</i>] _{K-12} (λ ^S) pLysS [<i>T7p20 ori_{p15A}</i>] (Cm ^R)
<i>E. coli</i> BL21-CodonPlus (DE3) -RP	B F ⁻ <i>ompT hsdS_B (r_B - m_B -)</i> <i>dcm⁺ Tetr gal λ (DE3) endA Hte [argU proL Cam^R]</i>
<i>E. coli</i> BL21-CodonPlus (DE3) -RIL	B F ⁻ <i>ompT hsdS_B (r_B - m_B -)</i> <i>dcm⁺ Tetr E. coli gal λ (DE3) endA Hte [argU ileY leuW Cam^R]</i>
<i>E. coli</i> BL21 (DE3) pT-GROE	B F ⁻ <i>ompT gal dcm lon hsdS_B (r_B⁻m_B⁻)</i> λ (DE3) PT7-groESL ColE1 Cm ^R

Fonte: Tapia, 2017

4.3 Indução por IPTG

O processo de indução foi realizado em diferentes condições. As transformantes foram cultivadas *overnight* na presença de 170µg/ml de cloranfenicol e 20µg/ml de ampicilina, a 37°C com agitação de 250 rpm em 3ml de meio de cultura LB [1% (p/v) NaCl, 0.5% (p/v) extrato de levedura e 1% (p/v) Triptona]. Posteriormente foi adicionado 10% (300µl) do pré inoculo em um novo meio de cultura contendo 3ml de meio de cultura LB [1% (p/v) NaCl, 0.5% (p/v) extrato de levedura e 1% (p/v) Triptona] sem adição de antibióticos, e deixadas para cultivo durante 2 horas a 37°C. As culturas foram divididas em dois grupos. As células do grupo T1 (temperatura 1) foram cultivadas a 30°C e as células do grupo T2 (temperatura 2) cultivadas a 37°C, e ambos os grupos foram deixados em indução por 3 horas com agitação de 250rpm. As culturas foram centrifugadas a 12.000 RPM e adicionado a elas 100µl de H₂O deionizada. Para lise celular as amostras foram submetidas a sonicação no equipamento *Tissue laser* em uma frequência de 30Hz/s durante 1 minuto, em frequência direta. Obtido assim o extrato proteico, as amostras foram centrifugadas e retirado seu sobrenadante para análise de solubilização, e o pellet correspondente suspenso em 100µl de H₂O deionizada.

4.4 Análise de expressão

As análises de expressão foram realizadas em gel de poliacrilamida desnaturante SDS-PAGE, com gel de empacotamento 4% (TRIS 0,5M pH 6,8, 10% de SDS, 30% de acrilamida, TMED, 10% de APS e água deionizada) e gel de corrida 12% (TRIS 1,5M pH 8,8, 10% de SDS, 30% de acrilamida, TEMED, 10% de APS). Para as amostras foram utilizados 10µl de extrato proteico, 5µl de tampão de amostra (glicerol, B-mercaptoetanol, 10% SDS, TRIS 0,5M pH 6,8) com tampão de corrida, a 100V em uma cuba eletroforética. O processo de coloração do gel ocorreu utilizando o corante de proteínas *comassie blue R-250* em período *overnight*.

4.5 Análise de solubilização

As análises de solubilização foram realizadas em gel de poliacrilamida desnaturante SDS-PAGE, com gel de empacotamento 4% (TRIS 0,5M pH 6,8, 10% de SDS, 30% de acrilamida, TMED, 10% de APS e água deionizada) e gel de corrida 12% (TRIS 1,5M pH 8,8, 10% de SDS, 30% de acrilamida, TEMED, 10% de APS). Para as amostras foram utilizados 10µl de amostra. A amostra corresponde ao extrato

proteico foi centrifugada a fim de separar o sobrenadante do sedimento, e então a porção do sobrenadante foi alíquotada e utilizada como amostra. Foi adicionado a amostra 5µl de tampão de amostra (glicerol, B-mercaptoetanol, 10% SDS, TRIS 0,5M pH 6,8) com tampão de corrida, a 100V em uma cuba eletroforética. O processo de coloração do gel ocorreu utilizando o corante de proteínas *comassie blue R-250* em período *overnight*.

5 RESULTADOS E DISCUSSÃO

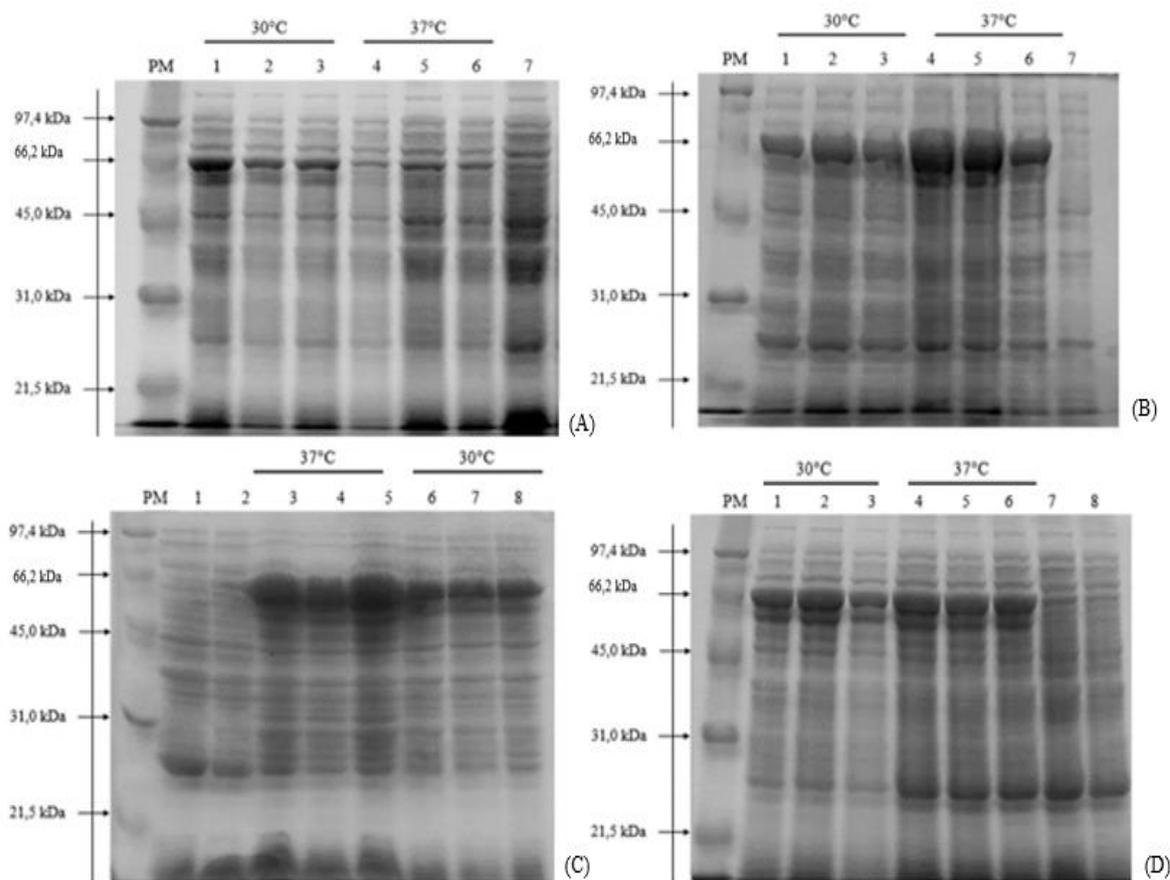
5.1 Expressão

A proteína de interesse foi expressa em todas as cepas testadas (quadro 1) com o tamanho esperado de aproximadamente 67kDa (transposase + *Glutathione S-transferases*), no entanto na linhagem *E. coli* BL21 (DE3) pT-GROE, ocorreu uma significativa diminuição na quantidade de expressão em relação a outras cepas de *E. coli* BL21 (DE3) testadas, como demonstrado na figura 4, onde vemos as induções das cepas de *E. coli*, em (a) *E. coli* BL21 (DE3) pT-GROE, em (b) *E. coli* BL21 (DE3) pLysS, e em (d) *E. coli* BL21-CodonPlus (DE3)-RIL onde as amostras de 1 a 3 foram induzidas com 1mM de IPTG, 0,5mM de IPTG e 0,1mM de IPTG respectivamente a 30°C e as amostras de 4 a 6 com 1mM de IPTG, 0,5mM de IPTG e 0,1mM de IPTG respectivamente a 37°C, a amostra 7 é correspondente a cepa sem indução. Em (c) *E. coli* BL21-CodonPlus (DE3)-RP as amostras de 1 a 3 foram induzidas com 1mM de IPTG, 0,5mM de IPTG e 0,1mM de IPTG respectivamente a 37°C e as amostras de 4 a 6 com 1mM de IPTG; 0,5mM de IPTG e 0,1mM de IPTG respectivamente a 30°C. Nossos resultados demonstraram de forma significativa que no caso da transposase do elemento *mos1*, as diferentes concentrações de indutor IPTG (0,1mM; 0,5mM e 1,0mM) não influenciam a eficiência de expressão desta proteína, o que corrobora com dados apresentados em outros trabalhos sobre expressão proteica que afirmam que a utilização de menores concentrações de IPTG são capazes de induzir mas sem influenciar a quantidade de proteína produzida (Larentis et al. 2011; Einsfeldt et al. 2011; Olaofe et al. 2010), tais estudos afirmam que existe uma tendência linear de aumento da atividade da enzima *β-galactosidase* variando a concentração de IPTG

até 1,0mM. Esse resultado foi atribuído provavelmente à titulação completa da proteína repressora LacI pelo IPTG, de tal modo que IPTG adicional não aumentaria a transcrição do gene lacZ de forma significativa (Olaofe et al. 2010). No entanto dados para a temperatura, demonstram que esta influenciou notavelmente na quantidade de proteínas expressa, como demonstrado figura 4 (A, B, C e D). Em estudos anteriores foi demonstrado que algumas cepas de bactérias, como exemplo *E. coli* HMS174, apresentou uma baixa taxa de expressão da proteína Est-AF (*alpha/beta hydrolase*) de *Archaeoglobus fulgidus* (Miyake et al. 2003), porém é eficiente na produção heteróloga de *cytosolic serine hydroxymethyltransferase* de *Danio rerio* (Chang et al. 2006). Isto demonstra que a ORF altera o nível de expressão em determinadas cepas de *E. coli*. Segundo trabalhos publicados recentemente (Stosch 2016), pode-se constatar que a temperatura influencia no nível de expressão proteico, porém não é o único fator responsável por diferentes expressões, dentre eles o pH, o tempo de indução e a disponibilidade de carbono também são determinantes para a eficiência de expressão de proteínas. Estudos realizados recentemente em linhagens de *Escherichia coli*, analisaram o efeito da temperatura e da concentração de IPTG na expressão de proteínas heterólogas e também no crescimento celular (Kim et al. 2016). No trabalho proposto por Kim e seus colaboradores, notou-se que o nível de expressão aumentou concomitantemente com o aumento da temperatura, e o melhor nível de expressão era observado aos 37° C, no entanto quando a temperatura excedia os 44°C havia uma queda de crescimento celular e nível de expressão. A concentração ótima de IPTG ficou próximo aos 0,4 mM, as taxas de expressão diminuía quando as concentrações de IPTG ficavam acima de 0,6 mM (Kim et al. 2016), informações que corroboram com nossos resultados, que demonstram que maiores níveis de expressão são notados a temperatura de 37°C.

Acredita-se que essa alteração seja devida ao tempo de divisão da bactéria, que em temperaturas inferiores leva um tempo maior do que em maiores temperaturas.

Figura 4 - Gel de poliacrilamida SDS PAGE 12% da indução das cepas de *Escherichia coli*. Em (A) *E. coli* BL21 (DE3) pT-GROE, em (B) *E. coli* BL21 (DE3) PlyS, (C) BL21 (DE3) RP e em (D) *E. coli* BL21 (DE3) RIL onde as amostras total das proteínas (sobrenadante e sedimento) foram induzidas com 0,1mM; 0,5mM e 1,0mM de IPTG respectivamente a 30°C e 37°C. As canaletas 7 de A e B, e as canaletas 1 e 2 de C e D correspondem a cepa sem indução a 37°C. Em (C) *E. coli* BL21 (DE3) RP, induzida a 37°C e 30°C.



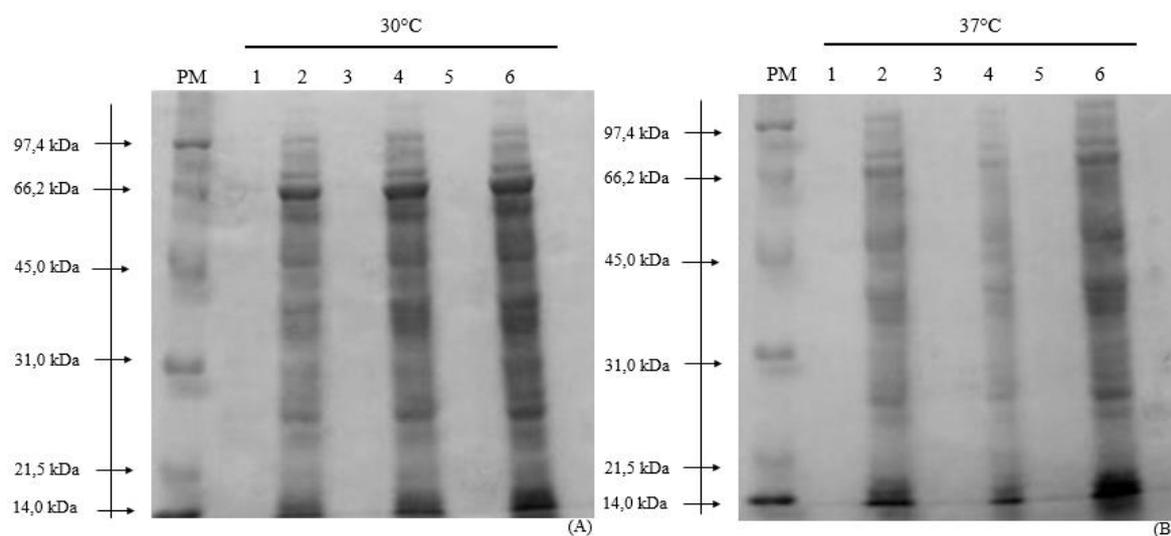
Outra explicação para dificuldade de expressão heteróloga de proteínas, é a modulação realizada por códons específicos, que pode causar erros de tradução (Robinson et al. 1984; apud Thangadurai et al. 2008), a escolha de um hospedeiro que sintetize tRNAs apropriados é importante para uma boa expressão, pois a presença de determinados códons faz com que o ribossomo cause a terminação prematura da cadeia e afete diretamente a produção proteica (Thangadurai et al. 2008). Nossos resultados demonstraram que todas as cepas de *E. coli* BL21 (DE3) pLysS, *E. coli* BL21-CodonPlus (DE3) –RP; *E. coli* BL21-CodonPlus (DE3) –RIL; *E. coli* BL21 (DE3) pT-GROE foram maquinarias eficazes para realizar a expressão heteróloga da transposase do elemento *mos1*. No entanto a cepa *E. coli* BL21 (DE3)

pT-GROE obteve um menor rendimento na expressão independente da temperatura de indução ou da concentração de IPTG utilizado, o que vai contra o que se acreditava que as linhagens com o vetor pT-groE expressaria maiores quantidades de proteínas solúveis, tendo em vista a presença da chaperonas responsável por o correto desdobramento das proteínas (Yasukawa et al. 1995) (Herrmann et al. 2004). As concentrações de IPTG não foram capazes de modificar o nível de expressão da proteína de todas as linhagens, no entanto as diferentes temperaturas interferem diretamente nos níveis de expressão, essas informações estão sumarizadas no quadro 2, no qual o nível de expressão é classificado com (+) para expressão satisfatória e (++) para níveis de expressão maiores que os demais. Os níveis de solubilidade foram classificados apenas entre solúvel ou insolúvel na análise para todas as linhagens, concentrações e temperaturas.

5.2 Solubilidade

O mecanismo pelo qual as proteínas se tornam solúveis não é totalmente compreendido. A solubilidade é o indicador de um dobramento correto de proteínas (Liddy et al. 2010) sendo relevante para a eficiência da purificação de uma proteína

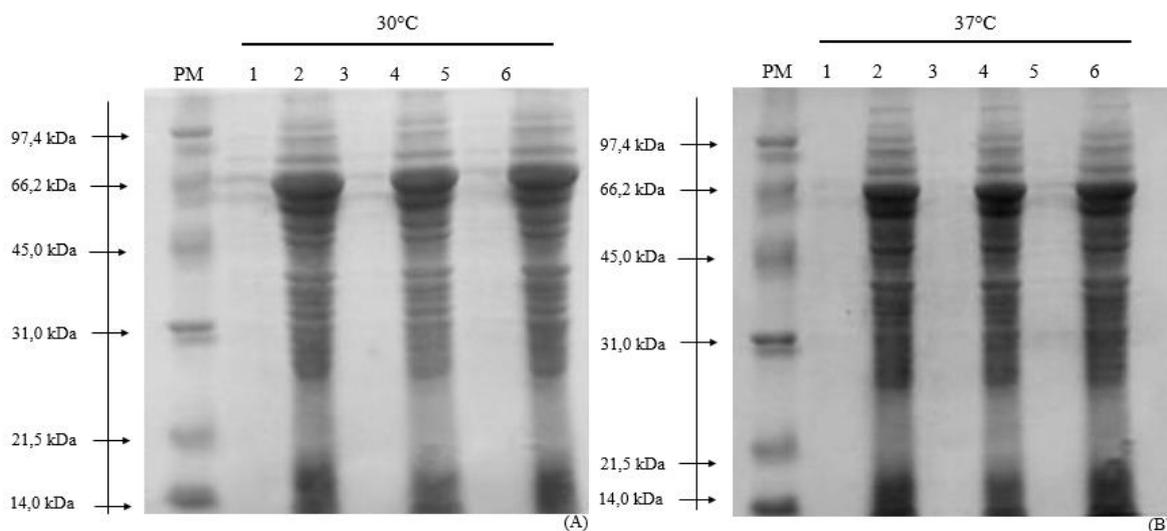
Figura 5 - Gel de poliacrilamida SDS PAGE 12% da cepa *E. coli* BL21 (DE3) pT-GROE. PM – Peso molecular. Na figura (A) as amostras foram induzidas a temperatura de 30°C e em (B) 37°C, intercalando sobrenadante e pellet induzidos por IPTG 1,0mM, 0,5mM e 0,1mM.



ativa. A proteína alvo deste estudo tem aproximadamente 67kDa, se demonstrou insolúvel, como demonstrado na figura 5 (A e B), figura 6 (A e B), figura 7 (A e B) e figura 8 (A e B).

As diferentes temperaturas na expressão (30°C e 37°C) não demonstraram influenciar a solubilidade da proteína, tão pouco as diferentes concentrações de IPTG (0,1mM; 0,5mM e 1,0mM). Este dado pode ser explicado com base na fusão intrínseca da sequência da transposase a uma *Glutathione S-transferases* (GST-tag), que facilita a purificação em coluna de HPLC (*High Performance Liquid Chromatography*) e o

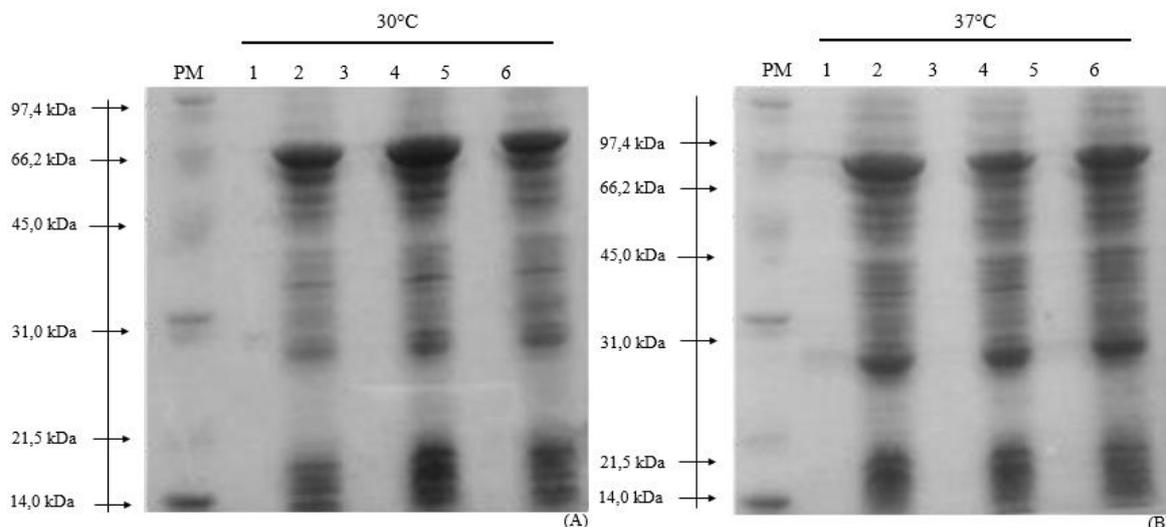
Figura 6 - Gel de poliacrilamida SDS PAGE 12% *E. coli* BL21-CodonPlus (DE3) –RIL. PM – Peso molecular. Na figura (a) as amostras foram induzidas a temperatura de 30°C e em (b) 37°C, intercalando sobrenadante e pellet induzidos com 0,1mM; 0,5mM e 1,0mM de IPTG.



resgate da proteína de interesse.

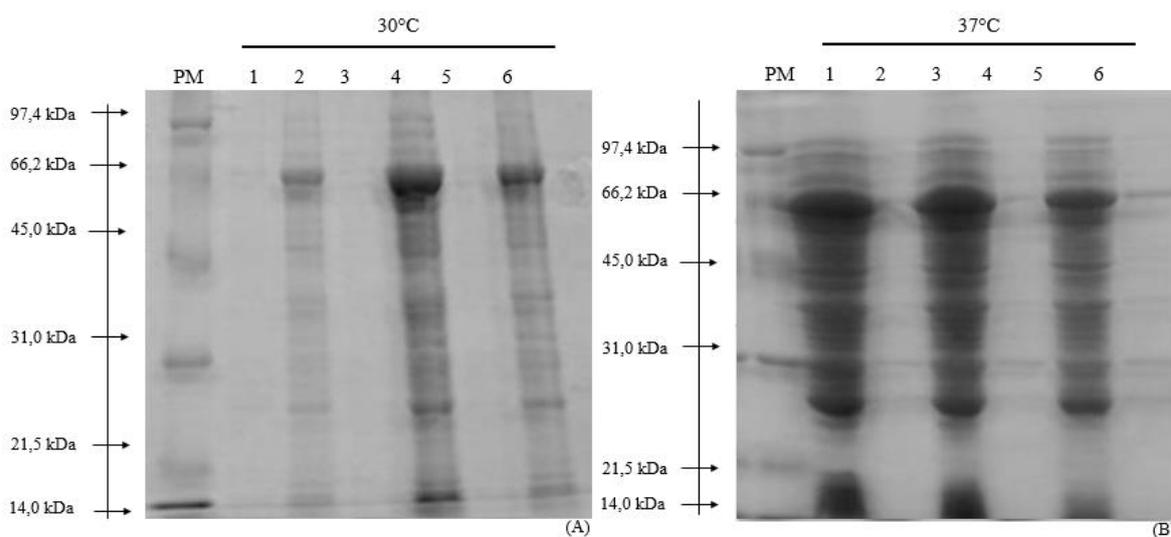
Um estudo publicado recentemente por Guerra e sua equipe em 2016, demonstrou que a solubilidade das proteínas é diretamente afetada pelo hospedeiro utilizado. Nesse experimento, diferentes cepas de *E.coli* foram usadas e concluiu-se

Figura 7- Gel de poliacrilamida SDS PAGE 12%; cepa *E. coli* BL21-CodonPlus (DE3) -RP PM: peso molecular. Na figura (a) as amostras foram induzidas a temperatura de 30°C e em (b) 37°C, intercalando sobrenadante e pellet induzidos com 0,1mM; 0,5mM e 1,0mM de IPTG.



que a linhagem BL21- CódonPlus que codifica RNAt raros e a BL21- pG-KJE8 que super expressa chaperonas apresentavam expressão de proteínas insolúveis, independentemente das temperaturas utilizadas na expressão bem como as concentrações de IPTG mesmo com diferentes temperaturas e diferentes concentrações de IPTG.

Figura 8 - Gel de poliacrilamida SDS PAGE 12%; cepa BL21(DE3) PlysS. PM - Peso molecular. Na figura (A) as amostras foram induzidas a temperatura de 30°C e em (B) 37°C, intercalando sobrenadante e pellet induzidos por IPTG 1,0mM, 0,5mM e 0,1mM



Quadro 2 - Análise da expressão das cepas de *E.coli*

Cepa de <i>Escherichia coli</i>	Temperatura	Concentração de IPTG	Nível de Expressão	Solubilidade
E. coli BL21 (DE3) pT-GROE	30°C	0,1mM	(+)	Insolúvel
E. coli BL21 (DE3) pT-GROE	30°C	0,5mM	(+)	Insolúvel
E. coli BL21 (DE3) pT-GROE	30°C	1,0mM	(+)	Insolúvel
E. coli BL21 (DE3) pT-GROE	37°C	0,1mM	(+)	Insolúvel
E. coli BL21 (DE3) pT-GROE	37°C	0,5mM	(+)	Insolúvel
E. coli BL21 (DE3) pT-GROE	37°C	1,0mM	(+)	Insolúvel
E. coli BL21 (DE3) pLysS	30°C	0,1mM	(+)	Insolúvel
E. coli BL21 (DE3) pLysS	30°C	0,5mM	(+)	Insolúvel
E. coli BL21 (DE3) pLysS	30°C	1,0mM	(+)	Insolúvel
E. coli BL21 (DE3) pLysS	37°C	0,1mM	(++)	Insolúvel
E. coli BL21 (DE3) pLysS	37°C	0,5mM	(++)	Insolúvel
E. coli BL21 (DE3) pLysS	37°C	1,0mM	(++)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RIL	30°C	0,1mM	(+)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RIL	30°C	0,5mM	(+)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RIL	30°C	1,0mM	(+)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RIL	37°C	0,1mM	(++)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RIL	37°C	0,5mM	(++)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RIL	37°C	1,0mM	(++)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RP	30°C	0,1mM	(+)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RP	30°C	0,5mM	(+)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RP	30°C	1,0mM	(+)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RP	37°C	0,1mM	(++)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RP	37°C	0,5mM	(++)	Insolúvel
E. coli BL21-CodonPlus (DE3) -RP	37°C	1,0mM	(++)	Insolúvel

Fonte: Tapia, 2017

CAPÍTULO II

Manuscrito

O artigo intitulado: “Design and construction of *Escherichia coli* expression vector: addressing to for soluble proteins” está disposto na forma na qual será submetido para o periódico Protein Expression and Purification (ISSN: 1046-5928).

Design and construction of *Escherichia coli* expression vector: looking for soluble proteins

Jéssica Silva Tapia^{1,2}, Luiz Fernando Duarte da Silva¹, Thalita Fonseca de Araújo¹, Juliano Tomazzoni Boldo², Gabriel da Luz Wallau³, Paulo Marcos Pinto^{1,2,*}.

¹Laboratório de Proteômica Aplicada, Campus São Gabriel, Unipampa.

²Programa de Pós-Graduação em Ciências Biológicas, Campus São Gabriel, Unipampa.

³Fundação Oswaldo Cruz, Departamento de Entomologia - Centro de Pesquisas Aggeu Magalhães.

* Author for corresponding: paulopinto@unipampa.edu.br. Phone and Fax. +55 (55) 3237-0850. Av. Antonio Trilha, 1847. São Gabriel – RS – Brazil. Zip code: 97300-000.

ABSTRACT

Expressing heterologous proteins has become increasingly indispensable; the search for these proteins has been shown to be increasing, in view of the need to dispose of these in sufficient quantities to elucidate questions around its function and structure. The use of recombinant proteins such as insulin has aroused interest in the search for new proteins in the pharmaceutical industry or clinically with therapeutic interest. The most common host for heterologous protein production is the *Escherichia coli* bacterium, due to its innumerable advantages, rapid growth, low cost cultivation, the ability to introduce exogenous DNA using a simple transformation method in addition to enabling production of heterologous proteins in large scale. However, the expression of biologically active and soluble proteins present some difficulties. In this process, subsequent cloning is required which ultimately increases the cost and time of the experiments. Thus, the objective of this work is to develop an expression vector

that does not require subcloning, using only one cloning and transformation event of the expression bacterium. The expression construct was designed in silico primarily from an IPTG responsive Lac promoter followed by a cell export signal (peptide-signal), a Strep-tagII fragment for purification and a site for thrombin, allowing the cleavage of the protein of the merged tags. Two sites for restriction endonucleases (BamHI and XbaI) were inserted which allow the cleavage of the cloned gene in the construct. In this way, we enable the expression of heterologous proteins, from the presented expression vector that may provide soluble and purified proteins more efficiently.

Key words: Heterologous proteins, expression vector, soluble proteins and cloning.

INTRODUCTION

The expression of heterologous proteins is an important biotechnological tool, in view of the increasing demand to produce sufficient proteins to study its function and structure, as well as its industrial applications (Schmidt 2004). The use of heterologous proteins has grown since the discovery of insulin production in a heterologous way, and the search for new expression systems capable of expressing proteins in the biologically active form with low costs and high yields becomes increasingly large (Vaz, 2008).

Several hosts are feasible for expression of heterologous proteins, such as fungi, yeasts, single-celled algae and bacteria or mammalian (Rosano & Ceccarelli 2014). Among these, the most common is the bacillary bacterium *Escherichia coli*, due to the vast knowledge about its characteristics and the advantages they present in relation to the others for this type of process (Rosano & Ceccarelli 2014). *Escherichia coli* ideally serves as a host for the production of heterologous proteins at low cost and high yields because it can easily grow to high densities and because the host-vector systems based there have been more advanced with the concomitant development of many high-Level expression vectors. *E. coli* host-vector systems are therefore most frequently used for expression of heterologous genes. (Ishii & Yura 2003).

E. coli present several advantages compared to other organisms due to its extremely rapid growth cycle, under ideal environmental conditions, it double the population size in approximately 20 minutes (Sezonov et al. 2007), as well as the low cost needed for bacterium maintenance. The ability to transform with exogenous DNA using a simple transformation method, in addition to allowing the production of

heterologous proteins in large quantity, considering that high cellular densities can be achieved (Jong Hyun Choi et al. 2006).

However, heterologous expression of proteins may be limited by numerous factors. Producing a recombinant protein requires analyzing several aspects, such as ease of host maintenance, but this choice depends on protein properties and final destination, protein solubility and stability, as well as yield needed (Jonasson et al. 2002; Williamson et al. 2000). The main steps for the production of a recombinant protein are simple in the theoretical scope, since they comprise the choice of the gene of interest, the choice of the expression system, cloning and expression, and then the protein is ready for purification and characterization. (Miyamoto 2013). Although several improvements in heterologous protein expression happened in the last years as highlighted above, in practice the different steps can be more complicated than anticipated. Several subsequent cloning may be needed that ultimately increase the cost and time of the experiments, the protein obtained may not be soluble or active and even the failure to obtain any protein at all are some of the problems often encountered.

Considering that in many cases protein expression is a challenging, in spite of the many options. Because of this, our aim is to establish a new technology for the expression of biologically active proteins and in soluble form using *E. coli* as a host. Our purpose was to develop an expression vector for the production of useful proteins in a low-cost, high-yielding way in order to facilitate protein cloning, expression and purification processes.

MATERIALS AND METHODS

Designing the expression construct

The construct was designed based on the pUC19 vector, from a T7 promoter (5' TAATACGACTCACTATAG 3') which uses the T7 phage polymerase which is up to 20 times more active than the bacterial one, being able to offer a very high level of protein expression (Stano & Patel 2004), followed by a Lac operator (5' 'GGAATTGTGAGCGGATAACAATTCC 3') whose function is to allow the induction of expression in the presence of IPTG when it binds the repressor protein, a ribosomal binding site (RBS) (5' 'AAGGAG 3') responsible for initiating translation with the 16s subunit binding of the ribosome. It has also been necessary to add an initial ATG

shortly after the RBS to ensure protein synthesis from the mRNA, and the signal peptide (5'AAA AAG ATT TGG CTG GCG CTG GCT GGT TTA GTT TTA GCG TTT AGC GCA TCG GCG GCG 3') for extracellular exportation of the protein, i.e., this sequence has the function of marking the proteins be exported. For purification of the protein a small sequence of 8-amino acids (WRHPQFGG) with low molecular weight called Strep-tagII was added (5 'TGGAGCCACCCGCAGTTCGAAAAG 3'), because of its high intrinsic stability, protease resistance, low non-specific interaction, is ideal for protein purification processes (Schmidt & Skerra 2007). A thrombin site (5'-CTGGTGCCACGCGGTAGT 3 ') was fused to allow cleavage of the protein of interest from the fused tags. Two sites for XbaI restriction endonucleases (TCTAGA) and BamHI (GGATCC), respectively, were inserted which flank the gene of interest and allow cleavage of the gene in the construct. In figure 1, it is possible to see the complete sequence of the construct.

Primers Design

The specific forward and reverse primers were designed to allow fusion between the construct and the pUC19 vector. The primer was then synthesized with the appropriate sequences. The primers were designed as each forward primer contained EcoRI (GAATTC) a restriction site, and reverse primers contained HindIII (AAGCTT) restriction site. The primers were designed using the plasmid sequence and 20 nucleotides of designing insert, being forward primer pUC19EcoRI Fw 5' AAA CGA CGG CCA GTG AAT TCT AAT A 3' ; and reverse primer pUC19HindIII Rv 5'ACC ATG ATT ACG CCA AGC TTT CAG 3'.(Quan & Tian 2009).

Synthesis of the construct

The fragment (which has a total of 852bp) was amplified by the Polymerase Chain Reaction (PCR) technique. The sequence was synthesized according to the desired fragments. That first PCR creates a linear insert with complementary sequences of the plasmid at both ends. These extensions subsequently allow the strands of the PCR product to act as a pair of oversized primers on the vector fragment. Using the Phusion™ High-Fidelity DNA Polymerase (Thermo-Scientific) on a SureCycler 8800 (Agilent) thermocycler, as instructed by manufacturer, using forward and reverse primers. Amplification was achieved with an initial denaturation

step at 94°C for 5 min, followed by 25 cycles of denaturation at 94°C for 30s, annealing at 60°C for 30s and extension at 72°C for 30s. The amplification reaction was performed in a final volume of 25µl of 5,0µl 10X sample buffer (50 mM Tris-HCl, pH 8.0, 50mM KCl, 1mM DTT, 0.1mM EDTA, stabilizers and 50% V/v) glycerol (Sigma), 2.5µl 5mM dNTP (Invitrogen), 0.75µl DMSO, 1.25µl of each specific primer, 1µl of the synthesized construct sample, 0.25µl Phusion™ High-Fidelity DNA Polymerase (Thermo-Scientific) and 14 µl sterile ultrapure water. PCR products and DNA fragments were routinely separated on 1 % (w/v) agarose (Biomax - Vetec). The samples were stained with GelRed and subjected to electrophoresis at 100 volts. The gels were photographed in L-pix loccus biotechnology.

Preparation of the plasmid

Plasmid pUC19 was chosen to be the vector where the insert will fuse. The molecule is a small double-stranded circle, 2686 base pairs in length, and has a high copy number. The pUC19 carries a 54 base pair multiple cloning site polylinker (5'GAA TTC GAG CTC GGT ACC CGG GGA TCC TCT AGA GTC GAC CTG CAG GCA TGC AAG CTT3') that contains unique sites for 13 different hexanucleotide-specific restriction endonucleases. The plasmid can be access in by the Assession number L09137 in GENBANK. For the construct to be fused to the vector, it was digested using the enzyme BamHI (GGATCC) that has cleavage site in vector pUC19 as can be seen in the map shown in figure 2. Following the protocol of use (Invitrogen) use 7µL (concentration) of plasmid pUC19, 1µL BamHI, 2µL of 10X Buffer K (200 mM Tris-HCl; pH 8.5; 100mM MgCl₂; 10mM Dithiothreitol; 1,000mM KCl) and 10µL sterile water. The reaction was subjected to 37°C for 16 hours. For inactivation of the restriction enzyme, there was a thermal step at 65°C for 20 Minutes.

Expression vector synthesis

The vector and insert were PCR amplified using Phusion DNA Polymerase. About 5µl of purified PCR product (purified insert with PEG13%) (Range 50ng in total) and 40ng of the appropriately linearized pUC19 vector, corresponding 2µl, at conditions 4:1. The amplification reaction was performed in a final volume of 25µl of 2,0µl 10X sample buffer (50mM Tris-HCl; pH 8.0; 50mM KCl; 1mM DTT; 0.1mM EDTA, stabilizers and 50% w/v) glycerol (Sigma); 2,0µl 5mM dNTP (Invitrogen); 0,6µl DMSO; 1,25µl of each specific primer; 0,50µl AccuPrime™ Taq enzyme DNA Polymerase Phusion (Sigma) and 13µl sterile ultrapure water. On a SureCycler 8800 (Agilent) thermocycler. Amplification was achieved with an initial denaturation step at 98°C for 10 seconds, followed by 25 cycles of denaturation at 98°C for 30 s, annealing at 56°C for 30 seconds and extension at 72 °C for 3 minutes. PCR products and DNA fragments were routinely separated on 1 % (w/v) agarose (Biomax - Vetec), the samples were stained with GelRed and subjected to electrophoresis at 100 volts. The gels were photographed in L-pix loccus biotechnology.

Transformation of Escherichia coli

Ligation mixtures were transformed into *E. coli* DH5α electrocompetent cells and the extraction process was carried out according to the manufacturer's instructions for the Invisorb Spin Plasmid Mini Kit (Invitex). The recombinant plasmids were confirmed by restriction digestion by BamHI, following the protocol of use (Invitrogen). Use 7µL (concentration) of plasmid pUC19, 1µL BamHI; 2µL of 10X Buffer K (200mM Tris-HCl; pH 8.5; 100mM MgCl₂; 10mM Dithiothreitol and 1,000 mM KCl) and 10µL Sterile water. The reaction was subjected to 37°C for 16 hours. And then treated with RNase (1µL for 1 hour). The transformed *Escherichia coli* were grown in 25ml Luria-Bertani (1% (w/v) NaCl, 0.5% (w/v) yeast extract, and 1% (w/v) Tryptone) culture media with 120µL (20µg/ml) ampicillin, and incubated at 37°C for 17 hours to select the transformed. Transformation efficiencies were expressed as the number of transformants per microgram of DNA (colony-forming unit, CFU).

RESULTS AND DISCUSSION

Designing the expression construct

The designed construct has a total of 852bp, which includes all given sequences and some spacers. As we can see in figure 1, the designing used a promoter T7 to drive high-level expression of the cloned fusion protein. The T7 promoter utilizes the T7 phage RNA polymerase, with the advantage of more active RNA polymerase than the host bacterial polymerase, and a much higher level of protein expression (Stano & Patel 2004). Followed by a regulatory sequence, the Lac operator, the Lac promoter is similar to that used by the wild type strains of *E. coli*, and is repressed by the lacI protein in the absence of inducer and induced in its presence (Lodish et al., 2000). An RBS site was also added followed by initial ATG, a signal peptide site and Strep-tagII sequence for purification. The Strep-tag provides a reliable tool in recombinant protein chemistry, which has proven particularly useful for the parallel isolation and functional analysis of multiple gene products in proteome gene, and versatile expression vectors are available for various host organisms especially for *Escherichia coli* (Schmidt & Skerra 2007). However, the Strep-tag II is largely resistant to cellular proteases, it can be used in the presence of mild detergents and it is biochemically almost inert. This gives a clear advantage over the His6 tag (Skerra & Schmidt 2000). Other vectors of expression already proven use the Strep-tag system for purification and present positive results (Weeks et al. 2007). Likewise a thrombin cleavage site and two endonuclease sites flanking the gene of interest, XbaI restriction sites were designed in the upstream and BamHI restriction sites were designed in the downstream presence. The gene were demonstrated by the FJL733 gene, but which can be replaced by any other gene using restriction enzymes of Circular PCR (Quan & Tian 2009).

Synthesis of the expression vector

First, the insert is PCR-amplified with the chimeric primers so that the final PCR product has overlapping regions with the vector. In addition, the result of PCR products and DNA fragments were routinely separated on 1 % (w/v) agarose (Biomax - Vetec), the samples were stained with GelRed and subjected to electrophoresis at 100 volts (figure 3). The gels were demonstrated in figure 4, we used 5µL of amplification of the construct. Then, vector (figure 4) and insert are mixed, denatured and annealed; the hybridized insert then is extended by Phusion DNA polymerase using vector as a

template until polymerase reaches 5' end of the insert. Using the polymerase extension mechanism to extend double- stranded overlapping insert and vector to form a complete plasmid. After several PCR cycles, the new plasmid with two nicks (one on each strand) is accumulated as a product. As shown in another paper, after denaturation and annealing, the insert and the vector will hybridize and extend using each other as a template to form a complete double-stranded plasmid, leaving only one nick in each strand (Quan & Tian 2009). Quan and his team added sequences at both ends of the lacZa gene to overlap the ends of the linearized vector. The linearized vector with the lacZa gene was not mixed without including the PCR primers in a typically normal PCR mixture. The expression vector can be seen in the figure 3, in this map we can see all the sequences of insert and the basic structure of pUC19. The gel electrophoresis results showed the formation of a significant amount of the insert insert fusion product and this appears to increase with the reaction cycles, results which corroborate with those presented by Quan, where there was a proportional increase after 2 and 5 cycles of reaction (Quan & Tian 2009).

Cloning

The expression vector constructed was transfected into *E. coli* DH5 α electrocompetent prepared following protocol (Sambrook & Russel, 2000) and grown in media Luria-Bertani (1% (w/v) NaCl, 0.5% (w/v) yeast extract, and 1% (w/v) Tryptone) culture media with 120 μ L (20 μ g/ml) ampicillin, and incubated at 37°C for 17 h to select the transformed. An examination of the large numbers of transforming colonies, indicating minimal cloning error or carry-over of intact vectors. The plasmid was extracted following the miniprep procedure using the Invisorb Spin Plasmid Mini Kit (Invitex) according to the manufacturer's instructions and the samples were precipitated using PEG13%. The linear vector can be seen in figure 4, and the samples were clivated with BamHI, thus it is possible to see the molecular weight of 3486bp (figure 5) corresponding to 852bp (figure 3) of the designed construct fused to the 2686bp of vector pUC19 (figure 4). A similar analysis done in another study shows that the efficiency of this type of cloning is almost 100% (Quan & Tian 2009).

FINAL REMARKS

The rapid and efficient production of recombinant proteins requires general "single-size" application methodologies (Weeks et al. 2007). However, the production

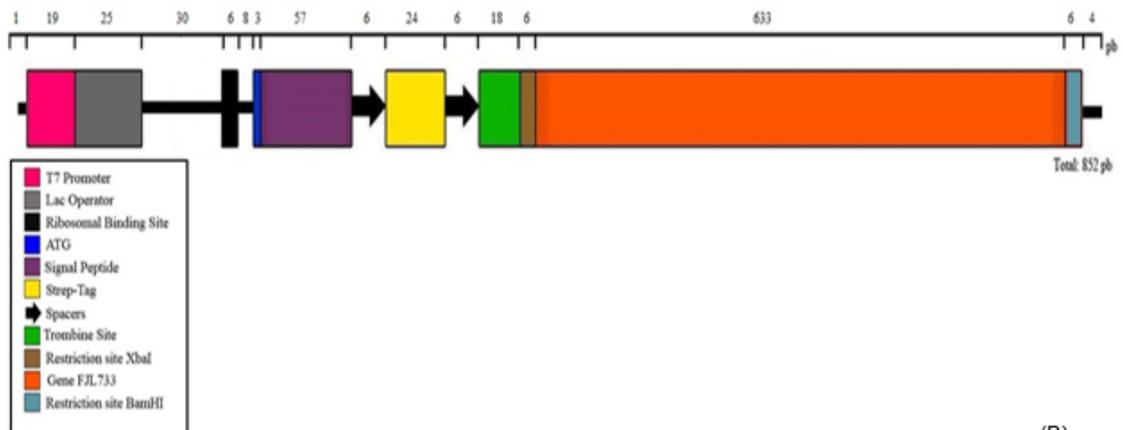
of a protein varies with its sequence and by which endonucleases that may be used during cloning are defined, the amino acid sequence determines the physical properties of the protein and thus determines the purification procedure (Weeks et al. 2007). It can be limited by several factors (Miroux & Walker 1996, apud Dumon-seignovert et al. 2004) and depend on several subclonations that end up making the process more expensive (Nascimento et al. 2003). In addition, many heterologous proteins, especially eukaryotic proteins, associate with the cytoplasm and form insoluble biologically inactive aggregates known as inclusion bodies when expressed at high levels in *E. coli*, which hinder the solubility of the protein (Ishii & Yura 2003).

In this present study independent protein sequence techniques were developed, a vector that in addition to expression addresses cloning and purification individually. It is known the real biotechnological, industrial and clinical importance of the production of heterologous proteins in a soluble and purified form, since its applicability is ample (Cereghino & Cregg 2000; Gustafsson et al. 2004). We focused on constructing vectors for the expression of proteins in a heterologous manner in a facilitated manner. Using as a base the vector pUC19, chosen because of its relatively small size and therefore suitable for modification and for cloning uses, is stable in addition to its number of copies considered. The establishment of this technology for the expression of biologically active proteins in the soluble form using *E. coli* is would be a tool for the mass production of useful proteins and opens precedents for fields of in-depth structural and functional studies of proteins as well as production for industrial and clinical use. In a recent future, it is expected to test the viability of vector expression, to demonstrate the functionality of an expression vector that makes the process much more effective than others available do.

Figure 1 - Schematic map of the construct, and corresponding sequence. The figure depicts the major features present in construct. In (A) We demonstrate the sequences of the construct with their identification below the underline, and in (B) a schematic with the fragments of the construct represented by colors and the respective value of base pairs in the upper part of the figure.

5' CTAATACGACTCACTATAGGGGAATTTGTGAGCGGATAACAATTCCCCTCCAGAAATAATTT
 T7 promoter Lac Operator
 TGTTTAACTTTAAGAAGGAGATATACATATGAAAAAGATTTGGCTGGCGCTGGCTGGTTTTAGT
 Ribosomal Biding Site Signal-Peptide
TTTAGCGTTTTAGCGCATCGGCCGCCCAAGCTGGAGCCACCCGCAGTTTCGAAAAAGGGTGCAC
 Strep-tagII
GGTGCCACGCGGTAGTTCTAGAATGAGTGACCGCTATTTAGAAACAAAGGATTAGTATCAAATT
 Trombin site XbaI
 TTGCGTGAAATTGAACAAGTCTGCAAGTGAGACCCACCATCTTTTAAAAGAAACTTATGGGGA
 TGAAGTCATGTCAAGGGCCAGAGTTTTTACTGGCACAAAAGGTTTAAAGAAGGACGGGAAGA
 TGTTCGAGATGATGCCCGAAGTGGCGTCCAGTCACCCACCGAACAGATGACAATATCCAGAA
 GGTCAAGGACTTGGTTTTGTTCAAACAGGCAGTTAACCGTGAGGATGATGGCTGAAGAGTTAAA
 TTTAGACAAAAGAACTGTTAGGCTCATTTTGAAAAGAAAACCTTGAACATGAGGAAGATTTCTGC
 AAAAGTTATTTCCGGTGTTTTGAAGGGTGAGCCTAAACCACGAAAACCTTGACTTTCCGGTCCGA
 TCTTTCAAAGGAACTAGGAAAAATAGCTCATGTTTGAGGAAAAAGGTAACAGGTTCTGAAAC
 ATGGAGTTATCTCCAGGGTGAAGCTGGTGGGGAAATGCCCTTGCCGGTATCCCATCCCAGAGT
 CCACTACTCTGCCAGTCAGCTTCTGCAGGCGTCATCTTCAACAAGCCTTCCCCCAGGGTAGC
 TGAGAATTGGTTTACCCCATGGTGAGGATCCTGA3'
 BamHI

(A)



(B)

Figure 2 - Schematic representation of the expression vector. Schematic representation of the expression vector with the restriction sites for the endonucleases (*XbaI*, *EcoRI*, *BamHI* and *HindIII*) selection gene, origin of replication derived from plasmid pMB1 (rep pMB1), sites of thrombin and Strep-tagII.

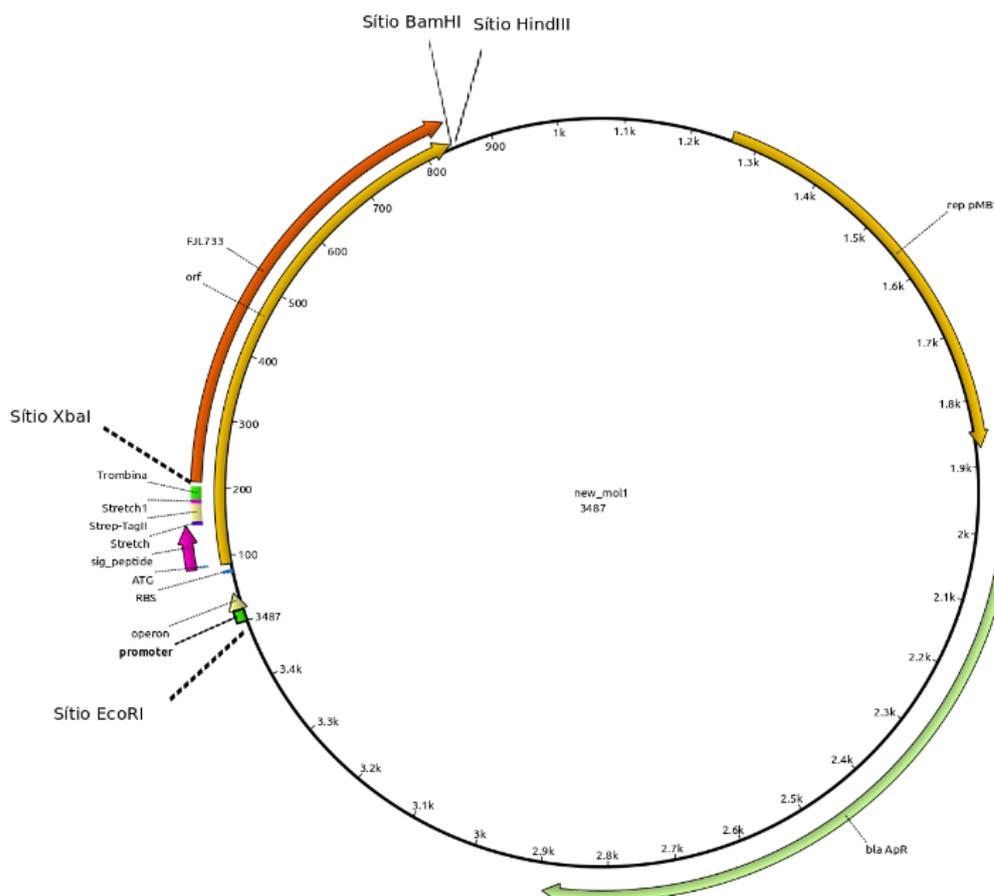


Figure 3 - Electrophoresis 1% (w / v) agarose of construct. Lane 1 corresponds to the molecular weight marker described in detail on the right side of the image; the lanes 1 through 2 correspond to the amplification of the expression construct with the corresponding primers, totaling 852 bp.

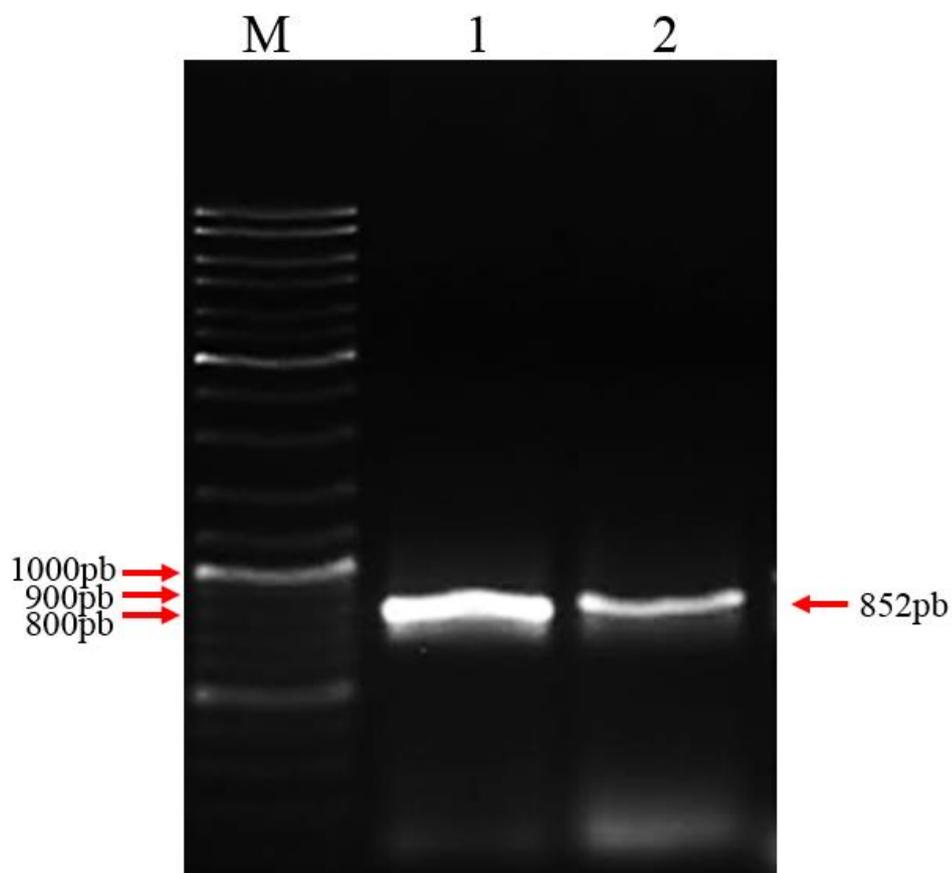


Figure 4- Electrophoresis 1% (w / v) agarose of pUC19 linearized. The lane 1 of the gel corresponds to the molecular weight marker described in detail on the right side of the image; the next lane corresponds to the extracted pUC19 vector cleaved with the BamHI restriction endonuclease, for linearization.

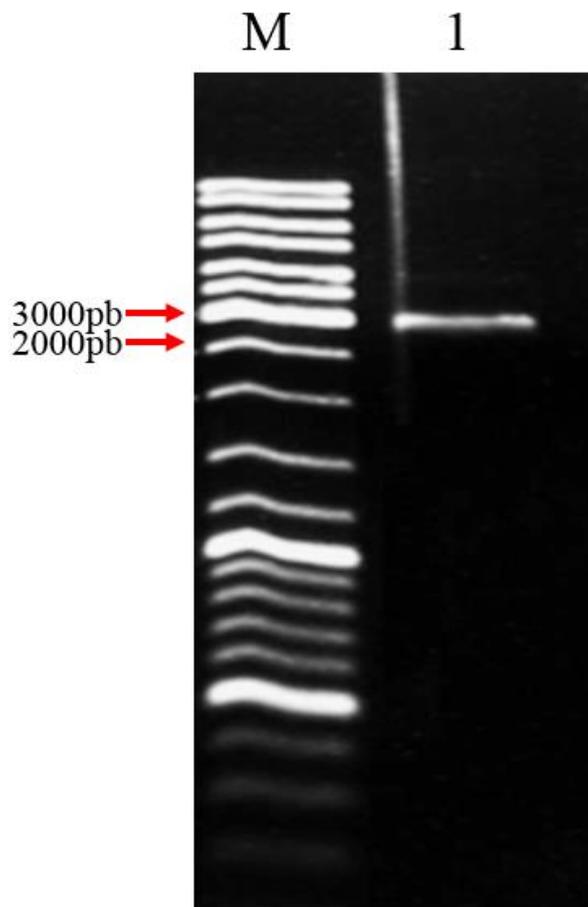
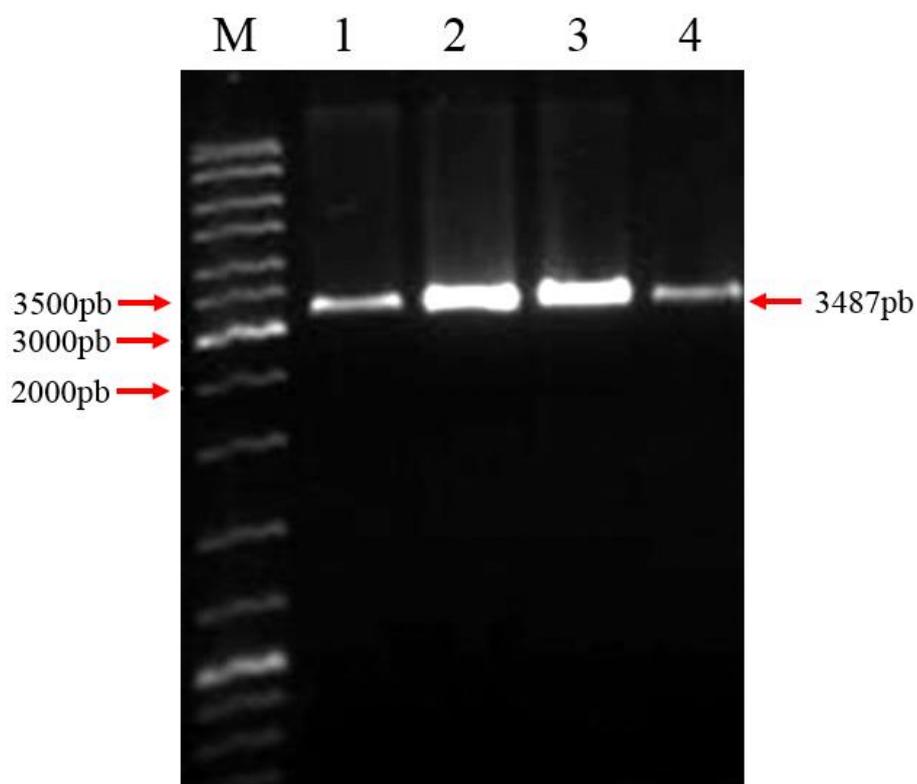


Figure 5 - Expression vector – Expression vector (pUC19 + insert) clivated with the BamHI endonuclease and precipitated with PEG13%. The vector used to receive the insert was pUC19 with 2686 base pairs, which will total after the addition of the insert 3538pb



6 CONCLUSÃO GERAL

Nossos resultados demonstraram que todas as cepas de *E. coli* BL21 (DE3) pLysS, *E. coli* BL21-CodonPlus (DE3) –RP; *E. coli* BL21-CodonPlus (DE3) –RIL; *E. coli* BL21 (DE3) pT-GROE apresentaram mecanismos eficientes na expressão heteróloga da transposase do elemento *mos1*. Todavia a cepa *E. coli* BL21 (DE3) pT-GROE obteve um menor rendimento na expressão, independente da temperatura de indução ou da concentração de IPTG utilizado. Além desta diferença significativa da expressão, as diferentes concentrações de IPTG não foram capazes de modificar o nível de expressão heteróloga da proteína de todas as linhagens, porém quando induzidas a uma temperatura de 37°C as expressões apresentaram visualmente uma expressão “melhor” se comparadas as de temperatura 30°C. As análises de solubilidade, visivelmente não apresentaram diferenças entre as cepas testadas em ambos fatores, temperatura e concentração de indutor IPTG. Estas análises permitem novos trabalhos a fim de melhor elucidar as funções, a estrutura molecular e o mecanismo de transposição da transposase deste TE. Tendo em vista este resultado de proteínas insolúveis, mesmo na maquinaria de BL21 (DE3) pT-GROE, que possui um mecanismo “ideal” para o desdobramento correto das proteínas, abre-se o debate da necessidade de ferramentas capazes de fornecer proteínas solúveis de forma eficiente.

Desta forma infere-se que este processo de expressão pode ser otimizado pelo uso do vetor de expressão construído neste presente estudo, que não necessita de subclonagens, utilizando da própria bactéria para expressão, ou seja, lançando mão do processo de clonagem para extração do plasmídeo e as demais etapas para adição no vetor de expressão, de maneira a maximizar o procedimento. Este vetor, portanto, um promotor lac responsivo a IPTG, seguido logo de um sinal de exportação que permita a exportação da proteína, um fragmento Strep-tagII o qual irá permitir a purificação das proteínas de forma mais eficaz. Além de possuir um sítio de trombina e dois sítios de clivagem que permitem a clivagem do gene inserido no constructo, tornando-o uma ferramenta versátil para diversos trabalhos. Além de permitir a obtenção das proteínas em conformação ativa e solúvel.

Frente a isto, pretende-se desenvolver novos testes com o vetor de expressão desenhado a fim de comprovar sua funcionalidade e testá-lo, analisar a transposase do elemento *mos1* com essa nova ferramenta. Com isso, permitindo o uso desta nova tecnologia para outros trabalhos de produção e análise de proteínas.

REFERÊNCIAS

- Adrio, J., Demain, A.L. **Recombinant organisms for production of industrial products**. Landes Bioscience. Institute for Scientists Emeriti, Madison, United States of America vol. 1(2), pp.116–131. 2010
- Alberts, B.; Johnson, A.; Raff, M.; Roberts, K.; Walter, P. **Biologia molecular da Célula**. Ed. Artmed. São Paulo. 2010.
- Arkhipova, I.R. & Meselson, M. **Diverse DNA transposons in rotifers of the class Bdelloidea**. PNAS. Harvard University, Cambridge. 102(33).2005.
- Auge, C., Brillet, B. & Bigot, Y. **Assembly of the mariner *mos1* Synaptic Complex**. Molecular and Cellular Biology. Tours, France. 25(7), pp.2861–2870. 2005.
- Bächli, G. **Taxodros: The database on taxonomy of Drosophilidae**, v1, 03. Database 2008/01. Disponível em < <http://www.taxodros.uzh.ch/>>. Acesso em: 8 dezembro. 2009.
- Biémont, C.; Vieira, C.,. **Junk DNA as an evolutionary force**. News & views feature. Nature. Villeurbanne Cedex, France. ed. 443(October), pp.521–524. 2006.
- Bire, S.; Casteret, S.; Arnaoty, A.; Piégu, B.; Lecomte, T.; Bigot, Y. **Transposase concentration controls transposition activity : Myth or reality ?** Gene. Tours, France. pp.165–171.2013.
- Böhne, A., Brunet, F.; Galiana A., Delphine S.; C.; Volff, J. **Transposable elements as drivers of genomic and biological diversity in vertebrates**. Chromosome Research. France. 16:203Y215 DOI: 10.1007/s10577-007-1202-6.2008
- BROWN, T. A., **Clonagem gênica e análise de DNA**. 4ª edição, Editora Artmed, Porto Alegre: 2003. p. 376
- Bulla, L. M. C.; Polonio, J. C.; Bredow, C.; Sacco, V. A.; Rhoden, S. A.; Azevedo, J. L.; Pamphile, J. A.; **Elementos transponíveis microbianos e a etiquetagem gênica**. Revista UNINGÁ Review. Maringá. 18, pp.22–27.2014
- Capy, E., David, J.R.I. & Hartl, D.L. **Evolution of the transposable element mariner in the Drosophila melanogaster species group**. Genetica. Washington. pp.37–46.1992

Carvalho, H. & Meneghini, R. **Increased expression and purification of soluble iron-regulatory protein 1 from *Escherichia coli* co-expressing chaperonins GroES and GroEL.** Brazilian Journal of Medical and Biological Research. São Paulo. pp.270–276.2008

Cereghino, J.L. & Cregg, J.M. **Heterologous protein expression in the methylophilic yeast *Pichia pastoris*.** FEMS Microbiology Reviews. Oregon. 2000.

Chang, W. et al., **Cloning , expression , purification , and characterization of zebrafish cytosolic serine hydroxymethyltransferase.** Protein Expression and Purification. Taiwan. 46, pp.212–220.2006

Clark, A.G.; Eisen, M.B.; Smith, D.R.; Bergman, C.M.; Oliver, B.; Markow, T.A.; Kaufman, T.C.; Kellis, M; Gelbart, W. **Evolution of genes and genomes on the *Drosophila* phylogeny.** Nature. v. 450, p. 203-218. 2007.

Crénés, G.; Moundras, C.; Demattei, M.; Bigot, Y.; Petit, A.; Renault, S., **Target site selection by the mariner -like element , *mos1*.** Genetica. France. pp.509–517.2010.

Davison, J. **Review Genetic Tools for Pseudomonads , Rhizobia , and Other Gram-Negative Bacteria.** BioTechniques. France. 401(February).2010

Dawson, A. & Finnegan, D.J. **Excision of the *Drosophila* Mariner Transposon *mos1* : Comparison with Bacterial Transposition and V (D) J Recombination.** Molecular Cell. 11(D), pp.225–235. Edinburgh. 2003.

Demain, A.L. & Vaishnav, P. **Production of recombinant proteins by microbes and higher organisms.** Biotechnology Advances. India. 27(3), pp.297–306.2009

Dumon-seignovert, L., Cariot, G. & Vuillard, L.,. **The toxicity of recombinant proteins in *Escherichia coli* : a comparison of overexpression in BL21 (DE3), C41 (DE3), and C43 (DE3).** Elsevier - Protein Expression & Purification. France. 37, pp.203–206. 2004

Durfee, T.; Nelson, R.; Baldwin, S.; Iii, G P.; Burland, V.; Mau, B.; Petrosino, J.F; Qin, X.; Muzny, D. M.; Ayele, M.; Gibbs, R. A.; Weinstock, G. M.; Blattner, F. R. **The Complete Genome Sequence of *Escherichia coli* DH10B : Insights into the**

Biology of a Laboratory Workhorse. Journal of Bacteriology. Hungary 190(7), pp.2597–2606. 2008.

Einsfeldt, K.; Baptista, J.; Júnior, S.; Paula, A.; Argondizzo, C.; Alberto, M.; Lívio, T.; Alves, M.; Volcan, R.; Leites, A. **Cloning and expression of protease ClpP from *Streptococcus pneumoniae* in *Escherichia coli*: Study of the influence of kanamycin and IPTG concentration on cell growth, recombinant protein production and plasmid stability.** Vaccine. Rio de Janeiro. 29(41), pp.7136–7143. 2011.

Feschotte, C.; Pritham, E.J. **DNA transposons and the evolution of eukaryotic genomes.** Annual review of genetics. Texas. 41, pp.331–68.2007

Filée, J.; Rouault, J.; Harry, M.; Hua-van, A.; **Mariner transposons are sailing in the genome of the blood-sucking bug *Rhodnius prolixus*.** BMC Genomics. France. 2015.

Gabrielczyk, J., Kluitmann, J. & Dammeyer, T. **Protein Expression and Purification Effects of ionic strength on inclusion body refolding at high concentration.** Elsevier Protein Expression and Purification. Germany. 130, pp.100–106. 2017

Gustafsson, C., Govindarajan, S. & Minshull, J. **Codon bias and heterologous protein expression.** Trends in Biotechnology. California, United States of America. 22(7), pp.346–353.2004

Herrmann, S.; Ma, Q.; Johnson, M. S.; Repik, A. V.; Taylor, B. **PAS Domain of the Aer Redox Sensor Requires C-Terminal Residues for Native-Fold Formation and Flavin Adenine Dinucleotide Binding.** Journal of bacteriology. California. 186(20), pp.6782–6791. 2004.

Hickman, A. B.; Dyda, F. **Mechanisms of DNA Transposition.** Microbiology Spectrum. United States of America. p. 1–22, 2015.

Ishii, S. & Yura, T. **Method for producing a soluble protein with bacteria.** European Patent Office. Application number: 96116359.9. Date of publication and mention of the grant of the patent: 26.03.2003 Bulletin 2003/13. 2003

Jacobson, J.W., Medhora, M.M. & Hartl, D.L. **Structure of a Somaticallly Unstable Transposable Element in *Drosophila*.** National Academy of Sciences. United States of America. 83(22), pp.8684–8688. 2014

Jana, S.; Deb, J.K., **Strategies for efficient production of heterologous proteins in *Escherichia coli***. Applied Microbiology and Biotechnology. India. pp.289–298.2005

Jeong, H., Kim, J. & Lee, J., **Complete Genome Sequence of *Escherichia coli* Strain BL21**. GenomeA. Korea. 2015.

Jonasson, P.; Liljeqvist, S.; Nygren, P. A.; Stahl, S.; Ståhl, S. **Genetic design for facilitated production and recovery of recombinant proteins in *Escherichia coli***. Biotechnol Appl Biochem. Sweden. 35(Pt 2), p.91–105. 2002.

Jong Hyun Choi, Keum, K.C. & Lee, S.Y. **Production of recombinant proteins by high cell density culture of *Escherichia coli***. Chemical Engineering Science. Korea. 61, pp.876–885. 2006

Kapitonov, V. V & Jurka, J. **Rolling-circle transposons in eukaryotes**. PNAS. California. vol. 98.n.15. 2001.

Kazazian, H.H. **Mobile elements and disease**. Current Opinion in Genetics & Development. Philadelphia. ed.8, pp.343–350. 1998.

Kidwell, M.G. & Lisch, D.R. **Perspective: transposable elements, parasitic dna, and genome evolution**. BioOne Research Evolved. Arizona. 55(1), pp.1–24. 2001.

Kim, J.; Kim, S.; I.; Hong, E.; Ryu, Y. **Strategies for increasing heterologous expression of a thermostable esterase from *Archaeoglobus fulgidus* in *Escherichia coli***. Protein Expression and Purification. Korea. 127, pp.98–104. 2016

Korpimäki, T., Kurittu, J. & Karp, M. **Surprisingly fast disappearance of β -lactam selection pressure in cultivation as detected with novel biosensing approaches**. Journal of Microbiological Methods. Finland. 53(1), pp.37–42.2003.

Larentis, A. L.; Argondizzo, A. P. C.; Esteves, G. S.; Jessouron, E.; Galler, R.; Medeiros, M. A. **Cloning and optimization of induction conditions for mature PsaA (*pneumococcal surface adhesin A*) expression in *Escherichia coli* and recombinant protein stability during long-term storage**. Protein Expression and Purification. Rio de Janeiro. 78(1), pp.38–47. 2011.

Liddy, N.; Molloy, P. E.; Bennett, A. D.; Boulter, J. M.; Jakobsen, B. K. **Production of a Soluble Disulfide Bond-Linked TCR in the Cytoplasm of *Escherichia coli* *trxB* Mutants.** Mol Biotechnol. Reino Unido. pp.140–149. 2010

Long, X.; Gou, Y.; Luo, M.; Zhang, S.; Zhang, H.; Bai, L. **Soluble expression, purification, and characterization of active recombinant human tissue plasminogen activator by auto-induction in *E. coli*.** BMC Biotechnology. China. pp.1–9.2015.

Madigan, M.; Martinko, J.; Bender, K.; Buckley, D.; Stahl, D. **Brock biology of microorganisms.** Porto Alegre. 14ª Editora Artmed, ed.14. 2016

Makrides, S.C. **Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*.** Microbiological reviews. Massachusetts. 60(3), pp.512–538. 1996

Mamoun, C. B.; Gluzman, I. Y.; Beverley, S. M.; Goldberg, D. E. **Transposition of the *Drosophila* element mariner within the human malaria parasite *Plasmodium falciparum*.** Elsevier - molecular and Biochemical parasitology> United States of America. pp.405–407. 2000

Mergulhão, F.J.M.; Summers, D.K.; Monteiro, G.A. **Recombinant protein secretion in *Escherichia coli*.** ELSEVIER - Biotechnology Advances. United Kingdom 23, pp.177–202.2005

Miroux, B.; Walker, J.E.. **Over-production of Proteins in *Escherichia coli* : Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels.** Journal of Molecular Biology. United Kingdom. pp.289–298. 1996.

Miyake, O.; Hashimoto, W.; Murata, K. **An exotype alginate lyase in *Sphingomonas sp. A1*: overexpression in *Escherichia coli*, purification, and characterization of alginate lyase IV (A1-IV).** Protein Expression and Purification. Japan. 29, pp.33–41.2003

Miyamoto, C.A. **Produção de Proteínas Recombinantes em *Escherichia coli*.** Revista Conexão Eletrônica. São Paulo. pp.220–233. 2013

Nascimento, A. A. C.; Espreafico, E. M.; Larson, M. L. P.; Monesi, N.; Rossi, N. M. M.; Rodrigues, V.i **Tecnologia do DNA recombinante.** Universidade de São Paulo Faculdade de Medicina de Ribeirão Preto tecnologia. São Paulo. pp.1–51. Disponível em: <http://biocelfmrp.com.br/sites/default/files/apostilatd_2005.pdf. 2003.> Acessado em: 20 de dezembro de 2014. 2003.

Nowacki, M.; Higgins, B. P.; Maquilan, G. M.; Swart, E. C.; Doak, T. G.; Landweber, L. F. **A functional role for transposases in a large eukaryotic genome.** Science. New York, N.Y. 324(1996), pp.935–938. 2009

Olaofe, O. A.; Burton, S. G.; Cowan, D. A.; Harrison, S. T L **Improving the production of a thermostable amidase through optimising IPTG induction in a highly dense culture of recombinant *Escherichia coli*.** Biochemical Engineering Journal. South Africa. 52(1), pp.19–24. 2010

Picardeau, M. **Transposition of fly mariner elements into bacteria as a genetic tool for mutagenesis.** Genetica. France. pp.551–558.2010

Pritham, E.J. **Transposable elements and factors influencing their success in eukaryotes.** Journal of Heredity. Texas. 100(5), pp.648–655.2009.

Pritham, E.J., Putliwala, T. & Feschotte, C. **Mavericks , a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses.** Elsevier - Gene Evolutionary genomics. Texas. 390, pp.3–17. 2007.

Quan, J. & Tian, J. **Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways.** Plos One. United States of America.2009.

Reed, J.L. **Model driven analysis of *Escherichia coli* metabolism.** Electronic Theses and Dissertations UC San Diego. University of California. Disponivel em: <<http://escholarship.org/uc/item/4qn471t6>>. Acessado em: 10 de Janeiro de 2017. 2005.

Reys, L.F., Macedo, J.N.A. & Damalio, J.C.P. **Dogma Central da Biologia Molecular e Introdução à Bioinformática.** W Educacional Editora e Cursos Ltda. Brasilia. 2011

Robinson, M.; Lilley, R.; Little, S.; Emtage, J.S.; Yarranton, G.; Stephens, P.; Millican A.; Eaton, M.; Humphreys, G. **Codon usage can affect efficiency of translation of genes In *Escherichia coli*.** Nucleic Acids Research. United Kingdon. 12(17), pp.6663–6671.1984.

Rosano, G.L.; Ceccarelli, E.A. **Recombinant protein expression in *Escherichia coli*: Advances and challenges.** Frontiers in Microbiology. Argentina. 5(APR), pp.1–17. 2014

Sambrook, J. & Russel, D, W. **Molecular Cloning: A Laboratory Manual**. Ed. CSHL PRESS New York.Vo3, p.999. 2010

Schmidt, F.R.**Recombinant expression systems in the pharmaceutical industry**. Application Microbiol Biotechnology. Germany. pp.363–372.2004

Schmidt, T.G.M. & Skerra, A. **The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins**. Nature protocols. Germany. 2(6), pp.1528–1535. 2007

Sezonov, G., Joseleau-Petit, D. & D'Ari, R. **Escherichia coli physiology in Luria-Bertani broth**. Journal of Bacteriology. France. 189(23), pp.8746–8749. 2007.

Skerra, B.A.; Schmidt, T.O.M., **Use of the Strep-Tag and Streptavidin for detection and purification of recombinant proteins**. Methods in Enzymology. Vol. 36, 2000

Sinzelle, L.; Izsvák, Z.; Ivics, Z. **Molecular domestication of transposable elements: From detrimental parasites to useful host genes**. Celular and Molecular Life Science. Hungary. 2008

Sörensen, H.P.; Mortensen, K.K.**Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli**. Microbial Cell Factories. Denmark., pp.1–8. 2005.

Stano, N.M.;Patel, S.S.**T7 Lysozyme Represses T7 RNA Polymerase Transcription by Destabilizing the Open Complex during Initiation**. Journal of Biological Chemistry. New Jersey. 279(16), pp.16136–16143.2004.

Stosch, M. Von.,**Toward Intensifying Design of Experiments in Upstream Bioprocess Development : An Industrial Escherichia coli Feasibility Study**. American Institute of Chemical Engineers. Canadá. DOI 10.1002/btpr.2295 2016.

Tatum, E.L.; Lederberg, J. **Gene recombination in the bacterium**. Journal of Bacteriology. Connecticut. 63(6), pp.673–684. 1947.

Tellier, M., Bouuaert, C.C.; Ronald, C. **Mariner and the ITm Superfamily of Transposons**. Microbiology Spectrum American Society for Microbiology Press.

United Kingdom. pp.1–19.2014.

Terpe, K. **Overview of bacterial expression systems for heterologous protein production : from molecular and biochemical fundamentals to commercial systems.** Appl Microbiol Biotechnol. Germany. pp.211–222. 2006.

Thangadurai, C.; Suthakaran, P.; Barfal, P.; Anandaraj, B.; Pradhan, S. N.; Boneya, H. K. **Biochemical and Biophysical Research Communications Rare codon priority and its position specificity at the 5' of the gene modulates heterologous protein expression in *Escherichia coli*.** Biochemical and Biophysical Research Communications. India. 376(4), pp.647–652. 2008.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** Nucleic Acids Res. 25: 4876-4882.1997

Tosi, L.R.O.; Beverley, S.M. **Cis and trans factors affecting *mos1* mariner evolution and transposition in vitro , and its potential for functional genomics.** Nucleic Acids Research. Washington. 28(3), pp.784–790.2000.

Vaz, M.R.F. **Estudo do cultivo de dois clones de *Escherichia coli* recombinantes (eIF, LACK) para a expressão de antígenos da *Leishmania Chagasi*.** Dissertação de Mestrado. Universidade Federal do Rio Grande do Norte - Departamento de Engenharia Química. Natal. 2008.

Venner, S., Feschotte, C.; Biéumont, C. **Transposable elements dynamics: toward a community ecology of the genome.** NIH Public Access. National Institutes of Health. France. 25(7), pp.317–323.2010.

Wallau, L. **História evolutiva de elementos transponíveis da superfamília tc1 - mariner em *Drosophilídeos*.** DISSERTAÇÃO DE MESTRADO. Universidade Federal de Santa Maria. Santa Maria. pp.1–85. 2010

Weeks, S.D., Drinker, M.; Loll, P.J. **Ligation independent cloning vectors for expression of SUMO fusions.** Protein, Expression and Purification. Philadelphia. , 53, pp.40–50.2007.

Lima W. N. J. **Produção de proteínas recombinantes utilizando *Escherichia coli* em cultivos em alta densidade celular.** Tese de Doutorado. Universidade estadual de campinas. Campinas. 2004

Lodish H, Berk A, Zipursky SL. **Molecular Cell Biology.** Bacterial Gene Control: The

Jacob-Monod Model. 4th edition. New York. 2000.

Williamson, P. T.; Roth, J. F.; Haddingham, T.; Watts, A. **Expression and purification of recombinant neurotensin in *Escherichia coli***. Protein expression and purification. United Kingdom. 19(2), pp.271–5.. 2000.

Yasukawa, T.; Kanei-Ishii, C.; Maekawa, T.; Fujimoto, J.; Yamamoto, T.; Ishii, S. **Increase of Solubility of Foreign Proteins in *Escherichia coli* by Coproduction of the Bacterial Thioredoxin ***. The journal of biological chemistry. Japan. 1995.

Yuan, Y. & Wessler, S.R. **The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies**. PNAS. Washington. 2011.

Yeates, D.K.; Wiegmann, B.M. **Congruence and controversy: Towards a higher-level phylogeny of Diptera**. Annu. Rev. Entomol., v. 44, p. 397-428. 1999.

Zaha, A.; Schrank, A.; Ferreira, H.B.; Schrank, I.S.; Rodrigues, J.J.S.; Regner, L.P.; Passaglia, L.M.P.; Rossetti, M.L.R.; Raupp, R.M.; Silva, S.C.; Gaiesky, V.L.V. **Biologia Molecular Básica**. Porto Alegre: Mercado Aberto,. 336p. 2003.