

**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO**

PATRÍCIA REGINA MOROMISATO DE OLIVEIRA

**CLONAGEM, EXPRESSÃO E PURIFICAÇÃO
DAS PROTEÍNAS RECOMBINANTES GP120 E
GP41 DO HIV-1**

**RIBEIRÃO PRETO
NOVEMBRO - 2018**

PATRÍCIA REGINA MOROMISATO DE OLIVEIRA

CLONAGEM, EXPRESSÃO E PURIFICAÇÃO DAS PROTEÍNAS RECOMBINANTES GP120 E GP41 DO HIV-1

Trabalho apresentado como requisito para a conclusão do curso de Bacharelado em Ciências Biomédicas – Ênfase em Biotecnologia em Saúde ministrado pela Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo.

Orientadora: Elisa Maria de Sousa Russo

**RIBEIRÃO PRETO
NOVEMBRO - 2018**

A minha mãe,
que nunca hesitou em me ajudar, me apoiar,
e trabalhar para dar o melhor para os filhos.

AGRADECIMENTOS

Agradeço primeiramente a Deus por ter me dado força e coragem para terminar esse curso. Agradeço a minha família, principalmente minha mãe, que sempre esteve me apoiando e me ajudando em tudo que precisei. Agradeço também meus irmãos, Luís e Vinícius e meu namorado Ricardo por ouvirem meus lamentos, minhas alegrias, reclamações, enfim, obrigada por confiarem em mim e me apoiarem. Sem vocês eu não chegaria ao fim! Agradeço imensamente a orientadora mais maravilhosa que pude ter, prof^a Dr^a Elisa Russo, que sempre teve paciência e compreensão. Agradeço as conversas pessoais, dicas e conselhos. Muito obrigada pelos ensinamentos!

A todo o pessoal do laboratório, prof^a Cleni, Giovanna, Natália, Luiz, Gabriel, Carlos e Alana o meu muito obrigada! Obrigada pelos momentos compartilhados, por toda ajuda com os experimentos (que foram muitas), pelas conversas e sorrisos. Todos vocês fizeram a minha ênfase (e minha vida) melhores e mais felizes.

As mentoras incríveis Bruna e Aline, sem vocês eu não saberia o que fazer. Obrigada pelos ensinamentos, pelos conselhos, obrigada por compartilharem comigo todo aprendizado de vocês. Vocês são demais! Não tenho a menor dúvida que vocês serão ótimas pesquisadoras!

Agradeço a equipe do prof^o Hamilton, Rafael, Natália e demais alunos, por estarem sempre dispostos a nos ajudar.

A toda equipe do laboratório do prof^o Sérgio Akira (principalmente ao Emerson), e da prof^a Andréia Leopoldino (principalmente a Cristiane).

Obrigada Nadine e prof^a Angela Kaysel por todo tempo dispensado a mim.

Agradeço também a minha amiga Helena Francelin a quem dividiu comigo os momentos de tristeza e os momentos de alegria ao longo de toda essa graduação. E a veterana mais maravilhosa que pude conhecer, Gláucia Almeida, obrigada por toda a ajuda e por sanar minhas dúvidas.

Por fim, agradeço também a Agência de fomento FAPESP pelo apoio financeiro a este trabalho.

“Em tudo há sempre uma coisa capaz de deixar a gente alegre; a questão é descobri-la.”

Pollyanna

RESUMO

O HIV, Vírus da Imunodeficiência Humana, é o retrovírus causador da Síndrome da Imunodeficiência Adquirida (AIDS). É um vírus linfotrópico com preferência principalmente para linfócitos T CD4+. Ao infectar tais células, o HIV causa imunocomprometimento no indivíduo, o que favorece o surgimento de doenças oportunistas em uma fase mais avançada da infecção. Os primeiros anticorpos a serem produzidos na infecção por HIV são da classe IgM contra a proteína gp41, produzidos cerca de 13 dias após a detecção do RNA viral no plasma. Outra proteína que é um importante alvo de anticorpos é a glicoproteína gp120, detectada a partir de 28º dia após a detecção de RNA viral. Segundo a UNAIDS, estima-se que em 2016 havia cerca de 36,7 milhões de pessoas vivendo com HIV/AIDS no mundo, sendo no Brasil, no mesmo ano, aproximadamente 830.000. Sendo assim, a infecção por HIV é vista hoje como um dos grandes problemas de saúde pública, principalmente porque sua incidência vem aumentando nos últimos anos. O diagnóstico precoce para infecção por HIV é fundamental para que ocorra o tratamento dos indivíduos e o controle da epidemia, diminuindo assim as chances de disseminação do vírus. Dessa forma, é importante que os testes diagnósticos rápidos sejam continuamente melhorados a fim de que sejam amplamente disponíveis. Em vista disso, o objetivo do presente estudo foi produzir as proteínas recombinantes gp41 e gp120 do HIV-1 através da tecnologia do DNA recombinante para serem empregadas futuramente em teste diagnóstico. Inicialmente foi escolhido o sistema de expressão procarioto *E. coli*. Porém, não obteve-se sucesso com a expressão das proteínas em BL21(DE3)pLYSs, pois foi observado uma grande quantidade de códons raros nas sequências. Desse modo, foi trocada a linhagem bacteriana para Rosetta(DE3), pois esta linhagem codifica os códons raros. Entretanto, também não foi possível obter a expressão das proteínas. Como alternativa, foi trocado o sistema de expressão procarioto para o sistema eucarioto *Pichia pastoris*. O gene das proteínas de interesse foram clonados no vetor pPICZαA e foi feito o teste de expressão em menor escala (*screening*) para identificação da colônia mais produtora. Após a indução em menor escala foi realizado a purificação do sobrenadante de uma colônia escolhida para cada proteína em cromatografia de afinidade a níquel. Por fim, as proteínas foram identificadas por *dot blotting*. Com os resultados obtidos até o momento a produção das proteínas gp41 e gp120 do HIV-1 mostra-se promissora em *Pichia pastoris*.

Palavras-chave: HIV-1, gp41, gp120, proteínas recombinantes, *E. coli*, *Pichia pastoris*

ABSTRACT

HIV, the Human Immunodeficiency Virus, is the retrovirus that causes Acquired Immunodeficiency Syndrome (AIDS). It is a lymphotropic virus with preference mainly for CD4 + T lymphocytes. By infecting such cells, HIV causes immunocompromising, which favors the emergence of opportunistic diseases at a later stage of infection. The first antibodies produced in HIV infection are of the IgM class against the gp41 protein, produced about 13 days after viral RNA detection in plasma. Another protein that is an important target for antibodies is glycoprotein gp120, detected from day 28 after detection of viral RNA. According to UNAIDS, it is estimated that in 2016 there were about 36.7 million people living with HIV / AIDS in the world, and in Brazil, in the same year, approximately 830,000. Thus, HIV infection is today seen as one of the major public health problems, mainly because its incidence has been increasing in recent years. Early diagnosis for HIV infection is essential for treatment of individuals and control of the epidemic, thus reducing the chances of spreading the virus. Therefore, it is important that rapid diagnostic tests are continuously improved so that they are widely available. In view of this, the aim of the present study was to produce recombinant HIV-1 gp41 and gp120 proteins by recombinant DNA technology to be employed in the diagnostic test in the future. Initially, the *E. coli* prokaryotic expression system was chosen. For this system the proteins were cloned into the vector pET28a. However, there was no success with the expression of the proteins in *BL21 (DE3) pLYSs*, since a large number of rare codons were observed in the sequences. Thus, the bacterial line was changed to *Rosetta (DE3)*, since this lineage codifies the rare codons. However, the expression of the proteins could not be obtained either. As an alternative, the prokaryotic expression system was changed to the *Pichia pastoris* eukaryote system. The gene of the proteins of interest were cloned into the pPICZαA vector and the small scale expression test was performed to identify the most productive colony. After the small scale induction, purification of the supernatant of a colony chosen for each protein in nickel affinity chromatography was performed. Finally, the proteins were identified by dot blotting. With the results obtained so far the production of HIV-1 gp41 and gp120 proteins is shown to be promising in *Pichia pastoris*.

Key words: HIV-1, gp41, gp120, recombinant proteins, *E. coli*, *Pichia pastoris*

LISTA DE FIGURAS

Figura 1 - Estrutura do HIV-1	4
Figura 2 - Organização do genoma do HIV-1	5
Figura 3 - Domínios das proteínas gp120 e gp41	6
Figura 4 - Mecanismo de entrada do HIV-1 na célula hospedeira	7
Figura 5 - Ciclo de replicação do HIV-1	8
Figura 6 - Dogma central da biologia	13
Figura 7 - Fluxograma da metodologia utilizada.....	20
Figura 8 - Mapa do vetor pAbT4674 (ATCC)	21
Figura 9 - Mapa do vetor pET28a (+) (Novagen)	24
Figura 10 - Mapa do vetor pPICZαA (Invitrogen)	25
Figura 11 - Gel de agarose 1% mostrando a amplificação das sequências gp41 e gp120 a partir do vetor pAbT4674 com os primers da tabela 1.....	39
Figura 12 - Gel agarose 1% mostrando a amplificação de DNA proveniente de colônias recombinantes que receberam o gene da gp41.....	40
Figura 13 - Gel agarose 1% mostrando a amplificação de DNA proveniente de colônias recombinantes que receberam o gene da gp120.....	40
Figura 14 - Análise dos aminoácidos da gp41 recombinante.....	41
Figura 15 - Gel de poliacrilamida 12% da indução de 4 horas da BL21(DE3)pLYSs_ pET28a _gp41 a 37°C.....	43
Figura 16 - Gel de poliacrilamida 12% da indução de 4 horas da BL21(DE3)pLYSs_ pET28a _gp120 a 37°	43
Figura 17 - Localização dos códons raros na gp41	44
Figura 18 - Localização dos códons raros na gp120.....	46
Figura 19 - Gel agarose 1% mostrando PCR de colônia para confirmação da transformação dos plasmídeos pET28a_gp41 e pET28a_gp120 em Rosetta(DE3).....	48
Figura 20 - Gel de poliacrilamida 12% da indução de 4 horas da Rosetta(DE3)_pET28a _gp41 a 37°	49
Figura 21 - Gel de poliacrilamida 12% da indução de 4 horas da Rosetta(DE3)_pET28a_ gp120 a 37°C	50
Figura 22 - Gel de agarose 1% mostrando a amplificação das sequências gp41 e gp120 a partir do vetor pAbT4674 com os primers da tabela 2	52
Figura 23 - Gel de agarose 1% mostrando a reação de restrição com as enzimas EcoRI e Sall	52
Figura 24 - Gel de poliacrilamida 12% do Screening das colônias gp41.....	54
Figura 25 - Gel de poliacrilamida 12% do Screening das colônias gp120.....	54
Figura 26 - Gel de poliacrilamida 12% da purificação por cromatografia de afinidade gp41.....	55
Figura 27 - Gráfico da purificação da proteína recombinante gp41.....	56
Figura 28 - Gel de poliacrilamida 12% da purificação por cromatografia de afinidade gp120.....	56
Figura 29 - Gráfico da purificação da proteína recombinante gp120.....	57
Figura 30 - <i>Dot Blotting</i> do screening das proteínas recombinantes gp41 e gp120.....	58

LISTA DE TABELAS

Tabela 1 - Sequência dos primers utilizados na clonagem dos genes das proteínas gp41 e gp120 em vetor pET28a	22
Tabela 2 - Sequência dos primers utilizados na clonagem dos genes das proteínas gp41 e gp120 em vetor pPICZαA	22
Tabela 3 - Diluição dos anticorpos utilizados no <i>Western Blotting</i>	36
Tabela 4 - Diluição dos anticorpos utilizados no <i>Dot Blotting</i>	37
Tabela 5 - Identidade da proteína gp41.....	41
Tabela 6 - Frequência dos códons raros na sequência gp41.....	45
Tabela 7 - Frequência dos códons raros na sequência gp120.....	46
Tabela 8 - Estudo de códons raros.....	47
Tabela 9 - Identidade da proteína gp41.....	53

LISTA DE ABREVIATURAS E SIGLAS

AIDS/SIDA - Síndrome da imunodeficiência adquirida

AP - Fosfatase alcalina

CDC - *Centers for Disease Control*

CRF - *Circulating recombinant form*

DNA - Ácido Desoxirribonucleico

dNTPs - Desoxiribonucleotídeos

DTT - Ditioneitol

EDTA - Ácido etileno-diamino tetra-acético

ELISA - *Enzyme-Linked Immunosorbent Assay*

g - grama

HAART - *Highly active antiretroviral therapy*

HEPES - ácido N- (2-hidroxietil) piperazina-N`-2- etanossulfônico

HIV-1 - Vírus da Imunodeficiência Humana do tipo 1

HIV-2 - Vírus da Imunodeficiência Humana do tipo 2

HTLV-III - *Human T-Lymphotropic Virus* ou Vírus Linfotrópico da célula T Humana do tipo III

ICTVdB - *International Committee on Taxonomy of Viruses Database*

Ig - Imunoglobulina

IPTG - isopropil- β -D-tiogalactopirósídeo

LAV - *Lymphadenopathy Associated Virus* ou Vírus Associado a Linfadenopatia

LB - Caldo Luria Bertani

M - Molar

MCS - *Multiple cloning site* ou sítio de múltipla clonagem

mg - miligrama

mM - milimolar

nº - número

PCR - Reação em cadeia da polimerase

RBS - sítio de ligação ao ribossomo

RNA - ácido ribonucleico

SDS - Dodecil sulfato de sódio

SDS-PAGE - Eletroforese em gel de poliacrilamida contendo SDS

TEMED - N,N,N',N' - Tetrametiletenodiamina

TAE - Tris-acetato-EDTA

TBS-T - Tampão Salina Tris com Tween 20

UNAIDS - Joint United Nations Program on HIV/AIDS

SUMÁRIO

DEDICATÓRIA	
AGRADECIMENTOS	
EPÍGRAFE	
RESUMO	i
ABSTRACT	ii
LISTA DE FIGURAS	iii
LISTA DE TABELAS	iv
LISTA DE ABREVIATURAS E SIGLAS	v
SUMÁRIO	vii
1 INTRODUÇÃO	1
1.1 Vírus da Imunodeficiência Humana Tipo 1 (HIV-1)	1
1.2 Estrutura e organização genômica	3
1.3 Ciclo de replicação viral	6
1.4 Fases da infecção pelo HIV-1	9
1.5 Diagnóstico	10
1.6 Produção de proteínas recombinantes	12
1.7 Vetores plasmidiais de clonagem e expressão	14
1.8 Sistema de expressão procaríoto	15
1.9 Sistema de expressão eucarioto	17
2 JUSTIFICATIVA	18
3 OBJETIVOS	19
3.1 Objetivo geral	19
3.2 Objetivos específicos	19
4. MATERIAL E MÉTODOS	20
4.1 Fluxograma da metodologia utilizada	20
4.2 Clonagem e expressão	20
4.2.1 Obtenção do material genético do HIV-1	20
4.2.2 Desenho dos oligonucleotídeos iniciadores (primers)	21
4.2.3 Sistemas de expressão	23
4.2.4 Vetores de clonagem e expressão	23
4.2.5 Expansão clonal	26
4.2.6 Extração plasmidial (Mini prep)	26
4.2.7 Reação em cadeia da polimerase (PCR)	27
4.2.8 Eletroforese em gel de agarose	28
4.2.9 Purificação da reação de PCR e quantificação	28
4.2.10 Digestão dos plasmídeos	28

4.2.11 Construção dos vetores pET28a_gp120, pET28a_gp41, pPICZαA_gp41 e pPICZαA_gp120.....	30
4.2.12 Preparo de células de competentes.....	30
4.2.13 Transformação por choque térmico de células competentes.....	31
4.2.14 Transformação por eletroporação em <i>Pichia pastoris</i> competentes.....	32
4.2.15 Seleção das colônias transformadas.....	32
4.2.16 Sequenciamento.....	33
4.2.17 Expressão das proteínas recombinantes gp41 e gp120.....	33
4.2.18 Eletroforese em gel de poliacrilamida	35
4.2.19 Coloração dos géis de poliacrilamida com Coomassie® Brilliant Blue R 250	36
4.2.20 Detecção das proteínas recombinantes por <i>Western blotting</i>	36
4.2.21 Purificação das proteínas recombinantes por cromatografia de afinidade.....	37
4.2.22 Detecção das proteínas recombinantes por <i>Dot Blotting</i>	37
5. RESULTADOS E DISCUSSÃO.....	38
5.1 Da clonagem no sistema de expressão procarioto <i>E. coli</i>	38
5.1.1 Clonagem dos fragmentos gênicos codificantes das proteínas gp41 e gp120 em vetor pET28a.....	38
5.1.2 Análise do Sequenciamento.....	41
5.1.3 Análise da Expressão das proteínas recombinantes gp41 e gp120 em BL21(DE3)pLYSs.....	42
5.1.4 Estudo de códons raros.....	44
5.1.5 Análise da Expressão das proteínas recombinantes gp41 e gp120 em Rosetta(DE3).....	48
5.2 Da clonagem no Sistema de expressão eucarioto <i>P. pastoris</i>	51
5.2.1 Clonagem dos fragmentos gênicos codificantes das proteínas gp41 e gp120 em vetor pPICZαA.....	51
5.2.2 Análise do Sequenciamento.....	53
5.2.3 Análise da Expressão das proteínas recombinantes gp41 e gp120.....	55
5.2.4 Purificação das proteínas recombinantes gp41 e gp120.....	55
5.2.5 Detecção das proteínas recombinantes por <i>Dot Blotting</i>	57
6. CONCLUSÃO.....	58
7. PERSPECTIVAS FUTURAS.....	59
8. REFERÊNCIAS.....	60

1. INTRODUÇÃO

1.1 Vírus da Imunodeficiência Humana Tipo 1 (HIV-1)

O HIV, Vírus da Imunodeficiência Humana ou do inglês, Human Immunodeficiency Virus pertence a classe dos retrovírus, família Retroviridae e subfamília dos lentivírus. É um vírus linfotrópico com preferência para linfócitos T CD4+, mas que também pode infectar outras células como linfócitos T helper, macrófagos e células dendríticas, devido a presença de receptores do tipo CD4 (MANUAL TÉCNICO PARA O DIAGNÓSTICO DA INFECÇÃO PELO HIV, MINISTÉRIO DA SAÚDE, 2014).

Foi identificado em 1981 pela primeira vez em indivíduos que compunham os chamados grupos de risco, que são: homossexuais (GOTTLIEB et al., 1981; MASUR et al., 1981; SIEGAL et al., 1981), hemofílicos (Centers for Disease Control - CDC 1982), usuários de drogas (GOEDERT, 1984) e crianças (RAQNI et al., 1985). Em 1983, Luc Montagnier, na França, isolou pela primeira vez o HIV, denominado até o momento de LAV (Lymphadenopathy Associated Virus ou Vírus Associado a Linfadenopatia). Ainda em 1983, Robert Gallo, nos Estados Unidos, caracterizou o vírus, que era chamado em seu país de HTLV-III (Human T-Lymphotropic Virus ou Vírus Linfotrópico da célula T Humana do tipo III) e que posteriormente foi nomeado para HIV-1 (BARRE-SINOUSSE et al., 1983; POPOVIC, 1983). Foi então que o HIV foi associado como o agente etiológico da Síndrome da Imunodeficiência Adquirida (SIDA, ou em inglês, Acquired Immunodeficiency Syndrome, AIDS), caracterizada por um conjunto de sintomas e pelo comprometimento do sistema imunológico devido a diminuição de linfócitos TCD4. Em 1986, um segundo tipo de HIV foi identificado no Oeste da África, denominado HIV-2, por ser também um retrovírus e por possuir algumas características semelhantes ao HIV-1, como o modo de transmissão, ciclo de replicação e condição clínica (CLAVEL et al., 1986). Enquanto o HIV-1 é o tipo mais difundido em todo o mundo, o HIV-2 é menos prevalente, possui uma menor capacidade de transmissão e é menos patogênico, caracterizando uma

menor probabilidade de progressão para AIDS (JAFFAR et al., 2004; NYAMWEYA et al., 2013).

A nível molecular, estudos filogenéticos demonstram que existem vários subtipos de HIV-1 e 2. O HIV-1 é classificado em três grupos: M (Main), O (Outlier) e N (Non-M, non-O). E o HIV-2 é classificado em grupos de A até H, sendo os mais comuns os subtipos A e B, que circulam principalmente na Costa do Marfim (BUTLER et al., 2007). O grupo M é responsável por 95% das infecções por HIV. Esse grupo é dividido em subgrupos: A, B, C, D, F, G, H, J e K e pelos menos 43 formas recombinantes (circulating recombinant forms - CRFs) que circulam em áreas geográficas distintas. No Brasil, encontra-se predominantemente o HIV-1 do grupo M, subgrupo B e as formas recombinantes B e F1 (MANUAL TÉCNICO PARA O DIAGNÓSTICO DA INFECÇÃO PELO HIV, MINISTÉRIO DA SAÚDE, 2014).

A partir de observações epidemiológicas foram identificadas as vias de transmissão do HIV, que pode ser por via sexual (relações sexuais sem o uso do preservativo) ; via sanguínea (transfusão de sangue ou hemoderivados que possuem o vírus) ; compartilhamento de agulhas e seringas e por transmissão vertical (transmissão mãe/criança durante a gestação, parto ou aleitamento materno) (LEVY et al. 2003).

A partir da metade dos anos de 1990 observou-se que a epidemia assumiu outro perfil que não os grupos de risco citados anteriormente. A principal via de transmissão passou a ser a transmissão heterossexual. Atualmente, há uma tendência do crescimento da infecção em jovens de 15 a 24 anos e em adultos com 50 anos ou mais, tanto em homens quanto em mulheres.

De 1980 a junho de 2016 foram notificados 842.710 casos de AIDS no Brasil, segundo o último boletim epidemiológico HIV/AIDS 2016 emitido pelo Ministério da Saúde em 2017. No ano de 2016, segundo a UNAIDS (Programa Conjunto das Nações Unidas sobre HIV/AIDS), no Brasil, estima-se que tenham ocorrido aproximadamente 48.000 novas infecções pelo HIV e que havia cerca de 830.000 pessoas vivendo com o vírus. O número de mortes relacionadas à AIDS no Brasil em 2015 foi de 15.000. Ainda segundo o Boletim

Epidemiológico HIV/AIDS 2016, nos últimos 10 anos observou-se um aumento da taxa de detecção do HIV, principalmente entre jovens de 15 a 24 anos. Segundo dados divulgados pelo relatório da UNAIDS em 2016, havia cerca de 36,7 milhões de pessoas vivendo com HIV/AIDS no mundo. Estima-se que no mesmo ano houve aproximadamente 2,1 milhões de novas infecções por HIV e que cerca de 1,1 milhão de pessoas morreram em decorrência da AIDS.

Espera-se que até 2030 a epidemia de AIDS seja controlada no mundo, contanto que os indivíduos infectados sejam diagnosticados e tratados precocemente, diminuindo assim a disseminação do HIV. O HIV possui uma alta taxa de mutação, que leva a formação de novas partículas virais recombinantes circulantes (circulating recombinant forms – CRFs) o que pode ocasionar uma baixa sensibilidade ao tratamento anti-retroviral e que pode conduzir para o desenvolvimento da AIDS em pouco tempo (BLACKARD et al., 2002; KILLIAN; LEVY, 2011; KOURI et al., 2015; McCUTCHAN, 2000; PEETERS; SHARP, 2000; ROBERTSON et al., 1995). Nesse contexto, o diagnóstico precoce e desenvolvimento de tratamentos ainda mais eficientes são fundamentais para o controle da epidemia de HIV.

1.2 Estrutura e organização genômica

O HIV-1 é um vírus com material genético constituído de duas fitas simples de RNA e pertence ao grupo de retrovírus citopáticos e não-oncogênicos (ICTVdB, 2015). Seu diâmetro é 100 nm e possui um tamanho de aproximadamente 9 kb. A partícula do HIV está envolvida por um envelope lipídico, composto por uma bicamada fosfolipídica obtida a partir da membrana da célula hospedeira em que ocorreu a replicação viral. Na parte interna do envelope localizam-se as proteínas de matriz p17 ou MA, que revestem todo o capsídeo viral (CA) que é formado pelas proteínas p24 (Figura 1) (FRANKEL; YOUNG, 1998; FREED, 1998). No interior do capsídeo, encontra-se o material genético que está associado às proteínas do nucleocapsídeo (NC ou p7) juntamente com as enzimas protease (p11),

transcriptase reversa (p66) e integrase (p32), que são necessárias para a replicação viral (FREED, 2001).

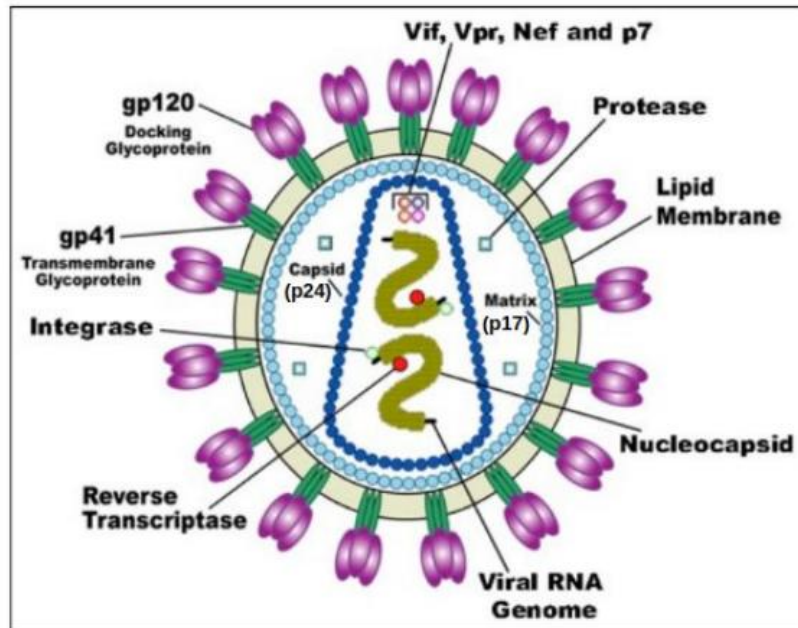


Figura 1 – Estrutura do HIV-1. Fonte:

<https://web.stanford.edu/group/virus/retro/2005gongishmail/HIV.html>

O genoma do HIV-1 é composto por nove genes, sendo eles: gag, pol, vif, vpr, vpu, env, tat, rev e nef (Figura 2). Os genes env e gag possuem a informação necessária para a síntese de proteínas estruturais para as novas partículas virais, enquanto que o restante dos genes codificam proteínas que regulam a transcrição e replicação viral. O gene env codifica a glicoproteína gp160 (Figura 3), que posteriormente, é clivada em outras duas proteínas: a gp120 e a gp41. Essas glicoproteínas formam o complexo gp120-gp41 que é responsável pela ligação do vírus à célula hospedeira, permitindo a fusão de suas membranas (EARL; MOSS; DOMS, 1991).

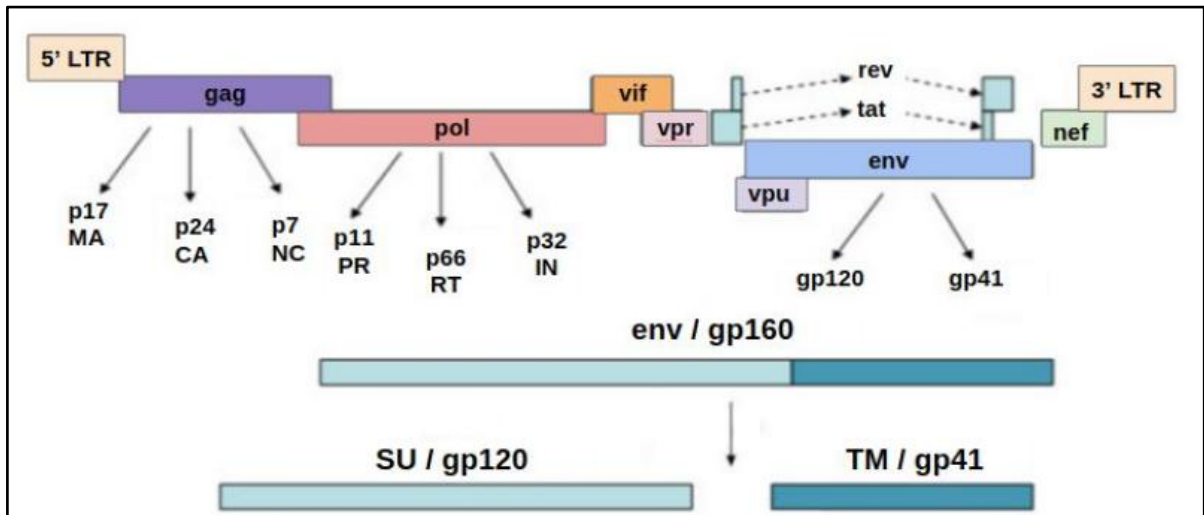


Figura 2 – Organização do genoma do HIV-1. Representação dos nove genes codificados. Os genes gag, pol e env codificam proteínas precursoras (pr55, pr160 e gp160, respectivamente) que serão clivadas em proteínas necessárias para a formação do vírus maduro. Em azul claro temos a proteína de superfície (SU) gp120. Em azul escuro temos a proteína transmembrana (TM) gp41. (Adaptado de COSTIN, J. 2007).

A proteína gp41 possui uma longa cauda citoplasmática, uma região hidrofóbica transmembrana e um ectodomínio, onde se liga a gp120. Na região do ectodomínio existem alguns determinantes funcionais que estão envolvidos na fusão das membranas viral e celular do hospedeiro. Esses determinantes são duas regiões helicoidais denominados HR1 e HR2 (GOMEZ; HOPE, 2005; MONTERO et al., 2008). A gp120 possui cinco regiões conservadas, de C1 a C5, que são separadas por cinco regiões variáveis de V1 a V5 (Figura 3) (POIGNARD et al., 2001).

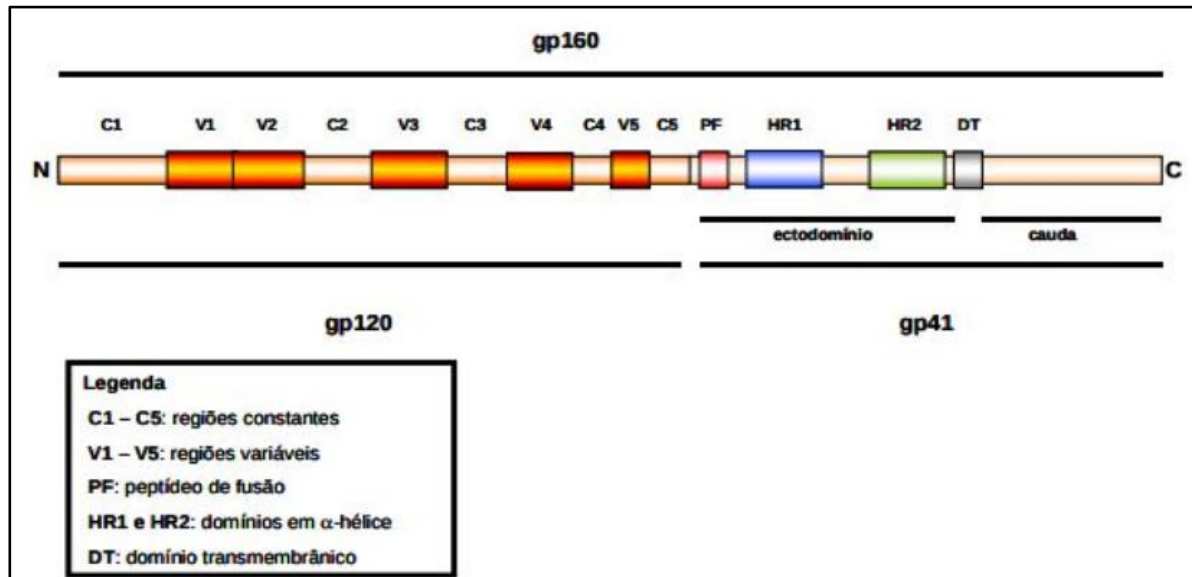


Figura 3 – Proteína precursora gp160, que ao ser clivada gera as glicoproteínas gp120 e gp41 e suas respectivas regiões e domínios. Fonte:

<http://repositorio.unb.br/bitstream/10482/3228/1/Dissert%20Thiago%20Machado%20Mello%20de%20Sousa.pdf>

1.3 Ciclo de replicação viral

A entrada da partícula viral na célula hospedeira ocorre através da ligação da glicoproteína gp120, a duas moléculas da célula hospedeira: um receptor específico da superfície celular, CD4+ e um co-receptor, que é membro da família de receptores de quimiocinas, CCR5 ou CXCR4. A gp120 ao se ligar ao receptor CD4+ altera sua conformação, expondo sua alça V3. Com a alça V3 exposta, ela consegue se ligar ao co-receptor CCR5 ou CXCR4 (CLAPHAM; WEISS, 1997; GOODNOW; COLLMAN, 2006). A partir de então, ocorre a mudança conformacional na gp41, o que permite que suas regiões helicoidais formem um complexo que auxiliam no término da fusão entre as membranas viral e celular (Figura 4 e etapa 1 da figura 5) (GOMEZ; HOPE, 2005).

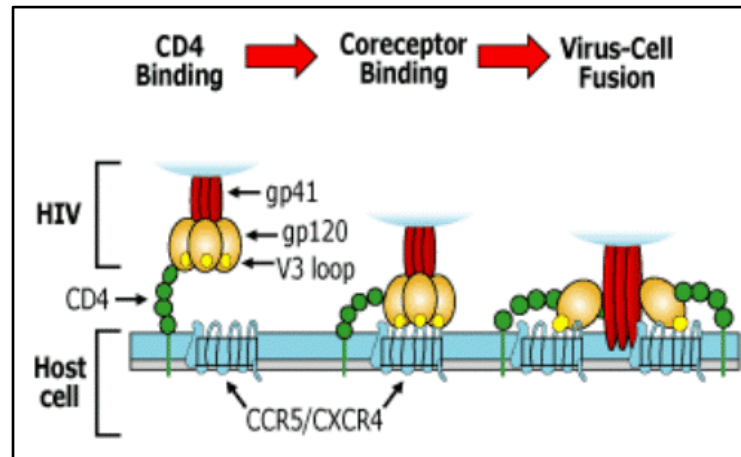


Figura 4 - Mecanismo de entrada do HIV-1 a célula hospedeira. Fonte: <http://chem3513-2007.pbworks.com/w/page/15648421/HIV%20Treatment>

O capsídeo do vírus é liberado para o citoplasma da célula hospedeira (etapa 2) e inicia-se a transcrição do RNA viral em DNA complementar, através da ação da atividade da enzima transcriptase reversa (etapa 3 e 4). Ocorre então o transporte desse DNA complementar para o núcleo da célula (etapa 5), onde a enzima integrase pode atuar fazendo com que aconteça a integração no genoma celular (etapa 6), denominado provírus, que pode permanecer em latência por vários anos (KATZ; SKALKA, 1994) ou permanecer em forma circular, isoladamente. O molde para a síntese de proteínas estruturais de novas partículas virais será o provírus, uma vez que ele produz o RNA mensageiro viral, que será transportado para o citoplasma da célula (etapa 7 e 8).

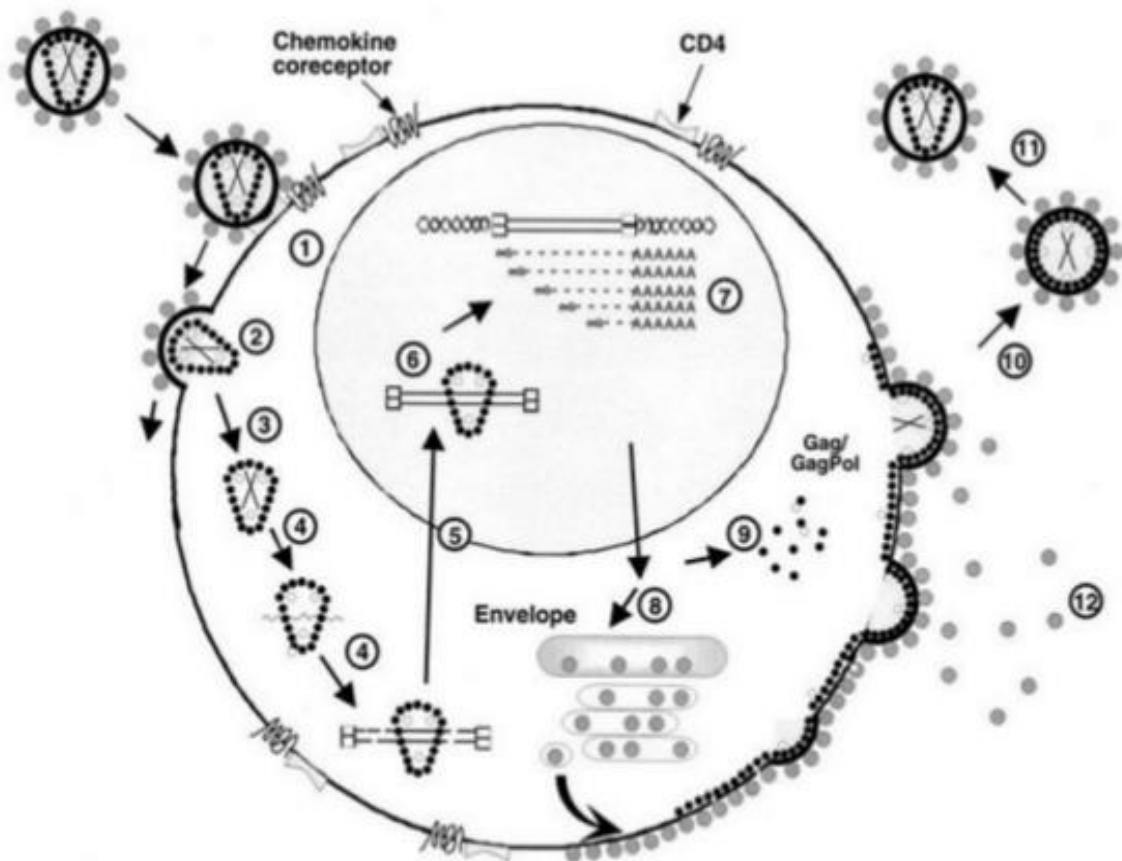


Figura 5 – Ciclo de replicação do HIV-1. Fonte:

<https://basicmedicalkey.com/human-immunodeficiency-viruses-replication-2/>

Em seguida, as proteínas do gene env são produzidas e transportadas até a membrana plasmática da célula. Duas cópias de RNA viral simples são recrutadas pelas proteínas que foram codificadas pelo gene gag e também são direcionadas à membrana plasmática (etapa 9). No citoplasma, inicia-se a montagem estrutural do núcleo do vírus com as proteínas produzidas pelo gene gag. Na membrana celular, esse complexo se junta com as proteínas de superfície gp120 e a transmembranar gp41 que são expressas na membrana e o processo de brotamento é iniciado liberando novas partículas virais. O processo de maturação viral se dá após o brotamento (etapa 10), que irá formar as proteínas p24, que se agrupam para formar o capsídeo (etapa 11). Ao ser liberada, a nova partícula viral pode infectar outras células ou permanecer no meio extracelular. As proteínas produzidas pelo gene pol agora regulam a síntese de novos vírus (CONNOR, 1992; FREED, 1998; GANSER- PORNILLOS; YEAGER ; SUNDQUIST, 2008).

Atualmente os antirretrovirais utilizados interferem em algum passo do ciclo replicativo do HIV. Existem duas principais classes de antirretrovirais: os inibidores da transcriptase reversa, que atuam impedindo a multiplicação viral e os inibidores da protease, que impedem a liberação de novas partículas virais.

1.4 Fases da infecção pelo HIV-1

A infecção pelo HIV pode ser dividida em 3 fases, sendo a primeira a fase aguda ou soroconversão; a segunda fase é denominada fase assintomática ou latência clínica e, por fim, a terceira fase é a fase sintomática. A fase aguda é caracterizada pela incubação do vírus, período denominado eclipse, que ocorre logo após a infecção, onde o vírus fica latente até que os primeiros sintomas da doença apareçam. Estes, podem aparecer entre 3 a 6 semanas após a infecção e desenvolvem-se em cerca de 50% a 70% dos adultos infectados. O organismo leva de 30 a 60 dias após a infecção para iniciar a produção de anticorpos anti-HIV, desse modo, marcadores virais não são identificados nesse período. Após esse tempo, ocorre a disseminação pelo tecido linfóide e sanguíneo. O tempo em que o organismo leva para responder a infecção é denominado janela imunológica e compreende desde o período de eclipse, que dura em torno de 7 dias, até o 20º dia após a infecção.

Durante esse período de janela imunológica apenas duas moléculas podem ser detectadas: o RNA viral e a proteína p24, portanto, caso seja feito testes para HIV nesse período há uma chance do resultado ser falso-negativo (COHEN et al., 2010; FIEBIG et al., 2003, HASSAN, 2013). Posteriormente ao período de janela imunológica, inicia-se a fase onde os anticorpos anti-HIV-1 começam a ser detectados, essa fase é chamada de soroconversão.

Na fase assintomática, o principal evento é a interação entre o vírus e linfócitos TCD4+. É nessa fase que o HIV-1 passa por mutações contínuas e rápidas, porém, tais eventos ainda não são capazes de deprimir o organismo a ponto de permitir com que doenças oportunistas se manifestem. A resposta imune é então montada e pode suprimir a viremia por anos,

período denominado latência clínica. Segundo Rachid (2005), a progressão da fase aguda até a assintomática pode ser de até 10 anos (RACHID; SCHECHTER, 2005). Com o passar do tempo, a viremia reaparece, em razão do aumento da replicação viral e diminuição da resposta imune (WONG-STAAAL, 1991).

A fase sintomática, última fase da infecção, é inicialmente definida pela grande diminuição de linfócitos T CD4+ acompanhada do aparecimento dos sintomas da AIDS como febre, diarreia, suores noturnos e emagrecimento (DEPARTAMENTO DE VIGILÂNCIA, PREVENÇÃO E CONTROLE DAS IST, DO HIV/AIDS E DAS HEPATITES VIRAIS, 2017). Devido a isso, a imunodeficiência pode favorecer o aparecimento de doenças oportunistas, como herpes viral, pneumonia, tuberculose, meningite e micoses, pois o organismo fica cada vez mais vulnerável a tais infecções (GRANT; ZOLOPA, 2012; NELSON et al., 2011).

1.5 Diagnóstico

Os testes para detecção do HIV tem por objetivo diagnosticar precocemente a infecção, a fim de que o indivíduo inicie o mais rapidamente a terapia anti-retroviral altamente ativa (highly active antiretroviral therapy – HAART), o que permite o controle dos indivíduos infectados. Além disso, o diagnóstico também é importante para que não ocorra a disseminação do vírus e nem o aparecimento de doenças oportunistas. Após a etapa de soroconversão, os primeiros anticorpos produzidos são da classe IgM contra a proteína gp41, produzidos cerca de 13 dias após a detecção do RNA viral no plasma. Por conter várias regiões imunogênicas, a gp41 pode desencadear uma forte produção de anticorpos contra ela (MONTERO et al., 2008). Dessa forma, a gp41 é considerada o antígeno mais importante na detecção de HIV-1 (MAHBOUDI et al., 2006). Em seguida, a produção de anticorpos muda de classe (class-switching), passando de IgM para IgG e IgA (TOMARAS et al., 2008). Outra proteína que é um importante alvo de anticorpos é a glicoproteína gp120 (POIGNARD et al., 2001). Ela pode ser detectada a partir de 28º dia após a detecção de RNA viral (TOMARAS; HAYNES, 2009). Além dessas proteínas, há ainda a proteína p24,

encontrada em altos níveis no sangue do indivíduo tanto nos estágios iniciais quanto nas fases mais avançadas da infecção (TANG et al., 2010). Os anticorpos anti-p24 podem ser detectados após 18 dias da detecção do RNA viral (TOMARAS; HAYNES, 2009).

Atualmente, os testes de diagnóstico de HIV-1 podem ser classificados como ensaios de triagem, que possuem alto grau de sensibilidade e testes confirmatórios que possuem alto grau de especificidade. Os ensaios de triagem foram desenvolvidos para detectar todos os indivíduos infectados. Dessa forma, eles podem detectar a presença de anticorpos e/ou antígenos virais, principalmente p24, por imunoenaios, como por exemplo pelo método ELISA (do Inglês, Enzyme-Linked Immunosorbent Assay) ou podem amplificar o genoma do vírus através da técnica denominada Reação em cadeia da polimerase (PCR). Enquanto que os testes confirmatórios, principalmente Immunoblotting (ou Western Blotting) e Imunofluorescência indireta, foram desenvolvidos para identificar os indivíduos que não estão infectados, mas que por alguma razão no ensaio de triagem seu resultado foi reativo (MANUAL TÉCNICO PARA O DIAGNÓSTICO DA INFECÇÃO PELO HIV, MINISTÉRIO DA SAÚDE, 2014). Existem ainda os testes rápidos, que são imunoenaios simples que podem ter duração de até 30 minutos. Eles são capazes de identificar anticorpos IgG e IgM contra a proteína gp41 e/ou gp120 em amostras de sangue, seja plasma ou soro, ou em fluidos orais (DASKALAKIS, 2011). Com o surgimento dos testes rápidos pode-se ampliar o acesso ao diagnóstico de HIV, pois eles são de fácil manipulação, não necessitando de uma infraestrutura laboratorial, equipamentos e profissionais especializados. Além disso, eles podem ser utilizados em regiões de difícil acesso, como por exemplo, onde vivem a população ribeirinha. Eles são mais baratos comparados a outros exames laboratoriais e permitem que um grande número de pessoas sejam avaliadas em função da rapidez do resultado. Os testes rápidos podem ser utilizados em situações que necessitam de um rápido diagnóstico, como nos prontos socorros e na maternidade, onde não se sabe se a gestante fez o teste de HIV durante o pré natal. Outra vantagem dos testes rápidos é que eles permitem o diagnóstico rápido da infecção possibilitando um tratamento precoce, diminuindo assim as chances de disseminação do vírus (MANUAL TÉCNICO PARA O

DIAGNÓSTICO DA INFECÇÃO PELO HIV, MINISTÉRIO DA SAÚDE, 2014; GREENWALD et al, 2006). Há vários tipos desses testes, mas os mais utilizados são dispositivos (ou tiras) de Imunocromatografia (ou fluxo lateral), Imunocromatografia de dupla migração (DPP) e dispositivos de imunoconcentração (MANUAL TÉCNICO PARA O DIAGNÓSTICO DA INFECÇÃO PELO HIV, MINISTÉRIO DA SAÚDE, 2014).

No Brasil, em 17 de dezembro de 2013 o Ministério da Saúde publicou a Portaria nº 29 que normatiza a testagem para HIV, onde recomenda-se que seja realizado pelo menos dois testes para confirmar o resultado: um ensaio de triagem e um ensaio confirmatório. Nesta portaria são apresentados 6 fluxogramas que permitem o diagnóstico seguro da infecção. Através desses fluxogramas é possível o diagnóstico precoce de AIDS. Estima-se que cerca de 25% dos novos casos de infecção aconteçam pelo contágio por indivíduos que estão na fase aguda, ou seja, que ainda não sabem que estão infectados e que possuem uma alta carga viral circulante, facilitando a chance de disseminação do vírus (MANUAL TÉCNICO PARA O DIAGNÓSTICO DA INFECÇÃO PELO HIV, MINISTÉRIO DA SAÚDE, 2014).

1.6 Produção de proteínas recombinantes

Proteínas recombinantes ou proteínas heterólogas são proteínas produzidas através da tecnologia do DNA recombinante, onde se clona o gene da proteína de interesse em um sistema de expressão diferentemente do produtor original.

O dogma central da biologia molecular possibilitou a produção de proteínas recombinantes, pois proporcionou o entendimento dos processos de replicação, transcrição e tradução que a célula utiliza para que a informação genética seja expressa (Figura 6).

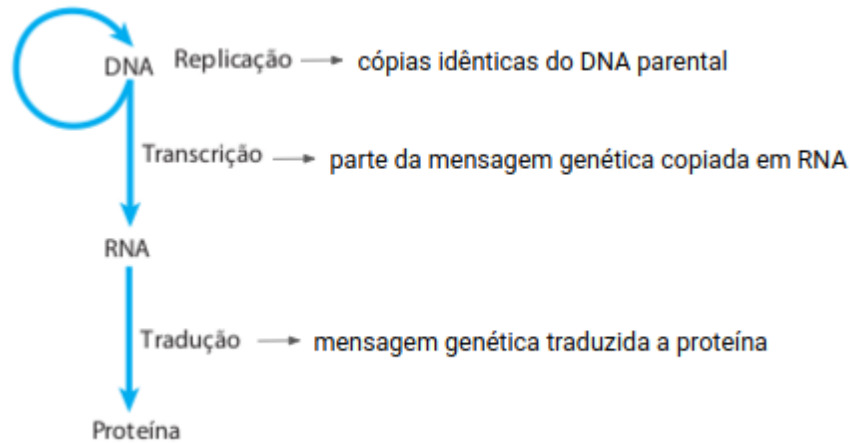


Figura 6 - Dogma central da biologia molecular. Fonte: MIYAMOTO, C. A. Produção de proteínas recombinantes em *Escherichia coli*, 2013.

A tecnologia do DNA recombinante foi introduzida em 1970 com a finalidade de produzir proteínas de maneira mais rápida e fácil em comparação a produção natural.

A produção de proteínas recombinantes é utilizada em várias áreas, dentre elas, na área da saúde. Suas aplicações são as mais diversas, desde medicamentos, vacinas, testes diagnósticos, terapia gênica a produção de anticorpos (DEMAIN; VAISHNAV,2009).

Para produção de proteínas recombinantes, inicialmente é necessário obter a sequência gênica que expressa a proteína de interesse. Em seguida, realiza-se a clonagem molecular dessa sequência em vetores de expressão, com o auxílio de enzimas de restrição e, insere-se o mesmo em um sistema de expressão, que pode ser procarioto ou eucarioto, para que a proteína possa ser produzida (DEMAIN; VAISHNAV,2009).

Os vetores de expressão e a célula hospedeira que serão escolhidos irão depender do objetivo do estudo, porém, alguns fatores são importantes na hora da escolha, bem como: as taxas de crescimento da célula hospedeira, eficiência na produção, necessidade de modificações pós traducionais, funcionalidade do sistema, local onde a proteína será produzida, atividade biológica da proteína, entre outros fatores (CHOI, 2006; GOEDDEL, 1990).

1.7 Vetores plasmidiais de clonagem e expressão

Vetores plasmidiais são moléculas de DNA circular fita dupla que tem como característica a capacidade de se replicar dentro da célula hospedeira, de modo independente, produzindo muitas cópias com o DNA recombinante inserido. Dessa forma, funcionam como veículos de transporte do gene para o interior da célula (NASCIMENTO et al. 1999). A transferência desse vetor de uma célula para outra é realizada através de transfecção, microinjeção e eletroporação em sistemas de expressão eucarióticos. Já em sistema de expressão utilizando células procarióticas a transferência dos vetores plasmidiais ocorrem por transformação ou eletroporação (MADIGAN, 2016).

Os vetores de clonagem e expressão mais utilizados são plasmídeos de bactérias. Os plasmídeos possuem elementos que são responsáveis pelo controle da produção da proteína recombinante. O promotor controla a taxa de transcrição do gene de interesse e os principais plasmídeos de expressão utilizam os promotores Lac, Tac e T7. Dentre os plasmídeos que utilizam o promotor T7 estão os do sistema pET de expressão (plasmid for expression by T7 RNA polimerase). Alguns vetores comerciais desse sistema apresentam o promotor T7lac, onde a expressão da proteína fica sob controle do promotor lac reduzindo seus níveis na ausência do indutor. Com a adição do indutor Isopropil- β -D-tiogalactopirosídeo (IPTG), o gene da T7 RNA polimerase, RNA polimerase do fago T7, é traduzido e se liga na região promotora lac, fazendo com que o gene de interesse seja transcrito e a proteína expressa (NASCIMENTO et al. 1999; Manual pET System, Novagen). A vantagem desse promotor é que a RNA polimerase do fago T7 é mais ativa que a RNA polimerase bacteriana, dessa forma o nível de expressão proteica é maior (NASCIMENTO et al. 1999).

A origem de replicação (ORI) controla a taxa de replicação do plasmídeo. Os marcadores de seleção, por sua vez, são genes que conferem resistência a determinado antibiótico, como por exemplo a ampicilina e kamicina, auxiliando na seleção das colônias que receberam o plasmídeo (ROSANO, 2014). O sítio de múltipla clonagem (Multiple cloning site - MCS) é o local que permite a inserção do gene de interesse. É nesse local que encontram-se os sítios

de clivagem para as enzimas de restrição, também chamadas de endonucleases, onde o inserto, DNA de interesse, é inserido. O plasmídeo deve ter apenas um local para o reconhecimento onde a enzima de restrição irá clivar, pois caso não tenha esse local, o plasmídeo não pode ser aberto para inserção do gene de interesse. Existem inúmeras enzimas de restrição e a escolha de quais utilizar irá depender de alguns fatores, como por exemplo não cortar o DNA inserido (NASCIMENTO et al, 1999). Além desses elementos, os plasmídeos utilizados em bactérias ainda podem conter um sítio de ligação ao ribossomo (Shine-Dalgarno), que irá controlar a tradução, um iniciador: a sequência ATG e um códon para finalizar a tradução (stop codon) (BERLEC, 2013).

Alguns vetores de expressão, como o pET28a, possuem características que auxiliam no processo de purificação, como a cauda de histidina que sai em fusão com a proteína, que são sequências composta por 6 histidinas que ajudam na purificação e também na solubilidade das proteínas heterólogas (COSTA, 2014; MERGULHÃO et al., 2005).

Um outro exemplo de vetor plasmidial é o vetor pPICZαA (EasySelect *Pichia* Expression Kit). Esse vetor é utilizado como vetor de clonagem em bactérias e vetor de expressão em *Pichia pastoris*. Os principais elementos desse vetor são: o promotor AOX1, que permite a indução da proteína heteróloga por metanol; o fator alfa, uma sequência derivada do *Saccharomyces cerevisiae* que possibilita a secreção da proteína heteróloga; o sítio de múltipla clonagem, local onde se insere a sequência de interesse e localizam-se os sítios das enzimas de restrição; a região terminadora de transcrição do gene AOX1; o gene de resistência ao antibiótico zeocina e a cauda de histidina, que auxilia no processo de purificação (EasySelect *Pichia* Expression Kit; CEREGHINO et al., 2000).

1.8 Sistema de expressão procarioto

Existem muitos organismos modelos que podem ser usados na produção de proteínas recombinantes. Porém, um dos mais utilizados na área da biotecnologia é a bactéria *Escherichia coli*, devido ao seu rápido cultivo, fácil manuseio, baixo custo, linhagens bem caracterizadas e fácil de ser manipulada (REDDI et al 2002; GOPAL; KUMAR, 2013).

Em *E. coli* a proteína recombinante expressa pode ser enviada para três locais distintos: 1) meio extracelular, podendo resultar em altas quantidades de proteína produzida e ter sua conformação nativa, entretanto, durante o cultivo, as bactérias podem produzir endotoxinas e proteases que podem degradar as proteínas produzidas (COSTA et al. 2014; SORENSEN, 2005); 2) periplasma bacteriano, que tem um ambiente oxidativo com enzimas que permitem a conformação natural da proteína, porém a expressão tem um baixo rendimento (MAKRIDES, 1996) e 3) citoplasma bacteriano, onde a proteína é expressa em grandes quantidades, mas com grande frequência em corpos de inclusão, forma agregada da proteína. Até 90% da proteína de interesse pode estar em corpos de inclusão, devido ao alto nível de expressão da proteína heteróloga, gerando um pool de proteínas mal enoveladas que se agregam (COSTA et al. , 2014; REDDI et al 2002). Para sua obtenção solúvel é necessário o uso de reagentes desagregantes como ureia ou guanidina. Pode-se também tentar otimizar o processo, diminuindo a temperatura durante a indução da expressão, pois assim ocorre uma diminuição do grau de transcrição, tradução e das forças de interação hidrofóbicas que contribuem para o dobramento incorreto da proteína. Uma outra estratégia que pode auxiliar no dobramento correto é a co-expressão de chaperonas (ROSANO; CECCARELLI, 2014).

Uma dificuldade na expressão de proteínas heterólogas é o padrão de utilização dos códons. O gene de interesse pode apresentar códons que são pouco utilizados ou não são reconhecidos pelo organismo usado como sistema de expressão, pois existe uma preferência no uso de códons o qual corresponde a moléculas de RNAt específicas.

Os códons pouco utilizados são denominados códons raros (EVANS et al., 2011). Para que não ocorra esse problema uma estratégia que pode ser adotada é a otimização dos códons do gene de interesse. Esse processo consiste em trocar os códons raros por códons preferenciais ao organismo usado no sistema de expressão (ROSANO; CECCARELLI, 2014). Outra estratégia é usar uma linhagem de *E. coli* que produza os códons raros, como as linhagens Rosetta, Rosetta (DE3), BL21 (DE3) CodonPlus-RIL e BL21 (DE3) CodonPlus-RP (KAUR et al. 2018).

Outra limitação que a *E. coli* apresenta é a falta de realização de modificações pós traducionais, como a glicosilação, por exemplo. Isso ocorre devido a falta de maquinaria sofisticada para realizar tais processos. Nesse caso, utilizar um sistema de expressão eucariótico é a melhor escolha para a produção da proteína que requer tais modificações (DALY e HEARN, 2005; KAUR et al. 2018).

1.9 Sistema de expressão eucarioto

Existem vários sistemas de expressão eucarióticos que podem ser usados na produção de proteínas recombinantes (YIN et al. 2007). Alguns exemplos são as células de mamíferos, células de plantas, células de insetos e leveduras.

As vantagens que o sistema eucarioto apresenta em comparação ao sistema procarioto são diversas, tais como: a secreção de proteínas para o meio extracelular, a realização de modificações pós traducionais, a ausência de formação de corpos de inclusão e o correto enovelamento das proteínas produzidas. Além dessas vantagens, há ainda a produção estável de clones recombinantes, devido a integração dos vetores de expressão ao genoma da levedura, o que aumenta a estabilidade das linhagens utilizadas na expressão (TORRES, MORAES, 2000), o rendimento relativamente alto de proteínas e custo inferior comparado a sistemas eucariotos multicelulares (ÇELIK et al., 2012; DEMAIN et al., 2009;).

Dentre esses sistemas, um dos mais utilizados são as leveduras, que são fungos unicelulares. Muitas espécies com características diferentes podem ser usadas na produção heteróloga de proteínas, como *Saccharomyces cerevisiae*, *Pichia pastoris*, *Kluyveromyces lactis* e *Hanselua polimorpha* (ÇELIK et al., 2012).

Por ser de fácil manipulação genética, apresentar um satisfatório crescimento e secretar altos níveis da proteína produzida, a levedura *Pichia pastoris* tem sido muito utilizada como sistema de expressão (CEREGHINO, CREGG; 2000).

Pichia pastoris é uma levedura anaeróbia facultativa e metilotrófica, ou seja, capaz de metabolizar o metanol como única fonte de carbono. Essa reação de metabolização ocorre inicialmente nos peroxissomos e depois no citoplasma. Quando as células são cultivadas

em metanol, encontram-se nos peroxissomos as enzimas álcool oxidases (AOX) 1 e 2, que oxidam o metanol em formaldeído e peróxido de hidrogênio. Por apresentar toxicidade para a célula o peróxido de hidrogênio é rapidamente decomposto em água e oxigênio pela enzima catalase (CREGG et al., 1989; CEREGHINO, CREGG; 2000). O formaldeído produzido é então utilizado para obtenção de energia para o crescimento da levedura e metabolizado no citoplasma, formando dióxido de carbono por desidrogenases citoplasmáticas (CEREGHINO, CREGG; 2000).

Uma característica importante é que a *Pichia pastoris* não é considerada uma levedura muito fermentadora, dessa forma ela não produz altos níveis de etanol como metabólito intermediário, evitando assim que atinja níveis tóxicos para as células (TORRES, MORAES, 2000).

2. JUSTIFICATIVA

Desde a descoberta da AIDS, por volta da década de 80 até junho de 2016 foram notificados no Brasil cerca de 843.000 casos de AIDS e mais de 300.000 óbitos. Nos últimos cinco anos, o país tem registrado uma média de 41,1 mil novos casos de AIDS por ano. Segundo o Boletim epidemiológico HIV/AIDS 2016, nos últimos dez anos, observou-se um aumento da taxa de detecção de HIV/AIDS no Brasil, principalmente entre jovens de 15 a 24 anos (BOLETIM EPIDEMIOLÓGICO HIV/AIDS, 2016). Sendo assim, a infecção por HIV é vista hoje como um dos grandes problemas de saúde pública, principalmente porque sua incidência vem aumentando nos últimos anos. Dentro deste contexto, diagnosticar precocemente a infecção é fundamental, para que se inicie o tratamento anti-retroviral o mais rápido possível, diminuindo assim as chances de disseminação do vírus. Com o surgimento dos testes rápidos pode-se ampliar o acesso ao diagnóstico de HIV, pois eles podem ser utilizados em regiões de difícil acesso, onde não há uma estrutura laboratorial. Entretanto, a maioria dos testes rápidos existentes no mercado atualmente detectam anticorpos que só podem ser identificados em fases mais avançadas da infecção. Testes rápidos que detectam anticorpos na fase mais inicial podem auxiliar no diagnóstico precoce,

permitindo assim um tratamento adequado. Desse modo, é necessário que os testes rápidos sejam continuamente melhorados a fim de que sejam amplamente disponíveis, apresentem baixo custo, sejam sensíveis e específicos. Investir no melhoramento de testes rápidos é ampliar o acesso ao diagnóstico de HIV de uma maneira simples e eficaz.

3. OBJETIVOS

3.1 Objetivo geral

O objetivo do presente estudo foi produzir as proteínas gp41 e gp120 do HIV-1 por tecnologia do DNA recombinante com a finalidade de serem empregadas futuramente em teste de diagnóstico.

3.2 Objetivos específicos

- Clonar as proteínas gp41 e gp120 em vetor plasmidial
- Expressar as proteínas recombinantes gp41 e gp120 do HIV-1
- Purificar as proteínas expressas por cromatografia de afinidade a níquel
- Confirmar a identidade das proteínas através do sequenciamento
- Verificar se as proteínas são reconhecidas por anticorpos

4. MATERIAL E MÉTODOS

4.1 Fluxograma da metodologia utilizada

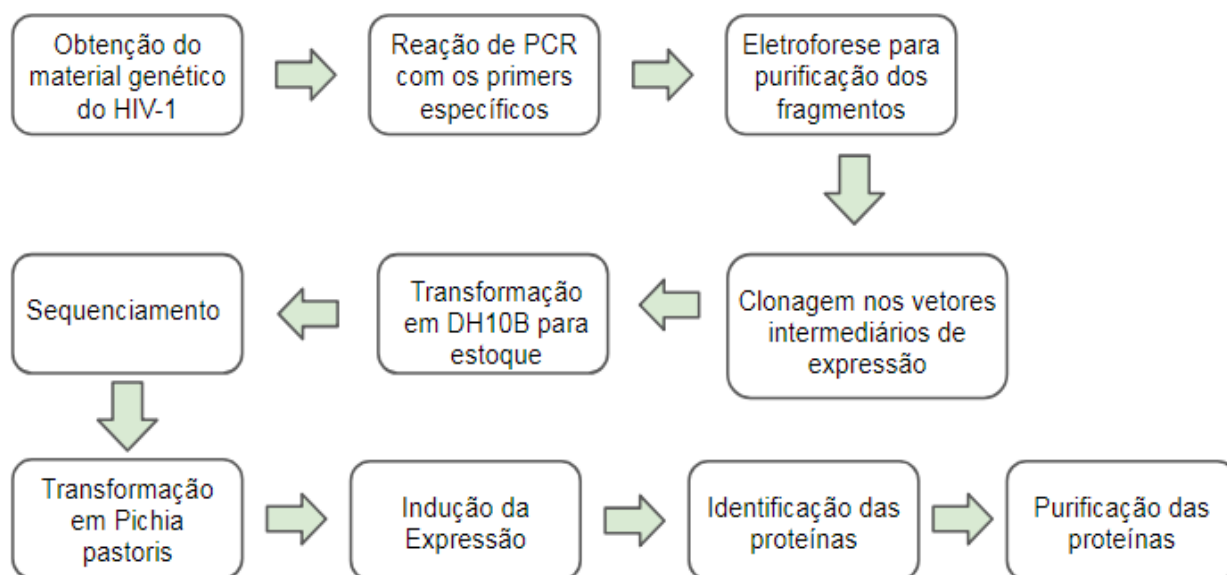


Figura 7 - Fluxograma da metodologia utilizada

4.2 Clonagem e expressão

4.2.1 Obtenção do material genético do HIV-1

As sequências das proteínas gp41 e gp120 foram clonadas a partir do vetor que carregava tais sequências, o vetor pABT4674 (ATCC) (Figura 8).

pAbT4674 is a vector for the insertion and expression of HIV-1 env (strain BH10) and gag-pol (strain HXB2) in vaccinia. pAbT4674 contains the HIV-1 env gene under the control of the vaccinia D1 promoter and the HIV-1 gag-pol genes under the control of the vaccinia 40K promoter flanked by vaccinia DNA for directing recombination into the vaccinia HindIII M region, a portion of the vaccinia 29K host-range gene for selection of recombinants, and a bacterial replicon and ampicillin-resistance gene for growth and selection in *E. coli*.

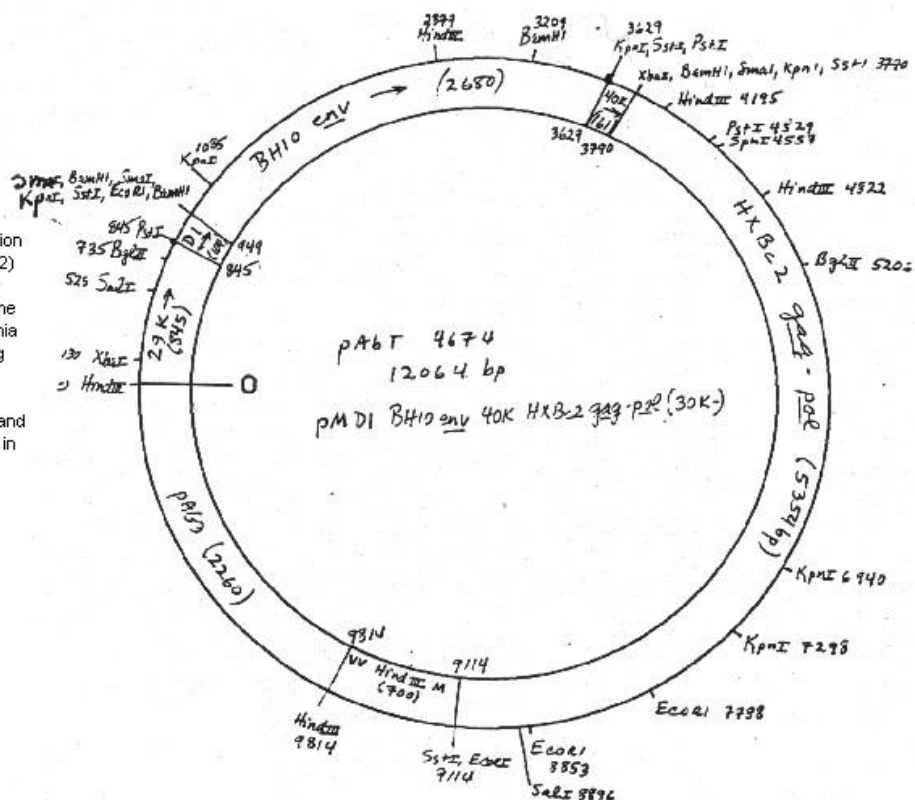


Figura 8 - Mapa do vetor pAbT4674 (ATCC).

4.2.2 Desenho dos oligonucleotídeos iniciadores (primers)

Após a análise do genoma do HIV-1, disponível no GenBank com número de acesso K03455.1 e, a partir do gene env que codifica as proteínas de interesse, os primers para clonagem no vetor pET28a foram desenhados. Eles foram desenhados com sítio de restrição para as enzimas NcoI (CCATGG) e a XhoI (CTCGAG) para amplificação das proteínas gp41 e gp120. A escolha das enzimas de restrição ocorreu pela vantagem da permanência da cauda de histidina, facilitando o processo de purificação, e também pela permanência do sítio de ligação ao ribossomo (RBS), importante para síntese das proteínas. Além disso, foram adicionados três nucleotídeos na extremidade 5' de cada primer, que serviram como facilitadores para ancoragem das enzimas de restrição para digestão do produto de PCR. Os primers foram sintetizados pela empresa Sinapse Biotecnologia (tabela 1).

Primer	Sequência
gp41 - Forward	5' AA CCATGG GGATGGGCGCAGCCTCAATG 3'
gp41 - Reverse	5' TTT CTCGAG TCCCCTCGGGGTTGGG 3'
gp120 - Forward	5' AA CCATGG GGTGTGGGTCACAGTCTATTATGGG 3'
gp120 - Reverse	5' TTT CTCGAG TCTTTTTTCTCTCTGCCACCACTCT 3'

Tabela 1 - Sequência dos primers utilizados na clonagem dos genes das proteínas gp41 e gp120 em vetor pET28a. Em negrito encontram-se o sítio das enzimas de restrição NcoI (CCATGG) nas sequências *forwards* e a XhoI (CTCGAG) nas sequências *reverses*.

Para a clonagem das proteínas gp41 e gp120 no vetor pPICZαA foram desenhados os primers descritos na tabela 2. Eles foram desenhados com sítio de restrição para as enzimas EcoRI (GAATTC) e a Sall (GTCGAC). A escolha das enzimas de restrição ocorreu também pela vantagem da permanência da cauda de histidina e pela permanência do fator alfa do vetor pPICZαA que possibilita que a proteína heteróloga seja secretada para o sobrenadante. Além disso, também foram adicionados três nucleotídeos na extremidade 5' de cada primer, que serviram como facilitadores para ancoragem das enzimas de restrição para digestão do produto de PCR. Os primers foram sintetizados pela empresa Exxtend Soluções em Oligos (tabela 3).

Primer	Sequência
gp41 - Forward	5' AA GAATTC ATGGGCGCAGCCTCAATG 3'
gp41 - Reverse	5' TTT GTCGAC TCCCCTCGGGGTTGGG 3'
gp120 - Forward	5' AA GAATTC AGTGCTACAGAAAAATTGTGG 3'
gp120 - Reverse	5' TTT GTCGAC TCTTTTTTCTCTCTGCACCAC 3'

Tabela 2 - Sequência dos primers utilizados na clonagem dos genes das proteínas gp41 e gp120 em vetor pPICZαA. Em negrito encontram-se o sítio das enzimas de restrição EcoRI (GAATTC) nas sequências *forwards* e a Sall (GTCGAC) nas sequências *reverses*.

4.2.3 Sistemas de expressão

Para o presente estudo as bactérias *E. coli* da linhagem DH10 β foram utilizadas para receber os plasmídeos após a clonagem e realizar a multiplicação dos mesmos.

Para expressão heteróloga das proteínas gp41 e gp120 foram utilizados dois sistemas de expressão: o procarioto *E. coli* e o eucarioto *Pichia pastoris*. No sistema procarioto foram utilizadas duas linhagens: a BL21(DE3)plysS e a Rosetta (DE3), ambas com resistência ao antibiótico cloranfenicol. No sistema eucarioto a linhagem utilizada foi a KM71H da levedura *Pichia pastoris*. Essa linhagem foi escolhida pois não possui genes que codificam algumas proteases, o que garante uma menor destruição da proteína secretada (CEREGHINO, 2000; DALY e HEARN, 2005).

4.2.4 Vetores de clonagem e expressão

Para clonagem e expressão de proteínas recombinantes em *E. coli* o vetor escolhido foi o pET28a(+) (Novagen; Figura 9). Este vetor apresenta resistência ao antibiótico kanamicina, que auxilia na seleção dos clones recombinantes. Após a clonagem das sequências de interesse no vetor pET28a(+), os plasmídeos foram inseridos em bactérias DH10 β e armazenados no freezer -80°C.

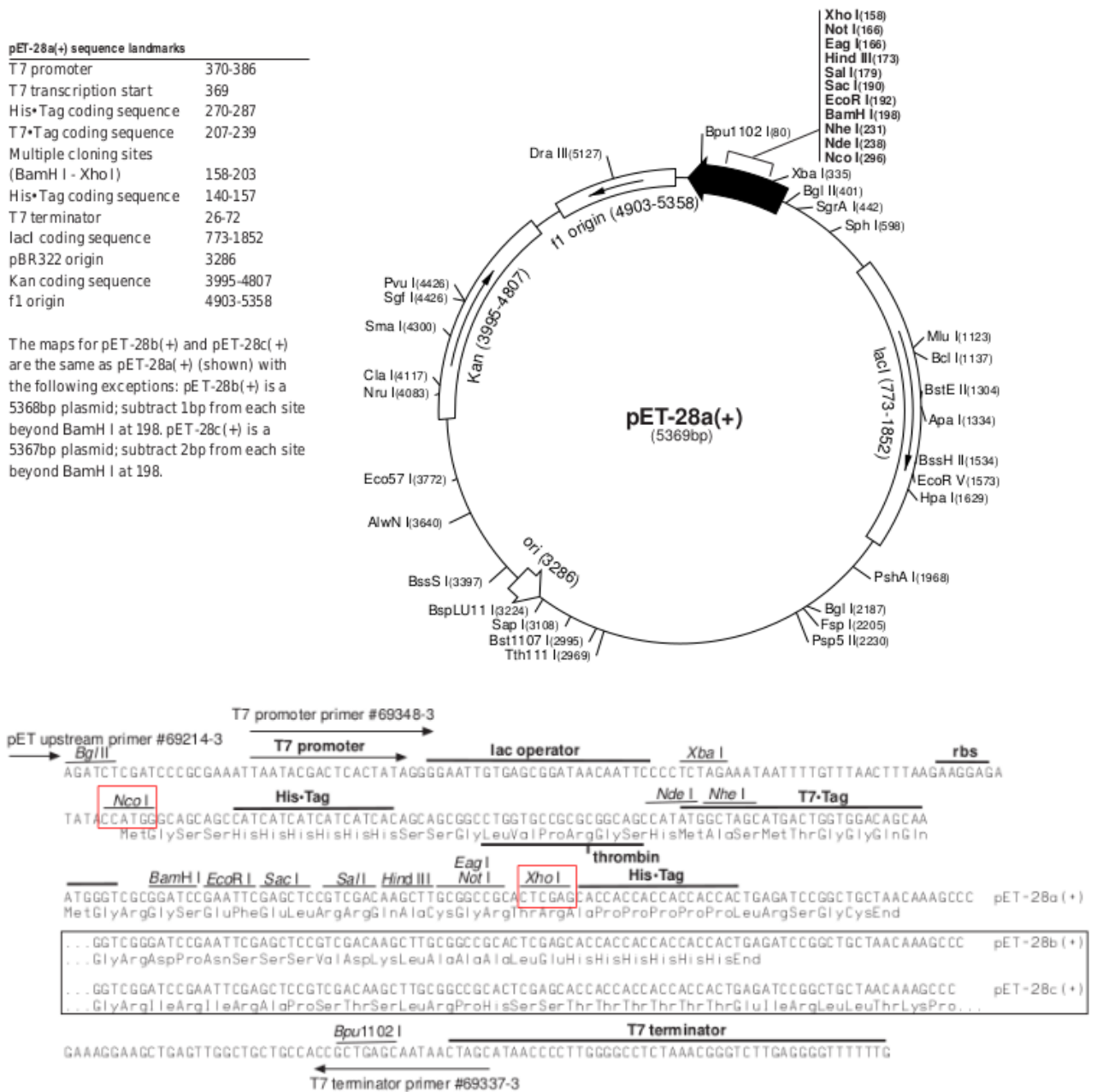


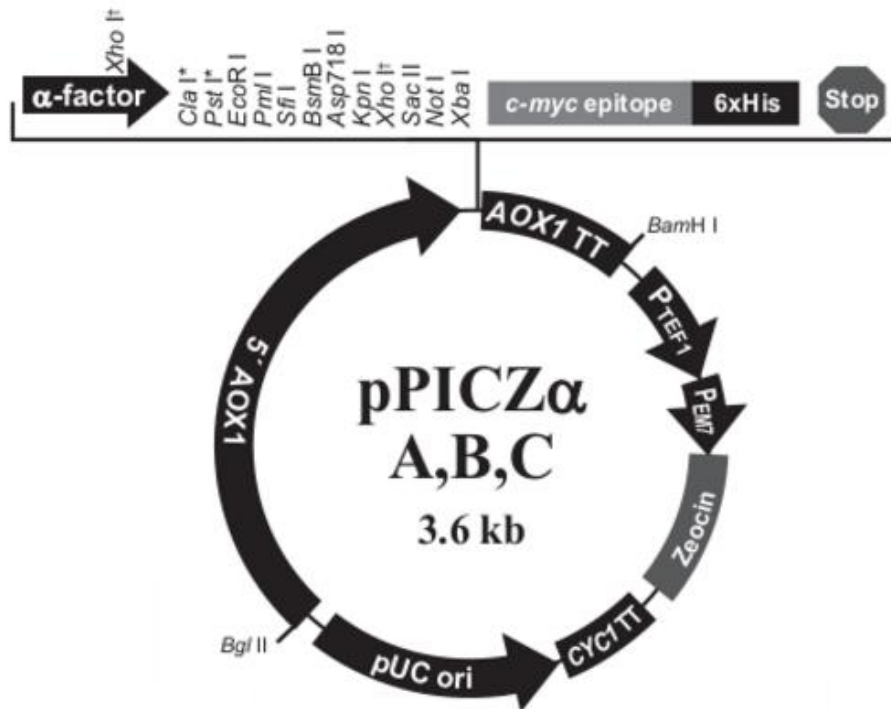
Figura 9 - Mapa do vetor pET28a (+). Nos retângulos em vermelho estão os sítios de restrição das enzimas NcoI e XhoI, utilizadas na clonagem. Fonte: Adaptado de

<https://www.staff.ncl.ac.uk/p.dean/pET28.pdf>

Para a clonagem e expressão na *Pichia pastoris* o vetor escolhido foi o pPICZaA (Figura 10; Invitrogen). O vetor pPICZaA é integrado no genoma da levedura juntamente com o promotor AOX1 presente no mesmo. Quando adicionado ao meio de cultivo o indutor metanol, a levedura produz grande quantidade da proteína recombinante, devido a regulação do promotor AOX1. Além disso, esse vetor apresenta resistência ao antibiótico

zeocina, que auxilia na seleção dos clones recombinantes. Outra característica desse vetor é que ele permite que a proteína recombinante seja fusionada a seis histidinas o que facilita o processo de purificação.

Após a clonagem das sequências de interesse no vetor citado anteriormente, os plasmídeos foram inseridos em bactéria *E. coli* da linhagem DH10 β e armazenados no freezer -80°C.



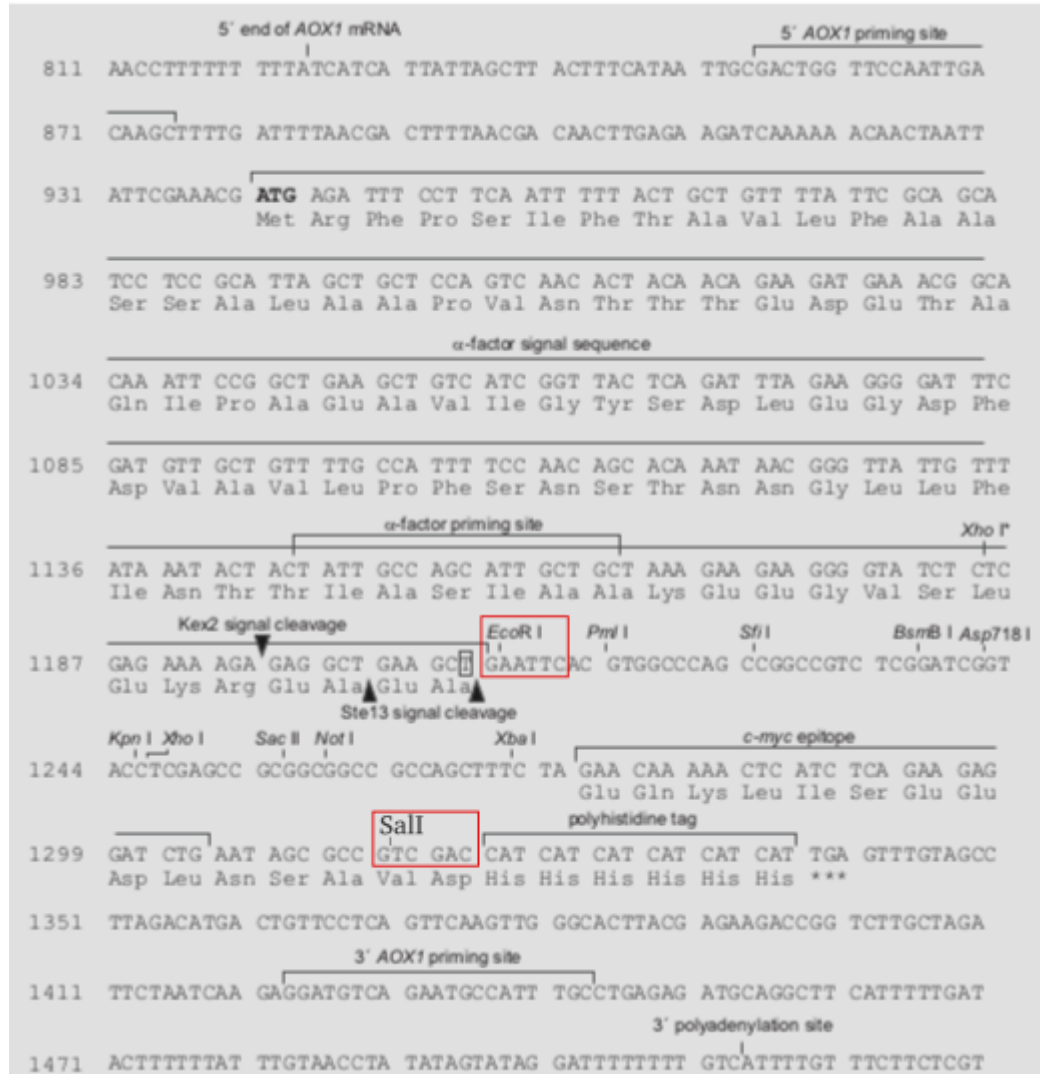


Figura 10 - Mapa do vetor pPICZαA. Nos retângulos em vermelho estão os sítios de restrição das enzimas EcoRI e Sall, utilizadas na clonagem. Fonte: Adaptado de EasySelect™ Pichia Expression Kit, Invitrogen.

4.2.5 Expansão clonal

Em uma placa de Petri foi colocado 15 mL de meio Luria - Bertoni (LB) ágar (1 g de extrato de levedura, 2 g de triptona, 2 g de NaCl e 200 mL de água) e adicionado 15 µL do antibiótico ampicilina (100 mg/mL). A *E. coli* DH10β transformada com o vetor pAbT4674 foi plaqueada e colocada em estufa, 37°C, overnight.

4.2.6 Extração plasmidial (Mini prep)

Para a extração do DNA plasmidial, foram escolhidos alguns clones transformantes que foram inoculados em 5 mL de meio LB com 5 µL de ampicilina (100 mg/mL) e incubados de

16 a 18h horas a 25°C no período noturno e 37°C no início da manhã, a 200 rpm. A extração foi realizada através do kit Illustra plasmid Prep Mini Spin Kit (GE HealthCare) de acordo com o protocolo recomendado pelo fabricante. Após a extração, as preparações de DNA plasmidiais obtidas foram quantificadas através de métodos espectrofotométricos (MANIATIS et al. 1989) pelo equipamento NanoVue Plus (GE Healthcare) ou Nanodrop 1000 (Thermo Fisher Scientific) e armazenadas a -20°C.

4.2.7 Reação em cadeia da polimerase (PCR)

Para construção dos vetores de expressão, os fragmentos de DNA foram amplificados a partir do vetor pAbT4674 pela reação de polimerização em cadeia (PCR) catalisada pela taq polimerase (SAIKI et al., 1988). Foram utilizados os primers específicos (Tabela 1 e 2) para cada sequência.

A reação de PCR para a sequência gp41, após a padronização, foi composta pelos seguintes reagentes: 2,5 µL de Buffer 10x (Ludwig Biotec), 0,75 µL de MgCl₂ (Ludwig Biotec), 1 µL de mix de dNTPs (10 mM) (Ludwig Biotec), 1 µL de primer Forward gp41 (10 mM) (Tabela 1 para *E. coli*, Síntese Biotecnologia e Tabela 2 para *Pichia pastoris*, Exxtend Soluções em Oligos), 1 µL de primer Reverse gp41 (10 mM) (Tabela 1 para *E. coli*, Síntese Biotecnologia e Tabela 2 para *Pichia pastoris*, Exxtend Soluções em Oligos), 0,5 µL da enzima taq polimerase (Ludwig Biotec), 17,25 µL de água autoclavada e 1 µL de DNA (39 ng/µL), totalizando um volume final de 25 µL. O termociclador utilizado foi o Thermal Cycler T100 BIORAD com um ciclo de 95°C por 5 minutos, 35 ciclos de 92°C por 45 segundos, 54°C por 40 segundos, 72°C por 50 segundos e um ciclo de 72°C por 5 minutos.

A reação de PCR para a sequência gp120, após a padronização, foi composta pelos seguintes reagentes: 2,5 µL de Buffer 10x (Ludwig Biotec), 0,75 µL de MgCl₂ (Ludwig Biotec), 1 µL de mix de dNTPs (10 mM) (Ludwig Biotec), 1 µL de primer Forward gp120 (10 mM) (Tabela 1 para *E. coli*, Síntese Biotecnologia e Tabela 2 para *Pichia pastoris*, Exxtend Soluções em Oligos), 1 µL de primer Reverse gp120 (10 mM) (Tabela 1 para *E. coli*, Síntese Biotecnologia e Tabela 2 para *Pichia pastoris*, Exxtend Soluções em Oligos), 0,5 µL da

enzima taq polimerase (Ludwig Biotec), 17,25 μL de água autoclavada e 1 μL de DNA (39 ng/ μL), totalizando um volume final de 25 μL . O termociclador utilizado foi o Thermal Cycler T100 BIORAD com um ciclo de 95°C por 5 minutos, 35 ciclos de 92°C por 45 segundos, 53°C por 40 segundos, 72°C por 1 minuto e 28 segundos e um ciclo de 72°C por 10 minutos.

4.2.8 Eletroforese em gel de agarose

Os fragmentos de DNA que foram amplificados por PCR foram separados e analisados por eletroforese em gel de agarose. A eletroforese em gel de agarose também foi usada para visualização da digestão dos plasmídeos e para purificação dos fragmentos de interesse.

Para a análise dos fragmentos de DNA amplificados pela PCR foram feitos géis de agarose a 1% e para a purificação dos fragmentos a 0,5% ou a 0,8%, todos feitos em tampão de corrida TAE (Tris-Base 0,04 mol/L, ácido acético glacial 1 mol/L, EDTA 50 mmol/L, pH 8,0). A eletroforese foi realizada em cuba Owl EasyCast B1 (Thermo Scientific) contendo tampão TAE, com voltagem fixa de 90V e amperagem livre. O tampão de amostra (azul de bromofenol 0,25% e sacarose 40%) contendo Gel Red (Biotium) foi adicionado nas amostras a serem analisadas. Os resultados foram visualizados em Transluminador spectroline UV (SPECTROLINE), para verificação da pureza e integridade das amostras e documentados pelo fotodocumentador ChemiDoc™ MP System (Bio Rad).

4.2.9 Purificação da reação de PCR e quantificação

Os produtos de PCR foram purificados utilizando o kit illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) de acordo com o protocolo recomendado pelo fabricante. Em seguida, as amostras foram quantificadas pelo equipamento NanoVue Plus (GE Healthcare).

4.2.10 Digestão dos plasmídeos

A digestão dos plasmídeos foi realizada pelas enzimas de restrição NcoI e XhoI com a finalidade de gerar extremidades coesivas que facilitem a reação de ligação com o vetor

pET 28a. O vetor pET28a também foi digerido com tais enzimas. Para reação de restrição do pET28a foram utilizados 12 μL de água autoclavada, 3 μL de Buffer (Thermo), 1,5 μL da enzima Fast Digest NcoI (Thermo), 1,5 μL da enzima Fast Digest XhoI (Thermo) e 12 μL de pET28a (136 ng/ μL), totalizando 30 μL de volume final. Para reação de restrição da sequência gp41 foram utilizados 3 μL de Buffer (Thermo), 1 μL da enzima Fast Digest NcoI (Thermo), 1 μL da enzima Fast Digest XhoI (Thermo) e 25 μL de DNA (10,6 ng/ μL), totalizando 30 μL de volume final. Para reação de restrição da sequência gp120 foram utilizados 6 μL de água autoclavada, 3 μL de Buffer (Thermo), 1,5 μL da enzima Fast Digest NcoI (Thermo), 1,5 μL da enzima Fast Digest XhoI (Thermo) e 18 μL de DNA (67 ng/ μL), totalizando 30 μL de volume final. As reações foram incubadas em banho a 37°C por 30 minutos e inativadas em banho a 80°C por 15 minutos.

Para clonagem no vetor pPICZ α A a digestão dos plasmídeos foi realizada com as enzimas de restrição EcoRI e Sall. O vetor pPICZ α A também foi digerido com tais enzimas. Para reação de restrição do pPICZ α A foram utilizados 20 μL de água autoclavada, 3 μL de Buffer O (Thermo), 1 μL da enzima EcoRI (Thermo), 1 μL da enzima Sall (Thermo) e 5 μL do vetor pPICZ α A (324 ng/ μL), totalizando 30 μL de volume final. Para reação de restrição da sequência gp41 foram utilizados 3 μL de Buffer O (Thermo), 1 μL da enzima EcoRI (Thermo), 1 μL da enzima Sall (Thermo) e 25 μL de DNA (15 ng/ μL), totalizando 30 μL de volume final. Para reação de restrição da sequência gp120 foram utilizados 3 μL de Buffer O (Thermo), 1 μL da enzima EcoRI (Thermo), 1 μL da enzima Sall (Thermo) e 25 μL de DNA (36 ng/ μL), totalizando 30 μL de volume final. Todas as reações foram incubadas em banho a 37°C por 2 horas e inativadas em banho a 65°C por 20 minutos.

Para verificação dos produtos digeridos foi feita a eletroforese em gel de agarose. Posteriormente, foi realizado a purificação do gel utilizando o kit illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) de acordo com o protocolo recomendado pelo fabricante. As amostras foram quantificadas pelo equipamento NanoVue Plus (GE Healthcare).

4.2.11 Construção dos vetores pET28a_gp120, pET28a_gp41, pPICZαA_gp41 e pPICZαA_gp120

Após a digestão dos plasmídeos, a reação de ligação foi realizada respeitando a proporção 3:1, ou seja, três moléculas de inserto para cada uma molécula de vetor. Para a clonagem do inserto gp41 no vetor pET28a montou-se a seguinte reação: 2,44 µL de água autoclavada, 2 µL de Buffer Ligase 10X (Biolabs), 1 µL da enzima T4 Ligase (Biolabs), 3,09 µL do inserto gp41 (7ng/uL) e 11,47 µL do vetor pET-28 (6,1 ng/uL). Para a clonagem do inserto gp120 no vetor pET28a utilizou-se 9 µL de água autoclavada, 2 µL de Buffer Ligase 10X (Biolabs), 1 µL da enzima T4 Ligase (Biolabs), 2,5 µL do inserto gp120 (17ng/uL) e 5,5 µL do vetor pET-28 (10 ng/uL).

Para a clonagem do inserto gp41 no vetor pPICZαA montou-se a seguinte reação: 9,84 µL de água autoclavada, 2 µL de Buffer Ligase 10X (Biolabs), 1 µL da enzima T4 Ligase (Biolabs), 3,55 µL do inserto gp41 (9 ng/uL) e 3,61 µL do vetor pPICZαA (18 ng/uL). Para o inserto gp120 utilizou-se 3,62 µL de água autoclavada, 2 µL de Buffer Ligase 10X (Biolabs), 1 µL da enzima T4 Ligase (Biolabs), 10,88 µL do inserto gp120 (5 ng/uL) e 2,5 µL do vetor pPICZαA (18 ng/uL).

Todas as reações de ligação foram incubadas a 16°C, overnight no termociclador (BIORAD).

4.2.12 Preparo de células de competentes

Foram preparadas células competentes de *E. coli* das seguintes linhagens: DH10β, BL21(DE3)pLYSs e Rosetta (DE3).

Para o preparo de bactérias competentes DH10β para a transformação, foi inoculado em 5 mL de meio LB com 5 µL de kanamicina (100 mg/mL) uma colônia de DH10β e incubado de 16 a 18h horas a 25°C no período noturno e 37°C no início da manhã, a 200 rpm.

Para o preparo de bactérias competentes BL21(DE3)pLYSs e Rosetta (DE3) foi inoculado em 5 mL de meio LB com 5 µL de kanamicina (100 mg/mL) e 170 µL cloranfenicol (1 mg/mL) uma colônia de cada linhagem separadamente e incubado de 16 a 18h horas a 25°C no período noturno e 37°C no início da manhã, a 200 rpm. No dia seguinte, 2 mL do pré-inóculo

realizado no dia anterior foi inoculado em um erlenmeyer de 500 mL contendo 100 mL do meio LB e incubado em agitador orbital a 200 rpm a 37°C por aproximadamente 2 horas, até atingir uma densidade óptica a 600nm (D.O. 600nm) de aproximadamente 0,4. A cultura então foi colocada em gelo e transferida para tubos cônicos de 50 mL. Em seguida, foi centrifugada por 7 minutos a 3095 x g a 4°C. O sobrenadante foi descartado, o pellet foi ressuspensão cuidadosamente no gelo em 35 mL de cloreto de cálcio (CaCl₂, 0,1 M, autoclavado e gelado) e foi deixado em gelo por mais 20 minutos. Após esse tempo, foi levado para centrifugação novamente por 7 minutos a 3095 x g a 4°C. O sobrenadante foi descartado e o pellet ressuspensão cuidadosamente no gelo em uma solução contendo 8 mL de CaCl₂, 0,1M e 2 mL de glicerol 50%. Alíquotas de 200 µL foram colocadas em eppendorfs estéreis e armazenadas a -80°C até momento do uso.

O preparo de *Pichia pastoris* competente foi feito de acordo com o manual "EasySelect™ *Pichia* Expression kit" (Invitrogen). A metodologia é baseada no crescimento prévio das células até que elas atinjam a fase logarítmica de crescimento celular. Após o crescimento, as células são transferidas para meio nutriente YEPD (Extrato de levedura 1%, Peptona 2% e Ágar 2%) adicionado de HEPES (ácido N- (2-hidroxietil) piperazina-N'-2- etanossulfônico) e DTT (Ditiotreitol), que tem a função de evitar a formação de pontes de sulfeto intracelular. Em seguida, foi realizado lavagens nas células com água gelada. Após as lavagens as células foram transferidas para uma solução de Sorbitol 1M.

4.2.13 Transformação por choque térmico de células competentes

As células competentes foram retiradas do freezer a -80°C e descongeladas em gelo por 15 minutos. Posteriormente, foi adicionado o DNA plasmidial e as células foram levadas ao banho-maria a 42°C por 90 segundos precisamente e em seguida, em gelo por 2 minutos. Em seguida ao choque térmico, foram adicionado 800 µL de meio LB no tubo de células competentes e este foi levado a banho-maria a 37°C por 1 hora. Utilizou-se 200 µL de células competentes DH10β e 20 µL da reação de ligação. A mistura foi então semeada em placas de Petri contendo meio LB e kanamicina (100 mg/mL) para as bactérias com os

vetores pET28a_gp41 e pET28a_gp120, LB low salt contendo zeocina (25 µg/mL) para as bactérias com os vetores pPICZαA_gp41 e pPICZαA_gp120 e LB contendo kanamicina (100 mg/mL) e cloranfenicol (1 mg/mL) para as linhagens BL21(DE3)pLYSs e Rosetta (DE3). As placas foram deixadas overnight na estufa a 37°C.

4.2.14 Transformação por eletroporação em *Pichia pastoris* competentes

A transformação das células competentes *Pichia pastoris* foi realizada com os plasmídeos recombinantes pPICZαA_gp41 e pPICZαA_gp120. Aproximadamente 1 µg dos plasmídeos purificados foi linearizado com a enzima de restrição PmeI (New England Biolab). A reação de linearização foi incubada a 37°C em banho maria overnight. No dia seguinte, a enzima foi inativada a 65°C por 20 minutos e os DNAs foram precipitados utilizando o kit illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare).

Com os vetores linearizados e precipitados, cerca de 1 µg foi adicionado em 40 µL de células competentes. As células foram incubadas em gelo por 5 minutos e submetidas a pulsos elétricos utilizando o equipamento Gene Pulser (Bio-Rad) em cubeta de 0,2 cm, com os parâmetros: 1,5 Kv, 25 µF, 200 Ω, seguindo as instruções do fabricante (Bio-Rad). Em seguida, foi adicionado 1 mL de sorbitol 1M e incubado a 30°C por 2 horas sem agitação. As células transformadas foram plaqueadas em meio de cultivo YEPDS (Extrato de levedura 1%, Peptona 2%, Dextrose 2%, Ágar 2% e Sorbitol 1M) contendo o antibiótico zeocina em duas concentrações distintas, 100 e 500 µg/mL por 48 horas a 30°C, dessa forma, clones recombinantes contendo as sequências de interesse foram identificados.

4.2.15 Seleção das colônias transformadas

As colônias selecionadas nas placas de Petri foram avaliadas quanto a presença dos plasmídeos recombinantes por meio da PCR de colônia ou reação de restrição, nas condições descritas anteriormente. Para as colônias positivas foram feitos estoques em glicerol 50% e armazenados a -80°C.

4.2.16 Sequenciamento

Os plasmídeos recombinantes foram sequenciados pelo método de Sanger (Sanger et al. 1977), método da terminação de cadeia por dideoxynucleotídeos marcados com fluorocromos (kit BigDye Terminator v3.1 Cycle Sequencing). O sequenciamento foi realizado de acordo com as recomendações do fabricante em um sequenciador automático ABI 3500 xL Genetic Analyzer com tecnologia Applied Biosystems (Life Technologies). Os primers utilizados na reação de sequenciamento estão localizados no vetor pET28a, T7 forward (Invitrogen) e T7 Reverse (Invitrogen). Para essas reações foram utilizados 2 µL de BigDye Terminator 3.1, 2 µL de tampão de sequenciamento, 1 µL de primer, 1 µL de DNA com concentração de 100 a 200 ng e água autoclavada para um volume final de 10 µL. Para as sequências clonadas no vetor pPICZαA os primers utilizados foram AOX1 Forward (5'-GACTGGTTCCAATTGACAAGC-3') e AOX1 Reverse (5'-GCAAATGGCATTCTGACATCC-3'). Para essas reações de sequenciamento foram utilizados 2 µL de BigDye Terminator 3.1, 3 µL de tampão de sequenciamento, 1 µL de cada primer, 8,36 µL de DNA da gp120 e 12,77 µL de DNA da gp41, com concentração final de aproximadamente 400 ng e água autoclavada para um volume final de 20 µL.

As sequências obtidas foram analisadas através do programa Bioedit Sequence Alignment Editor vc7.2.5 (Hall, 2013) que faz o alinhamento das sequências geradas e forma a sequência consenso. A sequência consenso é então colocada no programa BLAST (Basic Local Alignment Search Tool, NCBI) que procura em bancos de dados internacionais (GenBank: <http://www.ncbi.nlm.nih.gov>) genes e proteínas similares.

4.2.17 Expressão das proteínas recombinantes gp41 e gp120

Para expressão em sistema *E. coli*, no dia anterior a expressão, foram feitos os pré-inóculos respectivos de cada sequência de interesse e o meio LB (0,5 g de extrato de levedura, 0,5 g de triptona, 1 g de NaCl e 100 mL de água) foi preparado em um erlenmeyer de 500 mL e levado para ser autoclavado. Foi colocado em um pré-inóculo de 5 mL de meio LB 5 µL de kanamicina (100 mg/mL) e 170 µL de cloranfenicol (1 mg/mL), em seguida foi inoculado uma

colônia de BL21(DE3)pLYSs_pET28a_gp41 ou Rosetta(DE3)_pET28a_gp41 em um tubo falcon e em outro falcon uma colônia BL21(DE3)pLYSs_pET28a_gp120 ou Rosetta(DE3)_pET28a_gp120. Por fim, esses pré-inóculos foram incubados de 16 a 18h horas a 25°C no período noturno e 37°C no início da manhã, a 200 rpm.

No meio LB autoclavado foi adicionado 100 μ L de kanamicina (100 mg/mL), 3400 μ L de cloranfenicol (1 mg/mL) e 400 μ L do pré-inóculo (proporção de 4 mL de pré-inóculo para cada 1L de meio LB). Posteriormente, foi levado para o shaker em agitação orbital a 200 rpm por aproximadamente 3 horas até atingir uma D.O. 600 nm de 0,4 a 0,6. Quando atingiu a D.O. esperada, a cultura foi dividida em falcons de 50 mL, sendo colocado em cada falcon 10 mL de cultura. A indução foi realizada em três diferentes temperaturas: 37°C, 30°C e 25°C. Foram utilizadas concentrações crescentes de IPTG em cada falcon: sem IPTG, 0,1 mM de IPTG, 0,5 mM de IPTG e 1 mM de IPTG. E as alíquotas foram retiradas após 4 e 8 horas de indução.

Logo após, as alíquotas foram centrifugadas a 3095 x g, a 4°C por 10 minutos. O sobrenadante foi descartado e o pellet ressuscitado em 600 μ L de tampão de lise (100 mM de NaCl). Em seguida, as amostras foram sonicadas quatro vezes com pulsos de 10 segundos e intervalos do mesmo tempo no gelo entre cada pulso. Após a sonicação, as amostras foram levadas para serem centrifugadas por 20 minutos a 16000 x g 4°C. Por fim, foram separados o sobrenadante do pellet e o pellet foi ressuscitado em 600 μ L uréia 8M. As amostras devidamente identificadas foram armazenadas a -20°C.

Antes de iniciar a expressão em *Pichia pastoris* foi realizado um teste de expressão em pequena escala (screening) para identificar a colônia mais produtora de cada proteína. Foram escolhidas 5 colônias recombinantes da proteína gp41 e 5 colônias recombinantes da proteína gp120. Essas colônias foram mantidas em meio de cultivo BMGY (1% de Extrato de Levedura, 2% Peptona, 100 mM de Tampão Fosfato pH 6,0, 1,34% de YNB, 4 x 10⁻⁵% Biotina), adicionado glicerol 1% como única fonte de carbono por 48 horas a 30°C em falcons contendo aproximadamente 10 mL do meio em cada um. Logo após, foi realizada a indução com o meio BMMY (1% de Extrato de Levedura, 2% Peptona, 100 mM de Tampão

Fosfato pH 6,0, 1,34% de YNB, 4×10^{-5} % Biotina) contendo metanol como única fonte de carbono numa concentração final de 0,75%, com aproximadamente 7 mL de meio em cada falcon. Foi adicionado diariamente ao meio BMMY metanol absoluto 0,75% e retirado 100 μ L do meio (0, 24, 48, 72, 96, 120 e 144 horas) para posterior análise em SDS-PAGE. O sobrenadante foi separado do pellet por centrifugação a 21°C, 16000 x g por 5 minutos e armazenados a -20°C. Foi aplicado no gel de poliacrilamida 20 μ L de cada amostra. A colônia da levedura transformada que mostrou maior produção de cada proteína recombinante foi selecionada para o processo de indução em maior volume.

Após o teste de expressão em pequena escala e da identificação de uma colônia produtora de cada proteína recombinante será feito a produção das mesmas em maior escala.

4.2.18 Eletroforese em gel de poliacrilamida

Para eletroforese em gel de poliacrilamida, as amostras foram retiradas do -20°C e deixadas em gelo para descongelar. Foi realizado uma solução para a eletroforese SDS-PAGE com 20 μ L de cada amostra e 10 μ L de tampão de amostra. Em seguida, essas soluções foram incubadas a 100°C por 10 minutos. Foi aplicado no gel de poliacrilamida contendo SDS (Dodecil Sulfato de Sódio) 10 μ L de cada solução e submetido a eletroforese a amperagem constante de 80 V por aproximadamente 2 horas.

O gel de empilhamento foi composto por Tris-HCl 0,5 M pH 6,8, SDS 10%, acrilamida e bis-acrilamida (30:8), TEMED e persulfato de amônio 10%. Foi feito o gel de separação a 12% e 15% e foi composto por Tris-HCl pH 8,8, SDS 10%, acrilamida e bis-acrilamida (30:8), TEMED e persulfato de amônio 10%. A eletroforese foi realizada em equipamento Mini Protean Tetra System (Bio Rad) utilizando tampão de corrida composto por Tris 0,125 M, glicina 0,96 M, água e SDS 0,1%. A voltagem foi mantida constante a 80 V por tempo suficiente para separação das frações proteicas.

4.2.19 Coloração dos géis de poliacrilamida com Coomassie® Brilliant Blue R 250

Após a eletroforese os géis foram corados com solução corante com Coomassie Blue (Coomassie Brilliant Blue R 250 (Bio Rad), etanol 100%, ácido acético e água q.s.p. 800 mL) overnight. No dia seguinte, os géis foram descorados com solução descorante (Ácido acético e etanol 100%) até apresentarem coloração apenas nas proteínas.

4.2.20 Detecção das proteínas recombinantes por *Western blotting*

Após a separação das proteínas por eletroforese em gel de poliacrilamida, as mesmas foram transferidas para a membrana de nitrocelulose (Bio-Rad). A transferência foi realizada em equipamento Mini Protean Tetra System (Bio Rad) utilizando tampão de transferência (2,5 g de Tris, 4,13 g de glicina, 165 mL de metanol e 810 mL de água). A amperagem foi mantida constante a 150 mA por aproximadamente 2 horas. Após a transferência, a membrana foi bloqueada por 1 hora com solução de bloqueio composta por TBS (Tris Buffered Saline) e leite em pó desnatado 5% para saturação dos sítios inespecíficos. Em seguida, a membrana foi lavada 3 vezes com TBS sob agitação por 5 minutos. A membrana foi incubada a 4°C com o anticorpo primário diluído em TBS overnight. No dia seguinte, a membrana foi lavada novamente 3 vezes com TBS sob agitação por 5 minutos. Para detecção do complexo antígeno anticorpo, a membrana foi incubada por 1 hora e 30 minutos com o anticorpo secundário conjugado com fosfatase alcalina (AP) diluído em TBS. Realizou-se a última lavagem com TBS 3 vezes sob agitação por 5 minutos e por fim, foi adicionado a Solução Reveladora BCIP/NBT Solution, premixed (Sigma-Aldrich) que contém o substrato da enzima.

Anticorpo	Diluição	Marca
Anti-gp120 (Mouse)	1:1300	Abcam
Anti-mouse conjugado com AP	1:5000	Bio-Rad

Tabela 3 - Diluição dos anticorpos utilizados no *Western Blotting*

4.2.21 Purificação das proteínas recombinantes por cromatografia de afinidade

A purificação foi realizada por cromatografia de afinidade em uma coluna de resina carregada com níquel HisTrap™ HP 1 mL (GE Healthcare), onde as proteínas ficam aderidas à coluna pela cauda de histidina por afinidade química. Inicialmente a coluna foi equilibrada com solução tampão (Tris 100 mM, NaCl 100 mM, NaH₂PO₄ 17,4 mM). Em seguida, 5 mL da amostra filtrada do sobrenadante da indução de 144 horas em *Pichia pastoris* foi colocada na coluna e uma lavagem com a solução tampão foi realizada para retirar o que não se ligou à coluna. Uma segunda lavagem com solução de Imidazol 5 mM (Tris 100 mM, NaCl 100 mM, NaH₂PO₄ 17,4 mM, Imidazol 5 mM) foi feita para retirar as proteínas que se ligaram fracamente a coluna. Para coletar a proteína recombinante a eluição foi realizada com uma solução de imidazol 250 mM (Tris 100 mM, NaCl 100 mM, NaH₂PO₄ 17,4 mM, Imidazol 250 mM). O imidazol compete com a histidina pelo níquel, fazendo com que a proteína seja solta da coluna.

4.2.22 Detecção das proteínas recombinantes por *Dot Blotting*

Um *Dot Blotting* foi realizado para detectar a proteína no sobrenadante produzido após a indução (antes da purificação). Para isso foram utilizados os anticorpos identificados na tabela 4.

Anticorpo	Diluição	Marca
Anti-gp120 (Mouse)	1:500	Abcam
Anti-gp41 (Mouse)	1:500	Abcam
Anti-mouse conjugado com AP	1:5000	Bio-Rad

Tabela 4 - Diluição dos anticorpos utilizados no *Dot Blotting*

O *Dot Blotting* foi realizado em uma membrana de nitrocelulose 0,45 µm (BioRad), onde com o auxílio de um lápis marcou-se duas regiões onde seria feito o blot. Nessas regiões foram pipetados 15 µL do sobrenadante de 144 horas da indução em *Pichia pastoris* e 15 µL da

fração com a proteína purificada. Em seguida, esperou-se a membrana secar e foi realizado o bloqueio dos sítios não específicos com leite em pó desnatado 5% diluído com TBS-T (0,05% Tween 20 em TBS) por cerca de 1 hora. Após o bloqueio, a membrana foi lavada com TBS-T 3 vezes e em seguida, incubada com o anticorpo primário dissolvido em TBS por 30 minutos a temperatura ambiente. Novamente, a membrana foi lavada 3 vezes com TBS-T e foi incubada com o anticorpo secundário anti-mouse conjugado com fosfatase alcalina (AP) por 30 minutos a temperatura ambiente. Por fim, realizou-se mais 3 lavagens com TBS-T e a revelação foi feita com a Solução Reveladora BCIP/NBT Solution, premixed (Sigma-Aldrich) de acordo com as recomendações do fabricante.

5. RESULTADOS E DISCUSSÃO

5.1 Da clonagem no sistema de expressão procarioto *E. coli*

5.1.1 Clonagem dos fragmentos gênicos codificantes das proteínas gp41 e gp120 em vetor pET28a

As ampliações das sequências gênicas codificadoras das proteínas gp41 e gp120 foram realizadas com primers específicos (tabela 1) e o resultado das reações está mostrado na figura 11. Observa-se nas colunas 1 e 2 que houve amplificação do fragmento com tamanho esperado de cerca de 591 pb do gene codificante da proteína gp41. A temperatura de anelamento para esse fragmento, 54°C, foi eficiente visto que houve a amplificação do gene de interesse nas duas concentrações de DNA molde testadas.

Nas colunas 3 e 4 podemos observar que houve a amplificação do fragmento de aproximadamente 1434 pb do gene da proteína gp120. Para este fragmento, a temperatura de anelamento foi 53°C, mostrando-se eficiente para amplificação do gene de interesse nas duas concentrações testadas. O vetor pAbT4674 foi usado como DNA molde para as ampliações citadas.

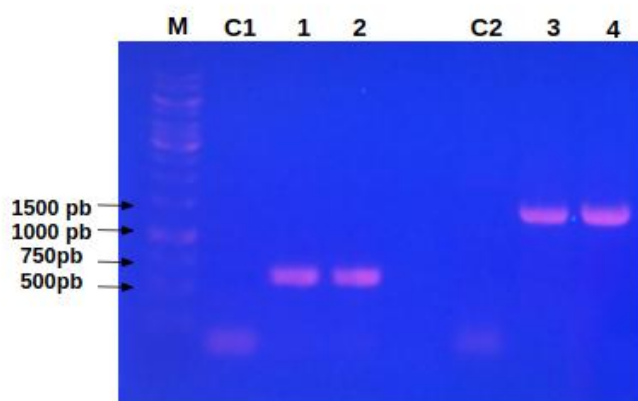


Figura 11 - Foto do gel de Agarose 1% corado com Gel Red mostrando a amplificação do fragmento de gp41 (591pb) a 54°C (Amostras 1 e 2) e da gp 120 (1434 pb) a 53°C (Amostras 3 e 4) a partir do vetor pAbT4674 (ATCC). M: Marcador de peso molecular GeneRuler 1Kb Ladder (Thermo Scientific); C1: Controle negativo gp41; 1: Amostra gp41 (78 ng/μL); 2: Amostra gp41 (7,8 ng/μL); C2: Controle negativo gp120; 3: Amostra gp120 (95,6 ng/μL); 4: Amostra gp120 (9,56 ng/μL).

As reações de amplificação dos fragmentos gênicos de interesse foram reunidas e purificadas. Realizou-se a reação de restrição do vetor de expressão e os DNA de interesse foram purificados do gel para que pudesse ocorrer a reação de ligação com o vetor pET28a. Após a reação de ligação, os plasmídeos recombinantes foram utilizados para transformar bactérias de *E. coli* da linhagem DH10β. As bactérias transformadas foram semeadas em meio LB ágar com o antibiótico kanamicina. A clonagem e geração dos plasmídeos recombinantes pET28_gp41 e pET28_gp120 foram confirmados por PCR de colônia (Figura 12 e 13). Observa-se na figura 11 amplificação de DNA provenientes de colônias recombinantes que receberam o gene da gp41. O tamanho do inserto esperado é de 591pb e apenas uma colônia não teve a amplificação esperada.

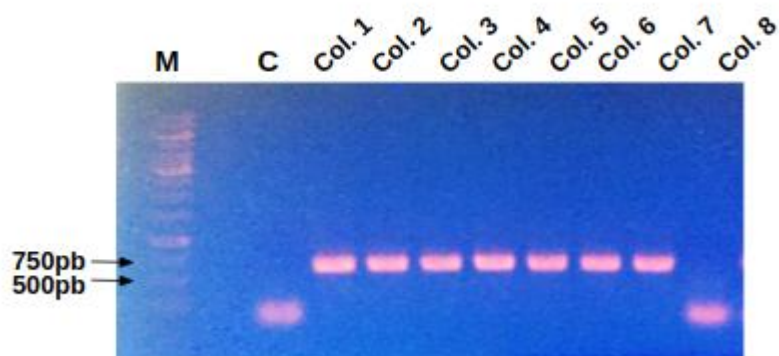


Figura 12 - Foto do gel de agarose 1% corado com Gel Red mostrando a amplificação de DNA proveniente de colônias recombinantes que receberam o gene da gp41. M: Marcador de peso molecular GeneRuler 1Kb Ladder (Thermo Scientific); C: Controle negativo; Resultado das reações de PCR direto das colônias 1, 2, 3, 4, 5, 6 e 7: colônias recombinantes; Colônia 8: colônia sem a amplificação esperada.

Na figura 13, observa-se a amplificação de DNA provenientes de colônias recombinantes que receberam o gene da gp120. O tamanho do inserto esperado é de 1434 pb e apenas uma colônia não teve a amplificação esperada.

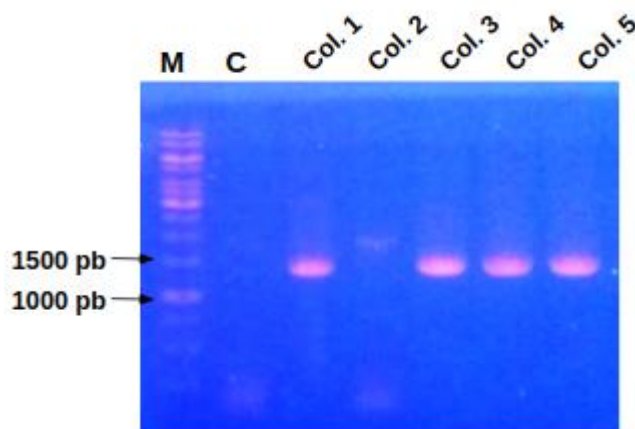


Figura 13 - Foto do gel de agarose 1% corado com GelRed mostrando a amplificação de DNA provenientes de colônias recombinantes que receberam o gene da gp120. M: Marcador de peso molecular GeneRuler 1Kb Ladder (Thermo Scientific); C: Controle negativo; Resultado das reações de PCR direto das colônias 1, 3, 4 e 5: colônias recombinantes; Colônia 2: colônia sem a amplificação esperada.

Desse modo, com a confirmação das colônias recombinantes através da PCR de colônia pode-se dar continuidade aos experimentos.

5.1.2 Análise do Sequenciamento

Após a construção dos plasmídeos recombinantes, os mesmos foram sequenciados via método de Sanger (Sanger et al. 1977). Foram usadas as plataformas BLAST e Bioedit para análises dos sequenciamentos. A tabela 5 mostra a identidade da proteína recombinante gp41 comparada com a sequência de referência através da plataforma BLAST.

Proteína	Identidade
gp41	99%

Tabela 5 - Identidade da proteína gp41.

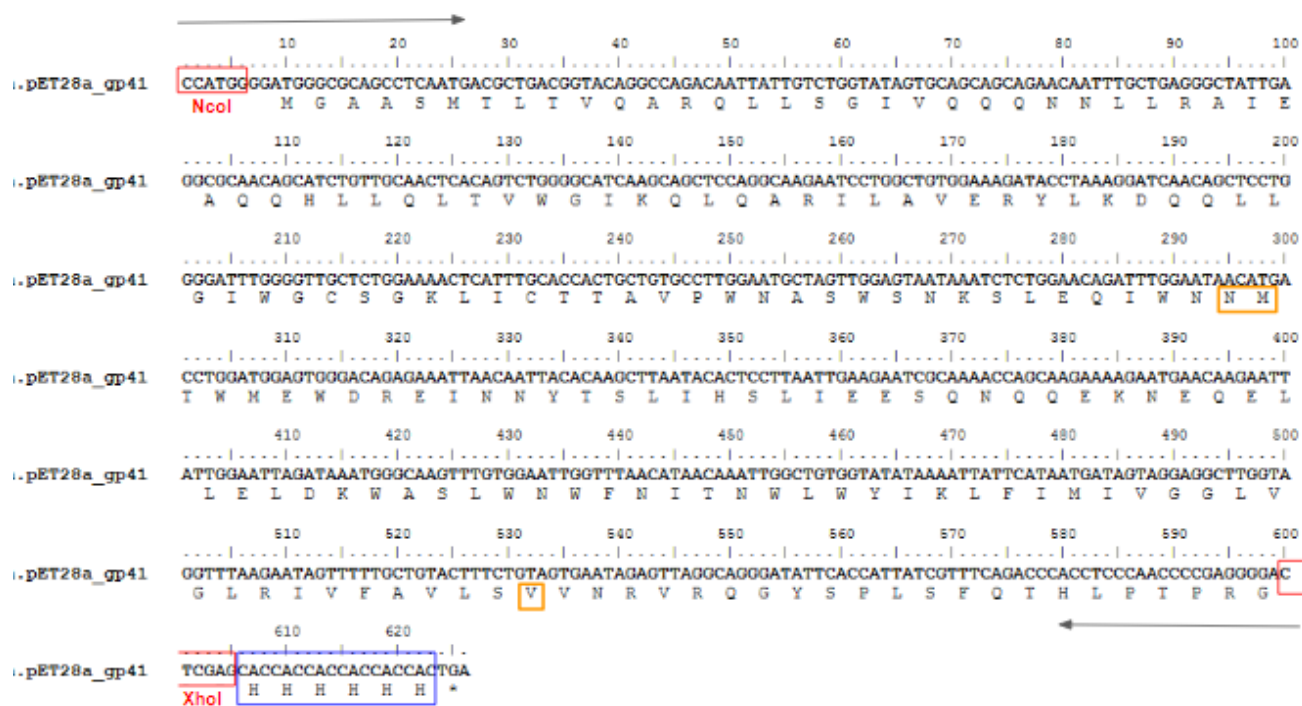


Figura 14 - Análise dos aminoácidos da proteína recombinante gp41. A caixa vermelha indica os sítios de restrição das enzimas NcoI e XhoI, que foram inseridos pelos primers, *forward* e *reverse*, respectivamente. As setas em preto indicam as regiões dos primers. A caixa azul mostra a cauda de histidina, em laranja os aminoácidos que foram trocados e o asterisco indica o stop códon.

Através da análise do sequenciamento pela plataforma BLAST do gene codificante para proteína gp41 observa-se que houve a troca de 3 nucleotídeos nas posições 295, 299 e 533, apresentando assim identidade de 99% com a proteína gp41 (Tabela 5). Após a tradução do

gene da proteína gp41 recombinante através do programa Bioedit (Figura 14) foi identificado a troca de 3 aminoácidos, resultante da troca dos 3 nucleotídeos. Os aminoácidos trocados foram histidina (H) por asparagina (N) na posição 295, treonina (T) por metionina (M) na posição 299 e isoleucina (I) por valina (V) na posição 533. Desse modo, a proteína recombinante gp41 mostrou 99% de identidade com a proteína gp41 do HIV-1. Portanto, a conclusão que se pode chegar dos resultados do sequenciamento é que a proteína recombinante gp41 foi clonada no vetor pET28a.

O gene da proteína gp120 também foi clonado no vetor pET28a, porém sendo o gene da proteína gp120 muito extenso (1434 pb) não foi possível identificar por completo a parte intermediária da sequência. Foram realizados vários sequenciamentos para tentar identificar melhor tal parte, contudo a técnica utilizada permite o sequenciamento de aproximadamente 700 nucleotídeos consecutivos (BROWN, 2003). O que podemos concluir desse sequenciamento é que o gene da gp120 foi clonado no vetor pET28a, visto que quando analisada pelo programa Bioedit observa-se tanto no início quanto no final da sequência os sítios de restrição das enzimas que foram inseridos nos primers e a cauda de histidina presente no vetor no final da sequência. Neste caso, para identificação da parte intermediária seria necessário desenhar primers específicos para essa região.

5.1.3 Análise da Expressão das proteínas recombinantes gp41 e gp120 em

BL21(DE3)pLYSs

Após a confirmação por sequenciamento que os fragmentos de interesse estavam clonados no vetor pET28a, células de BL21(DE3)pLYSs foram transformadas com estes plasmídeos recombinantes e a expressão foi induzida com IPTG a 37°C. Alíquotas após 4 horas ou 8 horas da indução foram retiradas para posterior análise em SDS-PAGE (Figura 15 e 16).

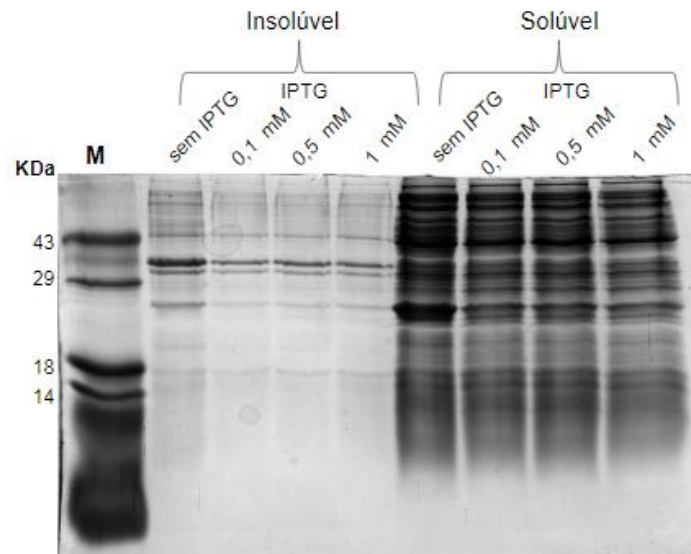


Figura 15 - Gel de poliacrilamida 15% corado com Comassie Blue mostrando as proteínas nas alíquotas após 4 horas da indução da BL21(DE3)pLYSs_pET28a_gp41 a 37°C. M: Marcador de peso molecular Protein Molecular Weight Standards (Life Technologies). À direita encontram-se as proteínas presentes na fração solúvel e à esquerda encontram-se as presentes na fração insolúvel. Indução com concentrações distintas de IPTG demonstradas na imagem.

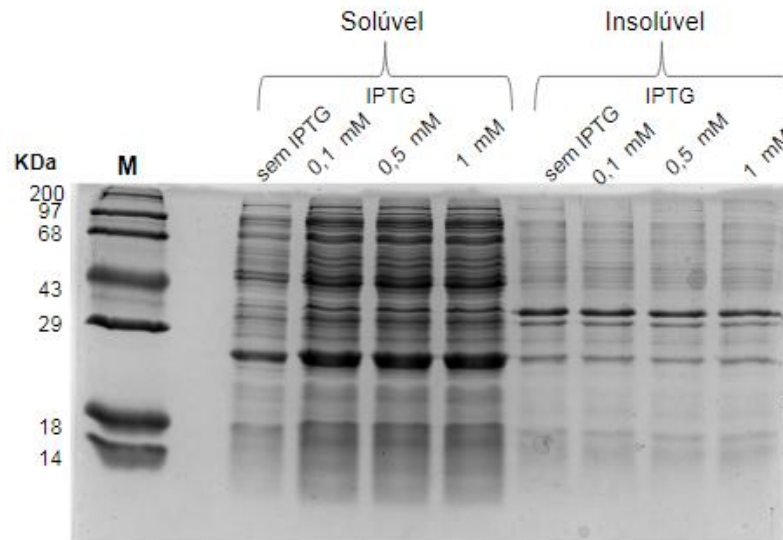


Figura 16 - Gel de poliacrilamida 12% corado com Comassie Blue mostrando as proteínas nas alíquotas após 4 horas da indução da BL21(DE3)pLYSs_pET28a_gp120 a 37°C. M: Marcador de peso molecular Protein Molecular Weight Standards (Life Technologies). À direita encontram-se as proteínas presentes na fração insolúvel e à esquerda encontram-se as presentes na fração solúvel. Indução com concentrações distintas de IPTG demonstradas na imagem.

Segundo Kaur (2018), as concentrações de IPTG de 0 a 1 mM não afetam a taxa de crescimento da *E. coli*, por isso essas concentrações foram escolhidas.

Nas imagens acima (figuras 15 e 16) pode-se observar que as bactérias BL21(DE3)pLYSs com o vetor recombinante pET28a não foram capazes de produzir as proteínas gp41 e gp120 a 37°C, como esperado, mesmo testando várias concentrações de IPTG. Vários fatores podem afetar a produção da proteína recombinante, bem como temperatura de indução, concentração do indutor IPTG, densidade óptica, entre outros. Dessa forma, para tentar otimizar a expressão das proteínas foram testadas outras temperaturas de cultivo após indução. Sabe-se que temperaturas mais baixas tendem a diminuir a formação de corpos de inclusão (KAUR, 2018). Testou-se as temperaturas 30°C e 25°C, porém após análise dos géis por SDS-PAGE (dados não mostrados) não foi visualizado nenhum aumento na faixa de massa molecular de expressão esperado de 22 kDa e 53 kDa para as proteínas recombinantes gp41 e gp120 respectivamente.

5.1.4 Estudo de códons raros

Após tentar sem sucesso expressar as proteínas em linhagem BL21(DE3)pLYSs foi realizado um estudo sobre códons raros presentes nas sequências de interesse. Para identificar e quantificar os códons raros nas sequências gp41 e gp120 foi utilizado a ferramenta disponível em <http://people.mbi.ucla.edu/sumchan/caltor.html> (by Edmund Ng/ UCLA). Na figura 17 estão destacado os códons raros presentes na sequência gp41 e na tabela 6 é mostrada a porcentagem de códons raros na mesma.

Amino Acid	Rare Codon
Arginine	CGA
	CGG
	AGG
	AGA
Glycine	GGA
	GGG
Isoleucine	AUA
Leucine	CUA
Proline	CCC
Threonine	ACG

gp41

ATG GGC GCA GCC TCA ATG **ACG** CTG **ACG** GTA CAG GCC AGA CAA TTA TTG TCT GGT **AUA** GTG CAG CAG CAG AAC AAT TTG CTG **AGG** GCT ATT GAG GCG CAA CAG CAT CTG TTG CAA CTC ACA GTC TGG GGC ATC AAG CAG CTC CAG GCA **AGA** ATC CTG GCT GTG GAA **AGA** TAC **CUA** AAG GAT CAA CAG CTC CTG **GGG** ATT TGG GGT TGC TCT **GGA** AAA CTC ATT TGC ACC ACT GCT GTG CCT TGG AAT GCT AGT TGG AGT AAT AAA TCT CTG GAA CAG ATT TGG AAT AAC ATG ACC TGG ATG GAG TGG GAC AGA GAA ATT AAC AAT TAC ACA AGC TTA **AUA** CAC TCC TTA ATT GAA GAA TCG CAA AAC CAG CAA GAA AAG AAT GAA CAA GAA TTA TTG GAA TTA GAT AAA TGG GCA AGT TTG TGG AAT TGG TTT AAC **AUA** ACA AAT TGG CTG TGG TAT **AUA** AAA TTA TTC **AUA** ATG **AUA** GTA **GGA** GGC TTG GTA GGT TTA **AGA** **AUA** GTT TTT GCT GTA CTT TCT GTA GTG AAT AGA GTT **AGG** CAG **GGA** TAT TCA CCA TTA TCG TTT CAG ACC CAC CTC CCA ACC CCG **AGG** **GGA**

Figura 17 - Localização dos códons raros presentes na sequência gp41 em colorido. Ao lado, a legenda de cada cor correspondente a tais códons.

Aminoácido	Códon Raro em <i>E. coli</i>	Nº de ocorrência de códons raros na sequência gp41	Nº de códons referente ao aminoácidos na sequência gp41	% de códons raros na sequência gp41	Frequência de códons raros a cada 1000 códons em <i>E. coli</i>*
Arginina	CGA	0	9	-	3,1
	CGG	0		-	4,6
	AGG	3 vezes		33,3%	1,4
	AGA	6 vezes		66,6%	2,1
Glicina	GGA	4 vezes	11	36,3%	7,0
	GGG	1 vez		9,09%	9,7
Isoleucina	AUA	7 vezes	14	50%	4,1
Leucina	CUA	1 vez	29	3,44%	3,2
Prolina	CCC	0	-	-	4,3
Treonina	ACG	2 vezes	10	20%	1,4

Tabela 6 - Frequência de códons raros na sequência gp41. *Frequência de códons (KANE, 1995; SHARP, P. M. & LI, W., 1986).

Com o estudo de códons raros para a proteína gp41 pode-se observar, por exemplo, que dos 29 códons que codificam para leucina presentes na sequência, apenas 1 deles, o que equivale a 3,44% dos códons de leucina, é considerado raro (CUA). Segundo Kane (1995), a frequência de códons raros a cada 1000 códons em *E. coli* é de 3,2 para leucina. Desse modo, a porcentagem 3,44 é superior a frequência identificada por Kane. Aplicando essa explicação para os outros códons raros presentes nessa sequência, observamos que todas as porcentagens são superiores aos valores das frequência a cada 1000 códons, o que indica que os códons raros estão em número significativo na sequência gp41 e,

consequentemente o sistema de expressão procariótico pode não estar produzindo a sequência devido a esse fator: a presença de muitos códons raros.

Na figura 18 estão destacado os códons raros presentes na sequência gp120 e na tabela 7 é mostrada a porcentagem de códons raros na mesma.



Figura 18 - Localização dos códons raros presentes na sequência gp120 em colorido. Ao lado, a legenda de cada cor correspondente a tais códons.

Aminoácido	Códon Raro em <i>E. coli</i>	Nº ocorrência de códons raros na sequência gp120	Nº de códons referente ao aminoácido na sequência gp120	% de códons raros na sequência gp120	Frequência de códons raros a cada 1000 códons em <i>E. coli</i> *
Arginina	CGA	0	22	-	3,1
	CGG	0		-	4,6
	AGG	2 vezes		9,09%	1,4
	AGA	20 vezes		90,9%	2,1

Glicina	GGA	17 vezes	25	68%	7,0
	GGG	7 vezes		28%	9,7
Isoleucina	AUA	16 vezes	32	50%	4,1
Leucina	CUA	5 vezes	27	18,5%	3,2
Prolina	CCC	6 vezes	22	27,2%	4,3
Treonina	ACG	3 vezes	31	7,14%	1,4

Tabela 7 - Frequência de códons raros na sequência gp120. *Frequência de códons (KANE,1995; SHARP, P. M. & LI, W., 1986).

A mesma linha de raciocínio aplicada na sequência gp41 pode ser utilizada para análise dos códons raros presentes na sequência gp120. Observando-se as duas últimas colunas da tabela 7 nota-se que a porcentagem de códons raros na sequência é superior a frequência de códons raros a cada 1000 códons, o que implica na presença de um número significativo de códons raros na sequência gp120. Dessa forma, uma possível explicação para que a bactéria *E. coli* da linhagem BL21(DE3)pLYSs não esteja produzindo as proteínas de interesse seja a presença de muitos códons raros nas sequências.

A partir da análise de códons raros pode-se construir a tabela 8.

Proteína	Nº total de códons	Nº total de códons raros	% de códons raros
gp41	197	24	12,1
gp120	478	76	15,8

Tabela 8 - Estudo de códons raros

De acordo com a tabela 7, o gene das proteínas recombinantes gp41 e gp120 apresentaram respectivamente 12,1% e 15,8% de códons raros para *E. coli*. Os principais códons raros para *E. coli* estão exemplificados nas tabelas 5 e 6 (KANE, 1995; MAKRIDES, 1996;

SHARP, P. M. & LI, W., 1986). A presença de códons raros pode limitar a produção das proteínas devido ao baixo número de RNAt para esses códons presentes na *E. coli* e resultar em uma baixa produção da proteína heteróloga (KANE, 1995; SORENSEN, 2005).

Quando há na sequência um agrupamento de códons raros, ou seja, códons raros consecutivos, erros de tradução podem ocorrer, como por exemplo, substituições de aminoácidos sem tradução ou terminação da tradução prematuramente (SOSENSEN, 2005).

Dessa maneira, para tentar solucionar o problema dos códons raros optou-se pela estratégia de mudança de linhagem bacteriana para uma que codifique os códons raros.

5.1.5 Análise da Expressão das proteínas recombinantes gp41 e gp120 em Rosetta(DE3)

A linhagem escolhida foi a Rosetta (DE3). Existem outras linhagens de *E. coli* que codificam códons raros como a linhagem BL21-CodonPlus, porém trabalhos anteriores mostraram que não foi possível produzir as proteínas nesta linhagem (DAVI, 2015).

Os plasmídeos recombinantes pET28a_gp41 e pET28a_gp120 foram extraídos da linhagem BL21(DE3)pLYSs e inseridos por transformação na Rosetta(DE3). Para confirmação da transformação foi realizada uma PCR de colônia (Figura 19).

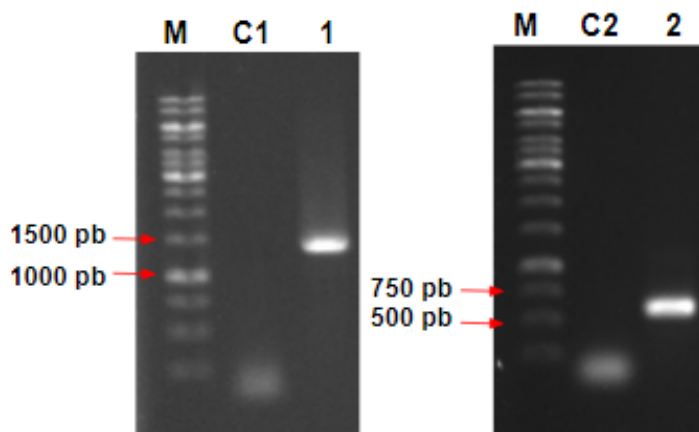


Figura 19 - Gel de agarose 1% corado com Gel Red mostrando a confirmação da transformação das bactérias com a presença dos plasmídeos pET28a_gp41 e pET28a_120 em Rosetta(DE3). Amplificação por PCR de colônia. M: Marcador de peso molecular GeneRuler 1Kb Ladder (Thermo

Scientific); C1: Controle negativo da gp120; 1: colônia recombinante gp120 (1434pb); C2: Controle negativo da g41; 2: colônia recombinante gp41(591pb).

Após a confirmação da transformação foi feita a expressão das proteínas conforme a descrição no item 4.2.16. Foram testadas concentrações crescentes de IPTG: sem IPTG, 0,1 mM, 0,5 mM e 1 mM. Alíquotas de 1mL foram retiradas após 4 e 8 horas de indução. Em seguida, essas alíquotas foram ressuspendidas em tampão de SDS-PAGE e aplicadas em gel de poliacrilamida. As massas moleculares esperadas são de aproximadamente 22 kDa para gp41 e 53 kDa para gp120. A expressão após 4 horas de indução das bactérias Rosetta(DE3)_pET28a_gp41 a 37°C está exemplificada na figura 20. Na figura 21 tem-se as alíquotas de 4 horas após a indução de Rosetta(DE3)_pET28a_gp120.

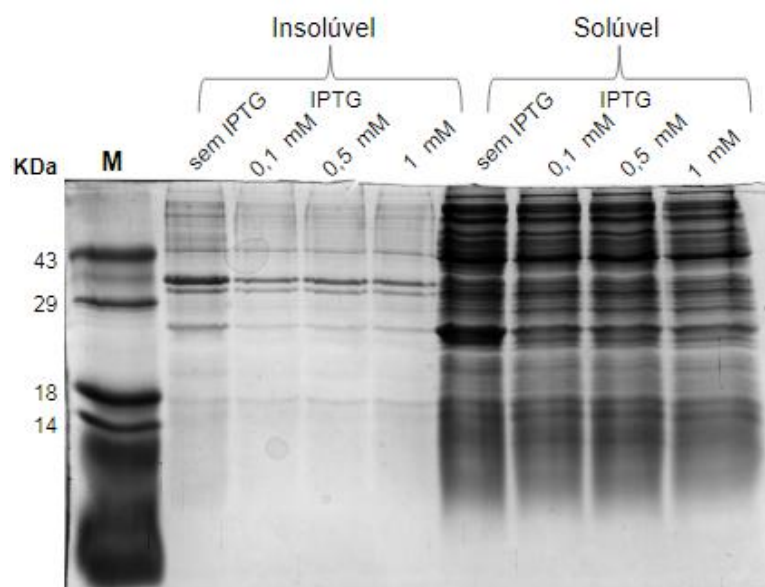


Figura 20 - Gel de poliacrilamida 15% corado com Comassie Blue mostrando as proteínas nas alíquotas após 4 horas da indução da Rosetta(DE3)_pET28a_gp41 a 37°C. M: Marcador de peso molecular Protein Molecular Weight Standards(Life Technologies). À direita encontram-se as proteínas presentes na fração solúvel e à esquerda encontram-se as presentes na fração insolúvel. Indução com concentrações distintas de IPTG demonstradas na imagem.

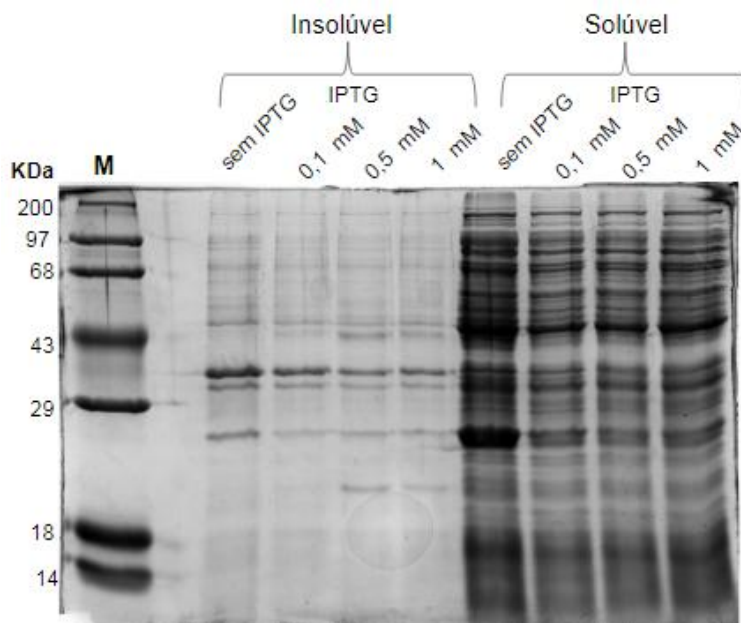


Figura 21 - Gel de poliacrilamida 12% corado com Comassie Blue mostrando as proteínas nas alíquotas após 4 horas da indução da Rosetta(DE3)_pET28a_gp120 a 37°C. M: Marcador de peso molecular Protein Molecular Weight Standards (Life Technologies). À direita encontram-se as proteínas presentes na fração solúvel e a esquerda encontram-se as presentes na fração insolúvel. Indução com concentrações distintas de IPTG demonstradas na imagem.

Bem como as bactérias da linhagem BL21(DE3)pLYSs, as bactérias da linhagem Rosetta(DE3) com os vetores recombinantes não foram capazes de produzir as proteínas gp41 e gp120 a 37°C como mostram as figuras 20 e 21. Outras temperaturas, como 30°C, 25°C e 20°C, foram testadas a fim de tentar otimizar o processo de indução, porém após análise dos géis por SDS-PAGE (dados não mostrados) não foi visualizado nenhum aumento de expressão com a massa molecular esperada. Mesmo não sendo possível a visualização de um aumento de expressão significativo foi realizado um *Western Blotting* na tentativa de detecção da proteína recombinante gp120, mesmo que em pequenas quantidades. O *Western* foi realizado com anticorpo específico (tabela 2), entretanto nenhuma marcação pode ser visualizada, o que indica que a proteína não estava sendo expressa.

Uma possível explicação para isso é que as sequências das proteínas gp41 e gp120 possuem muitos nucleotídeos C-G. A *E. coli* possui RNAt preferenciais para códons que não tenham nucleotídeos C e G. Dessa forma, a presença de códons raros juntamente com a

presença de muitos nucleotídeos C-G atrapalham a expressão das proteínas. Uma alternativa para isso seria utilizar a linhagem bacteriana CodonPlus-RP que é indicada para sequências que possuem uma grande quantidade de C-G. Uma outra estratégia seria fazer a otimização de códons, pois os códons raros seriam trocados por códons preferenciais pela *E. coli* e também seria otimizado a presença de C-G diminuindo assim seu conteúdo (KAUR, 2018). As sequências de interesse das proteínas gp41 e gp120 não foram otimizadas previamente, pois elas não foram sintetizadas e sim extraídas do vetor pABT4674 (Figura 8). Possivelmente, devido a esse fato a expressão dessas proteínas apresentou-se insatisfatória no sistema procarioto.

Uma outra hipótese para a não produção das proteínas recombinantes é que as proteínas estariam sendo degradadas por proteases do hospedeiro devido a falta de enovelamento correto ou o término da tradução antecipadamente. Esse mecanismo é utilizado pela *E. coli* com a finalidade de não haver acúmulo dessas proteínas mal dobradas (BAYNEX & MUJACIC, 2004).

Em vista disso, optou-se pela troca do sistema de expressão procarioto *E. coli* para o sistema de expressão eucarioto *Pichia pastoris*.

5.2 Da clonagem no Sistema de expressão eucarioto *P. pastoris*

5.2.1 Clonagem dos fragmentos gênicos codificantes das proteínas gp41 e gp120 em vetor pPICZ α A

As ampliações das sequências gênicas codificadoras das proteínas gp41 e gp120 foram realizadas com primers específicos (tabela 2) e o resultado das reações está mostrado na figura 22. Observa-se na coluna 1 que houve a amplificação do fragmento com tamanho esperado de cerca de 591 pb do gene da proteína gp41. Na coluna 2 podemos observar que houve a amplificação do fragmento de aproximadamente 1449 pb do gene da proteína gp120. Para ambos os fragmentos a temperatura de anelamento de 53°C mostrou-se eficiente, uma vez que houve a amplificação dos genes de interesse. As ampliações ocorreram a partir do vetor pABT4674.

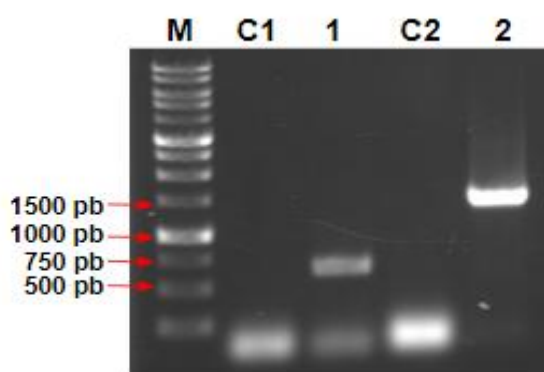


Figura 22 - Foto do gel de Agarose 1% corado com GelRed mostrando a amplificação do fragmento da gp41 (591pb) a 53°C (Amostra 1) e da gp 120 (1449pb) a 53°C (Amostra 2) a partir do vetor pAbT4674 (ATCC). M: Marcador de peso molecular GeneRuler 1Kb Ladder (Thermo Scientific); C1: Controle negativo gp41; 1: Amostra gp41 (22 ng/μL); C2: Controle negativo gp120; 2: Amostra gp120 (22 ng/μL).

As reações de amplificação dos fragmentos gênicos de interesse foram reunidas e purificadas. Realizou-se a reação de restrição no vetor pPICZαA e os DNA de interesse foram purificados do gel para que pudesse ocorrer a reação de ligação no vetor de expressão. Após a reação de ligação, os plasmídeos recombinantes foram utilizados para transformar bactérias de *E. coli* da linhagem DH10β. As bactérias transformadas foram semeadas em meio LB ágar low salt com o antibiótico zeocina. A clonagem e geração dos plasmídeos recombinantes pPICZαA_gp41 e pPICZαA_gp120 foram confirmados por restrição com as enzimas EcoRI e Sall (Figura 23).

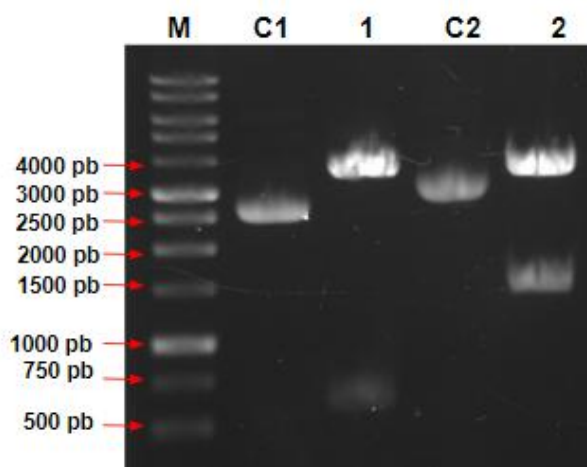


Figura 23 - Gel de agarose 1% corado com Gel Red mostrando os fragmentos digeridos pelas enzimas EcoRI e Sall. M: Marcador de peso molecular 1Kb DNA Ladder (KASVI). C1 e C2: Controle

negativo do plasmídeo pPICZ α A (3593pb) fechado com o gene da gp41 e gp 120, respectivamente.
 1: Restrição do plasmídeo com o gene da gp41 (591pb). 2: Restrição do plasmídeo com o gene da gp120 (1449pb).

5.2.2 Análise do Sequenciamento

A análise do sequenciamento dos genes das proteínas gp41 e gp120 clonado em vetor pPICZ α A foi realizada usando o programa BLAST e o programa Bioedit. A tabela 9 mostra a identidade da proteína recombinante gp41 comparada com a sequência de referência através da plataforma BLAST.

Proteína	Identidade
gp41	99%

Tabela 9 - Identidade da proteína gp41.

Através da análise do sequenciamento pela plataforma BLAST do gene codificante para proteína gp41 observa-se que houve a troca de 1 nucleotídeo na posição 286. Após a tradução do gene da proteína gp41 recombinante usando o programa Bioedit foi identificado que o nucleotídeo alterado gerou a troca de um aminoácido. O aminoácido trocado foi metionina (M) por treonina (T). Dessa forma, a proteína recombinante gp41 mostrou 99% de identidade com a proteína gp41 do HIV-1. Ainda pela análise realizada pelo Bioedit foi possível identificar os sítios de restrição das enzimas que foram inseridos pelos primers e a cauda de histidina presente no vetor. Portanto, a conclusão que se pode chegar dos resultados desse sequenciamento é que a proteína recombinante gp41 apesar de apresentar uma mutação foi clonada no vetor pPICZ α A.

Bem como no sequenciamento da gp120 no vetor pET28a, o sequenciamento da gp120 no vetor pPICZ α A também não possibilitou a identificação da parte intermediária da sequência. O que pode ser identificado nesse sequenciamento é que o sítio da enzima de restrição EcoRI estava presente no início do sequenciamento seguido do início da sequência gp120. No final do sequenciamento só pode ser identificado o final da sequência gp120, não foi identificado o sítio de restrição para enzima Sall e nem a cauda de histidina.

5.2.3 Análise da Expressão das proteínas recombinantes gp41 e gp120

Após a confirmação por sequenciamento que os fragmentos de interesse estavam clonados no vetor pPICZ α A, colônias de *Pichia pastoris* foram transformadas com os plasmídeos recombinantes. Foram escolhidas 5 colônias recombinantes de cada proteína para expressão em menor escala. A indução da expressão foi feita com metanol a 30°C por 144 horas. Alíquotas diárias de 1 mL foram retiradas para posterior análise em SDS-PAGE. As figuras 24 e 25 mostram a expressão das colônias escolhidas após 144 horas.

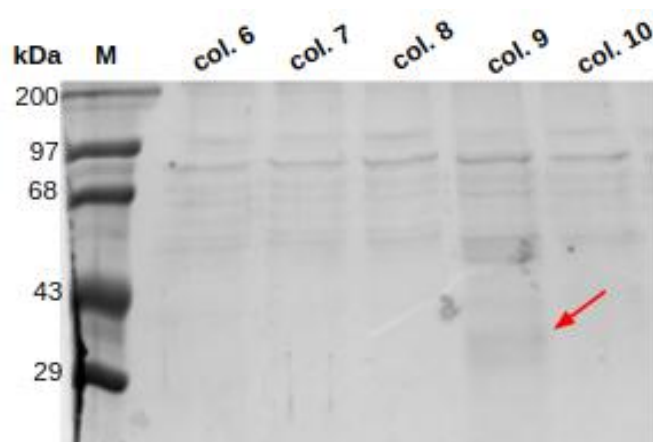


Figura 24 - Gel de poliacrilamida 12% corado com Comassie Blue mostrando as proteínas presentes no sobrenadante das colônias 6, 7, 8, 9 e 10, para o teste de expressão da gp41 recombinante em menor escala em *Pichia pastoris*. M: Marcador de peso molecular Protein Molecular Weight Standards (Life Technologies).

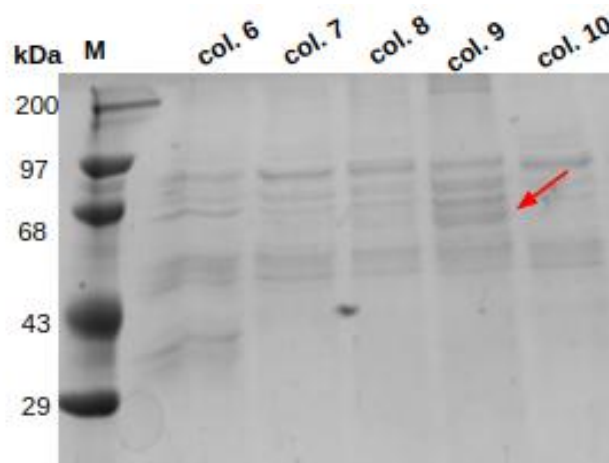


Figura 25 - Gel de poliacrilamida 12% corado com Comassie Blue mostrando as proteínas presentes no sobrenadante das colônias 6, 7, 8, 9 e 10, para o teste de expressão da gp120 recombinante em menor escala em *Pichia pastoris*. M: Marcador de peso molecular Protein Molecular Weight Standards (Life Technologies).

A expressão em menor escala mostrou que é possível identificar uma banda na colônia 9 com o tamanho esperado de 53 kDa, provavelmente referente a proteína gp120 (Figura 25). Apesar do peso esperado para proteína gp41 ser de 22 kDa, observa-se no gel (Figura 24), também na colônia 9, uma banda com marcação fraca um pouco acima de 29 kDa, que provavelmente é referente a gp41. Esse resultado indica que a proteína gp41 foi provavelmente glicosilada pela *Pichia pastoris*, o que é esperado devido a presença de sítios de O-glicosilação na sequência de aminoácidos da gp41.

Podemos observar nos géis acima que a coloração com *Comassie Blue* está fraca, isso pode ser devido a baixa quantidade de proteínas presentes nos géis. Dessa forma, tentar concentrar a amostra e fazer uma coloração por prata seria o ideal, visto que a coloração por prata é de 100 a 1000 vezes mais sensível que *Coomassie* (WINKLER, 2007).

5.2.4 Purificação das proteínas recombinantes gp41 e gp120

Após 144h de indução, o sobrenadante foi coletado e purificado em uma coluna cromatográfica com afinidade a níquel. As frações foram analisadas em gel de poliacrilamida 12% (Figuras 26 e 27) e lidas em espectrofotômetro a 280 nm. Com essas leituras pode-se fazer os gráficos 1 e 2 (Figuras 27 e 29).

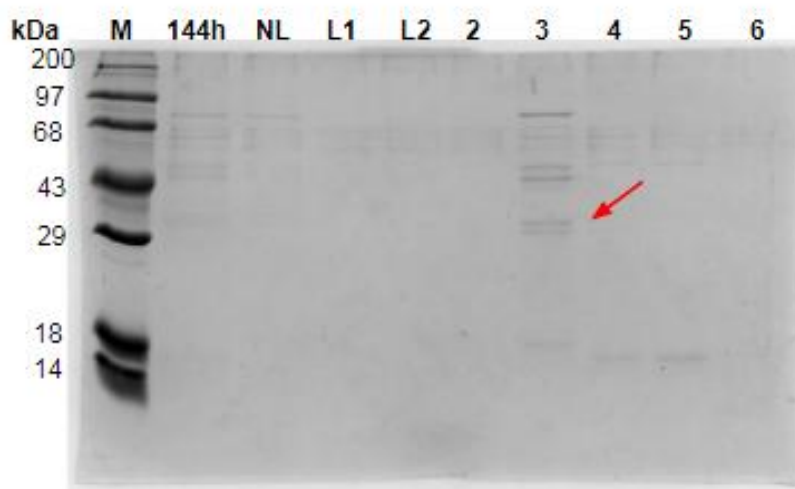


Figura 26 - Gel de poliacrilamida 12% corado com Comassie Blue mostrando a purificação da proteína recombinante gp41 na fração 3. M: Marcador de peso molecular Protein Molecular Weight Standards (Life Technologies); 144h: alíquota de 144 horas da indução; NL: fração não ligada; L1:

lavagem 1 com solução tampão; L2: lavagem com solução tampão com 5 mM de imidazol; 2 a 6: eluição das frações da purificação com 250 mM de imidazol.

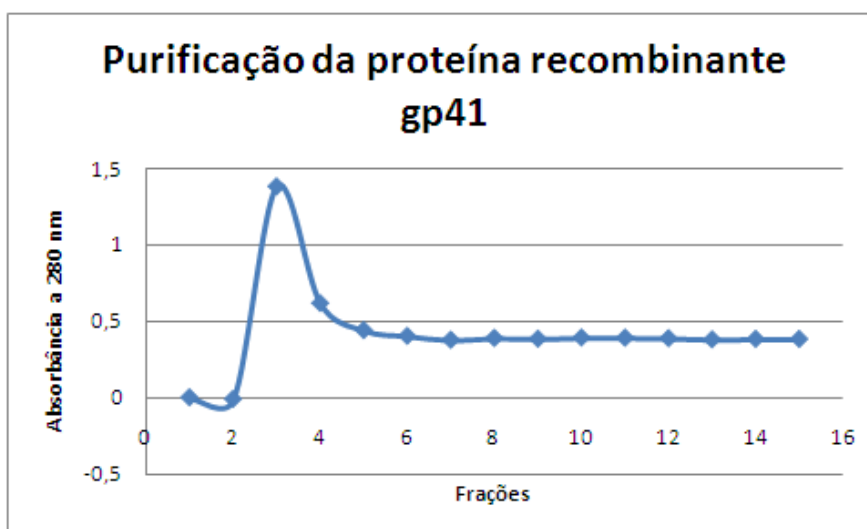


Figura 27 - Gráfico da purificação da proteína recombinante gp41, eluição da proteína na fração 3 com 250 mM de imidazol.

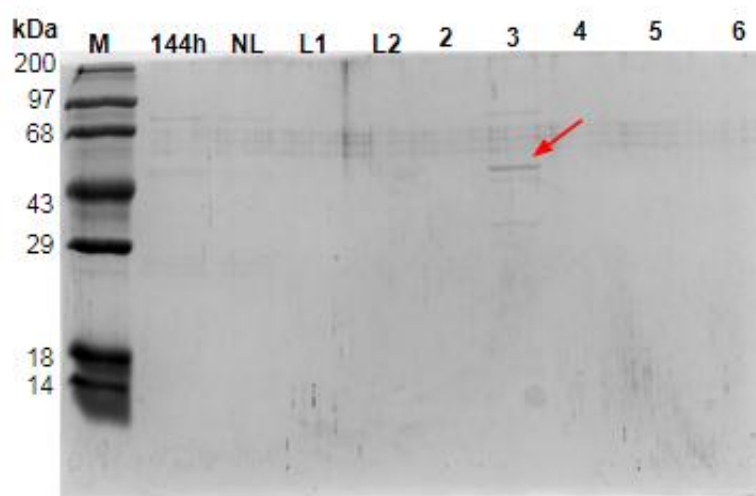


Figura 28 - Gel de poliacrilamida 12% corado com Comassie Blue mostrando a purificação da proteína recombinante gp120 na fração 3. M: Marcador de peso molecular Protein Molecular Weight Standards (Life Technologies); 144h: alíquota de 144 horas da indução; NL: fração não ligada; L1: lavagem 1 com solução tampão; L2: lavagem com solução tampão com 5 mM de imidazol; 2 a 6: eluição das frações da purificação com 250 mM de imidazol.

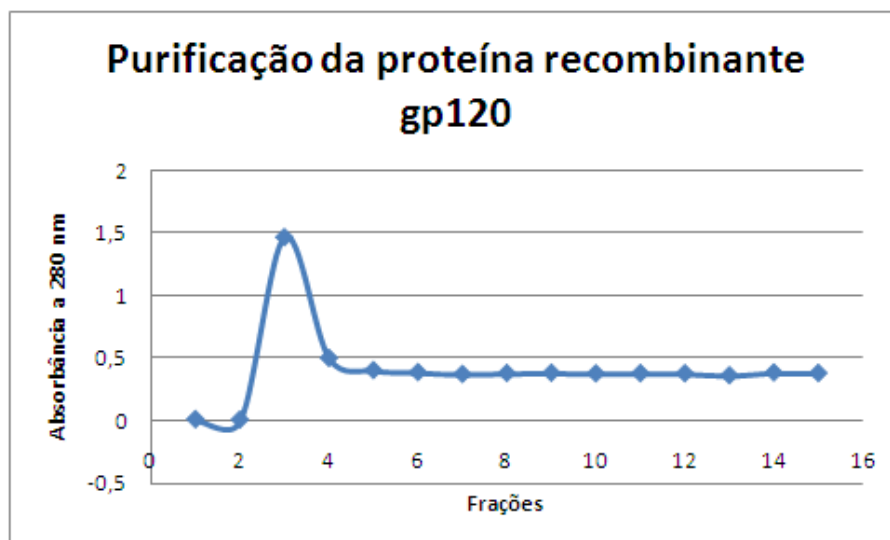


Figura 29 - Gráfico da purificação da proteína recombinante gp120, eluição da proteína na fração 3 com 250 mM de imidazol.

A purificação das proteínas recombinantes gp41 e gp120 (Figuras 26 e 28) foi realizada, entretanto ainda pode-se visualizar algumas outras proteínas presentes no gel, mostrando que a purificação não foi totalmente eficaz. Mesmo com a presença de outras proteínas pode-se visualizar um fragmento com o tamanho esperado de 53 kDa para proteína gp120. O fragmento da proteína gp41 está um pouco acima do tamanho de 22 kDa esperado, seu peso mostra-se perto dos 29 kDa. Como já dito anteriormente, o peso acima do esperado pode ser porque provavelmente a proteína gp41 foi glicosilada pela *Pichia pastoris*, devido a presença de sítios para O-glicosilação na sequência de aminoácidos da gp41.

5.2.5 Detecção das proteínas recombinantes por *Dot Blotting*

Para identificação das proteínas foi feito o *dot blotting* com as frações 3 de cada proteína e com o sobrenadante da indução da expressão de 144 horas. O *dot* foi realizado com os anticorpos específicos descritos na tabela 4 e o resultado está apresentado na figura 30.

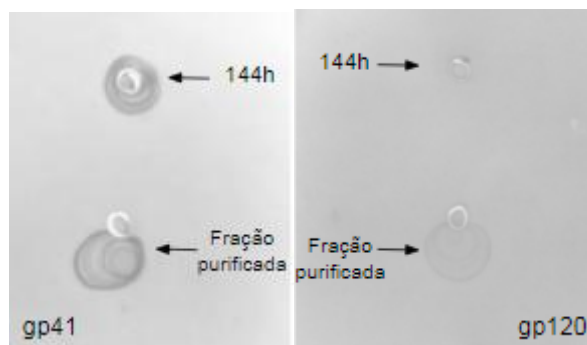


Figura 30 - *Dot Blotting* do screening das proteínas recombinantes gp41 e gp120.

A alíquota de 144 horas da indução da proteína recombinante gp41 foi marcada, porém, na alíquota de 144 horas da indução da proteína recombinante gp120 não houve marcação. Uma explicação para isso seria a baixa produção dessa proteína, visto que foi feito a expressão em menor escala, o que pode ser confirmado pela fraca marcação na fração purificada da gp120. A partir dos resultados demonstrados na figura 30 podemos confirmar que as proteínas recombinantes gp41 e gp120 foram reconhecidas pelos anticorpos específicos na fração purificada, mostrando assim que ocorreu a produção das proteínas recombinantes.

6. CONCLUSÃO

Os resultados apresentados neste trabalho permitem concluir que foi possível a clonagem dos genes das proteínas gp41 e gp120 do HIV-1 nos vetores de expressão para *E. coli* e *Pichia pastoris*. Entretanto, nas condições utilizadas não foi possível a expressão dessas proteínas em sistema procarioto. Mas, no sistema eucarioto *Pichia pastoris* foi possível identificar as proteínas recombinantes gp41 e gp120 em teste de expressão em menor escala.

7. PERSPECTIVAS FUTURAS

- Desenhar primer para região intermediária da gp120 e fazer novo sequenciamento
- Realizar a expressão das proteínas recombinantes gp41 e gp120 em maior escala
- Otimizar as condições de expressão para obter grandes quantidades de proteínas recombinantes

8. REFERÊNCIAS

- BANEYX, F.; MUJACIC, M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.*, Nova York, v.22, n.11, p.1399-1408, 2004.
- BARRE-SINOUSI, F. et al. Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science*, v. 220, p. 868-871, 1983.
- BERLEC, A.; STRUKELJ, B. Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells. *Journal of Industrial Microbiology and Biotechnology*, v. 40, n. 3, p. 257-74, 2013.
- BIRNBOIM HC; DOLY J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acid Research*, 7: 1513-23, 1979.
- BLACKARD, J. T. et al. Human Immunodeficiency Virus Superinfection and Recombination: Current State of Knowledge and Potential Clinical Consequences. *Clinical infectious diseases*, v. 34, n. 8, p.1108-14, 2002.
- BOLETIM EPIDEMIOLOGICO HIV/AIDS 2016. Disponível em: <http://portalarquivos.saude.gov.br/images/pdf/2017/janeiro/05/2016_034-Aids_publicacao.pdf>. Acessado em: 23 de agosto de 2017
- BRASIL. Ministério da Saúde. Gabinete do Ministro. Portaria no 29 de 17 de dezembro de 2013. Disponível em: <http://bvsms.saude.gov.br/bvs/saudelegis/svs/2013/prt0029_17_12_2013.html> Acessado em: 25 de agosto de 2017.
- BRANSON, B. M. Rapid Tests for HIV Antibody. *AIDS Reviews*, v. 2, p. 76-83, 2000.
- BROWN, T. A. Clonagem Gênica e Análise de DNA: Uma Introdução. Artmed Editora S.A. Porto Alegre, RS. 375p., 2003.
- BUTLER, I. et al. HIV Genetic Diversity: Biological and Public Health Consequences. *Current HIV Research, Oak Oark*, v. 5, p. 23-45, 2007.
- Centers for Disease Control (CDC). Update on acquired immune deficiency syndrome (AIDS)- United States. *Morbidity and Mortality Weekly Report*, v. 31, n. 37, p. 507-8, 513-4, 1982.
- ÇELIK, E.; ÇALIK, P. Production of recombinant proteins by yeast cells. *Biotechnology Advances*, Volume 30, Issue 5, p. 1108-1118, 2012.
- CEREGHINO, G. P. L.; CREGG, J. M. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiology Reviews*, 24: 45-66, 2000.
- CHOI, J. H.; KEUM, K. C.; LEE, S. Y. Production of recombinant proteins by high cell density culture of *Escherichia coli*. *Chemical Engineering Science*, v. 61, p. 876 – 885, 2006.

- CLAPHAM, P. R.; WEISS, R. A. Immunodeficiency viruses. Spoilt for choice of co-receptors. *Nature*, v.388, n. 6639, p. 230-231, 1997.
- CLAVEL, F., D. Guetard, F. Brun-Vezinet, S. Chamaret, M. A. Rey, M. O. Santos-Ferreira, A. G. Laurent, C. Dauguet, C. Katlama, C. Rouzioux, D. Klatzmann, J. L. Champalimaud, and L. Montagnier. Isolation of a new human retrovirus from West African patients with AIDS. *Science* 233: 343–346, 1986.
- COHEN, M. S. et al. The Detection of Acute HIV Infection. *The Journal of Infectious Diseases*, v. 202, p. S270-S277, 2010.
- CONNOR, R. I.; HO, D. D. Etiology of AIDS: biology of human retroviruses. In: Devita VT. Editor. *AIDS etiology, diagnosis, treatment and prevention*. Philadelphia: J. B. Lippincott Company, p. 13-85, 1992.
- COSTA, S. et al. Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. *Frontiers in Microbiology*, v. 5, n. 63, p. 1-20, 2014.
- COSTIN, J. M. Cytopathic mechanisms of HIV-1. *Virology* 18:4:100. 2007
- CREEG, J. M.; MADDEN, K. R.; BARRINGER, K. J.; THILL, G. P. e STILLMAN, C., A. Functional characterization of the two alcohol oxidase genes from the yeast *Pichia*. *Molecular and Cellular Biology*, 9: 1316-1323, 1989
- DALY, R.; HEARN, M. T. W. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production (Review). *Journal of Molecular Recognition*. v.18, p. 119-138, 2005.
- DASKALAKIS, D. HIV diagnostic testing: Evolving technology and testing strategies. *Topics in antiviral medicine*, v. 19, p.18-22, 2011.
- DAVI, E. V. Clonagem de fragmentos dos genes gag e env do HIV-1 e HTLV-1, expressão em *Escherichia coli* das proteínas gp21, p24 e gp46 do HTLV-1 e imunodeteção. Ribeirão Preto, 2015.
- DEMAÏN, A. L.; VAISHNAV, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnology Advances*, v. 27, p. 297–306, 2009.
- DEPARTAMENTO DE VIGILÂNCIA, PREVENÇÃO E CONTROLE DAS IST, DO HIV/AIDS E DAS HEPATITES VIRAIS, 2017. Disponível em: <http://www.aids.gov.br/pt-br/publico-geral/o-que-e-hiv/sintomas-e-fases-da-aids> Acessado em 24 de agosto de 2017.
- EARL, P. L; MOSS, B.; DOMS, R. W. Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. *Journal of Virology*, v. 65, n. 4, p. 2047-55, 1991.
- EVANS, T. C. Heterologous Gene Expression in *E. coli*, (Adjustment of Codon Usage Frequencies by Codon Harmonization Improves Protein Expression and Folding.

Disponível

em:

<http://www.lifesensors.com/pdf/2011_PeroutkaRJ_MethMolBiol_SUMO_fusion_technology_for_enhanced_protein_expression_and_purification_in_prokaryotes_and_eukaryotes.pdf>

- FIEBIG, E. W. et al. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. *AIDS*, v. 17, n. 13, 1871–1879, 2003.
- FRANKEL, A. D, YOUNG J. A. T. HIV-1: Fifteen proteins and an RNA. *Annual Review of Biochemistry*, v. 67, p.1-25, 1998.
- FREED, E. O. HIV-1 Gag Proteins: Diverse Functions in the Virus Life Cycle. *Virology*, v. 251, p. 1-15, 1998.
- FREED, E. O. HIV-1 Replication. *Somatic Cell and Molecular Genetics*, v. 26, p. 13-33, 2001.
- GANSER-PORNILLOS, B.; YEAGER, M.; SUNDQUIST, W. I. The structural biology of HIV assembly. *Current Opinion in Structural Biology*, v. 18, p. 203–217, 2008.
- GOEDERT, J. J. Recreational drugs: relationship to AIDS. *Annals of the New York Academy of Sciences*, v. 437, p. 192-199, 1984.
- GOEDEL, D. V. Systems for heterologous gene expression. *Methods of Enzymology*, v. 185, p. 3-7, 1990.
- GOMEZ, C.; HOPE, T. J. The ins and outs of HIV replication. *Cellular Microbiology*, v. 7, n.5, p. 621–626, 2005.
- GOODNOW, M. M.; COLLMAN, R. G. HIV-1 coreceptor preference is distinct from target cell tropism: a dual-parameter nomenclature to define viral phenotypes. *Journal of Leukocyte Biology*, v. 80, p. 965-972, 2006.
- GOPAL, G. J.; KUMAR, A. Strategies for the Production of Recombinant Protein in *Escherichia coli*. *The protein journal*, v. 32, p. 419–425, 2013.
- GOTTLIEB, M. S. et al. Pneumocystis carinii Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men — Evidence of a New Acquired Cellular Immunodeficiency. *New England Journal of Medicine*, v. 305, n. 24, p. 1425-31, 1981.
- GRANT, P. M.; ZOLOPA, A. R. When to Start ART in the Setting of Acute AIDS-Related. *Current HIV/AIDS reports.*, v. 9, p. 251–258, 2012.
- GREENWALD, J.L.; BURSTEIN G.R.; PINCUS J.; BRANSON B. A Rapid Review of Rapid HIV. *Current Infectious Disease Reports*; 8:125-131, 2006.
- HAMATAKE et al. A simple competitive RT-PCR assay for quantitation of HIV- 1 subtype B and non-B RNA in plasma. *Journal of Virological Methods*, v. 142, p. 113-117, 2007.

- HANAHAN D. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166(4): 1557-80, 1983.
- HASSAN M. Naif. Pathogenesis of HIV Infection, *Infectious Disease Reports*, v.5:s1e6. 2013.
- ICTVdB. The Universal Virus Database. Disponível em: <<http://ictvonline.org/>>. Acessado em 22 de agosto de 2017.
- JAFFAR, S. et al. The natural history of HIV-1 and HIV-2 infections in adults in Africa: a literature review. *Bull World Health Organ*, v. 82, p. 462–469, 2004.
- KANE, J. F., Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *SmithKline Beecham Pharmaceuticals, King of Prussia, USA, Current Opinion in Biotechnology* 1995, 6:494-500
- KATZ, R. A.; SKALKA, A. M. The retroviral enzymes. *Annual review of biochemistry*, v. 63, p. 133-173, 1994.
- KAUR, J.; KUMAR, A.; KAUR, J. Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements, *International Journal of Biological Macromolecules*, Vol. 106, p. 803-822, 2018.
- KILLIAN, M. S.; LEVY, J. A. HIV/AIDS: 30 years of progress and future challenges. *European journal of immunology*, v. 41, p. 3401-3411, 2011.
- KOURI, V. et al. CRF19_cpx is an Evolutionary fit HIV-1 Variant Strongly Associated With Rapid Progression to AIDS in Cuba. *EBioMedicine*, 2015.
- LAEMMLI, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, London, 227, 680-85, 1970.
- LEVY, J. A.; SCOTT, I & MCKEWICZ, C. Protection from HIV/AIDS: the importance of innate immunity. *Clin Immunol* . 108: 167-174; 2003.
- MADIGAN, Michael T. et al. *Microbiologia de Brock*. 14. ed. Porto Alegre: ArtMed, 2016.
- MAHBOUDI, F. et al. A serological screening assay of human immunodeficiency virus type 1 antibodies based on recombinant protein p24–gp41 as a fusion protein expressed in *Escherichia coli*. *Journal of Biotechnology*, v. 125, p. 295–303, 2006.
- MAKRIDES, S. C. Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*. *Microbiological Reviews*, v. 60, n. 3, p. 512–538, 1996.
- MANIATIS, T.; Fritschl, E.F.; Sambrook, J. *Molecular cloning, a laboratory manual*. New York: Spring Harbor, 1989. p. ES.
- MANUAL TÉCNICO PARA DIAGNÓSTICO DA INFECÇÃO PELO HIV, Ministério da Saúde, Departamento de DST, Aids e Hepatites Virais, Brasília, 2014. Disponível em: <http://bvsmms.saude.gov.br/bvs/publicacoes/manual_tecnico_diagnostico_infeccao_hiv.pdf> Acessado em 22 de agosto de 2017.

- MASUR, H. et al. An outbreak of community-acquired *Pneumocystis carinii* pneumonia: initial manifestation of cellular immune dysfunction. *New England Journal of Medicine*, v. 305, n. 24, p. 1431-1438, 1981.
- McCUTCHAN, F. E. Understanding the genetic diversity of HIV-1. *AIDS*, v. 14, p. 31-44, 2000.
- MERGULHÃO, F. J. M.; SUMMERS, D. K.; MONTEIRO, G. A. Recombinant protein secretion in *Escherichia coli*. *Biotechnology Advances*, v. 23, p. 177 – 202, 2005.
- MIYAMOTO, C. A. Produção de proteínas recombinantes em *Escherichia coli*. *Revista conexão eletrônica*, p. 220–233, 2013.
- MONTERO, M. et al. The Membrane-Proximal External Region of the Human Immunodeficiency Virus Type 1 Envelope: Dominant Site of Antibody Neutralization and Target for Vaccine Design. *Microbiology and Molecular Biology Reviews*, v. 72, n. 1, p. 54–84, 2008.
- NASCIMENTO, A. A. C.; ESPREAFICO, E. M.; LARSON, M. L. P.; MONESI, N.; ROSSI, N. M. M.; RODRIGUES, V. *Tecnologia do DNA recombinante*. Ribeirão Preto: Universidade de São Paulo, 1999. 85 p. Apostila
- NELSON, M. et al. British HIV Association and British Infection Association guidelines for the treatment of opportunistic infection in HIV-seropositive individuals. *HIV Medicine*, v. 12 Suppl 2, p. 100-140, 2011.
- NYAMWEYA, S., Hegedus, A., Jaye, A., Rowland-Jones, S., Flanagan, K. L., Macallan DC. Comparing HIV-1 and HIV-2 infection: Lessons for viral immunopathogenesis. *Rev Med Virol.*, v.23, p. 221-240, 2013.
- PANDORI, M. W. et al. Assessment of the ability of a fourth-generation immunoassay for human immunodeficiency virus (HIV) antibody and p24 antigen to detect both acute and recent HIV infections in a high-risk setting. *Journal of clinical microbiology*, v, 47, n. 8, p. 2639-2642, 2009.
- PEETERS, M.; SHARP, P. M. Genetic diversity of HIV-1: the moving target. *AIDS*, v. 14, p. 129- 40, 2000.
- POIGNARD, P. W. et al. GP120: Biologic Aspects of Structural Features. *Annual Reviews of Immunology*, v. 19, p. 253-274, 2001.
- POPOVIC, M. et al. Detection, Isolation, and Continuous Production of Cytopathic Retroviruses (HTLV-111) from Patients with AIDS and Pre-AIDS. *Science*, v.224, p. 497- 500, 1983.
- RACHID, M.; SCHECHTER, M. *Manual de HIV/AIDS*. 8a ed. Rio de Janeiro: Ed. Revinter, 2005.
- RAQNI, M. V. et al. Acquired immunodeficiency syndrome in the child of a haemophiliac. *Lancet.*, v.1, n. 8421, p. 133-5, 1985.

- RAVANSHAD, M. et al. Evaluation of a new dot blot assay for confirmation of human immunodeficiency virus type 1 and 2 infections using recombinant p24, gp41, gp120 and gp36 antigens. *Saudi medical journal*, v. 27, n. 1, p. 31-36, 2006.
- REDDI, H., Bhattacharya, A. and Kumar, V. (2002). The calcium-binding protein of *Entamoeba histolytica* as fusion partner for expression of peptides in *Escherichia coli*. *Biotechnol. Appl. Biochem.* 36: 213-8.
- ROBERTSON, D. L. et al. Recombination in HIV-1. *Nature*, v. 374, p.124–6, 1995.
- ROSANO, G. L.; CECCARELLI, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, v. 5, n. 172, p. 1-17, 2014.
- RUSSO-CARBOLANTE, E. M. S. Medeiros, L. ; Penteado, F. L et al. Clonagem e expressão da glicoproteína transmembrana do vírus linfotrópico de células T humanas em sistema procaríoto. *Revista da Sociedade Brasileira de Medicina Tropical*, v. 40, n. 3, p. 277-281,2007.
- SHARP, P. M., LI, W. Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for 'rare' codons. *Nucleic Acids Research*, Vol. 14 Nº 19, 1986.
- SAIKI, RK; Gelfand DH; Stoffel S; Schard SJ; Higuchi R; Horn GT; Mullis KB; Erlich HA. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239:487-91, 1988.
- SANGER F., Nicklen S., Coulson A. R. DNA sequencing with chain-terminating inhibitors. *Proc Nat Acad Sci, USA*, 74:5463-7, 1977.
- SIEGAL, F. P. et al. Severe acquired immunodeficiency in male homosexuals, manifested by chronic perianal ulcerative herpes simplex lesions. *New England Journal of Medicine*, v. 305, n. 24, p. 1439-1444, 1981.
- SORENSEN, H. P. et al. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology, Philadelphia*, v. 115, p. 113–128, 2005.
- TANG, S. et al. Characterization of Immune Responses to Capsid Protein p24 of Human Immunodeficiency Virus Type 1 and Implications for Detection. *Clinical and Vaccine Immunology*, v. 17, n. 8, p. 1244–1251, 2010.
- TOMARAS, G. D. et al. Initial B-Cell Responses to Transmitted Human Immunodeficiency Virus Type 1: Virion-Binding Immunoglobulin M (IgM) and IgG Antibodies Followed by Plasma Anti-gp41 Antibodies with Ineffective Control of Initial Viremia. *Journal of Virology*, v. 82, n. 24, p. 12449- 12463, 2008.
- TOMARAS, G. D.; HAYNES, B. F. HIV-1-specific antibody responses during acute and chronic HIV-1 infection. *Current opinion in HIV and AIDS*, v. 4, n. 5, p. 373– 379, 2009.

- TORRES, F. A. G.; MORAES, L. M. P. Proteínas recombinantes produzidas em leveduras. *Biotecnologia Ciência e Desenvolvimentos, Brasil*, v. 12, p. 20-22, 2000.
- UNAIDS. Estatísticas. Resumo Global da epidemia de AIDS, 2016 Disponível em: <<http://unaid.org.br/estatisticas/>>. Acessado em: 23 de Agosto de 2017.
- WINKLER, C. et al. Silver and Coomassie staining protocols: detection limits and compatibility with ESI MS. *Electrophoresis*, v. 28, n.12, p. 2095-9, 2007
- WONG-STAAAL, F. The AIDS Virus: What We Know and What We Can Do About It. *Biomedical Science*, v. 155, p. 481-487, 1991.
- YIN, J.; LI, G.; REN, X.; HERRLER, G.; Select what you need: A comparative evaluation of the advantages and limitations of frequently used expression system for foreign genes. *Journal of Biotechnology*, v. 127, p 335-347, 2007.