



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Biológicas e da Saúde

Programa de Pós Graduação em Genética Evolutiva e Biologia Molecular

**Produção heteróloga e caracterização de uma beta-glicosidase
identificada em sequências metagenômicas de um lago da região
amazônica**

Augusto Furio Balula

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos como parte do requisito para a obtenção do Título de Mestre em Genética Evolutiva e Biologia Molecular – Área de concentração em Bioquímica e Biologia Molecular

Orientador: Prof. Dr. Flavio Henrique Silva

São Carlos - SP

Orientador

Prof. Dr. Flávio Henrique Silva



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato Augusto Furio Balula, realizada em 15/02/2017



Prof. Dr. Flavio Henrique da Silva
UFSCar



Prof. Dr. Iran Malavazi
UFSCar



Prof. Dr. Richard Charles Garratt
USP

**"Procure ser um homem de valor, em
vez de ser um homem de sucesso"**

(Albert Einstein)

Dedicatória

Dedico este trabalho à minha família, fonte de
motivação onde construo minha vida.

AGRADECIMENTO

Agradeço a todos os membros de minha família, em especial minha mãe Sueli Furio Balula, meu pai Luiz Francisco Balula e meu irmão Bruno Furio Balula por me passarem bons exemplos e, principalmente, o apoio necessário nos momentos difíceis.

Agradeço ao meu orientador, Dr. Flávio Henrique Silva, que desde o momento em que entrei em seu laboratório sempre me ensinou muito sobre ciência e me ajudou muito no desenvolvimento deste trabalho. Em passagens por outros laboratórios, os elogios que recebi por conta das técnicas que domino vêm de seu ensinamento. Além dos ensinamentos acadêmicos também agradeço-o pela amizade ao longo destes anos.

Agradeço ao professor Dr. Sandro Marana, ao Vitor Medeiros e à Universidade de São Paulo (USP) por me ajudar na realização deste trabalho.

Ao CNPq pela bolsa concedida e fomento do projeto de pesquisa.

À Universidade Federal de São Carlos (UFSCar) e ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular (PPGGEV) pela infraestrutura oferecida para a realização deste trabalho. A todos os colegas e professores do Departamento de Genética e Evolução, pela boa convivência nesses anos.

Agradeço aos colegas e amigos do Laboratório de Biologia Molecular e da cidade de São Carlos.

A todos, meus sinceros agradecimentos

Resumo

A metagenômica permite estudar diretamente o material genético presente em uma amostra ambiental e quando aliada à bioinformática possibilita explorar o papel de novos genes e proteínas. A constante diminuição na quantidade de combustíveis fósseis e seus efeitos na economia global e no meio ambiente têm acelerado as pesquisas sobre combustíveis alternativos como, por exemplo, o etanol de segunda geração, o qual pode ser obtido a partir de biomassa vegetal. No entanto, o processo necessita que o material lignocelulolítico seja hidrolisado previamente por enzimas, para o fornecimento de açúcares fermentescíveis. As β -glicosidases são enzimas que participam da etapa final de degradação de celulose em glicose e são, portanto, consideradas passo limitante no processo. Muitas β -glicosidases já foram descritas, entretanto ainda há o interesse em encontrar enzimas que sejam resistentes à inibição por glicose e que exerçam sua atividade em temperaturas mais baixas. Neste sentido, o presente trabalho tratou da busca por β -glicosidases da família GH1 utilizando sequências obtidas a partir de um estudo metagenômico de rios e lagos da Amazônia, realizado em nosso laboratório. Foram encontradas 3 fases abertas de leitura (ORFs) correspondentes à esta classe de enzimas e uma delas foi selecionada para ser produzida em *E.coli* de forma recombinante e ser caracterizada bioquimicamente. A sequência que codifica a proteína denominada AmBgl1-LP foi clonada em vetor pET-28a e expressa em *E.coli*, rendendo uma enzima com massa molecular de 53,7 kDa. Os ensaios de atividade enzimática revelaram que a enzima é ativa em pH ótimo de 5,5 e temperatura ótima de 35 °C. Além disso, a enzima possui um K_i para glicose de 23 mM. A enzima aparentemente não realiza transglicosilação, frente aos ensaios com o substrato pNP β Gli. Aparentemente a enzima sofre inibição por este substrato em concentrações maiores que 10 mM. A AmBgl1-LP mostrou-se capaz de hidrolisar celobiose, além de pNP β Gal e pNP β Fuc. Desta forma, a enzima mostra-se promissora para utilização em coquetéis para degradação de biomassa.

Palavras-chave: β -glicosidase, tolerância à glicose, metagenoma, glicosil hidrolases, expressão recombinante, transglicosilação.

Abstract

Metagenomics studies allow the direct analysis of a genetic material in an environmental sample and when linked to bioinformatics it gives a powerful tool to explore the role of new genes and proteins not studied before. Constant decreases in the quantity of fossil fuels and their effects in the global economy and natural environment has accelerated researches in alternative fuels such as the second generation ethanol, which can be produced by vegetation biomass. However, this process demands previous hydrolysis of the lignocellulolytic material by hydrolytic enzymes to provide fermentable sugar. β -glucosidases are enzymes which plays an important role at the final step of cellulose breakdown to glucose, thus being considered the rate limiting enzyme in this process of biomass degradation. Many β -glucosidases are already known, however there is an interest to find new enzymes which are tolerant to glucose inhibition and which exhibits high activity at lower temperatures. In this study we searched for β -glucosidases (GH1) using sequences from a metagenomics database from rivers and lakes in the Amazon region developed in our laboratory. We found 3 complete open reading frames (ORFs) related to β -glucosidases and one of them was selected to be produced in *E.coli* in a heterologous way and to be biochemically characterized. The coding sequence of the protein named AmBgl1-LP was cloned in the plasmid pET-28a and produced an enzyme which has a molecular mass of 53,7 kDa. The enzymatic assays showed that the enzyme was active with an optimum pH of 5.5, optimum temperature of 35 °C and had a K_i for glucose of 23 mM. The enzyme does not apparently perform transglycosylation, according to the assays for pNP β Glu substrate. Supposedly, AmBgl1-LP suffers inhibition by pNP β Glu on concentrations higher than 10 mM. The enzyme showed to be capable of hydrolyzing cellobiose, pNP β Gal and pNP β Fuc. Thus, the enzyme is promising for use in cocktails for degradation of biomass.

Keywords: β -glucosidases, glucose tolerant, metagenome, glycosil hydrolases, heterologous expression, transglycosylation.

Lista de Figuras

Figura 1 : Mecanismo de hidrólise de ligações β -glicosídicas para as β glicosidasas.....	Pg 3
Figura 2: Alinhamento das sequências “seed” de GH1 do Pfam.....	Pg 14
Figura 3 : ORF e sequência de aminoácidos da <i>AmBgl1-LP</i>	Pg 16
Figura 4 : Alinhamento entre <i>AmBgl1-LP</i> e outras β -glicosidasas já descritas.....	Pg 17
Figura 5 : Amplificação da ORF <i>AmBgl1-LP</i>	Pg 18
Figura 6 : PCR de colônia para a ORF <i>AmBgl1-LP</i> no plasmídeo pET-28a em <i>E. coli</i> DH5 α	Pg 19
Figura 7 : Clivagem do vetor pET-28a contendo o gene <i>AmBgl1-LP</i>	Pg 20
Figura 8 : Teste de expressão da <i>AmBgl1-LP</i> em <i>E. coli</i>	Pg 21
Figura 9 : Purificação da <i>AmBgl1-LP</i> por cromatografia de afinidade molecular ao níquel.....	Pg 22
Figura 10 : Indução da <i>AmBgl1-LP</i> em <i>E. coli</i> a 20°C.....	Pg 23
Figura 11: Curva de pH para a <i>AmBgl1-LP</i>	Pg 25
Figura 12: Curva de temperatura e termoestabilidade para <i>AmBgl1-LP</i>	Pg 26
Figura 13: Curvas de $V_0/[S]$ para a <i>AmBgl1-LP</i> na presença do substrato celobiose e pNP β Gli.....	Pg 28
Figura 14: Modelo de reação proposto para a <i>AmBgl1-LP</i>	Pg 29
Figura 15: Gráfico de velocidades iniciais de formação de p-nitrofenolato e glicose utilizando o substrato pNP β Gli.....	Pg 30
Figura 16: Ensaios de inibição para a <i>AmBgl1-LP</i>	Pg 32
Figura 17: Alinhamento de sequências entre HiBgl, BglCell1A e <i>AmBgl1-LP</i>	Pg 35

Lista de Tabelas

Tabela 1: Dados cinéticos sobre a <i>AmBgl1-LP</i> para o substrato celobiose.....	Pg 29
Tabela 2: Dados cinéticos sobre a <i>AmBgl1-LP</i> para o substrato pNP β Gli tanto para o modelo de Michaelis-Menten quanto para o modelo de inibição pelo substrato.....	Pg 30

Lista de siglas e abreviações

BCA: Bicinchonic acid (ácido bicinchônico)

CAZy: carbohydrate-active enzymes database

DNA: ácido desoxirribonucléico

D.O.: Densidade óptica

E: enzima

ES: complexo enzima-substrato

et al.: e colaboradores, entre outros

g: Grama

GH1: glicosil hidrolases da família 1

h: hora

His: histidina

HMM: Hidden Markov Model (Modelo Oculto de Markov)

HPLC: Cromatografia líquida de alto desempenho

IPTG: isopropil- β -D-tiogalactopiranosídeo

Kb: quilobase

Kcat: constante catalítica

KDa: quilodalton

Ki: constante de inibição

Km: constante de Michaelis-Menten

Km: quilômetro

L: litro

L: Leucina

LB-Broth: Luria-Bertani ou Lysogeny broth – Meio de crescimento microbiano

LP: Lago Poraquê

M: molar

mg: miligrama

MgSO₄: sulfato de magnésio

min: minuto

mL: mililitro

mM: milimolar

mm: milímetro

MWCO: molecular weight cut-off (faixa de exclusão de peso molecular)

ng: nanograma

Ni²⁺: Níquel na forma iônica

nm: nanômetro

nmol: nanomol

NTA: ácido nitriloacético

ORF: *open reading frame* (fase aberta de leitura)

P: produto

P: prolina

Pb: par de base

PCR: Polymerase Chain Reaction (Reação em cadeia da polimerase)

pH: potencial hidrogeniônico

pmol: picomol

pNPβGlc: p-nitrofenil-β-D-glicopiranosídeo

pNPβGal: p-nitrofenil-β-D-galactopiranosídeo

pNPβFuc: p-nitrofenil-β-D-fucopiranosídeo

RNA: ácido ribonucléico

rRNA: ácido ribonucléico ribossômico

tRNA: ácido ribonucléico transportador

rpm: rotações por minuto

[S]: concentração de substrato

SDS: Sódio Dodecil Sulfate (dodecil sulfato de sódio)

SDS-PAGE: Sodium Dodecil Sulfate-Polyacrilamide Gel Eletrophoresis (eletroforese em gel de poliacrilamida sódio dodecil sulfato)

SES: complexo susbtrato-enzima-substrato

SIB: Swiss Institute of Bioinformatics (Instituto Suíço de Bioinformática)

SHF: Separate Hydrolysis and Fermentation (fermentação e hidrólise separadas)

SSF: Simultaneous Saccharification and Fermentation (sacarificação e fermentação simultânea)

Tris: tris(hidroximetil)aminometano

U: Unidade enzimática

V0: Velocidade Inicial

V: volume

Vmax: velocidade máxima

W: Triptofano

°C: graus Celsius

µg: micrograma

µL: microlitro

µm: micrômetro

µM: micromolar

SUMÁRIO

1 INTRODUÇÃO.....	1
1.1 Abordagem Metagenômica.....	1
1.2 Glicosil hidrolases da família 1: β -glicosidases e aplicações	2
2 OBJETIVOS.....	4
2.1 OBJETIVOS ESPECÍFICOS.....	4
3 MATERIAL E MÉTODOS.....	5
3.1 Busca por glicosil-hidrolases da família 1 em dados de metagenoma e escolha da AmBgl1-LP.....	5
3.2 Expressão heteróloga da AmBgl1-LP em <i>Escherichia coli</i>	7
3.2.1 Clonagem	7
3.2.2 Expressão e purificação da AmBgl1-LP	9
3.3 Caracterização enzimática da β -glicosidase	10
4 RESULTADOS E DISCUSSÃO	14
4.1 Identificação da AmBgl1-LP nas sequências metagenômicas.....	14
4.2 Expressão Heteróloga em <i>Escherichia coli</i>	18
4.3 Ensaio de atividade e caracterização enzimática.....	24
5 CONCLUSÕES.....	34
6 PERSPECTIVAS.....	34
7 REFERÊNCIAS BIBLIOGRÁFICAS	35
8 APÊNDICE	39

1 INTRODUÇÃO

1.1 Abordagem Metagenômica

A abordagem da metagenômica se baseia no estudo do material genético de microorganismos pela extração direta do DNA presente em uma amostra ambiental (Handelsman et al., 1998). Diferentemente da microbiologia e da genômica clássica para o sequenciamento de genomas microbianos, a metagenômica possibilita, por meio da análise dos genes de rRNA 16S ou 18S, produzir um perfil de microorganismos totais presentes em uma amostra ambiental sem a necessidade de cultivo dos mesmos (Hugenholtz et al., 1998; Handelsman et al., 1998). A metagenômica possui duas abordagens diferentes em relação aos objetivos de estudo: a análise baseada na sequência e a análise baseada na função (Handelsman et al., 2004).

A análise baseada na sequência utiliza de marcadores filogenéticos que indicam o grupo taxonômico dos organismos presentes em uma amostra por meio da análise de sequências conservadas como, por exemplo, a do gene de rRNA 16S; este tipo de abordagem produziu a primeira sequência genômica ligada ao gene de rRNA 16S de uma arqueobactéria não-cultivável pelo grupo de DeLong (Stein et al., 1996).

Além do foco baseado em marcadores filogenéticos também é possível realizar o sequenciamento do DNA total das amostras e uma busca funcional de genes. Este tipo de busca, quando realizado um trabalho massivo de anotação gênica por bioinformática, caracteriza a metagenômica funcional e representa uma abordagem poderosa na descoberta de novos genes e proteínas a partir de microorganismos não-cultiváveis (Steele & Streit, 2005) podendo ser aplicada a diversos ambientes ainda não explorados.

1.2 Glicosil hidrolases da família 1: β -glicosidases e aplicações

A imensa variedade de formas dos carboidratos é acompanhada também por uma grande quantidade de enzimas envolvidas em seu metabolismo. As glicosil-hidrolases são enzimas chave no metabolismo de carboidratos e são encontradas em arqueobactérias, eubactérias e eucariotos (Henrissat, 1991).

Com o surgimento crescente de milhares de glicosil hidrolases foi sugerida realizar uma classificação destas enzimas baseadas em suas sequências de aminoácidos (Henrissat, 1991), a qual foi atualizada diversas vezes (Henrissat & Bairoch, 1993, 1996) até que estas enzimas passaram a ter suas famílias classificadas e atualizadas no banco de dados do CAZy (carbohydrate-active enzymes database) (<http://www.cazy.org>), o qual atualmente conta com 135 famílias de glicosil hidrolases.

As glicosil hidrolases da família 1 (GH1) representam uma das maiores famílias desta classe de enzimas com cerca de 13228 sequências depositadas no CAZy. As GH1 compreendem enzimas com um grande número de atividades conhecidas, entre elas temos: β -glicosidases, β -galactosidases, 6-fosfo- β -galactosidase, 6-fosfo- β -glicosidases, β -manosidases, β -glucuronidases, β -xilosidases, exo- β -1,4-glucanase, entre várias outras funcionalidades; sendo as β -glicosidases e β -galactosidases as enzimas mais características dentro desta família.

O mecanismo de ação pelo qual as β -glicosidases hidrolisam seu substrato se inicia com sua ligação ao sítio ativo da enzima, que contém um par de ácidos carboxílicos conservados (McCarter & Withers, 1994). Primeiramente ocorre a etapa de glicosilação, na qual o glicosídeo é protonado por um dos ácidos, sendo um catalisador ácido/base, enquanto também sofre um ataque no centro anomérico pelo outro ácido carboxílico resultando na formação de um intermediário glicosil-enzima. Este intermediário glicosil-enzima sofre hidrólise em um segundo passo de deglicosilação por um ataque da água atuando como base no centro anomérico do substrato, a qual adiciona uma hidroxila em um dos resíduos de glicose restantes liberando-os da enzima (McCarter & Withers, 1994).

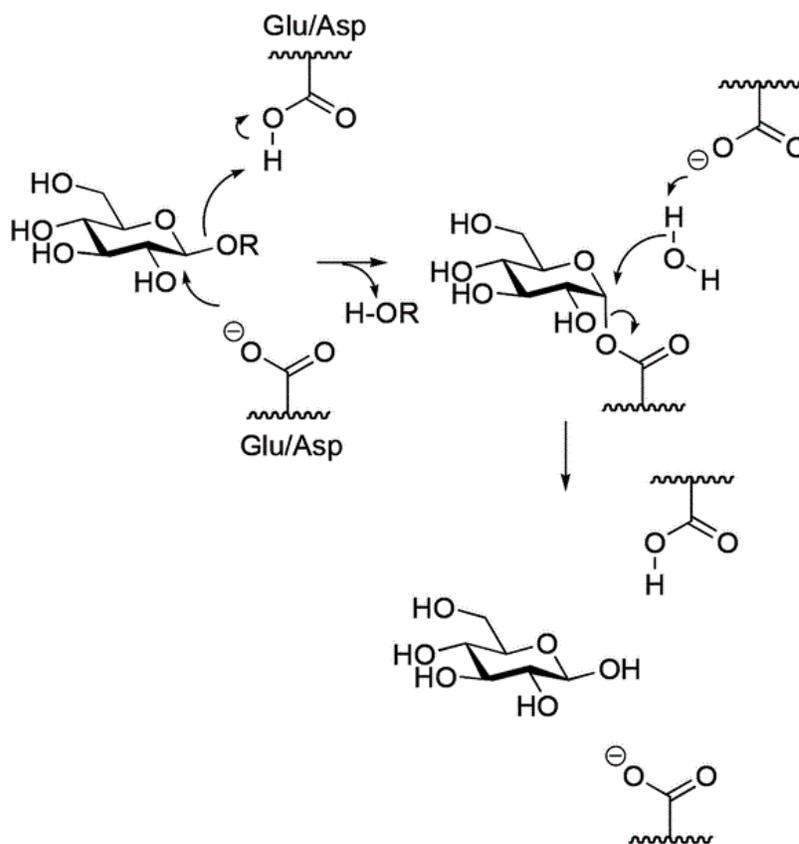


Figura 1: Mecanismo de hidrólise de ligações β-glicosídicas para as β-glicosidases, do tipo de retenção da configuração do substrato. (modificado de Li K. et al, 2014)

A constante diminuição na quantidade de combustíveis fósseis e seus efeitos na economia global e no meio ambiente tem acelerado as pesquisas em combustíveis alternativos como por exemplo o etanol de segunda geração. O bioetanol obtido de biomassa lignocelulósica vem sendo considerado uma fonte em potencial de combustível alternativo devido ao fato de ter a celulose como matéria prima, que é o componente principal das células de plantas, sendo renovável e abundante em nosso planeta (Lynd et al., 2002). A biomassa pode ser hidrolizada enzimaticamente para produzir glicose, a qual pode ser utilizada na produção de etanol a partir de processos fermentativos.

Existem muitas enzimas que são necessárias para a completa hidrólise da biomassa como por exemplo as celulasas, pectinases, ligninases, xilanases, entre outras, sendo que as celulasas desempenham a maior parte do trabalho já que a biomassa vegetal é composta de 40% de celulose (Singhania et al., 2013). O sistema celulolítico é um complexo multi-enzimático

composto basicamente por três enzimas diferentes: as exoglucanases, endoglucanases e β -glicosidases, as quais atuam conjuntamente para a completa hidrólise da celulose (Emert et al., 1974; Whitaker, 1971). As fibras de celulose são hidrolisadas primeiramente pelas endoglucanases, que conseguem atuar no meio das grandes cadeias deste oligosacarídeo liberando fragmentos menores com extremidades redutoras e não-redutoras. Após a hidrólise pelas endoglucanases as exoglucanases atuam nestes fragmentos menores liberando como produto a celobiose, a qual é um dissacarídeo que é hidrolisado em monômeros de glicose pelas β -glicosidases (Singhania et al., 2013).

Por participarem da etapa final de hidrólise da celulose, as β -glicosidases são consideradas enzimas de um passo limitante em todo o processo de hidrólise enzimática. Entretanto, um grande problema está na inibição das β -glicosidases pelo seu produto de reação que é a glicose, sendo que este processo de feedback negativo ocorre também com outras celulases. O acúmulo de glicose inibe as β -glicosidases, gerando um acúmulo de celobiose, a qual inibe as endoglucanases e exoglucanases, comprometendo assim todo o processo de hidrólise enzimática da celulose (Lynd et al., 2002). Devido à importância das β -glicosidases neste processo de hidrólise enzimática da celulose grandes esforços vêm sendo realizados na busca de enzimas tolerantes à inibição por glicose. Sendo assim o presente trabalho teve como foco identificar novas β -glicosidases da família GH1 pela abordagem metagenômica, caracterizá-las bioquimicamente e avaliar as suas tolerâncias perante à inibição por glicose.

2 OBJETIVOS

Dada à importância da metagenômica na identificação de organismos e de novos genes de interesse biotecnológico e da utilização das β -glicosidases em diversos processos industriais relacionados à degradação de celulose, o presente trabalho teve como objetivo geral realizar uma busca por uma β -glicosidase da família GH1 no metagenoma do Lago Poraquê da região amazônica, além de sua produção recombinante e estudos funcionais.

2.1 OBJETIVOS ESPECÍFICOS

- Busca por sequências de β -glicosidases em banco de dados local.
- Análise bioinformática das sequências de DNA e aminoácidos da AmBgl1-LP.

- Construção de vetor de expressão em *E.coli*.
- Expressão heteróloga da AmBgl1-LP em *E.coli*.
- Purificação da AmBgl1-LP.
- Caracterização enzimática da AmBgl1-LP, avaliando parâmetros como temperatura ótima, pH ótimo, Km, Vmax e Kcat para os substratos pNPβGli e celobiose.
- Avaliação de atividade de transglicosilação para a AmBgl1-LP.
- Avaliação da tolerância da AmBgl1-LP à inibição por glicose.

3 MATERIAL E MÉTODOS

3.1 Busca por glicosil-hidrolases da família 1 em dados de metagenoma e escolha da AmBgl1-LP

O trabalho foi realizado em parceria com a Petrobrás e é parte do trabalho de doutorado da Dra. Danyelle Toyama. O metagenoma foi realizado coletando-se amostras de 13 regiões diferentes entre o Rio Solimões, Rio Urucu, Rio Purus e lagos adjacentes entre os municípios de Manaus e Coari no estado do Amazonas (Toyama et al., 2016) (Santos-Júnior et al., 2017). Apenas as sequências obtidas do material coletado na região do Lago Poraquê (LP) foram utilizadas neste trabalho. A coleta no Lago Poraquê foi realizada na comunidade Esperança II da Costa de Santa Rosa no município de Coari, que se localiza a aproximadamente 400 km a montante de Manaus tendo o Rio Solimões como referência. O lago tem 1,8 m de profundidade sendo que as amostras foram coletadas na profundidade de 1 m para evitar a retirada de sedimento do fundo do mesmo. A temperatura da água era de 26,6°C e o pH de 5,0. O tipo de água neste ponto é chamada de “água preta”, a qual possui pouco sedimento em suspensão porém é rica em matéria orgânica dissolvida e possui pH ácido (entre 4 e 5). Para obter bactérias e archaeas, procedeu-se uma filtração prévia de algas e sedimentos da água com membranas AP 20 de microfibras de vidro com 142 mm de diâmetro (Millipore®) e Isopore 5 µm de policarbonato com 142 mm de diâmetro (Millipore®). As bactérias e archaeas foram obtidas pela passagem da água através de um filtro de PVDF (fluoruro de polivinilideno) de 0,22 µm (Sterivex GV Millipore®). O DNA destes microrganismos foi extraído com o kit *Metagenomic DNA from Environmental Water Samples* (Epicentre), e então sequenciado pelo uso da plataforma HiSeq1000 (Illumina). O sequenciamento da amostra de água do Lago Poraquê gerou 82.419.764 reads, 306.091 ORFs e 233.575 contigs com um valor de qualidade Phred médio de

36 por read. As sequências foram organizadas e anotadas em um banco de dados chamado *Amazon River Project* em nosso laboratório. A anotação foi baseada na funcionalidade dos genes.

A busca por domínios conservados de GH1 no metagenoma do Lago Poraquê iniciou-se utilizando a plataforma Pfam (Protein Families Database) (Finn et al., 2014) que é uma base de dados de famílias de proteínas que incluem suas anotações e alinhamentos de múltiplas sequências gerados pelo uso dos modelos ocultos de Markov (HMM). Para isto foi utilizado o número de acesso “PF00232” que é referente à esta família de proteínas e então foi realizado o *download* do alinhamento de 13 sequências “seed”, as quais são sequências já trimadas e no qual o perfil do modelo oculto de Markov para a família de proteínas é montado. O perfil de modelo oculto de Markov é basicamente uma análise estatística probabilística que aplicada às múltiplas sequências de proteínas em um banco de dados nos mostra os aminoácidos com maior probabilidade de estarem localizados em uma posição específica em uma cadeia polipeptídica pela comparação dessas sequências, desta forma gerando uma sequência consenso com os aminoácidos mais conservados.

Em seguida as sequências foram tratadas no software BioEdit (Tom Hall, 1999) para a retirada dos *gaps* e foram utilizadas em um alinhamento no software Clustal Omega, o qual possibilita realizar alinhamentos entre mais de duas sequências utilizando o perfil HMM, para gerar uma sequência consenso modelo que apresenta os aminoácidos mais conservados em proteínas da família GH1. O alinhamento no software Clustal Omega gerou uma sequência consenso chamada de “consensoGH1” a qual foi trimada manualmente. A sequência consenso final foi então utilizada como sequência *query* em um alinhamento de aminoácidos utilizando BlastP contra as ORFs traduzidas presentes no banco de dados de contigs do metagenoma do Lago Poraquê. As sequências com E-value menores que $1.e^{-05}$ foram selecionadas para análise da possível proteína traduzida utilizando a ferramenta Translate Tool do ExPASy (SIB) e então uma dentre 26 sequências encontradas, que se apresentava completa com códon de iniciação e terminação, foi escolhida para prosseguir com os experimentos de clonagem e recebeu o nome de *AmBglI-LP*.

3.2 Expressão heteróloga da AmBgl1-LP em *Escherichia coli*

3.2.1 Clonagem

Para a expressão heteróloga da ORF *AmBgl1-LP* em bactéria foi escolhido o vetor pET-28a (Novagen) que possui o promotor T7 em sua sequência, o qual propicia altos níveis de expressão gênica na presença da RNA polimerase T7. As enzimas de restrição utilizadas na clonagem do gene de interesse foram *Nde*-I e *Bam*H-I, as quais não possuem sítios de restrição na sequência de nucleotídeos do gene *AmBgl1-LP* e possibilitam a sua clonagem no vetor pET-28a em fusão com uma cauda de histidinas nas porções N-terminal ou C-terminal da proteína recombinante.

Os *primers* desenhados para a clonagem, AmBgl1-LPpETF e AmBgl1-LPpETR, possuem em suas sequências os sítios de restrição para *Nde*-I e *Bam*H-I respectivamente, além de três adeninas nas extremidades dos mesmos como facilitadores para a ancoragem das enzimas de restrição no processo de digestão do produto de PCR. Neste caso o códon terminador da tradução do gene da *AmBgl1-LP* foi mantido no *primer* reverso para que fosse codificada uma cauda de histidinas apenas na porção N-terminal da proteína recombinante. A construção final adicionou 20 aminoácidos na porção N-terminal da proteína recombinante, incluindo a cauda de histidinas e outros aminoácidos provenientes do vetor na seguinte sequência: MGSSHHHHHSSGLVPRGSH; aumentando a massa molar da AmBgl1-LP de 51,6 kDa para aproximadamente 53,7 kDa. Os primers utilizados na clonagem estão representados abaixo:

AmBgl1-LPpETF: AAACCATATGCGTAATTTTCCGAAAGGT

AmBgl1-LPpETR: AAAGGATCCTCAATCTAGAGCGTTACGCG

Para a obtenção da ORF da *AmBgl1-LP* foi realizada uma reação em cadeia da polimerase (PCR), utilizando 6 e 3 ng de DNA metagenômico do Lago Poraquê (LP), 0,6 U de *Pfu* DNA polimerase (Thermo Scientific), 0,2mM de dNTPs, 10 pMol dos primers AmBgl1-LPpETF e AmBgl1-LPpETR e tampão *Pfu* MgSO₄ plus 1X (Thermo Scientific) em um volume final de 50 µL. Após a amplificação, o produto de PCR foi submetido à uma eletroforese em gel de agarose 1% e purificado do mesmo com o Kit “Wizard SV Gel and PCR Clean-up System” (Promega).

O plasmídeo pET-28a e o produto de PCR contendo o gene da *AmBglI-LP* purificado foram clivados com as enzimas *Nde*-I e *Bam*H-I (Thermo Scientific) em reações separadas contendo 3 µg de DNA, 5 U das enzimas de restrição e tampão Orange™ 1X (Thermo Scientific) em um volume final de 50 µL. As reações foram incubadas em banho a 37°C por 16h e após este período as enzimas foram inativadas por aquecimento a 65°C por 20 minutos.

Em seguida, os produtos das reações foram submetidos à eletroforese em gel de agarose 1% (Invitrogen) para confirmação da clivagem e as bandas referentes ao vetor clivado (5,3 Kb) e ao gene *AmBglI-LP* (1,4Kb) foram recuperadas do gel com auxílio do Kit “Wizard SVGel and PCR Clean-up System” (Promega). A reação de ligação foi preparada seguindo a relação molar de inserto-plasmídeo na proporção de 5:1. Para isto foi utilizado 1 U da enzima T4 DNA Ligase (Invitrogen), tampão da ligase 1X (Invitrogen), 50 ng de plasmídeo e 65 ng de inserto em um volume final de 20 µL de reação, a qual foi incubada a 4°C por 16 horas. Posteriormente, a mistura de ligação foi utilizada para transformar células de *Escherichia coli* linhagem DH5α CaCl₂ competentes. Para tanto, 2 µL da reação de ligação foram adicionados a 200 µL de células químio-competentes, seguido de incubação a 42°C por 90 segundos e posterior resfriamento em gelo por 1 minuto para a efetivação do choque térmico. Em seguida foram adicionados 800 µL de meio de cultura LB-Broth às células e estas foram incubadas em banho a 37°C por 45 minutos. Para o plaqueamento, 200 µL da mistura foram dispensados sobre placas de Petri contendo 15 mL de meio LB-Broth ágar e 25 µg/mL de antibiótico kanamicina. As placas foram incubadas por 16 horas a 37°C em estufa e 6 colônias isoladas foram escolhidas para buscar colônias recombinantes por PCR de colônia.

Três clones positivos foram inoculados em 5 mL de meio LB Broth líquido por aproximadamente 16 horas a 37 °C para a extração do DNA plasmidial, a qual foi realizada por lise alcalina utilizando o auxílio do kit “QIAprep® Spin Miniprep kit” (Qiagen). A confirmação da inserção do gene de interesse em correta fase de leitura foi feita por sequenciamento pelo método de terminação de cadeia desenvolvido por Sanger (Sanger, 1977) no equipamento MegaBace 1000 (GE Healthcare Life Science) utilizando o kit DYEnamic™ ET Dye Terminator Kit (MegaBACE™). Para este procedimento foram utilizados aproximadamente 300 ng de plasmídeo e 1 pMol de primers T7 promoter e T7 terminator em reações separadas.

3.2.2 Expressão e purificação da AmBgl1-LP

As células de *E. coli* linhagem Rosetta (DE3), utilizadas para expressão heteróloga, possuem o fragmento DE3 que codifica a RNA polimerase do bacteriófago T7, a qual se liga fortemente ao promotor T7 do vetor pET-28a, garantindo uma alta expressão gênica. Esta cepa também possui o gene da T7 RNA polimerase regulado pelo repressor lac, uma molécula naturalmente produzida pela bactéria que inibe a transcrição do gene da T7 RNA polimerase no sítio DE3. Este repressor, porém, se complexa com a alolactose ou análogos como o IPTG (Isopropil- β -D-tiogalactopiranosídeo) quando presentes no meio, e isto faz com que o repressor perca a afinidade ao operador da região controladora do gene da T7 RNA polimerase, garantindo assim a sua transcrição e início da expressão do gene de interesse sob efeito do promotor T7. Além destes elementos, esta cepa também possui genes para tRNAs codificados em plasmídeos resistentes ao antibiótico cloranfenicol, específicos para os códons AUA, AGG, AGA, CUA, CCC e GGA, que são raramente utilizados pela *E. coli*, permitindo a expressão de proteínas heterólogas que possuam estes códons neste hospedeiro.

O procedimento de transformação química das células competentes de *E. coli* Rosetta (DE3), utilizando o plasmídeo pET-28a_AmBgl1-LP, ocorreu da mesma maneira descrita para a linhagem DH5 α e, após o crescimento de colônias na placa, uma delas foi inoculada em 5 mL de meio LB Broth acrescido de 25 μ g/mL dos antibióticos kanamicina e cloranfenicol. Este pré-inóculo foi mantido em shaker sob agitação constante de 250 rpm a 37°C por 16 horas. Após este período 500 μ L do pré-inóculo crescido foram adicionados a 50 mL do inóculo contendo meio LB Broth, o qual foi mantido sob agitação constante de 200 rpm a 37°C por aproximadamente 2 horas, etapa em que as células estavam na fase exponencial de crescimento com densidade óptica (D.O) de 0,6 no comprimento de onda de 600 nm, medida em espectrofotômetro U-5100 Hitachi. Neste momento foi coletada uma alíquota de 1 mL do inóculo como controle não induzido e então foi adicionado IPTG (Isopropil- β -D-tiogalactopiranosídeo) ao mesmo em uma concentração final de 0,4 mM para iniciar a indução da AmBgl1-LP. A partir de então, alíquotas de 1 mL foram coletadas do inóculo nos períodos de 1, 2, 3, 4 e 21 horas de indução, realizada a 37°C com 200 rpm, para avaliação da produção da proteína recombinante em SDS-PAGE 12% (Laemmli, 1970). Após o período de 21 horas de indução, as células foram coletadas, centrifugadas a 15000 g a 4°C por 5 minutos em centrífuga Sorvall RC-5C Plus, ressuspendidas em tampão de lise pH 8,0 (NaCl 100 mM, NaH₂PO₄ 50 mM, Tris 10 mM) e

submetidas à lise celular por meio de sonicação em 8 repetições de 59,9 segundos, com intervalos de 30 segundos à uma amplitude de 14%, utilizando o sonicador Sonic Desmembrator 500 (Fisher Scientific).

Em seguida, o produto de lise celular foi novamente centrifugado a 20000 g a 4°C por 15 minutos em centrífuga Sorvall RC-5C Plus para separação das frações solúveis e insolúveis de proteínas, sendo coletados 500 µL de sobrenadante e 500 µL do precipitado ressuspendido em 5 mL de tampão de lise pH 8,0, os quais foram utilizados para posterior análise em um teste de solubilidade em SDS-PAGE 12%.

Após lise celular, a fração de proteínas solúveis correspondente a 21h de indução foi submetida ao processo de purificação por afinidade em coluna com resina de níquel “Ni-NTA Superflow” (Qiagen). O processo de purificação por afinidade molecular nesta coluna se baseia na ligação entre a proteína alvo e o ligante imobilizado na resina. A proteína recombinante expressa em pET-28a possui em sua região N-terminal uma cauda com 6 histidinas, as quais possuem um anel de imidazol que possui afinidade pelo ligante imobilizado na coluna, o Ni-NTA (níquel – ácidonitrilacético). Isto implica em uma forte ligação da proteína recombinante na coluna e que permite a sua purificação em uma única etapa. Para a eluição da proteína foi utilizado tampão de lise pH 8,0 com crescentes concentrações de imidazol (10, 25, 50, 75, 100 e 250 mM), o qual compete com a proteína pela ligação ao ligante na coluna, liberando a proteína na forma pura. Em seguida estas frações foram analisadas em SDS-PAGE 12% e as correspondentes à proteína pura foram submetidas à diálise em membranas de celulose (Pierce SnakeSkin™ 3,500 MWCO) em tampão PBS pH 8,0 (NaCl 137 mM, KCl 2,7 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 1,8 mM), sendo realizadas 3 trocas de tampão em uma proporção de proteína – tampão de 1:1000 (V/V). O processo ocorreu sob agitação suave a 4 °C por aproximadamente 20 horas. Após o processo de diálise a proteína foi filtrada em filtros com porosidade de 0,22 µM (Milipore) e quantificada pelo método do ácido bicinônico (BCA) (Walker, 1994) utilizando o kit BCA protein assay reagent (Thermo Scientific).

3.3 Caracterização enzimática da β-glicosidase

Os ensaios enzimáticos foram realizados utilizando p-nitrofenil-β-D-glicopiranosídeo (pNPβGli) e celobiose como substratos. O pNPβGli é similar à celobiose que é um dos substratos naturais das β-glicosidases, com a diferença de que este possui uma molécula

de p-nitrofenolato ligada à glicose por uma ligação β -1,4, ao invés de dois resíduos de glicose como no substrato natural. Quando hidrolisado o p-nitrofenolato apresenta cor amarelada em pH básico e pode ter sua absorvância lida a 400-420 nm em espectrofotômetro.

Inicialmente, para verificar se a enzima estava ativa, foi realizado um ensaio qualitativo no qual foi preparada uma reação em microtubo de 1,5 mL adicionando 25 μ L de pNP β GLi 10 mM, 5 μ L da enzima AmBgl1-LP com massa final de 100 ng e 70 μ L de tampão acetato de sódio pH 5,0. A reação foi incubada por 30 minutos a 50 °C e após este tempo foram adicionados 100 μ L de glicina NaOH 0,4 M pH 10,8 para interromper a reação. Como a coloração da mistura mudou de transparente para amarelo em comparação com o controle sem enzima, nós consideramos que a mesma estava ativa hidrolisando o pNP β GLi.

Após verificar que a enzima estava ativa foi avaliado o seu pH ótimo de atividade. Para isto foram preparadas reações em placas de 96 poços com fundo reto colocando 25 μ L de pNP β GLi 10 mM, 5 μ L de enzima AmBgl1-LP com massa total de 100 ng e 70 μ L de tampão universal de Davies (ácido cítrico 110 mM, potássio dihidrogeno fosfato 100 mM, tetraborato de sódio 95 mM, tris hidroximetil aminometano 100 mM, cloreto de potássio 100 mM), variando o pH de 3 a 9 em intervalos de 0,5 a cada medida (Davies, 1959). As reações foram incubadas a 30 °C por 15 minutos e interrompidas com a adição de 100 μ L de glicina-NaOH 0,4 M pH 10,80. A absorvância resultante foi lida em espectrofotômetro PerkinElmer Victor3™ com comprimento de onda de 420 nm. Todas as medidas foram feitas em quintuplicata. O pH que apresentou maiores valores de absorvância foi definido como tendo atividade relativa de 100%.

Para o ensaio de temperatura ótima foram preparadas reações em placas de 96 poços com fundo reto adicionando 25 μ L de pNP β GLi 10 mM, 5 μ L de enzima AmBgl1-LP com massa total de 100 ng e 70 μ L de tampão universal de Davies pH 5,5 (Davies, 1959). As reações foram incubadas em diferentes temperaturas, variando de 30 °C a 70 °C em intervalos de 5 °C, por 15 minutos e interrompidas com a adição de 100 μ L de glicina-NaOH 0,4 M pH 10,80. A absorvância resultante foi lida em espectrofotômetro PerkinElmer Victor3™ com comprimento de onda de 420 nm. Todas as medidas foram feitas em quintuplicata. A temperatura que apresentou maiores valores de absorvância foi definida como tendo atividade relativa de 100%.

No ensaio de termoestabilidade foram avaliadas as seguintes temperaturas: 20, 30, 35, 40, 50 e 60 °C; nos tempos de 0, 5, 10, 15, 20, 30, 45 e 60 minutos. Para isto alíquotas de 200 µL de AmBgl1-LP foram pré-incubados nestas temperaturas e tempos descritos acima. Após isto foram realizados ensaios de atividade da seguinte maneira: 50 µL de AmBgl1-LP, equivalente a 100 ng de enzima, mais 50 µL de substrato pNPβGli 4 mM, diluído em tampão citrato-fosfato de sódio pH 6,0 (ácido cítrico 18 mM, fosfato de sódio dibásico 62 mM), foram pipetados em placas de 96 poços de fundo reto e incubados a 35 °C por 20 minutos. A cada 5 minutos, até o tempo final de 20 minutos, um grupo de reações era interrompido pela adição de 100 µL de carbonato de sódio 0,5 M e tinham suas absorvâncias resultantes lidas em espectrofotômetro PerkinElmer Victor3™ com comprimento de onda de 420 nm. Os coeficientes angulares das retas formadas pelas medidas de quatro tempos (5, 10, 15 e 20 minutos) de atividade, de cada tempo e temperatura de pré-incubação, foram pontuados em um gráfico de atividade residual por tempo de incubação para cada temperatura analisada. A temperatura de pré-incubação em conjunto com o tempo de pré-incubação que apresentaram maiores valores de absorvância foram definidos como tendo atividade residual de 100%.

Para os ensaios de cinética enzimática as reações foram realizadas utilizando 50 µL de concentrações crescentes de pNPβGli (0,05; 0,1; 0,15; 0,25; 0,4; 0,5; 1; 1,5; 2; 2,5; 3; 3,5; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15; 16; 17; 18; 19 e 20 mM final) preparado em tampão citrato-fosfato de sódio pH 6,0 e 50 µL de AmBgl1-LP (87 ng). A mistura foi incubada a 35°C por 40 minutos, com interrupções a cada 10 minutos pela adição de 100 µL de carbonato de sódio 0,5 M e tiveram suas absorvâncias resultantes lidas em espectrofotômetro ELx800 (BioTek) com comprimento de onda de 415 nm (Tamaki et al., 2014). Os coeficientes angulares das retas formadas pelas medidas de quatro tempos (10, 20, 30 e 40 minutos) de atividade para cada concentração de substrato foram pontuados em um gráfico de velocidade iniciais. Os parâmetros cinéticos como K_{cat} , V_{max} e K_m foram determinados pelo ajuste dos dados das curvas de $v_0/[S]$ utilizando a equação de Michaelis-Menten.

Ensaio com celobiose também foram executados para a β-glicosidase. Para isto as reações foram preparadas com 50 µL de concentrações crescentes de celobiose (0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; 0,9; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 12; 14; 16; 18; 20; 24; 28; 32; 36 e 40 mM) em tampão citrato-fosfato de sódio pH 6,0, mais 50 µL de AmBgl1-LP (87 ng). A mistura foi

incubada a 35°C por 16 minutos, com interrupções a cada 4 minutos por fervura. Após decorrido este tempo a detecção da glicose liberada foi feita utilizando o kit Glicose-PP (Analisa™ - Gold Analisa Diagnóstica, Belo Horizonte, MG) em espectrofotômetro ELx800 (BioTek) com comprimento de onda de 490 nm. Os coeficientes angulares das retas formadas pelas medidas de quatro tempos (4, 8, 12 e 16 minutos) de atividade para cada concentração de substrato foram pontuados em um gráfico de velocidade iniciais. Os parâmetros cinéticos como K_{cat} , V_{max} e K_m foram determinados pelo ajuste dos dados das curvas de $v_0/[S]$ utilizando a equação de Michaelis-Menten utilizando os softwares Enzfitter (Leatherbarrow, 1987), GraphPad Prism v. 6.00 e OriginPro v. 8.00.

Ensaio para avaliar a possível ocorrência de transglicosilação foram efetuados medindo a liberação de p-nitrofenolato e glicose a partir do substrato pNPβGli. Para isto as reações foram preparadas da seguinte maneira: em placas de 96 poços de fundo reto foram adicionados 50 μL de concentrações crescentes de pNPβGli (0,1; 0,25; 0,4; 0,5; 1; 2; 3; 4; 5; 7; 9; 11; 13; 15; 17 e 20 mM), dissolvido em tampão citrato-fosfato de sódio pH 6,0 e 50 μL de AmBgl1-LP (87 ng). As placas foram incubadas a 35 °C por 20 minutos, com interrupções a cada 5 minutos realizadas de duas maneiras diferentes. Em um grupo de placas foi medida a formação de p-nitrofenolato com a adição de 100 μL carbonato de sódio 0,5 M e leitura da absorbância resultante em 415 nm em espectrofotômetro ELx800 (BioTek). No outro grupo de placas foi medida a formação de glicose a partir do mesmo substrato, após interrupção da reação por fervura por 4 minutos, porém utilizando o kit de Glicose-PP (Analisa™) seguindo as instruções do fabricante. Os coeficientes angulares das retas formadas pelas medidas de quatro tempos (5, 10, 15 e 20 minutos) de atividade para cada concentração de substrato foram pontuados em um gráfico de velocidade iniciais.

Após a determinação dos parâmetros cinéticos, iniciou-se a etapa de avaliação da resistência das enzimas à inibição por glicose. A glicose, que é produto final da reação de clivagem da celobiose, é um monossacarídeo que geralmente inibe as β-glicosidases e a busca por enzimas que sejam resistentes à inibição é de grande interesse comercial (Singhania et al., 2013).

Os ensaios de inibição por glicose foram realizados em placas de 96 poços de fundo reto adicionando 50 μL de concentrações crescentes de pNPβGli (0,1; 0,15; 0,25; 0,4; 0,5; 1; 2; 3; 4 e 5 mM), em tampão citrato-fosfato de sódio pH 6,0 com concentrações crescentes de

D-glicose (90; 280; 444; 660 e 800 mM) e 50 μ L de AmBgl1-LP (87 ng). As placas foram incubadas a 35 $^{\circ}$ C por 20 minutos, com interrupções a cada 5 minutos pela adição de 100 μ L de carbonato de sódio 0,5 M. A absorbância resultante de cada reação foi lida em espectrofotômetro ELx800 (BioTek) com $\lambda = 415$ nm. Os coeficientes angulares das retas formadas pelas medidas de quatro tempos (5, 10, 15 e 20 minutos) de atividade para cada concentração de substrato e de inibidor foram pontuados em um gráfico de velocidade iniciais. Os valores obtidos em diferentes tempos e concentrações de pNP β Gli e D-glicose foram utilizados para determinar o K_i da β -glicosidase para a D-glicose utilizando o software Enzfitter, GraphPad Prism v. 6.00 e OriginPro v. 8.00.

4 RESULTADOS E DISCUSSÃO

4.1 Identificação da AmBgl1-LP nas sequências metagenômicas

Após o alinhamento realizado no *software* Clustal Omega e a “trimagem” manual de dados foi possível obter uma sequência consenso de aminoácidos que foi utilizada para encontrar proteínas da família GH1 no banco de dados de ORFs traduzidas do metagenoma do Lago Poraquê (8LP) utilizando o BlastP. A figura 2 que mostra o alinhamento de sequências “seeds”, assim como a sequência “consensoGH1” utilizada na identificação das ORFs de GH1 no banco de dados local estão representados abaixo:

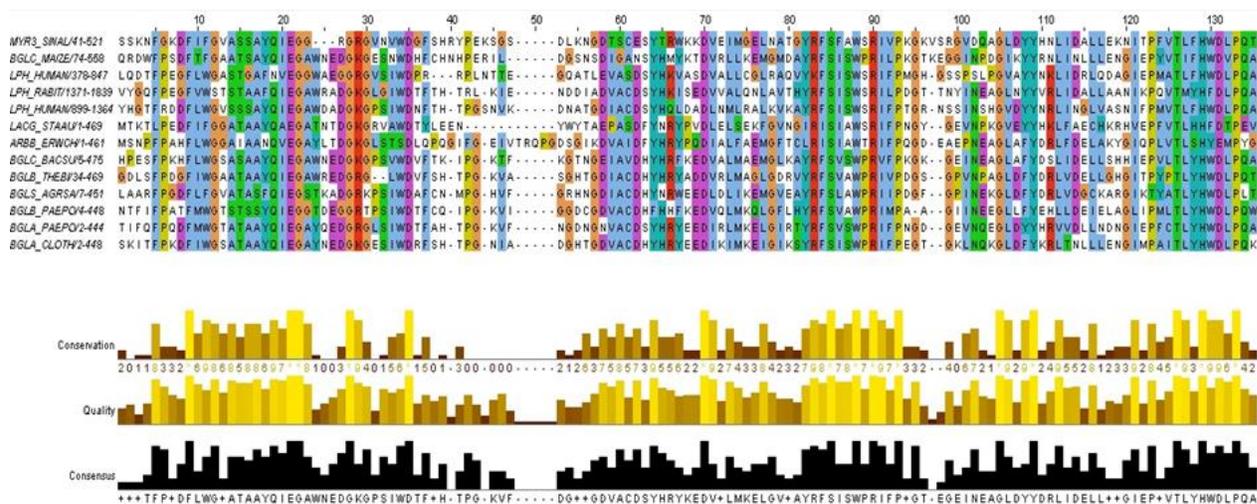


Figura 2: Alinhamento das sequências “seed” de GH1 do Pfam para gerar uma sequência consenso utilizada na busca de β -glicosidasas no banco de dados local em nosso laboratório.

>ConsensoGH1

```
LMSSTDFGKTFPEKDFLWGATVATAAYQIEGAWNEDGKGPSIWDTFCSTHTPGKVFDDHNTGDVACD
SYHRYKEDVARLMKELGVKTAYRFSISWPRIFPDKGTEGEINEAGLDYYDRLIDELLAANGIEPFMYVTLY
HWDLPQALQDEKYGGWELNRSTIDADFKDEYADETVCFKRFGRVDFKFNWILTFNPPWVAVASFLWYLG
YGTGVHAPGIKDPGDGNSERTVAPYEIQRTAAHHNLLLAHARAVHKLYREHKMVYRGQDKGQIGIVLNLS
WAYPASPSGKNRPEDVEAAERAMSDQFHNGWFLDPVFKRNGDYPEEMREILQWVGRSSEKLQGSRLP
EFFTEEDKELIAKGPYDFLGLNYYTSR FILRIVKRNDPPGNSENTTGDGAKGPSSTYDNGFLPGVEREIVN
PGSLWPKSTVTDMMGWIIIMYPEGLRDRLLMNRLKEDYGNPPIYITENGAGFIYKDEVEESNGAKQTVEND
DTKRIDYLRQSHLNQVHKAIREDDGVDLRGYFAWSLLDNFEWANGYSKRFGVLVHYVDFDTSSLQERTPKE
SAYWYKEVIAEKRSTNNGF
```

O alinhamento, utilizando o método de BlastP, entre a sequência consensoGH1 e o banco de dados de contigs do metagenoma do Lago Poraquê (8LP) mostrou 42 sequências que poderiam ser possíveis glicosil-hidrolases da família 1, entretanto apenas 26 foram selecionadas para análise da sequência de aminoácidos devido aos seus valores de E-value menores que $1.e^{-5}$. Das 26 sequências analisadas apenas uma mostrou ORF completa para β -glicosidase da família GH1 ao realizar um alinhamento no banco de dados do NCBI (National Center for Biotechnology Information) utilizando o método BlastP. A sequência escolhida para prosseguir com os experimentos de clonagem, 8LP_636.9967, apresentou identidade de 41% com a sequência consensoGH1, um E-value de $2e^{-76}$ e score de 284. A sequência de nucleotídeos da ORF do contig 8LP_636.9967 e seus aminoácidos da proteína codificada, agora chamada de AmBgl1-LP, podem ser visualizados na figura 3:

```

ATG CGT AAT TTT CCG AAA GGT TTC ACT TTT GGG GCA GCA ACG GCT GCC TAT CAA ATT GAA GGC GCT TAT CAC GAG GAT GGT AAG GGA CTT
M R N F P K G F T F G A A T A A Y Q I E G A Y H E D G K G L
TCC ATT TGG GAC ACT TTT AGC CAT ATT TCG GGC AAG GTA GAC AAC AAC GAT AAC GGG GAT ATT GCC TGC GAT CAC TAC CAT CGA TAC AAA
S I W D T F S H I S G K V D N N D N G D I A C D H Y H R Y K
GAA GAT ATT TCA TTG CTG AAA GAA TTA GGC ATT AAG AAT TAT CGA ATG AGC ATC GCC TGG ACT CGG CTC TTT CCT ACT GGT CGG GTA GAT
E D I S L L K E L G I K N Y R M S I A W T R L F P T G R V D
TCG CGA AAT CAA AAA GGC TTT GAC TAC TAC AAC GAT CTC ATT GAC ACT CTA CTT GAG GCC GGA ATA GAA CCC TTC GTT ACT TTA TAT CAC
S R N Q K G F D Y Y N D L I D T L L E A G I E P F V T L Y H
TGG GAT TTA CCA CAA GCC CTT GAA GAT TTA GGC GGC TGG AAA AAT CGG GAA ATT CTT ACC TGG TTT GGT GAA TAC GCA GCT GAA GTG GCA
W D L P Q A L E D L G G W K N R E I L T W F G E Y A A E V A
AAG GCA TTT GGT GAT CGA GTC ACG TAT TTC TCC CCT ATT AAT GAG CCC TGG GTT GTC GCC TGG CTA GGA CAT GGA ATG GGA ATT CAT GCA
K A F G D R V T Y F S P I N E P W V V A W L G H G M G I H A
CCA GGA GAC AAA GAT CGT GCG AAA GCT TTT GCT GCA GCG CAT CAC ACC GTT GTT GCG CAT GGC GTC GCG AAT CGT GCA ATG AAA GCT GTC
P G D K D R A K A F A A A H H T V V A H G V A N R A M K A V
TAC CCA CAT CTG AAA ATT GGT CCA GTT CTT AAC CAA ACA AAT TTT CCT ATT GAT GAT TCA ACT AAT GCA GAG TTG AAG CGT GCT CAG GAT
Y P H L K I G P V L N Q T N F P I D D S T N A E L K R A Q D
ATT TTG GAC GCG AAT CAA AAT CGA TTT TGG ATT GAT GCA ATC TTT CAC AAG AGA TAT CCA GAG ATC CTA ATT GAA AAT TTT CCT GAA GAT
I L D A N Q N R F W I D A I F H K R Y P E I L I E N F P E D
ATT CTT CCA AAT ATT CAT GAG GGG GAT TTT GAT TTA GCA GCT ACG CCA AAT GAT TGG CTT GGA ATC AAT TAC TAT TTT GAT ACA CCT ATG
I L P N I H E G D F D L A A T P N D W L G I N Y Y F D T P M
AAA GCG AGC TCT CGA CCT GTA GCA ACG GCC TTC GAT CCG GCG GCG CTT CTA GGA TTG AAT GTT GAC CAG ACG CCT CCA CCA CCT CTT ACT
K A S S R P V A T A F D P A A L L G L N V D Q T P P P P L T
GAT ATG GGT TGG CCA ATT TCT CCA GAT GGG TTA AGT AAT TTA CTT ATT CGT TGG CAT AAA GAG ATT GGT AAC CGT TTA CCG CCG ATC TAC
D M G W P I S P D G L S N L L I R W H K E I G N R L P P I Y
ATC ACT GAA AAT GGA GTG GCA TAC GAT GAC GGT CCT GGC GAA GAT GGA CTT ATT CAC GAC CAG CGC CGA ATC GAT TAT TTG CGA GGG CAT
I T E N G V A Y D D G P G E D G L I H D Q R R I D Y L R G H
CTT GCA GCC CTC CTT GAT GCT TGT GAA GTT GGA GTA GAT GTC AGG GGA TAT TTT CAG TGG TCC CTT CTT GAT AAC TTT GAA TGG GCT CTG
L A A L L D A C E V G V D V R G Y F Q W S L L D N F E W A L G
GGC TAT TCC AAG CGT TTT GGC ATC GTT CAC GTT GAC TAC CAA ACC CTT AAG CGC ACC CCA AAG GAT TCG GCG CTC TGG TAC AGC GGA GTA
G Y S K R F G I V H V D Y Q T L K R T P K D S A L W Y S G V
ATT GCG CGT AAC GCT CTA GAT TGA
I A R N A L D *

```

Figura 3 : ORF e sequência de aminoácidos da *AmBgl1-LP*. A figura indica a sequência de nucleotídeos que codifica a proteína AmBgl1-LP e seus aminoácidos correspondentes; nas caixas vermelhas estão indicados os códons iniciadores e terminadores respectivamente.

Como podemos observar na Figura 3 a ORF *AmBgl1-LP* possui 1374 nucleotídeos e a proteína codificada 457 aminoácidos; análises de parâmetros protéicos realizados utilizando o programa ProtParam no portal de bioinformática ExPASy do SIB (Swiss Institute of Bioinformatics) revelaram que a proteína possui massa molecular estimada de 51,6 kDa e um ponto isoelétrico teórico de 5,29. Além das análises de massa molecular e pI foi avaliado se a proteína possuía algum sinal de secreção utilizando a ferramenta SignalP4.1, a qual não revelou nenhum peptídeo sinal para secreção.

Para identificar os domínios conservados, regiões de sítio catalítico e com quais organismos a sequência da AmBgl1-LP tinha mais identidade, a mesma foi alinhada com outras

β -glicosidasas da família GH1 utilizando o software Multialin (Corpet, 1988); o resultado do alinhamento pode ser visto na Figura 4:

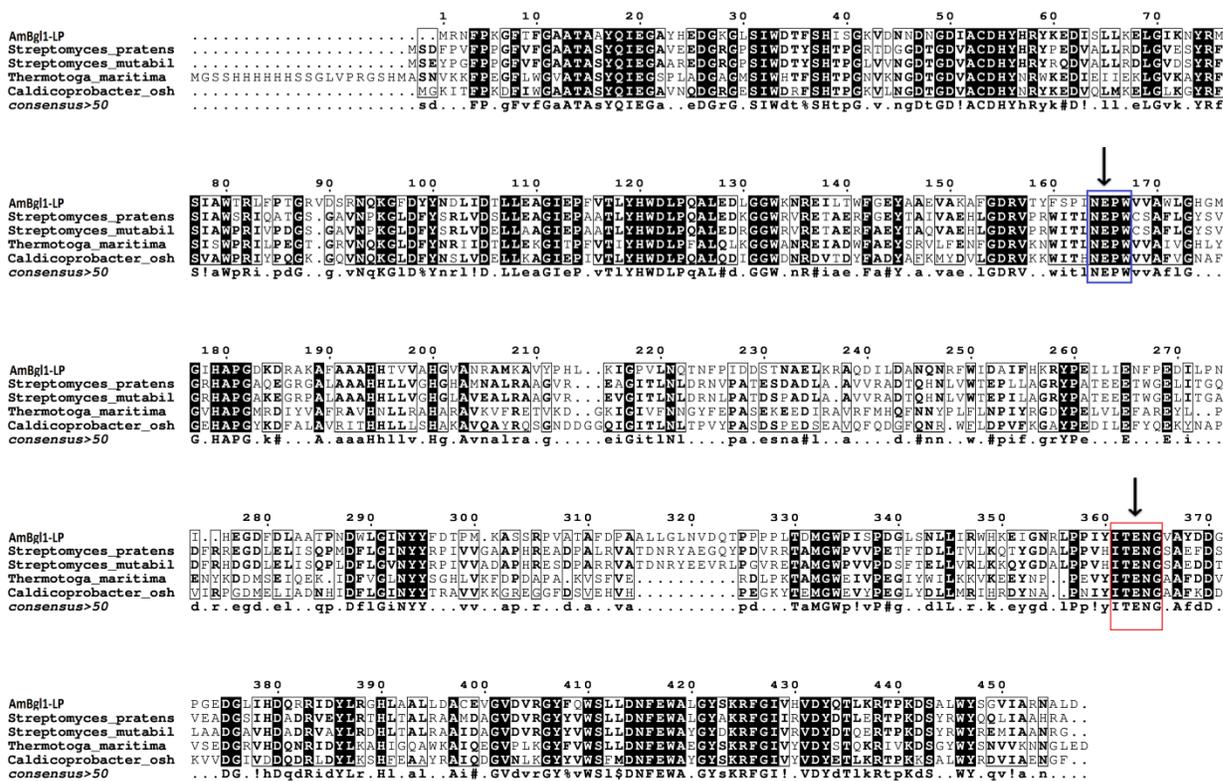


Figura 4 : Alinhamento entre AmBgl1-LP e outras β -glicosidasas. Na figura as regiões com fundo preto representam resíduos conservados na família de GH1, sendo que o resíduo de glutamato (E) 363, indicado por uma seta dentro de um retângulo vermelho, é o nucleófilo catalítico; já o resíduo de glutamato (E) 165, indicado em um retângulo azul, representa o ácido/base geral; ambos os resíduos participam diretamente da reação de catálise rompendo a ligação β 1-4 entre dissacarídeos. As β -glicosidasas utilizadas no alinhamento vêm dos organismos *Streptomyces mutabilis* (KFG76903.1), *Streptomyces pratensis* (WP_014153373.1), *Thermotoga maritima* (WP_015646202.1) e *Caldicoprobacter oshimai* (WP_025747479.1).

A partir do alinhamento mostrado na Figura 4 utilizando seqüências de β -glicosidasas com 48% a 51% de identidade com a AmBgl1-LP, que foram os valores mais altos encontrados, foi possível identificar as regiões mais conservadas da enzima e também procurar pelos resíduos catalíticos, que já foram amplamente descritos nesta classe de enzimas (Withers, 1990), com o auxílio do banco de dados do CAZy (Carbohydrate-Active Enzymes), como sendo o glutamato 363 o nucleófilo catalítico e o glutamato 165 sendo o ácido/base geral na AmBgl1-

LP. Além disto, os organismos com sequências mais próximas da AmBgl1-LP foram os do gênero *Streptomyces*.

Com sua sequência de DNA e aminoácidos determinadas e provada a sua classificação como uma β -glicosidase da família GH1 a partir de múltiplos alinhamentos teve início a etapa de clonagem do gene para a produção heteróloga da AmBgl1-LP.

4.2 Expressão Heteróloga em *Escherichia coli*

Inicialmente foi realizada a amplificação do gene *AmBgl1-LP* a partir de DNA metagenômico total do Lago Poraquê, a qual amplificou um fragmento específico de tamanho esperado de aproximadamente 1400 pb como pode ser visualizado na Figura 5:

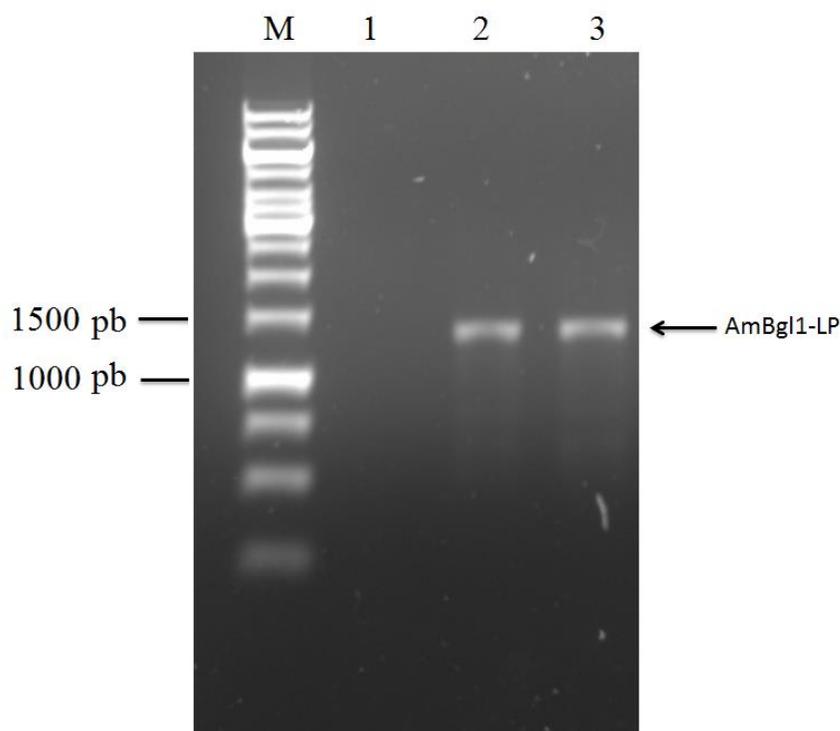


Figura 5 : Amplificação da ORF *AmBgl1-LP*. M: marcador de massa molecular GeneRuler 1Kb (Fermentas); 1: controle negativo da reação (sem molde); 2 e 3: amostras amplificadas, utilizando 3 e 6 ng de DNA metagenômico do Lago Poraquê respectivamente.

Após a amplificação do gene os fragmentos foram purificados do gel, clivados com as enzimas de restrição *Nde*-I e *Bam*H-I e utilizados em uma reação de ligação com o

plasmídeo pET-28a clivado com as mesmas enzimas. Em seguida foram transformadas células de *E. coli* linhagem DH5 α e a busca por clones recombinantes foi analisada por PCR de colônia e seu resultado pode ser visualizado na Figura 6, a seguir:

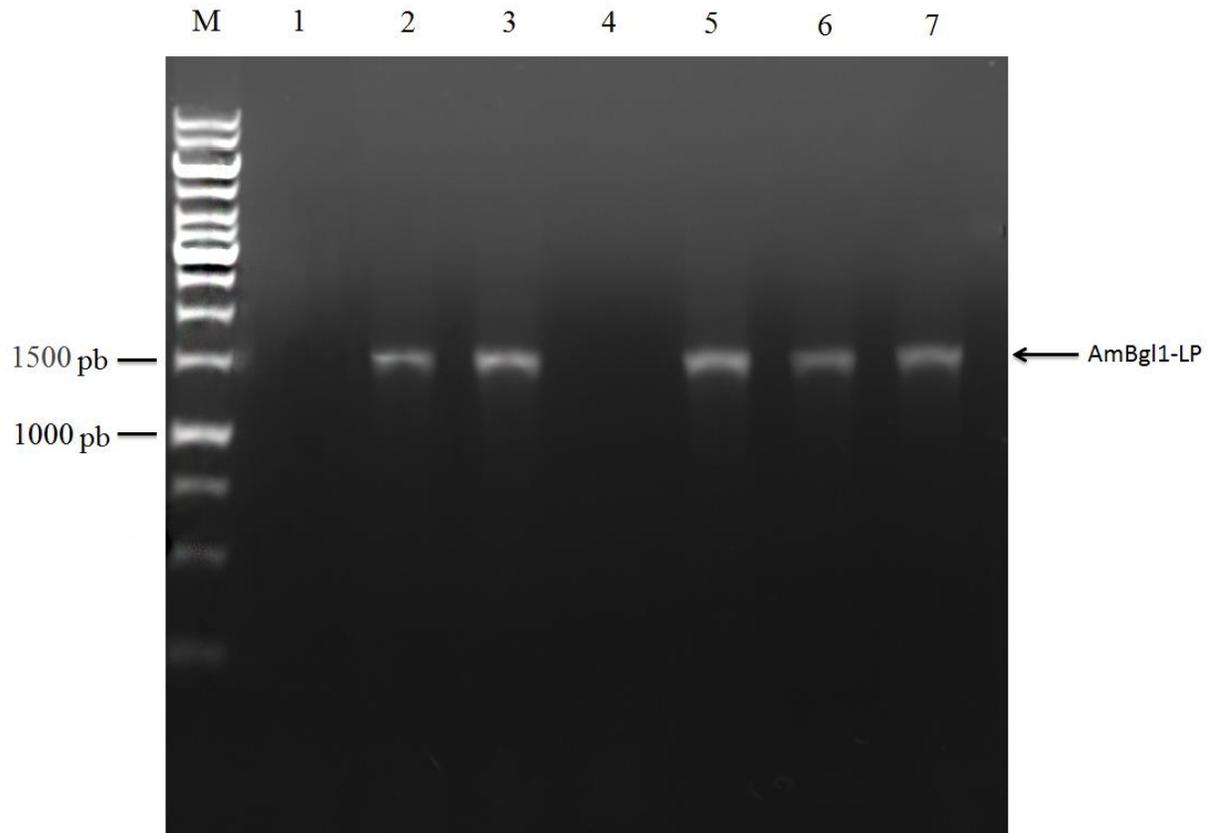


Figura 6 : PCR de colônia para a ORF *AmBgl1-LP* no plasmídeo pET-28a em *E. coli* DH5 α . M: marcador de peso molecular GeneRuler 1 Kb; 1: controle negativo; 2, 3, 5, 6 e 7: clones recombinantes; 4: amostra não amplificada.

Analisando a Figura 6 é possível observar que as bandas amplificadas apresentaram tamanho ligeiramente superior a 1500 pares de base, isto ocorreu devido ao uso de primers específicos para as sequências do promotor e terminador T7, as quais flanqueiam o gene de interesse inserido no plasmídeo pET-28a, aumentando assim o tamanho do produto amplificado.

Posteriormente extraiu-se o DNA plasmidial de 3 clones recombinantes e um destes foi submetido à uma clivagem com as mesmas enzimas utilizadas na clonagem (*Nde*-I e *Bam*H-I) para mais uma validação da inserção do gene *AmBgl1-LP* no vetor de expressão pET-28a, o resultado da clivagem está representado na Figura 7:

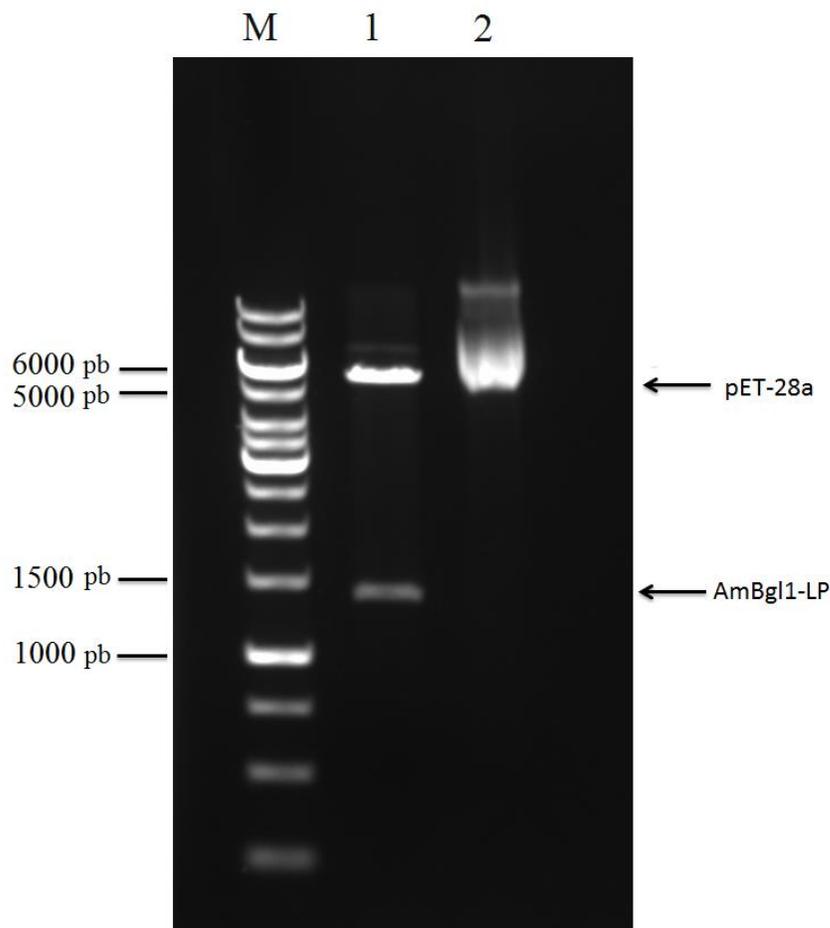


Figura 7 : Clivagem do vetor pET-28a contendo o gene *AmBgl1-LP*. M: marcador de peso molecular GeneRuler 1 Kb (Fermentas); 1: vetor recombinante clivado; 2: vetor recombinante não clivado.

É possível observar na coluna 1 da Figura 7 que a clivagem gerou bandas nas alturas de aproximadamente 6743 (fraca), 5369 e 1374 pb que são referentes aos fragmentos de vetor recombinante linearizado, plasmídeo pET-28a sem inserto e gene *AmBgl1-LP*, respectivamente. A banda fraca referente ao plasmídeo linearizado pode ter surgido porque a enzima *BamH*-I apresenta uma eficiência de reação de apenas 20 a 50% de sua atividade máxima no tampão utilizado na reação de restrição, Orange™ 1X (Thermo Scientific), significando que apenas a enzima *Nde*-I, que apresenta atividade de 100% neste tampão, conseguiu clivar completamente o vetor recombinante na reação.

Após a verificação, por sequenciamento, da correta inserção do gene de interesse em correta fase de leitura no plasmídeo pET-28a em um dos clones, transformou-se células

competentes de *E. coli* linhagem Rosetta (DE3) e iniciou-se a etapa de expressão da proteína recombinante. A indução foi realizada em um período total de 21 horas e alíquotas de 1 mL da cultura foram coletadas nos períodos de 1, 2, 3, 4 e 21 horas de indução para avaliação da produção da proteína recombinante, a qual foi avaliada por SDS-PAGE 12% como pode ser visto na Figura 8:

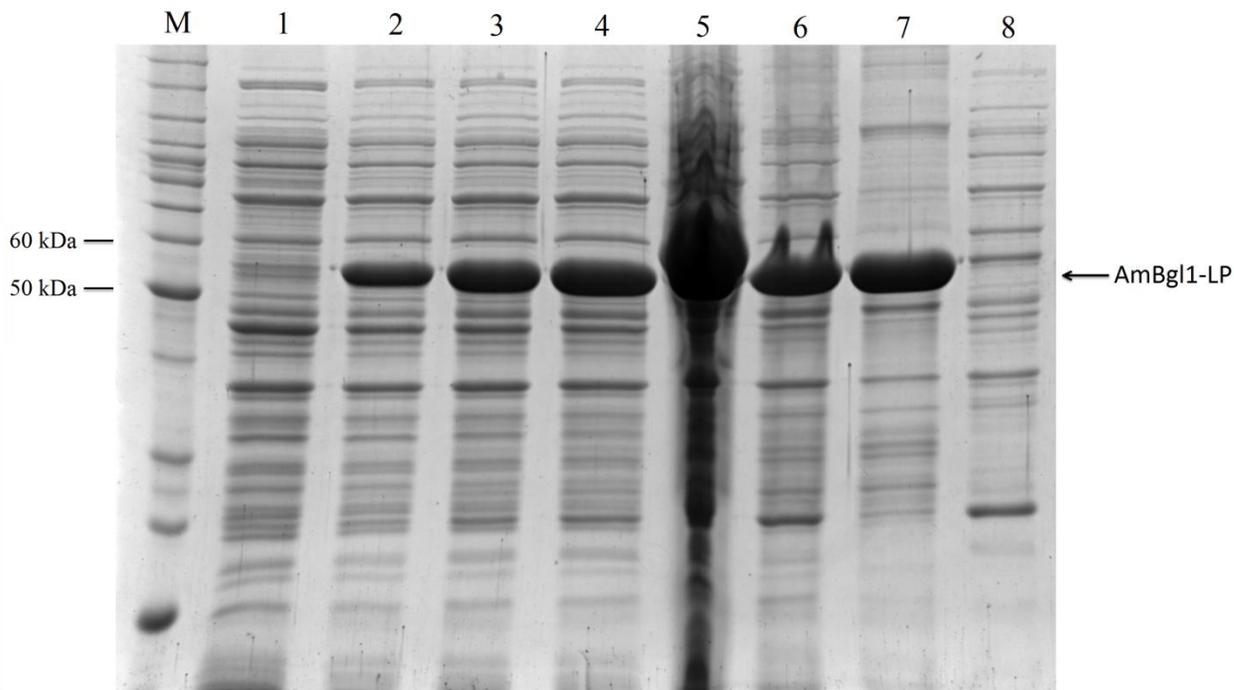


Figura 8 : Teste de expressão da AmBgl1-LP em *E. coli*. SDS-PAGE 12% a 37 °C mostrando M: marcador de peso molecular BenchMark (Invitrogen); 1: amostra não induzida; 2: amostra após 1 h de indução; 3: amostra após 2 h de indução; 4: amostra após 3 h de indução; 5: amostra após 4 h de indução; 6: amostra após 21 h de indução; 7: amostra da fração de proteínas insolúveis; 8: amostra da fração de proteínas solúveis.

Como pode ser visto na figura 8 a proteína recombinante foi expressa com sucesso no tamanho esperado de aproximadamente 53,7 kDa. A expressão se mostrou eficiente com rendimento alto a partir de 3 horas de indução, não sendo necessário continuar com a mesma até um período de 21 horas. Um problema identificado nesta indução foi que a maioria da AmBgl1-LP se apresentou na fração de proteínas insolúveis, fato que dificulta os passos posteriores de purificação da proteína, entretanto, decidiu-se realizar a purificação por afinidade molecular da fração solúvel de proteínas para avaliar se era possível utilizar apenas esta fração para isolar a proteína recombinante, os resultados da purificação podem ser vistos na Figura 9, a seguir:

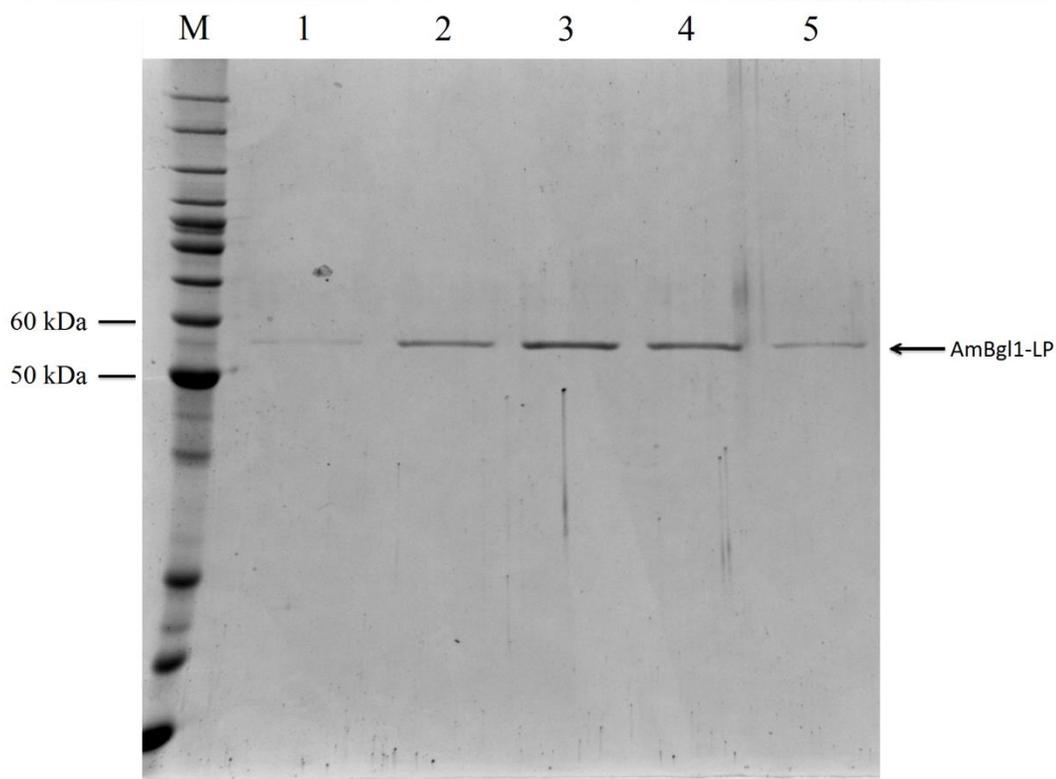


Figura 9 : Purificação da AmBgl1-LP por cromatografia de afinidade molecular ao níquel. SDS-PAGE 12% mostrando em M: marcador de peso molecular BenchMark (Invitrogen); 1: proteína eluída em 75 mM de imidazol; 2 e 3: proteína eluída em 100 mM de imidazol; 4 e 5: proteína eluída em 250 mM de imidazol.

Analisando a Figura 9 é possível concluir que a fração de proteínas solúveis continha a AmBgl1-LP, mesmo que em pequenas quantidades, e a mesma foi purificada com sucesso sem a presença de contaminantes nas eluições contendo 75 a 250 mM de imidazol em tampão de lise.

Com o intuito de solucionar o problema de baixa solubilidade da proteína recombinante foi realizada uma nova indução a 20 °C. Em menores temperaturas há uma redução na taxa de síntese de proteínas o que proporciona uma menor formação de corpos de inclusão e aumento da solubilidade e a qualidade conformacional das mesmas (Schein & Noteborn, 1988; Vera et al., 2007). Alíquotas de 1 mL de cultura foram retiradas nas mesmas condições descritas na indução anterior, o resultado deste novo experimento pode ser visualizado na Figura 10, a seguir:

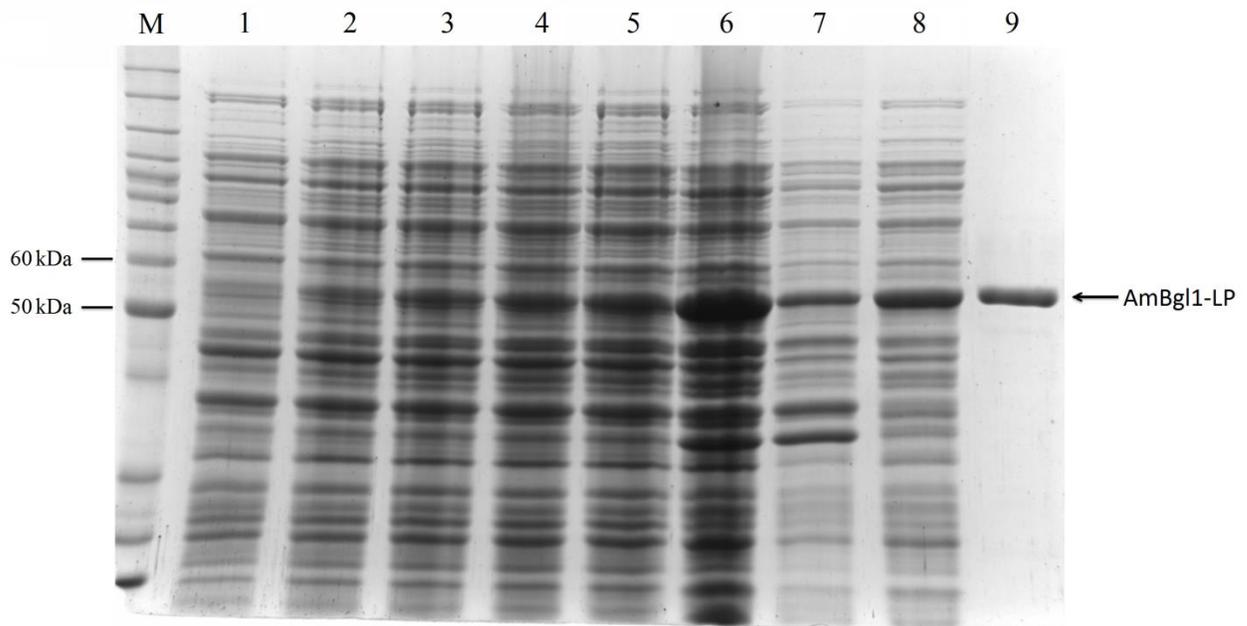


Figura 10 : Indução da AmBgl1-LP em *E. coli* a 20°C. SDS-PAGE 12% mostrando em M: marcador de peso molecular BenchMark; 1: amostra não induzida; 2: amostra após 1 h de indução; 3: amostra após 2 h de indução; 4: amostra após 3 h de indução; 5: amostra após 4 h de indução; 6: amostra após 21 h de indução; 7: fração de proteínas insolúveis; 8: fração de proteínas solúveis; 9: AmBgl1-LP purificada.

A Figura 10 mostra que a indução em 20°C produziu uma quantidade menor de proteína recombinante se comparado com a indução em 37°C, entretanto, a quantidade de proteína na fração solúvel aumentou consideravelmente, o que compensa a redução na produção e propicia maiores níveis de AmBgl1-LP para ser purificada por cromatografia de afinidade. O rendimento final de proteína purificada foi de aproximadamente 5 mg/L na indução em 37°C e aproximadamente 20 mg/L a 20°C quando quantificada pelo método do BCA.

Ensaio iniciais com a enzima eluída em 100-250 mM de imidazol revelaram que a mesma estava ativa, por meio de uma análise qualitativa na qual foi observado que a reação de hidrólise enzimática apresentou coloração amarelada após seu término com adição de tampão glicina NaOH 0,4 M pH 10,8, se comparada com um controle de reação sem coloração no qual a enzima recombinante não foi adicionada. Entretanto, após a diálise realizada em tampão PBS, para retirada do imidazol, foi observado que a enzima perdeu sua atividade embora ela tenha sido quantificada pelo teste de BCA e detectada por SDS-PAGE 12%. Diante deste cenário foram preparados outros tampões de diálise considerando a hipótese de que o contato da enzima recombinante com o tampão PBS a inativou de alguma maneira. Para isto foram realizadas novos

experimentos de diálise nos quais 25 μ L de enzima eram colocados sobre uma membrana de nitrocelulose por 30 minutos a temperatura ambiente sobre uma solução dentre as seguintes opções de tampões: fosfato de sódio 50 mM pH 6,5, fosfato de sódio 50 mM pH 8,0, Tris HCl 20 mM pH 7,5 e tampão de lise pH 8,0 (NaCl 100 mM, NaH_2PO_4 50 mM, Tris 10 mM). Após concluída a diálise nestas circunstâncias foram realizados novos ensaios enzimáticos, utilizando o mesmo volume de 25 μ L de enzima proveniente das membranas de nitrocelulose em tampões diferentes, tendo a absorbância resultante de cada tratamento lida em espectrofotômetro PerkinElmer Victor3™ com comprimento de onda de 420 nm (dados não mostrados). O melhor tampão para a diálise foi o tampão de lise pH 8,0 o qual apresentou maiores valores de absorbância se comparado com o controle não dialisado (dados não mostrados), assim passamos a utilizá-lo para fazer a diálise da AmBgl1-LP.

4.3 Ensaios de atividade e caracterização enzimática

Os ensaios de pH ótimo revelaram que a AmBgl1-LP tem maior atividade na faixa de pH entre 4,5 e 6,0, tendo atividade máxima em pH 5,5. A figura 11 abaixo mostra o comportamento da enzima em diferentes pHs entre 3 e 9:

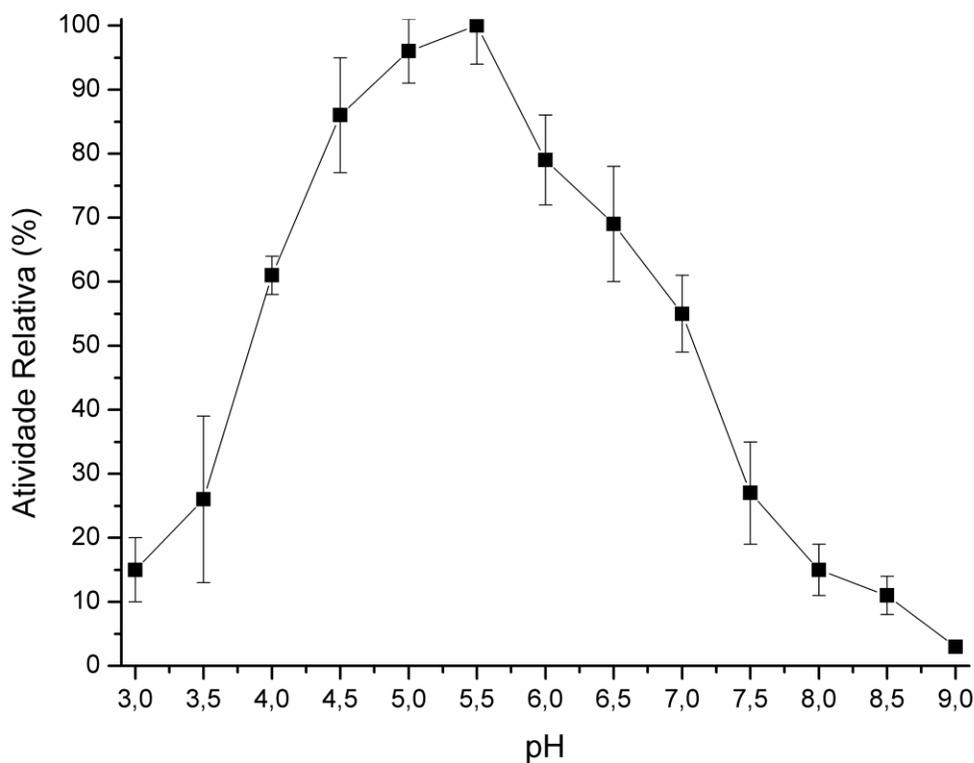


Figura 11: Curva de pH para a AmBgl1-LP. O gráfico mostra a liberação de produto p-nitrofenolato, em diferentes reações conduzidas no mesmo tampão com pHs diferentes, através da mensuração da absorbância resultante deste composto em espectrofotômetro com $\lambda = 420\text{nm}$.

A faixa de pH ótimo entre 4,5 e 5,5 é comum para as β -glicosidases e representa uma característica favorável ao explorar seu uso na produção de biocombustíveis a partir da degradação de celulose pelo processo de fermentação e sacarificação simultânea (SSF), já que as leveduras utilizadas neste processo se desenvolvem em pHs baixos na faixa de 4,4 a 5,3 (Kádár et al., 2003).

Os ensaios de temperatura mostraram que a AmBgl1-LP tem sua atividade máxima na temperatura de 35 °C, entretanto os dados de termoestabilidade mostram que é mais adequado trabalhar em temperaturas mais baixas como 30 °C, como pode ser visto na figura 12:

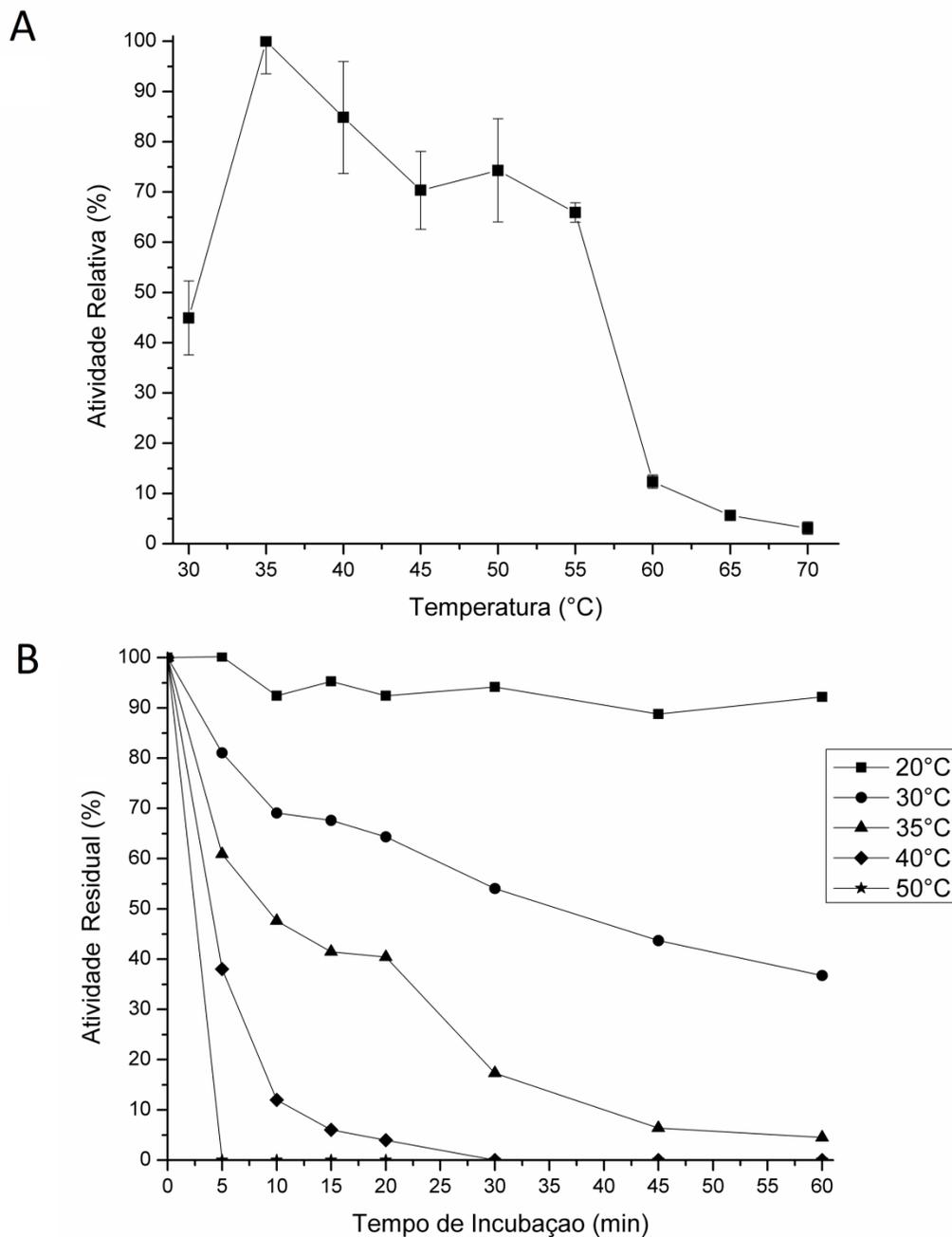


Figura 12: Curva de temperatura e termoestabilidade para AmBgl1-LP. A) Os experimentos de temperatura foram realizados utilizando o tampão Davies pH 5,5 na faixa de temperatura entre 30 e 70 °C com um máximo de atividade observado em 35 °C. Em B) foi avaliada a termoestabilidade da AmBgl1-LP, experimento o qual mostra que a enzima é relativamente estável em temperaturas menores que 20 °C para o tempo analisado. Na temperatura de 50 °C a enzima é inativada rapidamente e não possui atividade após 5 minutos de incubação.

Em 35 °C observa-se uma maior atividade da enzima, entretanto nesta temperatura a velocidade de inativação térmica é grande, portanto seria mais interessante utilizar a AmBgl1-LP em temperaturas menores, como 30 °C por exemplo. Esta temperatura é interessante para a aplicação desta β -glicosidase no processo de fermentação e sacarificação simultânea (SSF), já que as leveduras na dorna de fermentação desempenham melhor sua atividade em torno de 30 a 35 °C (Kádár et al., 2003).

Os ensaios de cinética enzimática para o substrato pNP β Gl_i e celobiose apresentaram o seguinte comportamento como visto na figura 13:

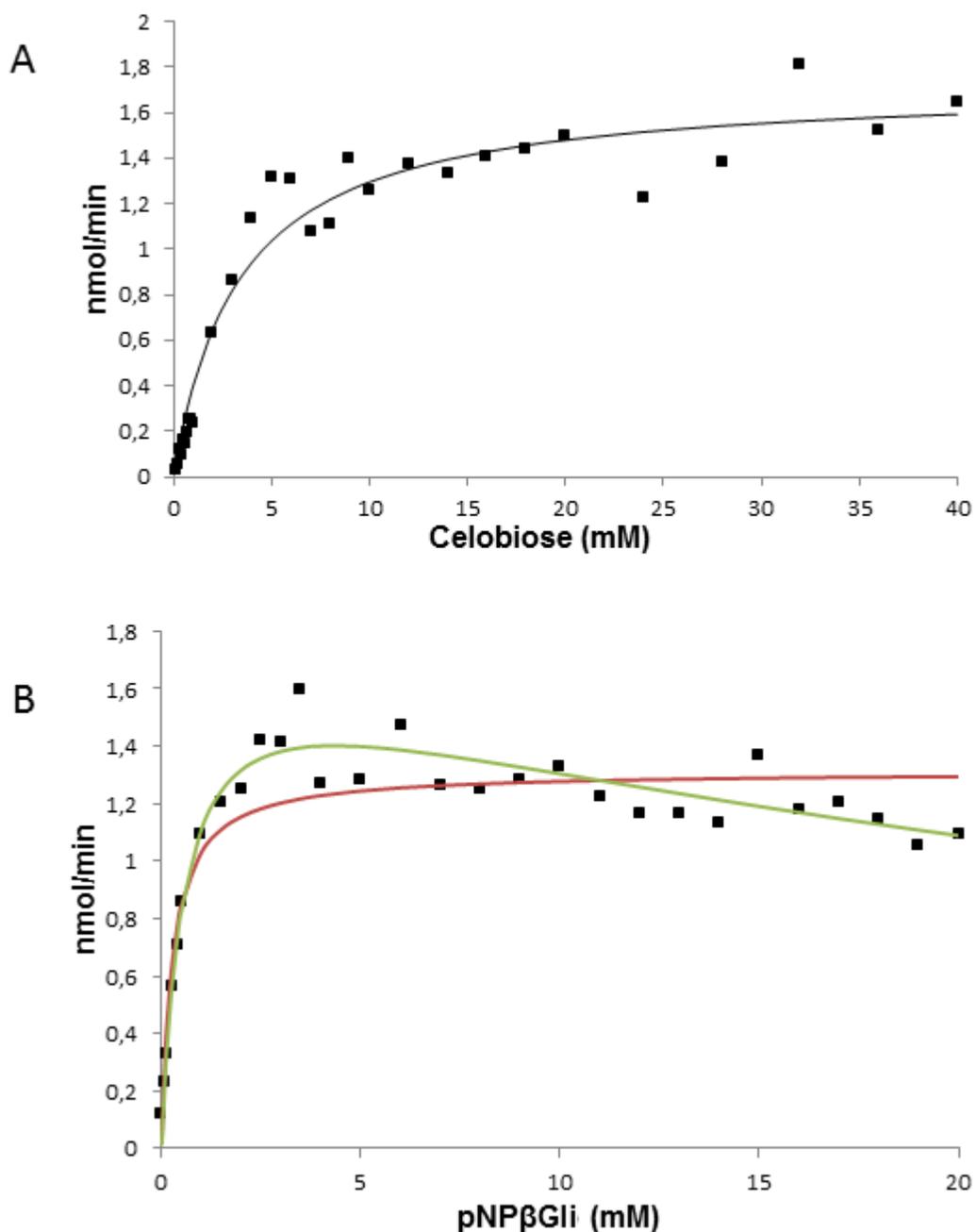


Figura 13: A) Curvas de $V_0/[S]$ para a AmBgl1-LP na presença do substrato celobiose. A cinética enzimática para este substrato segue o padrão de Michaelis-Menten já que a velocidade de reação aumenta conforme aumenta-se a concentração inicial de substrato até que a curva atinja um *plateau* máximo. B) Curvas de $V_0/[S]$ para a AmBgl1-LP na presença do substrato pNPβGli. A presença de duas curvas teóricas no gráfico mostra uma possível inibição da enzima pelo substrato (curva verde), que teve um ajuste melhor dos dados, além da curva de hidrólise que segue o modelo de Michaelis-Menten (vermelha).

Kcat e Ki para o substrato pNPβGli, calculados a partir do ajuste realizado pelo software Enzfitter, os quais podem ser vistos na tabela 2 abaixo:

	Michaelis-Menten	Inibição pelo substrato
Km	2,73±0,05 mM	0,55±0,08 mM
Vmax	1,31 ± 0,036 nmol/min	1,76±0,09 nmol/min
Kcat	0,00109 min ⁻¹	0,00109 min ⁻¹
Ki	-	34±6 mM

Tabela 2: Dados cinéticos sobre a AmBgl1-LP para o substrato pNPβGli tanto para o modelo de Michaelis-Menten quanto para o modelo de inibição pelo substrato.

Embora a velocidade de reação tenha diminuído com o aumento da concentração de substrato, considerando a família GH1 das β-glicosidases, outra hipótese para este evento seria de que a enzima possa realizar transglicosilação (Hancock, 2006). Neste processo ao invés de seguir o caminho de hidrólise as glicosil-hidrolases transferem um resíduo glicosídico de um doador para um aceptor, formando uma cadeia glicosídica maior que a inicial (Mangas-Sánchez, Adlercreutz, 2015). Para testar esta hipótese foram quantificados os dois produtos gerados na hidrólise de pNPβGli, o p-nitrofenolato e glicose, como mostrado na figura abaixo:

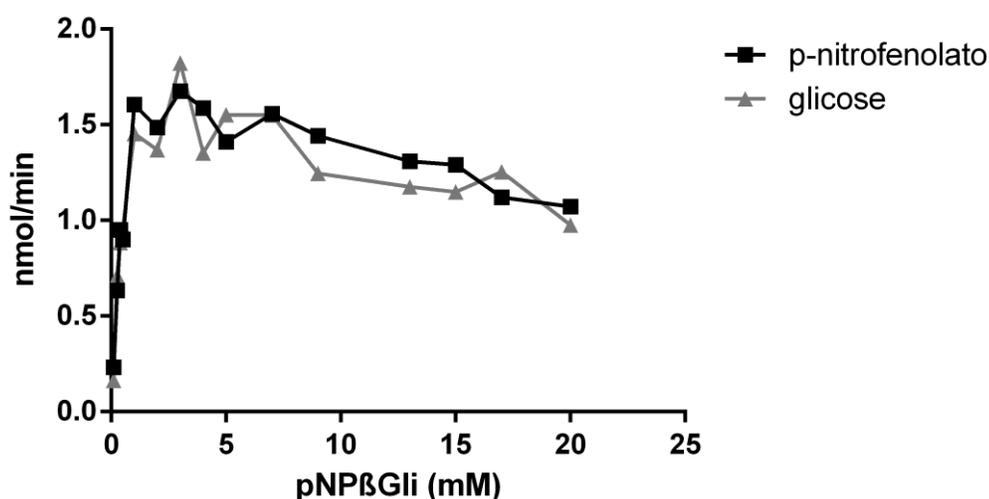


Figura 15: Gráfico de velocidades iniciais de formação de p-nitrofenolato e glicose utilizando o substrato pNPβGli. A distância entre os dois grupos de dados permite avaliar a ocorrência de transglicosilação para a AmBgl1-LP.

Como pode ser visto na figura 15 a velocidade de produção de glicose e p-nitrofenolato é praticamente a mesma ao longo de toda a faixa de concentração do substrato pNP β Gli. Nas reações de transglicosilação a velocidade da formação de glicose deve ser menor que a de formação de p-nitrofenolato, já que a glicose proveniente de hidrólise é transferida para o complexo ES (enzima-pNP β Gli) formando o complexo ESP (enzima-pNP β Gli-glicose) (Frutuoso & Marana, 2013). Para uma análise mais apurada dos dados, foi realizado um teste estatístico não paramétrico de Kolmogorov-Smirnov com intervalo de confiança de 95% no software GraphPad Prism v 6.00. O teste resultou em um P-valor de 0.9251. Assumindo $P < 0.05$ para a distância entre os grupos de dados ser significativa o resultado sugere que não ocorre transglicosilação até a concentração de substrato analisada. Este resultado é importante pois dá mais suporte à hipótese de que a diminuição da velocidade de reação observada em altas concentrações de substrato seja realmente ocasionada por uma inibição intrínseca do substrato sobre a enzima como proposto na figura 14. Pode-se supor isto porque caso a AmBgl1-LP estivesse realizando transglicosilação a enzima estaria deixando de realizar hidrólise intermitentemente, o que ocasionaria em uma diminuição da velocidade de formação de p-nitrofenolato, como este evento não acontece, a diminuição de velocidade de reação observada na figura 13B realmente deve ser proveniente de uma inibição intrínseca pelo substrato. Vale ressaltar ainda que o método definitivo para avaliar a ocorrência de transglicosilação seria realizar uma cromatografia líquida de alto desempenho (HPLC) do produto de reação da AmBgl1-LP para avaliar o aparecimento de moléculas glicosídicas com mais de 2 resíduos em suas cadeias.

Após a determinação dos parâmetros cinéticos da AmBgl1-LP para o substrato celobiose foram realizados ensaios para determinar a cinética de inibição da enzima por glicose. Para isto foram medidas as velocidades iniciais de hidrólise de diferentes concentrações de substrato (pNP β Gli) na presença de concentrações fixas do inibidor. Este experimento permite determinar o padrão de inibição e realizar o cálculo de K_i .

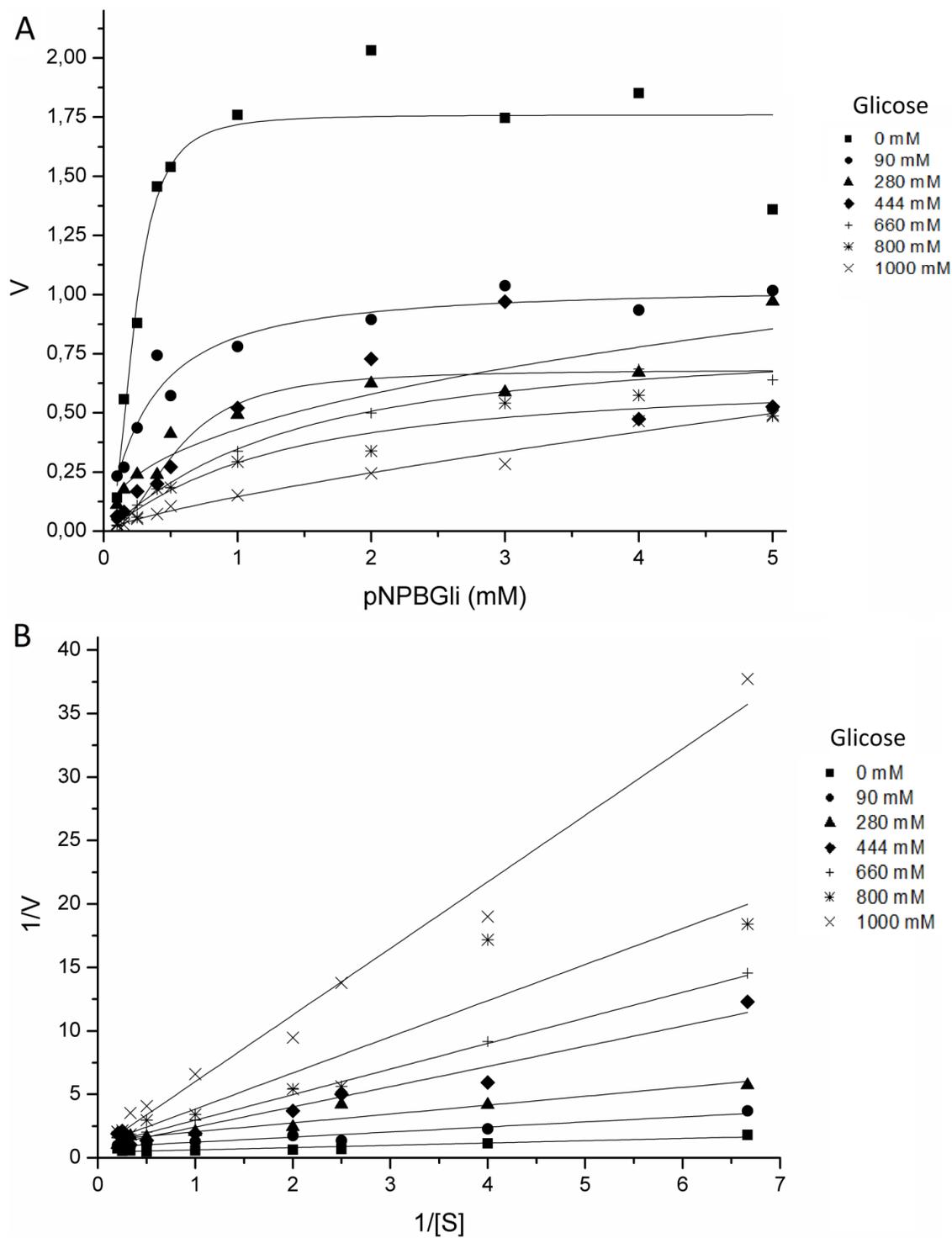


Figura 16: Ensaio de inibição para a AmBgl1-LP. A) Curvas de velocidades para a AmBgl1-LP utilizando o substrato pNPβGli em diferentes concentrações de glicose. B) Gráfico do tipo Lineweaver-Burk para a inibição por glicose, este gráfico é importante pois permite inferir o tipo de inibição ocorrida.

A figura 16 mostra que a glicose inibe de fato a AmBgl1-LP de forma que concentrações crescentes do inibidor diminuem as velocidades iniciais da reação de hidrólise de pNP β Gli com um K_i calculado de 23 mM. Do ponto de vista de aplicações industriais para a enzima ela não é adequada para o processo de hidrólise separada da fermentação (SHF), porém seu uso é adequado para o processo de fermentação e sacarificação simultânea (SSF) já que não há acúmulo de glicose no reator de fermentação neste processo.

Analisando a figura 16B é possível sugerir que este processo de inibição trata-se de uma inibição competitiva, já que as retas dos inversos de velocidade de Lineweaver-Burk tem valores próximos de intercepto no eixo Y (mínimo de 0,44 e máximo de 1,35 considerando a faixa de 0 a 1000 mM de inibidor) e tem inclinação crescente em função do aumento de [I]. A variação observada nos valores de intercepto provavelmente trata-se de uma variação experimental.

5 CONCLUSÕES

A busca por β -glicosidases em nosso banco local de sequências do metagenoma do Lago Poraquê revelou apenas uma sequência completa para β -glicosidases da família GH1, a qual foi chamada de AmBgl1-LP e foi utilizada neste trabalho para seguir com os experimentos de produção heteróloga e caracterização enzimática.

As análises teóricas de bioinformática da sequência da AmBgl1-LP como tamanho predito da proteína e sua classificação como uma β -glicosidase se mostraram eficientes já que estas informações foram confirmadas com os experimentos conduzidos ao longo deste trabalho.

A construção de um vetor de expressão heteróloga em bactéria, assim como a expressão da proteína recombinante e sua purificação se mostraram eficientes utilizando *E.coli*, sendo que o rendimento médio final de proteína após o aperfeiçoamento na temperatura de indução foi de aproximadamente 20 mg/L.

Os experimentos de caracterização enzimática revelaram que características como o pH ótimo da AmBgl1-LP se assemelham ao local no qual ela foi extraída, o Lago Poraquê, já que seu pH ótimo é próximo de 5,0. O mesmo não foi observado para os resultados de

temperatura ótima, entretanto os resultados de termoestabilidade sugerem o uso da AmBgl1-LP em temperaturas mais baixas do que 35 °C para que a mesma tenha uma vida útil mais elevada.

A AmBgl1-LP se mostrou eficiente ao hidrolisar substratos como pNPβGli e celobiose, sendo que este último é produto de degradação de biomassa vegetal, permitindo assim seu uso na indústria.

A AmBgl1-LP aparentemente não realiza transglicosilação. Um resultado interessante para o seu uso na indústria de degradação de biomassa.

A AmBgl1-LP não se mostrou tolerante à inibição por glicose, com um K_i baixo de 23 mM. Este resultado, entretanto, não descarta seu uso na indústria de degradação de biomassa, já que em processos como o SSF o acúmulo de glicose não ocorre ao final do processo de hidrólise da celulose.

6 PERSPECTIVAS

Em um estudo de Guo et al, uma mutação pontual em apenas dois aminoácidos da BglCel1A (L167W e P172L), também chamada de BglII de *Trichoderma reesei*, aumentou sua tolerância à inibição por glicose em 6x. Foi feito então um alinhamento de sequências da AmBgl1-LP, BglCel1A e HiBgl, a última vinda de *Humicola insolens* e que neste estudo serviu de modelo para a mutação na BglCel1A, para verificar a similaridade de aminoácidos na região onde foram substituídos os dois aminoácidos que aumentaram a tolerância à inibição por glicose na BglCel1A significativamente, o qual pode ser visto na figura 17:

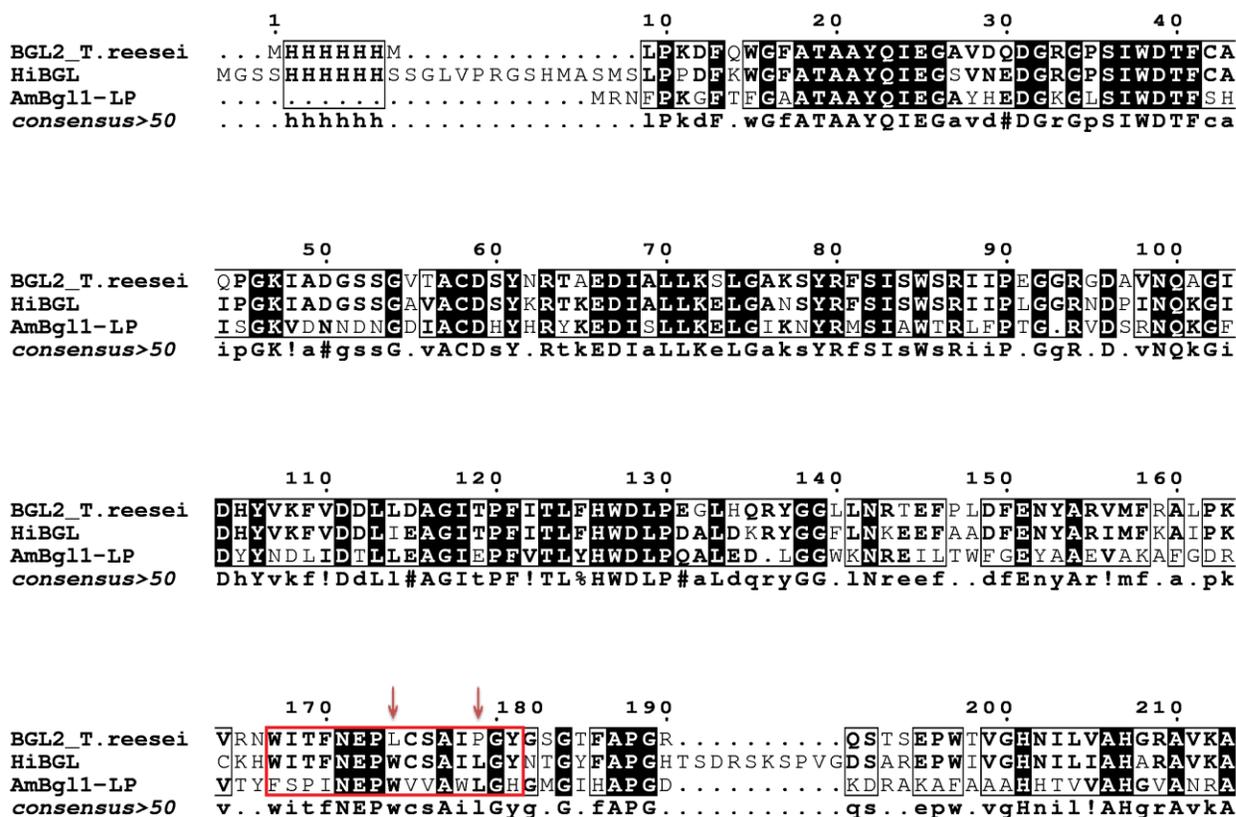


Figura 17: Alinhamento de seqüências entre HiBgl, BglCell1A e AmBgl1-LP. O retângulo vermelho representa toda a região a sofrer mutação proposta neste trabalho, as seta vermelhas indicam os dois aminoácidos que sofreram mutações na BglCell1A no trabalho de Guo et al e aumentaram a tolerância à inibição por glicose desta enzima em 6x.

Como pode ser visto na figura 17 os dois aminoácidos que conferiram resistência à inibição por glicose no trabalho de Guo et al permanecem idênticos na AmBgl1-LP se comparados com ao da HiBgl que é tolerante à inibição. Entretanto toda a região adjacente a esses resíduos é muito diferente na AmBgl1-LP, desta forma nós decidimos fazer uma mutação em toda a região do retângulo vermelho da figura da AmBgl1-LP para que a mesma fique idêntica à HiBgl. Esta construção está em andamento.

7 REFERÊNCIAS BIBLIOGRÁFICAS

CAZy – Carbohydrate Active Enzymes. Acesso on-line em 20/01/2017 pelo endereço: www.cazy.org

- Davies, M.T. A universal buffer solution for use in ultra-violet spectrophotometry. *Analyst*, v. 84, 248-251, 1959.
- Fang, Z.; Fang, W.; Liu, J.; Hong, Y.; Peng, H.; Zhang, X.; Sun, B.; Xiao, Y. Cloning and Characterization of a β -Glucosidase from Marine Microbial Metagenome with Excellent Glucose Tolerance. *J. Microbiol. Biotechnol.*;20(9): 1351~135, 2010.
- Finn, R.D.; Bateman, A.; Clements, J.; Coghill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; Sonnhammer, E.L.L.; Tate, J.; Punta, M. The Pfam protein families database. *Nucleic Acids Research, Database Issue 42:D222-D230*, 2014.
- Frutuoso, M. A.; Marana, S. R. A Single Amino Acid Residue Determines the Ratio of Hydrolysis to Transglycosylation Catalyzed by β -Glucosidases. *Protein and Peptide Letters*, v. 20, p. 102-106, 2013.
- Ghai, R.; Rodriguez-Valera, F.; McMahon, K. D; Toyama, D.; Rinke, R. et al. Metagenomics of the water column in the pristine upper course of the Amazon river. *PloS one*. v. 6 (8) p. e23785, 2011.
- Guo, B.; Amano, Y.; Nozaki, K. Improvements in Glucose Sensitivity and Stability of *Trichoderma reesei* β -Glucosidase Using Site-Directed Mutagenesis. *PLoS ONE* 11(1): e0147301, 2016.
- Hancock, S. M.; Vaughan, M.D.; Withers, S.G. Engineering of glycosidases and glycosyltransferases *Curr. Opin. Chem. Biol.*, 10, 509-519, 2006.
- Handelsman, J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews: MMBR*, v. 68 (4) p. 669-85, 2004.
- Handelsman, J.; Rondon, M. R.; Brady, S. F.; Clardy, J.; Goodman, R. M. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & Biology*, v. 5 (10) p. R245-9, 1998.
- Henrissat, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. *The Biochemical Journal*, v. 280 (Pt 2) p. 309-16, 1991.

Hugenholtz, P.; Goebel, B. M.; Pace, N. R. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *Journal of Bacteriology*, v. 180 (18) p. 4765-74, 1998.

Kádár, Zs.; Szengyel Zs.; Réczey K. Simultaneous saccharification and fermentation (SSF) of industrial wastes for the production of ethanol. *Industrial Crops and Products* 20 103–110, 2004.

Leatherbarrow, R. J. Enzfitter: a non-linear regression data analysis program for the IBM PC and true compatibles; manuel. Biosoft, 1987.

Li, K.; Jianbing, J.; Witte, M. D.; Kallemeijn, W. W.; Donker-Koopman, W. E.; Boot, R. G.; Aerts, J. M. F. G.; Codée, J. D. C.; van der Marel, G. A.; Overkleeft, H. S. Exploring functional cyclophellitol analogues as human retaining beta-glucosidase inhibitors. *Org. Biomol. Chem.* vol: 12 (39) pp: 7786-7791, 2014.

Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnology* 8:91, 2008.

Lynd, L. R.; Weimer, P. J.; Zyl, W. H.; Pretorius, I. S. Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiology and Molecular Biology Reviews*, v. 66 n.3 p. 506–577, 2002.

Mangas-Sánchez, J.; Adlercreutz, P. Enzymatic preparation of oligosaccharides by transglycosylation: A comparative study of glucosidases. *Journal of Molecular Catalysis B: Enzymatic* 122 51–55, 2015.

McCarter, J. D.; Withers, S. G. Mechanisms of enzymatic glycoside hydrolysis. *Current Opinion in Structural Biology*, v. 4 (6) p. 885-92, 1994.

Pei, J.; Pang, Q.; Zhao, L.; Fan, S.; Shi, H. *Thermoanaerobacterium thermosaccharolyticum* β -glucosidase: a glucose-tolerant enzyme with high specific activity for cellobiose. *Biotechnology for Biofuels*, v. 5 (1) p. 31, 2012.

Pei, X.; Zhao, J.; Cai, P.; Sun, W.; Ren, J.; Wu, Q.; Zhang, S.; Tian, C. Heterologous expression of a GH3 b-glucosidase from *Neurospora crassa* in *Pichia pastoris* with high purity and its

application in the hydrolysis of soybean isoflavone glycosides. *Protein Expression and Purification* 119, 75e84, 2016.

Santos-Júnior, C.D.; Kishi, L.T.; Toyama, D.; Soares-Costa, A.; Oliveira, T.C.S.; De Miranda, F.P.; Henrique-Silva, F. Metagenome Sequencing of Prokaryotic Microbiota Collected from Rivers in the Upper Amazon Basin. *Genome Announcements*, v. 5, p. e01450-16, 2017.

Schein, C. H.; Noteborn, M. H. M. Formation of Soluble Recombinant Proteins in *Escherichia Coli* is Favored by Lower Growth Temperature. *Nature Publishing Company*, v. 6 (3) p. 291-294, 1988.

Segel, I.H. *Enzyme Kinetics. Behaviour and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*. John Wiley and Sons, NY, 1993.

Singhania, R. R.; Patel, A. K.; Sukumaran, R. K.; Larroche, C.; Pandey, A. Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production. *Bioresource Technology*, v. 127 p. 500-7, 2013.

Steele, H. L.; Streit, W. R. *Metagenomics: advances in ecology and biotechnology*. FEMS Microbiology Letters. v. 247 (2) p. 105-11, 2005.

Stein, J. L.; Marsh, T. L.; Wu, K. Y.; Shizuya, H.; DeLong, E. F. Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. *Journal of Bacteriology*, v. 178 (3) p. 591-9, 1996.

Tamaki, F.K.; Textor, L.C.; Polikarpov, I.; Marana, S.R. Sets of Covariant Residues Modulate the Activity and Thermal Stability of GH1 β -Glucosidases. *PLoS ONE* 9(5): e96627., 2014.

Toyama, D.; Kishi, L.T.; Santos-Júnior, C.D.; Soares-Costa, A.; De Oliveira, T.C.S.; De Miranda, F.P.; Henrique-Silva, F. Metagenomics Analysis of Microorganisms in Freshwater Lakes of the Amazon Basin. *Genome Announcements*, v. 4, p. e01440-16, 2016.

Vera, A.; González-Montalbán, N.; Arís, A.; Villaverde, A. The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnology and Bioengineering*, v. 96 (6) p. 1101-6, 2007.

Walker, J. M. The bicinchoninic acid (BCA) assay for protein quantitation. *Methods in Molecular Biology*, v. 32 p. 5-8, 1994.

Withers, S. G., Warren, R. A. J., Street, I. P., Rupitz, K., Kempton, J. B. & Aebersold, R. Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of a retaining glycosidase. *Journal of the American Chemical Society*, v. 112 (15) p. 5887 – 5889, 1990.

8 APÊNDICE

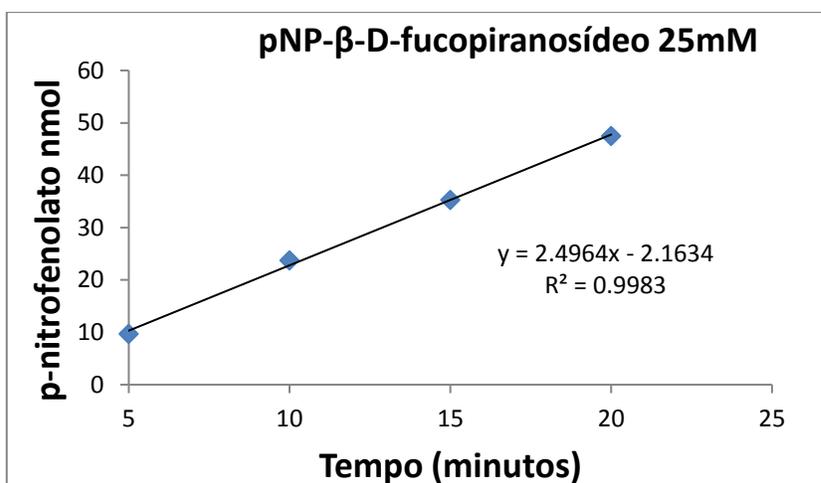
>AmBgl1-LP (sequência de DNA)

ATGCGTAATTTTCCGAAAGGTTTCACTTTTGGGGCAGCAACGGCTGCCTATCAAATTGAAGG
 CGCTTATCACGAGGATGGTAAGGGACTTTCCATTTGGGACACTTTTAGCCATATTTTCGGGCAA
 GGTAGACAACAACGATAACGGGGATATTGCCTGCGATCACTACCATCGATACAAAGAAGAT
 ATTTTCATTGCTGAAAGAATTAGGCATTAAGAATTATCGAATGAGCATCGCCTGGACTCGGCT
 CTTTCCTACTGGTCGGGTAGATTCGCGAAATCAAAAAGGCTTTGACTACTACAACGATCTCAT
 TGACACTCTACTTGAGGCCGGAATAGAACCCTTCGTTACTTTATATCACTGGGATTTACCACA
 AGCCCTTGAAGATTTAGGCGGCTGGAAAAATCGGGAAATTCTTACCTGGTTTGGTGAATACG
 CAGCTGAAGTGGCAAAGGCATTTGGTGATCGAGTCACGTATTTCTCCCCTATTAATGAGCCCT
 GGGTTGTGCCTGGCTAGGACATGGAATGGGAATTCATGCACCAGGAGACAAAGATCGTGC
 GAAAGCTTTTGCTGCAGCGCATCACACCGTTGTTGCGCATGGCGTCGCGAATCGTGCAATGA
 AAGCTGTCTACCCACATCTGAAAATTGGTCCAGTTCTTAACCAAACAAATTTTCCTATTGATG
 ATTCAACTAATGCAGAGTTGAAGCGTGCTCAGGATATTTTGGACGCGAATCAAAAATCGATTT
 TGGATTGATGCAATCTTTCACAAGAGATATCCAGAGATCCTAATTGAAAATTTTCCTGAAGAT
 ATTCTTCCAAATATTCATGAGGGGGATTTGATTTAGCAGCTACGCCAAATGATTGGCTTGGA
 ATCAATTACTATTTTGATACACCTATGAAAGCGAGCTCTCGACCTGTAGCAACGGCCTTCGAT
 CCGGCGGCGCTTCTAGGATTGAATGTTGACCAGACGCCTCCACCACCTCTTACTGATATGGGT
 TGGCCAATTTCTCCAGATGGGTAAAGTAATTTACTTATTCGTTGGCATAAAGAGATTGGTAAC
 CGTTTACCGCCGATCTACATCACTGAAAATGGAGTGGCATAACGATGACGGTCTGGCGAAGA
 TGGACTTATTCACGACCAGCGCCGAATCGATTATTTGCGAGGGGCATCTTGCAGCCCTCCTTGA
 TGCTTGTGAAGTTGGAGTAGATGTCAGGGGATATTTTCAGTGGTCCCTTCTTGATAACTTTGA
 ATGGGCTCTGGGCTATTCCAAGCGTTTGGCATCGTTCACGTTGACTACCAAACCCTTAAGCG
 CACCCCAAAGGATTCGGCGCTCTGGTACAGCGGAGTAATTGCGCGTAACGCTCTAGATTGA

>AmBgl1-LP (sequência de aminoácidos com His-Tag do vetor pET-28a)

MGSSHHHHHSSGLVPRGSHMRNFPKGFTFGAATAAYQIEGAYHEDGKGLSIWDTFSHISGKVD
 NNDNGDIACDHYHRYKEDISLLKELGIKNYRMSIAWTRLFPTGRVDSRNQKGFDDYNDLIDTLL
 AGIEPFVTLYHWDLPQALEDLGGWKNREILTWFGGEYAAEVAKAFGDRVTYFSPINEPWVVAWL
 GHGMGIHAPGDKDRAKAFAAAHTTVVAHGVANRAMKAVYPHLKIGPVLNQTNFPIDSTNAEL
 KRAQDILDANQNRFWIDAIFHKRYPEILNFPEDILPNIHEGDFDLAATPNDWLGINYFDTPMK
 ASSRPVATAFDPAALLGLNVDQTPPPPLTDMGWPISPDGLSNLLIRWHKEIGNRLPPIYITENGVA
 YDDGPGEDGLIHDQRRIDYLRGHLAALLDACEVGVVDVRGYFQWSLLDNFEWALGYSKRFGIVH
 VDYQTLKRTPKDSALWYSGVIARNALD

Ensaio de 4 tempos de hidrólise de pNP-β-D-fucopiranosídeo:



Ensaio de 4 tempos de hidrólise de pNP-β-D-galactopiranosídeo

