

UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Biológicas e da Saúde
Departamento de Genética e Evolução
Laboratório de Bioquímica e Genética Aplicada

OLINDA SOARES ATHAIDE ALCOBAÇA

Expressão recombinante de endo-beta-1,4 xilanase identificada no meta-transcriptoma de cupins da espécie *Heterotermes tenuis* para produção de bioetanol

Orientador: Prof. Dr. Anderson Ferreira da Cunha

SÃO CARLOS - SP
2022

OLINDA SOARES ATHAIDE ALCOBAÇA

Expressão recombinante de endo-beta-1,4 xilanase identificada no meta-transcriptoma de cupins da espécie *Heterotermes tenuis* para produção de bioetanol

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, para obtenção do título de mestre em Genética Evolutiva e Biologia Molecular.

Orientador: Prof. Dr. Anderson Ferreira da Cunha

SÃO CARLOS - SP
2022

Alcobaça, Olinda Soares Athaide

Expressão recombinante de endo-beta-1,4 xilanase identificada no meta-transcriptoma de cupins da espécie *Heterotermes tenuis* para produção de bioetanol / Olinda Soares Athaide Alcobaça -- 2022.
69f.

Dissertação (Mestrado) - Universidade Federal de São Carlos, campus São Carlos, São Carlos
Orientador (a): Anderson Ferreira da Cunha
Banca Examinadora: Prof. Dr. Anderson Ferreira da Cunha, Profa. Dra. Andrea Soares da Costa Fuentes, Prof. Dr. Gleidson Silva Teixeira
Bibliografia

1. Enzima Recombinante. 2. Bioetanol. 3. *Heterotermes tenuis*. I. Alcobaça, Olinda Soares Athaide. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática (SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Ronildo Santos Prado - CRB/8 7325



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

Defesa de Dissertação de Mestrado do candidato Olinda Soares Athaide Alcobaça, realizada em 22/02/2022.

Comissão Julgadora:


Prof. Dr. Anderson Ferreira da Cunha (UFSCar)


p/ Profa. Dra. Andrea Soares da Costa Fuentes (UFSCar)


p/ Prof. Dr. Gleidson Silva Teixeira (UNICAMP)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular.

Dedico este trabalho a Deus, por conduzir os meus caminhos até aqui. À minha família, pelas orações e apoio incondicional.

AGRADECIMENTOS

Agradeço a Deus, por sua infinita graça e amor, por todo cuidado e proteção durante as adversidades, pela sabedoria concedida e por ter me sustentado até aqui.

Aos meus pais, Jakyson e Ivone, pelo amor incondicional, incentivo e compreensão durante todo esse tempo que estive longe.

Ao meu companheiro, Edesio, por toda paciência e parceria, pelo suporte emocional e por continuar me encorajando a correr atrás dos meus maiores sonhos.

Ao meu sogro José e a minha cunhada Mayla, por sempre estarem ao meu lado, por todo afeto e carinho.

Ao professor Dr. Anderson Ferreira da Cunha, pela orientação e ensinamentos compartilhados. Obrigado pela confiança e por me incluir ao grupo de pesquisa. Tem sido enriquecedor!

À professora Dra. Andrea Soares da Costa Fuentes, por todos os conselhos e instruções, além de permitir meu acesso ao Laboratório de Biologia Vegetal, local que me proporcionou um profundo aprendizado nos últimos meses.

À Prof. Dra. Ana Maria Costa-Leonardo, por ter gentilmente cedido os cupins que foram utilizados neste trabalho e por todo carinho durante a coleta.

À Emeline, ao professor Dr. Caio Freire e a equipe do laboratório de Biologia Computacional, pela ajuda nas análises do meta-transcriptoma e das sequências.

À Iara, Sâmara, Isabela e Karina, por todo conhecimento compartilhado, pela cumplicidade e companhia durante os almoços e por todo auxílio durante os experimentos. Sou muito grata pela amizade de vocês!

Aos colegas do Laboratório de Bioquímica e Genética Aplicada, pelo companheirismo e suporte, especialmente ao Mateus, Hosana, Jéssica, Emeline, Grazi e o João que sempre estiveram por perto, dispostos a ajudar.

Aos professores e colegas do Departamento de Genética e Evolução (DGE).

Ao apoio financeiro e bolsa concedida pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Aos meus familiares e amigos que de alguma forma ajudaram na realização desse trabalho. Muito obrigada!

*“... seja forte e corajoso! Não se apavore, nem se desanime,
pois o Senhor, o seu Deus, estará com você por onde você andar”.*

Josué 1:9

RESUMO

A procura por fontes energéticas alternativas ao consumo do petróleo e a preocupação com o meio ambiente tem despertado para o uso de novas matérias-primas na fabricação de biocombustíveis. Alinhado a isso tem-se o etanol de segunda geração (E2G), que pode ser produzido a partir da biomassa lignocelulolítica, oriunda principalmente de resíduos agrícolas. Para que ser convertido em etanol esse material precisa passar por hidrólise de cada um dos componentes e essa é uma das etapas mais desafiadoras, pois é nela que ocorre a conversão do material lignocelulósico em açúcares fermentescíveis. Apesar de ser eficiente, a utilização de complexos enzimáticos encarece todo o processo de produção. Desta forma, a bioprospecção de novas enzimas lignocelulolíticas é importante pois traria alternativas para o processo de produção de bioetanol, aumentando sua viabilidade logística e econômica. Xilanases são hemicelulases que degradam a xilana, principal componente da hemicelulose, formado principalmente por pentoses. Essa enzima pode ser aplicada para diversos fins industriais, tais como: branqueamento de papel, clarificação de bebidas e na produção de bioetanol. Nesse último caso, as xilanases otimizam a ação do complexo celulolítico e consequentemente o rendimento da produção de bioetanol. Neste estudo, prospectamos uma endo-beta-1,4-xilanase (HtpXyl) do meta-transcriptoma de cupins da espécie *Heterotermes tenuis*, espécie de cupim que é considerada praga de cana-de-açúcar. A caracterização estrutural indicou a presença dos dois sítios ativos necessários para a atividade catalítica. Desta forma, a HtpXyl foi clonada no vetor pET28a e no pPICZαA, para expressão em *Escherichia coli* e *Pichia pastoris*, respectivamente. A proteína recombinante apresentou massa molecular de aproximadamente 20 kDa, conforme predito e em ambos os sistemas apresentou atividade xilanolítica quando submetida ao teste em ágar xilana 1% corado com vermelho do Congo 0,2%. A HtpXyl foi capaz de degradar o substrato presente no ágar tanto a 30°C, quanto a 40°C. Estes resultados contribuem para estabelecer a HtpXyl como um candidato promissor para estudos bioquímicos mais aprofundados para posterior uso em processos industriais de degradação de xilana, dentre eles, a produção de bioetanol.

Palavras-chave: *Heterotermes tenuis*; xilanase; proteína recombinante; hemicelulase; bioetanol.

ABSTRACT

The search for alternative energy sources to the consumption of petroleum and the attention to the environment has awakened the use of new bases in biofuels. In line with this is second-generation ethanol (E2G), which can be produced from lignocellulosic biomass, mainly derived from agricultural residues. For this material to be converted, it needs to undergo hydrolysis of each component. This is one of the stages in which they are most studying, as it is in this stage, the lignocellulosic material is converted into fermentable sugars. Despite being efficient, the use of enzymatic complexes faces the entire production process. In this way, bioprospecting for new lignocellulolytic enzymes is essential because it would bring alternatives for the bioethanol production process, increasing its logistical and economic viability. Xylanases are hemicellulases that degrade xylan, the main component of hemicellulose, formed mainly by pentoses. This enzyme can be applied to several industries, such as paper bleaching, beverage clarification, and bioethanol production. In the latter, xylanases optimize the action of the cellulolytic complex and consequently the yield of bioethanol production. In this study, we prospected an endo-beta-1,4-xylanase (HtpXyl) from the metatranscriptome of termites of the species *Heterotermes tenuis*, a termite species that is considered a pest of sugarcane. The characterization indicated the presence of two structural sites for catalytic activity. Thus, HtpXyl was cloned into the pET28a vector and pPICZαA for expression in *Escherichia coli* and *Pichia pastoris*, respectively. A recombinant protein showed approximately 20kDa, as both systems showed xylanolytic x while tested at 0% activity on 1% Congo red-stained agar, 2%. HtpXyl degraded the substrate on agar both at 30°C and 40°C. These results for collaborative studies to produce an Htp as a biochemical candidate more in-depth for industrial processes of degradation of xylan, among them, of bioethanol.

Keywords: *Heterotermes tenuis*; xylanase; recombinant protein; hemicellulase; bioethanol.

LISTA DE FIGURAS

Figura 1. Representação dos principais componentes da biomassa lignocelulósica	5
Figura 2. Estrutura química dos principais componentes da biomassa lignocelulósica	6
Figura 3. Estrutura química da xilana	7
Figura 4. Processo de produção de etanol 2G	9
Figura 5. Degradação enzimática da cadeia de celulose	12
Figura 6. Degradação enzimática da xilana	13
Figura 7. Ciclo de vida dos cupins	14
Figura 8. Degradação da lignocelulose feita por cupins inferiores e seus simbioses intestinais.....	16
Figura 9. Vetor pGEM_HtpXyl	19
Figura 10. Vetor pET28a_HtpXyl	20
Figura 11. Vetor pPICZαA_HtpXyl	20
Figura 12. Sequência de nucleotídeos e aminoácidos da β-1,4-endo-xilanase de <i>H. tenuis</i>	35
Figura 13. Resultado do alinhamento da xilanase obtido após análise BLASTn	35
Figura 14. Alinhamento da sequência CfXyn2-2 pela análise no BLAST contra a sequência da xilanase	36
Figura 15. Peptídeo Sinal predito da xilanase	37
Figura 16. Predição da estrutura secundária da xilanase	38
Figura 17. Resultado da avaliação da xilanase utilizando o software InterProScan	39
Figura 18. Predição da estrutura terciária da xilanase	39
Figura 19. Predição dos sítios de glicosilação e fosforilação da xilanase	40
Figura 20. Avaliação da integridade do RNA extraído por gel desnaturante	41
Figura 21. Amplificação do fragmento da β-actina	42
Figura 22. Amplificação do fragmento que codifica a xilanase	43
Figura 23. PCR das colônias transformantes (pGEM_HtpXyl)	43
Figura 24. PCR das colônias transformantes (pPICZαA_HtpXyl)	44
Figura 25. Comparação entre a sequência obtida pelo sequenciamento e a sequência da xilanase inicialmente escolhida do meta-transcriptoma	45
Figura 26. Resultados da sequência da xilanase obtida por sequenciamento após análise BLASTn	45

Figura 27. Alinhamento da sequência CfXyn3-1 pela análise no BLAST contra a sequência da xilanase que foi obtida no sequenciamento	46
Figura 28. Avaliação da HtpXyl utilizando o software InterProScan	47
Figura 29. Predição dos sítios de glicosilação e fosforilação da HtpXyl	48
Figura 30. Predição da estrutura terciária da HtpXyl	49
Figura 31. Comparação entre a xilanase inicialmente proposta e a HtpXyl	49
Figura 32. Linearização do vetor pPICZαA_HtpXyl	52
Figura 33. Expressão da HtpXyl em <i>P. pastoris</i>	53
Figura 34. Expressão da HtpXyl em <i>E. coli</i>	54
Figura 35. Teste de solubilidade da HtpXyl expressa em <i>E. coli</i>	55
Figura 36. Purificação da HtpXyl expressa em <i>E. coli</i> por cromatografia de afinidade	56
Figura 37. Avaliação da atividade xilanolítica da HtpXyl em ágar xilana	57

LISTA DE TABELAS

Tabela 1. Composição lignocelulósica de diferentes plantas e resíduos agroindustriais	5
Tabela 2. Produção de etanol 2G no mundo	8
Tabela 3. Sequências de <i>primers</i> utilizados	21
Tabela 4. <i>Screening</i> da indução da expressão da xilanase em meio com metanol	29
Tabela 5. Análise dos aminoácidos diferentes entre as duas xilanases	50

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

AOX	gene que codifica a álcool oxidase
BLAST	ferramenta de busca por alinhamento (<i>Basic Local Alignment Search Tool</i>)
BMGY	meio complexo contendo glicerol
BMMY	meio complexo contendo metanol
CaCl₂	cloreto de cálcio
CBH	celobiohidrolase
cDNA	DNA complementar
DEPC	dietilpirocarbonato
DNA	ácido desoxirribonucleico
dNTPs	desoxirribonucleotídeos trifosfato
D.O.	densidade óptica
DTT	ditiotreitól
E1G	etanol de primeira geração
E2G	etanol de segunda geração
E.C.	Classificação Enzimática (<i>Enzyme Classification</i>)
EDG	endoglucanase
EDTA	ácido etilenodiaminotetracético
GH	Glicosil-hidrolase
HCl	ácido clorídrico
HNO₃	ácido nítrico
HtpXyl	clone da endo-β-1,4-xilanase de protista simbiote de <i>Heterotermes tenuis</i>
H₂SO₄	ácido sulfúrico
IPTG	Isopropil β-D-1-tiogalactopiranosídeo
KCl	cloreto de potássio
kDa	QuiloDalton
KV	QuiloVolt
LB	Luria Bertani
LS	<i>Low Salt</i>
M	molar
MCS	Sítio de Múltipla Clonagem (<i>Multi Cloning Site</i>)
Mg	miligrama
MgCl₂	cloreto de magnésio
ML	megalitro

mL	microlitro
mM	Milimolar
NaCl	Cloreto de sódio
NaH₂PO₄	Fosfato monobásico de sódio
NCBI	<i>National Center for Biotechnology Information</i>
Ng	nanograma
Nm	nanômetro
ORF	Fase aberta de leitura (<i>Open Read Frame</i>)
PAGE	Eletroforese em gel de poliacrilamida (<i>Polyacrylamide Gel Electrophoresis</i>)
pb	Pares de base
PCR	Reação em cadeia da polimerase (<i>Polymerase Chain Reaction</i>)
pH	Potencial hidrogeniônico
Pmol	picomol
PO₄	grupo fosfato
PVDF	polivinilideno difluorido
RNA	ácido ribonucleico
Rpm	rotação por minuto
rRNA	RNA ribossômico
SDS	dodecil sulfato de sódio (<i>Sodium Dodecyl Sulfate</i>)
SHF	Hidrólise e Fermentação Separadas (<i>Separate Hydrolysis and Fermentation</i>)
SSF	Sacarificação e Fermentação Simultânea (<i>Simultaneous Saccharification and Fermentation</i>)
TAE	Tris-Acetato-EDTA
U	unidade enzimática
V	volts
X-gal	5-Bromo-4-cloro-indol-β-D-galactopiranosídeo
XIL	xilanase
YNB	Base de Nitrogênio de Levedura (<i>Yeast Nitrogen Base</i>)
βG	β-glicosidase
βX	β-xilosidase
μL	microlitro
μg	micrograma
μF	microfaraday
°C	graus Celsius
°GL	grau Lussac
Ω	ohm

SUMÁRIO

1.	INTRODUÇÃO	3
2.	FUNDAMENTAÇÃO TEÓRICA	4
2.1.	BIOMASSA LIGNOCELULÓSICA	4
2.1.1.	XILANA	7
2.2.	ETANOL A PARTIR DA BIOMASSA (E2G)	8
2.3.	ENZIMAS LIGNOCELULÓSICAS	11
2.3.1.	CELULASES	11
2.3.2.	XILANASES	12
2.4.	CUPINS	14
2.4.1.	<i>HETEROTERMES TENUIS</i>	16
3.	OBJETIVOS	17
3.1.	OBJETIVO GERAL	17
3.2.	OBJETIVOS ESPECÍFICOS	17
4.	MATERIAIS E MÉTODOS	17
4.1.	MATERIAIS	17
4.1.1.	CUPINS	17
4.1.2.	LINHAGENS DE <i>ESCHERICHIA COLI</i> UTILIZADAS	18
4.1.3.	LINHAGENS DE <i>PICHIA PASTORIS</i> UTILIZADAS	18
4.1.4.	PLASMÍDEOS	18
4.1.4.1.	PGEM-T EASY	18
4.1.4.2.	PET28A	19
4.1.4.3.	PPICZAA	20
4.1.4.4.	OLIGONUCLEOTÍDEOS INICIADORES	21
4.2.	MÉTODOS	21
4.2.1.	ANÁLISES <i>IN SILICO</i> DAS SEQUÊNCIAS	21
4.2.2.	EXTRAÇÃO E QUANTIFICAÇÃO DE RNA	22
4.2.3.	SÍNTESE DE CDNA	23
4.2.4.	REAÇÃO EM CADEIA DA POLIMERASE (PCR) E PURIFICAÇÃO DO FRAGMENTO	23
4.2.5.	PREPARO DE CÉLULAS QUIMIOCOMPETENTES DE <i>E. COLI</i> DH5-A	24
4.2.6.	REAÇÃO DE LIGAÇÃO DO INSERTO AO VETOR PGEM [®] -T EASY	24
4.2.7.	TRANSFORMAÇÃO EM <i>E. COLI</i> DH5-A	24
4.2.8.	SEQUENCIAMENTO	25
4.3.	EXPRESSÃO RECOMBINANTE DA XILANASE EM <i>P. PASTORIS</i>	26
4.3.1.	SUBCLONAGEM DO FRAGMENTO QUE CODIFICA A XILANASE NO VETOR PPICZAA	26

4.3.2.	PREPARO DE CÉLULAS ELETROCOMPETENTES DE <i>P. PASTORIS</i> X-33	27
4.3.3.	TRANSFORMAÇÃO EM <i>P. PASTORIS</i> X-33	27
4.3.4.	EXPRESSÃO EM PEQUENA ESCALA DA XILANASE EM <i>P. PASTORIS</i> X-33	28
4.3.5.	EXPRESSÃO DA XILANASE EM <i>P. PASTORIS</i> X-33 EM MAIOR ESCALA	29
4.3.6.	PURIFICAÇÃO POR CROMATOGRAFIA DE AFINIDADE	30
4.4.	EXPRESSÃO RECOMBINANTE DA XILANASE EM <i>E. COLI</i>	30
4.4.1.	SUBLONAGEM DO FRAGMENTO QUE CODIFICA A XILANASE NO VETOR PET28A	30
4.4.3.	EXPRESSÃO RECOMBINANTE DA XILANASE EM <i>E. COLI</i> ROSETTA (DE3)	32
4.4.4.	TESTE DE SOLUBILIDADE	33
4.4.5.	PURIFICAÇÃO DA XILANASE POR CROMATOGRAFIA DE AFINIDADE	33
4.4.6.	TESTE DE ATIVIDADE XILANOLÍTICA EM ÁGAR XILANA	34
5.	RESULTADOS E DISCUSSÃO	34
5.1.	ANÁLISES <i>IN SILICO</i> DA SEQUÊNCIA	34
5.2.	EXTRAÇÃO DE RNA E SÍNTESE DE CDNA	41
5.3.	AMPLIFICAÇÃO DO FRAGMENTO QUE CODIFICA A XILANASE	42
5.4.	CLONAGEM DO FRAGMENTO CODIFICANTE DA XILANASE NO VETOR PGEM-T <i>EASY</i>	43
5.5.	SUBCLONAGEM DO FRAGMENTO QUE CODIFICA A XILANASE NO VETOR PPICZAA	44
5.6.	ANÁLISE DO SEQUENCIAMENTO	44
5.7.	EXPRESSÃO RECOMBINANTE DE HTPXYL EM <i>P. PASTORIS</i> X-33	51
5.8.	EXPRESSÃO RECOMBINANTE DE HTPXYL EM <i>E. COLI</i> ROSETTA (DE3)	54
5.9.	TESTE DE ATIVIDADE XILANOLÍTICA EM ÁGAR XILANA	56
6.	CONCLUSÃO	59
7.	PERSPECTIVAS	60
8.	ANEXOS	60
9.	REFERÊNCIAS BIBLIOGRÁFICAS	64

1. INTRODUÇÃO

O crescimento da conscientização ambiental aliada às preocupações com a elevação dos preços do petróleo, despertaram o interesse por fontes renováveis de energia. Dentro desse contexto, surgiu o cenário dos biocombustíveis, que tem sido objeto de crescente interesse por se tratar de uma fonte de energia mais limpa, que emite menos gases ligados ao efeito estufa e conseqüentemente, apresenta menor contribuição para o aquecimento global (de Abreu & Nascimento, 2016; Rosseto et al., 2017).

No entanto, a matéria prima utilizada para a produção dos biocombustíveis exige uma extensa área de plantio, como a cana e o milho. A cana-de-açúcar já é usada para a produção de etanol de primeira geração (E1G), processo que gera resíduos lignocelulósicos, que hoje são utilizados para queima em caldeiras para produção de energia, mas que podem também ser aproveitados na produção de etanol de segunda geração (E2G) (Dias, Schlindwein, da Silva, & Ruviaro, 2015; Espíndola, 2015).

O bagaço de cana-de-açúcar é um dos resíduos mais ricos em compostos lignocelulolíticos e é produzida em abundância pela indústria sucroalcooleira. Atrelado a isso, tem-se a palha da cana que é removida na colheita e acumulada nos canaviais. Antes esse material era queimado, com a intenção de facilitar a colheita e diminuir os acidentes com animais peçonhentos, como aranhas e escorpiões, porém foi regulamentada a Lei da Queima da Cana N° 11.241/02 que proibiu a queima desse material, favorecendo ainda mais o seu acúmulo (Brinkman et al., 2018; de Oliveira Bordonal et al., 2018) .

A produção de etanol de segunda geração (E2G) visa o aproveitamento principalmente dos resíduos gerados na produção de E1G, mas essa tecnologia também pode ser aplicada para outros resíduos lignocelulósicos que são gerados pela agroindústria (Furtado, Hekkert, & Negro, 2020).

Por definição, biomassa lignocelulósica é toda matéria orgânica vegetal que armazena energia na forma de ligações químicas. Constituída principalmente por polissacarídeos, como a celulose e hemicelulose, além de compostos fenólicos, como a lignina que são arranjados de maneira recalcitrante (de Siqueira et al., 2010).

A celulose é formada por subunidades de D-glicose, unidas por ligações glicosídicas. Já a hemicelulose é constituída por diferentes tipos de pentoses, hexoses e açúcares ácidos unidos entre si. E a lignina é associada a parede celular vegetal em um arranjo complexo e

organizado conferindo à planta maior rigidez e impermeabilidade (Bhatia et al., 2019; Jagtap et al., 2012).

Esses componentes da biomassa precisam ser hidrolisados em açúcares fermentescíveis para posteriormente serem convertidos a etanol. A despolimerização pode ser feita com a aplicação de reagentes químicos ou utilização de coquetéis enzimáticos, sendo que este último é mais caro e depende fundamentalmente da evolução biotecnológica (Sorn, Chang, Phitsuwan, Ratanakhanokchai, & Dong, 2019).

Enquanto a maioria dos estudos investigam principalmente as enzimas que atuam na degradação da celulose, esse trabalho teve como foco de estudo uma endo-beta-1,4-xilanase, que é uma hemicelulose que hidrolisa xilana, um dos principais componentes da hemicelulose.

As xilanases possuem diferentes aplicações industriais, tais como, na fabricação de bebidas e na indústria de papel e celulose, mas dentro da cadeia produtiva de bioetanol ela é importante não só pela degradação de uma parte da hemicelulose, mas também por proporcionar um melhor aproveitamento das pentoses e otimizar a ação das celulases, aumentando a eficiência da hidrólise e conseqüentemente a produção de açúcares fermentescíveis (Bibra, Kunreddy, & Sani, 2018; X. Liu et al., 2018).

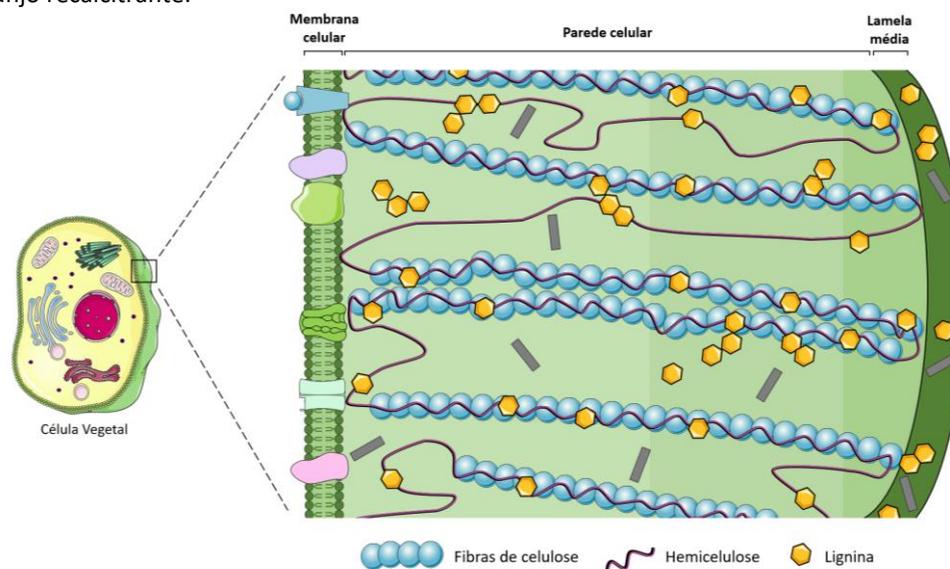
Dessa maneira, a bioprospecção de enzimas lignocelulósicas eficientes e estáveis em condições fermentativas é importante para que haja um melhor aproveitamento de todos os componentes da biomassa e conseqüentemente a otimização de processos industriais e biotecnológicos, dentre eles, a produção do etanol 2G.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. Biomassa lignocelulósica

A biomassa é composta por polissacarídeos como a celulose, hemicelulose em associação com moléculas aromáticas, como a lignina, sendo unidos por ligações covalentes ou pontes de hidrogênio (Figura 1). Esses constituintes estão presentes na parede celular vegetal e estão dispostos em um arranjo complexo, formando estruturas resistentes à choques mecânicos, ação microbiana e enzimática, o que dificulta a conversão da biomassa em produtos de valor agregado (de Siqueira et al., 2010; Zhang, Donaldson, & Ma, 2012).

Figura 1 – Representação dos principais componentes da biomassa lignocelulósica. Ela é composta por três componentes principais (celulose, hemicelulose e lignina) encontrados na parede celular vegetal e organizados em um arranjo recalcitrante.



Adaptado de (Achyuthan et al., 2010)

Os polissacarídeos podem ser hidrolisados em açúcares fermentescíveis e a lignina é usada na produção de produtos químicos. A composição de cada um desses componentes varia de entre as espécies de planta e entre os resíduos agroindustriais (Tabela 1). Além disso, em uma mesma planta pode haver variações de conteúdo devido a diferentes fatores, tais como a idade, condições de cultivo e época de colheita (M. Kim & Day, 2011).

Tabela 1 – Composição lignocelulósica de diferentes plantas e resíduos agroindustriais.

Material Lignocelulósico	Celulose (%)	Hemicelulose (%)	Lignina (%)	Referência
Bagaço de cana-de-açúcar	40-45	30-35	20-30	(Cardona & Quintero, 2010)
Bagaço de sorgo doce	45	27	21	(M. Kim & Day, 2011)
Eucalipto	44,9	28,9	26,2	(Muranaka et al., 2017)
Seringueira	39,5	28,4	27,5	(Khan et al., 2018)
Pinho	45,6	24	26,8	(Yu, Paterson, Blamey, & Millan, 2017)
Folhas de bambu	34,1	25,5	35	(Huang, Chiueh, & Lo, 2016)
Palha de trigo	33-40	20-25	15-20	(Talebnia, Karakashev, & Angelidaki, 2010)
Palha de arroz	38	32	12	(Lu & Hsieh, 2012)
Palha de soja	34	16	22	(Wan, Zhou, & Li, 2011)
Palha de milho	43,9	29,8	21,8	(Huang et al., 2016)
Espiga de milho	41	31	12	(Chen, Zhang, Singh, Yu, & Yang, 2010)
Casca de avelã	30	23	38	(Álvarez et al., 2018)
Talo de algodão	67	16	13	(H. Kim, Ahn, & Kwak, 2016)

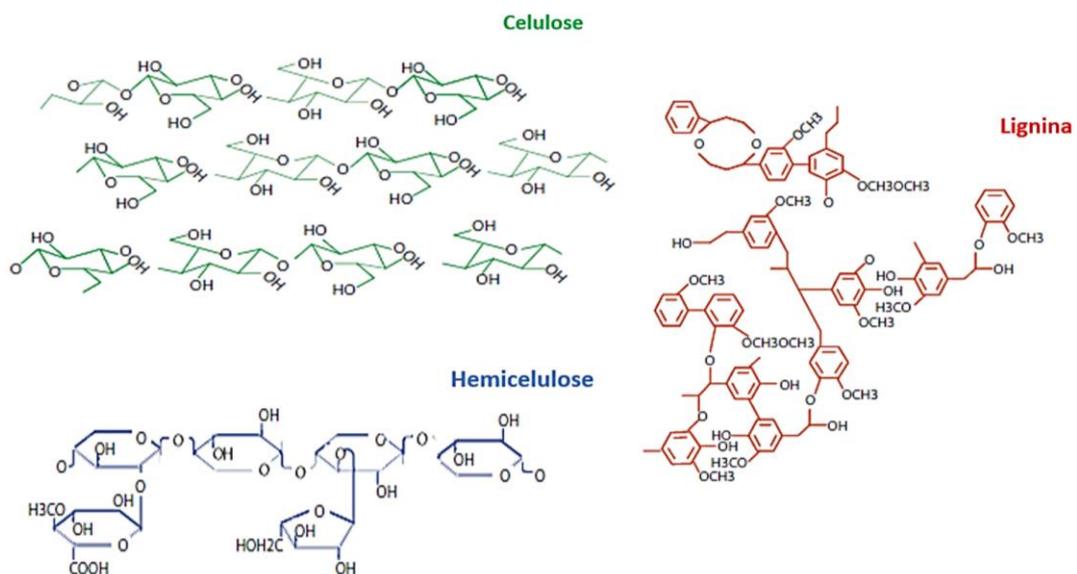
Elaborado pela autora.

Todo ano são gerados milhões de toneladas de material vegetal não comestível, tais como bagaço de cana de açúcar, palha de milho e de arroz, restos de madeiras lenhosas, entre outros. Todos esses resíduos lignocelulósicos que são entulhados ou queimados poderiam servir como matéria-prima para produção de etanol de segunda geração (Florencio, Badino, & Farinas, 2017).

O principal constituinte da biomassa é a celulose. Trata-se de um homopolímero linear com 100 a 140.000 unidades de celobiose (dissacarídeo de glicose) unidos por ligações glicosídicas do tipo β -1,4 (Figura 2). Além disso, cada fibra de celulose está ligada a outra por pontes de hidrogênio, formando um complexo alinhado e estável, com grau de cristalinidade variável, com regiões amorfas e regiões mais compactadas (Karimi & Taherzadeh, 2016; Santos et al., 2012).

Já a hemicelulose é um heteropolímero de cadeia curta (cerca de 50-300 monossacarídeos) com estrutura complexa, cuja composição depende da fonte. É formada principalmente por pentoses e hexoses, podendo apresentar cadeias ramificadas de arabinose, manose, ácido β -D-glucurônico, D-4-O-metilglucurônico, entre outros (Figura 2). Ela está ligada de forma não covalente às fibras de celulose e atua como uma barreira física, limitando o acesso às fibras de celulose. Entretanto, devido sua estrutura mais amorfa, ela seja mais suscetível a agentes externos (Peng, Peng, Xu, & Sun, 2012; Wyman et al., 2005).

Figura 2 – Estrutura química dos principais componentes da biomassa lignocelulósica. A celulose é formada por unidades de glicose; a hemicelulose é constituída por pentoses, hexoses e açúcares ácidos associados, enquanto a lignina apresenta compostos fenólicos em sua composição.



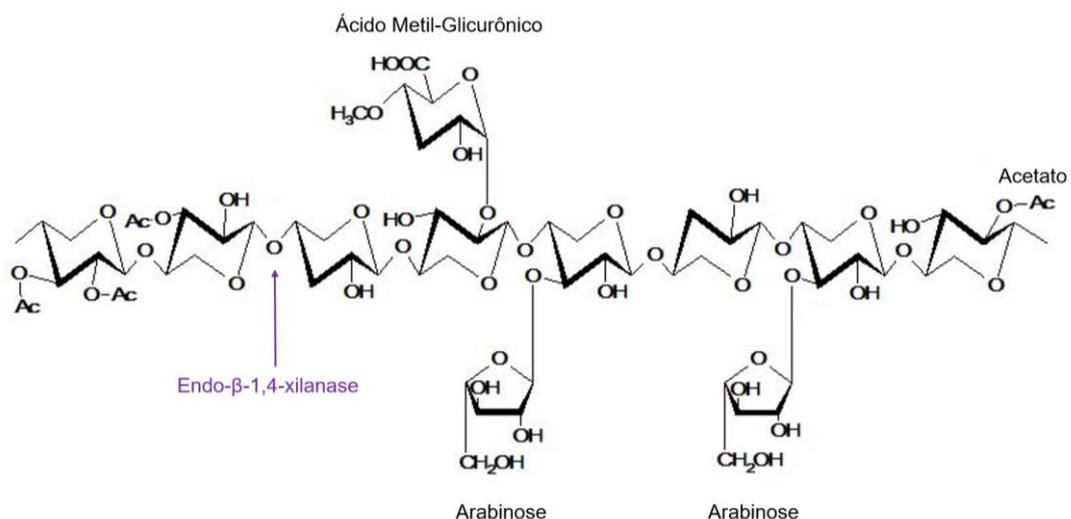
A lignina é um composto fenólico, formada por três diferentes monômeros: álcool cumárico, álcool coniferílico e álcool sinapílico (Figura 2). Por não possuir açúcares em sua composição, ela não é utilizada diretamente na produção de etanol, mas devido sua natureza aromática, é apreciada pela indústria química para obtenção de diferentes insumos. Na estrutura da biomassa, ela atua como uma cola ao redor das fibras de celulose e hemicelulose, formando uma estrutura recalcitrante, cuja principal função é a resistência mecânica e microbiana (Bhatia et al., 2019; Bichot et al., 2018)

2.1.1. Xilana

A xilana é o principal e mais abundante componente da hemicelulose. Trata-se de um heteropolissacarídeo formado por uma cadeia principal composta de resíduos de D-xilose unidos por ligações glicosídicas do tipo β -1,4 (Figura 3). Pode ter diferentes grupos associados à sua cadeia principal, tais como arabinose, grupos O-acetil, ácidos urônico, ferúlico, entre outros (Bibra et al., 2018; X. Liu et al., 2018).

Comumente são classificadas de acordo com as cadeias laterais que apresenta, como por exemplo, as glucuronoxilanas (apresentam resíduos de ácidos 4-O-metilglucurônicos) ou as arabinoxilanas (apresentam unidades de L-arabinofuranosil) (Kumar et al., 2017; Peng et al., 2012).

Figura 3 – Estrutura química da xilana. A molécula de xilana é composta principalmente por pentoses, podendo ter diferentes cadeias laterais associadas a ela.



Adaptado de Collins, Gerday & Feller, 2005.

2.2. Etanol a partir da biomassa (E2G)

Diferentes resíduos florestais e agrícolas têm sido implementados para produção de bioetanol (Tabela 2). Nota-se que no Brasil, os principais resíduos utilizados são aqueles gerados a partir da produção de etanol de primeira geração, o bagaço e a palha da cana-de-açúcar (Jansen et al., 2017).

Tabela 2 – Produção de etanol 2G no mundo. Dados dos EUA e Canadá refletem o cenário de maio de 2017 e os dados dos outros países ilustram o status em 2016.

Empresa/fábrica	País (estado)	Matéria-prima	Capacidade ML/ano ⁻¹
DuPont Cellulosic Ethanol LLC—Nevada	EUA (IA)	Palha de milho	113,6
Poet-DSM Advanced Biofuels LLC—Projeto Liberty	EUA (IA)	Espigas de milho/palha de milho	75,7
Planta de etanol celulósico de Quad County	EUA (IA)	Fibra de milho	7,6
Fábrica de Demonstração Fiberight	EUA (VA)	Fluxo de resíduos	1,9
Biorrefinaria Celulósica Integrada Piloto da ICM Inc.	EUA (MO)	Culturas de biomassa	1,2
American Process Inc.—Thomaston Biorefinery	EUA (GA)	De outros	1,1
ZeaChem Inc.—fábrica de demonstração	EUA (OU)	Culturas de biomassa	1,0
Enerkem Alberta Biofuels LP	Canadá (AB)	Resíduos sólidos urbanos	38,0
Enerkem Inc.—Westbury	Canadá (QC)	Biomassa lenhosa	5,0
Corporação Iogen	Canadá (ON)	Resíduos de colheita	2,0
Woodlands Biofuels Inc.—fábrica de demonstração	Canadá (ON)	Biomassa lenhosa	2,0
GranBio	Brasil	Bagaço	82,4
Raizen	Brasil	Bagaço de cana/palha	40,3
Longlive Bio-technology Co. Ltd —demonstração comercial	China	Espigas de milho	63,4
Mussi Chemtex/Beta Renováveis	Itália	<i>Arundo donax</i> , palha de arroz, palha de trigo	75,0
Borregaard Industries AS—ChemCell Etanol	Noruega	Resíduos de polpação de madeira	20,0

Adaptado de Jansen et al., 2017.

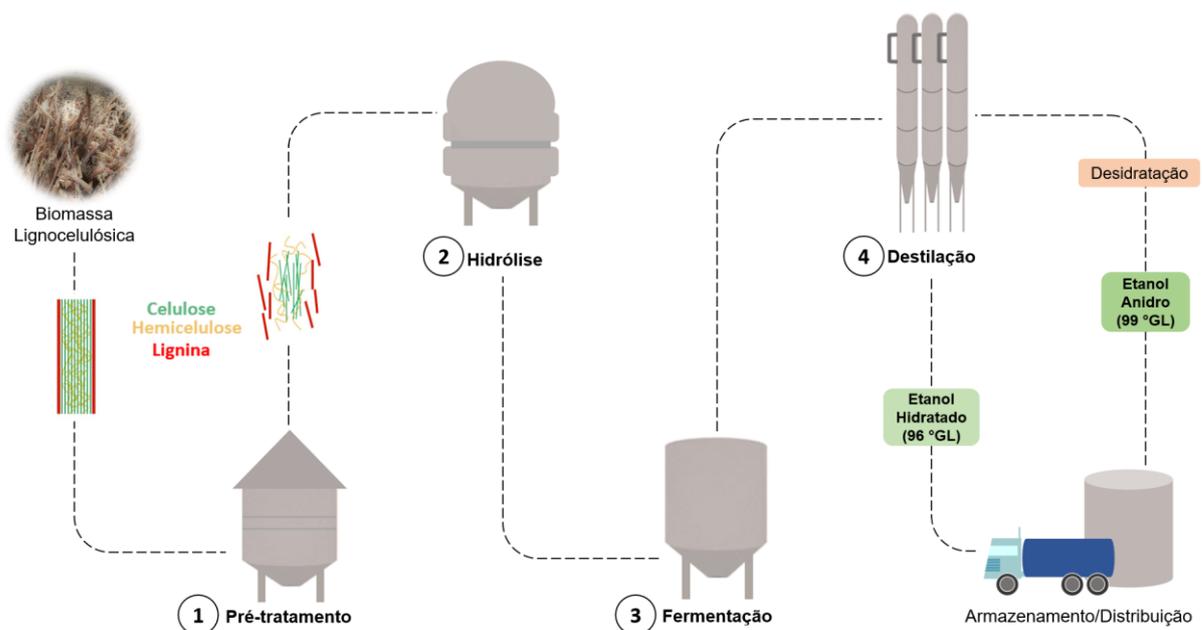
A produção de etanol de primeira geração ocorre através do sistema tradicional, através da fermentação da sacarose presente no caldo de cana. Por outro lado, o etanol de segunda geração produzido em nosso país é obtido a partir do aproveitamento da palha e do bagaço desse produto. Além de representar uma alternativa para o uso a biomassa, a produção de E2G traz consigo outros benefícios, tais como: uma vantagem logística de utilização de

insumos já disponíveis, aumento na produção de etanol sem ampliação da área de cultivo e possibilidade de produção de biocombustível durante a entressafra da cana (Brinkman et al., 2018; Khoo, 2015).

Basicamente o processo de produção de etanol 2G envolve quatro etapas: pré-tratamento da biomassa lignocelulósica, hidrólise, fermentação e destilação (Figura 4) (Sá, Cammarota, & LEITÃO, 2014; Santos et al., 2014).

A primeira etapa é o pré-tratamento, cujo objetivo é despolimerizar toda a estrutura vegetal, deixando seus constituintes mais expostos. Todavia, esse também é o principal gargalo na produção de E2G, devido ao complexo arranjo desse material (Rosseto et al., 2017).

Figura 4 – Processo de produção de etanol 2G. O processo de produção de E2G é resumido em quatro etapas: (1) o pré-tratamento para desestruturar a biomassa, expondo seus constituintes; (2) a hidrólise dos polissacarídeos em açúcares fermentescíveis; (3) a fermentação onde os açúcares serão convertidos em etanol; (4) a destilação para separação do etanol hidratado de qualquer impureza oriunda do processo. Se o objetivo for a obtenção do etanol anidro, mais uma etapa é incluída na produção, a desidratação.



Elaborado pela autora.

Um pré-tratamento eficiente envolve a decomposição da matriz de lignina e hemicelulose, bem como o rompimento das pontes de hidrogênio estabelecidas entre as fibras da celulose, deixando esses componentes mais fluidos e aumentando a superfície de contato e eficiência da etapa de hidrólise. Para isso podem ser empregados diferentes métodos: físicos (trituração e moagem); químicos (utilização H_2SO_4 , HCl ou HNO_3); físico-químicos (explosão a vapor) ou biológicos (Chen et al., 2010; Saini, Saini, & Tewari, 2015).

A escolha do melhor processo depende principalmente do material a ser tratado, tendo em vista que cada matéria-prima apresenta uma composição lignocelulósica diferente. Entretanto, é preciso ponderar toda a logística e custo do processo, pois a utilização de um método muito brando e econômico pode resultar em uma baixa eficiência na hidrólise devido a pouca exposição dos substratos, assim como um método intenso pode ocasionar degradação dos polissacarídeos e/ou formação de compostos inibidores da hidrólise e fermentação (Baruah et al., 2018; H. Kim et al., 2016; Wan et al., 2011).

Após o pré-tratamento acontece a hidrólise dos componentes lignocelulósicos, que pode ser ácida, com a utilização de ácidos fortes ou enzimática, catalisada por enzimas lignocelulolíticas para produção de açúcares fermentescíveis. Esses açúcares serão degradados pelas leveduras presentes nas dornas de fermentação, produzindo álcool. (Furtado et al., 2020; Khoo, 2015; Talebnia et al., 2010).

No Brasil, a maioria dos produtores de etanol adotam o processo de fermentação *Melle-Boinot*, caracterizado pela recuperação da levedura no final da cadeia fermentativa através de centrifugação. Essas leveduras recuperadas são tratadas com agentes bactericidas (H_2SO_4 em pH 2,5–3,0) e retornam ao processo fermentativo, sendo que em alguns casos uma parte é desidratada e comercializada como composto para ração animal (Lopes et al., 2016; Neitzel et al., 2020).

Ao final da fermentação, obtém-se o chamado vinho, que por sua vez é encaminhado para as colunas de destilação, cujo principal objetivo é realizar a separação do etanol hidratado dos seus componentes secundários, como álcool propílico, isoamílico, entre outros (Brinkman et al., 2018).

Basicamente são comercializados dois tipos de álcool: o etanol hidratado (96°GL) utilizado como combustível e o álcool anidro (99°GL) que é adicionado à gasolina principalmente para elevação do índice de octanagem do combustível. Para elevar a concentração do álcool são utilizados métodos de desidratação por meio do uso de ciclohexano, monoetilenoglicol, etc (Domínguez-Bocanegra, Torres-Muñoz, & López, 2015; Lopes et al., 2016).

Embora o E2G seja obtido a partir de resíduos, os custos de produção ainda são elevados, tornando a adesão a esse processo pouco atrativa. Por esse motivo, a integração de

etapas da cadeia de produção é vista como uma alternativa para redução de custos, economia de energia e água, bem como a otimização do processo.

Na literatura, alguns autores apontam a Sacarificação e Fermentação Simultânea (SSF - do inglês *Simultaneous Saccharification and Fermentation*) como uma dessas alternativas, visto que esse processo propõe a realização da hidrólise enzimática e da fermentação em um mesmo reator, ou seja, a levedura produz etanol à medida que os açúcares vão sendo formados. Seu emprego reduz os riscos de contaminação microbiana, minimiza a inibição enzimática devido a altas concentrações de glicose e ocasiona economia de água e energia, quando comparado à Hidrólise e Fermentação Separadas (SHF – do inglês *Separate Hydrolysis and Fermentation*) (Chohan, Aruwajoye, Sewsynker-Sukai, & Kana, 2020; Saha et al., 2015).

O principal impasse para implementação do processo SSF é que as enzimas lignocelulósicas e as leveduras precisam atuar dentro de uma mesma temperatura. As enzimas comercialmente disponíveis apresentam uma boa temperatura de atividade em torno de 50-60°C, enquanto as leveduras fermentam com eficiência a 30°C. Logo, se faz necessário a bioprospecção de enzimas lignocelulolíticas que atuem em temperaturas mais baixas e/ou a identificação de leveduras termotolerantes (Prado et al., 2020; Saha, Nichols, & Cotta, 2013)

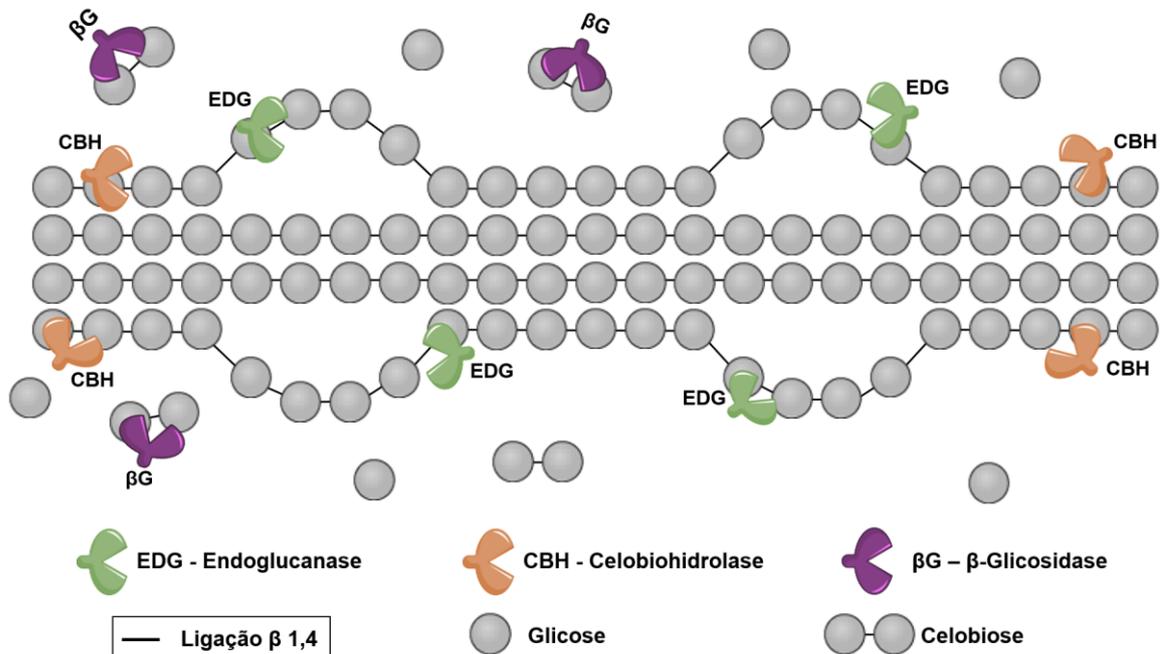
2.3. Enzimas lignocelulósicas

2.3.1. Celulases

O termo “celulases” é usado para definir um conjunto de enzimas que atuam de maneira sinérgica para hidrolisar o polímero de celulose. Esse complexo engloba as endoglucanases, celobiohidrolases/exoglucanases e as β -glicosidases (Florencio et al., 2017; Silva, 2016).

As endoglucanases (E.C. 3.2.1.4) clivam randomicamente as ligações β -1,4 das regiões amorfas da celulose, diminuindo seu grau de polimerização. Já as celobiohidrolases (E.C. 3.2.1.91) atuam nas extremidades reductoras e não-reductoras da cadeia de celulose liberando celobiose. E as β -glicosidases (E.C. 3.2.1.21) tem a função de hidrolisar as ligações das moléculas de celobiose, liberando glicose (Figura 5) (Horn, Vaaje-Kolstad, Westereng, & Eijsink, 2012).

Figura 5 – Degradação enzimática da cadeia de celulose. O sinergismo entre as celulases garante uma hidrólise eficiente da celulose. As endoglucanases (EDG) quebram a fibra de celulose em cadeias menores que tem suas extremidades hidrolisadas pelas celobiohidrolases (CBH), liberando moléculas de celobiose que são degradadas em glicose pelas β -glicosidases (β G).



Elaborado pela autora.

2.3.2. Xilanases

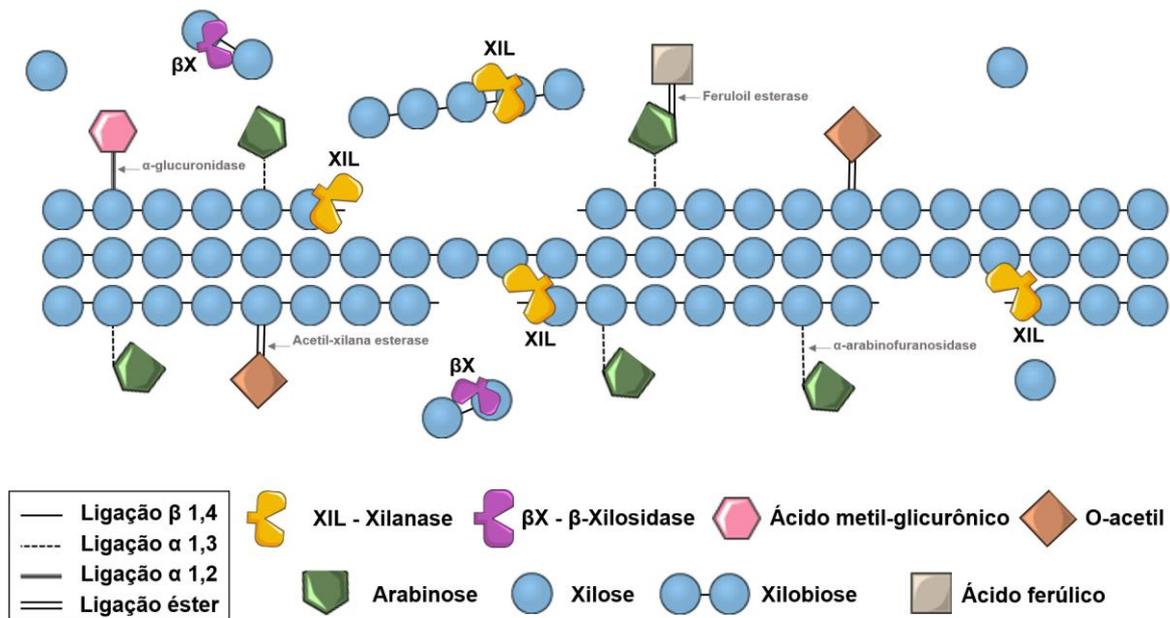
As xilanases (E.C. 3.2.1.8) catalisam a quebra das ligações β -1,4 da cadeia principal da xilana em xilooligossacarídeos e D-xilose. Entretanto, o sistema xilanólítico é formado por um complexo de enzimas hidrolíticas que atuam de maneira cooperativa com as xilanases para hidrolisar a xilana, tais como: β -xilosidases (E.C. 3.2.1.37), acetil-xilana esterases (E.C. 3.2.1.72), α -arabinofuranosidase (E.C. 3.2.1.55), α -glucuronidase (E.C. 3.2.1.139), ácido ferúlico esterase (E.C. 3.2.1.73), entre outras (Figura 6) (Collins, Gerday, & Feller, 2005).

De acordo com o banco de dados CAZy, as xilanases estão relacionadas com as famílias 5, 8, 10, 11, 16, 26, 30 e 62 das GHs. Sendo que as mais estudadas são aquelas dentro das famílias 10 e 11, classificadas com base na análise de seus domínios catalíticos e na similaridade das sequências de aminoácidos. As de família 10 apresentam topologia de *TIM barril*, massa molar > 30 kDa e menor especificidade por diferentes tipos de xilanas. Já as xilanases de família 11,

apresentam forma de *jelly-roll*, são chamadas de xilanases verdadeiras, apresentam massa molar < 30 kDa e atuam especificamente sobre a xilana (Bibra et al., 2018; Collins et al., 2005).

As enzimas pertencentes às famílias 5, 7, 10 e 11 apresentam um mecanismo catalítico baseado na ação de dois resíduos de ácidos glutâmicos presentes em seus sítios ativos, onde um atua como catalisador ácido e protona o oxigênio da ligação glicosídica e o outro realiza um ataque nucleofílico (Arakawa, Watanabe, Yamasaki, Maekawa, & Tokuda, 2009; Paës, Berrin, & Beaugrand, 2012).

Figura 6 – Degradação enzimática da xilana. As xilanases (XIL) quebram as ligações β -1,4 da cadeia principal da xilana, originando xilooligossacarídeos que são hidrolisados pela β -xilosidase (β X) em xilose. Diferentes enzimas acessórias (em cinza) atuam na remoção das cadeias laterais que podem estar associadas à cadeia principal desse heteropolissacarídeo.



Elaborado pela autora.

Na produção de etanol 2G as xilanases são utilizadas para degradação da xilana presente na hemicelulose e conseqüentemente proporcionar um melhor aproveitamento das pentoses presentes na biomassa. Além disso, seu uso otimiza a ação das celulases, devido maior exposição do polímero de celulose (Nielsen, Zacchi, Galbe, & Wallberg, 2017).

Além disso, as xilanases possuem uma gama de aplicações: na panificação, para melhorar a qualidade e estrutura da massa; na indústria de papel e celulose, no branqueamento da polpa de madeira; na clarificação de bebidas; na fabricação de ração animal, aumentando a

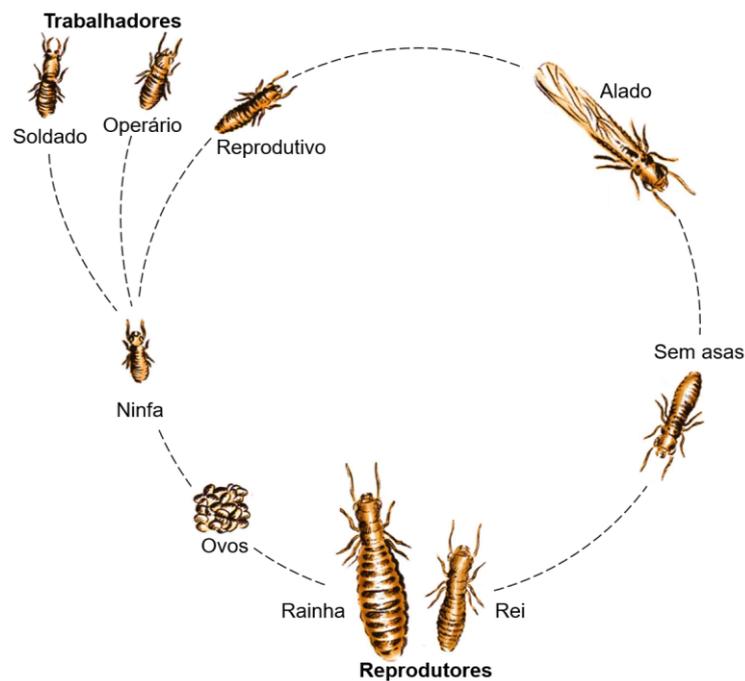
digestibilidade da mesma; na obtenção de xilooligossacarídeos, como o adoçante xilitol, dentre outras (Kumar et al., 2017; Nielsen et al., 2017; Paës et al., 2012).

Por possuírem uma gama de aplicações, enzimas lignocelulolíticas podem ser encontradas em diferentes organismos, desde bactérias, fungos até insetos, como os cupins, por exemplo.

2.4. Cupins

São insetos sociais, habitam em colônias compostas por indivíduos especializados morfologicamente de acordo com a casta que ocupam. A cada casta é atribuída uma função específica fundamental para a sobrevivência do cupinzeiro, criando assim uma interdependência entre elas (Castro et al., 2015; Haifig & Costa-Leonardo, 2016).

Figura 7 – Ciclo de vida dos cupins. A rainha realiza a postura dos ovos, de onde eclodem as ninfas que podem se diferenciar na casta trabalhadora ou reprodutiva. Os cupins reprodutivos apresentam asas que são perdidas no processo de revoada, para estabelecimento de uma nova colônia.



Adaptado de Rocha & Wiederhecker, 2018.

Os cupins são descritos como hemimetábulos, ou seja, apresentam durante seu desenvolvimento um processo incompleto de metamorfose (Figura 7). Do ovo eclode a ninfa (forma jovem) que ao longo de várias transformações cresce e assume um papel em uma das castas da colônia: soldados, responsáveis pela proteção e defesa do cupinzeiro; operários, encarregados pela construção de túneis, pela busca de alimentos e cuidado com os ovos e

reprodutores (rei e rainha), incumbidos pelo acasalamento e postura de ovos (Casarin, Costa-Leonardo, & Arab, 2008; Lima & Costa-Leonardo, 2007).

Operários e soldados são estéreis devido ao desenvolvimento sexual incompleto. Enquanto soldados apresentam cabeça e mandíbulas bem desenvolvidas, os operários possuem tamanho variado e são os mais numerosos. Já os reprodutores possuem asas frágeis que não permitem um voo longo. Deste modo ao caírem no solo buscam um local apropriado para formar uma nova colônia (Roisin & Korb, 2010).

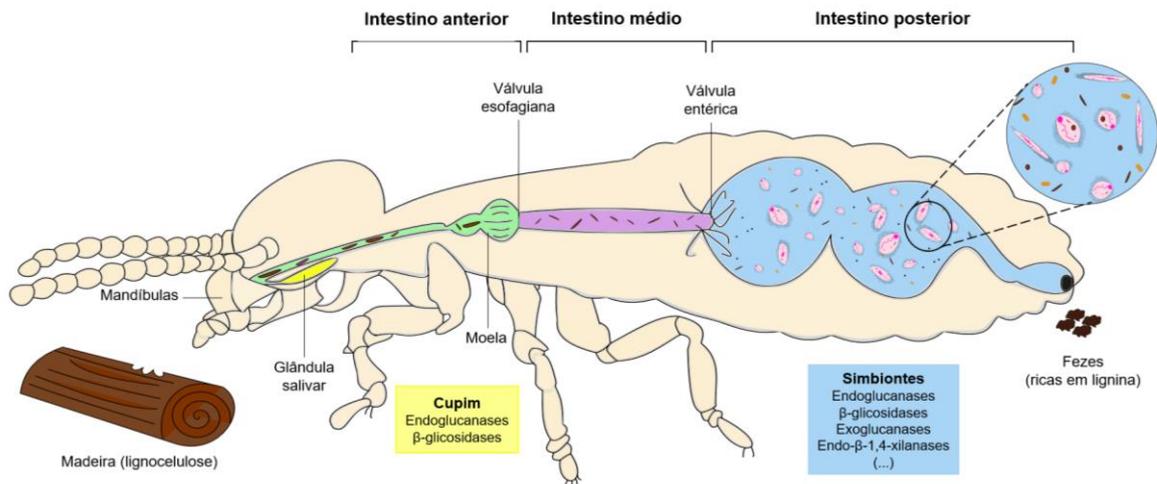
Os cupins são classificados em inferiores e superiores com base em três principais fatores: na presença de endossimbiontes, nos hábitos de forrageamento/alimentação e no caminho para a diferenciação das castas. Os cupins superiores apresentam um desenvolvimento bifurcado, onde há um caminho para formação dos reprodutores e outro para a formação de soldados e operários. Já os cupins inferiores apresentam apenas uma linha de desenvolvimento, de onde todas as castas se diferenciam (Campanini et al., 2021; Korb & Belles, 2017; Scharf, 2015).

Os superiores (família Termitidae) apresentam uma alimentação diversa: madeira, húmus, esterco animal, matéria orgânica presente no solo, entre outros. Por outro lado, os cupins inferiores (família Mastotermitidae, Stolotermitidae, Hodotermitidae, Archotermopsidae, Kalotermitidae, Stylotermitidae, Serritermitidae, e Rhinotermitidae) se alimentam principalmente de madeira (Bourguignon et al., 2016; Brune, 2014; Scharf & Tartar, 2008).

Além disso, as bactérias e archaeas estão presentes no trato intestinal de todas as espécies de cupins. Contudo, os protistas só são encontrados nas linhagens mais basais, os cupins inferiores (Lo & Eggleton, 2010; Warnecke et al., 2007).

Embora alguns cupins superiores e baratas aparentemente tenham lignocelulases endógenas, algumas outras espécies de cupins, principalmente os inferiores recebem a cooperação de archaeas, bactérias e protozoários flagelados presentes no seu trato intestinal posterior para degradar a lignocelulose (Figura 8) (Brune, 2014; Jang & Kikuchi, 2020; Scharf & Tartar, 2008).

Figura 8 – Degradação da lignocelulose feita por cupins inferiores e seus simbiontes intestinais. No intestino anterior as partículas da lignocelulose que foram trituradas pelas mandíbulas são misturadas com enzimas secretadas pelas glândulas salivares desses insetos. No intestino médio ocorre a absorção da glicose que já foi produzida e as partículas parcialmente digeridas passam para o intestino posterior, onde os simbiontes (archaeas, bactérias e protistas flagelados) hidrolisam os polissacarídeos restantes utilizando seu arsenal enzimático. Os resíduos de lignina não são digeridos e posteriormente são eliminados pelas fezes (Brune, 2014).



Adaptado de Brune, 2014.

2.4.1. *Heterotermes tenuis*

Heterotermes tenuis é uma espécie de cupim subterrânea nativa do Brasil, pertencente à ordem Isoptera e à família Rhinotermitidae sendo caracterizados como cupins inferiores. Bem adaptado a ambientes de clima tropical e subtropical, consegue suportar ambientes mais secos e a sua colônia é formada por uma rede de galerias subterrâneas que requer umidade para manutenção do seu metabolismo (Eaton, Jones, & Jenkins, 2016; Griebenow, Jones, & Eaton, 2017).

É considerada uma praga agrícola e urbana com importante impacto econômico por causar danos em canaviais e infestação em edifícios por se alimentar de madeira e qualquer outro material celulósico (Bourguignon et al., 2016).

O nosso grupo de pesquisa avaliou o meta-transcriptoma de *Heterotermes tenuis* buscando a identificação de genes biotecnológicos e envolvidos em diferenciação de casta. O fato de que insetos não são termorreguladores e forrageiam a noite, quando a temperatura é mais amena, sugere que as enzimas que atuam nesse cupim possuem atividade em temperaturas menores do que as comercialmente disponíveis, configurando-se como um

importante alvo para potencial aplicação na produção de E2G e processo de sacarificação e fermentação simultânea (Campanini et al., 2021).

3. OBJETIVOS

3.1. Objetivo Geral

O presente projeto tem como objetivo a clonagem e expressão heteróloga de uma endo- β -1,4 xilanase identificada no meta-transcriptoma de cupins da espécie *Heterotermes tenuis* para possível uso na produção de etanol de segunda geração.

3.2. Objetivos específicos

- i. Definição da endo-beta-1,4 xilanase com base nos dados do meta-transcriptoma previamente feito pelo grupo de pesquisa;
- ii. Análise *in silico* da sequência (predição de peptídeo sinal, sítios de fosforilação, glicosilação, entre outras);
- iii. Clonagem e expressão da proteína na bactéria *Escherichia coli* e na levedura *Pichia pastoris*;
- iv. Purificação da proteína por cromatografia de afinidade em resina de níquel;
- v. Teste de atividade em ágar xilana a 30°C e 40°C;

4. MATERIAIS E MÉTODOS

4.1. Materiais

4.1.1. Cupins

Os cupins utilizados na realização deste projeto foram gentilmente cedidos pela Profa. Dra. Ana Maria Costa Leonardo da Universidade Estadual Paulista “Júlio Mesquita” (UNESP) Campus Rio Claro. Vale ressaltar que apenas os operários da espécie *Heterotermes tenuis* foram utilizados e essa escolha foi pautada em análises de

bioinformática e ensaios preliminares realizados em laboratório, atrelados ao conhecimento prévio da função que esta casta desempenha, que é a busca de alimento.

4.1.2. Linhagens de *Escherichia coli* utilizadas

Para clonagem e propagação do DNA exógeno, foi utilizada neste estudo a linhagem bacteriana eletrocompetente de *E. coli* DH5- α ($F^- \phi 80lacZ\Delta M15 \Delta(lacZYA-argF)U169 \text{ } recA1 \text{ } endA1 \text{ } hsdR17(r_K^-, m_K^+) \text{ } phoA \text{ } supE44 \text{ } \lambda^- \text{ } thi-1 \text{ } gyrA96 \text{ } relA1$) e para expressão da proteína recombinante no sistema pET28a foi utilizada a linhagem de *E. coli* Rosetta (DE3) $F^- \text{ } ompT \text{ } hsdSB(rB - mB -) \text{ } gal \text{ } dcm \text{ } (DE3) \text{ } pRARE2 \text{ } (Cam^R)$.

4.1.3. Linhagens de *Pichia pastoris* utilizadas

A levedura *P. pastoris* é metilotrófica, ou seja, é capaz de crescer utilizando metanol como única fonte de carbono, devido a ação da enzima álcool oxidase. Em seu genoma essa levedura possui dois genes que codificam essa enzima: *AOX1* e *AOX2*, sendo o *AOX1* regulador da maior parte da atividade da enzima álcool oxidase na célula.

Neste projeto foi utilizada a linhagem de *P. pastoris* X-33 (*wild type*), que possui o gene *AOX* íntegro e, portanto, apresenta um crescimento rápido em metanol, quando comparada com linhagens que possuem o gene *AOX1* mutado, como a KM71H, que por esse motivo apresenta um crescimento mais lento, dependente do gene *AOX2* que é menos ativo (Cregg, Cereghino, Shi, & Higgins, 2000).

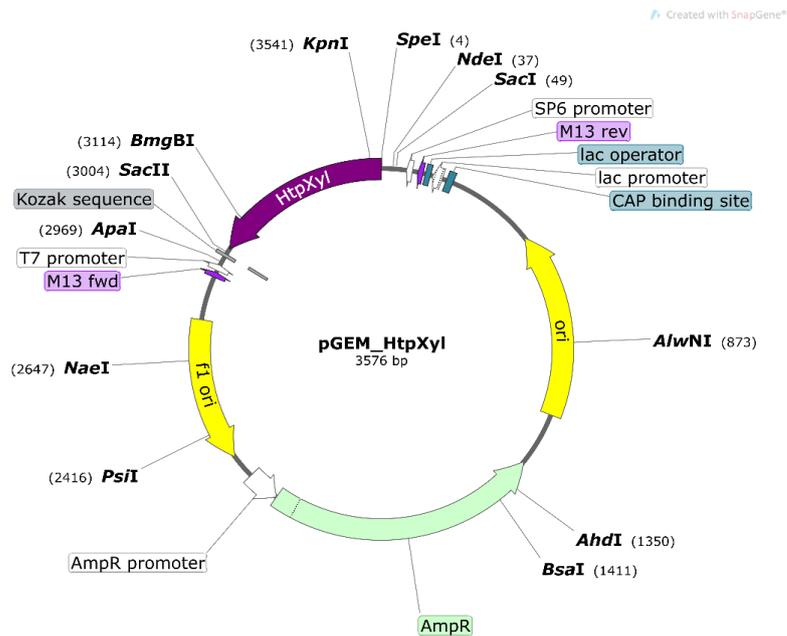
4.1.4. Plasmídeos

4.1.4.1. pGEM-T *easy*

Plasmídeo comercial pré-linearizado utilizado para clonagem de produtos de PCR que são gerados por polimerases que adicionam uma deoxiadenosina a cada extremidade do inserto amplificado. Essas extremidades são compatíveis com a do vetor, que possui extremidades 3'-T (Robles & Doers, 1994).

Apresenta como marcador de seleção o gene de resistência a ampicilina e possui os promotores de RNA polimerase T7 e SP6, flanqueando o sítio de múltipla clonagem (MCS – *Multiple Cloning Site*) que está localizado dentro da região codificadora da β -galactosidase (Figura 9), o que permite a seleção de clones recombinantes pela triagem azul/branco (Robles & Doers, 1994).

Figura 9 – Vetor pGEM_HtpXyl.

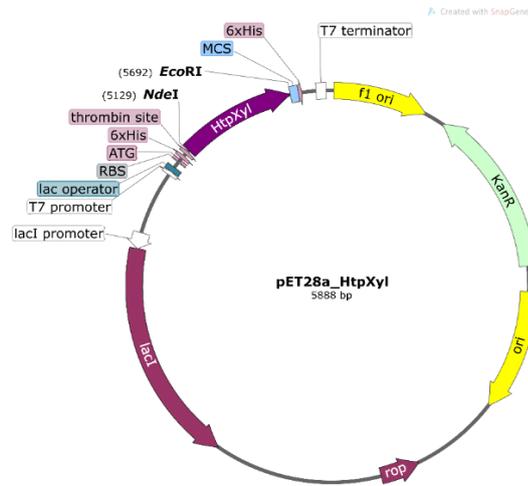


4.1.4.2. pET28a

Plasmídeo utilizado para expressão de proteínas recombinantes em *E. coli*. Possui o promotor T7-lac que é induzido na presença de IPTG e o gene de resistência a canamicina como marcador de seleção. Além disso apresenta sítio de trombina e sequência que codifica 6 resíduos de histidina (6x His-Tag) tanto na região N-terminal quanto na C-terminal. Ela é essencial para purificação de proteínas recombinantes em colunas que contenham resina com íons metálicos (Figura 10) (Dubendorf & Studier, 1991; Studier & Moffatt, 1986).

Na estratégia de clonagem que foi adotada neste trabalho, a His-Tag está na região N-terminal da proteína e as enzimas de restrição utilizadas foram *Nde I* e *EcoR I*.

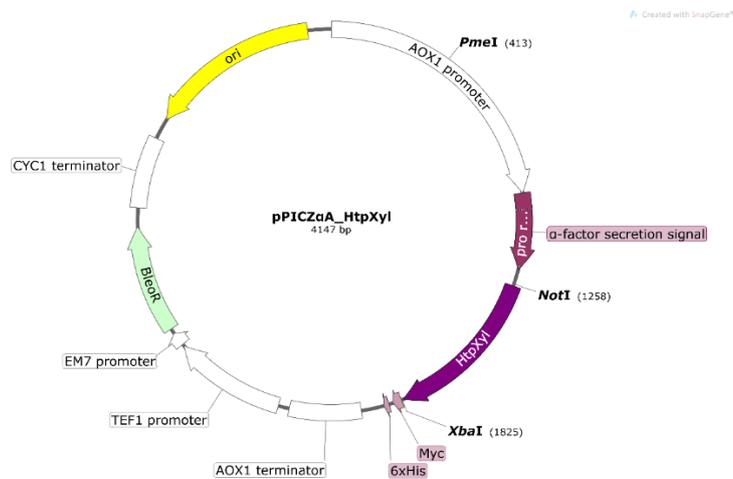
Figura 10 – Vetor pET28a_HtpXyl.



4.1.4.3. pPICZαA

Plasmídeo utilizado para expressão de proteínas recombinantes na levedura *P. pastoris*. Ele contém o promotor da enzima AOX1 que permite a expressão do gene de interesse de maneira regulada e induzida por metanol. Apresenta como marcador de seleção o gene de resistência a zeocina e o sinal de secreção do fator α que direciona a proteína produzida para o meio extracelular. Além disso apresenta a His-Tag e o epítipo c-myc para detecção e purificação da proteína (Figura 11).

Figura 11 – Vetor pPICZαA_HtpXyl.



As enzimas *Xba* I e *Not* I foram utilizadas como estratégia de clonagem e direcionamento do gene de interesse no vetor e a enzima *Pme* I foi utilizada como enzima de restrição de sítio único na etapa de linearização do plasmídeo.

4.1.4.4. Oligonucleotídeos iniciadores

Todos os oligonucleotídeos iniciadores (*primers*) utilizados nesse trabalho estão na Tabela 3. Para amplificação da ORF codificadora da xilanase foram desenhados iniciadores com base na sequência escolhida na biblioteca de cDNA previamente construída pelo nosso grupo de pesquisa e sequenciada através de sequenciamento de última geração. Além disso, para a clonagem foi retirado o peptídeo sinal previamente predito e o códon de parada da sequência. Os nucleotídeos em minúsculo são as sequências das enzimas de restrição utilizadas para direcionar a clonagem.

Tabela 3: Sequências de *primers* utilizados

<i>Primer</i>	Sequência 5' → 3'
β-actina_Ht F	TGGTATCGTGCTGGACTCTG
β-actina_Ht R	CGCAGGATTCCATACCCAAG
pET_HtpXyl F	catatgTCGAGTCAGACAGGGACAGACAAC
pET_HtpXyl R	gaattcTTAACTCACCGTCACATCTGAACTCCC
pPIC_HtpXyl F	AAATATg _{cggccgc} TCGAGTCAGACAGGGACAGACAAC
pPIC_HtpXyl R	GCA _{tctaga} GCACTCACCGTCACATCTGAACTCCC

Elaborado pela autora.

4.2. Métodos

4.2.1. Análises *in silico* das sequências

Foram empregadas diferentes ferramentas, tais como: o PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) para predição de estrutura secundária das enzimas; o SignalP-5.0 (<http://www.cbs.dtu.dk/services/SignalP/>) para predição de sequência de peptídeo sinal; o NetPhos 3.1 (<https://services.healthtech.dtu.dk/service.php?NetPhos-3.1>) para predição de possíveis sítios de fosforilação; e os softwares NetNGlyc 1.0 (<https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0>), o NetOglyc 4.0 (<https://services.healthtech.dtu.dk/service.php?NetOGlyc-4.0>) e o GlycoEP

(<https://webs.iiitd.edu.in/raghava/glycoep/submit.html>) para identificação de possíveis sítios de glicosilação na proteína que é objeto de estudo deste projeto. Além disso, foi feita a análise da sequência no ProtParam (<https://web.expasy.org/protparam/>) que permite avaliar diferentes parâmetros físico-químicos da proteína, como por exemplo sua massa molecular. Outro software utilizado foi o InterProScan (<https://www.ebi.ac.uk/interpro/search/sequence/>) para identificação de domínios, sítios ativos e família que a proteína pertence, com base em outras proteínas já depositadas em diferentes bancos de dados.

4.2.2. Extração e quantificação de RNA

Os cupins foram separados por castas em tubos de 1,5 mL com aproximadamente 50 indivíduos cada, que foram pulverizados com nitrogênio líquido. A um tubo contendo um *pool* de cupins operários foi adicionado 1mL de TRIzol® (Invitrogen) e os cupins foram macerados com auxílio de pistilos para extração total do RNA. Após homogeneização, a amostra foi centrifugada 12.000 x g a 4°C por 10 minutos e o sobrenadante foi transferido para um novo tubo contendo 200 µL de clorofórmio (Synth). A solução foi submetida a agitação vigorosa em equipamento tipo vórtex e incubada por 3 minutos em temperatura ambiente e posteriormente foi centrifugada novamente 12.000 x g a 4°C por 15 minutos. O sobrenadante foi transferido para novo tubo contendo 200 µL de clorofórmio, sendo novamente agitado vigorosamente em equipamento tipo vórtex e centrifugado 12.000 x g a 4°C por 15 minutos. Após ser transferido para outro tubo, foram acrescentados 500 µL de isopropanol (Synth) ao sobrenadante para a precipitação do RNA e a amostra foi agitada em equipamento tipo vórtex, sendo posteriormente incubada por 10 minutos a temperatura ambiente. O tubo contendo a amostra foi submetido a centrifugação 12.000 x g 4°C por 10 minutos. O sobrenadante foi descartado e ao precipitado foi acrescentado 1 mL de etanol 75% (v/v água), sendo agitado e centrifugado 7.500 g a 4°C por 5 minutos. O sobrenadante foi descartado e o precipitado foi ressuscitado em água ultrapura (dietilpirocarbonato, DEPC 0,2% v/v). Para determinação de concentração de RNA total, a amostra foi quantificada em espectrofotômetro NanoVue™ (GE Healthcare) e para verificação da integridade do RNA após extração, foi realizada eletroforese em gel de agarose desnaturante 1,2%.

4.2.3. Síntese de cDNA

Após a verificação da integridade, aproximadamente 1µg de RNA da amostra foi utilizada para a síntese do cDNA usando o Kit *SuperScript™ III Reverse Transcriptase* (Thermo Fischer Scientific), seguindo as especificações do fabricante, com modificações. A reação foi feita com 1µg de RNA total, 1 µL de Oligo DTV (100 pmol), 1 µL de *Random Primers* (100mM), 2 µL de dNTP Mix (100mM), 5x cDNA *Synthesis Buffer*, 2 µL de DTT (0,1M) e 1 U/µL de *SuperScript™ III Reverse Transcriptase*, em um volume final de 20 µL. O ciclo utilizado foi de 25°C – 10 min; 50°C – 50 min; 85°C – 5 min; 4°C. A eficiência da síntese foi verificada pela amplificação de um fragmento do gene endógeno β-actina, cujos iniciadores estão descritos na Tabela 1.

4.2.4. Reação em Cadeia da Polimerase (PCR) e purificação do fragmento

A ORF codificante da xilanase foi obtida por meio de PCR utilizando o cDNA sintetizado e os *primers* pPIC_HtpXyl F e pPIC_HtpXyl R (Tabela 1) que contém os sítios de restrição da *Not I* e *Xba I*, respectivamente. A reação foi feita com: 6 µL de cDNA, 10 pmol de *primer* F e R, 5 µL de *Buffer + KCl* (10x), 1 µL de dNTP (10mM), 2,5 µL de MgCl₂ (25mM) e 1 U/µL de Taq DNA Polimerase (Thermo Fischer Scientific), completando com água autoclavada até o volume final de 50 µL. A reação foi colocada em termociclador C1000 *Touch Thermal Cyclers* (Bio-Rad®) com o seguinte ciclo: 95°C – 3 min; 35X (95°C – 30 seg; 68°C – 30 seg; 72°C – 1 min); 72°C – 3 min; 8°C. Para confirmar a amplificação do gene de interesse, foi realizada eletroforese em gel de agarose 1% com brometo de etídio no tampão de corrida Tris-Acetato-EDTA (TAE) 1X; 90 V por 30 minutos. Após a corrida, a visualização foi feita em fotodocumentador Gel Doc™ XR+ (Bio-Rad®).

A partir da amostra em gel de agarose utilizou-se o *PureLink™ Quick Gel Extraction Kit* (Thermo Fisher Scientific) seguindo as recomendações do fabricante. Após a purificação, a amostra foi ressuspensa em 30 µL de água ultrapura tratada e quantificada a fim de delimitar qual a concentração ideal a ser utilizada na montagem do vetor recombinante. A eficiência da purificação foi confirmada por eletroforese em gel de agarose 1% com brometo de etídio, onde foi aplicado 1 µL da amostra purificada.

4.2.5. Preparo de células quimiocompetentes de *E. coli* DH5- α

A linhagem de *E. coli* DH5- α foi previamente inoculada em LB ágar por 24 horas a 37°C. Desta placa, uma única colônia isolada foi inoculada em 30 mL de meio Luria-Bertani (LB – 1,25g extrato de levedura; 2,48g NaCl; 2,5g triptona; 250 mL de água) a 37°C sob agitação constante, 200 rpm, por aproximadamente 2 horas até atingir o valor de densidade óptica (D.O.) entre 0,3_{nm} e 0,4_{nm}. Após atingir a D.O. ideal, o inóculo foi incubado em banho de gelo para parar o crescimento das células, seguido de centrifugação a 4000 rpm, por 7 minutos a 4°C. O sobrenadante foi descartado e o precipitado ressuscitado cuidadosamente em 10,5 mL de CaCl₂ 0,1 M gelado. Todo o volume foi centrifugado a 3500 rpm por 7 minutos. O sobrenadante foi descartado e as células ressuscitadas em 2,4 mL de CaCl₂ 0,1 M e glicerol 50% gelado. Cerca de 100 μ L dessa suspensão celular foi transferida para microtubos de 1,5 mL e imediatamente colocadas em nitrogênio líquido, para posterior armazenamento em freezer a -80°C.

4.2.6. Reação de ligação do inserto ao vetor pGEM[®]-T *Easy*

O produto de PCR purificado foi ligado ao vetor pGEM[®]-T *Easy*, segundo as recomendações do fabricante. A reação de ligação foi feita com 220 ng da xilanasase purificada, 50 ng do pGEM[®]-T *Easy*, 3 μ L do *Buffer* T4 Ligase (10x), 1 μ L da enzima T4 DNA Ligase e água autoclavada até o volume final de 10 μ L. A reação ficou a 4°C, *overnight*. Esse vetor é muito usado para clonar produtos de PCR gerados por polimerases que adicionam uma desoxiadenosina às extremidades do fragmento amplificado. De forma que pGEM[®]-T *Easy* contém extremidades 3'-T no local de inserção que é compatível para os produtos de PCR, otimizando assim a eficiência da ligação (Robles e Doers, 1994).

4.2.7. Transformação em *E. coli* DH5- α

Todo o volume da reação de ligação (10 μ L) foi utilizado para transformar as células de *E. coli* DH5- α quimiocompetentes, pelo método de utilização de cloreto de cálcio (CaCl₂), conforme descrito por Sambrook e Green (2001). O vetor recombinante foi adicionado ao tubo contendo as células competentes, mantendo sempre em gelo. Esse tubo contendo a reação de transformação foi colocado em banho-maria a 42°C por 1 minuto e 30 segundos e

imediatamente colocado no gelo para proporcionar o choque térmico. À reação foi adicionado 800 µL de meio LB e incubadas em câmara incubadora com agitador orbital (*shaker*) em baixa agitação (80 rpm) a 37°C por 45 minutos. As bactérias transformadas foram plaqueadas com auxílio da alça de Drigalski em meio LB ágar contendo 100 µg/µL de ampicilina, 20mg/mL de X-gal (5-bromo-4-cloro-indolyl-β-D-galactopiranosídeo) e 0,4mM de IPTG (isopropil β-D-1-tiogalactopiranosídeo) que é um análogo da lactose, capaz de induzir a transcrição do operon *Lac*. As placas foram incubadas a 37°C, *overnight*.

A seleção dos clones recombinantes acontece através do crescimento das colônias na presença de ampicilina, que é o marcador de seleção presente no pGEM[®]-T *Easy* e por meio do *screening* de colônias brancas e azuis, visto que esse vetor possui o gene que codifica a β-galactosidase e a região de múltipla clonagem é no meio dele. Quando ocorre a produção da β-galactosidase funcional o X-gal é hidrolisado em galactose e 5-bromo-4-cloro-indoxil, que espontaneamente oxida e é convertido em 5,5'-dibromo-4,4'-dicloro-índigo, que possui coloração azul. Caso o gene de interesse seja inserido corretamente no vetor, a expressão da β-galactosidase é interrompida, o X-gal não é clivado e as colônias serão brancas (Padmanabhan, Banerjee, Mandi, & Biology, 2011).

Para confirmação dos clones recombinantes foi realizada PCR de colônias, utilizando os mesmos parâmetros descritos no item 4.2.4. A diferença está no molde utilizado na reação, no caso cada colônia foi inoculada em seu respectivo tubo com o auxílio de palitos autoclavados.

Após confirmação dos clones recombinantes, uma colônia foi cultivada em tubo cônico contendo 5 mL de meio LB e 5 µL de ampicilina (100 µg/µL). Os inóculos foram incubados sob agitação em aparelho tipo *shaker* a 37°C, 200 rpm, *overnight*. Após crescimento, foi realizado a extração do plasmídeo recombinante pGEM_HtpXyl com o kit *Fast-n-Easy Plasmid Mini Prep* (Cellco) segundo o protocolo do fabricante. As amostras foram analisadas em gel de agarose 1% para avaliação da integridade do DNA e a quantificação foi feita em espectrofotômetro.

4.2.8. Sequenciamento

Os dois plasmídeos (pGEM_HtpXyl e pPICZαA_HtpXyl) foram enviados para sequenciamento, de acordo com Sanger e colaboradores (1977) para confirmação da

sequência e da clonagem. Os dados foram comparados com a sequência de xilanase escolhida como objeto deste estudo, obtida no sequenciamento realizado pelo nosso grupo anteriormente.

4.3. Expressão recombinante da xilanase em *P. pastoris*

4.3.1. Subclonagem do fragmento que codifica a xilanase no vetor pPICZαA

O plasmídeo pGEM_HtpXyl foi clivado com as enzimas de restrição *Not* I e *Xba* I para subclonagem no pPICZαA, vetor para expressão de proteína em *P. pastoris*. A reação foi feita com 300 ng do pGEM_HtpXyl, 1 U/μL de *Not* I, 1 U/μL de *Xba* I, 10 μL de Tango Buffer (2x) e água autoclavada, totalizando um volume de 50 μL. A reação ficou a 37°C em banho seco, *overnight*. Vale ressaltar que para a subclonagem, o vetor pPICZαA vazio também foi tratado com as mesmas enzimas e condições. A confirmação da clivagem foi feita por meio de gel de agarose 1% com brometo de etídio no tampão de corrida Tris-Acetato-EDTA (TAE) 1X; 70 V por 30 minutos. Após a corrida, a visualização foi feita em foto documentador Gel Doc™ XR+ (BioRad®). O fragmento da xilanase foi purificado do gel de agarose utilizando o kit *PureLink™ Quick Gel Extraction* (Thermo Fisher Scientific) seguindo as recomendações do fabricante.

A reação de ligação entre o fragmento codificante da xilanase e o vetor pPICZαA foi realizada utilizando 220 ng da xilanase purificada, 50 ng do pPICZαA clivado, 3 μL do *Buffer* T4 Ligase (10X), 1 U/μL da enzima T4 DNA Ligase e água autoclavada até o volume final de 30 μL. Essa reação ficou a 4°C, *overnight*. No dia seguinte, foi feita a transformação de *E. coli* DH5-α quimiocompetentes utilizando o método de choque térmico (Sambrook & Green, 2001).

Neste caso, as colônias recombinantes foram selecionadas em placas contendo meio *Low Salt* (LS – 1g extrato de levedura; 2g triptona; 200 mL de água; pH 7,5) com 25 μg/mL de Zeocina™ (Invitrogen), que é o marcador de seleção presente no pPICZαA. As placas foram incubadas a 30°C, *overnight*. Concentrações > 0,5% de NaCl podem inibir a atividade da zeocina, por esse motivo foi utilizado um meio de cultura com baixa concentração de sal (Cregg & Higgins, 2007).

Para confirmação dos clones recombinantes foi realizada PCR de colônias, conforme descrito no item 4.2.7. Uma colônia foi cultivada em tubo cônico contendo 5 mL de meio LS e 1,25 μL de zeocina (0,25 μg/mL). Os inóculos foram incubados sob agitação em aparelho tipo

shaker a 30°C, 200 rpm, *overnight*. Após esse período, foi feita a obtenção do plasmídeo recombinante pPICZαA_HtpXyl com o kit *Fast-n-Easy Plasmid Mini Prep* (Cellco) segundo o protocolo do fabricante. As amostras foram analisadas em gel de agarose 1% para avaliação da integridade do DNA e a quantificação foi feita em espectrofotômetro.

4.3.2. Preparo de células eletrocompetentes de *P. pastoris* X-33

Em um tubo cônico contendo 5 mL de meio YEPD (2g extrato de levedura; 4g peptona; 180 mL de água; 4g de dextrose em 20 mL de água autoclavada) foi preparado um pré-inóculo com uma quantidade generosa de células de *P. pastoris* X-33 e este foi incubado em aparelho tipo *shaker* (30°C; 250 rpm) por 8 horas. Deste pré-inóculo foi inoculado 40 µL em 50 mL de YEPD líquido que foi incubado sob agitação (250 rpm) a 30°C, *overnight*. No dia seguinte a D.O. foi medida em 600_{nm} até alcançar 1,4 (ideal é 1,0 – 1,5). A cultura foi colocada no gelo por 10 minutos e centrifugada 1500 x G, 4°C, 10 minutos. Posteriormente o sobrenadante foi descartado. O precipitado foi ressuscitado com uma solução contendo 10 mL de YEPD e 2 mL HEPES 1,0M, pH 7,0 em movimentos circulares. Foi adicionado 250 µL de DTT 1,0 M e incubado a 30°C, por 15 minutos, sem agitação. Foi acrescido água autoclavada gelada até o volume de 50 mL e centrifugado a 1500 x G, a 4°C por 10 minutos. O sobrenadante foi descartado. O precipitado foi ressuscitado em 25 mL de água autoclavada gelada e centrifugado novamente sob as mesmas condições anteriores, descartando o sobrenadante. O precipitado foi ressuscitado em 2 mL de sorbitol 1,0 M gelado e a amostra foi centrifugada a 1500 x G, a 4°C por 10 minutos. Por fim, o precipitado foi ressuscitado em 100 µL de sorbitol 1,0 M gelado e foram feitas alíquotas de 40 µL para serem usadas na transformação (Cregg & Higgins, 2007).

4.3.3. Transformação em *P. pastoris* X-33

Para a transformação o plasmídeo pPICZαA_HtpXyl foi linearizado com *Pme* I. A reação de linearização foi feita com 2 µg de DNA, 2U de enzima *Pme* I, 5 µL de Tampão B (10x) e água até um volume final de 50 µL. A reação foi deixada a 37°C, *overnight*. A confirmação da linearização foi feita por meio de gel de agarose 1% com brometo de etídio no tampão de corrida Tris-Acetato-EDTA (TAE) 1X; 90 V por 30 minutos. Após a corrida, a visualização foi

feita em foto documentador Gel Doc™ XR+ (BioRad®). A construção foi precipitada com 1/10 do volume de reação de 3 M de acetato de sódio pH 5,2 e 130 µL de etanol puro gelado. A solução foi incubada durante 20 minutos a -80°C e depois centrifugada a 13000 rpm durante 20 minutos. Em seguida, o sobrenadante foi descartado, o precipitado foi lavado com 500 µL de etanol 80% gelado e centrifugado novamente. Após a evaporação do etanol, a amostra foi ressuspensa em 10 µL de água Milli-Q autoclavada.

Todo o DNA linearizado (10 µL) foi misturado em 40 µL de células eletrocompetentes e posteriormente transferidos para as cubetas de eletroporação de 0,2 cm, que foram mantidas no gelo por 5 minutos. As células foram transformadas por eletroporação utilizando o equipamento *GenePulser Unit* (Bio-Rad®). A eletroporação ocorreu com 1,5 KV, 25 µF, 200 Ω. As células transformadas foram ressuspensas em 1 mL de sorbitol 1,0 M, transferidas delicadamente para um tubo cônico e incubadas a 30°C por 2 horas, sem agitação.

As colônias recombinantes foram selecionadas em placas de meio YEPD com sorbitol (YEPDS - 2g extrato de levedura; 4g peptona; 36,44g de sorbitol; 4g de ágar; 180 mL de água; 4g de dextrose em 20 mL de água autoclavada) contendo crescentes concentrações de zeocina (100 µg/mL, 250 µg/mL e 500 µg/mL). As placas foram incubadas a 30°C por 48 horas.

O uso da seleção de colônias com maior resistência à zeocina é um meio eficiente para identificar cepas recombinantes contendo múltiplas cópias do plasmídeo. Isso porque essa resistência geralmente também corresponde a múltiplas cópias do gene heterólogo e há a possibilidade de se obter níveis mais altos de expressão da proteína recombinante (Cregg & Higgins, 2007).

4.3.4. Expressão em pequena escala da xilanase em *P. pastoris* X-33

As colônias que cresceram na placa de 500 µg/mL foram usadas para *screening* em placa de 24 poços. Inicialmente, cada colônia foi inoculada em um poço contendo 3 mL de meio BMGY (63 mL de base contendo 0,9g extrato de levedura e 1,8g peptona; 9 mL tampão fosfato 1,0M, pH 6,0; 9 mL YNB 10x; 9 mL glicerol 10%; 180 µL biotina 0,02%) com 25 µg/mL de zeocina. A placa foi incubada por 48 horas, 30°C a 250 rpm. Após esse tempo, as culturas foram centrifugadas na placa e o sobrenadante foi descartado. Em seguida, o precipitado de cada poço foi ressuspensa com 2 mL de meio BMMY (48 mL de base contendo 0,6g extrato de levedura e 1,2g peptona; 6 mL tampão fosfato 1,0M, pH 6,0; 6 mL YNB 10x; 120 µL biotina

0,02%; 300 µL de metanol absoluto filtrado) com 25 µg/mL de zeocina. Foi retirado uma alíquota de 200 µL de cada um dos poços (amostra não-induzida) e aos mesmos foi adicionado 13,5 µL de metanol absoluto (concentração final - 0,75%) para indução do promotor AOX1 (Cregg & Higgins, 2007).

A placa foi incubada a 30°C por 144h a 250 rpm, e a cada 24 horas, uma alíquota de 200 µL foi retirada de cada poço e diariamente uma quantidade de metanol era adicionado ao meio para manutenção da concentração de 0,75%, como é possível observar na Tabela 4. As alíquotas retiradas diariamente foram centrifugadas a 1500 × g por 5 minutos, e o sobrenadante foi utilizado para análise da expressão proteica por SDS-PAGE 12% (Laemmli, 1970).

Tabela 4: *Screening* da indução da expressão da xilanase em meio com metanol

Tempo de Indução	0h	24h	48h	72h	96h	120h	144h
Meio BMMY	1800 µL	1600 µL	1400 µL	1200 µL	1000 µL	800 µL	600 µL
Metanol (0,75%)	13,5 µL	12 µL	10,5 µL	9 µL	7,5 µL	6 µL	4,5 µL

Elaborado pela autora.

4.3.5. Expressão da xilanase em *P. pastoris* X-33 em maior escala

Uma colônia recombinante que expressou proteína no *screening* foi utilizada em uma nova expressão em maior escala. Para isso a colônia foi inoculada em um tubo cônico contendo 10 mL de meio BMGY (7 mL de base contendo 0,1g extrato de levedura e 0,2g peptona; 1 mL tampão fosfato 1,0M, pH 6,0; 1 mL YNB 10x; 1 mL glicerol 10%; 20 µL biotina 0,02%) e 10 µL de zeocina (100mg/mL) por 24 horas, 30°C a 250 rpm. Posteriormente, 10 mL desse inóculo foi transferido para 500 mL de meio BMGY (350 mL de base contendo 5g extrato de levedura e 10g peptona; 50 mL tampão fosfato 1,0M, pH 6,0; 50 mL YNB 10x; 50 mL glicerol 10%; 1 mL biotina 0,02%) e incubado por 16 horas, 30°C a 250 rpm.

A cultura foi centrifugada por 5 minutos, 24°C a 1500 x G. Foi retirado uma alíquota de 200 µL após centrifugação (amostra não-induzida), o sobrenadante foi descartado e as células foram ressuspendidas em um erlenmeyer contendo 100 mL de meio BMMY (79 mL de base contendo 1g extrato de levedura e 2g peptona; 10 mL tampão fosfato 1,0M, pH 6,0; 10 mL YNB 10x; 1 mL de metanol absoluto filtrado e 200 µL biotina 0,02%). Foi adicionado à cultura

750 µL de metanol absoluto filtrado para indução da expressão e ela foi incubada por 24 horas a 30°C, 250 rpm. Após esse período a cultura foi centrifugada por 10 minutos, 4°C a 6000 rpm. O sobrenadante foi filtrado em membrana de PVDF 0,45 µm (Millipore) e utilizado para análise da expressão proteica por SDS-PAGE 12% (Laemmli, 1970).

4.3.6. Purificação por cromatografia de afinidade

Para purificação foi utilizada a coluna *Ni-NTA Superflow* (QIAGEN) que contém íons de níquel ligados à resina de agarose. A intenção é que a proteína recombinante fique retida na coluna devido à afinidade que a His-tag apresenta aos íons de níquel. A purificação da proteína expressa em *P. pastoris* foi feita a partir do sobrenadante coletado da expressão em maior escala. A coluna foi equilibrada com um tampão de lise contendo 10 mM de imidazol (Tris-HCl 50 mM, NaCl 300 mM, pH 8,0).

Em seguida, a amostra previamente filtrada foi adicionada à coluna contendo 3 mL de resina, e o que passou por gravidade, ou seja, aquilo que não apresentou afinidade aos íons presentes na coluna foi coletado e denominado eluato. Em seguida, foi feita uma lavagem com 3 volumes do tampão de lise sem imidazol (9 mL) para retirar as proteínas que não estavam ligadas à coluna.

Posteriormente, iniciou-se a passagem de frações crescentes de imidazol em duplicata para eluição da proteína recombinante (10 mM; 25 mM; 50 mM; 75 mM; 100 mM; 250 mM; 500 mM; 750 mM e 1M) para eluição de proteínas. Apenas a fração de 1 M foi passada em triplicata. O imidazol funciona como um competidor, visto que ele também apresenta afinidade pelos íons de níquel que estão presentes na resina, fazendo com que a proteína recombinante tenha uma menor afinidade e seja eluída completamente. Alíquotas de todas as etapas foram retiradas para análise por SDS-PAGE a 12% (Laemmli, 1970).

4.4. Expressão recombinante da xilanase em *E. coli*

4.4.1. Subclonagem do fragmento que codifica a xilanase no vetor pET28a

O fragmento da xilanase foi amplificado por PCR utilizando os *primers* pET_HtpXyl F e pET_HtpXyl R (Tabela 1) que contém os sítios de restrição da *Nde* I e *EcoR* I, respectivamente. A reação foi feita utilizando como molde 10 ng do plasmídeo pGEM_HtpXyl, 10 pmol de *primer*

F e R, 5 µL de *Buffer* + KCl (10x), 1 µL de dNTP (10mM), 2,5 µL de MgCl₂ (25mM) e 1 U/µL de Taq DNA Polimerase (Thermo Fischer Scientific), completando com água autoclavada até o volume final de 50 µL. A reação foi colocada em termociclador C1000 *Touch Thermal Cycler* (Bio-Rad®) com o seguinte ciclo: 95°C – 3 min; 35X (95°C – 30 seg; 65°C – 30 seg; 72°C – 1 min); 72°C – 3 min; 8°C. Para confirmar a amplificação do gene de interesse, foi realizada eletroforese em gel de agarose 1% com brometo de etídio. A partir da amostra em gel de agarose utilizou-se o *PureLink™ Quick Gel Extraction Kit* (Thermo Fisher Scientific) seguindo as recomendações do fabricante. Após a purificação, a amostra foi ressuspensa em 30 µL de água ultrapura tratada e quantificada a fim de delimitar qual a concentração ideal a ser utilizada na montagem do vetor recombinante. A eficiência da purificação foi confirmada por eletroforese em gel de agarose 1% com brometo de etídio, onde foi aplicado 1 µL da amostra purificada.

O vetor pET28a vazio foi clivado com *Nde* I e *EcoR* I em uma reação que contou com 900 ng do pET28a vazio, 1 U/µL de *Nde* I, 1 U/µL de *EcoR* I, 5 µL de tampão O (10x) e água autoclavada, totalizando um volume de 50 µL. A reação ficou a 37°C em banho seco, *overnight*. A confirmação da clivagem foi feita por meio de gel de agarose 1% com brometo de etídio no tampão de corrida Tris-Acetato-EDTA (TAE) 1X; 90 V por 30 minutos. Após a corrida, a visualização foi feita em foto documentador Gel Doc™ XR+ (Bio-Rad®).

A reação de ligação entre o fragmento codificante da xilanase e o vetor pET28a foi realizada utilizando 300 ng da xilanase purificada, 60 ng do pET28a clivado, 3 µL do *Buffer* T4 Ligase (10X), 1 U/µL da enzima T4 DNA Ligase e água autoclavada até o volume final de 30 µL. Essa reação ficou a 4°C, *overnight*. O fragmento da xilanase com extremidades *Nde* I e *EcoR* I foi subclonado no vetor pGEM_T *Easy* conforme o item 4.2.6 e o plasmídeo pET28a_HtpXyl foi transformado e propagado em *E. coli* DH5α, conforme o item 4.2.7.

4.4.2. Transformação da xilanase em *E. coli* Rosetta (DE3)

Foi utilizado 10 ng do plasmídeo pET28a_HtpXyl para transformar células de *E. coli* Rosetta (DE3) quimiocompetentes através o método de choque térmico e cloreto de cálcio. O vetor recombinante foi adicionado ao tubo contendo as células competentes, mantendo sempre em gelo. Esse tubo contendo a reação de transformação foi colocado em banho-maria a 42°C por 1 minuto e 30 segundos e imediatamente colocado no gelo para proporcionar o choque

térmico. À reação foi adicionado 800 µL de meio LB e incubadas em aparelho tipo *shaker* em baixa agitação (80 rpm) a 37°C por 45 minutos. As bactérias transformadas foram plaqueadas com auxílio da alça de Drigalski em meio LB ágar contendo 25 mg/mL de canamicina e 25 mg/mL de cloranfenicol. As placas foram incubadas a 37°C, *overnight* (Sambrook & Green, 2001).

Para confirmação dos clones recombinantes foi realizada PCR de colônias, utilizando os mesmos parâmetros descritos no item 4.2.14. A diferença está no molde utilizado na reação, no caso cada colônia foi inoculada em seu respectivo tubo com o auxílio de palitos autoclavados. Após confirmação dos clones recombinantes, uma colônia foi cultivada em tubo cônico contendo 5 mL de meio LB e 25mg/mL de canamicina e de cloranfenicol. Os inóculos foram incubados sob agitação em aparelho tipo *shaker* a 37°C, 200 rpm, *overnight*. Após crescimento, foi realizado a extração do plasmídeo recombinante pET28a_HtpXyl com o kit *Fast-n-Easy Plasmid Mini Prep* (Cellco) segundo o protocolo do fabricante. As amostras foram analisadas em gel de agarose 1% para avaliação da integridade do DNA e a quantificação foi feita em espectrofotômetro.

4.4.3. Expressão recombinante da xilanase em *E. coli* Rosetta (DE3)

Em um tubo cônico contendo 7,5 mL de meio LB contendo 25mg/mL de canamicina e de cloranfenicol foi feito um pré-inóculo a partir de uma colônia recombinante. Ele foi incubado a 37°C, *overnight*. Posteriormente, 5 mL do pré-inóculo foi adicionado a 500 mL de meio LB contendo 25mg/mL de canamicina e de cloranfenicol e essa cultura foi incubada em aparelho tipo *shaker* a 37°C e 200 rpm, até atingirem uma densidade óptica (D.O.) de 0,6, a 600nm. Foi retirada uma alíquota da cultura (amostra não-induzida). Em seguida, foi adicionado à cultura 0,4 mM de IPTG para indução da expressão e esta foi incubada por 16 horas a 37°C, 200 rpm.

Durante esse período foram retiradas alíquotas em um intervalo de 1, 2, 3, 4, 5 e 16 horas de indução. As alíquotas foram centrifugadas a 14000 rpm por 10 minutos e ao fim o sobrenadante foi descartado. Os precipitados foram utilizados para análise da expressão proteica por SDS-PAGE 12% (Laemmli, 1970). A cultura foi centrifugada por 15 minutos a 4°C, 7000 rpm e o sobrenadante foi descartado. As células foram ressuscitadas em tampão de lise (10Mm Tris, 50Mm NaH₂PO₄, 300 mM NaCl, pH 8,0) e sonicadas por 6 minutos, em pulsos

de 59,9 s e intervalos de 30 segundos entre cada pulso. O lisado foi centrifugado por 13000 g, a 4°C por 10 minutos e o sobrenadante foi filtrado em membrana de PVDF 0,45 µm (Millipore). A solubilidade da proteína foi avaliada em gel de poliacrilamida, SDS -PAGE 12% (Laemmli, 1970).

4.4.4. Teste de solubilidade

O teste de solubilidade foi feito por meio da avaliação do sobrenadante e do precipitado obtidos após 4 horas de indução com IPTG, através de SDS-PAGE 12%. A proteína é considerada solúvel quando presente apenas no sobrenadante. Caso esteja parcialmente solúvel, será encontrada tanto no sobrenadante quanto no precipitado. Da mesma forma, se a proteína estiver presa em corpos de inclusão (insolúvel), ela será encontrada somente no precipitado.

4.4.5. Purificação da xilanase por cromatografia de afinidade

Para a purificação foi utilizada a coluna *Ni-NTA Superflow* (QIAGEN) que contém íons de níquel ligados à resina de agarose. A intenção é que a proteína recombinante fique retida na coluna devido à afinidade que a His-tag apresenta aos íons de níquel. A purificação da proteína expressa em *E. coli* foi feita a partir do lisado celular filtrado, tendo em vista que a expressão nesse sistema é intracelular.

Em seguida, a amostra previamente filtrada foi adicionada à coluna contendo 3 mL de resina, e o que passou por gravidade, ou seja, aquilo que não apresentou afinidade aos íons presentes na coluna foi coletado e denominado eluato. Em seguida, foi feita uma lavagem com 3 volumes do tampão de lise sem imidazol (9 mL) para retirar as proteínas que não estavam ligadas à coluna.

Posteriormente, iniciou-se a passagem de frações crescentes de imidazol em duplicata para eluição da proteína recombinante (10 mM; 25 mM; 50 mM; 75 mM; 100 mM; 250 mM; 500 mM; 750 mM e 1M) para eluição de proteínas. Apenas a fração de 1 M foi passada em triplicata. O imidazol funciona como um competidor, visto que ele também apresenta afinidade pelos íons de níquel que estão presentes na resina, fazendo com que a proteína recombinante tenha uma menor afinidade e seja eluída completamente. Alíquotas de todas as etapas foram retiradas para análise por SDS-PAGE a 12% (Laemmli, 1970).

4.4.6. Teste de atividade xilanólítica em ágar xilana

A atividade enzimática da xilanase produzida nos dois sistemas de expressão foi avaliada em placas contendo ágar xilana (1% xilana de bétula; 2% ágar bacteriológico). Para isso, 15 µL de extrato proteico (~ 5 µg/µL) e diluições seriadas (1:2; 1:4; 1:8) foram adicionados a cada placa. O extrato proteico expresso em *P. pastoris* foi concentrado (10X) devido à baixa expressão observada no SDS-PAGE como será visto nos resultados apresentados.

Os controles negativos para a proteína expressa em *E. coli* foram: meio LB; células de Rosetta (DE3) sem plasmídeo pET28a_HtpXyl (NT) e a amostra antes da indução (NI) com IPTG. Quanto à expressão proteica em *P. pastoris*, os controles negativos foram: o meio BMMY, células X-33 sem o plasmídeo pPICZαA_HtpXyl (NT), e a amostra antes da indução (NI) com metanol. As placas foram incubadas por 24 h em duas temperaturas diferentes: 30°C e 40°C. Após esse período, o meio foi corado com vermelho do Congo 0,2% por 15 minutos e depois lavado com NaCl 1M por 15 minutos. O aparecimento de zonas mais claras (halos) indica a presença de atividade xilanólítica (Kasana, Salwan, Dhar, Dutt, & Gulati, 2008; Yoon et al., 2007).

5. Resultados e Discussão

5.1. Análises *in silico* da sequência

A partir do meta-transcriptoma de cupins desenvolvido pelo nosso grupo de pesquisa foi possível a identificação de diferentes enzimas que vem sendo empregadas em diversos trabalhos em nosso laboratório. Este projeto especificamente tem o foco em uma endo-β-1,4 xilanase. Importante ressaltar que o nosso grupo de pesquisa trabalha com três espécies de cupins neotropicais, sendo uma delas o *H. tenuis*. Para essa espécie há dois meta-transcriptomas montados, sendo um construído apenas com a casta operária e outro montado somente com os soldados (Campanini et al., 2021).

Sendo assim, foi feita uma busca nos meta-transcriptomas dessa espécie a fim de encontrar transcritos codificantes dessa enzima. A análise retornou 6 transcritos de cupins operários e 4 transcritos de soldados. Destes, apenas 2 transcritos de operários e 1 de soldado estavam completos, com ATG inicial e o códon de parada. Dentre as duas sequências de

operários, foi escolhida aquela que apresentou os dois sítios ativos, necessários para que a proteína tenha atividade xilanolítica (Takita et al., 2019; Vardakou et al., 2008).

Essa parte do trabalho foi realizado em parceria com o Prof. Dr. Caio César de Melo Freire e o grupo de Bioinformática do Laboratório de Biologia Computacional da Universidade Federal de São Carlos (UFSCar). A sequência escolhida possui 609 nucleotídeos, correspondendo a 202 resíduos de aminoácidos (Figura 12).

Figura 12 - Sequência de nucleotídeos e aminoácidos da β -1,4-endo-xilanase de *H. tenuis*.

```

atggtgcttgacgctagcggggttggtgagttcaggggtgctgagctcgagtcagaca
M V L A A L A G L V S S G C L S S S Q T
gggacagacaacgggtactactactcgttctggacggacggaggaggatcaataagtttc
G T D N G Y Y Y S F W T D G G G S I S F
tgcctggagagtgagggtcggtagcgtcgcagtgagcaatgtaggaaactggtagga
C L E S G G R Y T S Q W S N V G N W V G
gggaaggggtggtcgcaggggagtgcaagggcgtaacatactcagggacgttcaactcg
G K G W S T G S A K A V T Y S G T F N S
ccagggaacgggtacttgcgctgtatgggtggacgaagagcccgtggtggagtattac
P G N G Y L S L Y G W T K S P L V E Y Y
atagtggacagctggggagctaccgtccgaggaggagaaggggaatcaggggacagtg
I V D S W G S Y R P P G G E G N Q G T V
acggcagatggagcgcgtacgacatctacaagcgcagcagtgtaaccagccgagtatc
T A D G A T Y D I Y K T Q R V N Q P S I
gaggggacagcaacattctatcagttctggagtgtgcgtcaggcgaagcgggtaggagga
E G T A T F Y Q F W S V R Q A K R V G G
acgataacgacggggactcatttcacggcgtggcggggaaggggatgaatttggatcg
T I T T G T H F T A W A G K G M N L G S
cagcactacatgatcctggcgtcggaggggtaccagagcagtgggagttcagatgtgacg
H D Y M I L A S E G Y Q S S G S S D V T
gtgagttga
V S -

```

A sequência da xilanase foi comparada com outras sequências depositadas em bancos de dados utilizando a ferramenta BLASTn (*Nacional Center for Biotechnology Information*, NCBI) e curiosamente os cinco primeiros alinhamentos foram com sequências do protista *Holomastigotoides mirabile* (Figura 13).

Figura 13 – Resultado do alinhamento da xilanase obtido após análise BLASTn. Os cinco alinhamentos com maior similaridade foram com genes do protista *H. mirabile*, comumente encontrado na microbiota intestinal de cupins inferiores.

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input type="checkbox"/>	Holomastigotoides mirabile CfXyn2-2 mRNA for endo-beta-1.4-xylanase, complete cds	Holomastigotoides mirabile	647	647	99%	0.0	87.66%	792	AB469375.1
<input type="checkbox"/>	Holomastigotoides mirabile CfXyn2-1 mRNA for endo-beta-1.4-xylanase, complete cds	Holomastigotoides mirabile	641	641	99%	2e-179	87.48%	780	AB469374.1
<input type="checkbox"/>	Holomastigotoides mirabile CfXyn3-1 mRNA for endo-beta-1.4-xylanase, complete cds	Holomastigotoides mirabile	614	614	97%	4e-171	87.02%	757	AB469376.1
<input type="checkbox"/>	Holomastigotoides mirabile CfXyn1-2 mRNA for endo-beta-1.4-xylanase, complete cds	Holomastigotoides mirabile	608	608	97%	2e-169	86.97%	718	AB469373.1
<input type="checkbox"/>	Holomastigotoides mirabile CfXyn1-1 mRNA for endo-beta-1.4-xylanase, complete cds	Holomastigotoides mirabile	592	592	97%	2e-164	86.42%	712	AB469372.1
<input type="checkbox"/>	Cellulomonas flavigena xyIIB gene for xylanase XIIB	Cellulomonas flavigena	122	122	63%	3e-23	73.26%	1380	AM182259.2
<input type="checkbox"/>	Cellulomonas sp. PSBB021, complete genome	Cellulomonas sp. PSBB021	111	111	63%	7e-20	72.70%	3794193	CP021430.1
<input type="checkbox"/>	Streptomyces sp. RTd22, complete genome	Streptomyces sp. RTd22	108	108	69%	9e-19	72.41%	11142275	CP015726.1

Os cupins são divididos basicamente em duas categorias: cupins inferiores, que apresentam em protozoários em sua microbiota intestinal e os cupins superiores, caracterizado por espécies independentes, devido à ausência desses flagelados. O *H. tenuis* é uma espécie de cupim inferior e, portanto, é esperado que ele possua alguma interdependência com protistas relacionada à degradação de lignocelulose (Campanini et al., 2021; Scharf, 2015).

O alinhamento com maior identidade mostrou 87,66% de similaridade com a sequência denominada CfXyn2-2 do protista *H. mirabile* (NCBI Number: AB469375.1) que foi encontrada no intestino do cupim *Coptotermes formosanus* (Figura 14), que é um cupim inferior subterrâneo da família Rhinotermitidae, assim como o *H. tenuis*.

Figura 14 – Alinhamento da sequência CfXyn2-2 pela análise no BLAST contra a sequência da xilanase. Nota-se identidade de aproximadamente 88% da com a sequência CfXyn2-2 de *H. mirabile*. O *Query* indica a sequência da xilanase e *Subject*, a sequência do protozoário. O alinhamento se inicia no par de base 1 na sequência do meta-transcriptoma e no par de base 56 da sequência CfXyn2-2.

Holomastigotoides mirabile CfXyn2-2 mRNA for endo-beta-1,4-xylanase, complete cds

Sequence ID: [AB469375.1](#) Length: 792 Number of Matches: 1

Range 1: 56 to 612 [GenBank](#) [Graphics](#)

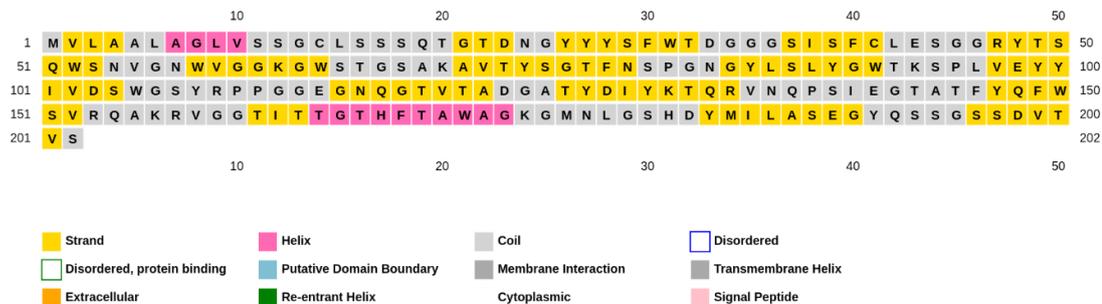
[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Identities	Gaps	Strand
647 bits(350)	0.0	490/559(88%)	4/559(0%)	Plus/Plus
Query 1	TCGAGTCAGACAGGGACAGACAACGGGTACTACTACTCGTTCTGGACGGACGGAGGAGGA	60		
Sbjct 56	TCAAGTCAAACGGGACGGACAACGGGTATTACTTCTCGTTCTGGACGGATGGAGGAGGA	115		
Query 61	TCAATAAGTTTTCGCCTGGAGAGTGGAGGTCGGTACACGTCGCAGTGGAGCAATGTAGGG	120		
Sbjct 116	TCAATAAGCTACTGCAACGAGGCGGAGGTCGGTACACGTCGCAGTGGAGCAATGTGGGG	175		
Query 121	AACTGGGTAGGAGGGAAGGGTGGTCGACGGGGAGTGCGAAGGCGGTAACTACTCAGGG	180		
Sbjct 176	AACTGGGTAGGTGGGAAGGGATGGTCGACGGGGAGTGCGAAGGCGGTGACGACTCAGGG	235		
Query 181	ACGTTCAACTCGCCAGGGAACGGGTACTTGTGCTGTATGGGTGGACGAAGAGCCCGCTG	240		
Sbjct 236	ACGTTCAAGTTCGCCGGGAACGGGTACTTGTGCTGTATGGGTGGACGAAGAGTCCGTTG	295		
Query 241	GTTGGAGTATTACATAGTGGACAGCTGGGGGAGCTACCGTCCGCCAGGAGGAGAAGGGAAT	300		
Sbjct 296	GTTGGAGTATTACATAGTGGATAGCTGGGGAACGTACCGTCCGCCGGGCGGAGAAGGGTTC	355		
Query 301	CAGGGGACAGTGACG-GCAGATGGAGCGACGTACGACATCTACAAGACGCAGCGAGTGAA	359		
Sbjct 356	CAGGGGACGGTGTGAGC-GACGGAGGGACGTACGACATCTACAAGACGCAGCGGGTGAA	414		
Query 360	CCAGCCGAGTATCGAGGGGACAGCAACATTCTATCAGTTCTGGAGTGTGCGTCAGGCGAA	419		
Sbjct 415	CCAGCCGAGCATCGAGGGGACGGCGACGTTCTATCAGTTTTGGAGTGTGCGTCAGGACAA	474		
Query 420	GCGGGTAGGAGGAACGATAACGACGGGGACTCATTTACGGCGTGG-GCGGGGAAGGGGA	478		
Sbjct 475	GCGAGTAGGGGGGACGATAACGACGGGGAACTTTACGGCGTGGAGC-GGGGAAGGGGA	533		
Query 479	TGAATTTGGGATCGCAGACTACATGATCTGGCGTCCGGAGGGGTACCAGAGCAGTGGGA	538		
Sbjct 534	TGAGTTTGGGGTCCGCAACTACATGATAGTGGCGACGGAGGGGTATCAGAGCAGCGGGA	593		
Query 539	GTTTCAGATGTGACGGTGAG	557		
Sbjct 594	GTTTCGGATATAACGGTGAG	612		

recombinante pode ser atrapalhada. Por esse motivo, optou-se por retirar essa sequência da xilanase em estudo, de modo que após essa remoção a sequência ficou com 558 nucleotídeos e 186 aminoácidos (Johnson, Powis, & High, 2013).

Foi feita a predição da estrutura secundária utilizando o programa PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) (Figura 16), mostrando um predomínio de β -folhas e *random coils*, além de duas α -hélices. Além disso, utilizando a ferramenta ProtParam (<https://web.expasy.org/protparam/>) foi predito que a xilanase possui massa molecular de aproximadamente 20 kDa.

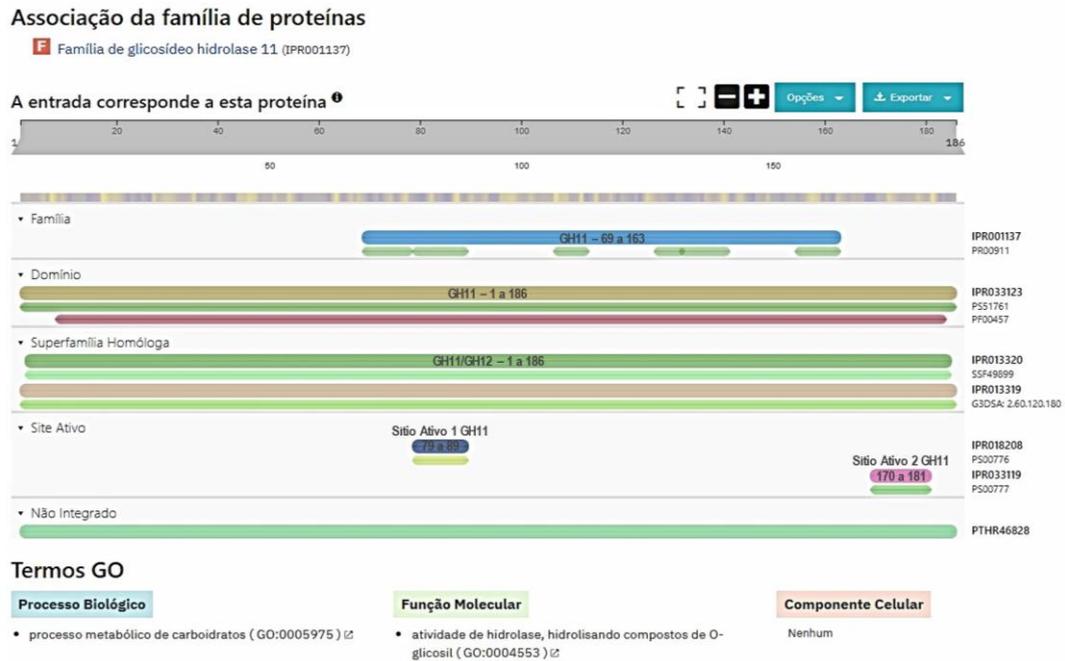
Figura 16 – Predição da estrutura secundária da xilanase. Observa-se a predominância de folhas β na estrutura da xilanase e *random coils*, além de duas α -hélices.



A sequência de aminoácidos foi avaliada no InterProScan (<https://www.ebi.ac.uk/interpro/search/sequence/>) para identificação de domínios e sítios ativos, com base em outras proteínas já depositadas em diferentes bancos de dados (Figura 17).

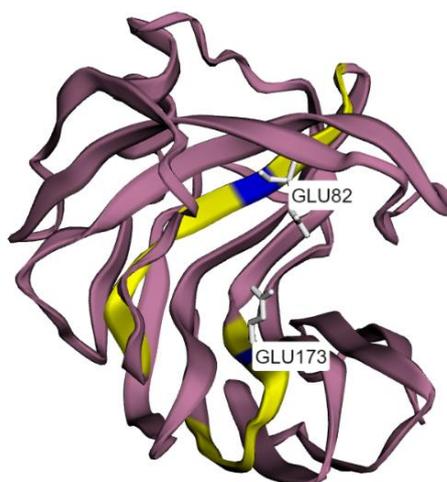
A xilanase deste trabalho pertence à família das GH11 e sua sequência apresenta o domínio dessa família (Figura 17). O grupo das GH11 compreende apenas enzimas com atividade de xilanase, são as ditas xilanases verdadeiras, que anteriormente eram conhecidas como família G das celulases. O domínio GH11 se dobra em forma de *jelly-roll*, similar a uma mão segurando um sanduíche. Além disso, a xilanase estudada apresenta os dois sítios ativos das GH11 que são reconhecidos como necessários para que a enzima apresente atividade e segundo a análise ela está envolvida na degradação de carboidratos, apresentando atividade de hidrolase, conforme esperado (Arakawa et al., 2009; Paës et al., 2012).

Figura 17 – Resultado da avaliação da xilanase utilizando o software InterProScan. Foi constatado que a xilanase pertence à família das GH11 e apresenta o domínio e os dois sítios ativos desse grupo de enzimas.



Também foi feita uma predição 3D utilizando o programa Phyre2 e o software EzMol (Figura 18), onde a imagem é obtida com base na estrutura terciária de proteínas semelhantes já depositadas em bancos de dados públicos (Kelley, Mezulis, Yates, Wass, & Sternberg, 2015).

Figura 18 – Predição da estrutura terciária da xilanase. A estrutura da xilanase é mostrada em rosa. Em amarelo são evidenciados os dois sítios ativos dessa proteína e em azul são destacados os dois resíduos de ácidos glutâmicos dispostos em direções opostas, nas posições 82 e 173 da sequência.

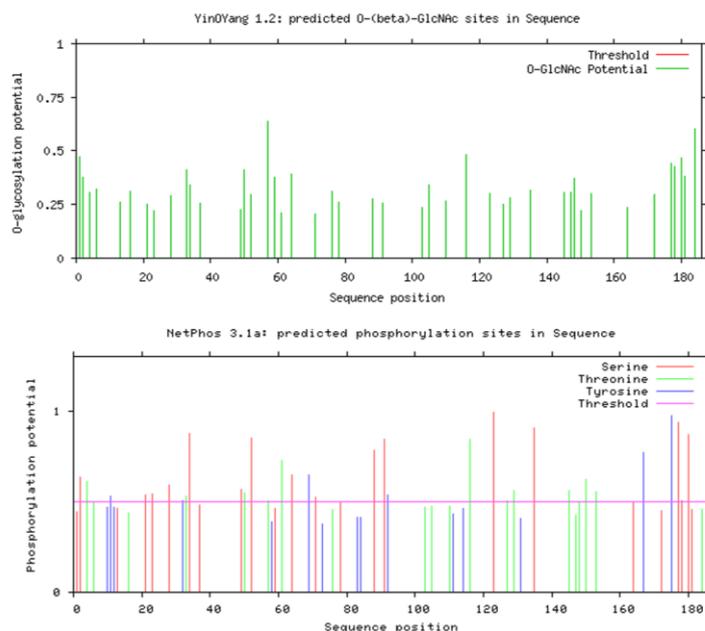


Nas xilanases da família GH11 os sítios ativos são caracterizados pela presença de dois resíduos de ácidos glutâmicos (GLU) catalíticos que se posicionam em lados opostos do sítio do sulco de degradação do substrato. A reação de hidrólise ocorre por um mecanismo de duplo deslocamento, onde um resíduo GLU atua como catalisador ácido/base e o outro age como nucleófilo. A importância dos dois resíduos de ácidos glutâmicos para atividade xilanolítica foi investigado primeiramente por Ko e colaboradores (1992), através de mutagênese sítio-dirigida nos resíduos GLU dos sítios ativos da xilanase de *Bacillus pumilus* o que resultou uma redução drástica da atividade dessa enzima (Arakawa et al., 2009; Ko et al., 1992; Kumar et al., 2017).

Foi feita a predição de possíveis sítios de glicosilação e fosforilação da xilanase, conforme a Figura 19. Glicosilação é a adição de monossacarídeos em aminoácidos específicos da proteína recém-formada. Pode ser do tipo N, que geralmente acontece na cadeia lateral de resíduos de asparagina, ou do tipo O, comumente identificadas em resíduos de serina e treonina. É um processamento pós-traducional importante pois auxilia no enovelamento correto da proteína em sua conformação ativa e pode influenciar no estabelecimento de interações com ligantes, receptores ou mesmo outras proteínas (Arey, 2012).

Figura 19 – Predição dos sítios de glicosilação e fosforilação da xilanase. Observa-se a presença de aproximadamente 44 sítios possíveis de glicosilação do tipo O e 28 sítios de fosforilação (acima do *threshold*) na sequência de aminoácidos da xilanase. Não foram encontrados sítios de N-glicosilação na sequência.

SSQTGIDNGYYSFWTDGGSSISFCLSEGGRYTSQSNVGNWVGGKGWSTGSAKAVIYSGIFNSPGNGYLSLYGWIKSPLVEYYIVDSWGSYRPPGEGN
QGIIVADGATYDIYKIQRVNQPSIEGIAIFYQFWSVRQAKRVGGIITIGTHFAWAGKGMNLGSHDYMILASEGYOSSGSSDVTVS



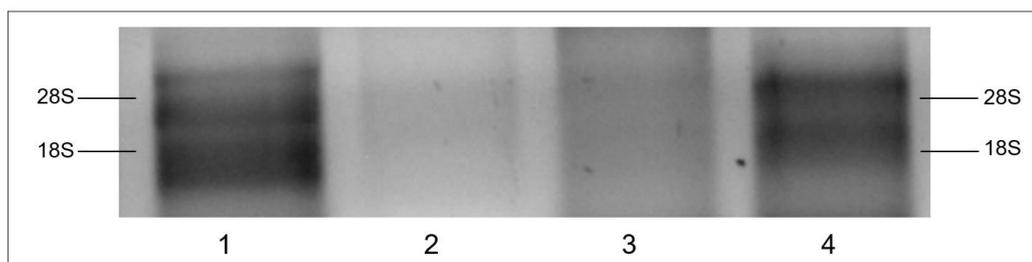
A fosforilação é a adição de um grupo fosfato (PO_4) geralmente em resíduos de serina, treonina e tirosina. Por ser fortemente negativo, a adição desse grupo PO_4 pode causar uma alteração significativa na conformação da proteína, auxiliando nos processos de reconhecimento para interação proteína-proteína e na modulação de diferentes receptores e enzimas (Zhang, 2002).

A escolha do sistema de expressão adequado envolve a análise das características da proteína a ser produzida. No caso, proteínas que apresentem em sua estrutura modificações pós-traducionais, como glicosilações ou fosforilações, na maioria dos casos carecem de um sistema capaz de realizar tais alterações para serem expressas, por exemplo, um sistema eucarioto (Nevalainen & Peterson, 2014).

5.2. Extração de RNA e síntese de cDNA

A extração de RNA foi feita conforme descrito na metodologia. A concentração dos ácidos ribonucleicos foi medida em espectrofotômetro e, para avaliar a integridade, foi feita eletroforese em gel de agarose desnaturante a 1,2% para observação dos fragmentos 18S e 28S do rRNA, sobre o qual infere-se a integridade das outras frações de RNA (Figura 20). A concentração ($\text{ng}/\mu\text{L}$) do RNA foi feita em espectrofotômetro NanoVue™ (GE Healthcare) e o grau de pureza foi avaliado através da razão $260_{\text{nm}}/280_{\text{nm}}$ que é indicada pelo próprio aparelho e o ideal é que seja menor que 2,0. Caso essa razão esteja fora dessa margem, pode indicar a presença de contaminantes como fenol, proteínas ou outros. No caso, as amostras de RNA extraídas apresentaram uma boa qualidade, com razão 260/280 entre 1,81 - 1,86 (Green & Sambrook, 2012).

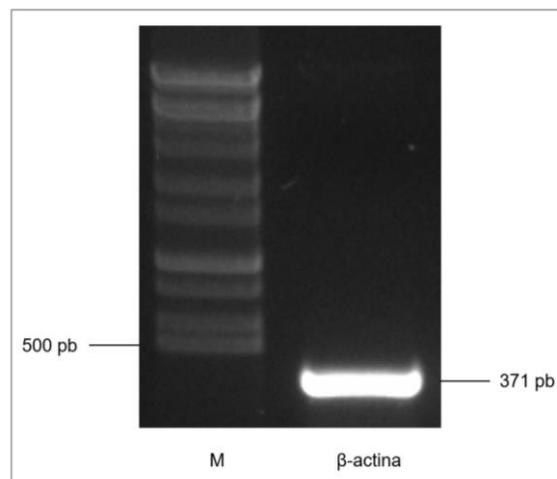
Figura 20 - Avaliação da integridade do RNA extraído por gel desnaturante. Nota-se a presença das porções 28S e 18S íntegras nas amostras 1 e 4, e estas foram utilizadas para síntese de cDNA.



Estudos demonstraram que amostras de ácidos ribonucleicos de insetos podem apresentar uma “quebra oculta” do fragmento 28S em dois fragmentos de tamanhos parecidos que migram juntamente com a porção 18S e não devem ser interpretados como degradação, trata-se apenas de uma particularidade dos RNA de insetos (Macharia, Ombura, & Aroko, 2015; Winnebeck, Millar, & Warman, 2010).

Os RNAs que apresentaram melhor concentração e integridade foram transcritos. Para avaliar a eficiência da transcrição o cDNA foi testado pela amplificação de um fragmento da β -actina, que é um gene constitutivo (Figura 21).

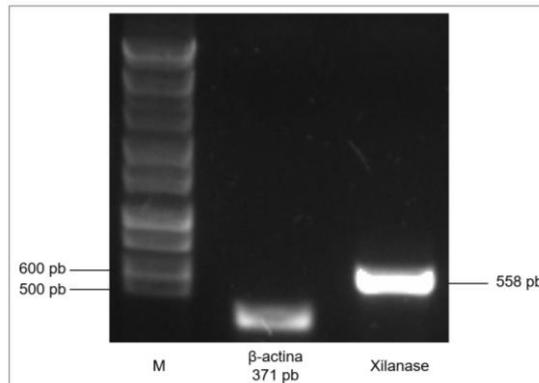
Figura 21 – Amplificação do fragmento da β -actina. Amplificação de um fragmento da β -actina (371 pb) com amostra de cDNA sintetizada, demonstrando a eficiência da síntese. M: marcador de peso molecular *High Range DNA Ladder* (Cellco).



5.3. Amplificação do fragmento que codifica a xilanase

Para amplificação da xilanase o cDNA sintetizado foi utilizado como molde da reação de PCR, conforme descrito no item 4.2.2. Foi gerado um fragmento único de aproximadamente 558 pb (Figura 22), conforme o esperado. Esse fragmento foi purificado para posterior clonagem.

Figura 22 – Amplificação do fragmento que codifica a xilanase. M: marcador de peso molecular *High Range DNA Ladder* (Cellco). A amplificação do fragmento da β -actina foi utilizada como um controle positivo da reação de PCR.

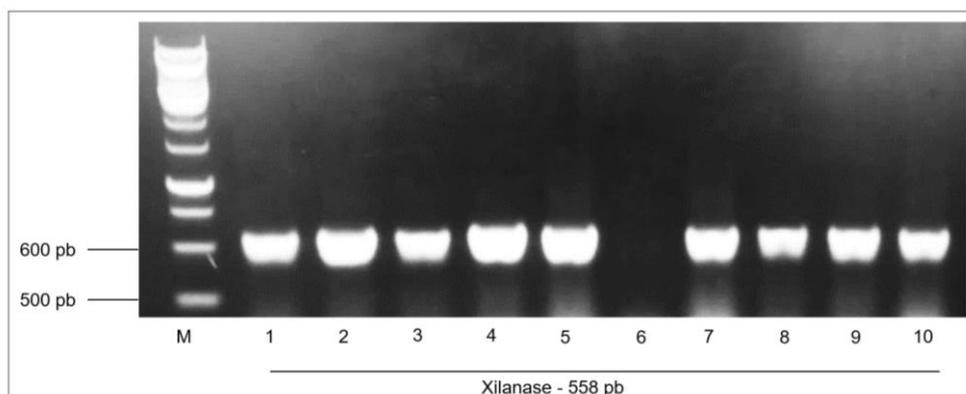


5.4. Clonagem do fragmento codificante da xilanase no vetor pGEM-T *Easy*

O fragmento purificado da xilanase foi ligado ao pGEM-T *Easy*. Foi feita a transformação de células de *E. coli* DH5- α e o *screening* de colônias brancas/azuis, onde as colônias brancas (que provavelmente possuem o plasmídeo pGEM-T_HtpXyl) foram avaliadas por PCR de colônia. O resultado mostrou apenas uma colônia negativa (Figura 23). Foi feita a extração do DNA plasmidial da colônia recombinante número 4.

É possível realizar o *screening* de colônias brancas/azuis com o vetor pGEM-T *Easy* pois o mesmo apresenta a região de clonagem localizada dentro da região codificadora da β -galactosidase, permitindo a seleção dos clones recombinantes em placas contendo X-gal e IPTG, além da ampicilina, que é o marcador de seleção presente nesse plasmídeo (Robles & Doers, 1994).

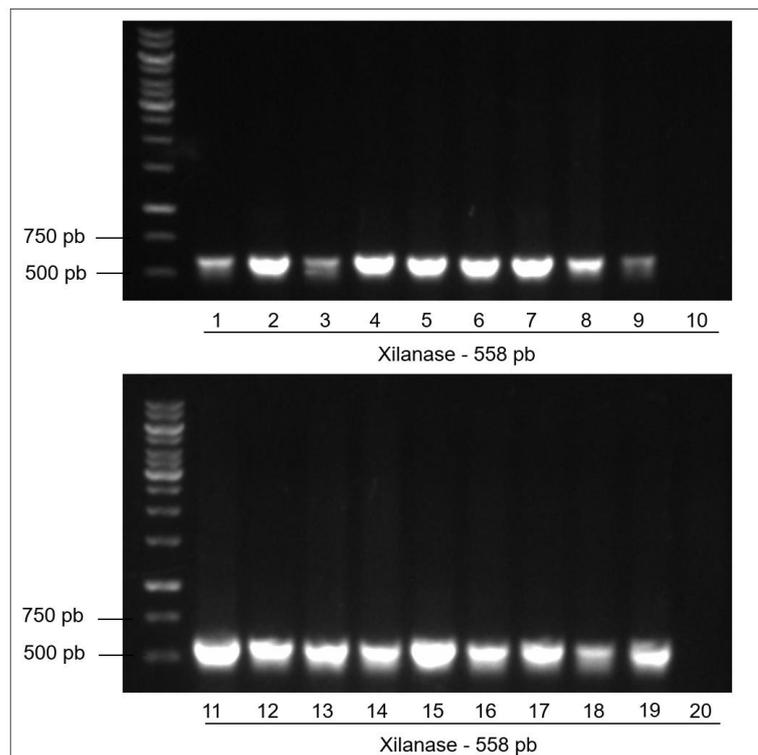
Figura 23 – PCR das colônias transformantes (pGEM_HtpXyl). M: marcador de peso molecular *High Range DNA Ladder* (Cellco); 1 a 10: colônias transformantes avaliadas, onde apenas a colônia 6 não apresentou amplificação.



5.5. Subclonagem do fragmento que codifica a xilanase no vetor pPICZ α A

O plasmídeo pGEM-T_HtpXyl e o vetor pPICZ α A foram digeridos com as enzimas de restrição *Not* I e *Xba* I. O fragmento codificante da HtpXyl foi subclonado no vetor de expressão pPICZ α A. Esse plasmídeo foi utilizado para transformação de células competentes de *E. coli* DH5- α . As colônias transformantes foram avaliadas por meio de PCR de colônia. Como é possível observar na Figura 24, apenas duas colônias apresentaram resultado negativo quanto a presença do fragmento da HtpXyl. Posteriormente foi feita a extração plasmidial da colônia recombinante número 4, conforme descrito na metodologia.

Figura 24 – PCR das colônias transformantes (pPICZ α _HtpXyl). M: marcador de peso molecular *GeneRuler 1kb DNA Ladder* (Thermo Scientific); 1 a 20: colônias transformantes que foram avaliadas, onde observa-se que apenas as colônias 10 e 20 não apresentaram amplificação.

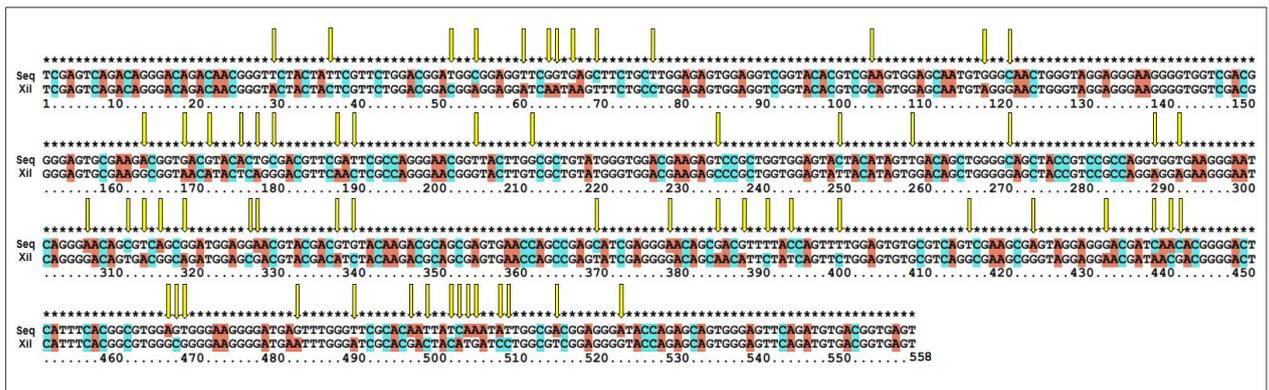


5.6. Análise do sequenciamento

O fragmento purificado da HtpXyl e os dois plasmídeos pGEM-T_HtpXyl e o pPICZ α -HtpXyl foram enviados para sequenciamento (Sanger, Nicklen, & Coulson, 1977). Como evidenciado na Figura 25, a sequência obtida apresentou 66 nucleotídeos diferentes quando

comparada com a sequência de xilanase obtida inicialmente do meta-transcriptoma. Essas bases diferentes resultam na mudança de 19 aminoácidos na sequência da proteína.

Figura 25 – Comparação entre a sequência obtida pelo sequenciamento e a sequência da xilanase inicialmente escolhida do meta-transcriptoma. As setas em amarelo evidenciam as diferenças encontradas entre as duas sequências nucleotídicas mostrando as 66 bases diferentes.



Comparando as duas sequências elas são 87,7% similares e a que foi sequenciada possui 98,18% de similaridade com um Contig437 do meta-transcriptoma. Vale ressaltar que a sequência da xilanase presente nos dois plasmídeos sequenciados são idênticas ao fragmento da xilanase purificada, evidenciando que não houve mutações durante as etapas de clonagem do inserto. Além disso, a obtenção dessa sequência diferente da inicialmente proposta pode ser devido a uma variedade entre os cupins, pelo fato de que o meta-transcriptoma foi construído com base em um *pool* de indivíduos e essa nova sequência apresenta alta semelhança com outra sequência presente nessa construção (Campanini et al., 2021).

A sequência obtida através do sequenciamento foi verificada por BLASTn e constatou-se ser uma xilanase (Figura 26).

Figura 26 – Resultados da sequência da xilanase obtida por sequenciamento após análise BLASTn. Os três primeiros alinhamentos com maior similaridade foram com genes do protista *H. mirabile*, comumente encontrado na microbiota intestinal de cupins inferiores.

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input type="checkbox"/>	Holomastigotoides mirabile CfXyn3-1 mRNA for endo-beta-1.4-xylanase, complete cds	Holomastigotoides mirabile	608	608	97%	2e-169	86.84%	757	AB469376.1
<input type="checkbox"/>	Holomastigotoides mirabile CfXyn1-2 mRNA for endo-beta-1.4-xylanase, complete cds	Holomastigotoides mirabile	597	597	97%	4e-166	86.58%	718	AB469373.1
<input type="checkbox"/>	Aspergillus wentii DTO 134E9 uncharacterized protein (ASPWEDRAFT_39760), mRNA	Aspergillus wentii DTO 134E9	58.4	58.4	9%	9e-04	86.54%	1020	XM_040835174.1
<input type="checkbox"/>	Holomastigotoides mirabile CfXyn1-1 mRNA for endo-beta-1.4-xylanase, complete cds	Holomastigotoides mirabile	547	547	97%	4e-151	84.93%	712	AB469372.1

O alinhamento com maior identidade mostrou 86,84% de similaridade com a sequência denominada CfXyn3-1 do protista *H. mirabile* (NCBI Number: AB469376.1) que foi encontrada no intestino do cupim *Coptotermes formosanus* (Figura 27), cupim inferior semelhante ao *H. tenuis*.

Esse resultado demonstra que a nova sequência obtida provavelmente é desse protista flagelado presente na microbiota intestinal de cupins inferiores, de maneira que haja uma relação de interdependência entre esse simbiote e seu hospedeiro na digestão de materiais lignocelulósicos (Arakawa et al., 2009; Ni & Tokuda, 2013).

Figura 27 – Alinhamento da sequência CfXyn3-1 pela análise no BLAST contra a sequência da xilanase que foi obtida no sequenciamento. Nota-se identidade de aproximadamente 87% da com a sequência CfXyn3-1 de *H. mirabile*. O *Query* indica a sequência da xilanase e *Subject*, a sequência do protozoário. O alinhamento se inicia no par de base 1 na sequência do meta-transcriptoma e no par de base 59 da sequência CfXyn3-1.

Holomastigotoides mirabile CfXyn3-1 mRNA for endo-beta-1,4-xylanase, complete cds

Sequence ID: [AB469376.1](#) Length: 757 Number of Matches: 1

Range 1: 59 to 603 [GenBank](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
608 bits(329)	2e-169	475/547(87%)	4/547(0%)	Plus/Plus
Query 1	TCGAGTCAGACAGGGACAGACAACGGGTTCTACTATTCGTTCTGGACGGATGGCGGAGGT	60		
Sbjct 59	TCAAGTCAGACTGGGACGGATAATGGGTACTACTATTCGTTTGGACGGATGGAGGAGGA	118		
Query 61	TCGGTGAGCTTCTGCTTGGAGAGTGGAGGTCGGTACACGTCGAAGTGGAGCAATGTGGGC	120		
Sbjct 119	TCGATAAGCTTTTGCCTGGAGAGTGGAGGTCGGTACACGTCGCAATGGAGCAATGTAGGG	178		
Query 121	AACTGGGTAGGAGGGAAGGGTGGTGCAGCGGGAGTGCGAAGACGGTGACGTACACTGCG	180		
Sbjct 179	AACTGGGTAGGAGGGAAGGGATGGTCAACGGGAAGTGCGAAGCGGTGACGTATTCAGGG	238		
Query 181	ACGTTTCGATTCGCCA-GGGAACGGTTACTTGGCGCTGTATGGGTGGACGAAGAGTCCGCT	239		
Sbjct 239	ACGTTTC-AGTCCAAACGGCAATGGTTACCTGTCCCTGTATGGGTGGACGAAGAGTCCGTT	297		
Query 240	GGTGGAGTACTACATAGTTGACAGCTGGGGCAGCTACCGTCCGCCAGGTGGTGAAGGGAA	299		
Sbjct 298	GGTGGAGTACTACATAGTAGACAGCTGGGGCAGCTACCGTCCGCCGGGTGGAGAGGGGAA	357		
Query 300	TCAGGGAACAGCGTCAGCGGATGGAGGAACGTACGACGTGTACAAGACGCAGCGAGTGAA	359		
Sbjct 358	TCAGGGAACATTGACAGCGGATGGAGCGACATATGACATCTACAAGACGCAGCGAGTGAA	417		
Query 360	CCAGCCGAGCATCGAGGGAACAGCGACGTTTTTACCAGTTTTGGAGTGTGCGTCAGTCGAA	419		
Sbjct 418	CCAGCCGAGCATAGAGGGGACGGCGACATTTTACCAATTTTGGAGTGTGCGTCAGGCGAA	477		
Query 420	GCGAGTAGGAGGGACGATCAACACGGGACTCATTTACGGCGTGGAGTGGGAAGGGGAT	479		
Sbjct 478	GCGAGTAGGAGGAACGATCACAACGGGACGCACTTTACGGGTGGAGCGGGAAGGGGAT	537		
Query 480	GAGTTTGGGTTTCGCACAATTATCAA-ATATTGGCGACGGAGGGATACCAGAGCAGTGGGA	538		
Sbjct 538	GAGTCTGGGGTTCGCACAACATA-CATGATATTGGCGACGGAGGGCTATCAGAGCAGCGGGA	596		
Query 539	G TTCAGA 545			
Sbjct 597	G TTCAGA 603			

A sequência da xilanase obtida por sequenciamento foi denominada de HtpXyl endo- β -1,4-xilanase-like de protista simbiote não caracterizado de *Heterotermes tenuis* (será chamada daqui em diante de HtpXyl) (N° NCBI de acesso: OL692627).

Todas as análises *in silico* foram refeitas para essa nova sequência, para avaliar se ela continuava sendo uma boa candidata para expressão recombinante, assim como a inicialmente proposta. Como a intenção deste trabalho é a busca por uma xilanase funcional para aplicação biotecnológica, a sequência foi analisada utilizando o InterProScan (Figura 28), com o objetivo principal de identificar se ela apresenta os dois sítios ativos necessários para sua atividade, bem como a predição de sítios de glicosilação e fosforilação (Figura 29) e da estrutura terciária (Figura 30).

Figura 28 – Avaliação da HtpXyl utilizando o software InterProScan. Foi constatado que a HtpXyl pertence à família das GH11 e apresenta o domínio e os dois sítios ativos desse grupo de enzimas, assim como a xilanase inicialmente proposta.

Associação da família de proteínas

F Família de glicosídeo hidrolase 1.1 (IPR001137)

A entrada corresponde a esta proteína ¹



Termos GO

Processo Biológico

- processo metabólico de carboidratos (GO:0005975) [⌵](#)

Função Molecular

- atividade de hidrolase, hidrolisando compostos de O-glicosil (GO:0004553) [⌵](#)

Componente Celular

Nenhum

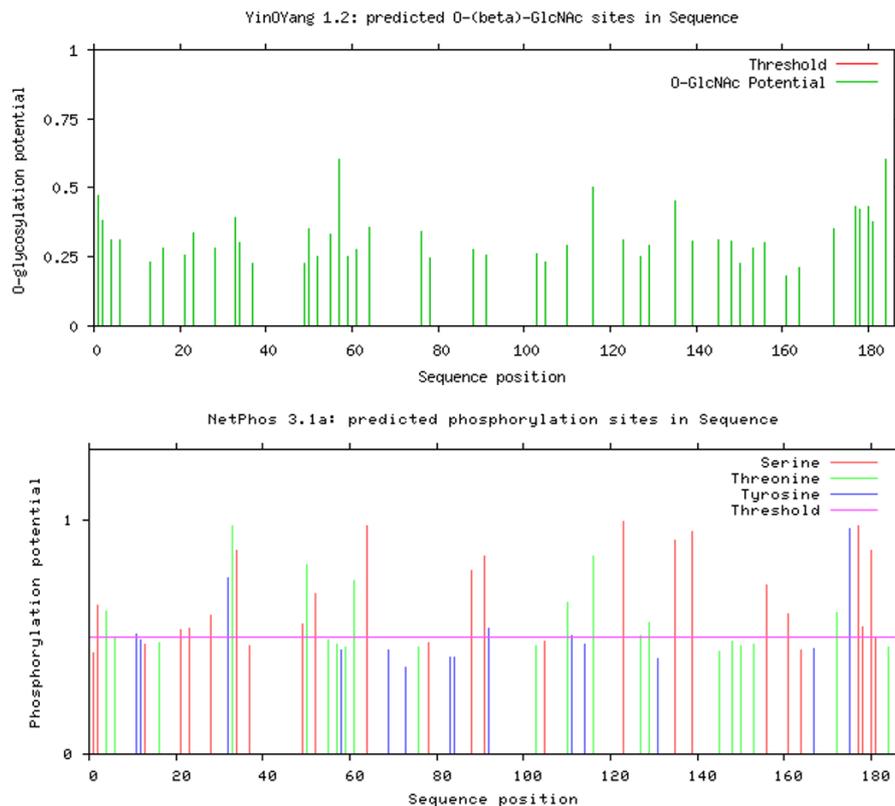
Assim como a xilanase que foi inicialmente proposta, a HtpXyl pertence à família das GH11 e sua sequência apresenta o domínio que configura a estrutura em forma de *jelly-roll* dessa família (Figura 28). Ela também possui os sítios ativos das GH11 necessários para a sua atividade, se consolidando como uma boa candidata para expressão recombinante, tendo em vista o objetivo deste trabalho (Arakawa et al., 2009; Paës et al., 2012).

A Figura 29 mostra a predição dos sítios de glicosilação e fosforilação da HtpXyl que podem ser necessários para que ela apresente sua conformação nativa (Arey, 2012; Nevalainen & Peterson, 2014).

Figura 29 – Predição dos sítios de glicosilação e fosforilação da HtpXyl. Observa-se a presença de aproximadamente 46 sítios possíveis de glicosilação do tipo O e 30 sítios de fosforilação (acima do *threshold*) na sequência de aminoácidos da HtpXyl. Não foram encontrados sítios de N-glicosilação na sequência.

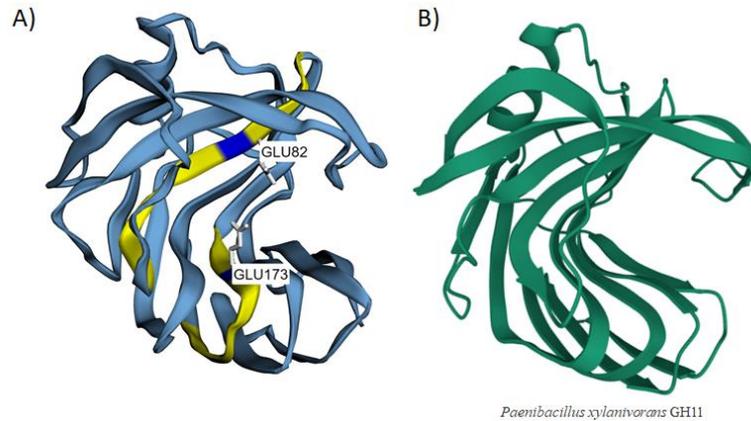
SSQIGTIDNGFYYSFWIDGGGSVSFCLESGGRYTSKWSNVGNWVGGKGWSIGSAKTVIYIATFDSPGNGYLALYGWIKSPLVEYYIVDSWGSYRPPGGEGN

QGIASADGGTYDVYKIQRVNQPSIEGTAIFYQFWSVRQSKRVGGIINIGIHFIAWSGKGMSLGSHNYQILATEGYQSSGSSDDIVS



A estrutura 3D da HtpXyl também foi predita utilizando o programa Phyre2 e o software EzMol (Figura 30). Como a imagem é obtida com base na estrutura terciária de proteínas semelhantes que já foram depositadas em bancos de dados públicos, o software utilizou a mesma estrutura que foi usada para prever a conformação terciária da xilanase inicialmente proposta, evidenciando os dois sítios ativos com a presença dos dois resíduos de ácidos glutâmicos (GLU) posicionados em lados opostos do sulco de degradação do substrato (Kelley et al., 2015).

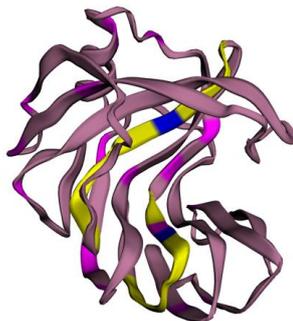
Figura 30 – Predição da estrutura terciária da HtpXyl. Em (A) tem-se a estrutura da xilanase mostrada em azul claro. Em amarelo são evidenciados os dois sítios ativos dessa proteína e em azul escuro são destacados os dois resíduos de ácidos glutâmicos dispostos em direções opostas, nas posições 82 e 173 da sequência. Em (B) é mostrada uma xilanase de *Paenibacillus xylanivorans* (PDB DOI: 10.2210/pdb7KV0/pdb), da família GH11, assim como a HtpXyl. Nota-se que estruturalmente as duas proteínas são similares.



A HtpXyl foi comparada com a xilanase que foi inicialmente proposta a partir do meta-transcriptoma (Figura 31) e apesar de possuírem algumas diferenças na sequência elas continuam apresentando similaridades estruturais.

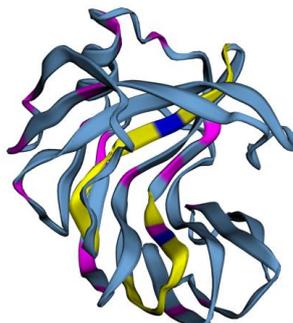
Figura 31 – Comparação entre a xilanase inicialmente proposta e a HtpXyl. Em amarelo são evidenciados os dois sítios ativos das proteínas, sendo que em azul escuro são evidenciados os resíduos de ácidos glutâmicos. Em rosa pink são destacados os aminoácidos que são diferentes nas duas sequências.

A) Xilanase inicialmente obtida do meta-transcriptoma



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
S	S	Q	T	G	T	D	N	G	F	Y	Y	S	F	W	T	D	G	G	G	S	I	S	F	C	L	E	S	G	G	R	Y
33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
T	S	K	W	S	N	V	G	N	W	V	G	G	K	G	W	S	T	G	S	A	K	A	V	T	Y	S	G	T	F	N	S
65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96
P	G	N	G	Y	L	S	L	Y	G	W	T	K	S	P	L	V	E	Y	Y	I	V	D	S	W	G	S	Y	R	P	P	G
97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
G	E	G	N	Q	G	T	V	T	A	D	G	A	T	Y	D	I	Y	K	T	Q	R	V	N	Q	P	S	I	E	G	T	A
129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
T	F	Y	Q	F	W	S	V	R	Q	A	K	R	V	G	G	T	I	T	T	G	T	H	F	A	W	A	G	K	G	M	
161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186						
N	L	G	S	H	D	Y	M	I	L	A	S	E	G	Y	Q	S	S	G	S	S	D	V	T	V	S						

B) Xilanase sequenciada (HtpXyl)



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
S	S	Q	T	G	T	D	N	G	F	Y	Y	S	F	W	T	D	G	G	G	S	I	S	F	C	L	E	S	G	G	R	Y
33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
T	S	K	W	S	N	V	G	N	W	V	G	G	K	G	W	S	T	G	S	A	K	A	V	T	Y	S	G	T	F	N	S
65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96
P	G	N	G	Y	L	A	L	Y	G	W	T	K	S	P	L	V	E	Y	Y	I	V	D	S	W	G	S	Y	R	P	P	G
97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
G	E	G	N	Q	G	T	A	S	A	D	G	G	T	Y	D	V	Y	K	T	Q	R	V	N	Q	P	S	I	E	G	T	A
129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
T	F	Y	Q	F	W	S	V	R	Q	S	K	R	V	G	G	T	I	N	T	T	G	T	H	F	A	W	S	G	K	G	M
161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186						
S	L	G	S	H	N	Y	Q	I	L	A	T	E	G	Y	Q	S	S	G	S	S	D	V	T	V	S						

Conforme visto na Figura 31, observa-se em rosa pink os 19 aminoácidos que são diferentes entre as proteínas. Tendo em vista que mudanças em aminoácidos podem implicar em alterações na proteína, foi realizada a predição de possíveis alterações físico-químicas que podem ser causadas devido a troca desses aminoácidos utilizando o software TreeSAAP, conforme visto na Tabela 5 (Woolley, Johnson, Smith, Crandall, & McClellan, 2003).

Tabela 5: Análise dos aminoácidos diferentes entre as duas xilanases. A predição das propriedades físico-químicas foi feita por meio do software TreeSAAP. Valores de 6 a 8 representam a intensidade da mudança, sendo 8 de maior intensidade; os valores decimais representam o nível da significância. Quanto menor o valor de p , mais significativo é o resultado. Em amarelo destaca-se a mudança que está presente no sítio ativo 2, onde não houve predição de nenhuma alteração.

Posição	Aminoácido na Xilanase inicial	Aminoácido na HtpXyl	Polaridade da cadeia lateral	Propriedades físico-químicas potencialmente alteradas
10	Y (Tyr)	F (Phe)	Polar → Apolar	Sem predições
22	I (Ile)	V (Val)	Apolares	(8, $p = 0.05$) Constante de equilíbrio
35	Q (Gln)	K (Lys)	Polar → Polar básico	(6, $p = 0.05$) Ponto Isoelétrico
55	A (Ala)	T (Thr)	Apolar → Polar	(6, $p = 0.01$) Tendência a α -hélices
59	S (Ser)	T (Thr)	Polares	Sem predições
60	G (Gly)	A (Ala)	Apolares	Sem predições
63	N (Asn)	D (Asp)	Polar → Polar ácido	(6, $p = 0.001$) Carga da região C-terminal
71	S (Ser)	A (Ala)	Polar → Apolar	(6, $p = 0.01$) Tendência a α -hélices (6, $p = 0.05$) Tendência de <i>coils</i> (6, $p = 0.001$) Tendência a torção
104	V (Val)	A (Ala)	Apolares	Sem predições
105	T (Thr)	S (Ser)	Polares	Sem predições
109	A (Ala)	G (Gly)	Apolares	Sem predições
113	I (Ile)	V (Val)	Apolares	(8, $p = 0.05$) Constante de equilíbrio
139	A (Ala)	S (Ser)	Apolar → Polar	(6, $p = 0.01$) Tendência a α -hélices (6, $p = 0.05$) Tendência de <i>coils</i> (6, $p = 0.001$) Tendência a torção
147	T (Thr)	N (Asn)	Polares	Sem predições
156	A (Ala)	S (Ser)	Apolar → Polar	(6, $p = 0.01$) Tendência a α -hélices (6, $p = 0.05$) Tendência de <i>coils</i> (6, $p = 0.001$) Tendência a torção
161	N (Asn)	S (Ser)	Polares	Sem predições
166	D (Asp)	N (Asn)	Polar ácido → Polar	(6, $p = 0.001$) Carga da região C-terminal
168	M (I)	Q (Gln)	Apolar → Polar	Sem predições
172	S (Ser)	T (Thr)	Polares	Sem predições

O software TreeSAAP prediz a influência propriedades estruturais e bioquímicas de aminoácidos com base em diferenças fixas entre resíduos. Todos os aminoácidos possuem uma cadeia lateral característica que define suas propriedades físico-químicas e por consequência, as da proteína que eles integram. De acordo com a polaridade da cadeia lateral eles podem ser agrupados naqueles que apresentam a cadeia lateral apolar (hidrofóbicos) e o grupo com cadeia lateral polar (hidrofílicos), que são subdivididos em polares neutros, com carga positiva (básicos) ou com carga negativa (ácidos) (Taylor, 1986; Woolley et al., 2003).

Certos aminoácidos são mais ou menos prováveis de serem encontrados formando α -hélices ou folhas β pregueadas. Por exemplo, uma cadeia polipeptídica com muitos resíduos de lisina ou arginina próximos provavelmente a formação de uma α -hélice será impedida, devido a carga positiva das suas cadeias laterais, que irão se repelir. De igual modo, comumente são encontrados resíduos de fenilalanina, tirosina ou triptofano em folhas β pregueadas. Além disso, a prolina difere dos demais aminoácidos por apresentar uma cadeia lateral cíclica, que envolve o grupo amino e a impede de formar hélices (Hol, 1985; Smith, Withka, & Regan, 1994).

É possível observar na Tabela 5, destacado em amarelo, a troca de uma Serina (S) por uma Treonina (T) na posição 172 que fica no sítio ativo 2. A partir das análises feitas não houve predição de nenhuma alteração físico-química nesse ponto da proteína. Isso provavelmente se deve ao fato de que as cadeias laterais desses aminoácidos apresentam propriedades similares e ambos são classificados como polares (hidrofílicos).

Como não foi constatado nenhuma mudança de propriedades devido essa substituição no sítio ativo 2 que fosse significativa ou que implicasse problema na atividade xilanolítica, prosseguiu-se com a expressão recombinante da proteína HtpXyl.

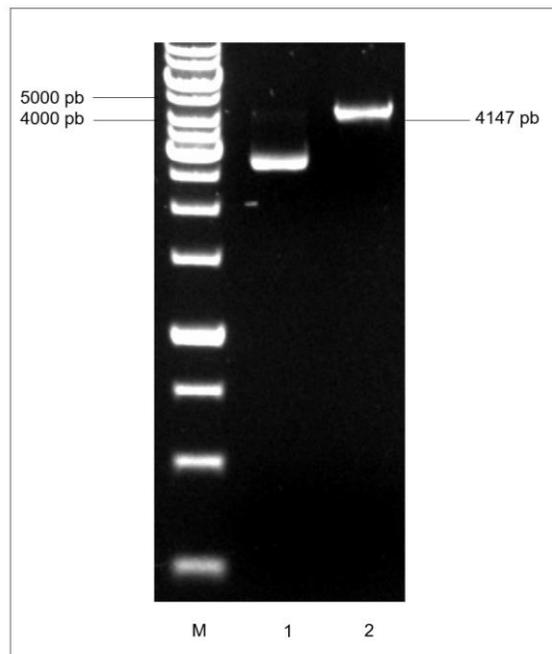
5.7. Expressão recombinante de HtpXyl em *P. pastoris* X-33

Para a transformação o vetor pPICZ α A_HtpXyl foi linearizado com a enzima *Pme* I (Figura 32). Nota-se uma banda no tamanho de 4147 pb, correspondente ao tamanho do plasmídeo pPICZ α A_HtpXyl.

A etapa de linearização é feita em proteínas que serão expressas em *P. pastoris* para facilitar a recombinação homóloga do cassete de expressão no genoma da levedura, visto que

ela apresenta propensão para esse tipo de evento. Dessa forma, utiliza-se uma enzima de restrição de sítio único para linearizar o plasmídeo de maneira que a construção contendo a proteína recombinante fique flanqueada pelas regiões do promotor *AOX1*, presente no vetor de expressão (Cregg et al., 2000; Cregg & Higgins, 2007).

Figura 32 – Linearização do vetor pPICZ α A_HtpXyl. M: marcador de peso molecular *GeneRuler 1kb DNA Ladder* (Thermo Scientific); 1: plasmídeo fechado; 2: vetor linearizado com a enzima *Pme I*, apresentando uma banda com tamanho esperado (4147 pb).



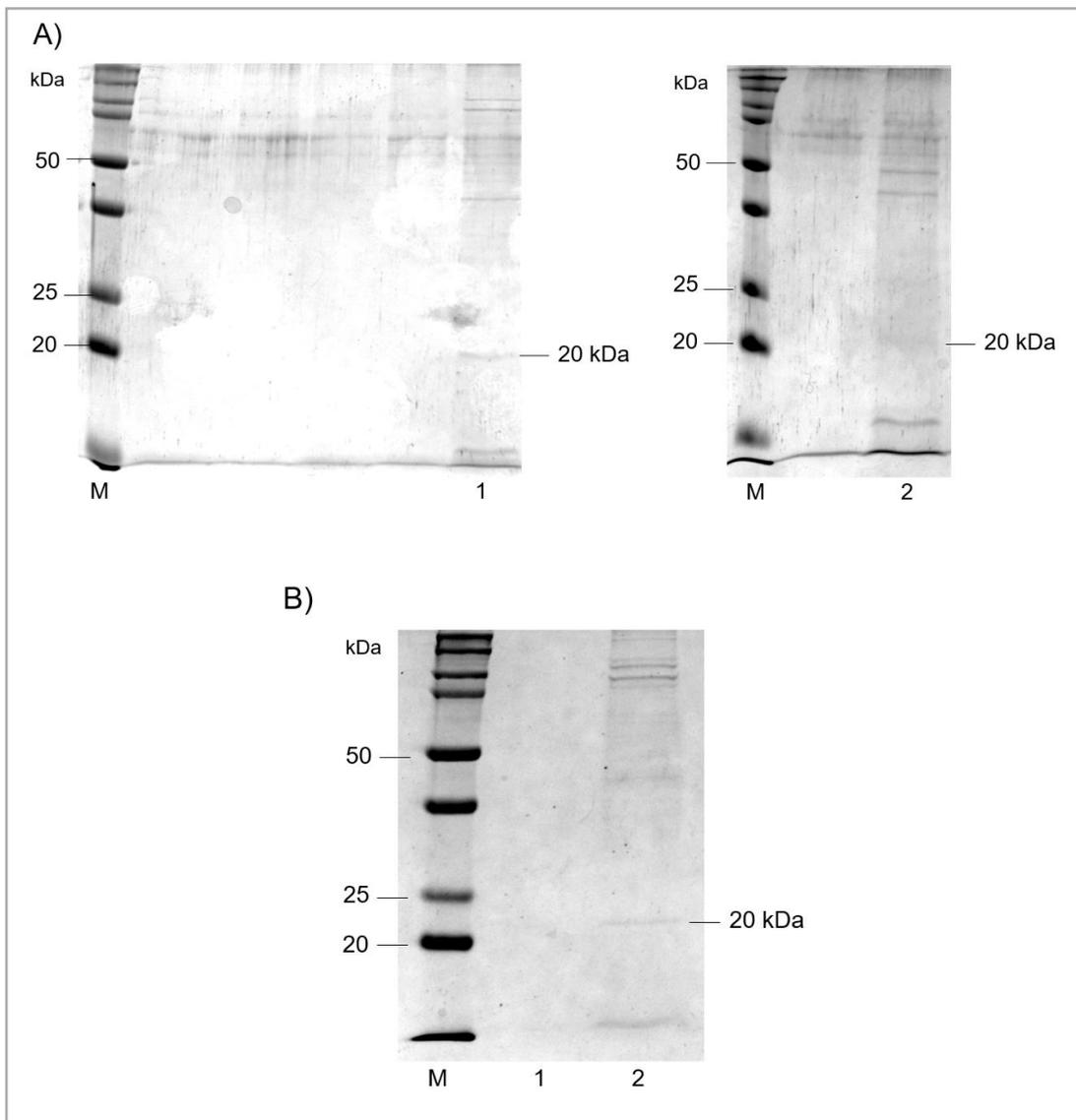
As colônias que cresceram na placa de 500 $\mu\text{g}/\text{mL}$ de zeocina foram usadas para *screening* em placa de 24 poços por 144h em meio induzido metanol. Conforme descrito na metodologia, durante o *screening* foram coletadas alíquotas a cada 24 horas e todos os respectivos sobrenadantes foi analisado por meio de SDS-PAGE 12%. Deste modo, para cada clone havia 6 alíquotas para analisar (24, 48, 72, 96, 120 e 144 horas).

No *screening* obteve-se a expressão da proteína HtpXyl em 24 e 48 horas de indução, conforme verificado em SDS-PAGE 12%, onde observa-se a presença de bandas correspondentes a aproximadamente 20 kDa, conforme predito em análises preliminares (Figura 33A) (Laemmli, 1970).

Esse clone foi utilizado para expressão da HtpXyl em maior escala por 24 horas, que foi considerado o melhor tempo de indução, tendo em vista o resultado obtido durante o *screening*. Nota-se a presença de uma banda no tamanho esperado para a HtpXyl (~20 kDa)

no sobrenadante analisado (Figura 33B), porém a purificação por cromatografia de afinidade em resina contendo íons de níquel não foi bem-sucedida e por esse motivo para os ensaios de atividade foi utilizado o sobrenadante obtido na indução em maior escala.

Figura 33 – Expressão da HtpXyl em *P. pastoris*. Em (A) tem-se o resultado do *screening* em placa de 24 poços. 1: HtpXyl expressa após 24 horas de indução; 2: HtpXyl expressa após 48 horas de indução. Em (B) tem-se o resultado da expressão em maior escala da HtpXyl. 1: amostra não induzida com metanol; 2: HtpXyl expressa em maior escala durante 24 horas de indução. Em ambos os casos se tem M: marcador de massa molecular *1kb DNA Ladder* (Promega). Análise feita em SDS-PAGE 12%.

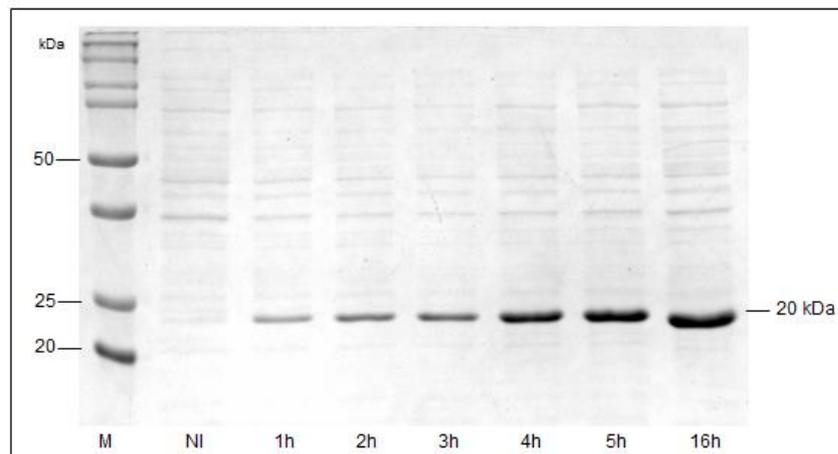


A levedura *P. pastoris* é considerada um sistema de expressão barato e eficaz. É de fácil manipulação, apresenta altos índices de expressão proteica, meios de cultura baratos e a vantagem de secretar proteínas para o meio extracelular (Cregg et al., 2000).

5.8. Expressão recombinante de HtpXyl em *E. coli* Rosetta (DE3)

A expressão em sistema procarioto ocorreu em paralelo a expressão em *P. pastoris*. Para tal, o inserto da HtpXyl foi obtido a partir do plasmídeo pGEM-T_HtpXyl, conforme descrito na metodologia. O cassete de expressão pET28a_HtpXyl foi propagado em *E. coli* DH5 α , transformado em *E. coli* Rosetta (DE3) e induzido em meio LB contendo 0,4Mm de IPTG por 16 horas. O resultado foi avaliado em SDS-PAGE a 12% e pode ser observado na Figura 34.

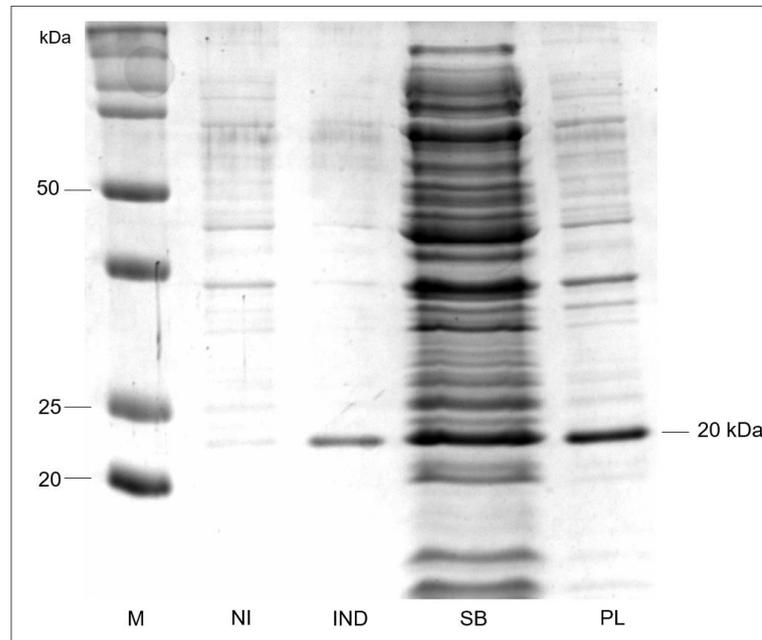
Figura 34 – Expressão da HtpXyl em *E. coli*. O perfil de expressão da HtpXyl em *E. coli* foi traçado durante 16 horas de indução, com a retirada de alíquotas a cada hora. NI: amostra não induzida; posteriormente tem-se as horas de indução em que as alíquotas foram coletadas. M: marcador de massa molecular *1kb DNA Ladder* (Promega). Análise feita em SDS-PAGE 12%.



Observa-se a presença da banda de expressão no tamanho de aproximadamente 20 kDa desde a primeira hora de indução que foi aumentando gradualmente ao passar do tempo, comparando com o controle não induzido. Com base no resultado, foi considerado que o tempo ideal de expressão de 4 horas.

Foi feito o teste de solubilidade da proteína HtpXyl expressa em *E. coli* (Figura 35) e foi constatado que ela foi expressa parcialmente solúvel, ou seja, havia proteína presente tanto no precipitado (corpos de inclusão) quanto no sobrenadante do lisado celular que foi posteriormente filtrado.

Figura 35 – Teste de solubilidade da HtpXyl expressa em *E. coli*. Análise da solubilidade da proteína HtpXyl em tampão de vivo, pH 8,0. M: marcador de massa molecular *1kb DNA Ladder* (Promega); NI: amostra não induzida; IND: amostra induzida (4h de indução); SB: sobrenadante do lisado celular; PL: precipitado. Análise feita em SDS-PAGE 12%. Nota-se que a proteína está presente tanto no sobrenadante quanto no precipitado, caracterizando um perfil parcialmente solúvel.

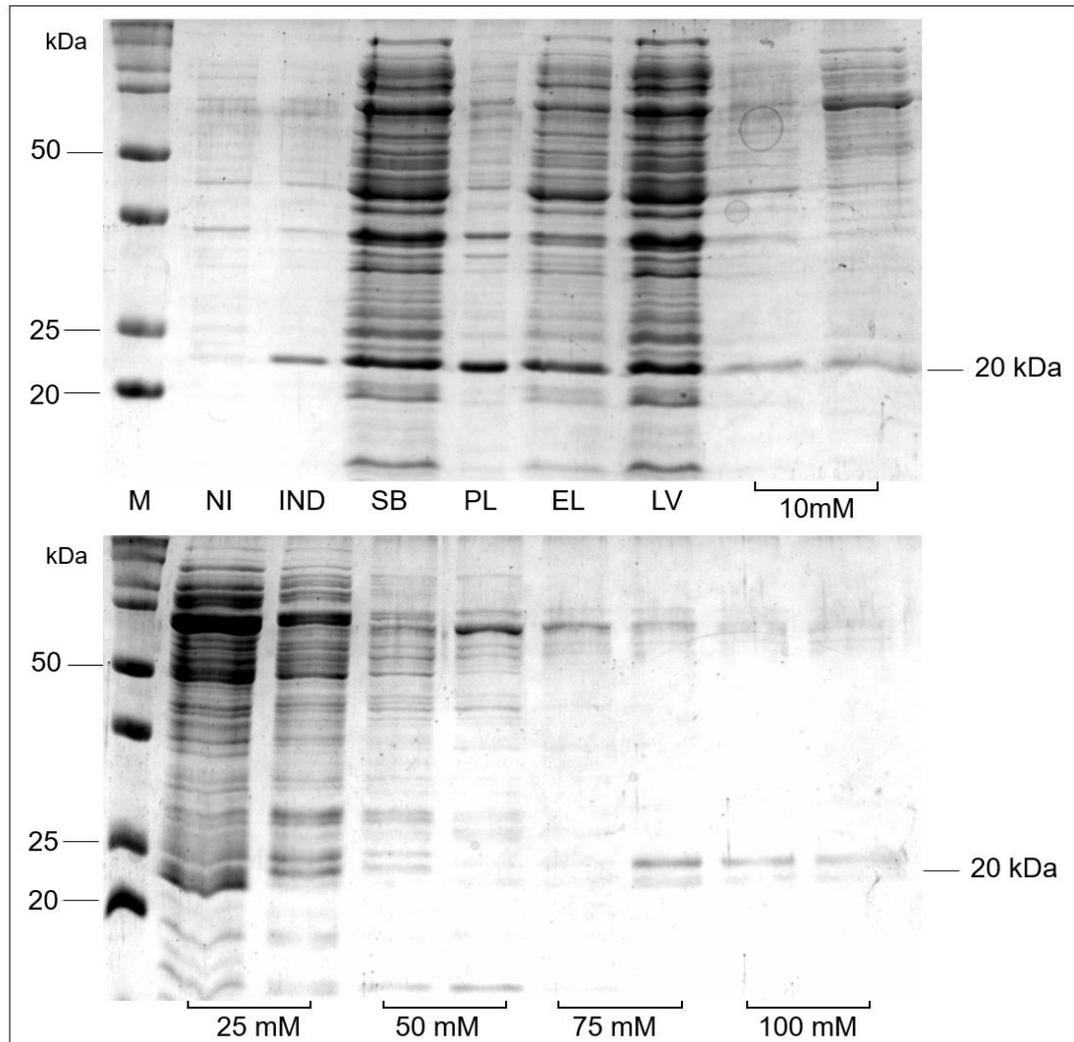


A formação de corpos de inclusão é algo corriqueiro na expressão de proteínas em *E. coli* e pode ser ocasionada por altas taxas de expressão proteica que levam a um dobramento incorreto da proteína. Nesses casos recomenda-se a diminuição da temperatura do experimento ou tentativa de solubilização das mesmas (Kaur, Kumar, & Kaur, 2018; Singh, Upadhyay, Upadhyay, Singh, & Panda, 2015).

O sobrenadante contendo a parte solúvel da HtpXyl foi purificado por cromatografia de afinidade com resina contendo íons de níquel (Figura 36) conforme descrito anteriormente. É possível observar a presença da proteína em todas as frações, sendo que as frações com maior pureza são as de 75mM e 100 mM.

Com base na purificação por cromatografia de afinidade verifica-se que há uma perda de proteína no eluato e na lavagem, ou seja, proteína que não ficou retida na coluna (não apresentou afinidade suficiente pelo níquel). Entretanto, obteve-se proteína parcialmente purificada (observa-se ainda alguns contaminantes) de forma solúvel nas eluições contendo 50mM, 75 mM e 100 Mm de imidazol.

Figura 36 – Purificação da HtpXyl expressa em *E. coli* por cromatografia de afinidade. M: marcador de massa molecular *1kb DNA Ladder* (Promega); NI: amostra não induzida; IND: amostra induzida (4 horas de indução); SB: sobrenadante do lisado celular; PL: precipitado; EL: fração contendo as proteínas que não tiveram afinidade com os íons de níquel presentes na resina (eluato); LV: fração contendo tampão de lise e passado na coluna para retirar o excesso de proteínas que não se ligaram à resina (lavagem); posteriormente são as amostras contendo diferentes concentrações de imidazol (10 mM; 25 mM; 50 mM; 75 mM; 100 mM). Análise feita em SDS-PAGE 12%.



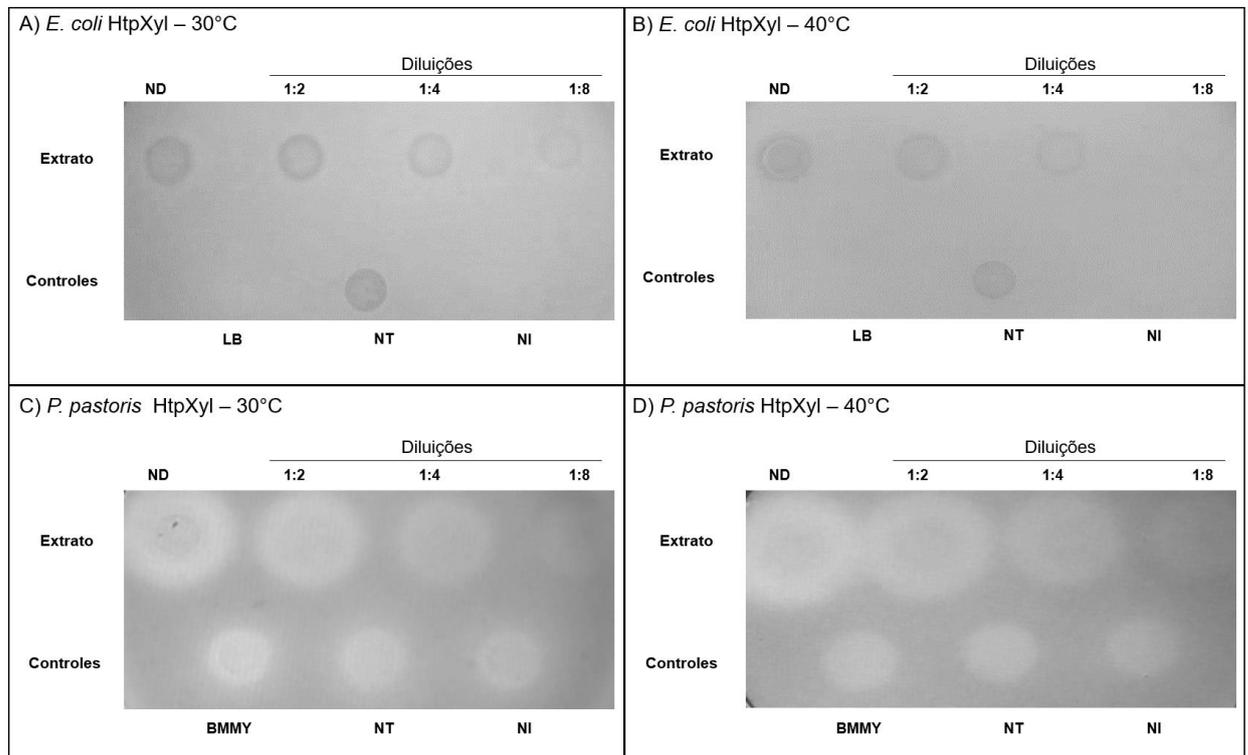
5.9. Teste de atividade xilanolítica em ágar xilana

A atividade de HtpXyl foi avaliada por ensaio qualitativo em placas contendo 1% de ágar xilana. Para isso, foram utilizados os extratos de proteína bruta recombinante (~ 5µg de proteína total) de ambos os sistemas de expressão, ou seja, a amostra da expressão em *E. coli* foi o sobrenadante filtrado do lisado celular e a amostra da expressão em *P. pastoris* foi o sobrenadante obtido na expressão de maior escala.

A formação de halos está relacionada à afinidade do vermelho do Congo por carboidratos, ligando-se fortemente a eles. Caso não haja mais o substrato, devido ação xilanolítica, o corante não se liga e é removido na etapa de lavagem com NaCl 1 M (Kasana et al., 2008; Teather & Wood, 1982).

Após o tratamento com vermelho do Congo 0,2%, foi possível observar a formação de zonas mais claras (halos) na HtpXyl expressa em ambos os sistemas e nas duas temperaturas testadas (Figura 38).

Figura 37 – Avaliação da atividade xilanolítica da HtpXyl em ágar xilana. A atividade enzimática da xilanase produzida nos dois sistemas de expressão foi avaliada em placas contendo ágar xilana. Para isso, 15 µL de extrato protéico (~ 5 µg/µL) e diluições seriadas (1:2; 1:4; 1:8) foram adicionados a cada placa. Controles negativos para a proteína expressa em *E. coli*: meio LB; células de Rosetta (DE3) sem plasmídeo pET28a_HtpXyl (NT) e amostra antes da indução (NI) com IPTG. Quanto à expressão proteica em *P. pastoris*, os controles negativos foram: o meio BMMY, células X-33 sem o plasmídeo pPICZαA_HtpXyl (NT), e a amostra antes da indução (NI) com metanol. Após o período de incubação o meio foi corado com vermelho do Congo 0,2% por 15 minutos e depois lavado com NaCl 1M por 15 minutos. O aparecimento de zonas mais claras (halos) indica a presença de atividade xilanolítica.



A HtpXyl foi expressa em sua forma ativa, degradando o substrato de forma eficiente tanto em *E. coli* quanto em *P. pastoris*. Comparando a atividade entre os dois sistemas, a expressão em levedura mostrou um aumento consistente na degradação da atividade xilana em todas

as diluições testadas e para ambas as temperaturas. (Figura 38C e 38D). Essa degradação mais eficiente do substrato pode ser devido as modificações pós-traducionais que esse sistema eucariótico realiza. A presença de halos nos controles NT E NI pode ter sido causada pela ação de hidrolases endógenas secretadas por *P. pastoris* X-33 ou componentes produzidos por *E. coli* Rosetta DE3. Enquanto a formação de halos no controle contendo apenas BMMY pode ter ocorrido devido à ação indireta dos compostos presentes neste meio complexo, como o extrato de levedura (Cregg et al., 2000).

A expressão heteróloga de proteínas é um dos primeiros passos para testar sua aplicação industrial. Por esta razão, neste estudo foram utilizados dois sistemas de expressão diferentes para a produção de uma xilanase recombinante, com o objetivo futuro de sua aplicação biotecnológica e industrial. Além dos biocombustíveis, essas enzimas podem participar de outros processos industriais como branqueamento da polpa de papel, produção de ração animal para melhorar a digestibilidade, indústria alimentícia no processo de clarificação de bebidas, e obtenção do xilitol, um adoçante de baixo índice glicêmico (Meilany, Anugeraheni, Aziz, Kresnowati, & Setiadi, 2020; Thomas, Ushasree, & Pandey, 2014; Zhu et al., 2012).

A escolha de *E. coli* como sistema de expressão de proteínas recombinantes deve-se ao seu crescimento simples, ciclo de vida curto e facilidade de manipulação genética. No entanto, os sistemas procarióticos são limitados quanto às modificações pós-traducionais necessárias, o que em alguns casos pode resultar em uma proteína sem atividade (Gopal & Kumar, 2013; Kaur et al., 2018; Rosano & Ceccarelli, 2014)

Nesses casos, a expressão da proteína é realizada em outro sistema mais complexo, como o *P. pastoris*. Quando comparado a outros sistemas eucarióticos, a expressão neste sistema é considerada simples, apresenta vetores de expressão bem caracterizados e capacidade de realizar diferentes modificações pós-traducionais, como glicosilação, metilação, processamento proteolítico, entre outras. Além disso, devido ao fator α (presente em alguns vetores), as proteínas podem ser produzidas e secretadas para o meio extracelular fazendo com que processos de purificação sejam opcionais (Ahmad, Hirz, Pichler, & Schwab, 2014; M.-q. Liu, Huo, Xu, & Weng, 2017; Weinacker et al., 2013)

Como visto na Figura 38, a enzima recombinante foi capaz de degradar o substrato nas duas temperaturas utilizadas (30°C e 40°C). Na produção de bioetanol, para diminuir os custos do processo, é desejável que as enzimas utilizadas na etapa de hidrólise atuem em temperaturas próximas às utilizadas para a fermentação das leveduras, geralmente 30°C. Uma alternativa para esse problema seria prospectar enzimas que atuem na faixa de 40°C e sejam utilizadas em associação com leveduras termotolerantes (Prado et al., 2020; Soccol et al., 2019).

Mediante o exposto, tem-se que foi realizada a identificação de uma sequência codificante de xilanase partir de dados do meta-transcriptoma de cupins da espécie *H. tenuis*. Essa sequência posteriormente foi denominada HtpXyl, clonada e expressa em sistema procariótico e eucariótico na sua forma ativa, visto que foi capaz de degradar a xilana em duas temperaturas testadas (30°C e 40°C).

É possível afirmar que a HtpXyl é uma enzima candidata potencialmente aplicável em processos biotecnológicos. Contudo, esses resultados são iniciais e novos estudos deverão ser feitos para otimização dos ensaios de expressão e para a caracterização bioquímica e funcional dessa enzima, a fim de avaliar a fundo essa premissa. Além disso, o fato de que a HtpXyl foi encontrada na microbiota associada a uma espécie de cupim com hábitos noturnos de forrageamento sugere que, provavelmente essa enzima atua em temperaturas mais baixas, uma vez que insetos não são termorreguladores, instaurando a possibilidade de uso desta enzima no processo SSF futuramente, reduzindo os custos de produção do E2G.

6. Conclusão

- i. Uma sequência de uma endo-beta-1,4 xilanase de foi identificada no meta-transcriptoma de cupins da espécie *Heterotermes tenuis*;
- ii. Após as etapas de clonagem essa sequência foi submetida a sequenciamento pelo método de Sanger, onde obteve-se uma sequência com 87,7% de similaridade com a xilanase que foi inicialmente proposta;
- iii. As análises *in silico* dessa sequência demonstraram se tratar de uma possível nova xilanase de um protozoário endossimbionte (*H. mirabile*) já conhecido de *H. tenuis*;

- iv. A sequência foi denominada “HtpXyl endo- β -1,4-xilanase-like de protista simbiote não caracterizado de *Heterotermes tenuis* - HtpXyl” (N° NCBI de acesso: OL692627);
- v. A HtpXyl foi clonada e expressa em dois sistemas heterólogos: na bactéria *Escherichia coli* e na levedura *Pichia pastoris*;
- vi. A proteína expressa em *E. coli* apresentou um perfil parcialmente solúvel e a purificação da proteína por cromatografia de afinidade não foi bem-sucedida em nenhum dos dois sistemas recombinantes;
- vii. O teste qualitativo de atividade utilizando o extrato proteico total foi realizado em ágar xilana e a proteína expressa em ambos os sistemas demonstrou atividade a 30°C e 40°C.

7. Perspectivas

- Otimização dos ensaios de expressão em *E. coli* e *P. pastoris*
- Caracterização bioquímica e funcional da HtpXyl
- Expressão da HtpXyl em *Saccharomyces cerevisiae*
- Expressão e caracterização de outras enzimas lignocelulósicas, como a lacase, por exemplo.

8. Anexos

Em anexo estão duas produções científicas que a autora teve a oportunidade de contribuir.

- Artigo publicado na revista *Insect Molecular Biology* - Campanini, E. B., Pedrino, M., Martins, L. A., Athaide Neta, O. S., Carazzolle, M. F., Ciancaglini, I., ... & da Cunha, A. F. (2021). *Expression profiles of neotropical termites reveal microbiota-associated, caste-biased genes and biotechnological targets*. **Insect Molecular Biology**, 30(2), 152-164.
- Artigo aceito para publicação na revista *Microorganisms – Identification of a new Endo- β -1,4-xylanase prospected from the microbiota of the termite *Heterotermes tenuis*.*

Title:

Expression profiles of neotropical termites reveal microbiota-associated, caste-biased genes and biotechnological targets

Running title:

Meta-transcriptomes of neotropical termites

List of Authors and Affiliations:

Emeline Boni Campanini¹, Matheus Pedrino¹, Leonardo do Amaral Martins¹, Olinda Soares Athaide Neta¹, Marcelo Falsarella Carazzolle², Iara Ciancaglioni¹, Iran Malavazi¹, Ana Maria Costa-Leonardo³, Caio César de Melo Freire¹, Francis Morais Franco Nunes^{1*}, Anderson Ferreira da Cunha¹

¹ Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, SP, Brazil

² Departamento de Genética, Evolução, Microbiologia e Imunologia, UNICAMP, Campinas, SP, Brazil

³ Laboratório de Cupins, Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), Campus de Rio Claro, Rio Claro, SP, Brazil

*** Corresponding author:**

Francis Morais Franco Nunes

Departamento de Genética e Evolução, Universidade Federal de São Carlos, Rod. Washington Luis - km 235, São Carlos-SP, 13565-905, Brazil

Email address: francis.nunes@ufscar.br

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the [Version of Record](#). Please cite this article as doi: [10.1111/imb.12684](https://doi.org/10.1111/imb.12684)

This article is protected by copyright. All rights reserved.

Abstract

Termites are well recognized by their complex development trajectories, involving dynamic differentiation process between non-reproductive castes, workers and soldiers. These insects are associated with endosymbiotic microorganisms, which help in lignocellulose digestion and nitrogen metabolism. Aiming to identify genes harboring biotechnological potential, we analyzed workers and soldiers RNA-Seq data of three neotropical termites: *Heterotermes tenuis* (Isoptera: Rhinotermitidae), *Velocitermes heteropterus* (Isoptera: Termitidae) and *Cornitermes cumulans* (Isoptera: Termitidae). We observed differences in the microbiota associated with each termite family, and found protists' genes in both Termitidae species. We found an opposite pattern of caste-biased gene expression between *H. tenuis* and the termitids studied. Moreover, the two termitids are considerably different concerning the number of differentially expressed genes. Functional annotation indicated considerable differences in caste-biased gene content between *V. heteropterus* and *C. cumulans*, even though they share similar diet and biological niche. Among the most differentially expressed genes, we highlighted those involved in caste differentiation and cellulose digestion, which are attractive targets for studying more efficient technologies for termite control, biomass digestion and other biotechnological applications.

Keywords: Rhinotermitidae; Termitidae; meta-transcriptomes; cellulolytic genes.



Article

Identification of a new Endo- β -1,4-xylanase prospected from the microbiota of the termite *Heterotermes tenuis*

Olinda S. A. Alcobaça ^{1†}, Emeline B. Campanini ^{1†}, Iara Ciancaglini ^{2†}, Sâmara V. Rocha ¹, Iran Malavazi ¹, Caio C. M. Freire ¹, Francis M. F. Nunes ¹, Andrea S. C. Fuentes ¹, and Anderson Ferreira da Cunha ^{1,*}

¹ Department of Genetics and Evolution, Federal University of São Carlos, São Carlos, 13565-905, São Paulo, Brazil

² Brazilian Biorenewables National Laboratory, Brazilian Center of Research in Energy and Materials, Campinas, 13083-100, São Paulo, Brazil

* Correspondence: anderf@ufscar.br;

† These authors contributed equally

Abstract: Xylanases are hemicellulases that break down xylan to soluble pentoses. They are used for industrial purposes, such as paper whitening, beverage clarification, and biofuel production. The second-generation bioethanol production is hindered by the enzymatic hydrolysis step of the lignocellulosic biomass, due to the complex arrangement established among its constituents. Xylanases can potentially increase the production yield by improving the action of the cellulolytic enzyme complex. We prospected endo- β -1,4-xylanases from meta-transcriptomes of the termite *Heterotermes tenuis*. *In silico* structural characterization and functional analysis of an endo- β -1,4-xylanase from a symbiotic protist of *H. tenuis* indicate two active sites and a substrate-binding groove needed for the catalytic activity. No N-glycosylation sites were found. This endo- β -1,4-xylanase was recombinantly expressed in *Pichia pastoris* and *Escherichia coli* cells, presenting a molecular mass of approximately 20 kDa. Enzymatic activity assay using recombinant endo- β -1,4-xylanase was also performed on 1% xylan agar stained with Congo red at 30 °C and 40 °C. The enzyme expressed in both systems was able to hydrolyze the substrate xylan, becoming a promising candidate for further analysis aiming to determine its potential for application in industrial xylan degradation processes.

Keywords: *Heterotermes tenuis*; meta-transcriptome; symbiotic protist; hemicellulase; xylanase; second-generation bioethanol

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Microorganisms* 2022, 10, x. <https://doi.org/10.3390/xxxxx>

Academic Editor: Firstname Lastname

Received: date

Accepted: date

Published: date

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bioethanol is a biofuel with the greatest potential to replace oil-based fuels worldwide. The second-generation bioethanol is distinguished from the first-generation by the use lignocellulosic biomass as raw material in its production. The conversion of lignocellulosic biomass (cellulose, hemicellulose, and lignin) into monomeric sugars is laborious and involves several steps. Enzymatic hydrolysis is one of the most critical steps in this process because of costs and low catalytic efficiency of enzymes at currently used temperatures [1,2]. Therefore, finding new and more efficient enzymes are would reduce production costs and increase the efficiency of second-generation bioethanol.

Termites are one of the most efficient lignocellulose decomposers in nature due to an association with symbiont gut microorganisms. This ability has been investigated towards several biotechnological applications, especially in biofuel production [3,4]. The association results in a dual enzymatic system comprising endogenous and exogenous carbohydrate-active enzymes (CAZymes) [5,6]. *Heterotermes tenuis* (Isoptera: Rhinotermitidae) is a lower species of termites, presenting a gut microbiota composed of flagellate protists and several species of bacteria and archaea [7]. Since workers partake in the initial

9. Referências Bibliográficas

- Achyuthan, K. E., Achyuthan, A. M., Adams, P. D., Dirk, S. M., Harper, J. C., Simmons, B. A., & Singh, A. K. (2010). Supramolecular self-assembled chaos: polyphenolic lignin's barrier to cost-effective lignocellulosic biofuels. *Molecules*, *15*(12), 8641-8688.
- Ahmad, M., Hirz, M., Pichler, H., & Schwab, H. (2014). Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Applied microbiology and biotechnology*, *98*(12), 5301-5317.
- Álvarez, A., Cachero, S., González-Sánchez, C., Montejo-Bernardo, J., Pizarro, C., & Bueno, J. L. (2018). Novel method for holocellulose analysis of non-woody biomass wastes. *Carbohydrate polymers*, *189*, 250-256.
- Arakawa, G., Watanabe, H., Yamasaki, H., Maekawa, H., & Tokuda, G. (2009). Purification and molecular cloning of xylanases from the wood-feeding termite, *Coptotermes formosanus* Shiraki. *Bioscience, biotechnology and biochemistry*, *73*(3), 710-718.
- Arey, B. J. (2012). The role of glycosylation in receptor signaling. *Glycosylation*, *10*, 50262.
- Baruah, J., Nath, B. K., Sharma, R., Kumar, S., Deka, R. C., Baruah, D. C., & Kalita, E. (2018). Recent trends in the pretreatment of lignocellulosic biomass for value-added products. *Frontiers in Energy Research*, *6*, 141.
- Bhatia, S. K., Gurav, R., Choi, T.-R., Han, Y. H., Park, Y.-L., Park, J. Y., . . . Kim, S.-H. J. B. t. (2019). Bioconversion of barley straw lignin into biodiesel using *Rhodococcus* sp. YHY01. *289*, 121704.
- Bibra, M., Kunreddy, V. R., & Sani, R. K. (2018). Thermostable Xylanase Production by *Geobacillus* sp. Strain DUSELR13, and Its Application in Ethanol Production with Lignocellulosic Biomass. *6*(3), 93.
- Bichot, A., Delgenès, J.-P., Méchin, V., Carrère, H., Bernet, N., & García-Bernet, D. (2018). Understanding biomass recalcitrance in grasses for their efficient utilization as biorefinery feedstock. *Reviews in Environmental Science Bio/Technology*, *17*(4), 707-748.
- Bourguignon, T., Lo, N., Šobotník, J., Sillam-Dussès, D., Roisin, Y., & Evans, T. A. (2016). Oceanic dispersal, vicariance and human introduction shaped the modern distribution of the termites *Reticulitermes*, *Heterotermes* and *Coptotermes*. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1827), 20160179.
- Brinkman, M. L., da Cunha, M. P., Heijnen, S., Wicke, B., Guilhoto, J. J., Walter, A., . . . reviews, s. e. (2018). Interregional assessment of socio-economic effects of sugarcane ethanol production in Brazil. *88*, 347-362.
- Brune, A. (2014). Symbiotic digestion of lignocellulose in termite guts. *Nature Reviews Microbiology*, *12*(3), 168-180.
- Campanini, E. B., Pedrino, M., Martins, L., Athaide Neta, O., Carazzolle, M. F., Ciancaglini, I., . . . Nunes, F. J. I. M. B. (2021). Expression profiles of neotropical termites reveal microbiota-associated, caste-biased genes and biotechnological targets. *30*(2), 152-164.
- Cardona, C., & Quintero, J. (2010). Production of bioethanol from sugarcane bagasse: status and perspectives. *Bioresource technology*, *101*(13), 4754-4766.
- Chen, S., Zhang, X., Singh, D., Yu, H., & Yang, X. (2010). Biological pretreatment of lignocellulosics: potential, progress and challenges. *Biofuels*, *1*(1), 177-199.
- Chohan, N. A., Aruwajoye, G., Sewsynker-Sukai, Y., & Kana, E. G. (2020). Valorisation of potato peel wastes for bioethanol production using simultaneous saccharification and

- fermentation: process optimization and kinetic assessment. *Renewable Energy*, 146, 1031-1040.
- Cleveland, L. R. (1924). The physiological and symbiotic relationships between the intestinal protozoa of termites and their host, with special reference to *Reticulitermes flavipes* Kollar. *The Biological Bulletin*, 46(5), 203-227.
- Collins, T., Gerday, C., & Feller, G. (2005). Xylanases, xylanase families and extremophilic xylanases. *FEMS microbiology reviews*, 29(1), 3-23.
- Cregg, J. M., Cereghino, J. L., Shi, J., & Higgins, D. R. J. M. b. (2000). Recombinant protein expression in *Pichia pastoris*. 16(1), 23-52.
- Cregg, J. M., & Higgins, D. R. (2007). *Pichia protocols* (Vol. 268): Springer.
- de Abreu, Y. V., & Nascimento, H. R. (2016). A produção da cana-de-açúcar e de etanol nas novas fronteiras agrícolas: o estado do Tocantins. *Revista Liberato*, 17(27).
- De Bona, P., Deshmukh, L., Gorbatyuk, V., Vinogradova, O., Kendall, D. A. J. P. S., Function., & Bioinformatics. (2012). Structural studies of a signal peptide in complex with signal peptidase I cytoplasmic domain: The stabilizing effect of membrane-mimetics on the acquired fold. 80(3), 807-817.
- de Oliveira Bordonal, R., Carvalho, J. L. N., Lal, R., de Figueiredo, E. B., de Oliveira, B. G., & La Scala, N. J. A. f. S. D. (2018). Sustainability of sugarcane production in Brazil. A review. 38(2), 1-23.
- Dias, E. F., Schlindwein, M. M., da Silva, L. F., & Ruviaro, C. F. (2015). A situação da cadeia produtiva do etanol no Brasil em Mato Grosso do Sul a partir da crise mundial 2008. *Desenvolvimento Socioeconômico em Debate*, 1(1), 112-129.
- Domínguez-Bocanegra, A. R., Torres-Muñoz, J. A., & López, R. A. (2015). Production of bioethanol from agro-industrial wastes. *Fuel*, 149, 85-89.
- Dubendorf, J. W., & Studier, F. W. J. J. o. m. b. (1991). Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. 219(1), 45-59.
- Eaton, T. D., Jones, S. C., & Jenkins, T. M. (2016). Species diversity of Puerto Rican Heteroterms (Dictyoptera: Rhinotermitidae) revealed by phylogenetic analyses of two mitochondrial genes. *Journal of Insect Science*, 16(1).
- Espíndola, A. A. (2015). Processo de certificação do etanol brasileiro. *Revista de Ciências Gerenciais*, 13(17), 113-130.
- Florencio, C., Badino, A. C., & Farinas, C. S. (2017). Current challenges on the production and use of cellulolytic enzymes in the hydrolysis of lignocellulosic biomass. *Química Nova*, 40(9), 1082-1093.
- Furtado, A. T., Hekkert, M. P., & Negro, S. O. (2020). Of actors, functions, and fuels: Exploring a second generation ethanol transition from a technological innovation systems perspective in Brazil. 70, 101706.
- Gopal, G. J., & Kumar, A. (2013). Strategies for the production of recombinant protein in *Escherichia coli*. *The protein journal*, 32(6), 419-425.
- Green, M. R., & Sambrook, J. (2012). Molecular cloning: a laboratory manual. In *Molecular cloning: a laboratory manual* (pp. 1890-1890).
- Griebenow, Z. H., Jones, S. C., & Eaton, T. D. (2017). Seeking quantitative morphological characters for species identification in soldiers of Puerto Rican Heteroterms (Dictyoptera, Blattaria, Termitoidea, Rhinotermitidae). *ZooKeys*(725), 17.
- Hol, W. G. (1985). The role of the α -helix dipole in protein function and structure. *Progress in biophysics and molecular biology*, 45(3), 149-195.
- Horn, S. J., Vaaje-Kolstad, G., Westereng, B., & Eijsink, V. (2012). Novel enzymes for the degradation of cellulose. *Biotechnology for biofuels*, 5(1), 45.

- Huang, Y.-F., Chiueh, P.-T., & Lo, S.-L. (2016). A review on microwave pyrolysis of lignocellulosic biomass. *Sustainable Environment Research*, 26(3), 103-109.
- Jagtap, S. S., Dhiman, S. S., Jeya, M., Kang, Y. C., Choi, J.-H., & Lee, J.-K. J. B. t. (2012). Saccharification of poplar biomass by using lignocellulases from *Pholiota adiposa*. 120, 264-272.
- Jang, S., & Kikuchi, Y. (2020). Impact of the insect gut microbiota on ecology, evolution, and industry. *Current Opinion in Insect Science*, 41, 33-39.
- Jansen, M. L., Bracher, J. M., Papapetridis, I., Verhoeven, M. D., de Bruijn, H., de Waal, P. P., . . . Pronk, J. T. (2017). *Saccharomyces cerevisiae* strains for second-generation ethanol production: from academic exploration to industrial implementation. *FEMS yeast research*, 17(5).
- Johnson, N., Powis, K., & High, S. J. B. e. B. A.-M. C. R. (2013). Post-translational translocation into the endoplasmic reticulum. 1833(11), 2403-2409.
- Karimi, K., & Taherzadeh, M. J. (2016). A critical review on analysis in pretreatment of lignocelluloses: Degree of polymerization, adsorption/desorption, and accessibility. *Bioresource technology*, 203, 348-356.
- Kasana, R. C., Salwan, R., Dhar, H., Dutt, S., & Gulati, A. J. C. m. (2008). A rapid and easy method for the detection of microbial cellulases on agar plates using Gram's iodine. 57(5), 503-507.
- Kaur, J., Kumar, A., & Kaur, J. (2018). Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements. *International Journal of Biological Macromolecules*, 106, 803-822.
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. J. N. p. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. 10(6), 845-858.
- Khan, A. S., Man, Z., Bustam, M. A., Nasrullah, A., Ullah, Z., Sarwono, A., . . . Muhammad, N. (2018). Efficient conversion of lignocellulosic biomass to levulinic acid using acidic ionic liquids. *Carbohydrate polymers*, 181, 208-214.
- Khoo, H. H. (2015). Review of bio-conversion pathways of lignocellulose-to-ethanol: Sustainability assessment based on land footprint projections. *Renewable Sustainable Energy Reviews*, 46, 100-119.
- Kim, H., Ahn, Y., & Kwak, S.-Y. (2016). Comparing the influence of acetate and chloride anions on the structure of ionic liquid pretreated lignocellulosic biomass. *Biomass and Bioenergy*, 93, 243-253.
- Kim, M., & Day, D. F. (2011). Composition of sugar cane, energy cane, and sweet sorghum suitable for ethanol production at Louisiana sugar mills. *Journal of industrial microbiology and biotechnology*, 38(7), 803-807.
- Ko, E., Akatsuka, H., Moriyama, H., Shinmyo, A., Hata, Y., Katsube, Y., . . . Okada, H. (1992). Site-directed mutagenesis at aspartate and glutamate residues of xylanase from *Bacillus pumilus*. *Biochemical Journal*, 288(1), 117-121.
- Korb, J., & Belles, X. (2017). Juvenile hormone and hemimetabolism eusociality: a comparison of cockroaches with termites. *Current Opinion in Insect Science*, 22, 109-116.
- Kumar, D., Kumar, S. S., Kumar, J., Kumar, O., Mishra, S. V., Kumar, R., & Malyan, S. K. (2017). Xylanases and their industrial applications: a review. *Biochem Cell Arch*, 17(1), 353-360.
- Laemmli, U. K. J. n. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. 227(5259), 680-685.
- Liu, M.-q., Huo, W.-k., Xu, X., & Weng, X.-y. (2017). Recombinant *Bacillus amyloliquefaciens* xylanase A expressed in *Pichia pastoris* and generation of

- xylooligosaccharides from xylans and wheat bran. *International Journal of Biological Macromolecules*, *105*, 656-663.
- Liu, X., Liu, Y., Jiang, Z., Liu, H., Yang, S., & Yan, Q. J. F. c. (2018). Biochemical characterization of a novel xylanase from *Paenibacillus barengoltzii* and its application in xylooligosaccharides production from corncobs. *264*, 310-318.
- Lo, N., & Eggleton, P. (2010). Termite phylogenetics and co-cladogenesis with symbionts. In *Biology of termites: a modern synthesis* (pp. 27-50): Springer.
- Lopes, M. L., Paulillo, S. C. d. L., Godoy, A., Cherubin, R. A., Lorenzi, M. S., Giometti, F. H. C., . . . Amorim, H. V. d. (2016). Ethanol production in Brazil: a bridge between science and industry. *Brazilian Journal of Microbiology*, *47*, 64-76.
- Lu, P., & Hsieh, Y.-L. (2012). Preparation and characterization of cellulose nanocrystals from rice straw. *Carbohydrate polymers*, *87*(1), 564-573.
- Macharia, R. W., Ombura, F. L., & Aroko, E. O. J. J. o. N. A. (2015). Insects' RNA profiling reveals absence of "hidden break" in 28S ribosomal RNA molecule of onion thrips, *Thrips tabaci*. *2015*.
- Meilany, D., Anugeraheni, D., Aziz, A., Kresnowati, M. T. A. P., & Setiadi, T. (2020). The Effects of Operational Conditions in Scaling Up of Xylanase Enzyme Production for Xylitol Production. *Reaktor*, *20*(1), 32-37.
- Muranaka, Y., Nakagawa, H., Hasegawa, I., Maki, T., Hosokawa, J., Ikuta, J., & Mae, K. (2017). Lignin-based resin production from lignocellulosic biomass combining acidic saccharification and acetone-water treatment. *Chemical Engineering Journal*, *308*, 754-759.
- Neitzel, T., Lima, C. S., Biazi, L. E., Collograi, K. C., da Costa, A. C., dos Santos, L. V., & Ienczak, J. L. (2020). Impact of the Melle-Boinot process on the enhancement of second-generation ethanol production by *Spathaspora passalidarum*. *Renewable Energy*, *160*, 1206-1216.
- Nevalainen, H., & Peterson, R. J. F. i. m. (2014). Making recombinant proteins in filamentous fungi-are we expecting too much? , *5*, 75.
- Ni, J., & Tokuda, G. (2013). Lignocellulose-degrading enzymes from termites and their symbiotic microbiota. *Biotechnology Advances*, *31*(6), 838-850.
- Nielsen, F., Zacchi, G., Galbe, M., & Wallberg, O. (2017). Sequential Targeting of Xylose and Glucose Conversion in Fed-Batch Simultaneous Saccharification and Co-fermentation of Steam-Pretreated Wheat Straw for Improved Xylose Conversion to Ethanol. *Bioenergy research*, *10*(3), 800-810.
- Padmanabhan, S., Banerjee, S., Mandi, N. J. M. C. S. A. i. M., & Biology. (2011). Screening of bacterial recombinants: Strategies and preventing false positives. p1-19.
- Paës, G., Berrin, J.-G., & Beaugrand, J. (2012). GH11 xylanases: structure/function/properties relationships and applications. *Biotechnology Advances*, *30*(3), 564-592.
- Peng, F., Peng, P., Xu, F., & Sun, R.-C. (2012). Fractional purification and bioconversion of hemicelluloses. *Biotechnology Advances*, *30*(4), 879-903.
- Prado, C. D., Mandrujano, G. P., Souza, J. P., Sgobbi, F. B., Novaes, H. R., da Silva, J. P., . . . Giudici, R. (2020). Physiological characterization of a new thermotolerant yeast strain isolated during Brazilian ethanol production, and its application in high-temperature fermentation. *Biotechnology for biofuels*, *13*(1), 1-15.
- Robles, J., & Doers, M. (1994). pGEM®-T Vector Systems troubleshooting guide. *Promega Notes*, *45*, 19-20.
- Roisin, Y., & Korb, J. (2010). Social organisation and the status of workers in termites. In *Biology of termites: a modern synthesis* (pp. 133-164): Springer.
- Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in microbiology*, *5*, 172.

- Rosseto, R. E., de Souza Madalena, L. C., de Oliveira, A. F., Chang, P., Primieri, B. F., Frigo, E. P., & Santos, R. F. (2017). Panorama do etanol brasileiro. *Acta Iguazu*, 6(5), 13-22.
- Sá, L., Cammarota, M., & LEITÃO, V. (2014). Produção de hidrogênio via fermentação anaeróbia-aspectos gerais e possibilidade de utilização de resíduos agroindustriais brasileiros. *Química Nova*, 37(5), 857-867.
- Saha, B. C., Nichols, N. N., & Cotta, M. A. (2013). Comparison of separate hydrolysis and fermentation versus simultaneous saccharification and fermentation of pretreated wheat straw to ethanol by *Saccharomyces cerevisiae*. *Journal of Biobased Materials and Bioenergy*, 7(3), 409-414.
- Saha, B. C., Nichols, N. N., Qureshi, N., Kennedy, G. J., Iten, L. B., & Cotta, M. A. (2015). Pilot scale conversion of wheat straw to ethanol via simultaneous saccharification and fermentation. *Bioresource technology*, 175, 17-22.
- Saini, J. K., Saini, R., & Tewari, L. (2015). Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. *Biotech*, 5(4), 337-353.
- Sambrook, J., & Green, M. R. (2001). Molecular Cloning: A Laboratory Manual, 3^o ed. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, V.1*.
- Sanger, F., Nicklen, S., & Coulson, A. R. J. P. o. t. n. a. o. s. (1977). DNA sequencing with chain-terminating inhibitors. *74*(12), 5463-5467.
- Santos, F. A., Queiróz, J. H. d., Colodette, J. L., Fernandes, S. A., Guimarães, V. M., & Rezende, S. T. (2012). Potencial da palha de cana-de-açúcar para produção de etanol. *Química Nova*, 35(5), 1004-1010.
- Santos, F. A., Queiroz, J. H. d., Colodette, J. L., Manfredi, M., Queiroz, M. E. L., Caldas, C. S., & Soares, F. E. (2014). Otimização do pré-tratamento hidrotérmico da palha de cana-de-açúcar visando à produção de etanol celulósico. *Química Nova*, 37(1), 56-62.
- Scharf, M. E. (2015). Termites as targets and models for biotechnology. *Annual review of entomology*, 60, 77-102.
- Scharf, M. E., & Tartar, A. (2008). Termite digestomes as sources for novel lignocellulases. *Biofuels, Bioproducts and Biorefining*, 2(6), 540-552.
- Silva, D. S. G. d. (2016). *Isolamento de xilanas do bagaço de cana-de-açúcar integrado à hidrólise enzimática da celulose residual*. Universidade de São Paulo,
- Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M., & Panda, A. K. (2015). Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microbial cell factories*, 14(1), 1-10.
- Smith, C. K., Withka, J. M., & Regan, L. (1994). A thermodynamic scale for the. beta.-sheet forming tendencies of the amino acids. *Biochemistry*, 33(18), 5510-5517.
- Soccol, C. R., Faraco, V., Karp, S. G., Vandenberghe, L. P., Thomaz-Soccol, V., Woiciechowski, A. L., & Pandey, A. (2019). Lignocellulosic bioethanol: current status and future perspectives. In *Biofuels: Alternative feedstocks and conversion processes for the production of liquid and gaseous biofuels* (pp. 331-354): Elsevier.
- Sorn, V., Chang, K.-L., Phitsuwan, P., Ratanakhanokchai, K., & Dong, C.-D. J. B. t. (2019). Effect of microwave-assisted ionic liquid/acidic ionic liquid pretreatment on the morphology, structure, and enhanced delignification of rice straw. *293*, 121929.
- Studier, F. W., & Moffatt, B. A. J. J. o. m. b. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *189*(1), 113-130.
- Takita, T., Nakatani, K., Katano, Y., Suzuki, M., Kojima, K., Saka, N., . . . technology, m. (2019). Increase in the thermostability of GH11 xylanase XynJ from *Bacillus* sp. strain 41M-1 using site saturation mutagenesis. *130*, 109363.

- Talebnia, F., Karakashev, D., & Angelidaki, I. (2010). Production of bioethanol from wheat straw: an overview on pretreatment, hydrolysis and fermentation. *Bioresource technology*, *101*(13), 4744-4753.
- Taylor, W. R. (1986). The classification of amino acid conservation. *Journal of theoretical Biology*, *119*(2), 205-218.
- Teather, R. M., & Wood, P. J. (1982). Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *43*(4), 777-780.
- Thomas, L., Ushasree, M. V., & Pandey, A. (2014). An alkali-thermostable xylanase from *Bacillus pumilus* functionally expressed in *Kluyveromyces lactis* and evaluation of its deinking efficiency. *Bioresource technology*, *165*, 309-313.
- Vardakou, M., Dumon, C., Murray, J. W., Christakopoulos, P., Weiner, D. P., Juge, N., . . . Flint, J. E. J. J. o. m. b. (2008). Understanding the structural basis for substrate and inhibitor recognition in eukaryotic GH11 xylanases. *375*(5), 1293-1305.
- Wan, C., Zhou, Y., & Li, Y. (2011). Liquid hot water and alkaline pretreatment of soybean straw for improving cellulose digestibility. *Bioresource technology*, *102*(10), 6254-6259.
- Warnecke, F., Luginbühl, P., Ivanova, N., Ghassemian, M., Richardson, T. H., Stege, J. T., . . . Aboushadi, N. (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*, *450*(7169), 560-565.
- Weinacker, D., Rabert, C., Zepeda, A. B., Figueroa, C. A., Pessoa, A., & Farías, J. G. (2013). Applications of recombinant *Pichia pastoris* in the healthcare industry. *Brazilian Journal of Microbiology*, *44*, 1043-1048.
- Winnebeck, E. C., Millar, C. D., & Warman, G. R. (2010). Why does insect RNA look degraded? *Journal of Insect Science*, *10*(1), 159.
- Woolley, S., Johnson, J., Smith, M. J., Crandall, K. A., & McClellan, D. A. (2003). TreeSAAP: selection on amino acid properties using phylogenetic trees. *Bioinformatics*, *19*(5), 671-672.
- Wyman, C. E., Dale, B. E., Elander, R. T., Holtzapple, M., Ladisch, M. R., & Lee, Y. (2005). Coordinated development of leading biomass pretreatment technologies. *Bioresource technology*, *96*(18), 1959-1966.
- Yoon, J. H., Park, J. F., Suh, D. Y., Hong, S. B., Ko, S. J., & Kim, S. H. J. M. (2007). Comparison of dyes for easy detection of extracellular cellulases in fungi. *35*(1), 21-24.
- Yu, J., Paterson, N., Blamey, J., & Millan, M. (2017). Cellulose, xylan and lignin interactions during pyrolysis of lignocellulosic biomass. *Fuel*, *191*, 140-149.
- Zhang, Z.-Y. (2002). Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development. *Annual review of pharmacology toxicology*, *42*(1), 209-234.
- Zhu, Y., Li, X., Sun, B., Song, H., Li, E., & Song, H. (2012). Properties of an Alkaline-Tolerant, Thermostable Xylanase from *Streptomyces chartreusis* L1105, Suitable for Xylooligosaccharide Production. *Journal of food science*, *77*(5), C506-C511.