



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE BIOLOGIA

MARIA LORENZA LEAL MOTTA

**CARACTERIZAÇÃO BIOQUÍMICA DE UMA NOVA
ARABINOFURANOSIDASE DA FAMÍLIA GH54 RELACIONADA
A DEGRADAÇÃO DE HEMICELULOSE EM TRICHODERMA
HARZIANUM**

**“BIOCHEMICAL CHARACTERIZATION OF A NEW
ARABINOFURANOSIDASE OF THE GH54 FAMILY RELATED
TO DEGRADATION OF HEMICELLULOSIS IN TRICHODERMA
HARZIANUM”**

CAMPINAS

2021

MARIA LORENZA LEAL MOTTA

**CARACTERIZAÇÃO BIOQUÍMICA DE UMA NOVA
ARABINOFURANOSIDASE DA FAMÍLIA GH54 RELACIONADA
A DEGRADAÇÃO DE HEMICELULOSE EM TRICHODERMA
HARZIANUM**

**“BIOCHEMICAL CHARACTERIZATION OF A NEW
ARABINOFURANOSIDASE OF THE GH54 FAMILY RELATED
TO DEGRADATION OF HEMICELLULOSIS IN TRICHODERMA
HARZIANUM”**

Dissertação apresentada ao Instituto de Biologia da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestra em Genética e Biologia Molecular, na área de Genética de Microrganismos.

Dissertation presented to the Institute of Biology of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Genetics and Molecular Biology, in the area of Genetics of Microorganisms.

Orientador: Profa. Dra. Anete Pereira de Souza

ESTE ARQUIVO DIGITAL CORRESPONDE À VERSÃO
FINAL DA DISSERTAÇÃO DEFENDIDA PELA ALUNA
MARIA LORENZA LEAL MOTTA E ORIENTADA PELA
PROFESSORA DOUTORA ANETE PEREIRA DE SOUZA.

CAMPINAS

2021

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

Motta, Maria Lorenza Leal, 1995-
M858c Caracterização bioquímica de uma nova arabinofuranosidase da família GH54 relacionada a degradação de hemicelulose em *Trichoderma harzianum* / Maria Lorenza Leal Motta. – Campinas, SP : [s.n.], 2021.

Orientador: Anete Pereira de Souza.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Alfa-L-arabinofuranosidase. 2. Glicosídeo hidrolases. 3. *Trichoderma harzianum*. I. Souza, Anete Pereira de, 1962-. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Biochemical characterization of a new arabinofuranosidase of the GH54 family related to degradation of hemicellulosis in *Trichoderma harzianum*

Palavras-chave em inglês:

Alpha-L-arabinofuranosidase

Glycoside hydrolases

Trichoderma harzianum

Área de concentração: Genética de Microorganismos

Titulação: Mestra em Genética e Biologia Molecular

Banca examinadora:

Anete Pereira de Souza [Orientador]

Adriano Rodrigues Azzoni

André da Silva Santiago

Data de defesa: 26-02-2021

Programa de Pós-Graduação: Genética e Biologia Molecular

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0003-3833-1561>

- Currículo Lattes do autor: <http://lattes.cnpq.br/4910429951890751>

Campinas, 26 de fevereiro de 2021.

COMISSÃO EXAMINADORA

Profa. Dra. Anete Pereira de Souza

Dr. Adriano Rodrigues Azzoni

Dr. André da Silva Santiago

Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa de Pós-Graduação em Genética e Biologia Molecular do Instituto de Biologia.

*Dedico as três mulheres da minha vida;
Vanessa, Sheila e Maria, ao meu avô
Valdo Longo e ao companheiro de alma Vinícius
por todo amor e apoio que me deram ao longo da minha vida.*

AGRADECIMENTOS

Gostaria de agradecer em primeiro lugar a minha família por todo o apoio que me deram não só durante o meu mestrado mas também em minha vida.

A minha mãe Vanessa que sempre esteve ao meu lado torcendo, me aconselhando, chorando com minhas perdas e comemorando minhas conquistas. Você é o verdadeiro significado de amor. A Sheila que com seu jeito único de ser me faz sentir mais amada a cada dia. A minha avó Maria que além de me dar o nome me deu sua força para viver. Ao meu avô Valdo Longo com o jeito simples me ensinou a cultivar e cuidar de minhas raízes. Ao meu namorado Vinícius que foi meu porto seguro nesta jornada e na vida.

Ao meu amigo Dr. Jaire Alves F. Filho que foi um dos maiores presentes do mestrado. Obrigada por tudo que fez por mim. Sempre serei eternamente grata.

A minha orientadora Dra Anete Pereira de Souza por ter me aceito em seu grupo de pesquisa e por todo conhecimento não só científico, mas sobre a vida. Agradeço pela oportunidade de vivenciar uma experiência única em minha vida. Ao meu coorientador Dr Clelton Aparecido dos Santos que foi fundamental para que esse trabalho fosse concluído. A todos os colegas de trabalho do LAGM/CBMEG – Barracão por me ajudarem durante diversos momentos. Aos técnicos do laboratório Danilo, Aline e Patrícia por toda ajuda.

A todos os meus amigos Amanda, Maria Carolina, Gessica, Caio, Jessica, Isabela, Gabriela, Isabela, entre vários outros. Vocês fazem parte da minha vida.

Ao Programa de Pós-Graduação em Genética e Biologia Molecular e à Universidade Estadual de Campinas – UNICAMP.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. À CAPES (Programa Biologia Computacional – 88887.176241/2018-00 e Programa de Excelência Acadêmica – PROEX – 88882.329483/2019-01) pela bolsa de mestrado concedida. À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (2015/09202-0) pelo auxílio na pesquisa.

Enfim, sou grata a todos que tornaram de alguma forma a realização do meu projeto de mestrado

*“Quem não sabe para onde vai não
chega a parte nenhuma.”*

Carlos Ruiz Záfon

RESUMO

Os fungos pertencentes ao gênero *Trichoderma* possuem grande potencial biotecnológico, com aplicações na indústria de biocombustíveis e no desenvolvimento de biofungicidas. O uso da biomassa vegetal para a produção de bioprodutos de alto valor agregado, incluindo o bioetanol, tem motivado a busca por novas enzimas/proteínas a serem utilizadas nesses processos. No entanto, uma ampla gama de enzimas que atuam sinergicamente é necessária para a decomposição eficiente da matéria lignocelulósica. Dessa forma, enzimas mais eficientes que atuam em diferentes tipos de açúcares podem ser uma solução para a obtenção de coquetéis melhores elaborados para esses fins. As α -L-arabinofuranosidases (ABFs) são enzimas que, além de atuarem sobre a L-arabinose, também podem exercer esse papel multifuncional. Alguns ABFs da família da glicosil hidrolase 54 (GH54) possuem esse mecanismo de atuação em diferentes substratos. Assim, este projeto bioprospecção *in silico* um ABF pertencente à família GH54 de *Trichoderma harzianum* dependente de metal com atividade expandida para outros substratos. Também foi possível verificar que a família GH54 possui excelente diversidade funcional. Esses resultados demonstram que o ABF estudado pode ser aplicado nas indústrias sucroalcooleiras para a produção de etanol 2G e demonstram o potencial enzimático da família GH54.

Palavras-chave: α -L-arabinofuranosidase; GH54; *Trichoderma harzianum*.

ABSTRACT

Fungi belonging to the genus *Trichoderma* have great biotechnological potential, with applications in the biofuels industry and the development of biofungicides. The use of vegetable biomass to produce bioproducts with high added value, including bioethanol, has motivated the search for new enzymes/proteins to be used in these processes. However, a wide range of enzymes synergistically acting is required for the efficient breakdown of lignocellulosic matter. In this way, more efficient enzymes that act in different types of sugars can be a solution to obtain better cocktails designed for these purposes. α -L-arabinofuranosidases (ABFs) are enzymes that, in addition to acting on L-arabinose, can also play this multifunctional role. Some ABFs of the glycosyl hydrolase 54 (GH54) family have this acting mechanism on different substrates. Thus, this project bioprospecting *in silico* an ABF belonging to the GH54 family of *Trichoderma harzianum* dependent metal with expanded activity for other substrates. It was also possible to verify that the GH54 family has excellent functional diversity. These results demonstrate that the ABF studied can be applied in the sugar and alcohol industries for the production of 2G ethanol and demonstrate the enzymatic potential of the GH54 family.

Keywords: α -L-arabinofuranosidase; GH54; *Trichoderma harzianum*.

LISTA DE ILUSTRAÇÕES

Revisão bibliográfica

- Figura 1-** Composição da biomassa vegetal (RUBIN,2008). 15
Figura 2- Tipos de açúcares e ácidos encontrados na estrutura da hemicelulose (MORAIS; NASCIMENTO; MELO, 2005). 16
Figura 3- Fluxograma do processo de produção de etanol de segunda geração através da biomassa vegetal (AZEVEDO et al. 2016). 17
Figura 4- Enzimas necessárias para a degradação da celulose (a) Enzimas que participam da degradação da hemicelulose (b).(FLORENCIO et al. 2017)..... 18

Artigo

- Figure 1-** Phylogenetic analysis and sequence domains of the GH54 family..... 25
Figure 1- Alignment and structure prediction of ThABF..... 26
Figure 2- SDS-PAGE in 12% gels showing purification results and graphs of secondary and tertiary folding analyses..... 27
Figure 3- Relative activity of ThABF in the presence of different metallic ions..... 29
Figure 4- Biochemical characterization of ThABF..... 30
Table 1- Relative activity of ThABF against different synthetic substrates..... 29
Table 2- Kinetic parameters of ThABF against selected substrates..... 31

SUMÁRIO

1 INTRODUÇÃO GERAL	12
2 REVISÃO BIBLIOGRÁFICA	14
2.1 Panorama geral da produção de etanol	14
2.2 Biomassa vegetal e o gênero <i>Trichoderma spp.</i>.....	14
2.3 Hidrólise enzimática	16
2.4 α-L-arabinofuranosidase.....	19
3 OBJETIVOS	20
3.1 Objetivos gerais.....	20
3.2 Objetivos específicos	20
4 ARTIGO	21
“Novel fungal metal-dependent GH54 α -L-arabinofuranosidase: expanded substrate specificity and potential use for plant biomass degradation”	21
5 CONCLUSÕES GERAIS	41
6 PERSPECTIVAS.....	42
7 REFERÊNCIAS BIBLIOGRÁFICAS	43
ANEXO I – DECLARAÇÃO CIBio	47
ANEXO II – DECLARAÇÃO DE AUTORIA	48

1 INTRODUÇÃO GERAL

Os biocombustíveis surgiram devido às crises enfrentadas nos setores econômico e ambiental (NAKANISHI *et al.*, 2018). O etanol de segunda geração (2G) tem como uma das vertentes o aproveitamento máximo de insumos industriais que em sua maioria são descartados como a biomassa vegetal (CHANDEL *et al.*, 2018). No Brasil o bagaço de cana-de-açúcar é a principal fonte de biomassa para a produção de etanol 2G (SANTOS; MAGRINI, 2018). São necessárias diversas etapas para a produção do etanol 2G que podem ser resumidas basicamente em pré-tratamento, sacarificação, fermentação e destilação (BORIN *et al.*, 2017).

A sacarificação é fundamental nesse processo pois é onde ocorre a hidrólise enzimática da biomassa lignocelulósica. É necessário um complexo de múltiplas enzimas atuando de maneira sinérgica para conversão dos carboidratos complexos em monômeros de açúcares (GUPTA *et al.*, 2016; HORTA *et al.*, 2018). Estas enzimas podem ser encontradas em diversos microrganismos como nos fungos ascomicetos filamentosos do gênero *Trichoderma* (DRUZHININA *et al.*, 2018). As espécies desse gênero podem ser encontradas em diferentes habitats e por isso, podem ser aplicados industrialmente, seja ela no ramo de produtos alimentícios, na produção de biofungicidas ou biocombustíveis (DRUZHININA *et al.*, 2018). O *T.harzianum* é uma espécie promissora para a produção de enzimas hidrolíticas, podendo ser aplicada na produção de etanol 2G (HORTA *et al.*, 2014). Essas enzimas produzidas por estes fungos, entre outros, são fundamentais para a degradação da celulose, lignina e hemicelulose (YEO *et al.*, 2019).

Polissacarídeos como a hemicelulose possuem uma grande complexidade estrutural devido a presença de diversos açúcares ligados a sua cadeia principal (CHANDEL *et al.*, 2018). Para que ocorra sua degradação é necessário uma grande variedade de enzimas. Neste contexto, a sacarificação da hemicelulose é um dos desafios enfrentados na produção de etanol 2G (BIELY; SINGH; PUCHART, 2016). Uma das soluções seria o emprego de enzimas que atuam em uma grande variedade de açúcares presentes nas cadeias laterais (CHEN *et al.*, 2018; FRY, 2016).

As α -L-arabinofuranosidases (ABFs) é um grupo de enzimas com um importante papel na hidrólise da hemicelulose (JORDAN; BRAKER, 2011). Elas agem nas ligações glicosídicas de L-arabinose da estrutura da arabinoxilano e tem capacidade de atuar

sinergisticamente com outras enzimas, além de algumas serem bifuncionais (atuando em dois ou mais substratos) (BENOCCI *et al.*, 2018; HU *et al.*, 2018; WANG *et al.*, 2014). As ABFs podem ser encontradas em diferentes famílias de glicosil hidrolases (GH), sendo uma delas a GH54 (LAGAERT *et al.*, 2014). Este grupo de enzimas tem membros com atividade em diferentes substratos e bifuncionais, no entanto, ela é pouco explorada. Desta forma, as ABFs desta família são potenciais candidatas com capacidade de catalisar diferentes reações enzimáticas, sendo aptas a hidrolisar uma variedade de substratos (YANG *et al.*, 2017).

Diante do exposto, as ABFs pertencentes a família GH54 de *Trichoderma harzianum* são alvos promissores para serem estudados. Neste estudo, foi selecionada uma ABF multifuncional através da bioprospecção *in silico* e sua atividade em diferentes substratos foi determinada. Os resultados obtidos neste trabalho ampliam as informações sobre a atuação enzimática da família GH54 além de poder ser uma nova enzima a ser aplicada no aprimoramento da tecnologia da produção de etanol 2G.

2 REVISÃO BIBLIOGRÁFICA

2.1 Panorama geral da produção de etanol

A produção de etanol no mundo emanou-se de uma crescente necessidade de novas fontes energéticas renováveis, do esgotamento das reservas de petróleo e das crises geopolíticas (NAKANISHI *et al.*, 2018). No Brasil, as crises no setor açucareiro e a dependência de petróleo importado impulsionaram o crescimento das indústrias sucroalcooleiras (DOS SANTOS CASTRO *et al.*, 2014; NAKANISHI *et al.*, 2018).

Existem basicamente dois tipos de etanol produzidos no Brasil. O de primeira geração que é feito a partir do caldo de cana fermentado e o segunda (2G) que é produzido através da biomassa vegetal (TURSI, 2019). Há diversas fontes de matérias orgânicas que são utilizadas na produção do etanol 2G, como o bagaço de cana, palha de milho, arroz e de trigo, sendo que, no Brasil o bagaço de cana-de-açúcar tem uma maior participação (POPP *et al.*, 2016).

O Brasil é o maior produtor mundial de cana, sendo presente em 24 dos 27 estados brasileiros e cinco destes concentram 88% da produção nacional de açúcar e etanol (SANTOS; MAGRINI, 2018). A produção de etanol de segunda geração (2G) representa uma opção coerente, segura e ecológica, que faz aproveitamento ao máximo deste tipo de insumo industrial (DÍAZ PÉREZ *et al.*, 2018).

2.2 Biomassa vegetal e o gênero *Trichoderma spp.*

A biomassa lignocelulósica é bastante diversa e pode provir de diferentes fontes como madeira lenhosa, gramínea e resíduos agrícolas (ISIKGOR; BECER, 2015). Este último é produzido em abundância no mundo todo e em sua maioria é descartado ou queimado para gerar energia nas usinas como o bagaço de cana (SAINI; SAINI; TEWARI, 2015). Esses descartes lignocelulósicos produzidos na agricultura e na indústria alimentícia podem ser matéria-prima em biorrefinarias para a produção de etanol 2G (FRANCOCCI; RECA, 2016).

A biomassa vegetal é composta principalmente por lignina, celulose e hemicelulose (Figura 1) (TURSI, 2019). Estes polissacarídeos são formados por diversos polímeros de açúcares fermentáveis que mudam a concentração e a composição de acordo com o tipo de biomassa, condições climáticas e cultivo (TURSI, 2019). A hemicelulose

é o polissacarídeo que tem maior diversidade em sua composição (ZOGHLAMI; PAËS, 2019). A sua cadeia principal é formada de D-xilose com açúcares laterais de L-arabinose, D-manoze, D-glicose, D-galactose, L-ramnoze, L-frutose e ácido D-glucurônico, que podem ser acetilados ou metilados (Figura 2) (CHANDEL *et al.*, 2018).

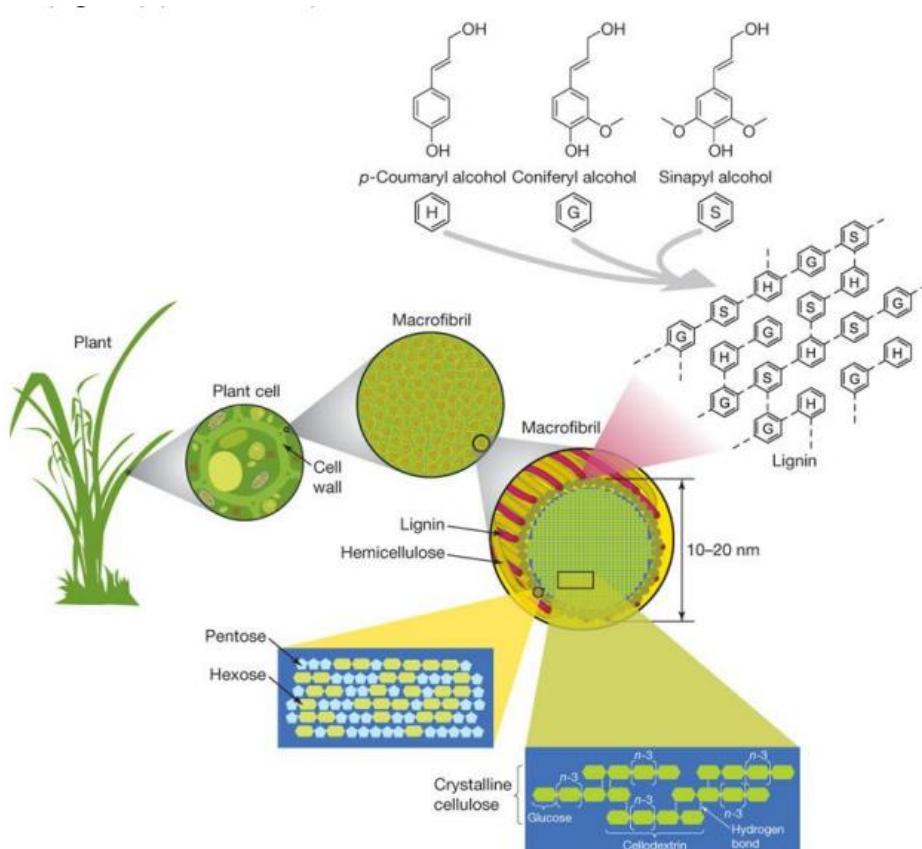


Figura 1- Composição da biomassa vegetal (RUBIN, 2008).

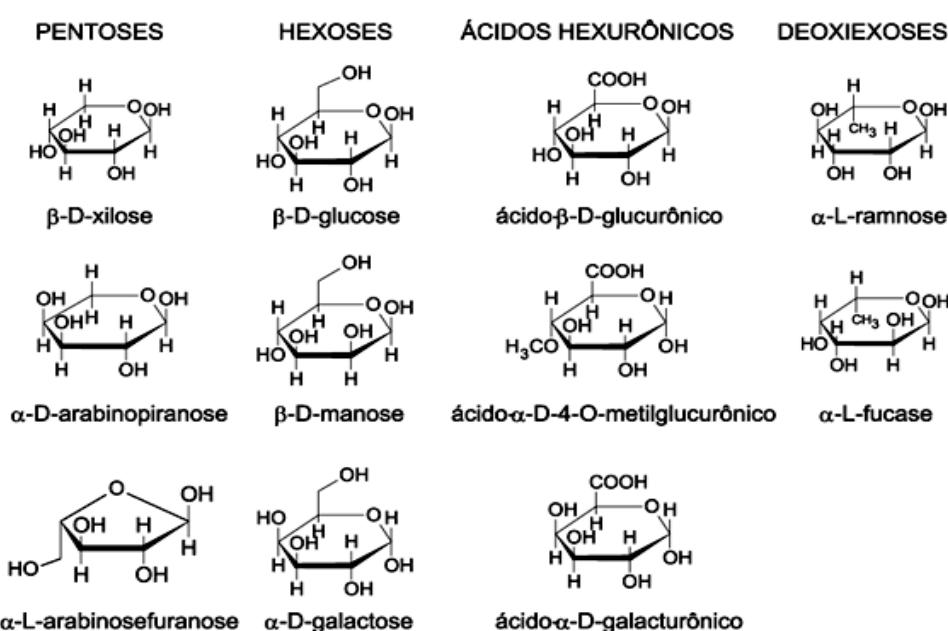


Figura 2- Tipos de açúcares e ácidos encontrados na estrutura da hemicelulose (MORAIS; NASCIMENTO; MELO, 2005).

Além do etanol a hemicelulose também pode gerar outros produtos com alto valor agregado como o xilitol, 2,3-butanodiol, ácido poli-hidroxibutírico, biopigmentos, entre outros (HASSAN; WILLIAMS; JAISWAL, 2018). Entretanto ela precisa ser despolimerizada para obter açúcares fermentáveis e para isso, é necessário um conjunto de enzimas (MELATI *et al.*, 2019). Outro aspecto importante da bioconversão da hemicelulose é o pré-tratamento. Esta etapa é imprescindível para sua despolimerização, pois expõe os biopolímeros para que as enzimas possam atuar sobre eles (MELATI *et al.*, 2019).

As enzimas utilizadas nesses processos provém de microrganismos que são encontrados na natureza. Os fungos são excelentes produtores de enzimas hidrolíticas devido a diversidade de habitats onde são encontrados (YEO *et al.*, 2019). Para que possam se alimentar eles utilizam coquetéis de enzimas que quebram uma variedade de biopolímeros (DRUZHININA *et al.*, 2018; KUBICEK; STARR; GLASS, 2014). O gênero *Trichoderma* pode ser encontrado em diversos habitats. Existem vários trabalhos que abordam a versatilidade funcional dos membros deste gênero que vão desde aplicações em indústrias sucroalcooleiras até desenvolvimento de biofungicidas (KUBICEK; STARR; GLASS, 2014).

A espécie do *T.harzianum* tem demonstrado ter um grande capacidade de produzir enzimas hidrolíticas (FERREIRA FILHO *et al.*, 2017). O potencial desta espécie foi demonstrado no trabalho de Horta *et al.* (2014), que publicou primeiro transcriptoma de *T. harzianum* IOC-3844 em diferentes condições (bagaço, celulose e lactose) demonstrando sua capacidade para degradação da biomassa vegetal. Além deste, outros trabalhos mostram o potencial dessa espécie na produção de enzimas hidrolíticas, podendo ser empregada na produção de enzimas e aplicada em diferentes processos biotecnológicos (MELATI *et al.*, 2019; ROCHA *et al.*, 2016).

2.3 Hidrólise enzimática

A produção do etanol de segunda geração é um processo que envolve basicamente três etapas: o pré-tratamento físico e/ou químico da biomassa vegetal que auxilia na remoção de impurezas e a expor os açúcares complexos; a sacarificação (hidrólise enzimática) que envolve a quebra de polissacarídeos complexos em monossacarídeos; a

fermentação, que é responsável pela produção de etanol a partir dos monossacarídeos e por fim, a destilação que obtém-se o etanol puro (Figura 3) (BORIN *et al.*, 2017).

O processo de hidrólise enzimática é uma das etapas fundamentais e mais críticas na produção de biocombustíveis, pois é através dela que a biomassa vegetal será convertida em açúcares fermentáveis (BASU, 2010). A composição da matéria lignocelulósica é um dos fatores que irão influenciar neste processo, pois as concentrações dos açúcares podem variar (ZOGHLAMI; PAËS, 2019).

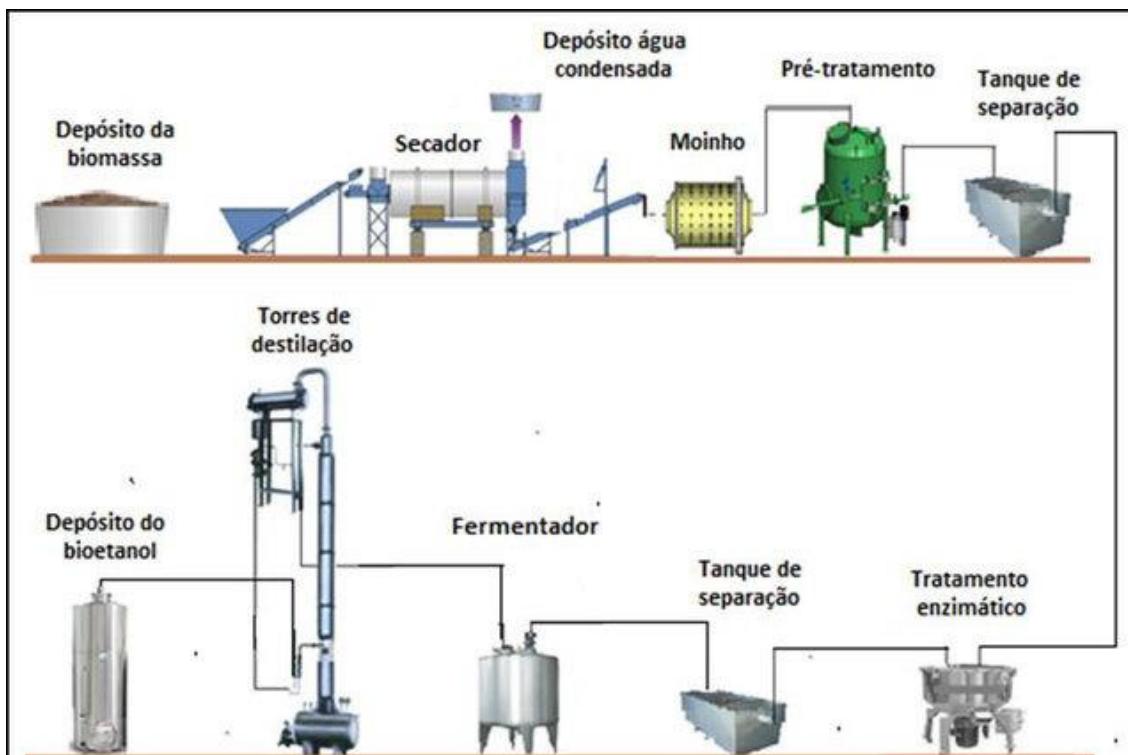


Figura 3- Fluxograma do processo de produção de etanol de segunda geração através da biomassa vegetal (AZEVEDO *et al.*, 2016).

As enzimas que participam deste processo estão categorizadas de acordo com sua função no banco de dados *Carbohydrate-Active enzymes* (CAZy -<http://www.cazy.org>), sendo elas, as hidrolases glicosídicas (GHs) que hidrolisam as ligações glicosídicas de carboidratos complexos, as glicosiltransferases (GTs) que atuam na transferência de moléculas de açúcar, os polissacarídeos liases (PLs) que clivam as cadeias polissacáridicas contendo ácido urônico, os carboidratos esterase (CEs) que fazem a hidrólise das ligações éster em carboidratos e por último, as de atividades auxiliares (AAs) que promovem as reações de oxidação-redução na degradação da biomassa vegetal (CONTESINI *et al.*, 2017; GUPTA *et al.*, 2016).

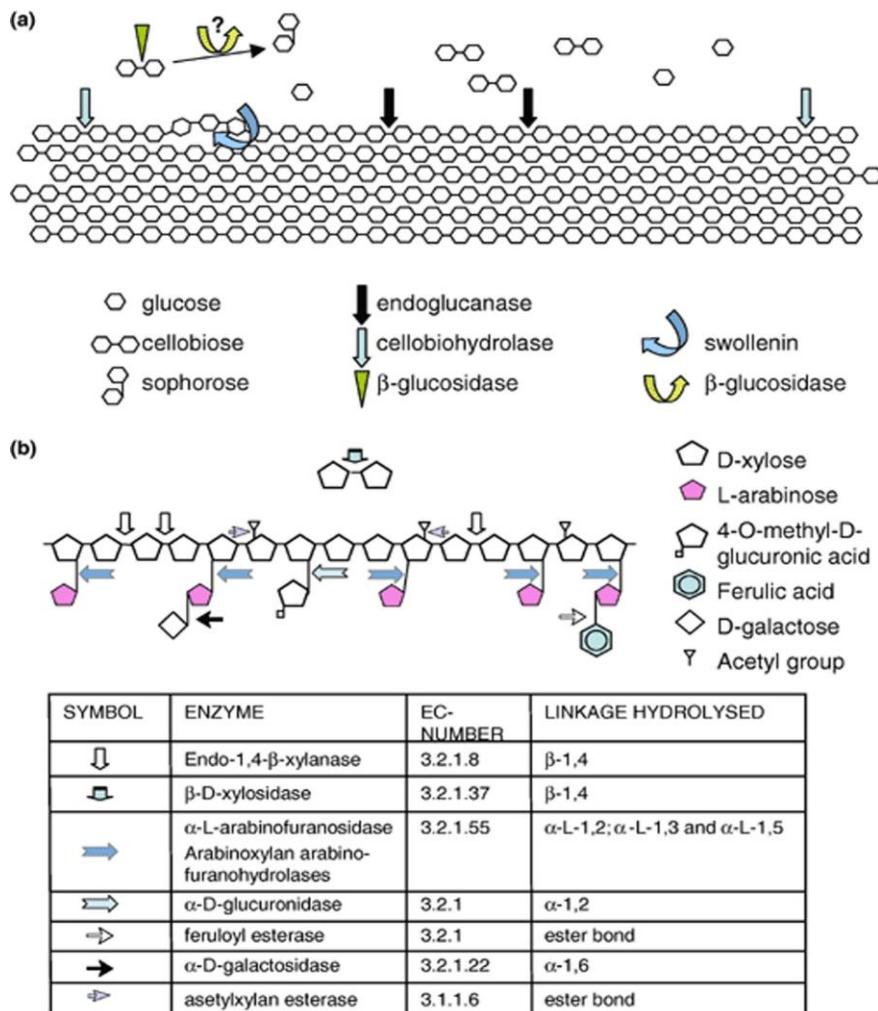


Figura 4- Enzimas necessárias para a degradação da celulose (a) Enzimas que participam da degradação da hemicelulose (b).(FLORENCIO *et al.*, 2017)

Há três classes enzimáticas que estão envolvidas na degradação da celulose: as endo- β -1,4-glucanases (EC 3.2.1.4), celobiohidrolases (EC 3.2.1.91/176), e β -glicosidases (EC 3.2.1.21). Já na hidrólise da hemicelulose é necessário vários grupos de enzimas, como: endo-1,4- β -xilanases (EC 3.2.1.8), β -xilosidases (EC 3.2.1.37), β -mananases (EC 3.2.1.78), arabinofuranosidases (EC 3.2.1.55), acetil-xilano esterases (EC 3.1.1.72), feruloil esterases (EC 3.1.1.73), α -glucuronidases (EC 3.2.1.139), entre outras (Figura 4) (MAITAN-ALFENAS; VISSER; GUIMARÃES, 2015).

As reações de hidrólise dependem das proporções das enzimas envolvidas e de suas características em que a interação enzimática será diferente para cada tipo de substrato (VAN DYK; PLETSCHKE, 2012). Sendo assim, faz-se necessário entender

melhor as relações entre as enzimas e seus papéis no processo de degradação da biomassa vegetal.

2.4 α -L-arabinofuranosidase

As α -L-arabinofuranosidases (ABFs) agem na hidrólise nas ligações glicosídicas das cadeias laterais de α -l-arabinofuranose no arabinoxilano, além de atuarem de maneira sinérgica com outras enzimas, como por exemplo as acetil xilano esterases, feruloil esterases e α -glucuronidases (BENOCCI *et al.*, 2018; HU *et al.*, 2018; WANG *et al.*, 2014).

As ABFs são excelentes candidatas para analisar sistemas bifuncionais pois grande parte destas enzimas são multifuncionais com capacidade atuar sobre diferentes substratos (YANG *et al.*, 2017). Essas enzimas são um interessante alvo para estudar o aperfeiçoamento da degradação da hemicelulose uma vez que podem atuar sobre vários açúcares desta (JORDAN; BRAKER, 2011).

As ABFs estão presentes nas famílias das hidrolases glicosídicas (GHs) 3, 43, 51, 54 e 62, e elas, mais as xilosidases são enzimas cruciais para a degradação intensiva da arabinoxilano (LAGAERT *et al.*, 2014). Na literatura existem estudos relacionados a algumas ABFs bifuncionais que pertencem principalmente às GHs 3, 43, 51 e 54.

A família GH54 possui principalmente arabinofuranosidases e xilanases. Alguns membros desta família têm atividade em diferentes substratos e também bifuncionais. Existem até o atual estudo 311 sequências depositadas nesta família sendo que apenas 19 foram caracterizadas e destas 6 testadas com diferentes substratos e duas apresentaram atividade em vários destes (MARGOLLES-CLARK *et al.*, 1996; RAVANAL; EYZAGUIRRE, 2015).

Assim, se faz necessário o estudo mais aprofundado dos membros desta, uma vez que demonstram ter potenciais características para serem aplicadas em processos biotecnológicos como a produção de biocombustíveis (LAGAERT *et al.*, 2014).

3 OBJETIVOS

3.1 Objetivos gerais

Bioprospectar *in silico* uma ABF da família GH54 de *Trichoderma harzianum* e realizar análises funcionais e caracterização bioquímica da enzima.

3.2 Objetivos específicos

- Realizar busca em bancos de dados por ABFs caracterizadas e mapear os reads de RNA-seq de *Trichoderma harzianum* para obter a sequência da proteína alvo;
- Fazer filogenia da sequência alvo com a família GH54;
- Analisar os domínios, fazer predição de estrutura e alinhamento com sequências caracterizadas;
- Realizar a clonagem, expressão e purificação da proteína alvo;
- Analisar os componentes de enovelamentos secundário e terciário;
- Caracterizar bioquimicamente a enzima expressa e purificada;
- Avaliar sua atividade na presença de diferentes íons metálicos e substratos sintéticos;
- Realizar cinética enzimática com os substratos que a enzima teve atividade.

4 ARTIGO

“Novel fungal metal-dependent GH54 α -L-arabinofuranosidase: expanded substrate specificity and potential use for plant biomass degradation”

Maria Lorenza Leal Motta¹, Jaire Alves Ferreira Filho¹, Ricardo Rodrigues de Melo²,

Letícia M. Zanphorlin², Clelton A. Santos², Anete P. Souza^{1*}

¹Center for Molecular Biology and Genetic Engineering (CBMEG), University of Campinas (UNICAMP), Campinas, SP, Brazil

²National Biorenewables Laboratory, National Research Center for Energy and Materials, Campinas, SP, Brazil

***Corresponding author**

Profa. Anete Pereira de Souza
Dept. de Biologia Vegetal
Universidade Estadual de Campinas
CEP 13083-875
Campinas, São Paulo, Brazil
Tel.: +55-19-3521-1132
Email: anete@unicamp.br

Abstract

Trichoderma genus fungi present great potential for the production of carbohydrate-active enzymes (CAZYmes), including glycoside hydrolase (GH) family members. From a renewability perspective, CAZYmes can be biotechnologically exploited to convert plant biomass into free sugars for the production of advanced biofuels and other high-value chemicals. GH54 is an attractive enzyme family for biotechnological applications because many GH54 enzymes are bifunctional. Thus, GH54 enzymes are interesting targets in the search for new enzymes for use in industrial processes such as plant biomass conversion. Herein, a novel metal-dependent GH54 arabinofuranosidase (ThABF) from the cellulolytic fungus *Trichoderma harzianum* was identified and biochemically characterized. Initial *in silico* searches were performed to identify the GH54 sequence. Next, the gene was cloned and heterologously overexpressed in *Escherichia coli*. The recombinant protein was purified, and the enzyme's biochemical and biophysical properties were assessed. The GH54 members show wide functional diversity and specifically remove plant cell decorations including arabinose and galactose, in the presence of a metallic cofactor. Plant cell wall decoration have a major impact on lignocellulosic substrate conversion into high-value chemicals. These results expand the known functional diversity within the GH54 family, showing the potential of a novel arabinofuranosidase for plant biomass degradation.

Introduction

Arabinofuranosidases (ABFs) (EC 3.2.1.55) are enzymes that are capable of cleaving residues of l-arabinofuranosyl present in various oligosaccharides and in polysaccharides such as hemicellulose; for this reason, they are interesting targets for biotechnological applications¹. These enzymes are part of a group of glycosidases necessary for the degradation of polymeric substrates, such as arabinane, arabinoxylan and other polysaccharides, that constitute the walls of plant cells². ABFs are grouped in glycosidic hydrolase (GH) families 3, 43, 51, 54 and 62, where GH54 includes some enzymes that are described as bifunctional³. However, few studies focusing on the functional specificity of this family have been conducted.

Ravanal *et al.*⁴ characterized an α -L-arabinofuranosidase from *Penicillium purpurogenum* and demonstrated that this family includes enzymes that are capable of acting synergistically with others, both releasing arabinose from different sources and acting on other substrates. Thus, given the features presented by this group of enzymes, further studies are still needed to better understand and biotechnologically explore the enzymes in this family, especially with regard to their application for the conversion of lignocellulosic substrates into bioproducts, such as biofuels.

Vegetable biomass is a reservoir of sugars organized in complex regular structures represented by cellulose, hemicellulose and lignin⁵. Such structures need to be deconstructed for the release of sugar-free monomers, which can then be converted into high-added-value chemicals such as second-generation ethanol⁶. This process requires the enzymatic conversion of biomass, and hemicellulose is a biopolymer that requires a large complex of enzymes for its conversion due to its structural composition⁷.

For the enzymes to access the main chain of hemicellulose, it is necessary to remove side-chain decorations, as they are one of the main reasons for the recalcitrance of plant biomass^{8,9}. Thus, the search for enzymes that target hemicellulose decorations is a promising approach for improving the saccharification of lignocellulolytic material¹⁰. It is of interest to apply ABFs for this purpose because in addition, removing l-arabinose, some of these enzymes can act on other sugars involved in decorations¹.

These enzymes are particularly important in the genus *Trichoderma*, where they play a crucial physiological role¹¹. This genus is composed of heterotrophic filamentous ascomycetous fungi that are capable of growing on different substrates and under different environmental conditions. There are several studies that have addressed the versatility of this genus, which range from applications in sugar and alcohol industries to

the development of biofungicides¹². *Trichoderma reesei* is of particular prominence within the group because of its recognized potential to produce several hydrolytic enzymes; however, other species of the genus, especially *Trichoderma harzianum*, have been shown to be as efficient as *T. reesei* in this regard¹³.

Therefore, in this study, we conducted the *in silico* bioprospecting of a new GH54 from *T. harzianum*. For this purpose, several approaches were used, such as RNA-seq, in addition to sequence analysis. The enzyme was also characterized biochemically with different substrates and metallic salts to determine their influence on the enzyme.

Results

Obtaining the sequence of ThABF and its in silico characterization

Through database searches, a sequence from *Trichoderma koningii* (AAA81024.1) was selected and used in the subsequent steps detailed in the methodology. With the resulting BLAST sequence, the *T. harzianum* CBS 226.95 genome was mapped the *T. harzianum* IOC-3844 RNA-seq reads and obtained the target protein sequence (MT439956). With this target protein sequence, a BLAST search was performed against the CAZy database, where our data suggest that ThABF is more similar to the GH54 family and that this family has several carbohydrate-binding models (CBMs; 1, 2 , 6, 13 and 42), as shown in Supplementary Table S1.

A phylogenetic analysis of ThABF against sequences of the GH54 family (represented mainly by fungal and bacterial sequences) revealed great diversity between the sequences (Fig. 1a). The sequences used in the analysis are in Supplementary Table S2. The phylogenetic tree was divided into five groups that best explained this diversity: a Fungi group, containing only fungi enzymes, and Bacterium I to IV groups, containing bacterial enzymes.

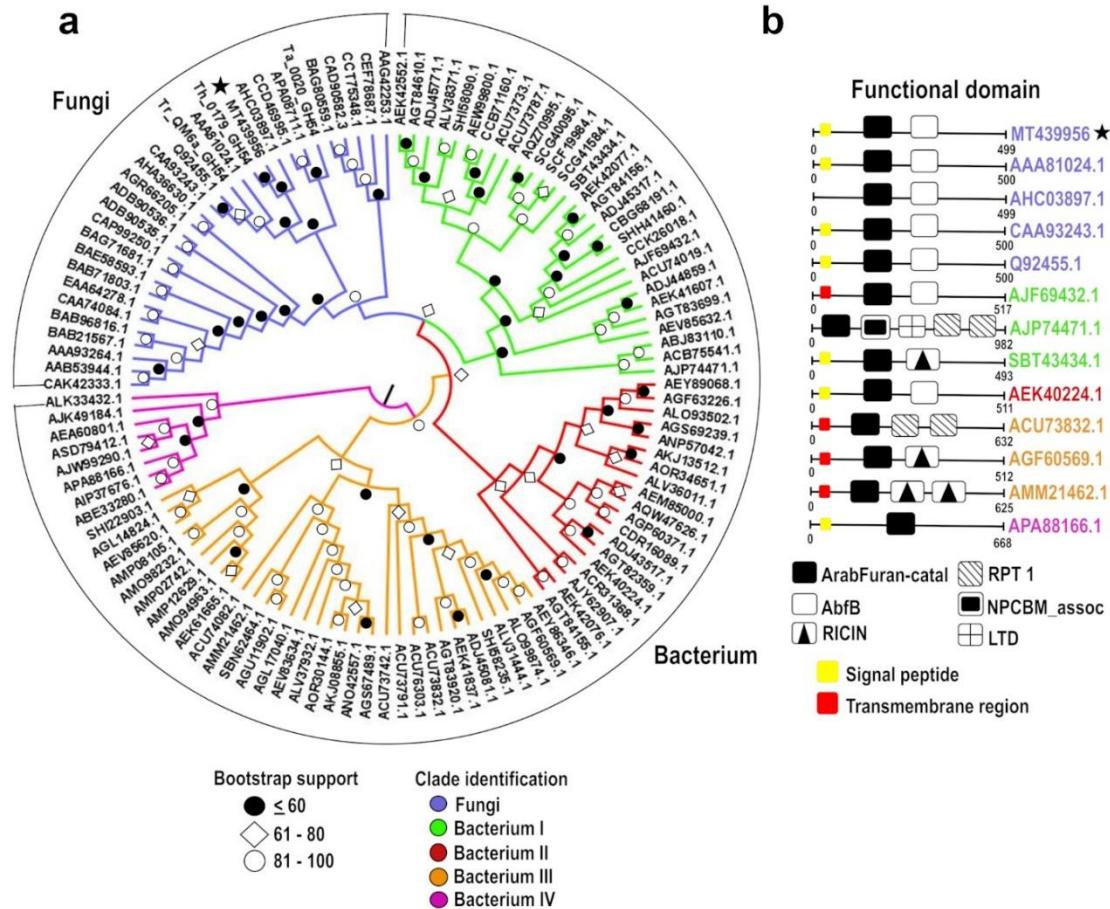


Figure 1. Phylogenetic analysis and sequence domains of the GH54 family. (a) Phylogenetic tree of the sequences. The star indicates the studied protein. Below the tree is a legend with the bootstrap values divided into ≤ 60%, 61-80 and 81-100. (b) Differences in the composition of domains in the sequences distributed in the groups. The identification of the groups is the same for A and B.

The analysis of the domains of the sequences in the phylogenetic tree (Supplementary Table S2) showed that all members of Fungi contained ArabFuran-catal (α -L-arabinofuranosidase B catalytic) and AbfB (α -L-arabinofuranosidase B) domains, in addition to CBM 42 (Fig. 1b). In Bacterium I, the presence of CBMs 13 and 42 was observed, in addition to sequences with the following combinations of domains: ArabFuran-catal and AbfB; ArabFuran-catal, RICIN (ricin B lectin domain), LTD (lamin tail domain), NPCBM_assoc (NEW3 domain of α -galactosidase) and RPT1; and ArabFuran-catal and RICIN. All Bacterium II sequences exhibited only CBM 42 and ArabFuran-catal and AbfB domains.

The Bacterium III group sequences only CBM 13 and the following combinations of domains: ArabFuran-catal and two RPT1 domains; ArabFuran-catal and RICIN; and ArabFuran-catal and two RICIN domains. Bacterium IV was the smallest group; its sequences included no CBMs and were the most distinct from those of the other groups. The Bacterium IV members included only the ArabFuran-catal domain.

To analyze the similarity of the ThABF sequence with some arabifuranosidases characterized as GH54 members in *Trichoderma*, alignment was performed (Fig. 2a). It was found that the target sequence was most similar to a sequence from *T. virens*, with a percent identity of 92.38%, followed by sequences from *T. reesei* (88.38%) and *T. koningii* (87.78%). The structure of the ThABF protein was also predicted, and the model that was used was an *Aspergillus kawachii* α-L-arabinofuranosidase B with a C-score of 0.95, a TM-score of 0.84 ± 0.08 and an RMSD of 5.3 ± 3.4 Å (Fig. 2b).

a

MT439956 - <i>T. harzianum</i> AHC03897.1 - <i>T. virens</i> AAA81024.1 - <i>T. koningii</i> CAA93243.1 - <i>T. reesei</i> Q92455.1 - <i>T. reesei</i>	-MHSRVSKALGLVATSSLVAAGPCDIYSSGGTPCVAHSTTRALYSAYTGPLYQVKRGSDGATTITISPLSSGVANAAAQQSFCAGTTCLITIYYDQSGRGNHLTQAPPGGFGNPESENGYD -MLPPIRTALGLVATGSLATAGPCDIYSSGGTPCVAHSTTRALYSAYTGPLYQVKRGSDGATTITISPLSSGVANAAAQQSFCAGTTCLITIYYDQSGRGNHLTQAPPGGFGNPESENGYD MLSNARIIAACGIAAGSLVAAGPCDIYSSGGTPCVAHSTTRALYSAYTGPLYQVKRGSDGATTITISPLSSGVANAAAQQSFCAGTTCLITIYYDQSGRGNHLTQAPPGGFGNPESENGYD MLSNARIIAACGIAAGSLVAAGPCDIYSSGGTPCVAHSTTRALYSAYTGPLYQVKRGSDGATTITISPLSSGVANAAAQQSFCAGTTCLITIYYDQSGRGNHLTQAPPGGFGNPESENGYD MLSNARIIAACGIAAGSLVAAGPCDIYSSGGTPCVAHSTTRALYSAYTGPLYQVKRGSDGATTITISPLSSGVANAAAQQSFCAGTTCLITIYYDQSGRGNHLTQAPPGGFGNPESENGYD	119 119 120 120 120
MT439956 - <i>T. harzianum</i> AHC03897.1 - <i>T. virens</i> AAA81024.1 - <i>T. koningii</i> CAA93243.1 - <i>T. reesei</i> Q92455.1 - <i>T. reesei</i>	MLASAIAGPVTLNGQKAYGVFISPGTGYRNNAAAGSTAKGDAEGLYAVLDGTHYNGACCFDYGNMAETSRDTGNGHEAIZYGDSTWGTGSGNGPWIADLENLFGSGAVHNNAQDPS MLASAIAGPVMLNGQKAYGVFISPGTGYRNNAAAGSTAKGDAEGLYAVLDGTHYNGACCFDYGNMAETSRDTGNGHEAIZYFGDTWGTGSGGPWIMADLENLFGSGAVHNNAQDPS MLASAIAGPVTLNGQKAYGVFISPGTGYRNNAAAGSTAKGDAEGLYAVLDGTHYNGACCFDYGNMAETSRDTGNGHEAIZYFGDTWGTGSGKPWIMADLENLFGSGSPGNNAQDPS MLASAIAGPVTLNGQKAYGVFISPGTGYRNNAAAGSTAKGDAEGLYAVLDGTHYNGACCFDYGNMAETSRDTGNGHEAIZYFGDTWGTGSGKPWIMADLENLFGSGSPGNNAQDPS MLASAIAGPVTLNGQKAYGVFISPGTGYRNNAAAGSTAKGDAEGLYAVLDGTHYNGACCFDYGNMAETSRDTGNGHEAIZYFGDTWGTGSGKPWIMADLENLFGSGSPGNNAQDPS	239 239 248 248 248
MT439956 - <i>T. harzianum</i> AHC03897.1 - <i>T. virens</i> AAA81024.1 - <i>T. koningii</i> CAA93243.1 - <i>T. reesei</i> Q92455.1 - <i>T. reesei</i>	TSYRFVTAIAVKGQPQNQWIRGGNSASGSLSTFYNGARPASGYNPMSKGATLITGIGDONSNGAQGTYEGWMTSGVPDSATENSVQANIVAAKYAVTLSLTSGPALTGSVSLRATTAC TSYRFVTAIAVKGQPQNQWIRGGNSASGSLSTFYSGARPVTSGYNPMSKEGAIILITGIGDONSNGAQGTYEGWMTSGVPDSATENSVQANIVAAKYAVTLSLTSGPALTGSVSLRATTAC TSYRFVTAIAIKGQPQNQWIRGGNSASGSLSTFYSGARPQVSGYNPMSKEGAIILITGIGDONSNGAQGTYEGWMTSGVPDSATENSVQANIVAAKYAVAPLTSGPALTGSVSLRATTAC TSYRFVTAIAIKGQPQNQWIRGGNSASGSLSTFYSGARPQVSGYNPMSKEGAIILITGIGDONSNGAQGTYEGWMTSGVPDSATENSVQANIVAAKYAVAPLTSGPALTGSVSLRATTAC TSYRFVTAIAIKGQPQNQWIRGGNSASGSLSTFYSGARPQVSGYNPMSKEGAIILITGIGDONSNGAQGTYEGWMTSGVPDSATENSVQANIVAAKYAVAPLTSGPALTGSVSLRATTAC	359 359 366 366 366
MT439956 - <i>T. harzianum</i> AHC03897.1 - <i>T. virens</i> AAA81024.1 - <i>T. koningii</i> CAA93243.1 - <i>T. reesei</i> Q92455.1 - <i>T. reesei</i>	CTTRYIATHGSTVNTQVSSSSATAPKQQA5NTRVAGLGNSCGCFSESKDTPGSVIRHNFFVLLNANDGSQFAEDATFCQAGLNGQGSSLRAWGPTRYFRHYDNVLVASNGGVHT CTTRYIAHTGSTVNTQVSSSSAAALKQQA5NTRVTLGLNSCGCFSESKDTPGSVIRHNFFVLLNANDGSQFAEDATFCQAGLNGQGSSLRAWGPTRYFRHYNNILYASNGGVHT CTTRYIAHGSTVNTQVSSSSATALKQQA5NTRVAGLANNACFSFESQDTGSYVIRHSNFGFLVNANDGSKLFAEDATFCQAGINGQGSSIRSWSYPTRYFRHYNNTLYIASNGGVHV CTTRYIAHGSTVNTQVSSSSATALKQQA5NTRVAGLANNACFSFESRDTGSYVIRHSNFGFLVNANDGSKLFAEDATFCQAGINGQGSSIRSWSYPTRYFRHYNNTLYIASNGGVHV CTTRYIAHGSTVNTQVSSSSATALKQQA5NTRVAGLANNACFSFESRDTGSYVIRHSNFGFLVNANDGSKLFAEDATFCQAGINGQGSSIRSWSYPTRYFRHYNNTLYIASNGGVHV	479 479 480 480 480
MT439956 - <i>T. harzianum</i> AHC03897.1 - <i>T. virens</i> AAA81024.1 - <i>T. koningii</i> CAA93243.1 - <i>T. reesei</i> Q92455.1 - <i>T. reesei</i>	FDATGFSNDDDSVFSVIASAF A 499 FDATSFNDDDSVFSVIASAF A 499 FDATAAFNDDDSVFSVSGGA 500 FDATAAFNDDDSVFSVSGGA 500 FDATAAFNDDDSVFSVSGGA 500	ArabFuran-catal AbfB

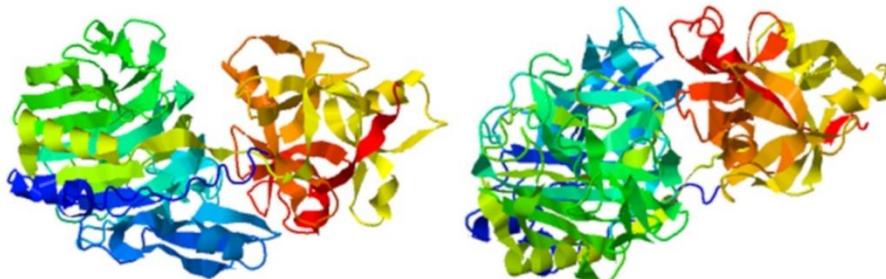
b

Figure 2. Alignment and structure prediction of ThABF. (a) Alignment of ThABF with other *Trichoderma* sequences belonging to the GH54 family. (b) Prediction of the ThABF structure by I-TASSER.

Production of recombinant protein and evaluation of the folding components

With the sequence obtained from ThABF *in silico*, cloning and expression were performed. The enzyme was present in the insoluble fraction (Fig. 3a), and it was necessary to use the refolding approach described in the methodology, through which a protein of 53.44 kDa was purified (Fig. 3b). The recombinant ThABF sequence was used to evaluate the secondary and tertiary folding components of the enzyme since it was obtained through refolding. On the basis of the obtained circular dichroism profile, we identified rich α helices in the ThABF structure (Fig. 3c). In the size exclusion chromatography (SEC), the protein was eluted in a single peak, presenting the characteristics of a monomer in solution (Fig. 3d). Thus, these analyses demonstrated that it was possible to obtain a protein with folding components through the refolding method used.

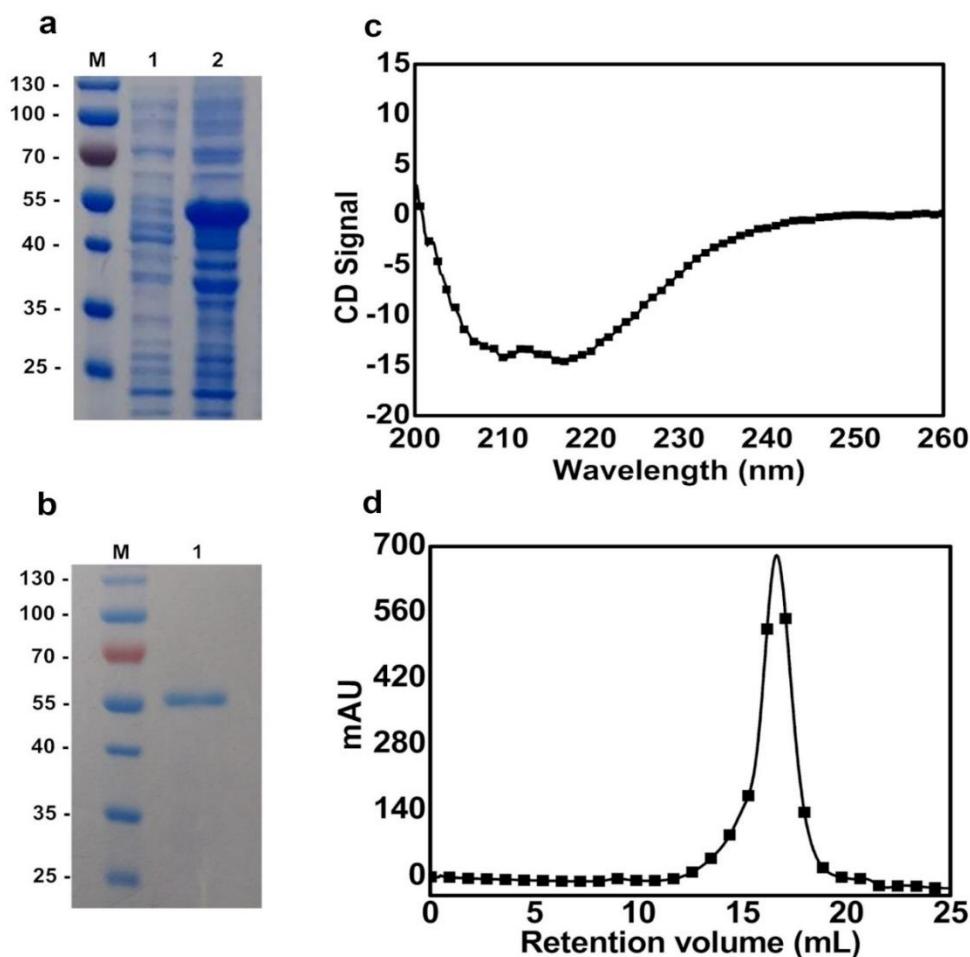


Figure 3. SDS-PAGE in 12% gels showing purification results and graphs of secondary and tertiary folding analyses. (a) Gel containing the soluble (1) and insoluble (2) fractions; M indicates the marker. (b) Gel containing purified ThABF, with a size of approximately 53.44 kDa (1), and a marker (M). Both gels a and b were stained for 30 minutes in a dye solution with coomassie blue. (c) Analysis of secondary folding components of the recombinant enzyme by circular dichroism. (d) Analytical size exclusion chromatography of ThABF.

Determination of enzyme specificity and tests with metal ions

In the first experiment, enzyme activity was tested with the pNPAra substrate because *in silico* analysis indicated that ThABF would present a relatively high affinity with this substrate. However, the detected activity was not significant, leading to the hypothesis that a cofactor is needed to activate the catalytic site of the enzyme. Thus, we evaluated the effects of different metals on the activity of ThABF (Fig. 4). It was observed that the enzyme requires a metal ion cofactor to exert its catalytic activity, since the control exhibited a profile similar to that of ethylene diamine tetraacetic acid (EDTA). CuCl₂ had no effect on the enzyme's activity. MgCl₂ had the greatest effect on the enzyme's activity among the eleven tested metal salts, followed by MnCl₂, CoCl₂, CaCl₂ and NiCl₂. Despite the demonstration that ThABF requires a cofactor, its activity against α -D-arabinofuranoside (pNPAra) was not significantly increased; thus, a panel of chromogenic substrates derived from nitrophenol was tested to obtain information on ThABF enzymatic specificity (Table 1). The enzyme exhibited the greatest activity against β -D-galactopyranoside (pNPG), followed by α -D-arabinopyranoside (pNPAp), β -D-fucopyranoside (pNPF) and finally pNPAra.

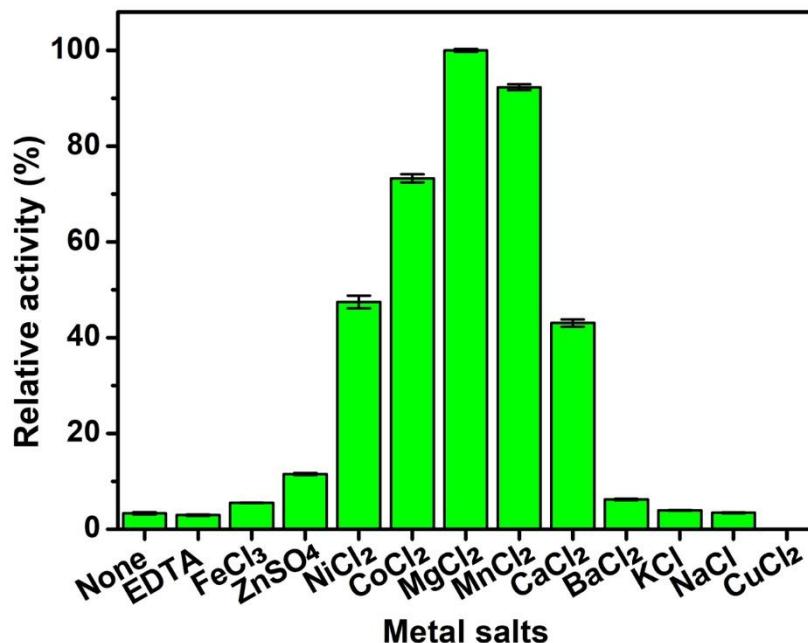


Figure 4. Relative activity of ThABF in the presence of different metallic ions.

Synthetic substrates	Relative activity (%)
β -D-galactopyranoside	100 \pm 0
α -D-arabinopyranoside	61.72 \pm 0
β -D-fucopyranoside	21.79 \pm 0.15
α -D-arabinofuranoside	4.97 \pm 0.13
α -D-glucopyranoside	1.12 \pm 0.03
α -L-fucopyranoside	0.69 \pm 0.60
α -L-ramnopyranoside	0.44 \pm 0.53
β -D-glucopyranoside	0.34 \pm 0.09
α -D-xylopyranoside	0.26 \pm 0.05
α -D-mannopyranoside	0.25 \pm 0.05
α -D-galactopyranoside	0.16 \pm 0.09
β -D-cellobioside	0.09 \pm 0.04
β -D-xylopyranoside	0.07 \pm 0.09
β -D-mannopyranoside	0 \pm 0.05

Table 1. Relative activity of ThABF against different synthetic substrates.

Optimization of pH and temperature conditions and enzymatic kinetics

Following the selection of ThABF-specific substrates and ions, the optimization of pH and temperature was performed under different conditions to identify the ideal performance of the enzyme. As shown in Fig. 5a and b, maximum ThABF activity was observed at 55°C and pH 6.5. The enzyme exhibited greater than 50% activity between temperatures of 45 and 60°C. However, at lower temperatures, in the range of 20 to 40°C, its activity was below 40%. ThABF exhibited a wide range of performance levels at pH values from 6.0 to 9.0, and the optimal pH was shown to be 6.5 in Na₂HPO₄ phosphate buffer. At pH 4 to 5.5, enzyme activity decreased significantly. Experiments were carried out to examine the kinetics of ThABF with the substrates selected in the specificity tests; the parameters Vmax, Km, Kcat and Kcat/Km were calculated, and the results are shown in Table 2. The curves constructed with the substrates pNPG and pNP Ara are shown in Fig. 5c and d.

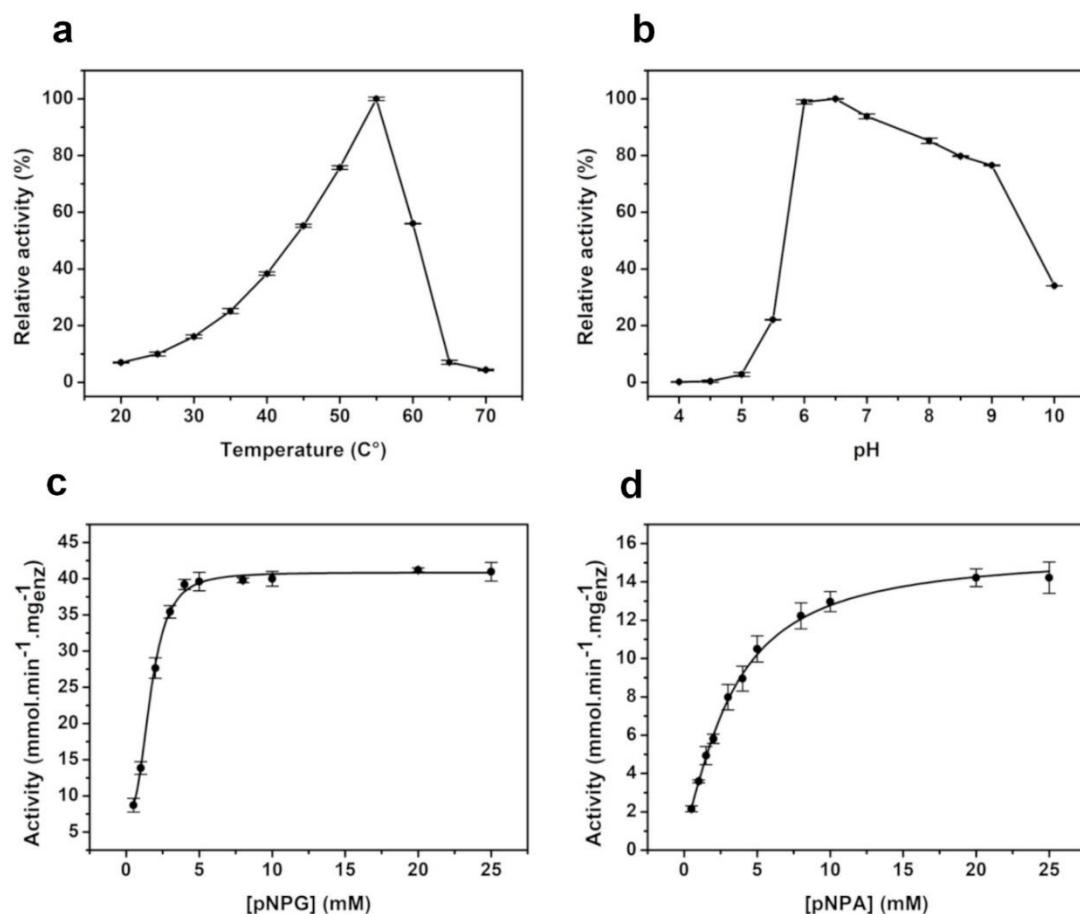


Figure 5. Biochemical characterization of ThABF. Relative activity of the recombinant protein at temperatures of 20 to 70°C (a) and in the pH range of 4 to 10 (b). Enzymatic kinetics of TiBgal54A with the substrates pNPG (c) and pNPA (d).

	Vmax (mmol.min⁻¹.mg⁻¹_{enz})	km (mM)	Kcat (s⁻¹)	Kcat/km (M⁻¹s⁻¹)
pNPG	40.85 ± 0.27	1.68 ± 0.12	35.53	2.11. 10 ⁴
pNPAp	15.47 ± 0.37	3.30 ± 0.16	13.45	4.07. 10 ³
pNPF	41.03 ± 4.02	6.51 ± 1.00	35.68	5.48. 10 ³
pNP Ara	1.85 ± 0.43	6.75 ± 3.24	1.61	2.38. 10 ²

Table 2. Kinetic parameters of ThABF against selected substrates.

Discussion

Arabinofuranosidases are enzymes that, in addition to acting on L-arabinose, can also degrade other types of sugars and can be applied in various industrial processes^{1,14}. In this work, we bioprospected an arabifuranosidase from the metal-dependent GH54 family from *T. harzianum* and demonstrated that it showed expanded activity against other substrates. This enzyme can be introduced in enzymatic cocktails to assist in the hydrolysis of hemicellulose, since this biopolymer is present in approximately 50% of the most diverse sources of plant biomass, and its degradation requires highly diverse enzymes due to its structural complexity¹⁵⁻¹⁸.

Through the *in silico* analyses carried out in this study, great functional diversity was revealed in the GH54 family. Some studies have demonstrated that enzymes of this family exert activity against other types of substrates, in addition to those that are commonly identified, such as β-D-galactofuranoside, α-L-arabinopyranoside and β-D-xylopyranoside^{4,19}. Evidence supporting this conclusion was found in a phylogenetic analysis in which we observed the presence of LTD, RICIN and NPCBM domains (related to the cleavage of simple sugars such as galactose and lactose) in several not yet characterized sequences²⁰⁻²³. We also identified several types of CBMs related mainly to enzymes that act on hemicellulose from both the phylogenetic and BLAST results. CBMs 42 and 13 are the most prevalent in GH54 sequences, possibly because they are present in some arabinofuranosidases and xylanases^{24,25}.

Through the alignment and prediction of the structure of ThABF, we showed that the target protein identified in this study has the characteristics of an arabinofuranosidase, particularly regarding the high conservation of amino acids in the ArabFuran-catal and AbfB catalytic domains^{24,26,27}. The ArabFuran-catal domain was present in all sequences

analyzed in the family. This domain targets several L-arabinose bonds present in hemicellulose, as does AbfB^{24,26}. For the purification of ThABF, it was necessary to use a refolding method adapted to recover the protein from inclusion bodies in its best oligomeric state²⁸. The final protein yield following purification was relatively low, at 0.51 mg/ml, which made it difficult to carry out other types of biochemical tests; however, the method allowed us to characterize the enzyme and study it, as demonstrated in other studies^{28,29}.

In the tests with different substrates, we demonstrated that ThABF exerted activity on the substrates pNPG, pNPAp, pNPF and pNP Ara. As previously described, some GH54 enzymes exhibit activity against substrates other than those reported in this family; however, there have been only two studies addressing this topic^{4,19}. At the time of writing the present report, 19 enzymes had been characterized in this family, among which only 6 have been tested on different types of synthetic substrates^{4,19,30,31}. In addition, as demonstrated via *in silico* analyses, there are several unstudied proteins of this family harboring domains related to the hydrolysis of galactose, among other types of sugars. This demonstrates the limitation of studies involving enzymes of this family of GHs. The data obtained in the tests with metallic cofactors showed that the enzyme was dependent on metals. This role of Mg²⁺ has been reported for other enzymes; there are several hypotheses regarding how magnesium acts on these enzymes, one of which is that it helps maintain the conformation of the active site³². The analysis of the optimal conditions of the enzyme showed that ThABF presents characteristics similar to those of other fungal arabifuranosidases of the GH54 family, making it quite adaptable to changes in the conditions imposed in several types of biotechnological processes, such as the hydrolysis of plant biomass¹.

Based on the results presented above, ThABF has the potential to be used in the saccharification of lignocellulosic material, mainly because it acts on different types of sugars that constitute the side chains of hemicellulose^{33,34}. These side chains contribute to the recalcitrancy of plant biomass due to its diverse structural composition^{2,35}. The removal of hemicellulose makes it less recalcitrant, improving the hydrolysis process^{8,9}. For this purpose, enzymes such as arabinofuranosidases that act on these sugars are necessary³⁴.

In conclusion, the present study reports the *in silico* bioprospection, expression, purification and biochemical characterization of a GH54-dependent metal with expanded substrate specificity that acts in a wide pH range. These characteristics indicate the

potential to use this enzyme in industrial bioprocesses, such as the saccharification of plant biomass.

Methods

In silico data mining, RNA-Seq read mapping and phylogenetic analyses of ThABF

A search for sequences of bifunctional *Trichoderma* enzymes that act on hemicellulose was carried out in the NCBI nonredundant protein³⁶, UniProt and CAZy databases. With the selected sequence, a BLASTP search was performed versus the genome of *T. harzianum* CBS 226.95 (TaxID: 983964) in the NCBI database. The most similar sequence was used to map the RNA-Seq reads of the *T. harzianum* IOC-3844 data generated in the work of Horta *et al.*³⁷ and, thus, obtain the sequence of the specific protein of the target lineage. For mapping, CLC *Genomics Workbench software* (CLC bio - v4.0; Finlandsgade, Dk) was used. Analyses of physical-chemical parameters were also performed with ProtParam; the presence of signal peptides was evaluated on the SignalP-5.0 server; domain prediction was conducted with SMART; and protein structure prediction was conducted with I-TASSER.

A BLASTP search of the sequence resulting from mapping was performed versus the CAZy database, and different enzymes of the GH54 family were then selected to perform phylogenetic analysis. The sequences were aligned using *ClustalW*³⁸ implemented in *Molecular Evolutionary Genetics Analysis software*, version 7.0³⁹. Gaps were removed manually, and incomplete or difficult-to-align sequences were excluded from the analysis. Phylogenetic analyses were performed with *MEGA7* using *maximum likelihood* (ML)⁴⁰ inference based on the Jones-Taylor-Thornton (JTT) model with 1000 *bootstrap*⁴¹ repetitions for each analysis. The tree, drawn to scale, was obtained automatically by applying the *neighbor-joining* BioNJ algorithms to a matrix of distances in pairs estimated using the JTT model and the topology with the highest log probability values. The trees were visualized and edited using the program *Figtree*⁴².

Heterologous production and refolding of TiBgal54A

The ThABF gene was cloned into the pET-28a (+)vector using the standard gene cloning protocol of Sambrook⁴³ with the aid of the forward primer TAAGAATTCTGGGCCCTGTGA and the reverse primer

TGGTCGACTTAAGCAAACTGG. The recombinant plasmids were transformed into *Escherichia coli* Rosetta (Novagen, Darmstadt, Germany) for protein overexpression.

Transformed *E. coli* Rosetta cells were grown to an optical density of approximately 0.8 at 800 nm. To trigger the transcription of the ThABF gene, it was inserted in a culture containing isopropyl β -D-1-thiogalactopyranoside (IPTG) at 0.4 mM. Then, the cells were incubated overnight at 25°C, harvested by centrifugation and resuspended in 50 mL of buffer A (250 mM NaCl, 40 mM sodium phosphate, 20 mM imidazole, 3 mM MgCl₂, 1 mM EDTA and pH 8.0), 1 mg/mL lysozyme and 1 mM phenylmethylsulfonyl fluoride (PMSF) for 30 min with shaking on ice. The cells were disrupted by sonication, and the soluble fraction was obtained by centrifugation (16,000 rpm, 40 min, 4°C).

After rupturing the cells and resuspending the inclusion bodies, the refolding protocol of Santos *et al.*²⁸ was applied. The pellet was resuspended in 15 mL of buffer A containing 1 M urea. Then, sonication and centrifugation (16,000 rpm, 15 min, 4°C) were performed 6 times. Dialysis was subsequently performed under agitation *overnight* with buffer A plus 0.4 M arginine, followed by centrifugation (10,000 rpm, 10 min, 4°C). The concentrations of the purified proteins were determined spectroscopically using the molar extinction coefficient (ϵ) predicted on the basis of the amino acid sequence, and sample purity was estimated by polyacrylamide gel electrophoresis with 12% sodium dodecyl sulfate (SDS-PAGE).

Analysis of folding components

The far-UV CD spectra of ThABF were collected using a Jasco model J-810 spectropolarimeter from the National Biorenewables Laboratory (LNBR) coupled to a Peltier control system (PFD 425S-Jasco). The CD spectra were generated using the purified recombinant protein at a concentration of approximately 1.024 mg/ml in 10 mM sodium phosphate buffer, pH 8.0. A total of 12 accumulations were recorded within the range of 260 to 208 nm at a rate of 50 nm min⁻¹ using a quartz cuvette with a travel length of 1 mm, and the results were averaged.

The tertiary folding components of the recombinant refolded ThABF were evaluated by analytical SEC using a Superdex 200 10/300 GL column (GE Healthcare, Uppsala, Sweden). The protein sample at a concentration of approximately 0.56 mg/ml was dialyzed against buffer A, and gel filtration was then performed at a flow rate of 0.5 mL·min⁻¹. The elution fractions from each chromatographic run were collected and analyzed by 12% SDS-PAGE.

Biochemical characterization

The substrate specificity of ThABF was tested with the following pNP glycosides: α -D-arabinopyranoside (pNPAp), β -D-galactopyranoside (pNPG), α -D-arabinofuranoside (pNP Ara), β -D-fucopyranoside (pNPF), α -D-glucopyranoside, α -D-xylopyranoside, α -D-mannopyranoside, α -L-ramnopyranoside, α -L-fucopyranoside, β -D-glucopyranoside, α -D-galactopyranoside, β -D-xylopyranoside, β -D-celllobioside and β -D-mannopyranoside. A 10 mM concentration of pNPs was used in the tests, which were conducted in 100 mM phosphate buffer (HPO_4), pH 6.5, at 55°C for 10 min. The reactions were stopped with 100 mM calcium carbonate (CaCO_3), and the released pNP was quantified spectrophotometrically at 410 nm.

After selecting the substrates against which ThABF exhibited activity, its activity was tested with different 1 mM concentrations of different metal salts (CaCl_2 , KCl , NaCl , MgCl_2 , MnCl_2 , CoCl_2 , ZnSO_4 , CuCl_2 , NiCl_2 , FeCl_3 and BaCl_2) and EDTA. Before these tests, any metal ions that could interfere with the analysis were removed via two dialysis steps performed on the recombinant protein overnight under agitation. First, 15 mM EDTA was added to chelate the metals present in the sample, and the protein was then transferred to a buffer containing 10 mM sodium phosphate. Subsequently, the tested chemicals were individually incubated with the recombinant enzyme; the reactions occurred at 55°C at pH 6.5 for 10 min with pNPG as the substrate.

After the selection of the ion with the greatest effect on the enzyme, it was used to optimize pH and temperature conditions. The optimal pH was determined by incubating 0.12 mg/ml of purified ThABF with 10 mM pNPG in buffers with pH levels ranging from 3.0 to 10.0 [CH₃COONa sodium acetate buffer (100 mM, pH 3.0 to 5.0), Na₂HPO₄ phosphate buffer (100 mM, pH 6.0 to 8.0) and Tris-NaCl (100 mM, pH 7.0 to 9.0)]. The mixtures were incubated at 55°C for 10 min, and the released pNP was quantified spectrophotometrically at 410 nm. The same conditions were applied to investigate the ideal temperature; these tests were conducted with Na₂HPO₄ phosphate buffer (pH 6.5, 100 mM) at a temperature range of 20 to 70°C. Subsequently, the pH test was repeated at the identified temperature.

After performing the above tests, assays of enzymatic kinetics were performed at pH 6.5 at 55°C for 10 min, with pNPG, pNPAp, pNPF and pNP Ara as substrates. The parameters Km, Vmax, and Kcat and the Kcat/Km ratio were obtained by plotting using the Lineweaver-Burk method, where the plots were constructed by plotting the substrate

concentration on the x axis and the speed of the enzymatic reaction on the y axis using the OriginPro 8.5.0 program. The experiments were performed in triplicate at an ideal temperature and pH. A unit of enzymatic activity was defined as the amount of enzyme required to release 1 µmol of p-nitrophenol per minute under the tested conditions.

References

1. Poria, V., Saini, J. K., Singh, S., Nain, L. & Kuhad, R. C. Arabinofuranosidases: characteristics, microbial production, and potential in waste valorization and industrial applications. *Bioresour. Technol.* **304**, 123019 (2020).
2. Basu, P. Biomass characteristics in *Biomass gasification design handbook* (ed. Basu, P.) 27-63 (Elsevier Inc., 2010).
3. Lagaert, S., Pollet, A., Courtin, C. M. & Volckaert, G. β -Xylosidases and α -l-arabinofuranosidases: accessory enzymes for arabinoxylan degradation. *Biotechnol. Adv.* **32**, 316-332 (2014).
4. Ravanal, M. C. & Eyzaguirre, J. Heterologous expression and characterization of α -l-arabinofuranosidase 4 from *Penicillium purpurogenum* and comparison with the other isoenzymes produced by the fungus. *Fungal Biol.* **119**, 641-647 (2015).
5. Tursi, A. A review on biomass: importance, chemistry, classification, and conversion. *Biofuel Res. J.* **6**, 962-979 (2019).
6. Francocci, F. & Reca, I. B. Composition of plant biomass biotech engineering of cell wall to optimize biofuel production in *Biotransformation of agricultural waste and by-products* (eds. Francocci, F. & Reca, I. B.) 219-236 (Elsevier Inc., 2016).
7. Chen, H. Lignocellulose biorefinery conversion engineering in *Lignocellulose biorefinery engineering* (ed. Chen, H.) 87-124 (Elsevier, 2015).
8. Zoghlami, A. & Paës, G. Lignocellulosic biomass: understanding recalcitrance and predicting hydrolysis. *Front. Chem.* **7**, 874 (2019).
9. Melati, R. B. *et al.* Key factors affecting the recalcitrance and conversion process of biomass. *Bioenergy Res.* **12**, 1-20 (2019).
10. Chandel, A. K. *et al.* Bioconversion of hemicellulose into ethanol and value-added products: commercialization, trends, and future opportunities in *Advances in sugarcane biorefinery* (eds. Chandel, A. K. *et al.*) 97-134 (Elsevier, 2018).

11. Adav, S. S. & Sze, S. K. Trichoderma secretome: an overview in *Biotechnology and biology of trichoderma* (eds. Adav, S. S. & Sze, S. K.) 103-114 (Elsevier, 2014).
12. Druzhinina, I. S. *et al.* Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. *PLoS Genet.* **14**, 1-33 (2018).
13. Horta, M. A. C. *et al.* Transcriptome profile of *Trichoderma harzianum* IOC-3844 induced by sugarcane bagasse. *PLoS One* **9**, 1-17 (2014).
14. Dimarogona, M. & Topakas, E. Regulation and heterologous expression of lignocellulosic enzymes in *Aspergillus* in *New and future developments in microbial biotechnology and bioengineering: aspergillus system properties and applications* (eds. Dimarogona, M. & Topakas, E.) 171-190 (Elsevier B.V., 2016).
15. Yeo, J. Y., Chin, B. L. F., Tan, J. K. & Loh, Y. S. Comparative studies on the pyrolysis of cellulose, hemicellulose, and lignin based on combined kinetics. *J. Energy Inst.* **92**, 27-37 (2019).
16. Biely, P., Singh, S. & Puchart, V. Towards enzymatic breakdown of complex plant xylan structures: state of the art. *Biotechnol. Adv.* **34**, 1260-1274 (2016).
17. Fry, S. C. Cell walls and fibers in *Encyclopedia of applied plant sciences* (ed. Fry, S. C.) 75-87 (Elsevier, 2003).
18. Chen, D. *et al.* Investigation of biomass torrefaction based on three major components: Hemicellulose, cellulose, and lignin. *Energy Convers. Manag.* **169**, 228-237 (2018).
19. Margolles-Clark, E., Tenkanen, M., Nakari-Setälä, T. & Penttilä, M. Cloning of genes encoding alpha-L-arabinofuranosidase and beta-xylosidase from *Trichoderma reesei* by expression in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **62**, 3840-3846 (1996).
20. Hazes, B. The (QxW)3 domain: a flexible lectin scaffold. *Protein Sci.* **5**, 1490-1501 (1996).
21. Spinette, S. *et al.* Ufd2, a novel autoantigen in scleroderma, regulates sister chromatid separation. *Cell Cycle* **3**, 1612-1618 (2004).
22. Hirabayashi, J., Dutta, S. K. & Kasai, K. I. Novel galactose-binding proteins in annelida. *J. Biol. Chem.* **273**, 14450-14460 (1998).
23. Naumoff, D. G. Phylogenetic analysis of -galactosidases of the GH27 family. *Mol. Biol.* **38**, 388-400 (2004).

24. Miyanaga, A. *et al.* The family 42 carbohydrate-binding module of family 54 alpha-L-arabinofuranosidase specifically binds the arabinofuranose side chain of hemicellulose. *Biochem. J.* **399**, 503-511 (2006).
25. Li, S. *et al.* Family 13 carbohydrate-binding module of alginate lyase from *Agarivorans* sp. L11 enhances its catalytic efficiency and thermostability, and alters its substrate preference and product distribution. *FEMS Microbiol. Lett.* **362**, 1-7 (2015).
26. Gielkens, M. *et al.* The abfB gene encoding the major α -L-arabinofuranosidase of *Aspergillus nidulans*: nucleotide sequence, regulation and construction of a disrupted strain. *Microbiology* **145**, 735-741 (1999).
27. Koseki, T. *et al.* Role of Two α -l-arabinofuranosidases in arabinoxylan degradation and characteristics of the encoding genes from shochu koji molds, *Aspergillus kawachii* and *Aspergillus awamori*. *J. Biosci. Bioeng.* **96**, 232-241 (2003).
28. Santos, C. A. & Souza, A. P. Solubilization, folding, and purification of a recombinant peptidoglycan-associated lipoprotein (PAL) expressed in *Escherichia coli*. *Curr. Protoc. Protein Sci.* **92**, e53 (2018).
29. Santos, C. A. *et al.* Characterization of the TolB–Pal trans-envelope complex from *Xylella fastidiosa* reveals a dynamic and coordinated protein expression profile during the biofilm development process. *Biochim. Biophys. Acta - Proteins Proteom.* **1854**, 1372-1381 (2015).
30. de Wet, B. J. M., Matthew, M. K. A., Storbeck, K. H., van Zyl, W. H. & Prior, B. A. Characterization of a family 54 α -l-arabinofuranosidase from *Aureobasidium pullulans*. *Appl. Microbiol. Biotechnol.* **77**, 975-983 (2008).
31. Wan, C. F., Chen, C. T., Li, Y. K. & Huang, L. Expression, purification and characterization of a bifunctional α -l-arabinofuranosidase/ β -D-xylosidase from *Trichoderma koningii* G-39. *J. Chin. Chem. Soc.* **54**, 109-116 (2007).
32. Höhn, S., Virtanen, S. & Boccaccini, A. R. Protein adsorption on magnesium and its alloys: a review. *Appl. Surf. Sci.* **464**, 212-219 (2019).
33. Madeira, J. V. *et al.* Agro-industrial residues and microbial enzymes: an overview on the eco-friendly bioconversion into high value-added products in *Biotechnology of microbial enzymes: production, biocatalysis and industrial applications* (eds. Brahmachari, G., Demain, A. L. & Adrio, J. L.) 475-511 (Elsevier Inc., 2017).

34. Liu, X. & Kokare, C. *Microbial enzymes of use in industry. Biotechnology of microbial enzymes: production, biocatalysis and industrial applications* (Elsevier Inc., 2017).
35. Ebringerová, A., Hromádková, Z. & Heinze, T. Hemicellulose in *Polysaccharides I* (ed. Heinze, T.) 1-67 (Springer-Verlag, 2005).
36. Pearson, W. R. An introduction to sequence similarity ("homology") searching. *Curr. Protoc. Bioinform. Chapter 3*, Unit 3.1 3.1.1-7 (2013).
37. Horta, M. A. C. *et al.* Network of proteins, enzymes and genes linked to biomass degradation shared by *Trichoderma species*. *Sci. Rep.* **8**, 1-11 (2018).
38. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680 (1994).
39. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870-1874 (2016).
40. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275-282 (1992).
41. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Soc. Study Evol.* **39**, 1-15 (1985).
42. Ferreira Filho, J. A., Horta, M. A. C., Beloti, L. L., Dos Santos, C. A. & de Souza, A. P. Carbohydrate-active enzymes in *Trichoderma harzianum*: a bioinformatic analysis bioprospecting for key enzymes for the biofuels industry. *BMC Genom.* **18**, 1-12 (2017).
43. Dong, L., Lv, L. B. & Lai, R. *Molecular cloning. Dong wu xue yan jiu = Zoological research / "Dong wu xue yan jiu" bian ji wei yuan hui bian ji* (Yunnan Ren Min Chu Ban She, 2012).

Acknowledgements

This work was supported by grants from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 2015/09202-0 and 2018/19660-4), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) Computational Biology Program (CBP: Process number 88882.160095/2013-01) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq: Process number 312777/2018-3). MLLM received a Ms fellowship from CAPES (CBP - 88887.200427/2018-00) ; JAFF received a PhD fellowship from CNPq (170565/2017-3) and a PD fellowship from CAPES (CBP - 88887.334235/2019-00); RRM received a PD fellowship from FAPESP (17/14253-9); LMZ recipient of a research fellowship from FAPESP (19/08855-1); CAS received a PD fellowship from FAPESP (2016/19775-0) and an SWE PD fellowship from CAPES (CBP); and APS is the recipient of a research fellowship from CNPq (312777/2018-3). We thank the National Biorenewables Laboratory (LNBR).

Author contributions

M.L.L.M. carried out all the experiments and wrote the manuscript. J.A.F.F. helped with the *in silico* analyses and wrote the manuscript. R.R.M. and L.M.Z. helped with the biochemical characterization. C.A.S. designed the experiments and wrote the manuscript. A.P.S. directed the general study and wrote the manuscript. All authors read and approved the manuscript.

Accession Codes

The protein sequence studied in the study is deposited at the national center for biotechnology information (NCBI) with the access code MT439956.

Additional information

Supplementary Table 1. ThABF Blastp against CAZy bank.

Supplementary Table 2. GH54 sequences used in the phylogenetic tree.

Competing interests

The authors declare that they have no competing interests.

5 CONCLUSÕES GERAIS

De modo geral, este estudo identificou uma ABF de *Trichoderma harzianum* com atuação enzimática diferente das demais enzimas encontradas da família GH54. Nas análises *in silico* foi possível observar que esta família tem uma grande diversidade funcional pouco explorada. Através da caracterização bioquímica a ABF apresentou uma ampla faixa de atuação de pH e temperatura. Isso a torna uma potencial enzima a ser utilizada em processos industriais como a sacarificação da biomassa vegetal. No entanto, mais estudos devem ser conduzidos para compreender melhor a capacidade hidrolítica da ThABF em substratos mais complexos.

6 PERSPECTIVAS

Os resultados obtidos nesta dissertação envolvendo diferentes abordagens biotecnológicas devem contribuir para novos estudos envolvendo a busca e caracterização bioquímica de proteínas.

Este estudo fornece uma base para trabalhos que almejam realizar a bioprospecção *in silico* de sequências alvos e purificá-las utilizando métodos de purificação como o refolding. Além disso, a metodologia empregada no trabalho pode auxiliar no estudo de proteínas não conhecidas como também para o aprimoramento ou modificação de enzimas envolvidas na degradação da biomassa vegetal. Estes processos são fundamentais para tornar mais eficiente o processo de hidrólise enzimática e, consequentemente aumentar o rendimento da produção de bioetanol de segunda geração a partir da biomassa lignocelulósica do Brasil.

7 REFERÊNCIAS BIBLIOGRÁFICAS

- BASU, P. Biomass Characteristics. In: **Biomass Gasification Design Handbook**. First Edit ed. [s.l.] © 2010 Elsevier Inc., 2010. p. 27–63.
- BENOCCI, T. et al. ARA1 regulates not only l-arabinose but also d-galactose catabolism in *Trichoderma reesei*. **FEBS Letters**, v. 592, n. 1, p. 60–70, 2018.
- BIELY, P.; SINGH, S.; PUCHART, V. Towards enzymatic breakdown of complex plant xylan structures: State of the art. **Biotechnology Advances**, v. 34, n. 7, p. 1260–1274, 2016.
- BORIN, G. P. et al. Comparative transcriptome analysis reveals different strategies for degradation of steam-exploded sugarcane bagasse by *Aspergillus niger* and *Trichoderma reesei*. **BMC Genomics**, v. 18, n. 1, p. 1–21, 2017.
- CHANDEL, A. K. et al. Bioconversion of Hemicellulose Into Ethanol and Value-Added Products: Commercialization, Trends, and Future Opportunities. In: **Advances in Sugarcane Biorefinery**. [s.l]: s.n.], p. 97–134.
- CHEN, D. et al. Investigation of biomass torrefaction based on three major components: Hemicellulose, cellulose, and lignin. **Energy Conversion and Management**, v. 169, n. February, p. 228–237, 2018.
- CONTESINI, F. J. et al. Structural and functional characterization of a highly secreted α-L-arabinofuranosidase (GH62) from *Aspergillus nidulans* grown on sugarcane bagasse. **Biochimica et Biophysica Acta - Proteins and Proteomics**, v. 1865, n. 12, p. 1758–1769, 2017.
- DÍAZ PÉREZ, Á. A. et al. Thermodynamic and economic evaluation of reheat and regeneration alternatives in cogeneration systems of the Brazilian sugarcane and alcohol sector. **Energy**, v. 152, p. 247–262, 2018.
- DOS SANTOS CASTRO, L. et al. Comparative metabolism of cellulose, sophorose and glucose in *Trichoderma reesei* using high-throughput genomic and proteomic analyses. **Biotechnology for Biofuels**, v. 7, n. 1, 2014.

- DRUZHININA, I. S. et al. Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus Trichoderma from its plant-associated hosts. **PLOS Genetics**, v. 14, n. 4, p. 1–33, 2018.
- FERREIRA FILHO, J. A. et al. Carbohydrate-active enzymes in Trichoderma harzianum: A bioinformatic analysis bioprospecting for key enzymes for the biofuels industry. **BMC Genomics**, v. 18, n. 1, p. 1–12, 2017.
- FRANCOCCI, F.; RECA, I. B. Composition of Plant Biomass Biotech Engineering of Cell Wall to Optimize Biofuel Production. In: **Biotransformation of Agricultural Waste and By-Products**. [s.l.] Elsevier Inc., 2016. p. 219–236.
- FRY, S. C. **Cell Walls**. Second Edi ed. [s.l.] Elsevier, 2016. v. 1
- GUPTA, V. K. et al. The Post-genomic Era of Trichoderma reesei: What's Next? **Trends in Biotechnology**, v. 34, n. 12, p. 970–982, 2016.
- HASSAN, S. S.; WILLIAMS, G. A.; JAISWAL, A. K. Emerging technologies for the pretreatment of lignocellulosic biomass. **Bioresource Technology**, v. 262, n. April, p. 310–318, 2018.
- HORTA, M. A. C. et al. Transcriptome profile of Trichoderma harzianum IOC-3844 induced by sugarcane bagasse. **PLoS ONE**, v. 9, n. 2, p. 1–17, 2014.
- HORTA, M. A. C. et al. Network of proteins, enzymes and genes linked to biomass degradation shared by Trichoderma species. **Scientific Reports**, v. 8, n. 1, p. 1–11, 2018.
- HU, X. et al. Cloning and expression of a novel α -1,3-arabinofuranosidase from Penicillium oxalicum sp. 68. **AMB Express**, v. 8, p. 2–10, 2018.
- ISIKGOR, F. H.; BECER, C. R. Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. **Polymer Chemistry**, v. 6, n. 25, p. 4497–4559, 2015.
- JORDAN, D. B.; BRAKER, J. D. Opposing influences by subsite - 1 and subsite + 1 residues on relative xylopyranosidase/arabinofuranosidase activities of bifunctional β -D-xylosidase/ α -L-arabinofuranosidase. **Biochimica et Biophysica Acta - Proteins and Proteomics**, v. 1814, n. 12, p. 1648–1657, 2011.

KUBICEK, C. P.; STARR, T. L.; GLASS, N. L. Plant Cell Wall–Degrading Enzymes and Their Secretion in Plant-Pathogenic Fungi. **Annual Review of Phytopathology**, v. 52, n. 1, p. 427–451, 2014.

LAGAERT, S. et al. β -Xylosidases and α -L-arabinofuranosidases: Accessory enzymes for arabinoxylan degradation. **Biotechnology Advances**, v. 32, n. 2, p. 316–332, 2014.

MAITAN-ALFENAS, G. P.; VISSER, E. M.; GUIMARÃES, V. RIA M. Enzymatic hydrolysis of lignocellulosic biomass: Converting food waste in valuable products. **Current Opinion in Food Science**, v. 1, n. 1, p. 44–49, 2015.

MARGOLLES-CLARK, E. et al. Cloning of genes encoding α -L-arabinofuranosidase and β -xylosidase from *Trichoderma reesei* by expression in *Saccharomyces cerevisiae*. **Applied and Environmental Microbiology**, v. 62, n. 10, p. 3840–3846, 1996.

MELATI, R. B. et al. Key Factors Affecting the Recalcitrance and Conversion Process of Biomass. **Bioenergy Research**, v. 12, n. 1, p. 1–20, 2019.

NAKANISHI, S. C. et al. Comparative material balances and preliminary technical analysis of the pilot scale sugarcane bagasse alkaline pretreatment to 2G ethanol production. **Industrial Crops and Products**, v. 120, n. April, p. 187–197, 2018.

POPP, J. et al. Biofuels and their co-products as livestock feed: Global economic and environmental implications. **Molecules**, v. 21, n. 3, p. 1–26, 2016.

RAVANAL, M. C.; EYZAGUIRRE, J. Heterologous expression and characterization of α -l-arabinofuranosidase 4 from *Penicillium purpurogenum* and comparison with the other isoenzymes produced by the fungus. **Fungal Biology**, v. 119, n. 7, p. 641–647, 2015.

ROCHA, V. A. L. et al. Characterization of the cellulolytic secretome of *Trichoderma harzianum* during growth on sugarcane bagasse and analysis of the activity boosting effects of swollenin. **Biotechnology Progress**, v. 32, n. 2, p. 327–336, 2016.

SAINI, J. K.; SAINI, R.; TEWARI, L. Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. **3 Biotech**, v. 5, n. 4, p. 337–353, 2015.

SANTOS, V. E. N.; MAGRINI, A. Biorefining and industrial symbiosis: A proposal for regional development in Brazil. **Journal of Cleaner Production**, v. 177, p. 19–33, 2018.

TURSI, A. A review on biomass: Importance, chemistry, classification, and conversion. **Biofuel Research Journal**, v. 6, n. 2, p. 962–979, 2019.

VAN DYK, J. S.; PLETSCHKE, B. I. A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes-Factors affecting enzymes, conversion and synergy. **Biotechnology Advances**, v. 30, n. 6, p. 1458–1480, 2012.

WANG, W. et al. Elucidation of the molecular basis for arabinoxylan-debranching activity of a thermostable family GH62 α -L-arabinofuranosidase from *Streptomyces thermophilic*. **Applied and Environmental Microbiology**, v. 80, n. 17, p. 5317–5329, 2014.

YANG, Y. et al. A novel bifunctional acetyl xylan esterase/arabinofuranosidase from *Penicillium chrysogenum* P33 enhances enzymatic hydrolysis of lignocellulose. **Microbial Cell Factories**, v. 16, n. 1, p. 1–12, 2017.

YEO, J. Y. et al. Comparative studies on the pyrolysis of cellulose, hemicellulose, and lignin based on combined kinetics. **Journal of the Energy Institute**, v. 92, n. 1, p. 27–37, 2019.

ZOGHLAMI, A.; PAËS, G. Lignocellulosic Biomass: Understanding Recalcitrance and Predicting Hydrolysis. **Frontiers in Chemistry**, v. 7, n. December, 2019.

ANEXO I – DECLARAÇÃO CIBio



**CIDADE UNIVERSITÁRIA "ZEFERINO VAZ",
10 DE ABRIL DE 2021.**

CIBIO: 02 / 2021

IDENTIFICAÇÃO

Mestrado: Maria Lorenza Leal Motta

Programa de Pós-Graduação: Genética e Biologia Molecular – Instituto de Biologia – UNICAMP

PROJETO

" Caracterização bioquímica de uma nova arabinofuranosidase da família GH54 relacionada a degradação de hemicelulose em *Trichoderma harzianum*. "

PARECER

Projeto aprovado pela CIBio / CBMEG sob número 01 / 2010 – Expressão de proteínas de interesse biotecnológico e seu estudo bioquímico, funcional e estrutural.

Coordenador: Profa. Dra. Anete Pereira de Souza.

PROFA. DRA. MÔNICA BARBOSA DE MELO
Presidente da CIBio / CBMEG – UNICAMP

ANEXO II – DECLARAÇÃO DE AUTORIA

Declaração

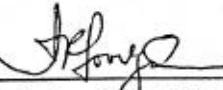
As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **CARACTERIZAÇÃO BIOQUÍMICA DE UMA NOVA ARABINOFURANOSIDASE DA FAMÍLIA GH54 RELACIONADA A DEGRADAÇÃO DE HEMICELULOSE EM TRICHODERMA HARZIANUM**, não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 12 de abril de 2021

Assinatura : 

Nome do(a) autor(a): **Maria Lorenza Leal Motta**

RG n.º 1.611.086

Assinatura : 

Nome do(a) orientador(a): **Anete Pereira de Souza**

RG n.º