

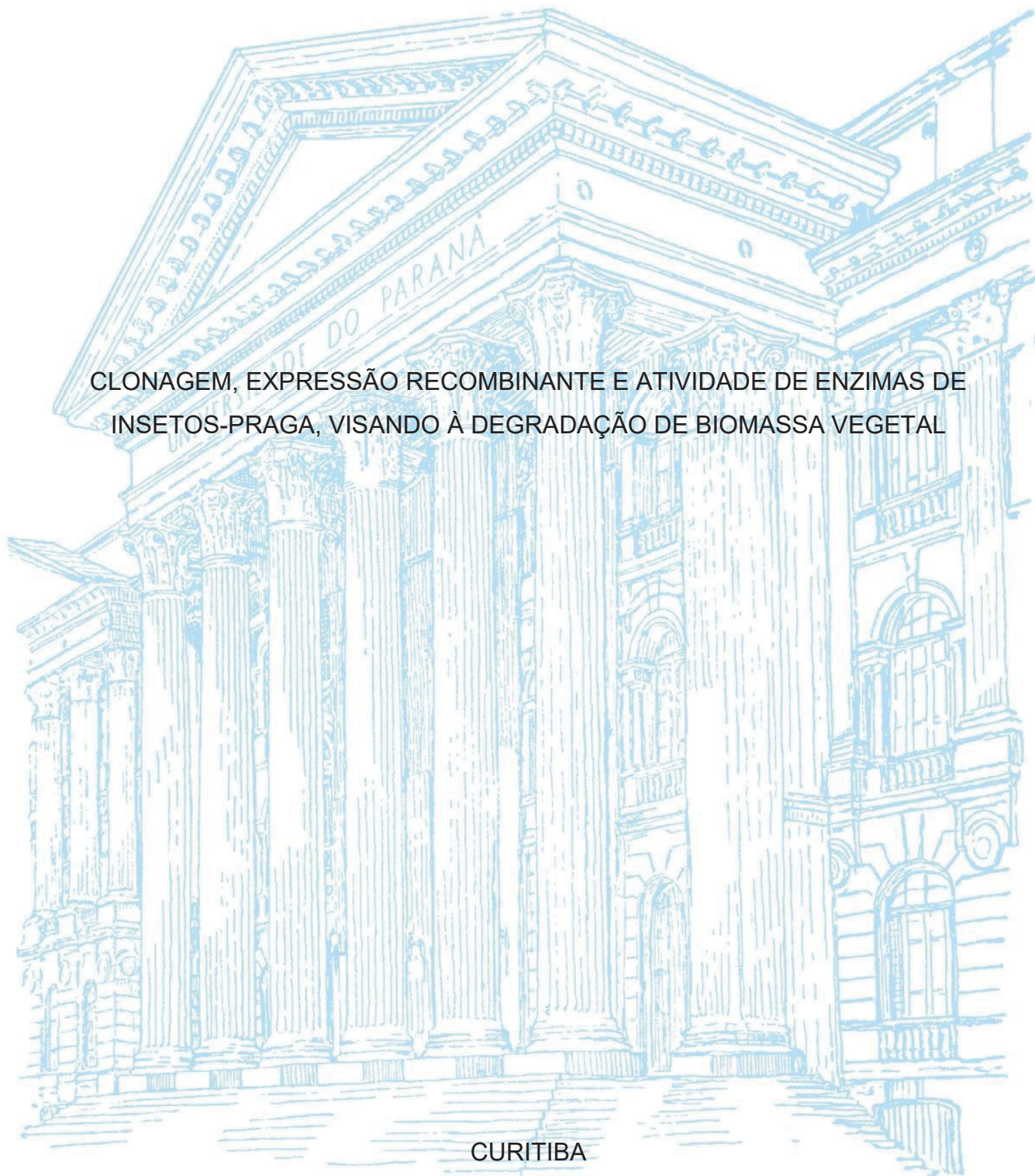
UNIVERSIDADE FEDERAL DO PARANÁ

LIZ NATHALIA IBARRA DUARTE

CLONAGEM, EXPRESSÃO RECOMBINANTE E ATIVIDADE DE ENZIMAS DE  
INSETOS-PRAGA, VISANDO À DEGRADAÇÃO DE BIOMASSA VEGETAL

CURITIBA

2020



LIZ NATHALIA IBARRA DUARTE

CLONAGEM, EXPRESSÃO RECOMBINANTE E ATIVIDADE DE ENZIMAS DE  
INSETOS-PRAGA, VISANDO À DEGRADAÇÃO DE BIOMASSA VEGETAL

Tese apresentada ao Programa de Pós-Graduação em Engenharia de Bioprocessos, Setor de Tecnologia da Universidade Federal do Paraná, como requisito parcial para obtenção do título de Doutor em Engenharia de Bioprocessos e Biotecnologia.

Orientadora: Prof<sup>a</sup>. Dra. Maria Fatima Grossi-de-Sa

Co-orientadores: Prof. Dr. José Dijair Antonino de Souza Júnior.  
Prof. Dr. Érico Augusto Rosas de Vasconcelos.

CURITIBA

2020

Catálogo na Fonte: Sistema de Bibliotecas, UFPR  
Biblioteca de Ciência e Tecnologia

D812c

Duarte, Liz Nathalia Ibarra

Clonagem, expressão recombinante e atividade de enzimas de insetos-praga, visando à degradação de biomassa vegetal [recurso eletrônico] / Liz Nathalia Ibarra Duarte. – Curitiba, 2020.

Tese - Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-Graduação em Engenharia de Bioprocessos, 2020.

Orientadora: Maria Fatima Grossi de Sa.

Coorientadores: Dijair Antonino de Souza Júnior, Érico Augusto Rosas de Vasconcelos.

1. Bicudo-do-algodoeiro. 2. Enzimas. 3. Biomassa vegetal. I. Universidade Federal do Paraná. II. Sa, Maria Fatima Grossi de. III. Souza Júnior, Dijair Antonino de. IV. Vasconcelos, Érico Augusto Rosas de. V. Título.

CDD: 660.28449

Bibliotecária: Vanusa Maciel CRB- 9/1928



MINISTÉRIO DA EDUCAÇÃO SETOR DE  
TECNOLOGIA UNIVERSIDADE  
FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO ENGENHARIA DE  
BIOPROCESSOS E BIOTECNOLOGIA - 40001016036P8

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA DE BIOPROCESSOS E BIOTECNOLOGIA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **LIZ NATHALIA IBARRA DUARTE** intitulada: **CLONAGEM, EXPRESSÃO RECOMBINANTE E ATIVIDADE DE ENZIMAS DE INSETOS-PRAGA, VISANDO A DEGRADAÇÃO DE BIOMASSA VEGETAL**, sob orientação da Profa. Dra. MARIA FATIMA GROSSI DE SA, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 13 de Outubro de 2020.

Assinatura Eletrônica  
15/10/2020 16:35:29.0  
MARIA FATIMA GROSSI DE SA  
Presidente da Banca Examinadora

Assinatura Eletrônica  
14/10/2020 19:43:57.0  
GUILHERME SOUZA PRADO  
Avaliador Externo (EMBRAPA ARROZ E FEIJÃO (CNPAF))

Assinatura Eletrônica  
14/10/2020 18:43:05.0  
MARIA EUGÊNIA LISEI DE SÁ  
Avaliador Externo (EMBRAPA)

Assinatura Eletrônica  
14/10/2020 17:46:55.0  
CARLOS RICARDO SOCCOL  
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

---

Universidade Federal do Paraná- Centro Politécnico - CURITIBA - Paraná - Brasil

CEP 81531-990 - Tel: (41) 3361-3695 - E-mail: [secretaria.pb@ufpr.br](mailto:secretaria.pb@ufpr.br)

Documento assinado eletronicamente de acordo com o disposto na legislação federal Decreto 8539 de 08 de outubro de 2015.

Gerado e autenticado pelo SIGA-UFPR, com a seguinte identificação única: 55656

Para autenticar este documento/assinatura, acesse <https://www.prppg.ufpr.br/siga/visitante/autenticacaoassinaturas.jsp> e insira o código 55656

## RESUMO

O sequenciamento dos transcriptomas dos insetos-praga *Anthonomus grandis* (bicudo-do-algodoeiro) e *Telchin licus licus* (broca-gigante da cana-de-açúcar) possibilitou a identificação e caracterização de novas moléculas potencialmente envolvidas na degradação de biomassa vegetal. Após analisar os *contigs* destes transcriptomas, foram selecionadas para expressão heteróloga moléculas com similaridade de sequências de aminoácidos para enzimas envolvidas na degradação de biomassa vegetal. As moléculas selecionadas foram: (1) uma endo-beta-1,4-glucanase [(EGase), EC 3.2.1.4], da família glicosil-hidrolase 45 (GHF45), proveniente do *A. grandis*, a qual denominou-se AgraGH45-1; (2) uma Aldo-ceto Redutase da família 1 (AKR1), proveniente da *T. licus licus*, denominada TlicAKR-1; e (3,4) duas  $\beta$ -frutofuranosidases (sacarase/invertase) também de *T. licus licus*, que foram denominadas de TlicSuc1 e TlicSuc2. Para validar a função das proteínas codificadas por esses transcritos, foi utilizado um sistema de expressão baseado na levedura *Pichia pastoris* para a expressão heteróloga das proteínas correspondentes estudadas. A atividade de AgraGH45-1 purificada foi detectada no ensaio de degradação de CMC (Carboximetilcelulose) e HEC (Hidroxietilcelulose), nas condições ideais de pH 5,0 a 50° C. Quando comparado à uma celulase comercial de *Aspergillus niger*, a enzima AgraGH45-1 foi 30% mais eficiente para degradar o substrato HEC, nas condições testadas. No caso das sacarases, ambas demonstraram que possuem capacidade de converter sacarose e rafinose em glicose e frutose, com medições realizadas a partir do ensaio de redução de açúcares com DNS. Já para TlicAKR-1, foi realizado o ensaio de atividade avaliando-se diferentes substratos disponíveis, dos quais uma atividade significativa foi encontrada utilizando 2-nitrobenzaldeído; não obstante, a enzima conseguiu uma atividade mínima com aldoses (glicose e xilose), utilizando o cofator NADPH. Os dados obtidos com estas enzimas são promissores, e fazem destas enzimas interessantes candidatas para dar continuidade às pesquisas, visando a suas aplicações na produção de etanol 2G.

Palavras-chave: *Anthonomus grandis*. *Telchin licus licus*. Enzimas. *Pichia pastoris*.  
Expressão heteróloga.

## ABSTRACT

The sequencing of transcriptomes of insect pests *Anthonomus grandis* (boll weevil) and *Telchin licus licus* (sugarcane giant borer) allows the identification and characterization of new molecules potentially used in the degradation of plant biomass. After analyzing the transcripts' contigs, molecules with similar amino acid sequences were selected for enzymes used in the degradation of plant biomass. The selected molecules were: (1) an endo-beta-1,4-glucanase [(EGase), EC 3.2.1.4], from the glycosyl hydrolase family 45 (GHF45), from *A. grandis*, which was named AgraGH45-1; (2) an Aldo Keto Reductase from family 1 (AKR1), from *T. licus licus* that was called TlicAKR-1; and (3.4) two  $\beta$ -fructofuranosidases (sucrase / invertase) from *T. licus licus* as well, which were called TlicSuc1 and TlicSuc2. In order to validate the function of the proteins encoded by these transcripts, an expression system based on the yeast *Pichia pastoris* was used for the heterologous expression of the studied proteins. The activity of purified AgGH45-1 was detected in the degradation assay with CMC (Carboxymethylcellulose) and HEC (Hydroxyethylcellulose), under ideal conditions of pH 5.0 at 50 ° C. Compared to the commercial cellulase of *Aspergillus niger*, AgraGH45-1 was 30% more efficient to degrade HEC substrate under tested conditions. In the case of sucrases, both demonstrate the ability to convert sucrose and raffinose into glucose and fructose, with measurements made from the DNS sugar reduction assay. In the case of TlicAKR-1, activity assay was performed evaluating different available substrates, of which a significant activity with 2-nitrobenzaldehyde was found; nevertheless, an enzyme achieved a moderate activity with aldoses (glucose and xylose), using the NADPH cofactor. The data obtained with these enzymes are promising and make these enzymes interesting candidates to continue the research, aimed at its applications in the production of 2G ethanol.

Keywords: *Anthonomus grandis*. *Telchin licus licus*. Enzymes. *Pichia pastoris*. Heterologous expression.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	8
<b>2</b>	<b>REVISÃO DE LITERATURA</b>	10
2.1	BIOCOMBUSTÍVEIS	10
2.1.1	Biomassa vegetal, lignocelulose e cana-de-açúcar	12
2.2	ENZIMAS ENDÓGENAS DE INSETOS PARA FINS BIOTECNOLÓGICOS	13
2.2.1	Celulases (Endoglucanases)	16
2.2.2	Beta-fructofuranosidases	20
2.2.3	Aldo-Ceto redutases	21
2.3	<i>PICHA PASTORIS</i> ( <i>KOMAGATAELLA PHAFFII</i> ) COMO SISTEMA DE EXPRESSÃO HETERÓLOGA	26
<b>3</b>	<b>JUSTIFICATIVA</b>	28
<b>4</b>	<b>OBJETIVOS</b>	29
4.1	OBJETIVO GERAL	29
4.2	OBJETIVOS ESPECÍFICOS	29
<b>5</b>	<b>MATERIAL E MÉTODOS</b>	30
5.1	MATERIAL BIOLÓGICO	30
5.2	EXTRAÇÃO DE RNA E SÍNTESE DE cDNAs	30
5.3	CLONAGEM E SEQUENCIAMENTO DOS GENES	31
5.4	ALINHAMENTO DAS SEQUÊNCIAS E ANÁLISES DE DADOS	32
5.5	CONSTRUÇÃO DA ÁRVORE FILOGENÉTICA DE <i>TlicAKR-1</i> E ALINHAMENTO DE SEQUÊNCIAS	32
5.6	CONSTRUÇÃO DOS PLASMÍDEOS DE EXPRESSÃO	33
5.7	TRANSFORMAÇÃO GENÉTICA DE <i>P. PASTORIS</i> E ANÁLISE DA INTEGRAÇÃO DOS GENES	34
5.8	EXPRESSÃO EM PEQUENA ESCALA DAS PROTEÍNAS EM <i>P. PASTORIS</i> E WESTERN BLOT	35

5.9	PURIFICAÇÃO DAS PROTEÍNAS .....	36
5.10	ENSAIOS EM PLACA DE DIFUSÃO RADIAL PARA DETERMINAÇÃO DE ATIVIDADE CELULÁSICA PELO AGRAGH45-1 .....	38
5.11	DETEÇÃO DE AÇÚCARES REDUTORES LIBERADOS PELO AGRAGH45-1 .....	38
5.12	AVALIAÇÃO DO PH NA ATIVIDADE DO AGRAGH45-1 E COMPARAÇÃO DA ATIVIDADE E ESTABILIDADE TÉRMICA CONTRA A CELULASE COMERCIAL DE <i>A. NIGER</i> .....	39
5.13	ATIVIDADE ENZIMÁTICA DAS SACARASES .....	40
5.14	ATIVIDADE ENZIMÁTICA DA TlicAKR-1 .....	40
<b>6</b>	<b>RESULTADOS E DISCUSSÃO</b> .....	<b>41</b>
6.1	ANÁLISES DAS SEQUÊNCIAS E ALINHAMENTO DE DADOS.....	41
6.1.1	AgraGH45-1.....	41
6.1.2	TlicAKR-1.....	43
6.1.3	TlicSuc1 e TlicSuc2 .....	47
6.2	ANÁLISES DE EXPRESSÃO GÊNICA POR PCR EM TEMPO REAL .....	49
6.2.1	AgraGH45-1.....	49
6.2.2	TlicAKR-1.....	51
6.2.3	TlicSuc1 e TlicSuc2 .....	52
6.3	CLONAGEM DOS GENES .....	54
6.4	TRANSFORMAÇÃO EM <i>P. PASTORIS</i> , IDENTIFICAÇÃO DAS PROTEÍNAS EXPRESSAS E ATIVIDADE ENZIMÁTICA .....	57
6.4.1	AgraGH45-1.....	58
6.4.2	TlicAKR-1.....	62
6.4.3	TlicSuc1 e TlicSuc2 .....	66
<b>7</b>	<b>CONCLUSÃO</b> .....	<b>69</b>
	<b>REFERÊNCIAS</b> .....	<b>71</b>

## 1 INTRODUÇÃO

A indústria de produção de biocombustíveis vem sendo periodicamente aperfeiçoada com inovações tecnológicas para aumentar a produção, eficiência, e obtenção de energia mais limpa e renovável. Os principais biocombustíveis produzidos no mundo são o etanol de primeira geração (1G) e de segunda geração (2G) (LOPES *et al.*, 2016). As pesquisas envolvendo engenharia enzimática têm se tornado foco das indústrias de produção de álcool. As pesquisas envolvendo o uso de novas celulasas, sacarases e outras enzimas, que possam formar parte do coquetel enzimático que auxilia na produção de etanol 2G (LOPES *et al.*, 2016).

Existem vários estudos direcionados à conversão de polissacarídeos da parede celular de plantas em etanol de segunda geração, desde o desenvolvimento de plantas geneticamente modificadas (GM), com a finalidade de obter parede celular vegetal mais suscetível à degradação enzimática, até a engenharia de enzimas aplicadas para degradar biomassa vegetal, visando a produção de biocombustíveis. (DOBLIN *et al.*, 2014; SOUZA *et al.*, 2018). Grande parte dessas enzimas é isolada de microrganismos, como bactérias e fungos. Os simbiontes, que formam parte do intestino dos insetos, também têm sido alvo de muitos estudos, principalmente para a identificação de potenciais enzimas endógenas de insetos. Alguns relatos reportam que enzimas endógenas de insetos possuem alta atividade após serem expressas em células de insetos, ou seja, em sistema homólogo (LEE *et al.*, 2004; WEI *et al.*, 2006; KIM *et al.*, 2008; WILLIS *et al.*, 2011).

O bicudo-do-algodoeiro (*Anthonomus grandis* Boheman; Coleoptera: Curculionidae) e a broca-gigante da cana-de-açúcar (*Telchin licus licus*, Lepidoptera: Castniidae) podem ser considerados fontes de potenciais genes envolvidos na degradação de biomassa vegetal, dados os danos que estes insetos-praga causam quando se alimentam da biomassa dessas culturas. Já foram descritos genes codificadores deste tipo de enzimas nos genomas de insetos da ordem Lepidoptera e Coleoptera, por exemplo (KUMAR *et al.*, 2008; GLASS *et al.*, 2013). Tais enzimas podem ser utilizadas para pesquisas, visando à produção de biocombustíveis de segunda geração.

A expressão heteróloga de enzimas endógenas de insetos é um fato inovador, porém, tem sido pouco relatado. São necessários mais estudos que envolvam prospecção e avaliação do potencial destas enzimas, focando principalmente em ensaios de durabilidade e eficiência em relação ao tempo de produção e obtenção do biocombustível, utilizando misturas de enzimas de características desejáveis. Como as enzimas de inseto provêm de organismos eucariotos, muitos pesquisadores preferem expressá-las em células de inseto. Porém algumas leveduras, como *Pichia pastoris*, também são excelentes sistemas de expressão para a produção de moléculas eucarióticas recombinantes. As leveduras, diferentemente das bactérias, são capazes de realizar modificações pós-traducionais, como a glicosilação, uma das características desejáveis, já que muitas proteínas provenientes de eucariotos precisam desta modificação para se tornarem funcionais (WEI *et al.*, 2005; WEI *et al.*, 2006). Ainda em comparação com a expressão heteróloga em células de insetos, as leveduras possuem um mecanismo mais simples e menos oneroso de transformação, expressão e manutenção das células, além de serem mais resistentes e de crescerem e se desenvolverem de forma satisfatória em biofermentadores (WU e LETCHWORTH, 2004; MACAULEY-PATRICK *et al.*; 2005; OLIVEIRA *et al.*, 2020).

Com os estudos envolvendo o sequenciamento dos transcriptomas de *T. licus licus* e *A. grandis*, foi possível se ter acesso a uma importante fonte de genes que codificam enzimas capazes de degradar componentes de biomassa vegetal. (FIRMINO *et al.*, 2013; FONSECA *et al.*, 2015). Com base nisso, o presente estudo descreve a clonagem, expressão e caracterização dos genes que codificam uma endo-beta-1,4-glucanase, proveniente do *A. grandis*, e de uma Aldo-ceto Redutase e duas  $\beta$ -frutofuranosidasas, provenientes da *T. licus licus*.

## 2 REVISÃO DE LITERATURA

### 2.1 BIOCOMBUSTÍVEIS

Biocombustível pode ser definido como todo combustível líquido, sólido ou gasoso, produzido a partir de organismos vivos; entre estes cabe destacar o etanol, o biodiesel, o biogás e a lenha. O etanol é um dos principais combustíveis consumido no mundo. Até o momento, existem três tipos de etanol classificados quanto ao processo de produção, (a) o de primeira geração (1G), produzido principalmente a partir de açúcares fermentáveis e amido; (b) etanol de segunda geração (2G), produzido a partir de biomassa vegetal não-fermentável (celulose, hemicelulose e lignina); e (c) etanol de terceira geração (3G), produzido a partir de algas e outros microrganismos. (CARDONA e SÁNCHEZ, 2007; UNICA, 2016).

O etanol é produzido principalmente por meio de processo microbiológico, denominado fermentação alcoólica ou etanólica (biocombustível de 1G). Durante esse processo, os açúcares digeridos pelos microrganismos são convertidos em etanol, energia, biomassa celular, CO<sub>2</sub> e outros subprodutos (MONCEAUX, 2009). Normalmente, é utilizada a levedura *Saccharomyces cerevisiae* para realizar essa função. Açúcares prontamente fermentáveis como sacarose, glicose e frutose são convertidos diretamente em etanol durante a fermentação. No caso dos polissacarídeos como amido, celulose e hemicelulose, estes precisam ser hidrolisados em açúcares simples a serem fermentados. Para converter açúcares em etanol, as leveduras gastam cerca de 7% da energia contida nas moléculas de açúcar, o que significa que aproximadamente 93% da energia, presente nos açúcares, é convertida em etanol (BHATIA, 2014). Esses açúcares mencionados podem ser provenientes de diferentes matérias-primas e resíduos agrícolas. No Brasil, a principal matéria-prima é a cana-de-açúcar, enquanto nos Estados Unidos (EUA) o etanol é produzido principalmente a partir de milho.

Segundo o Instituto de Economia Agrícola (IEA-2020), o Brasil continua liderando o *ranking* como produtor mundial de açúcar e cana-de-açúcar, sendo o responsável pela maior parte do açúcar comercializado no mundo. O país é o segundo maior produtor de etanol no mundo, seguido dos EUA; contudo é o primeiro em exportação do biocombustível. Estas duas nações produtoras de etanol são

responsáveis por aproximadamente 90% da produção mundial deste combustível, desde o início do ano 2000. Este importante comércio internacional contribui para o aumento da segurança energética e ambiental no mundo (MME, 2014).

Há 44 anos, o Brasil vem desenvolvendo tecnologias e gerando avanços científicos que contribuíram significativamente para a evolução da indústria de etanol. Essas importantes contribuições aumentaram nosso conhecimento e compreensão sobre as fermentações do etanol 1G e, mais recentemente, do etanol 2G. Atualmente se tem acesso a novas tecnologias para a produção de etanol a partir de cana-de-açúcar, milho e outras matérias-primas. Foram desenvolvidas e melhoradas as condições de fermentação para reduzir o estresse nas células de leveduras reutilizadas, assim como para também diminuir a contaminação por bactérias e leveduras selvagens. (LOPES *et al.*, 2016). O país conta com uma alta e diversificada produção de *commodities*, as quais têm grande potencial para a geração de energia sustentável por meio do aproveitamento da biomassa. “Biomassa pode ser considerada como qualquer produto biodegradável, o que inclui: 1) Resíduos vegetais da agricultura, silvicultura, indústria de alimentos, etc.; 2) Subprodutos da indústria de polpa e papel; 3) Resíduos municipais separados (IEA, 2001). Muitos resíduos agrícolas, florestais e industriais produzem uma alta densidade de biomassa que podem ser aproveitados, como o caso da cana-de-açúcar, uma das culturas economicamente mais importantes para o país. O processo da produção de açúcar de mesa libera uma grande quantidade de material lignocelulósico a partir dos resíduos industriais, que podem ser eficientemente aproveitados ao converter os polissacarídeos das folhas, das palhas e do bagaço (que representam dois terços da biomassa total) a etanol. O país também conta com a produção de outras importantes culturas, como madeira de eucalipto e cedro, e os das lavouras de café e arroz. Os resíduos dessas *commodities* geram o equivalente em volume de combustível fóssil. (PROTASIO *et al.*, 2013).

Até pouco tempo atrás, a produção do país era limitada à fermentação do melaço de cana. Contudo, desde 2007 os investimentos em pesquisas sobre etanol 2G estimularam a participação de algumas empresas na produção de etanol bagaço de cana e lixo vegetal; vários projetos de financiamento à pesquisa em etanol 2G foram estabelecidos, com a fundação de novos laboratórios e instituições de pesquisa. Esses projetos foram concretizados com o início da produção em larga escala de

etanol 2G, quando em 2014, as três primeiras usinas comerciais de etanol 2G entraram em operação (SOUTINHO *et al.*, 2014). No entanto, a indústria brasileira de etanol 2G vem enfrentado novos desafios científicos e tecnológicos para melhorar o processo de fermentação da produção de etanol no Brasil. Marques (2018) menciona alguns desses obstáculos, como o custo elevado das enzimas lignocelulíticas, que é crucial para melhorar a eficiência e redução do custo das mesmas, a hidrólise de pentoses e tratamento da lignina, a produção de contaminantes para a levedura durante a fermentação, além dos problemas na parte mecânica do processo e nos equipamentos do pré-tratamento.

No entanto, a construção de uma ponte entre a Ciência e a indústria exige investimentos em pesquisa, desenvolvimento e transferência de novas tecnologias para a indústria, além de pessoal qualificado para lidar com novos desafios tecnológicos. Nesse sentido, é comum ocorrer, quando são testadas rotas tecnológicas, um processo de afunilamento até chegar ao processo mais eficiente. (MARQUES, 2018).

### 2.1.1 Biomassa vegetal, lignocelulose e cana-de-açúcar

A biomassa vegetal é produzida por meio da fixação do dióxido de carbono do ar pelas plantas por meio da fotossíntese. Essa fixação consome, anualmente, 15% do carbono da atmosfera, e, desse carbono assimilado, metade entra na composição da lignocelulose, que compõem a parede celular das plantas (SANDGREN *et al.*, 2005). Por conseguinte, a lignocelulose é um material orgânico e o maior componente estrutural da parede celular de todas as plantas. É constituída por frações poliméricas firmes e fibrosas: a celulose, que é fração polissacarídica mais abundante da parede celular vegetal; entremeadas a ela e em menor proporção se encontram a hemicelulose, a lignina e pectinas, além de uma baixa porcentagem de proteínas (Castro, 2006). Essas frações mais importantes representam 97-99% de toda massa seca das plantas. Dependendo do tipo de planta e de seu estágio de crescimento, 35 a 50% do seu peso seco é composto por celulose, 25 a 35% de hemicelulose e 5 a 30% de lignina (LYND *et al.*, 2002). O bagaço de cana-de-açúcar contém aproximadamente 46% de celulose, 25% de hemicelulose, e 21% de lignina, como principais componentes (ABRAMOVAY, 2009).

O bagaço da cana-de-açúcar é um resíduo lignocelulósico da indústria sucro-alcooleira, que compõe aproximadamente 28% do peso seco da planta. Anteriormente, no Brasil, o bagaço era considerado apenas um lixo industrial, mais aproveitado para cogeração de energia ao queimar esses resíduos nas caldeiras das usinas. Depois de começar o sistema mecanizado nas colheitas, a quantidade de tais materiais tem aumentado cada vez mais, pois a indústria sucro-alcooleira tem se tornado uma das principais fontes de renda do país. As indústrias produtoras de açúcar consomem cerca de 340 milhões de toneladas de cana-de-açúcar, e gera como subproduto dessa atividade o bagaço de cana, que chega a totalizar até 95 milhões de toneladas por ano (BUCKERIDGE e GOLDMAN, 2011).

Em 1999, foi lançado no Brasil o projeto “Genoma Cana”, financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), um projeto com fins de melhorar e inovar a produção de etanol 2G. Uma pesquisa realizada por esta instituição, com a finalidade de melhor compreender a estrutura e os componentes da parede celular da cana-de-açúcar, descobriu que o colmo e as folhas da planta contém alto teor de açúcares, em maior proporção para hemiceluloses (50%) do que celulosas (30%), além das pectinas, que representam 10% do peso seco. Os avanços na tecnologia para a produção de etanol de segunda geração estão mais focados na sacarificação da celulose; contudo, são necessárias mais pesquisas e alternativas sobre a degradação de hemicelulose, que possui maior dificuldade de degradação (OLIVEIRA, 2012). As hemiceluloses ficam entrelaçadas com as celulosas, obstruindo o acesso das enzimas celulolíticas. É necessário, ao menos, degradar parcialmente a hemicelulose e desestruturá-la para facilitar a ação destas enzimas, e aumentar a eficiência da degradação da celulose (FARINAS, 2011). A maioria das hemiceluloses está constituída de pentoses como açúcares principais. As xilanas são o tipo de hemicelulose mais abundante. A estrutura destas é composta por moléculas de xilose, principalmente. Estas pentoses, entretanto, não são processadas pelos microrganismos fermentadores (OLIVEIRA, 2012).

## 2.2 ENZIMAS ENDÓGENAS DE INSETOS PARA FINS BIOTECNOLÓGICOS

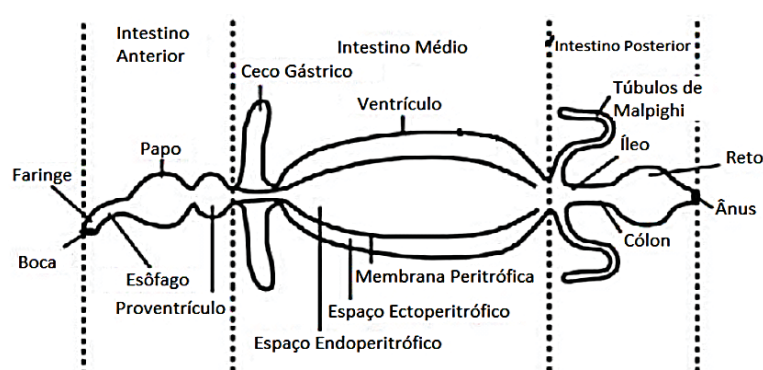
Atualmente, existem várias pesquisas, direcionadas na descoberta e caracterização de enzimas envolvidas na degradação de biomassa vegetal, visando à

produção de etanol. O Laboratório de Interação Molecular Planta-Praga (LIMPP), da Embrapa Recursos Genéticos e Biotecnologia, vem desenvolvendo projetos com esse foco, iniciados com a obtenção de transcriptomas de insetos-praga. As enzimas utilizadas neste estudo provêm de genes escolhidos a partir das sequências nucleotídicas encontradas nos transcriptomas de *T. licus licus* e *A. grandis*. Esses insetos foram alvo de estudo porque, em seu genoma, existem famílias gênicas que codificam enzimas que degradam componentes de biomassa vegetal, o que explica a grande capacidade desses insetos praga de destruírem culturas agrícolas. O bicudo-do-algodoeiro (*A. grandis*.) é a principal praga das plantas de algodão, considerada um dos insetos-praga de maior relevância na cotonicultura, possuindo enzimas que degradam pectina e celulose. O inseto praga *T. licus licus* (broca-gigante-da-cana) é considerado a principal praga da cana-de-açúcar no Nordeste brasileiro, e possui enzimas que degradam oligossacarídeos que são subprodutos das degradações dos polissacarídeos (FIRMINO *et al.*, 2013; FONSECA *et al.*, 2015; OLIVEIRA, 2020, REF). Este inseto é uma praga com uma ascendente importância, pois ataca culturas economicamente importantes, incluindo banana e cana-de-açúcar. A ocorrência desta praga é muito prejudicial em países das Américas do Sul e Central, entre estes, Brasil, Bolívia, Colômbia, Panamá, Trinidad e Tobago e Guiana (NEGRISOLI JR., 2015). No Brasil, é encontrada principalmente nas regiões Norte e Nordeste, nos estados de Alagoas, Pernambuco, Paraíba, Rio Grande do Norte, Pará e Amapá, além de já ter sido detectada, no ano de 2007, em lavouras de Limeira, no estado de São Paulo (BENEDINI e CONDE, 2008). As lagartas desta espécie conseguem chegar a dez centímetros de comprimento. Possuem um ciclo de vida de dez meses, se desenvolvendo nos colmos da cana-de-açúcar, criando galerias à medida que avançam e se alimentam. Quando atacam as plantas novas, causam a morte da gema apical, sintoma conhecido por coração morto, o qual causa apodrecimento e secagem dos tecidos da planta. Além disso, os danos indiretos provocados pela entrada de microrganismos oportunistas, principalmente fungos (*Fusarium moniliforme* e/ou *Colletotrichum falcatum*), que causam perdas às indústrias, como a inversão da sacarose e diminuição da pureza do caldo, resultam em menor rendimento na produção de açúcar e etanol, além da redução da biomassa (GALLO *et al.*, 2002).

O bicudo-do-algodoeiro, *A. grandis*, é considerado a praga de maior risco e habilidade para prejudicar a produção de fibras de algodão (MIRANDA e

RODRIGUES, 2015). Esta espécie é reconhecida pelos danos que causa, pelas dificuldades de seu controle e alta capacidade reprodutiva (BELOT, 2015). Detectado pela primeira vez no México em 1843, de onde se espalhou pelos EUA e outros países até chegar no Brasil, em 1983. Esta praga introduzida no Brasil, desde então, causa enormes prejuízos econômicos (MIRANDA e RODRIGUES, 2015). O problema se inicia com a oposição da fêmea nos botões florais, a qual precisa se alimentar de pólen para poder desenvolver seus ovos. Após a eclosão dos ovos, as larvas sobrevivem na estrutura floral, alimentando-se dessa estrutura e causando prejuízos. Chegando no estágio adulto, os besouros emergem e continuam alimentando-se dos botões florais (AZAMBUJA e DEGRANDE, 2014).

É suposto que ambos os insetos (a broca-gigante-da-cana e o bicudo-do-algodoeiro) possuem uma grande capacidade de degradar biomassa vegetal. Já no processo mastigatório, começa a degradação mecânica do material lignocelulósico, e, logo depois, no sistema digestivo, mediante a ação das enzimas lignocelulíticas (TERRA e FERREIRA, 1994). Terra e Ferreira (1994), explicam sobre o sistema digestivo generalizado dos insetos, o qual consiste no intestino anterior (cavidade bucal, faringe, esôfago, papo, pró-ventrículo); intestino médio (ceco gástrico, ventrículo, membrana peritrófica, espaço ectoperitrófico e endoperitrófico); e intestino posterior (túbulos de Malpighi, íleo, cólon, reto e ânus) (FIGURA 1)



**Figura 1.** Modelo generalizado do intestino dos insetos. A imagem apresenta as três divisões da estrutura digestiva de um inseto, consistindo em intestino anterior, intestino médio e intestino posterior (Figura adaptada de Terra e Ferreira, 1994).

Dentro do sistema digestivo, o intestino médio é o lugar onde são sintetizadas enzimas endógenas lignocelulolíticas, especialmente as  $\beta$ -1,4-endoglucanases. Estas

enzimas foram confirmadas em algumas ordens taxonômicas de insetos, especialmente a ordem Coleoptera (MARTIN, 1983; SCRIVENER *et al.*, 1997; GIRARD e JOUANIN, 1999; WATANABE e TOKUDA, 2001; NAKASHIMA *et al.*, 2002; SUGIMURA *et al.*, 2003; WEI *et al.*, 2006; TOKUDA e WATANABE, 2007; ZHOU *et al.*, 2007; Li *et al.*, 2012).

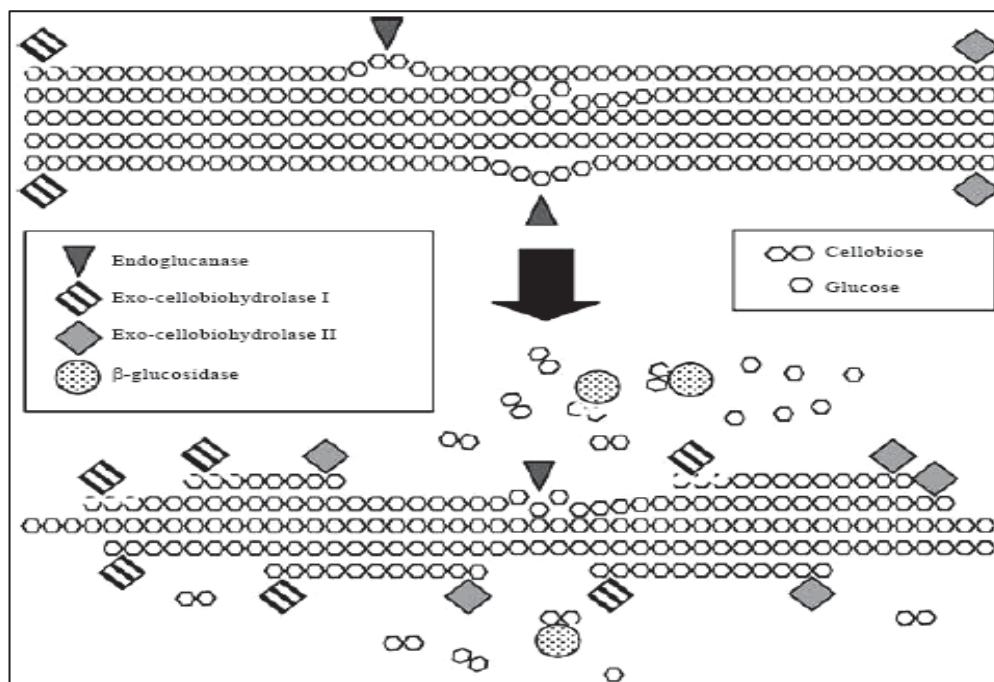
Existem evidências de bactérias, fungos e protistas, que possuem enzimas celulíticas, além de alguns animais vertebrados e invertebrados, que possuem microrganismos simbióticos no intestino, capazes de degradar biomassa (WATANABE e TOKUDA, 2010). No caso dos insetos, antes se acreditava que estas enzimas provinham apenas de bactérias e protozoários simbiontes do trato intestinal, porém, existem registros de várias enzimas endógenas, produzidas pelo mecanismo simbiose–independente, ou seja, produzidas pelo próprio inseto (CALDERÓN-CORTÉS *et al.*, 2012); e que provavelmente foram inseridos por transferência horizontal de genes (HGT) (LI *et al.*, 2011; ZHU *et al.*, 2011). A razão pela qual insetos, em especial alguns lepidópteros, têm uma habilidade rara de adquirir e fixar material genético de outros organismos ainda não é clara. No entanto, genes transferidos horizontalmente por bactérias podem ter contribuído para novas funções para os insetos hospedeiros, para se adaptar a várias dietas e nichos (LI *et al.*, 2011). Várias enzimas endógenas de insetos foram descobertas e avaliadas por meio da disponibilidade de dados genômicos, como as bibliotecas de cDNAs e ESTs (Expressed sequence tag). Tendo acesso a essas informações, é importante a continuidade das pesquisas na descoberta e caracterização de novas enzimas-alvo. (WILLIS *et al.*, 2010).

### 2.2.1 Celulases (Endoglucanases)

A parede celular vegetal é composta principalmente por celulose (20 – 40%), o polímero de carboidrato mais abundante produzido pelas plantas. A celulose é constituída por unidades repetitivas de glicose, unidas entre si por ligações  $\beta$ -1,4-glucosídicas (WEI *et al.*, 2006). Segundo dados da literatura, existem dois tipos de mecanismos de degradação da parede celular vegetal: o sistema hidrolítico e o sistema oxidativo (SCHARF e TARTAR, 2008). No caso do sistema hidrolítico, são

necessárias três enzimas do tipo celulases para quebrar essas ligações, até atingir a degradação completa da celulose: endoglucanases (EC 3.2.1.4), exoglucanases (EC 3.2.1.91), e celobiasas (EC 3.2.1.21) (WEI *et al.*, 2006). O sistema oxidativo, que será explicado mais adiante, é o responsável pela despolimerização da lignina e diminuição da recalcitrância da holocelulose.

As endoglucanases, denominadas 1,4- $\beta$ -D-glucana-4-glucanoidrolases por seu nome sistemático, segundo a IUBMB – International Union of Biochemistry and Molecular Biology -, são as celulases pioneiras na digestão da celulose. Estas agem em regiões randômicas da cadeia, hidrolisando as ligações glicosídicas  $\beta$ -1,4 intramoleculares, diminuindo o grau de polimerização e deixando exposta a região cristalina da celulose. As exoglucanases são o segundo tipo de celulase que age nessa região cristalina exposta, liberando celobiose, desde os extremos terminais, podendo ser exoglucanases, que clivam extremos redutores e as que clivam os extremos não-redutores. Finalmente, as  $\beta$ -glucosidases hidrolisam essas moléculas de celobiose e liberam monômeros de glicose, que são usados na fermentação e, conseqüentemente, na produção de etanol (FIGURA 2). As exoglucanases e as celobiasas são inibidas por seus próprios produtos de hidrólise: celobiose e glicose, respectivamente (ZHANG *et al.*, 2006).



**Figura 2.** Completa degradação da celulose até a obtenção de glicose. Primeiramente, endoglucases catalisam as regiões mais superficiais e amorfas (representadas por triângulos grafite), enquanto as exoglucanases de Tipo I (losangos listrados) nos extremos redutores e as de Tipo II (losangos cinza-claros) nos extremos não redutores, degradam a região cristalina da celulose, que ficou exposta pelas endoglucanases. As celobioses (representadas por círculos com pontos) catalisam moléculas de celobiose, liberadas pelas exoglucanases, convertendo-as em glicose (CARVALHO, 2009).

As enzimas celulolíticas, que participam na sacarificação da PCV, são classificadas com base na similaridade das sequências de DNA e comparações feitas no banco de dados, denominado Carbohydrate-Active EnZymes (Cazy). A família de glicosil hidrolases (GHs) contém o maior número de enzimas, que formam parte do processo da degradação da parede. (CANTAREL *et al.*, 2009). Muitas enzimas endógenas de insetos pertencem à família GH9. Estas famílias de celulases possivelmente provêm de um ancestral comum entre os invertebrados, que se refere à transferência vertical de genes (FISCHER *et al.*, 2013), pois está amplamente distribuída entre várias ordens de insetos. Também cabe ressaltar a grande diferença entre sequências de GH9 de insetos e sequências de celulases de micro-organismos simbiotes. Não é o caso das celulases das famílias GH5 e GH45, que diferem estruturalmente das GH9, e têm sido encontradas apenas na ordem dos coleópteros (LEE *et al.*, 2004; WEI *et al.*, 2005; FISCHER *et al.*, 2013). Porém ainda não existe uma certeza do mecanismo de herança das famílias GH5 e GH45, podendo ter ocorrido através de um ancestral comum dos coleópteros, ou por herança genética de

outros organismos, mais distantes filogeneticamente (transferência horizontal de genes). (FISCHER *et al.*, 2013).

As endoglucanases são enzimas monômeras em sua maioria, podem ter glicosilação baixa ou nula. Geralmente reagem com pH favorável entre 4,0 e 5,0 e temperatura entre 50 e 70°C (CASTRO, 2006). O mesmo autor menciona que estas enzimas possuem sítio ativo em formato de fenda, o que contribui para atuar ao longo da cadeia com mais eficiência, especialmente nas regiões de menor organização estrutural. Estas regiões não possuem ligações tão fortes quanto as regiões mais cristalinas e internas da celulose, que têm interações intermoleculares de hidrogênio bem firmes. Geralmente, a estrutura das endoglucanases está aderida por dois domínios distintos, o domínio de ligação à celulose (CBD - *Cellulose Binding Domain*) e o domínio catalítico (CD - *Catalytic Domain*), que abriga o sítio ativo. Ambos os domínios estão ligados por meio de uma cadeia polipeptídica flexível (*linker*) O-glicosilada e rica em prolinas, serinas e treoninas. No domínio catalítico, onde se encontra o sítio ativo, acontece a hidrólise das ligações glicosídicas da celulose. Os sítios ativos de cada classe de celulase podem variar, e dessa forma acontece a hidrólise de ligações, que se encontram em regiões distintas do substrato. O domínio de ligação à celulose possibilita manter uma concentração de celulasas na superfície da celulose, por meio de interações não-covalentes, como ligações de hidrogênio, forças eletrostáticas e interações hidrofóbicas (MARTINS, 2005). Existem celulasas que contêm apenas o domínio catalítico sem CBD, e estas conseguem se ligar a celulose com menor afinidade, diferentemente das que possuem os dois domínios, onde a afinidade é mais eficiente. Estudos relatam que o comprimento da cadeia polipeptídica flexível é importante para a atividade enzimática. Esta promove flexibilidade e permite independência dos dois domínios. (SRISODSUK *et al.*, 1993). As celulasas em geral possuem grande relevância nas indústrias de produção. Ocupam o terceiro lugar de enzimas produzidas industrialmente, e são necessárias no processamento de algodão, reciclagem de papel, formulação de detergente, extração de sucos, bioestonagem e biopolimento de tecidos, extração de óleo vegetal, por exemplo (WILSON, 2009).

Os organismos mais utilizados na produção destas enzimas são os fungos filamentosos, principalmente aqueles dos gêneros *Aspergillus*, *Trichoderma*, *Humicola*, *Penicillium* e *Phanerochaete* (SINGHANIA *et al.*, 2010). As empresas

Novozymes e Genencor Danisco, reconhecidas na produção comercial de enzimas, produzem celulasas para tratamento biológico da biomassa. O desafio com este sistema para tratar a biomassa é o custo muito alto destas celulasas comerciais, além da dificuldade para serem produzidas em grandes quantidades; o que muitas vezes não compensa o custo de produção do etanol 2G, podendo chegar a 18% do custo operacional de uma planta (SINGHANIA *et al.*, 2010).

Por muitos anos têm sido estudadas soluções para estes inconvenientes. Pesquisas vêm sendo realizadas na busca por tecnologias melhoradas. Uma das medidas necessárias seria encontrar novas enzimas com probabilidade de serem mais eficientes e produtivas, com características desejáveis para a indústria e, especialmente, conseguir produzi-las em abundância, de forma rápida e simples.

### 2.2.2 Beta-fructofuranosidases

No processo da fotossíntese, um dos principais produtos da planta é a sacarose, um dissacarídeo conformado por glicose e frutose, unidas por ligações  $\alpha$ -1,2-glicosídicas. Este açúcar fornece energia aos organismos vivos, podendo-se assimilar e aproveitar facilmente, sendo de vital importância para o metabolismo. A sacarose pode ser hidrolisada por dois tipos de enzimas sacarases,  $\alpha$ -glucosidase (EC 3.2.1.20) e  $\beta$ -frutofuranosidase ( $\beta$ -FFase, EC 3.2.1.26). (GAN *et al.*, 2018). A diferença entre estas está no mecanismo de hidrólise, enquanto as  $\beta$ -frutofuranosidases são específicas para a porção  $\beta$ -frutosil, as  $\alpha$ -glicosidases são específicas para a região  $\alpha$ -glicosil do substrato (TERRA e FERREIRA, 1994). As  $\beta$ -FFase ou invertases formam parte da família das glicosil-hidrolase - GH32, que se caracterizam por três regiões conservadas, cada uma contendo seus resíduos chave essenciais: Asn-Asp-Pro-Asn-Gly (NDPNG), Arg-Asp-Pro (RDP) e Glu-Cys (EC) (PONS *et al.*, 2004). Estas enzimas não só hidrolisam sacarose, como também rafinose; ou seja, substratos de  $\beta$ -D-frutofuranoides. Desta forma, a enzima libera  $\beta$ -frutose, a partir dos terminais não-redutores das cadeias. (LINDE *et al.*, 2009). Existem registros de  $\beta$ -FFases caracterizadas em microrganismos e plantas, mas a presença da enzima no reino animal é pouco conhecida. Foram encontradas em lepidópteros (SANTOS e TERRA, 1986; SUMIDA *et al.*, 1994; CARNEIRO *et al.*, 2004); e só mais recentemente se descobriu transcritos que codificam  $\beta$ -frutofuranosidases no intestino

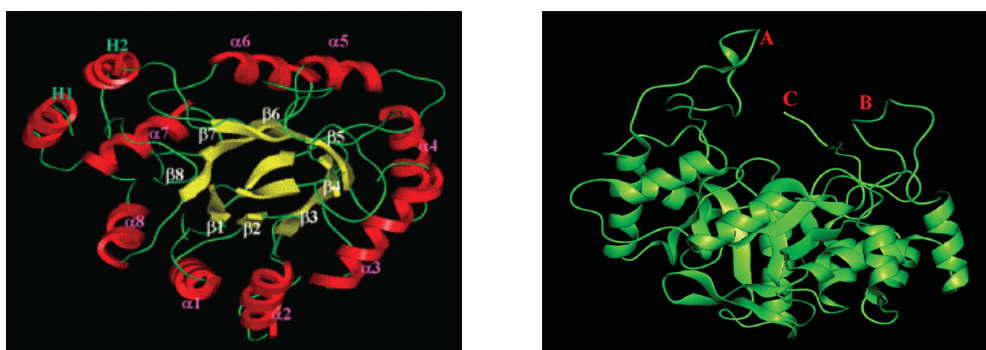
médio da *Helicoverpa armigera* (PAUCHET *et al.*, 2008), assim como no intestino de *Bombyx mori* e na glândula de seda das larvas (DAIMON *et al.*, 2008).

Posteriormente, outras  $\beta$ -frutofuranosidases foram descobertas no transcriptoma de outra lepidóptera, *Manduca sexta* (PAUCHET *et al.*, 2010), e no coleóptero *Dendroctonus ponderosae* (KEELING *et al.*, 2013). E mais tarde, a partir de projetos realizados do genoma e transcriptoma de *Agrius planipennis* (ZHAO *et al.*, 2014), e com pesquisas desenvolvidas por Peduzzi e colaboradores, no besouro *Sphenophorus levis*. (PEDIZZI *et al.*, 2014). Esses estudos confirmaram a presença de uma  $\beta$ -FFase endógena, diferente de outros relatos que não confirmaram se a atividade sacarase acontecia no sistema digestivo próprio do inseto ou da sua microbiota. Além disso, existe a hipótese que a hidrólise de sacarose dos ensaios enzimáticos pude ter acontecido com  $\alpha$ -glicosidases, já que não foram feitos ensaios específicos para  $\beta$ -frutofuranosidase. (DEVI e SINGH, 2011). Daimon e colaboradores (2008) relataram que a atividade de enzimas sacarases acontece principalmente por  $\alpha$ -glicosidases. Como as  $\beta$ -FFases foram encontradas apenas em coleópteros e lepidópteros, é possível que esta enzima seja produzida no organismo destes insetos, como mecanismo de adaptação alimentar.

### 2.2.3 Aldo-Ceto redutases

As AKRs são uma superfamília de enzimas oxirredutases, de estrutura similar entre seus membros que provêm de um ancestral comum. Caracterizam-se pelo mecanismo de reação, que catalisa a oxido-redução, dependente de nicotinamida adenina dinucleotídeo (fosfato) (NAD(P)H), de grupos carbonila e seus respectivos metabólitos alcoólicos. Essas reações redox estão envolvidas na biossíntese, metabolismo intermediário e desintoxicação; podendo atuar com uma ampla classe de substratos como glicose, esteroides, produtos finais de glicosilação, produtos de peroxidação lipídica e poluentes ambientais. (JEZ *et al.*, 1997; BARSKI *et al.*, 2008). Estas proteínas contêm um  $(\alpha\beta)_8$  ou motivo TIM-barrel (triosfosfato isomerase) (FIGURA 3A), onde acontece o reconhecimento e ligação com o (s) substrato (s). O sítio ativo das AKRs se encontra na extremidade C-terminal do barril. Devido a esta estrutura, as AKRs permitem prender uma gama diversa de substratos carbonílicos

(ELLIS, 2002). Esta ligação com o substrato acontece por meio da interação com aminoácidos dentro dos três *loops*, que se projetam para fora da estrutura do barril (Figura 3B). Esses *loops* dão uma característica especial a cada sequência de AKR, pois a conservação da sequência dos membros é mais baixa nesses *loops*, em comparação com as  $\alpha\beta$ -hélices que formam a própria estrutura do barril. (MINDNICH e PENNING, 2009).



**Figura 3.** A) Estrutura da  $(\alpha\beta)_8$  ou motivo TIM-barrel de uma AKR. B) três loops A, B e C do TIM-barrel. Dados extraídos do banco de dados de membros existentes e potenciais da Superfamília AKR. <https://hosting.med.upenn.edu/akr/barrel-motif/>

A Escola de Medicina Perelman da Universidade de Pensilvânia estabeleceu um banco de dados online de membros existentes e potenciais da Superfamília AKR, com dados atualizados até o presente ano. Segundo esses dados, a definição da nomenclatura das AKRs utiliza um padrão que compara as sequências de aminoácidos de cada membro. Sendo assim, os novos nomes de AKR são: o símbolo raiz "AKR" para Aldo-Keto Reductase; um número árabe que designa a família; uma letra indicando a subfamília, quando existem várias subfamílias; e um numeral arábico representando a sequência proteica única. Sob esse sistema, a proteína AKR1A1 seria a primeira AKR da família 1, subfamília A e, neste caso, corresponde à aldeído-redutase humana. Os membros de uma família AKR devem ter <40% de identidade de aminoácidos com qualquer outra família. Dentro de uma determinada família, as subfamílias podem ser definidas por uma identidade >60% na sequência de aminoácidos entre os membros da subfamília. Por essa definição, 9 das 14 famílias de AKR incluem várias subfamílias. Como exemplo, a família AKR1 inclui as seguintes subfamílias: A) aldeído-redutase de mamífero; B) aldose redutase de mamífero; C) desidrogenases hidroxisteróides (HSDs); e D) 4-3-cetosteróide-5 $\beta$ -redutase. A

numeração dos membros conhecidos de cada subfamília foi atribuída de forma arbitrária. Por exemplo, AKR1A1, AKR1A2 e AKR1A3 são as aldeído-redutases de humanos, porcos e ratos, respectivamente. Quaisquer novas adições a uma subfamília são numeradas cronologicamente.

As AKRs catalisam várias reações metabólicas em diferentes organismos, podendo ter mais de uma função específica entre elas: degradação de carboidratos, desintoxicação xenobiótica, degradação de éteres  $\beta$ -arílicos da lignina e várias aplicações industriais e clínicas (BARSKI *et al.*, 2008; FELDMAN *et al.*, 2015). Como reagem com uma grande diversidade de substratos, o que inclui a maioria dos aldeídos biológicos, uma das funções características destas enzimas é a conversão e desintoxicação de aldeídos e cetonas, produzidas pelo próprio organismo ou obtida no ambiente em forma de nutriente, medicamento ou toxina. (BACHUR, 1976). As AKRs são um tipo de enzimas PADs (Pró-oxidantes, Antioxidantes e Detoxificantes), que promovem a degradação e/ou eliminação de toxinas endógenas e exógenas, assim como a formação ou degradação de espécies reativas de oxigênio (EROs). As enzimas superóxido dismutases, peroxidases, catalases, P450, e aldeído-desidrogenases, também são exemplos de PADs (LIU *et al.*, 2016). Existem estudos que confirmam a capacidade das AKRs para degradar carboidratos, além de promover a degradação da lignina, bem como as aryl-álcool desidrogenases, nesse caso, estas enzimas formam parte do “sistema oxidativo” de degradação da parede celular vegetal (VASCONCELOS *et al.*, 2013). Dessa forma, estas classes de enzimas são candidatas para a realização de novos estudos aplicados a biorrefinarias de etanol 2G, visto que estas PADs catalisam a redução dos aldeídos como o furfural aos seus respectivos álcoois, que são menos tóxicos e menos reativos para leveduras fermentadoras e coquetéis enzimáticos. (GOUVEIA *et al.*, 2009; WANG *et al.*, 2014).

Um dos principais inconvenientes da produção de etanol 2G é a geração de intermediários tóxicos, resultantes do pré-tratamento da lignocelulose. No pré-tratamento da lignocelulose, como a autohidrólise, são produzidos vários subprodutos químicos que inibem microrganismos fermentativos e coquetéis enzimáticos (JÖNSSON e MARTÍN, 2016). Entre esses subprodutos estão os aldeídos, ácidos alifáticos, derivados de furano e compostos fenólicos, ácido hidroxibenzóico, furfural e hidroximetilfurfural (HMF). O furfural, por exemplo, pode danificar a parede celular e a membrana das leveduras fermentadoras, inibindo significativamente o crescimento,

assim como a interrupção da atividade enzimática por danos ao DNA e sínteses de RNA e proteínas (LIU *et al.*, 2009). Jönsson e Martín, (2016), mencionam que estudos envolvendo métodos físico-químicos e biológicos podem diminuir esses efeitos inibitórios nos fermentadores de etanol 2G. As lacases e peroxidases são as mais estudadas para estes objetivos, porém outros produtos ou enzimas PADs não foram relatados com muita frequência. Pesquisas sobre mecanismos oxidorreduativos para ajudar na sacarificação da lignocelulose vem sendo estudadas. A exemplo, sabe-se que algumas lacases, peroxidases e outras enzimas oxidorreduases, são envolvidas na hidrólise de biomassa. Estas enzimas agem através de oxidação direta ou indireta da holocelulose, e dessa forma atuam na recalcitrância de materiais lenhosos. (PECIULYTE *et al.*, 2015).

Existem relatos sobre AKR da família 1 (AKR1) caracterizadas em insetos da ordem Blattodea (cupins) e besouros, que se alimentam de madeira (*Anoplophora glabripennis*), reconhecidas como enzimas PADs endógenas, expressas para auxiliar na digestão da lignocelulose da qual se alimentam. Essas PADs são expressas no intestino desses insetos, juntamente com enzimas celulasas e hemicelulasas (carbohydrate-active enzymes: CAZymes), e atuam sinergicamente com CAZymes de microrganismos simbióticos. Foi confirmado que transcritos de genes que codificam AKR são expressos significativamente no intestino do cupim *Coptotermes gestroi*, nas castas dos 'trabalhadores', os quais têm sua rotina alimentar baseada em madeira apenas. Isto sugere que algumas AKRs de *C. gestroi* são altamente expressas em resposta à dieta, consistente em alimentos lignocelulósicos. (MASON *et al.*, 2016; CAIRO *et al.*, 2016; TRAMONTINA *et al.*, 2017; TRAMONTINA *et al.*, 2019). As AKR1 caracterizadas nesses insetos podem ser adicionadas a coquetéis enzimáticos para a melhoria da produção de etanol 2G, o qual pode aumentar consideravelmente com a ajuda destas enzimas, por meio de dois mecanismos propostos: 1) a diminuição de produtos tóxicos inibidores da fermentação, 2) fomentar as interações sinérgicas enzimáticas com glicosil-hidrolases, mediante a produção de espécies reativas de oxigênio (EROs) como H<sub>2</sub>O<sub>2</sub>, os quais promovem a desconstrução/despolimerização da lignocelulose. (TRAMONTINA *et al.*, 2017). As AKRs são expressas por vários organismos, incluindo procariotos e eucariotos. Podem-se encontrar relatos de AKRs expressas em leveduras, bactérias, humanos, entre outros. Porém, são escassos os dados moleculares em invertebrados. Existem estudos desenvolvidos com AKRs em

algumas ordens de insetos, particularmente em lepidópteras e dípteras (<https://hosting.med.upenn.edu/akr/existing/>). Entre as AKRs reportadas e caracterizadas em insetos, as que mais se sobressaem são as enzimas catalisadoras das reações do hormônio ecdisona, Ecdysone oxidase (OE) e 3-dehydroecdysone reductase (3DE-3 $\alpha$ -redutase), e as Aldose redutases (ARs).

As OE e 3DE-3 $\alpha$ -redutase pertencem à família AKR2, subfamília E (AKR2E). Estas são envolvidas na regulação do hormônio de muda e metamorfoses (ecdisteroide), um dos hormônios mais importantes nos insetos. O processo de síntese e inativação do ecdisteróide controla essa transformação dos insetos. Uma das formas para inativar a ecdisona é a conversão desta em 3-desidroecdisona, mediante a enzima ecdisona oxidase (EO) e, seguidamente, em 3-epiecdisona por meio da 3DE-3 $\alpha$ -redutase. (SUN *et al.*, 2012). As Aldose redutases (AR, EC: 1.1.1.21) pertencem à Família AKR1, que inclui: A) aldeído redutase de mamífero; B) aldose redutase de mamífero; C) desidrogenases hidroxisteróides (HSDs); e D) A4-3-cetosteróide-5 $\beta$ -redutase. Estas enzimas catalisam especificamente aldoses ou cetonas em álcool, na presença do cofator NADPH. A enzima é expressa em forma de monômero em vários tecidos de organismos, sendo característicos o peso molecular de 30-40 kDa, ponto isoelétrico levemente ácido, uma estrutura tridimensional representativa, contendo oito  $\alpha$ -hélices em torno de um núcleo de oito cadeias e orientação paralela, que forma uma bolsa elíptica hidrofóbica (WANG *et al.*; 2018). Apesar de que não foi estabelecida sua real função fisiológica, esta enzima está envolvida nas complicações diabéticas e galactosemicas em humanos, catalisando a redução da glucose para transformá-la em sorbitol, e galactose em galactitol; por isso são alvo de pesquisa os inibidores de AR, que já são utilizados em pacientes diabéticos. (BARSKI *et al.*, 2008). As ARs possuem especificidade para vários substratos, podendo ter maior preferência por do que outro. Sugere-se que uma das funções desta família de redutases possa ser a desintoxicação de xenobióticos, contendo carbonil (VANDER JAGT *et al.*, 1995). Já foram identificados AKR1 em insetos, especialmente algumas aldose-redutases (AR), mas a função específica destas enzimas nesses organismos ainda é desconhecida. Existem evidências que possam estar envolvidas no crescimento e desenvolvimento fisiológico, além da desintoxicação e inibição de defensinas de plantas (STAFFORD-BANKS *et al.*, 2014; WANG *et al.*, 2018).

### 2.3 *PICHIA PASTORIS* (*KOMAGATAELLA PHAFFII*) COMO SISTEMA DE EXPRESSÃO HETERÓLOGA

A levedura *Pichia pastoris* é um microrganismo metilotrófico, amplamente utilizado na expressão de proteínas heterólogas tanto em pesquisa laboratoriais como em uso industrial. Muitas proteínas já foram expressas com sucesso por meio deste sistema (CREGG *et al.*, 2001; ARAÚJO, 2012; OLIVEIRA, 2020). O gênero *Pichia* tem sido estudado com maior profundidade nas análises filogenéticas, e atualmente foi transferido para o gênero *Komagataella*, abrangendo três espécies: *K. pastoris*, *K. phaffii* e *K. pseudopastoris* (SCHUTTER *et al.*, 2009). Nesse trabalho será utilizado o nome científico *P. pastoris* para fazer referência à marca comercial da levedura utilizada nos experimentos de expressão heteróloga.

Este sistema de expressão possui características desejáveis para uma expressão eficiente de proteínas heterólogas, incluindo: manipulação genética simples, modificações pós-traducionais, como formação de pontes dissulfeto, O- e N-glicosilação e processamento de peptídeos-sinal; desta forma é possível superexpressar proteínas eucarióticas. (HOHENBLUM *et al.*, 2004). Além disso, é um sistema que pode ser utilizado de forma rápida, simples e econômica, comparando com outros sistemas eucarióticos superiores, como células de mamíferos e insetos (GELLISSEN, 2000).

Os resultados mais promissores até agora obtidos com expressão em *P. pastoris* são devido a sua alta taxa de proliferação (podendo atingir até 200 g/L de densidade celular em uma fermentação), baixo custo e capacidade de realizar modificações pós-traducionais com acurácia (LIU *et al.*, 2012). Esse mesmo autor menciona que há mais de 20 anos tem se utilizado as leveduras como sistemas de expressão heteróloga de proteínas para uso humano (ex.: insulina e vacina da hepatite B), e com outras aplicações (LIU *et al.*, 2012).

*Pichia pastoris* também possui certa vantagem, comparando com *Saccharomyces cerevisiae*, quando se trata de expressão de proteínas; ambos sistemas têm alta capacidade de N-glicosilação do tipo manose. A diferença é que a *S. cerevisiae* adiciona entre 50 a 150 resíduos de manose à cadeia, causando uma hiperglicosilação, enquanto que *P. pastoris* adiciona entre 8 a 14 resíduos apenas.

Além disso, *P. pastoris* não adiciona  $\alpha$ -1,3-manose terminais como a *S. cerevisiae*, sendo isso melhor para expressão de proteínas de humanos, já que  $\alpha$ -1,3-manose é desfavorável, pois transformam as proteínas em fortes determinantes antigênicos, tornando-as inadequadas para utilização terapêutica. (TSCHOPP *et al.*, 1987; GRINNA e TSCHOPP, 1989).

O promotor do gene codificador da enzima álcool oxidase 1 (AOX1) é o mais utilizado em vetores de expressão comerciais para conduzir a expressão de genes em *P. pastoris*. A expressão do AOX1 é regulada em células crescidas em metanol ~5%, mas fica totalmente bloqueado quando as células crescem na presença de glicose (TSCHOPP *et al.*, 1987; CREGG *et al.*, 2002). O promotor gliceraldeído-3-fosfato desidrogenase (pGAP) também tem sido utilizado para expressão constitutiva de muitas proteínas heterólogas. Oferece um sistema de expressão constitutiva dependente de fontes de carbono como ativador. Este promotor apresenta algumas vantagens frente ao promotor AOX1, pois não precisa da adição de metanol como fonte de carbono, o qual é muito tóxico e economicamente inviável para ser utilizado na indústria. Além disso, consegue produzir uma quantidade razoável (g/L) de peso seco de células em menos tempo, em comparação com o promotor AOX1, que precisa de mais do dobro do tempo; isso resulta em um nível ligeiramente superior de expressão observado no GAP em contrapartida do AOX1 (MACAULEY-PATRICK *et al.*, 2005).

Uma técnica para melhorar a expressão de proteínas em *P. pastoris* que está sendo muito utilizada, atualmente, é a otimização de códons do gene de interesse para o *codon usage* da levedura. Alguns exemplos de estudos realizados com esta tecnologia incluem uma endoglucanase 1 (E1) proveniente de *Acidothermus cellulolyticus*, muito utilizada nas pesquisas sobre celulases, que teve os códons correspondentes otimizados para expressão em *P. pastoris* e atingiu 550 mg/L no sobrenadante da cultura (SHUMIAO *et al.*, 2010). Mesmo que se tenham escassos exemplos de expressão de proteínas de  $\geq 10$  g/L, existem muitos relatos de expressão de  $\geq 1$  g/L em *P. pastoris*. Isso demonstra que *P. pastoris* é um dos mais eficientes sistemas de expressão eucariótica disponíveis (MACAULEY-PATRICK *et al.*, 2005), sendo por isso utilizado com maior frequência nas pesquisas, e também uma excelente alternativa para a superexpressão de genes heterólogos para fins biotecnológicos.

### 3 JUSTIFICATIVA

A linha de pesquisa que envolve a produção de combustíveis é muito ampla e abrangente. Cada vez são mais surgem tecnologias e alternativas a respeito dessa área, focadas em tecnologias mais eficientes, econômicas e amigáveis para o meio ambiente. No caso do etanol 2G, um dos focos é a identificação e comprovação de enzimas novas e eficientes para serem utilizadas no coquetel enzimático, assim como celulases de alto potencial. Pode-se dizer o mesmo para o etanol 1G, pois é relevante a utilização de enzimas que degradam diretamente a sacarose para a obtenção de glicose. Outras enzimas envolvidas indiretamente na melhoria da produção de etanol 2G são as AKRs, pouco relatadas nas pesquisas, pois são enzimas que ainda estão começando a serem exploradas nas pesquisas em etanol 2G.

Este estudo apresenta a validação de três tipos diferentes de enzimas, provenientes de insetos-praga, importantes na agricultura. Enzimas endógenas de insetos também podem ser consideradas novidades na aplicação da bioprospecção de moléculas, pois além de poder aportar possíveis alternativas para a produção de etanol, enzimas endógenas de insetos-praga também podem ser alvos de estudos para outros objetivos, como na aplicação no controle de pragas agrícolas.

## 4 OBJETIVOS

### 4.1 OBJETIVO GERAL

Validar genes de insetos-praga, codificadores de enzimas endógena, visando identificar moléculas com atividade potencial para a degradação de biomassa vegetal.

### 4.2 OBJETIVOS ESPECÍFICOS

- Identificar *contigs*, dentro do banco de dados dos transcriptomas de *A. grandis* e de *T. licus licus*, que possuam similaridade com sequências de genes de enzimas que degradam biomassa vegetal;
- Isolar os genes de interesse, a partir de cDNAs desses insetos;
- Expressar os genes de interesse em sistema heterólogo (*P. pastoris*);
- Avaliar e caracterizar as proteínas expressas, mediante ensaios de atividade enzimática.

## 5 MATERIAL E MÉTODOS

### 5.1 MATERIAL BIOLÓGICO

Diferentes estágios de *A. grandis* (ovos, larvas e adultos) foram obtidos do Laboratório de Bioecologia e Semioquímicos de Insetos da Embrapa Recursos Genéticos e Biotecnologia, Brasília-DF; já espécimes de *T. licus licus* (ovos e lagartas) foram coletados de canaviais da Usina Triunfo, em Boca da Mata-AL. Ambas as espécies levadas para o Laboratório de Interação Molecular Planta-Praga da Embrapa Recursos Genéticos e Biotecnologia, Brasília-DF, onde foram alimentados com dieta artificial. Para *A. grandis*, a dieta consistiu em: 1500 mL de água, 40 g de ágar, 60 g de levedo de cerveja, 100 g de proteína de soja, 60 g de gérmen de trigo, 40 g de pharmamedia, 60 g de açúcar, 2 g de nipagin, 2,4 g de ácido sórbico, 10 g de sais minerais, 20 g de ácido ascórbico e 10 mL de solução vitamínica. Para *T. licus*, a dieta consistiu em: 20 g de caseína, 10 g de extrato de levedura, 60 g de açúcar branco, 10 g de ácido ascórbico, 11 mL de solução vitamínica, 7,5 g de sais de Wesson, 300 mg de colesterol, 300 mg de benzoato de sódio, 80 mg de ampicilina, 800.000 UI de nistatina, 2 mL de cloreto de colina (50%) e 800 mL de água. Os insetos foram mantidos a  $26 \pm 2$  °C, umidade relativa de  $70 \pm 10\%$  e fotoperíodo de 12 horas. Para a expressão gênica das enzimas de interesse, os insetos foram alimentados com cana-de-açúcar e avaliados durante todos os estágios do ciclo de vida (ovo, lagarta, pupa e adultos de ambos os sexos).

### 5.2 EXTRAÇÃO DE RNA E SÍNTESE DE cDNAs

Foram escolhidos quatro genes com sequências completas, que codificam as enzimas endo- $\beta$ -1,4-glucanase de *A. grandis*, e AKR e duas  $\beta$ -frutofuranosidases de *T. licus licus*, com base no banco de dados: transcriptomas do intestino de ambos insetos (Firmino *et al.*, 2013; Fonseca *et al.*, 2015, NORIEGA *et al.*, 2020). Para a obtenção destes genes, foi isolado o RNA total do intestino e carcaça de *A. grandis* (larva e adulto), e do intestino, carcaça, corpo gorduroso e túbulos de Malpighi de *T. licus licus* (ovo e lagarta). Esses tecidos escolhidos, previamente dissecados, foram

macerados em N<sub>2</sub> líquido, utilizando almofariz e pistilo. Para homogeneizar as amostras foi adicionado o reagente Trizol (Invitrogen Life Technologies™ Califórnia, EUA), segundo as instruções do fabricante. As amostras foram tratadas com 2U de DNase I RNase-free (Ambion, Invitrogen Life Sciences™ Califórnia, EUA) por 30 minutos a 37 °C, de acordo com as instruções do fabricante. O material foi purificado e extraído, utilizando o kit RNeasy (QIAGEN, Hilden, Alemanha). Posteriormente, o produto foi avaliado em gel de agarose a 1,5%. A concentração de RNA foi estimada em espectrofotômetro NanoVue (GE HealthCare, Chicago, Illinois, EUA). Os cDNAs foram sintetizados, utilizando 2 µg de RNA total, utilizando o kit Superscript II™ First-Strand Synthesis SuperMix (Invitrogen) e o oligonucleotídeo NV-d(T)30-AP.

### 5.3 CLONAGEM E SEQUENCIAMENTO DOS GENES

Para a amplificação dos genes em questão, a partir do cDNA total do insetos, foram desenhados os iniciadores (*primers*) e a partir das sequências dos genes encontrados nos transcriptomas, utilizando as ferramentas Primer 3 (<http://frodo.wi.mit.edu/>) e Tm calculator (<https://www.thermofisher.com/us/en/.../tm-calculator.html>). As sequências dos *primers* para cada gene foram denominadas “Forward” (Fwd) e “Reverso” (Rev) (TABELA 1). Foram adicionados 0,3 µL da Taq Polimerase (Ludwig™), 10 µmol de cada primer e 2 µL de cDNA diluído 20 vezes, para cada reação de PCR. O programa do termociclador consistiu em uma desnaturação inicial de 5 min a 95 °C, seguido de 30 ciclos de desnaturação (30s a 95°C), anelamento (30 s a 58 °C), extensão (1 min a 72°C), e extensão adicional (5 min a 72°C). Os produtos de PCR foram purificados do gel de agarose 1% utilizando o QIAquick Gel Extraction Kit, (Qiagen, Hilden, Alemanha). Os produtos purificados foram clonados no plasmídeo pGEM-T Easy (Promega™, Madison, EUA), e as construções foram usadas para transformar *E. coli* OmniMAX™ 2 T1 R. Os transformantes (pGEM-T easy:AgEG1 e pGEM-T easy:TIXR1) foram distribuídos em placas com meio LB (10 g/l triptona, 10 g/l NaCl, 5 g/l de extrato de levedura, 15 g/l de ágar, com pH 7.5) suplementadas com ampicilina (100µg/ml), IPTG 100 mM e X-Gal 50 µg/ml, utilizando o método de seleção “sistema branco-azul”. Os clones positivos foram comprovados por digestão com a enzimas *EcoRI* e por sequenciamento das extremidades 5' e 3', em sequenciador automático (ABI 3130xl Genetic Analyzer).

## 5.4 ALINHAMENTO DAS SEQUÊNCIAS E ANÁLISES DE DADOS

As relações filogenéticas das sequências foram realizadas utilizando estas e outras sequências conservadas, recuperadas do GenBank – (<http://www.ncbi.nlm.nih.gov>), utilizando as ferramentas de bioinformática “BLASTn” e “BLASTx”. As análises foram feitas com as sequências proteicas dos genes para limitar a busca no grupo de eucariontes da classe Insecta, visando a encontrar similaridades filogenéticas com outras enzimas endógenas. As sequências proteicas preditas foram obtidas utilizando o ExPASy Translate (<http://web.expasy.org/translate/>). O ponto isoelétrico (pI) e o peso molecular (MW) foram preditos, mediante a ferramenta pI/MW tool ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)). Os softwares DictyOGlyc 1.1 (<http://www.cbs.dtu.dk/services/DictyOGlyc>) e NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) foram utilizados para determinar os sítios de O- e N-glicosilação, respectivamente. O SignalP 4.1 Server (<http://www.cbs.dtu.dk/services/SignalP/>) contribuiu para determinar possíveis sequências de peptídeo-sinal.

O alinhamento de todas as sequências foi realizado com a ferramenta ClustalW para o alinhamento múltiplo de sequências, usando parâmetros padrão, ajustados manualmente.

## 5.5 CONSTRUÇÃO DA ÁRVORE FILOGENÉTICA DE *TlicAKR-1* E ALINHAMENTO DE SEQUÊNCIAS

Foi realizada uma análise evolutiva com o gene *TlicAKR-1* especificamente, pois pertence à superfamília das AKR, que inclui uma ampla variedade de genes. Estas análises, representadas em árvores filogenéticas, foram realizadas utilizando o software MEGAX64, utilizando o método MUSCLE para alinhamento das sequências e Maximun Likelihood para a construção da árvore. Este processo de filogenia foi feito, primeiramente com todas as sequências existentes da superfamília AKR, para deduzir

em qual família se encontra o gene de interesse. Posteriormente, foi realizada a comparação com as sequências da família a que pertence o gene *TlicAKR-1* para poder agrupar este a uma subfamília. As sequências dos membros das AKRs foram localizadas em <https://hosting.med.upenn.edu/akr/existing/>. Também foi realizado um estudo com base no alinhamento de sequências de aminoácidos para comparar o gene *TlicAKR-1* com outras AKRs, já caracterizadas na literatura, utilizando o software CLUSTALW para obter um alinhamento múltiplo.

## 5.6 CONSTRUÇÃO DOS PLASMÍDEOS DE EXPRESSÃO

As regiões codificadoras dos genes sem peptídeo-sinal, foram amplificadas por PCR usando a Taq Polimerase (Ludwig). Foram utilizados os nucleotídeos iniciadores, desenhados com os sítios para as enzimas de restrição: *EcoRI* e *Sall* para *AgEG-1*; *EcoRI* e *NotI* para *TlicAKR-1*; *EcoRI* e *Sall* para *TlicSuc1*; e *NotI* e *Sall* para *TlicSuc2*. Tais enzimas foram escolhidas, utilizando a ferramenta NEBcutter V2.0 (<http://nc2.neb.com/NEBcutter2/>), para confirmar que as sequências não possuem os mesmos sítios de restrição utilizados para clonagem e, portanto, não seriam clivadas indesejadamente. Os parâmetros para os ciclos da PCR foram os mesmos mencionados anteriormente. Os produtos de PCR foram purificados (QIAquick Gel Extraction Kit, Qiagen) e clonados no vetor pGEM-T Easy (Promega™, Madison, EUA). As construções obtidas foram transformadas em *E. coli* OmniMAX™. Os plasmídeos foram extraídos de *E. coli* e digeridos com as enzimas de restrição apropriadas para excisão dos insertos, e confirmar a clonagem. O vetor de expressão pGAPZαB (Invitrogen, Califórnia, EUA), foi digerido com as respectivas enzimas de restrição para cada construção. Com ajuda da T4 ligase obteve-se a ligação dos insertos purificados e o vetor pGAPZαB, gerando assim os cassetes de expressão pGAPZαB:*AgEG-1*, pGAPZαB:*AKR1-1*, pGAPZαB:*TlicSuc1* e pGAPZαB:*TlicSuc2*. Os produtos da ligação foram transformados em *E. coli* OmniMAX™, e os transformantes foram selecionados em placas de meio LB contendo 25µg/ml de Zeocina™ (Invitrogen, Califórnia, EUA). Procedeu-se à extração dos plasmídeos de *E. coli*, e a confirmação dos genes por sequenciamento de DNA.

## 5.7 TRANSFORMAÇÃO GENÉTICA DE *P. PASTORIS* E ANÁLISE DA INTEGRAÇÃO DOS GENES

Para transformação estável de *P. pastoris* foi utilizado o vetor pGAPZ $\alpha$ B para permitir uma expressão constitutiva, utilizando o promotor (PGAP) do gene gliceraldeído-3-fosfato desidrogenase (GAPDH). A 6xHis Tag foi adicionada aos genes para facilitar a subsequente purificação das proteínas de interesse. Foram escolhidas duas cepas de *P. pastoris* (SMD1168 H e X33) (Invitrogen, Califórnia, EUA) para os testes de expressão. A primeira é uma levedura sem atividade de proteinases, que foi selecionada para diminuir o risco da degradação proteolítica das proteínas de interesse. A segunda é uma cepa selvagem com alta eficiência de crescimento e expressão. Foram utilizados 10  $\mu$ g de DNA correspondente, que foram linearizados com as enzimas de restrição *Bgl*II para pGAPZ $\alpha$ B:*AgEG-1*, *Bsp*HI para pGAPZ $\alpha$ B:*AKR1-1*, e *Dra*I para pGAPZ $\alpha$ B:*TlicSuc1* e pGAPZ $\alpha$ B:*TlicSuc2*, para facilitar a integração do DNA no cromossomo de *P. pastoris*. As células foram transformadas por eletroporação, utilizando o protocolo da MicroPulser™ Xcell™ Electroporation System (Bio-Rad™, Hercules, USA). As células competentes foram crescidas nas condições de 28°C a 220 rpm, no meio YPD (1% de extrato de levedura, 2% de peptona e 2% de dextrose), até atingir uma densidade de 5-7 x 10<sup>7</sup> células/ml. Após centrifugação, o precipitado de células foi tratado com Tampão HEPES (ácido N-(2-hidroxietil) piperazina-N'-2-etanossulfônico) (pH 8.0) e DTT (1,4-Dithiothreitol), 1 M, antes das lavagens com sorbitol (1 M). 40  $\mu$ l de células competentes foram adicionadas para cada DNA e foram transferidas para cubetas de 0,2 cm para a eletroporação, utilizando o parâmetro "Pic" (2 kV). As células foram incubadas por duas horas a 30°C antes de serem distribuídas em placas de YPD/sorbitol contendo ágar e 100  $\mu$ g/ml de zeocina (Invitrogen, Califórnia, EUA), e foram crescidas durante quatro dias a 28°C.

A integração dos genes no cromossomo da levedura foi confirmada por PCR de colônias, utilizando os oligonucleotídeos "forward" e "reverse" de cada gene. Foram selecionadas as colônias desenvolvidas após uma semana e foram submetidas à lise celular por repetidos choques-térmicos, consistindo em água em ebulição e nitrogênio líquido. Após centrifugação, foram extraídos 2  $\mu$ l do sobrenadante contendo DNA para ser utilizado nas PCRs.

## 5.8 EXPRESSÃO EM PEQUENA ESCALA DAS PROTEÍNAS EM *P. PASTORIS* E WESTERN BLOT

Foram escolhidos dez transformantes positivos de cada gene, baseando-se na presença de bandas amplificadas na PCR. As colônias selecionadas foram submetidas aos ensaios de expressão em meio YPG, com 5 ml de meio em tubos Falcon de 50 ml. Todos os meios de cultura preparados com YPG líquido e sólido foram preparados contendo o antibiótico cefotaxima (100 µg/ml), tanto para o meio utilizado no preparo das células competentes, como para os ensaios de expressão e placas contendo ágar. A expressão das proteínas consistiu em deixar crescer as células de levedura até atingir a fase log (4 dias), extraíndo alíquotas de 1 ml a cada 24 horas, para analisar a expressão relativa ao tempo. Posteriormente, foram extraídas as proteínas intra e extracelulares, como descrito a seguir.

Para a extração das proteínas intracelulares, foram utilizadas microesferas de vidro para lisar as células, seguindo o protocolo do manual da Invitrogen. Para a extração das proteínas do meio extracelular, o sobrenadante foi separado das células por centrifugação, e posteriormente o conteúdo livre de células foi precipitado com 10% ácido tricloroacético (TCA) em acetona, durante 2 horas a -20°C. Uma vez precipitadas as proteínas, estas foram centrifugadas e lavadas três vezes com acetona gelada. Depois de secar e evaporar totalmente a acetona, o conteúdo de proteínas foi solubilizado em água MilliQ e a concentração de proteínas foi determinada pelo método de Bradford (1976). As amostras foram ressuspensas em 30 µL de tampão de carregamento amostra contendo SDS.

Um micrograma do material ressuspendido foi aplicado em SDS-PAGE 12% para eletroforese com as seguintes condições para a corrida: 10 miliampères (mA) no primeiro gel (5%) e 20 mA no segundo gel (12%). A coloração foi efetuada submergindo os géis em uma solução com 0,25% de "Coomassie Brilliant Blue G 250", 45% de metanol e 8% de ácido acético e mantendo-o sob agitação por aproximadamente uma hora. Posteriormente, os géis foram descorados em uma solução com 40% de metanol e 10% de ácido acético glacial, sendo realizadas trocas sucessivas até ser obtida a transparência desejada.

Foi realizado o *Western blotting*, utilizando o anticorpo específico Anti-His (Invitrogen, Califórnia, EUA) (Towbin et al., 1979). O gel e a membrana primeiramente foram incubados por 10 min em Tampão de Transferência, contendo Tris base 25 mM, glicina 192mM e metanol 20%, e mantidos sob agitação lenta. A membrana utilizada foi de nitrocelulose Hybond-C (G.E. Healthcare). Os procedimentos de eletrotransferência foram realizados em sistema Trans-blot SD Semi-Dry Transfer Cell (Bio-Rad), a 10 mA e durante 20 minutos. Posteriormente, a membrana foi incubada com a solução bloqueio, contendo gelatina a 1% e PVA a 0,5%, durante 16 horas. Posteriormente, a membrana foi incubada com o Anticorpo monoclonal anti-His (Thermo Fisher Scientific, Waltham, Massachusetts, USA) conjugado a Fosfatase Alcalina, na concentração 1:3.000, por 4 horas. A reação de revelação foi realizada com o Kit Colorimétrico Immun-Blot® AP (Bio-Rad) (Brasileiro e Carneiro, 1998).

## 5.9 PURIFICAÇÃO DAS PROTEÍNAS

Para a purificação das proteínas foram testados vários métodos cromatográficos para determinar a melhor metodologia. Inicialmente, foi realizada a cromatografia de afinidade por íons metálicos imobilizados (IMAC) para facilitar a correta purificação da proteína, marcada com His tag na coluna de ácido nitrilotriacético (NTA) revestida com Níquel. Foi escolhida uma colônia de *P. pastoris* X-33 para cada gene transformado com pGAPZ $\alpha$ -B. As colônias foram inoculadas em 100 mL de YPG nas condições de 28 °C e 225 rpm durante 3 dias. Desse material pré-inoculado, 10 mL foi adicionado a 1 L de YPG, onde as células cresceram a 28 °C e 225 rpm por 3 dias. Após isso, o sobrenadante foi separado da massa celular por centrifugação, filtrado em filtro de 0,2  $\mu$ m e diluído no tampão fosfato de sódio pH 7,4 (1:1), até uma concentração final de 20 mM. Foi adicionado cloreto de sódio a uma concentração final de 0,5 M. Essa amostra foi passada na coluna HisTrap FF de 5 mL (G.E. Healthcare, Chicago, Illinois, EUA), previamente equilibrada com tampão de ligação (fosfato de sódio 20 mM, pH 7,4, NaCl 0,5 M). A coluna foi lavada com o tampão de lavagem, contendo imidazol 20 mM. Posteriormente, a proteína recombinante foi eluída com tampão de ligação, contendo imidazol 0,5 M. As frações lavadas e eluídas foram dialisadas com ajuda de uma membrana de celulose,

regenerada de 10K MWCO, colocadas em 5L de água MilliQ e deixada a 4°C por 16h. Em seguida, as amostras foram liofilizadas e guardadas para ensaios posteriores.

Para a cromatografia de troca iônica, as células foram crescidas da mesma forma que o protocolo já mencionado. O sobrenadante foi filtrado e diluído no tampão Acetato de Sódio (1:1) para uma concentração final de 40 mM, pH 4,0, sendo posteriormente aplicado à coluna de agarose HiTrap SP de 1mL (G.E. Healthcare, Chicago, Illinois, EUA), previamente equilibrada com o tampão Acetato de Sódio 40 mM pH 4.0. O material retido foi eluído mediante um gradiente de 0 – 100 % do tampão Acetato de Sódio 40 mM e Cloreto de Sódio 1M. As amostras eluídas de cada pico e o extrato lavado foram dialisados, liofilizados e estocados para futuros experimentos.

Para o caso especial da AKR expressa pelo gene *TlicAKR-1*, foi testado um protocolo de cromatografia de afinidade, utilizando a coluna HisTrap Blue Sepharose de 1 mL (G.E. Healthcare, Chicago, Illinois, EUA), seguindo o protocolo de Wermuth et al. (1982). Todas as etapas desta purificação foram realizadas a 4 °C na presença de ditioneitol (DTT) a 1 mM. O extrato bruto (sobrenadante) foi submetido a um primeiro processo de purificação na coluna de troca iônica HiTrap SP. Uma vez coletado o pico contendo a proteína eluída, a mesma foi diluída em tampão fosfato de sódio 10 mM, pH 6,2. Esta amostra foi aplicada a uma coluna de afinidade HisTrap Blue Sepharose, já equilibrada com tampão de fosfato de sódio 10 mM, pH 6,2. A coluna foi lavada com o mesmo tampão, seguido de um volume de fosfato de sódio 100 mM, pH 6,2 e, novamente, um volume do original, do primeiro tampão. A proteína foi eluída utilizando um gradiente de 0-150 pM de NADPH em 10 mM de tampão fosfato de sódio. As frações obtidas foram dialisadas, secadas e guardadas para realizar outras análises.

#### 5.10 ENSAIOS EM PLACA DE DIFUSÃO RADIAL PARA DETERMINAÇÃO DE ATIVIDADE CELULÁSICA PELO AGRAGH45-1

Para o gene *AgraGH45-1* foram realizados dois tipos de ensaios em placa, o primeiro antes da expressão das proteínas em meio líquido, para fazer um *screening* das colônias transformantes e ativas. As colônias selecionadas foram crescidas em placas, contendo meio sólido YPG/CMC 0,2 %, por quatro dias a 28° C. A placa foi imersa em uma solução de Vermelho Congo 0,1 %, deixando repousar a solução por 20 minutos sob leve agitação, posteriormente foi lavada com NaCl 1M por 15 minutos sob leve agitação, seguindo protocolo modificado de Kasana et al., (2008).

No segundo ensaio foram utilizadas alíquotas da expressão de três dias no meio líquido. As placas foram preparadas com YPG sólido, carboximetilcelulose (CMC) 1 % e vermelho congo a 1% pH 5,5 (ajustado com Acetado de Sódio 50 mM). Foram realizadas pequenas perfurações na placa para adicionar as amostras: 50 µL da expressão controle (vetor vazio) e 50 µL de sobrenadante das expressões das proteínas recombinantes. O material foi incubado a 45 °C, por 16 horas. Para a revelação do resultado, foi adicionando ácido acético a 10% sob leve agitação durante 1 hora, a temperatura ambiente. A hidrólise do CMC foi analisada pela formação de halos brancos ao redor dos orifícios. (JIMENEZ *et al.*, 2014).

#### 5.11 DETECÇÃO DE AÇÚCARES REDUTORES LIBERADOS PELO AGRAGH45-1

A atividade foi determinada por ensaio colorimétrico para detecção de açúcares redutores, gerados pela hidrólise do substrato CMC. Para a detecção de açúcares redutores em solução foi utilizado o reagente Acido Dinitrossalicílico (ADNS) (MILLER, 1959). Os ensaios de atividade foram realizados nas seguintes condições: 50 µL de CMC 2% (p/v), em tampão acetato de sódio 200 mM (pH 5,5) e 50 µL de amostra contendo a enzima (sobrenadante da cultura). A mistura foi incubada a 50°C por 30-120 minutos e em seguida adicionaram-se 400 µL de reagente DNS. As amostras foram fervidas por 10 minutos, e 100 µL foram diluídos em 800 µL de água e então a absorbância foi lida a 540 nm em espectrofotômetro. As leituras de absorbância foram convertidas em quantidade equivalente de glicose, por meio da curva de calibração

obtida pela leitura da absorbância de diferentes concentrações de glicose (10 a 100 µg) com o reagente DNS. Todos os ensaios foram realizados em triplicata, utilizando como branco a solução contendo 50 µL do tampão de reação e 50 µL do substrato (CMC 2%); tratados como descrito anteriormente. Como controle da concentração inicial de açúcares redutores no meio de cultivo, foi realizado também reação com 50 µL da amostra de enzima e 50 µL de tampão de reação sem o substrato CMC e submetidos ao mesmo procedimento. Uma unidade de atividade foi definida como a quantidade de enzima necessária para gerar 1 µmol de glicose por minuto de reação.

#### 5.12 AVALIAÇÃO DO PH NA ATIVIDADE DO AGRAGH45-1 E COMPARAÇÃO DA ATIVIDADE E ESTABILIDADE TÉRMICA CONTRA A CELULASE COMERCIAL DE *A. NIGER*

A atividade da AgraGH45-1 testado em diferentes faixas de pH foi realizado com o substrato hidroxietilcelulose (HEC) conjugado com o cromóforo vermelho Ostazin Brilliant-Red (OBR-HEC). Cada ensaio foi preparado com 25 µL do sobrenadante de expressão de três dias com diferentes tampões (50 mM de Citrato de sódio pH 3,0; 50 mM de acetato de sódio pH 5,0; 50 mM de fosfato de sódio pH 7,0; e 50 mM de Tris / HCl pH 9,0) incubado por 15 min a 45° C. A reação foi iniciada pela adição de 50 µL de OBR-HEC 5 mg / mL em um volume final de 150 µL, continuando a 45 ° C por 30 minutos a mais. A reação foi parada pela adição de 900 mL de etanol e depois centrifugada a 10.000 × g. O sobrenadante foi extraído para leitura da absorbância a 550 nm. Como controle para cada reação foram utilizados 25 µL do sobrenadante de uma expressão inoculada com *P. pastoris* transformada com vetor pGAPZα-B vazio. Cada reação foi realizada em triplicata.

Foram também realizados ensaios para comparar a atividade e estabilidade térmica de *AgraGH45-1* e uma celulase comercial de *A. niger* (Sigma Aldrich, EUA). Utilizando OBR-HEC em pH 5,0 e temperatura variando de 40 a 60 °C. Portanto, 10 µg de ambas enzimas foram dissolvidas em 25 µL de citrato de sódio 50 mM pH 5,0 e incubadas por 15 min a 45 ° C. A reação foi iniciada pela adição de 50 µL de OBR-HEC 5 mg / mL em um volume final de 150 µL. A reação foi iniciada pela adição de 50 µL de OBR-HEC 5 mg / mL em um volume final de 150 µL, continuando a 45 ° C por 30 minutos a mais. A reação foi parada pela adição de 900 mL de etanol e depois

centrifugada a  $10.000 \times g$ . O sobrenadante foi extraído para leitura da absorbância a 550 nm. Os mesmos procedimentos foram repetidos para incubação a 50 e 60 °C. Cada reação foi realizada em triplicata. (JIMENEZ *et al.*, 2014).

### 5.13 ATIVIDADE ENZIMÁTICA DAS SACARASES

Para este ensaio, também foi utilizado o método com DNS para determinação de açúcares redutores, como descrito anteriormente (Miller, 1959), usando glicose como padrão para construção de curvas de calibração. Todos os ensaios foram realizados em triplicata e tiveram controles negativos com vetor vazio. A mistura consistiu em 100 mM de substrato (sacarose, rafinose e maltose), 10 mM de tampão fosfato de sódio (pH 7,0), e 20  $\mu$ L de sobrenadante da enzima expressa em três dias, para um volume final de 50  $\mu$ L, incubando por 20 min a 30° C. Para bloquear a reação da enzima, 100  $\mu$ L de DNS foram adicionados, seguidos por 10 min de ebulição e 5 min de resfriamento no gelo. Os açúcares redutores foram aferidos a 540 nm.

### 5.14 ATIVIDADE ENZIMÁTICA DA TlicAKR-1

A atividade da AKR foi testada com a diminuição da absorção de NADPH por intervalo de tempo, a temperatura ambiente. O ensaio foi conduzido utilizando 250 mM de tampão fosfato de sódio (pH 6,5); 0,5 mM de cofator (NADPH e NADH); 100  $\mu$ L de sobrenadante contendo a enzima; e 100 mM de substrato, (glicose, xilose, frutose, arabinose, manose, manitol, glicerol, 2-nitrobenzaldehyde ou metilglioxal) para um volume final de 500  $\mu$ L. A enzima foi incubada com NADPH por 2,5 minutos antes da adição do substrato para iniciar a reação. Foram avaliadas alíquotas a cada 5 minutos de reação ocorrida desde o tempo zero, durante 45 minutos. Uma unidade de atividade enzimática foi definida como a quantidade de enzima que catalisa a oxidação de 1  $\mu$ mol de NADPH / min a 25 ° C. (BRUINENBERG *et al.*, 1982).

## 6 RESULTADOS E DISCUSSÃO

### 6.1 ANÁLISES DAS SEQUÊNCIAS E ALINHAMENTO DE DADOS

Nos transcriptomas de ambos os insetos-praga, *A. grandis* e *T. licus licus*, obtidos pela equipe do Laboratório de Interação Molecular Pranta – Praga (LIMPP) da EMBRAPA Recursos Genéticos e Biotecnologia (FIRMINO *et al.*, 2013; FONSECA *et al.* 2015; NORIEGA *et al.*, 2020), foram encontrados vários “*contigs*” (sequências montadas), que apresentaram sequências parciais ou totais de genes endógenos, relacionados à degradação da biomassa vegetal. As sequências de enzimas envolvidas na degradação de biomassa foram identificadas no NCBI e comparadas com as sequências encontradas nos nossos transcriptomas. Foram identificados vários *contigs* e escolhidos aqueles com maior escore e e-value (dados não mostrados). Foram confirmadas as sequências completas de quatro genes: um semelhante a uma  $\beta$ -1-4-endoglucanase, que foi denominada *AgraGH45-1*; outra semelhante a uma Aldo-Ceto redutase, denominada *TlicAKR-1*; e duas similares a  $\beta$ -frutofuranosidasas (sacarase/invertase), que foram chamadas de *TlicSuc1* e *TlicSuc2*. Os resultados do sequenciamento dos genes, após a clonagem, são similares aos *contigs*, com exceção de algumas bases, que foram substituídas por outras em algumas regiões das cadeias, especialmente no gene *TlicAKR-1*, que apresentou mudanças mais evidentes; estas variações ao longo das cadeias de DNA, se deve, provavelmente, a erros no sequenciamento ou a polimorfismos.

#### 6.1.1 *AgraGH45-1*

O resultado do alinhamento de sequências no BlastX demonstra que este gene possui alta semelhança (>90%) com outras sequências de genes de diferentes coleópteros, encontradas nos bancos de dados. Alguns desses genes já estão caracterizados como  $\beta$ -1,4-endoglucanases endógenas de insetos, pertencentes à família GH45. Além disso, este gene contém o sitio de N-glicosilação conservado, próprio de celulasas da família GH45 de coleópteros. Estes resíduos de N-glicosilação são essenciais para a atividade enzimática, especialmente para as  $\beta$ -1,4-

endoglucanases (WEI *et al.*, 2005 e 2006). A proteína predita possui um peso molecular estimado em 21,4 kDa, sem contar com a cauda de histidina, que foi adicionada artificialmente após a clonagem no vetor de expressão. O ponto isoelétrico (pI) predito foi de 5,16. Foi previsto um peptídeo-sinal na sequência proteica, o qual é um dos indicativos que o gene codifica uma proteína secretada no meio extracelular. Este fragmento foi excluído durante a clonagem do gene no vetor de expressão, pois o vetor já possui o próprio peptídeo-sinal. Na região 3'-UTR foi encontrado um local de poliadenilação, seguido de uma cauda poli-A, que são características de ser mRNA eucariótico, ou seja, endógeno do inseto. A sequência de *AgraGH45-1* compartilha um alto nível de identidade com outras sequências da família glicosil-hidrolases 45 de coleópteros (FIGURA 4), escolhidas para a análise: uma  $\beta$ -1,4-endoglucanase de *Oncideres albomarginata* (Oa-EGase II) [GenBank: GU001942] e uma  $\beta$ -1,4-endoglucanase de *Rhynchophorus ferrugineus* (Rf-GH45) [GenBank: ANU06045.1]. A proteína predita apresentou 74% de identidade com *Rf-GH45* e 56% com *Oa-EGase II*.

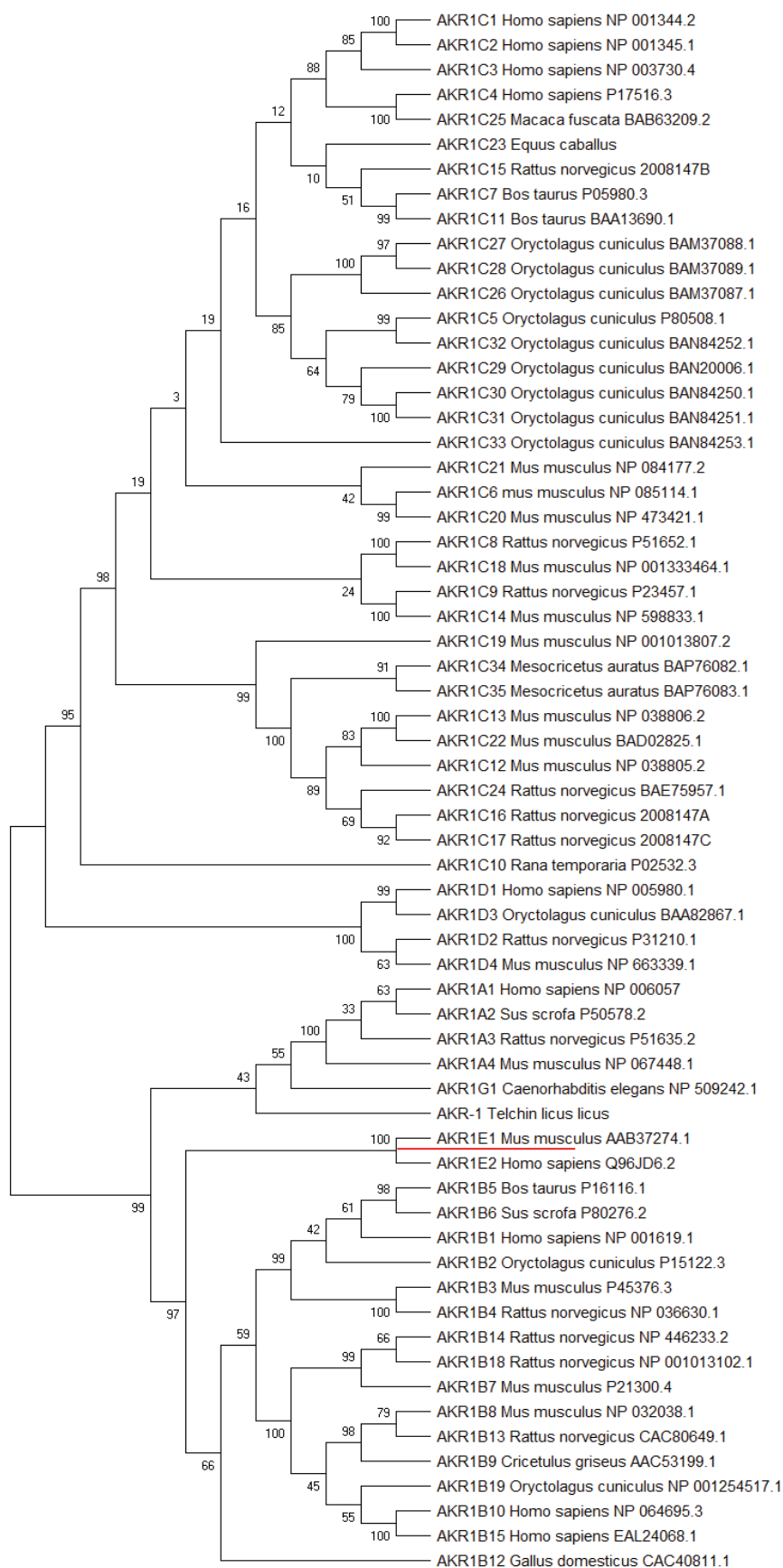
As endoglucanases são agrupadas em 13 famílias de glicosil hidrolases (GH), agrupadas segundo as semelhanças de estruturas primárias e terciárias, especificidades de substratos, bem como mecanismos catalíticos (CANTAREL *et al.*, 2009). Tais propriedades são importantes quando uma enzima recombinante é uma potencial candidata para fins industriais. O processo de conversão de biomassa vegetal utiliza padrões de temperatura e pH, em que as estruturas primárias e terciárias das enzimas precisam se submeter, mantendo a estabilidade e atividade enzimática aperfeiçoada. A conservação da estrutura terciária em uma  $\beta$ -1,4-endoglucanase recombinante também é fundamental para a eficiência enzimática. Este gene também conserva o sítio catalítico próprio das GH45, "TRYWDCKPSC" (LEE *et al.*, 2004 e 2005). Gao e colegas relataram que é necessário apenas um movimento de *loop* para acontecer a catálise em uma  $\beta$ -1,4-endoglucanase do fungo *Thielavia terrestris*, ordem Sordariales, família Chaetomiaceae, a qual compartilha o mesmo mecanismo catalítico que o *AgraGH45-1* (GAO *et al.*, 2017). Esta reação consiste em um resíduo de ácido aspártico (D), próximo à extremidade N-terminal, que serve como ácido catalítico. Este ácido aspártico catalítico é conservado entre as sequências dos *AgraGH45-1*, *Rf-GH45* e *Oa-EGase II* (FIGURA 4). Nos domínios ativos das três  $\beta$ -1,4-endoglucanases destes insetos, as sequências mostram



A análise no BlastX demonstrou que o gene teve na sua maioria, mais de 95% de identidade com outras enzimas endógenas, provenientes de insetos. Dentro desse grupo, todas as sequências pertencem à superfamília das Aldo-ceto redutases (AKR), onde se podem destacar, especialmente, sequências de Aldose redutases (AR) de lepidópteros. A primeira árvore filogenética foi construída para determinar a qual família pertence o gene *TlicAKR-1*. Para isso, a sequência de interesse foi analisada com todas as sequências das diferentes famílias das AKRs, resultando de uma maior aproximação com a Família 1. Para confirmar melhor o resultado, foi feito o alinhamento das sequências com o algoritmo MUSCLE e se encontrou que a proteína predita possui <40% de identidade de aminoácidos com qualquer outra família com exceção da AKR-1 (dados-não mostrados). A família AKR1 inclui as seguintes subfamílias: A) aldeído redutase de mamífero; B) aldose redutase de mamífero; C) desidrogenases hidroxisteróides (HSDs); e D) A 4-3-cetosteróide-5 $\beta$ -redutase. (<https://hosting.med.upenn.edu/akr/existing/>). Uma segunda árvore foi feita para classificar *TlicAKR-1* em uma subfamília, para isso foram utilizadas apenas as sequências das AKR1. O resultado confirmou maior semelhança com a subfamília A (Aldeído redutase) e B (Aldose redutase) (FIGURA 5), porém, a análises do alinhamento entre as sequências apresentou que o gene é ainda mais semelhante com membros da AKR1B, chegando inclusive até 56% com AR de humano, zebu e coelho. Mesmo que o alinhamento das sequências apresentou maior porcentagem de identidade com esta subfamília, o gene não atingiu >60% de identidade esperado (<https://hosting.med.upenn.edu/akr/nomenclature/>). Desta forma não é possível classificar *TlicAKR-1* nesta subfamília. A enzima apresentou também similitudes com todas as outras subfamílias das AKR1, isto deve-se à estrutura ( $\alpha\beta$ )<sub>8</sub> TIM-barrel das AKRs, que permite prender uma gama diversa de substratos carbonílicos. (ELLIS, 2002). Mas a sequência mantém uma porcentagem de identidade maior com as AKR1B, por causa dos loops do TEM-barrel, que dão uma característica especial a cada sequência de AKR, já que a conservação da sequência dos membros é mais baixa nesses *loops* (MINDNICH e PENNING, 2009).

Este estudo de filogenia foi reforçado com outra análises de alinhamento múltiplo entre sequências de outras AKRs encontradas na literatura (FIGURA 6): uma Aldose redutase de *B. mori* (BmAR) [GenBank: XP\_004925118]; uma 3-dehydroecdysone reductase (3DE-3 $\alpha$ -redutase) de *B. mori* (AKR2E4) [GenBank:

AB701383]; Uma Aldo-Ceto redutase Família 1 de *C. gestroi* (CGAKR1) [GenBank: KU686221]; e uma Aldose redutase de *Homo sapiens* (AKR1B1) [GenBank: NP\_001619.1]. Pode-se observar maior similitude entre as sequências das duas Aldose redutase e o *TlicAKR-1*: 56% com AR de humano e 83% para AR de *B. mori*; mas também não teve diferenças significativas contra as outras sequências mais distantes: 41% para AKR1 de cupim e 43% para 3DE-3 $\alpha$ -redutase de *B. mori*. O alinhamento mostra que a enzima de interesse possui quase todos os resíduos que interagem com substratos aldoses (amarelo), assim como também conserva os sítios para o cofator NADPH. A sequência de *BmAR* é a mais idêntica de *TlicAKR-1*, e estas duas possuem também uma identidade alta com AR de humano. Existem vários resíduos diferentes na região C-terminal entre as sequências de AKR das lagartas e AR de humano, provavelmente indicando as diferenças estruturais das proteínas. Mas pode-se observar que os resíduos que se combinam ao NADPH para participar da catálise e a maior parte dos resíduos para os substratos aldoses (marcados com triângulos pretos e iluminados em amarelo) são bastante consistentes, sugerindo que a função das AKR das lagartas e humanos sejam basicamente as mesmas. Por outro lado, *TlicAKR-1* também possui alguns aminoácidos que participam na ligação a substratos esteroides (azul). Como as AKRs têm a capacidade de catalisar uma ampla gama de substratos diferentes, é provável que *TlicAKR-1* possa interagir com aldoses preferencialmente, mas também com aldeídos e inclusive esteroides.



**Figura 5.** Relações evolutivas dos genes da subfamília AKR1 das Aldo-Ceto redutases, utilizando o método Maximum Likelihood. Cada ramo da árvore possui um número que identifica a distância filogenética existente entre cada sequência

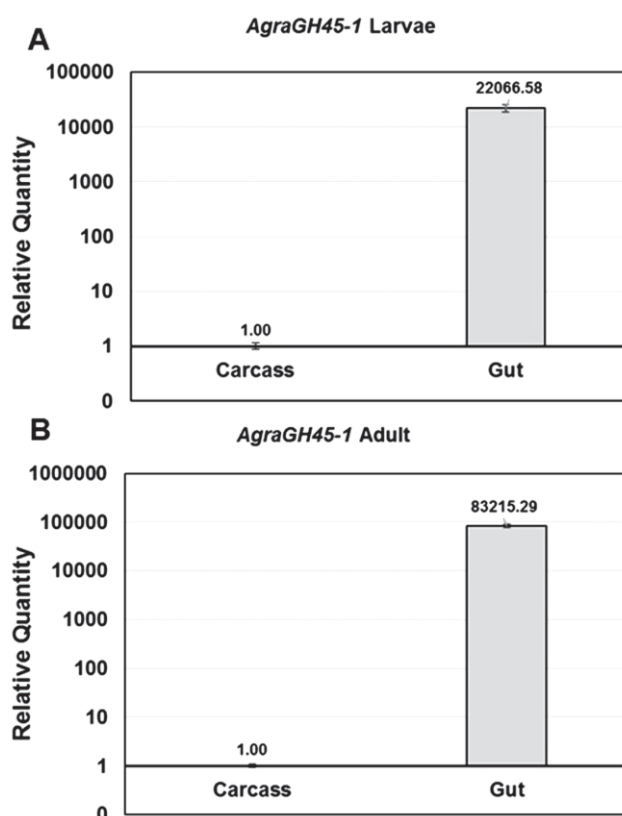


sequências proteicas preditas, a sequência predita da enzima TlicSuc1 corresponde a 58,2 kDa e pI de 5,75; e para a enzima TlicSuc2 o tamanho predito foi de 59.88 kDa, com pI de 4,63. Foi predito um peptídeo-sinal nas sequências proteicas, indicando que se trata de proteínas extracelulares. Esses peptídeos-sinal foram deletados para a clonagem dos genes no vetor de expressão. Igualmente dos genes *AgraGH45-1* e *TlicAKR-1*, as sequências das sacarases apresentaram locais de poliadenilação, e uma cauda poli-A, demonstrando ser mRNA eucarióticos.

Atualmente existem poucos relatos sobre  $\beta$ -FFase em animais, e em insetos foi encontrado em apenas lepidópteros e coleópteros (SANTOS e TERRA, 1986; SUMIDA *et al.*, 1994; CARNEIRO *et al.*, 2004; PAUCHET *et al.*, 2008; DAIMON *et al.*, 2008; PAUCHET *et al.*, 2010; KEELING *et al.*, 2013; PEDIZZI *et al.*, 2014). Os resultados no BlastX mostraram a alta identidade das sequências das sacarases com outros ortólogos de Lepidoptera, porém estas sequências ainda não foram bem caracterizadas. Portanto, foram escolhidas sequências de  $\beta$ -FFases já validadas para realizar o alinhamento múltiplo com as  $\beta$ -FFases da *T. licus licus*: *B. mori* (BmSuc1) [GenBank: NP\_001119721.1]; e *Bacillus aryabhatai* [GenBank: WP\_043979130.1]. O resultado do alinhamento mostrou que entre *TlicSuc1* e *TlicSuc2* existe 45% de identidade, e ambas apresentaram perto dessa proporção ~40 – 45 % também contra as outras sequências; sendo que apenas *TlicSuc2* resultou mais próxima de *BmSuc1* (65%). Os resíduos dos sítios catalíticos (tríade de aminoácidos) permanecem conservados em todas as sequências das sacarases (FIGURA 7), o que sugere que estas enzimas poderiam ter a mesma função com substratos utilizados por  $\beta$ -FFases. Foi confirmado que TlicSuc1 e TlicSuc2 são enzimas produzidas pelo próprio organismo da *T. licus licus*. Existe a possibilidade que a lagarta tenha integrado esses genes por transferência horizontal de genes (HGT) (LI *et al.*, 2011; ZHU *et al.*, 2011). Muitos insetos, especialmente alguns lepidópteros, conseguem integrar material genético de outros organismos no próprio genoma. Até o momento não se tem certeza do motivo desse fenômeno, porém genes transferidos horizontalmente por bactérias desenvolveram novas funções nesses insetos, que ajudaram a se adaptar a condições ambientais e alimentares (LI *et al.*, 2011).



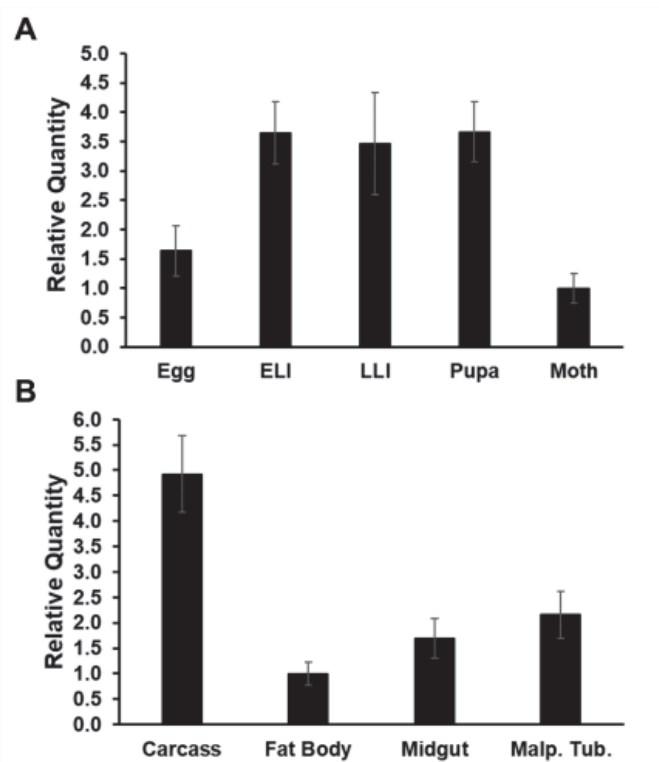
carcaça (FIGURA 8). O intestino médio é a região mais importante para a digestão de carboidratos no sistema digestivo de insetos (TERRA e FERREIRA, 1994). Estes resultados levam a uma hipótese de que este gene possa codificar uma enzima funcional. De fato, coleópteros das famílias Curculionidae e Chrysomelidae têm um grande número de glicosil-hidrolases em seus genomas, incluindo endoglucanases das famílias GH5, GH9, GH45 e GH48. No entanto, a validação funcional *in vivo* sempre será necessária, pois a enzima pode desenvolver outra função no organismo do inseto, que possa não se equiparar com a literatura e outros ensaios, além da aquisição de novas funções, que permitam ter especificidade para um substrato diferente. (BUSCH *et al.*, 2019).



**Figura 8.** Quantificação relativa do gene *AgraGH45-1* na carcaça e no intestino de *A. grandis*. Quantificação em larva (A) e inseto adulto (B).

### 6.2.2 TlicAKR-1

Para a realização da qPCR foram utilizados *TlicGAPDH* e *TlicRPS18*, como genes de referência (Fonseca et al., 2015). O resultado confirmou que a lagarta *T. licus licus* expressa este gene em todos os estádios ovo, larva, pupa e adulto (borboleta), mas em maior proporção nos estágios de larva e pupa (FIGURA 9). O estágio larval é a fase em que o inseto praga causa danos nas plantações da cana de açúcar, alimentando-se do interior dos colmos. O gene *TlicAKR-1* é expresso em todos os tecidos selecionados para o ensaio (carcaça, corpo gorduroso, intestino médio e túbulos de Malpighi), sendo bem notória a expressão na carcaça do inseto (FIGURA 9). Özsü e Monteiro, 2017 determinaram a presença de uma Aldose redutasa, com função não identificada, no tecido epidérmico de uma pupa de *Bicyclus anynana* (lepidóptera). Eles utilizaram apenas a região epidérmica, que forma os halos das asas do adulto. (ÖZSU e MONTEIRO, 2017). Um AR de *B. mori*, que tem 83% de identidade com *TlicAKR-1*, é também expresso em maior quantidade nos estádios de larva e em todos os tecidos. No entanto, *BmAR* exibiu a mais alta expressão nos testículos, ovário, traqueia e intestino médio, e um baixo nível de expressão na epiderme (carcaça) (WANG et al., 2018). Até o momento não foi encontrado outro relato de gene *AKR* expresso em maior quantidade na carcaça, além do *TlicAKR-1*.

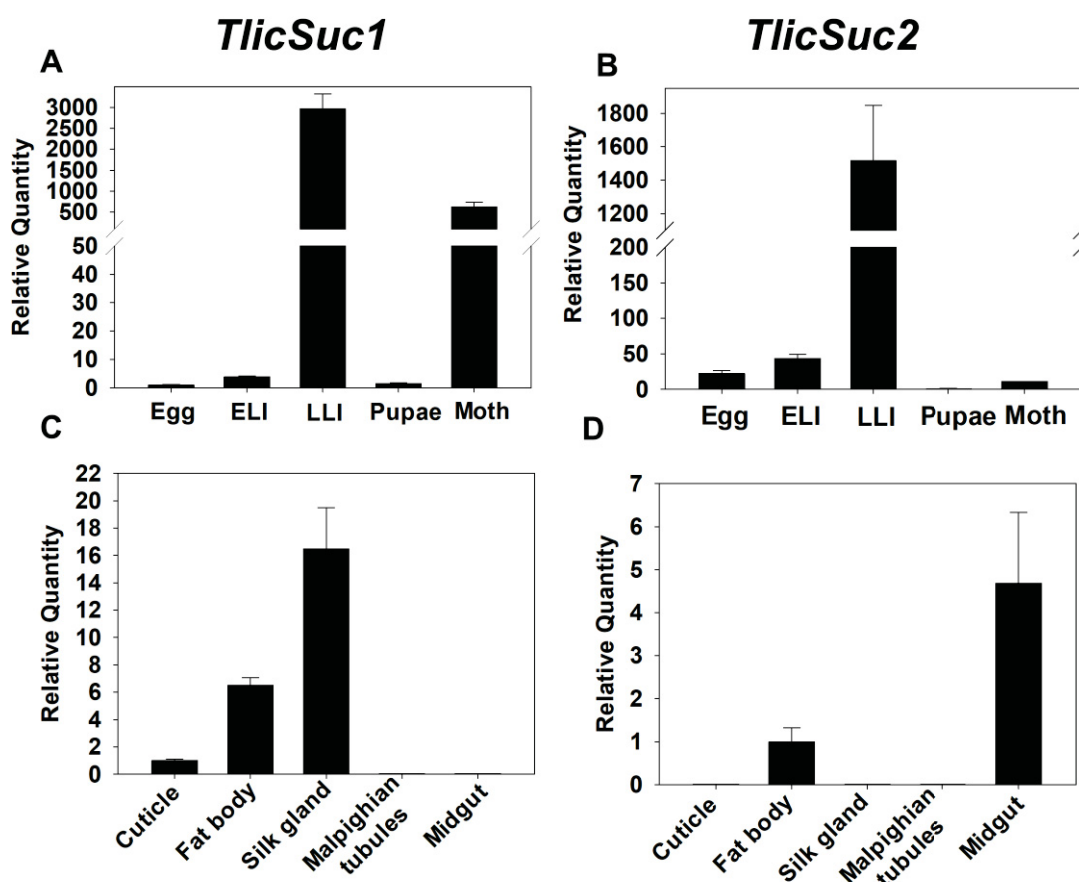


**Figura 9. Análise de expressão da enzima TlicAKR-1 de *T. licus licus* por RT-qPCR.** A) Expressão do gene durante o desenvolvimento de insetos, desde o ovo até o inseto adulto (mariposa). ELI - Instares Larvares Precoces; LLI - Instares larvares tardios, Pupa e Adulto. B) Expressão do gene em diferentes tecidos larvais: Carcaça, Corpo gorduroso, Intestino médio e Túbulos de Malpighi.

### 6.2.3 TlicSuc1 e TlicSuc2

Tanto o gene *TlicSuc1* como *TlicSuc2* apresentaram níveis mais altos de expressão no estágio final de lagartas, depois com quase nula nas pupas e só *TlicSuc1* demonstrou um aumento significativo no estágio borboleta (FIGURA 10A e 10B). Coincidentemente, outros relatos de  $\beta$ -frutofuranosidases de *B. mori* (*BmSuc1* e *BmSuc2*) também apresentaram expressão mais alta no estágio final de lagarta, ou seja, no quinto instar (DAIMON *et al.*, 2008; GAN *et al.*, 2018), assim como em fase tardia de larvas de besouro (PEDEZI *et al.*, 2014). Estes resultados sugerem que as de  $\beta$ -frutofuranosidases possuem um rol importante no desenvolvimento das larvas destes insetos. Curiosamente, a expressão de *TlicSuc1* foi maior nas glândulas de seda da lagarta, a diferença de *TlicSuc2* que foi mais expressa no intestino médio (FIGURA 10C e 10D). No bicho de seda, esta enzima está envolvida no processo de formação das proteínas da seda como serina e fibroína (GAN *et al.*, 2018), mas a real

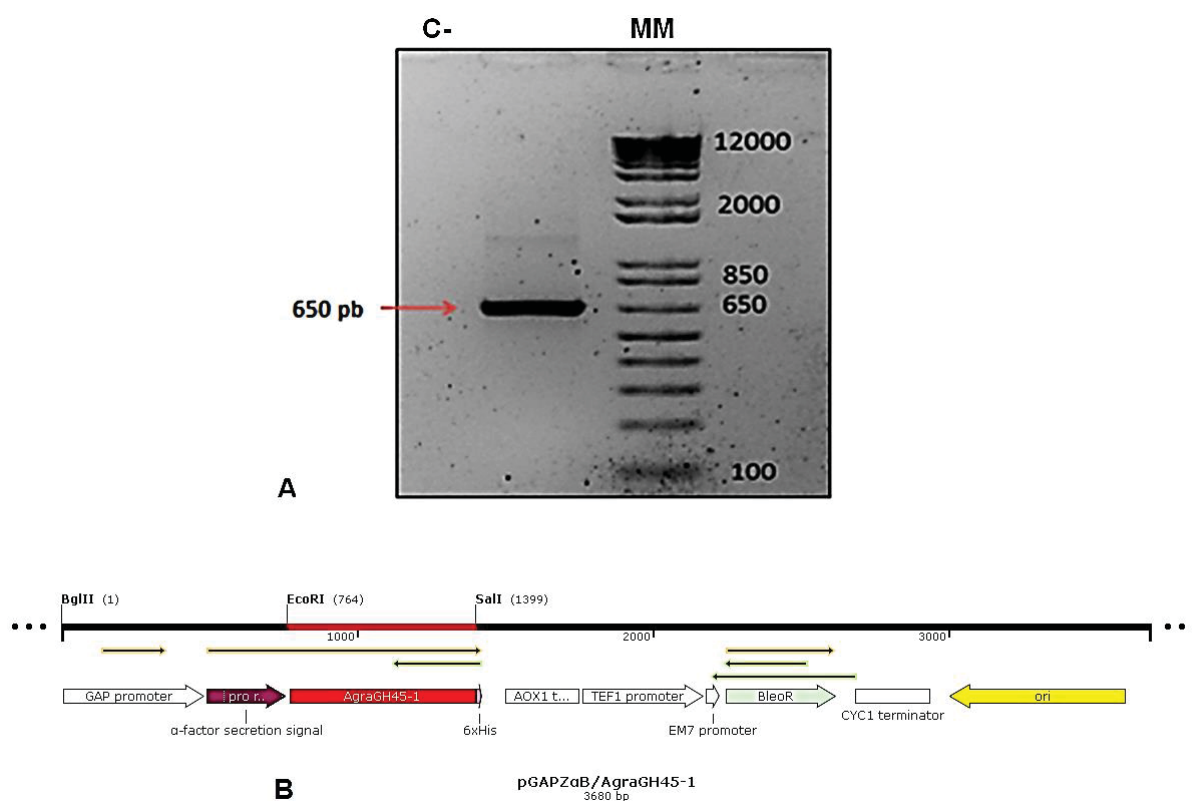
função dela nas glândulas de seda ainda é inédita. A alta expressão de  $\beta$ -fructofuranosidase no intestino de *B. mori* provavelmente deva-se a um mecanismo de adaptação alimentar, já que este inseto consome principalmente amoreira. Esta planta contém alcaloides de estrutura parecida com açúcares e também atuam como inibidores específicos de  $\alpha$ -glucosidase, dessa forma, a lagarta produz  $\beta$ -fructofuranosidase que substitui as  $\alpha$ -glucosidases (DAIMON *et al.*, 2008). O besouro *Sphenophorus levis* também é uma praga importante da cana de açúcar, assim como a *T. licus licus*, que se alimentam dos colmos desta cultura. Pedezzi e colaboradores caracterizaram uma  $\beta$ -fructofuranosidase neste besouro, com alta expressão no intestino. Este estudo relata a possibilidade de que esta  $\beta$ -fructofuranosidase também seja produzida como uma forma de adaptação alimentar, dado que o melaço da cana contém flavonoides que inibem a atividade da enzima  $\alpha$ -glucosidase e  $\alpha$ -amilase (PEDEZZI *et al.*, 2014; HOLT *et al.*, 2003). Sabe-se que a cana-de-açúcar, com melhoria agrônômica, contém pelo menos 18,1% de sacarose em base ao total da biomassa utilizada nas indústrias (PAPINI-TERZI *et al.*, 2009), o que poderia propiciar a presença deste tipo de sacarase no genoma da *T. licus licus* e do *S. levis*.



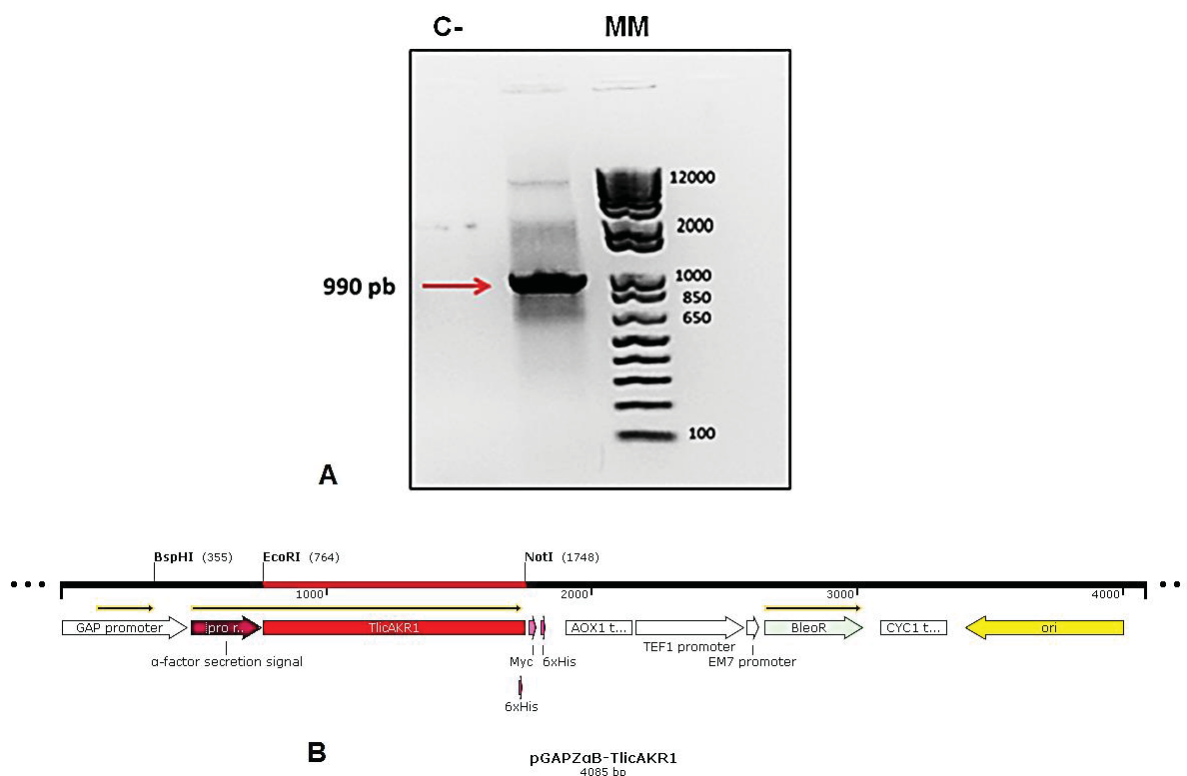
**Figura 10. Análise da expressão dos genes *TlicSuc1* e *TlicSuc2* durante os diferentes de estágios de vida e tecidos de *T. licus licus* por RT-qPCR. Expressão dos genes *TlicSuc1* (A) e *TlicSuc2* (B) durante os principais estágios de vida, ovo, lagartas em instares iniciais LII), lagartas em instares finais (LIF), pupa e adultos; Expressão dos genes *TlicSuc1* (C) e *TlicSuc2* (D) em diferentes tecidos das lagartas em instares finais do inseto, Cutícula, corpos gordurosos, glândula de seda, túbulos de Malpighi e intestino médio. Os genes *TlicGAPDH* e *TlicRPS18* foram utilizados como genes referência.**

### 6.3 CLONAGEM DOS GENES

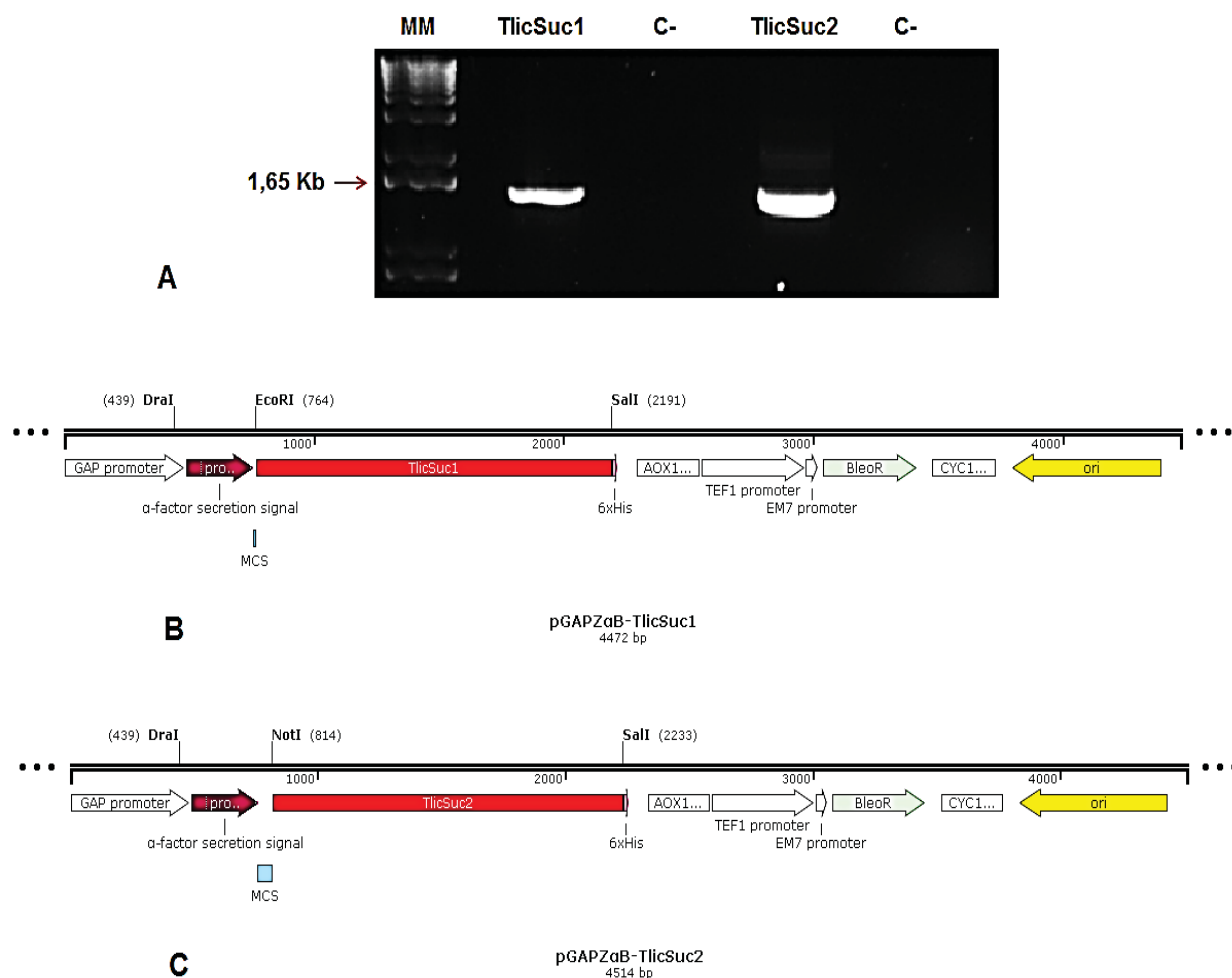
As PCRs realizadas confirmaram as bandas dos genes de 650 pb para *AgGH45-1* (FIGURA 11A), 990 pb para *TlicAKR-1* (FIGURA 12A), 1650 pb tanto para *TlicSuc1* e *TlicSuc2* (FIGURA 13A), amplificados a partir de cDNAs de intestino da fase larval de ambos insetos, confirmando que todos os genes são transcritos durante a fase larval do inseto. Os resultados do sequenciamento também validaram as sequências completas destes genes, após clonagem no vetor *pGEM-T Easy* e subclonagem no vetor de expressão *pGAPZ $\alpha$ -B* (FIGURAS 11B, 12B, 13B1 e 13B2).



**Figura 11.** A) Eletroforese em gel de agarose 1% do produto da amplificação por PCR do gene *AgraGH45-1* (650 pb) obtido a partir do cDNA do intestino da larva do *A. grandis*. B) Representação esquemática da inserção do gene da  $\beta$ -1,4-endoglucanase de *A. grandis* (*AgraGH45-1*) no vetor de expressão *pGAPZ $\alpha$ -B*. O gene *AgraGH45-1* foi subclonado entre os sítios de restrição, *EcoRI* e *Sall*, inserido entre o sinal de secreção do fator  $\alpha$  e sob controle do promotor constitutivo de GAP. No mapa da sequência também aparece o sítio da enzima *BglII*, que foi utilizado para linearizar o gene para inserção no genoma de *P. pastoris*.



**Figura 12.** A) Eletroforese em gel de agarose 1% do produto da amplificação por PCR do gene *TlicAKR-1* (990 pb), obtido a partir do cDNA do intestino das lagartas de *T. licus licus*. B) Representação esquemática da inserção do gene *TlicAKR-1* no vetor de expressão *pGAPZ $\alpha$ -B*. O gene foi subclonado entre os locais de restrição *EcoRI* e *NotI* inserido entre o sinal de secreção do fator  $\alpha$  e sob controle do promotor constitutivo de GAP. No mapa da sequência também aparece o sítio da enzima *BspHI*, que foi utilizado para linearizar o gene para inserção no genoma da *P. pastoris*.



**Figura 13.** A) Eletroforese em gel de agarose 1% do produto da amplificação por PCR dos dois genes *TlicSuc1* e *TlicSuc2* com o mesmo tamanho de bandas (1650 pb), a partir dos cDNAs de lagartas de *T. licus*. B) Representação esquemática da inserção dos genes *TlicSuc1* (B1) e *TlicSuc2* (B2) no vector de expressão *pGAPZα-B*. Os genes foram subclonados entre os sítios de restrição, *EcoRI* e *NotI*, inseridos entre o sinal de secreção do fator  $\alpha$  e sob controle do promotor constitutivo de GAP. No mapa da sequência também aparece o sítio da enzima *DraI*, que foi utilizado para linearizar o gene para inserção no genoma da *P. pastoris*.

#### 6.4 TRANSFORMAÇÃO EM *P. PASTORIS*, IDENTIFICAÇÃO DAS PROTEÍNAS EXPRESSAS E ATIVIDADE ENZIMÁTICA

O sistema *P. pastoris* foi o escolhido para expressão de proteínas heterólogas solúveis e eucarióticas funcionais, já que esta levedura é capaz de realizar pós-tradução modificações, como a N-glicosilação, necessárias para o dobramento

funcional das proteínas, além de outras características desejáveis deste sistema (HOHENBLUM *et al.*, 2004; GELLISSEN, 2000). A cepa X-33 deu um melhor resultado levando em consideração a quantidade de transformantes/placa. Provavelmente, se deva ao potencial da cepa e ao protocolo de transformação utilizado, o qual consistiu em eletroporação e tratamento prévio com DTT (1,4-Dithiothreitol). Trabalhos reportados evidenciam maior eficiência na transformação em *P. pastoris* utilizando DTT como pré-tratamento das células antes da eletroporação. Geralmente, células de *Pichia* são transformadas através da eletroporação ou pelo método de esferoplastos, ambos os métodos podem gerar resultados similares (103–104 transformantes/ $\mu\text{g}$  de DNA). Uns dos problemas da transformação são as dificuldades técnicas dos métodos e também está na escolha de um método, sem algum tipo de pré-tratamento, pois a eficiência na transformação pode ser muito baixa. Isto pode ser devido ao cromossomo da levedura ter apenas um sítio para a integração do gene exógeno, o que dificulta sua inserção (WU e LETCHWORTH, 2004; SUGA e HATAKEYAMA, 2001; GANEVA *et al.*, 1995).

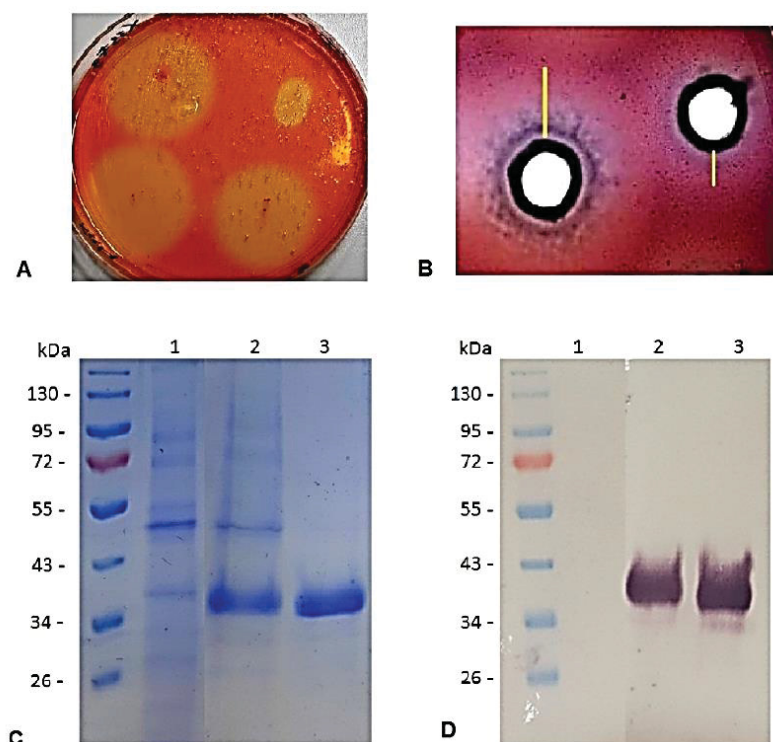
Apesar da cepa SMD1168H ter a vantagem de ser livre de proteases, esta teve um rendimento menor do que X-33. Na preparação de células competentes já se podia observar que a cepa X33 possuía um crescimento celular muito mais rápido e atingia uma densidade (células/ml) ideal, em comparação com a SMD1168H. Cereghino e Cregg (2000) mencionam que, geralmente, as cepas mutadas não possuem o potencial e a resistência das cepas selvagens, a exceção das que contém mutações AOX, que em alguns casos podem ser melhores do que as selvagens. As proteinases deficientes têm menor viabilidade, apresentam taxas de crescimento inferior e são mais difíceis de transformar (CEREGHINO e CREGG, 2000).

#### 6.4.1 AgraGH45-1

A seleção de colônias transformantes ajuda a facilitar os testes de expressão posteriores, para detecção de atividade endoglucanase. Esta enzima tem a vantagem de poder ser detectada facilmente em placas de YPG/CMC, podendo-se testar várias colônias transformantes. Como pode ser observado na figura 14A, alguns clones analisados apresentaram halos de hidrólise de tamanhos similares, confirmando assim atividade celulolítica nos transformantes analisados. O controle negativo (vetor

sem o gene) apresentou uma pequena formação de halo produzida pelas endoglucanases nativas da própria levedura (FIGURA 14A). As placas coradas para o teste de difusão radial também demonstraram a presença de atividade celulolítica através da formação de halos brancos (FIGURA 14B), produzidos a partir da endoglucanase expressa em meio líquido, após três dias. Existem relatos atuais de várias enzimas  $\beta$ -1,4-endoglucanase foram expressas com sucesso em *P. pastoris* (GAO *et al*, 2017; YAN *et al*, 2016; LIU *et al*, 2016).

A proteína heteróloga AgraGH45-1 foi purificada eficientemente por cromatografia de afinidade e detectado por *western blotting*, com anticorpo Anti-6xHis (FIGURA 14C e 14D). A enzima recombinante foi estável por cinco dias sob condições de expressão (28 ° C e 225 rpm). No entanto, mesmo que seu peso molecular previsto seja de 23 kDa, por meio da detecção por *western blotting*, o peso ficou cerca de ~35 kDa. Isso pode ser explicado pela mutação no local de clivagem Kex2 (YANG *et al*, 2013), que poderia impedir a excisão do peptídeo sinal após a secreção de proteínas. A retenção do peptídeo sinal do fator  $\alpha$  pôde ter aumentado o peso molecular da proteína recombinante em 10 kDa, além da cauda His que já aumenta 2,5KDa a mais. A hiperglicosilação também é comumente conhecida por aumentar o peso molecular das proteínas recombinantes expressas em *P. Pastoris* (CHAHED *et al*, 2018). A enzima AgraGH45-1 possui a presença de sítios de glicosilação, e essa modificação pós-traducional é crucial para sua atividade (WEI *et al.*, 2005 e 2006).



**Figura 14. Ensaio para detecção da proteína AgGH45-1 expressa em *P. pastoris*.** A) *Screening* de colônias transformantes que expressam endoglucanase vs uma colônia transformada com o vetor sem o gene, feito em placa contendo CMC 0,2%. B) Ensaio de difusão radial em placa de CMC 1% utilizando sobrenadante de expressão de três dias tanto do AgraGH45-1 como o controle negativo (vetor vazio), pode-se observar a diferença de tamanho dos halos (linhas amarelas). C) SDS-PAGE a 12% do processo de expressão e purificação por cromatografia de troca iônica da proteína AgraGH45-1: (1) sobrenadante da expressão do vetor vazio (3 dias). (2) sobrenadante da expressão de AgraGH45-1 (3 dias). (3) 1 µg de AgraGH45-1 purificado por cromatografia de afinidade. D) *western blotting* com Anti-His, a partir de amostras em C. Pode se observar o teste positivo para a proteína de interesse e o controle negativo sem aparecer.

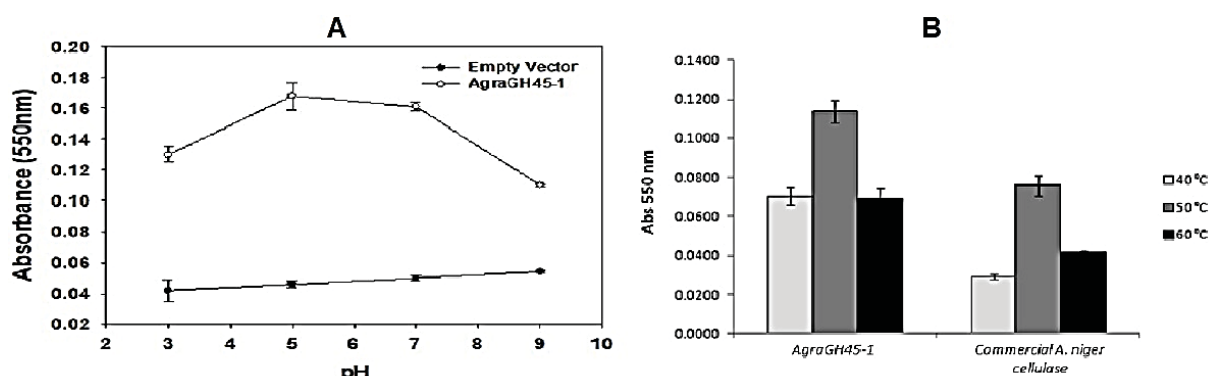
O alto custo da produção de celulases é uma das principais inconvenientes para a indústria e o mercado de bioetanol, a partir da biomassa de lignoceluloses (SINGHANIA *et al.*, 2010). Nas indústrias, o pH e temperatura são parâmetros-chave para melhorar a etapa de sacarificação enzimática nos processos de bioconversão. Uma enzima exposta a esses parâmetros, que consiga manter uma estabilidade também contribui para uma produção de bioetanol econômica e eficiente. A maioria das celulases comerciais, utilizadas para produção de etanol, tem uma atividade máxima em pH, variando de 4 a 5 e em temperaturas próximas a 50 ° C (NIEVES *et al.*, 1998). Um estudo sobre duas  $\beta$ -1,4-endoglucanases recombinantes, expressas em *P. pastoris*, foi desenvolvido para testar o pH, temperatura e estabilidade ideais (AnCel5A de *A. niger* e TtCel45A do fungo termofílico, *Theilavia terrestris* (GAO *et al.*, 2017). Ambas enzimas apresentaram atividade ótima em pH e temperatura,

convenientes de acordo com os requisitos industriais, como também foi alcançado pelo AgraGH45-1. A enzima de interesse mantém a atividade sempre alta em uma faixa de pH ampla. A atividade enzimática de AgraGH45-1 sobre o substrato hidroxietilcelulose (HEC) foi maior em pH 5,0, quando analisado em uma faixa de pH de 3 a 9 (FIGURA 15A). O gráfico mostra que a eficiência do AgraGH45-1 sobre o HEC permanece estável em pH, variando de 5,0 a 7,0, diminuindo apenas 5,5%. Por outro lado, a faixa de pH para maior atividade das enzimas mencionadas anteriormente (AnCel5A e TtCel45A) foi de 4,0 a 5,0. A enzima AgraGH45-1 parece manter sua atividade em uma faixa de pH mais ampla, o que é coincidente com a faixa necessária para o bioprocessamento de etanol. Outros relatos também indicaram atividade da  $\beta$ -1,4-endoglucanase de insetos em pH, variando de 4,5 a 5,5, como *Diabrotica virgifera* e *Apriona germari* (JIMÉNEZ *et al*, 2014; LEE *et al*, 2004), e até mesmo para o molusco *Ampullaria crossean* (GUO *et al*, 2008).

A eficiência de AgraGH45-1 também foi avaliada quanto à melhor temperatura, variando de 40 a 60 °C, e comparada à celulase comercial de *A. niger* (FIGURA 15B). O fungo *A. niger* é a fonte mais importante de celulases comerciais e tem sido utilizado em diversos processos industriais de sacarificação da celulose. Seu genoma contém cerca de 170 genes envolvidos na degradação de biomassa vegetal (PEL *et al*, 2007). A atividade máxima do AgraGH45-1 foi encontrada a 50 °C e foi 1,3 vezes maior do que a da celulase comercial de *A. niger*. Também foi observado que a 40 e 60 °C a atividade da enzima de interesse caiu 38% e 39%, respectivamente, em comparação com a atividade máxima; por outro lado, a endoglucanase de *A. niger* caiu 61% e 45%, respectivamente, em comparação com sua atividade máxima a 50 °C. Portanto, esses resultados sugerem que o AgraGH45-1 é mais eficiente que a celulase comercial de *A. niger*, e sua atividade pode resistir levemente a variação de temperatura. Essas características de tolerar faixas de pH e temperatura poderia ser aperfeiçoado para sua aplicabilidade a um bioprocessamento eficiente para conversão de biomassa.

A atividade de AgraGH45-1 sobre CMC e HEC demonstra a capacidade dessa enzima recombinante em degradar especificamente esses substratos análogos à celulose. Entretanto, sabe-se que a eficiência de algumas enzimas sobre CMC e HEC às vezes não são reproduzidas sobre substratos complexos de biomassa (SU *et al*, 2013). Portanto, seria importante estudar ainda mais a capacidade da enzima

AgraGH45-1 de degradar substratos complexos derivados de biomassa, e quando adicionado a um coquetel de enzimas celulolíticas.

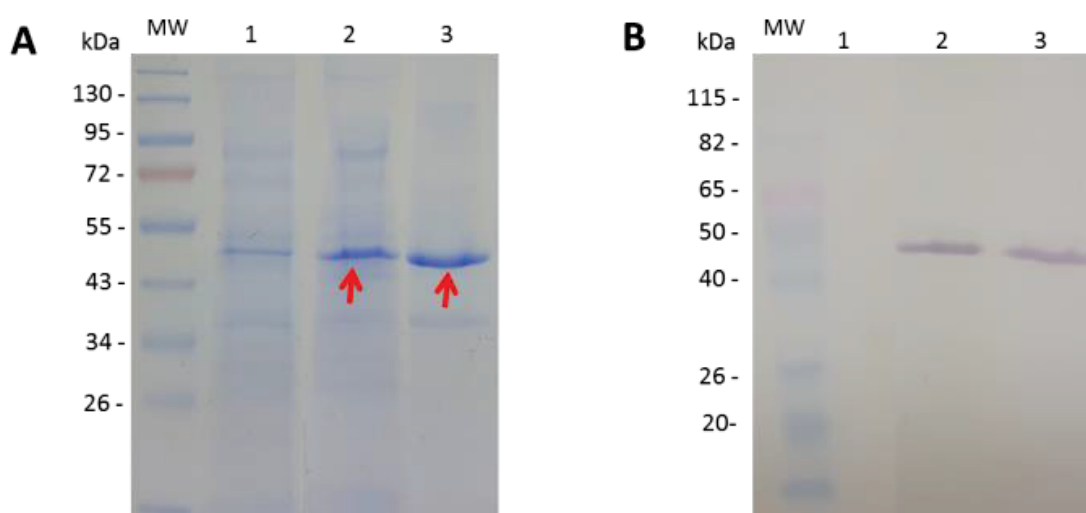


**Figura 15. Caracterização da eficiência da endoglucanase de *A. grandis*.** A) Avaliação da atividade enzimática do AgraGH45-1 em pH variado de 3,0 a 9,0. A reação enzimática foi iniciada pela adição de 50 ml de OBR-HEC (5 mg / ml) como substrato. Cada ponto na curva é a média de três medições. O aumento no Abs 550 nm é proporcional à degradação do substrato. B) Eficiência de AgraGH45-1 em comparação com a celulase comercial de *A. Niger*. A enzima AgraGH45-1 foi 1,3 vezes mais eficiente que a celulase comercial de *A. niger* para hidrolisar OBR-HEC 5 mg / mL a 50 ° C. Quando analisado a temperaturas de 40 e 60 ° C, a queda da atividade de AgraGH45-1 foi menor que a celulase comercial de *A. niger*.

#### 6.4.2 TlicAKR-1

Após a clonagem e sequenciamento foi encontrado que a proteína TlicAKR-1 não possui peptídeo sinal, confirmando ser uma proteína intracelular no organismo da *T. licus licus*. Mesmo assim, esta foi expressa com sucesso no sistema *P. pastoris* e secretada mediante o fator  $\alpha$  do vetor *pGAPZ $\alpha$ B*. A proteína predita de 33 kDa foi expressa e confirmada por *western blot* com um peso molecular de ~45 KDa (FIGURA 16). A proteína teve um peso molecular de ~10KDa a mais, da mesma forma que AgraGH45-1.

A proteína TlicAKR-1 foi purificada parcialmente por cromatografia, apenas pelo método de troca iônica (FIGURA 16). Apesar da presença de um His-Tag, não foi possível purificar a proteína na coluna de níquel, nem mesmo utilizando condições desnaturantes, com a adição de ureia 6M nas soluções. A proteína também não foi purificada na coluna de cibacron azul.

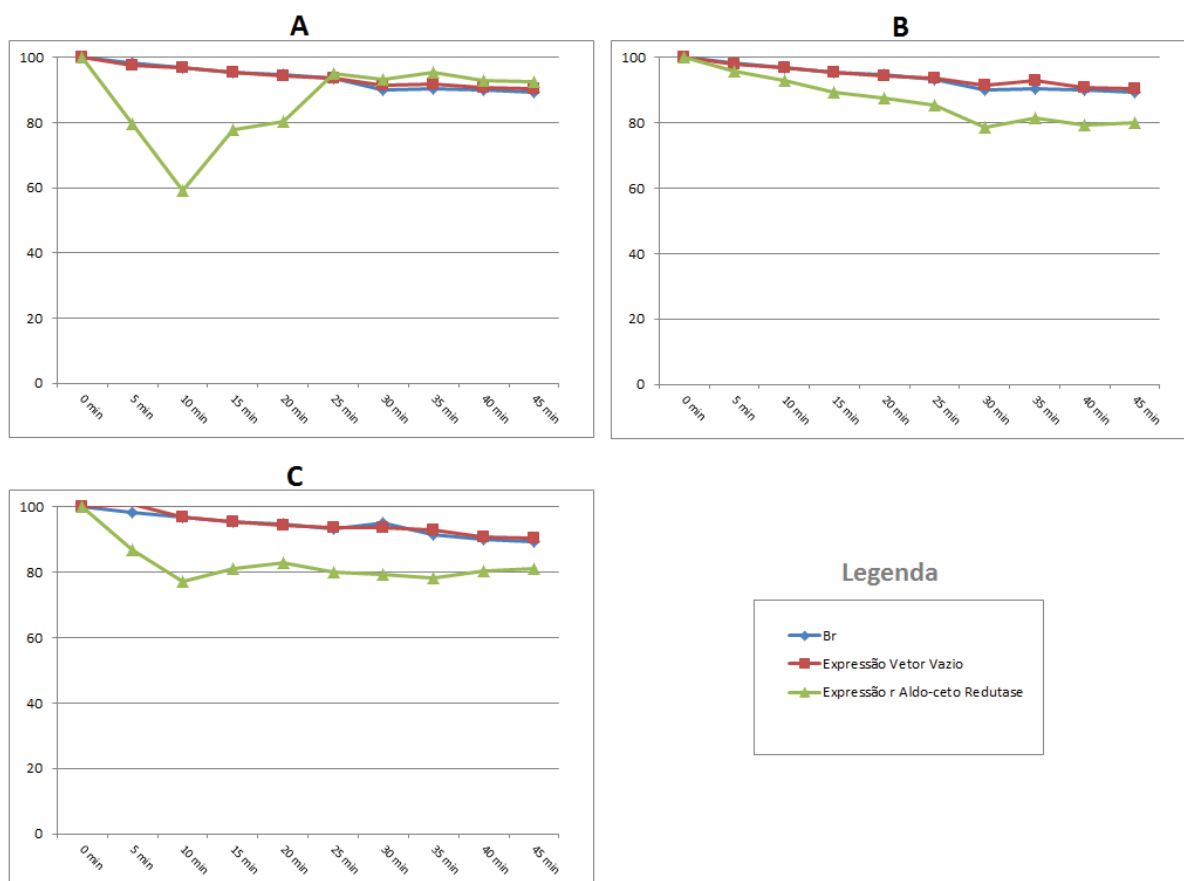


**Figura 16. Eletroforese e *western blotting* do processo de expressão e purificação por cromatografia de troca iônica da proteína TlicAKR-1.** (A) SDS-PAGE a 12%: 1) sobrenadante da expressão do vetor vazio (4 dias). 2) sobrenadante da expressão de TlicAKR-1 (4 dias). 3) 1 µg de TlicAKR-1 purificado por cromatografia de troca iônica (B) *western blotting* com Anti-His a partir de amostras em A.

Amostras do sobrenadante da expressão de TlicAKR-1 (quatro dias) foram utilizadas para os testes de atividade enzimática, pois a enzima do extrato purificado por troca iônica não apresentava sinais de atividade em nenhum substrato. O extrato bruto contendo a enzima apresentou máxima atividade sob o substrato 2-nitrobenzaldeído, uns dos substratos mais comuns das AKR (YAMAMOTO e WILSON, 2013), obtendo-se ~40% de atividade em 10 minutos em comparação com os controles (FIGURA 17A). A enzima já começou a mostrar atividade a partir dos primeiros minutos de reação, caindo gradativamente após 10 minutos. Uma AKR da família 1 proveniente do cupim *C. gestroi* também apresentou atividade enzimática mais alta com este substrato, além de outros aldeídos utilizados e numa temperatura e pH semelhantes (25 °C e pH 6,5) (TRAMONTINA *et al.*, 2017), porém utilizando a enzima purificada a diferença da TlicAKR-1. Ao igual que a AKR1 do cupim, TlicAKR-1 apresentou atividade na reação quando foi utilizado NADPH como cofator, e atividade não detectável com NADH. Geralmente as AKR preferem NADPH ao NADH (BARSKI *et al.*, 2008); a forte ligação que estas enzimas têm com o NADPH, permite algumas AKRs (por exemplo, aldose redutase), uma vantagem termodinâmica para executar a reação oxiredutiva, sem colocar muita demanda energética no substrato, pois a maior parte da energia para reduzir o grupo carbonila do substrato provem do

nucleotídeo do cofator. Dessa forma, mesmo os substratos que se ligam fracamente aos resíduos do sítio ativo, são reduzidos com alta eficiência. Como resultado, a aldose redutase reduz uma grande gama de aldeídos (GRIMSHAW, 1992).

Dentre as aldoses utilizadas como substratos para os ensaios de atividade da TlicAKR-1, apenas D-Glicose e DL-Xilose apresentaram sinais de atividade, os outros substratos (frutose, arabinose, manose, manitol, glicerol, metilglicoxal) não apresentaram resultados diferenciais dos controles negativos). Para D-glicose a curva mostrou que a enzima teve atividade de pelo menos 20% aos 10 minutos de reação, e a enzima se manteve funcional mesmo após os 45 minutos (FIGURA 17C). Com DL-Xilose a curva foi aumentando lentamente e só aos 30 minutos se mostrou significativa, mantendo-se dessa forma até o tempo final da reação (FIGURA 17B). A enzima TlicAKR-1 parece ter características para ser uma verdadeira Aldose redutase, além das análises filogenéticas feitas, a atividade sobre glicose e xilose também comprovam esta hipótese, apesar de ter excelente resultado com um aldeído também. Sabe-se que muitas aldose redutases estão envolvidas nas complicações diabéticas e galactosemias (diabetes), já que catalisa a glicose para transformá-la em sorbitol, e galactose em galactitol; por isso é são alvo de pesquisa os inibidores de AR que já são utilizados em pacientes com estas doenças. (BOHREN *et al.*, 1991). As ARs possuem especificidade por vários substratos, podendo ser maior a preferência por um mais do que outro, que também podem coincidir com a especificidade de outras AKR; o que sugere que uma das funções desta família de redutases pode ser a desintoxicação de xenobióticos contendo carbonil. (VANDER JAGT *et al.*, 1995). Não se tem certeza das funções pontuais da AR nos insetos, mas existem evidências que possa estar envolvida no crescimento e desenvolvimento, além de desintoxicação e inibição de defesas de plantas. (STAFFORD-BANKS *et al.*, 2014; WANG *et al.*, 2018).



**Figura 17.** Ensaio de atividade de TlicAKR-1, com queda gradual do cofator NADPH (0,5 mM) em função do tempo, com leitura da absorbância a 340 nm, pH6,5, temperatura ambiente e diferentes substratos: A) 100 mM de 2-nitrobenzaldeido. B) 100 mM de DL-Xilose. C) 100 mM de D-Glicose. Os gráficos são apresentados com curvas do branco, atividade do vetor vazio e atividade do TlicAKR-1.

Nos humanos, as complicações diabéticas acontecem pela superexpressão de AR, que causar acúmulo de sorbitol no organismo (RANSOHOFF e CARDONA, 2010). Esta enzima é significativamente expressa em pacientes diabéticos, e um dos sintomas clínicos geralmente é o ganho de peso; isto último quer dizer que se a AR é expressa em níveis normais, é uma enzima importante no crescimento e desenvolvimento de organismos (DREL *et al.*, 2006). Wang e colaboradores encontraram na epiderme de uma *B. mori* mutada, alguns genes envolvidos no metabolismo da glicose, incluindo um que codifica uma AR (BmAR). Esta AR é expressa em quantidades muito baixas, ao contrário de uma *B. mori* selvagem, sem mutações, onde a AR é expressa em quantidades normais, e esta envolvida no bom crescimento e desenvolvimento da lagarta (WANG *et al.*, 2017).

Esta mutação causa uma cor roxa na epiderme da lagarta de seda, além de deficiências no crescimento e perda de peso. Foi comprovado que BmAR produz

sorbitol no organismo da lagarta, a partir da reação com glicose e NADPH (WANG *et al.*, 2018). Esta enzima possui uma sequência de aminoácidos de identidade alta com TlicAKR-1 e ambas enzimas possuem a maioria de aminoácidos essenciais, que participam na catalises de aldoses.

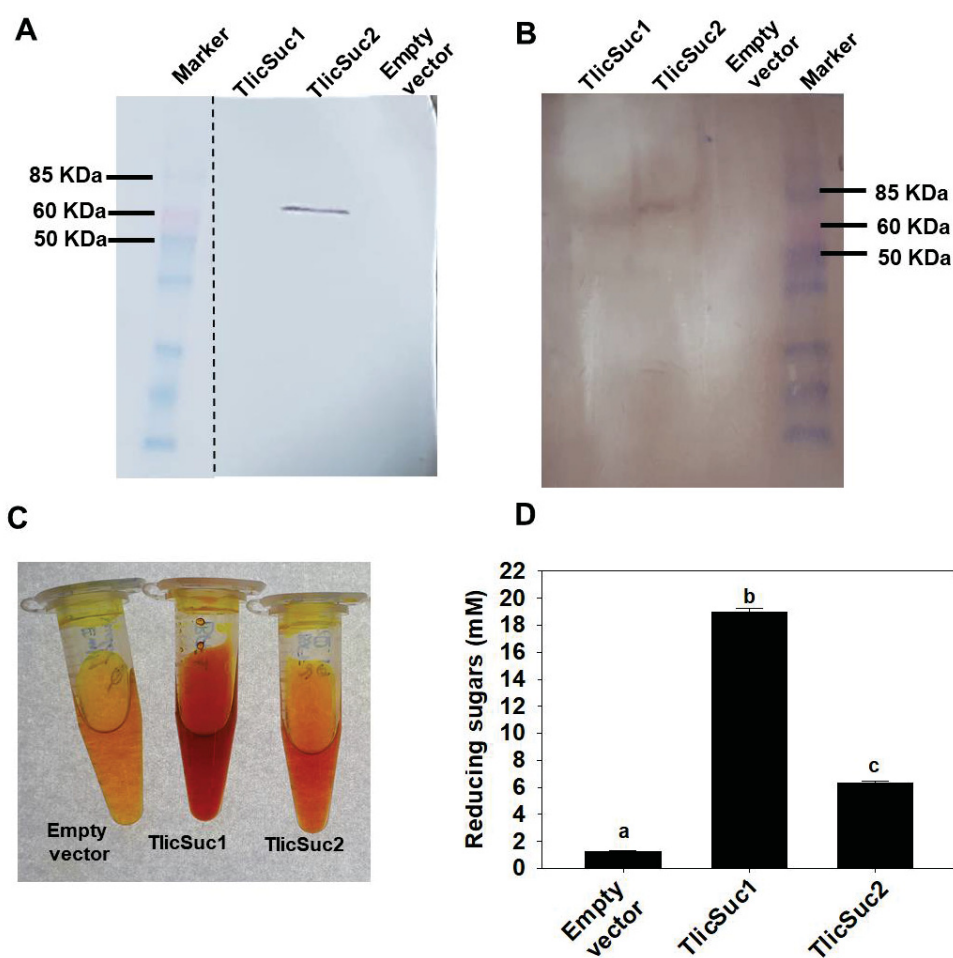
Não entanto, TlicAKR-1 também demonstrou afinidade com o substrato nitrobenzaldeído, igual ao que a AKR do cupim e Aldehide redutase de humano (AKR1A1) (STEWART *et al.*, 2011). Este composto é um dos muitos inibidores de fermentação na produção de etanol 2G. Segundo Wahlbom, estes compostos podem causar o esgotamento de cofatores redox durante a fermentação, resultando em uma produção mais lenta e baixa de etanol (WAHLBOM e HAHN-HÄGERDAL, 2002). Os inibidores podem ser mais ou menos prejudiciais, dependendo do grupo funcional que os compõem, sendo que os aldeídos são mais tóxicos do que os outros (JÖNSSON e MARTÍN, 2016). A Aldo-ceto redutase do cupim *C. gestroi* (*CgAKR-1*) também tem capacidade de reagir com outros inibidores da fermentação, como o seringadeído, hidroxibenzaldeído, HMF e furfural (TRAMONTINA *et al.*, 2017).

#### 6.4.3 TlicSuc1 e TlicSuc2

As sacarases expressas em *P. pastoris* foram secretadas no sobrenadante da cultura e esse estrato bruto foi utilizado para a realização dos ensaios de detecção das proteínas e atividade enzimática. Assim como a TlicAKR-1, as sacarases também não puderam ser purificadas na coluna de níquel, apesar de ser adicionado a His-Tag terminal. Primeiramente, a técnica de *western blot* confirmou a expressão da TlicSuc2 (~60 KDa), mas não se observou sinais da TlicSuc1 (FIGURA 18A). Após o tratamento das amostras com ureia 6 M, foi observada a aparição da banda correspondente a TlicSuc1, também com um peso molecular ~60 KDa (FIGURA 18B). Provavelmente a TlicSuc1 manteve sua estrutura nativa com a cauda 6xHis dobrada para dentro, impedindo a detecção no *western blot*, mas o tratamento com ureia desnaturou a proteína, deixando-a solúvel. Para a aplicação das amostras no gel de policrilamida é necessário que as proteínas estejam solúveis, ou seja, disponíveis na sua estrutura primária. O SDS é um reagente desnaturante, que provoca o desdobramento da estrutura quaternária, mas muitas vezes essa estrutura permanece firme, não sendo possível quebrar essas interações não-covalentes e pontes dissulfeto. A adição de

Ureia, Thioureia e outros detergentes são usados como agentes solubilizantes e desnaturantes para o tratamento das proteínas (RABILLOUD, 2009).

A especificidade do substrato das duas  $\beta$ -FFases foi testada usando sacarose como substrato. Ambas sacarases demonstraram atividade sobre o substrato testado, que possui os resíduos não redutores de  $\beta$ -D-frutofuranosídeos; este resultado indica que os genes *TlicSuc1* e *TlicSuc2* codificam a enzima  $\beta$ -frutofuranosidase (EC 3.2.1.26). Os ensaios de atividade com sacarose, utilizando DNS, exibiram uma coloração laranja mais escura para *TlicSuc1*, indicando a quebra do substrato e liberação de glicose reduzida, já a *TlicSuc2* apresentou uma cor mais clara, indicando uma atividade mais baixa do que a outra sacarase (FIGURA 18C). Este ensaio quantificado pelo espectrofotômetro confirmou que *TlicSuc1* liberou muito mais açúcares reduzidos que *TlicSuc2*, indicando maior potencial enzimático com sacarose (FIGURA 18D), em pH 7 e temperatura de 30 °C. *TlicSuc1* se desenvolveu bem em pH 7 e temperatura de 30 °C, o que coincide com a  $\beta$ -FFase de *B. mori* (GAN *et al.*, 2018; DAMON *et al.*, 2008), porém difere da  $\beta$ -FFase “ácida” do besouro *S. levis*, o qual teve uma ótima atividade em pH 5. É possível que *TlicSuc2* não tenha mostrado um melhor resultado de atividade pelas condições de pH e temperatura, sendo possível desenvolver melhor em outras condições. São necessários mais ensaios destas sacarases de *T. licus licus* para comprovar a eficiência das mesmas em diferentes pH e temperaturas e melhor caracterizá-las como enzima purificadas.



**Figura 18. Expressão heteróloga dos genes *TlicSuc1* e *TlicSuc2* em *P. pastoris* e atividade enzimática das proteínas expressas.** (A) *Western blot* de amostras de proteínas totais do sobrenadante da cultura, mostrando sinal para a proteína *TlicSuc2* apenas. (B) *Western blot* com amostras de proteínas totais do sobrenadante da cultura, tratadas com ureia 6 M, mostrando sinal para ambas as proteínas. (C). Teste de DNS mostrando a reação de quebra de sacarose pela ação das proteínas *TlicSuc1* e *TlicSuc2* presentes nas proteínas totais do sobrenadante de ambas culturas. (D) Quantificação da atividade hidrolítica com o substrato sacarose por espectrometria. Os valores foram analisados estatisticamente por análise de variância e teste de Tukey para comparação de médias ( $P < 0,05$ ). Em todos os passos as amostras foram comparadas sempre com a expressão do vetor vazio.

## 7 CONCLUSÃO

O presente trabalho contribuiu no avanço dos estudos envolvendo a engenharia de enzimas, potencialmente aplicáveis na produção de etanol. Os estudos com enzimas endógenas de insetos ainda precisam ser mais aprofundados, buscando as diferentes possibilidades e alternativas de organismos e enzimas candidatas, visando uma produção eficiente do álcool.

As enzimas reportadas neste estudo podem ser consideradas prováveis candidatas para suas aplicações em misturas enzimáticas, capazes de agir sinergicamente na degradação de biomassa vegetal. Todos os genes que codificam essas enzimas foram isolados, subclonados e expressos em *P. pastoris*, com sucesso. A enzima  $\beta$ -1,4-endoglucanase do inseto-praga *A. grandis* (AgraGH45-1) foi capaz de degradar eficientemente os substratos análogos da celulose, CMC e HEC. Esta demonstrou ser mais eficiente que a enzima  $\beta$ -1,4-endoglucanase comercial de *A. niger*, e sua atividade permaneceu estável, quando analisada na faixa de 40 °C - 60 °C, o que indica ser uma enzima com características desejáveis para a indústria de etanol 2G. Contudo, ainda é necessário comprovar a eficácia desta enzima sobre substratos complexos, derivados de biomassas.

Em relação a enzima Aldo-ceto redutase da *T. licus licus* (TlicAKR-1), esta demonstrou afinidade por 2-nitrobenzaldeído e pelas aldoses D-glicose e DL-xilose. Ainda não está totalmente esclarecido qual seria o melhor substrato para a redutase TlicAKR-1, pois serão necessários testes no estado purificado da enzima, e em condições variadas de pH e temperatura. A função desta enzima no organismo da lagarta ainda é desconhecida. Diferentes estudos apontam que as AKRs de insetos são importantes para o desenvolvimento destes organismos, porém, os estudos são ainda escassos, necessitando mais estudos aprofundados sobre estas enzimas.

No caso das sacarases de *T. licus licus* (TlicSuc1 e TlicSuc2), estas comprovaram ser verdadeiras  $\beta$ -fructofuranosidases, conseguindo catalisar o substrato específicos da enzima. Apesar dos poucos relatos desta enzima no reino animal, pode-se presumir que seja importante entre algumas espécies de Lepidóptera e Coleóptera, como uma das capacidades de adaptação à sobrevivência. Estas sacarases também poderiam ser prováveis candidatas para coquetéis enzimáticos,

porém, será fundamental a realização de novos experimentos para melhor comprovação, além da necessidade de purificar ambas as enzimas, realizar os testes de condicionamento, além de otimizar a atividade enzimática de TlicSuc2.

## REFERÊNCIAS

- ABRAMOVAY, R. Biocombustíveis – a energia da controvérsia. Editora Senac. 1º Edição. 2009.
- AZAMBUJA, R; DEGRANDE, P.E. Trinta anos do bicudo-do-algodoeiro no Brasil. Programa de Pós-graduação em Entomologia e Conservação da Biodiversidade; Faculdade de Ciências Biológicas e Ambientais; Universidade Federal da Grande Dourados. 2014.
- BACHUR, N.R. Cytoplasmic aldo-keto reductases: a class of drug metabolizing enzymes. Science. PubMed: 959821. 193:595–597, 1976.
- BARSKI, O.A; TIPPARAJU, S.M; BHATNAGAR, A. The Aldo-Keto Reductase Superfamily and its Role in Drug Metabolism and Detoxification. Institute of Molecular Cardiology, University of Louisville, Louisville, KY. NIH Public Access, Drug Metab. Vol. 40(4): 553–624. doi:10.1080/03602530802431439, 2009.
- BELOT, J.L. O bicudo-do-algodoeiro (*Anthonomus grandis*) nos cerrados brasileiros: biologia e medidas de controle. Instituto Mato-grossense do Algodão. ISBN: 978-85-66457-07-0. 254 p, 2015.
- BENEDINI, M.S; CONDE, A. J. Broca gigante: nova praga da cana-de-açúcar na região centro-sul. Revista Coplana. 25: 24-25, 2008.
- BHATIA, S.C. Ethanol. In: Bhatia SC, ed. Advanced renewable energy systems. Part I. New Delhi: Woodhead Publish in gIndia PVT LTD. 523–572, 2014.
- BOHREN, K. M; PAGE, L; SHANKAR, R; HENRY, P; GABBAY, H. Expression of Human Aldose and Aldehyde Reductases. The Journal of Biological Chemistry. The American Society for Biochemistry and Molecular Biology, Vol. 266. No. 35. 15 p, 24031-24037, 1991.
- BRUINENBERG, P.M; VAN DIJKEN, J.P; SCHEFFERS, A. An Enzymic Analysis of NADPH Production and Consumption in *Candida utilis*. Laboratory of Microbiology, University of Technology, Netherlands. Journal of General Microbiology 129, 965-971, 1982.
- BUCKERIDGE, M. S; GOLDMAN, G. H. Routes to cellulosic ethanol. Editora Springer. 1º Edição. 2011.
- BUSCH, A; DANCHIN, E. G. J; PAUCHET, Y. Functional diversification of horizontally acquired glycoside hydrolase family 45 (GH45) proteins in *Phytophaga* beetles. BMC Evol. Biol. 19, 100, 2019.

CANTAREL, B. L.; COUTINHO, P. M.; RANCUREL, C.; BERNARD, T.; LOMBARD, V.; HENRISSAT, B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Research*, Oxford. Vol 37, p. 233-238, 2009.

CANTAREL, B. L. et al. The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* 37, D233–D238, 2009.

CAIRO, J.P.L.; CARAZZOLLE, M.F.; LEONARDO, F.C.; MOFATTO, L.S.; BRENELLI, L.; GONÇALVES, T.A.; UCHIMA, C.; DOMINGUES, R.R.; ALVAREZ, T.M.; TRAMONTINA, R.; VIDAL, R.; COSTA, F.F.; LEONARDO, A.M.; LEME, A.F.; PEREIRA, G.A.G.; SQUINA, F.M. Expanding the Knowledge on Lignocellulolytic and Redox Enzymes of Worker and Soldier Castes from the Lower Termite *Coptotermes gestroi*. *Frontiers in Microbiology*. Vol 7. DOI: 10.3389/fmicb.01518, 2016.

CASTRO, A. M. Produção e propriedades de celulases de fungos filamentosos obtidas a partir de celulignina de bagaço de cana-de-açúcar (*Saccharum* spp). Dissertação de Mestrado, Universidade Federal do Rio de Janeiro. 2006.

CHAHED, H.; BOUMAIZA, M.; EZZINE, A.; MARZOUKI, M. N. Heterologous expression and biochemical characterization of a novel thermostable *Sclerotinia sclerotiorum* GH45 endoglucanase in *Pichia pastoris*. *Int. J. Biol. Macromol.* 106, 629–635, 2018. Disponível em, <https://doi.org/10.1016/j.ijbiomac.2017.08.062>

CREGG, J. M.; CEREGHINO, L.; SHI, J.; HIGGINS, D. R. Recombinant protein expression in *Pichia pastoris*. *Mol. Biotechnol* 16, 23–52, 2000.

CREGG, J. M.; CEREGHINO, G. P. L.; CEREGHINO, J. L.; ILGEN, C. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Curr. Opin. Biotechnol* 13: 329-332, 2002.

DAIMON, T.; TAGUCHI, T.; MENG, Y.; KATSUMA, S.; MITA, K.; SHIMADA, T. B-fructofuranosidase genes of the silkworm, *Bombyx mori*: insights into enzymatic adaptation of *B. mori* to toxic alkaloids in mulberry latex. *J. Biol. Chem.* 283, 15271 e 15279. 2008.

DEVI, M.B.; SINGH, S.N. Physiology of digestion in mango weevil, *Sternochetus gravis* (Curculionidae: Coleoptera). *J. Exp. Sci.* 2 (12), 07–09, 2011.

DREL, V.R.; MASHTALIR, N.; ILNYTSKA, O.; SHIN, J.; LI, F.; LYZOGUBOV, V.V.; OBROSOVA, I.G. The leptin-deficient (ob/ob) mouse: a new animal model of peripheral neuropathy of type 2 diabetes and obesity. *Diabetes* 55, 3335, 2006.

DOBLIN, M. S., JOHNSON, K. L., HUMPHRIES, J., NEWBIGIN, E. J. & BACIC, A. Are designer plant cell walls a realistic aspiration or will the plasticity of the plant's metabolism win out. *Curr. Opin. Biotechnol* 26, 108–114, 2014.

ELLIS, E.M. Microbial aldo-keto reductases. *FEMS Microbiol. Lett* 216:123–131. PubMed:12435492. 2002.

FARINAS, C.S. *A Parede Celular Vegetal E As Enzimas Envolvidas Na Sua Degradação*. 1 ed. Embrapa Instrumentação, São Paulo. 2011.

FELDMAN, D; KOWBEL, D.J; GLASS, N.L; YARDEN, O; HADAR, Y. Detoxification of 5-hydroxymethylfurfural by the *Pleurotus ostreatus* lignolytic enzymes aryl alcohol oxidase and dehydrogenase. *Biotechnol Biofuels* 8:63, 2015.

FISCHER, R; OSTAFE, R; TWYMAN, R. M. Cellulases from insects. *Advances in Biochemical Engineering/biotechnology*, 136, 51–64. doi:10.1007/10\_2013\_206, 2013.

GAN, Q; ZHANG, X; ZHANG, D; SHI, L; ZHOU, Y; SUN, T. BmSUC1 is essential for glycometabolism modulation in the silkworm, *Bombyx mori*. *BBA - Gene Regulatory Mechanisms*, 1861(6), 543–553, 2018. Disponível em, <http://doi.org/10.1016/j.bbagr.2018.04.002>

GANEVA, V.; GALUTZOV, B.; TEISSIE, J. Electric field mediated loading of macromolecules in intact yeast cells is critically controlled at the wall level. *Biochim. Biophys. Acta* 1240:229-236, 1995.

GAO, J. et al. Characterization and crystal structure of a thermostable glycoside hydrolase family 45 1,4- $\beta$ -endoglucanase from *Thielavia terrestris*. *Enzyme Microb. Technol.* 99, 32–37, 2017. Disponível em, <https://doi.org/10.1016/j.enzmictec.2017.01.005>

GELLISSEN, G. Heterologous protein production in methylotrophic yeast. *Appl. Microbiol. Biotechnol* 54: 741-750, 2000.

GIRARD, C; JOUANIN, L. Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, *Phaedon cochleariae*. *Insect Biochem. Mol. Biol.* 29:1129–42, 1999

GOUVEIA, E.R; NASCIMENTO, R.T; SOUTO-MAIOR, A.M; ROCHA, G.J.M. Validação de metodologia para a caracterização química de bagaço de cana-de-açúcar. *Quim Nova* 32:1500-03, 2009.

GRINNA, L.S; TSCHOPP, J.F. Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, *Pichia pastoris*. *Yeast* 5: 107–115, 1989.

GRIMSHAW CE. Aldose reductase: model for a new paradigm of enzymic perfection in detoxification catalysts. PubMed: 1420136. *Biochemistry* 31:10139–10145, 1992.

HOHENBLUM, H; GASSER, B; MAURER, M; BORTH, N; MATTANOVICH, D. Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol. Bioeng.* 85: 367-375, 2004.

Instituto de Economia Agrícola (IEA). “Bioenergy agreement: Task 32 – Biomass combustion and cofiring. 2001. Disponível em [http://www.ieabcc.nl/meetings/32\\_01\\_meeting\\_Zuerich.pdf](http://www.ieabcc.nl/meetings/32_01_meeting_Zuerich.pdf).

Instituto de Economia Agrícola (IEA), Secretaria de Agricultura e Abastecimento. Artigos: Cana-de-açúcar: produção e processamento em 2019. 2020. Disponível em <http://www.iea.sp.gov.br/out/TerTexto.php?codTexto=14767>.

JEZ, J.M; FLYNN, T.G; PENNING, T.M. A nomenclature system for the aldo-keto reductase superfamily. *Adv. Exp. Med. Biol.* Vol. 414 pp. 579–600, 1997.

JIMENEZ, A.V; WANG, H; SIEGFRIED, B.D. Expression and Characterization of a Recombinant Endoglucanase From Western Corn Rootworm, in *Pichia pastoris*. *Journal of Insect Science.* i. 14(242), 2014. DOI: 10.1093/jisesa/ieu104

JÖNSSON, L.J; MARTÍN, C. Pretreatment of lignocellulose: formation of inhibitory by-products and strategies for minimizing their effects. *Bioresour Technol* 199:103–12, 2016.

KASANA, R.C; SALWAN, R; DHAR, H; DUTT, S; GULATI, A. A Rapid and Easy Method for the Detection of Microbial Cellulases on Agar Plates Using Gram’s Iodine. *Curr Microbiol* 57:503–507, 2008. DOI 10.1007/s00284-008-9276-8

KIM, N; CHOO, Y.M; LEE, K.S; HONG, S.J; SEOL, K.Y; JE, Y.H; SOHN, H.D; JIN, B.R. Molecular cloning and characterization of a glycosyl hydrolase family 9 cellulase distributed throughout the digestive tract of the cricket *Teleogryllus emma*. *Comparative Biochemistry and Physiology* 150B, 368–376, 2008.

LEE, S. J; KIM, S. R; YOON, H. J; KIM, I; LEE, K. S; JE, Y. H; JIN, B. R. cDNA cloning, expression, and enzymatic activity of a cellulase from the mulberry longicorn beetle, *Apriona germari*. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, 139 (1), 107–16, 2004.

LEE, S. J.; LEE, K. S.; KIM, S. R.; GUI, Z. Z.; KIM, Y. S.; YOON, H. J.; JIN, B. R. A novel cellulase gene from the mulberry longicorn beetle, *Apriona germari*: gene structure, expression, and enzymatic activity. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, 140(4), 551–60. <http://doi.org/10.1016/j.cbpc.2004.12.003>, 2005.

LI, Z.W; SHEN, Y.H; XIANG, Z.H; ZHANG, Z. Pathogen-origin horizontally transferred genes contribute to the evolution of Lepidopteran insects. *BMC Evolutionary Biology* 11:356, 2011.

LI, H; LU, J; MO, J. Physiochemical lignocellulose modification by the formosan subterranean termite *Coptotermes formosanus* Shiraki (Isoptera: Rhinotermitidae) and its potential uses in the production of biofuels. *BioResources* 7:675-85, 2012.

LINDE, D; MACIAS, I; FERNANDEZ-ARROJO, L; PLOU, F.J; JIMENEZ, A; FERNANDEZ-LOBATO, M. Molecular and biochemical characterization of a  $\beta$ -fructofuranosidase from *Xanthophyllomyces dendrorhous*. *Appl. Environ. Microbiol* 75, 1065–1073, 2009.

LIU, Z.L; MA, M; SONG, M. Evolutionarily engineered ethanologenic yeast detoxifies lignocellulosic biomass conversion inhibitors by reprogrammed pathways. *Mol Genet Genomics* 282:233–44, 2009.

LIU, D; ZHANG, R; YANG, X; ZHANG, Z; SONG, S; MIAO, Y; SHEN, Q. Characterization of a thermostable beta-glucosidase from *Aspergillus fumigatus* Z5, and its functional expression in *Pichia pastoris* X33. *Microb. Cell Fact.* 11, 25, 2012.

LIU, X; XU, W; MAO, L; ZHANG, C; YAN, P; XU, Z; et al. Lignocellulosic ethanol production by starch-base industrial yeast under PEG detoxification. *Sci Rep.* 6:20361, 2016.

LIU, G. *et al.* Functional and structural analyses of a 1,4–endoglucanase from *Ganoderma lucidum*. *Enzyme Microb. Technol.* 86, 67–74, 2016. Disponível em, <https://doi.org/10.1016/j.enzmictec.2016.01.013>

LOPES, M.L; PAULILLO, S.C; GODOY, A; CHERUBIN, R.A; LORENZI, M.S; CARVALHO, F.H; DOMINGOS, C; AMORIM, H.B; AMORIM, H.V. Ethanol Production in Brazil: A Successful History Ethanol Production : Impact in Brazil Fermentation Process : Main Characteristics, 44–47, 2006.

LYND, L. R; WEIMER, P. L; VAN ZYL, W. H; PRETORIUS, I. S. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol.* 66: 506-577, 2002.

MACAULEY-PATRICK, S; FAZENDA, M. L; MCNEIL, B; HARVEY, L. M. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast.* Chichester, England, 22 (4), p. 249–70, 2005. Disponível em: <http://doi.org/10.1002/yea.1208>.

MARQUES, F. Bioenergia: Obstáculos no caminho. *Pesq. FAPESP* 268, 58–63, 2018. Disponível em, <https://revistapesquisa.fapesp.br/2018/06/18/obstaculos-no-caminho/>

MARTIN, M.M. Cellulose digestion in insects. *Comparative Biochemistry and Physiology* 75, 426–428, 1983.

MARTINS, L. F. Caracterização do complexo celulásico de *Penicillium echinulatum*. Dissertação (Mestrado em Química Orgânica) – Universidade Federal do Paraná, Curitiba, PR. 2005; 121 p.

MASON, C.J; SCULLY, E.D; GEIB, S.M; HOOVER, K. Contrasting diets reveal metabolic plasticity in the tree-killing beetle, *Anoplophora glabripennis* (Cerambycidae: Lamiinae). Scientific Reports 6:33813, 2016. DOI: 10.1038/srep33813

MILLER, L. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. 31: 426-428, 1959.

MINDNICH, R.D; PENNING, T.M. Aldo-keto reductase (AKR) superfamily: Genomics and annotation. Department of Pharmacology, University of Pennsylvania, School of Medicine, Philadelphia, USA. Henry Stewart publications 1479–7364. Human genomics Vol 3. NO 4. 362–370, 2009.

Ministério de Minas e Energia (MME). Boletim mensal dos combustíveis renováveis. 2014. Disponível em [www.mme.gov.br/spg/menu/publicacoes.html](http://www.mme.gov.br/spg/menu/publicacoes.html).

MIRANDA, J. E; RODRIGUES, S. M. M. História do bicudo no Brasil Capítulo 1, p. 10-44, Embrapa algodão. 2015. Disponível em, <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/141686/1/Historia-do-bicudo-no-brasil.pdf>

MONCEAUX, D.A. Alternative feedstocks for fuel ethanol production. The alcohol textbook. A reference for the beverage, fuel and industrial alcohol industries. 5th ed. Nottingham: Nottingham University Press; 2009:47–57.

NAKASHIMA, K; WATANABE, H; SAITOH, H; TOKUDA, G; AZUMA, J.I. Dual cellulose digesting system of the wood-feeding termite, *Coptotermes formosanus* Shiraki. Insect Biochem. Mol. Biol. 32, 777–784, 2002.

NEGRISOLI JUNIOR, AS; BALDANI, JI; GROSSI- DE- SÁ, MF; SILVA, MCM; MACEDO, LLP; FONSECA, FCA; BARBOSA, CRC; GUZZO, NEC. Manejo da broca-gigante da cana-de-açúcar (*Telchin licus*) (Drury) (Lepidoptera: Castniidae) no nordeste do Brasil. Aracaju: Embrapa Tabuleiros Costeiros. Documentos / Embrapa Tabuleiros Costeiros, ISSN 1678-1953; 198, 2015.

NIEVES, R. A; EHRMAN, C. I; ADNEY, W. S; ELANDER, R. T; HIMMEL, M. E. Technical Communication: survey and analysis of commercial cellulose preparations suitable for biomass conversion to ethanol. World J. Microbiol. Biotechnol. 14, 301–304, 1998.

OLIVEIRA, M. Entre açúcares e genes: Aplicação do conhecimento científico da cana deverá servir ao desenvolvimento de novas ferramentas para a produção de etanol. Tecnologia, Biocombustíveis, p. 86-91, 2012.

OLIVEIRA, A.E. Expressão de celulasas recombinantes de insetos praga em *Pichia pastoris*. Tese de doutorado em biotecnologia e biodiversidade. Universidade de Brasília. Rede centro-oeste de pós-graduação, pesquisa e inovação. 2020.

ÖZSU, N; MONTEIRO, A. Wound healing, calcium signaling, and other novel pathways are associated with the formation of butterfly eyespots. *BMC Genomics* 18:788, 2017. DOI 10.1186/s12864-017-4175-7

PAUCHET, Y; MUCK, A; SVATOS, A; HECKEL, D.G; PREISS, S. Mapping the larval midgut lumen proteome of *Helicoverpa armigera*, a generalist herbivorous insect. *J. Proteome Res* 7. 1629-1639, 2008.

PAUCHET, Y; WILKINSON, P; VOGEL, H; NELSON, D.R; REYNOLDS, S.E; HECKEL, D.G; FFRENCH-CONSTANT, R.H. Pyrosequencing the *Manduca sexta* larval midgut transcriptome: messages for digestion, detoxification and defence. *Insect Mol. Biol.* 19 Vol (1) 61-75, 2010.

PEDEZZI, R; FONSECA, F.P; JÚNIOR, C.D; KISHI, L.T; TERRA, W. R; SILVA, F.H. A novel  $\beta$ -fructofuranosidase in Coleoptera: Characterization of a  $\beta$ -fructofuranosidase from the sugarcane weevil, *Sphenophorus levis*, *Insect Biochemistry and Molecular Biology*. Science Direct. 55, 31–38, 2014. Disponível em, <http://doi.org/10.1016/j.ibmb.2014.10.005>

PECIULYTE, A; KARLSTRÖM, K; LARSSON, P.T; OLSSON, L. Impact of the supramolecular structure of cellulose on the efficiency of enzymatic hydrolysis. *Biotechnol Biofuels* 8:56, 2015.

PEL, H. J. *et al.* Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat. Biotechnol.* 25, 221–231, 2007.

PONS, T; NAUMOFF, D.G; MARTINEZ-FLEITES, C; HERNANDEZ, L. Three acidic residues are at the active site of a beta-propeller architecture in glycoside hydrolase families 32, 43, 62, and 68. *Proteins* 54, 424–432, 2004.

RABILLOUD, T. Solubilization of Proteins in 2DE. PubMed, *Molecular Biology* 519:19-30, 2009. Disponível em, DOI: 10.1007/978-1-59745-281-6\_2

RANSOHOFF, R.M; CARDONA, A.E. The myeloid cells of the central nervous system parenchyma. *Nature* 468, 253–262, 2010.

SANDGREN, M; STAHLBERG, J; MITCHINSON, C. Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. *Progr. Biophys. Mol. Biol.* 3: 249–291, 2005.

SCHARF, M.E; TARTAR, A. Termite digestomes as sources for novel lignocellulases. *Biofuels Bioprod Biorefining.* 2:540-52, 2008.

SCHUTTER DE, K; LIN, Y; TIELS, P; VAN HECKE, A; GLINKA, S; WEBER-LEHMANN, J; ROUZE, P; VAN DE PEER, Y; CALLEWAERT, N. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat. Biotechnol.* 27:561-566, 2009.

SCRIVENER, A. M; WATANABE, H; NODA, H. Diet and carbohydrate digestion in the yellow-spotted longicorn beetle *Psacotheta hilaris*. J. Insect Physiol, p. 43:1039–52, 1997.

SHUMIAO, Z; HUANG, J; ZHANG, C; DENG, L; HU, N; LIANG, Y. High-level expression of an *Aspergillus niger* endo-beta-1,4-glucanase in *Pichia pastoris* through gene codon optimization and synthesis. J. Microbiol. Biotechnol. 20(3): 467-473, 2010.

SINGHANIA, R. R; SUKUMARAN, R. K; PATEL, A. K; LARROCHE, C; PANDEY, A. Advancement and comparative profiles in the production Technologies using solid-state and submerged fermentation for microbial cellulases. Enzyme Microbial. Technol. 46: 541-549, 2010.

SRISODSUK, M; REINIKAINEN, T; PENTTILÄ, TEERI, T. T. Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase-I in its interaction with crystalline cellulose. J. Biol. Chem. 268: 20756-20761, 1993.

SOUTINHO, L; ROSARIO, F; LIMA, A. Mapping of technological routes of the Brazilian energetic suction sector technological change in ethanol production derived from sugar cane. Rev. Econ. Pol. Desen. 6(18), 80–89, 2014.

SOUZA, W. R. *et al.* Suppression of a single BAHD gene in *Setaria viridis* causes large, stable decreases in cell wall feruloylation and increases biomass digestibility. New Phytol. 218, 81–93, 2018. Disponível em <https://doi.org/10.1111/nph.14970>

STAFFORD-BANKS, C. A; ROTENBERG, D; JOHNSON, B. R; WHITFIELD, A. E; ULLMAN, D. E. Analysis of the Salivary Gland Transcriptome of *Frankliniella occidentalis*. 9(4), 1–10, 2014. Disponível em <http://doi.org/10.1371/journal.pone.0094447>

STEWART AD, ANAND RR, LAIRD JS, VERRALL M, RYAN CG, DE JONGE MD, *et al.* Distribution of metals in the termite *Tumulitermes tumuli* (Froggatt): two types of Malpighian tubule concretion host Zn and Ca mutually exclusively. PLoS ONE 6:27578, 2011.

SU, L. J. *et al.* Evaluation of cellulolytic activity in insect digestive fluids. Gen and Mol Res 12, 2432–2441, 2013.

SUGA, M; HATAKEYAMA, T. High efficiency transformation of *Schizosaccharomyces pombe* pretreated with thiol compounds by electroporation. Yeast 18:1015-1021, 2001.

SUGIMURA, M; WATANABE, H; LO, N; SAITO, H. Purification, characterization, cDNA cloning and nucleotide sequencing of a cellulase from the yellow-spotted longicorn beetle, *Psacotheta hilaris*. Eur. J. Biochem. 270:3455–60, 2003.

SUN, W; YI-HONG, S; QI, D; XIANG, Z; ZHANG, Z. Molecular Cloning and Characterization of Ecdysone oxidase and 3-dehydroecdysone-3 $\alpha$ -reductase involved

in the Ecdysone Inactivation Pathway of Silkworm, *Bombyx mori*. International Journal of Biological Sciences. 8(1):125-138, 2012.

TERRA, WR; FERREIRA, C. Insect digestive enzymes: properties, compartmentalization and function. Comparative Biochemistry and Physiology. 109B: 1-62, 1994.

TERRA, W.R; FERREIRA, C. Biochemistry and molecular biology of digestion. In: Gilbert, L.I. (Ed.), Insect Molecular Biology and Biochemistry. Elsevier, Amsterdam, pp. 365-418, 2011.

TOKUDA, G; WATANABE, H; LO, N. Does correlation of cellulase gene expression and cellulolytic activity in the gut of termite suggest synergistic collaboration of cellulases gene. Eur. J. Biochem. 401, 131–134, 2007.

TOWBIN, H., STAEGELIN, T. AND GORDON, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets - procedure and some applications. Proceedings of the National Academy of Science, v.76, p.4350-4354. 1979.

TSCHOPP, J. F; SVERLOW, G; KOSSON, R; CRAIG, W; GRINNA, L. High Level Secretion of Glycosylated Invertase in the Methylophilic Yeast *Pichia pastoris*. Bio/Technology 5, 1305-1308, 1987.

TRAMONTINA, R; CAIRO, J.P; Liberato, M.V; Mandelli, F; Sousa, A; Santos, S; Rabelo, S.C; Campos, B; Ienczak, J; Ruller, R; Damásio, A.R.L; Squina, F.M. The *Coptotermes gestroi* aldo–keto reductase: a multipurpose enzyme for biorefinery applications. Biotechnol Biofuels. 10:4, 2017. DOI 10.1186/s13068-016-0688-6

TRAMONTINA, R; Brenelli, L.B; Sousa, A; Alves,R; Arenas, A.M.Z; Nascimento, V.M; Rabelo, S.C; Freitas, S; Ruller, R; Squina, F.M. Designing a cocktail containing redox enzymes to improve hemicellulosic hydrolysate fermentability by microorganisms. Science Direct. Enzyme and Microbial Technology. 2019.

União de Indústrias de Cana de Açúcar (UNICA). Notícias: Etanol. Consumo de etanol hidratado cresceu mais de 37% em 2015. 2016. Disponível em: <http://www.unica.com.br/noticia/14883647920325965467/consumo-de-etanol-hidratado-cresceu-mais-de-37-por-cento-em-2015/>

VANDER JAGT, D. L; KOLB, N.S; VANDER JAGT, T.J; CHINO, J; MARTINEZ, F.J; HUNSAKER, L.A; ROYER, R.E. Substrate specificity of human aldose reductase: identification of 4-hydroxynonenal as an endogenous substrate. University of New Mexico, School of Medicine, Albuquerque, NM 87131, USA. 1995.

VASCONCELOS, S.M DE; SANTOS, A.M; ROCHA, G.J; SOUTO-MAIOR, A.M. Diluted phosphoric acid pretreatment for production of fermentable sugars in a sugarcane-based biorefinery. Bioresour Technol. 135:46-52, 2013.

WAHLBOM, C.F; HAHN-HÄGERDAL, B. Furfural, 5-hydroxymethyl furfural, and acetoin act as external electron acceptors during anaerobic fermentation of xylose in recombinant *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 78:172–8, 2002.

WANG, Q; WANG, Z; SHEN, F; HU, J; SUN, F; LIN, L. Pretreating lignocellulosic biomass by the concentrated phosphoric acid plus hydrogen peroxide (PHP) for enzymatic hydrolysis: evaluating the pretreatment flexibility on feedstocks and particle sizes. *Bioresour Technol.* 166:420-28, 2014.

WANG, P; BI, S; WEI, W; QIU, Z; XIA, D; SHEN, X; ZHAO, Q. Downregulation of aldose reductase is responsible for developmental abnormalities of the silkworm purple quail-like mutant (q-lp). *Science Direct. Gene*, 665, 96–104, 2018. Disponível em, <http://doi.org/10.1016/j.gene.2018.05.010>

WANG, P; QIU, Z; XIA, D; TANG, S; SHEN, X; ZHAO, Q. Transcriptome analysis of the epidermis of the purple quail-like (q-lp) mutant of silkworm, *Bombyx mori*. *PLOS ONE*. 2017, Disponível em, <https://doi.org/10.1371/journal.pone.0175994>

WATANABE, H; TOKUDA, G. Animal cellulases. *Cell. Mol. Life Sci.* 58: 1167–78, 2001.

WATANABE, H; TOKUDA, G. Cellulolytic systems in insects. *Annual Review of Entomology.* 55, 609–32, 2010. Doi:10.1146/annurev-ento-112408-085319

WEI, Y.D; LEE, S.J; LEE, K.S; GUI, Z.Z; YOON, H.J. N-glycosylation is necessary for enzymatic activity of a beetle (*Apriona germari*) cellulase. *Biochem. Biophys. Res. Commun.* 329:331–36, 2005.

WEI, Y. D; LEE, K. S; GUI, Z. Z; YOON, H. J; KIM, I; ZHANG, G. Z; JIN, B. R. Molecular cloning, expression, and enzymatic activity of a novel endogenous cellulase from the mulberry longicorn beetle, *Apriona germari*. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, 145 (2), 220–9, 2006.

WEI, Y. D; LEE, K. S; GUI, Z. Z; YOON, H. J; KIM, I; JE, Y. H; JIN, B. R. N-linked glycosylation of a beetle (*Apriona germari*) cellulase Ag-EGase II is necessary for enzymatic activity. *Insect Biochemistry and Molecular Biology*, 36 (6), 435–441, 2006.

WERMUTH, B; BURGISSER, H; BOHREN, K; VON WARTBURG, J.P. Purification and Characterization of Human-Brain Aldose Reductase. *Medizinisch-chemisches Institut der Universität, Bern. Eur. J. Biochem.* 127, 279-284, 1982.

WILSON, D. B. Cellulases and biofuels. *Curr. Opin. Biotechnol.* 20: 295-299, 2009.

WILLIS, J. D; OPPERT, C; JURAT-FUENTES, J. L. Methods for discovery and characterization of cellulolytic enzymes from insects. *Insect Science*, 17 (3), 184–198, 2010. Doi:10.1111/j.1744-7917.2010.01322.x,

WILLIS, J. D; OPPERT, B; OPPERT, C; KLINGEMAN, W. E; JURAT-FUENTES, J. L. Identification, cloning, and expression of a GHF9 cellulase from *Tribolium castaneum* (Coleoptera: Tenebrionidae). *Journal of Insect Physiology*, 57(2), 300–6, 2011. Doi:10.1016/j.jinsphys.2010.11.019

WU, S; LETCHWORTH, G. J. High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol. *BioTechniques*, 36 (1), 152–154, 2004. Disponível em, <http://doi.org/10.2144/3601A0152>

YAMAMOTO, K; WILSON, D.K. Identification, characterization, and crystal structure of an aldo–keto reductase (AKR2E4) from the silkworm *Bombyx mori*. *Science Direct, Archives of Biochemistry and Biophysics* 538, 156–163, 2013. Disponível em, <http://dx.doi.org/10.1016/j.abb.2013.08.018>

YAN, J. *et al.* Functional and structural analysis of *Pichia pastoris*-expressed *Aspergillus niger* 1,4- $\beta$ -endoglucanase. *Biochem. Biophys. Res. Commun.* 475 (1), 8–12, 2016, Disponível em, <https://doi.org/10.1016/j.bbrc.2016.05.012>

YANG, S. *et al.* Enhanced production of recombinant secretory proteins in *Pichia pastoris* by optimizing Kex2 P1' site. *PLOS ONE* 8. 2013.

ZHANG, Y. H. P; HIMMEL, M. E; MIELENZ, J. R. Outlook for cellulase improvement: Screening and selection strategies. *Biotechnology Advances*, New York, v. 24, p. 452-481, 2006.

ZHAO, C; DOUCET, D; MITTAPALLI, O. Characterization of horizontally transferred  $\beta$ -fructofuranosidase (ScrB) genes in *Agrius planipennis.*, *Insect Molecular Biology*. 2014. Doi:10.1111/imb.12127

ZHOU, X; SMITH, J.A; OI, F.M; KOEHLER, P.G; BENNETT, G.W; SCHARF, M.E. Correlation of cellulase gene expression and cellulolytic activity throughout the gut of the termite *Reticulitermes flavipes*. *Gene*. 395, 29–39, 2007.

ZHU, B; LOU, M.M; XIE, G.L; ZHANG, G.Q; ZHOU, X.P; LI, B; JIN, G.L. Horizontal gene transfer in silkworm, *Bombyx mori*. *BMC Genomics* 12:248, 2011.