



**THABATA MARIA ALVAREZ**

**“DESENVOLVIMENTO DE UMA BIBLIOTECA DE  
ENZIMAS A PARTIR DE METAGENOMA DE SOLO”**

***“LIBRARY GENERATION FOR BIOMASS CONVERSION  
ENZYMES FROM SOIL METAGENOME”***

**CAMPINAS  
2013**





UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE BIOLOGIA

**THABATA MARIA ALVAREZ**

**“DESENVOLVIMENTO DE UMA BIBLIOTECA DE ENZIMAS A  
PARTIR DE METAGENOMA DE SOLO”**

**Orientador: Prof. Dr. Fabio Marcio Squina**

***“LIBRARY GENERATION FOR BIOMASS CONVERSION ENZYMES  
FROM SOIL METAGENOME”***

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Biologia Funcional e Molecular do Instituto de Biologia da Universidade Estadual de Campinas para obtenção do título de Doutora em Biologia Funcional e Molecular, na área de Bioquímica.

*Doctorate thesis presented to the Functional and Molecular Biology Postgraduation Programme of the Biology Institute of the University of Campinas to obtain the ph.D. grade in Functional and Molecular Biology, in Biochemistry.*

Este exemplar corresponde à redação  
final da TESE defendida pela candidata

*Thabata Maria Alvarez*

e aprovada pela Comissão Examinadora.

CAMPINAS  
2013

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

AL86d Alvarez, Thabata Maria, 1986-  
Desenvolvimento de uma biblioteca de enzimas a partir de metagenoma de solo / Thabata Maria Alvarez. – Campinas, SP : [s.n.], 2013.

Orientador: Fabio Marcio Squina.  
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Metagenômica. 2. Glicosídeo hidrolases. I. Squina, Fabio Marcio. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Library generation for biomass conversion enzymes from soil metagenome

**Palavras-chave em inglês:**

Metagenomics

Glycoside hydrolases

**Área de concentração:** Bioquímica

**Titulação:** Doutora em Biologia Funcional e Molecular

**Banca examinadora:**

Fabio Marcio Squina [Orientador]

Carmen Verissima Ferreira Halder

Marie Anne Van Sluys

Igor Polikarpov

Hiroshi Aoyama

**Data de defesa:** 28-08-2013

**Programa de Pós-Graduação:** Biologia Funcional e Molecular

Campinas, 28 de agosto de 2013

**BANCA EXAMINADORA**

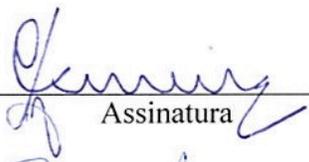
Dr. Fabio Marcio Squina (Orientador)



---

Assinatura

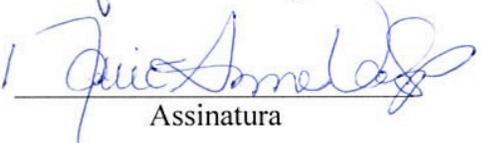
Profa. Dra. Carmen Verissima Ferreira Halder



---

Assinatura

Profa. Dra. Marie Anne Van Sluys



---

Assinatura

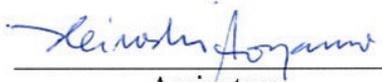
Prof. Dr. Igor Polikarpov



---

Assinatura

Prof. Dr. Hiroshi Aoyama



---

Assinatura

Dr. Celso Eduardo Benedetti

---

Assinatura

Dr. Itamar Soares de Melo

---

Assinatura

Prof. Dr. Marcelo Lancellotti

---

Assinatura

## RESUMO

Devido à necessidade do desenvolvimento de fontes de energias renováveis é de grande interesse a descoberta de novas enzimas envolvidas na desconstrução da parede celular vegetal para a produção de biocombustíveis. A metagenômica é uma poderosa ferramenta para a descoberta de novos genes em comunidades microbianas que não são passíveis de cultivo pelas técnicas tradicionais.

Neste contexto, o objetivo desta tese foi o desenvolvimento de estratégias metagenômicas para prospecção de novas enzimas atuantes na degradação da biomassa vegetal no metagenoma de solo de canavial bem como a caracterização funcional das mesmas.

A biblioteca metagenômica construída com DNA extraído de um consórcio microbiano especializado na degradação de bagaço de cana-de-açúcar explodido a vapor e deslignificado foi empregada nos experimentos de triagem funcional de alto desempenho. Como resultado, foram identificados três clones positivos com atividade celulolítica e dois clones com atividade xilanolítica. A análise dos insertos de cada um dos clones resultou na localização de ORFs cujas sequências de aminoácidos apresentaram identidade com domínios conservados de glicosil hidrolases da família 5 (celulases E-1 e E-2), família 6 (celulase E-3), família 10 (xilanase X-1) e família 16 (glicosil hidrolase X-2). A celulase E-1 apresentou em sua estrutura além do domínio catalítico, E-1 Cat, um domínio de ligação a carboidratos, denominado E-1 CBM, que não apresentou identidade de sequência com domínios conservados conhecidos. A análise funcional do E-1 CBM revelou tratar-se de um CBM específico para cadeias de glucano com grau de polimerização mínimo de cinco unidades de glicose.

Ensaio de atividade enzimática em diferentes substratos mostraram que E-1 Cat atuou especificamente na hidrólise das ligações glicosídicas do tipo  $\beta(1,4)$  entre resíduos de glicose. Os maiores valores de atividade enzimática foram obtidos em pH 7,0 e temperatura de 50°C. Os parâmetros cinéticos calculados em CMC foram  $K_m$  igual a  $6,05 \pm 0,37$  mg/mL,  $V_{max}$  de  $42,51 \pm 1,2$   $\mu\text{mol}/\text{min}/\text{mg}$  e eficiência catalítica  $k_{cat}/K_m$  de  $4,06$  mL/mg/s. A enzima apresentou termoestabilidade a 40°C por cinco horas. A atividade enzimática de E-1 Cat em celulose cristalina e bagaço de cana-de-açúcar

explodido a vapor resultou na liberação de açúcares solúveis, evidenciando sua potencial aplicação em processos de conversão da biomassa vegetal.

Ensaio de atividade em diferentes substratos mostraram que X-1 apresentou maior atividade enzimática em xilana não ramificada, nas condições de pH e temperatura de 6,0 e 45°C, respectivamente. Os parâmetros cinéticos calculados utilizando como substrato xilana de madeira de faia foram  $K_m$  de  $2,18 \pm 0,13$  mg/mL,  $V_{max}$  de  $1.435 \pm 30,4$   $\mu$ mol/min/mg e  $k_{cat}/K_m$  de 496,32 mL/mg/s. Em relação à termoestabilidade, a enzima se manteve estável a 40°C e 50°C por seis horas. A hidrólise de substratos complexos com X-1 resultou na liberação de xilo-oligossacarídeos, xilobiose e xilose, que são compostos que apresentam potencial aplicação nas indústrias alimentícias e de biocombustíveis.

Os resultados obtidos neste estudo validaram a abordagem metagenômica desenvolvida para a descoberta de novos genes codificantes para glicosil hidrolases. Além disso, a estratégia descrita nesta tese pode ser estendida para a descoberta de uma miríade de bioprodutos de interesse biotecnológico.

## ABSTRACT

Due to the necessity of development of renewable sources of energy, it is of great interest the discovery of novel enzymes involved in plant cell wall deconstruction for biofuels production. Metagenomics is a powerful tool for the discovery of novel genes in microbial communities that are not liable to cultivation by traditional techniques.

In this context, the aim of this thesis was the development of metagenomic strategies for prospection of novel enzymes involved in plant biomass degradation in sugarcane field soil metagenome and functional characterization of the identified enzymes.

The metagenomic library constructed with DNA extracted from a microbial consortium specialized in degradation of steam exploded delignified sugarcane bagasse was used in the experiments of high-performance functional screening. As a result, we identified three positive clones with cellulolytic activity and two clones with xylanolytic activity. The analysis of the inserts from each clone resulted in the location of ORFs whose amino acid sequences showed identity to conserved domains of glycoside hydrolase family 5 (cellulases E-1 and E-2), family 6 (cellulase E-3), family 10 (xylanase X-1) and family 16 (glycoside hydrolase X-2). Cellulase E-1 exhibited in addition to the catalytic domain, E-1 Cat, a carbohydrate binding module, called E-1 CBM, which showed no sequence identity with known conserved domains. Functional analysis of E-1 CBM showed that it is a CBM specific for glucan chains with a degree of polymerization of at least five units of glucose.

Assays with a set of different substrates revealed that E-1 Cat hydrolyzed specifically  $\beta(1,4)$  glycoside bonds between glucose residues. The highest value of enzymatic activity was obtained at pH 7.0 and temperature of 50°C. The kinetic parameters  $K_m$ ,  $V_{max}$  and catalytic efficiency  $k_{cat}/K_m$  calculated using CMC were  $6.05 \pm 0.37$  mg/mL,  $42.51 \pm 1.2$   $\mu$ mol/min/mg and 4.06 mL/mg/s, respectively. The enzyme showed thermal stability at 40°C for five hours. The enzymatic activity of E-1 Cat in crystalline cellulose and steam exploded sugarcane bagasse resulted in the release of soluble sugars, demonstrating its potential application in processes of biomass conversion.

The xylanase X-1 showed higher enzyme activity in debranched xylan, in reactions conducted in pH 6.0 and temperature of 45°C. The kinetic parameters  $K_m$ ,  $V_{max}$  and catalytic efficiency  $k_{cat}/K_m$  calculated using beechwood xylan were  $2.18 \pm 0.13$  mg/mL,  $1,435 \pm 30.4$   $\mu$ mol/min/mg and 496.32 mL/mg/s, respectively. In relation to thermal stability, the enzyme was stable at 40°C and 50°C for six hours. The hydrolysis of complex substrates resulted in the release of xylo-oligosaccharides, xylobiose and xylose, which are compounds that have potential application in food and biofuels industries.

The results of this study validated the metagenomic approach developed for the discovery of novel genes coding for glycoside hydrolases. Moreover, the strategy described in this work can be extended to the discovery of a myriad of byproducts of biotechnological interest.

## SUMÁRIO

<b>INTRODUÇÃO GERAL.....</b>	<b>1</b>
<b>CAPÍTULO 1 - REVISÃO BIBLIOGRÁFICA .....</b>	<b>5</b>
1.1. O Uso da Biomassa como Fonte de Energia Renovável .....	5
1.2. Composição da Parede Celular Vegetal.....	7
1.3. Desconstrução da Parede Celular Vegetal .....	11
1.3.1. Desconstrução enzimática da parede celular vegetal.....	12
1.3.2. Domínios de ligação a carboidratos .....	16
1.4. Metagenômica.....	19
1.4.1. Metagenômica de solo .....	20
1.4.2. Prospecção de novos genes em metagenomas .....	21
1.4.2.1. Prospecção de glicosil hidrolases .....	24
<b>CAPÍTULO 2 – OBJETIVOS .....</b>	<b>27</b>
2.1 Objetivos Gerais.....	27
2.2 Objetivos Específicos .....	27
<b>CAPÍTULO 3 – DESENVOLVIMENTO DE ESTRATÉGIAS PARA A IDENTIFICAÇÃO DE NOVAS ENZIMAS ATIVAS NA DEGRADAÇÃO DA PAREDE CELULAR VEGETAL.....</b>	<b>29</b>
3.1. Introdução .....	29
3.2. Objetivos do Capítulo .....	29
3.3. Materiais e Métodos .....	29
3.3.1. Construção das bibliotecas metagenômicas .....	29
3.3.1.1. Coleta das amostras de solo.....	29
3.3.1.2. Consórcio microbiano especializado na degradação de bagaço de cana-de-açúcar.....	30
3.3.1.3. Extração do DNA metagenômico.....	30
3.3.1.4. Quantificação de DNA.....	31
3.3.1.5. Digestão do DNA metagenômico.....	31
3.3.1.6. Purificação do DNA metagenômico digerido .....	32
3.3.1.7. Preparo do vetor pUC19 .....	32
3.3.1.8. Reação de ligação .....	33
3.3.1.9. Preparo de células eletrocompetentes .....	34
3.3.1.10. Transformação bacteriana por eletroporação.....	34
3.3.1.11. Maxipreparação dos plasmídeos recombinantes .....	35
3.3.1.12. Análise do tamanho médio dos insertos .....	35
3.3.1.13. Coleta e estocagem dos clones.....	36
3.3.2. Prospecção de novos genes através de abordagens dirigidas por homologia de sequência nucleotídica.....	36
3.3.2.1. Reação de PCR para amplificação de fragmentos gênicos de xilanases com oligonucleotídeos iniciadores degenerados .....	36
3.3.2.2. Clonagem dos insertos amplificados com oligonucleotídeos iniciadores degenerados.....	38

3.3.2.3.	Confirmação e sequenciamento dos fragmentos clonados em vetor pGEM® T-Easy .....	39
3.3.2.4.	Identificação de genes no metagenoma de solo de canavial .....	40
3.3.2.5.	Reação de PCR para amplificação dos genes alvos .....	41
3.3.2.6.	Clonagem dos fragmentos no vetor pJET1.2/blunt .....	42
3.3.3.	Prospecção de novos genes através de abordagem dirigida por triagem funcional.....	44
3.3.3.1.	Triagem funcional da biblioteca metagenômica .....	44
3.3.3.2.	Coleta dos clones positivos identificados.....	44
3.3.3.3.	Confirmação da atividade enzimática dos clones positivos.....	45
3.3.3.4.	Avaliação da atividade enzimática por eletroforese capilar .....	46
3.3.3.5.	Mapeamento das ORFs nos clones positivos identificados .....	46
3.4	Resultados e Discussão .....	49
3.4.1.	Construção das bibliotecas metagenômicas.....	49
3.4.2.	Prospecção de novos genes através da abordagem dirigida por homologia de sequência nucleotídica.....	54
3.4.2.1.	Reação de PCR para amplificação de fragmentos gênicos de xilanasases com oligonucleotídeos iniciadores degenerados .....	54
3.4.2.2.	Amplificação de genes alvos .....	59
3.4.3.	Prospecção de novos genes através da abordagem dirigida por triagem funcional.....	61
3.4.4.	Mapeamento dos insertos dos clones positivos identificados na triagem funcional.....	68
3.4.4.1.	Clone E-1 .....	68
3.4.4.2.	Clone E-2.....	71
3.4.4.3.	Clone E-3.....	73
3.4.4.4.	Clone X-1 .....	75
3.4.4.5.	Clone X-2.....	77
3.5.	Conclusões .....	80

## **CAPÍTULO 4- CARACTERIZAÇÃO BIOQUÍMICA DA CELULASE E-1 E DA XILANASE X-1 E AVALIAÇÃO DA APLICAÇÃO EM PROCESSOS BIOTECNOLÓGICOS .....**

4.1.	Introdução .....	83
4.2.	Objetivos do Capítulo .....	83
4.3.	Materiais e Métodos .....	83
4.3.1.	Desenho dos oligonucleotídeos iniciadores para amplificação dos genes e clonagem em vetor de expressão .....	83
4.3.2.	Amplificação dos genes por PCR .....	85
4.3.3.	Purificação dos fragmentos amplificados.....	85
4.3.4.	Digestão dos fragmentos de PCR .....	85
4.3.5.	Preparo do vetor pET-28a .....	86
4.3.6.	Reação de ligação.....	87
4.3.7.	Confirmação da clonagem em pET-28a.....	87
4.3.8.	Cepas utilizadas .....	88
4.3.9.	Preparo de células cálcio competentes.....	89
4.3.10.	Transformação por choque térmico .....	89

4.3.11.	Testes de expressão.....	89
4.3.12.	Purificação proteica.....	91
4.3.13.	Zimografia.....	92
4.3.14.	Ensaio de ligação a substratos insolúveis.....	92
4.3.15.	Preparo da amostra para análise em LC-MS/MS .....	93
4.3.16.	Caracterização bioquímica .....	95
4.3.16.1.	Quantificação de proteínas .....	95
4.3.16.2.	Determinação da atividade enzimática .....	96
4.3.16.3.	Determinação dos valores de pH e temperatura ótimos.....	97
4.3.16.4.	Determinação da energia de ativação .....	97
4.3.16.5.	Determinação de parâmetros cinéticos .....	98
4.3.16.6.	Estabilidade térmica .....	98
4.3.16.7.	Efeitos de íons, sais e agentes quelantes .....	99
4.3.16.8.	Avaliação da atividade enzimática em substratos complexos ....	99
4.3.16.9.	Eletroforese capilar.....	100
4.4.	Resultados e Discussão .....	100
4.4.1.	Celulase E-1.....	100
4.4.1.1.	Avaliação da capacidade da celulase E-1 em se ligar a substratos insolúveis .....	103
4.4.1.2.	E1-Cat.....	107
4.4.1.3.	Caracterização bioquímica E1-Cat .....	109
4.4.1.4.	Xilanase X-1 .....	115
4.4.1.5.	Caracterização bioquímica X-1 .....	117
4.5.	Conclusões .....	126

## **CAPÍTULO 5 – ESTUDOS FUNCIONAIS DO DOMÍNIO DE LIGAÇÃO A CARBOIDRATOS PRESENTE NA CELULASE E-1 (E1-CBM)..... 129**

5.1.	Introdução.....	129
5.2.	Objetivos .....	129
5.3.	Materiais e Métodos .....	130
5.3.1.	Clonagem em pET-28a e pET-41a .....	130
5.3.2.	Reação de PCR para amplificação do gene alvo .....	131
5.3.3.	Preparo do vetor pET-41a .....	132
5.3.4.	Reação de ligação.....	132
5.3.5.	Transformação por choque térmico .....	132
5.3.6.	Confirmação da clonagem em pET-41a.....	132
5.3.7.	Testes de expressão.....	133
5.3.8.	Purificação proteica.....	134
5.3.9.	Calorimetria de titulação isotérmica .....	135
5.3.10.	Remoção da cauda de GST.....	136
5.3.11.	Preparo da amostra para os testes de cristalização .....	137
5.4.	Resultados e Discussão .....	138
5.4.1.	Clonagem em pET-28a .....	138
5.4.2.	Clonagem em pET-41a .....	139
5.4.3.	Testes de expressão com o vetor pET-28a .....	140
5.4.4.	Teste de expressão com o vetor pET-41a .....	141
5.4.5.	Calorimetria de titulação isotérmica.....	142

5.4.6. Remoção da cauda de GST.....	145
5.4.7. Testes de cristalização.....	147
5.5. Conclusões .....	147
CONCLUSÕES GERAIS DA TESE.....	149
REFERÊNCIAS BIBLIOGRÁFICAS.....	151
ANEXO 1 – ARTIGO PUBLICADO (1ª AUTORA) NA REVISTA <i>PLoS ONE</i> EM JULHO DE 2013.....	173
ANEXO 2 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA <i>BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS</i> EM NOVEMBRO DE 2010.....	174
ANEXO 3 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA <i>BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS</i> EM FEVEREIRO DE 2011.....	175
ANEXO 4 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA <i>BIOTECHNOLOGY FOR BIOFUELS</i> EM NOVEMBRO DE 2011.....	176
ANEXO 5 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA <i>JOURNAL OF STRUCTURAL BIOLOGY</i> EM DEZEMBRO DE 2011.....	177
ANEXO 6 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA <i>BIORESOURCE TECHNOLOGY</i> EM MAIO DE 2012.....	178
ANEXO 7 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA <i>PLoS ONE</i> EM DEZEMBRO DE 2012.....	179
ANEXO 8 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA <i>LETTERS IN APPLIED MICROBIOLOGY</i> EM ABRIL DE 2013.....	180
ANEXO 9 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA <i>BIOENERGY RESEARCH</i> EM MAIO DE 2013.....	181
ANEXO 10 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA <i>BIOCHIMICA et BIOPHYSICA ACTA</i> EM FEVEREIRO DE 2013.....	182
ANEXO 11 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA <i>MOLECULAR BIOTECHNOLOGY</i> EM JUNHO DE 2013.....	183
ANEXO 12 - DECLARAÇÃO DE BIOÉTICA E BIOSSEGURANÇA.....	185
ANEXO 13 - COMPROVANTE DE AUTORIZAÇÃO DE BIOÉTICA E BIOSSEGURANÇA.....	187

## **AGRADECIMENTOS**

*Primeiramente a Deus por sempre iluminar meu caminho;*

*Ao meu orientador, Dr. Fabio Squina, pela oportunidade de desenvolver esse trabalho, e que sempre me motivou com sua participação ativa e incentivo constante;*

*Aos meus pais, Sergio e Rosana, meu irmão João Paulo e o nosso querido Tobias, por todo amor e por sempre acreditarem em mim, não há palavras que expressem o amor e gratidão que sinto por vocês;*

*Ao Péricles, pois o seu amor e companheirismo foram fundamentais para eu chegar até aqui;*

*Aos meus avôs, tios, tias, primos e primas, que sempre me apoiaram e torceram por mim;*

*A todos meus amigos de Marília e Curitiba, porque com cada um de vocês eu aprendi a ser uma pessoa melhor;*

*Aos meus amigos do CTBE, em especial ao João, que me acompanha nessa caminhada desde o começo e com quem aprendi tantas coisas, obrigada pela sua amizade sincera. À Ana, por ser uma amiga tão verdadeira e presente em minha vida. À Rebeca, por todas as histórias e risadas com certeza inesquecíveis. À Isabela, por ter me ajudado sempre com muita paciência e disponibilidade e pelos momentos de descontração. À Lívia, Dani, Fernanda B., Fernanda M, Zaira, Carla, Camilinha, Aline, Thuany, Thiago, Marcelo, Damásio, Júnio, Deise, Cleiton, Diogo, Beto, Evandro, Cris, Thamy, obrigada pelos ótimos momentos;*

*A todos que participaram e ajudaram na execução deste trabalho, em especial ao Douglas e Rodrigo, pelo suporte fundamental em todas as etapas e à Rosana e Fer B, pela disponibilidade em ajudarem na execução*

*de parte deste trabalho;*

*Aos pesquisadores Rolf Prade, Adriana Franco Paes Leme e Roberto Ruller pelas conversas e sugestões que agregaram valor ao trabalho;*

*Ao Dr. Harry Gilbert, por ter me recebido tão bem em seu laboratório e pela ótima experiência proporcionada;*

*Aos membros titulares da banca: Profa. Dra. Marie Anne Van Sluys, Profa. Dra. Carmen Verissima Ferreira Halder, ao Prof. Dr. Igor Polikarpov e ao Prof. Dr. Hiroshi Aoyama pelas valiosas críticas e sugestões;*

*Ao CNPEM e ao CTBE pela infraestrutura e suporte técnico disponibilizado para a realização deste trabalho;*

*À Universidade Estadual de Campinas, através do Instituto de Biologia e a pós-graduação em Biologia Funcional e Molecular;*

*À FAPESP pelo apoio financeiro para realização deste trabalho.*

## LISTA DE ABREVIATURAS E SIGLAS

- ADNS – Ácido 3,5-dinitrosalicílico
- Amp – Ampicilina
- APTS - Ácido 1,3,6-trissulfônico- 8-aminopireno
- BCB - Biblioteca metagenômica construída com o DNA metagenômico extraído do consórcio microbiano degradador de bagaço de cana-de-açúcar
- BEX – Bagaço de cana-de-açúcar explodido
- BG –  $\beta$ -glucano
- BLAST - *Basic Local Alignment Sequence Tool* (Ferramenta para alinhamento básico local)
- BP – Biblioteca metagenômica construída com DNA metagenômico extraído diretamente do solo de canavial de Paulínia
- BSA – Albumina de soro bovino
- CAZy - *Carbohydrate-Active EnZymes* (Enzimas ativas em carboidratos)
- CB – Consórcio microbiano degradador de bagaço de cana – de – açúcar
- CBM - Domínio de ligação a carboidratos
- CDD - *Conserved Domain Database* (Banco de dados de domínios conservados)
- cm – Centímetro
- CMC – Carboximetilcelulose
- C2 – Celobiose
- C3 – Celotriose
- C4 – Celotetrose
- C5 – Celopentose
- C6 - Celohexose
- DMSO – Dimetilsulfóxido
- DNA - Ácido desoxirribonucleico
- DNTPs - Desoxirribonucleotídeos trifosfatos
- $E_a$  – Energia de ativação
- EDTA - Ácido etilenodiamino tetra-acético
- g – Grama

Gb – Gigabases  
GH – Glicosil hidrolase  
GST – Glutathione – S – Transferase  
ITC - Calorimetria de titulação isotérmica  
IPTG - Isopropil  $\beta$ -D-1-tiogalactopiranoside  
K<sub>a</sub> - Constante de associação  
K<sub>D</sub> - Constante de dissociação  
kb- Mil pares de base  
kDa – Mil Daltons  
IMAC - Cromatografia de afinidade com metal imobilizado  
L – Litro  
LB - Luria-Bertani  
m - Massa  
M – Molar (mol/L)  
Mb – Mega base  
Mini-prep – Mini-preparação de plasmídeos  
MCB – Metagenoma extraído do consórcio degradador de bagaço  
mL – Mililitro  
MSC - Metagenoma extraído do solo de canavial de São Carlos  
MSP - Metagenoma extraído do solo de canavial de Paulínia  
N - Estequiometria da reação  
NCBI - National Center for Biotechnology Information (Centro Nacional de  
Informação Biotecnológica)  
nm – Nanometro  
ORF - *Open reading frame* (Fase aberta de leitura)  
PCR - Reação em cadeia da polimerase  
PDB – *Protein Data Bank*  
PMSF – Fluoreto de fenilmetano sulfonila  
RNA - Ácido ribonucleico  
rRNA - RNA ribossomal  
rpm – Rotações por minuto

SDS – Dodecil sulfato de sódio

SDS – PAGE – Eletroforese em gel de poliacríamida – SDS

TA – Temperatura ambiente

TAE – Tris-acetato-EDTA

U – Unidade de atividade enzimática (quantidade de enzima capaz de formar 1  $\mu\text{mol}$  de produto por minuto nas estabelecidas condições de ensaio)

U\* - Unidade de atividade enzimática (quantidade de enzima capaz de formar 1  $\text{nmol}$  de produto por minuto nas estabelecidas condições de ensaio)

UV – Ultravioleta

v- Volume

V- Volt

X-gal- 5-bromo-4-cloro-3-indolil-  $\beta$ -D-galactopiranosídeo

XB – Xilana *beechwood* (madeira de faia)

XG – Xiloglucano

XXXG - Heptassacarídeo  $\text{X}_3\text{Glc}_4$

X- Xilose

X2 – Xilobiose

X3 – Xilotriose

X4 – Xilotetrose

X5 – Xilopentose

X6 – Xilohexose

x g – Força centrífuga relativa

$\mu\text{A}$  – Microampére

## INTRODUÇÃO GERAL

A dependência em relação aos recursos naturais não renováveis para o suprimento da demanda energética mundial constitui o principal estímulo para o desenvolvimento de tecnologias que permitam o aproveitamento de recursos naturais renováveis para geração de energia (TENGERDY & SZAKACS, 2003). A crise do petróleo em 1970 incentivou as primeiras pesquisas na área de aproveitamento de recursos renováveis para a produção de biocombustíveis. Porém, nos anos seguintes, a queda no preço do petróleo voltou a estimular o uso dos combustíveis fósseis (LOVINS, 2004). As estimativas do aumento da demanda energética em 50% até 2025, principalmente devido ao rápido crescimento de várias nações em desenvolvimento, impulsionaram novamente as buscas por fontes de energia renováveis (RAGAUSKAS et al., 2006).

A biomassa é considerada a principal alternativa para a substituição dos combustíveis fósseis (ÖZÇIMEN & KARAOSMANOĞLU, 2004, JEFFERSON, 2006), por três motivos principais. O primeiro deles é o fato de tratar-se de uma fonte renovável, seguido pela característica de apresentar balanço de emissão de gás carbônico nulo, pois o dióxido de carbono liberado a partir de sua queima foi previamente absorvido pelas plantas em seu desenvolvimento. Por último, o uso da biomassa como fonte de energia apresenta potencial econômico, tendo em vista a previsão do aumento no preço dos combustíveis fósseis no futuro (CADENAS & CABEZUDO, 1998).

O termo biocombustível se refere aos combustíveis líquidos ou gasosos empregados no setor de transporte e que são predominantemente produzidos a partir de biomassa (DEMIRBAS, 2007). Os biocombustíveis apresentam como pontos fortes o fato de derivarem de fontes ilimitadas e serem biodegradáveis, resultando em menores impactos ambientais (DEMIRBAS, 2007).

O bioetanol é um biocombustível derivado de fontes renováveis (como trigo, milho e cana-de-açúcar), produzido pelo processo de fermentação anaeróbica, em que açúcares são convertidos em etanol pela ação de micro-organismos, principalmente as

leveduras. Em geral a biomassa vegetal, incluindo diversos resíduos agroindustriais, é rica em polissacarídeos dos quais podem ser obtidos açúcares fermentescíveis (DEMIRBAS, 2007).

O principal desafio para produção de biocombustíveis em larga escala é o desenvolvimento de tecnologias que superem a recalcitrância da parede celular vegetal à degradação para obtenção de açúcares (FARINAS, 2011).

Atualmente, os processos químicos empregados na degradação da biomassa vegetal envolvem tratamentos ácidos e térmicos, que apresentam alto custo e baixa eficiência (MOSIER et al., 2005). Além disso, muitos desses tratamentos produzem compostos furfurais, que atuam como inibidores nas etapas posteriores de fermentação (RUBIN, 2008).

A hidrólise enzimática da parede celular apresenta vantagens em relação às rotas químicas, principalmente no que diz respeito a não liberação de compostos tóxicos que podem ser inibitórios para as etapas posteriores de fermentação (NAIK et al., 2010). Porém o alto custo e baixo rendimento associados aos processos enzimáticos são os principais entraves para que sejam economicamente viáveis (XIA et al., 2013).

Neste sentido, a metagenômica, que consiste no direto isolamento de ácidos nucleicos de comunidades microbianas sem a necessidade de cultivo e isolamento, auxilia na identificação de novas e potenciais enzimas para serem empregadas no processo de desconstrução da parede celular vegetal (XIA et al., 2013, VERMA et al., 2013, ILMBERGER & STREIT, 2010, LI et al., 2009).

Nesse contexto, o presente trabalho foi desenvolvido. A tese se encontra dividida em capítulos, visando dar ênfase aos objetivos propostos e resultados obtidos.

Abaixo, se encontra uma breve descrição de cada um dos capítulos:

- Capítulo 1 – Revisão bibliográfica apresentando as informações que contextualizam e justificam a realização do trabalho;
- Capítulo 2 - Objetivos gerais e específicos que motivaram cada etapa do trabalho;
- Capítulo 3 – Desenvolvimento de metodologias para prospecção de novos genes de glicosil hidrolases por abordagens metagenômicas;

- Capítulo 4 – Clonagem, expressão e caracterização bioquímica de duas enzimas isoladas pela triagem funcional da biblioteca metagenômica;
- Capítulo 5 – Investigação e estudos funcionais de um domínio de ligação a carboidratos identificado em uma celulase de origem metagenômica.



## **CAPÍTULO 1 - REVISÃO BIBLIOGRÁFICA**

### **1.1. O Uso da Biomassa como Fonte de Energia Renovável**

De acordo com o Centro Nacional de Referência em Biomassa (CENBIO) (disponível em: <http://cenbio.iee.usp.br/>) e do ponto de vista energético, a biomassa é definida como o conjunto de todos os recursos renováveis provenientes de matéria orgânica, tanto de origem animal como vegetal, que podem ser empregados na geração de energia. Um tipo em particular de biomassa é a lignocelulósica, que corresponde a mais de 60% de toda a biomassa vegetal produzida no mundo (TENGERDY & SZAKACS, 2003). A biomassa vegetal armazena energia solar captada pelo processo de fotossíntese na forma de polímeros na parede celular vegetal. Esta energia pode ser disponibilizada através da queima ou processos de bioconversão (TENGERDY & SZAKACS, 2003).

As projeções para o fim das reservas naturais da principal fonte energética utilizada atualmente, o petróleo, e o conseqüente incremento da demanda por recursos energéticos sustentáveis demonstram a importância do desenvolvimento de fontes alternativas de energias denominadas limpas e renováveis (JOÃO, 2011).

Em relação ao uso de energia renovável, o Brasil apresenta 90% de sua matriz energética baseada na hidroeletricidade (BRASIL, 2009), além do domínio da produção de bioetanol a partir da cana-de-açúcar, por meio do processo de fermentação alcoólica da sacarose, denominado de etanol de primeira geração (JOÃO, 2011). O Brasil é o segundo maior produtor de etanol do mundo, atrás apenas dos Estados Unidos, que utiliza o milho como matéria prima para produção do etanol (SIMS et al., 2010).

No Brasil, a produção de etanol se destacou na década de 1970 incentivada pela criação do Programa Nacional do Álcool – Proálcool. O Proálcool foi criado em virtude da crise energética mundial, com o objetivo de diminuir a dependência em relação ao petróleo (JOÃO, 2011). Porém, até o ano 2000 o etanol necessitava de subsídios do governo para a inserção no mercado nacional (CERQUEIRA LEITE, 2009). O etanol se tornou um combustível líquido altamente competitivo a partir de 2003, com o

desenvolvimento dos veículos *flex fuel*, também chamados de veículos de combustível duplo (CERQUEIRA LEITE, 2009). Em 2012, aproximadamente 87% dos veículos leves licenciados no país possuíam a tecnologia *flex fuel* e apenas 7,5% eram movidos exclusivamente à gasolina (ANFAVEA, 2013).

No Brasil, a tendência é o crescimento contínuo da produção de etanol, devido aos diversos investimentos feitos na área, como a construção de novas destilarias e o emprego de estratégias que visam elevar a produtividade através do aproveitamento dos resíduos gerados na produção do etanol de primeira geração (MARTINELLI & FILOSO, 2008).

Nas usinas esses resíduos correspondem principalmente ao bagaço de cana – de - açúcar, material resultante da extração do caldo após esmagamento da cana nas moendas (PANDEY et al., 2000), e pela palha, que deixada sobre solo após a colheita sofre decomposição por micro-organismos (CORTEZ et al., 2008). Anualmente no Brasil, a produção de bagaço de cana-de-açúcar é de aproximadamente 186 milhões de toneladas e aproximadamente 92% desse total é queimado nas usinas para geração de energia (SOCCOL et al., 2010).

O bagaço apresenta significativa heterogeneidade morfológica e consiste em feixes de fibras e outros elementos estruturais como parênquima e células epiteliais (SANJUAN et al., 2001). Em relação à sua composição, apresenta de 19-24% de lignina, 27-32% de hemicelulose, 32-44% de celulose e 4,5-9,0% de cinzas (JACOBSEN & WYMAN, 2002). A celulose e a hemicelulose correspondem às frações de polissacarídeos que podem ser extraídas, fermentadas e transformadas em etanol, dando origem ao chamado etanol de segunda geração (CARDONA et al., 2010).

No Brasil, a produção de etanol a partir do bagaço é bastante favorecida, pois se trata de uma matéria prima disponível em grande quantidade e com baixo valor agregado (CARDONA & SANCHEZ, 2007). Além disso, as plantas de etanol de segunda geração podem ser construídas no mesmo local das usinas de primeira geração, reduzindo assim os custos com implantação, suprimento energético e transporte de matérias primas (SOCCOL et al., 2010). Estima-se que o aproveitamento da palha e do bagaço de cana-de-açúcar resultaria em um aumento de cerca de 40% da produtividade de etanol sem a necessidade do aumento da área cultivada

(MACEDO, 2005).

Apesar do otimismo em relação ao etanol de segunda geração, existem dificuldades técnicas para que o processo de desconstrução da parede celular vegetal seja viável economicamente (RUBIN, 2008). A principal dificuldade do processo está relacionada à recalcitrância da parede celular vegetal, devido à alta associação entre os três principais componentes: celulose, hemicelulose e lignina (SOCCOL et al., 2010). A organização da parede celular vegetal é descrita em detalhes no tópico abaixo.

## **1.2. Composição da Parede Celular Vegetal**

Nas células vegetais, a parede celular tem como função envolver o conteúdo celular, evitar o estresse mecânico, manter a morfologia da célula e proteger contra o ataque de patógenos. Trata-se de uma estrutura flexível para suportar forças de cisalhamento e permeável, permitindo a troca de pequenas moléculas entre o meio intra e extracelular (QUIROZ-CASTAÑEDA & FOLCH-MALLOL, 2011). A parede celular vegetal é dividida em três camadas, denominadas lamela média, parede primária e parede secundária (CARPITA & GIBEAUT, 1993, COSGROVE, 2001).

A lamela média é uma camada rica em pectina, que compõe a interface entre duas células vegetais adjacentes e é responsável por manter as células unidas (COSGROVE, 2005).

A parede primária geralmente é a camada mais fina e flexível, formada na fase de crescimento celular. É composta predominantemente por polissacarídeos e em menores quantidades por glicoproteínas, ésteres fenólicos, compostos minerais e enzimas. Os principais polissacarídeos da parede celular primária são a celulose, hemicelulose e pectina (COSGROVE, 2005). As fibras de celulose são envoltas por uma matriz constituída por pectina e hemicelulose, como pode ser observado na Figura 1.

Em muitos tipos de células, após o fim do crescimento celular, camadas adicionais são depositadas internamente à parede primária, originando a parede secundária (COSGROVE, 2005). Essas camadas são denominadas S1, S2 e S3 e resultam na formação de uma estrutura espessa e resistente (CARVALHO et al., 2009). A parede secundária é composta principalmente por celulose, xilana (um tipo de

hemicelulose) e lignina.

A Figura 2 apresenta um desenho esquemático das três camadas que compõem a parede celular vegetal em uma célula com crescimento secundário.

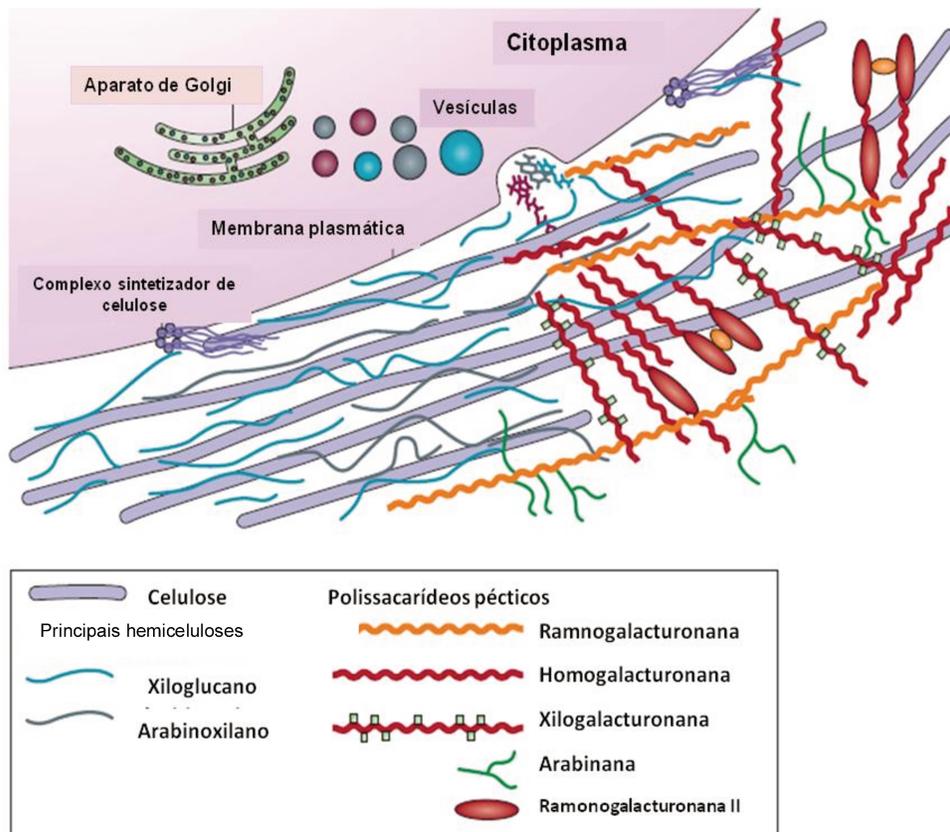


Figura 1 – Estrutura da parede celular vegetal primária. Adaptada de COSGROVE, 2005.

As pectinas constituem um grupo de polissacarídeos estruturais encontrados na parede celular primária das células e na lamela média. São extraídos da parede celular vegetal por tratamentos térmicos, tratamentos com ácidos diluídos e pelo uso de quelantes de cálcio. As pectinas incluem os polissacarídeos: homogalacturananas, ramnogalacturonanas I e II, galactanas e arabinanas (COSGROVE, 2005).

A celulose é o polímero natural mais abundante no planeta e na parede celular vegetal, constituindo cerca de 50% do peso seco em madeiras (QUIROZ-CASTAÑEDA & FOLCH-MALLOL, 2011). Seu grau de polimerização difere de acordo com sua origem e pode variar de 2.000 a 25.000 monômeros de glicose unidos por ligações glicosídicas

do tipo  $\beta(1\rightarrow4)$  (LEVY et al., 2002, HILDEN & JOHANSSON, 2004). A celulose é encontrada na natureza na forma de filamentos, chamados de microfibrilas, que contêm de 36 a 1.200 cadeias de celulose, com diâmetro de 5 a 15 nm (LEVY et al., 2002). A interação entre as microfibrilas de celulose ocorre por pontes de hidrogênio e por forças de Van der Waals, resultando em uma estrutura cristalina e organizada, intercalada por regiões amorfas (5 a 20%), as quais são as mais suscetíveis à hidrólise enzimática (RUBIN, 2008, BRETT, 2000, FUJITA et al., 2004).

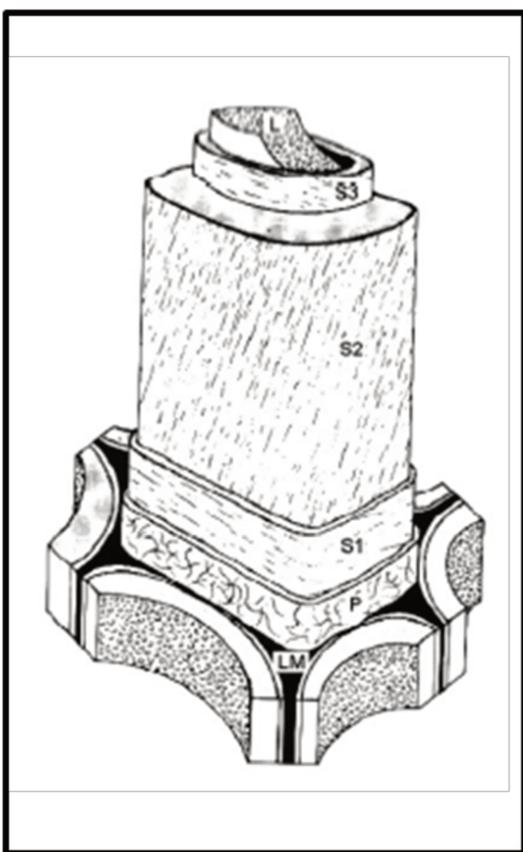


Figura 2 – Representação esquemática das camadas da parede celular de uma célula vegetal com crescimento secundário: LM – Lamela média, P – Parede celular primária, S1 – Parede secundária externa, S2 – Parede secundária média, S3 – Parede secundária interna, L – Lúmen. Adaptada de CARVALHO et al., 2009.

A hemicelulose constitui de 25 a 30% do peso seco das madeiras em geral. Trata-se de um heteropolissacarídeo composto principalmente por pentoses (D-xilose e L-arabinose) e hexoses (D-glucose, D-manose e D-galactose), normalmente acetilados,

formando cadeias ramificadas, e também por ácidos 4-O-metilglucurônico, D-galacturônico e glucurônico, condensados por ligações do tipo  $\beta(1\rightarrow4)$  e ocasionalmente  $\beta(1\rightarrow3)$ . As cadeias laterais, formadas por diferentes açúcares tornam a hemicelulose menos resistente à hidrólise (PEREZ et al., 2002).

A xilana é o tipo de hemicelulose mais abundante. Ela é constituída por unidades de D-xiloses condensadas por ligações do tipo  $\beta(1\rightarrow4)$ , podendo apresentar substituições, originando arabinoxilana (no caso de substituição por arabinose), glucuronoxilana e glucuronoarabinoxilana (se substituída por glicose ou glicose-arabinose, respectivamente). As xilanas podem ainda conter outros tipos de substituições, como: ácido glucurônico, ácido 4-O-metil éter glucurônico, ácido acético, ferúlico e p-cumárico (QUIROZ-CASTAÑEDA & FOLCH-MALLOL, 2011). Outros tipos de hemicelulose são mananas, galactomananas e glucomananas, compostas principalmente por resíduos de manose unidos por ligação  $\beta(1\rightarrow4)$ , podendo apresentar ramificações de galactose no caso das galactomananas, e ramificações de glicose nas glucomananas (QUIROZ-CASTAÑEDA & FOLCH-MALLOL, 2011).

O terceiro componente, a lignina, é um dos polímeros mais abundantes na natureza. Trata-se de uma macromolécula altamente resistente à degradação química e biológica, que promove suporte estrutural para a parede celular, diminuindo a permeabilidade e proporcionando resistência ao ataque de micro-organismos (RUBIN, 2008). Estruturalmente, a lignina se organiza na forma de um heteropolímero reticular, insolúvel em água, resultante da polimerização de três tipos de álcoois p-hidróxi-cinâmílicos precursores: p-cumarílico, coniferílico e sinapílico (CARVALHO et al., 2009). Trata-se do componente da parede celular vegetal que apresenta a maior resistência à degradação, o que limita a sua aplicação e também o acesso aos polissacarídeos protegidos por ela (QUIROZ-CASTAÑEDA & FOLCH-MALLOL, 2011). A lignina se encontra covalentemente ligada à celulose e xilana. Ligações cruzadas podem ocorrer por diversas formas, entre elas: pontes de hidrogênio, ligações iônicas com íons  $\text{Ca}^{2+}$ , ligações covalentes do tipo éster e éter e forças de Van der Waals (LAUREANO-PEREZ et al., 2005).

A estrutura resultante da interação entre a celulose, hemicelulose e lignina pode ser visualizada na Figura 3.

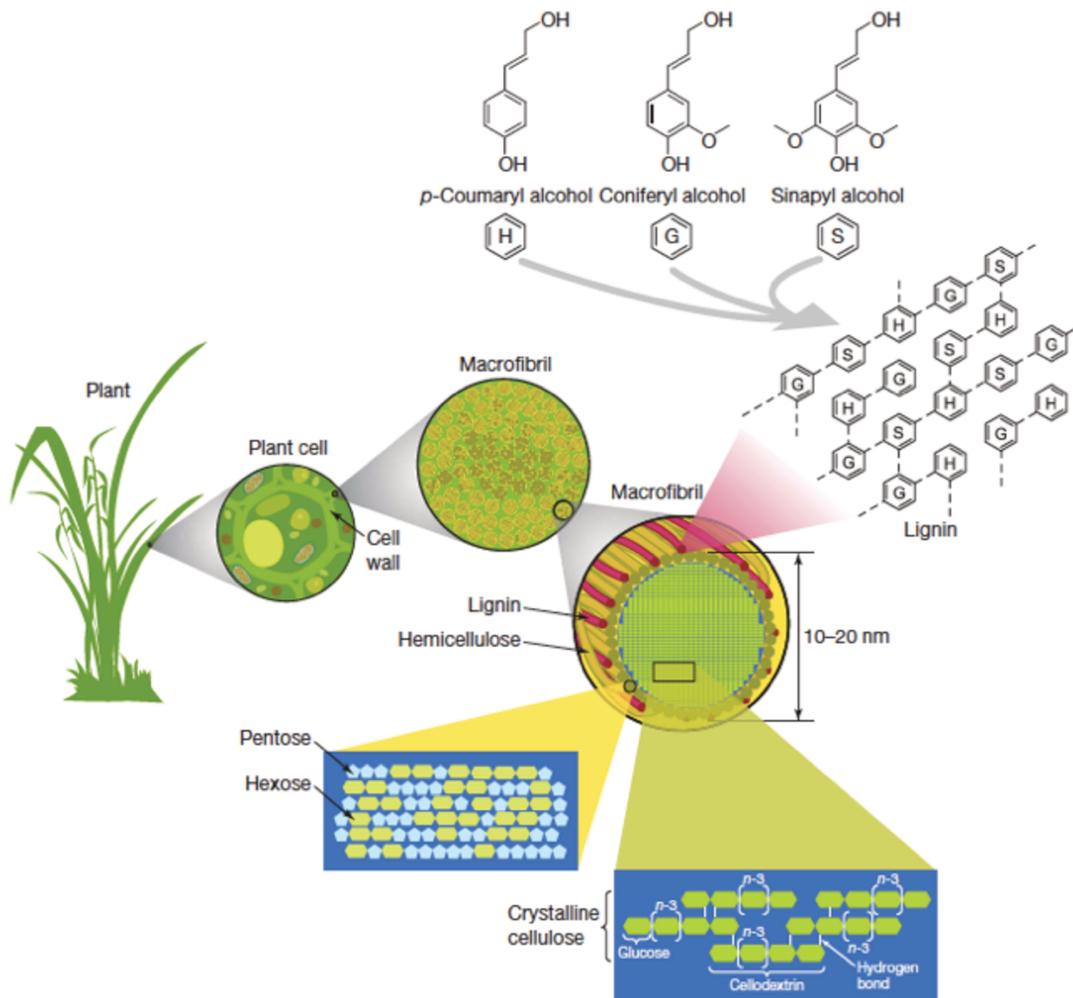


Figura 3 – Figura esquemática da arquitetura da parede celular vegetal da biomassa lignocelulósica. Fonte: RUBIN, 2008.

### 1.3. Desconstrução da Parede Celular Vegetal

O conhecimento da maquinaria enzimática utilizada pelos micro-organismos para a desconstrução da biomassa lignocelulósica oferece novas alternativas para o desenvolvimento de processos que podem ser aplicados em escala industrial para a degradação da biomassa vegetal (RUBIN, 2008). Os micro-organismos lignocelulolíticos, principalmente fungos e bactérias, são capazes de degradar os

componentes da parede celular vegetal através da secreção de um conjunto de enzimas que atuam de forma sinérgica (HIMMEL et al., 2010).

Exemplos de tais micro-organismos são as bactérias aeróbicas celulolíticas do filo Actinobacteria que se desenvolvem em solos, amostras de água e resíduos agrícolas, como bagaço de cana-de-açúcar e restos de folhas (DOI, 2008). Bactérias anaeróbicas do filo Firmicutes geralmente encontradas em solo, biomassa em decomposição, rúmen de animais e intestino de cupins, expressam um complexo enzimático celulolítico, denominado celulosomo (DOI, 2008). Ambientes extremos também abrigam uma série de micro-organismos capazes de degradar celulose, como por exemplo, a bactéria *Pseudoalteromonas haloplanktis*, isolada na Antártida (SONAN et al., 2007).

Os fungos são os principais organismos responsáveis pela degradação de material lignocelulósico (SANCHEZ, 2009). Até mesmo o grupo mais primitivo, os quitridiomycetos, apresentam representantes que atuam na degradação de celulose no intestino de animais ruminantes (LYND et al., 2002). Em geral, o processo de degradação ocorre extracelularmente de duas formas: pelo sistema enzimático composto por hidrolases, que atuam na degradação de polissacarídeos, e pela via oxidativa, que atua principalmente na degradação de lignina (SANCHEZ, 2009). A espécie *Trichoderma reesei* e seus mutantes são de grande interesse industrial devido à secreção de um potente coquetel celulolítico, que inclui exoglucanases, endoglucanases e  $\beta$ -glicosidases (LYND et al., 2002). Já o fungo *Phanerochaete chrysosporium* é o organismo alvo de estudo para a degradação de material lignocelulósico por secretar um sistema enzimático complexo que atua eficientemente na degradação de lignina (SANCHEZ, 2009).

### **1.3.1. Desconstrução enzimática da parede celular vegetal**

O emprego da hidrólise enzimática para a desconstrução da parede celular vegetal depende da combinação de atividades catalíticas das chamadas hidrolases glicolíticas (RUBIN, 2008, FARINAS, 2011). A Figura 4 ilustra a ação das principais

enzimas que atuam no processo de degradação da celulose e hemicelulose.

As celulases são O-glicosídeo hidrolases que atuam na quebra das ligações  $\beta(1\rightarrow4)$  da celulose (QUIROZ-CASTAÑEDA & FOLCH-MALLOL, 2011). São predominantemente encontradas em fungos e procariotos (HILDEN & JOHANSSON, 2004). Elas são classificadas de acordo com sua atividade enzimática em três grupos principais: exoglucanases, endoglucanases e  $\beta$ -glicosidases (GILBERT, 2010).

As endoglucanases clivam as ligações internas da celulose principalmente nas regiões amorfas liberando oligossacarídeos de vários tamanhos e criam novas extremidades para o ataque de outras enzimas, como as exoglucanases (ARO et al., 2005). As exoglucanases, também chamadas de celobiohidrolases, catalisam a hidrólise da cadeia de celulose em ambas as extremidades, redutora e não redutora, resultando na liberação de moléculas de celobiose. Geralmente, elas compõem cerca de 40 a 70% de sistemas enzimáticos celulolíticos e são capazes de hidrolisar celulose cristalina (ARO et al., 2005). A quebra da celobiose e celo-oligossacarídeos em monômeros de glicose requer a atuação das  $\beta$ -glicosidases (ARO et al., 2005). Quando atuam em conjunto, as exoglucanases, endoglucanases e  $\beta$ -glicosidases apresentam um rendimento maior do que a soma dos rendimentos individuais (CASTRO & PEREIRA Jr., 2010). Trata-se do fenômeno conhecido como sinergia (LYND et al., 2002).

Ainda em relação à degradação da celulose, existem proteínas conhecidas como expansinas, que ao se ligarem às microfibrilas de celulose, enfraquecem as ligações de hidrogênio entre as cadeias, facilitando assim o acesso das celulases (LEE et al., 2001).

Apesar da hemicelulose apresentar estrutura complexa, seu mecanismo de degradação é bem conhecido. A degradação completa da xilana, principalmente em xilose e arabinose, requer a cooperação de várias enzimas hidrolíticas, como as endo-xilanases,  $\beta$ -xilosidases e proteínas acessórias (PEREZ et al., 2002). As endo-xilanases atuam na hidrólise das ligações internas da cadeia principal de xilana, liberando predominantemente xilo-oligossacarídeos, que são posteriormente convertidos em xilose pela ação das  $\beta$ -xilosidases (SANCHEZ, 2009).

As proteínas acessórias estão envolvidas na remoção das cadeias laterais que facilitam o acesso das enzimas responsáveis pela hidrólise das ligações internas e externas da cadeia principal (ARO et al., 2005, DE VRIES & VISSER, 2001). Entre elas estão as xilana esterases, feruloil e cumaril esterases,  $\alpha$ - arabinofuranosidades,  $\alpha$ -4-metil glucuronidasas, entre outras que atuam de forma sinérgica (ARO et al., 2005, PEREZ et al., 2002). As feruloil esterases são enzimas que separam a hemicelulose da lignina, pela quebra da ligação éster entre o ácido ferúlico e a hemicelulose (KOSEKI et al., 2006). Foi observado que a atividade da feruloil esterase de *Aspergillus sp.* resultou no aumento da liberação de açúcares de materiais lignocelulósicos pela remoção dos resíduos de ácido ferúlico, promovendo a desestabilização da estrutura da hemicelulose, tornando-a mais susceptível à ação das enzimas hidrolíticas (HERMOSO et al., 2004, RAMIREZ et al., 2008).

A degradação da lignina envolve a ação de enzimas oxidativas que liberam produtos altamente instáveis, que são utilizados em reações de oxidação (BOERJAN et al., 2003). As peroxidases se dividem em dois grupos: lignina peroxidases e peroxidases manganês dependentes, sendo que ambas são oxiredutases que catalisam as oxidações dependentes de peróxido que envolvem compostos fenólicos e não fenólicos, essenciais para a degradação da lignina (PEREZ et al., 2002). As lacases são polifenol-oxidases com quatro íons de cobre em seu sítio ativo que catalisam a oxidação de vários compostos fenólicos e não fenólicos. Elas são capazes de oxidar a lignina e gerar radicais aromáticos altamente instáveis que favorecem a despolimerização através da quebra das ligações éter C4, quebra de anéis aromáticos e desmetoxilação (SANCHEZ, 2009).

Em 1991, foi proposto por Henrissat e colaboradores, um sistema de classificação das enzimas ativas em carboidratos (CAZymes), organizado em um banco de dados denominado CAZy *Database* (disponível em <http://www.cazy.org/Glycoside-Hydrolases.html>). A classificação foi realizada com base na similaridade das seqüências de aminoácidos, enovelamento proteico e mecanismo de catálise das enzimas (GILBERT, 2010, LEVASSEUR et al., 2013, HENRISSAT & BAIROCH, 1993, HENRISSAT, 1991). Atualmente o banco de dados do CAZy contempla mais de 300 famílias incluindo glicosil hidrolases, glicosil transferases, polissacarídeo liases,

carboidrato esterases e domínios de ligação a carboidratos (CBMs) (LEVASSEUR et al., 2013). Na última atualização ocorrida em 2013, foi criada uma nova classe, que agrupa as proteínas com atividades auxiliares, incluindo as que atuam na degradação de lignina (LEVASSEUR et al., 2013).

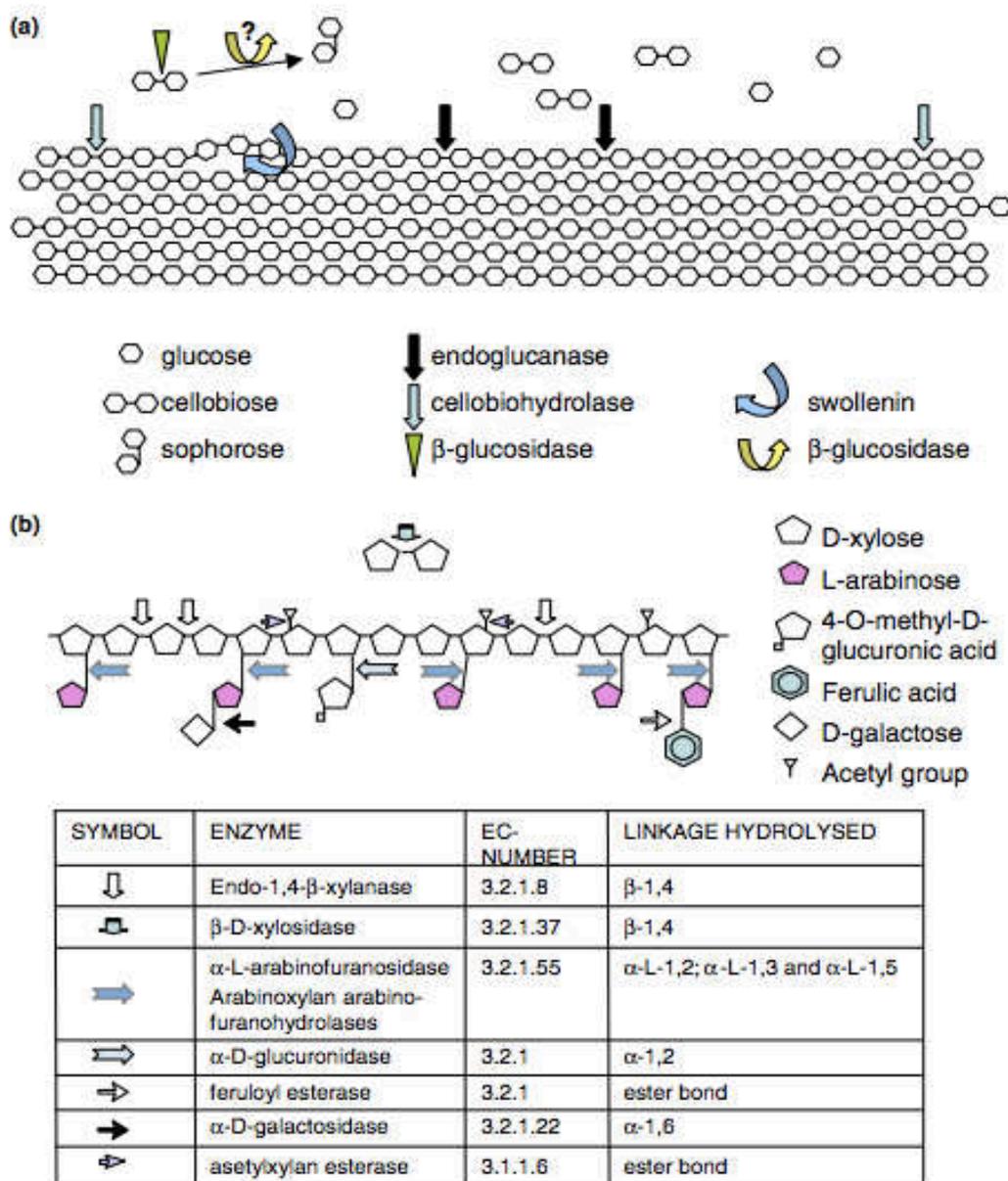


Figura 4 – (a) Esquema da hidrólise da celulose pelo sistema celulolítico composto por endoglucanases, celobiohidrolases, β-glicosidases e pela proteína acessória expansina (b) Esquema da hidrólise de arabinoxilana por hemicelulases e proteínas acessórias. Fonte: ARO et al, 2005.

### 1.3.2. Domínios de ligação a carboidratos

A natureza insolúvel de muitos polissacarídeos que compõem a parede celular vegetal representa um desafio para a atividade das enzimas que atuam na degradação da biomassa (LYND et al., 2002). Várias enzimas ativas em carboidratos apresentam-se estruturalmente divididas em módulos, denominados de domínios catalíticos e domínios de ligação a carboidratos (CBM) (FREELOVE et al., 2001, BLACK et al., 1996), geralmente unidos por um *linker* flexível (GILBERT, 2010), como ilustrado na Figura 5.

A função mais comumente aceita para o CBM é a de ligação aos polissacarídeos, promovendo a aproximação do domínio catalítico ao substrato, facilitando sua degradação. Uma função secundária do CBM está relacionada à movimentação da enzima ao longo do substrato, de forma que a hidrólise se processe de forma contínua (GUILLEN et al., 2010). Exemplos de polissacarídeos reconhecidos pelos CBMs são: celulose, quitina,  $\beta$ -glucanos, amido, glicogênio, inulina e pululana. Alguns CBMs são específicos para um tipo de ligante, enquanto outros reconhecem uma gama de diferentes substratos (BORASTON et al., 2004).

Existem muitos relatos na literatura que descrevem a excisão proteolítica dos CBMs resultando em significativa perda da atividade enzimática nos substratos insolúveis (BORASTON et al., 2004), porém a atividade nos substratos solúveis normalmente não sofre alteração (GUILLEN et al., 2010).

O elemento conformacional mais importante na estrutura secundária dos CBMs é a folha- $\beta$  (GUILLEN et al., 2010), sendo o  $\beta$ -sanduíche o tipo de enovelamento mais comum (BORASTON et al., 2004).

A topologia dos sítios de ligação dos CBMs está relacionada aos tipos de ligantes que o CBM reconhece e é utilizada como critério de classificação em três diferentes tipos (BORASTON et al., 2004). O tipo A apresenta superfície hidrofóbica e plana composta por resíduos de aminoácidos aromáticos, sendo capaz de interagir com a superfície plana de polissacarídeos cristalinos, como celulose e quitina. Por outro lado, os CBMs do tipo B apresentam uma espécie de fenda na qual os aminoácidos aromáticos interagem com cadeias livres e únicas de polissacarídeos. Estes tipos de CBMs reconhecem  $\beta$ -1,3-glucanos,  $\beta$ -1,3-1,4-glucanos,  $\beta$ -1,4-mananas, glucomananas

e galactomananas. O terceiro tipo é o C, que se liga apenas aos mono-, di- ou trissacarídeos devido ao impedimento estérico que ocorre no sítio de ligação (BORASTON et al., 2004).

Os resíduos de aminoácidos envolvidos nos sítios de ligação podem ser identificados através de estudos de ressonância nuclear magnética (RMN), mutagênese sítio dirigida e dados estruturais (GUILLEN et al., 2010). Em geral, os aminoácidos aromáticos, principalmente o triptofano e a tirosina, são os responsáveis pela interação. Estes resíduos de aminoácidos interagem com os anéis das estruturas dos açúcares através de forças de Van der Waals que estabilizam o complexo. A estabilização também ocorre por pontes de hidrogênio formadas entre as cadeias laterais dos resíduos de aminoácidos e dos açúcares. Muitos estudos sugerem que a orientação dos resíduos aromáticos é determinante para as diferentes especificidades apresentadas pelos CBMs (GUILLEN et al., 2010).

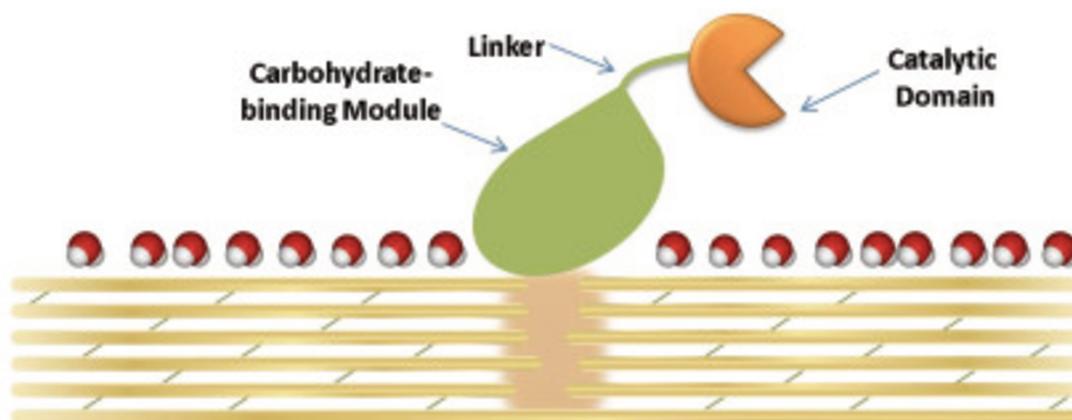


Figura 5 - Esquema da divisão de uma celulase que possui em sua estrutura um domínio catalítico e um domínio de ligação a carboidratos, unidos por um peptídeo flexível, chamado de *linker*. Fonte: ARANTES & SADDLER, 2010.

A termodinâmica da interação do CBM com diferentes ligantes pode ser avaliada através da técnica de calorimetria de titulação isotérmica (ITC) (LAMMERTS VAN BUEREN & BORASTON, 2004), a qual permite medir a variação de entalpia ( $\Delta H$ ) que ocorre durante a ligação entre duas moléculas a temperatura constante (LADBURY & DOYLE, 2004).

O ponto inicial para tais estudos termodinâmicos é a determinação experimental da constante de associação ( $K_a$ ), que corresponde ao inverso da constante de dissociação ( $K_D$ ). A partir do valor de  $K_a$ , a variação da energia livre de ( $\Delta G$ ) pode ser calculada através da equação 1.1 (LADBURY & DOYLE, 2004):

$$\Delta G = -RT \ln K_a$$

Equação 1.1 - Equação para o cálculo de  $\Delta G$  a partir do valor de  $K_a$ .

Onde T corresponde à temperatura absoluta em Kelvin e R é a constante universal dos gases.

Através do valor de  $\Delta H$  e  $\Delta G$  é possível calcular a variação de entropia ( $\Delta S$ ) usando a equação 1.2 (LADBURY & DOYLE, 2004):

$$\Delta S = (\Delta H - \Delta G)/T$$

Equação 1.2 - Equação para o cálculo da variação de  $\Delta S$  a partir dos valores de  $\Delta H$  e  $\Delta G$ .

Valores negativos de  $\Delta G$  demonstram que a reação em questão ocorre de forma espontânea. Os valores de  $\Delta H$  expressam a medida do conteúdo energético das ligações que são quebradas e formadas durante a interação, principalmente as ligações de hidrogênio. Valores negativos de  $\Delta H$  indicam que a variação de entalpia foi favorável à interação entre as moléculas avaliadas. Em relação à entropia, valores positivos de  $\Delta S$  indicam se tratar de reações entropicamente dirigidas e a principal contribuição neste caso é dada pelas interações hidrofóbicas (LADBURY & DOYLE, 2004).

De forma semelhante às glicosil hidrolases, os CBMs são classificados em famílias com base na similaridade de sequências de aminoácidos (BORASTON et al., 2004). Em julho de 2009, o banco de dados do CAZy era constituído 55 famílias de CBMs (GUILLEN et al., 2010), porém atualmente este número subiu para 67.

Um fator importante para constante inclusão de novas famílias ao banco de dados do CAZy é a quantidade de informação gerada pelos trabalhos de genômica e metagenômica, principalmente através do emprego de tecnologias de sequenciamento

de segunda e terceira geração (YOOSEPH et al., 2007, FREELOVE et al., 2001).

A descoberta de novas famílias contribui significativamente para o entendimento das estratégias através das quais os CBMs reconhecem e se ligam aos substratos (GUILLEN et al., 2010). Tais informações podem ser utilizadas para o desenvolvimento de tecnologias que explorem os mecanismos de reconhecimento de carboidratos (BORASTON et al., 2004, GUILLEN et al., 2010). Atualmente existem várias patentes depositadas por empresas de grande porte descrevendo o uso de CBMs em processos de conversão da biomassa vegetal, em indústrias alimentícias, como forma de retardar o endurecimento de massas e também em indústrias têxteis, pelo fato da ligação dos CBMs às fibras dos tecidos resultar no aumento da resistência desses materiais (CHRISTENSEN & SCHNORR, 2007).

#### **1.4. Metagenômica**

Uma das primeiras indicações de que os micro-organismos cultivados não representavam uma parcela significativa do universo microbiano foi em decorrência da chamada “grande anomalia em contagem de placas”. Neste caso, era verificada uma discrepância muito grande entre o tamanho das populações microbianas estimadas por contagem microscópica direta de células e a obtida por contagem em placas de cultivo (STALEY & KONOPKA, 1985).

Estima-se que o número total de células procarióticas no planeta Terra seja de 4 a  $6 \times 10^{30}$  (WHITMAN et al., 1998) e que desse total cerca de 95 a 99,9% não sejam passíveis de cultivo em condições de laboratório (COWAN et al., 2005). Os micro-organismos não cultivados são essenciais em inúmeros processos que ocorrem na natureza e representam uma fonte ilimitada de aplicações biotecnológicas (RIESENFELD et al., 2004).

Como forma de superar as dificuldades no cultivo dos organismos que compõem a microbiota de diversos ambientes naturais, métodos moleculares indiretos baseados no isolamento e análise dos ácidos nucleicos foram desenvolvidos (HANDELSMAN, 2004). O estudo dos ácidos nucleicos recuperados diretamente das amostras ambientais é denominado de metagenômica. A coleção de material genético isolada de um determinado ambiente é denominada de metagenoma (DANIEL, 2005).

A clonagem de DNA metagenômico foi descrita pela primeira vez em 1985 por Pace et al. Neste trabalho foi descrita a construção de uma biblioteca metagenômica em vetor do tipo fago com DNA extraído de uma comunidade microbiana derivada de amostra de grama seca enriquecida em laboratório (PACE et al., 1985). O trabalho que consolidou este campo de pesquisa foi desenvolvido pelo grupo do pesquisador DeLong, no qual foi reportada a identificação de um clone contendo um gene de RNAr 16S relacionado a um organismo classificado no domínio Archaea nunca antes cultivado (STEIN et al., 1996).

Desde então, a metagenômica tem sido utilizada com sucesso na avaliação de genes específicos em amostras ambientais, determinação de vias metabólicas complexas, descrição do metabolismo de bactérias desconhecidas bem como estudos de comunidades microbianas (ENTCHEVA et al., 2001, KNIETSCH et al., 2003).

A diversidade biológica é comumente avaliada com base nos genes RNAr 16S (SCHMEISSER et al., 2007), que permitem a identificação e inserção de organismos em um contexto evolutivo. Esta prática é possível devido à existência de sequências conservadas que podem ser amplificadas através de reações de PCR com oligonucleotídeos iniciadores universais (STEIN et al., 1996). Em seguida, os produtos de PCR são clonados em vetores e submetidos ao sequenciamento. A análise filogenética dos dados gerados é feita por comparação com bancos de dados especializados em sequências de RNAr 16S, como o *Ribosomal Database Project* (RDP- <http://rdp.cme.msu.edu/>), *European Ribosomal Database* (<http://www.psb.ugent.be/rRNA/>) e *Silva Database* (<http://www.arb-silva.de/>).

Desde o fim dos anos 90, abordagens metagenômicas têm sido empregadas por empresas, como BASF ([www.corporate.basf.com](http://www.corporate.basf.com)), B.R.A.I.N ([www.brain-biotech.de](http://www.brain-biotech.de)), Invitrogen ([www.invitrogen.com](http://www.invitrogen.com)), Genencor ([www.genencor.com](http://www.genencor.com)), na busca por novas enzimas e compostos ativos (COWAN et al., 2005).

#### **1.4.1. Metagenômica de solo**

Acredita-se que as comunidades microbianas encontradas no solo sejam as mais desafiadoras em relação ao tamanho e diversidade genética. Estima-se que um grama

de solo contenha  $10^9$  células procarióticas e cerca de 2.000 a 8.000 genomas (DANIEL, 2005). Um dos motivos para tal diversidade está relacionado à heterogeneidade espacial, natureza multifásica, além das propriedades químicas e biológicas do solo (DANIEL, 2005). A versatilidade dos micro-organismos que habitam esses ambientes é de grande interesse industrial, e tem sido a principal fonte de produtos naturais (OSBURNE et al., 2000).

Muitos trabalhos descrevem protocolos para a extração de DNA metagenômico do solo (HOLBEN et al., 1988, STEFFAN et al., 1988, JACOBSEN & RASMUSSEN, 1992). O maior desafio no processo de extração é a remoção de compostos que se ligam às moléculas de DNA, como por exemplo, ácidos húmicos, que podem interferir nas etapas posteriores de fragmentação, clonagem e até mesmo em análises que envolvam reações de PCR para a amplificação de sequências específicas (DANIEL, 2005). Colunas Sephadex G-200 constituem uma das maneiras mais efetivas para a eliminação dos contaminantes. A técnica de eletroforese de campo pulsado (PFGE, Pulse Field Gel Electrophoresis), utilizando gel de agarose bi-fásico, com uma das fases composta por polivinilpirolidona (PVP), também é utilizada para a remoção de ácidos húmicos (QUAISER et al., 2002).

Novas enzimas foram identificadas através do emprego da triagem funcional de bibliotecas metagenômicas construídas a partir de DNA extraído de amostras de solo (GUPTA et al., 2002, HENNE et al., 2000, RICHARDSON et al., 2002, VOGET et al., 2006, HU et al., 2008). Outros trabalhos descrevem a aplicação bem sucedida da metagenômica no entendimento da composição filogenética e diversidade funcional de amostras de solo de diferentes ambientes (ANDREOTE et al., 2012, FIERER et al., 2012, LI et al., 2013).

#### **1.4.2. Prospecção de novos genes em metagenomas**

Pelo fato de muitas vezes os genes de interesse representarem uma fração reduzida no conteúdo total de material genético, técnicas de pré-enriquecimento da amostra são comumente utilizadas (COWAN et al., 2005). O número de genes identificados aumenta significativamente através da aplicação de técnicas como: marcação com 5-bromo-2'-deoxiuridina (BrdU), hibridização subtrativa supressiva

(SSH), *phage display*, captura por afinidade e exposição da comunidade microbiana à pressão seletiva, que pode ser realizada pela adição de substratos específicos a serem metabolizados ou ainda exposição a determinadas condições físicas, químicas e nutricionais (COWAN et al., 2005, LI et al., 2009).

Exemplos das vantagens da realização da etapa de enriquecimento são descritos na literatura (GABOR et al., 2004, VOGET et al., 2003, SEBAT et al., 2003, REES et al., 2003). O trabalho conduzido por Voget em 2003 justificou em função da etapa de enriquecimento, a elevada frequência de clones celulolíticos identificados na triagem funcional de uma biblioteca metagenômica de DNA de solo. O trabalho de Rees em 2003 demonstrou que o cultivo de amostras de água em meio mínimo adicionado de CMC e óleo de oliva, separadamente, resultou no aumento da frequência genes de celulases e esterases, respectivamente, em comparação com as amostras que não foram submetidas ao enriquecimento com tais substratos (REES et al., 2003).

Duas estratégias complementares são utilizadas para a prospecção de novos genes em metagenomas, a baseada na triagem funcional de bibliotecas metagenômicas e a baseada em homologia de sequência nucleotídica (LI et al., 2009). Em ambos os casos, há a necessidade de obtenção de amostras de DNA metagenômico de alto peso molecular, livre de contaminantes e que sejam representativas das comunidades microbianas (DANIEL, 2005). O procedimento de extração pode ser feito de forma direta, quando as células são lisadas na amostra, ou indireta, no qual as células são removidas e posteriormente lisadas. A comparação destas duas técnicas mostrou não existir diferença no espectro de diversidade obtido (COURTOIS et al., 2003). Normalmente, maiores quantidades de DNA com menor grau de pureza são obtidas utilizando-se o método de lise direta (TRINGE & RUBIN, 2005).

A prospecção de genes através da triagem funcional envolve primeiramente a construção de bibliotecas através da clonagem direta do DNA metagenômico, que envolve os mesmos procedimentos de clonagem de DNA genômico: fragmentação do DNA através de digestão com enzimas de restrição ou por quebra mecânica e ligação dos fragmentos em vetor de clonagem seguida por transformação em hospedeiros apropriados (KNIETSCH et al., 2003, DANIEL, 2005).

Em relação ao tamanho do inserto utilizado, as bibliotecas podem ser

classificadas como de pequeno inserto (menor que 15 kb), nas quais são utilizados vetores do tipo plasmídeo, ou de grande inserto (acima de 40 kb), em que se empregam vetores do tipo cosmídeos, fosmídeos ou BACs (cromossomos artificiais de bactérias). Bibliotecas de pequeno inserto são adequadas para o isolamento de genes ou operóns relacionados a novas funções metabólicas. Para a prospecção de clusters gênicos ou grandes fragmentos de DNA, recomenda-se o uso de bibliotecas de grandes insertos (DANIEL, 2005).

No caso da triagem funcional, a detecção de novos genes se baseia na identificação de clones que expressam atividades enzimáticas relacionadas aos genes de interesse (LI et al., 2009). As vantagens deste tipo de abordagem são a identificação de genes fracamente ou totalmente dissociados de outros conhecidos, análise simultânea de milhares de clones e possibilidade de localização de genes inteiros e funcionais (DANIEL, 2005). Por outro lado, existe a desvantagem em relação à dependência da expressão dos genes pelo organismo hospedeiro que podem tanto não serem expressos como inativados (DANIEL, 2005). O principal hospedeiro utilizado é a bactéria *Escherichia coli*, porém muitos genes apresentam requerimentos para expressão que não são disponibilizados pela *E.coli*, podendo se utilizar *Pseudomonas putida* e *Streptomyces lividans* como hospedeiros alternativos (DANIEL, 2005).

A PCR e a hibridização utilizando sondas alvo-específicas são as técnicas mais empregadas na abordagem dirigida por homologia de sequência nucleotídica (COURTOIS et al., 2003, KNIETSCH et al., 2003). Ambas as abordagens requerem o uso de sondas e oligonucleotídeos iniciadores apropriados que derivam de regiões conservadas de genes conhecidos. Alternativamente, através de ferramentas de bioinformática, os genes podem ser identificados em banco de dados criados a partir das informações provenientes do sequenciamento em larga escala de metagenomas (LI et al., 2009). O sequenciamento em larga escala aliado ao uso de ferramentas de bioinformática é uma ferramenta amplamente utilizada para acessar a informação genética contida no genoma dos micro-organismos não cultivados (DANIEL, 2005, SIMON & DANIEL, 2011, VENTER et al., 2004). As plataformas de sequenciamento mais utilizadas em estudos de metagenômica são 454 FLX da Roche, Solexa da Illumina e SOLiD da Applied Biosystems (YU & ZHANG, 2012, SMITH et al., 2012,

ZHOU et al., 2010).

A abordagem baseada em homologia de sequência nucleotídica apresenta como principal desvantagem o fato de estar relacionada à descoberta de genes homólogos a outros já conhecidos. Porém, diferentemente das estratégias baseadas na triagem funcional, permite o resgate de genes sem depender de sistemas de expressão de células hospedeiras e da integralidade da sequência alvo (LI et al., 2009).

#### **1.4.2.1. Prospecção de glicosil hidrolases**

Um dos principais trabalhos que descreve a identificação de genes que codificam glicosil hidrolases por abordagem dirigida por homologia de sequência nucleotídica foi publicado em 2011 por Hess e colaboradores. Neste trabalho, através da tecnologia *Illumina Hi-seq* foram sequenciados e analisados 268 Gb da amostra de DNA metagenômico extraída da comunidade microbiana aderida às fibras vegetais presentes no rúmen bovino. Foram identificados 27.755 prováveis genes que codificavam enzimas ativas em carboidratos, sendo que 43% deles apresentaram identidade menor que 50% com genes conhecidos, distribuídos em 21 famílias de glicosil hidrolases distintas. Desse total de genes, noventa foram expressos e aproximadamente a metade foi confirmada em relação à atividade enzimática (HESS et al., 2011).

A Tabela 1 apresenta alguns exemplos de trabalhos que identificaram novos genes de glicosil hidrolases através de abordagens dirigidas por homologia de sequência nucleotídica.

A abordagem dirigida por triagem funcional foi empregada com sucesso em diversos trabalhos para a identificação de novas glicosil hidrolases (Tabela 2). Em 2006, Voget e colaboradores descreveram o isolamento de uma celulase no metagenoma extraído de um consórcio microbiano derivado de solo. A enzima apresentou características interessantes em relação à estabilidade térmica, tolerância a elevadas concentrações salinas e atividade em ampla faixa de pH (VOGET et al., 2006). Outro exemplo é a xilanase recuperada do metagenoma de solo de compostagem, que apresentou potencial aplicação na indústria de papel e celulose, em virtude da alta atividade catalítica em condições de temperatura elevada e pH alcalino (VERMA et al., 2013).

Tabela 1 – Exemplos de trabalhos que utilizaram abordagens dirigidas por homologia de sequência nucleotídica para identificação de novos genes que codificam glicosil hidrolases.

Metagenoma de origem/ Referência	Número de megabases sequenciadas	Número de genes preditos	Genes preditos com identidade para glicosil hidrolases
Intestino humano/ (GILL et al., 2006)	36,3	46.503	705
Amostra de água do mar/ (KUNIN et al., 2008b)	84,3	135.922	786
Amostra de água enriquecida com formaldeído/ (KALYUZHAYAYA et al., 2008)	57,6	89.729	397
Intestino de rato/ (TURNBAUGH et al., 2006)	6,5	8.510	119
Intestino de cupim/ (WARNECKE et al., 2007)	62	83.225	1.267
Rúmen bovino/ (HESS et al., 2011)	268.000	2.547.270	27.755
Consórcio microbiano termofílico/ (XIA et al., 2013)	1.200	31.499	253

Tabela 2 – Exemplos de trabalhos que utilizaram as abordagens dirigidas por triagem funcional para identificação de novos genes que codificam glicosil hidrolases.

Metagenoma de origem/ Referência	Enzima alvo	Vetor utilizado	Tamanho do inserto (kb)	Número de clones positivos identificados / número de clones avaliados
Solo/ (VOGET et al., 2003)	Agarase	Cosmídeo	25-40	1/127
Solo/ (VOGET et al., 2003)	Amilase	Cosmídeo	25-40	1/1.523
Amostra de água/ (REES et al., 2003)	Celulase	Lambda	2-10	1/28.500
Amostra de água (COTTRELL et al., 1999)	Quitinase	Lambda	1.8-4.2	1/37.500
Intestino de inseto/ (BRENNAN et al., 2004)	Xilanase	Lambda	3-6	1/250.000
Solo/ (VOGET et al., 2003)	Pectato liase	Cosmídeo	25-40	1/762



## **CAPÍTULO 2 – OBJETIVOS**

### **2.1 Objetivos Gerais**

O objetivo geral desta tese de doutorado foi o desenvolvimento de estratégias metagenômicas para a identificação de novas enzimas que atuam na desconstrução da parede celular vegetal, bem como a caracterização funcional das mesmas com ênfase na aplicação em processos de conversão da biomassa vegetal para geração de bioprodutos.

### **2.2 Objetivos Específicos**

- Desenvolvimento de metodologias para a prospecção de novos genes que codificam enzimas ativas em carboidratos que compõem a parede celular vegetal através de abordagens metagenômicas dirigidas por homologia de sequência nucleotídica e triagem funcional;
- Clonagem, expressão e purificação das enzimas identificadas visando à caracterização bioquímica e determinação do potencial de aplicação das enzimas em processos de conversão da biomassa vegetal.



## **CAPÍTULO 3 – DESENVOLVIMENTO DE ESTRATÉGIAS PARA A IDENTIFICAÇÃO DE NOVAS ENZIMAS ATIVAS NA DEGRADAÇÃO DA PAREDE CELULAR VEGETAL**

### **3.1. Introdução**

Um dos objetivos principais do emprego da metagenômica é a prospecção de novos genes e produtos gênicos nos genomas de micro-organismos não cultivados (FERRER et al., 2009, DANIEL, 2005). Para atingir tal objetivo, estratégias baseadas em similaridade de sequências de nucleotídeos e na triagem funcional de bibliotecas metagenômicas são empregadas de forma complementar (SIMON & DANIEL, 2011).

Com base nisso, neste capítulo são descritas as metodologias, bem como os resultados e discussões acerca das estratégias metagenômicas desenvolvidas neste trabalho.

### **3.2. Objetivos do Capítulo**

As atividades descritas neste capítulo tiveram como objetivo o desenvolvimento de estratégias metagenômicas para a identificação de novas enzimas que atuam na degradação da biomassa vegetal no metagenoma de solo de canavial.

### **3.3. Materiais e Métodos**

#### **3.3.1. Construção das bibliotecas metagenômicas**

##### **3.3.1.1. Coleta das amostras de solo**

As coletas de amostras de solo foram realizadas em canaviais localizados nos municípios de São Carlos (GPS: S22°04'77.982" O 48°35'29.9", altitude 600 m) e Paulínia (GPS: 22°28'35.03"; -47°6'17.68"; altitude 700m), ambos no Estado de São Paulo, Brasil.

As amostras de solo foram coletadas a uma profundidade de aproximadamente dois centímetros e armazenadas em tubos estéreis do tipo Falcon. No laboratório, as amostras foram liofilizadas, pulverizadas e estocadas em ultrafreezer a - 80°C.

### **3.3.1.2. Consórcio microbiano especializado na degradação de bagaço de cana-de-açúcar**

O procedimento para o desenvolvimento do consórcio microbiano especializado na degradação de bagaço de cana-de-açúcar foi baseado no trabalho de Paixão et al., 2010. Em linhas gerais, aproximadamente um grama da amostra de solo foi usado como inóculo em um frasco tipo Erlenmeyer de 250 mL contendo 100 mL de meio mineral Bushnell-Haas (BUSHNELL & HAAS, 1941) adicionado de 5,0 g/L de bagaço de cana-de-açúcar previamente submetido ao tratamento de explosão a vapor e deslignificação (composição: 76,0% de celulose, 9,1% de hemicelulose, 14,8% de lignina e 1,9% de cinzas), gentilmente cedido pelo professor Dr. George Jackson (CTBE). O frasco contendo o consórcio microbiano foi mantido a 30°C sob agitação de 150 rpm.

Semanalmente o repique era realizado pela transferência de 1 mL do consórcio para um novo frasco contendo meio de cultura fresco. Na 15ª semana de cultivo, o conteúdo do frasco foi deixado em repouso por cerca de duas horas. Em seguida o sobrenadante foi removido e centrifugado por 30 minutos a 5.000 rpm. O *pellet* obtido foi utilizado para extração do DNA metagenômico.

### **3.3.1.3. Extração do DNA metagenômico**

O método empregado para a extração de DNA metagenômico das amostras foi o de lise direta através do FastDNA® SPIN Kit for soil (MP - Biomedicals) e do homogeneizador FastPrep® Instrument (MP - Biomedicals), de acordo com o protocolo recomendado pelo fabricante (Disponível em: <http://www.bio101.com/protocols/dna-kits/p-fastdna-soil.html>).

As amostras de DNA metagenômico extraídas foram denominadas:

- MSC - Metagenoma extraído do solo de canavial de São Carlos;

- MSP - Metagenoma extraído do solo de canavial de Paulínia;
- MCB - Metagenoma extraído do consórcio microbiano degradador de bagaço de cana – de – açúcar.

Os produtos das extrações foram avaliados por eletroforese em gel de agarose 1,0% (m/v) em tampão TAE 1,0 X (Tris-Acetato 40 mM, EDTA 1mM, pH 8,3) a 90 V por 1 hora, utilizando como padrão de peso molecular o marcador *1 kb plus DNA Ladder* (Invitrogen). As bandas foram visualizadas em transiluminador de luz UV, após tratamento com brometo de etídeo 0,25 µg/mL por 15 minutos. As imagens foram capturadas através do fotodocumentador Gel Logic 6000Pro da marca Carestream.

#### **3.3.1.4. Quantificação de DNA**

Os ácidos nucleicos foram quantificados através do método espectrofotométrico por medidas de absorvância a 260 nm realizadas no espectrofotômetro NanoDrop modelo 2000c (Thermo Scientific).

#### **3.3.1.5. Digestão do DNA metagenômico**

A enzima de restrição empregada no processo de digestão do DNA metagenômico foi a *Sau3AI*, que por reconhecer um sítio de restrição de apenas quatro nucleotídeos (GATC) cliva uma sequência nucleotídica em média a cada 256 pares de base (ROHWER et al., 2001). Controlando-se o tempo de digestão, é possível a fragmentação do DNA dentro de uma faixa de tamanho de interesse e com ambas as extremidades possuindo sequências conhecidas.

Inicialmente foi necessária a determinação do tempo de digestão suficiente para a obtenção de fragmentos na faixa de tamanho de 2 a 5 kb. As reações foram realizadas em microtubos e eram compostas por: 3 µL de DNA metagenômico (concentração aproximada de 200 ng/µL), 1 µL tampão 10 X (NEBuffer 1, New England Biolabs ®), 0,01 U da enzima *Sau3AI* (NewEngland Biolabs ®), 0,1 µL de BSA 100 X e água ultrapura estéril até o volume final de 10 µL. As reações foram incubadas à 37°C e em intervalos de tempos regulares os microtubos eram retirados e submetidos à inativação a 65°C por 20 minutos.

Os produtos das digestões foram avaliados por eletroforese em gel de agarose 1,0% em tampão TAE 1,0 X, a 90 V por 1 hora, utilizando como padrão de peso molecular o marcador *1 kb plus DNA Ladder* (Invitrogen). As bandas foram visualizadas em transiluminador de luz UV, após tratamento com brometo de etídeo 0,25 µg/mL por 15 minutos. As imagens foram capturadas através do fotodocumentador Gel Logic 6000Pro da marca Carestream.

Tendo-se determinado a condição ideal de digestão para obtenção de fragmentos na faixa de tamanho de interesse, a escala da reação foi ampliada mantendo as mesmas proporções da reação descrita acima.

### **3.3.1.6. Purificação do DNA metagenômico digerido**

A técnica de visualização de DNA em gel de agarose que envolve o uso do brometo de etídeo e exposição à luz ultravioleta pode causar danos à estrutura da molécula de DNA. Desta maneira, para o isolamento dos fragmentos de DNA metagenômico resultantes da digestão enzimática foi empregado um artifício que não envolveu o uso de nenhum dos dois elementos citados acima.

Primeiramente o marcador de peso molecular *Ladder 1 kb Plus* (Invitrogen) foi aplicado em ambas as extremidades do gel e a amostra de DNA digerido foi aplicada na canaleta central. Após a eletroforese, as extremidades do gel foram cortadas, submetidas à coloração com brometo de etídeo e visualizadas através de transiluminador com luz UV. As extremidades foram utilizadas como molde para o corte da região correspondente à amostra de DNA digerido na canaleta central. Em seguida o DNA foi purificado através do kit *GFX<sup>tm</sup> PCR DNA and gel band purification* da GE Healthcare.

### **3.3.1.7. Preparo do vetor pUC19**

O vetor utilizado foi o plasmídeo pUC19 (Fermentas - Thermo Scientific, USA) (Figura 6). Ele foi adquirido na forma linearizada (por ação da enzima de restrição *Bam*HI) e desfosforilada.

A clonagem dos fragmentos foi possível, pois as extremidades do vetor pUC19 resultantes da clivagem enzimática com *Bam*HI são compatíveis com as extremidades geradas pela fragmentação do DNA metagenômico com *Sau*3AI.

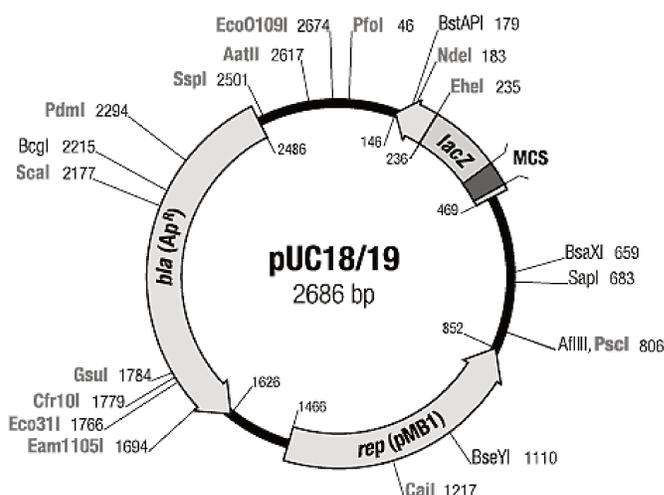


Figura 6 - Mapa do vetor pUC18/19.

### 3.3.1.8. Reação de ligação

Esta etapa consistiu na ligação dos insertos de DNA metagenômico digeridos ao vetor pUC19. Considerando-se um volume final de reação de 15  $\mu$ L e razão molar de inserto: vetor de 3:1 (equação 2), foram adicionados 2 U da enzima T4-DNA ligase (Promega), 1,5  $\mu$ L do vetor linearizado (concentração de aproximadamente 15 ng/ $\mu$ L), 5  $\mu$ L do DNA metagenômico digerido e purificado (concentração de aproximadamente 18 ng/ $\mu$ L), 1,5  $\mu$ L do tampão 10 X (Promega) e água ultrapura estéril para completar o volume final. A reação foi incubada a 8°C por 16 horas.

$$\frac{\text{ng de vetor} \times \text{tamanho do inserto (kb)}}{\text{tamanho do vetor (kb)}} \times \text{razão molar inserto: vetor} = \text{ng de inserto}$$

Equação 2 – Cálculo da quantidade de inserto necessária para a reação de ligação.

### 3.3.1.9. Preparo de células eletrocompetentes

A linhagem de *E.coli* SURE® (*Agilent Technologies - Stratagene Products*) foi crescida em 300 mL de meio TY (20 g/L triptona, 5 g/L extrato de levedura) até atingir valores de absorvância a 600 nm entre 0,4 e 0,5. Em seguida o volume do frasco foi dividido em seis tubos tipo Falcon de 50 mL, centrifugados a 5.000 x g, por 5 minutos a 4°C. Os *pellets* referentes às células precipitadas foram lavados três vezes com água ultrapura estéril gelada e por duas vezes com glicerol 10%. Na última lavagem com glicerol, uma parte do sobrenadante foi descartado conservando-se um volume de aproximadamente 1 mL dentro do tubo para que as células fossem solubilizadas. As células foram armazenadas no gelo até o momento de sua utilização.

### 3.3.1.10. Transformação bacteriana por eletroporação

A técnica de transformação por eletroporação consiste em submeter às células a um campo elétrico de intensidade controlada, durante um curto período de tempo. Na reação foram utilizados 50 µL de células eletrocompetentes e 1,0 µL do produto da reação de ligação. A mistura foi transferida para uma cubeta de eletroporação de 0,2 cm (BioRad) previamente resfriada a -20°C. As amostras foram submetidas ao choque elétrico de 2,5 kV durante aproximadamente 5 milissegundos no eletroporador *Eppendorf Multiporator*®. Em seguida, 1 mL de meio SOC (triptona 20 g/L, extrato de levedura 5 g/L, NaCl 10 mM, KCl 2,5 mM, MgCl<sub>2</sub> 20 mM e glicose 20 mM) foi adicionado às cubetas e o conteúdo de cada uma delas foi transferido para novos tubos, que foram incubados a 37°C por cerca de uma hora sob agitação de 200 rpm.

Após o período de incubação, as culturas bacterianas transformadas foram plaqueadas em placa de Petri, contendo meio LB sólido com 100 µg/mL ampicilina (LB+Amp) e 80 µg/mL X-gal (substrato da enzima β-galactosidase). Em seguida, as placas foram incubadas a 37°C por 16 horas.

### **3.3.1.11 Maxipreparação dos plasmídeos recombinantes**

No caso do MSP, optou-se por armazenar os clones obtidos após a transformação na forma de uma mistura de plasmídeos recombinantes.

Para avaliação da qualidade dos clones obtidos, cerca de 50 µL de células transformadas foram plaqueadas em meio contendo LB + Amp + X-gal. O restante da transformação foi inoculado em frasco tipo Erlenmeyer de 1 L contendo 250 mL de meio líquido LB+Amp e incubado a 37°C sob agitação de 200 rpm por 16 horas.

Transcorrido esse período, a cultura bacteriana foi submetida à maxipreparação para a extração dos plasmídeos, através do kit comercial *Wizard Plus Maxipreps DNA Purification System* (Promega).

### **3.3.1.12 Análise do tamanho médio dos insertos**

Após o crescimento das colônias nas placas, foi verificada a presença de colônias brancas e azuis. A confirmação da clonagem de diferentes insertos foi realizada através da seleção de colônias brancas aleatórias, que foram coletadas com auxílio de palitos de madeira estéreis, inoculadas em meio líquido LB+Amp e incubadas por aproximadamente 16 horas a 37°C, 200 rpm de agitação.

As culturas bacterianas crescidas foram submetidas à miniprep através do kit comercial *Wizard Plus SV Minipreps DNA Purification System* (Promega).

Os plasmídeos extraídos pela miniprep foram digeridos com as enzimas de restrição *EcoRI* e *HindIII*, que flanqueiam o sítio *BamHI* utilizado na clonagem, tornando possível a liberação completa do inserto clonado. Nas reações foram adicionados 3 µL de DNA (na concentração média de 55 ng/µL), 2 µL tampão 10 X (NEBuffer 2, New England Biolabs®), 10 U enzima *HindIII*, 6 U da enzima *EcoRI* (ambas da New England Biolabs®) e 16 µL de água ultrapura estéril. A reação foi incubada a 37°C por 1 hora e 30 minutos.

Os produtos das digestões enzimáticas foram avaliados por eletroforese em gel de agarose 1,0% em tampão TAE 1 X, a 90 V por 1 hora, utilizando como padrão de

peso molecular o marcador *1 kb plus DNA Ladder* (Invitrogen). As bandas foram visualizadas em transiluminador de luz UV, após tratamento com brometo de etídeo 0,25 µg/mL por 15 minutos. As imagens foram capturadas através do fotodocumentador Gel Logic 6000Pro da marca Carestream.

### **3.3.1.13 Coleta e estocagem dos clones**

As colônias brancas foram coletadas e inoculadas com auxílio do robô *Colony Picker* (Genetix Corporation) em placas de 384 poços contendo meio líquido LB+Amp. As placas foram incubadas em estufa a 37°C por aproximadamente 16 horas.

Posteriormente, com o auxílio do pipetador automatizado *epMotion® 5075 TMX systems* (Eppendorf), foi adicionado glicerol estéril em cada um dos poços, na concentração final de 20% para armazenagem dos clones em ultrafreezer a -80°C.

## **3.3.2. Prospecção de novos genes através de abordagens dirigidas por homologia de sequência nucleotídica**

### **3.3.2.1. Reação de PCR para amplificação de fragmentos gênicos de xilanases com oligonucleotídeos iniciadores degenerados**

A mistura de plasmídeos recombinantes resultantes da clonagem dos fragmentos do MSP em pUC19 foi utilizada como alvo na reação de PCR com oligonucleotídeos iniciadores degenerados para uma região conservada de genes que codificam xilanases da família GH10 (RADOMSKI et al., 1998).

As sequências dos oligonucleotídeos iniciadores utilizados são descritas abaixo:

Oligonucleotídeo iniciador direto:

5' CGS GGS CAC ACS XTS XTS TGG 3'

Oligonucleotídeo iniciador reverso:

5' GTT GTA GTC GTT GWX GXA SA 3'

Tais nucleotídeos iniciadores são degenerados, possuindo em suas sequências posições ambíguas representadas pelas letras S (C ou G), W (A ou T) e X (inosina).

Utilizou-se a técnica de PCR *touchdown*, que consiste basicamente em uma modificação da PCR convencional, com o intuito de reduzir as amplificações não específicas. Esta técnica envolve a utilização de temperaturas de anelamento superiores às temperaturas de anelamento calculadas para os oligonucleotídeos iniciadores nos primeiros dez ciclos de reação (DON et al., 1991). Posteriormente, a cada ciclo a temperatura de anelamento diminui 1°C até que a chamada temperatura *touchdown*, que corresponde à temperatura de anelamento calculada para os oligonucleotídeos iniciadores, seja atingida.

Na preparação da reação de PCR foram adicionados 34,9 µL de água ultrapura estéril, 5 µL do tampão de reação 10 X, 0,4 µL de dNTP 25 mM, 1,5 µL de MgCl<sub>2</sub> 50 mM, 3,0 µL de cada oligonucleotídeo iniciador 200 pmol/µL, 2 µL de DNA molde (40 ng/µL) e 1 U da Platinum®*Taq* DNA polimerase. A reação de PCR foi realizada em termociclador MasterCycler pro-S – Eppendorf, através dos seguintes ciclos: 60 segundos a 94°C, 10 vezes os ciclos de 30 segundos a 94°C, 60 segundos a 65°C (temperatura de anelamento) e 60 segundos a 72°C, sendo que a cada ciclo a temperatura de anelamento decrescia em um 1°C, até atingir a temperatura de 55°C. Posteriormente foram realizados 25 ciclos de 30 segundos a 94°C, 60 segundos a 55°C e 60 segundos a 72°C e um último ciclo de 10 minutos a 72°C.

O produto de PCR foi submetido à separação por eletroforese em gel de agarose 1,0% em tampão TAE 1 X, a 90 V por 1 hora, utilizando como padrão de peso molecular o marcador *1 kb plus DNA Ladder* (Invitrogen). As bandas foram visualizadas em transiluminador de luz UV, após tratamento com brometo de etídeo 0,25 µg/mL por 15 minutos. As imagens foram capturadas através do fotodocumentador Gel Logic 6000Pro da marca Carestream.

A banda amplificada foi excisada do gel e purificada com o kit *GFX<sup>tm</sup> PCR DNA and gel band purification* (GE Healthcare).

### 3.3.2.2. Clonagem dos insertos amplificados com oligonucleotídeos iniciadores degenerados

Os fragmentos purificados foram clonados através do sistema pGEM® T-Easy da Promega (Figura 7). Tal sistema consiste basicamente em um vetor com timinas nas extremidades, permitindo a clonagem de produtos de PCR amplificados com DNA polimerases que possuem atividade de desoxinucleotidil-transferase, resultando na adição de adeninas extras nas extremidades 3' dos fragmentos de DNA amplificados. Dessa forma, ocorre o pareamento das adeninas nas extremidades do inserto com as timinas nas extremidades do vetor, possibilitando a clonagem.

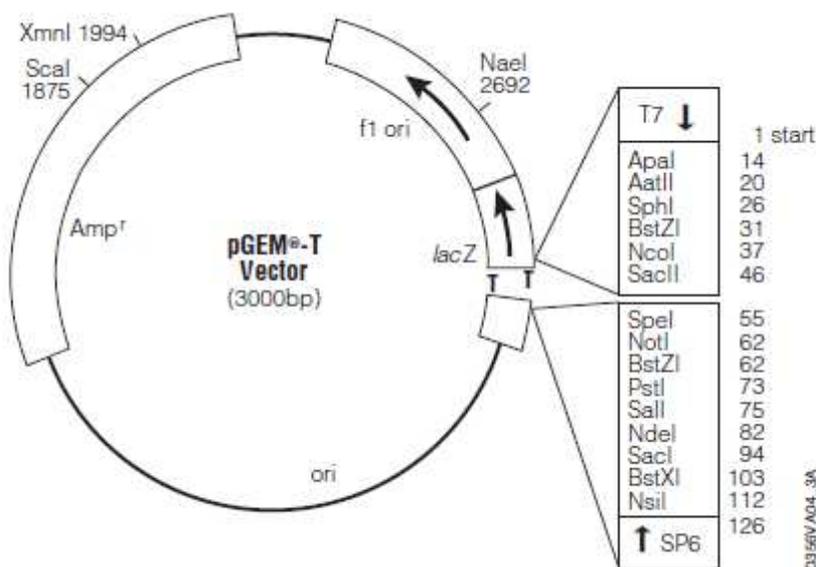


Figura 7 - Mapa do vetor pGEM® T-Easy.

A reação de ligação foi feita utilizando-se 5 µL do 2 X Rapid Ligation Buffer, 1 µL do vetor pGEM® T-Easy (50 ng/µL), 3 µL do produto de PCR e 2 U da T4 DNA Ligase. As ligações foram incubadas por 16 horas a 8°C.

Posteriormente seguiu-se a transformação, realizada como descrito no item 3.3.1.10.

### 3.3.2.3. Confirmação e sequenciamento dos fragmentos clonados em vetor pGEM® T-Easy

A coleta dos clones transformados foi realizada como descrita no tópico 3.3.1.12. Os plasmídeos isolados foram submetidos à digestão enzimática com *EcoRI* para liberação dos insertos clonados, visto que os sítios de restrição desta enzima flanqueiam ambas as extremidades do sítio de policlonagem do vetor pGEM® T-Easy.

As reações de digestão consistiram em: 1 µL de tampão 10 X (NEBuffer U, New England Biolabs®), 10 U de *EcoRI* (New England Biolabs®), 5,5 µL de água ultrapura estéril e 3 µL da solução de plasmídeos extraídos (concentração aproximada de 55 ng/µL).

Os produtos das digestões enzimáticas foram avaliados por eletroforese em gel de agarose 1,0% em tampão TAE 1 X, a 90 V por 1 hora, utilizando como padrão de peso molecular o marcador *1 kb plus DNA Ladder* (Invitrogen). As bandas foram visualizadas em transiluminador de luz UV, após tratamento com brometo de etídeo 0,25 µg/mL por 15 minutos. As imagens foram capturadas através do fotodocumentador Gel Logic 6000Pro da marca Carestream.

Os clones confirmados em relação à presença do inserto foram sequenciados com o uso dos oligonucleotídeos iniciadores M13R e M13F (M13R - 5' GCGGATAACAATTTTCACACAGG 3' / M13F - 5' GTAAAACGACGGCCAGT 3'), cujas sequências complementares flanqueiam o sítio de policlonagem do vetor pGEM® T-Easy. O sequenciamento foi realizado nas instalações do Laboratório Nacional de Biociências (LNBio).

As reações foram feitas com o reagente *Dye Terminator Cycle Sequencing Kit* (Amersham Biosciences/GE Healthcare), em reação de 10 µL, contendo 4 µL de Dye Terminator, 2 a 3 pmol do oligonucleotídeo iniciador e 300 a 600 ng de DNA plasmidial. As reações foram incubadas a 95°C por 1 minuto e submetidas a 35 ciclos de 10 segundos a 96°C, 10 segundos a 50°C e 4 minutos a 60°C em termociclador *PCR System 9700* (Perkin Elmer). As amostras foram analisadas no sequenciador automático DNA *ABI PRISM 3130xl Genetic Analyser* (Applied Biosystems).

Após o sequenciamento, os dados obtidos foram analisados pelo programa *Chromas* a fim de gerar sequências no formato FASTA. As sequências FASTA foram comparadas com o banco de dados do NCBI (*National Center for Biotechnology Information*) através da ferramenta BLASTx (disponível em: [http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx&BLAST\\_PROGRAMS=blastx&PAGE\\_TYPE=BlastSearch&SHOW\\_DEFAULTS=on&LINK\\_LOC=blasthome](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx&BLAST_PROGRAMS=blastx&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome)).

#### **3.3.2.4. Identificação de genes no metagenoma de solo de canavial**

O MSC foi submetido ao sequenciamento em larga escala através da plataforma 454 GS FLX da Roche, no Laboratório ENGECORE situado na Universidade da Carolina do Sul, nos Estados Unidos. As análises dos dados gerados, que incluíram a triagem das sequências em relação à qualidade das bases, presença de homopolímeros e sequências contaminantes (KUNIN et al., 2008a), bem como a montagem das sequências em regiões contíguas, foram realizadas através do *Newbler Assembler Software (version 2.0)* da 454 Life Sciences.

No presente trabalho, os dados obtidos a partir do sequenciamento em larga escala do MSC foram utilizados especificamente com a finalidade de identificar genes alvos para amplificação por PCR. As análises da diversidade funcional e filogenética do MSC são objetos de estudo de outros trabalhos em andamento.

A predição de genes foi feita por dois programas especialmente empregados para análises metagenômicas. Para as sequências contíguas as análises foram realizadas com o programa *Metagenemark* (ZHU et al., 2010), enquanto que para as leituras de sequenciamento, que não foram organizadas em sequências contíguas, as predições foram feitas utilizando o *FragGeneScan* (RHO et al., 2010).

Entre os genes preditos, seis foram selecionados por apresentarem identidade com genes que codificam enzimas ativas em carboidratos. As sequências de aminoácidos correspondentes aos genes selecionados foram avaliadas em relação à presença de possíveis domínios catalíticos por comparação com o banco de dados do

*Pfam* (<http://pfam.sanger.ac.uk>), que contempla uma larga coleção de famílias proteicas.

Visando a amplificação desses genes, foram desenhados oligonucleotídeos iniciadores que além da região complementar à sequência alvo continham também sítios de restrição, visando a posterior clonagem dos produtos de PCR no vetor de expressão pET-28a, e cinco nucleotídeos anteriores aos sítios para a ancoragem das enzimas de restrição. Os sítios de restrição foram escolhidos após análise na ferramenta *NEBufferCutter* (<http://tools.NEBuffer.com/NEBUFFERcutter2/>), que identifica os sítios de restrição presentes nas sequências a serem clonadas.

Os oligonucleotídeos iniciadores desenhados (Tabela 3) foram avaliados quanto à possibilidade de formação de dímeros ou grampos, teor de citosina e guanina (%CG) e temperatura de anelamento através do programa *IDT OligoAnalyzer*.

Tabela 3 - Oligonucleotídeos iniciadores desenhados para amplificação de genes alvos identificados pelo sequenciamento do MSC.

Identificação	Sequência F (5' - 3') / Temperatura de anelamento	Sequência R (5' - 3') / Temperatura de anelamento
Endoglucanase_1	TATATGCTAGCGATGCGGAGCGCAAG / 54,4°C	ATATGGATCCTTAGTAGATCGTGATGAATGCCG / 53,3°C
Esterase	TATATGCTAGCTACGTGCGGAGCCC / 54,8°C	ATATGGATCCTTAGTAGCCGCCCTCCG / 55,9°C
Beta-manosidase	TATATGCTAGCCTGCTCCGGGCCG / 56,9°C	ATATGGATCCTTAGTGCCAGACGCTCCAG / 56,2°C
Amilase	TATATGCTAGCCCCTCCAAGAAGTACCAG / 54,5°C	ATATGGATCCTTAGTACCCGGCATAGATGC / 52,7°C
Liquenase	TATATGCTAGCAAGGCCGACGGCTAC / 54,9°C	ATATGGATCCTTAGATCGTGTGGTGAAGTAGTTG / 43,6°C
Beta-glicosidase	TATATGCTAGCCTCCTGCTCTCGCAGC / 55,8°C	ATATGGATCCTTAGTACGGCGGTCCGCTC / 56,3°C

### 3.3.2.5. Reação de PCR para amplificação dos genes alvos

O DNA utilizado como alvo na reação correspondia a uma fração da amostra do MSC que foi submetida ao sequenciamento na plataforma 454.

As condições das reações foram as seguintes: 30 µL de água ultrapura estéril, 10 µL de tampão 5X Phusion HF (Finzymes®), 1 µL dNTP 10 mM, 2,5 µL de cada oligonucleotídeo iniciador 10 pmol/ µL, 2 µL de DNA (concentração aproximada de 50 ng/ µL), 1,5 µL de DMSO e 1 U da polimerase Phusion HF (Finzymes®). A reação de PCR foi realizada em termociclador MasterCycler pro-S – Eppendorf, através dos seguintes ciclos: 90 segundos a 98°C, 10 vezes os ciclos de 50 segundos a 98°C, 60 segundos na temperatura de anelamento (específica para cada gene alvo) e 30 segundos a 72°C, sendo que a cada ciclo a temperatura de anelamento decrescia em um 1°C, até que se atingisse a temperatura de *touchdown*. Posteriormente foram realizados 35 ciclos de 50 segundos a 98°C, 60 segundos na temperatura de *touchdown* e 60 segundos a 72°C, seguido por um último ciclo a 72°C por 10 minutos.

Os produtos das PCR foram avaliados por eletroforese em gel de agarose 1,0% em tampão TAE 1 X, a 90 V por 1 hora, utilizando como padrão o marcador *1 kb plus DNA Ladder* (Invitrogen). As bandas foram visualizadas em transiluminador de luz UV, após tratamento com brometo de etídeo 0,25 µg/mL por 15 minutos. As imagens foram capturadas através do fotodocumentador Gel Logic 6000Pro da marca Carestream.

### **3.3.2.6. Clonagem dos fragmentos no vetor pJET1.2/blunt**

O kit de clonagem CloneJET™ (Thermo Scientific) é indicado para clonagem de produtos de PCR com extremidades abruptas, amplificados pela ação de DNA polimerases com atividade revisora 3'-5', também conhecidas como "*proofreading enzymes*", com é o caso da enzima Phusion HF.

A clonagem ocorre por meio do vetor pJET1.2 (Figura 8). Este vetor apresenta um sistema de seleção positiva, que consiste na presença de um gene letal que é interrompido pela ligação do inserto de DNA no sítio de clonagem. Dessa forma, apenas as células transformadas com o plasmídeo recombinante conseguem se propagar.

As reações eram compostas por 2 µL de tampão de ligação 10 X, 1 U de T4 DNA ligase, vetor e inserto na proporção molar de 1:3 e água ultrapura estéril para completar o volume final de 20 µL. As reações foram incubadas a 22°C por 16 horas, seguindo a recomendação do fabricante para a obtenção do número máximo de transformantes.

A transformação foi feita por eletroporação como descrito no item 3.3.1.10, com a exceção de que neste caso não havia necessidade das placas com LB + Amp conterem X-gal.

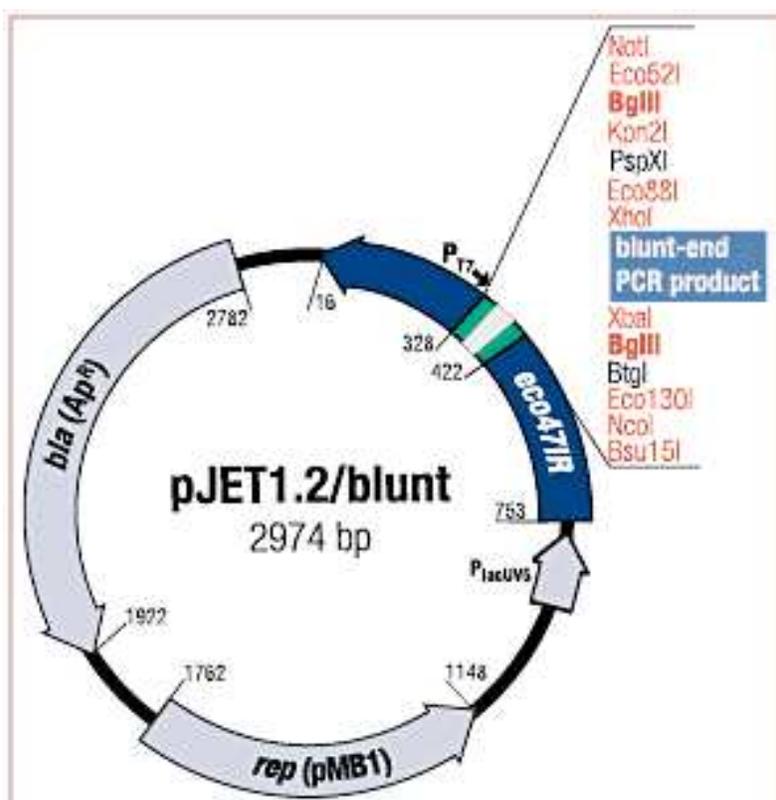


Figura 8 - Mapa do vetor pJET1.2/blunt.

Alguns clones foram selecionados aleatoriamente para extração do DNA plasmidial. Os plasmídeos isolados foram sequenciados como descrito no tópico 3.3.2.3, com os oligonucleotídeos iniciadores pJET 1.2 *Forward Sequencing Primer* (5' CGACTCACTATAGGGAGAGCGGC 3') e pJET 1.2 *Reverse Sequencing Primer* (5' AAGAACATCGATTTTCCATGGCAG 3'). A análise das sequências geradas foi feita

através da comparação com o banco de dados do NCBI, pela ferramenta BLASTx, como descrito no item 3.3.2.3.

### **3.3.3. Prospecção de novos genes através de abordagem dirigida por triagem funcional**

#### **3.3.3.1. Triagem funcional da biblioteca metagenômica**

Primeiramente, os clones estocados nas placas de 384 poços foram repicados em placas (Genetix Q-trays) de LB sólido contendo 100 µg/mL de ampicilina e 100 mM de IPTG (LB-AGAR-Amp-IPTG). Os repiques dos clones foram realizados com o auxílio do robô *Colony Picker* (Genetix Corporation).

As placas foram incubadas por cerca de 16 horas a 37°C para o crescimento e formação das colônias. Em seguida, as placas foram recobertas com 100 mL de meio contendo o substrato de interesse dissolvido em uma solução de tampão acetato de sódio 100 mM, pH 6,0, NaCl 180 mM e 1% (m/v) de ágar. Após a solidificação do meio, as placas foram incubadas a 50°C por aproximadamente 16 horas. Transcorrido o período de incubação, os clones positivos foram identificados através da visualização halos de degradação em torno das colônias.

#### **3.3.3.2. Coleta dos clones positivos identificados**

Os clones positivos foram coletados com o auxílio de palitos de madeira estéreis diretamente das placas nas quais foram identificados e inoculados em tubos de ensaios estéreis contendo meio LB+Amp. Em seguida os tubos foram incubados a 37°C por aproximadamente 16 horas. As culturas crescidas foram submetidas à miniprep para isolamento dos plasmídeos.

### **3.3.3.3. Confirmação da atividade enzimática dos clones positivos**

Como forma de confirmar que a atividade enzimática apresentada pelos clones positivos era conferida por um gene localizado no inserto de DNA metagenômico clonado, realizou-se a retransformação dos plasmídeos isolados dos clones positivos.

Em alguns casos, os clones obtidos na retransformação foram incubados na presença de determinado substrato para a confirmação da formação dos halos de hidrólise. Alternativamente, os clones foram inoculados em tubos de ensaio contendo meio LB+Amp e incubados a 37°C por aproximadamente 16 horas. Posteriormente, as culturas foram centrifugadas por 15 minutos a 10.000 x g. O sobrenadante foi coletado e utilizado para confirmação da atividade enzimática.

As reações foram feitas em um volume final de 100 µL, sendo 40 µL de tampão acetato de sódio 100 mM, pH 6,0, 50 µL de substrato 0,5% (m/v) dissolvido em água ultrapura estéril e 10 µL do sobrenadante proveniente das culturas bacterianas. As reações correspondentes aos controles negativos correspondiam à reação descrita acima exceto pela não adição do sobrenadante, substituído por 10 µL de água ultrapura estéril. Os controles negativos tiveram a função de confirmar que a hidrólise do substrato ocorria exclusivamente pela ação das enzimas presentes nos sobrenadantes das culturas bacterianas.

Os ensaios foram preparados em placas de PCR de 96 poços no gelo, posteriormente incubadas em termociclador (MasterCycler pro-S – Eppendorf) por 20 minutos a 40°C. Transcorrido esse período, a cada reação foram adicionados 100 µL de ADNS (ácido 3,5-dinitrosalicílico) (MILLER, 1959), responsável pela revelação dos açúcares redutores liberados durante a reação, através do aparecimento de coloração alaranjada. Nas reações controle, após a adição do ADNS, foram ainda adicionados 10 µL do sobrenadante, como forma de confirmar a ausência de açúcares redutores nos sobrenadantes. Nesses casos, não ocorria atividade enzimática pelo fato da enzima ser rapidamente inativada pelo contato com o ADNS.

#### **3.3.3.4. Avaliação da atividade enzimática por eletroforese capilar**

Os carboidratos utilizados como substratos nas reações foram previamente marcados com o fluoróforo APTS (ácido 1,3,6-trissulfônico- 8-aminopireno), que se liga à extremidade redutora dos açúcares, por meio de aminação redutiva (EVANGELISTA et al., 1995). A presença do APTS pode dificultar a atividade de algumas enzimas que atacam as cadeias de oligossacarídeos pela extremidade redutora. Por isso, alternativamente, os produtos da hidrólise enzimática em substratos não marcados foram submetidos à marcação com APTS.

A eletroforese capilar dos oligossacarídeos foi realizada em um sistema P/ACE MDQ (Beckman Coulter) com detector de fluorescência induzida por laser. Um capilar de sílica fundida (TSP050375, Polymicro Technologies) com diâmetro interno de 50 µm e 31 cm de comprimento foi usado como coluna de separação dos oligossacarídeos. As amostras foram injetadas através da aplicação de uma pressão de 0,5 psi por 0,5 segundo. A corrida foi realizada em tampão fosfato de sódio 40 mM pH 2,5 e as condições de corrida foram 15 kV / 70-100 µA, com o cátodo na entrada, em temperatura controlada de 20°C. Entre as corridas, o capilar foi lavado com NaOH 0,1 M seguido de lavagem com tampão para evitar a ocorrência de contaminação cruzada. Os oligômeros marcados com APTS foram excitados a 488 nm e a emissão foi coletada em filtro com banda de 520 nm.

A identificação dos produtos de hidrólise foi possível pela comparação dos tempos de retenção de cada pico gerado com os tempos de retenção dos padrões de carboidratos.

#### **3.3.3.5. Mapeamento das ORFs nos clones positivos identificados**

O primeiro passo para identificação das ORFs consistiu na determinação do tamanho do inserto de cada um dos clones positivos. Isso foi possível através da digestão enzimática dos plasmídeos isolados com as enzimas *EcoRI* e *HindIII*. Nestas

reações foram adicionados 3 µL de DNA (com uma concentração média de 70 ng/µL), 2 µL tampão 10 X (NEBuffer 2, New England Biolabs®), 10 U de *Hind*III, 6 U de *Eco*RI (ambas da New England Biolabs®) e 16 µL de água ultrapura estéril. As reações foram incubadas a 37°C por 2 horas com subsequente inativação a 65°C por 20 minutos.

O resultado da digestão foi analisado através de eletroforese em gel de agarose 1,0% em tampão TAE 1,0X a 90 V por 1 hora, utilizando como padrão de peso molecular o marcador *1Kb plus DNA Ladder* (Invitrogen) e alternativamente o marcador *GeneRuler 1 kb DNA Ladder* (Fermentas). As bandas foram visualizadas em transiluminador de luz UV, após tratamento com brometo de etídeo 0,25 µg/mL por 15 minutos e as imagens capturadas através do fotodocumentador Gel Logic 6000Pro da marca Carestream.

Em seguida, o sequenciamento dos clones foi realizado no sequenciador Applied Biosystems 3500xL Genetic Analyzer, localizado no Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE). As reações foram feitas com reagente *BigDye Terminator Cycle Sequencing Kit* (Applied Biosystems), em uma reação de 10 µL, contendo 4 µL de Dye Terminator, 2 a 3 pmol do oligonucleotídeo iniciador e 300 a 600 ng de DNA plasmidial. As reações foram incubadas a 95°C por um minuto e submetidas a 35 ciclos de 10 segundos a 96°C, 10 segundos a 50°C e 4 minutos a 60°C em termociclador *PCR System 9700* (Perkin Elmer).

As primeiras etapas de sequenciamento foram realizadas com os oligonucleotídeos iniciadores M13F e M13R, resultando no sequenciamento das extremidades dos insertos. Devido ao fato do sequenciamento pelo método de Sanger gerar leituras de sequenciamento com tamanho aproximado entre 800-900 pares de base, a obtenção das seqüências completas dos genes de interesse foi realizada através da construção de novos oligonucleotídeos iniciadores, desenhados com base nas primeiras seqüências geradas, visando ampliar o sequenciamento para as regiões internas dos insertos.

No caso de um clone específico, que apresentou um inserto de aproximadamente 7 kb, optou-se pela construção de uma sub-biblioteca em pUC19. Para isso, o plasmídeo isolado do clone positivo foi parcialmente digerido com a enzima de restrição *Sau*3AI (como descrito no item 3.3.1.5) para a obtenção de fragmentos na

faixa de 1 a 2 kb. Os testes foram realizados a partir da seguinte reação: 2 µL de DNA plasmidial isolado do clone de interesse (concentração aproximada de 200 ng/µL), 1 µL tampão 10 X (NEBuffer 1, NewEngland Biolabs ®), 0,01 U da enzima *Sau3AI* (NewEngland Biolabs ®), 0,1 µL de BSA 100 X e água ultrapura estéril para completar o volume final de 10 µL. Os tempos de incubação avaliados foram 30, 60, 90 e 120 minutos a 37°C. Posteriormente, as reações foram inativadas a 65°C por 20 minutos.

Após a determinação da condição ideal de digestão, a reação foi ampliada mantendo as proporções da reação descrita acima. O produto da reação de digestão foi submetido à eletroforese em gel de agarose 1,0% e os fragmentos na faixa de tamanho de interesse foram excisados do gel e purificados. A reação de ligação e a transformação foram realizadas de acordo com o descrito no tópico 3.3.1.8 e 3.3.1.10, respectivamente.

Os clones obtidos na transformação foram submetidos à triagem funcional no substrato de interesse. Os clones positivos identificados foram inoculados em meio líquido LB + Amp, para o isolamento dos plasmídeos por miniprep. Em seguida os plasmídeos foram submetidos à digestão enzimática com *EcoRI* e *HindIII* (segundo a reação descrita acima) e sequenciados.

Todas as sequências geradas foram reunidas em regiões contíguas com auxílio do programa *Geneious Pro 4.8.5*. A partir das sequências consenso geradas foram identificadas as ORFs presentes em cada uma das seis fases de leituras. As ORFs foram avaliadas individualmente por comparação com o banco de dados do NCBI através da ferramenta BLASTx, que compara as sequências de nucleotídeos traduzidas em aminoácidos com o banco de proteínas do NCBI. As ORFs também foram avaliadas em relação à presença de domínios conservados, através da ferramenta BLASTp que realiza a comparação das sequências aminoácidos com o banco de dados de domínios conservados, o CDD, (disponível em: <http://www.ncbi.nlm.nih.gov/cdd>). O CDD reúne informações de vários outros bancos de dados, como o Pfam, SMART (disponível em <http://smart.embl-heidelberg.de/e>) e COG (disponível em <http://www.ncbi.nlm.nih.gov/COG/>).

As sequências das ORFs traduzidas em aminoácidos que codificavam as enzimas de interesse foram ainda comparadas com o banco de dados do PDB (*Protein*

*Data Bank*), que reúne informações sobre as estruturas tridimensionais de moléculas biológicas. Esta etapa teve como objetivo, a identificação de proteínas com estruturas tridimensionais resolvidas e que apresentaram homologia de sequência com as enzimas identificadas.

As ORFs identificadas foram ainda avaliadas em relação à presença de peptídeo sinal, o qual sinaliza que a proteína deve ser exportada para uma região específica, onde irá exercer sua função. As análises em relação à presença de peptídeo sinal foram feitas através da ferramenta Signal-BLAST (disponível em: <http://sigpep.services.came.sbg.ac.at/signalblast.html>).

### **3.4 Resultados e Discussão**

#### **3.4.1. Construção das bibliotecas metagenômicas**

A extração de DNA metagenômico de alto peso molecular e com alto grau de pureza é importante para a construção de bibliotecas metagenômicas de qualidade (DANIEL, 2005). Como se pode observar na Figura 9, as extrações realizadas através do kit comercial FastDNA® SPIN Kit for soil (MP - Biomedicals) e do homogeneizador FastPrep® Instrument (MP - Biomedicals) resultaram na obtenção de amostras de DNA metagenômico íntegras e com alto peso molecular.

No caso do MCB e do MSP, que foram utilizados para a construção das bibliotecas metagenômicas, foi necessária a determinação do tempo de digestão enzimática para a obtenção de fragmentos com tamanho entre 2 e 5 kb. Esta faixa de tamanho está de acordo com o tipo de vetor escolhido (plasmídeo pUC19) que apresenta boa estabilidade para insertos de até 5 kb. O plasmídeo pUC19 apresenta protocolo simples de manipulação e por ser de alta cópia, facilita a detecção da atividade enzimática de genes que são pouco expressos (DANIEL, 2005).

A Figura 10 apresenta o perfil de digestão obtido para cada uma das amostras. A digestão do DNA metagenômico em fragmentos na faixa de 2 a 5 kb foi observada após 20 e 60 minutos de incubação, para o MCB e MSP, respectivamente.

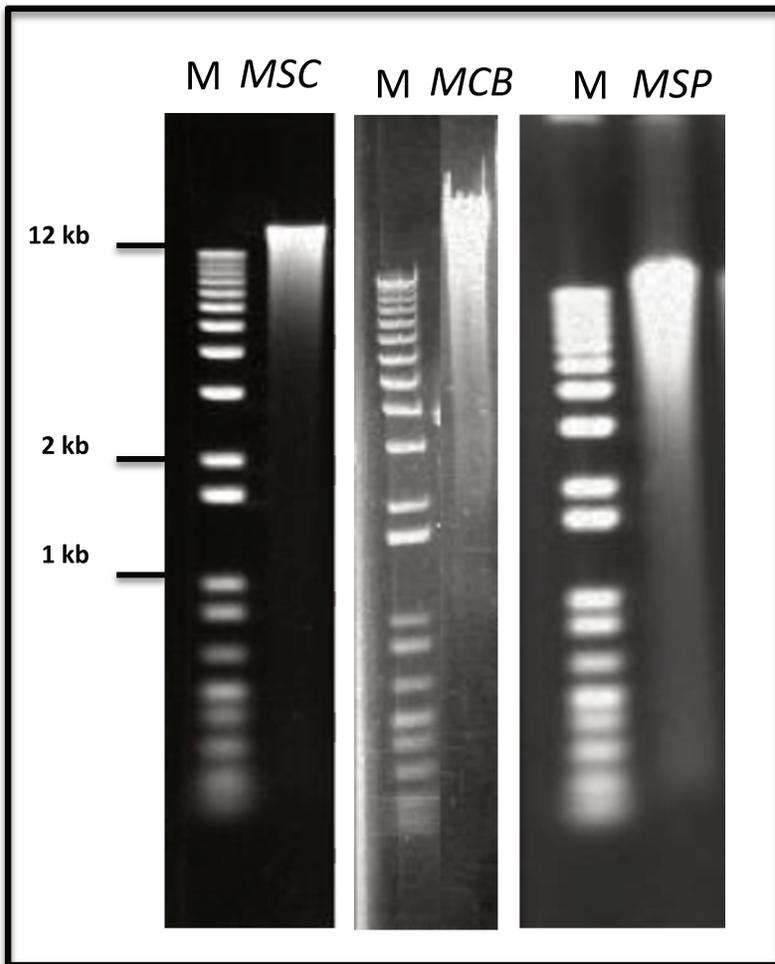


Figura 9 – Eletroforese em gel de agarose 1,0% das amostras de DNA metagenômico extraídas de solo coletado no município de São Carlos (MSC), do consórcio microbiano degradador de bagaço de cana-de-açúcar (MCB) e do solo coletado no município de Paulínia (MSP). M – marcador de peso molecular 1 kb.

Tendo se determinado as melhores condições de digestão, maiores quantidades de DNA foram digeridas. As etapas seguintes de excisão e purificação dos fragmentos digeridos foram bem sucedidas, como pode ser observado na Figura 11.

A partir das amostras de DNA metagenômico digeridas e purificadas foram construídas bibliotecas metagenômicas no vetor pUC19, denominadas:

- **Biblioteca metagenômica derivada do solo de Paulínia (BP):** Construída com DNA metagenômico extraído diretamente do solo de canavial de Paulínia (MSP), armazenada em ultrafreezer a  $-80^{\circ}\text{C}$  na forma de uma mistura de plasmídeos

recombinantes. Foi empregada para a prospecção de genes através de abordagem dirigida por homologia de sequência nucleotídica;

- **Biblioteca metagenômica derivada do consórcio microbiano (BCB):** Construída com o DNA metagenômico extraído do consórcio microbiano degradador de bagaço de cana-de-açúcar (MCB), estocada em ultrafreezer a  $-80\text{ }^{\circ}\text{C}$  na forma de clones isolados em placas de 384 poços. Foi empregada na prospecção de genes através da abordagem dirigida por triagem funcional.

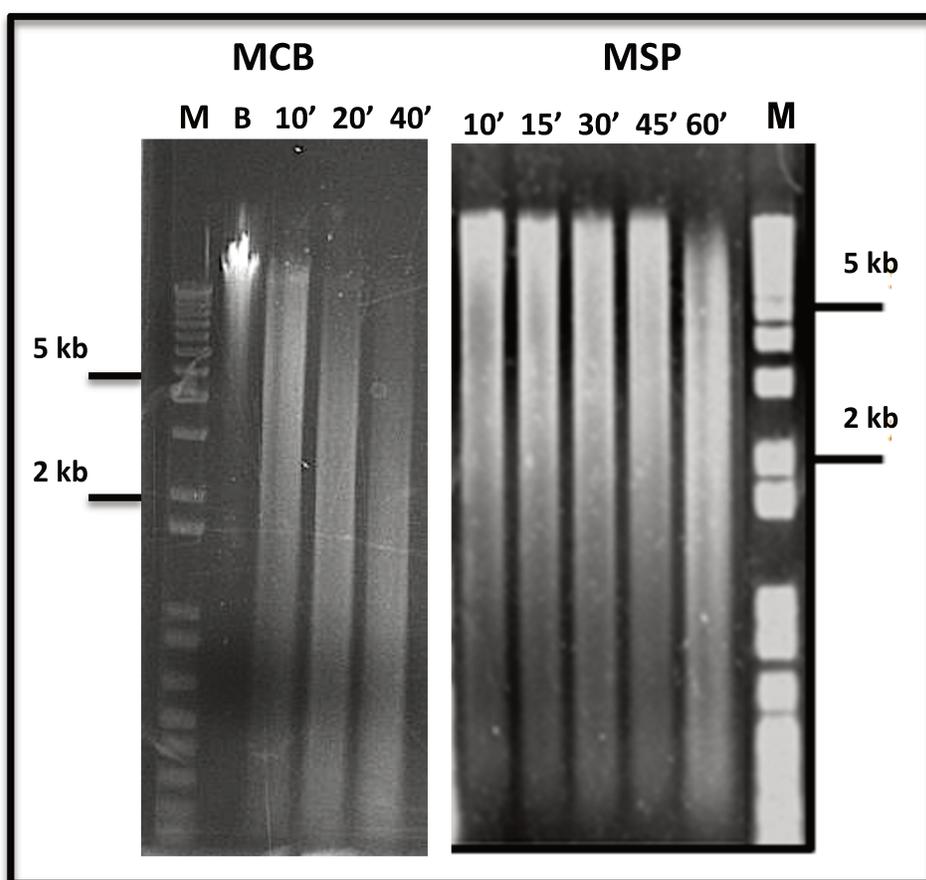


Figura 10 - Eletroforese em gel de agarose 1,0% dos produtos dos testes de digestões enzimáticas com *Sau3AI* de amostras do MCB (por 10, 20 e 40 minutos) e do MSP (por 10, 15, 30, 45 e 60 minutos). M = marcador de peso molecular 1 kb, B = amostra de DNA metagenômico controle (não submetida à reação de digestão).

No processo de clonagem em pUC19, a inserção dos fragmentos de DNA ocorreu na região do gene *lacZ*, responsável pela síntese da enzima  $\beta$ -galactosidase.

Esta enzima atua na hidrólise do substrato X-gal que resulta no aparecimento de coloração azulada nas colônias bacterianas. Sendo assim, as colônias brancas correspondiam aos clones que receberam o inserto e conseqüentemente apresentavam o gene da  $\beta$ -galactosidade truncado, enquanto as colônias azuis correspondiam aos clones sem inserto, e, portanto com a  $\beta$ -galactosidade ativa (LANGLEY et al., 1975).

Visando caracterizar o tamanho médio dos insertos clonados e determinar a qualidade e representatividade das bibliotecas construídas, colônias brancas aleatórias coletadas das duas bibliotecas foram submetidas à digestão enzimática, como pode ser observado na Figura 12.

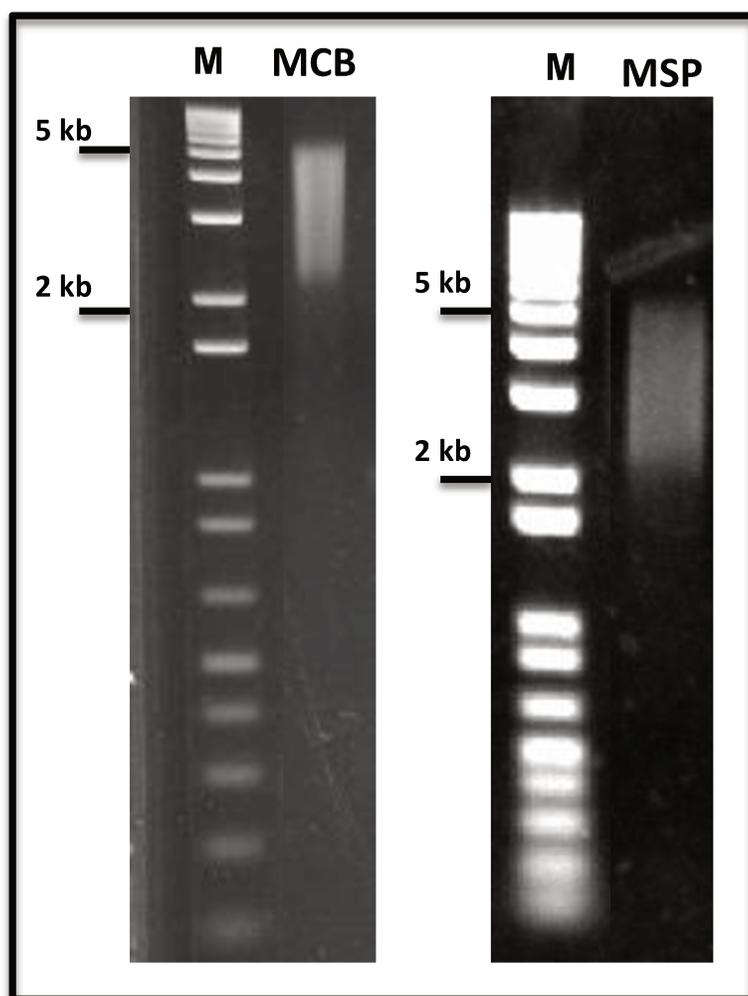


Figura 11 - Eletroforese em gel de agarose 1,0% referente às purificações das amostras de DNA do MCB e MSP digeridas com *Sau3A*I. M - marcador de peso molecular 1 kb.

Em relação à BP, a análise do perfil de digestão do DNA plasmidial demonstrou que todos os clones coletados aleatoriamente continham insertos com tamanho médio de 2,3 kb (com insertos na faixa de 1,2 a 3,1 kb). A eficiência máxima de transformação foi de aproximadamente  $10^7$  clones por  $\mu\text{g}$  de DNA transformado. Considerando a transformação do volume total da ligação (10  $\mu\text{L}$ ), estimou-se um valor total de aproximadamente 67 mil clones. Considerando esta estimativa, o tamanho médio dos insertos e a porcentagem de clones com inserto, calculou-se que a BP apresentou cobertura de 154,1 Mb.

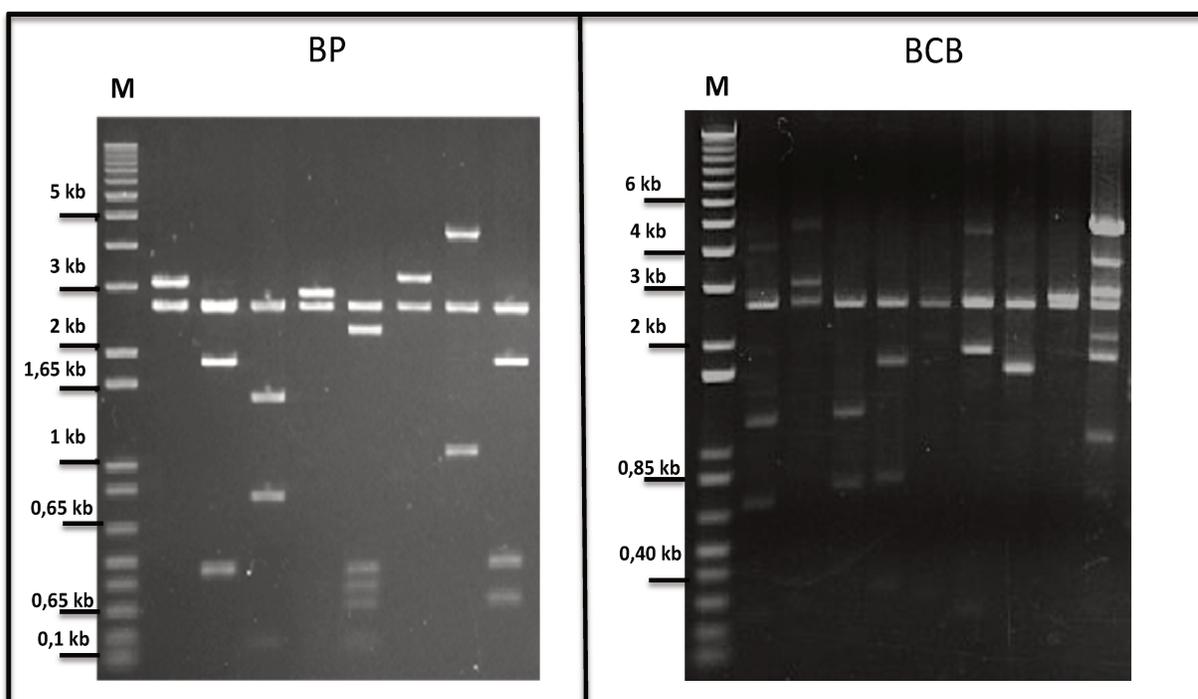


Figura 12 - Eletroforese em gel de agarose 1,0%. Padrões de restrições dos plasmídeos isolados de clones aleatórios da BP e BCB. M – Marcador de peso molecular 1 kb.

A média de tamanho dos insertos clonados na BCB foi de aproximadamente 3,5 kb (com insertos na faixa de 2 a 8 kb). Dentre os oito clones coletados (colônias brancas) e submetidos à digestão enzimática, apenas um apresentava vetor vazio, demonstrando que 87,5% das colônias brancas avaliadas continham insertos. A eficiência máxima de transformação obtida foi de  $10^7$  clones por  $\mu\text{g}$  de DNA

transformado. Como resultado da transformação de 4 µL do produto da reação de ligação foram obtidas cerca de 30.000 colônias brancas, sendo que aproximadamente 26.900 delas foram isoladas com o auxílio do robô *Colony Picker* e estocadas em placas de 384 poços. O cálculo da cobertura resultou em um valor aproximado de 82,4 Mb.

As características das bibliotecas metagenômicas construídas estão resumidas na Tabela 4.

Tabela 4 – Descrição das bibliotecas metagenômicas BP e BCB.

	Tipo de solo	Município de Origem	Consórcio microbiano	Forma de estoque	Total de clones	Média de inserto	Extensão do tamanho dos insertos	Estimativa da porcentagem de clones com inserto	Estimativa de cobertura da biblioteca
<b>BP</b>	Canavial	Paulínia - SP	Não	Mistura de plasmídeos transformados	67 mil*	2,3 kb	1,2-3,1 kb	100%	54,1 Mb
<b>BCB</b>	Canavial	São Carlos - SP	Sim	Clones isolados em placas de 384 poços	26.900	3,5 kb	2,0-8,0 kb	87,5%	82,4 Mb

\* valor do número de clones estimados considerando a eficiência de transformação de  $10^7$  clones por µg de DNA transformado.

### 3.4.2. Prospecção de novos genes através da abordagem dirigida por homologia de sequência nucleotídica

#### 3.4.2.1. Reação de PCR para amplificação de fragmentos gênicos de xilanases com oligonucleotídeos iniciadores degenerados

A reação de PCR com o par de oligonucleotídeos iniciadores degenerados para amplificação de regiões conservadas de xilanases, utilizando como alvo a BP (armazenada na forma de uma mistura de plasmídeos recombinantes) resultou na

amplificação de produtos de PCR com tamanho esperado de 0,3 kb (Figura 13).

O sequenciamento dos plasmídeos isolados de doze clones aleatórios, resultantes da clonagem do produto de PCR purificado permitiu a identificação de nove sequências nucleotídicas distintas apresentando identidade menor que 85% com genes de xilanases conhecidos (Tabela 5).

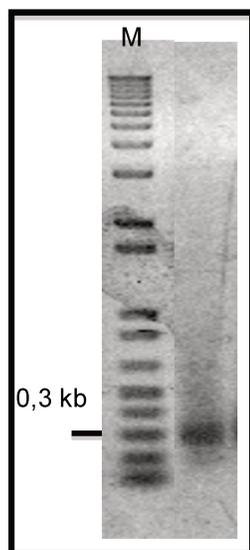


Figura 13 – Eletroforese em gel de agarose 1,0% dos produtos de PCR com o par de oligonucleotídeos iniciadores degenerados. M – Marcador de peso molecular 1 kb.

As sequências de aminoácidos correspondentes aos fragmentos gênicos amplificados foram alinhadas com sequências de xilanases conhecidas (Figura 14). Pôde-se observar que todas as sequências analisadas apresentaram resíduos de aminoácidos conservados em posições específicas (destacados em preto e cinza) e considerável grau de variabilidade em outras regiões.

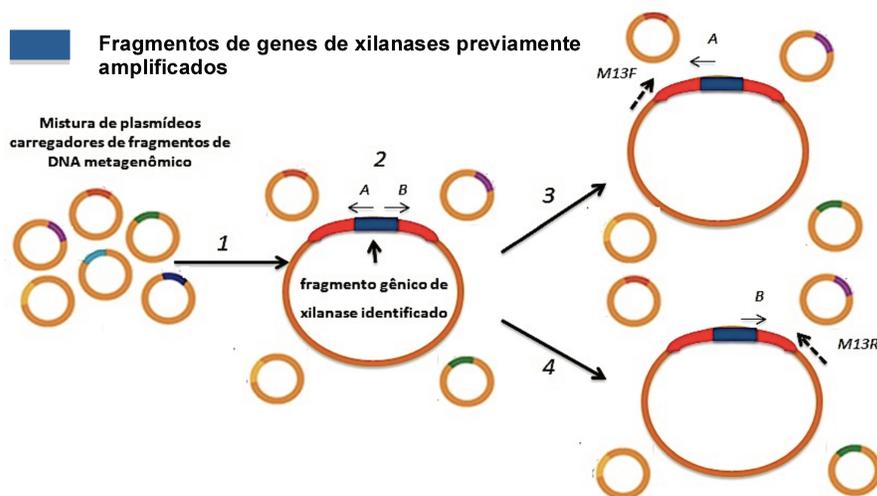
A próxima etapa visou à obtenção das sequências completas dos genes. Para isso foi desenvolvida uma estratégia baseada na técnica de *genome walking* que consiste na identificação de sequências nucleotídicas desconhecidas a partir de trechos conhecidos (ACEVEDO et al., 2008). A Figura 15 descreve a estratégia empregada, que consistiu no desenho de novos oligonucleotídeos iniciadores baseados nas sequências de ambas as extremidades dos fragmentos de xilanases previamente

amplificados. Esses oligonucleotídeos foram então combinados em reações de PCR com oligonucleotídeos iniciadores do vetor, M13F e M13R, visando à amplificação das sequências desconhecidas dos genes de interesse. Posteriormente, as sequências obtidas seriam alinhadas para formação de sequências contíguas para a busca de possíveis ORFs.

Tabela 5 – Comparação das sequências correspondentes aos fragmentos amplificados com o par de oligonucleotídeos degenerados com o banco de dados do NCBI pela ferramenta BLASTx.

<b>Amostra</b>	<b>Tamanho (kb)</b>	<b>Proteína / Bactéria relacionada</b>	<b>Acesso GenBank</b>	<b>E-value</b>	<b>Identidade (%)</b>
<b>Xil_F</b>	0,3	Provável xilanase (Organismo não cultivável)	AAP87472	1,16E-35	66
<b>Xil_G</b>	0,28	Glicosil hidrolase família 10 ( <i>Ktedonobacter racemifer</i> DSM 44963)	ZP_06970944	1,00E-39	76
<b>Xil_J</b>	0,31	Provável xilanase (Organismo não cultivável)	AAP87504	3,20E-38	70
<b>Xil_M</b>	0,29	Provável xilanase (Organismo não cultivável)	AAP87451	2,25E-23	79
<b>Xil_3</b>	0,27	Glicosil hidrolase família 10 ( <i>Verrucomicrobiae bacterium</i> DG1235)	ZP_05058163	7,20E-22	57
<b>Xil_9</b>	0,28	Provável xilanase (Organismo não cultivável)	AAP87452	3,78E-39	77
<b>Xil_10</b>	0,28	Enzima bifuncional xilanase-arabinofuranosidase ( <i>Streptomyces chattanoogensis</i> )	AAD32559.2	3,00E-36	78
<b>Xil_11</b>	0,3	Provável xilanase (Organismo não cultivável)	AAP87502	7,00E-47	84
<b>Xil_12</b>	0,3	Glicosil hidrolase família 10 ( <i>Streptomyces sviceps</i> ATCC 29083)	ZP_06914483	5,98E-32	77





#### Etapas

1 = A PCR com oligonucleotídeos iniciadores degenerados resultou na identificação de fragmentos gênicos com homologia para xilanases

2 = Desenho de oligonucleotídeos iniciadores específicos (A e B), em ambas extremidades, para os fragmentos gênicos identificados

3= PCR com os oligonucleotídeos iniciadores M13F e A

4= PCR com os oligonucleotídeos iniciadores M13R e B

Figura 15 - Estratégia empregada na tentativa de amplificação das sequências completas dos genes de xilanases identificados por PCR com o par oligonucleotídeos iniciadores degenerados.

Porém, através do sequenciamento dos fragmentos obtidos nas reações de PCR com os oligonucleotídeos iniciadores M13F e A e M13R e B foi verificado que as sequências nucleotídicas não apresentavam homologia com genes de xilanases. Alguns trabalhos reportaram a dificuldade de se empregar métodos baseados em PCR para amplificação de alvos específicos diretamente de amostras de DNA metagenômico (GONZALEZ et al., 2012, ROSEN et al., 2012, UCHIYAMA & WATANABE, 2006, ACEVEDO et al., 2008). Esta dificuldade pode estar relacionada ao fato das sequências de interesse representarem uma fração muito pequena na amostra alvo (GONZALEZ et al., 2012). Além disso, pelo fato das amostras de DNA metagenômico apresentarem elevada diversidade de sequências nucleotídicas, os oligonucleotídeos iniciadores podem anelar em segmentos de DNA que não correspondem à sequência alvo e gerarem produtos de ampliações não específicas (ACEVEDO et al., 2008).

### 3.4.2.2. Amplificação de genes alvos

O pirosequenciamento do MSC na plataforma 454 GS FLX da Roche resultou no sequenciamento de um total de 241.632.546 pares de base organizadas em 576.702 leituras de sequenciamento com tamanho médio de 425 pares de base, sendo que a menor e maior leitura de sequenciamento apresentaram 40 e 768 pares de base, respectivamente. Aproximadamente 1% das leituras de sequenciamento foi organizada em sequências contíguas, com tamanho variando entre 100 e 1.392 pares de base e tamanho médio de 599 pares de base.

Através de ferramentas de bioinformática foi realizada a predição de genes que codificavam enzimas ativas em carboidratos nas leituras de sequenciamento e sequências contíguas obtidas através do pirosequenciamento. Após a predição, seis genes foram selecionados para o desenho de oligonucleotídeos iniciadores visando à amplificação de tais genes por reações de PCR utilizando como alvo o MSC. A Tabela 6 apresenta as características dos genes selecionados bem como os respectivos domínios conservados identificados por comparação com o banco de dados do Pfam. A cor verde do modelo esquemático indica que a identidade com domínios conservados conhecidos foi significativa. Em rosa estão destacados os sítios ativos preditos.

As reações de PCR funcionaram apenas para os seguintes alvos: endoglucanase\_1, liquenase e beta-manosidase (Figura 16). No caso da endoglucanase\_1 o produto de PCR apresentou várias bandas, com tamanhos em torno de 0,65, 0,5 e 0,3 kb. Em relação à liquenase, o produto de PCR também era composto por várias bandas, porém a principal delas apresentou o tamanho esperado de aproximadamente 0,3 kb. Em relação à beta-manosidase, foram amplificadas duas bandas com tamanhos aproximados de 0,3 e 0,5 kb. As bandas com tamanho próximo ao esperado (indicadas pelas setas na Figura 16) foram excisadas do gel, purificadas, clonadas no vetor pJET1.2 e submetidas ao sequenciamento.

Como pode ser observado na Tabela 7, a partir do sequenciamento, verificou-se que os fragmentos amplificados não apresentaram identidade com os genes esperados. Esse resultado evidenciou mais uma vez a dificuldade em se resgatar sequências nucleotídicas específicas em uma amostra complexa, como o MSC.

Tabela 6 - Genes alvos escolhidos para amplificação por PCR. A cor verde indica que a identidade com domínios conservados conhecidos foi significativa. Em rosa estão destacados os sítios ativos preditos.

Gene alvo	BLASTx / Número de acesso NCBI	Número de aminoácidos	Família do Pfam	Identificação do Pfam	E - Value	Domínios/ sítios ativos identificados
<b>Endoglucanase_1</b>	Endoglucanase/ YP_005147986	132	Celulase	PF00150	1,1E-15	
<b>Esterase</b>	Provável esterase/ WP_007573064	97	Esterase	PF00756	1,6E-5	
<b>Beta-manosidase</b>	Beta- manosidase/ WP_002628632	156	Glicosil Hidrolase Família 2	PF02836	0,00018	
<b>Amilase</b>	Provável glucano hidrolase/ WP_010850249	151	Alfa-amilase	PF00128	0,00032	
<b>Beta-glicosidase</b>	Glicosil hidrolase/ WP_008512821	146	Glicosil Hidrolase Família 1	PF00232	5,7E-23	
<b>Liquenase</b>	Endo-beta-1,3- glucanase/ WP_006124082	104	Glicosil Hidrolase Família 16	PF00722	1,7E-24	

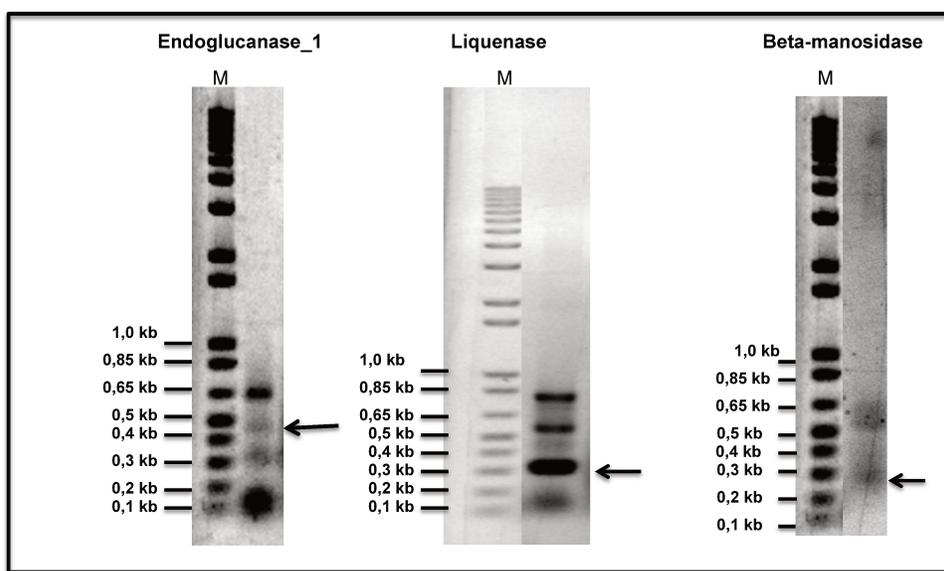


Figura 16 - Eletroforese em gel de agarose 1,0% dos produtos das reações de PCR visando à amplificação de genes alvos (as setas indicam as bandas que foram excisadas do gel, purificadas e clonadas no vetor pJET1.2). M – marcador de peso molecular 1 kb.

Tabela 7 – Comparação das sequências correspondentes aos produtos de PCR clonados no vetor pJET1.2 com o banco de dados do NCBI pela ferramenta BLASTx.

<i>Amostra</i>	<i>Proteína / Bactéria relacionada</i>	<i>Acesso GenBank</i>	<i>E-value</i>	<i>Identidade (%)</i>
<i>Endoglucanase_1</i>	RNA helicase/ <i>Escherichia coli</i>	WP_001419243	2,0E-104	100
<i>Beta-manosidase</i>	Helicase/ <i>Shigella flexneri</i>	WP_005117492	2,0E-54	100
<i>Liquease</i>	Nenhuma similaridade identificada	-	-	-

### 3.4.3. Prospecção de novos genes através da abordagem dirigida por triagem funcional

Visando aumentar as chances da identificação de genes que codificavam para enzimas ativas em carboidratos que compõem a parede celular vegetal, a biblioteca

construída com o DNA metagenômico extraído do consórcio microbiano enriquecido com bagaço de cana-de-açúcar (BCB) foi empregada nos experimentos de triagem funcional.

Os substratos utilizados nesta etapa do trabalho foram escolhidos por apresentarem características que permitiam a identificação de clones capazes de expressar atividades enzimáticas de interesse. A Tabela 8 apresenta a relação dos substratos selecionados bem como o modo de detecção dos clones positivos.

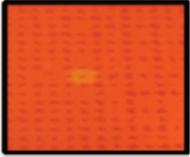
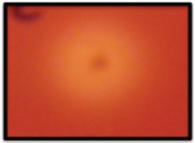
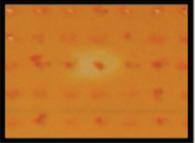
Tabela 8 – Substratos utilizados na triagem funcional da biblioteca metagenômica BCB.

Enzima alvo	Substrato / Concentração de uso	Modo de revelação	Referência
<b>Celulase</b>	CMC (Sigma) / 1,0%	Vermelho congo (Sigma)	KIM et al., 2000
<b>Xilanase</b>	Xilana de madeira de Faia (Sigma) / 0,5%	Vermelho congo (Sigma)	HAARHOFF et al., 1999
<b>Pectinase</b>	Pectina cítrica (Sigma) / 1,0%	Vermelho congo (Sigma)	KAUR et al., 2011
<b>Pectato Liase</b>	Azo-ramnogalacturonana (Megazyme) / 0,1%	Formação de halo de degradação ao redor da colônia	SOLBAK et al., 2005
<b>Amilase</b>	Amido de batata (Sigma) / 2,0%	Lugol (Sigma)	YUN et al., 2004
<b><math>\beta</math>-glucosidase</b>	Esculina (Sigma) / 0,1%	Formação de halo de degradação ao redor da colônia	FANG et al., 2010
<b><math>\beta</math>-glucosidase</b>	pnpG (pNP- $\beta$ -D-glucopiranosídeo) (Sigma) / 0,5%	Formação de halo de degradação ao redor da colônia	DEL POZO et al., 2012
<b>Mananase</b>	Manana de goma de alfarroba (Sigma) / 0,1%	Vermelho congo (Sigma)	STALBRAND et al., 1995
<b>Mananase</b>	Azo-galactomanana de alfarroba (Megazyme) / 0,2%	Formação de halo de degradação ao redor da colônia	BRAITHWAITE et al., 1995
<b>Esterase</b>	Alfa-Naftil-acetato (Sigma) / 80 $\mu$ g/mL	Corante Fast Blue RR	KIM et al., 2006
<b>Esterase</b>	Tween 80 (Sigma) / 1,0%	CaCl <sub>2</sub> .2H <sub>2</sub> O	PANCHAL & INGLE, 2011
<b>Lacase</b>	ABTS (Ácido 2,2"-Azino-bis 3-etilbenzotiazolino-6-sulfônico) (Sigma) / 0,5 mg/L	Formação de halo de degradação ao redor da colônia	PRIYADARSINI et al., 2011
<b><math>\beta</math>-Xilosidase</b>	MUX (4-O-metil umbeliferil- $\beta$ -D-xilopiranosídeo) (Sigma) / 10 mg/L	Formação de halo de degradação ao redor da colônia (visível sob luz UV)	UNGCHAITHAM et al., 2001

A triagem funcional dos 26.900 clones da BCB nos diferentes substratos resultou na identificação de dez clones positivos, porém apenas cinco (Tabela 9) foram confirmados em relação à atividade enzimática esperada. Os demais correspondiam a clones falsos positivos, cujo aparecimento pode ser atribuído a ocorrência de contaminações nas placas, em virtude da extensa manipulação a qual foram

submetidas.

Tabela 9 – Clones identificados através da triagem funcional da biblioteca metagenômica BCB.

Substrato	Identificação do clone	Clone positivo identificado
<i>CMC</i>	<i>E-1</i>	
<i>CMC</i>	<i>E-2</i>	
<i>CMC</i>	<i>E-3</i>	
<i>Xilana de madeira de faia</i>	<i>X-1</i>	
<i>Xilana de madeira de faia</i>	<i>X-2</i>	

A confirmação da atividade enzimática do clone E-1, foi possível através de ensaios com o sobrenadante da cultura em  $\beta$ -glucano e CMC (Figura 17A, na qual o aparecimento da coloração alaranjada indicou a ocorrência de reação enzimática). Além disso, a análise por eletroforese capilar dos produtos da hidrólise gerados pela ação do sobrenadante de E-1 em celopentose marcada com APTS, revelou um típico padrão de endo- $\beta$ -1,4-glucanase ao atuar na quebra das ligações internas e gerar celotriose como principal produto (Figura 17B).

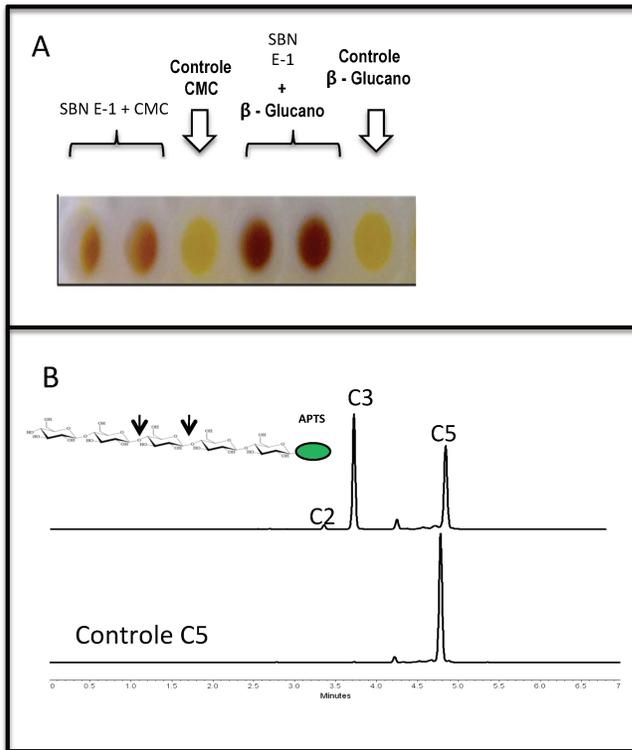


Figura 17 – A) Confirmação da atividade enzimática do clone E-1 em CMC e  $\beta$ -glucano revelada pelo método do ADNS. B) Ensaio de eletroforese capilar com o sobrenadante da cultura do clone E-1. C5, C3 e C2, correspondem a celopentose, celotriose e celobiose, respectivamente. As setas indicam às ligações glicosídicas hidrolisadas durante a reação enzimática com C5.

Ensaio de atividade enzimática com o sobrenadante do clone E-2 em  $\beta$ -glucano e CMC permitiram a confirmação da atividade catalítica (Figura 18).

Os clones obtidos pela retransformação dos plasmídeos isolados do clone positivo E-3 foram confirmados quanto à formação de halo de hidrólise na presença de CMC (Figura 19).

Os ensaios de atividade enzimática do sobrenadante da cultura crescida do clone X-1, resultaram na hidrólise das ligações tipo  $\beta(1\rightarrow4)$  entre os resíduos de xilose da cadeia principal de xilana (Figura 20A). A análise por eletroforese capilar dos produtos da hidrólise gerados pela ação do sobrenadante de X-1 em xilohexose marcada com APTS, revelou um típico padrão de endo- $\beta$ -1,4-xilanase, ao atuar na quebra das ligações internas e gerar xilotetose, xilotriose e xilobiose como produtos

principais (Figura 20B).



Figura 18- Ensaio de atividade enzimática com o sobrenadante do clone E-2 em (A)  $\beta$ -glucano; (B) CMC, revelados pelo método do ADNS.

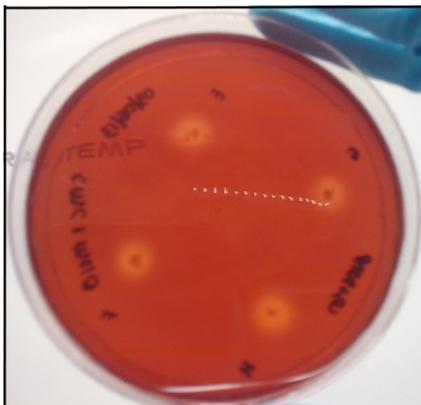


Figura 19 – Confirmação da atividade enzimática do clone E-3 em placa contendo CMC, revelada com solução de vermelho congo 0,1%.

A atividade catalítica do sobrenadante do clone X-2 foi confirmada em xilana, após incubação por 16 horas a 45°C (dados não mostrados). Curiosamente, o sobrenadante do clone X-2 também apresentou capacidade de hidrolisar os polissacarídeos liquenana e  $\beta$ -glucano (Figura 21).

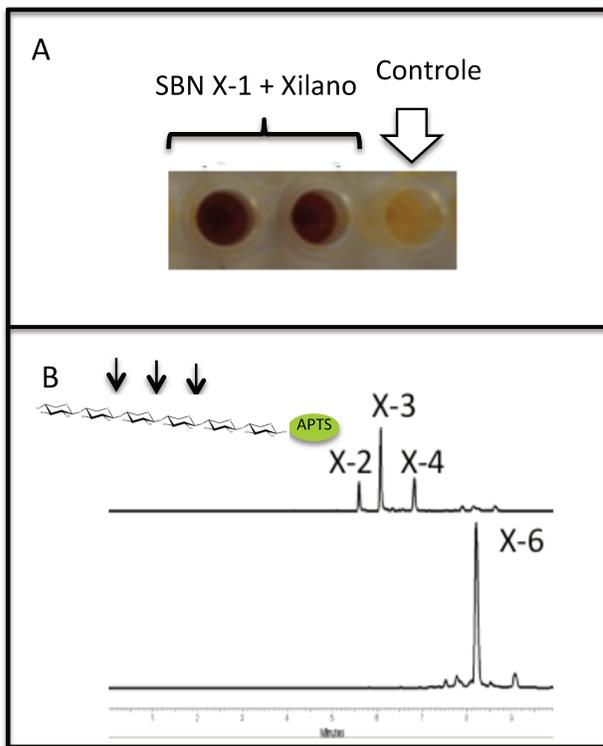


Figura 20 – A - Confirmação da atividade enzimática do clone X-1 em xilana de madeira de faia, revelada pelo método do ADNS. B - Ensaio de eletroforese capilar com o sobrenadante da cultura do clone X-1. X-6, X-4, X-3 e X-2 correspondem a xilohexose, xilotetrose, xilotriose e xilobiose, respectivamente. As setas indicam às ligações glicosídicas hidrolisadas durante a reação enzimática de X-6.

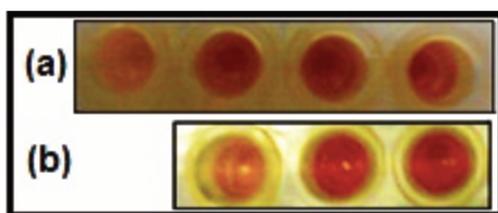


Figura 21 – Atividade enzimática com o sobrenadante do clone X-2 em (a)  $\beta$ -glucano e (b) liquenana, revelada pelo método do ADNS.

A frequência de clones positivos identificados em relação ao número de clones avaliados foi de 1: 8.967 na triagem em CMC e 1: 13.450 na triagem em xilana. Em virtude dos trabalhos que descrevem a busca de novas enzimas por abordagem funcional utilizarem diferentes tipos de vetores, tamanhos médios de insertos e números de clones, a frequência de clones positivos obtida foi comparada com a de

outros trabalhos pela relação entre o número de clones positivos identificados pela quantidade total, em megabases Mb, de DNA metagenômico submetida à triagem (Tabela 10). Para a obtenção deste valor, multiplicou-se o número de clones totais da biblioteca pelo tamanho médio dos insertos.

Tabela 10 – Comparação das frequências de clones positivos identificados através da triagem funcional de bibliotecas metagenômicas.

Enzima-alvo	Metagenoma de origem	Tipo de vetor	Amostra enriquecida / fonte de enriquecimento	DNA total (Mb)	Frequência de clones positivos / Mb de DNA	Referência
Celulase	Amostra de solo de canavial	Plasmídeo	Sim / bagaço de cana-de-açúcar	82,4	1/ 27,5	Presente trabalho
<b>Celulase</b>	Microbiota do intestino grosso de suínos em crescimento	Plasmídeo	Sim / Solka-Floc celulose comercial	430	1/ 39,1	WANG et al., 2012
<b>Celulase</b>	Amostra de água do lago Nakuru	Lambda	Sim/ CMC	60	1/ 60	REES et al., 2003
<b>Celulase</b>	Microbiota do ceco de coelhos	Cosmídeos	Sim / gramínea	1.141	1/ 103,7	FENG et al., 2007
<b>Celulase</b>	Amostra de compostagem contendo estrume de porco e resíduos de cogumelos	Fosmídeo	Não	433,3	1/ 216,7	KWON et al., 2010
<b>Celulase</b>	Amostra de solo de compostagem	Cosmídeo	Não	3.300	1/ 825	PANG et al., 2009
<b>Xilanase</b>	Amostra de solo de canavial	Plasmídeo	Sim / bagaço de cana-de-açúcar	82,4	1/ 41,2	Presente trabalho
<b>Xilanase</b>	Amostra de compostagem contendo estrume de porco e resíduos de cogumelo	Fosmídeo	Não	433,3	1/ 86,7	KWON et al., 2010
<b>Xilanase</b>	Amostra de solo	Plasmídeo	Não	204	1/ 204	HU et al., 2008
<b>Xilanase</b>	Amostra de solo	Cosmídeo	Sim / palha de arroz	306	1/ 306	MO et al., 2010
<b>Xilanase</b>	Microbiota de rúmen de vaca	Fosmídeo	Não	960	1/ 960	CHENG et al., 2012

Como pode ser observado na Tabela 10, a quantidade de clones positivos identificados por Mb de DNA na BCB foi superior ao encontrado em outros trabalhos tanto para triagem de celulases como de xilanases. Isso pode ser atribuído à etapa de enriquecimento realizada, que possivelmente promoveu à especialização da comunidade microbiana e o enriquecimento dos genes de interesse.

### **3.4.4. Mapeamento dos insertos dos clones positivos identificados na triagem funcional**

#### **3.4.4.1. Clone E-1**

A digestão enzimática do DNA plasmidial isolado do clone E-1, resultou na identificação de um inserto com tamanho aproximado de 3,1 kb (Figura 22).

O sequenciamento do inserto permitiu a localização de duas ORFs na fase de leitura +2, denominadas ORF1(E1) e ORF2(E1). Elas foram analisadas através da ferramenta BLASTx e as informações referentes às maiores identidades de sequências encontradas estão detalhadas na Figura 23. O fato das ORFs adjacentes apresentarem diferentes classificações taxonômicas, pode estar relacionado com a ocorrência da transferência horizontal de genes, um evento comum em bactérias, através do qual elas podem compartilhar genes que oferecem vantagens adaptativas (TASSE et al., 2010, MATTEOTTI et al., 2011).

A comparação das sequências de aminoácidos das ORFs com o banco de dados do CDD, através da ferramenta BLASTp, permitiu a identificação de domínios conservados, como pode ser observado na Figura 24.

A proteína codificada pela ORF1(E-1), denominada celulase E-1, apresentou identidade com o domínio conservado da família 5 de glicosil hidrolases (GH5), pertencente à superfamília das celulases. A identidade neste caso, foi observada na região N-terminal de E-1 se estendendo do aminoácido 16 ao 259. Já na região C-terminal de E-1 não foi encontrada identidade com domínios conservados conhecidos.

A família GH5 é uma das maiores famílias de celulases e contém enzimas de origem bacteriana e fúngica. As enzimas da família GH5 são tipicamente endoglucanases com ampla atuação em diferentes substratos, porém atuam

preferencialmente em  $\beta$  - glucano e CMC. Elas ainda apresentam baixa atividade enzimática em substratos cristalinos (HALL et al., 1995).

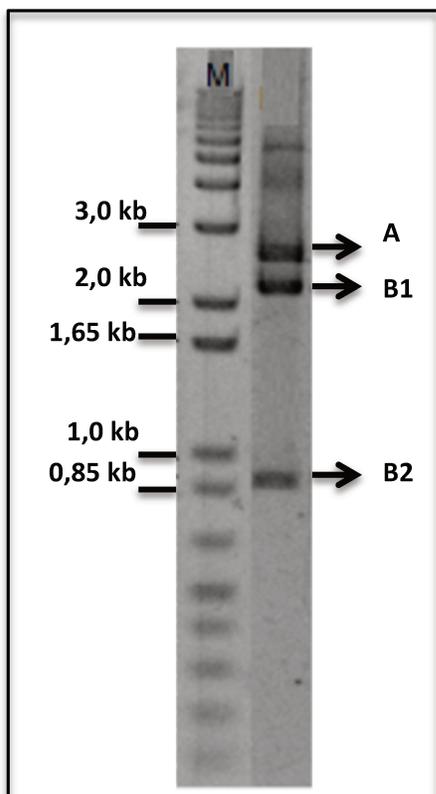
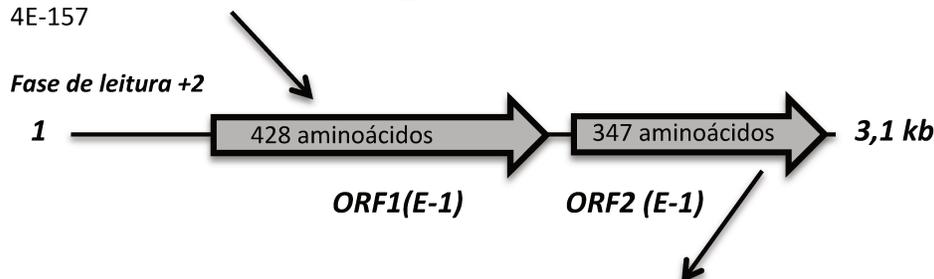


Figura 22 – Eletroforese em gel de agarose 1,0% do produto do perfil de restrição do clone E-1. A - Banda corresponde ao vetor pUC19 linearizado. B1 e B2 - bandas referentes ao inserto liberado. M marcador de peso molecular 1 kb.

Analisando a sequência de aminoácidos de E-1 foi observado que adjacente ao domínio conservado da família GH5 existia uma região rica em resíduos de serina. Tais regiões, comumente encontradas em enzimas ativas em carboidratos de origem procariótica, atuam como regiões espaçadores (*linkers*) separando domínios catalíticos e CBMs (HOWARD et al., 2004). Além disso, na sequência de aminoácidos da porção C-terminal de E-1 foram identificados três resíduos de triptofano e três resíduos de tirosina. Esses resíduos de aminoácidos, que apresentam cadeias laterais com características hidrofóbicas, são comumente encontrados em CBMs (SHOSEYOV et al., 2006, GUILLEN et al., 2010) e estão envolvidos no processo de reconhecimento e

ligação dos CBMs aos polissacarídeos (RUBINI et al., 2010). Tais informações sugeriram a existência de um CBM na região C-terminal da celulase E-1 e esta hipótese foi analisada e descrita em detalhes nos capítulos 4 e 5.

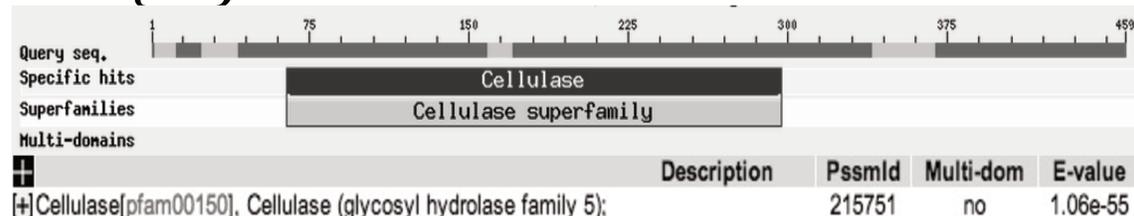
A ORF 1 apresentou 70% de cobertura de sequência e 79% de identidade com endo-1,4-beta-glicanase Cel5b de *Cellvibrio japonicus* UEDA107 (YP\_001983438) – E value = 4E-157



A ORF 2 apresentou 98% de cobertura de sequência e 51% de identidade com glicosil-transferase de *Acidovorax* sp. JS42 (YP\_984516) – E value = 2E-98

Figura 23 – Esquema da organização de duas ORFs localizadas no inserto de DNA clonado em E-1. Através da ferramenta BLASTx as sequências de nucleotídeos traduzidas foram comparadas com o banco de sequências de aminoácidos do NCBI. O tamanho do inserto está indicado em kb, e as setas indicam a orientação das ORFs.

### **ORF1(E-1)**



### **ORF2(E-1)**

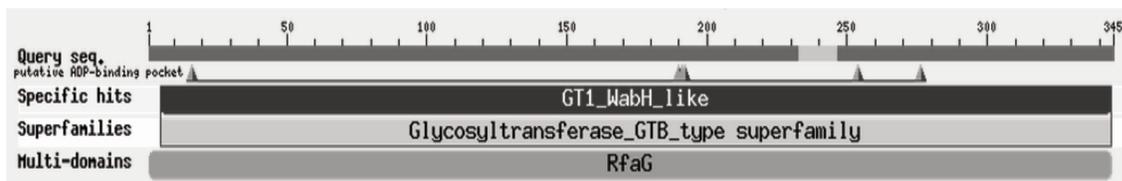


Figura 24 – Domínios conservados encontrados nas sequências de aminoácidos correspondentes a ORF1(E-1) e ORF2(E-1) após comparação com o banco de dados do CDD através da ferramenta BLASTp.

Através da ferramenta Signal-Blast, foi identificado um peptídeo sinal predito na região entre o primeiro e o sétimo aminoácido da sequência de aminoácidos da sequência correspondente a E-1.

Em relação à comparação com o banco de dados do PDB, verificou-se que a sequência de aminoácidos de E-1 apresentou 66% de identidade com a sequência de aminoácidos da celulase Cel5 de *Erwinia Chrysanthemi*, com estrutura tridimensional identificada no PDB como 1EGZ\_A.

Já a proteína codificada pela ORF2(E-1) apresentou identidade com a família GT1 das glicosil-transferases, sendo membro da superfamília de glicosil-transferases do tipo GTB. As glicosil transferases são enzimas responsáveis pela biossíntese de dissacarídeos, oligossacarídeos e polissacarídeos, através da catálise da reação de formação das ligações glicosídicas (LAIRSON et al., 2008).

#### **3.4.4.2. Clone E-2**

O tamanho do inserto metagenômico clonado em E-2, determinado pela digestão enzimática, foi de aproximadamente 2,3 kb (Figura 25).

O sequenciamento permitiu a identificação de uma ORF na fase de leitura -2, denominada ORF1(E-2), como pode ser visualizado na Figura 26. A ORF1(E-2) apresentou maior identidade de sequência (69%) com a celulase da família GH5 de *Mycobacterium rhodesiae*.

Por comparação com o banco de dados do CDD, observou-se que assim como E-1, a proteína codificada pela ORF1(E-2), denominada celulase E-2, apresentou em sua sequência um domínio conservado de celulases da família GH5 (Figura 27). Também foi identificado um segundo domínio conservado, denominado Calx-beta, que são domínios comuns em trocadores de cálcio e sódio (SCHWARZ & BENZER, 1997).

Em relação à comparação com o banco de dados do PDB, verificou-se que a sequência de aminoácidos correspondente à E-2 apresentou 59% de identidade com a sequência de aminoácidos da endocelulase de *Acidothermus Cellulolyticus*, com estrutura tridimensional identificada no PDB como 1ECE\_A.

Através da ferramenta Signal-Blast, foi identificado um peptídeo sinal predito na

região correspondente aos primeiros trinta e seis aminoácidos da sequência correspondente a E-2.

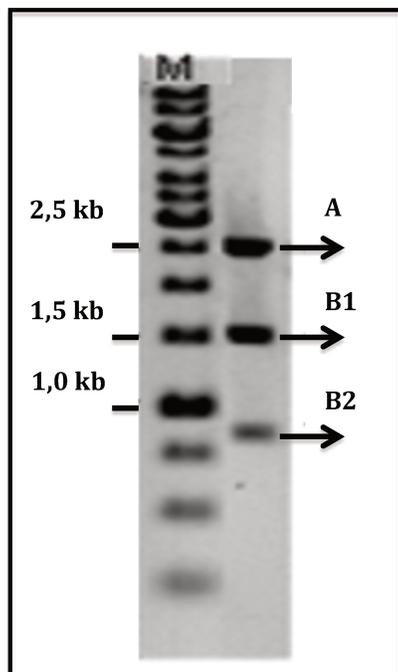


Figura 25 - Eletroforese em gel de agarose 1,0% do perfil de restrição do clone E-2. A - Banda correspondente ao vetor pUC19 linearizado. B1 e B2 - bandas referentes ao inserto liberado. M - Marcador de peso molecular 1 kb.

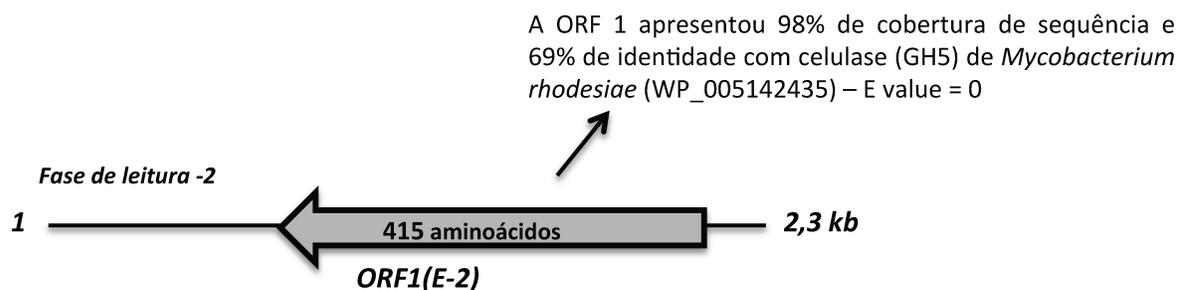
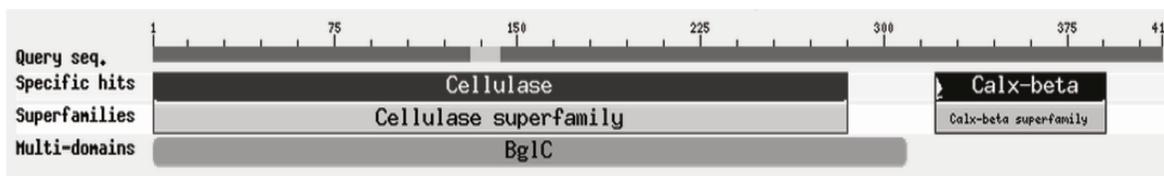


Figura 26 - Esquema da organização da ORF localizada no inserto de DNA presente no clone E-2. Através da ferramenta BLASTx a sequência de nucleotídeos traduzida foi comparada com o banco de sequências de aminoácidos do NCBI. O tamanho do inserto está indicado em kb, e as setas indicam a orientação da ORF.



List of domain hits				
	Description	PssmId	Multi-dom	E-value
[+]	Cellulase[pfam00150], Cellulase (glycosyl hydrolase family 5);	215751	no	3.48e-38
[+]	Calx-beta[pfam03160], Calx-beta domain;	202559	no	8.45e-22

Figura 27 - Domínio conservado identificado na sequência de aminoácidos correspondente à ORF1(E-2), após comparação com o banco de dados do CDD através da ferramenta BLASTp.

### 3.4.4.3. Clone E-3

O tamanho do inserto metagenômico clonado em E-3, determinado pela digestão enzimática, foi de aproximadamente 1,5 kb (Figura 28).

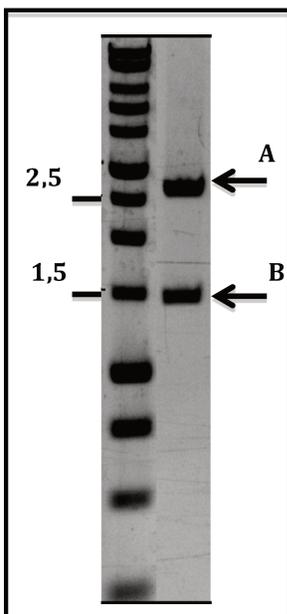


Figura 28 – Eletroforese em gel de agarose 1,0% do perfil de restrição do clone E-3. A - Banda correspondente ao vetor pUC19 linearizado. B - banda referente ao inserto liberado. M - Marcador de peso molecular 1 kb.

O sequenciamento e análise do inserto de DNA clonado no clone E-3, permitiu a identificação de apenas uma ORF, na fase de leitura +2, denominada ORF1(E-3), apresentando 49% de identidade com celobiohidrolase de *Mycobacterium sp.* JDM 601 (Figura 29).

A comparação da sequência de aminoácidos da proteína codificada pela ORF1(E-3), denominada E-3, com o banco de dados do CDD, resultou na identificação de um domínio conservado para glicosil hidrolases da família 6 (GH6) (Figura 30). Os representantes da família GH6 são enzimas do tipo endoglucanases e celobiohidrolases (LIU et al., 2010).

Em relação à comparação com o banco de dados do PDB, verificou-se que a sequência de aminoácidos correspondente à E-3 apresentou 46% de identidade com a sequência de aminoácidos da endoglucanase Cel6 de *Mycobacterium Tuberculosis*, com estrutura tridimensional identificada no PDB como 1UP3\_A.

Através da ferramenta Signal-Blast, foi identificado um peptídeo sinal predito na região correspondente aos primeiros trinta aminoácidos da sequência correspondente a E-3.

A ORF 1 apresentou 89% de cobertura e 49% de identidade com celobiohidrolase CelA de *Mycobacterium sp.* JDM 601 (YP\_004523900) E value = 1E-67

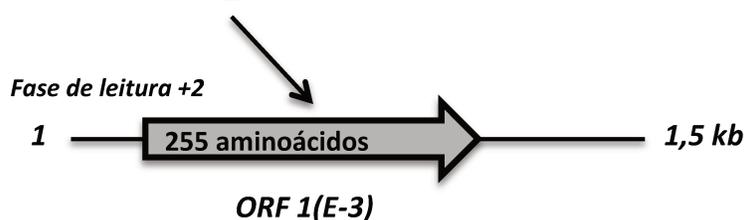


Figura 29 - Esquema da organização da ORF localizada no inserto de DNA presente no clone E-3. Através da ferramenta BLASTx a sequência de nucleotídeos traduzida foi comparada com o banco de sequências de aminoácidos do NCBI. O tamanho do inserto está indicado em kb, e a seta indica a orientação da ORF.

Existe apenas um relato na literatura (BELOQUI et al., 2010) que descreve o isolamento de uma enzima pertencente à família GH6 a partir da triagem funcional de uma biblioteca metagenômica construída a partir de metagenoma de minhoca. A

maioria das celulasas identificadas a partir da triagem funcional de bibliotecas metagenômicas apresentou identidade com glicosil hidrolases da família 5 e 9 (GH5 e GH9) (DUAN & FENG, 2010). As razões para isso podem estar relacionadas ao fato dessas enzimas serem mais facilmente expressas em *Escherichia coli* e também pela abundância dos genes que codificam GH5 e GH9 no genoma de bactérias (KRAUSE et al., 2003, TAYLOR et al., 2006, XIE et al., 2007).

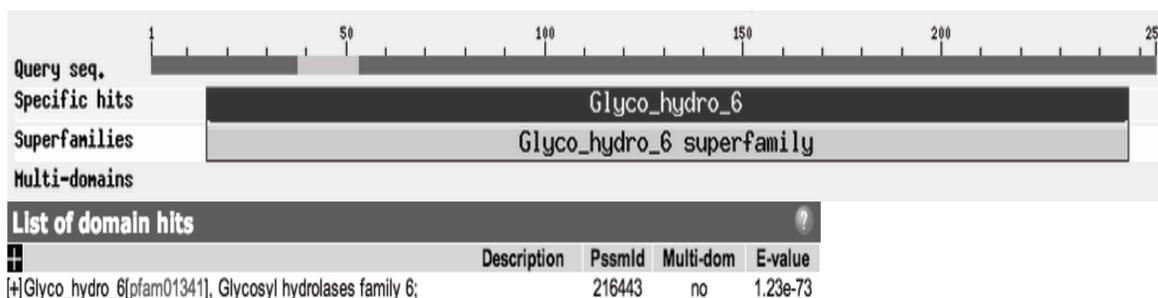


Figura 30 – Domínio conservado identificado na sequência de aminoácidos correspondente à ORF1(E-3), após comparação com o banco de dados do CDD pela ferramenta BLASTp.

#### 3.4.4.4. Clone X-1

A digestão do DNA plasmidial extraído do clone X-1, indicou a presença de um inserto com aproximadamente 7,0 kb (Figura 31).

Visando facilitar as etapas de sequenciamento para identificação da ORF codificante para xilanase no inserto de aproximadamente 7,0 kb, foi construída uma sub-biblioteca contendo clones com insertos na faixa de 1 a 2 kb. A triagem funcional dos clones da sub-biblioteca permitiu identificação de um clone com atividade enzimática, contendo um inserto de aproximadamente 2,2 kb.

O sequenciamento deste inserto resultou na identificação de uma ORF, na fase de leitura +2, denominada ORF1(X-1), com 61% de identidade com a endo-1,4-beta-xilanase de *Cellvibrio mixtus*, como pode ser verificado na Figura 32. Neste caso, a proteína com estrutura tridimensional depositada no PDB com maior identidade em relação à sequência de aminoácidos da ORF1(X-1) é a própria endo-1,4-beta-xilanase

de *C. mixtus* (identificação no PDB - 2CNC\_A).

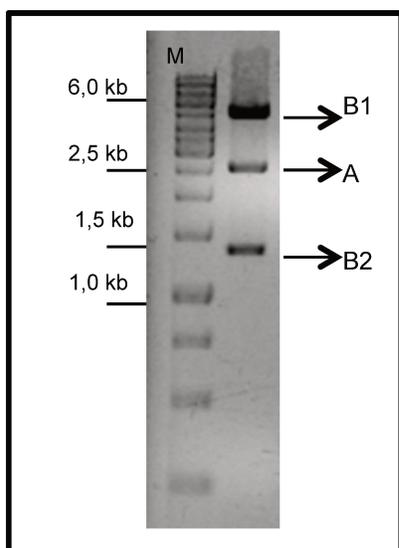


Figura 31 – Eletroforese em gel de agarose 1,0% do perfil de restrição do clone obtido pela sub-clonagem de X-1. A - Banda correspondente ao vetor pUC19 linearizado. B1 e B2 - bandas referentes ao inserto liberado. M - Marcador de peso molecular 1 kb.

A ORF 1 apresentou 93% de cobertura e 61% de identidade com endo-1,4-beta-xilanase de *Cellvibrio mixtus* (AAD09439) – E value = 3E-163

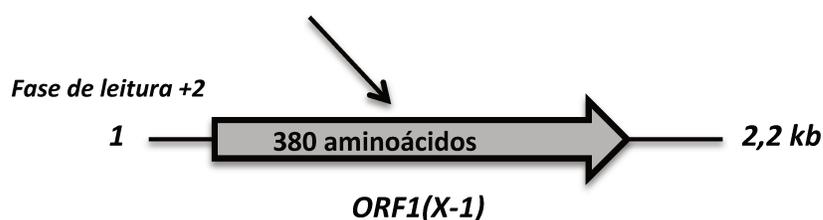


Figura 32 – Esquema da organização da ORF localizada no inserto de DNA presente no clone X-1. Através da ferramenta BLASTx a sequência de nucleotídeos traduzida foi comparada com o banco de sequências de aminoácidos do NCBI. O tamanho do inserto está indicado em kb, e a seta indica a orientação da ORF.

Em relação à análise da presença de domínios conservados, por comparação com o banco de dados do CDD, foi identificado um domínio conservado para glicosil hidrolases da família 10 (GH10) (Figura 33), cujos representantes são enzimas do tipo

endo-1,4- $\beta$ -xilanasas, endo-1,3- $\beta$ -xilanasas e celobiohidrolases (SHI et al., 2013). A enzima codificada pela ORF1(X-1) foi denominada de xilanase X-1.

Através da ferramenta Signal-Blast, foi identificado um peptídeo sinal predito na região correspondente aos primeiros trinta e seis aminoácidos da sequência correspondente à X-1.

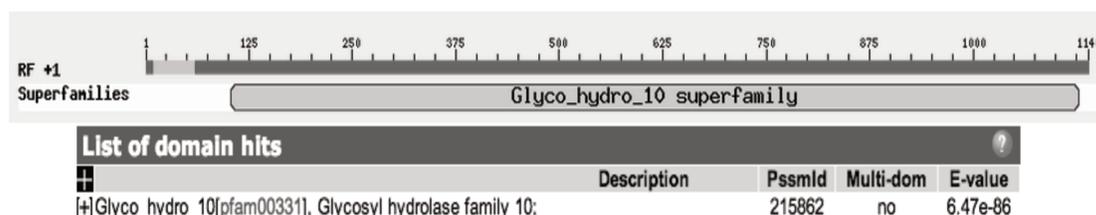


Figura 33 – Domínio conservado identificado na sequência de aminoácidos correspondente à ORF1(X-1), após comparação com o banco de dados do CDD pela ferramenta BLASTp.

#### 3.4.4.5. Clone X-2

A digestão do DNA plasmidial extraído do clone X-2, indicou a presença de um inserto com aproximadamente 3,5 kb (Figura 34).

O sequenciamento e análise do inserto, permitiu a identificação de duas ORFs (Figura 35), denominadas ORF1(X-2) e ORF2(X-2). Neste caso, ambas as ORFs apresentaram identidade com proteínas sintetizadas pela espécie *Sphingomonas elodea*.

A Figura 36 ilustra os domínios conservados identificados em cada uma das ORFs através da comparação com o banco de dados do CDD.

A proteína codificada pela ORF1(X-2) apresentou identidade com domínios conservados do tipo ACT. Esses domínios estão presentes em proteínas envolvidas no controle do metabolismo, transporte de solutos e transdução de sinal, e cuja atividade é regulada por diferentes concentrações de aminoácidos específicos (GRANT, 2006).

A proteína codificada pela ORF2(X-2), denominada X-2, apresentou identidade com o domínio conservado presente na família 16 de glicosil hidrolases (GH16). Alguns exemplos de enzimas classificadas na família GH16 são xiloglucanases, endo-1,4- $\beta$ -

galactosidases, endo-1,3-β-glucanases e liquenases, que atuam na hidrólise de ligações do tipo β – 1,4 e β – 1,3 em glucanos e galactanos (EKLOF & BRUMER, 2010).

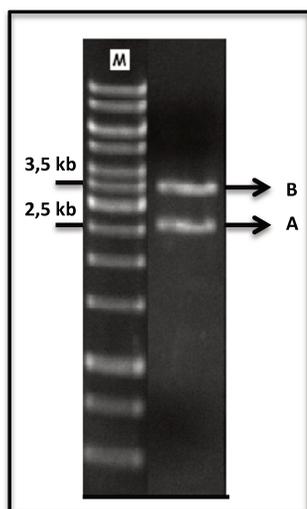
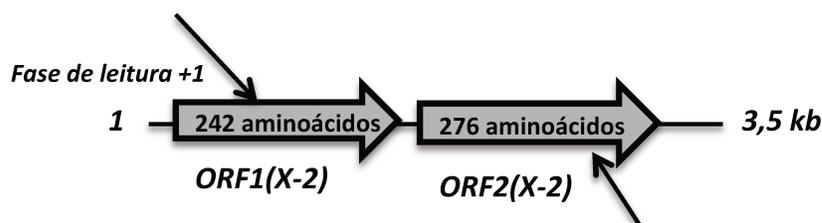


Figura 34 – Eletroforese em gel de agarose 1,0% do perfil de restrição do clone X-2. (A) Banda correspondente ao vetor pUC19 linearizado, (B) banda correspondente ao inserto liberado e (M) Marcador de peso molecular 1 kb.

A ORF 1 apresentou 99% de cobertura de sequência e 82% de identidade com uridil transferase de *Sphingomonas elodea* (WP\_010544327) – Evalue = 7E-30



A ORF 2 apresentou 94% de cobertura de sequência e 70% de identidade com glicosil hidrolase de *Sphingomonas elodea* (WP\_010544485) --Evalue = 4E-138

Figura 35 – Esquema da organização das ORFs localizadas no inserto de DNA presente no clone X-2. Através da ferramenta BLASTx as sequências de nucleotídeos traduzidas foram comparadas com o banco de sequências de aminoácidos do NCBI. O tamanho do inserto está indicado em kb, e as setas indicam a orientação das ORFs.

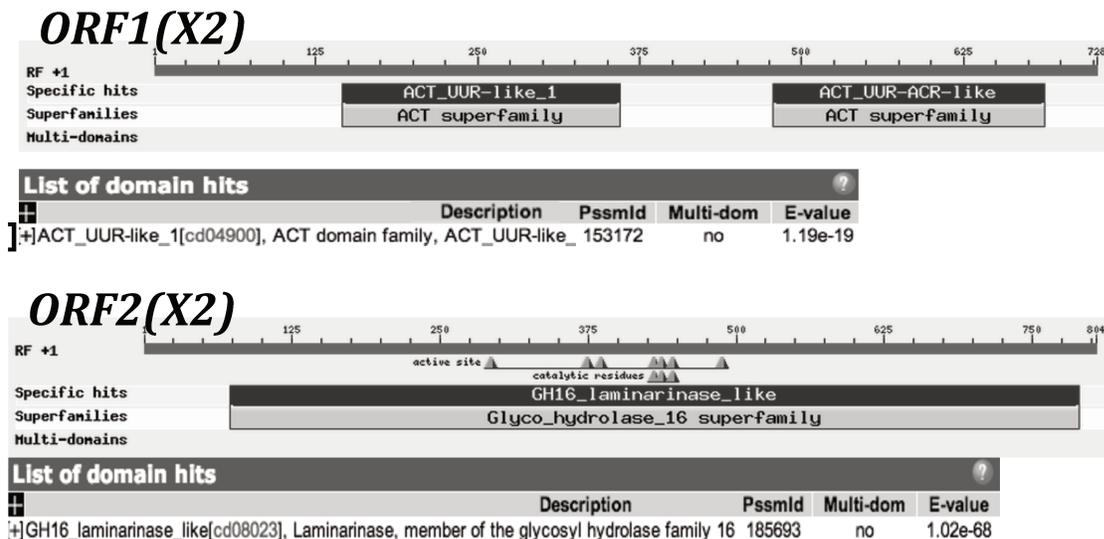


Figura 36 – Domínios conservados identificados nas sequências de aminoácidos correspondentes à ORF1(X-2) e ORF2(X-2) após comparação com o banco de dados do CDD pela ferramenta BLASTp.

Em relação à comparação com o banco de dados do PDB, verificou-se que a sequência de aminoácidos correspondente à X-2 apresentou 48% de identidade com a sequência de aminoácidos da laminarase de *Rhodothermus Marinus*, com estrutura tridimensional identificada no PDB como 3ILN\_A.

Através da ferramenta Signal-Blast, foi identificado um peptídeo sinal predito na região correspondente aos primeiros vinte e três aminoácidos da sequência correspondente à X-2.

Curiosamente, apesar do clone X-2 ter sido identificado através da formação de halo de hidrólise em xilana, não foi encontrada nenhuma ORF com identidade para xilanases.

Como descrito anteriormente, foi verificada atividade enzimática do clone X-2 tanto em xilana, que contém predominantemente ligações glicosídicas do tipo  $\beta(1\rightarrow4)$  entre resíduos de xilose, em  $\beta$ -glucano e liquenana, compostos pela alternância de ligações glicosídicas do tipo  $\beta(1\rightarrow3)$  e  $\beta(1\rightarrow4)$  em diferentes proporções entre resíduos de glicose. Isso sugere que a enzima X-2, apresentou característica bifuncional, ou seja, capaz de atuar em substratos com características distintas. Algo semelhante foi

verificado para liquenase de *Thermomonospora sp.*, com um único sítio ativo capaz de hidrolisar liquenana e xilana (ANISH & RAO, 2007).

### 3.5. Conclusões

Neste capítulo foram descritas as estratégias metagenômicas desenvolvidas visando à identificação de novos genes codificantes para enzimas envolvidas na degradação dos componentes da parede celular vegetal no metagenoma de solo de canavial.

As abordagens baseadas em homologia de sequência nucleotídica, que envolveram a realização de reações de PCR para a amplificação de genes alvos, não permitiram o isolamento dos genes de interesse. Uma das razões para isso pode estar relacionada à elevada diversidade de sequências nucleotídicas presente nas amostras de DNA metagenômico utilizadas como alvo, o que favoreceu a ocorrência de ampliações inespecíficas.

Por outro lado, através da triagem funcional dos clones da biblioteca metagenômica derivada da etapa de enriquecimento (BCB), foi possível a identificação de três clones (E-1, E-2 e E-3) com atividade celulolítica e dois clones (X-1 e X-2) com atividade xilanolítica. A análise dos dados gerados pelo sequenciamento dos insertos de cada um dos clones resultou na identificação de ORFs com identidade para genes de diferentes famílias de glicosil hidrolases.

O clone E-1 apresentou uma ORF com 79% de identidade com a endo-1,4-beta-glucanase de *Cellvibrio japonicus* UEDA 107. A proteína codificada por esta ORF recebeu o nome de celulase E-1. A região N-terminal de E-1 apresentou identidade com o domínio conservado de celulases da família GH5. Já em relação à região C-terminal foram encontrados indícios da existência de um CBM, principalmente pela identificação de um provável *linker* rico em resíduos de serina e de aminoácidos aromáticos nesta região.

O clone E-2 apresentou uma ORF com 69% identidade com a celulase de *Mycobacterium rhodesiae*. A sequência de aminoácidos da celulase codificada por esta ORF, denominada celulase E-2, também apresentou identidade com o domínio conservado da família GH5.

O clone E-3, apresentou uma ORF com 49% de identidade com celobiohidrolase de *Mycobacterium sp.* JDM 601. A sequência de aminoácidos da enzima codificada por esta ORF, denominada celulase E-3, apresentou identidade com domínios conservados da família GH6.

A ORF identificada no clone X-1 apresentou 61% de identidade com a endo-1,4-beta-xilanase de *Cellvibrio mixtus*. A xilanase X-1, codificada por esta ORF, apresentou em sua sequência um domínio conservado com identidade para xilanases da família GH10.

O clone X-2 apresentou uma ORF com 70% de identidade com glicosil hidrolase de *Sphingomonas elodea*. A análise da sequência de aminoácidos da proteína codificada por esta ORF permitiu a identificação de um domínio conservado comum às glicosil hidrolases da família 16 (GH16).

As enzimas identificadas nesta etapa do trabalho apresentam-se como potenciais alvos de caracterização bioquímica e estrutural, visando à avaliação da potencial aplicação das mesmas em processos de conversão da biomassa vegetal.

Comumente, a frequência de clones positivos identificados por meio da triagem funcional de bibliotecas metagenômicas é baixa, principalmente devido às dificuldades inerentes ao método como: necessidade da clonagem de genes completos, ocorrência da transcrição de RNA mensageiro, síntese proteica em quantidades suficientes para que a atividade enzimática possa ser detectada associada ao correto enovelamento proteico e secreção para o meio extracelular. No presente trabalho a frequência de clones positivos identificados por megabase de DNA triado foi superior a outros valores descritos na literatura. Este fato pode estar diretamente relacionado à etapa de enriquecimento que deu origem ao consórcio microbiano do qual foi extraído o DNA metagenômico empregado na construção da biblioteca submetida à triagem funcional.

Por fim, as bibliotecas metagenômicas descritas neste trabalho podem ser empregadas na busca por genes que codificam enzimas envolvidas em diferentes processos metabólicos, em virtude da diversidade funcional presente nas amostras de DNA metagenômico das quais derivaram.



# **CAPÍTULO 4- CARACTERIZAÇÃO BIOQUÍMICA DA CELULASE E-1 E DA XILANASE X-1 E AVALIAÇÃO DA APLICAÇÃO EM PROCESSOS BIOTECNOLÓGICOS**

## **4.1. Introdução**

A triagem funcional da biblioteca metagenômica construída a partir do DNA extraído do consórcio microbiano resultou na identificação de novas enzimas com habilidade de degradar polissacarídeos presentes na parede celular vegetal.

As ORFs dos clones E-1 e X-1 que apresentaram identidade, respectivamente, com celulases da família GH5 e xilanases da família GH10 foram clonadas e caracterizadas bioquimicamente.

As informações obtidas nesta etapa do trabalho foram importantes para a caracterização da potencial aplicação destas enzimas em processos biotecnológicos.

## **4.2. Objetivos do Capítulo**

- Clonagem, expressão heteróloga e caracterização bioquímica da celulase E-1 e da xilanase X-1;

- Avaliação da atividade enzimática de ambas as enzimas em substratos complexos, visando sua aplicação em processos de conversão de biomassa vegetal em bioprodutos.

## **4.3. Materiais e Métodos**

### **4.3.1. Desenho dos oligonucleotídeos iniciadores para amplificação dos genes e clonagem em vetor de expressão**

Os oligonucleotídeos iniciadores foram desenhados visando à clonagem dos genes de interesse no vetor de expressão pET-28a (Novagen) em fusão com cauda de histidina na região N-terminal. Primeiramente, as ORFs foram analisadas

individualmente em relação aos sítios de restrição presentes nas sequências, através da ferramenta *NEBufferCutter* (disponível em: <http://tools.NEBuffer.com/NEBUFFERcutter2/>). Os sítios de restrição foram escolhidos de acordo com os sítios disponíveis para clonagem no vetor e ausentes nas ORFs analisadas.

A ORF1(E-1) e ORF1(X-1), excluindo as regiões correspondentes aos peptídeos-sinal preditos, foram utilizadas como molde para o desenho dos oligonucleotídeos iniciadores, que foi realizado com o auxílio da ferramenta *OligoAnalyzer* ([www.idtdna.com](http://www.idtdna.com)). Além da sequência complementar ao DNA alvo e dos sítios de restrição, foram adicionadas bases extras nas extremidades 5' dos oligonucleotídeos iniciadores diretos e reversos, necessárias para ancoragem das enzimas de restrição. O número de bases adicionais foi determinado de acordo com a tabela disponível em: [http://www.NEBuffer.com/NEBuffercomm/tech\\_reference/restriction\\_enzymes/cleavage\\_olignucleotides\\_old.pdf](http://www.NEBuffer.com/NEBuffercomm/tech_reference/restriction_enzymes/cleavage_olignucleotides_old.pdf).

A Tabela 11 apresenta os oligonucleotídeos iniciadores desenhados e utilizados nas reações de PCR. A temperatura de anelamento ( $T_a$ ) se refere à sequência que se anela ao gene molde.

Tabela 11 - Oligonucleotídeos iniciadores utilizados nas reações de PCR. As cores indicam os sítios de restrição enzimática: amarelo – *NdeI* e azul – *BamHI*. **E-1** = Celulase E-1, **E-1 Cat**= Apenas o domínio catalítico da celulase E-1, **X-1**= xilanase X-1. **Ta** indica a temperatura de anelamento dos oligonucleotídeos iniciadores construídos.

Genes Alvos	Oligonucleotídeos iniciadores diretos (5'→ 3')	Oligonucleotídeos iniciadores reversos (5'→ 3')
<b>E-1</b>	TATATATCATATGTCGCGACCCATTACTACCAGC $T_a = 58^\circ\text{C}$	ATACGATCCTTACCAGTTATCGAACTTCACATTA $T_a = 52^\circ\text{C}$
<b>E-1 Cat</b>	TATATATCATATGTCGCGACCCATTACTACCAGC $T_a = 58^\circ\text{C}$	ATACGATCCTTACGGCCAACCGGAAATAAT $T_a = 53^\circ\text{C}$
<b>X-1</b>	TATATATCATATGGCCAAAGCCGCGAGCGC $T_a = 50,1^\circ\text{C}$	ATACGATCCTTACTTCTTCAAATCCAGCAA $T_a = 49,9^\circ\text{C}$

### 4.3.2. Amplificação dos genes por PCR

Todas as PCRs foram realizadas utilizando como alvo a solução dos plasmídeos extraídos dos clones positivos. As condições de reações foram: 34,5 µL de água ultrapura estéril, 10 µL de tampão Phusion® HF 5X (Finzymes), 1 µL de dNTP 10 mM, 1,0 µL de cada um dos oligonucleotídeos iniciadores (direto e reverso) 10 pmol/µL, 1,0 µL do DNA molde (concentração entre 50 e 100 ng/µL), 1,0 µL de DMSO e 1 U da Taq Phusion® HF DNA polimerase (Finzymes), totalizando um volume de reação final de 50 µL. As reações foram realizadas no termociclador MasterCycler pro-S da Eppendorf.

Todos os oligonucleotídeos iniciadores apresentaram temperatura de anelamento próxima de 55°C, o que possibilitou a amplificação de todos os genes a partir das seguintes condições: 1 ciclo de 1 minuto a 98°C, 30 ciclos de 20 segundos a 98°C, 30 segundos a 55°C e 1 minuto a 72°C e um último ciclo de 10 minutos a 72°C.

A confirmação da amplificação dos genes de interesse foi realizada por eletroforese em gel de agarose 1,0%, 90 V por 60 minutos, utilizando como padrão de peso molecular o marcador 1Kb *plus DNA Ladder* (Invitrogen). As bandas foram visualizadas em transiluminador de luz UV, após tratamento com brometo de etídeo 0,25 µg/mL por 15 minutos e as imagens capturadas através do fotodocumentador Gel Logic 6000Pro da marca Carestream.

### 4.3.3. Purificação dos fragmentos amplificados

A purificação dos fragmentos amplificados por PCR foi feita através do kit *GFX<sup>tm</sup> PCR DNA and gel band purification* da GE Healthcare.

### 4.3.4. Digestão dos fragmentos de PCR

Os produtos de PCR purificados foram submetidos à digestão com as enzimas de restrição adequadas, através da seguinte reação: 12 µL do fragmento purificado (com concentração de aproximadamente 30 ng/ µL), 2 µL de tampão 10 X (NEBuffer 2 - New England Biolabs), 10 U de cada uma das enzimas de restrição *NdeI* e *BamHI* (ambas da New England Biolabs) e 4 µL de água, completando o volume para 20 µL. A

reação foi incubada por 2 horas a 37°C e posteriormente inativada por 20 minutos a 65°C. Em seguida, a reação enzimática foi diretamente purificada através do kit *GFX<sup>tm</sup> PCR DNA and gel band purification* da GE Healthcare.

### 4.3.5. Preparo do vetor pET-28a

O vetor pET-28a (Novagen – Figura 37), é comumente utilizado para expressão heteróloga de enzimas em sistemas bacterianos.

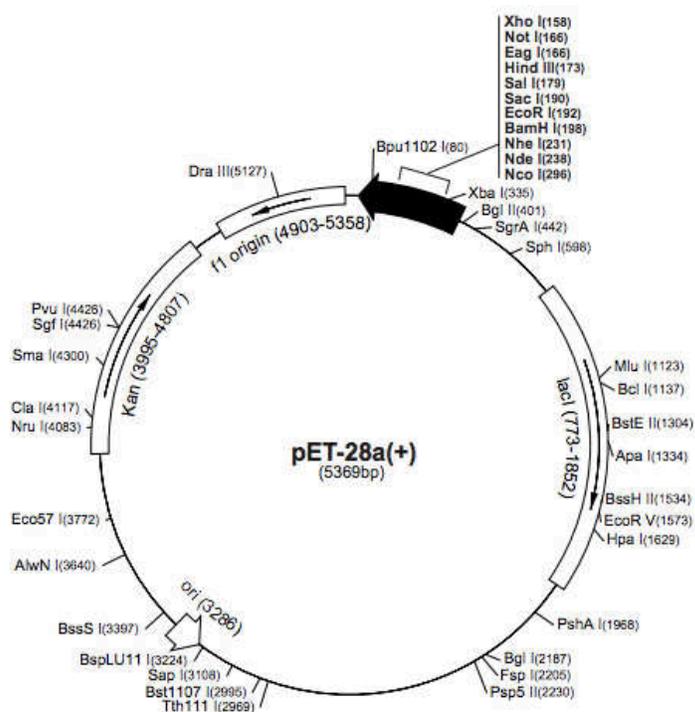


Figura 37 - Mapa do vetor de expressão pET-28a.

O vetor foi preparado da seguinte forma: 40 µL do vetor (concentração aproximada de 100 ng/µL), 10 µL do tampão 10 X (NEBuffer 2 - New England Biolabs), 10 U da cada enzima (*NdeI* e *BamHI*, ambas da New England Biolabs) e água ultrapura estéril para completar o volume final de 100 µL. A reação foi incubada a 37°C por 2 horas e posteriormente inativada a 65°C por 20 minutos. Como forma de prevenir a religação do vetor, nos casos em que a dupla digestão não tenha sido completa, a

amostra contendo o vetor digerido foi submetida à desfosforilação. Para isso, pelo fato do tampão NEBuffer 2 ser compatível com o uso da enzima CIP (*calf intestinal phosphatase* - New England Biolabs), apenas foram adicionados 10 U da enzima ao tubo contendo a reação de digestão previamente inativada. A reação foi incubada a 37°C por 45 minutos e posteriormente inativada por 20 minutos a 65°C.

O produto da reação foi purificado por eletroforese em gel de agarose 0,6%, em tampão TAE 1 X. A banda referente ao plasmídeo foi recuperada do gel e purificada com o kit *GFX<sup>™</sup> PCR DNA and gel band purification* (GE Healthcare).

#### 4.3.6. Reação de ligação

Esta etapa consistiu na ligação dos insertos digeridos e purificados ao vetor pET-28a digerido e desfosforilado, seguindo as recomendações do fabricante, disponível em: [http://lifeserv.bgu.ac.il/wb/zarivach/media/protocols/Novagen\\_pET\\_system\\_manual.pdf](http://lifeserv.bgu.ac.il/wb/zarivach/media/protocols/Novagen_pET_system_manual.pdf). As reações consistiram em 0,2 pmol do inserto purificado (a conversão de µg/µL para pmol foi feita através da ferramenta disponível em: <http://www.promega.com/techserv/tools/biomath/calc01.htm>), 50 ng do vetor pET-28a digerido e desfosforilado, 2 µL do tampão 10 X da ligase (Promega), 1 µL da T4 DNA ligase e o volume foi completado com água ultrapura estéril para 20 µL. A reação foi então incubada a 16°C por aproximadamente 16 horas.

#### 4.3.7. Confirmação da clonagem em pET-28a

As clonagens foram confirmadas através de reações de digestões enzimáticas, contendo: 5 µL do DNA plasmidial, 2 µL do tampão 10 X (NEBuffer 2 - NewEngland Biolabs), 10 U µL de cada uma das enzimas (*NdeI* e *BamHI*) e água para completar o volume final de 15 µL.

Como forma de avaliar se as sequências clonadas estavam corretas em relação às sequências de nucleotídeos originais, os plasmídeos extraídos foram sequenciados com os oligonucleotídeos iniciadores específicos do vetor (T7 “*sequencing primer*” – 5’ TAATACGACTCACTATAGGG 3’, T7 “*terminator primer*” – 5’ CTAGTTATTGCTCAGCGGTG 3’).

#### 4.3.8. Cepas utilizadas

Primeiramente os produtos das reações de ligação foram transformados em *E.coli* linhagem DH5 $\alpha$  com o objetivo de multiplicação dos plasmídeos recombinantes. Em seguida os plasmídeos foram transformados nas seguintes linhagens de *E.coli*: BL21(DE3), ArticExpress(DE3) (ambas da Strategene), Rosetta2(DE3) e Rosetta-gami2(DE3)pLysS (ambas da Novagen), que são destinadas à expressão heteróloga de proteínas. A designação DE3 indica que a célula hospedeira é um lisógeno do fago  $\lambda$ DE3, e que, portanto carrega uma cópia cromossomal do gene da T7 RNA polimerase, sob o controle do promotor *lacUV5*, que sofre indução por IPTG. A designação pLysS indica a presença do plasmídeo que codifica a lisozima T7, um inibidor natural da T7 RNA polimerase, que auxilia na supressão da expressão basal da T7 RNA polimerase antes da fase de indução. Todas as linhagens citadas são compatíveis com vetores pET e outros tipos de vetores cuja expressão é controlada pelo promotor T7.

A linhagem BL21(DE3) é normalmente a primeira a ser avaliada nos testes de expressão de proteínas heteróloga em *E.coli*, pois possui baixo nível de produção de proteases, envolvidas na degradação da proteína alvo. Já a linhagem ArticExpress(DE3) é derivada da linhagem BL21-Gold, e é adequada para contornar os casos em que a superexpressão resulta no enovelamento incorreto das proteínas fazendo com que elas se aglomerem nos chamados corpos de inclusão. O aumento da solubilidade ocorre pelo fato de que essa linhagem é capaz de expressar chaperoninas, cuja função é auxiliar no correto enovelamento das proteínas. A linhagem Rosetta2(DE3), também derivada da linhagem BL21, é recomendada para a expressão de proteínas que contém códons raramente utilizados em *E.coli*. Esta linhagem dispõe de genes no plasmídeo pRARE que codifica tRNAs para sete códons raros em *E.coli*. Ao oferecer estes códons raros, a expressão passa a não ser mais limitada pelos códons utilizados especificamente pela *E.coli*. Por último, a linhagem Rosetta-gami2(DE3)pLysS contém genes que codificam tRNAs para códons raros e possui características que facilitam a formação de pontes dissulfeto durante a expressão heteróloga (todas as informações apresentadas foram fornecidas pelos fabricantes).

#### **4.3.9. Preparo de células cálcio competentes**

Todas as células utilizadas nas etapas de transformação foram preparadas da seguinte forma: uma colônia da linhagem de interesse foi inoculada em tubo de ensaio estéril contendo 10 mL de meio LB e incubada a 37°C por cerca de 16 horas com agitação de 250 rpm. No dia seguinte, o conteúdo do tubo foi transferido para um frasco tipo Erlenmeyer de 1 L contendo 400 mL de meio LB e incubado a 37°C sob agitação de 200 rpm. Tendo se atingido a absorbância a 600 nm entre 0,3 e 0,4, a cultura foi dividida em oito tubos tipo Falcon (50 mL) e centrifugada a 4.000 x g por 10 minutos, a 4°C. O sobrenadante foi descartado e os *pellets* foram ressuspensos em 4 mL de solução de CaCl<sub>2</sub> 0,1 M contendo 10% de glicerol, e incubados por 15 minutos no gelo. Transcorrido este período, os tubos foram novamente centrifugados nas mesmas condições anteriores, e os sobrenadantes descartados. Nesta etapa, os *pellets* foram ressuspensos em 500 µL da solução de CaCl<sub>2</sub> 0,1 M com 10% de glicerol e aliquotados em frações de 100 µL, imediatamente congeladas em nitrogênio líquido e armazenadas em ultrafreezer a - 80°C.

#### **4.3.10. Transformação por choque térmico**

As transformações por choque térmico foram realizadas pela adição de 5 µL do produto da reação de ligação a 50 µL da suspensão de células cálcio competentes. A mistura foi incubada em gelo por 30 minutos. Posteriormente, a mistura foi incubada a 42°C por 90 segundos, e em seguida rapidamente resfriada a 4°C. Para a recuperação das células transformadas foram adicionados 900 µL de meio SOC, seguido por incubação a 37°C e 250 rpm de agitação por 1 hora. Após o período de recuperação, as células foram plaqueadas em meio LB sólido, contendo 50 µg/mL do antibiótico canamicina. As placas foram então incubadas a 37°C por aproximadamente 16 horas.

#### **4.3.11. Testes de expressão**

Os plasmídeos extraídos de *E.coli* linhagem DH5α foram transformados nas linhagens de *E.coli* BL21(DE3), ArticExpress(DE3), Rosetta2(DE3) e Rosetta-

gami2(DE3)pLysS.

Foram realizados testes de expressão em diferentes condições para avaliar o nível de expressão e solubilidade das enzimas em cada condição. Primeiramente colônias resultantes das transformações foram inoculadas em 5 mL de meio LB com canamicina (50 µg/mL) no caso das linhagens Rosetta2(DE3) e BL21(DE3), enquanto que para a ArticExpress(DE3) o pré-inóculo foi feito em meio LB com canamicina (50 µg/mL) e gentamicina (12 µg/mL). Em relação à cepa Rosetta-gami2(DE3)pLysS, o pré-inóculo foi feito em meio LB contendo canamicina (50 µg/mL), estreptomicina (50 µg/mL), tetraciclina (12,5 µg/mL) e cloranfenicol (34 µg/mL). Todos os inóculos foram incubados a 37°C com agitação de 200 rpm por cerca de 16 horas. Posteriormente 5 mL de cada pré-inóculo foi inoculado em 50 mL de LB com canamicina (50 µg/mL) no caso da Rosetta2(DE3), Rosetta-gami2(DE3)pLysS e BL21(DE3), e sem nenhum antibiótico no caso da ArticExpress(DE3). Em relação às três primeiras linhagens, após o inóculo, os frascos foram incubados a 37°C sob agitação de 250 rpm até atingir a absorbância a 600 nm de 0,6, com posterior adição do indutor IPTG na concentração final de 0,5 mM. No caso da ArticExpress(DE3), os frascos inoculados foram incubados a 30°C por 3 horas sob agitação de 220 rpm e em seguida foram adicionados de IPTG na concentração final de 1 mM.

No caso Rosetta2(DE3), as induções foram realizadas a 18°C por 16 horas, 180 rpm de agitação e a 30°C por 4 horas, 200 rpm de agitação. As expressões nas linhagens Rosetta-gami2(DE3)pLysS e BL21(DE3) foram realizados à 37°C por 4 horas, 200 rpm. Em relação à ArticExpress(DE3), a indução foi avaliada a 12°C por aproximadamente 16 horas sob agitação de 180 rpm.

Transcorrido o tempo de indução, as culturas foram centrifugadas a 5.800 x g por 30 minutos. Os *pellets* resultantes da expressão foram solubilizados em 5 mL de tampão A (20 mM de fosfato de sódio, 100 mM de cloreto de sódio e 5 mM de imidazol, pH 7,4) contendo lisozima e PMSF na concentração final de 0,5 mg/mL e 1 mM, respectivamente, e incubados sob agitação por 30 minutos para lise das células. Posteriormente as células foram sonicadas em ultrasonicador (Sonics Vibra-cell) com 40% de amplitude máxima, durante 10 segundos e intervalos de 50 segundos por sete vezes, a 4°C. Após a sonicação, as células foram centrifugadas a 5.800 x g por 40

minutos a 4°C.

Os testes de expressão foram avaliados por SDS-PAGE (12,5% gel de separação e 4% gel de empilhamento), onde 20 µL das amostras foram misturadas com 10 µL do tampão Laemmli 3 X e aquecidas por 5 minutos a 99°C. Após a corrida, os géis foram corados em solução de *coomassie brilliant blue* e descorados com a solução *destain* (10% de ácido acético, 40% de etanol).

#### **4.3.12. Purificação proteica**

A purificação proteica foi realizada por cromatografia de afinidade com metal imobilizado (IMAC), neste caso o Ni<sup>2+</sup>, que apresenta alta afinidade pela cauda de histidina, clonada em fusão com as sequências das proteínas de interesse.

As amostras foram injetadas a 0,5 mL/min em uma coluna de 5 mL His-Trap-Ni-NTA (GE Healthcare), previamente equilibrada com tampão A. A eluição foi feita através de gradiente de 0 a 100% do tampão B (20 mM de fosfato de sódio, 200 mM de cloreto de sódio e 500 mM de imidazol, pH 7,4) em 20 volumes de coluna (20 VC), com fluxo de 1,0 mL/min e recolhimento das frações de 2 mL. As frações cromatográficas equivalentes aos picos de absorbância a 280 nm com intensidade maior que 150 mAU foram submetidas à atividade enzimática em substratos apropriados visando à confirmação da presença das proteínas de interesse nas frações avaliadas.

As frações positivas para os testes de atividade enzimática foram analisadas por SDS-PAGE.

No caso da necessidade de realização de mais uma etapa cromatográfica, as frações foram concentradas em concentrador do tipo VivaSpin com 10 kDa de corte (Ge Healthcare) até o volume final de 1 mL. A concentração teve como objetivo o preparo da amostra para a cromatografia de filtração em gel onde as proteínas são separadas em função do peso molecular. Nesta etapa, 1 mL da amostra foi aplicado na coluna de filtração em gel Superdex 10/300 75 GL (Gel Filtration Column – GE Healthcare), previamente equilibrada com 36 mL de tampão fosfato de sódio 20 mM e NaCl 50 mM, pH 7,4. A eluição foi realizada em um volume final de 36 mL, a 0,5 mL/min e recolhimento de frações de 2 mL. As frações equivalentes aos picos de absorbância a 280 nm com intensidade maior que 150 mAU foram submetidas à atividade enzimática

em substrato adequado visando confirmar a presença da proteína de interesse na fração avaliada. O grau de pureza foi avaliado a partir de SDS-PAGE.

Nos casos em que não houve a necessidade da realização da cromatografia de filtração em gel, as amostras foram dialisadas para a remoção do imidazol em concentrador tipo VivaSpin (GE).

Em alguns casos, visando proteger as amostras purificadas da degradação por proteólise, foram adicionados 10 µL de uma solução contendo uma mistura de diferentes inibidores de proteases 100 X concentrada (*Protease Inhibitor Cocktail* – Amresco) para cada 1 mL de amostra proteica.

#### **4.3.13. Zimografia**

A zimografia é uma técnica eletroforética que consiste na copolimerização do gel de poliacrilamida com um substrato de interesse, visando à detecção de determinada atividade enzimática em uma amostra proteica (CHOI et al., 2009).

Neste caso, foi utilizado gel de poliacrilamida 12% contendo 0,1% (m/v) de CMC. Antes de serem aplicadas no gel, as amostras foram misturadas com tampão Laemmli 3 X e aquecidas a 99°C por 5 minutos, para a desnaturação parcial das enzimas. A corrida em SDS-PAGE foi feita por 2 horas a 110 V. Terminado esse tempo, o gel foi lavado com uma solução 0,1% (v/v) de Triton X-100 por 30 minutos. Essa lavagem teve como objetivo a remoção do SDS, facilitando o reenovelamento das enzimas para que fossem capazes de recuperar sua conformação ativa. Posteriormente, o gel foi submetido a duas novas lavagens, de 30 minutos cada, com tampão acetato de sódio 0,1 M, pH 5,5, adequado para a atividade enzimática. Em seguida, o gel foi incubado a 50°C, em solução de vermelho congo 0,1% por 30 minutos, para visualização da formação de halos de degradação.

#### **4.3.14. Ensaio de ligação a substratos insolúveis**

Como forma de avaliar a presença de um provável domínio de ligação a carboidratos (CBM) na celulase E-1, foram realizados ensaios de ligação a substratos insolúveis de acordo com o proposto por Okazaki et al., 2002, com algumas

modificações. O experimento consistiu basicamente na incubação de aproximadamente 1 µg de proteína com 200 µL de solução contendo 30 mg/mL de substrato insolúvel dissolvido em tampão acetato de amônio 25 mM, pH 5,0, por 20 minutos, a 8°C, sob agitação constante de 1.000 rpm. Transcorrido o período de incubação, o tubo contendo a reação foi centrifugado a 12.000 x g por 15 minutos. O sobrenadante foi coletado e concentrado a vácuo (*Concentrator plus* – Eppendorf), até que o volume de 20 µL fosse atingido. Posteriormente, foram adicionados 10 µL de tampão Laemmli (3X), e a amostra foi aquecida por 5 minutos a 99°C para avaliação por SDS-PAGE. A fração insolúvel foi lavada três vezes com 1 mL de tampão acetato de amônio 25 mM, pH 5,0 contendo 1 M de NaCl. Após as lavagens, foram adicionados 100 µL do tampão Laemmli (3X) ao precipitado para a desnaturação das proteínas aderidas através de incubação a 99°C com agitação de 1.000 rpm por 5 minutos. Após a incubação a mistura foi centrifugada por 10 minutos a 12.000 x g e o sobrenadante coletado para ser analisado por SDS-PAGE.

A capacidade da enzima em se ligar ou não ao substrato insolúvel foi avaliada pelo aparecimento da banda correspondente à enzima na fração solúvel (não ligada ao substrato) ou na fração insolúvel (ligada ao substrato).

#### **4.3.15. Preparo da amostra para análise em LC-MS/MS**

A espectrometria de massas é uma técnica muito utilizada para a identificação de moléculas, como por exemplo, proteínas, puras ou em misturas, através da ionização e posterior separação pela razão massa/carga ( $m/z$ ) desses íons.

Neste trabalho, a plataforma utilizada foi a de espectrometria de massas em tandem (MS/MS) acoplada à cromatografia líquida de ultra-performance (uPLC-MS/MS ou uPLC-QUAD-TOF).

Para análise de proteínas por espectrometria de massas, primeiramente a amostra proteica foi submetida à separação por SDS-PAGE. Após a eletroforese e visualização das bandas através da coloração com *coomassie blue*, recortou-se as regiões do gel correspondente às proteínas de interesse (cortes de aproximadamente 1 mm). As bandas foram incubadas com solução descorante (50% metanol, 2,5% ácido acético) por três horas à temperatura ambiente (TA). Posteriormente, a solução

descolorante foi removida e o gel desidratado através de duas lavagens com 200  $\mu\text{L}$  de acetonitrila (100%) por cinco minutos. Em seguida, foi realizada a etapa de redução com a adição de 30  $\mu\text{L}$  de ditionitrosol (DTT) 50 mM dissolvido em bicarbonato de amônio 100 mM e incubação por 30 minutos à TA. Posteriormente, a solução de DTT foi removida, seguida por adição de 30  $\mu\text{L}$  de solução de iodoacetamida (IAA) 50 mM dissolvida em bicarbonato de amônio 100 mM, e incubada por 30 minutos à TA. A solução de IAA foi removida, e o gel foi lavado com 100  $\mu\text{L}$  de bicarbonato de amônio 100 mM por 10 minutos. Em seguida, a solução foi removida e o gel novamente desidratado com 200  $\mu\text{L}$  de acetonitrila (100%) por 5 minutos à TA. Após a remoção da acetonitrila, o gel foi reidratado com 200  $\mu\text{L}$  de bicarbonato de amônio 100 mM por 10 minutos. A solução foi removida e o gel novamente desidratado com 200  $\mu\text{L}$  de acetonitrila (100%) por 5 minutos a TA. Este passo foi repetido mais uma vez. O restante da acetonitrila foi evaporado e em seguida foram adicionados 30  $\mu\text{L}$  da solução de tripsina (20 ng/ $\mu\text{L}$  dissolvida em bicarbonato de amônio 50 mM) para incubação por 30 minutos em gelo. Esta etapa teve como função a reidratação, permitindo o acesso da tripsina às proteínas presentes no gel. Transcorrido esse tempo, o excesso de tripsina foi removido e o gel foi coberto com 20  $\mu\text{L}$  de bicarbonato de amônio 50 mM e incubado a 37°C por aproximadamente 16 horas.

Em seguida, foram adicionados aos tubos 10  $\mu\text{L}$  da solução 1 de extração (5% ácido fórmico em água ultrapura estéril) por 10 minutos à TA, seguido por um *spin* rápido, coleta e passagem do sobrenadante para um novo tubo. Posteriormente, foram adicionados 12  $\mu\text{L}$  da solução 2 de extração (5% ácido fórmico em acetonitrila 50%) por 10 minutos à TA, seguido de um *spin* rápido, coleta e passagem do sobrenadante para o tubo que continha o extrato do passo anterior. Este passo foi repetido por mais uma vez. Os extratos obtidos foram evaporados até atingir o volume final de aproximadamente 1  $\mu\text{L}$ .

Para aplicação das amostras tripsinizadas no LC-MS/MS cada amostra foi ressuspensa em 12  $\mu\text{L}$  de ácido fórmico 0,1%. Uma alíquota de 4,5  $\mu\text{L}$  da mistura de peptídeos foi injetada no cromatógrafo RP-nanoUPLC (nanoAcquity, Waters). A cromatografia foi realizada em coluna C18 (100  $\mu\text{m}$  x 100 mm), de hidrofobicidade, equilibrada em tampão com 0,1% de ácido fórmico. O gradiente de eluição foi de 2% a

90% de acetonitrila em 0,1% de ácido fórmico, com um fluxo de 0,6 µL/min. À medida que os peptídeos eram eluídos da coluna, os mesmos eram injetados no espectrômetro quadrupolo-tempo de vôo Q-Tof (Ultima Mass Spectrometer Waters) com fonte de ionização do tipo elétron spray durante 60 minutos. O instrumento foi operado no modo “top three- MS e MS/MS”, onde para cada espectro de MS adquirido, os três peptídeos não monocarregados (precursores) mais abundantes foram selecionados para posterior fragmentação (gerando séries y e b) e sequenciados gerando um espectro MS/MS de cada peptídeo.

Os espectros foram obtidos usando o software *MassLynx v.4.1* (Waters-Milford, MA, EUA) e os dados brutos foram convertidos para o formato “peak list format (mgf)”, pelo software *Mascot Distiller v.2.3.2.0 2009* (Matrix Science Ltd.). Esses resultados foram processados pelo software *Mascot v.2.3.0.1 Engine* (Matrix Science Ltd.), utilizando como banco de dados a sequência de aminoácidos da proteína alvo para a digestão e geração dos espectros de MS/MS teóricos. Os parâmetros utilizados nesse processamento foram: carbamidometilação como uma modificação fixa, oxidação da metionina como modificação variável, um erro de clivagem para a tripsina e tolerância altíssima de 0,1 Da de razão  $m/z$  para o íon precursor (selecionado no modo MS) e o íon fragmentado (selecionado para fragmentação no modo MS/MS). Somente peptídeos com no mínimo cinco aminoácidos e com *score* dentro da probabilidade de não ser um evento aleatório ( $p < 0,05$ ) foram selecionados como produto da clivagem peptídica (ou seja, são parte de uma proteína).

As análises de LC-MS/MS foram realizadas no Laboratório de Espectrometria de Massas, situado no Laboratório Nacional de Biociências (LNBio), sob supervisão da Dra. Adriana Paes Leme.

### **4.3.16. Caracterização bioquímica**

#### **4.3.16.1. Quantificação de proteínas**

As quantificações proteicas foram realizadas através da lei de *Lambert-Beer* (equação 3):

$$A=e.b.c \text{ (equação 3)}$$

Onde **A** se refere à absorvância a 280 nm (medição feita no equipamento Nanodrop, modelo 2000c), **e** equivale à absortividade molar em  $L.mol^{-1}cm^{-1}$ , **b** é o comprimento do caminho que a luz tem que atravessar na solução em centímetros e **c** é a concentração do elemento que absorve na solução, em  $mol.L^{-1}$ . Os valores de absortividades molares foram calculados através da ferramenta ProtParam (disponível em: <http://web.expasy.org/protparam/>).

#### 4.3.16.2. Determinação da atividade enzimática

A avaliação da atividade enzimática foi realizada em diferentes polissacarídeos naturais e sintéticos. Os substratos testados foram: CMC, xilana de madeira de faia, manana, pectina (todos da Sigma-Aldrich),  $\beta$ -glucano, galactomanana, xiloglucano, arabinoxilana de centeio, arabinoxilana de trigo, liquenana e laminarina (todos da Megazyme).

Os ensaios foram realizados em triplicata em reações com volume final de 100  $\mu L$ , sendo: 40  $\mu L$  de tampão acetato de sódio 100 mM pH 5,5, 50  $\mu L$  do substrato 0,5% (m/v) dissolvido em água ultrapura estéril e 10  $\mu L$  da enzima na concentração de aproximadamente 10 ng/ $\mu L$ .

Cada reação possuía seu respectivo controle, e todos os ensaios foram realizados em placas de PCR de 96 poços no gelo. As reações foram incubadas em termociclador (MasterCycler pro-S-Eppendorf) por 30 minutos a 40°C. Posteriormente, foram adicionados 100  $\mu L$  de ADNS para a revelação dos açúcares redutores e 10  $\mu L$  da solução enzimática nas reações controle. Em seguida, as reações foram incubadas a 99°C por 5 minutos. Transcorrido esse tempo, 100  $\mu L$  de cada reação foram transferidos para placas de cultura de fundo chato de 96 poços, para medição da absorvância a 540 nm. Através da curva padrão de glicose ( $R^2= 0,999$ ) foi possível o cálculo da quantidade de açúcares totais liberados após a reação enzimática.

### **4.3.16.3. Determinação dos valores de pH e temperatura ótimos**

Os ensaios de determinação do pH e temperatura ótimos foram realizados em CMC 0,5% no caso da celulase e xilana de madeira de faia 0,5% para a xilanase. Na determinação do valor de pH ótimo, foram utilizados tampões citrato (100 mM) – fosfato (200 mM) (MCILVAINE, 1921), com valores de pH variando entre 2,0 e 8,0. As reações eram compostas por 40 µL de tampão citrato-fosfato no pH a ser avaliado, 50 µL do substrato 0,5% (m/v) dissolvido em água ultrapura estéril e 10 µL da enzima na concentração aproximada de 10 ng/µL. As reações foram incubadas a 37°C por 20 minutos no caso da celulase e por 5 minutos a 37°C no caso da xilanase. Em seguida os açúcares redutores liberados foram quantificados pelo método do ADNS.

Os ensaios para determinação dos valores de temperatura ótima foram realizados em 40 µL de tampão acetado de sódio 100 mM no pH ótimo de atuação, 50 µL do substrato e 10 µL da enzima na concentração aproximada de 10 ng/µL. As reações foram incubadas em temperaturas variando de 15°C a 90°C por 20 minutos no caso da celulase e de 10°C a 90°C por 5 minutos no caso da xilanase.

### **4.3.16.4. Determinação da energia de ativação**

Utilizando-se a parte ascendente do gráfico de velocidade em função da temperatura foram determinadas as energias de ativação ( $E_a$ ) das reações enzimáticas através da equação de Arrhenius (equação 4):

$$k = A.e^{(-E_a/RT)} \text{ (equação 4)}$$

Onde,

$k$  = constante de velocidade

$E_a$  = Energia de ativação

$R$  = Constante real dos gases (8,314472 J.K<sup>-1</sup>. mol<sup>-1</sup>)

T = Temperatura absoluta (K)

A equação 4 pode ser transformada em:

$$\text{Log } k = \text{log}A - (E_a/2,3R)/T$$

Sendo assim, através do gráfico de Arrhenius ( $\log k$  em função de  $T^{-1}$ ) e determinação da inclinação da reta obtida foi possível o cálculo do valor de  $E_a$ .

#### **4.3.16.5. Determinação de parâmetros cinéticos**

Entre os parâmetros cinéticos calculados estão  $K_m$  e  $k_{cat}$ .  $K_m$  equivale à concentração de substrato na qual a velocidade inicial ( $V_0$ ) é igual à metade da velocidade máxima ( $V_{m\acute{a}x}$ ). Este parâmetro indica a afinidade da enzima pelo substrato, sendo que quanto menor o valor de  $K_m$  maior é a afinidade da enzima pelo substrato. A constante  $k_{cat}$ , definida como  $V_{m\acute{a}x}/[E_t]$  onde  $[E_t]$  representa a concentração de enzima total, descreve a velocidade limitante de qualquer reação catalisada por uma enzima nas condições de saturação. Por fim, a relação  $k_{cat}/K_m$ , chamada de eficiência catalítica, relaciona a constante catalítica da enzima com a afinidade pelo substrato.

As reações foram realizadas como descrito no item 4.3.16.2, nos valores de temperatura e pH ótimos, em diferentes concentrações de substratos, com tempo de incubação de 5 minutos.

Os parâmetros foram calculados com auxílio do programa *GraphPad Prism 6*, através da construção de uma curva de regressão não linear.

#### **4.3.16.6. Estabilidade térmica**

Os experimentos de avaliação da estabilidade térmica das enzimas recombinantes foram realizados através da incubação da enzima em temperaturas variando de 80°C a 40°C em uma faixa de trinta segundos a seis horas. Em determinados intervalos de tempo uma alíquota da enzima era retirada e submetida à reação enzimática com o substrato adequado, nas condições de pH e temperatura

ótimos.

Dessa forma, foi possível a determinação da perda da atividade enzimática em função da temperatura e do tempo de incubação.

#### **4.3.16.7. Efeitos de íons, sais e agentes quelantes**

Os efeitos de diferentes íons e sais foram avaliados pela adição de soluções desses compostos às reações na concentração final de 1 mM no caso da celulase (FERCHAK & PYE, 1983) e 5 mM no caso da xilanase (MO et al., 2010). Os agentes quelantes EDTA e EGTA foram avaliados na concentração final de 10 mM.

#### **4.3.16.8. Avaliação da atividade enzimática em substratos complexos**

No caso da celulase E-1, a atividade enzimática foi avaliada em bagaço de cana-de-açúcar explodido a vapor (BEX) e Avicel PH-101 (Sigma-Aldrich), um tipo de celulose microcristalina. Uma amostra de BEX foi gentilmente cedida pelo professor Dr. George Jackson, sendo composta por: 51,7% de celulose, 8,9% de hemicelulose, 34,3% de lignina e 5,5% de cinzas. As reações foram feitas em triplicada, em microtubos de 2 mL, contendo 1 mL de tampão fosfato de sódio 200 mM, pH 7,0, 10 µg da proteína recombinante e 1,0% (m/v) do substrato insolúvel. As reações foram incubadas por 200 minutos à 50°C, em agitação constante de 1.000 rpm. Após a incubação, os tubos foram submetidos à centrifugação e o sobrenadante foi separado para quantificação dos açúcares liberados pelo método do ADNS. Nesses ensaios, 1 U de atividade foi definida como a quantidade de enzima que converte 1 nmol de substrato por minuto.

No caso da xilanase X-1, a atividade enzimática foi avaliada em bagaço de cana-de-açúcar previamente submetido ao tratamento com ácido fosfórico 2,5% para afrouxamento das fibras (gentilmente cedido pelo Prof. Dr. George Jackson), com a seguinte composição: 48,5% de celulose, 17,0% de hemicelulose, 28,8% de lignina e 3,8% de cinzas. Neste caso, as reações feitas em triplicata, contendo 1,0% de bagaço (massa/volume) em 1 mL de tampão acetato de sódio 100 mM, pH 6,0, foram

incubadas com 10 µg da proteína recombinante, a 40°C, por 24 horas, sob agitação constante. Transcorrido o período de incubação, o impacto da atividade enzimática de X-1 no aumento da digestibilidade das fibras de celulose foi avaliado através da adição de 1,4 µg do complexo celulolítico comercial ACCELERASE® 1500 (Genencor). As reações foram incubadas a 40°C por 23,5 horas. Posteriormente o sobrenadante foi separado da biomassa residual e os açúcares redutores liberados foram quantificados pelo método do ADNS.

Os tempos e temperaturas de incubação das reações foram determinados de acordo com os resultados obtidos nos experimentos de estabilidade térmica.

#### **4.3.16.9. Eletroforese capilar**

Os ensaios de eletroforese capilar foram realizados de acordo com a metodologia descrita no item 3.3.3.4.

### **4.4. Resultados e Discussão**

#### **4.4.1. Celulase E-1**

Através da reação de PCR foi amplificado um fragmento de DNA com o tamanho esperado de aproximadamente 1,2 kb (Figura 38A). Como se pode observar na Figura 38B, por digestão enzimática foi possível a confirmação da clonagem do fragmento de interesse no vetor pET-28a. O sequenciamento do inserto garantiu que nenhuma mutação ocorreu no gene e que ele foi clonado na fase de leitura correta.

Através da ferramenta *ProtParam* foram determinados alguns parâmetros bioquímicos e biofísicos da celulase E-1 (Tabela 12), como: peso molecular, ponto isoelétrico, coeficiente de extinção molar e índice de hidrofobicidade (onde valores negativos indicam que a proteína apresenta possibilidade de ser solúvel). Os dados foram gerados a partir da sequência de aminoácidos com a adição do segmento do vetor pET-28a que também foi traduzido.

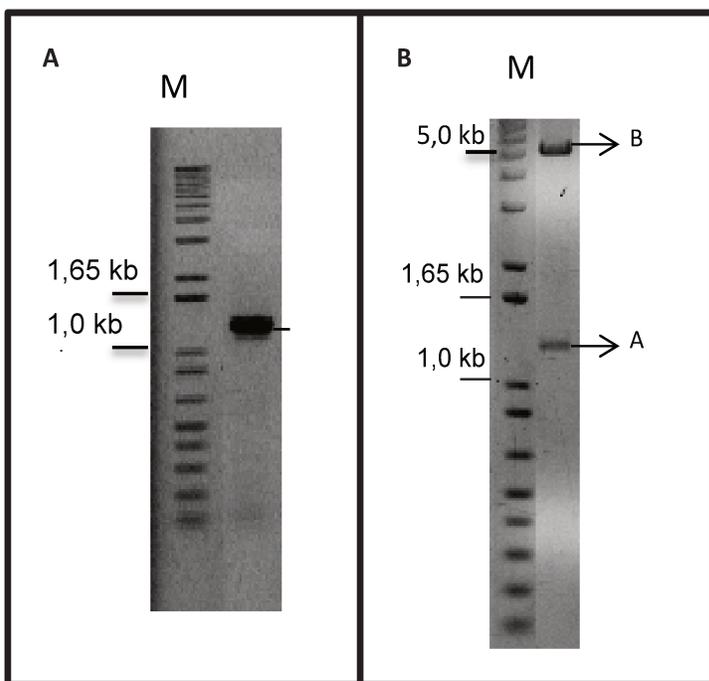


Figura 38 – A) Amplificação do fragmento correspondente ao gene da celulase E-1 por PCR. B) Confirmação da clonagem em pET-28a por digestão com enzimas de restrição, onde (A) corresponde ao fragmento liberado e (B) vetor pET-28a linearizado. M – Marcador de peso molecular 1 kb.

Tabela 12 - Parâmetros bioquímicos e biofísicos de E-1 calculados a partir do ProtParam.

Proteína	Número de aminoácidos	Peso molecular	Ponto isoelétrico	Coefficiente de Extinção Molar	Índice de hidrofobicidade
E-1	440	46,3 kDa	8,3	106340 M <sup>-1</sup> . cm <sup>-1</sup>	-0,266

Tendo-se confirmado a clonagem em pET-28a, os plasmídeos recombinantes foram transformados nas linhagens: BL21(DE3), Rosetta2 (BL21) e ArticExpress(DE3).

Após a expressão e lise das células para liberação das proteínas, observou-se por SDS-PAGE que as expressões em BL21(DE3) e Rosetta2(DE3) não apresentaram resultados satisfatórios, pois nestes casos a maior parte da enzima foi encontrada na

fração insolúvel enquanto uma pequena porção permaneceu na fração solúvel (dados não mostrados).

O melhor resultado em relação à solubilidade da proteína de interesse foi obtido na expressão com a linhagem ArticExpress(DE3). Neste caso, a expressão foi acompanhada por alíquotas retiradas no tempo zero (To na Figura 39, antes da adição do IPTG) e no tempo final (Tf na Figura 39, após 4 horas de expressão). O resultado do teste de expressão foi avaliado por SDS-PAGE, onde além das amostras de To e Tf também foram aplicadas as amostras correspondentes à fração solúvel do lisado (canaleta S na Figura 39) e insolúvel do lisado (canaleta IS na Figura 39). A análise do gel de SDS-PAGE resultou na identificação de uma banda de expressão com aproximadamente 45 kDa na fração solúvel, com tamanho próximo do predito pela ferramenta ProtParam (Tabela 12).

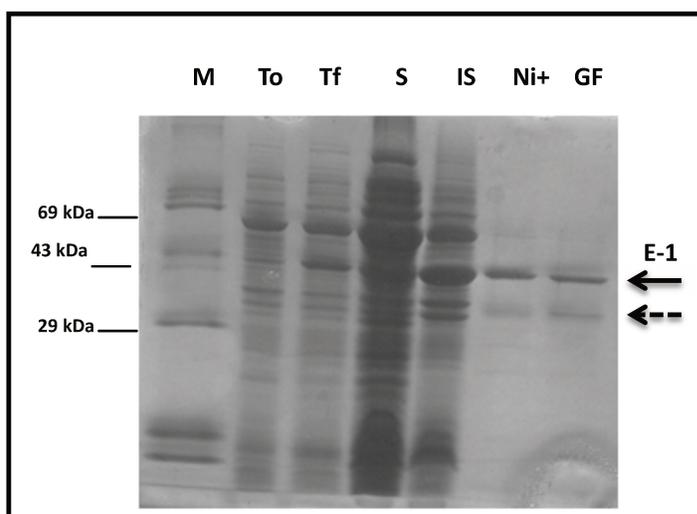


Figura 39 - Análise por SDS-PAGE do teste de expressão em ArticExpress(DE3) e o resultado da purificação da celulase E-1. M - marcador de peso molecular. To - tempo zero da expressão. Tf - tempo final da expressão. S - fração solúvel do lisado. IS - fração insolúvel do lisado. Ni<sup>+</sup> - amostra resultante da cromatografia de afinidade. GF - amostra resultante da cromatografia de gel filtração. A seta pontilhada indica a banda de aproximadamente 38 kDa presente nas amostras de cromatografia de afinidade e gel filtração.

O passo seguinte foi a expressão em um volume de 2 L, visando a obtenção de maiores quantidades de enzima. Em seguida, a purificação proteica foi realizada por

cromatografia de afinidade em matriz de níquel (Ni<sup>2+</sup>- IMAC). A análise em SDS-PAGE da fração resultante da purificação (canaleta Ni<sup>+</sup> na Figura 39) resultou na identificação de outras bandas além da banda de interesse, demonstrando a necessidade da realização de uma etapa adicional de purificação. As frações resultantes da cromatografia de afinidade foram concentradas e aplicadas na coluna Superdex 10/300 75 GL. As frações recolhidas foram confirmadas quanto à atividade enzimática, concentradas em VivaSpin (corte de 10 kDa) e avaliadas por SDS-PAGE (canaleta GF na Figura 39).

Como se pode observar na Figura 39, apesar do processo de gel filtração ter auxiliado na purificação, uma banda de aproximadamente 38 kDa continuou presente na amostra (seta pontilhada na Figura 39).

#### **4.4.1.1. Avaliação da capacidade da celulase E-1 em se ligar a substratos insolúveis**

Uma maneira de verificar a existência de CBMs em proteínas é através da avaliação da capacidade de ligação a substratos insolúveis. Para isso, a amostra purificada de E-1 foi incubada com os seguintes substratos insolúveis: bagaço de cana-de-açúcar *in natura* e Sigmacell 20 (Sigma-Aldrich, tamanho de partícula de 20 µm), um tipo de celulose cristalina. Como forma de validar a metodologia, utilizou-se uma endoglucanase sem CBM (isolada de *Coptotermes gestroi*), como controle negativo.

Através de gel de SDS-PAGE (Figura 40), observou-se que a maior parte da enzima E-1 adicionada (correspondente à banda de aproximadamente 45 kDa) permaneceu na fração insolúvel, demonstrando a capacidade da enzima em se ligar ao substrato. Uma pequena fração da enzima foi detectada na fração solúvel, possivelmente devido à quantidade de proteína adicionada em excesso no experimento. Como esperado, na reação controle a enzima sem CBM permaneceu na fração solúvel, já que não possuía em sua estrutura um domínio que permitisse sua ligação ao substrato insolúvel.

Além da banda de 45 kDa corresponde a E-1, se observou na fração solúvel a presença de outra banda com tamanho aproximado de 38 kDa (indicada pela seta na

Figura 40). Como forma de investigar a função de ambas as bandas, uma amostra de E-1 foi submetida à zimografia em gel copolimerizado com CMC. Pelo SDS-PAGE convencional (canaleta 1, Figura 41) pôde-se observar a presença de duas bandas com tamanhos aproximados de 45 kDa (banda A), 38 kDa (banda B) e uma terceira (banda C) com tamanho menor que 13 kDa. Através do zimograma em gel de CMC (Figura 41, canaleta 2) foi observado que ambas as bandas, de 38 kDa e 45 kDa, apresentavam atividade celulolítica.

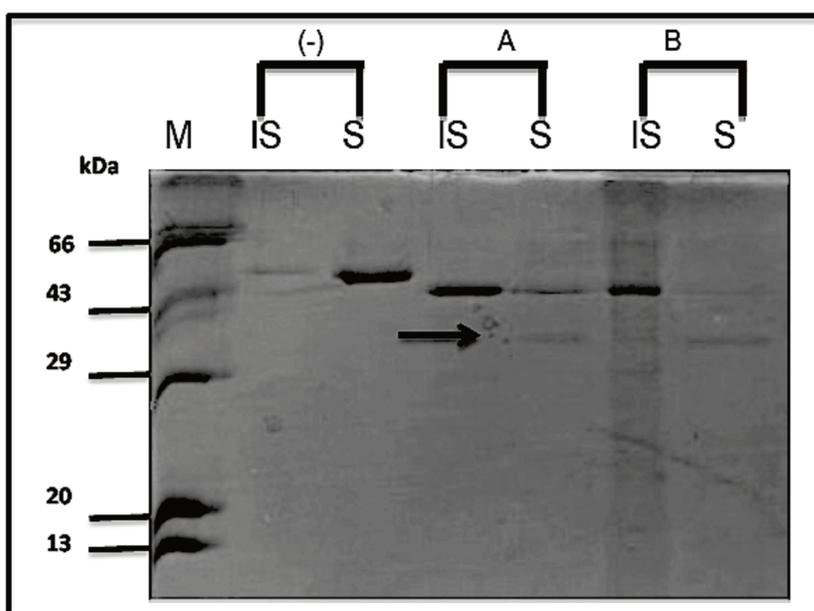


Figura 40 – Análise por SDS-PAGE dos ensaios de ligação de E-1 a substratos insolúveis. M – marcador de peso molecular, (-) controle negativo (endoglucanase sem CBM) em Sigmacell. A – Ensaio com E-1 em Sigmacell, (B) Ensaio com E-1 em bagaço de cana-de-açúcar *in natura*, IS - Proteínas identificadas na fração insolúvel e S - Proteínas identificadas - fração solúvel. A seta indica a presença de uma banda com aproximadamente 38 kDa na fração solúvel.

Cada uma das três bandas foi analisada separadamente por espectrometria de massas. A Figura 42, mostra o resultado obtido através da comparação dos peptídeos identificados em cada banda (em vermelho) com a sequência de aminoácidos de E-1 (em preto). Em cinza está destacada a região rica em resíduos de serina.

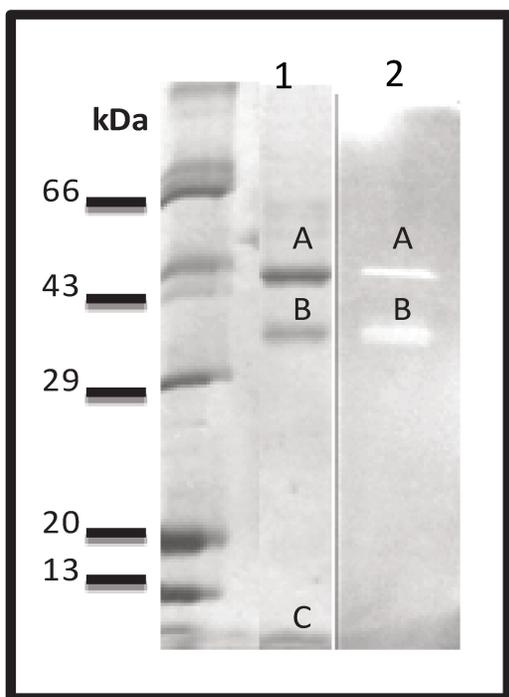


Figura 41 – 1 – Análise da amostra purificada da proteína E-1 por SDS-PAGE , 2 – Análise da amostra purificada da proteína E-1 por zimografia em gel com CMC.

A banda de aproximadamente 45 kDa (banda A na Figura 41), correspondente à enzima E-1, apresentou peptídeos com identidade para a região N-terminal e C-terminal da proteína. Já a banda de 38 kDa (banda B na Figura 41), apresentou peptídeos com identidade apenas para a região N-terminal da proteína, na qual foi encontrado o domínio conservado da família GH5 e que compreende o domínio catalítico de E-1. A banda C (Figura 41) apresentou identidade com sequências da região C-terminal de E-1, correspondente ao provável CBM. Esses resultados justificaram a detecção de duas bandas com atividade catalítica no zimograma, já que o domínio catalítico de E-1 foi identificado nas bandas A e B.

Já foi verificado que os *linkers* ricos em resíduos de serina são susceptíveis ao ataque de proteases, como tripsina e quimiotripsina, por serem regiões essencialmente desordenadas (WARD et al., 2004, FONTANA et al., 2004, STROH et al., 2005). Sendo assim, a separação entre os domínios, como verificada por zimografia, pode ser explicada pela ocorrência de proteólise na região correspondente ao *linker*, separando o domínio catalítico (banda de 38 kDa) do provável CBM (banda com tamanho inferior a

13 kDa). Foi observado que a adição de um coquetel de inibidores de protease à amostra de E-1 purificada resultou na diminuição da proteólise.

Ainda assim, a frequente ocorrência de proteólise na região rica em resíduos de serina, dificultou o estudo da enzima E-1 em sua forma completa. Como forma de avaliar as características do domínio catalítico (E-1 Cat) e do provável CBM (E-1 CBM) oligonucleotídeos iniciadores foram desenhados para a amplificação e clonagem desses domínios separadamente. Os dados referentes aos estudos com E-1 CBM estão descritos no capítulo 5.

Banda de 38 kDa				
MASAHAAVAP	ITTSGNKVLF	GGQQGSIAGN	SFFWSNTGNG	GEKYYNAQTV
AWLKSDWKSS	LVRRAAMGVDE	SGGYITDSYN	KTRVTTVVDA	AIANNMYVII
DWHSHHAEQY	QSQAIAFFKE	MATKYGNNNN	VIYEIYNEPL	QVSWSSVIKP
YATAVIAEIR	KIDPDNLIVV	GTPTWSQDVD	VAANDPITGY	ANIAYTLHFX
AGTHGQSLRN	KASTALSKGI	PLFVTEWGSV	NADGGGSVAT	AETNSWVSPM
KTNNISNANW	ALNDKAEAGS	ALVSGAXSAN	GGWTSSQLTA	SGTLAKSIIS
GWPGTNQNSS	SSSSSSKSSS	SIASSVKSSV	PSSSSSASCG	SGNPNKTAAK
GVEPSAVAGD	CIKYNKSSGT	LQIGSWTGVA	SSYNITSGPQ	GITNTGNGWT
TVANAANGDL	YIKIVSASRS	FNVKFDNW		
Banda com tamanho inferior a 13 kDa				
MASAHAAVAP	ITTSGNKVLF	GGQQGSIAGN	SFFWSNTGNG	GEKYYNAQTV
AWLKSDWKSS	LVRRAAMGVDE	SGGYITDSYN	KTRVTTVVDA	AIANNMYVII
DWHSHHAEQY	QSQAIAFFKE	MATKYGNNNN	VIYEIYNEPL	QVSWSSVIKP
YATAVIAEIR	KIDPDNLIVV	GTPTWSQDVD	VAANDPITGY	ANIAYTLHFX
AGTHGQSLRN	KASTALSKGI	PLFVTEWGSV	NADGGGSVAT	AETNSWVSPM
KTNNISNANW	ALNDKAEAGS	ALVSGAXSAN	GGWTSSQLTA	SGTLAKSIIS
GWPGTNQNSS	SSSSSSKSSS	SIASSVKSSV	PSSSSSASCG	SGNPNKTAAK
GVEPSAVAGD	CIKYNKSSGT	LQIGSWTGVA	SSYNITSGPQ	GITNTGNGWT
TVANAANGDL	YIKIVSASRS	FNVKFDNW		
Banda de 45 kDa				
MASAHAAVAP	ITTSGNKVLF	GGQQGSIAGN	SFFWSNTGNG	GEKYYNAQTV
AWLKSDWKSS	LVRRAAMGVDE	SGGYITDSYN	KTRVTTVVDA	AIANNMYVII
DWHSHHAEQY	QSQAIAFFKE	MATKYGNNNN	VIYEIYNEPL	QVSWSSVIKP
YATAVIAEIR	KIDPDNLIVV	GTPTWSQDVD	VAANDPITGY	ANIAYTLHFX
AGTHGQSLRN	KASTALSKGI	PLFVTEWGSV	NADGGGSVAT	AETNSWVSPM
KTNNISNANW	ALNDKAEAGS	ALVSGAXSAN	GGWTSSQLTA	SGTLAKSIIS
GWPGTNQNSS	SSSSSSKSSS	SIASSVKSSV	PSSSSSASCG	SGNPNKTAAK
GVEPSAVAGD	CIKYNKSSGT	LQIGSWTGVA	SSYNITSGPQ	GITNTGNGWT
TVANAANGDL	YIKIVSASRS	FNVKFDNW		

Figura 42 - Peptídeos identificados (em vermelho) a partir da análise por espectrometria de massas das bandas de 38 kDa, 45 kDa e da banda com tamanho inferior a 13 kDa, em relação a sequência completa de E-1 (em preto). A região marcada em cinza representa o provável *linker*, constituído por uma região rica em resíduos de serina.

#### 4.4.1.2. E1-Cat

Através da reação de PCR foi amplificado um fragmento de DNA com o tamanho esperado de aproximadamente 0,885 kb (Figura 43A). Como se pode observar na Figura 43B, por digestão enzimática foi possível a confirmação da clonagem do fragmento de interesse no vetor pET-28a. O sequenciamento do inserto garantiu que nenhuma mutação ocorreu no gene de interesse e que ele foi clonado na fase de leitura correta.

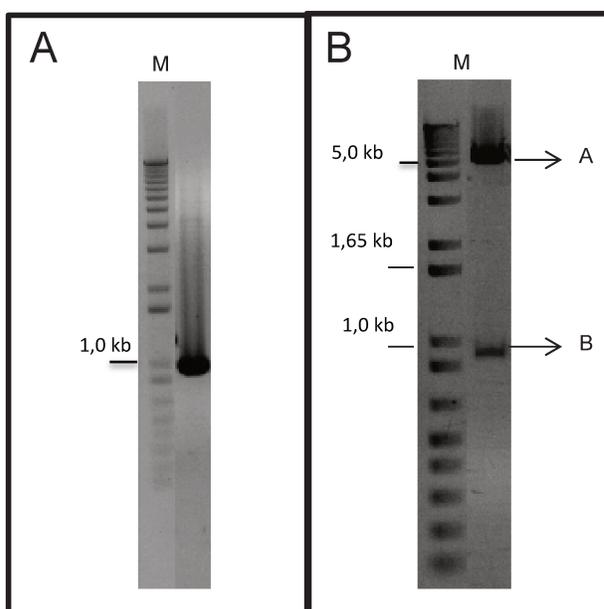


Figura 43 – (A) Amplificação do fragmento correspondente a gene que codifica E1-Cat por PCR. (B) Confirmação da clonagem em pET-28a por digestão com enzimas de restrição, onde (A) corresponde ao vetor pET-28a linearizado e (B) ao fragmento liberado.

Através da ferramenta *ProtParam* foram determinados alguns parâmetros bioquímicos e biofísicos da proteína E1-Cat (Tabela 13), como: peso molecular, ponto isoelétrico, coeficiente de extinção molar, índice de hidrofobicidade (onde valores negativos indicam que a proteína apresenta possibilidade de ser solúvel). Os dados foram gerados a partir da sequência clonada traduzida em aminoácidos com a adição do segmento do vetor pET-28a que também foi traduzido.

Tabela 13 - Parâmetros bioquímicos e biofísicos de E1-Cat calculados a partir do ProtParam.

<i>Proteína</i>	<i>Número de aminoácidos</i>	<i>Peso molecular</i>	<i>Ponto isoelétrico</i>	<i>Coefficiente de Extinção Molar</i>	<i>Índice de hidrofobicidade</i>
E-1-Cat	315	34 kDa	6,44	85370 M <sup>-1</sup> . cm <sup>-1</sup>	-0,241

Tendo-se confirmado a clonagem, os plasmídeos recombinantes foram transformados nas linhagens BL21(DE3), ArticExpress(DE3) e Rosetta-gami2(DE3)pLysS.

A melhor condição de expressão foi obtida com a linhagem Rosetta-gami2(DE3)pLysS com indução por 4 horas a 37°C. A Figura 44 mostra o resultado da purificação proteica em cromatografia de afinidade. Devido ao elevado grau de purificação obtido na primeira etapa de purificação, não foi necessária à realização da cromatografia de filtração em gel. As frações resultantes da cromatografia de afinidade foram dialisadas em tampão fosfato de sódio 50 mM, NaCl 20 mM, pH7,4, para a remoção do imidazol.

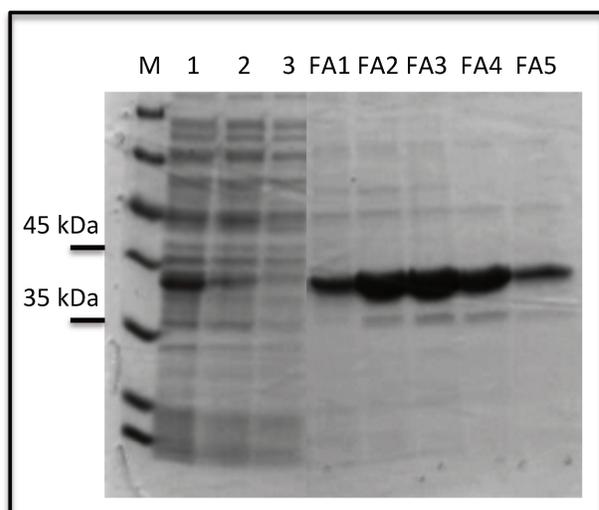


Figura 44 - Análise por SDS-PAGE da purificação de E1-Cat por cromatografia de afinidade. M - marcador de peso molecular. 1 – Extrato bruto, 2 – Amostra aplicada na coluna, 3 – Fração proteica que não ligou na coluna, FA-F5 – Frações correspondentes a E-1 Cat purificada.

#### 4.4.1.3. Caracterização bioquímica E1-Cat

Primeiramente, E1-Cat foi avaliada quanto à capacidade de hidrolisar diferentes ligações glicosídicas. Para isso, sua atividade enzimática foi testada frente a diferentes substratos: CMC,  $\beta$ -glucano, galactomanana, xiloglucano, xilana de madeira de faia, arabinoxilana de centeio, manana, liquenana, laminarina e pectina.

E1-Cat apresentou maior atividade em liquenana (atividade relativa de 100%),  $\beta$ -glucano (98,63%) seguido por CMC (33,82%). A liquenana e o  $\beta$ -glucano são polissacarídeos constituídos por ligações do tipo  $\beta(1\rightarrow4)$  e  $\beta(1\rightarrow3)$ , entre resíduos de glicose, na proporção aproximada de 2:1 em liquenana e de 2:1 ou 3:1 em  $\beta$ -glucano (informações fornecidas pelo fabricante). Já a CMC é constituída exclusivamente por ligações do tipo  $\beta(1\rightarrow4)$  entre resíduos de glicose. O fato de E1-Cat não apresentar atividade em laminarina mostrou que ela não foi capaz de hidrolisar ligações do tipo  $\beta(1\rightarrow3)$  e  $\beta(1\rightarrow6)$  entre resíduos de glicose.

O substrato CMC é extensamente utilizado para a caracterização enzimática de endoglucanases (JAVED et al., 2009, ZHOU & INGRAM, 2000, SADHU et al., 2013), por apresentar estrutura basicamente constituída por regiões amorfas (NGUYEN & LE, 2013), que são susceptíveis ao ataque enzimático por endoglucanases (FUJITA et al., 2004). Em virtude disso, a CMC foi escolhida como o substrato a ser empregado nas etapas seguintes de caracterização enzimática.

Os ensaios para determinação do pH ótimo foram realizados em CMC 0,5%. Através da Figura 45, pode se observar que o ótimo de atividade foi atingido em valores em torno do pH 7,0. Nos pHs 2,0, 3,0 e 12,0 a enzima sofreu desnaturação, justificando a ausência de atividade enzimática. Nos pHs 6,0 e 9,0 a enzima conseguiu manter a atividade acima de 70% de seu valor ótimo. Já foi reportado que a maioria das celulasas apresenta sua atividade ótima em faixas de pH que variam de meio levemente ácido (5,0, 6,0) até meio neutro (7,0) (FUKUMORI et al., 1986).

Através do ensaio de temperatura ótima (Figura 46), realizado em CMC 0,5% e pH 7,0, foi verificado que os maiores valores de atividade enzimática foram obtidos em temperaturas próximas a 50°C. A enzima foi capaz de manter 60% de sua atividade ótima na faixa de 15°C a 70°C. O valor de temperatura ótima encontrado para E-1

apresentou valores intermediários em comparação ao verificado para a celulase Cel5A de *Paenibacillus xylanilyticus* KJ-03 (PARK et al., 2012), com ótimo próximo a 40°C e para as celulases isoladas do metagenoma de digestor anaeróbico, que apresentaram ótimo entre 60°C e 70°C (HEALY et al., 1995).

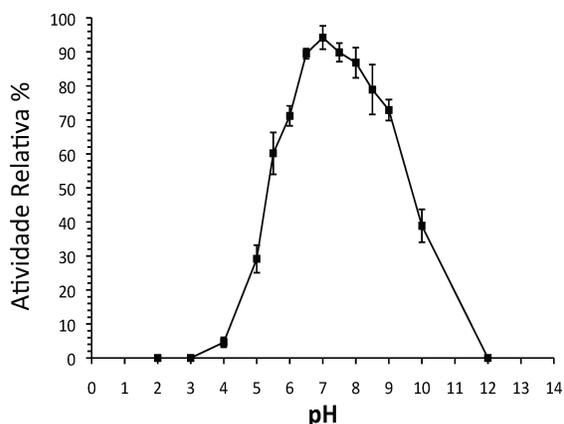


Figura 45 - Determinação do valor de pH ótimo de atividade de E1-Cat em reações contendo 0,5% CMC em tampão 100 mM citrato – 200 mM fosfato, com pHs variando entre 2,0 e 12,0. As reações foram incubadas a 37°C por 20 minutos e reveladas pelo método do ADNS.

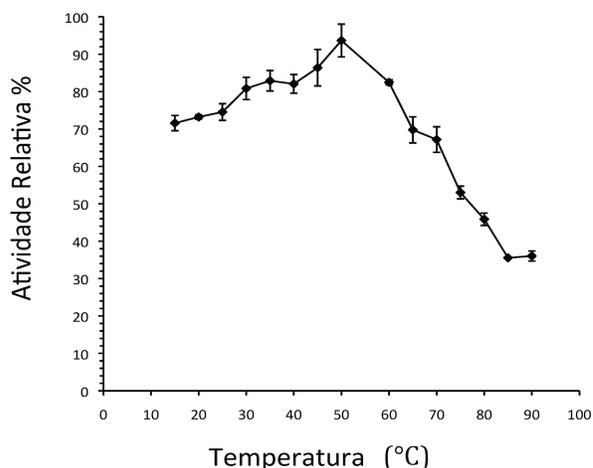


Figura 46 - Determinação do valor de temperatura ótima de atividade de E1-Cat em reações contendo 0,5% de CMC em tampão acetato de sódio 100 mM pH 7,0. As reações foram incubadas por 20 minutos em temperaturas variando de 15°C a 90°C e em seguida reveladas pelo método do ADNS.

A partir dos dados obtidos no experimento de temperatura ótima, e seguindo a equação de Arrhenius (equação 4) foi possível a determinação da energia de ativação ( $E_a$ ) envolvida na hidrólise de CMC pela atividade de E1-Cat. O valor de  $E_a$  calculado foi de  $5,3 \text{ kJ.mol}^{-1}$ , que está dentro da faixa de  $3\text{-}50 \text{ kJ.mol}^{-1}$  relacionada às energias de ativação dos processos de hidrólise enzimática da celulose (SHUAI & PAN, 2012). Os processos de degradação de celulose baseados na hidrólise ácida envolvem energias de ativação superiores, na faixa de  $170\text{-}180 \text{ kJ.mol}^{-1}$ , o que justifica as elevadas temperaturas empregadas em tais processos (SHUAI & PAN, 2012).

Os parâmetros cinéticos foram determinados a partir do cálculo da atividade enzimática em concentrações de CMC variando de  $0,625 \text{ mg/mL}$  a  $54 \text{ mg/mL}$ . Os valores obtidos foram:  $K_m$  de  $6,05 \pm 0,37 \text{ mg/mL}$ ,  $V_{max}$  de  $42,51 \pm 1,2 \text{ } \mu\text{mol/min/mg}$ ,  $k_{cat}$  de  $24,54 \text{ s}^{-1}$  e  $k_{cat}/K_m$  de  $4,06 \text{ mL/mg/s}$ . A Figura 47 exibe a curva de cinética enzimática obtida, evidenciando o perfil de Michaelis-Menten com um  $R^2$  de 0.95. A eficiência catalítica ( $k_{cat}/K_m$ ) obtida neste caso foi superior ao valor de  $2,0 \pm 0,5 \text{ mL/mg/s}$  obtido por Berlemont *et al* na caracterização da celulase GH5 Rbccl1 isolada de uma biblioteca metagenômica construída com DNA extraído de uma amostra de solo da Antártida (BERLEMONT *et al.*, 2009).

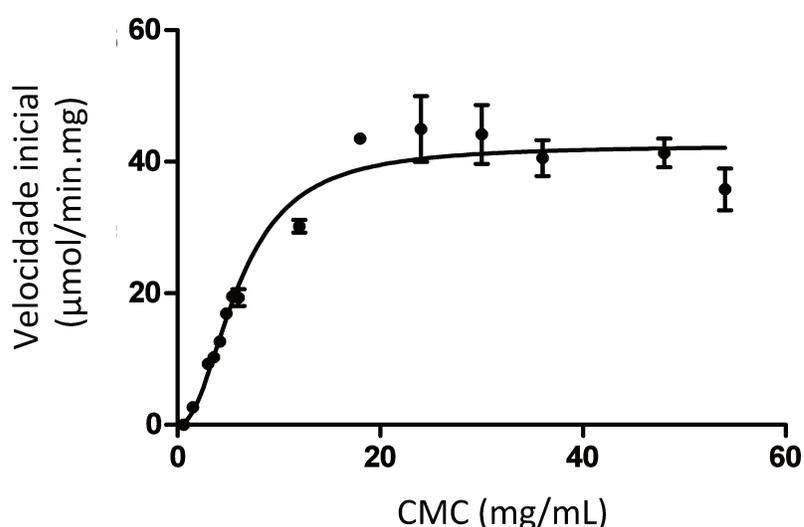


Figura 47 – Gráfico de Michaelis-Menten para E-1 Cat em presença de CMC como substrato, com concentrações variando de  $0,625 \text{ mg/mL}$  a  $54 \text{ mg/mL}$ . As reações foram realizadas em tampão acetato de sódio  $100 \text{ mM}$   $\text{pH } 7,0$ , incubadas por 5 minutos a  $50^\circ\text{C}$  e em seguida reveladas pelo método do ADNS.

Em relação à estabilidade térmica, através da Figura 48, podemos observar que a incubação a 60°C e 70°C resultou na completa inativação da enzima em menos de 60 minutos. Na temperatura de atividade ótima (50°C), a enzima perdeu expressivamente atividade após cerca de 180 minutos. Já na temperatura de 40°C, condição em que a enzima apresentou 80% de sua atividade ótima, os valores de atividade foram mantidos relativamente estáveis após 300 minutos de incubação.

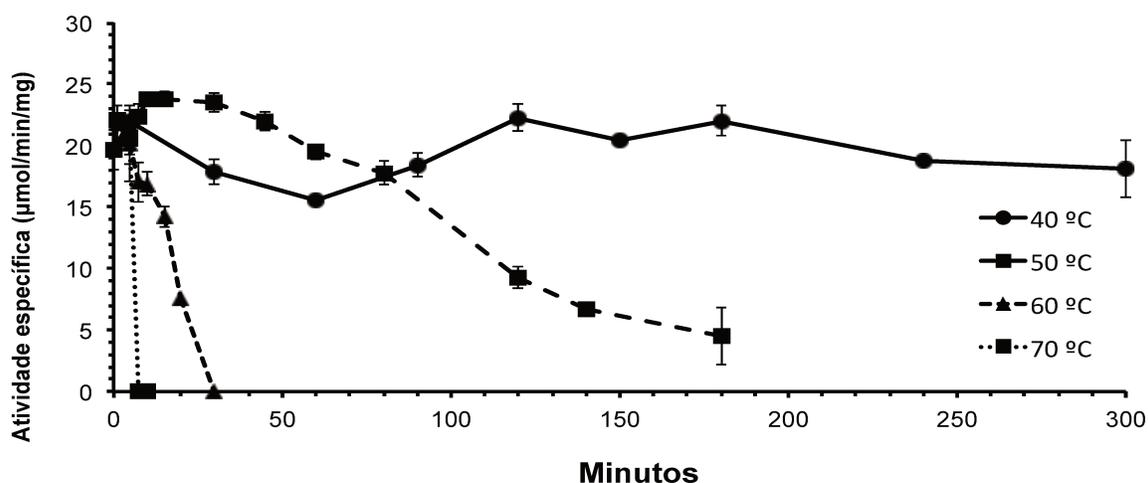


Figura 48 - Ensaio de estabilidade térmica de E1-Cat realizado com CMC 0,5% em tampão acetato de sódio 100 mM pH 7,0. Após a incubação, as reações foram reveladas pelo método do ADNS.

Os testes com íons e outros reagentes químicos (Tabela 14), mostraram que os íons  $\text{Cu}^{2+}$ ,  $\text{Fe}^{3+}$  e  $\text{Zn}^{2+}$  apresentaram efeitos altamente inibitórios. A inibição enzimática de celulasas da família GH5 na presença de íons metálicos, como  $\text{Cu}^{2+}$ ,  $\text{Fe}^{3+}$  e  $\text{Zn}^{2+}$  já foi reportada por outros trabalhos (BERA-MAILLET et al., 2000, RUBINI et al., 2010). Em geral, a inibição por cobre e zinco pode estar relacionada à interação dos íons com resíduos de cisteínas presentes nas proteínas (MEANS & FEENEY, 1971, KARLSTROM & LEVINE, 1991, PEI et al., 2000). No caso de E1-Cat, a inibição deve ocorrer por um mecanismo distinto, visto que a enzima não apresenta cisteína em sua sequência de aminoácidos. Já a presença do íon  $\text{Mn}^{2+}$  resultou em uma inibição parcial, reduzindo a atividade enzimática em aproximadamente 35%. Os demais íons avaliados,  $\text{Co}^{2+}$ ,  $\text{K}^+$ ,  $\text{Li}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Ni}^{2+}$  e  $\text{Ca}^{2+}$ , e o NaCl, não resultaram em inibição ou ativação significativas. A inibição enzimática parcial observada na presença de EDTA (quelante

de íons divalentes em geral) e EGTA (quelante específico de íons cálcio) sugere que a presença de íons divalentes nas reações resulta na ativação da enzima, mas que eles não são indispensáveis para a atividade enzimática (GONG et al., 2012, LOPRETE, 2002).

Tabela 14 - Efeito de íons e reagentes químicos na atividade enzimática de E1-Cat.

Íon ou Reagente químico	Concentração	Atividade relativa (%)
Controle	-	100
Co <sup>2+</sup>	5 mM	84,6 ± 3,6
K <sup>+</sup>	5 mM	88,7 ± 2,9
Li <sup>+</sup>	5 mM	81,7 ± 6,1
Mn <sup>2+</sup>	5 mM	65,6 ± 2,1
Mg <sup>2+</sup>	5 mM	90,6 ± 4,6
Ni <sup>2+</sup>	5 mM	88,0 ± 2,2
Cu <sup>2+</sup>	5 mM	4,2 ± 1,6
Zn <sup>2+</sup>	5 mM	18,7 ± 3,6
Fe <sup>3+</sup>	5 mM	4,6 ± 7,3
Ca <sup>2+</sup>	5 mM	105,1 ± 3,8
NaCl	5 mM	94,0 ± 4,6
EDTA	10 mM	76,4 ± 3,8
EGTA	10 mM	69,4 ± 3,1

Os produtos de hidrólise de celo-oligossacarídeos marcados com APTS foram avaliados eletroforese capilar (Figura 49). E-1 Cat atuou na hidrólise das ligações glicosídicas internas de celohexose, celopentose e celotetrose resultando na liberação de oligossacarídeos menores sendo que não foi observada a formação de glicose. Um padrão semelhante de clivagem foi observado para a celulase isolada do metagenoma do solo da Antártida (BERLEMONT et al., 2009).

Existem relatos na literatura de endoglucanases que não apresentam a capacidade de degradar celulose cristalina (HERR et al., 1978, MANSFIELD et al., 1998). Por outro lado, algumas endoglucanases apresentam a característica de processividade, ou seja, quebram as ligações internas da celulose cristalina e são

capazes de liberar oligossacarídeos solúveis (GILAD et al., 2003, REVERBEL-LEROY et al., 1997, TOMME et al., 1996).

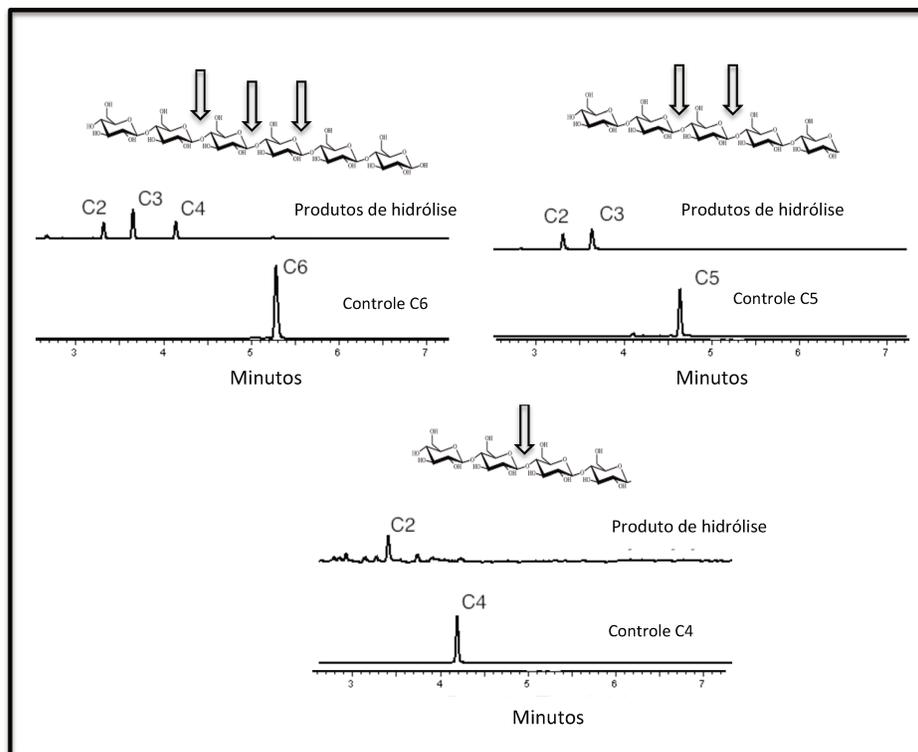


Figura 49 – Análise dos produtos de hidrólise de E1-Cat por eletroforese capilar. As setas indicam as ligações onde ocorreram os ataques enzimáticos, resultando na formação dos produtos observados. C6, C5, C4, C3 e C2 correspondem a celohexose, celopentose, celotetrose, celotriose e celobiose, respectivamente.

Com base nisso, E-1 foi avaliada em relação à capacidade de degradação de celulose cristalina e bagaço de cana-de-açúcar, um substrato complexo, no qual o acesso à fração de celulose é dificultado devido à presença de hemicelulose e lignina. A Tabela 15 apresenta os valores de atividades específicas obtidos na degradação dos substratos testados.

O valor de atividade específica encontrado em Avicel foi de 8,8 U\*/mg, superior ao valor encontrado para a celulase GH5, Cel5A, do fungo *Gloeophyllum trabeum*, que apresentou atividade específica de 4,5 U\*/mg (COHEN et al., 2005) e para a endoglucanase GH9 do cupim *Coptotermes gestroi*, que apresentou o valor de 1,1 U\*/mg (FRANCO CAIRO et al., 2013), ambas avaliadas em Avicel.

Com base nesses resultados, a atividade enzimática da celulase E-1 foi avaliada em bagaço de cana-de-açúcar explodido. O processo de explosão a vapor tem como objetivo facilitar o acesso das enzimas às fibras de celulose e hemicelulose (KAAR et al., 1998). O valor de atividade específica encontrado foi de 10,23 U\*/mg, indicando que a celulase foi capaz de liberar açúcares redutores diretamente da biomassa lignocelulósica. Trata-se de um resultado que demonstra a potencial aplicação desta enzima em processos de conversão de biomassa.

É importante ressaltar, que as atividades em substratos insolúveis foram determinadas apenas com o domínio catalítico da celulase E-1, sem o provável CBM. Considerando que a presença do CBM é constantemente associada ao aumento na atividade enzimática em substratos insolúveis (GUILLEN et al., 2010), estudos posteriores serão realizados como forma de avaliar se a construção correspondente à enzima E-1 completa apresenta valores de atividades específicas superiores aos encontrados para o domínio catalítico E1-Cat.

Tabela 15 – Atividade específica de E1-Cat em substratos insolúveis.

<b>Atividade específica (U*/mg)</b>	
<b>Avicel</b>	<b>Bagaço de cana-de-açúcar explodido</b>
8,8	10,23

U\* equivale à quantidade de equivalentes de nmols de glicose liberados por minuto.

#### **4.4.1.4. Xilanase X-1**

Através da reação de PCR foi amplificado um fragmento com o tamanho esperado de aproximadamente 1,06 kb (Figura 50A). Como se pode observar na Figura 50B, por digestão enzimática foi possível a confirmação da clonagem do fragmento de interesse no vetor pET-28a. O sequenciamento do inserto garantiu que nenhuma mutação ocorreu no gene e que ele foi clonado na fase de leitura correta.

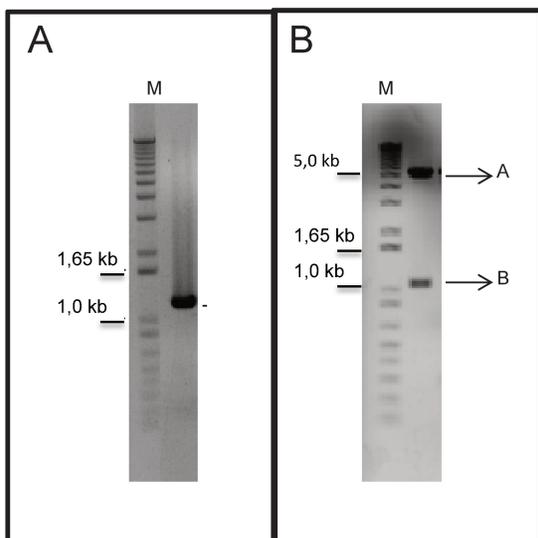


Figura 50 – A - Amplificação do fragmento correspondente ao gene da xilanase X-1 por PCR. B - Confirmação da clonagem em pET-28a por digestão com enzimas de restrição, onde (A) corresponde ao vetor pET-28a linearizado e (B) ao fragmento liberado.

Através da ferramenta *ProtParam*, foram determinados alguns parâmetros bioquímicos e biofísicos da proteína X-1 (Tabela 16), como: peso molecular, ponto isoelétrico, coeficiente de extinção molar, índice de hidrofobicidade (onde valores negativos indicam que a proteína apresenta possibilidade de ser solúvel). Os dados foram gerados a partir da sequência clonada traduzida em aminoácidos com a adição do segmento do vetor pET-28a que também é traduzido.

Tabela 16 - Parâmetros bioquímicos e biofísicos correspondentes a xilanase X-1 calculados a partir do ProtParam.

<i>Proteína</i>	<i>Número de aminoácidos</i>	<i>Peso molecular</i>	<i>Ponto isoelétrico</i>	<i>Coefficiente de Extinção Molar</i>	<i>Índice de hidrofobicidade</i>
X-1	315	45,5 kDa	8,58	72880 M <sup>-1</sup> . cm <sup>-1</sup>	-0,505

Tendo-se confirmado a clonagem, os plasmídeos recombinantes foram transformados nas linhagens BL21(DE3), ArticExpress(DE3) e Rosetta2(DE3), para realização dos testes de expressão. O melhor resultado foi obtido com a linhagem

Rosetta2(DE3) com indução por 4 horas a 30°C, onde foi verificada a expressão de elevadas quantidades da proteína de interesse na forma solúvel (Figura 51A).

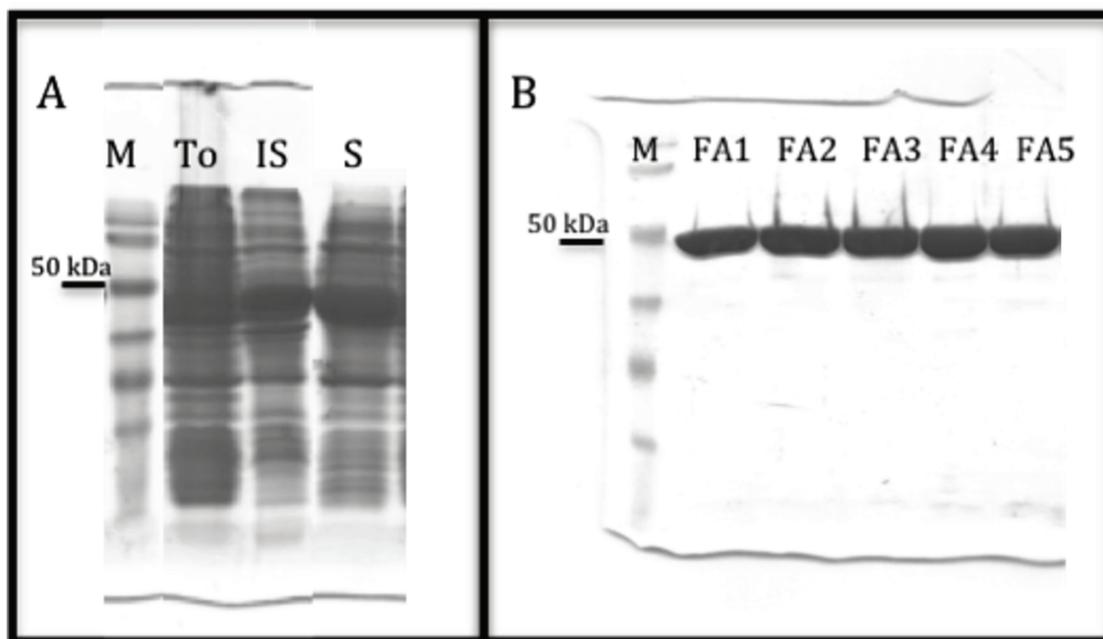


Figura 51 - Análise por SDS-PAGE do teste de expressão da xilanase X-1 em Rosetta2(DE3). M - marcador de peso molecular. (A) To - tempo zero da expressão, IS - fração insolúvel do lisado, S - fração solúvel do lisado. (B) FA1-FA5 - frações resultantes da cromatografia de afinidade correspondentes a xilanase X-1.

Através da Figura 51B, pode-se observar que as amostras resultantes da cromatografia de afinidade apresentaram alto grau de pureza, sem a necessidade da realização da etapa de gel filtração. As frações resultantes da cromatografia de afinidade foram dialisadas em tampão fosfato de sódio 50 mM, NaCl 20 mM, pH 7,4, para a remoção do imidazol.

#### 4.4.1.5. Caracterização bioquímica X-1

Primeiramente, a xilanase X-1 foi avaliada quanto à capacidade de hidrolisar diferentes ligações glicosídicas. Para isso, sua atividade foi testada frente a diferentes substratos: CMC,  $\beta$ -glucano, galactomanana, xiloglucano, xilana extraída de madeira de

faia, arabinoxilana de centeio, arabinoxilana de trigo e arabinana.

Observou-se que X-1 apresentou capacidade de hidrolisar as ligações glicosídicas do tipo  $\beta(1\rightarrow4)$  entre os resíduos de xilose presentes na cadeia de xilana de madeira de faia (atividade relativa 100%), arabinoxilana de centeio (41,3%) e trigo (1,1%). Nos demais substratos testados não foram identificados níveis de atividade enzimática significativos.

A xilana de madeira de faia apresenta em sua composição de açúcares 90% de xilose, a arabinoxilana de centeio apresenta 60% de xilose e 40% arabinose e a arabinoxilana de trigo contém 50% de xilose e 50% de arabinose (informações fornecidas pelo fabricante). Nos três substratos, a arabinose corresponde a decorações da cadeia principal, que é constituída por unidades de xilose unidas por ligação  $\beta(1\rightarrow4)$  (VARDAKOU et al., 2005).

Os dados de atividade relativa em cada um dos três substratos mostraram que pelo fato da xilana de madeira de faia apresentar menor número de decorações, o acesso da enzima à cadeia principal foi facilitado, o que justificou a maior atividade observada neste substrato. O mesmo foi observado pela comparação das atividades em arabinoxilana de trigo e de centeio, pois neste último, o menor número decorações da cadeia principal com arabinose também foi relacionado com maiores valores de atividade enzimática. Com base nesses resultados, os demais ensaios de caracterização foram realizados utilizando xilana de madeira de faia 0,5% como substrato.

Em relação ao pH ótimo de atuação, como pode ser observado na Figura 52, o maior valor de atividade enzimática foi obtido em valores próximos do pH 6,0. Nos pHs 4,5 e 7,0 a enzima reteve valores de atividade acima de 70% em relação ao valor ótimo.

O ensaio de temperatura ótima (Figura 53) mostrou que a maior atividade enzimática foi obtida em torno de 45°C. Além disso, a enzima conseguiu manter 60% de sua atividade ótima na faixa de 20°C a 55°C.

Os valores ótimos de pH e temperatura reportados para X-1 são comparáveis a outras xilanases da família GH10 descritas na literatura. No caso das xilanases isoladas de *Plectophaerella cucumerina* (XynZC) (ZHANG et al., 2007), *Clostridium cellulovorans* (XynB) (HAN et al., 2004) e *Trichoderma harzianum* T4 (FRANCO et al., 2004), a

temperatura ótima encontrada foi de 40°C. Para a xilanase isolada de metagenoma de solo de tundra (WANG et al., 2010), o pH ótimo foi observado na faixa de 6,0 a 6,5.

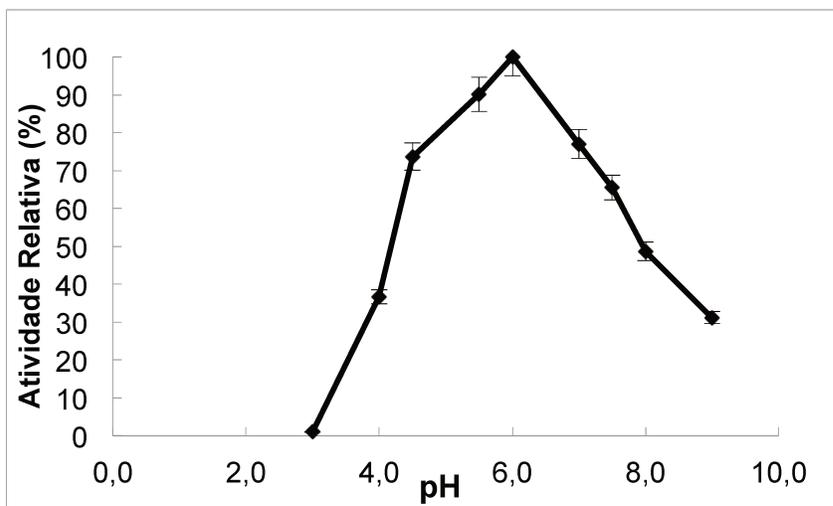


Figura 52 - Determinação do valor de pH ótimo de atividade de X-1 em reações contendo 0,5% de xilana de madeira de faia em tampão 100 mM citrato – 200 mM fosfato, com pHs variando entre 3,0 e 9,0. As reações foram incubadas a 37°C por cinco minutos e reveladas pelo método do ADNS.

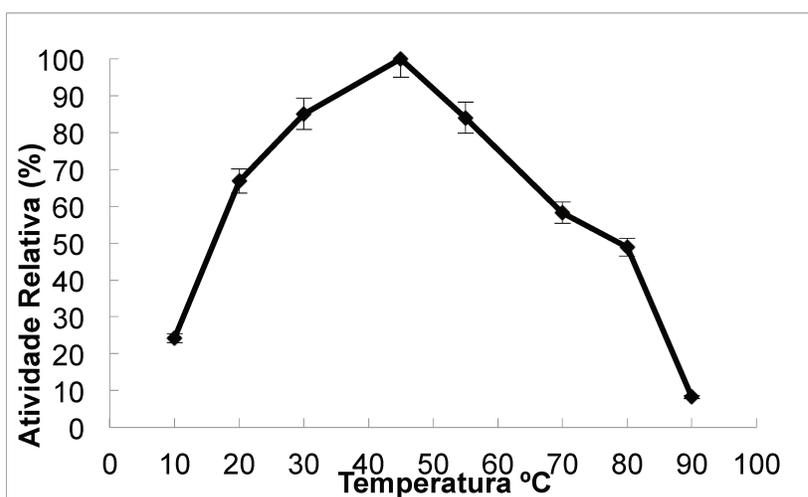


Figura 53 - Determinação do valor de temperatura ótima de atividade de X-1 em reações contendo 0,5% de xilana de madeira de faia em tampão acetato de sódio 100 mM pH 6,0. As reações foram incubadas por 5 minutos em temperaturas variando de 10°C a 90°C e em seguida foram reveladas pelo método do ADNS.

O valor de energia de ativação para hidrólise de xilana de madeira de faia calculado foi de  $31,4 \text{ kJ}\cdot\text{mol}^{-1}$ , semelhante ao verificado para xilanase isolada de *Bacillus subtilis* (BERNIER et al., 1983), cujo valor foi de  $35,1 \text{ kJ}\cdot\text{mol}^{-1}$  e bem abaixo do valor de  $878,6 \text{ kJ}\cdot\text{mol}^{-1}$ , descrito para a xilanase isolada do fungo *Scytalidium thermophilum* (GAUR et al., 2005).

Em relação à estabilidade térmica, através da Figura 54, podemos observar que a incubação a  $70^\circ\text{C}$  e  $80^\circ\text{C}$  resultou na completa inativação da enzima em menos de 10 minutos. Já quando incubada a  $60^\circ\text{C}$ , foi verificada a perda total de atividade depois de uma hora. Nas temperaturas de  $50^\circ\text{C}$  e  $40^\circ\text{C}$  a enzima mostrou-se bastante estável por uma hora e reteve cerca de 60% de sua atividade inicial durante seis horas (dados não mostrados no gráfico).

A Figura 55 exibe a curva de cinética enzimática obtida, evidenciando o perfil de Michaelis-Menten com um  $R^2$  de 0.96. Os parâmetros cinéticos foram obtidos pelo cálculo das atividades enzimáticas em concentrações de xilana de madeira de faia variando de  $0,625 \text{ mg/mL}$  a  $11,25 \text{ mg/mL}$ . Os valores obtidos foram:  $K_m$  de  $2,18 \pm 0,13 \text{ mg/mL}$ ,  $V_{\max}$  de  $1.435 \pm 30,4 \text{ }\mu\text{mol/min/mg}$ ,  $k_{\text{cat}}$  de  $1.083,47 \text{ s}^{-1}$  e  $k_{\text{cat}}/K_m$  de  $496,32 \text{ mL/mg/s}$ .

O valor de eficiência catalítica ( $k_{\text{cat}}/K_m$ ) de  $496,32 \text{ mL/mg/s}$  é superior ao obtido em outros trabalhos que também utilizaram xilana de madeira de faia como substrato. No caso da xilanase GH10 isolada de *Thermotoga thermarum* (SHI et al., 2013) o valor encontrado para  $k_{\text{cat}}/K_m$  foi de  $276,37 \text{ mL/mg/s}$ , enquanto que para a xilanase isolada da bactéria marinha *Glaciecola mesophila* este valor foi de  $56,56 \text{ mL/mg/s}$  (SHI et al., 2013, GUO et al., 2009).

Os testes com íons e outros reagentes químicos (Tabela 17), mostraram que os íons  $\text{Zn}^{2+}$ ,  $\text{Cu}^{2+}$  e  $\text{Fe}^{3+}$  apresentaram efeitos altamente inibitórios na atividade enzimática de X-1. A inibição por estes íons metálicos já havia sido previamente descrita para outras xilanases (JEONG et al., 2012, HU et al., 2008, KNOB et al., 2013, LIU et al., 2012). A inibição por  $\text{Cu}^{2+}$  e  $\text{Zn}^{2+}$  pode estar relacionada à ligação destes íons ao grupo sulfidríla da cisteína  $\text{C}^{78}$  de X-1, mesmo este não sendo um dos resíduos de aminoácidos catalíticos (ALVAREZ et al., 2013). A inibição pela ligação de íons nos resíduos de cisteína pode ocorrer mesmo que tais resíduos não sejam essenciais para

a atividade catalítica desempenhada pela enzima, como já foi observado para a protease do vírus HIV-1 (KARLSTROM & LEVINE, 1991). A presença dos íons  $\text{Co}^{2+}$ ,  $\text{Mn}^{2+}$  e  $\text{Ni}^{2+}$  resultaram na inibição parcial da atividade enzimática de X-1. Os demais íons avaliados,  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Li}^+$ ,  $\text{K}^+$  e os reagentes NaCl, EDTA e EGTA não influenciaram significativamente na atividade enzimática da xilanase.

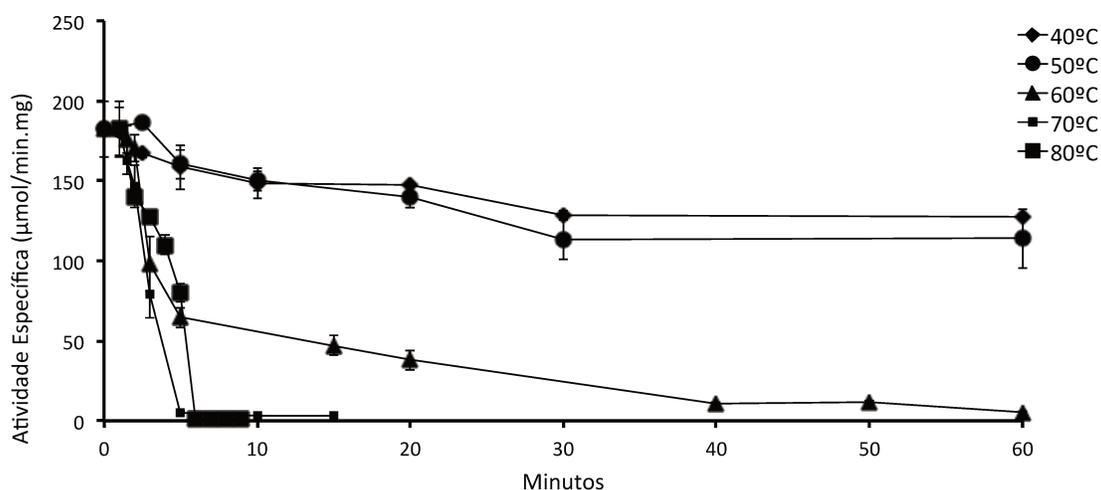


Figura 54 - Ensaio de estabilidade térmica de X-1 realizado com xilana de madeira de faia 0,5% em tampão acetato de sódio 100 mM pH 6,0. Após a incubação, as reações foram reveladas pelo método do ADNS.

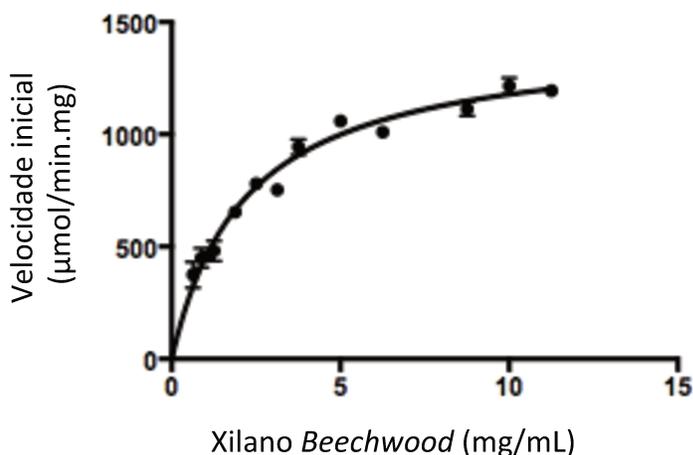


Figura 55 - Gráfico de Michaelis-Menten para X-1 em presença de xilana de madeira de faia como substrato, com concentrações variando de 0,625 mg/mL a 11,25 mg/mL. As reações foram realizadas em tampão acetato de sódio 100 mM pH 6,0, incubadas por 5 minutos a 45°C e em seguida foram reveladas pelo método do ADNS.

Tabela 17 - Efeito de íons e reagentes químicos na atividade enzimática de X-1.

<b>Íon ou Reagente químico</b>	<b>Concentração</b>	<b>Atividade Relativa (%)</b>
<b>Controle</b>	-	100,0
<b>CaCl<sub>2</sub></b>	5 mM	93,6 ± 9,9
<b>CoCl<sub>2</sub></b>	5 mM	48,1 ± 12,9
<b>ZnSO<sub>4</sub></b>	5 mM	0,00
<b>MnCl<sub>2</sub></b>	5 mM	50,2 ± 6,25
<b>MgCl<sub>2</sub></b>	5 mM	89,2 ± 7,8
<b>FeCl<sub>3</sub></b>	5 mM	4,3 ± 0,27
<b>CuSO<sub>4</sub></b>	5 mM	0,00
<b>NiCl<sub>2</sub></b>	5 mM	29,3 ± 5,7
<b>LiCl</b>	5 mM	99,0 ± 1,9
<b>KCl</b>	5 mM	107,1 ± 7,08
<b>NaCl</b>	5 mM	96,0 ± 4,0
<b>EDTA</b>	10 mM	92,0 ± 1,9
<b>EGTA</b>	10 mM	85,1 ± 2,5

A análise dos produtos de hidrólise por eletroforese capilar (Figura 56) mostrou que X-1 foi capaz de hidrolisar os xilo-oligossacarídeos testados, mas não foi capaz de hidrolisar xilobiose (X2) (dados não mostrados). Em relação à degradação dos xilo-oligossacarídeos, foi observada a formação de oligossacarídeos intermediários após curto período de hidrólise (30 minutos a 45°C). No caso da xilotriose, a formação desses intermediários foi discreta (Figura 56A). Como resultado da hidrólise por 16 horas a 45°C, verificou-se a conversão de xilohexose (X6), xilopentose (X5), xilotetose (X4) e xilotriose (X3) completamente em xilobiose e xilose (X) (Figura 56 A-D). Esses resultados mostraram que o ataque enzimático ocorreu preferencialmente nas ligações internas e em seguida nas ligações externas. Isso ficou evidente no caso da hidrólise de X4 (Figura 56B), pois primeiramente o ataque ocorreu predominantemente na ligação interna, liberando principalmente X2 e de forma secundária nas ligações

externas, resultando na liberação de pequenas quantidades de X3 e X.

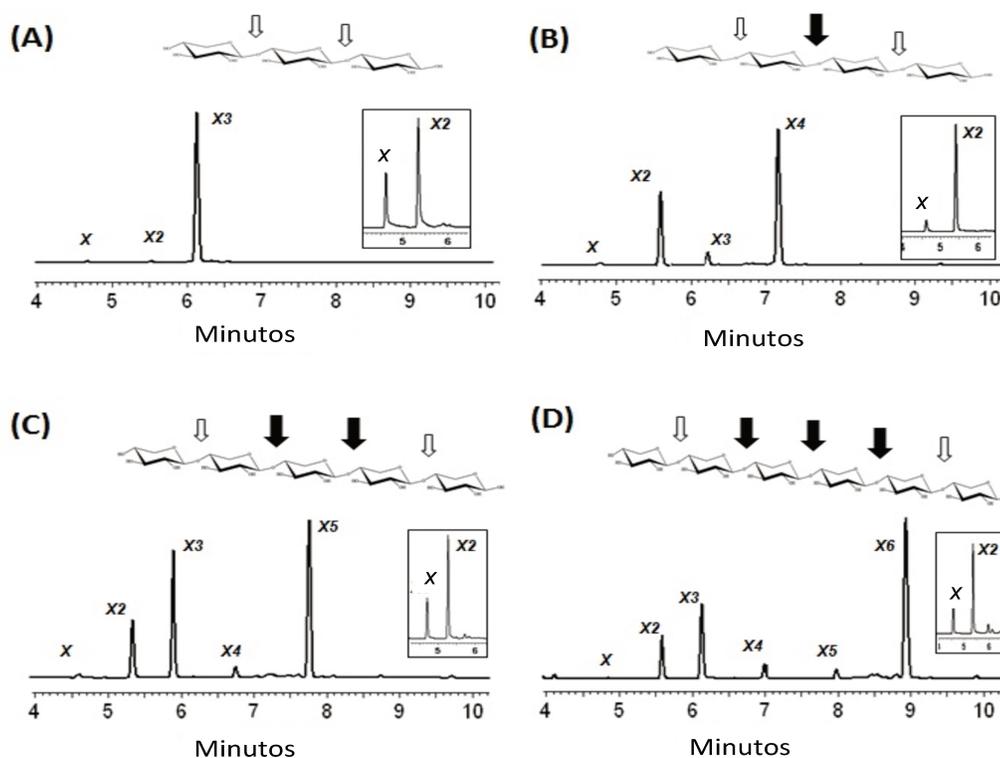


Figura 56 – Perfil de degradação de A) xilotriose (X3), B) xilotetose (X4), C) xilopentose (X5) e D) xilohexose (X6). Os produtos intermediários após incubação por 30 minutos estão representados nas figuras externas. Dentro dos retângulos estão representados os produtos finais de degradação. As setas pretas indicam as ligações onde os ataques enzimáticos ocorreram preferencialmente, e as setas brancas indicam as ligações onde ocorreram os ataques secundários.

Os produtos de hidrólise resultantes da atividade enzimática em xilana de madeira de faia, arabinxilana de trigo e bagaço de cana-de-açúcar pré-tratado com ácido fosfórico foram avaliados por eletroforese capilar. Em todos os casos os principais produtos de degradação foram xilo-oligossacarídeos, xilobiose e xilose (Figura 57). Os xilo-oligossacarídeos e a xilobiose são de grande interesse para a indústria alimentícia devido a sua aplicação como prebióticos e adoçantes (VÁZQUEZ et al., 2000). A xilose gerada a partir da hidrólise desses substratos pode ser empregada para a produção de xilitol, um adoçante com alto valor agregado (WINKELHAUSEN & KUZMANOVA, 1998). A xilose gerada pode ser ainda utilizada para a produção de etanol, através de

processos que envolvem a utilização de leveduras capazes de fermentar xilose (ZHANG & GENG, 2012, JIN et al., 2003).

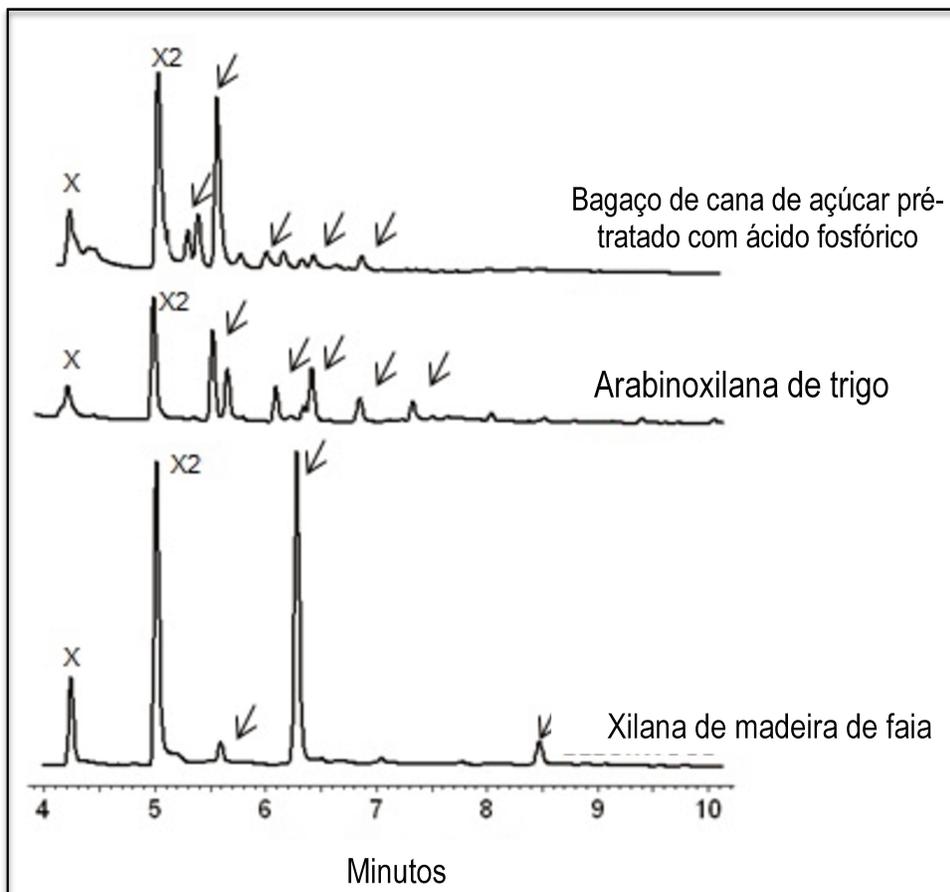


Figura 57 – Produtos de hidrólise resultantes da atividade enzimática de X-1 em bagaço de cana-de-açúcar pré-tratado com ácido fosfórico, arabinoxilana de trigo e xilana de madeira de faia. As setas indicam os xilo-oligosacarídeos liberados pela atividade enzimática.

Já foi observado que o pré-tratamento da biomassa vegetal pela atividade hidrolítica de xilanasas apresenta efeitos benéficos nas etapas seguintes de sacarificação da biomassa com coquetéis celulolíticos (JEONG et al., 2012). Com base nisso, avaliou-se o efeito da atividade enzimática da xilanase X-1 como pré-tratamento do bagaço de cana-de-açúcar (previamente tratado com ácido fosfórico), visando à degradação parcial da hemicelulose e o aumento do acesso das enzimas celulolíticas às fibras de celulose.

Não foi verificada a liberação de quantidades significativas de açúcares redutores após o tratamento enzimático do bagaço apenas com a xilanase X-1. Porém, após a adição do coquetel celulolítico ACCELERASE® 1500, houve um aumento de aproximadamente 65% na quantidade de açúcares redutores liberados na reação em que foi incluída a etapa de pré-tratamento com xilanase em comparação com a amostra de bagaço não tratada com xilanase (Figura 58).

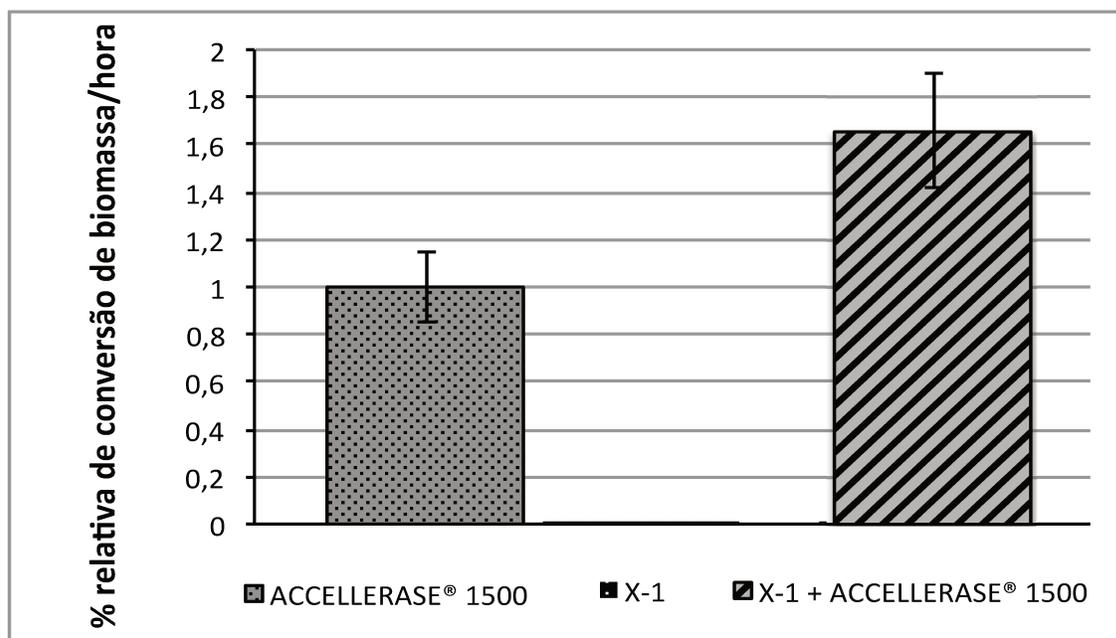


Figura 58 – Comparação do efeito da atividade enzimática da xilanase X-1 como etapa de tratamento do bagaço de cana-de-açúcar pré-tratado com ácido fosfórico anteriormente ao processo de sacarificação com o coquetel celulolítico ACCELERASE® 1500.

Todos os dados apresentados nesta tese em relação à caracterização bioquímica da xilanase X-1, juntamente estudos biofísicos e resolução da estrutura tridimensional realizados em colaboração com outros grupos de pesquisas, foram reunidos em um trabalho denominado *“Development and biotechnological application of a novel endoxylanase family GH10 identified from sugarcane soil metagenome”* publicado no periódico PLOS ONE.

## 4.5. Conclusões

Neste capítulo foram apresentados os resultados referentes à caracterização bioquímica e investigação da potencial aplicação biotecnológica da celulase E-1 e xilanase X-1, isoladas do metagenoma do consórcio microbiano degradador de bagaço de cana-de-açúcar.

Em relação à celulase E-1, ensaios de ligação a substratos insolúveis mostraram tratar-se de uma proteína com dois domínios. O primeiro deles, localizado na região N-terminal correspondia ao domínio catalítico, denominado E-1 Cat e o segundo, localizado na região C-terminal, tratava-se do domínio de ligação a carboidratos, E-1 CBM. Como descrito no capítulo 3, a sequência de aminoácidos do E-1 CBM não apresentou identidade com domínios conservados conhecidos, inclusive de famílias de CBMs descritas. Isto sugeriu que o E-1 CBM pode ser o membro de uma família de CBMs ainda não conhecida.

A caracterização bioquímica de E-1 Cat permitiu identificar que a enzima atuou especificamente na hidrólise de ligações glicosídicas do tipo  $\beta(1\rightarrow4)$  entre resíduos de glicose. Através de análises por eletroforese capilar foi visto ainda que E-1 Cat atuou na hidrólise das ligações glicosídicas internas das cadeias de celo-oligossacarídeos, resultando na formação de oligossacarídeos menores, não sendo verificada a formação de glicose. A enzima apresentou atividade ótima a 50°C e pH 7,0, e se mostrou termo-tolerante a 40°C por seis horas. Tais condições estão de acordo com as condições de hidrólise normalmente empregadas nos processos industriais de hidrólise enzimática. Como já descrito para outras celulasas, E-1 Cat sofreu inibição na presença dos íons  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$  e  $\text{Fe}^{3+}$ , e não apresentou aumento significativo da atividade enzimática com nenhum dos íons testados. O valor de eficiência catalítica  $k_{\text{cat}}/K_m$  de 4,06 mL/mg/s encontrado para a atividade de E-1 Cat em CMC foi superior ao descrito para outras celulasas da família GH5.

Em relação à aplicação biotecnológica, E-1 Cat apresentou capacidade de liberação de açúcares solúveis diretamente de celulose cristalina e bagaço de cana-de-açúcar explodido, demonstrando sua potencial aplicação como componente de coquetéis enzimáticos empregados em processos de conversão da biomassa vegetal.

Já em relação à xilanase X-1, ela apresentou especificidade na hidrólise de ligações glicosídicas do tipo  $\beta(1\rightarrow4)$  entre resíduos de xilose, apresentando maior atividade em cadeias de xilana não ramificadas. Em relação às condições de atuação, X-1 também apresentou características desejáveis para processos industriais, com temperatura e pH ótimos em torno de 45°C e 6,0, respectivamente, e termo-estabilidade a 40°C e 50°C por seis horas. O valor de eficiência catalítica em xilana de madeira de faia de 496,32 mL/mg/s foi superior ao encontrado para outras xilanases. X-1 apresentou forte inibição da atividade enzimática na presença dos íons  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$  e  $\text{Fe}^{3+}$ .

Através de ensaios de eletroforese capilar com diferentes xilo-oligossacarídeos, polissacarídeos e bagaço de cana-de-açúcar pré-tratado com ácido fosfórico, observou-se que X-1 foi capaz de atuar tanto na hidrólise das ligações glicosídicas internas como nas externas, resultando na formação de diversos xilo-oligossacarídeos, xilobiose e xilose. Todos esses produtos apresentam grande interesse industrial para a indústria de alimentos, para uso como prebióticos ou para produção de adoçantes. A xilose apresenta ampla aplicação na área de biocombustíveis, já que pode ser utilizada em processos de fermentação para a produção de etanol. Por fim, o tratamento de uma amostra de bagaço de cana-de-açúcar pré-tratado com ácido fosfórico com a xilanase X-1 antes da etapa de sacarificação resultou no aumento dos açúcares redutores liberados em comparação com a amostra que não foi submetida à hidrólise com X-1.

Os resultados obtidos através da caracterização bioquímica de E-1 Cat e X-1 permitiram observar que ambas as enzimas apresentaram características interessantes para aplicação em processos industriais, principalmente no que diz respeito à conversão da biomassa para a geração de bio-produtos.



## **CAPÍTULO 5 – ESTUDOS FUNCIONAIS DO DOMÍNIO DE LIGAÇÃO A CARBOIDRATOS PRESENTE NA CELULASE E-1 (E1-CBM)**

### **5.1. Introdução**

De forma geral, os CBMs são definidos como domínios proteicos presentes em proteínas envolvidas na degradação de carboidratos, que apresentam enovelamento próprio e capacidade de ligação a carboidratos (SHOSEYOV et al., 2006).

Como descrito anteriormente, a celulase E-1, identificada na triagem funcional da biblioteca metagenômica, apresentou evidências sobre a existência de um CBM em sua estrutura, principalmente em virtude da capacidade de ligação a substratos insolúveis.

Com bases nessas informações, neste capítulo são descritas em detalhes as estratégias e metodologias empregadas para avaliar as características da interação do E-1 CBM com diferentes carboidratos, visto que a sequência de aminoácidos correspondente a este domínio não apresentou identidade com nenhum membro de famílias de CBMs conhecidas.

Uma parte dos estudos descritos neste capítulo foi realizada no laboratório do professor Dr. Harry Gilbert, professor e pesquisador na Universidade de Newcastle, no Reino Unido, que já esteve envolvido em vários trabalhos de caracterização de novos CBMs (GILBERT et al., 1990, DURRANT et al., 1991, POOLE et al., 1993, RAGHOTHAMA et al., 2000, BOLAM et al., 2004). As principais atividades realizadas neste período foram: otimização da expressão heteróloga do CBM, experimentos de calorimetria de titulação isotérmica (ITC) e testes de cristalização visando a obtenção da estrutura tridimensional do E-1 CBM.

### **5.2. Objetivos**

Investigação e caracterização do domínio de ligação a carboidratos identificado na celulase E-1.

## 5.3. Materiais e Métodos

### 5.3.1. Clonagem em pET-28a e pET-41a

A sequência nucleotídica corresponde ao E-1 CBM foi clonada nos vetores pET-28a e pET-41a (Figura 59).

O vetor pET-41a é otimizado para obtenção de altos níveis de expressão de proteína heteróloga em virtude da fusão da proteína de interesse com a cauda altamente solúvel de GST (glutathione-S-transferase) na região N-terminal da construção (informações fornecidas pelo fabricante, Novagen). Neste tipo de vetor, existem sítios de clivagem para as proteases enteroquinase e trombina, que permitem a separação da proteína de interesse da cauda de GST. A clonagem em pET-41a também permite a fusão da cauda de histidina à proteína de interesse, permitindo que a purificação proteica seja feita por IMAC.

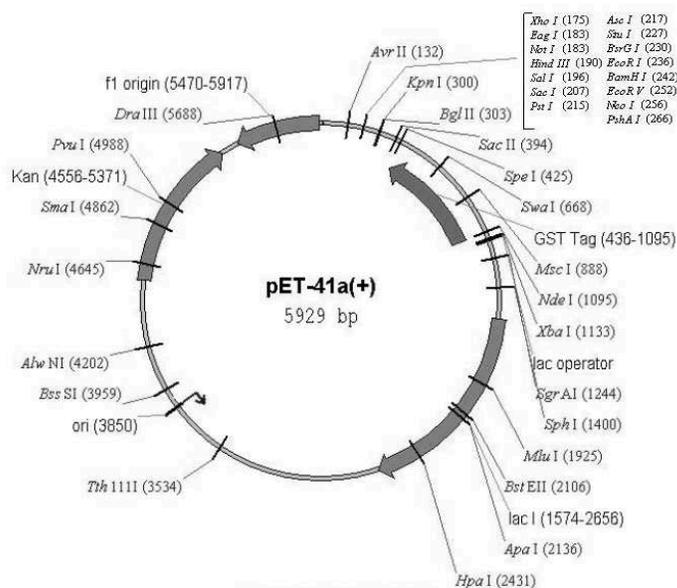


Figura 59 - Mapa do vetor de expressão pET-41a.

A Tabela 18 apresenta os oligonucleotídeos iniciadores (desenhados como descrito no tópico 4.3.1) utilizados nas reações de amplificação da sequência nucleotídica de interesse.

Tabela 18 - Oligonucleotídeos iniciadores utilizados nas reações de PCR. As cores indicam os sítios de restrição: amarelo – *NdeI*, azul – *BamHI* e verde - *XhoI*. **Ta** indica a temperatura de anelamento dos oligonucleotídeos iniciadores construídos.

	Oligonucleotídeo iniciador direto (5' → 3')	Oligonucleotídeo iniciador reverso (5' → 3')
E-1 CBM (clonagem em pET28a)	TATATATCATATGAGCGCATCATGCGGTAGC <i>Ta</i> = 58°C	ATACCATCTTACCAGTTATCGAACTTCACATTA <i>Ta</i> = 52°C
E-1 CBM (clonagem em pEt41a)	CTCGCGCCATCCAGCGCATCATGCGGTAGC <i>Ta</i> = 58,7°C	CGCGAGCTCGAGTACCAGTTATCGAACTTCAC <i>Ta</i> = 49,8°C

A reação de PCR, ligação, transformação e confirmação da clonagem em pET-28a foram realizadas como descrito no capítulo 4.

As etapas de clonagem em pET-41a foram realizada no laboratório do Prof. Dr. Harry Gilbert, no Reino Unido, e estão descritas entre nos tópicos 5.3.2, 5.3.3 e 5.3.4.

### 5.3.2. Reação de PCR para amplificação do gene alvo

A reação de PCR foi realizada utilizando-se como alvo a solução dos plasmídeos extraídos do clone positivo E-1. As condições de reações foram as seguintes: 32 µL de água ultrapura estéril, 5 µL de tampão KOD 10 X *PCR Buffer for KOD Hot Start DNA Polymerase* (Merck), 1 µL de dNTP 10 mM, 1,0 µL de cada um dos oligonucleotídeos iniciadores (direto e reverso) a 10 pmol/µL, 1,0 µL do DNA molde (concentração entre 50 e 100 ng/µL) e 1 U da *KOD Hot Start* polimerase (Merck), totalizando um volume de reação de 50 µL. As reações foram realizadas no termociclador *MyCycler™ Thermal Cycler* (BioRad).

As condições de amplificação compreenderam uma desnaturação inicial a 98°C por um minuto, seguida de 30 ciclos de 20 segundos a 98°C, 30 segundos a 55°C e 1 minuto a 72°C e um último ciclo de 10 minutos a 72°C.

O produto de PCR foi avaliado por eletroforese de gel de agarose 1,0% (90 V por 60 minutos). As bandas foram visualizadas em transiluminador de luz UV, após tratamento com brometo de etídeo 0,25 µg/mL por 15 minutos e as imagens capturadas através do fotodocumentador Gel Doc™ EZ Imager da Bio-Rad Life Sciences.

### 5.3.3. Preparo do vetor pET-41a

O vetor foi preparado da seguinte forma: 35 µL do vetor (concentração aproximada de 60 ng/µL), 10 µL do tampão TANGO 10X (Fermentas), 10 U de *Bam*HI e 4,0 µL de água. A reação foi incubada a 37°C por 1,5 horas e em seguida inativada a 80°C por 20 minutos. Posteriormente foram adicionados 10 U de *Xho*I na mesma reação, que foi novamente incubada a 37°C por 1,5 horas. A eficiência da digestão foi confirmada por eletroforese em gel de agarose 1,0%. Após a confirmação a reação foi purificada através do kit *QIAquick PCR Purification Kit* (Qiagen).

Como forma de prevenir a religação dos vetores nos casos em que a digestão não tenha sido eficiente, realizou-se a desfosforilação do vetor digerido. Para isso, 32 µL do vetor purificado foram adicionados de 5 µL do tampão 10 X da fosfatase, 7,5 U de fosfatase (*Antarctica Phosphatase* – New England Biolabs) e 11,5 µL de água. A reação foi incubada a 37°C por 1 hora e posteriormente inativada por 20 minutos a 65°C.

### 5.3.4. Reação de ligação

A reação de ligação foi feita de acordo com o que foi descrito no tópico 4.3.6.

### 5.3.5. Transformação por choque térmico

A transformação foi feita em *E.coli* linhagem DH5α de acordo com o que foi descrito no tópico 4.3.10.

### 5.3.6. Confirmação da clonagem em pET-41a

A clonagem foi confirmada por PCR de colônia. Neste caso, alguns clones foram selecionados aleatoriamente e 2 µL do inóculo de cada colônia foram adicionados à reação de PCR contendo os oligonucleotídeos iniciadores específicos para o vetor pET-41a (T7 “sequencing primer” – 5’ TAATACGACTCACTATAGGG 3’, T7 “terminator primer” – 5’ CTAGTTATTGCTCAGCGGTG 3’).

Os clones que foram confirmados quanto à presença do inserto após visualização do produto de PCR por eletroforese em gel de agarose 1,0%, foram selecionados para a extração do DNA plasmidial, através do kit *Wizard Plus SV Minipreps DNA Purification System* (Promega).

Como forma de avaliar se as sequências clonadas estavam corretas em relação à sequência de nucleotídeos original, os plasmídeos extraídos foram sequenciados na empresa Eurofins MWG Operon.

### **5.3.7. Testes de expressão**

A primeira etapa consistiu na otimização da expressão do E-1 CBM clonado em pET-28a (CBM-pET-28a). A Tabela 19, apresenta todas as cepas testadas e as condições de expressão avaliadas. Todas as expressões foram realizadas em 1 L de meio de cultura, a 220 rpm de rotação.

As características das linhagens BL21(DE3) e Rosetta2(DE3) foram citadas anteriormente no item 4.3.8. A linhagem C41(DE3) (Lucigen) é derivada da cepa BL21(DE3) e apresenta mutações que conferem tolerância à toxicidade (informações fornecidas pelo fabricante), pois em alguns casos a proteína expressa pode ser tóxica para a célula hospedeira.

A linhagem Tuner (DE3) (EMD Millipore) é um mutante deletério da cepa BL21, sensível a diferentes concentrações do indutor IPTG (informações fornecidas pelo fabricante).

O meio de autoindução, citado na Tabela 19, foi descrito por Studier, em 2005. Este meio contém além de outros elementos glicose, glicerol e lactose. A glicose representa a fonte primária de energia além de atuar como repressor do promotor envolvido na expressão heteróloga. A fonte secundária de energia é representada pelo glicerol. Após a metabolização do glicerol, as células passam a utilizar a lactose que também atua como o agente indutor da expressão (STUDIER, 2005). O benefício desta técnica se baseia no fato de que o processo de expressão ocorre naturalmente, sem a necessidade do monitoramento da absorbância a 600 nm para a adição do agente indutor da expressão. Além disso, a expressão ocorre com uma densidade de células superior (absorbância a 600 nm aproximadamente igual a 5), o que aumenta as

chances de obtenção de maiores níveis de expressão proteica (STUDIER, 2005).

Tabela 19 – Condições dos testes de expressão do CBM-pET-28a.

Cepa	Temperatura (°C)	Período de indução	Concentração de IPTG	Meio de cultura
BL21(DE3) Novagen	16°C	16 horas	1 mM	LB
BL21(DE3) Novagen	37°C	4 horas	1 mM	LB
BL21(DE3) Novagen	16°C	16 horas	-	Auto-indução
Rosetta2(DE3) Novagen	16°C	16 horas	1 mM	LB
C41(DE3) Lucigen	16°C	16 horas	1 mM	LB
Tuner(DE3) EMD Millipore	16°C	16 horas	1 mM	LB

Em relação à clonagem em pET-41a, os plasmídeos extraídos de *E. coli* linhagem DH5 $\alpha$  foram transformados na linhagem BL21(DE3) para realização dos testes de expressão em 1 L de meio LB. As células foram mantidas sob agitação de 250 rpm a 37°C até que o valor de absorbância a 600 nm atingisse o valor de 0.6. A expressão proteica foi então induzida pela adição de 1 mM de IPTG em duas condições a 16°C por 16 horas e a 37°C por 4 horas.

### 5.3.8. Purificação proteica

Após a expressão e centrifugação da cultura bacteriana a 5.800 x *g* por 30 minutos, o *pellet* foi ressuspensão em tampão Talon, contendo Tris-HCl 20 mM pH 8,0 e NaCl 100 mM. Posteriormente, a amostra foi lisada por ultrasonificação e centrifugada a 5.800 x *g* por 40 minutos. O sobrenadante foi coletado e purificado por IMAC em resina de cobalto do tipo Talon.

Após a injeção da amostra, a resina foi lavada uma vez com tampão Talon contendo 10 mM de imidazol, seguida por duas lavagens com tampão Talon com 100 mM de imidazol, através das quais era esperada a recuperação da proteína de interesse.

A análise da pureza da amostra resultante da purificação foi feita através de SDS-PAGE 12,5%.

### 5.3.9. Calorimetria de titulação isotérmica

As medições de ITC foram feitas a 25°C, seguindo procedimentos padrões em um calorímetro de titulação isotérmica Microcal Omega (PELL et al., 2003). A amostra proteica foi extensamente dialisada em tampão HEPES 50 mM, pH 8,0. O ligante foi dissolvido no mesmo tampão para minimizar os calores de diluição. Durante as titulações, aproximadamente 2 mL de proteína na concentração 100  $\mu$ M foram adicionados à célula de reação, mantida a 25°C e homogeneizada a 300 rpm. Aproximadamente 10  $\mu$ L do ligante foram injetados sucessivamente, em um total de 25 injeções, com intervalos de 200 segundos. Os ligantes avaliados foram: celohexose (C6), celopentose (C5), celotetrose (C4), manohexose (M6), xilohexose (X6), heptassacarídeo X<sub>3</sub>Glc<sub>4</sub> (XXXG) (todos fornecidos pela Megazyme),  $\beta$ -glucano de baixa viscosidade (BG), xiloglucano (XG) e xilana beechwood (de madeira de faia) (XB) (todos fornecidos pela Sigma Aldrich). Os ligantes C6, C5, C4, M6, X6 e XXXG foram utilizados na concentração de 5 mM, enquanto que a concentração de BG foi 6 mg/mL, e as de XG e XB foram 10 mg/mL.

A celohexose, celopentose e celotetrose correspondem a 6, 5 e 4 unidades de glicose, respectivamente, unidas por ligações glicosídicas do tipo  $\beta(1\rightarrow4)$ . A manohexose, consiste de 6 unidades de manose unidas por ligações glicosídicas do tipo  $\beta(1\rightarrow4)$  e a xilohexose de 6 unidades de xilose unidas por ligações glicosídicas do tipo  $\beta(1\rightarrow4)$ . O  $\beta$ -glucano apresenta uma cadeia mista de glicoses unidas por ligações glicosídicas dos tipos  $\beta(1\rightarrow3)$  e  $\beta(1\rightarrow4)$  na proporção de 1:2 ou 1:3 (informações fornecidas pelo fabricante). O xiloglucano é formado por uma cadeia principal de glicoses unidas por ligações glicosídicas do tipo  $\beta(1\rightarrow4)$ , com substituições, geralmente na posição 6, de xilose, galactose, arabinose ou outros açúcares (HENSEL, 1993). O

heptassacarídeo XXXG apresenta quatro resíduos consecutivos de glicoses unidos por ligações  $\beta(1\rightarrow4)$  sendo que três deles estão ligados a unidades de xilose (MARCUS et al., 2008).

A concentração molar dos sítios de ligação nos polissacarídeos (BG, XB e XG) foi determinada através da alteração da concentração do ligante para o ajuste da curva, até que número de sítios de ligação na proteína (N) fosse igual a 1 (SZABO et al., 2001).

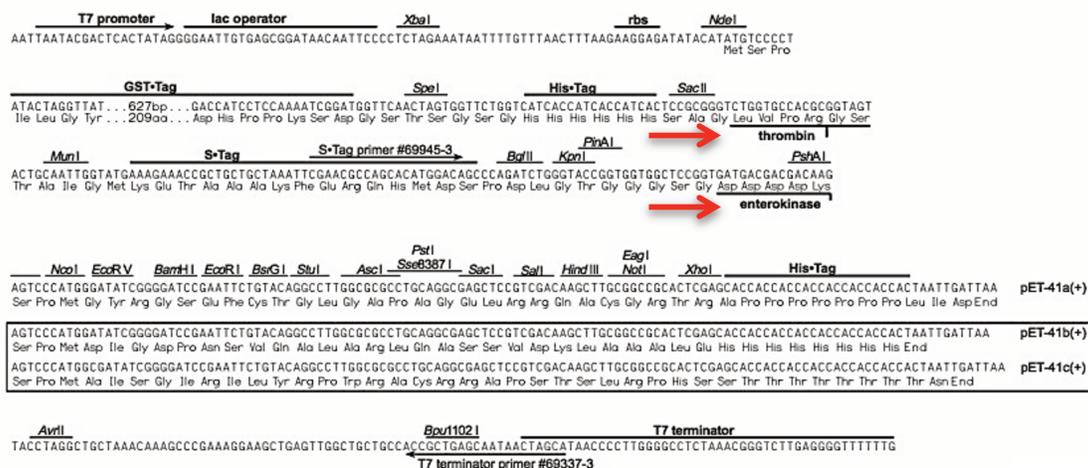
A integração dos valores correspondentes à energia liberada em cada uma das titulações foi feita por regressão não linear, através do modelo de sítio de ligação único (Microcal Origin, versão 2.9), que permitiu a determinação da constante de associação  $K_a$  e a entalpia de ligação  $\Delta H$ . Os demais parâmetros termodinâmicos,  $\Delta G$  e  $\Delta S$  foram calculados de acordo com as equações 1.1 e 1.2 (Capítulo 1).

### **5.3.10. Remoção da cauda de GST**

Para a realização dos testes de cristalização foi necessária à remoção da cauda de GST expressa em fusão com a proteína de interesse. Como descrito anteriormente o vetor pET-41a apresenta sítios de clivagem para enteroquinase e trombina (Figura 60).

Para os experimentos de cristalização é desejável que seja removida a maior quantidade possível de aminoácidos que não pertencem à sequência nativa da proteína. Em virtude disso, como pode ser visualizado na Figura 60, a enzima mais adequada para a remoção da cauda de GST é a enteroquinase, pois é a que apresenta o sítio de clivagem mais próximo do início da sequência da proteína de interesse (sítio *Bam*HI).

A eficiência de remoção da cauda de GST com enteroquinase foi avaliada por testes nos quais se variava a quantidade de enzima adicionada. As reações eram constituídas por 5  $\mu$ L do tampão de reação (10 X *rEK Cleavage/Capture Buffer*, Merck Millipore), 10  $\mu$ L da proteína alvo de clivagem (10  $\mu$ g/ $\mu$ L), água ultrapura estéril para o volume final de 50  $\mu$ L e diferentes quantidades de enteroquinase (0,2 U, 0,5 U e 1,0 U). As reações foram incubadas a 21°C por aproximadamente 16 horas, e o resultado foi avaliado por SDS-PAGE.



### pET-41a(+) cloning/expression regions

Figura 60 – Detalhe da região de clonagem do vetor de expressão pET-41a. As setas em vermelho indicam os sítios de clivagem proteolítica com enteroquinase e trombina, para a remoção da cauda de GST expressa em fusão com a proteína de interesse.

A clivagem com trombina (*Biotinylated Thrombin*, Merck Millipore) também foi avaliada, seguindo as mesmas condições descritas acima, com a diferença do tampão utilizado na reação, que no caso foi o *Thrombin Cleavage/Capture Buffer* - Merck Millipore.

Posteriormente, as reações foram escalonadas, mantendo as proporções descritas, para a digestão de maiores quantidades da proteína alvo.

### 5.3.11. Preparo da amostra para os testes de cristalização

Após a remoção da cauda de GST, a reação teve que ser purificada para a separação da proteína de interesse da cauda de GST clivada. A purificação foi feita por IMAC, visto que a cauda de GST continha a cauda de histidina capaz de se ligar à resina TALON. A fração correspondente ao CBM, que não interagiu com a coluna, foi recuperada após a injeção da amostra.

Posteriormente, a fração correspondente ao E-1 CBM foi concentrada em concentrador do tipo VivaSpin 2 (GE Healthcare) com corte de 3 kDa. Em seguida, uma

solução contendo 6,25 mM de celohexose foi adicionada à amostra, com o objetivo de se obter cristais do E-1 CBM complexados com o oligossacarídeo.

Os testes de cristalização foram realizados pelo método de matriz esparsa, através do *nL Mosquito liquido handling robot* através dos seguintes kits comerciais: Structure Screen 1 & 2 HT-96, JCSG-plus™ Screen e Pact premier™, todos da Molecular Dimensions.

## 5.4. Resultados e Discussão

### 5.4.1. Clonagem em pET-28a

Através da reação de PCR foi amplificado um fragmento correspondente à região C-terminal de E-1, correspondente ao E-1 CBM, com o tamanho esperado de aproximadamente 0,28 kb (Figura 61A). Como se pode observar na Figura 61B, por digestão enzimática foi possível a confirmação da clonagem do fragmento de interesse no vetor pET-28a. O sequenciamento do inserto garantiu que nenhuma mutação ocorreu no gene e que ele foi clonado na fase de leitura correta.

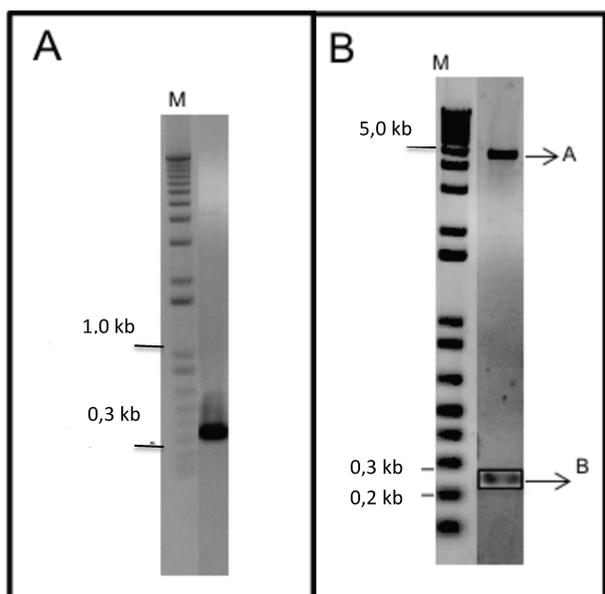


Figura 61 – A - Amplificação do fragmento correspondente a E-1 CBM por PCR. B - Confirmação da clonagem em pET-28a por digestão com enzimas de restrição, onde (A) corresponde ao vetor pET-28a linearizado e (B) ao fragmento liberado.

Através da ferramenta *ProtParam* foram determinados alguns parâmetros bioquímicos e biofísicos de E-1 CBM (Tabela 20), como: peso molecular, ponto isoelétrico, coeficiente de extinção molar, índice de hidrofobicidade (onde valores negativos indicam que a proteína apresenta possibilidade de ser solúvel). Os dados foram gerados a partir da sequência clonada traduzida em aminoácidos com a adição do segmento do vetor pET-28a que também é traduzido.

Tabela 20 - Parâmetros bioquímicos e biofísicos de E1- CBM clonado em pET-28a calculados a partir do ProtParam.

<i>Proteína</i>	<i>Número de aminoácidos</i>	<i>Peso molecular</i>	<i>Ponto isoelétrico</i>	<i>Coeficiente de Extinção Molar</i>	<i>Índice de hidrofobicidade</i>
E-1-CBM	113	12 kDa	9,34	20970 M <sup>-1</sup> . cm <sup>-1</sup>	-0,378

#### 5.4.2. Clonagem em pET-41a

Através da reação de PCR foi amplificado um fragmento com o tamanho esperado de aproximadamente 0,28 kb (Figura 62A). Como se pode observar na Figura 62B, por PCR de colônia foi possível a confirmação da clonagem do fragmento de interesse no vetor pET-41a. O sequenciamento do inserto garantiu que nenhuma mutação ocorreu no gene de interesse e que ele foi clonado na fase de leitura correta.

Através da ferramenta *ProtParam* alguns parâmetros bioquímicos e biofísicos da proteína E-1 CBM fusionada à cauda de GST (Tabela 21) foram determinados, como: peso molecular, ponto isoelétrico, coeficiente de extinção molar, índice de hidrofobicidade (onde valores negativos indicam que a proteína apresenta possibilidade de ser solúvel). Os dados foram gerados a partir da sequência clonada traduzida em aminoácidos com a adição do segmento do vetor pET-41a que também é traduzido e que inclui a proteína GST, que isoladamente apresenta cerca de 27,7 kDa.

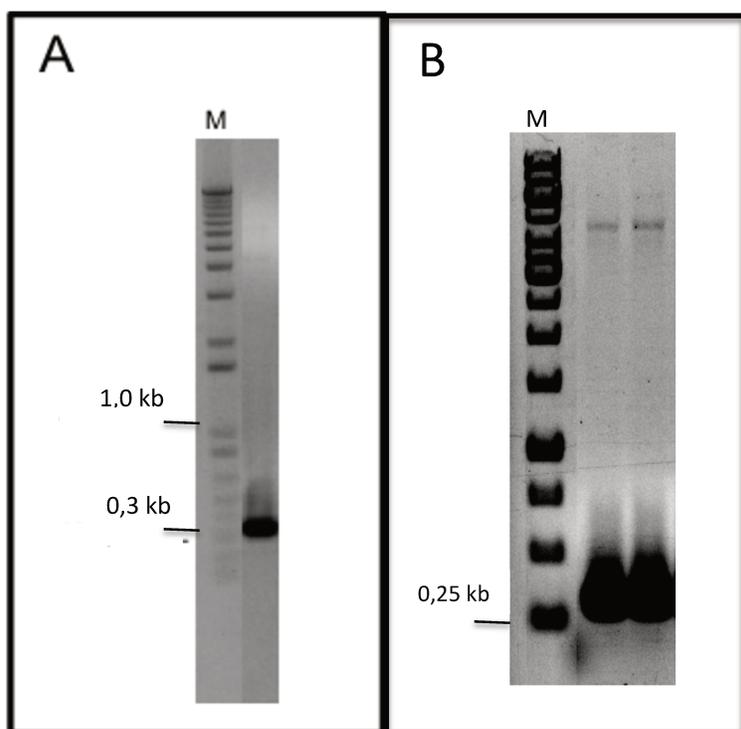


Figura 62 -(A) Amplificação do fragmento correspondente ao gene E-1 CBM por PCR. (B) Confirmação da clonagem em pET-41a por PCR de colônia. M – marcador de peso molecular 1 kb.

Tabela 21 - Parâmetros bioquímicos e biofísicos do E-1 CBM clonado em pET-41a calculados a partir do ProtParam.

<i>Proteína</i>	<i>Número de aminoácidos</i>	<i>Peso molecular</i>	<i>Ponto isoelétrico</i>	<i>Coefficiente de Extinção Molar</i>	<i>Índice de hidrofobicidade</i>
E-1-CBM	378	41,8 kDa	6,5	65695 M <sup>-1</sup> . cm <sup>-1</sup>	-0,444

### 5.4.3. Testes de expressão com o vetor pET-28a

Os resultados dos testes de expressão podem ser visualizados na Figura 63. Podemos observar que nas condições A, B, D, E e F, a maior parte da proteína resultante da expressão foi encontrada na fração insolúvel (em azul), enquanto que uma fração muito pequena foi observada na fração solúvel (em vermelho). Na condição C, foi observada uma pequena quantidade de proteína na fração solúvel.

Em virtude da necessidade de grandes quantidades de proteína para os experimentos de ITC e cristalização, optou-se por testar a expressão do E-1 CBM em pET-41, na tentativa de obtenção de maiores níveis de expressão proteica.

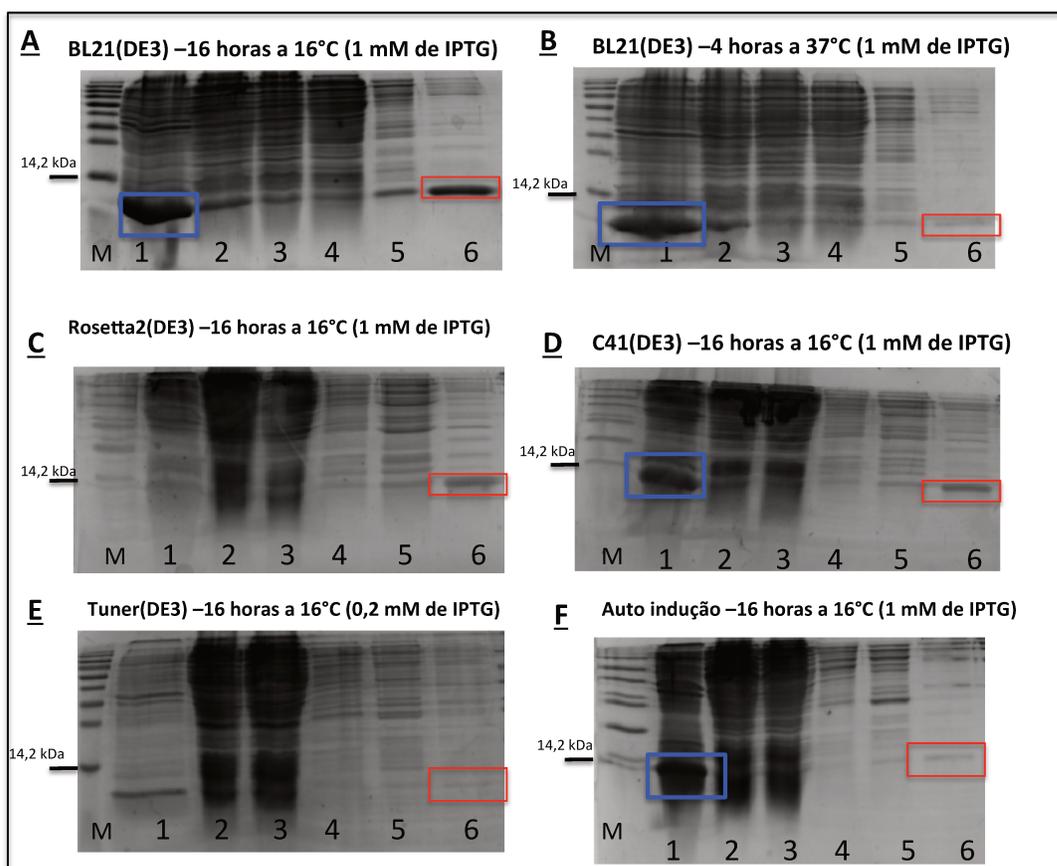


Figura 63 - Testes de expressão do E-1 CBM clonado em pET-28a. 1 - fração insolúvel, 2- fração solúvel, 3 - fração não ligada à resina de cobalto, 4 - lavagem com tampão TALON contendo 10 mM de imidazol, 5 - 1ª lavagem com tampão TALON contendo 100 mM de imidazol, 6 - 2ª lavagem com tampão TALON contendo 100 mM de imidazol. M – Marcador de peso molecular.

#### 5.4.4. Teste de expressão com o vetor pET-41a

A Figura 64, apresenta o resultado da expressão em BL21(DE3) do E-1 CBM clonado pET-41a nas condições testadas. No caso da expressão a 37°C, grande parte da proteína foi expressa na forma insolúvel (em azul). Por outro lado, a expressão a

16°C por 16 horas resultou na obtenção de grande quantidade da proteína de interesse na fração solúvel (em vermelho). Em virtude disso, essa condição foi definida como sendo a mais apropriada para a expressão heteróloga do E-1 CBM.

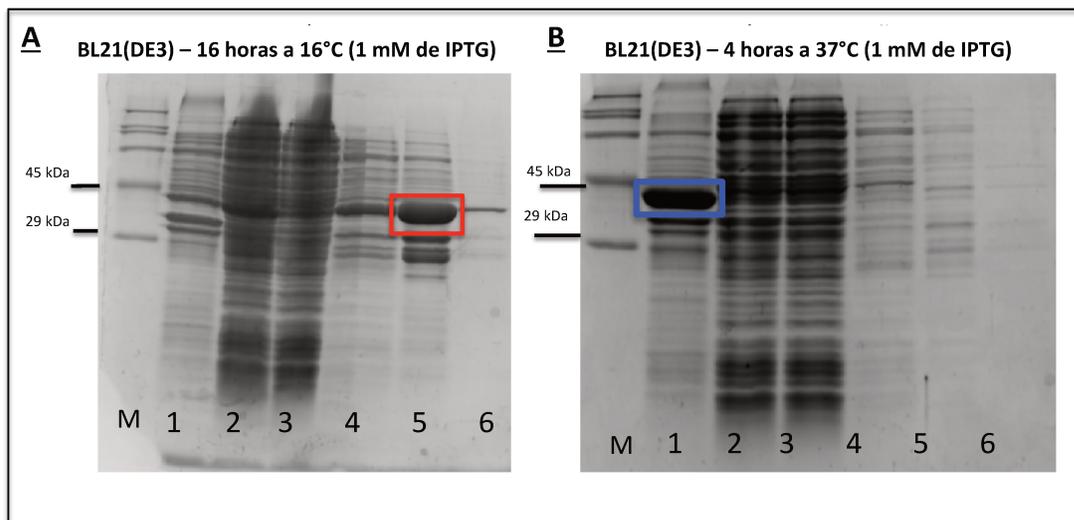


Figura 64 - Testes de expressão do E-1 CBM clonado em pET-41a. 1 - fração insolúvel, 2 - fração solúvel, 3 - fração não ligada à resina de cobalto, 4 - lavagem com tampão TALON contendo 10 mM de imidazol, 5 - 1ª lavagem com tampão TALON contendo 100 mM de imidazol, 6 - 2ª lavagem com tampão TALON contendo 100 mM de imidazol. M – Marcador de peso molecular.

#### 5.4.5. Calorimetria de titulação isotérmica

A capacidade e afinidade de ligação do CBM a diversos substratos foram avaliadas por meio da técnica de ITC. A cauda de GST não apresenta interação com os tipos de ligantes testados e por isso sua remoção não foi necessária. Como forma de garantir que o calor de diluição do ligante no tampão não interferia nos dados obtidos, primeiramente foram realizadas as titulações dos ligantes no tampão HEPES 50 mM, pH 8,0. Através desses experimentos foi confirmado que não ocorria liberação significativa de energia (dados não mostrados).

A Figura 65 mostra alguns exemplos de titulações, enquanto a Tabela 22 apresenta os parâmetros termodinâmicos calculados para cada uma das titulações

realizadas. Foi observado que o E-1 CBM apresentou maior afinidade de ligação com celohexose e  $\beta$ -glucano, com constantes de associação ( $K_a$ ) na ordem de  $10^4 \text{ M}^{-1}$ , seguido por xiloglucano e celopentose com  $K_a$  aproximadamente 10 vezes menor.

Os dados da Tabela 22 demonstram ainda que de forma similar a outros CBMs capazes de ligar a substratos solúveis (BORASTON et al., 2002, BOLAM et al., 2004, HENSHAW et al., 2006), a interação do E-1 CBM com polissacarídeos e oligossacarídeos foi dirigida por mudanças na entalpia ( $\Delta H$ ) e que a diminuição do valor entropia ( $\Delta S$ ) teve impacto negativo na afinidade.

O E-1 CBM foi classificado como tipo B, em função da capacidade de se ligar a cadeias únicas e livres de polissacarídeos solúveis e oligossacarídeos.

Como já descrito em outros trabalhos, a afinidade do CBM por oligossacarídeos está diretamente associada ao tamanho da cadeia (BORASTON et al., 2003, BOLAM et al., 2004), o que corrobora os resultados que mostraram que a afinidade de ligação do E-1 CBM com celopentose foi cerca de 10 vezes inferior à celohexose. Além disso, o fato do E-1 CBM não apresentar capacidade de ligação à celotetrose, indicou a necessidade de uma cadeia com pelo menos 5 unidades de glicoses para a interação do CBM com celo-oligossacarídeos.

Através dos dados em XXXG, foi verificado que o E-1 CBM não foi capaz de se ligar nos pontos onde ocorrem substituições na cadeia principal. Porém, a capacidade de ligação ao xiloglucano, que apresenta de 60 a 75% dos resíduos de glicose substituídos na posição 6 (FRY, 1989), mostrou que a presença de substituições não interfere no reconhecimento da cadeia principal de glucano. Além disso, E-1 CBM foi capaz de se ligar aos polissacarídeos com mais de um tipo de ligação entre as unidades de glicose, como no caso do  $\beta$ -glucano.

A não interação do E-1 CBM com manohexose, xilohexose e xilana *beechwood* demonstrou sua habilidade em reconhecer exclusivamente cadeias de glucano. Essa característica é compatível com o tipo de substrato degradado pelo domínio catalítico ao qual o E-1 CBM está ligado em sua forma nativa.

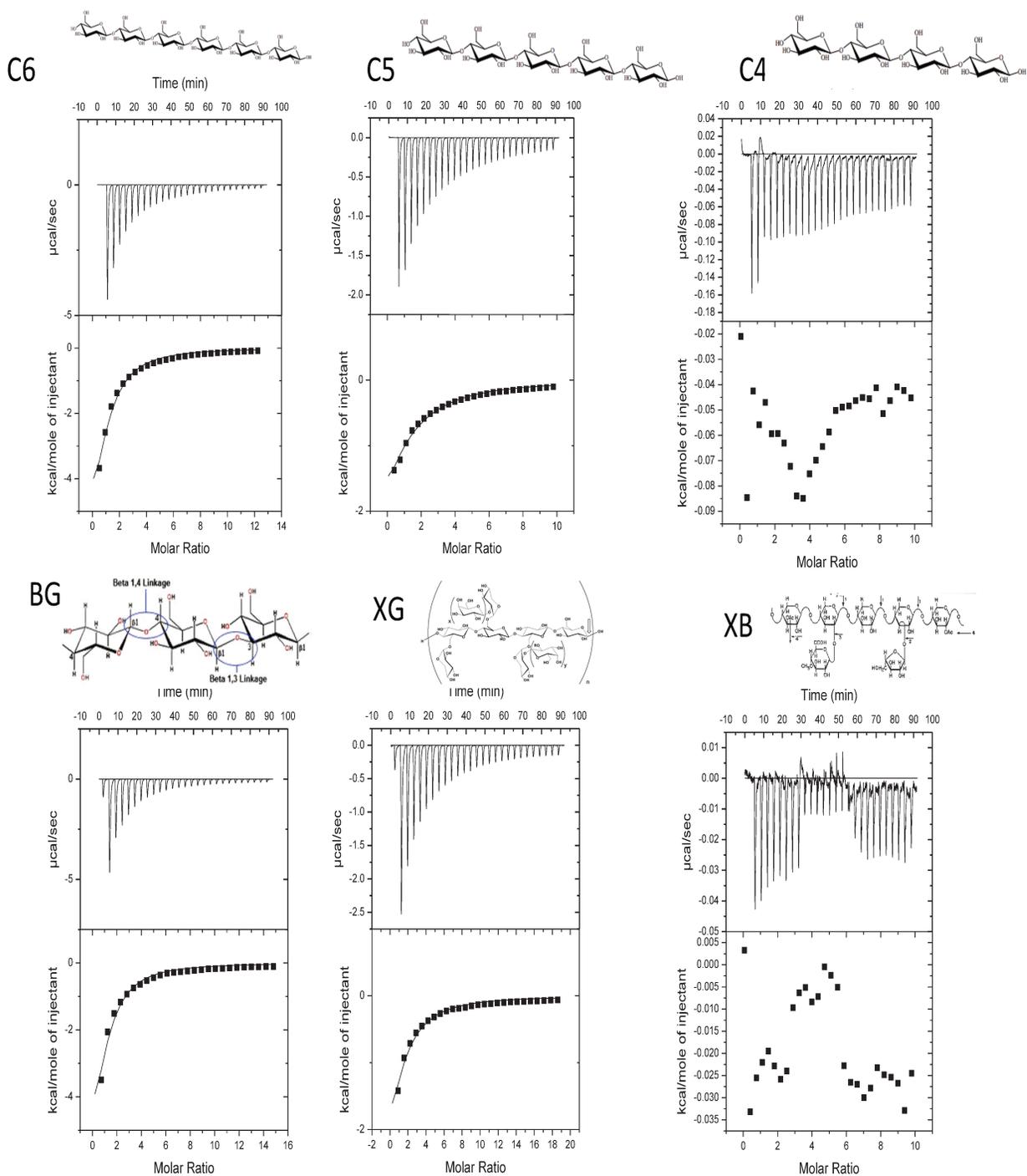


Figura 65 – Resultados dos experimentos de ITC do E-1 CBM titulado com celohexose (C6), celopentose (C5), celotetrose (C4),  $\beta$ -glucano (BG), xiloglucano (XG) e xilana de madeira de faia (XB).

Tabela 22 – Afinidade e parâmetros termodinâmicos da ligação do E-1 CBM com diferentes polissacarídeos e oligossacarídeos determinados através de experimentos de ITC.

Ligante	N	$K_a$ ( $M^{-1}$ )	$\Delta H$ ( $kcal.mol^{-1}$ )	$T\Delta S$ ( $kcal.mol^{-1}$ )	$\Delta G$ ( $kcal.mol^{-1}$ )
celohexose	$0,995 \pm 0,172$	$1,23 \times 10^4 \pm 2 \times 10^3$	$-8,25 \pm 1,64$	-2,67	-5,58
celopentose	$0,995 \pm 0,285$	$4,33 \times 10^3 \pm 0,634$	$-5,1 \pm 1,67$	-0,115	-4,99
celotetrose	<i>Não houve interação</i>				
$\beta$ -glucano	$1,00 \pm 0,304$	$1,11 \times 10^4 \pm 2,7 \times 10^3$	$-8,78 \pm 3,01$	-3,25	-5,53
Xilo-glucano	$1,03 \pm 0,474$	$5,03 \times 10^3 \pm 1,01 \times 10^3$	$-5,56 \pm 2,76$	-0,51	-5,05
XXXG	<i>Não houve interação</i>				
manohexose	<i>Não houve interação</i>				
xilohexose	<i>Não houve interação</i>				
Xilano Beechwood	<i>Não houve interação</i>				

#### 5.4.6. Remoção da cauda de GST

Em relação à digestão com enteroquinase, a Figura 66 mostra que não houve a clivagem da cauda de GST em nenhuma das condições testadas. Isso pode ser explicado pela possibilidade do sítio de clivagem se localizar em uma região de difícil acesso para a enteroquinase em função de características do enovelamento proteico.

Por outro lado, os resultados dos testes de clivagem com trombina resultaram na separação entre a cauda de GST e o E-1 CBM (Figura 67), resultando no aparecimento de outras duas bandas menores, com tamanhos próximos do esperado de 27,7 kDa para a cauda de GST e 14,1 kDa para E-1 CBM. O valor de 14,1 kDa foi calculado levando em consideração a sequência de aminoácidos correspondente ao E-1 CBM e os aminoácidos adicionais codificados pela sequência do vetor e que não são clivados pela trombina, em virtude do sítio de clivagem estar relativamente distante do sítio de clonagem. Considerando que o tamanho predito com base na sequência de

aminoácidos do E-1 CBM é de 9,5 kDa, observou-se que uma quantidade elevada de resíduos de aminoácidos permaneceu na sequência após a clivagem. A presença desses resíduos adicionais poderia interferir negativamente nas etapas de cristalização proteica, visto que não fazem parte da sequência de aminoácidos original do E-1 CBM.

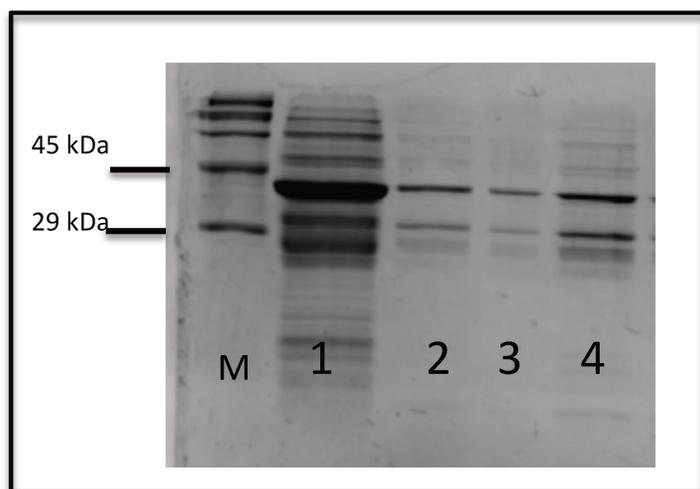


Figura 66 – Digestão do E-1 CBM fusionado à cauda de GST com enteroquinase. 1 - Reação controle sem enteroquinase, 2 - Reação com 0,2 U de enteroquinase, 3 - Reação com 0,5 U de enteroquinase e 4 - Reação com 1,0 U de enteroquinase. M – Marcador de peso molecular.

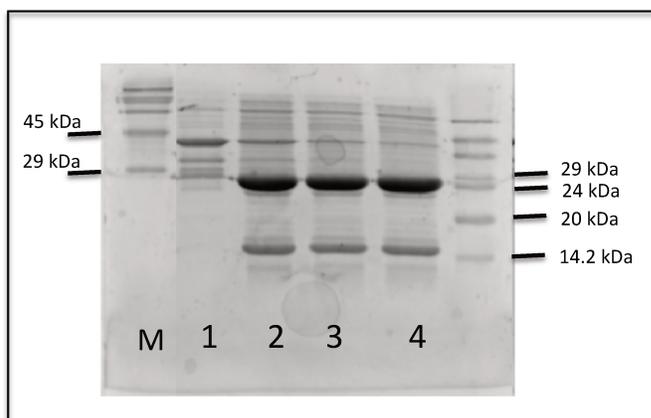


Figura 67 - Digestão do E-1 CBM fusionado à cauda de GST com trombina. 1 - Reação controle sem trombina, 2 - Reação com 0,2 U de trombina, 3 - Reação com 0,5 U de trombina e 4 - Reação com 1,0 U trombina. M – marcador de peso molecular.

### 5.4.7. Testes de cristalização

Após a clivagem da cauda de GST, a amostra foi submetida à purificação por IMAC. A amostra que não interagiu com a coluna por não possuir a cauda de histidina, correspondente ao E-1 CBM, foi coletada e concentrada, atingindo concentração final aproximada de 8,6 mg/mL. A amostra (Figura 68) foi submetida aos testes de cristalização, porém não foi verificada a formação de cristais em nenhuma das condições testadas.

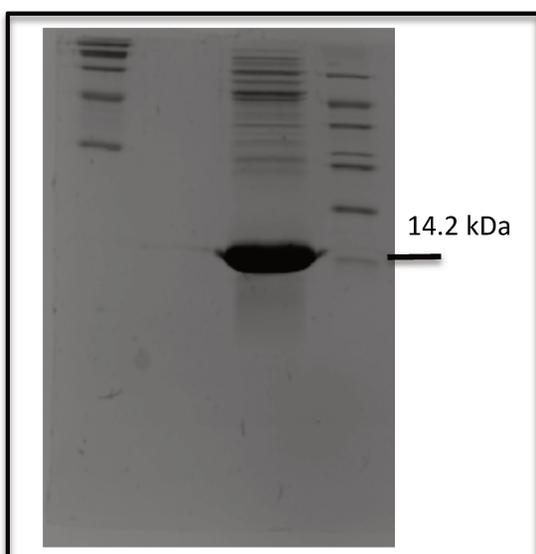


Figura 68 – Amostra de E-1 CBM empregada nos testes de cristalização.

## 5.5. Conclusões

Os resultados obtidos nesta etapa do trabalho confirmaram a identidade do domínio proteico E-1 CBM como sendo um típico CBM, pois ao ser clonado isoladamente do domínio catalítico continuou apresentando a capacidade de se ligar a diversos polissacarídeos. O fato de E-1 CBM não apresentar identidade de sequência com domínios conservados de famílias de CBMs conhecidas tornou necessário o estudo das características de interação deste CBM com diferentes ligantes.

Através dos experimentos de ITC, foi observado que o E-1 CBM apresentou especificidade de ligação a cadeias de glucano, o que é consistente com o substrato alvo de degradação do domínio catalítico ao qual o CBM se encontra ligado. As afinidades de ligação seguiram a seguinte ordem decrescente: celohexose,  $\beta$ -glucano, xiloglucano e celopentose. Em relação à ligação com celo-oligossacarídeos, verificou-se que o E1-CBM foi capaz de interagir com celohexose e celopentose, mas não com celotetose, indicando a necessidade de pelo menos 5 unidades de glicose para a ocorrência da interação.

A capacidade do E-1 CBM de interagir com  $\beta$ -glucano mostrou que ele é capaz de reconhecer cadeias de glucano que apresentem em sua composição ligações glicosídicas do tipo  $\beta(1\rightarrow3)$  e  $\beta(1\rightarrow4)$  intercaladas.

A interação de E-1 CBM com xiloglucano foi avaliada em detalhes através do ensaio realizado com XXXG, que constitui a principal unidade repetitiva da cadeia de xiloglucano. Neste caso não houve interação do E1-CBM com o ligante, demonstrando que a interação do E-1 CBM com a cadeia de xiloglucano não ocorre nesses pontos de substituição, mas que a presença deles não impede o reconhecimento da cadeia principal de glucano.

Além disso, a capacidade do E-1 CBM em interagir com polissacarídeos solúveis e oligossacarídeos, permitiu sua classificação como CBM do tipo B.

Em relação aos parâmetros termodinâmicos foi observado que a reação de ligação foi dirigida por mudanças na entalpia e que a diminuição da entropia teve impactos negativos na afinidade de interação, em concordância com o que já foi descrito para outros CBMs capazes de se ligar a ligantes solúveis.

Estudos futuros em relação à estrutura do E-1 CBM serão realizados através da técnica de ressonância nuclear magnética. Tais estudos serão importantes para avaliação do mecanismo de interação do E-1 com diferentes ligantes, de forma a contribuir com o entendimento do mecanismo de reconhecimento dos CBMs a diferentes carboidratos.

## CONCLUSÕES GERAIS DA TESE

Neste trabalho descrevemos com sucesso a identificação de novos genes codificantes para glicosil hidrolases através da triagem funcional de uma biblioteca metagenômica construída com DNA extraído de um consórcio microbiano derivado de amostra de solo de canavial.

A avaliação das sequências de aminoácidos das proteínas codificadas por tais genes resultou na identificação de duas celulases da família GH5, E-1 e E-2, uma celulase da família GH6, E-3, uma xilanase da família GH10, X-1, e uma glicosil hidrolase com identidade para a família GH16, denominada X-2.

A identificação de celulases e xilanases das famílias GH5 e GH10, respectivamente, pela triagem funcional de bibliotecas metagenômicas é frequentemente descrita na literatura. Por outro lado, a identificação de enzimas da família GH6 e GH16 é pouco comum. Isso demonstra que as metodologias empregadas para a construção e triagem funcional das bibliotecas metagenômicas permitiram a identificação de membros de diferentes famílias de glicosil hidrolases inclusive daquelas cujas atividades enzimáticas não são comumente expressas em bibliotecas metagenômicas.

Além disso, a frequência de clones positivos identificados em relação ao número de megabases submetidos à triagem funcional foi superior ao encontrado em outros trabalhos de metagenômica. Isso pode estar diretamente relacionado à etapa de enriquecimento que deu origem ao consórcio microbiano, resultando no aumento da frequência dos genes de interesse.

Em relação à celulase E-1 foi constatado tratar-se de uma proteína com dois domínios, sendo um deles o domínio catalítico, E-1 Cat, e o outro o domínio de ligação a carboidratos, E-1 CBM. Enquanto E-1 Cat apresentou identidade com domínios conservados das celulases da família GH5, no caso do E-1 CBM não foi observada identidade com nenhum domínio proteico conhecido, inclusive de famílias de CBMs conhecidas. Dessa forma, nesse trabalho propomos que o E-1 CBM deve ser o membro de uma nova família de CBMs. A caracterização funcional do E-1 CBM, permitiu sua classificação como sendo do tipo B, pois foi capaz de se ligar a oligossacarídeos e

polissacarídeos solúveis. O E-1 CBM apresentou especificidade de ligação a cadeias de glucano com grau de polimerização de no mínimo cinco unidades de glicose, com ligações do tipo  $\beta(1\rightarrow4)$  ou ligações mistas do tipo  $\beta(1\rightarrow4)$  e  $\beta(1\rightarrow3)$ . Estudos posteriores serão realizados visando à determinação da estrutura tridimensional deste CBM de forma a desvendar o mecanismo de reconhecimento e interação com diferentes ligantes.

A celulase E-1 Cat e a xilanase X-1 foram submetidas às etapas de caracterização bioquímica e os resultados obtidos mostraram que ambas apresentaram condições de atuação referentes à temperatura, pH e estabilidade térmica adequadas para aplicação em processos industriais.

E-1 Cat apresentou ainda elevada eficiência catalítica e foi capaz de liberar açúcares solúveis diretamente de celulose cristalina e bagaço de cana-de-açúcar explodido, evidenciando seu potencial uso como um componente de coquetéis enzimáticos utilizados na conversão da biomassa vegetal.

Em relação à xilanase X-1, foi observada sua capacidade de hidrolisar ligações glicosídicas do tipo  $\beta(1\rightarrow4)$  internas e externas entre resíduos de xilose presentes em polissacarídeos complexos, como bagaço de cana-de-açúcar pré-tratado com ácido fosfórico, xilana de madeira de faia e arabinoxilana de trigo. A hidrólise desses polissacarídeos resultou na liberação de xilo-oligossacarídeos, xilobiase e xilose, que são compostos empregados em indústrias alimentícias e de biocombustíveis para a geração de bio-produtos. Além disso, a etapa de hidrólise de bagaço de cana-de-açúcar pré-tratado com ácido fosfórico com X-1 antes da etapa de sacarificação com coquetel celulolítico resultou no aumento da quantidade de açúcares solúveis liberada.

Dessa forma, conclui-se que as bibliotecas metagenômicas construídas neste trabalho se mostraram adequadas para a prospecção de novos genes codificantes de glicosil hidrolases. Pelo fato das bibliotecas metagenômicas terem sido construídas com DNA metagenômico de solo, ambiente caracterizado pela enorme diversidade genética, elas podem ser utilizadas para a prospecção de genes envolvidos em diversas atividades metabólicas, constituindo uma fonte ilimitada para prospecção de novas biomoléculas.

## REFERÊNCIAS BIBLIOGRÁFICAS

ACEVEDO, J. P., REYES, F., PARRA, L. P., SALAZAR, O., ANDREWS, B. A. & ASENJO, J. A. 2008. Cloning of complete genes for novel hydrolytic enzymes from Antarctic sea water bacteria by use of an improved genome walking technique. *Journal of Biotechnology*, 133, 277-286.

ALVAREZ, T. M., GOLDBECK, R., DOS SANTOS, C. R., PAIXAO, D. A., GONCALVES, T. A., FRANCO CAIRO, J. P., ALMEIDA, R. F., DE OLIVEIRA PEREIRA, I., JACKSON, G., COTA, J., BUCHLI, F., CITADINI, A. P., RULLER, R., POLO, C. C., DE OLIVEIRA NETO, M., MURAKAMI, M. T. & SQUINA, F. M. 2013. Development and Biotechnological Application of a Novel Endoxylanase Family GH10 Identified from Sugarcane Soil Metagenome. *PloS ONE*, 8, e70014.

ANDREOTE, F. D., JIMENEZ, D. J., CHAVES, D., DIAS, A. C., LUVIZOTTO, D. M., DINI-ANDREOTE, F., FASANELLA, C. C., LOPEZ, M. V., BAENA, S., TAKETANI, R. G. & DE MELO, I. S. 2012. The microbiome of Brazilian mangrove sediments as revealed by metagenomics. *PloS ONE*, 7, e38600.

ANFAVEA. 2013. Estatísticas. Disponível em: <http://www.anfavea.com.br/tabelas.html%3E>. Acesso em 04 de março de 2013.

ANISH, R. & RAO, M. 2007. Biochemical characterization of a novel beta-1-3,1-4 glucan 4-glucanohydrolase from *Thermomonospora* sp. having a single active site for lichenan and xylan. *Biochimie*, 89, 1489-1497.

ARO, N., PAKULA, T. & PENTTILA, M. 2005. Transcriptional regulation of plant cell wall degradation by filamentous fungi. *FEMS Microbiology Reviews*, 29, 719-739.

BELOQUI, A., NECHITAYLO, T. Y., LOPEZ-CORTES, N., GHAZI, A., GUAZZARONI, M. E., POLAINA, J., STRITTMATTER, A. W., REVA, O., WALICZEK, A., YAKIMOV, M. M., GOLYSHINA, O. V., FERRER, M. & GOLYSHIN, P. N. 2010. Diversity of glycosyl hydrolases from cellulose-depleting communities enriched from casts of two earthworm species. *Applied and Environmental Microbiology*, 76, 5934-5946.

BERA-MAILLET, C., ARTHAUD, L., ABAD, P. & ROSSO, M. N. 2000. Biochemical characterization of MI-ENG1, a family 5 endoglucanase secreted by the root-knot nematode *Meloidogyne incognita*. *European Journal of Biochemistry / FEBS*, 267, 3255-3263.

BERLEMONT, R., DELSAUTE, M., PIPERS, D., D'AMICO, S., FELLER, G., GALLEN, M. & POWER, P. 2009. Insights into bacterial cellulose biosynthesis by functional metagenomics on Antarctic soil samples. *The ISME Journal*, 3, 1070-1081.

BERNIER, R., DESROCHERS, M., JURASEK, L. & PAICE, M. G. 1983. Isolation and Characterization of a Xylanase from *Bacillus subtilis*. *Applied and Environmental Microbiology*, 46, 511-514.

BLACK, G. W., RIXON, J. E., CLARKE, J. H., HAZLEWOOD, G. P., THEODOROU, M. K., MORRIS, P. & GILBERT, H. J. 1996. Evidence that linker sequences and cellulose-binding domains enhance the activity of hemicellulases against complex substrates. *The Biochemical Journal*, 319, 515-520.

BOERJAN, W., RALPH, J. & BAUCHER, M. 2003. Lignin biosynthesis. *Annual Review of Plant Biology*, 54, 519-546.

BOLAM, D. N., XIE, H., PELL, G., HOGG, D., GALBRAITH, G., HENRISSAT, B. & GILBERT, H. J. 2004. X4 modules represent a new family of carbohydrate-binding modules that display novel properties. *The Journal of Biological Chemistry*, 279, 22953-22963.

BORASTON, A. B., BOLAM, D. N., GILBERT, H. J. & DAVIES, G. J. 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *The Biochemical Journal*, 382, 769-781.

BORASTON, A. B., NURIZZO, D., NOTENBOOM, V., DUCROS, V., ROSE, D. R., KILBURN, D. G. & DAVIES, G. J. 2002. Differential oligosaccharide recognition by evolutionarily-related beta-1,4 and beta-1,3 glucan-binding modules. *Journal of Molecular Biology*, 319, 1143-1156.

BORASTON, A. B., REVETT, T. J., BORASTON, C. M., NURIZZO, D. & DAVIES, G. J. 2003. Structural and thermodynamic dissection of specific mannan recognition by a carbohydrate binding module, TmCBM27. *Structure*, 11, 665-675.

BRAITHWAITE, K. L., BLACK, G. W., HAZLEWOOD, G. P., ALI, B. R. & GILBERT, H. J. 1995. A non-modular endo-beta-1,4-mannanase from *Pseudomonas fluorescens* subspecies cellulosa. *The Biochemical Journal*, 305, 1005-1010.

BRASIL. 2009. Balanço Energético Nacional (BEN) 2009: ano base 2008 - relatório final. *Ministério de Minas e Energia (MME)*.

BRENNAN, Y., CALLEN, W. N., CHRISTOFFERSEN, L., DUPREE, P., GOUBET, F., HEALEY, S., HERNANDEZ, M., KELLER, M., LI, K., PALACKAL, N., SITTENFELD, A., TAMAYO, G., WELLS, S., HAZLEWOOD, G. P., MATHUR, E. J., SHORT, J. M., ROBERTSON, D. E. & STEER, B. A. 2004. Unusual microbial xylanases from insect guts. *Applied and Environmental Microbiology*, 70, 3609-3617.

BRETT, C. T. 2000. Cellulose microfibrils in plants: biosynthesis, deposition, and integration into the cell wall. *International Review of Cytology*, 199, 161-199.

BUSHNELL, L. D. & HAAS, H. F. 1941. The Utilization of Certain Hydrocarbons by Microorganisms. *Journal of Bacteriology*, 41, 653-673.

CADENAS, A. & CABEZUDO, S. 1998. Biofuels as sustainable technologies: perspectives for less developed countries. *Technological Forecasting and Social Change*, 58, 83-103.

CARDONA, C. A., QUINTERO, J. A. & PAZ, I. C. 2010. Production of bioethanol from sugarcane bagasse: Status and perspectives. *Bioresource Technology*, 101, 4754-4766.

CARDONA, C. A. & SANCHEZ, O. J. 2007. Fuel ethanol production: process design trends and integration opportunities. *Bioresource Technology*, 98, 2415-2457.

CARPITA, N. C. & GIBEAUT, D. M. 1993. Structural models of primary cell walls in flowering plants: consistency of 11 structure with the physical properties of the walls during growth. *The Plant Journal : for Cell and Molecular Biology*, 3, 1-30.

CARVALHO, W., CANILHA, L., FERRAZ, A. & MILAGRES, A. M. F. 2009. Uma visão sobre a estrutura, composição e biodegradação da madeira. *Química Nova*, 32, 2191-2195.

CASTRO, A. M. & PEREIRA JR., N. 2010. Produção, propriedades e aplicação de celulases na hidrólise de resíduos agroindustriais. *Química Nova*, 33, 181-188.

CERQUEIRA LEITE, R. C. 2009. Bioetanol combustível: uma oportunidade para o Brasil. Brasília: Centro de Gestão e Estudos Estratégicos.

CHENG, F., SHENG, J., CAI, T., JIN, J., LIU, W., LIN, Y., DU, Y., ZHANG, M. & SHEN, L. 2012. A protease-insensitive feruloyl esterase from China Holstein cow rumen metagenomic library: expression, characterization, and utilization in ferulic acid release from wheat straw. *Journal of Agricultural and Food Chemistry*, 60, 2546-2553.

CHOI, N. S., KIM, B. H., PARK, C. S., HAN, Y. J., LEE, H. W., CHOI, J. H., LEE, S. G. & SONG, J. J. 2009. Multiple-layer substrate zymography for detection of several enzymes in a single sodium dodecyl sulfate gel. *Analytical Biochemistry*, 386, 121-122.

CHRISTENSEN, L. & SCHNORR, K. 2007. Carbohydrate-binding modules of a new family. United States Patent (Patent No. US 20070072185 A1).

COHEN, R., SUZUKI, M. R. & HAMMEL, K. E. 2005. Processive endoglucanase active in crystalline cellulose hydrolysis by the brown rot basidiomycete *Gloeophyllum trabeum*. *Applied and Environmental Microbiology*, 71, 2412-2417.

CORTEZ, L. A. B., LORA, E. E. S. & GÓMEZ, E. O. 2008. Biomassa para Bioenergia. *Campinas: Editora da Unicamp. Cap.13, 435-473.*

COSGROVE, D. J. 2001. Wall structure and wall loosening. A look backwards and forwards. *Plant Physiology*, 125, 131-134.

COSGROVE, D. J. 2005. Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology*, 6, 850-861.

COTTRELL, M. T., MOORE, J. A. & KIRCHMAN, D. L. 1999. Chitinases from uncultured marine microorganisms. *Applied and Environmental Microbiology*, 65, 2553-2557.

COURTOIS, S., CAPPELLANO, C. M., BALL, M., FRANCOU, F. X., NORMAND, P., HELYNCK, G., MARTINEZ, A., KOLVEK, S. J., HOPKE, J., OSBURNE, M. S., AUGUST, P. R., NALIN, R., GUERINEAU, M., JEANNIN, P., SIMONET, P. & PERNODET, J. L. 2003. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Applied and Environmental Microbiology*, 69, 49-55.

COWAN, D., MEYER, Q., STAFFORD, W., MUYANGA, S., CAMERON, R. & WITTEWER, P. 2005. Metagenomic gene discovery: past, present and future. *Trends in Biotechnology*, 23, 321-329.

DANIEL, R. 2005. The metagenomics of soil. *Nature Reviews Microbiology*, 3, 470-478.

DE VRIES, R. P. & VISSER, J. 2001. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microbiology and Molecular Biology Reviews: MMBR*, 65, 497-522.

DEL POZO, M. V., FERNANDEZ-ARROJO, L., GIL-MARTINEZ, J., MONTESINOS, A., CHERNIKOVA, T. N., NECHITAYLO, T. Y., WALISZEK, A., TORTAJADA, M., ROJAS, A., HUWS, S. A., GOLYSHINA, O. V., NEWBOLD, C. J., POLAINA, J., FERRER, M. & GOLYSHIN, P. N. 2012. Microbial beta-glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail. *Biotechnology for Biofuels*, 5, 73.

DEMIRBAS, A. 2007. Progress and recent trends in biofuels. *Progress in Energy and Combustion Science*, 33, 1-18.

DOI, R. H. 2008. Cellulases of mesophilic microorganisms: cellulosome and noncellulosome producers. *Annals of the New York Academy of Sciences*, 1125, 267-279.

DON, R. H., COX, P. T., WAINWRIGHT, B. J., BAKER, K. & MATTICK, J. S. 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research*, 19, 4008.

DUAN, C. J. & FENG, J. X. 2010. Mining metagenomes for novel cellulase genes. *Biotechnology Letters*, 32, 1765-1775.

DURRANT, A. J., HALL, J., HAZLEWOOD, G. P. & GILBERT, H. J. 1991. The non-catalytic C-terminal region of endoglucanase E from *Clostridium thermocellum* contains a cellulose-binding domain. *The Biochemical Journal*, 273, 289-293.

EKLOF, J. M. & BRUMER, H. 2010. The XTH gene family: an update on enzyme structure, function, and phylogeny in xyloglucan remodeling. *Plant Physiology*, 153, 456-466.

ENTCHEVA, P., LIEBL, W., JOHANN, A., HARTSCH, T. & STREIT, W. R. 2001. Direct cloning from enrichment cultures, a reliable strategy for isolation of complete operons and genes from microbial consortia. *Applied and Environmental Microbiology*, 67, 89-99.

EVANGELISTA, R. A., LIU, M. S. & CHEN, F. A. 1995. Characterization of 5-aminopyrene-1,4,6- trisulfonate-derivatized sugars by capillary electrophoresis with laser-induced fluorescence detection. *Analytical Chemistry*, 67, 2239-2245.

FANG, Z., FANG, W., LIU, J., HONG, Y., PENG, H., ZHANG, X., SUN, B. & XIAO, Y. 2010. Cloning and characterization of a beta-glucosidase from marine microbial metagenome with excellent glucose tolerance. *Journal of Microbiology and Biotechnology*, 20, 1351-1358.

FARINAS, C. S. 2011. A parede celular vegetal e as enzimas envolvidas na sua degradação. São Carlos: Embrapa Instrumentação.

FENG, Y., DUAN, C. J., PANG, H., MO, X. C., WU, C. F., YU, Y., HU, Y. L., WEI, J., TANG, J. L. & FENG, J. X. 2007. Cloning and identification of novel cellulase genes from uncultured microorganisms in rabbit cecum and characterization of the expressed cellulases. *Applied Microbiology and Biotechnology*, 75, 319-328.

FERCHAK, J. D. & PYE, E. K. 1983. Effect of cellobiose, glucose, ethanol, and metal ions on the cellulase enzyme complex of *Thermomonospora fusca*. *Biotechnology and Bioengineering*, 25, 2865-2872.

FERRER, M., BELOQUI, A., TIMMIS, K. N. & GOLYSHIN, P. N. 2009. Metagenomics for mining new genetic resources of microbial communities. *Journal of Molecular Microbiology and Biotechnology*, 16, 109-123.

FIERER, N., LEFF, J. W., ADAMS, B. J., NIELSEN, U. N., BATES, S. T., LAUBER, C. L., OWENS, S., GILBERT, J. A., WALL, D. H. & CAPORASO, J. G. 2012. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 21390-21395.

FONTANA, A., DE LAURETO, P. P., SPOLAORE, B., FRARE, E., PICOTTI, P. & ZAMBONIN, M. 2004. Probing protein structure by limited proteolysis. *Acta Biochimica Polonica*, 51, 299-321.

FRANCO CAIRO, J. P., OLIVEIRA, L. C., UCHIMA, C. A., ALVAREZ, T. M., CITADINI, A. P., COTA, J., LEONARDO, F. C., COSTA-LEONARDO, A. M., CARAZZOLLE, M. F., COSTA, F. F., PEREIRA, G. A. & SQUINA, F. M. 2013. Deciphering the synergism of endogenous glycoside hydrolase families 1 and 9 from *Coptotermes gestroi*. *Insect Biochemistry and Molecular Biology*, 43, 970-981.

FRANCO, P. F., FERREIRA, H. M. & FILHO, E. X. 2004. Production and characterization of hemicellulase activities from *Trichoderma harzianum* strain T4. *Biotechnology and Applied Biochemistry*, 40, 255-259.

FREELove, A. C. J., BOLAM, D. N., WHITE, P., HAZLEWOOD, G. P. & GILBERT, H. J. 2001. A novel carbohydrate-binding protein is a component of the plant cell wall-degrading complex of *Piromyces equi*. *The Journal of Biological Chemistry*, 276, 43010-43017.

FRY, S. C. 1989. The structure and functions of xyloglucan. *Journal of Experimental Biology*, 40, 1-11.

FUJITA, Y., ITO, J., UEDA, M., FUKUDA, H. & KONDO, A. 2004. Synergistic saccharification, and direct fermentation to ethanol, of amorphous cellulose by use of an engineered yeast strain codisplaying three types of cellulolytic enzyme. *Applied and Environmental Microbiology*, 70, 1207-1212.

FUKUMORI, F., KUDO, T., NARAHASHI, Y. & HORIKOSHI, K. 1986. Molecular cloning and nucleotide sequence of the alkaline cellulase gene from the alkalophilic *Bacillus* sp. strain 1139. *Journal of General Microbiology*, 132, 2329-2335.

GABOR, E. M., DE VRIES, E. J. & JANSSEN, D. B. 2004. Construction, characterization, and use of small-insert gene banks of DNA isolated from soil and enrichment cultures for the recovery of novel amidases. *Environmental Microbiology*, 6, 948-958.

GAUR, R., LATA & KHARE, S. K. 2005. Immobilization of xylan-degrading enzymes from *Scytalidium thermophilum* on Eudragit L-100. *World Journal of Microbiology & Biotechnology*, 21, 1123-1128.

GILAD, R., RABINOVICH, L., YARON, S., BAYER, E. A., LAMED, R., GILBERT, H. J. & SHOHAM, Y. 2003. Cell, a noncellulosomal family 9 enzyme from *Clostridium thermocellum*, is a processive endoglucanase that degrades crystalline cellulose. *Journal of Bacteriology*, 185, 391-398.

GILBERT, H. J. 2010. The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiology*, 153, 444-455.

GILBERT, H. J., HALL, J., HAZLEWOOD, G. P. & FERREIRA, L. M. 1990. The N-terminal region of an endoglucanase from *Pseudomonas fluorescens* subspecies

cellulosa constitutes a cellulose-binding domain that is distinct from the catalytic centre. *Molecular Microbiology*, 4, 759-767.

GILL, S. R., POP, M., DEBOY, R. T., ECKBURG, P. B., TURNBAUGH, P. J., SAMUEL, B. S., GORDON, J. I., RELMAN, D. A., FRASER-LIGGETT, C. M. & NELSON, K. E. 2006. Metagenomic analysis of the human distal gut microbiome. *Science*, 312, 1355-1359.

GONG, X., GRUNINGER, R. J., QI, M., PATERSON, L., FORSTER, R. J., TEATHER, R. M. & MCALLISTER, T. A. 2012. Cloning and identification of novel hydrolase genes from a dairy cow rumen metagenomic library and characterization of a cellulase gene. *BMC Research Notes*, 5, 566.

GONZALEZ, J. M., PORTILLO, M. C., BELDA-FERRE, P. & MIRA, A. 2012. Amplification by PCR artificially reduces the proportion of the rare biosphere in microbial communities. *PLoS ONE*, 7, e29973.

GRANT, G. A. 2006. The ACT domain: a small molecule binding domain and its role as a common regulatory element. *The Journal of Biological Chemistry*, 281, 33825-33829.

GUILLEN, D., SANCHEZ, S. & RODRIGUEZ-SANOJA, R. 2010. Carbohydrate-binding domains: multiplicity of biological roles. *Applied Microbiology and Biotechnology*, 85, 1241-1249.

GUO, B., CHEN, X. L., SUN, C. Y., ZHOU, B. C. & ZHANG, Y. Z. 2009. Gene cloning, expression and characterization of a new cold-active and salt-tolerant endo-beta-1,4-xylanase from marine *Glaciecola mesophila* KMM 241. *Applied Microbiology and Biotechnology*, 84, 1107-1115.

GUPTA, R., BERG, Q. K. & LORENZ, P. 2002. Bacterial alkaline proteases: molecular approaches and industrial applications. *Applied Microbiology and Biotechnology*, 59, 15-32.

HAARHOFF, J., MOES, C. J., CERFF, C., WYK, W. J., GERISCHER, G. & JANSE, N. J. H. 1999. Characterization and biobleaching effect of hemicellulases produced by thermophilic fungi. *Biotechnology Letters*, 21, 414-420.

HALL, J., BLACK, G. W., FERREIRA, L. M., MILLWARD-SADLER, S. J., ALI, B. R., HAZLEWOOD, G. P. & GILBERT, H. J. 1995. The non-catalytic cellulose-binding domain of a novel cellulase from *Pseudomonas fluorescens* subsp. cellulosa is important for the efficient hydrolysis of Avicel. *The Biochemical Journal*, 309, 749-756.

HAN, S. O., YUKAWA, H., INUI, M. & DOI, R. H. 2004. Isolation and expression of the xynB gene and its product, XynB, a consistent component of the *Clostridium cellulovorans* cellulosome. *Journal of Bacteriology*, 186, 8347-8355.

- HANDELSMAN, J. 2004. Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews: MMBR*, 68, 669-685.
- HEALY, F. G., RAY, R. M., ALDRICH, H. C., WILKIE, A. C., INGRAM, L. O. & SHANMUGAM, K. T. 1995. Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulose. *Applied Microbiology and Biotechnology*, 43, 667-674.
- HENNE, A., SCHMITZ, R. A., BOMEKE, M., GOTTSCHALK, G. & DANIEL, R. 2000. Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. *Applied and Environmental Microbiology*, 66, 3113-3116.
- HENRISSAT, B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *The Biochemical Journal*, 280, 309-316.
- HENRISSAT, B. & BAIROCH, A. 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *The Biochemical Journal*, 293, 781-788.
- HENSEL, A. 1993. Xyloglucan--structure, genesis, and functions of a widely distributed substance group. *Pharmazie in unserer Zeit*, 22, 228-234.
- HENSHAW, J., HORNE-BITSCHY, A., VAN BUEREN, A. L., MONEY, V. A., BOLAM, D. N., CZJZEK, M., EKBORG, N. A., WEINER, R. M., HUTCHESON, S. W., DAVIES, G. J., BORASTON, A. B. & GILBERT, H. J. 2006. Family 6 carbohydrate binding modules in beta-agarases display exquisite selectivity for the non-reducing termini of agarose chains. *The Journal of Biological Chemistry*, 281, 17099-17107.
- HERMOSO, J. A., SANZ-APARICIO, J., MOLINA, R., JUGE, N., GONZALEZ, R. & FAULDS, C. B. 2004. The crystal structure of feruloyl esterase A from *Aspergillus niger* suggests evolutive functional convergence in feruloyl esterase family. *Journal of Molecular Biology*, 338, 495-506.
- HERR, D., BAUMER, F. & DELLWEG, H. 1978. Purification and properties of an extracellular endo-1,4--glucanase from *Lenzites trabea*. *Archives of Microbiology*, 117, 287-292.
- HESS, M., SCZYRBA, A., EGAN, R., KIM, T. W., CHOKHAWALA, H., SCHROTH, G., LUO, S., CLARK, D. S., CHEN, F., ZHANG, T., MACKIE, R. I., PENNACCHIO, L. A., TRINGE, S. G., VISEL, A., WOYKE, T., WANG, Z. & RUBIN, E. M. 2011. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science*, 331, 463-467.
- HILDEN, L. & JOHANSSON, G. 2004. Recent developments on cellulases and carbohydrate-binding modules with cellulose affinity. *Biotechnology Letters*, 26, 1683-1693.

- HIMMEL, M., XU, Q., LUO, Y., DING, S. Y., LAMED, R. & BAYER, E. A. 2010. Microbial enzymes systems for biomass conversion: emerging paradigms. *Biofuels*, 1, 323-341.
- HOLBEN, W. E., JANSSON, J. K., CHELM, B. K. & TIEDJE, J. M. 1988. DNA probe method for the detection of specific microorganisms in the soil bacterial community. *Applied and Environmental Microbiology*, 54, 703-711.
- HOWARD, M. B., EKBORG, N. A., TAYLOR, L. E., HUTCHESON, S. W. & WEINER, R. M. 2004. Identification and analysis of polyserine linker domains in prokaryotic proteins with emphasis on the marine bacterium *Microbulbifer degradans*. *Protein science : a Publication of the Protein Society*, 13, 1422-1425.
- HU, Y., ZHANG, G., LI, A., CHEN, J. & MA, L. 2008. Cloning and enzymatic characterization of a xylanase gene from a soil-derived metagenomic library with an efficient approach. *Applied Microbiology and Biotechnology*, 80, 823-830.
- ILMBERGER, N. & STREIT, W. R. 2010. Screening for cellulase-encoding clones in metagenomic libraries. *Methods in Molecular Biology*, 668, 177-188.
- JACOBSEN, C. S. & RASMUSSEN, O. F. 1992. Development and application of a new method to extract bacterial DNA from soil based on separation of bacteria from soil with cation-exchange resin. *Applied and Environmental Microbiology*, 58, 2458-2462.
- JACOBSEN, S. E. & WYMAN, C. E. 2002. Xylose monomer and oligomer yields for uncatalyzed hydrolysis of sugarcane bagasse hemicellulose at varying solids concentration. *Industrial & Engineering Chemistry Research*, 41, 1454-1461.
- JAVED, M. R., RASHID, M. H., NADEEM, H., RIAZ, M. & PERVEEN, R. 2009. Catalytic and thermodynamic characterization of endoglucanase (CMCase) from *Aspergillus oryzae* cmc-1. *Applied Biochemistry and Biotechnology*, 157, 483-497.
- JEFFERSON, M. 2006. Sustainable energy development: performance and prospects. *Renewable Energy*, 31, 571-582.
- JEONG, Y. S., NA, H. B., KIM, S. K., KIM, Y. H., KWON, E. J., KIM, J., YUN, H. D., LEE, J. K. & KIM, H. 2012. Characterization of xyn10J, a novel family 10 xylanase from a compost metagenomic library. *Applied Biochemistry and Biotechnology*, 166, 1328-1339.
- JIN, Y. S., NI, H., LAPLAZA, J. M. & JEFFRIES, T. W. 2003. Optimal growth and ethanol production from xylose by recombinant *Saccharomyces cerevisiae* require moderate D-xylose kinase activity. *Applied and Environmental Microbiology*, 69, 495-503.
- JOÃO, I. S. P., G.S.; GALINA, S.C.R. 2011. A posição do Brasil na corrida pelo etanol celulósico: mensuração por indicadores C&T e programas de P&D. *Revista Brasileira de Inovação*, 1, 105-136.

KAAR, W. E., GUTIERREZ, C. V. & KINOSHITA, C. M. 1998. Steam explosion of sugarcane bagasse as a pretreatment for conversion to ethanol. *Biomass & Bioenergy*, 14, 277-287.

KALYUZHAYAYA, M. G., LAPIDUS, A., IVANOVA, N., COPELAND, A. C., MCHARDY, A. C., SZETO, E., SALAMOV, A., GRIGORIEV, I. V., SUCIU, D., LEVINE, S. R., MARKOWITZ, V. M., RIGOUTSOS, I., TRINGE, S. G., BRUCE, D. C., RICHARDSON, P. M., LIDSTROM, M. E. & CHISTOSERDOVA, L. 2008. High-resolution metagenomics targets specific functional types in complex microbial communities. *Nature Biotechnology*, 26, 1029-1034.

KARLSTROM, A. R. & LEVINE, R. L. 1991. Copper inhibits the protease from human immunodeficiency virus 1 by both cysteine-dependent and cysteine-independent mechanisms. *Proceedings of the National Academy of Sciences of the United States of America*, 88, 5552-5556.

KAUR, A., MAHAJAN, R., SINGH, A., GARG, G. & SHARMA, J. 2011. A novel and cost effective methodology for qualitative screening of alkalo-thermophilic cellulase free xylano-pectinolytic microorganisms using agricultural wastes. *World Journal of Microbiology and Biotechnology*, 27, 459-463.

KIM, Y. J., CHOI, G. S., KIM, S. B., YOON, G. S., KIM, Y. S. & RYU, Y. W. 2006. Screening and characterization of a novel esterase from a metagenomic library. *Protein Expression and Purification*, 45, 315-323.

KIM, Y. S., JUNG, H. C. & PAN, J. G. 2000. Bacterial cell surface display of an enzyme library for selective screening of improved cellulase variants. *Applied and Environmental Microbiology*, 66, 788-793.

KNIETSCH, A., WASCHKOWITZ, T., BOWIEN, S., HENNE, A. & DANIEL, R. 2003. Construction and screening of metagenomic libraries derived from enrichment cultures: generation of a gene bank for genes conferring alcohol oxidoreductase activity on *Escherichia coli*. *Applied and Environmental Microbiology*, 69, 1408-1416.

KNOB, A., BEITEL, S. M., FORTKAMP, D., TERRASAN, C. R. & DE ALMEIDA, A. F. 2013. Production, purification, and characterization of a major *Penicillium glabrum* xylanase using Brewer's spent grain as substrate. *BioMed Research International*, 2013, 728-735.

KOSEKI, T., TAKAHASHI, K., HANDA, T., YAMANE, Y., FUSHINOBU, S. & HASHIZUME, K. 2006. N-linked oligosaccharides of *Aspergillus awamori* feruloyl esterase are important for thermostability and catalysis. *Bioscience, Biotechnology and Biochemistry*, 70, 2476-2480.

KRAUSE, D. O., DENMAN, S. E., MACKIE, R. I., MORRISON, M., RAE, A. L., ATTWOOD, G. T. & MCSWEENEY, C. S. 2003. Opportunities to improve fiber

degradation in the rumen: microbiology, ecology, and genomics. *FEMS Microbiology Reviews*, 27, 663-693.

KUNIN, V., COPELAND, A., LAPIDUS, A., MAVROMATIS, K. & HUGENHOLTZ, P. 2008a. A bioinformatician's guide to metagenomics. *Microbiology and Molecular Biology Reviews* : *MMBR*, 72, 557-578.

KUNIN, V., RAES, J., HARRIS, J. K., SPEAR, J. R., WALKER, J. J., IVANOVA, N., VON MERING, C., BEBOUT, B. M., PACE, N. R., BORK, P. & HUGENHOLTZ, P. 2008b. Millimeter-scale genetic gradients and community-level molecular convergence in a hypersaline microbial mat. *Molecular Systems Biology*, 4, 198.

KWON, E. J., JEONG, Y. S., KIM, Y. H., KIM, S. K., NA, H. B., KIM, J., YUN, H. D. & KIM, H. 2010. Construction of a metagenomic library from compost and screening of cellulase- and xylanase-positive clones. *Journal of the Korean Society for Applied Biological Chemistry*, 53, 702-708.

LADBURY, J. E. & DOYLE, M. L. 2004. *Biocalorimetry 2 - Applications of Calorimetry in the Biological Sciences*, John Wiley & Sons Ltd.

LAIRSON, L. L., HENRISSAT, B., DAVIES, G. J. & WITHERS, S. G. 2008. Glycosyltransferases: structures, functions, and mechanisms. *Annual Review of Biochemistry*, 77, 521-555.

LAMMERTS VAN BUEREN, A. & BORASTON, A. B. 2004. Binding sub-site dissection of a carbohydrate-binding module reveals the contribution of entropy to oligosaccharide recognition at "non-primary" binding subsites. *Journal of Molecular Biology*, 340, 869-879.

LANGLEY, K. E., VILLAREJO, M. R., FOWLER, A. V., ZAMENHOF, P. J. & ZABIN, I. 1975. Molecular basis of beta-galactosidase alpha-complementation. *Proceedings of the National Academy of Sciences of the United States of America*, 72, 1254-1257.

LAUREANO-PEREZ, L., TEYMOURI, F., ALIZADEH, H. & DALE, B. 2005. Understanding factors that limit enzymatic hydrolysis of biomass: characterization of pretreated corn stover. *Applied Biochemical Biotechnology*, 121-124, 1081-1099.

LEE, Y., CHOI, D. & KENDE, H. 2001. Expansins: ever-expanding numbers and functions. *Current Opinion in Plant Biology*, 4, 527-532.

LEVASSEUR, A., DRULA, E., LOMBARD, V., COUTINHO, P. M. & HENRISSAT, B. 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnology for Biofuels*, 6, 41.

LEVY, I., SHANI, Z. & SHOSEYOV, O. 2002. Modification of polysaccharides and plant cell wall by endo-1,4-beta-glucanase and cellulose-binding domains. *Biomolecular Engineering*, 19, 17-30.

- LI, A., CHU, Y., WANG, X., REN, L., YU, J., LIU, X., YAN, J., ZHANG, L., WU, S. & LI, S. 2013. A pyrosequencing-based metagenomic study of methane-producing microbial community in solid-state biogas reactor. *Biotechnology for Biofuels*, 6, 3.
- LI, L. L., MCCORKLE, S. R., MONCHY, S., TAGHAVI, S. & VAN DER LELIE, D. 2009. Bioprospecting metagenomes: glycosyl hydrolases for converting biomass. *Biotechnology for Biofuels*, 2, 10.
- LIU, B., ZHANG, N., ZHAO, C., LIN, B., XIE, L. & HUANG, Y. 2012. Characterization of a recombinant thermostable xylanase from hot spring thermophilic *Geobacillus* sp. TC-W7. *Journal of Microbiology and Biotechnology*, 22, 1388-1394.
- LIU, Y., YOSHIDA, M., KURAKATA, Y., MIYAZAKI, T., IGARASHI, K., SAMEJIMA, M., FUKUDA, K., NISHIKAWA, A. & TONOZUKA, T. 2010. Crystal structure of a glycoside hydrolase family 6 enzyme, CcCel6C, a cellulase constitutively produced by *Coprinopsis cinerea*. *The FEBS Journal*, 277, 1532-1542.
- LOPRETE, D. M. 2002. Isolation and characterization of an endo-(1,4)- $\beta$ -glucanase secreted by *Achlya ambisexualis*. *Mycologia*, 94, 903-911.
- LOVINS, A. B. 2004. *Winning the oil endgame: innovation for profits, jobs and security*, Snowmass, Colo., Rocky Mountain Institute.
- LYND, L. R., WEIMER, P. J., VAN ZYL, W. H. & PRETORIUS, I. S. 2002. Microbial cellulose utilization: fundamentals and Biotechnology. *Microbiology and Molecular Biology Reviews: MMBR*, 66, 506-577.
- MACEDO, I., 2005. *A Energia da Cana-de-Açúcar* (Sugarcane's energy: twelve studies on Brazilian sugarcane agribusiness and its sustainability). UNICA. São Paulo, Brasil. [Online].
- MANSFIELD, S. D., SADDLER, J. N. & GUBITZ, G. M. 1998. Characterization of endoglucanases from the brown rot fungi *Gloeophyllum sepiarium* and *Gloeophyllum trabeum*. *Enzyme and Microbial Technology*, 23, 133-140.
- MARCUS, S. E., VERHERTBRUGGEN, Y., HERVE, C., ORDAZ-ORTIZ, J. J., FARKAS, V., PEDERSEN, H. L., WILLATS, W. G. & KNOX, J. P. 2008. Pectic homogalacturonan masks abundant sets of xyloglucan epitopes in plant cell walls. *BMC Plant Biology*, 8, 60.
- MARTINELLI, L. A. & FILOSO, S. 2008. Expansion of sugarcane ethanol production in Brazil: environmental and social challenges. *Ecological applications : a Publication of the Ecological Society of America*, 18, 885-898.
- MATTEOTTI, C., THONART, P., FRANCIS, F., HAUBRUGE, E., DESTAIN, J., BRASSEUR, C., BAUWENS, J., DE PAUW, E., PORTETELLE, D. & VANDENBOL, M. 2011. New glucosidase activities identified by functional screening of a genomic DNA

library from the gut microbiota of the termite *Reticulitermes santonensis*. *Microbiological Research*, 166, 629-642.

MCILVAINE, T. C. 1921. A buffer solution for colorimetric comparisons. *Journal Biology Chemistry*, 49, 185-186.

MEANS, G. E. & FEENEY, R. E. 1971. Chemical modification of proteins, São Francisco, Holden-Day Inc, EUA.

MILLER, G. L. 1959. Use of DNS reagent for determination of reducing sugars. *Analytical Chemistry*, 31, 426-428.

MO, X. C., CHEN, C. L., PANG, H., FENG, Y. & FENG, J. X. 2010. Identification and characterization of a novel xylanase derived from a rice straw degrading enrichment culture. *Applied Microbiology and Biotechnology*, 87, 2137-2146.

MOSIER, N., WYMAN, C., DALE, B., ELANDER, R., LEE, Y. Y., HOLTZAPPLE, M. & LADISCH, M. 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology*, 96, 673-686.

NAIK, S. N., GOUD, V. V., ROUT, P. K. & DALAI, A. K. 2010. Production of first and second generation biofuels: a comprehensive review. *Renewable and Sustainable Energy Reviews*, 14, 578-597.

NGUYEN, T. T. T. & LE, V. V. M. 2013. Effects of ultrasound on cellulolytic activity of cellulase complex. *International Food Research Journal*, 20, 557-563.

OKAZAKI, F., TAMARU, Y., HASHIKAWA, S., LI, Y. T. & ARAKI, T. 2002. Novel carbohydrate-binding module of beta-1,3-xylanase from a marine bacterium, *Alcaligenes* sp. strain XY-234. *Journal of Bacteriology*, 184, 2399-2403.

OSBURNE, M. S., GROSSMANN, T. H., AUGUST, P. R. & MACNEIL, I. A. 2000. Tapping into microbial diversity for natural products drug discovery. *ASM News*, 66, 411-417.

ÖZÇİMEN, D. & KARAOSMANOĞLU, F. 2004. Production and characterization of bio-oil and biochar from rapeseed cake. *Renewable Energy*, 29, 779-787.

PACE, N. R., STAHL, D. A., LANE, D. J. & OLSEN, G. J. 1985. Analyzing natural microbial populations by rRNA sequences. *ASM News*, 51, 4-12.

PAIXÃO, D. A. A., DIMITROV, M. R., PEREIRA, R. M., ACCORSINI, F. R., VIDOTTI, M. B. & LEMOS, E. G. M. 2010. Molecular analysis of the bacterial diversity in a specialized consortium for diesel oil degradation. *Revista Brasileira de Ciência do Solo*, 34, 773-781.

PANCHAL, H. & INGLE, S. 2011. Isolation and characterization of endophytes from the root of medicinal plant *Chlorophytum borivilianum* (Safed musli). *Journal of Advances in Developmental Research*, 2, 205-209.

PANDEY, A., SOCCOL, C. R., NIGAM, P. & SOCCOL, V. T. 2000. Biotechnological potential of agro-industrial residues. Part I. Sugarcane bagasse. *Bioresource Technology*, 74, 69-80.

PANG, H., ZHANG, P., DUAN, C. J., MO, X. C., TANG, J. L. & FENG, J. X. 2009. Identification of cellulase genes from the metagenomes of compost soils and functional characterization of one novel endoglucanase. *Current Microbiology*, 58, 404-408.

PARK, I. H., CHANG, J., LEE, Y. S., FANG, S. J. & CHOI, Y. L. 2012. Gene cloning of endoglucanase Cel5A from cellulose-degrading *Paenibacillus xylanilyticus* KJ-03 and purification and characterization of the recombinant enzyme. *The Protein Journal*, 31, 238-245.

PEI, D., KANG, T. & QI, H. 2000. Cysteine array matrix metalloproteinase (CA-MMP)/MMP-23 is a type II transmembrane matrix metalloproteinase regulated by a single cleavage for both secretion and activation. *The Journal of Biological Chemistry*, 275, 33988-33997.

PELL, G., WILLIAMSON, M. P., WALTERS, C., DU, H., GILBERT, H. J. & BOLAM, D. N. 2003. Importance of hydrophobic and polar residues in ligand binding in the family 15 carbohydrate-binding module from *Cellvibrio japonicus* Xyn10C. *Biochemistry*, 42, 9316-9323.

PEREZ, J., MUNOZ-DORADO, J., DE LA RUBIA, T. & MARTINEZ, J. 2002. Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. *International Microbiology : the Official Journal of the Spanish Society for Microbiology*, 5, 53-63.

POOLE, D. M., HAZLEWOOD, G. P., HUSKISSON, N. S., VIRDEN, R. & GILBERT, H. J. 1993. The role of conserved tryptophan residues in the interaction of a bacterial cellulose binding domain with its ligand. *FEMS Microbiology Letters*, 106, 77-83.

PRIYADARSINI, R. I., BHUVANESWARI, V. & KUMAR, K. S. 2011. Isolation, identification and phylogenetic analysis of white rot fungus and heterologous expression of gene encoding laccase. *Journal of Applied Sciences in Environmental Sanitation*, 6, 69-83.

QUAISER, A., OCHSENREITER, T., KLENK, H. P., KLETZIN, A., TREUSCH, A. H., MEURER, G., ECK, J., SENSEN, C. W. & SCHLEPER, C. 2002. First insight into the genome of an uncultivated crenarchaeote from soil. *Environmental Microbiology*, 4, 603-611.

QUIROZ-CASTAÑEDA, R. E. & FOLCH-MALLOL, J. L. 2011. Plant cell wall degrading and remodeling proteins: current perspectives. *Biotecnología Aplicada* 28, 205-215.

RADOMSKI, C. C. A., SEOW, K. T., WARREN, R. A. J. & YAP, W. H. 1998. Method for isolation of xylanase gene sequences from soil DNA, compositions useful in such method and compositions obtained thereby. United States Patent (Patent No. US 5,849,491).

RAGAUSKAS, A. J., WILLIAMS, C. K., DAVISON, B. H., BRITOVSEK, G., CAIRNEY, J., ECKERT, C. A., FREDERICK, W. J., JR., HALLETT, J. P., LEAK, D. J., LIOTTA, C. L., MIELENZ, J. R., MURPHY, R., TEMPLER, R. & TSCHAPLINSKI, T. 2006. The path forward for biofuels and biomaterials. *Science*, 311, 484-489.

RAGHOTHAMA, S., SIMPSON, P. J., SZABO, L., NAGY, T., GILBERT, H. J. & WILLIAMSON, M. P. 2000. Solution structure of the CBM10 cellulose binding module from *Pseudomonas xylanase A*. *Biochemistry*, 39, 978-984.

RAMIREZ, L., ARRIZON, J., SANDOVAL, G., CARDADOR, A., BELLO-MENDOZA, R., LAPPE, P. & MATEOS-DIAZ, J. C. 2008. A new microplate screening method for the simultaneous activity quantification of feruloyl esterases, tannases, and chlorogenate esterases. *Applied Biochemistry and Biotechnology*, 151, 711-723.

REES, H. C., GRANT, S., JONES, B., GRANT, W. D. & HEAPHY, S. 2003. Detecting cellulase and esterase enzyme activities encoded by novel genes present in environmental DNA libraries. *Extremophiles*, 7, 415-421.

REVERBEL-LEROY, C., PAGES, S., BELAICH, A., BELAICH, J. P. & TARDIF, C. 1997. The processive endocellulase CelF, a major component of the *Clostridium cellulolyticum* cellulosome: purification and characterization of the recombinant form. *Journal of Bacteriology*, 179, 46-52.

RHO, M., TANG, H. & YE, Y. 2010. FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Research*, 38, e191.

RICHARDSON, T. H., TAN, X., FREY, G., CALLEN, W., CABELL, M., LAM, D., MACOMBER, J., SHORT, J. M., ROBERTSON, D. E. & MILLER, C. 2002. A novel, high performance enzyme for starch liquefaction. Discovery and optimization of a low pH, thermostable alpha-amylase. *The Journal of Biological Chemistry*, 277, 26501-26507.

RIESENFELD, C. S., SCHLOSS, P. D. & HANDELSMAN, J. 2004. Metagenomics: genomic analysis of microbial communities. *Annual Review of Genetics*, 38, 525-552.

ROHWER, F., SEGURITAN, V., CHOI, D. H., SEGALL, A. M. & AZAM, F. 2001. Production of shotgun libraries using random amplification. *BioTechniques*, 31, 108-12, 114-6, 118.

- ROSEN, M. J., CALLAHAN, B. J., FISHER, D. S. & HOLMES, S. P. 2012. Denoising PCR-amplified metagenome data. *BMC Bioinformatics*, 13: 283.
- RUBIN, E. M. 2008. Genomics of cellulosic biofuels. *Nature*, 454, 841-845.
- RUBINI, M. R., DILLON, A. J., KYAW, C. M., FARIA, F. P., POCAS-FONSECA, M. J. & SILVA-PEREIRA, I. 2010. Cloning, characterization and heterologous expression of the first *Penicillium echinulatum* cellulase gene. *Journal of Applied Microbiology*, 108, 1187-1198.
- SADHU, S., SAHA, P., SEN, S. K., MAYILRAJ, S. & MAITI, T. K. 2013. Production, purification and characterization of a novel thermotolerant endoglucanase (CMCase) from strain isolated from cow dung. *SpringerPlus*, 2:10.
- SANCHEZ, C. 2009. Lignocellulosic residues: biodegradation and bioconversion by fungi. *Biotechnology Advances*, 27, 185-194.
- SANJUAN, R., ANZALDO, J., VARGAS, J., TURRADO, J. & PATT, R. 2001. Morphological and chemical composition of pith and fibres from Mexican sugarcane bagasse. *Holz als Roh-und Werkstoff*, 59, 447-450.
- SCHMEISSER, C., STEELE, H. & STREIT, W. R. 2007. Metagenomics, biotechnology with non-culturable microbes. *Applied Microbiology and Biotechnology*, 75, 955-962.
- SCHWARZ, E. M. & BENZER, S. 1997. Calx, a Na-Ca exchanger gene of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 10249-10254.
- SEBAT, J. L., COLWELL, F. S. & CRAWFORD, R. L. 2003. Metagenomic profiling: microarray analysis of an environmental genomic library. *Applied and Environmental Microbiology*, 69, 4927-4934.
- SHI, H., ZHANG, Y., LI, X., HUANG, Y., WANG, L., WANG, Y., DING, H. & WANG, F. 2013. A novel highly thermostable xylanase stimulated by  $\text{Ca}^{2+}$  from *Thermotoga thermarum*: cloning, expression and characterization. *Biotechnology for Biofuels*, 6, 26.
- SHOSEYOV, O., SHANI, Z. & LEVY, I. 2006. Carbohydrate binding modules: biochemical properties and novel applications. *Microbiology and Molecular Biology Reviews: MMBR*, 70, 283-295.
- SHUAI, L. & PAN, X. 2012. Hydrolysis of cellulose by cellulase-mimetic solid catalyst. *Energy & Environmental Science*, 5, 6889-6894.
- SIMON, C. & DANIEL, R. 2011. Metagenomic analyses: past and future trends. *Applied and Environmental Microbiology*, 77, 1153-1161.

SIMS, R. E., MABEE, W., SADDLER, J. N. & TAYLOR, M. 2010. An overview of second generation biofuel technologies. *Bioresource Technology*, 101, 1570-1580.

SMITH, R. J., JEFFRIES, T. C., ROUDNEW, B., FITCH, A. J., SEYMOUR, J. R., DELPIN, M. W., NEWTON, K., BROWN, M. H. & MITCHELL, J. G. 2012. Metagenomic comparison of microbial communities inhabiting confined and unconfined aquifer ecosystems. *Environmental Microbiology*, 14, 240-253.

SOCCOL, C. R., VANDENBERGHE, L. P., MEDEIROS, A. B., KARP, S. G., BUCKERIDGE, M., RAMOS, L. P., PITARELO, A. P., FERREIRA-LEITAO, V., GOTTSCHALK, L. M., FERRARA, M. A., DA SILVA BON, E. P., DE MORAES, L. M., ARAUJO JDE, A. & TORRES, F. A. 2010. Bioethanol from lignocelluloses: status and perspectives in Brazil. *Bioresource Technology*, 101, 4820-4825.

SOLBAK, A. I., RICHARDSON, T. H., MCCANN, R. T., KLINE, K. A., BARTNEK, F., TOMLINSON, G., TAN, X., PARRA-GESSERT, L., FREY, G. J., PODAR, M., LUGINBUHL, P., GRAY, K. A., MATHUR, E. J., ROBERTSON, D. E., BURK, M. J., HAZLEWOOD, G. P., SHORT, J. M. & KEROVUO, J. 2005. Discovery of pectin-degrading enzymes and directed evolution of a novel pectate lyase for processing cotton fabric. *The Journal of Biological Chemistry*, 280, 9431-9438.

SONAN, G. K., RECEVEUR-BRECHOT, V., DUEZ, C., AGHAJARI, N., CZJZEK, M., HASER, R. & GERDAY, C. 2007. The linker region plays a key role in the adaptation to cold of the cellulase from an Antarctic bacterium. *The Biochemical Journal*, 407, 293-302.

STALBRAND, H., SALOHEIMO, A., VEHEMAANPERA, J., HENRISSAT, B. & PENTTILA, M. 1995. Cloning and expression in *Saccharomyces cerevisiae* of a *Trichoderma reesei* beta-mannanase gene containing a cellulose binding domain. *Applied and Environmental Microbiology*, 61, 1090-1097.

STALEY, J. T. & KONOPKA, A. 1985. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology*, 39, 321-346.

STEFFAN, R. J., GOKSOYR, J., BEJ, A. K. & ATLAS, R. M. 1988. Recovery of DNA from soils and sediments. *Applied and Environmental Microbiology*, 54, 2908-2915.

STEIN, J. L., MARSH, T. L., WU, K. Y., SHIZUYA, H. & DELONG, E. F. 1996. Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. *Journal of Bacteriology*, 178, 591-599.

STROH, J. G., LOULAKIS, P., LANZETTI, A. J. & XIE, J. 2005. LC-mass spectrometry analysis of N- and C-terminal boundary sequences of polypeptide fragments by limited proteolysis. *Journal of the American Society for Mass Spectrometry*, 16, 38-45.

STUDIER, F. W. 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expression and Purification*, 41, 207-234.

SZABO, L., JAMAL, S., XIE, H., CHARNOCK, S. J., BOLAM, D. N., GILBERT, H. J. & DAVIES, G. J. 2001. Structure of a family 15 carbohydrate-binding module in complex with xylopentaose. Evidence that xylan binds in an approximate 3-fold helical conformation. *The Journal of Biological Chemistry*, 276, 49061-49065.

TASSE, L., BERCOVICI, J., PIZZUT-SERIN, S., ROBE, P., TAP, J., KLOPP, C., CANTAREL, B. L., COUTINHO, P. M., HENRISSAT, B., LECLERC, M., DORE, J., MONSAN, P., REMAUD-SIMEON, M. & POTOCKI-VERONESE, G. 2010. Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Research*, 20, 1605-1612.

TAYLOR, L. E., 2ND, HENRISSAT, B., COUTINHO, P. M., EKBORG, N. A., HUTCHESON, S. W. & WEINER, R. M. 2006. Complete cellulase system in the marine bacterium *Saccharophagus degradans* strain 2-40T. *Journal of Bacteriology*, 188, 3849-3861.

TENGERDY, R. P. & SZAKACS, G. 2003. Bioconversion of lignocellulose in solid substrate fermentation. *Biochemical Engineering Journal*, 13, 169-179.

TOMME, P., KWAN, E., GILKES, N. R., KILBURN, D. G. & WARREN, R. A. 1996. Characterization of CenC, an enzyme from *Cellulomonas fimi* with both endo- and exoglucanase activities. *Journal of Bacteriology*, 178, 4216-4223.

TRINGE, S. G. & RUBIN, E. M. 2005. Metagenomics: DNA sequencing of environmental samples. *Nature Reviews Genetics*, 6, 805-814.

TURNBAUGH, P. J., LEY, R. E., MAHOWALD, M. A., MAGRINI, V., MARDIS, E. R. & GORDON, J. I. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*, 444, 1027-1031.

UCHIYAMA, T. & WATANABE, K. 2006. Improved inverse PCR scheme for metagenome walking. *BioTechniques*, 41, 183-188.

UNGCHAITHAM, S., THANA, C., TALAWANICH, Y. & PINPHANICHAKARN, P. 2001.  $\beta$ -xylosidase from *Streptomyces* sp. CH7 and its gene cloning and expression in *Escherichia coli*. *Journal of Scientific Research*, 26, 107-116.

VARDAKOU, M., FLINT, J., CHRISTAKOPOULOS, P., LEWIS, R. J., GILBERT, H. J. & MURRAY, J. W. 2005. A family 10 *Thermoascus aurantiacus* xylanase utilizes arabinose decorations of xylan as significant substrate specificity determinants. *Journal of Molecular Biology*, 352, 1060-1067.

- VÁZQUEZ, M. J., ALONSO, J. L., DOMÍNGUEZ, H. & PARAJÓ, J. C. 2000. Xylooligosaccharides: manufacture and applications. *Trends in Food Science & Technology*, 11, 387-393.
- VENTER, J. C., REMINGTON, K., HEIDELBERG, J. F., HALPERN, A. L., RUSCH, D., EISEN, J. A., WU, D., PAULSEN, I., NELSON, K. E., NELSON, W., FOUTS, D. E., LEVY, S., KNAP, A. H., LOMAS, M. W., NEALSON, K., WHITE, O., PETERSON, J., HOFFMAN, J., PARSONS, R., BADEN-TILLSON, H., PFANNKOCHE, C., ROGERS, Y. H. & SMITH, H. O. 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, 304, 66-74.
- VERMA, D., KAWARABAYASI, Y., MIYAZAKI, K. & SATYANARAYANA, T. 2013. Cloning, expression and characteristics of a novel alkalistable and thermostable xylanase encoding gene (Mxyl) retrieved from compost-soil metagenome. *PloS ONE*, 8, e52459.
- VOGET, S., LEGGEWIE, C., UESBECK, A., RAASCH, C., JAEGER, K. E. & STREIT, W. R. 2003. Prospecting for novel biocatalysts in a soil metagenome. *Applied and Environmental Microbiology*, 69, 6235-6242.
- VOGET, S., STEELE, H. L. & STREIT, W. R. 2006. Characterization of a metagenome-derived halotolerant cellulose. *Journal of Biotechnology*, 126, 26-36.
- WANG, G., WANG, Y., YANG, P., LUO, H., HUANG, H., SHI, P., MENG, K. & YAO, B. 2010. Molecular detection and diversity of xylanase genes in alpine tundra soil. *Applied Microbiology and Biotechnology*, 87, 1383-1393.
- WANG, W., ARCHBOLD, T., KIMBER, M. S., LI, J., LAM, J. S. & FAN, M. Z. 2012. The porcine gut microbial metagenomic library for mining novel cellulases established from growing pigs fed cellulose-supplemented high-fat diets. *Journal of Animal Science*, 4, 400-402.
- WARD, J. J., SODHI, J. S., MCGUFFIN, L. J., BUXTON, B. F. & JONES, D. T. 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *Journal of Molecular Biology*, 337, 635-645.
- WARNECKE, F., LUGINBUHL, P., IVANOVA, N., GHASSEMIAN, M., RICHARDSON, T. H., STEGE, J. T., CAYOUILLE, M., MCHARDY, A. C., DJORDJEVIC, G., ABOUSHADI, N., SOREK, R., TRINGE, S. G., PODAR, M., MARTIN, H. G., KUNIN, V., DALEVI, D., MADEJSKA, J., KIRTON, E., PLATT, D., SZETO, E., SALAMOV, A., BARRY, K., MIKHAILOVA, N., KYRPIDES, N. C., MATSON, E. G., OTTESEN, E. A., ZHANG, X., HERNANDEZ, M., MURILLO, C., ACOSTA, L. G., RIGOUTSOS, I., TAMAYO, G., GREEN, B. D., CHANG, C., RUBIN, E. M., MATHUR, E. J., ROBERTSON, D. E., HUGENHOLTZ, P. & LEADBETTER, J. R. 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*, 450, 560-565.

- WHITMAN, W. B., COLEMAN, D. C. & WIEBE, W. J. 1998. Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 6578-6583.
- WINKELHAUSEN, E. & KUZMANOVA, S. 1998. Microbial conversion of D-xylose to xylitol. *Journal of Fermentation and Bioengineering*, 86, 1-14.
- XIA, Y., JU, F., FANG, H. H. & ZHANG, T. 2013. Mining of novel thermo-stable cellulolytic genes from a thermophilic cellulose-degrading consortium by metagenomics. *PLoS ONE*, 8, e53779.
- XIE, G., BRUCE, D. C., CHALLACOMBE, J. F., CHERTKOV, O., DETTER, J. C., GILNA, P., HAN, C. S., LUCAS, S., MISRA, M., MYERS, G. L., RICHARDSON, P., TAPIA, R., THAYER, N., THOMPSON, L. S., BRETTIN, T. S., HENRISSAT, B., WILSON, D. B. & MCBRIDE, M. J. 2007. Genome sequence of the cellulolytic gliding bacterium *Cytophaga hutchinsonii*. *Applied and Environmental Microbiology*, 73, 3536-3546.
- YOOSEPH, S., SUTTON, G., RUSCH, D. B., HALPERN, A. L., WILLIAMSON, S. J., REMINGTON, K., EISEN, J. A., HEIDELBERG, K. B., MANNING, G., LI, W., JAROSZEWSKI, L., CIEPLAK, P., MILLER, C. S., LI, H., MASHIYAMA, S. T., JOACHIMIAK, M. P., VAN BELLE, C., CHANDONIA, J. M., SOERGEL, D. A., ZHAI, Y., NATARAJAN, K., LEE, S., RAPHAEL, B. J., BAFNA, V., FRIEDMAN, R., BRENNER, S. E., GODZIK, A., EISENBERG, D., DIXON, J. E., TAYLOR, S. S., STRAUSBERG, R. L., FRAZIER, M. & VENTER, J. C. 2007. The Sorcerer II Global Ocean Sampling expedition: expanding the universe of protein families. *PLoS Biology*, 5, e16.
- YU, K. & ZHANG, T. 2012. Metagenomic and metatranscriptomic analysis of microbial community structure and gene expression of activated sludge. *PLoS ONE*, 7, e38183.
- YUN, J., KANG, S., PARK, S., YOON, H., KIM, M. J., HEU, S. & RYU, S. 2004. Characterization of a novel amylolytic enzyme encoded by a gene from a soil-derived metagenomic library. *Applied and Environmental Microbiology*, 70, 7229-7235.
- ZHANG, G. M., HUANG, J., HUANG, G. R., MA, L. X. & ZHANG, X. E. 2007. Molecular cloning and heterologous expression of a new xylanase gene from *Plectosphaerella cucumerina*. *Applied Microbiology and Biotechnology*, 74, 339-346.
- ZHANG, W. & GENG, A. 2012. Improved ethanol production by a xylose-fermenting recombinant yeast strain constructed through a modified genome shuffling method. *Biotechnology for Biofuels*, 5, 46.
- ZHOU, S. & INGRAM, L. O. 2000. Synergistic hydrolysis of carboxymethyl cellulose and acid-swollen cellulose by two endoglucanases (CelZ and CelY) from *Erwinia chrysanthemi*. *Journal of Bacteriology*, 182, 5676-5682.

ZHOU, X., REN, L., MENG, Q., LI, Y., YU, Y. & YU, J. 2010. The next-generation sequencing technology and application. *Protein & Cell*, 1, 520-536.

ZHU, W., LOMSADZE, A. & BORODOVSKY, M. 2010. Ab initio gene identification in metagenomic sequences. *Nucleic Acids Research*, 38, e132.



## Development and Biotechnological Application of a Novel Endoxylanase Family GH10 Identified from Sugarcane Soil Metagenome

Thabata M. Alvarez<sup>1,2,3</sup>, Rosana Goldbeck<sup>1,3</sup>, Camila Ramos dos Santos<sup>3,9</sup>, Douglas A. A. Paixão<sup>1</sup>, Thiago A. Gonçalves<sup>1,2</sup>, João Paulo L. Franco Cairo<sup>1,2</sup>, Rodrigo Ferreira Almeida<sup>1</sup>, Isabela de Oliveira Pereira<sup>1</sup>, George Jackson<sup>1</sup>, Junio Cota<sup>1</sup>, Fernanda Büchli<sup>1,2</sup>, Ana Paula Citadini<sup>1</sup>, Roberto Ruller<sup>1</sup>, Carla Cristina Polo<sup>3</sup>, Mario de Oliveira Neto<sup>4</sup>, Mário T. Murakami<sup>3\*</sup>, Fabio M. Squina<sup>1\*</sup>

**1** Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, São Paulo, Brasil, **2** Departamento de Bioquímica, Instituto de Biologia (IB), Universidade Estadual de Campinas (UNICAMP), Campinas, São Paulo, Brasil, **3** Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, São Paulo, Brasil, **4** Departamento de Física e Biofísica, Instituto de Biociências, Universidade Estadual Paulista (UNESP), Botucatu, São Paulo, Brasil

### Abstract

Metagenomics has been widely employed for discovery of new enzymes and pathways to conversion of lignocellulosic biomass to fuels and chemicals. In this context, the present study reports the isolation, recombinant expression, biochemical and structural characterization of a novel endoxylanase family GH10 (SCXyl) identified from sugarcane soil metagenome. The recombinant SCXyl was highly active against xylan from beechwood and showed optimal enzyme activity at pH 6.0 and 45°C. The crystal structure was solved at 2.75 Å resolution, revealing the classical ( $\beta/\alpha$ )<sub>8</sub>-barrel fold with a conserved active-site pocket and an inherent flexibility of the Trp281-Arg291 loop that can adopt distinct conformational states depending on substrate binding. The capillary electrophoresis analysis of degradation products evidenced that the enzyme displays unusual capacity to degrade small xylooligosaccharides, such as xylotriose, which is consistent to the hydrophobic contacts at the +1 subsite and low-binding energies of subsites that are distant from the site of hydrolysis. The main reaction products from xylan polymers and phosphoric acid-pretreated sugarcane bagasse (PASB) were xylooligosaccharides, but, after a longer incubation time, xylobiose and xylose were also formed. Moreover, the use of SCXyl as pre-treatment step of PASB, prior to the addition of commercial cellulolytic cocktail, significantly enhanced the saccharification process. All these characteristics demonstrate the advantageous application of this enzyme in several biotechnological processes in food and feed industry and also in the enzymatic pretreatment of biomass for feedstock and ethanol production.

**Citation:** Alvarez TM, Goldbeck R, Santos CRd, Paixão DAA, Gonçalves TA, et al. (2013) Development and Biotechnological Application of a Novel Endoxylanase Family GH10 Identified from Sugarcane Soil Metagenome. PLoS ONE 8(7): e70014. doi:10.1371/journal.pone.0070014

**Editor:** Mark R. Liles, Auburn University, United States of America

**Received:** April 22, 2013; **Accepted:** June 13, 2013; **Published:** July 29, 2013

**Copyright:** © 2013 Alvarez et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was financially supported by FAPESP: 2008/58037-9 (FMS), 2010/11469-1 (TMA), 2012/18559-5 (RG), 2011/20977-3 (JPLFC) The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist

\* E-mail: fabio.squina@bioetanol.org.br (FMS); mario.murakami@lnbio.cnpem.br (MTM)

<sup>9</sup> These authors contributed equally to this work.

### Introduction

The conversion of lignocellulosic biomass into renewable fuels has been considered a promising technology to replace fossil fuels and to attend the global need for clean energy. Although the recent worldwide growth in the production of biofuels from plant biomass, several technological bottlenecks still exist and this bioconversion process is not profitable yet [1]. The development of low cost biocatalysts and the improvement of the catalytic efficiency are considered the key to the future of biofuels [1,2,3,4].

Xylanases (E.C. 3.2.1.8) are hemicellulases responsible for breaking down xylan, the major hemicellulosic component of plant cell walls, into short xylooligosaccharides by a general acid–base mechanism involving two glutamic acid residues [5,6]. Typically, these enzymes can be classified into glycoside hydrolase (GH) families 10 and 11 based on amino-acid sequence similarities [7]. Recently, these enzymes have received much attention owing to

their use in degradation of lignocellulosic biomass for biofuels production [8,9].

The development of novel enzymes is primarily dependent on the effectiveness of the screening strategy and the variety of candidate microorganisms present in certain environmental location [10]. In addition, less than 1% of microorganisms from natural environments can be cultivated using traditional culture methods [11]. Thus, researchers have developed strategies to prospect genes through culture-independent methods [1]. Metagenomics is an advanced strategy to seek for novel functional genes and/or biologically active compounds, by means of direct extraction of all microbial genomic DNA from an environmental sample [12,13].

Novel xylanases with designed characteristics for biotechnological application have been identified by metagenomics approaches [14–16]. A novel GH11 was retrieved from a compost-soil

**ANEXO 2 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA  
BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS EM  
NOVEMBRO DE 2010.**

Biochemical and Biophysical Research Communications 403 (2010) 214–219



Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: [www.elsevier.com/locate/ybbrc](http://www.elsevier.com/locate/ybbrc)



**Thermal-induced conformational changes in the product release area drive the enzymatic activity of xylanases 10B: Crystal structure, conformational stability and functional characterization of the xylanase 10B from *Thermotoga petrophila* RKU-1**

Camila Ramos Santos<sup>a</sup>, Andreia Navarro Meza<sup>a</sup>, Zaira Bruna Hoffmam<sup>b</sup>, Junio Cota Silva<sup>b</sup>,  
Thabata Maria Alvarez<sup>b</sup>, Roberto Ruller<sup>b</sup>, Guilherme Menegon Giesel<sup>c</sup>, Hugo Verli<sup>c</sup>, Fabio Marcio Squina<sup>b</sup>,  
Rolf Alexander Prade<sup>d</sup>, Mario Tyago Murakami<sup>a,\*</sup>

<sup>a</sup>Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, SP, Brazil

<sup>b</sup>Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, SP, Brazil

<sup>c</sup>Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

<sup>d</sup>Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, USA

**ARTICLE INFO**

**Article history:**  
Received 27 October 2010  
Available online 9 November 2010

**Keywords:**  
Xylanase  
Glycoside hydrolase family 10  
Crystal structure  
Thermostability  
*Thermotoga petrophila* RKU-1

**ABSTRACT**

Endo-xylanases play a key role in the depolymerization of xylan and recently, they have attracted much attention owing to their potential applications on biofuels and paper industries. In this work, we have investigated the molecular basis for the action mode of xylanases 10B at high temperatures using biochemical, biophysical and crystallographic methods. The crystal structure of xylanase 10B from hyperthermophilic bacterium *Thermotoga petrophila* RKU-1 (TpXyl10B) has been solved in the native state and in complex with xylobiose. The complex crystal structure showed a classical binding mode shared among other xylanases, which encompasses the –1 and –2 subsites. Interestingly, TpXyl10B displayed a temperature-dependent action mode producing xylobiose and xylotriose at 20 °C, and exclusively xylobiose at 90 °C as assessed by capillary zone electrophoresis. Moreover, circular dichroism spectroscopy suggested a coupling effect of temperature-induced structural changes with this particular enzymatic behavior. Molecular dynamics simulations supported the CD analysis suggesting that an open conformational state adopted by the catalytic loop (Trp297-Lys326) provokes significant modifications in the product release area (+1,+2 and +3 subsites), which drives the enzymatic activity to the specific release of xylobiose at high temperatures.

© 2010 Elsevier Inc. All rights reserved.

**1. Introduction**

Xylanases ( $\beta$ -1,4-xylan xylanohydrolase, E.C. 3.2.1.8) are responsible for breaking down xylan, the major hemicellulosic component of plant cell walls, into short xylooligosaccharides by a general acid–base mechanism involving two glutamic acid residues [1,2]. Typically, these enzymes can be classified into glycoside hydrolase (GH) families 10 and 11 based on amino-acid sequence similarities [3]. However, xylanases belonging to families 5, 8

and 43 have also been identified [4,5]. The biotechnological applications of xylanases include food, feed, textile and paper industries [6], and recently they have received much attention owing to their use in degradation of lignocellulosic biomass for biofuels production [7,8]. In this context, enzymes exhibiting high enzymatic efficiency combined with hyper thermal and chemical tolerance are extremely desired. Although countless extremophilic enzymes have been characterized, the molecular basis for their mode of action at high temperatures is still poorly understood. Several structures of family 10 xylanases have been solved, from mesophilic to hyperthermophilic organisms enabling an in-depth comparative study in terms of thermophilicity, thermostability and action mechanism. Thus, in order to gain insights into the structural determinants of their mode of action at high temperatures we have performed a functional and structural characterization of a hyperthermostable xylanase 10B from the bacterium *Thermotoga petrophila* RKU-1 (TpXyl10B). Our studies addressed important

**Abbreviations:** TpXyl10B, xylanase 10B from *Thermotoga petrophila* RKU-1; TmXyl10B, xylanase 10B from *Thermotoga maritima* MSB8; GH, glycoside hydrolase; CD, circular dichroism; CZE, capillary zone electrophoresis; U, unit; MD, molecular dynamics; XTAL, X-ray crystallography.

\* Corresponding author. Address: Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Rua Giuseppe Maximo Scolfaro, 10000, Campinas, 13083-970 SP, Brazil. Fax: +55 19 3512 1004.

E-mail address: [mario.murakami@lnbio.org.br](mailto:mario.murakami@lnbio.org.br) (M.T. Murakami).

0006-291X/\$ - see front matter © 2010 Elsevier Inc. All rights reserved.  
doi:10.1016/j.bbrc.2010.11.010

# ANEXO 3 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS EM FEVEREIRO DE 2011.

Biochemical and Biophysical Research Communications 406 (2011) 590–594



Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: [www.elsevier.com/locate/ybbrc](http://www.elsevier.com/locate/ybbrc)



## Mode of operation and low-resolution structure of a multi-domain and hyperthermophilic endo- $\beta$ -1,3-glucanase from *Thermotoga petrophila*

Junio Cota<sup>a,d</sup>, Thabata M. Alvarez<sup>a</sup>, Ana P. Citadini<sup>a</sup>, Camila Ramos Santos<sup>b</sup>, Mario de Oliveira Neto<sup>c</sup>,  
Renata R. Oliveira<sup>b</sup>, Glucia M. Pastore<sup>d</sup>, Roberto Ruller<sup>a</sup>, Rolf A. Prade<sup>e</sup>, Mario T. Murakami<sup>b</sup>,  
Fabio M. Squina<sup>a,\*</sup>

<sup>a</sup> Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), do Centro Nacional de Pesquisa em Energia e Materiais, Campinas SP, Brazil

<sup>b</sup> e Laboratório Nacional de Biociências (LNBio), do Centro Nacional de Pesquisa em Energia e Materiais, Campinas SP, Brazil

<sup>c</sup> Instituto de Física de São Carlos (IFSC), Universidade de São Paulo, São Carlos SP, Brazil

<sup>d</sup> Faculdade de Engenharia de Alimentos (FEA), Universidade Estadual de Campinas, Campinas SP, Brazil

<sup>e</sup> Department of Microbiology and Molecular Genetics, Oklahoma State University (OSU), Stillwater, OK, USA

### ARTICLE INFO

#### Article history:

Received 16 February 2011

Available online 23 February 2011

#### Keywords:

Endo- $\beta$ -1,3-glucanase  
Glycoside hydrolase family 16  
Hyperthermostable enzyme  
Capillary zone electrophoresis  
Small angle X-ray scattering

### ABSTRACT

1,3- $\beta$ -Glucan depolymerizing enzymes have considerable biotechnological applications including biofuel production, feedstock-chemicals and pharmaceuticals. Here we describe a comprehensive functional characterization and low-resolution structure of a hyperthermophilic laminarinase from *Thermotoga petrophila* (TpLam). We determine TpLam enzymatic mode of operation, which specifically cleaves internal  $\beta$ -1,3-glucosidic bonds. The enzyme most frequently attacks the bond between the 3rd and 4th residue from the non-reducing end, producing glucose, laminaribiose and laminaritriose as major products. Far-UV circular dichroism demonstrates that TpLam is formed mainly by beta structural elements, and the secondary structure is maintained after incubation at 90 °C. The structure resolved by small angle X-ray scattering, reveals a multi-domain structural architecture of a V-shape envelope with a catalytic domain flanked by two carbohydrate-binding modules.

Crown Copyright © 2011 Published by Elsevier Inc. All rights reserved.

### 1. Introduction

Renewable structural polysaccharides are abundant resources in the biosphere and represent a valuable industrial substrate for many industrial processes, such as biofuels production, feedstock-chemicals and pharmaceuticals [1].  $\beta$ -1,3-Glucans are important polymers found mainly in yeast and filamentous fungi and are produced by plants (as callose) in response to tissue damage. This polymer is a major structural storage polysaccharide (laminarin) of the marine brown macro-algae of the genus *Laminaria*.  $\beta$ -1,3-glucan is also produced as insoluble exopolysaccharide by some fungal and bacterial species [2–5].

The  $\beta$ -1,3-glucanases are widely distributed among bacteria, fungi, and plants and are generally classified into two types, exo- $\beta$ -1,3-glucanases (EC 3.2.1.58) and endo- $\beta$ -1,3-glucanases (EC 3.2.1.6 and 3.2.1.39), which hydrolyze terminal or internal

glycoside linkages, respectively [28]. These enzymes are further classified as family members (GH) 16, 17, 55, 64, and 81 of the CaZy system, which is based on structural similarities of the amino acid sequence [28]. However, CaZy GH families display a dispersed range of enzymatic activity, and updating biochemical and biophysical information strengthens the CaZy purpose.

To date, crystal structures from family GH 16 endo- $\beta$ -1,3-glucanases (3.2.1.39) have been determined for *Pyrococcus furiosus* [6], *Rhodothermus marinus* [6], *Streptomyces sioyaensis* [7], and *Nocardiaopsis sp* [8]. The active site region of family GH16 consists of three acidic residues, two glutamic and one aspartic acid in the active site, and two conserved tryptophan residues, which promote substrate binding and transition-state stabilization [9]. The overall fold of these enzymes consists of a classical sandwich-like  $\beta$ -jelly-roll motif formed by the face-to-face packing of two antiparallel sheets containing seven and eight strands with a deep cavity [6–8,10].

In this work is reported a comprehensive functional characterization of an endo- $\beta$ -1,3-glucanase from *Thermotoga petrophila* (TpLam), including a detailed description of the hydrolytic mode of operation, along with structural insights of this hyperthermostable enzyme.

\* Corresponding author. Address: Rua Giuseppe Máximo Scolforo, 10.000 Polo II de Alta Tecnologia, Caixa Postal 6170, Campinas, SP 13083-970, Brazil. Fax: +55 19 35183104.

E-mail address: [fabio.squina@bioetanol.org.br](mailto:fabio.squina@bioetanol.org.br) (F.M. Squina).

## ANEXO 4 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA BIOTECHNOLOGY FOR BIOFUELS EM NOVEMBRO DE 2011.

Franco Cairo et al. *Biotechnology for Biofuels* 2011, 4:50  
<http://www.biotechnologyforbiofuels.com/content/4/1/50>



Biotechnology  
for Biofuels

RESEARCH

Open Access

# Functional characterization and target discovery of glycoside hydrolases from the digestome of the lower termite *Coptotermes gestroi*

João Paulo L Franco Cairo<sup>1,2</sup>, Flávia C Leonardo<sup>2,3</sup>, Thabata M Alvarez<sup>1</sup>, Daniela A Ribeiro<sup>1</sup>, Fernanda Büchli<sup>1</sup>, Ana M Costa-Leonardo<sup>4</sup>, Marcelo F Carazzolle<sup>2,5</sup>, Fernando F Costa<sup>3</sup>, Adriana F Paes Leme<sup>5</sup>, Gonçalo AG Pereira<sup>2,5\*</sup> and Fabio M Squina<sup>1</sup>

### Abstract

**Background:** Lignocellulosic materials have been moved towards the forefront of the biofuel industry as a sustainable resource. However, saccharification and the production of bioproducts derived from plant cell wall biomass are complex and lengthy processes. The understanding of termite gut biology and feeding strategies may improve the current state of biomass conversion technology and bioproduct production.

**Results:** The study herein shows comprehensive functional characterization of crude body extracts from *Coptotermes gestroi* along with global proteomic analysis of the termite's digestome, targeting the identification of glycoside hydrolases and accessory proteins responsible for plant biomass conversion. The crude protein extract from *C. gestroi* was enzymatically efficient over a broad pH range on a series of natural polysaccharides, formed by glucose-, xylose-, mannan- and/or arabinose-containing polymers, linked by various types of glycosidic bonds, as well as ramification types. Our proteomic approach successfully identified a large number of relevant polypeptides in the *C. gestroi* digestome. A total of 55 different proteins were identified and classified into 29 CAZY families. Based on the total number of peptides identified, the majority of components found in the *C. gestroi* digestome were cellulose-degrading enzymes. Xylanolytic enzymes, mannan- hydrolytic enzymes, pectinases and starch-degrading and debranching enzymes were also identified. Our strategy enabled validation of liquid chromatography with tandem mass spectrometry recognized proteins, by enzymatic functional assays and by following the degradation products of specific 8-amino-1,3,6-pyrenetrisulfonic acid labeled oligosaccharides through capillary zone electrophoresis.

**Conclusions:** Here we describe the first global study on the enzymatic repertoire involved in plant polysaccharide degradation by the lower termite *C. gestroi*. The biochemical characterization of whole body termite extracts evidenced their ability to cleave all types of glycosidic bonds present in plant polysaccharides. The comprehensive proteomic analysis, revealed a complete collection of hydrolytic enzymes including cellulases (GH1, GH3, GH5, GH7, GH9 and CBM 6), hemicellulases (GH2, GH10, GH11, GH16, GH43 and CBM 27) and pectinases (GH28 and GH29).

### Background

Fossil fuels are the world's primary energy source and a major issue regarding greenhouse gases emission. The key mitigation action is replacing petroleum and its derivatives with renewable energy sources [1]. In this context, lignocellulosic materials, which are an abundant

and low-cost source of stored energy in the biosphere, have been moved towards the forefront of the biofuel industry as a sustainable resource [2]. However, saccharification and production of bioproducts derived from plant cell walls are complex and usually lengthy processes. For example, cellulosic biomass must go through an intensive pretreatment step, in which enzymes are used to break down biomass into simple sugars suitable for fermentation and bioethanol production. The protein repertoire involved in biomass conversion includes

\* Correspondence: [goncalo@unicamp.br](mailto:goncalo@unicamp.br)

<sup>2</sup>Genomic and Expression Laboratory (LGE), Genetic, Evolution and Bioagents Department, State University of Campinas, Campinas, Brazil  
Full list of author information is available at the end of the article

# ANEXO 5 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA JOURNAL OF STRUCTURAL BIOLOGY EM DEZEMBRO DE 2011.

Journal of Structural Biology 177 (2012) 469–476



Contents lists available at SciVerse ScienceDirect

Journal of Structural Biology

journal homepage: [www.elsevier.com/locate/yjsbi](http://www.elsevier.com/locate/yjsbi)



## Molecular insights into substrate specificity and thermal stability of a bacterial GH5-CBM27 endo-1,4- $\beta$ -D-mannanase

Camila Ramos dos Santos<sup>a</sup>, Joice Helena Paiva<sup>a</sup>, Andreia Navarro Meza<sup>a</sup>, Junio Cota<sup>b</sup>, Thabata Maria Alvarez<sup>b</sup>, Roberto Ruller<sup>b</sup>, Rolf Alexander Prade<sup>c</sup>, Fabio Marcio Squina<sup>b</sup>, Mario Tyago Murakami<sup>a,\*</sup>

<sup>a</sup> Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, SP, Brazil

<sup>b</sup> Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, SP, Brazil

<sup>c</sup> Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, USA

### ARTICLE INFO

#### Article history:

Received 9 July 2011

Received in revised form 4 November 2011

Accepted 18 November 2011

Available online 3 December 2011

#### Keywords:

Mannan endo-1,4- $\beta$ -mannosidase

Glycoside hydrolase family 5

Carbohydrate binding module 27

*Thermotoga petrophila* RKU-1

Crystal structure

Substrate recognition

### ABSTRACT

The breakdown of  $\beta$ -1,4-mannoside linkages in a variety of mannan-containing polysaccharides is of great importance in industrial processes such as kraft pulp delignification, food processing and production of second-generation biofuels, which puts a premium on studies regarding the prospection and engineering of  $\beta$ -mannanases. In this work, a two-domain  $\beta$ -mannanase from *Thermotoga petrophila* that encompasses a GH5 catalytic domain with a C-terminal CBM27 accessory domain, was functionally and structurally characterized. Kinetic and thermal denaturation experiments showed that the CBM27 domain provided thermo-protection to the catalytic domain, while no contribution on enzymatic activity was observed. The structure of the catalytic domain determined by SIRAS revealed a canonical ( $\alpha/\beta$ )<sub>8</sub>-barrel scaffold surrounded by loops and short helices that form the catalytic interface. Several structurally related ligand molecules interacting with TpMan were solved at high-resolution and resulted in a wide-range representation of the subsites forming the active-site cleft with residues W134, E198, R200, E235, H283 and W284 directly involved in glucose binding.

© 2011 Elsevier Inc. All rights reserved.

### 1. Introduction

Mannan endo-1,4- $\beta$ -D-mannosidase or 1,4- $\beta$ -D-mannan mannohydrolase (EC 3.2.1.78), commonly referred as  $\beta$ -mannanase, catalyzes the hydrolysis of  $\beta$ -1,4-mannoside linkages in various mannan-containing polysaccharides, such as glucomannans and galactomannans (Stålbrand et al., 1993; de Vries and Visser, 2001). Degradation of these polysaccharides represents a key step for a number of industrial applications including delignification of kraft pulps (Tenkanen et al., 1997; Montiel et al., 2002), food processing (Sachslehner et al., 2000; Dhavan and Kaur, 2007) and production of second-generation biofuels (Dhavan and Kaur, 2007). In general, these biotechnological processes such as biomass pre-treatments, are performed under extreme environmental

conditions regarding pH, osmolarity and temperature. Thus,  $\beta$ -mannanases being stable and functional at high temperatures offer substantial techno-economical advantages.

In addition to their biotechnological relevance, mannan-degrading enzymes also participate in a number of biological processes such as fruit ripening (Pressey, 1989), seed germination (Black, 1996) and remodeling of plant cell walls (reviewed in Schröder et al., 2009). These enzymes have also been used in structural characterization of polysaccharides having  $\beta$ -mannosidic linkages and sequencing of heteropolysaccharides and carbohydrates attached to glycoproteins (Dhavan and Kaur, 2007).

*Thermotoga petrophila* strain RKU-1 (T) is a hyperthermophilic bacterium isolated from the Kubiki oil reservoir in Niigata (Japan) that grows optimally at 80 °C (Takahata et al., 2001). Some hyperthermostable enzymes produced by this microorganism have demonstrated great potential for industrial applications and served as models for investigating structure–function–stability relationships in multidomain glycosyl hydrolases (Santos et al., 2010, 2011; Squina et al., 2010; Cota et al., 2011).

The  $\beta$ -mannanase from *T. petrophila* RKU-1, TpMan, consists of a CaZy GH5 catalytic core connected to a CBM27 accessory domain by an 100-residue-long linker. To date, only six structures of GH5 endo- $\beta$ -1,4-mannanases have been solved: *Thermobifida fusca*

**Abbreviations:** Mannanases: TpMan, from *Thermotoga petrophila*; TfMan, from *Thermobifida fusca*; TrMan, from *Trichoderma reesei*; LeMan, from *Lycopersicon esculentum*; CmMan, from *Cellvibrio mixtus*; CjMan, from *Cellvibrio japonicus*; BaMan, from *Bacillus agaradhaerens*; BsMan, from *Bacillus subtilis*; VsMan, from *Vibrio* sp. strain MA-138.

\* Corresponding author. Address: Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Rua Giuseppe Maximo Scolfaro, 10000 Campinas, 13083-970 SP, Brazil. Fax: +55 19 3512 1004.

E-mail address: [mario.murakami@lnbio.org.br](mailto:mario.murakami@lnbio.org.br) (M.T. Murakami).

1047-8477/\$ - see front matter © 2011 Elsevier Inc. All rights reserved.  
doi:10.1016/j.jsb.2011.11.021

# ANEXO 6 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA BIORESOURCE TECHNOLOGY EM MAIO DE 2012.

Bioresource Technology 119 (2012) 293–299



Contents lists available at SciVerse ScienceDirect

Bioresource Technology

journal homepage: [www.elsevier.com/locate/biortech](http://www.elsevier.com/locate/biortech)



## Functional characterization and synergic action of fungal xylanase and arabinofuranosidase for production of xylooligosaccharides

T.A. Gonçalves<sup>a,1</sup>, A.R.L. Damásio<sup>a,1</sup>, F. Segato<sup>a,1</sup>, T.M. Alvarez<sup>a</sup>, J. Bragatto<sup>a</sup>, L.B. Brenelli<sup>a</sup>, A.P.S. Citadini<sup>a</sup>, M.T. Murakami<sup>b</sup>, R. Ruller<sup>a</sup>, A.F. Paes Leme<sup>b</sup>, R.A. Prade<sup>c</sup>, F.M. Squina<sup>a,\*</sup>

<sup>a</sup> Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, SP, Brazil

<sup>b</sup> Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, SP, Brazil

<sup>c</sup> Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, United States

### HIGHLIGHTS

- ▶ Hypersecretion of xylanase and arabinofuranosidase by fungal expression system.
- ▶ Biochemical and biophysical characterization of both enzymes.
- ▶ Depiction of the mode of operation of both enzymes.
- ▶ Synergistic breakdown of wheat xylan and sugar cane bagasse.
- ▶ Enzymatic production of xylooligosaccharides from hemicellulosic feedstock.

### ARTICLE INFO

#### Article history:

Received 3 February 2012

Received in revised form 10 May 2012

Accepted 11 May 2012

Available online 23 May 2012

#### Keywords:

Endo-xylanases

Arabinofuranosidases

Xylooligosaccharides

Hemicellulosic feedstock

Sugar cane bagasse

### ABSTRACT

Plant cell wall degrading enzymes are key technological components in biomass bioconversion platforms for lignocellulosic materials transformation. Cost effective production of enzymes and identification of efficient degradation routes are two economic bottlenecks that currently limit the use of renewable feedstocks through an environmental friendly pathway. The present study describes the hypersecretion of an endo-xylanase (GH11) and an arabinofuranosidase (GH54) by a fungal expression system with potential biotechnological application, along with comprehensive characterization of both enzymes, including spectrometric analysis of thermal denaturation, biochemical characterization and mode of action description. The synergistic effect of these enzymes on natural substrates such as sugarcane bagasse, demonstrated the biotechnological potential of using GH11 and GH54 for production of probiotic xylooligosaccharides from plant biomass. Our findings shed light on enzymatic mechanisms for xylooligosaccharide production, as well as provide basis for further studies for the development of novel enzymatic routes for use in biomass-to-bioethanol applications.

© 2012 Elsevier Ltd. All rights reserved.

### 1. Introduction

Lignocellulosic biomass has great potential as a renewable energy source and is considered the most promising feedstock for producing biofuels due to its versatility, availability and low cost. Biomass hydrolytic enzymes are key technological components for the efficient use of renewable feedstocks through an environmentally friendly bioconversion routes. Bottlenecks in this process include the cost effective production of enzymes and efficient deg-

radation of recalcitrant plant cell wall polymers into fermentable sugars (Pauly and Keegstra, 2008). Therefore, development of new routes for plant feedstock bioconversion into simple sugars, along with strategies for production of enzymes at high yield, are main foci in the biofuel research field.

Xylans can be broadly classified as homoxylans, arabinoxylans, glucuronoxylans, and arabinoglucuronoxylans (Polizeli et al., 2005). Arabinoxylans are the main hemicellulose forms in plant cell walls, especially in cereal grains such as wheat, and consist of a xylose backbone with arabinose residues linked to its O-2 or O-3 (Polizeli et al., 2005). One of the critically important enzymatic activities required to break down the xylan backbone are provided by endo-1,4- $\beta$ -xylanases. These enzymes cleave the  $\beta$ -1,4 glycosidic linkage between xylose residues in the backbone. Xylanases have been classified into glycoside hydrolases (GH) families 5, 7, 8, 10, 11, and 43 on the basis of their amino acid sequences,

\* Corresponding author. Address: Centro Nacional de Pesquisa em Energia e Materiais (CNPEM) Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE) Caixa Postal 6170, 13083-970, Campinas (SP), Brazil. Tel.: +55 19 3518 3111; fax: +55 19 35183104.

E-mail address: [fabio.squina@bioetanol.org.br](mailto:fabio.squina@bioetanol.org.br) (F.M. Squina).

<sup>1</sup> These authors contributed equally to this work.

## The *Penicillium echinulatum* Secretome on Sugar Cane Bagasse

Daniela A. Ribeiro<sup>1</sup>, Júnio Cota<sup>1</sup>, Thabata M. Alvarez<sup>1</sup>, Fernanda Brüchli<sup>1</sup>, Juliano Bragato<sup>1</sup>, Beatriz M. P. Pereira<sup>1</sup>, Bianca A. Pauletti<sup>1</sup>, George Jackson<sup>1</sup>, Maria T. B. Pimenta<sup>1</sup>, Mario T. Murakami<sup>2</sup>, Marli Camassola<sup>3</sup>, Roberto Ruller<sup>1</sup>, Aldo J. P. Dillon<sup>3</sup>, Jose G. C. Pradella<sup>1</sup>, Adriana F. Paes Leme<sup>1</sup>, Fabio M. Squina<sup>1\*</sup>

**1** Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, (CNPEM), Campinas, São Paulo, Brazil, **2** Laboratório de Espectrometria de Massas, Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, (CNPEM), Campinas, São Paulo, Brazil, **3** Instituto de Biotecnologia, Universidade de Caxias do Sul (UCS), Caxias do Sul, Rio Grande do Sul, Brazil

### Abstract

Plant feedstocks are at the leading front of the biofuel industry based on the potential to promote economical, social and environmental development worldwide through sustainable scenarios related to energy production. *Penicillium echinulatum* is a promising strain for the bioethanol industry based on its capacity to produce large amounts of cellulases at low cost. The secretome profile of *P. echinulatum* after grown on integral sugarcane bagasse, microcrystalline cellulose and three types of pretreated sugarcane bagasse was evaluated using shotgun proteomics. The comprehensive chemical characterization of the biomass used as the source of fungal nutrition, as well as biochemical activity assays using a collection of natural polysaccharides, were also performed. Our study revealed that the enzymatic repertoire of *P. echinulatum* is geared mainly toward producing enzymes from the cellulose complex (endoglucanases, cellobiohydrolases and  $\beta$ -glucosidases). Glycoside hydrolase (GH) family members, important to biomass-to-biofuels conversion strategies, were identified, including endoglucanases GH5, 7, 6, 12, 17 and 61,  $\beta$ -glycosidase GH3, xylanases GH10 and GH11, as well as debranching hemicellulases from GH43, GH62 and CE2 and pectinases from GH28. Collectively, the approach conducted in this study gave new insights on the better comprehension of the composition and degradation capability of an industrial cellulolytic strain, from which a number of applied technologies, such as biofuel production, can be generated.

**Citation:** Ribeiro DA, Cota J, Alvarez TM, Brüchli F, Bragato J, et al. (2012) The *Penicillium echinulatum* Secretome on Sugar Cane Bagasse. PLoS ONE 7(12): e50571. doi:10.1371/journal.pone.0050571

**Editor:** Gustavo Henrique Goldman, Universidade de Sao Paulo, Brazil

**Received:** July 27, 2012; **Accepted:** October 23, 2012; **Published:** December 5, 2012

**Copyright:** © 2012 Ribeiro et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This research was supported by grants from FAPESP (2008/58037-9) and CNPq (475022/2011-4 e 310177/2011-1). TMA received a fellowship from FAPESP (2010/11499-1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: fabio.squina@bioetanol.org.br

### Introduction

Plant structural polysaccharides are the most abundant and renewable biomass in the biosphere. Plant feedstocks are at the leading front of the biofuel industry based on the potential to promote economical, social and environmental development worldwide through sustainable scenarios related to energy production [1]. However, saccharification and bioproduct manufacturing from lignocellulose biomass are complex and lengthy processes. The current schemes for the biotechnological conversion of plant cell wall polysaccharides rely on first reducing biomass recalcitrance through a pretreatment step, and afterward, enzymatic cocktails are needed to breakdown biomass into more simple, fermentable saccharides, which could be fed into several bioprocesses, such as bioethanol production.

Despite the advantages of enzyme-catalyzed processes, *i.e.*, speed, specificity and mildness, the high cost of enzyme production and low catalytic efficiency are still major hindrances for cellulosic bioethanol. Thus, relevant biotechnological challenges in this field include the improvement of the catalytic efficiency of enzymes, the economic benefit and the synergy between the type of pretreat-

ment and enzymatic load, and reduction of the cost of enzyme production by filamentous fungi [2,3].

The enzymatic degradation of cellulosic materials by fungal enzyme systems, especially those produced by *Trichoderma reesei* and *Aspergillus* spp., has been extensively studied due to its effectiveness for the liberation of fermentable sugars for bioethanol production [4,5]. Several efforts, such as optimization of fermentation processes [6, 7, and 8] and genetic modifications of the microorganisms [9, 10, 11, and 12], are being targeted for improvement of fungal enzymatic systems. More recently, *Penicillium echinulatum* has been the focus of attention due to its potential to produce large amounts of cellulases at low costs; it has also been considered a promising strain for the bioethanol industry [7, 8, 9, 10, 13 and 14].

Advances in proteomics analysis have pushed forward secretome studies on filamentous fungi [15, 16, 17 and 18], mainly because they can highlight pathways as well as target genes for deletions or co-expression, to improve strains for biotechnological purposes [19]. In addition, secretome studies are appealing for basic research regarding the role of filamentous fungi not only as ubiquitous saprophytes in nature but also as cell factories to secrete

## ANEXO 8 – ARTIGO PUBLICADO (EM COLABORAÇÃO) NA REVISTA *LETTERS IN APPLIED MICROBIOLOGY* EM ABRIL DE 2013.

### ORIGINAL ARTICLE

## The characterization of a thermostable and cambialistic superoxide dismutase from *Thermus filiformis*

F. Mandelli<sup>1,2</sup>, J.P.L. Franco Cairo<sup>1</sup>, A.P.S. Citadini<sup>1</sup>, F. Büchli<sup>1</sup>, T.M. Alvarez<sup>1</sup>, R.J. Oliveira<sup>3</sup>, V.B.P. Leite<sup>3</sup>, A.F. Paes Leme<sup>4</sup>, A.Z. Mercadante<sup>2</sup> and F.M. Squina<sup>1</sup>

1 Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, Brazil

2 Departamento de Ciência de Alimentos da Faculdade de Engenharia de Alimentos, UNICAMP, Campinas, Brazil

3 Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, Brazil

4 Laboratório Nacional de Biociências (LNBio), do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, Brazil

**Significance and Impact of the Study:** This manuscript describes the expression and characterization of a superoxide dismutase (SOD) from *Thermus filiformis* with thermophilic and cambialistic characteristics. The SODs are among the most potent antioxidants known in nature, and their stability and pharmacokinetics can vary widely in accordance to their biological source. Although the currently clinical research work has been focused on human and bovine SODs, alternative sources may become more biotechnological attractive in the near future. Our study brings new insights for the research field of antioxidant enzymes with potential application on pharmaceutical, cosmetics and food formulations.

### Keywords

3-D model, cambialistic enzyme, circular dichroism, nitroblue tetrazolium, pyrogallol autoxidation.

### Correspondence

F.M. Squina, Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Rua Giuseppe Máximo Scolfaro, nº 10000, Campinas, SP-13083-970, Brazil. E-mail: fabio.squina@bioetanol.org.br

2012/2317: received 27 December 2012, revised 15 March 2013 and accepted 15 March 2013

doi:10.1111/lam.12071

### Abstract

The superoxide dismutase (TfSOD) gene from the extremely thermophilic bacterium *Thermus filiformis* was cloned and expressed at high levels in mesophilic host. The purified enzyme displayed approximately 25 kDa band in the SDS-PAGE, which was further confirmed as TfSOD by mass spectrometry. The TfSOD was characterized as a cambialistic enzyme once it had enzymatic activity with either manganese or iron as cofactor. TfSOD showed thermostability at 65, 70 and 80°C. The amount of enzyme required to inhibit 50% of pyrogallol autoxidation was 0.41, 0.56 and 13.73 mg at 65, 70 and 80°C, respectively. According to the circular dichroism (CD) spectra data, the secondary structure was progressively lost after increasing the temperature above 70°C. The 3-dimensional model of TfSOD with the predicted cofactor binding corroborated with functional and CD analysis.

### Introduction

Antioxidant enzymes such as superoxide dismutase (SOD) and catalase protect living cells from reactive oxygen and nitrogen species that are responsible for oxidative damage of essential components of cell structure (Pantazaki *et al.* 2002). Superoxide dismutases are a class of antioxidant defence metalloenzymes that disproportionate superoxide radical ion into molecular oxygen and hydrogen peroxide (Whittaker and Whittaker 1999). Due to its antioxidative effects, SOD has been widely applied in medical treatments,

as well as cosmetic, food, agricultural and chemical industries (Liu *et al.* 2011). SODs have recently found applications as supplementation to prevent or reverse the adverse effect of cardiovascular diseases, ageing, infertility, neurological disorders, ischaemia–reperfusion injury, transplant rejection, autoimmune diseases, rheumatoid arthritis, diabetes, asthma, septic shock-induced tissue injury and cancer, as well as in the pharmaceutical and cosmetic industries (Bafana *et al.* 2011).

Once thermal denaturation is a common cause of enzyme inactivation in the industry, one of the major

## Cellulase On-Site Production from Sugar Cane Bagasse Using *Penicillium echinulatum*

Beatriz Merchel Piovesan Pereira · Thabata Maria Alvarez ·  
Priscila da Silva Delabona · Aldo José Pinheiro Dillon ·  
Fabio Márcio Squina · José Geraldo da Cruz Pradella

© Springer Science+Business Media New York 2013

**Abstract** *Penicillium echinulatum* was evaluated as a cellulolytic enzyme producer in shaking flasks and bioreactor submerged culture using sugarcane bagasse as carbon source. Sodium hydroxide delignified steam-exploded pretreated bagasse (SDB) and hydrothermal pretreated bagasse had a maximum filter paper activity (FPase) of 2.4 and 2.6 FPU/mL, respectively. Delignified acid pretreated bagasse and Celufloc 200TM (CE) carbon sources displayed maximum FPase of 1.3 and 1.6 FPU/mL while in natura bagasse (INB) provided the lowest enzyme activity, ca. 0.4 FPU/mL. Measurement of surface specific area of lignocellulosic material and scanning electron microscopic images showed a possible correlation between fungal mycelia accessibility to lignocellulosic particles and obtained cellulolytic enzyme activity of fermentation broth. Fed-batch experiments performed in a controlled bioreactor attained the highest value of FPase of 3.7 FPU/mL, enzyme productivity of 25.7 FPU/L h, and enzyme yield from cellulose equal to 134 FPU/g with SDB. Enzyme hydrolysis of steam-pretreated bagasse accomplished with the obtained supernatant of fermentation broth (10 FPU/g of biomass and 5 % w/v) performed better than commercial cellulose complex. The results showed that *P. echinulatum* has potential to be used as an on-site enzyme platform aiming second bioethanol production from sugarcane lignocellulosic residue.

**Keywords** *Penicillium echinulatum* · Cellulase · Enzyme · Submerged fermentation · Second generation bioethanol

### Introduction

Over 40,000,000 tons of bagasse and leaves from sugarcane are generated every season in Brazil, and this material has become recognized worldwide as one of the most important lignocellulosic wastes for second generation bioethanol production. Its utilization could increase Brazilian ethanol production by 40 % for the same area of sugarcane crop [1].

Enzymatic hydrolysis is the preferred technique for release of fermentable sugars from lignocellulosic biomass. However, the main bottleneck for wider application of cellulases in second generation ethanol production is their cost, especially considering that large quantities of the enzymes are required [2]. The cost is a function of enzyme productivity (the amount of enzyme produced in a defined bioreactor volume according to time), enzyme yield from the carbon source ( $Y_{p/S}$ , the amount of enzyme produced according to the amount of carbon source consumed), enzyme product recovery, formulation, and transportation, as well as its tax for commercialization.

As a low-cost feedstock [3], cellulosic biomass is an ideal carbon source for enzyme production. It is therefore logical to try to produce, on-site, tailor-made enzymatic mixtures that are optimized for the specific feedstock to be processed. Furthermore, the use of the same material for enzyme production and hydrolysis at the same location should help to reduce the production costs of second generation ethanol since both steps of the operation could be undertaken in the same place and benefit from the installed infrastructure (including the provision of water, electricity, and steam). Other advantages are avoidance of downstream operations including unnecessary product formulations, transport from the manufacturer to the consumer, and storage for long periods [3, 4].

B. M. P. Pereira · T. M. Alvarez · P. da Silva Delabona ·  
F. M. Squina · J. G. da Cruz Pradella (✉)  
Brazilian Bioethanol Science and Technology Laboratory, Rua  
Giuseppe Maximo Scolfaro 10000,  
Pólo II de Alta Tecnologia, CP 6192,  
Campinas, São Paulo 13083-970, Brazil  
e-mail: jpradella@bioetanol.org.br

A. J. P. Dillon  
Biotechnology Institute, University of Caxias do Sul,  
Rua Francisco Getulio Vargas 1130,  
Caxias do Sul, Rio Grande do Sul 95010-550, Brazil



## Assembling a xylanase–lichenase chimera through all-atom molecular dynamics simulations



Junio Cota<sup>a,d,1</sup>, Leandro C. Oliveira<sup>a,b,1</sup>, André R.L. Damásio<sup>a</sup>, Ana P. Citadini<sup>a</sup>, Zaira B. Hoffmam<sup>a</sup>, Thabata M. Alvarez<sup>a</sup>, Carla A. Codima<sup>a</sup>, Vitor B.P. Leite<sup>b</sup>, Glaucia Pastore<sup>c</sup>, Mario de Oliveira-Neto<sup>d</sup>, Mario T. Murakami<sup>e</sup>, Roberto Ruller<sup>a</sup>, Fabio M. Squina<sup>a,\*</sup>

<sup>a</sup> Laboratório Nacional de Ciência e Tecnologia do Bioetanol – CTBE/CNPEM, Campinas, SP, Brazil

<sup>b</sup> Departamento de Física, IBILCE, Universidade Estadual Paulista - UNESP, São José do Rio Preto, SP, Brazil

<sup>c</sup> Faculdade de Engenharia de Alimentos, Universidade Estadual de Campinas, Campinas, SP, Brazil

<sup>d</sup> Departamento de Física e Biofísica, Instituto de Biociências, UNESP, Botucatu, São Paulo, Brazil

<sup>e</sup> Laboratório Nacional de Biociências – LNBio/CNPEM, Campinas, SP, Brazil

### ARTICLE INFO

#### Article history:

Received 8 October 2012

Received in revised form 23 January 2013

Accepted 20 February 2013

Available online 28 February 2013

#### Keywords:

Multifunctional enzyme

Small-angle X-ray scattering

Molecular dynamics

Computational characterization

Experimental validation

### ABSTRACT

Multifunctional enzyme engineering can improve enzyme cocktails for emerging biofuel technology. Molecular dynamics through structure-based models (SB) is an effective tool for assessing the tridimensional arrangement of chimeric enzymes as well as for inferring the functional practicability before experimental validation. This study describes the computational design of a bifunctional xylanase–lichenase chimera (XylLich) using the *xynA* and *bglS* genes from *Bacillus subtilis*. In silico analysis of the average solvent accessible surface area (SAS) and the root mean square fluctuation (RMSF) predicted a fully functional chimera, with minor fluctuations and variations along the polypeptide chains. Afterwards, the chimeric enzyme was built by fusing the *xynA* and *bglS* genes. XylLich was evaluated through small-angle X-ray scattering (SAXS) experiments, resulting in scattering curves with a very accurate fit to the theoretical protein model. The chimera preserved the biochemical characteristics of the parental enzymes, with the exception of a slight variation in the temperature of operation and the catalytic efficiency ( $k_{cat}/K_m$ ). The absence of substantial shifts in the catalytic mode of operation was also verified. Furthermore, the production of chimeric enzymes could be more profitable than producing a single enzyme separately, based on comparing the recombinant protein production yield and the hydrolytic activity achieved for XylLich with that of the parental enzymes.

© 2013 Elsevier B.V. All rights reserved.

### 1. Introduction

Plant biomass saccharification and biofuel production have been described as promising renewable alternatives to petroleum and natural gas. However, the polysaccharide network in plant cell walls is one of the most complex structures in nature, which jeopardizes the production of biofuels from plant biomass. First, biomass feedstock must go through a recalcitrance-reducing step (pretreatment). Then, enzymatic cocktails are used to precisely breakdown the polysaccharides into simple sugars suitable for several bioprocesses, such as fermentation to ethanol [1].

The enzymatic cocktail for plant biomass saccharification and biofuel production must include cellulolytic hydrolases, such as cellobiohydrolases, endo-glucanases and  $\beta$ -glucosidases. For hemicellulose degradation, synergistic action by hydrolytic enzymes is required at the polysaccharide backbone, side chains and decorating units. For instance, the hydrolysis of feedstock containing arabinoxylan requires several hydrolytic enzymes, such as endo-xylanases,  $\beta$ -xylosidases, arabinofuranosidases, ferulic acid esterases, glucuronidases and other enzymes [2].  $\beta$ -glucans, polysaccharides with  $\beta$ -1,3 and  $\beta$ -1,4 glucosidic linkages, are also abundant in many plant cell walls, especially in sugarcane [3]. The enzymatic depolymerization of 1,3-1,4- $\beta$ -glucans is catalyzed by 1,4- $\beta$ -D-glucan 4-glucanohydrolase (EC 3.2.1.4), 1,3- $\beta$ -D-glucan 3-glucanohydrolase (EC 3.2.1.39) and 1,3-1,4- $\beta$ -D-4-glucanohydrolase or lichenase (EC 3.2.1.73) [4].

The engineering of multifunctional proteins with a synergistic catalytic capacity has the potential to streamline biomass conversion strategies [5]. However, the unsupervised construction of enzyme

\* Corresponding author at: Laboratório Nacional de Ciência e Tecnologia do Bioetanol – CTBE/CNPEM, Caixa Postal 6170, 13083-970 Campinas, São Paulo, Brazil. Tel.: +55 19 3518 3111; fax: +55 19 3518 3104.

E-mail address: [fabio.squina@bioetanol.org.br](mailto:fabio.squina@bioetanol.org.br) (F.M. Squina).

<sup>1</sup> These authors contributed equally to this work.

## Characterization of a Hexameric Exo-Acting GH51 $\alpha$ -L-Arabinofuranosidase from the Mesophilic *Bacillus subtilis*

Zaira B. Hoffmam · Leandro C. Oliveira · Junio Cota · Thabata M. Alvarez ·  
José A. Diogo · Mario de Oliveira Neto · Ana Paula S. Citadini · Vitor B. P. Leite ·  
Fabio M. Squina · Mario T. Murakami · Roberto Ruller

© Springer Science+Business Media New York 2013

**Abstract**  $\alpha$ -L-Arabinofuranosidases ( $\alpha$ -L-Abfases, EC 3.2.1.55) display a broad specificity against distinct glycosyl moieties in branched hemicellulose and recent studies have demonstrated their synergistic use with cellulases and xylanases for biotechnological processes involving plant biomass degradation. In this study, we examined the structural organization of the arabinofuranosidase (GH51 family) from the mesophilic *Bacillus subtilis* (AbfA) and its implications on function and stability. The recombinant AbfA showed to be active over a broad temperature range with the maximum activity between 35 and 50 °C, which is desirable for industrial applications. Functional studies demonstrated

that AbfA preferentially cleaves debranched or linear arabinan and is an exo-acting enzyme producing arabinose from arabinohexaose. The enzyme has a canonical circular dichroism spectrum of  $\alpha/\beta$  proteins and exhibits a hexameric quaternary structure in solution, as expected for GH51 members. Thermal denaturation experiments indicated a melting temperature of 53.5 °C, which is in agreement with the temperature–activity curves. The mechanisms associated with the unfolding process were investigated through molecular dynamics simulations evidencing an important contribution of the quaternary arrangement in the stabilization of the  $\beta$ -sandwich accessory domain and other regions involved in the formation of the catalytic interface of hexameric Abfases belonging to GH51 family.

Zaira B. Hoffmam and Leandro C. Oliveira have contributed equally to this work.

**Electronic supplementary material** The online version of this article (doi:10.1007/s12033-013-9677-1) contains supplementary material, which is available to authorized users.

Z. B. Hoffmam · L. C. Oliveira · J. Cota ·  
T. M. Alvarez · J. A. Diogo · A. P. S. Citadini ·  
F. M. Squina · R. Ruller (✉)

Laboratório Nacional de Ciência e Tecnologia do Bioetanol, Centro  
Nacional de Pesquisa em Energia e Materiais, Rua Giuseppe  
Maximo Solfaro 10000, Campinas, SP 13083-970, Brazil  
e-mail: roberto.ruller@bioetanol.org.br

Z. B. Hoffmam  
e-mail: zairabh@gmail.com

L. C. Oliveira  
e-mail: leandro.oliveira@bioetanol.org.br

J. Cota  
e-mail: juniocs@gmail.com

T. M. Alvarez  
e-mail: thabata.alvarez@gmail.com

J. A. Diogo  
e-mail: josealberto2511@yahoo.com.br

**Keywords** Arabinofuranosidase · Glycosyl hydrolase  
family 51 · *Bacillus subtilis* · Quaternary structure

A. P. S. Citadini  
e-mail: ana.citadini@bioetanol.org.br

F. M. Squina  
e-mail: fabio.squina@bioetanol.org.br

Z. B. Hoffmam · J. Cota · T. M. Alvarez ·  
J. A. Diogo · F. M. Squina · M. T. Murakami · R. Ruller  
Universidade Estadual de Campinas (UNICAMP),  
Campinas, SP, Brazil

L. C. Oliveira · V. B. P. Leite  
Departamento de Física, Instituto de Biociências,  
Letras e Ciências Exatas, Universidade Estadual Paulista  
(IBILCE/UNESP), São José do Rio Preto, SP, Brazil

V. B. P. Leite  
e-mail: vleite@sjrp.unesp.br



## ANEXO 12 – DECLARAÇÃO DE BIOÉTICA E BIOSSEGURANÇA

Uso exclusivo da CIBio:

Número de projeto / processo: FMSA.3

### Formulário de encaminhamento de projetos de pesquisa para análise pela CIBio - Comissão Interna de Biossegurança da ABTLuS – Associação Brasileira de Tecnologia de Luz Síncrotron

Título do projeto: "Library Generation for Biomass Conversion Enzymes from Soil Metagenome"

Pesquisador responsável: Fabio Marcio Squina

Experimentador: Thabata Maria Alvarez

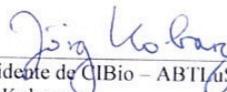
Nível do treinamento do experimentador: [ ]-Iniciação científica, [ ]-mestrado, [ ]-doutorado, [x]-doutorado direto, [ ]-pós-doutorado, [ ]-nível técnico, [ ]-outro, especifique: \_\_\_\_\_

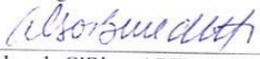
Resumo do projeto:

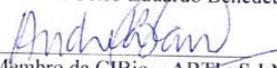
O projeto visa a através das ferramentas de metagenômica elaborar uma biblioteca de enzimas com potencial aplicação biotecnológica, auxiliando no entendimento da estrutura complexa da parede celular vegetal e a participação no processo de produção do etanol de segunda geração

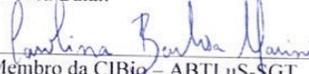
A CIBio analisou este projeto em reunião realizada no dia: 17-2-2011

Parecer final: [x]-projeto aprovado, [ ]-projeto recusado, [ ]-projeto com deficiências, favor comentários anexo.

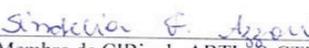
  
Presidente da CIBio – ABTLuS-LNBio  
Jörg Kobarg

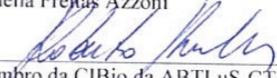
  
Membro da CIBio – ABTLuS-LNBio  
Prof. Dr. Celso Eduardo Benedetti

  
Membro da CIBio – ABTLuS-LNBio  
Andrea Balan

  
Membro da CIBio – ABTLuS-SGT  
Carolina Barbosa Marini

  
Membro da CIBio da ABTLuS-CTBE  
Fabio Squina

  
Membro da CIBio da ABTLuS-CTBE  
Sindelia Freitas Azzoni

  
Membro da CIBio da ABTLuS-CTBE  
Roberto Ruller



# ANEXO 13 – COMPROVANTE DE AUTORIZAÇÃO DE BIOÉTICA E BIOSSEGURANÇA

## DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação de Mestrado/tese de Doutorado intitulada “Desenvolvimento de uma Biblioteca de Enzimas a partir de Metagenoma de Solo”.

( ) não se enquadra no § 3º do Artigo 1º da Informação CCPG 01/08, referente a bioética e biossegurança.

Tem autorização da(s) seguinte(s) Comissão(ões):

( x ) CIBio – Comissão Interna de Biossegurança , projeto No. FMS1.3, Instituição: CTBE – Laboratório Nacional de Ciência e Tecnologia do Bioetanol / CNPEM – Centro Nacional de Pesquisa em Energia e Materiais.

( ) CEUA – Comissão de Ética no Uso de Animais , projeto No. \_\_\_\_\_, Instituição:

( ) CEP - Comissão de Ética em Pesquisa, protocolo No. \_\_\_\_\_, Instituição:

*\* Caso a Comissão seja externa ao IB/UNICAMP, anexar o comprovante de autorização dada ao trabalho. Se a autorização não tiver sido dada diretamente ao trabalho de tese ou dissertação, deverá ser anexado também um comprovante do vínculo do trabalho do aluno com o que constar no documento de autorização apresentado.*



Aluno Thabata Maria Alvarez

  
Orientador: Fabio Marcio Squina

Para uso da Comissão ou Comitê pertinente:

(x) Deferido ( ) Indeferido

Carimbo e assinatura

  
Prof. Dr. MARCELO LANCELLOTTI  
Presidente da Comissão Interna de Biossegurança  
Instituto de Biologia - UNICAMP

Para uso da Comissão ou Comitê pertinente:

( ) Deferido ( ) Indeferido

Carimbo e assinatura

