# UNIVERSIDADE ESTADUAL DE CAMPINAS

# INSTITUTO DE BIOLOGIA

## AGNES CRISTINA PIMENTEL

## CARACTERÍSTICAS BIOQUÍMICAS E ESTRUTURAIS DO DOMÍNIO CATALÍTICO DA CELULASE CelE2 E SUA RELAÇÃO COM O DOMÍNIO ACESSÓRIO CALX-β

## BIOCHEMICAL AND STRUCTURAL FEATURES OF THE CATALYTIC DOMAIN OF CELLULASE CelE2 AND ITS RELATION WITH THE CALX-β ACCESSORY DOMAIN

**CAMPINAS**

**2018**

**AGNES CRISTINA PIMENTEL**

**CARACTERÍSTICAS BIOQUÍMICAS E ESTRUTURAIS DO DOMÍNIO CATALÍTICO DA CELULASE CelE2 E SUA RELAÇÃO COM O DOMÍNIO ACESSÓRIO CALX-β**

**BIOCHEMICAL AND STRUCTURAL FEATURES OF THE CATALYTIC DOMAIN OF CELLULASE CelE2 AND ITS RELATION WITH THE CALX-β ACCESSORY DOMAIN**

> *Dissertação apresentada ao Instituto de Biologia da Universidade Estadual de Campinas como parte dos requisitos exigidos para obtenção título de Mestra em BIOLOGIA FUNCIONAL E MOLECULAR, na Área de BIOQUÍMICA.*

> *Dissertation presented to the Institute of Biology of the University of Campinas in partial fulfillment of the requirements for the degree of Master in FUNCTIONAL AND MOLECULAR BIOLOGY, in the area of BIOCHEMISTRY.*

*Supervisor/ Orientador: FABIO MARCIO SQUINA*
*Co-supervisora/ Coorientadora: THÁBATA MARIA ALVAREZ*

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELA ALUNA AGNES CRISTINA PIMENTEL E ORIENTADA PELO PROF. DR. FABIO MARCIO SQUINA

**CAMPINAS**

**2018**

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

Informações para Biblioteca Digital

# COMISSÃO EXAMINADORA

Dr. Fabio Marcio Squina (orientador)

Dr. Fernando Segato

Dra. Thamy Lívia Ribeiro Corrêa


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

# AGRADECIMENTOS

À Deus primeiramente, por sempre me guiar e me dar forças para enfrentar os desafios.

Aos meus pais, minha eterna gratidão por sempre estarem presentes e me apoiarem com muito amor e carinho.

Ao meu irmão Igor, por toda ajuda e carinho.

Ao meu namorado Gabriel, por toda compreensão, amor e ajuda em todos os momentos.

À toda minha família e amigos por estarem sempre presentes. Agradeço por fazerem parte da minha vida.

Aos meus orientadores Fabio e Thabata, pela oportunidade e por toda disponibilidade, apoio, paciência, atenção e incentivo. Sou plenamente grata por todo ensinamento.

Aos meus amigos e colegas do CTBE e da UNICAMP, Marcelo, Thiago, Fernanda, Lívia, Carol, João, Mariana, Gabriela, Emerson, Douglas, Cláudia, Rebeca, Geize, Robson, e todos os outros que me acompanharam e ajudaram desde o começo.

Ao grupo de pesquisa do professor Mario de Oliveira Neto de Botucatu, pela colaboração no trabalho.

Ao Centro Nacional de Pesquisa em Energia e Materiais, ao Laboratório Nacional de Ciência e Tecnologia do Bioetanol, e à Universidade Estadual de Campinas pelo suporte e infraestrutura.

**RESUMO**

As glicosil hidrolases (GHs) compõem uma classe de enzimas amplamente utilizadas nos processos enzimáticos por atuarem na hidrólise das ligações glicosídicas de polissacarídeos. Além do domínio catalítico, diversas GHs apresentam domínios acessórios que podem influenciar diversos aspectos bioquímicos e estruturais das enzimas. A celulase CelE2 de origem metagenômica, caracterizada recentemente, apresenta em sua estrutura um domínio catalítico N-terminal característico da família GH5 e um domínio C-terminal com identidade para o domínio Calx-β, previamente identificado em outras GHs mas sem função claramente definida nestes casos. Com o intuito de investigar as características bioquímicas e estruturais dos domínios de CelE2 separadamente, foram construídas derivações truncadas referentes ao domínio catalítico e ao domínio acessório Calx- β. Como resultados, observou-se que a excisão do domínio acessório Calx-β ($CelE2_{382-477}$) não provocou alterações nas características bioquímicas de CelE2, visto que o domínio catalítico ($CelE2_{1-381}$) apresentou valores semelhantes de atividade enzimática em β-glucano, carboximetilcelulose e liquenana, valores de pH e temperatura ótimos próximos de 5,3 e 45 ℃, respectivamente, estabilidade térmica a 40 ℃ e 50 ℃ por 360 minutos e aumento de atividade enzimática na presença de $CaCl_2$. O perfil de degradação e as evidências de transglicosilação também foram mantidas após a excisão do domínio acessório Calx-β. No entanto, o domínio catalítico apresentou um aumento de cerca de 4 vezes da atividade enzimática na hidrólise de Avicel, comparado à CelE2. As análises por dicroísmo circular (CD) de $CelE2_{1-381}$ indicaram um perfil predominante de α-hélice, diminuição de 4,6 ℃ da temperatura de desnaturação e um aumento destes valores na presença de $CaCl_2$, tanto para a construção completa, quanto para $CelE2_{1-381}$. Assim como para CelE2, as análises por espalhamento de raio-X de baixo ângulo (SAXS) mostraram que a adição de $CaCl_2$ resultou na diminuição da tendência de agregação proteica para $CelE2_{1-381}$. Os ensaios de avaliação da ligação a polissacarídeos solúveis e insolúveis indicaram que $CelE2_{382-477}$ não apresentou capacidade de ligação para os substratos testados. A estrutura tridimensional do domínio catalítico $CelE2_{1-381}$ foi determinada a uma resolução de 2,1 Å que permitiu a identificar um enovelamento do tipo barril (α/ β)$_8$. Os resultados obtidos mostraram que a presença do domínio acessório Calx-β não é fundamental para atividade enzimática de CelE2, de modo que a alteração mais significativa diz respeito ao aumento da atividade enzimática do domínio catalítico em Avicel em comparação à construção completa. Estes resultados contribuirão para estudos futuros acerca deste domínio pouco estudado entre as celulases e informações para a família 5 das GHs.

**ABSTRACT**

Glycoside hydrolases (GHs) comprise a class of enzymes widely used in the enzymatic processes, acting in the hydrolysis of the polysaccharides glycosidic bonds. In addition to the catalytic domain, several GHs present accessory domains, which may influence the biochemical and structural aspects of the enzymes. The CelE2, a metagenomic-derived cellulase, recently characterized, harbors in its structure an N-terminal catalytic domain characteristic of the GH5 family, and a C-terminal domain with identity for the Calx-$\beta$ domain, previously identified in other glycoside hydrolases but without clearly defined function in these cases. In order to investigate the biochemical and structural characteristics of the CelE2 domains separately, truncated derivations were constructed regarding the catalytic domain and the Calx-$\beta$ accessory domain. The results showed that the deletion of Calx-$\beta$ domain (CelE2$_{382-477}$) did not cause alterations in the biochemical characteristics of CelE2, since the catalytic domain (CelE2$_{1-381}$) presented similar values of enzymatic activity in $\beta$-glucan, carboxymethylcellulose and lichenan, optimal pH and temperature values near 5,3 and 45 °C, respectively, thermal stability at 40 °C and 50 °C up to 360 minutes, and increase of enzymatic activity in the presence of CaCl$_2$. The cleavage pattern and evidence of transglycosylation were also maintained after Calx-$\beta$ deletion from CelE2. However, the catalytic domain showed a 4-fold increase in the enzymatic activity of Avicel hydrolysis, compared to CelE2. Circular dichroism (CD) analysis of CelE2$_{1-381}$ indicated a predominant $\alpha$-helix profile, decrease in the denaturation temperature around 4.6 °C and an increase of these values in the presence of CaCl$_2$, both for the complete construct and for CelE2$_{1-381}$. As well as CelE2, small angle X-ray scattering (SAXS) analyzes showed that the addition of CaCl$_2$ resulted in a decrease in the protein aggregation tendency for CelE2$_{1-381}$. The binding assays for soluble and insoluble polysaccharides indicated that CelE2$_{382-477}$ did not show the binding ability for the substrates tested. The three-dimensional structure of the catalytic domain CelE2$_{1-381}$ was determined at a resolution of 2.1 Å which allowed to identify an $(\alpha/\beta)_8$ barrel fold. The results showed that the presence of the Calx-$\beta$ accessory domain is not fundamental for the enzymatic activity of CelE2, so that the most significant change concerns the increase of the enzymatic activity of the catalytic domain in Avicel, comparing to the whole construct. These results may contribute to future studies about this understudied domain among cellulases and information for the family 5 of glycoside hydrolases.

# SUMÁRIO

## CAPÍTULO 1 – INTRODUÇÃO

## 1. Revisão bibliográfica

### 1.1 Problemática energética

A capacidade energética de um país tem extrema influência sobre o seu desenvolvimento socioeconômico (ARAUJO; NAVARRO; SANTOS, 2013). A economia global cresce em média 3,4% ao ano e a estimativa é que a população mundial supere mais de 9 bilhões em 2040 (IEA, 2017a). O aumento da população mundial está associado ao aumento constante do consumo de energia desde o início do século 19 (SORRELL, 2015; WILSON; GRUBLER; WINTERFELDT, 2011).

No entanto, grande parte da energia mundial é proveniente das fontes fósseis, como o petróleo e o carvão mineral, e está claro que as reservas globais dessas fontes não serão suficientes para atender a demanda energética global, que juntamente com as grandes variações nos preços do mercado proporcionam um cenário instável (ARAUJO; NAVARRO; SANTOS, 2013; DEMAIN, 2009). Além disso, a utilização de combustíveis fósseis promove a liberação de gases de efeito estufa, principalmente de $CO_2$, gerando impactos ao meio ambiente, riscos à saúde e contribuindo para as mudanças climáticas globais. As emissões anuais de $CO_2$ decorrentes da combustão de combustíveis fósseis aumentaram drasticamente desde a revolução industrial (FARHAD; SAFFAR-AVVAL; YOUNESSI-SINAKI, 2008; IEA, 2017e). O setor energético representa dois terços do total de emissões de gases de efeito estufa (80% somente de $CO_2$), sendo que 24% das emissões de $CO_2$ no ano de 2015 foram referentes ao setor de transporte (Figura 1) (IEA, 2017b). Diante a este cenário, iniciativas e mudanças vêm sendo propostas pelos governos em todo o mundo para criar um sistema energético eficiente, limpo, economicamente viável e sustentável para a produção de combustíveis e produtos químicos (IEA, 2017e).
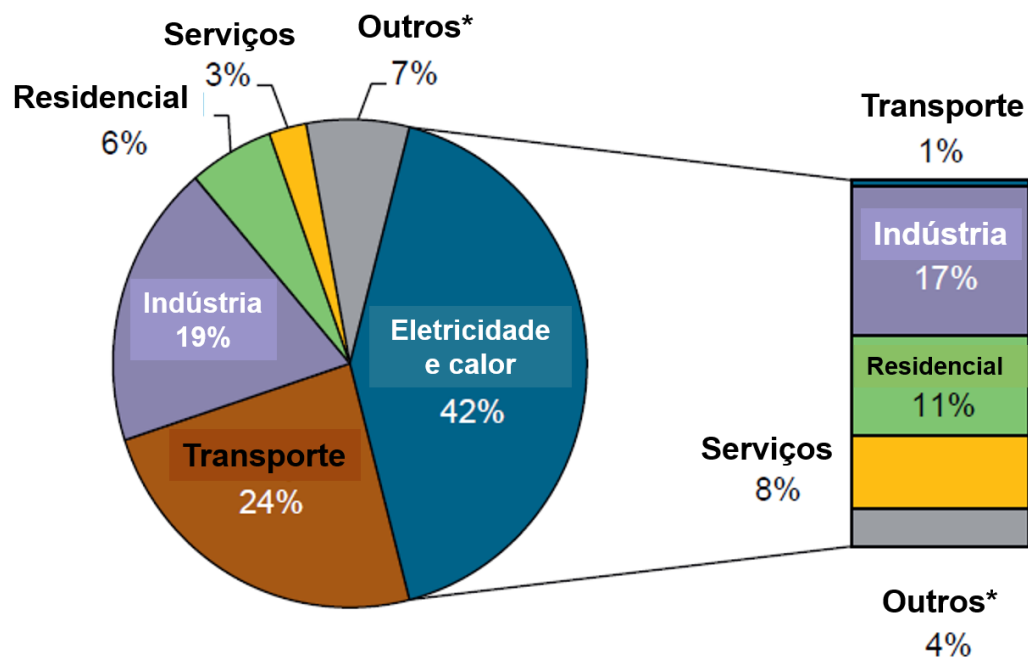
**Figura 1.** Emissões mundiais de $CO_2$ da combustão de combustíveis por setor em 2015 e a distribuição de eletricidade e calor para os setores de uso final. * incluem setores que não sejam a produção de eletricidade e calor, e outros usos não especificados, como agricultura/ silvicultura, pesca e indústrias de energia. Adaptada de IEA, (2017e).

No contexto global, a busca por fontes de energia renováveis e sustentáveis tem sido impulsionada por programas do governo através de subsídios e créditos fiscais para aumentar o custo competitivo com as fontes de energia fósseis. Dentre elas, a energia solar, eólica, geotérmica, hidrelétrica, bioenergia e energia oceânica são consideradas energias renováveis por serem provenientes de fontes naturais (como a luz solar e o vento) e inesgotáveis ou que podem ser reabastecidas em curtos períodos de tempo (BHATTACHARYA et al., 2016). No entanto, como o setor de transporte representa um dos maiores consumidores de petróleo, o interesse global nos biocombustíveis líquidos (principalmente bioetanol e biodiesel) tem se tornado cada vez mais acentuado devido à necessidade de mitigar as emissões de gases de efeito estufa (IEA, 2017c; WALTER et al., 2011).

1.2   Biocombustíveis

Os biocombustíveis são produzidos a partir de fontes renováveis de carbono e podem existir nas formas sólidas (como lenha e lascas de madeira), líquidas (como bioetanol e biodiesel) ou gasosas (como biogás e syngas), no entanto o bioetanol e biodiesel são mundialmente os mais utilizados no setor de transporte (GUO; SONG; BUHAIN, 2015; LEE; LAVOIE, 2013).

A produção de biocombustíveis está associada ao mercado de *commodities* e são geralmente classificados em categorias:

- Biocombustíveis de primeira geração: são produzidos a partir de biomassa com finalidade alimentícia constituída por açúcares, amido, gorduras animais e óleos. Encontram-se já estabelecidos nos mercados de combustíveis e o biodiesel, bioetanol, biogás, bioéteres, syngas e biocombustíveis sólidos, como madeira, serragem, lixo doméstico, carvão, resíduos agrícolas e esterco seco são os mais populares (BHARGAVA; FUENTES, 2018; KIMBLE; PASDELOUP; SPENCER, 2008).
- Biocombustíveis de segunda geração: são produzidos a partir de uma ampla gama de matérias-primas, como biomassa lignocelulósica não comestível e resíduos sólidos urbanos. Incluem a produção de biodiesel, bioetanol, combustíveis sintéticos e bio-hidrogênio, no entanto, ainda não são produzidos em grande escala devido às tecnologias para aprimoramento estarem em fase de desenvolvimento (BHARGAVA; FUENTES, 2018; KIMBLE; PASDELOUP; SPENCER, 2008; LEE; LAVOIE, 2013).
- Biocombustíveis de terceira geração: são produzidos através de biomassa marinha, como biomassa de algas, a partir do acúmulo de lipídeos e carboidratos. Esta biomassa pode ser utilizada para a produção de biodiesel, bioetanol, biogás, bio-hidrogênio e syngas. Contudo, esta tecnologia ainda não é economicamente viável (LEE; LAVOIE, 2013; SINGH; OLSEN, 2011).
- Biocombustíveis de quarta geração: sua produção é baseada na conversão direta da energia solar em combustível através de matérias primas baratas e renováveis para a produção de combustíveis solares fotobiológicos e em eletro-combustíveis. No entanto, esta tecnologia está em estado emergente e de estudos para aumento da produção do conteúdo de lipídeos (BHARGAVA; FUENTES, 2018).

Os biocombustíveis produzidos a partir de biomassa que não competem com alimentos, são chamados de biocombustíveis avançados e estão em estágio de desenvolvimento e aprimoramento (BHARGAVA; FUENTES, 2018).

Com isso, os biocombustíveis são considerados elementos chave para a mitigação das mudanças climáticas por compensar as emissões de gases de efeito estufa e o aquecimento global. Mesmo com a liberação de $CO_2$ na combustão dos biocombustíveis, o balanço positivo de carbono pode ocorrer através da reutilização de $CO_2$ para a produção de biomassa (NANDA et al., 2014). Além do balanço positivo de carbono em relação aos combustíveis fósseis, os biocombustíveis representam grande potencial econômico e sustentável para os países industrializados através da utilização de materiais renováveis locais (LEE; LAVOIE, 2013).

Os Estados Unidos e o Brasil lideram a produção mundial de biocombustíveis convencionais representando mais de 70% da produção mundial em 2016 (IEA, 2017d). Além disso, o Brasil possui vantagem para atender a demanda de biocombustíveis devido à abundância de terras e de recursos hídricos (MARIN et al., 2016).

## 1.3   Etanol

O etanol, bioetanol ou álcool etílico, é um combustível líquido derivado de biomassa, considerado o mais utilizado no mundo com mais de 64 países com programas ativos para o seu uso como combustível principal. Além do grande potencial para substituir os combustíveis fósseis, contribui para a redução das emissões de carbono e de partículas cancerígenas provenientes dos motores a gasolina (BIOENERGY AUSTRALIA, 2016; NANDA et al., 2014).

A Figura 2 mostra o ciclo de produção do etanol que inicia sua vida como carbono armazenado na biomassa, então é convertido em etanol que é queimado como combustível, liberando água e dióxido de carbono durante o processo. A fotossíntese converte o carbono em uma nova biomassa para ser utilizado no próximo ciclo de produção de etanol, evitando acumulação líquida de $CO_2$ na atmosfera (BIOENERGY AUSTRALIA, 2016).
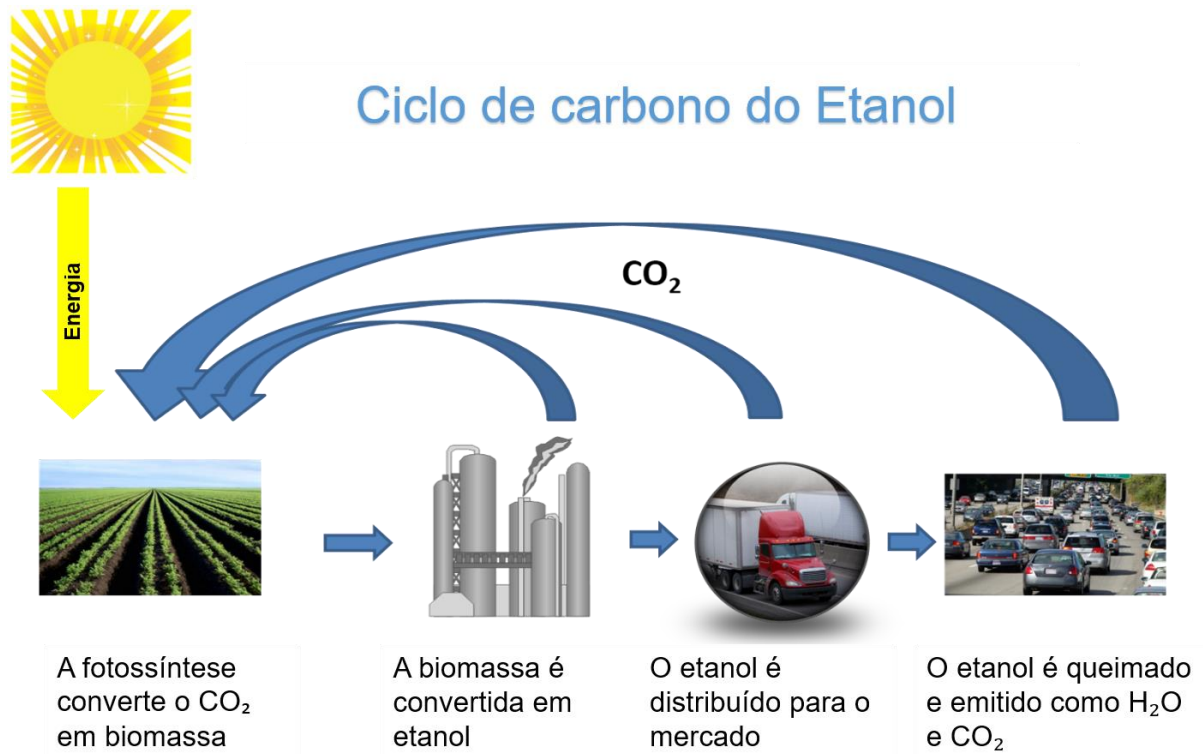
**Figura 2.** Ciclo do carbono no processo de produção até o consumo do etanol. Adaptada de Bioenergy australia (2016).

O etanol pode ser produzido a partir de matérias-primas açucaradas ou amiláceas como cana-de-açúcar, milho, trigo e mandioca (primeira geração), material lignocelulósico como resíduos agrícolas, florestais e industriais e culturas lenhosas (segunda geração) e biomassa de algas (terceira geração) (BAEYENS et al., 2015). O consumo de etanol vem crescendo cerca de 10% ao ano, com os Estados Unidos e o Brasil representando os maiores produtores possuindo quase 85% da produção mundial (Figura 3) (BIOENERGY AUSTRALIA, 2016; MARIN et al., 2016), com a produção principal de etanol de milho e etanol de cana-de-açúcar, respectivamente (WANG et al., 2012).
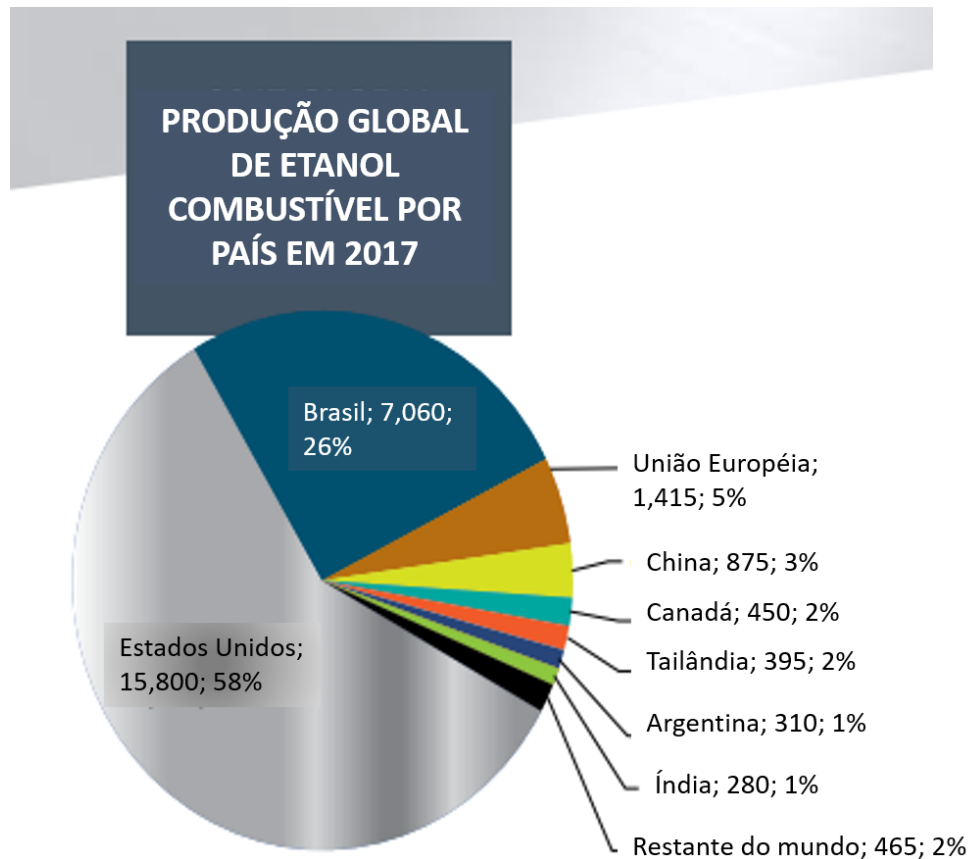
**Figura 3.** Produção global de etanol em 2017 por país, ordenado por: país, milhões de galões e porcentagem da produção global. Adaptada de RFA (2018).

O etanol pode ser misturado com a gasolina ou ser diretamente utilizado. Nos veículos de combustível flexível (flex), os motores são modificados para usar misturas de etanol e gasolina como o E10, E15 e o E85 nos Estados Unidos (máximo de 10, 15 e 85% de etanol anidro, respectivamente) e como o E25 e o E100 no Brasil (máximo de 25% de etanol anidro e o etanol hidratado puro, respectivamente) (BAEYENS et al., 2015; IEA, 2017d).

Para o Brasil, a disponibilidade de terras, clima, experiência a longo prazo, tecnologia comercial existente e o tamanho do mercado doméstico tornam condições favoráveis e de destaque para produção de etanol de cana-açúcar (WALTER et al., 2011). A produção estimada de cana-de-açúcar para a safra de 2018/19 é de 625,96 milhões de toneladas e para a produção de etanol é de 28,16 bilhões de litros (CONAB, 2018). A produção de etanol de primeira geração no Brasil ocorre a partir da fermentação dos açúcares do caldo e do melaço, provenientes da moagem da cana, através de leveduras clássicas ou OGMs (organismos geneticamente modificados) (*Saccharomyces cerevisiae*), seguida pela destilação. Este processo tem como principal subproduto o bagaço que em sua maioria é queimado para ser

aproveitado para a cogeração de energia. Cerca de 90% do bagaço gerado é queimado, e para cada tonelada de bagaço queimado é gerado aproximadamente 25 kg de cinzas (ASSAD, 2017; BASSO; BASSO; ROCHA, 2011; LEE; LAVOIE, 2013).

O bagaço e os resíduos agrícolas podem ser utilizados para a produção de etanol de segunda geração por se tratar de biomassa lignocelulósica que, no caso do bagaço de cana-de-açúcar, representa dois terços da energia produzida pela cana-de-açúcar. Assim, o etanol de segunda geração representa grande potencial de expansão da produção de etanol através do uso de material lignocelulósico que consiste em uma fonte abundante, barata e sustentável que não está diretamente relacionada com a produção de alimentos e que permite maior quantidade de açúcares fermentescíveis produzidos por hectare, evitando a expansão de áreas plantadas (DOS SANTOS et al., 2016; SANTOS; GÓMEZ; BUCKERIDGE, 2011; SRIVASTAVA et al., 2015).

Em relação à gasolina, o etanol de cana-de-açúcar e o etanol lignocelulósico podem reduzir em 78% e 86% as emissões de gases de efeito estufa, respectivamente, e ainda reduzir em 97% o uso de energia fóssil (WANG et al., 2008). A integração com a indústria de etanol de primeira geração representa uma das melhores formas para produção de etanol de segunda geração. No entanto, a produção de etanol a partir de biomassa lignocelulósica enfrenta alguns entraves que o torna ainda inviável comercialmente. As etapas de pré-tratamento, sacarificação enzimática e fermentação são tecnologias em desenvolvimento e apresentam possibilidade para significativos avanços tecnológicos. Com isso, o desenvolvimento de tecnologias para melhorar a eficiência da produção de etanol tem sido o foco de esforços científicos vigorosos (GUPTA; VERMA, 2015; LOSORDO et al., 2016; RODRIGUES MOTA et al., 2018).

Apesar da complexidade tecnológica de produção do etanol de segunda geração, trata-se de um processo que pode se tornar realidade com as primeiras unidades industriais. As usinas em operação atualmente iniciaram suas atividades em 2014, sendo duas delas no Brasil e três nos Estados Unidos. A produção de etanol de segunda geração consiste em uma conquista importante diante a uma era focada em políticas de descarbonização e preocupações ambientais (DOS SANTOS et al., 2016).

## 1.4    Biomassa lignocelulósica

A biomassa lignocelulósica (como resíduos agrícolas, culturas lenhosas, culturas herbáceas) é considerada a matéria prima biológica mais abundante do planeta, composta por moléculas altamente energéticas (SANTOS; GÓMEZ; BUCKERIDGE, 2011; YANG et al., 2011). Esta fonte de energia alternativa pode ser empregada para a produção de bioquímicos e biocombustíveis devido principalmente ao baixo custo, abundância e por constituir fonte de baixa emissão de carbono. Esta crescente demanda por biomassa representa possibilidade de redução da dependência de importação de fonte não-renováveis e aumento do emprego rural (BHATTACHARYA et al., 2016; ZHANG; PEI; WANG, 2016).

A parede celular vegetal é o principal componente da biomassa vegetal e consiste em uma estrutura altamente ordenada composta por uma malha de polissacarídeos, proteínas estruturais e compostos fenólicos que conferem proteção celular, resistência mecânica e bioquímica aos tecidos vegetais. O seu conteúdo de polissacarídeos pode variar entre as espécies, fase do desenvolvimento, tipo e tecido vegetal bem como condições ambientais. A parede celular primária é formada por células em crescimento, no qual os principais polissacarídeos encontrados são celulose, hemicelulose e pectina. E em algumas células, após cessar o desenvolvimento, há a formação da parede celular secundária com a deposição de lignina (GUERRIERO et al., 2016; TAIZ; ZEIGER, 2006).

No geral, a biomassa lignocelulósica é composta principalmente por celulose, hemicelulose e lignina (Figura 4) em uma rede mantida por ligações covalentes, pontes intermoleculares e forças de Van der Waals que a torna recalcitrante ao ataque enzimático para a liberação de açúcares fermentescíveis (NANDA et al., 2014; SINDHU; BINOD; PANDEY, 2016).
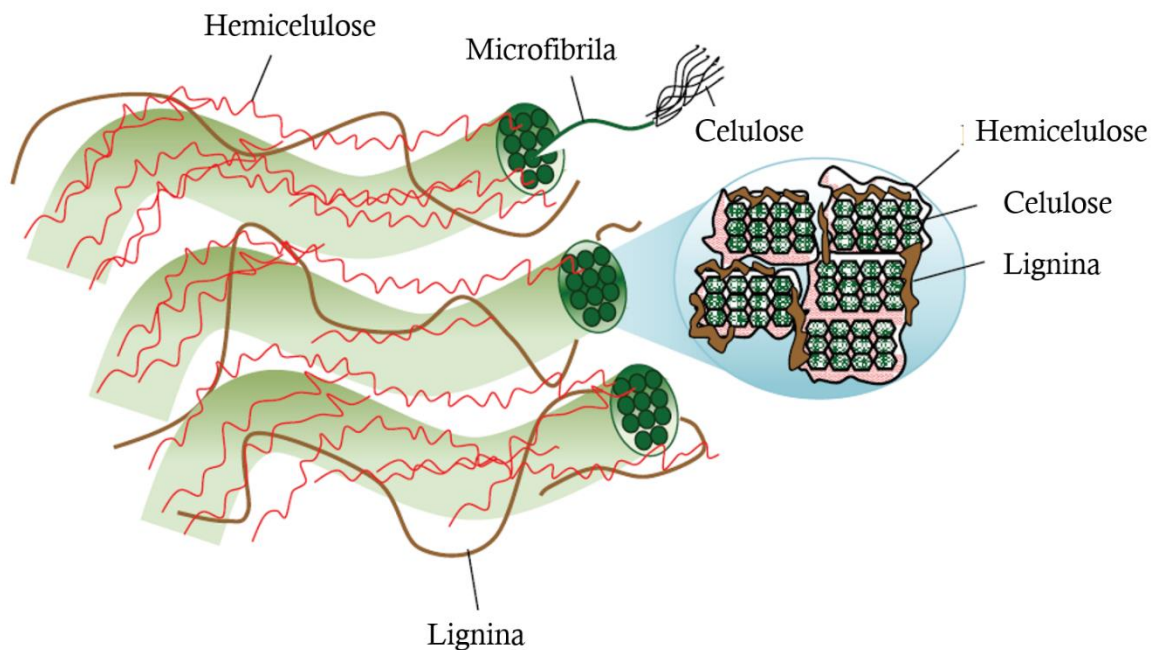
**Figura 4.** Representação esquemática da arquitetura da parede celular da biomassa lignocelulósica com seus principais componentes. Adaptada de Lee; Hamid; Zain (2014).

A celulose é o polímero mais abundante do planeta e representa cerca de 40-60% do total de massa seca da biomassa lignocelulósica (HAMELINCK; HOOIJDONK; FAAIJ, 2005). Constitui um homopolissacarídeo (não ramificado) formado por uma longa cadeia de D-glicose as quais são unidas por ligações glicosídicas β-1,4, formando as microfibrilas que se agregam por ligações de hidrogênio e forças de Van der Waals. Esta conformação cria estruturas lineares organizadas e altamente compactadas formando regiões cristalinas e regiões com menor grau de organização chamadas de amorfas. O resultado é uma estrutura resistente e insolúvel em água (Figura 5) (ABDUL KHALIL; BHAT; IREANA YUSRA, 2012; LEE; HAMID; ZAIN, 2014; PÉREZ et al., 2002; RUBIN, 2008). A celulose atrai a atenção das indústrias por constituir um recurso renovável que pode ser convertido em produtos de base biológica de interesses comerciais (SHAHZADI et al., 2014).
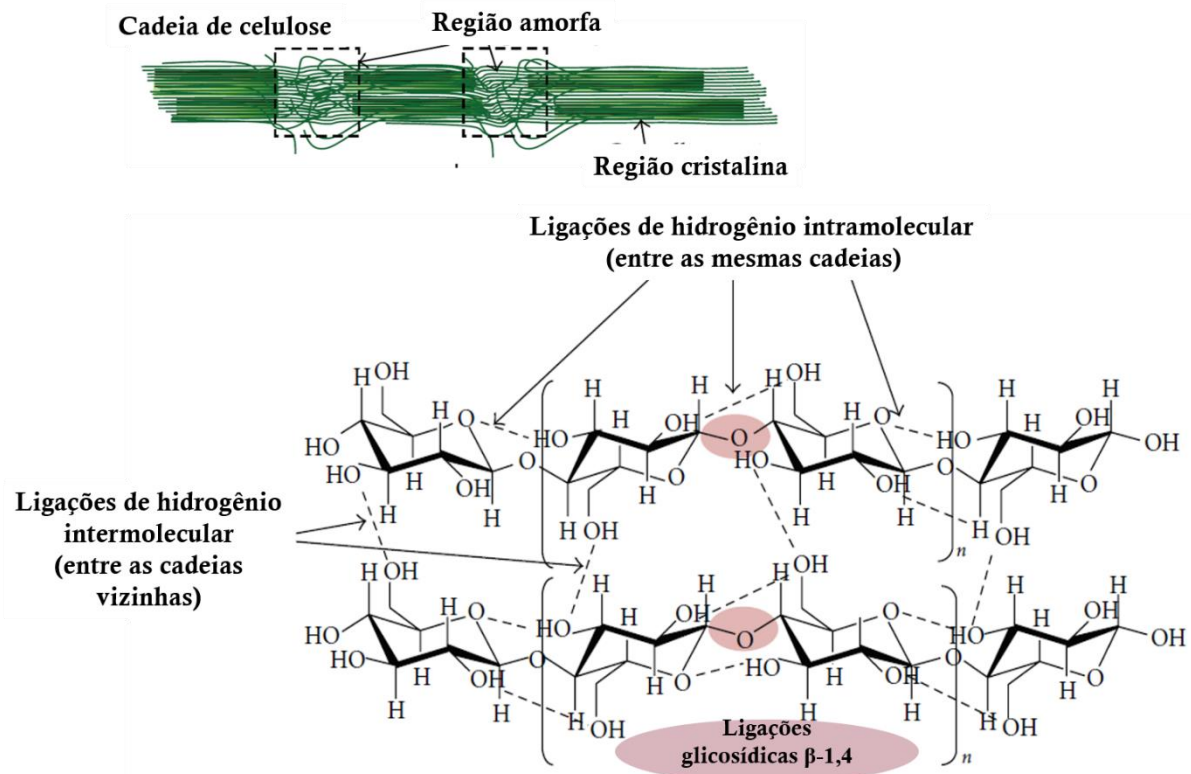
**Figura 5.** Microfibrilas de celulose com regiões cristalinas e amorfas distribuídas aleatoriamente ao longo de seu comprimento e estrutura química da cadeia de celulose. Adaptado de Lee; Hamid; Zain (2014).

A hemicelulose é um heteropolímero que constitui cerca de 20-40% do total de massa seca da biomassa lignocelulósica (HAMELINCK; HOOIJDONK; FAAIJ, 2005). É composta por ramificações laterais curtas que podem ser formadas por monômeros de pentoses (β-D-xilose, α-L-arabinose), de hexoses (β-D-manose, β-D-glicose, α-D-galactose) e/ou ácidos urônicos (α-D-glucuronico, α-D-4-O- metilgalacturonico, α-D-galacturonico) unidos por ligações β-1,3 e β-1,4 (Figura 6) (GÍRIO et al., 2010; SÁNCHEZ, 2009). A hemicelulose pode ser classificada como xilogucanos, xilanos, mananas, β-(1→3, 1→4) glucanos, galactanos, arabinanas e arabinogalactanos, dependendo do monômero constituinte (SCHELLER; ULVSKOV, 2010). O xilano é o tipo mais comum de polímero encontrado na família das hemiceluloses, o qual forma uma cadeia de polímero ramificado formado principalmente por monômeros de açúcar de cinco carbonos, xilose, unidos por ligações β-1,4 (HARMSEN et al., 2010). Além disso, os açúcares obtidos a partir da hemicelulose também podem ser convertidos em etanol (CARDONA; SÁNCHEZ, 2007).
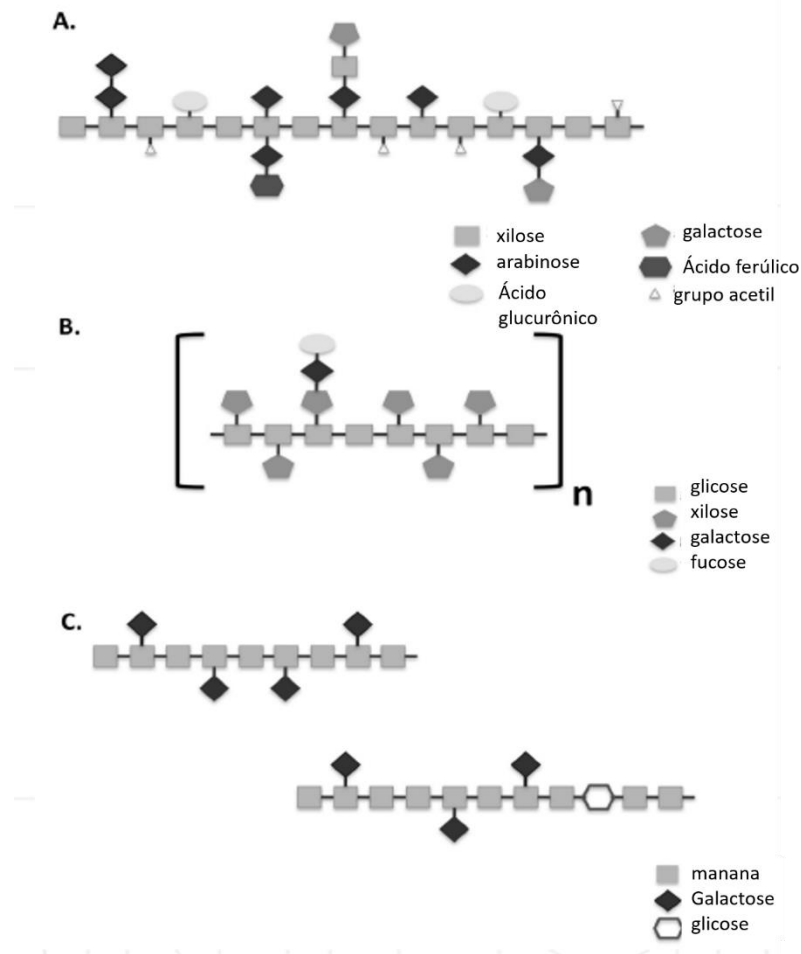
**Figura 6.** Representação esquemática das três estruturas mais comuns de hemicelulose. A: xilano. B: xiloglucano. C: galactomanana (superior) e galactoglucomana (inferior). Adaptada de De Souza (2013).

A lignina é um polímero complexo que compreende cerca de 10-25% do total de massa seca da biomassa lignocelulósica (HAMELINCK; HOOIJDONK; FAAIJ, 2005). Este heteropolímero tridimensional amorfo é sintetizado a partir dos precursores monoméricos álcool p-cumarílico, álcool coniferílico e álcool sinapílico (Figura 7), que originam as unidades de p-hidroxifenila, guaiacila e siringila, respectivamente, os quais são unidas por ligações do tipo éter ou éster e carbono-carbono. A lignina envolve a celulose e hemicelulose, formando uma barreira física na parede celular (PÉREZ et al., 2002; SÁNCHEZ, 2009). Sua presença confere suporte estrutural, impermeabilidade e resistência contra ataque microbiano e estresse oxidativo. No entanto, a lignina pode ser utilizada como combustível ou como plataforma para produtos químicos de valor agregado como ácidos orgânicos, fenóis e vanilina (HAMELINCK; HOOIJDONK; FAAIJ, 2005; VOLYNETS; EIN-MOZAFFARI; DAHMAN, 2017).
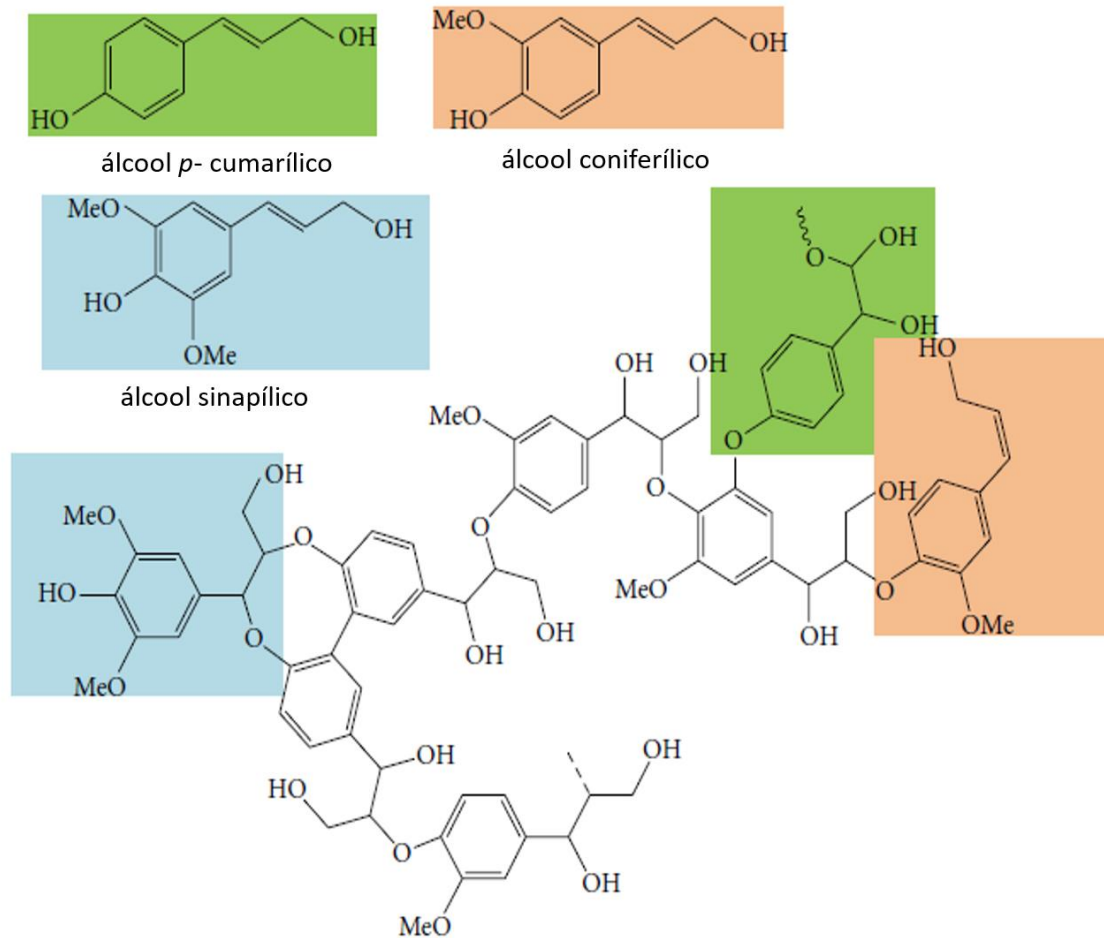
**Figura 7.** Componentes estruturais do polímero de lignina com os precursores sintéticos álcool p-cumarílico, álcool coniferílico e álcool sinapílico. Adaptado de Lee; Hamid; Zain (2014).

A recalcitrância da biomassa lignocelulósica à degradação é um grande obstáculo na separação dos seus polissacarídeos para as diferentes aplicações (LEE; HAMID; ZAIN, 2014). Para a sua conversão é preciso a deslignificação para a liberação da celulose e da hemicelulose, despolimerização dos carboidratos para produzir açúcares livres, e a fermentação em etanol. Para isso, são necessárias as etapas de pré-tratamento (como mecânico, termoquímico e biológico), sacarificação enzimática e a fermentação dos açúcares por leveduras e bactérias. No entanto são necessárias novas tecnologias para aumentar a eficiência e diminuir os custos destas etapas (ADITIYA et al., 2016; NANDA et al., 2014; VOLYNETS; EIN-MOZAFFARI; DAHMAN, 2017).

A etapa de sacarificação enzimática apresenta muitos desafios quanto à eficiência, dosagem, tempo de hidrólise e configuração do processo. Além disso, este processo de conversão das cadeias de carboidratos representa cerca de 20% dos custos da produção de

etanol. Diversos coquetéis enzimáticos foram desenvolvidos e estão sendo aprimorados para hidrolisar celulose e hemicelulose em açúcares monoméricos. Com isso, numerosos esforços estão sendo feitos para reduzir os custos da produção de enzimas (DOS SANTOS et al., 2016; GUPTA; VERMA, 2015; PHITSUWAN et al., 2013).

## 1.5 Enzimas ativas em carboidratos e a sacarificação da biomassa vegetal

A desconstrução dos polissacarídeos da parede celular vegetal envolve uma gama de enzimas denominadas de Enzimas Ativas em Carboidratos (CAZymes), as quais são reunidas no banco de dados do CAZy (Carbohydrate Active Enzymes database - http://www.cazy.org), classificadas inicialmente com base na comparação da sequência de aminoácidos. As famílias das glicosil hidrolases (GHs), glicosiltransferases (GTs), polissacarídeo liases (PLs), carboidrato esterases (CEs), atividade auxiliares (AAs) e dos módulos de ligação a carboidratos (CBMs) encontradas no CAZy, promovem respectivamente: hidrólise/rearranjo, formação, hidrólise não glicosídica, hidrólise de ésteres, atividade oxidativa, e adesão aos carboidratos (LOMBARD et al., 2014).

As GHs constituem uma classe de enzimas com diferentes mecanismos de ação, divididas pelo CAZy em clãs pela similaridade da estrutura terciária, resíduos catalíticos e mecanismo de ação. Os clãs são divididos em famílias pela similaridade da sequência de aminoácidos; e algumas famílias são divididas em subfamílias com base na análise filogenética (HENRISSAT; BAIROCH, 1996; LOMBARD et al., 2014). O mecanismo de hidrólise conservado nas GHs envolve a inversão ou retenção do carbono anomérico do substrato, e a ampla diversidade funcional das famílias de GHs é considerada de extrema importância para a desconstrução da biomassa lignocelulósica (ABOT et al., 2016; VAN WYK et al., 2017).

As GHs são conhecidas por possuírem arquitetura modular, e algumas funções podem não estar diretamente associadas ao sítio catalítico, mas sim aos módulos adjacentes. Enzimas modulares são constituídas frequentemente por ao menos um domínio catalítico e um ou mais domínios acessórios que podem desempenhar diferentes funções (BORASTON et al., 2004). Dentre os domínios acessórios, os domínios Calx-β podem ser encontrados em várias famílias de GHs, no entanto a sua função para estes casos não é completemente elucidada.

Em virtude da ampla variedade de GHs em diversos organismos e da importância dos estudos de domínios acessórios, a caracterização das enzimas e seus módulos individuais é

fundamental para o melhor entendimento acerca dos mecanismos de degradação e modos de atuação para melhoramento e aplicações biotecnológicas de enzimas.

As celulases são enzimas distribuídas entre algumas famílias de GHs (1-3, 5-12, 26, 30, 39, 44, 45, 48, 51, 74, 116, 124), consideradas importantes pelo alto potencial de aplicação em diferentes indústrias e por constituírem elemento chave para a degradação da celulose. Estas enzimas, promovem a hidrólise glicosídica das ligações β-1,4 da cadeia de celulose e constituem um sistema composto por pelo menos três tipos de atividade enzimática: endoglucanases, exoglucanases ou celobiohidrolases e as β-glicosidades ou celobiases. As endoglucanases (EC 3.2.1.4) clivam randomicamente as ligações internas da cadeia liberando extremidades redutoras e não redutoras; as exoglucanases clivam as extremidades redutoras (EC 3.2.1.176) e não redutoras (EC 3.2.1.91) da cadeia liberando celobiose; as β-glicosidades (EC 3.2.1.21) promovem a conversão da celobiose em glicose a partir da extremidade não redutora dos oligossacarídeos (Figura 8) (SHAHZADI et al., 2014; SHARADA et al., 2014; SILVA, 2016). A complexidade desse sistema enzimático e o alto potencial biotecnológico, fazem das celulases uma das enzimas mais comercialmente disponíveis para diferentes aplicações industriais como papel e celulose, têxteis, lavanderia e detergente (KUMAR et al., 2014; SHARMA et al., 2016). Dentre as famílias que contemplam as celulases, a família 5 das GHs foi a primeira descrita e constitui uma das maiores famílias com alta variedade de especificidade e ampla distribuição entre os organismos (ASPEBORG et al., 2012).

Além disso, as LPMOs (*Lytic Polysaccharide Monooxygenases*) e os CBMs (*Carbohydrate-Binding-Modules*) atuam em sinergismo com as celulases, auxiliado na conversão da celulose (Figure 8). As LPMOs constituem uma classe de metaloenzimas oxidativas presentes em cinco famílias de enzimas de atividade auxiliar (AA) (AA9, AA10, AA11, AA13 e AA14) (LEVASSEUR et al., 2013), que podem atuar na modificação da fibra de celulose diminuindo a recalcitrância das regiões cristalinas e tornando o substrato mais acessível para as celulases convencionais (VILLARES et al., 2017). O mecanismo de quebra da celulose pelas LPMOs ocorre por meio do cobre reduzido no sítio ativo, do oxigênio molecular e de um agente redutor, que promovem a oxidação do anel de piranose na posição C1 (produzindo ácidos aldônicos) ou C4 (produzindo 4-cetoaldose), podendo gerar extremidade de cadeia oxidadas e não oxidadas (CANNELLA et al., 2016; PAYNE et al., 2015). Já os CBMs, são considerados domínios estruturais que normalmente são encontrados conectados às CAZymes, os quais desempenham como principais papéis o reconhecimento e ligação à diferentes substratos promovendo diferentes funções como o aumento da eficiência

catalítica em substratos insolúveis, ruptura da estrutura polissacarídica e ancoragem da proteína (GUILLÉN; SÁNCHEZ; RODRÍGUEZ-SANOJA, 2010). São reunidos em 83 famílias no banco de dados CAZy, classificados com base na similaridade da sequência primária (<http://www.cazy.org>) e são agrupados de acordo com a estrutura e especificidade de ligação em três tipos: Tipo A (se ligam a polissacarídeos cristalinos), tipo B (se ligam internamente às cadeias de glucanos - endo) e tipo C (se ligam às extremidades terminais dos glucanos - exo) (GILBERT; KNOX; BORASTON, 2013). Em muitos casos, a ligação das celulases à celulose ocorre através dos CBMs, e por isso, as interações entre CBM-celulase têm sido amplamente estudadas (VÁRNAI et al., 2014).
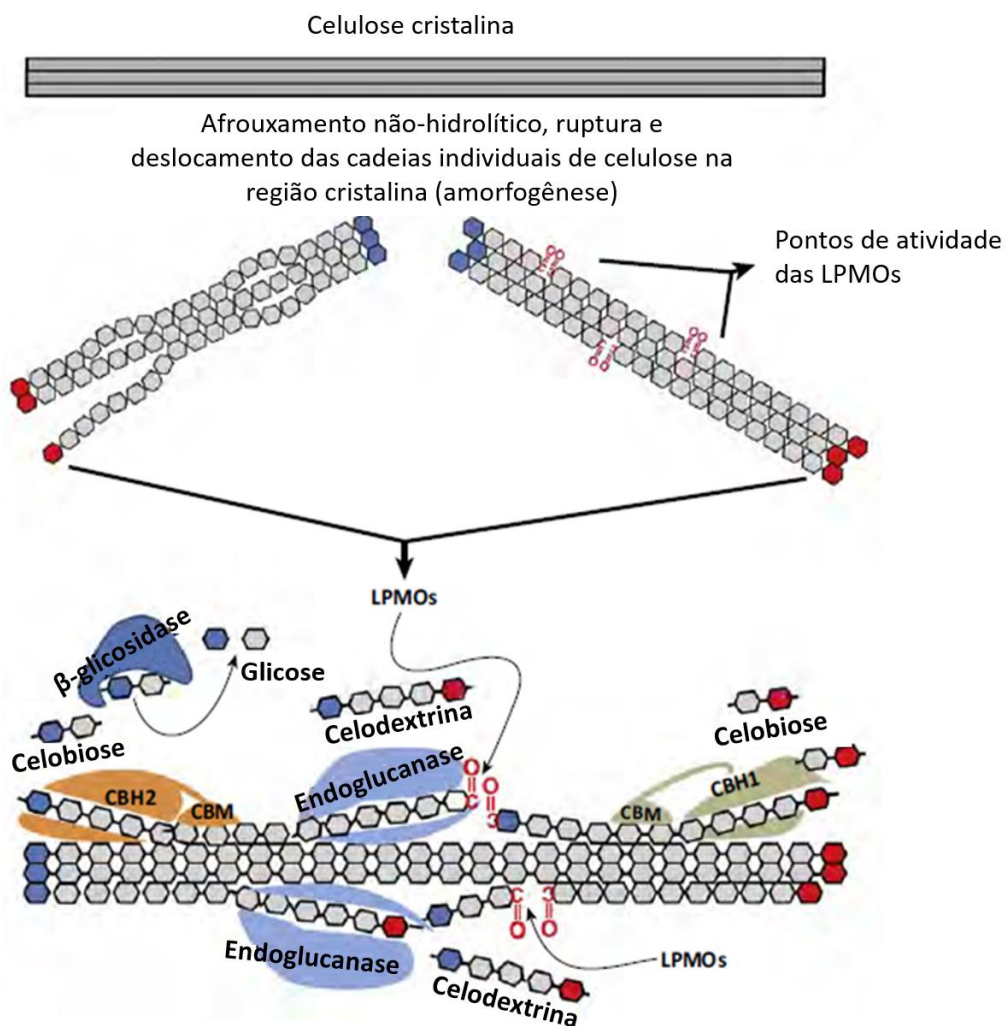


**Figura 8.** Representação esquemática da degradação da celulose com a clivagem oxidativa e hidrolítica. As LPMOs auxiliam na degradação da celulose atuando de modo oxidativo com a ruptura ou afrouxamento de regiões da celulose cristalina. As celulases clássicas (endoglucanases, celobiohidrolases e β-glicosidases) atuam em sinergismo na degradação da celulose com a clivagem aleatória, o ataque nas extremidades da cadeia e a liberação da glicose.

Os CBMs auxiliam as celulases com o direcionamento e ligação na celulose. Adaptada de Gupta *et al.* (2016) .

A degradação da hemicelulose envolve uma série de enzimas que dependem do tipo de hemicelulose a ser degradada, envolvendo endo-enzimas que atuam na clivagem interna da cadeia principal, exo-enzimas que liberam monômeros, e enzimas auxiliares que clivam as cadeias laterais liberando mono e/ou dissacarídeos. A maioria delas são GHs, porém existe as carboidrato esterases que atuam nas ligações ésteres. As xilanases são as principais enzimas que atuam na degradação da hemicelulose, com as endo-xilanases e as β-xilosidases atuando nas cadeias principais, e as α-glucuronidases e as acetil xilano esterases que podem contribuir com a degradação das cadeias laterais do xilano (ARO; PAKULA; PENTTILÄ, 2005; GÍRIO et al., 2010; QUIROZ-CASTAÑEDA; FOLCH-MALLOL, 2011; SÁNCHEZ, 2009).

A degradação da lignina ocorre através de enzimas oxidativas e não-específicas como a manganês peroxidase, lignina peroxidase com reações dependentes de $H_2O_2$, e as lacases que oxidam compostos fenólicos e reduzem o oxigênio molecular. Há também as enzimas que geram peróxidos essenciais para as peroxidases como a glioxal oxidase e a glicose-2-oxidase (ARO; PAKULA; PENTTILÄ, 2005).

As enzimas hidrolíticas possuem grande importância para muitos processos industriais como a fabricação alimentos, medicamentos, algodão para tecidos, e biocombustíveis. Com isso, a descoberta de novas linhagens de microrganismos, novos processos e tecnologias são essenciais para a descoberta de novas enzimas, redução de custos e aumento da eficiência catalítica (MANISHA; YADAV, 2017).

Os microrganismos constituem a principal fonte de prospecção de enzimas que atuam em diferentes carboidratos. Estes organismos oferecem grande potencial como fonte de enzimas com alta especificidade e estabilidade térmica, as quais são características de grande importância para a diminuição de custos e aumento de rendimento de processos biotecnológicos (YANG et al., 2011). No entanto, cerca de 99% dos microrganismos ainda permanecem inexplorados por não serem passíveis de isolamento e cultivo por técnicas tradicionais (XU, 2010).

Com isso, a metagenômica, uma técnica baseada em abordagens independentes de cultivo, tem possibilitado a identificação de novos genes codificantes de enzimas de interesse biotecnológico a partir de comunidades microbianas ou de microrganismos não cultiváveis (JÚNIOR et al., 2017; MANISHA; YADAV, 2017). O emprego da metagenômica permitiu a identificação e caracterização de diversas enzimas, novos genes, produtos naturais, compostos

bioativos, e bioprocessos a partir de diferentes fontes ambientais como solo, rúmen, solo de adubo e água, por exemplo. Algumas enzimas que foram obtidas por metagenômica como as celulases, lipases, xilanases, amilases e proteases são de grande importância para a indústria (NAZIR, 2016).

A prospecção de novos genes pode ser realizada a partir de suas abordagens: análise funcional ou por sequência. A abordagem orientada por função requer a fragmentação do DNA metagenômico e ligação em vetores, e a transformação em uma célula hospedeira para a criação da biblioteca metagenômica, e com isso, os clones são selecionados a partir de funções novas ou conhecidas, permitindo a identificação de novos genes. As abordagens guiadas por sequências utilizam de iniciadora de PCR e sondas de hibridização que derivam de regiões conservadas de genes conhecidos para identificar genes conhecidos e disponíveis nos bancos de dados (HANDELSMAN, 2004; NAZIR, 2016; RIESENFELD; SCHLOSS; HANDELSMAN, 2004).

## MOTIVAÇÃO E OBJETIVOS

A endoglucanase CelE2 foi obtida a partir de uma triagem funcional de uma biblioteca metagenômica de solo (ALVAREZ, 2013). A comparação da sequência de aminoácidos predita para a celulase CelE2 com o banco de dados do NCBI (NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION) através da ferramenta blastp (que permite a comparação com o CDD, banco de dados de domínios conservados) (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>), resultou na identificação de dois domínios conservados, sendo um deles que se estende do aminoácido 29 ao 348, apresentou identidade com o domínio conservado da família 5 de glicosil hidrolases (GH5) e o outro, que se estende do aminoácido 384 ao 454, com o domínio Calx-β.

Segundo a definição do Interpro, (<http://www.ebi.ac.uk/interpro/>), os motivos proteicos Calx-β estão presentes nos domínios citoplasmáticos Calx trocadores de sódio e cálcio, que são utilizados para a expulsão de cálcio das células, e que também se sobrepõem a domínios de ligação a cálcio e regulação. Uma busca no Uniprot (realizada em 03 de julho de 2017) relacionando os termos "Calx-beta" e "Glycoside Hydrolase" revelou 350 entradas, e com os termos "Calx-beta" e "Glycoside Hydrolase Family 5" revelou 7 entradas (FINN et al., 2017). Apesar de tratar de um domínio que ainda não possui função definida, as buscas revelam

que se trata de um domínio proteico já identificado em outras GHs. Dois trabalhos (HAYASE et al., 2008; MAEKAWA et al., 2006) buscaram elucidar a função dos domínios Calx-β na β-glucosidase (GH3), BglM1, de *Physarum polycephalum*. Em resumo estes trabalhos sugerem que os domínios Calx-β devem atuar na modulação da atividade da β-glucosidase em resposta a diferentes concentrações de cálcio bem como serem importantes para a manutenção da estrutura terciária da proteína.

A comparação da CelE2 com o banco de dados do CAZy, que reúne informações relativas aos domínios característicos das famílias de enzimas ativas em carboidratos (CAZymes) através da ferramenta dbCan (YIN et al., 2012), identificou dois domínios, sendo um deles, do aminoácido 27 ao 354, com identidade com o domínio conservado da família GH5, enquanto que o segundo domínio, do aminoácido 374 ao 451, apresentou identidade, embora com valor de Expect (E) value (0,0005) maior do que o recomendado (valor recomendado: E-value $< 1e^{-18}$), com a família 35 de domínios de ligação a carboidratos (CBM35). Um fator interessante foi observado em relação à família CBM35: em 2004, Bolam *et al.* descreveu a nova família de CBMs, denominada CBM35, com membros apresentando capacidade de ligação, por exemplo, a galactomanana, glucomanana e xilana de aveia, sendo que um dos representantes desta nova família é um dos primeiros a apresentar interação dependente de íon metálico divalente, como por exemplo, o cálcio.

Sendo assim, o objetivo deste trabalho consistiu em avaliar em detalhes a função dos domínios individuais da celulase CelE2, através de estudos bioquímicos e biofísicos do domínio catalítico e do domínio acessório Calx-β.

## CAPÍTULO 2 – DOCUMENTOS PUBLICADOS E/OU EM PREPARAÇÃO

Este capítulo apresenta dois documentos científicos relacionados ao projeto de mestrado, referentes à experimentos elaborados que buscaram elucidar as propriedades bioquímicas e biofísicas dos domínios individualizados da celulase CelE2 e a relação com o funcionamento enzimático. O primeiro documento trata-se do trabalho inicial, já publicado, que deu origem aos questionamentos e hipóteses levantados sobre o papel de cada domínio para CelE2. Este documento traz os resultados da caracterização bioquímica e de propriedades estruturais da CelE2 completa, os quais além de contribuir para os estudos acerca das GHs, mostrou a presença de um domínio acessório não-convencional e pouco estudado que poderia estar envolvido com ativação e estabilidade por cálcio ou promover outros efeitos sobre a atividade enzimática da enzima. O segundo documento constitui em um artigo científico em elaboração contendo os resultados referentes às análises dos domínios individuais de CelE2. O documento traz as propriedades bioquímicas referentes ao domínio catalítico, ensaios de ligação a carboidratos do domínio Calx-β, análise filogenética, análises estruturais por dicroísmo circular, análises por SAXS (colaboração com o grupo de pesquisa do professor Dr. Mario de Oliveira Neto - Botucatu), e resolução da estrutura tridimensional do domínio catalítico (colaboração com o Dr. Marcelo V. Liberato).

# DOCUMENTO 1 – BIOCHEMICAL AND BIOPHYSICAL PROPERTIES OF A METAGENOME-DERIVED GH5 ENDOGLUCANASE DISPLAYING AN UNCONVENTIONAL DOMAIN ARCHITECTURE

**Título: Propriedades bioquímicas e biofísicas de uma endoglucanase GH5 derivada de metagenômica apresentando arquitetura de domínios não-convencional**

**Resumo do documento:** As endoglucanases constituem enzimas chave na degradação da celulose que é o polímero mais abundante da Terra. O objetivo deste trabalho consistiu em realizar a caracterização bioquímica e biofísica de CelE2, uma endoglucanase derivada de metagenômica de solo. CelE2 possui um domínio conservado da família 5 das glicosil hidrolases e um domínio C-terminal com identidade para os domínios Calx-beta. A proteína recombinante CelE2 mostrou preferência por hidrólise de β-glucano, seguido por liquenana e carboximetilcelulose. Valores ótimos de atividade enzimática foram verificados a 45 ºC e pH 5,3, além de considerável termoestabilidade a 40ºC por 360 minutos. Em relação ao perfil de clivagem em polissacarídeos, a liberação de oligossacarídeos com grau de polimerização variados indicou característica de atividade de endoglucanase. Além disso, a análise dos produtos gerados a partir da hidrólise de celo-oligossacarídeos sugeriu que CelE2 exibiu atividade de transglicosilação. Interessantemente, a presença de $CaCl_2$ afetou positivamente CelE2, incluindo na presença de surfactantes. Os experimentos de espalhamento de raio-X de baixo ângulo (SAXS) revelaram informações importantes do efeito de $CaCl_2$ sobre a estabilidade, e os modelos de resíduos e de corpo rígido foram gerados. Para melhor entendimento, esta é a primeira caracterização bioquímica e biofísica de um endoglucanase da família 5 que exibe essa organização modular não-convencional.

# BIOCHEMICAL AND BIOPHYSICAL PROPERTIES OF A METAGENOME-DERIVED GH5 ENDOGLUCANASE DISPLAYING AN UNCONVENTIONAL DOMAIN ARCHITECTURE

Agnes C. Pimentel[a,b], Gabriela C. G. Ematsu[a], Marcelo V. Liberato[a], Douglas A. A. Paixão[a,b], João Paulo L. Franco Cairo[a,b], Fernanda Mandelli[a], Robson Tramontina[a,c], César A. Gandin[d], Mario de Oliveira Neto[d], Fabio M. Squina[a*], Thabata M. Alvarez[a,1*]

[a] Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Caixa Postal 6192, CEP 13083-970, Campinas, São Paulo, Brasil.

[b] Departamento de Bioquímica, Instituto de Biologia (IB), Universidade Estadual de Campinas (UNICAMP), R. Monteiro Lobato, 255 - Cidade Universitária, Campinas, São Paulo, Brasil

[c] Programa de Pós-Graduação em Biociências e Tecnologia de Produtos Bioativos (BTPB), Instituto de Biologia (IB) - CP 6109, Universidade Estadual de Campinas (UNICAMP) - 13083-970, Campinas, SP, Brasil.

[d] Departamento de Física e Biofísica, Instituto de Biociências, Universidade Estadual Paulista (UNESP), Distrito de Rubião Jr. s/n, Botucatu, SP, Brasil

*Corresponding authors. Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Caixa Postal 6192, CEP 13083-970, Campinas, São Paulo, Brasil and Universidade Positivo, Master in Industrial Biotechnology, R. Prof. Pedro Viriato Parigot de Souza, 5300, Cidade Industrial, 81280-330, Curitiba, Paraná, Brasil.

Email addresses: agnescristinap@gmail.com, gabriela.ematsu@bioetanol.org.br, marcelovliberato@gmail.com, douglas.paixao@bioetanol.org.br, jpcairo@gmail.com, fernanda.mandelli@bioetanol.org.br, robson.tramontina@gmail.com, cesargandin@gmail.com, mario.neto@ibb.unesp.br, fmsquina@gmail.com, thabata.alvarez@up.edu.br

**Abstract**

Endoglucanases are key enzymes in the degradation of cellulose, the most abundant polymer on Earth. The aim of this work was to perform the biochemical and biophysical characterization of CelE2, a soil metagenome derived endoglucanase. CelE2 harbors a conserved domain from glycoside hydrolase family 5 (GH5) and a C-terminal domain with identity to Calx-beta domains. The recombinant CelE2 displayed preference for hydrolysis of oat beta-glucan, followed by lichenan and carboxymethyl cellulose. Optimum values of enzymatic activity were observed at 45ºC and pH 5.3, and CelE2 exhibited considerable thermal stability at 40ºC for up to 360 minutes. Regarding the cleavage pattern on polysaccharides, the release of oligosaccharides with a wide degree of polymerization indicated a characteristic of

endoglucanase activity. Furthermore, the analysis of products generated from the cleavage of cellooligosaccharides suggested that CelE2 exhibited transglycosylation activity. Interestingly, the presence of $CaCl_2$ positively affect CelE2, including in the presence of surfactants. SAXS experiments provided key information on the effect of $CaCl_2$ on the stability of CelE2 and dummy residues and rigid-body models were generated. To the best of our knowledge this is the first biochemical and biophysical characterization of an endoglucanase from family GH5 displaying this unconventional modular organization.

**Key words**

Lignocellulosic biomass

Glycoside hydrolase family 5

Calx-beta domain

## 1    Introduction

The main drivers for the development of renewable energy sources are the high dependence on nonrenewable energy sources coupled with their price fluctuation and environmental concerns regarding the emission of greenhouse gases from burning of fossil fuels (DWIVEDI; ALAVALAPATI; LAL, 2009). In this context, lignocellulosic biomass represents a promising renewable source for the production of biofuels (ALONSO; BOND; DUMESIC, 2010; HIMMEL et al., 2007).

Lignocellulosic biomass is typically composed of 40-60% cellulose, 20-40% hemicellulose, 10-25% lignin and minor amounts of ash, acids and extractives (HAMELINCK; HOOIJDONK; FAAIJ, 2005). The production of biofuel from lignocellulosic biomass involves deconstruction of the plant cell wall and depolymerization of cellulose and hemicellulose into monomeric sugars through a combination of physicochemical pre-treatment and enzymatic hydrolysis (HAMELINCK; HOOIJDONK; FAAIJ, 2005; HENDRIKS; ZEEMAN, 2009). Glycoside hydrolase (GH) is a class of enzymes with fundamental importance for the hydrolysis of cellulose and hemicelluloses. Since the degradation of these components is part of many biological reactions in nature, a variety of organisms are able to produce enzymes with versatile specificities and biochemical properties (FRANCO CAIRO et al., 2013; SEGATO et al., 2014; YANG et al., 2011).

The classical scheme for hydrolytic cellulose degradation involves at least three types of enzyme activities working synergistically: (i) endo-1,4-β-glucanases, which randomly cleave internal bonds of the cellulose chain, (ii) exo-1,4-β-glucanases, which hydrolyze the external bonds of the reducing or non-reducing end of the chain and (iii) β-glucosidases, that

cleave cellobiose molecules to glucose (HORN et al., 2012). Recently, lytic polysaccharide monooxygenases (LPMOs) have been recognized as important players in the improvement of cellulose degradation by acting in synergy with cellulases (CORRÊA; DOS SANTOS; PEREIRA, 2016; HORN et al., 2012).

Representatives of endoglucanases are distributed in a variety of GH families accordingly to the CAZy (Carbohydrate Active Enzymes) database (LOMBARD et al., 2014). Of these, GH5 is one of the largest family in terms of number of members, taxonomic dispersion and substrate specificity, and has been the target of several studies, since it contemplates important activities for deconstruction of carbohydrates from the plant cell (ASPEBORG et al., 2012; CHEN et al., 2012).

Commonly, enzymes involved in plant cell wall degradation display multi-modular organization, which includes the catalytic domain and accessory domains such as carbohydrate binding modules (CBMs), the most well studied type of accessory domain, fibronectin type III domain (Fn3), immunoglobulin-like domain (Ig-like) and Calx-beta domains, whose functions are still not fully elucidated (BORASTON et al., 2004; HAYASE et al., 2008; KATAEVA et al., 2002; LIU et al., 2010; MAEKAWA et al., 2006; MEDIE et al., 2012).

In this context, the aim of this work was to perform comprehensive biochemical and biophysical characterization of CelE2, a modular cellulase derived from soil metagenome composed by an N-terminal GH5 domain and a C-terminal with identity to Calx-beta domains.

## 2 Materials and methods

### 2.1 *Sequence and architecture analysis*

The *cele2* gene was identified after functional screening on carboxymethyl cellulose of a soil-derived metagenomic library, as described previously (ALVAREZ et al., 2013a). Plasmid from the positive clone was extracted and submitted to DNA sequencing in an Applied Biosystems 3500xL Genetic Analyzer at the Brazilian Bioethanol Science and Technology Laboratory. The sequences generated were analyzed in Geneious Pro 5.6.2 (http://www.geneious.com) (KEARSE, M., MOIR, R., WILSON, A., STONES-HAVAS, S., CHEUNG, M., STURROCK, S., BUXTON, S., COOPER, A., MARKOWITZ, S., DURAN, C., THIERER, T., ASHTON, B., MENTJIES, P., DRUMMOND, 2012) for identification of the coding DNA sequence (CDS). The nucleotide sequence from the predicted CDS was compared to the NCBI database using the BLASTX tool and deposited in the NCBI GenBank

(accession number KU715986). The modular architecture of CelE2 was evaluated in further details by comparison with PFAM and SMART databases (FINN et al., 2015; LETUNIC; DOERKS; BORK, 2015; SCHULTZ et al., 1998). Predictions of the signal peptide and biochemical and biophysical parameters were performed via SignalP 4.1 (PETERSEN et al., 2011) and Protparam tool from ExPASy (GASTEIGER et al., 2005), respectively.

In order to perform orthologous group annotation, protein sequence was compared against the Eggnog database (HUERTA-CEPAS et al., 2016). For phylogenetic analysis, protein sequences from the respective orthologous group were downloaded and used for construction of the phylogenetic tree in Mega 6.06 (TAMURA et al., 2013) using the neighbor-joining method (SAITOU; NEI, 1987).

## 2.2    *Cloning, heterologous expression and protein purification*

The coding DNA sequence of CelE2 was PCR-amplified using the forward primer 5' TATA<u>GCTAGC</u>CTGAGTATCTCGGATGCC 3' and reverse primer 5' ATA<u>AAGCTT</u>TTAGGCGCTACCCACCCGATC 3', containing *Nhe*I and *Hind*III restriction sites (underlined in the primers, respectively). The amplified gene was cloned into expression vector pET-28a (Novagen), in fusion with an N-terminal 6-His tag. Subsequently to the transformation in *Escherichia coli* Rosetta-gami™ 2 (Novagen), heterologous expression were performed by cultivation of cells in LB broth containing 50 $\mu$g.mL$^{-1}$ kanamycin and 40 $\mu$g.mL$^{-1}$ chloramphenicol, at 37°C and 200 rpm until the $OD_{600}$ reached 0.6–0.8. Afterward, IPTG was added at a final concentration of 0.5 mM to induce expression and the temperature and agitation speed were reduced to 18°C and 180 rpm, respectively, for approximately 16 h. Then, the culture was centrifuged at 35,250 x *g* for 30 minutes at 4°C and the recovered cells were resuspended in 20 mM Tris-HCl (pH 7.5), 200 mM NaCl, 20% glycerol and 5 mM imidazole followed by cell disruption in an ultra sonicator (Sonics Vibra-Cell™). The first step of purification consisted of affinity chromatography that involved incubation of the recovered supernatant with TALON resin (Clontech) followed by elution of proteins with 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 20% glycerol and 100 mM imidazole. Next, the protein sample was further purified by size exclusion chromatography using a Superdex 75 10/300 GL column (GE Healthcare), equilibrated with 20 mM sodium phosphate buffer (pH 7.4) containing 50 mM NaCl. The purity of the collected protein fractions was assessed by SDS-PAGE (LAEMMLI, 1970) and quantification was performed by absorbance at 280 nm, considering the predicted extinction coefficient of CelE2 (123300 M$^{-1}$ cm$^{-1}$).

*2.3    Biochemical characterization*

In all steps of biochemical characterization the hydrolytic activity was determined by quantification of the amount of reducing sugar released according to the 3,5–dinitrosalicylic acid method (MILLER, 1959), and all the assays were carried out at least in triplicate.

First, the enzymatic activity was analyzed in reactions (100 µL) with different soluble polysaccharides 0.25% (w/v) (purchased from Megazyme and Sigma-Aldrich) in 40 mM of sodium acetate buffer pH 5.5, for 60 min at 40°C. For evaluation of temperature-dependence, reactions containing oat β-glucan 0.25% (w/v) (Megazyme) in 40 mM sodium acetate buffer pH 5.5 were incubated in a range of 15–90°C for 30 min. The influence of different pH values on enzymatic activity was evaluated in 40 mM sodium acetate buffer (pH 3.6-5.6) and 40 mM sodium phosphate buffer (pH 5.8-8.0), in reactions containing oat β-glucan 0.25% (w/v) followed by incubation at 45ºC for 60 min.

After these initial characterization steps, all assays were performed in oat β-glucan 0.25% (w/v) and at the optimal pH and temperature. Thermal stability was assessed by incubation of CelE2 in 20 mM sodium phosphate buffer (pH 7.4) containing 50 mM NaCl at temperatures ranging from 40ºC–90ºC, followed by the collection of aliquots at regular times for measurement of residual activity by incubation for 30 min at optimal conditions.

The effect of different metal ions on enzymatic activity was evaluated by addition of compounds to be tested in the final concentration of 5 mM (MO et al., 2010), followed by incubation for 60 min at 45ºC. Considering that small variations of the concentration of ions present on purified CelE2 may interfere in the assays of ion supplementation, two independent experiments were performed. For analysis of the chelating effect, an enzyme solution was treated with final concentrations of 1 mM, 5 mM and 10 mM EDTA (Ethylenediaminetetraacetic acid - Calbiochem) for 60 min at 5ºC. Next, reactions containing EDTA-treated enzyme were incubated for 30 min at 45ºC for determination of residual enzymatic activity. Additionally, 5 mM $CaCl_2$ were added in reactions containing EDTA-treated enzyme in order to evaluate the ability of this metal salt to restore the enzymatic activity. Lastly, the enzymatic activity of CelE2 was evaluated in the presence of 1% (v/v) Triton™ X-100 (Sigma-Aldrich) and 1 mM SDS (Sodium n-Dodecyl Sulfate - Merck), with incubation for 60 min at 45ºC. Reactions containing these agents were also evaluated with supplementation of 5 mM $CaCl_2$. Comparisons of enzymatic activity in the presence and absence of compounds

were assessed by performing the unpaired Student *t* test using GraphPad Prism version 7.00, GraphPad Software, La Jolla California USA, www.graphpad.com.

### 2.4   *Analysis of the cleavage pattern by capillary zone electrophoresis*

The cleavage pattern was determined by evaluation of the hydrolysis products released after overnight incubation (at 45ºC) with 5 mM of oligosaccharides (cellobiose - C2, cellotriose - C3, cellotetraose - C4 and cellopentaose - C5, all from Megazyme), 0.25% (w/v) oat β-glucan and 10% (w/v) Avicel PH-101 (Sigma Aldrich). Following incubation, the soluble fractions from the reactions were recovered and derivatized with 8-aminopyrene-1,3,6-trisulfonic acid (APTS) by reductive amination (EVANGELISTA; LIU; CHEN, 1995). Separations were performed using a P/ACE MQD Instrument (Beckman Coulter) equipped with laser induced fluorescence detection at the same conditions as described in Alvarez *et al.* (2015).

### 2.5   *Circular dichroism spectroscopy (CD) and thermal denaturation*

Circular dichroism spectroscopy and thermal denaturation were performed according to Mandelli *et al.* (2013). Far-UV CD spectra were recorded using the JASCO 815 spectropolarimeter (JASCO Inc., Tokyo, Japan), equipped with a Peltier temperature control unit, from 195 to 260 nm in a 1 mm path length quartz cuvette. Data collection considered a scanning speed of 100 nm.min$^{-1}$, spectral bandwidth of 1 nm and a response time of 0.5 s. The purified enzyme (0.2 mg.mL$^{-1}$), diluted in 20 mM of sodium phosphate buffer (pH 7.5) containing 50 mM NaCl, was used for collection of spectral data. Solvent spectra were subtracted in all experiments, and each spectrum was an average of 20 scans. Thermal denaturation analyzes of the recombinant enzyme were monitored by measuring the ellipticity changes at 219 nm while increasing the temperatures from 20 to 100°C at a rate of 1°C.min$^{-1}$. Residual activity was measured by recovering of protein sample from the cuvette followed by incubation for 30 min at optimal conditions in oat β-glucan 0.25% (w/v).

### 2.6   *Small-Angle X-ray Scattering*

SAXS (Small angle X-ray scattering) data were collected at the Brazilian Synchrotron Light Laboratory (LNLS – CNPEM) at the SAXS2 beamline. Scattering patterns were recorded using a Pilatus detector and the wavelength of the incident radiation was set to $\lambda = 1.54$ Å with a sample-detector distance of 908 mm, resulting in a scattering vector (defined

by $q = 4\pi \sin(\theta)/\lambda$, where $2\theta$ is the scattering angle) ranging from 0.012 < q < 0.436 $\text{Å}^{-1}$. SAXS analyzes were performed at 20°C, with enzyme concentrations varying from 1.0 mg.mL$^{-1}$ to 2.4 mg.mL$^{-1}$ in the presence and absence of 5 mM of CaCl$_2$. SAXS data were corrected regarding the detector response and scaled by the intensity of the beam and sample absorption. The background scattering was subtracted from the sample scattering and integration of the SAXS patterns was performed using the Fit2D software (HAMMERSLEY, 1997).

Fitting of the experimental curve and evaluation of the pair-distance distribution function *p(r)* were conducted using the Gnom software (SVERGUN, 1992) from the ATSAS package. The low-resolution dummy atoms model (DAM) was generated using the *ab initio* routine implemented in the program Dammin (SVERGUN, 1999). The molecular weight and oligomerization state were evaluated using SaxsMoW (FISCHER et al., 2010). High-resolution homologous models were generated from the sequence using the Swiss-Model (BIASINI et al., 2014), based on the templates PDB id 1VRX and 2DPK, respectively for catalytic and Calx-beta domains. The rigid body model (RBM) was generated using Bunch (PETOUKHOV; SVERGUN, 2005) from the two homologous domains. Crysol (SVERGUN et al., 1995) was used to simulate the scattering pattern and to evaluate the structural parameters of the RBM. The DAM and RBM were superimposed using the program SUPCOMB (KOZIN; SVERGUN, 2001).

## 3  Results and discussion

### 3.1  *Phylogenetic analysis of CelE2*

After functional screening of a metagenomic library derived from soil, a clone harboring an insert of approximately 2.3 kb with the ability to hydrolyze CMC was isolated. Sequence analysis of the insert resulted in identification of a CDS for a protein with 477 amino acids, referred to as CelE2, with no putative signal peptide. Comparison of the predicted protein sequencing with the NCBI-nr database indicated high identity values to a hypothetical protein from *Mycobacterium sp.* EPa45 (81% identity, accession number WP_052960145.1), glycoside hydrolase family 5 from *Mycobacterium rhodesiae* JS60 (79% identity, accession number EHB55226.1) and a hypothetical protein from *Mycobacterium rhodesiae* (79% identity, accession number WP_050950473.1).

The domain architecture evaluation performed by comparison of the CelE2 protein sequence to the Pfam database indicated the presence of a conserved domain from glycoside hydrolase family 5 (PF00150) in the N-terminal region, extending from amino acid 29 to 348

(e-value $1.5e^{-57}$, 100% coverage of HMM model), while in the C-terminal region, a Calx-beta domain was predicted from amino acid 360 to 454 (PF03160, e-value $2.9e^{-20}$, 89% coverage of HMM model). The same domain was predicted in the C-terminal end, from amino acid 350 to 454, when the analysis was performed against the SMART database (e-value $2.16e^{-15}$). The Calx-beta motif was first reported in 1997 by Schwarz and Benzer (SCHWARZ; BENZER, 1997) as a component of the Na-Ca exchanger from *Drosophila melanogaster*, also identified in mammalian integrin β4 and in proteins from *Synechocystis sp*. PCC6803 (SCHWARZ; BENZER, 1999, 1997). A recent search for the term "Calx-beta" associated with the term "glycoside-hydrolase" in the Uniprot database (THE UNIPROT CONSORTIUM, 2014) resulted in the identification of 251 entries. Maekawa *et al.* (2006) and Hayase *et al.* (2008) provided insights on the existence and importance of the Calx-beta domain in the β-glucosidase (GH3), BglM1, from *Physarum polycephalum,* suggesting that it may play a role in maintaining the protein tertiary structure.

Regarding the orthologous group annotation, the best classification of CelE2 (e-value 3.71 $e^{-169}$, score 566.5) was to ENOG4106R7G which referred to the category of carbohydrate transport and metabolism and was annotated as glycoside hydrolase family 5. The phylogenetic tree (Fig. 1) indicated a close association of CelE2 to endoglucanases from *Mycobacterium flavescens* and *Mycobacterium vanbaalenii* (Taxon IDs: 350054 and 350058, respectively), with both harboring Calx-beta domains according to the Pfam annotation.

**Fig. 1.** Phylogenetic analysis of CelE2 and orthologs of ENOG4106R7G constructed in Mega 6.06 using the neighbor-joining method (SAITOU; NEI, 1987; TAMURA et al., 2013). The evolutionary distances were computed using the p-distance method (NEI; KUMAR, 2000) and the Bootstrap values (n= 1,000) are shown as percentage at the nodes. Sequences are identified by their taxonomic ID (NCBI) followed by gene name from the Uniprot Database (THE UNIPROT CONSORTIUM, 2014).

Recombinant CelE2 was successfully overexpressed in the cytoplasmic fraction of the *E.coli* Rosetta-gami2 strain and displayed a high level of purity after two protein purification steps, as indicated by SDS-PAGE (Fig. 2), with an apparent molecular weight similar to the predicted value of 53.8 kDa.
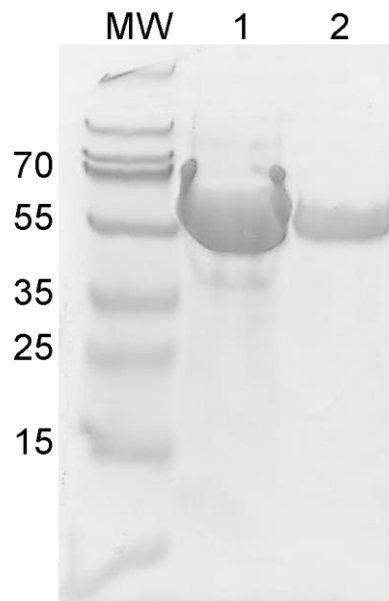
**Fig. 2.** SDS-PAGE of recombinant CelE2 protein purification performed first by affinity chromatography (column 1) followed by gel filtration chromatography (column 2). MW - Molecular weight marker (PageRuler Plus Prestained Protein Ladder 10-250 kDa, Thermo Scientific).

*3.2    Biochemical properties, substrate specificity and thermal denaturation*

CelE2 displayed a notably preference for hydrolysis of oat β-glucan (10.0 ± 0.95 U/ mg - relative activity 100%), followed by lichenan (3.6 ± 0.06 U/ mg - relative activity 35.9 ± 3.9%) and CMC (0.7 ± 0.30 U/ mg - relative activity 6.6 ± 2.8%) (Fig. 3A). Both oat β-glucan and lichenan are mixed β-1,3-1,4-glucans, however, in the case of lichenan there is a higher proportion of β-1,3 linkages (LAZARIDOU et al., 2004; WOOD; WEISZ; BLACKWELL, 1994). In the other soluble substrates tested composed of a variety of glycosidic bond types and monomeric units, no significant activity was observed (Fig. 3A). Specifically, the inability to cleave curdlan, a β-1,3-glucan polysaccharide, suggested that CelE2 is not able to cleave β-1,3 glycosidic bonds. Thus, the activity of CelE2 on both oat β-glucan and lichenan may be the result of hydrolysis of β-1,4 glycosidic bonds.

The evaluation of optimum temperature and pH using oat β-glucan as substrate revealed optimum values at 45 ℃ and pH 5.3 (Fig. 3B and 3C). From 15 ℃ to 45 ℃, CelE2 activity exhibited a crescent profile as a function of temperature increase, with a maximum at 45 ℃. At 50 ℃, CelE2 was able to retain approximately 82% of its maximum enzymatic activity; however, from 50-55 ℃ there was a drastic decrease in enzymatic activity to approximately 39%. Optimum temperature values of approximately 50 ℃ have already been

reported for other soil metagenome derived cellulases (ALVAREZ et al., 2013a; HUA et al., 2015). At pH values ranging from 4.8 and 5.6, CelE2 showed relative activity values greater than 55%, with the optimum at pH 5.3. This result indicated an acidophilic characteristic of CelE2, which is a common behavior in comparison to other cellulases derived from soil metagenome (HUA et al., 2015; LIU et al., 2011).

CelE2 exhibited considerable thermal stability (Fig. 3D) at 40 ℃ after 360 min, retaining approximately 68% of its initial enzymatic activity. After 360 minutes of incubation at 50 ℃ and 60 ℃, CelE2 was able to retain 38% and 20% of its enzymatic activity, respectively. CelE2 was virtually inactivated after 180 min of incubation at 70 ℃. At higher temperatures, 90 ℃ and 80 ℃, CelE2 was able to maintain approximately 18% and 57% of its activity, respectively, after 30 minutes of incubation.
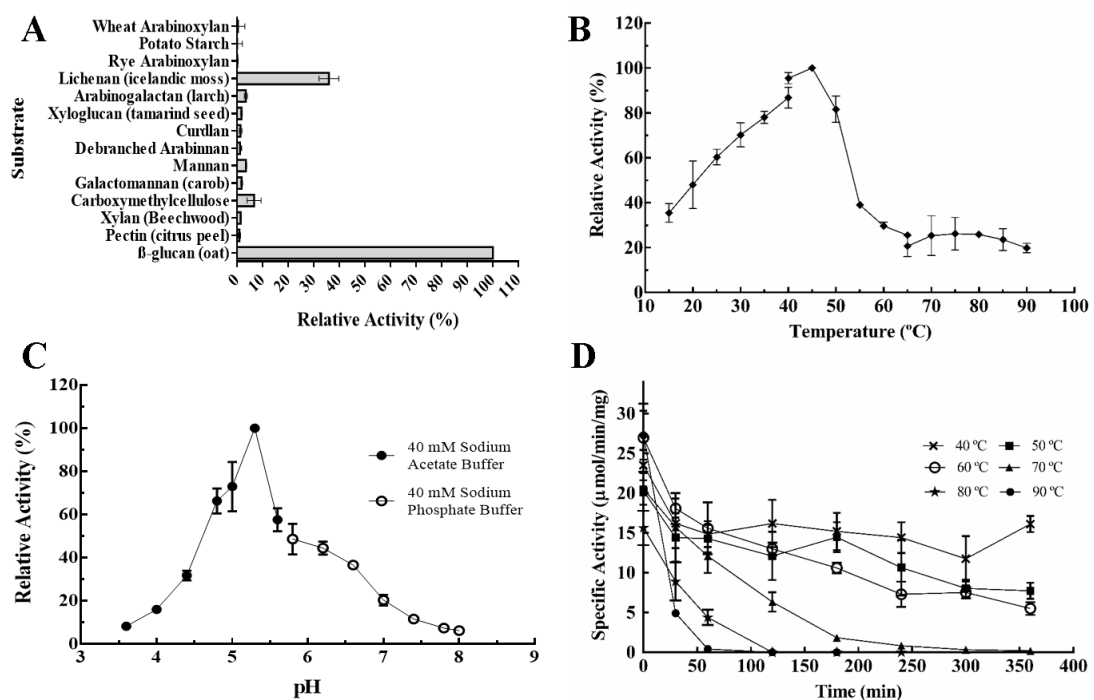


**Fig. 3. A:** Substrate specificity of CelE2 evaluated in substrates containing a variety of glycosidic bonds. **B and C**: Influence of temperature and pH on CelE2 activity, respectively. **D:** Thermal stability of CelE2 evaluated via enzymatic activity.

Thermal unfolding was assessed by circular dichroism spectroscopy. Firstly, the CelE2 far-UV CD spectra resulted in a profile with negative bands around 209 and 219 nm and a positive band around 190 nm (Fig. 4A), reveling a secondary structure consistent with alpha-beta proteins (GREENFIELD, 2007), which is similar to the profile of others members of the

GH5 family (ALVAREZ et al., 2013a; SANTOS et al., 2012; ZHENG; DING, 2013). Thermal denaturation of CelE2 was evaluated by analysis of temperature-induced changes in the CD signals at 219 nm. The melting temperature (Tm) estimated via sigmoidal fitting of these data for CelE2 was 50°C (Fig. 4B), which is consistent with the activity assays (Fig. 3D), since at higher temperatures the enzyme showed progressive loss of enzymatic activity over time. Interestingly, when incubated at 99 °C for 15 min, CelE2 did not lose its secondary structure (Fig. 4A) nor its activity (residual relative activity of approximately 26%), which explains the activity observed at 90 °C during approximately 50 minutes, as shown in Fig. 3D. However, when CelE2 was progressively heated (from 20 to 100°C at a rate of 1°C min$^{-1}$), it lost its secondary structure and was inactivated, as assessed by enzymatic reactions (relative activity of 0.78%).
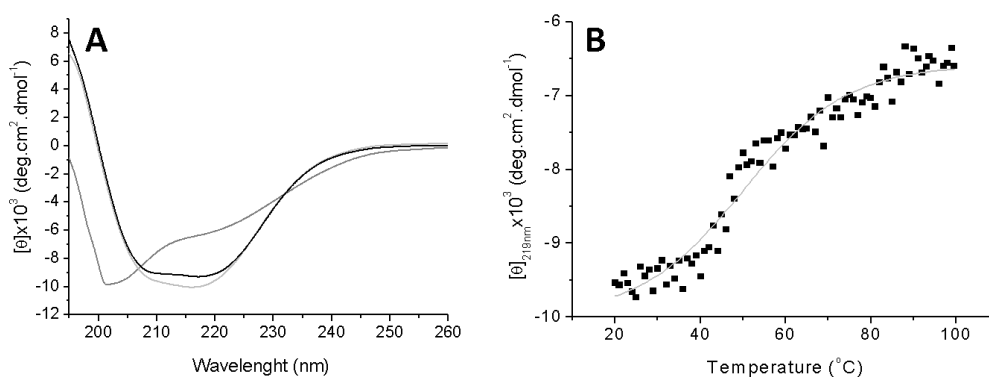


**Fig. 4**. Thermal denaturation of CelE2 monitored by circular dichroism. **A:** CD spectra in the far ultraviolet region. Black line represents the CelE2 spectra at 20 °C, light gray line is the CelE2 spectra after incubation at 99 °C for 15 min and gray line is the CelE2 spectra after being progressively heated from 20 to 100 °C (1°C min$^{-1}$). **B:** Ellipticity at 219 nm as a function of temperature (bold squares), the solid line shows the fitted denaturation curve using the Sigmoidal adjustment (y = -6583 + (-9980 + 6583)/ (1 + exp ((x-50)/12)) with R-squared of 0.96.

## 3.3  *Effect of metallic compounds, EDTA and surfactants on CelE2 activity*

As shown in Fig. 5A, apart from CaCl$_2$ and CoCl$_2$, all metallic salts evaluated resulted in varying degrees of inhibition to CelE2 enzymatic activity. Based on results described in literature, the effect of metal ions on endoglucanase activity is highly variable, however, inhibition by Cu$^{2+}$, Fe$^{3+}$, Fe$^{2+}$ and Zn$^{2+}$ have already been reported (BÉRA-MAILLET et al.,

2000; GHATGE et al., 2013; LEE; KIM, 1999; MCGAVIN; FORSBERG, 1988; RUBINI et al., 2010; SILVA et al., 2016; TEJIRIAN; XU, 2010) .

While 5 mM of $CoCl_2$ did not show a clear effect on enzymatic activity, the addition of 5 mM of $CaCl_2$ promoted an increase in CelE2 enzymatic activity. The effect of calcium on CelE2 enzymatic activity was further evaluated, first by the treatment of a diluted sample of CelE2 with three different concentrations of EDTA (1 mM, 5 mM and 10 mM). As shown in Fig. 5B, treatment with 5 mM and 10 mM promoted a significant reduction in CelE2 activity ($p$-value $< 0.05$), suggesting that CelE2 activity is influenced by the presence of divalent cations even without external metal ions supplementation. Interestingly, the addition of 5 mM of $CaCl_2$ to reactions containing the EDTA-treated enzyme was able to restore CelE2 activity (Fig. 5B). In a study performed with an endoglucanase from *B. succinogenes* (EG1), the presence of EDTA decreased enzymatic activity, but the addition of $CaCl_2$ was not able to completely restore enzyme activity (MCGAVIN; FORSBERG, 1988). On the other hand, Liu *et al.* 2011 reported a soil metagenome derived endoglucanase which presented approximately a 25% improvement on enzymatic activity in the presence of $CaCl_2$, although the treatment with 10 Mm EDTA showed only a 7% reduction in activity (LIU et al., 2011). These results suggest that CelE2 enzymatic activity may be modulated in the presence of $CaCl_2$.

Regarding the influence of the surfactants, Triton X-100 reduced CelE2 activity by approximately 26% ($\pm$ 8%) while SDS completely inactivated the enzyme. Interestingly, the addition of 5 mM $CaCl_2$ to the reactions with surfactants resulted in less susceptibility of CelE2 to these compounds, since CelE2 was able to retain around 98% ($\pm$ 2%) and 7% ($\pm$ 2.5%) of its enzymatic activity in Triton X-100 and SDS, respectively. Altogether, these results indicated that the presence of $CaCl_2$ may have an important influence on CelE2 activity.
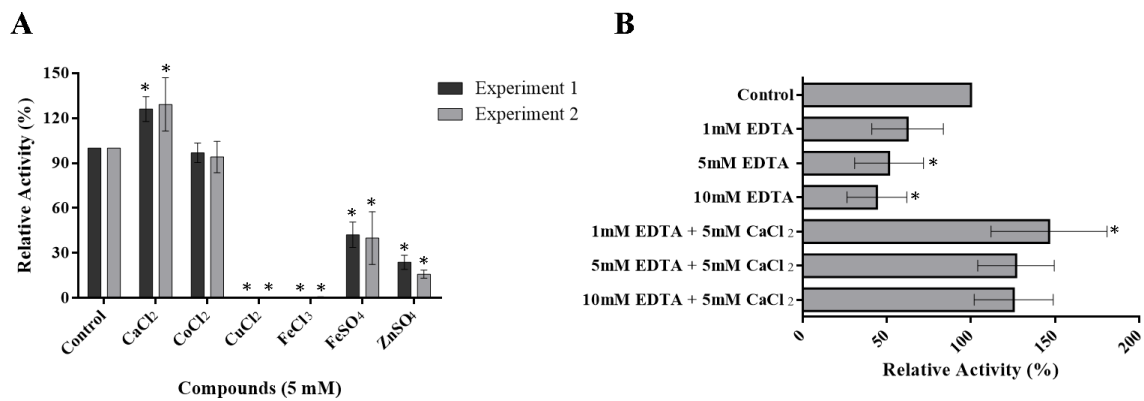
**Fig. 5.** Analysis of the effect of ions **(A)** and EDTA **(B)** on enzymatic activity. * indicates the statistically relevant results in relation to the control reactions by the Student's t-test with p-value $< 0.05$.

### 3.4 CelE2 displayed beta-1,4 endoglucanase behavior with evidences of transglycosilation activity

Capillary zone electrophoresis was used as a tool to characterize the cleavage pattern of CelE2. Degradation of oat β-glucan generated mainly di-, tri-, tetra- and pentasaccharides (Fig. 6A), while degradation of Avicel (Fig. 6A) resulted in accumulation of mainly di- and trisaccharides. Therefore, CelE2 displayed a behavior typical of endoglucanases, which randomly attack the polysaccharide chain resulting in the formation of oligosaccharides with a wide range of length (LYND et al., 2002).

Also, the cleavage pattern of CelE2 was analyzed by the products released from hydrolysis of cellooligosaccharides. CelE2 was able to cleave C5 and C4 (Fig. 6B), but not C3 and C2 (data not shown), suggesting that the minimal length of the glucan chain to be cleaved is four glucose units. Regarding hydrolysis of C4, it was observed an accumulation of di- and trisaccharides. However, hypothesizing that CelE2 would be able to cleave an external glycosidic bond of C4 to form C3, it should result in the release of glucose, which was not observed. The same occurred in the hydrolysis of C5 which generated tetra-, tri- and disaccharides, with no clear evidence of glucose release. In addition, in both cases there were evidences of formation of products with higher degree of polymerization than the substrates initially added (Fig. 6B). Together, these two results suggested the occurrence of transglycosylation activity, which is a characteristic displayed by many retaining glycoside hydrolases (SINNOTT, 1990), which is the mechanism exhibited by GH5 members (LOMBARD et al., 2014). Several members from the GH5 family were already described as

possessing transglycosylation activity (DILOKPIMOL et al., 2011; SCHRÖDER et al., 2006; TANABE et al., 2003), however, accordingly to Aspeborg *et al.* (2012), the transglycosylation activity does not correlate with subfamily classification and may be related to subtle peculiarities in protein structure.
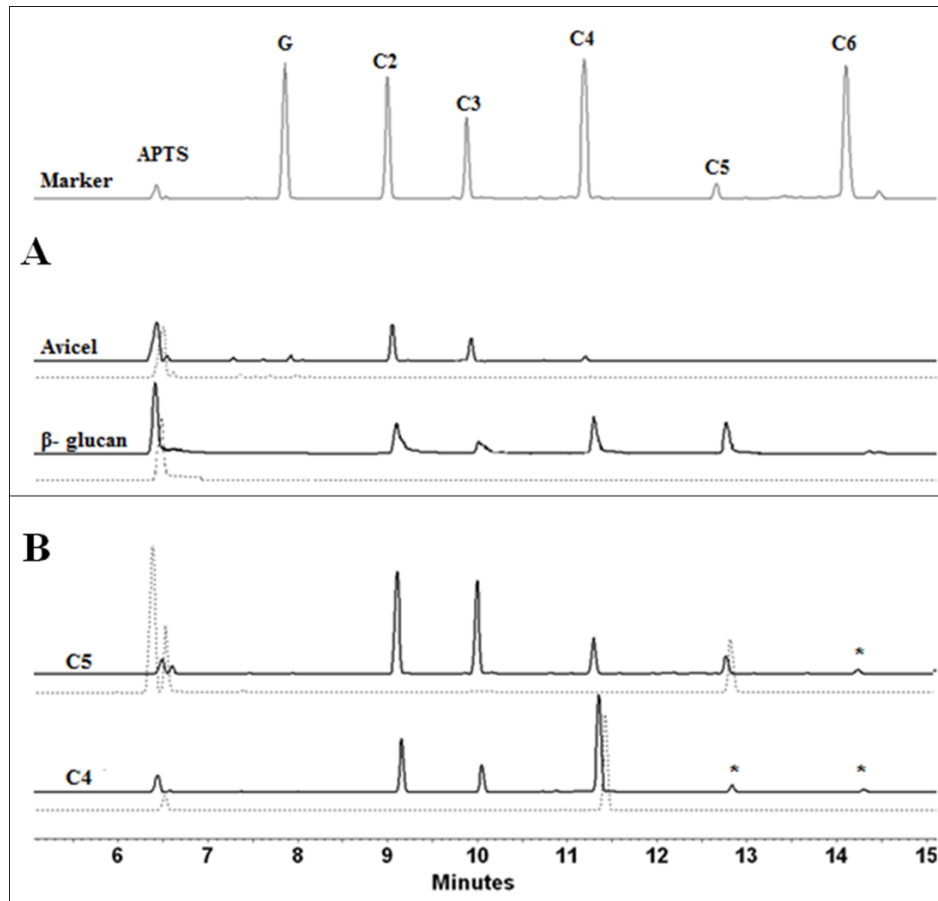


**Fig. 6.** Capillary zone electrophoresis for characterization of the cleavage pattern of CelE2. Solid lines indicate the products released after overnight hydrolysis of Avicel, oat β-glucan, cellopentaose (C5) and cellotetraose (C4). Dotted lines represent the controls reaction without the addition of enzyme. Gray line shows the marker containing: glucose (G), cellobiose (C2), cellotriose (C3), cellotetraose (C4), cellopentaose (C5) and cellohexaose (C6). * Indicates the formation of peaks with higher degree of polymerization than the oligosaccharide initially added to the reaction. APTS indicates excess of the fluorescent label 9-Aminopyrene-1,4,6-trisulfonic acid.

*3.5   Small-Angle X-ray Scattering*

In order to obtain the low-resolution molecular envelope of CelE2 and investigate the effect of calcium ions on enzyme stability, SAXS data were collected in the presence and

absence of 5 mM of $CaCl_2$. Since no interparticle correlation effect was detected, analyzes were conducted using data at the higher protein concentration (2.4 mg.mL$^{-1}$ in the presence and 2.0 mg.mL$^{-1}$ in the absence of calcium), due to better statistical results. Fig. 7 shows the experimental data for CelE2 in the presence and absence of $CaCl_2$. The zoom to the rectangle in the low angle region (Fig. 7A – Insert) reveals that the enzyme has a larger tendency to aggregate in the absence of calcium, indicated by the ascent of the data points in this region. The radius of gyration ($R_g$) obtained by Guinier approximation (Fig. 7B) respected the q.$R_g$ < 1.3 limit (GUINIER, A. AND FOURNET, 1955). $R_g$ obtained in the presence of calcium (31.1 Å) had a significant decrease compared to data without calcium ions (35.1 Å), quantitatively demonstrating the aggregation. Guinier's analysis showed that calcium ions provide CelE2 with greater stability.
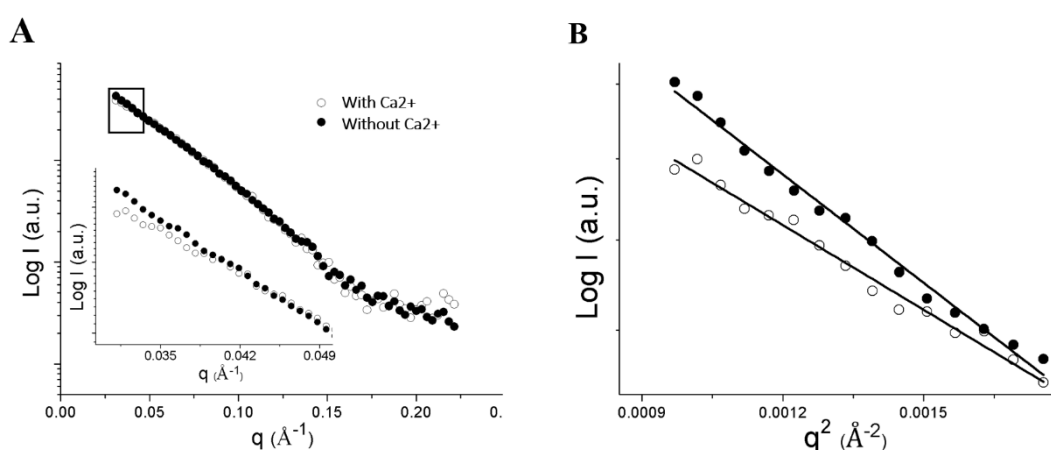


**Fig. 7.** Experimental SAXS curve of CelE2 in the presence and absence of $CaCl_2$ **(A)**. Zoom of the low angle region (Insert) reveals a tendency to aggregate in the absence of $CaCl_2$. Guinier's Approximation **(B)**.

Homologous models from the (1) catalytic domain was generated based on the crystallography structure of an endocellulase from *Acidothermus cellulolyticus* (BAKER et al., 2005) (PDB id: 1VRX), which presents a sequence identity of 53.91%, and from the (2) Calx-beta domain was generated based on the crystallography structure of a sodium/ calcium exchanger from *Canis lupus* (NICOLL et al., 2006) (PDB id: 2DPK), which presents a sequential identity of 59.02%.

Fitting of the experimental data and evaluation of the *p(r)* (Fig. 8 – Insert) were conducted for data in the presence of $CaCl_2$ due to the higher stability. The low-resolution

model of dummy atoms model was generated and compared with the rigid-body model generated based on the relative positions between the domains and the experimental data. Fig. 8 shows the fits against the SAXS experimental pattern. Structural parameters from the experimental SAXS data of CelE2 and simulated SAXS from the DAM and the RBM are presented in Table 1.
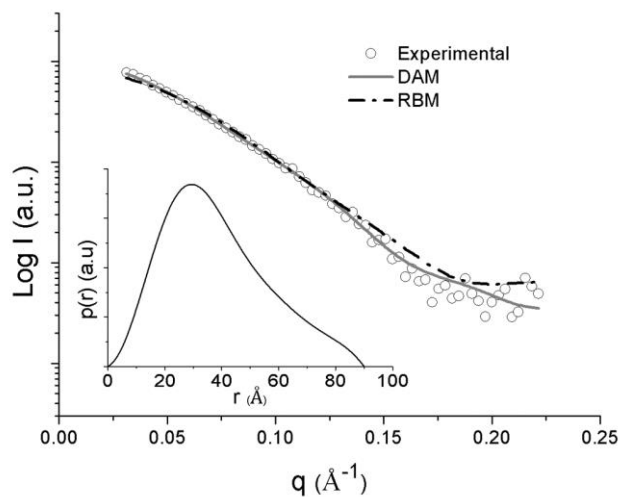


**Fig. 8.** Experimental data of CelE2 in the presence of 5 mM of $CaCl_2$ and fitting procedures (DAM: Dummy atoms model; RBM: Rigid-body model). Pair-distance distribution function p(r) (Insert).

**Table 1:** SAXS structural parameters of CelE2. [1]Exp, calculated from the experimental data. [2]DAM, parameters of the dummy atoms model. [3]RBM, parameters of the rigid-body model. [4]Resolution is calculated as $2\pi/q_{max}$.

| Parameters | Exp[1] | DAM[2] | RBM[3] |
|---|---|---|---|
| $R_g$ (Å)(Guinier) | 31.1 | - | - |
| $R_g$ (Å) | 29.7 | 29.3 | 25.0 |
| $D_{max}$ (Å) | 90.0 | 91.2 | 82.1 |
| SAXS resolution[4] (Å) | 28.4 | - | - |
| Molecular weight (kDa) | 67.4 | - | - |
| $\chi^2$ | - | 3.5 | 4.7 |

The molecular weight evaluated from SAXSMoW (67.4 kDa) compared with the theoretical value (53.8 kDa) shows that, in the experimental conditions CelE2 behaves mainly

as a monomer in solution. However, the parameters reveal a tendency of CelE2 to form oligomers, indicated by the higher molecular weight values evaluated from SAXS and increase in the $R_g$ parameter when evaluated only in the low $q$ region. Nevertheless, this tendency is so small that it does not influence the reliability of the results, as shown in the good agreement of the models (Fig. 9).
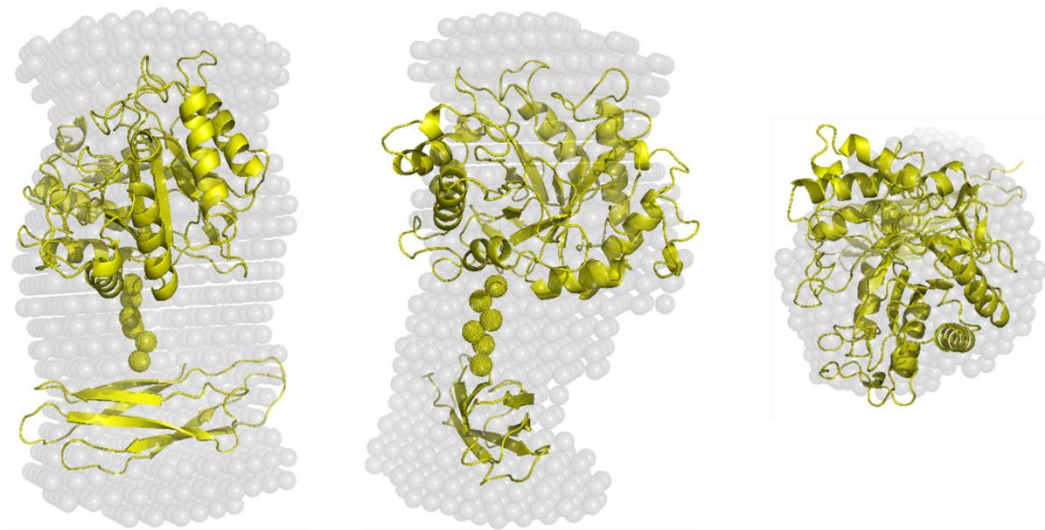


**Fig. 9.** Superposition of the rigid body-model (RBM) with the low-resolution dummy atoms model (DAM). The center and right models were rotated 90° around the y- axis and 90° around the x-axis from the orientation shown on the left panel.

## 4    Concluding Remarks

The cellulase (GH5) CelE2, displaying an unconventional domain architecture, was successfully heterologously expressed in *E.coli*, purified and submitted to biochemical characterization. CelE2 exhibited mesophilic and acidic characteristics, with substrate preference for mixed β-1,3-1,4-glucans and a typical cleavage pattern of endoglucanase with evidences of transglycosylation activity. The addition of $CaCl_2$ was able to not only promote an increase in CelE2 activity but also resulted in less susceptibility of inactivation in the presence of surfactants. SAXS experiments provided key information regarding the increased stability of the enzyme in the presence of $CaCl_2$, which possibly explains the higher activity observed in the presence of this compound. Influence of the Calx-beta domain on cellulase activity is a

key point to be elucidated, especially considering that this domain has already been identified in other glycoside hydrolases.

## Author Contributions

Conceived and designed the experiments: AGP, TMA, FMS, MVL. Performed the experiments: ACP, TMA, GCGE, FM, RT, CAG, MON. Analyzed the data: ACP, TMA, FM, CAG, MON, DAAP, FMS, JPLFC. Wrote the paper: ACP, TMA, FMS, FM, RT, MVL, JPLFC, CAG, MON.

**DOCUMENTO 2 – STRUCTURAL AND FUNCTIONAL PROPERTIES OF THE CATALYTIC GH5 AND CALX- β DOMAINS OF THE METAGENOME-DERIVED ENDO-BETA-1,4-GLUCANASE CelE2**

Trabalho em preparação.

**Título: Propriedades estruturais e funcionais dos domínios catalítico GH5 e Calx-β da endo-beta-1,4-glucanase CelE2 de origem metagenômica**

**Resumo do documento:** As celulases apresentam grande potencial de aplicação biotecnológica por promoverem a despolimerização da celulose, o polímero natural mais abundante do planeta. Além do domínio catalítico, a presença de domínios acessórios podem influenciar diretamente a atividade de enzimas envolvidas na degradação de carboidratos. A endoglucanase CelE2 de origem metagenômica possui em sua arquitetura modular um domínio N-terminal característico da família GH5 e um segundo domínio C-terminal com identidade para o domínio Calx-β. Com o intuito de avaliar as propriedades e funcionalidade dos domínios individualizados de CelE2, foram desenhadas duas possíveis construções para o domínio catalítico e para o domínio acessório, as quais foram estudadas quanto às propriedades bioquímicas e biofísicas. Os resultados mostraram que a deleção do domínio Calx-β (construção $CelE2_{382-477}$) ocasionou um aumento cerca de 4 vezes da atividade do domínio catalítico $CelE2_{1-381}$ no substrato insolúvel Avicel. No entanto, as principais características de CelE2, tais como substratos preferencias de atuação (β-glucano, carboximetilcelulose e liquenana), valores de pH e temperatura ótimos (próximos de 5,3 e 45 ℃, respectivamente), estabilidade térmica a 40 ℃ e 50 ℃ em 360 minutos e ativação por $CaCl_2$ não foram alteradas com a deleção do domínio Calx-β. As análises de termoestabilidade por dicroísmo circular (CD) revelaram uma diminuição dos valores de temperatura de desnaturação de cerca de 4,6 ℃ para o domínio catalítico (construção $CelE2_{1-381}$) em comparação a CelE2. Além disso, as análises por espalhamento de raio-X de baixo ângulo (SAXS) mostraram que a adição de $CaCl_2$ resultou na diminuição da tendência de agregação proteica para $CelE2_{1-381,}$ assim como já havia sido observado previamente para CelE2. Os ensaios de avaliação da ligação a polissacarídeos solúveis e insolúveis apontaram que $CelE2_{382-477}$ não possui afinidade para os substratos testados. A estrutura tridimensional do domínio catalítico $CelE2_{1-381}$ foi determinada a uma resolução de 2,1 Å que permitiu identificar um enovelamento do tipo barril $(α/β)_8$. Por fim, os resultados obtidos neste trabalho mostraram

que o domínio acessório não se apresentou essencial para atividade enzimática de CelE2, bem como para a manutenção de suas propriedades bioquímicas e biofísicas. Sua excisão resultou em alteração significativa da atividade do domínio catalítico em Avicel, visto que o valor encontrado foi superior ao observado para a construção completa, além de ter provocado a redução na temperatura de desnaturação conforme avaliação realizada por dicroísmo circular. Além das informações sobre os domínios individualizados e estrutura tridimensional do domínio catalítico, novos estudos poderão ser conduzidos a fim de verificar se os resultados obtidos neste trabalho em relação ao domínio Calx-β se estendem para outros representantes deste grupo, uma vez que este domínio já foi encontrado em outras celulase, e assim contribuir de forma mais ampla no entendimento da função deste domínio acessório para a atividade de glicosil hidrolases.

## Abstract

Cellulases present great potential of biotechnological application in cellulose depolymerization, the most abundant natural polymer in the Earth. In addition to the catalytic domain, the presence of accessory domains may directly influence the activity of enzymes involved in the carbohydrates degradation. The metagenome-derived endoglucanase CelE2 has in its modular architecture an N-terminal domain characteristic of the GH5 family, and a C-terminal domain with identity to the Calx-β domain. In this context, in order to evaluate the properties and functionality of the individualized domains of CelE2, two possible constructs were designed for the catalytic domain and for the accessory domain, which were studied for biochemical and biophysical properties. The results obtained showed that the deletion of the Calx-β domain (CelE2$_{382-477}$ construct) increased 4-fold the activity of the catalytic domain CelE2$_{1-381}$ on the insoluble substrate Avicel. However, the main properties of CelE2, such as substrate specificity (β-glucan, carboxymethyl cellulose and lichenan), optimal pH and temperature (close to 5.3 and 45 ℃, respectively), thermal stability at 40 ℃ and 50 ℃ for up to 360 minutes, and activation by $CaCl_2$ were not altered with the Calx-β domain deletion. Thermostability analyzes by circular dichroism (CD) revealed a decrease in the denaturation temperature values around 4.6 °C for the catalytic domain (CelE2$_{1-381}$ construct) in comparison to CelE2. In addition, Small Angle X-ray Scattering (SAXS) analyzes showed that the addition of $CaCl_2$ resulted in a decrease in the tendency of protein aggregation for CelE2$_{1-381,}$ as previously observed for CelE2. The binding assays with soluble and insoluble polysaccharides indicated that CelE2$_{382-}$

[477] has no affinity to the substrates tested. The three-dimensional structure of the catalytic domain CelE2$_{1-381}$ was determined at 2.1 Å resolution, which allowed the identification of a $(\alpha/\beta)_8$ barrel fold. Finally, the results obtained in this work showed that the accessory domain is not essential for the enzymatic activity of CelE2, as well as for the biochemical and biophysical properties. Its excision resulted in a significant alteration of the activity of the catalytic domain in Avicel, since the value found was superior to that observed for the complete construct, in addition to the reduction in the denaturation temperature, as assessed by circular dichroism. Besides to the results with of the individualized domains and the three-dimensional structure of the catalytic domain, further studies may be conducted to verify other representatives of Calx-β domain presents the same characteristics. Since this domain has already been found in other cellulases, it would bring a significant contribution to the understanding of the function of this domain for the activity of glycoside hydrolases.

## 1. Introduction

Cellulose is the most abundant biopolymer in the world, representing one of the most important component of plant cell wall. Since it is a renewable resource, this polymer has attracted attention due to the great potential in industrial applications, such as ethanol production, veterinary foods, wood and paper, fiber and clothes, cosmetic and pharmaceutical products (DE SOUZA LIMA; BORSALI, 2004; SHAHZADI et al., 2014; SHOKRI; ADIBKI, 2013).

Cellulases comprise a system with at least three different enzymatic activities that act on the hydrolysis of β-1,4 glycosidic bonds of cellulose chain. Endoglucanase (EC 3.2.1.4), exoglucanase/exo-cellobiohydrolase (EC 3.2.1.91/ 3.2.1.176), and β-glucosidase (EC 3.2.1.21), together with other non-hydrolytic enzymes with oxidative activity such as LPMOs, act in synergism to hydrolyze the cellulose polymer. Endoglucanases initiate the process with random attack to generate reducing and non-reducing ends, which are attacked by exo-cellobiohydrolases releasing cellobiose. Then cellobiose is converted into glucose monomers by β-glucosidases (GUPTA et al., 2016; SANDHU et al., 2018; SINGHANIA et al., 2017).

Due to the complexity of enzyme system and high biotechnological potential, cellulases are one of the major enzymes commercially available with applications in several industrial branches such as pulp and paper, textile, food, animal feed, laundry and detergent, pharmaceutical, extraction of bioactive compounds from plants, and others (KUMAR et al., 2014; SHARMA et al., 2016). Moreover, the increasing interest in second generation ethanol production as an alternative for replacing fossil fuels and reducing environmental impacts reinforces the importance of cellulases in the biofuels sector for lignocellulosic biomass degradation (KUMAR et al., 2014; YENNAMALLI et al., 2013).

Endoglucanases are distributed into different sequence-based families of glycoside hydrolases (GHs) in the Carbohydrate-Active Enzyme database (CAZy). Family 5 of glycoside hydrolases (GH5) constitute a large family which harbors multiple specificities and activities (LOMBARD et al., 2014).

Nature has evolved cellulolytic microbes that can produce a wide variety of enzymes that act on lignocellulosic biomass (YANG et al., 2011). In this context, metagenomics has been exploring non-cultured microorganisms, which represent the majority of organisms on Earth, resulting in the identification of coding genes for novel enzymes (ALVAREZ et al., 2013b, 2013a, 2015; CAMPOS et al., 2016; HUGENHOLTZ; TYSON, 2008; JIMENEZ et al., 2016; KIM et al., 2008; MARUTHAMUTHU et al., 2016).

Several glycoside hydrolases are known to be multi-modular, containing at least one catalytic domain which is responsible for the catalysis itself, and one or more accessory modules that may imply different functions (ASPEBORG et al., 2012). Carbohydrate-binding modules (CBMs) are the most commonly found and in general promotes proximity effect, targeting and disruptive function to the substrate (BORASTON et al., 2004). On the other hand, Calx-β domains are found in some glycoside hydrolases, including GH5 such as the enzyme from uncultured *Mycobacterium sp.* (UniProtKB - A0A1Y5PA04) and from *Cyanobacterium aponinum* (UniProtKB - K9Z7S6) (THE UNIPROT CONSORTIUM, 2017), but have no defined function in these cases. For β-glucosidase (GH3) BglM1 from *Physarum polycephalum*, Hayase *et al.* (2008) considered that Calx-β domain may play an important role in maintaining the tertiary structure of the protein. On the other hand, for Calx-β domains of the NCX Na+/Ca2+-exchangers, the structural information showed that this domain binds to $Ca^{2+}$ ions, unlike Calx-β domain of integrins β4 of cell-surface adhesion receptors, whose three-dimensional structure and other results showed that this domain does not bind $Ca^{2+}$ (ALONSO-GARCÍA et al., 2009; BESSERER et al., 2007; NICOLL et al., 2006). The structural and

biochemical characteristics that the accessory domains may confer to the catalytic domains are the main motivation for studies around individual modules (ARAKI et al., 2004).

Recently, Pimentel *et al.* (2017) reported the characteristics of a modular endoglucanase CelE2 derived from soil metagenome that harbors two conserved domains, the N-terminal region with identity to glycoside hydrolases family 5 (PF00150) and the C-terminal region with identity to Calx-β domain (PF03160). CelE2 exhibited activity against soluble substrate such as oat β-glucan, followed by lichenan and carboxymethyl cellulose, with indicative of transglycosylation activity in the hydrolysis of cellotetraose and cellopentaose. The optimal pH and temperature of enzymatic activity were observed at 45 ºC and pH 5.3, and CelE2 exhibited considerable thermal stability at 40 ºC for up to 360 minutes. Furthermore, the presence of $CaCl_2$ improved the activity and reduced the aggregation tendency.

In this context, the present study aimed to evaluate the biochemical and structural properties of the CelE2 domains individually in order to elucidate the influence of each domain on CelE2 enzymatic activity.

## 2. Materials and methods

### 2.1. Constructs analysis and delimitation

Full-length CelE2 sequence (GenBank - accession number KU715986) was compared to PDB database (Protein Data Bank) (BERMAN et al., 2000) using the NCBI BLASTp tool to delimit the regions to be cloned for the catalytic and accessory domain. Comparison of the CelE2 sequence to the proteins with the highest sequence identity resulted in the design of two possible constructs for the catalytic domain, called $cele2_{1-381}$ and $cele2_{1-363}$, and two for the Calx-β accessory domain, called $cele2_{382-477}$ and $cele2_{364-477}$.

### 2.2. Cloning, production and purification of the recombinant proteins

The sequences were PCR-amplified using the primers listed in Table 1 containing *Nhe*I and *Hind*III (New England Biolabs®) restriction sites. The reactions were described in supplementary Frame S1. The DNA of *cele2* was used as the template. Supplementary Fig. S1 shows a representation of the primers annealing for each construct.

**Table 1**. Primers and denaturation temperature used for the amplification of the constructs. F1F and F1R correspond to the forward and reverse primer oligonucleotides designed for *cele2* amplification. E1R and E2R correspond to the reverse primer oligonucleotides specific for the constructs *cele2$_{1-363}$* and *cele2$_{1-381}$* (catalytic domain), respectively. E3F and E4F correspond to the specific forward primer oligonucleotides for the constructs *cele2$_{364-477}$* and *cele2$_{382-477}$* (accessory domain), respectively.

| Primer Name | Sequence (5´-3´) | Tm |
|---|---|---|
| **F1F** | **TATA**<u>GCTAGC</u>CTGAGTATCTCGGATGCC | Tm = 51.9 ℃ |
| **F1R** | **ATA**<u>AAGCTT</u>TTAGGCGCTACCCACCCGATC | Tm = 61.3 ℃ |
| **E1R** | **TATA**<u>AAGCTT</u>TTAGTTGATGGTCCGCCAATC | Tm = 53.9 ℃ |
| **E2R** | **TATA**<u>AAGCTT</u>TTACGTTCCGTTGCCGCCGGTGTA | Tm = 64.4 ℃ |
| **E3F** | **TATA**<u>GCTAGC</u>CAGAACAAAATGGTCTATCTC | Tm = 48.6 ℃ |
| **E4F** | **TATA**<u>GCTAGC</u>GCCATGGCATCGTTCACCG | Tm = 59.6 ℃ |

The PCR-amplified fragments were purified and digested with restriction enzymes and then cloned into expression vector pET-28a (Novagen) in fusion with the N-terminal 6-His tag by heat shock transformation in *Escherichia coli* DH5α (NEBuffer – New England Biolabs®). The sequences were submitted to DNA sequencing in an Applied Biosystems 3500xL Genetic Analyzer at the Brazilian Bioethanol Science and Technology Laboratory. Physical and chemical parameters were predicted using the ProtParam tool (GASTEIGER et al., 2005) considering the 6-His tag and other fused amino acids. The recombinant plasmids were transform to *E. coli* Rosetta-gami2 (DE3) pLysS (Novagen) competent cells for protein expression. The heterologous expressions were conducted as described in Pimentel *et al.* (2017). Subsequently, the proteins were first purified from lysed cells by immobilized metal ion affinity chromatography using Talon resin (Clontech), with proteins elution in 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 20% glycerol and 100 mM imidazole (for CelE2$_{1-381}$ and CelE2$_{1-363}$) and in 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 20% glycerol and 300 mM imidazole (for CelE2$_{382-477}$ and CelE2$_{364-477}$). Then, the proteins were further purified by gel filtration using Superdex 75 10/300 (for CelE2$_{1-381}$ and CelE2$_{1-363}$) and Superdex 200 10/300 (for CelE2$_{382-477}$ and CelE2$_{364-477}$) GL (GE Healthcare), equilibrated with 20 mM sodium phosphate (pH 7.4) containing 50 mM NaCl.

For the structural analyzes of CD and SAXS in the presence of $CaCl_2$, the proteins were eluted in the gel filtration with 100 mM sodium acetate buffer (pH 5.3), since in this buffer the proteins presented lower precipitation tendency at high concentrations (0.2-2 mg.ml$^{-1}$).

The purity and integrity of the proteins were estimated by SDS-PAGE analysis (*sodium dodecyl sulfate-polyacrylamide gel electrophoresis*) (LAEMMLI, 1970) and the proteins quantification were performed by absorbance at 280 nm, considering the predicted extinction coefficient.

## 2.3. *Phylogenetic Analysis*

For the phylogenetic analysis of the catalytic domain and the Calx-β domain, two databases were constructed. All amino acid sequences of the GH5 family available in the database of active carbohydrate enzymes, dbCAN (YIN et al., 2012), were retrieved for the construction of the first database. The CelE2$_{1-381}$ amino acid sequence was aligned with the GH5 database using the Blastp tool. From this alignment, the GH5 sequences showing with e-value 1e$^{-30}$ and at least 50 % of identity were recovered for phylogenetic analysis.

Regarding to the second database, the CelE2$_{382-477}$ sequence was subjected to a search by sequence similarity in the EggNOG database (HUERTA-CEPAS et al., 2016). Only the sequences that presented a cut-off score of less than 1e$^{-30}$ were retrieved for selection of the region corresponding to the Calx-β domains (manual selection).

For each phylogenetic tree, the domain amino acid sequences were aligned using the ClustalW program (LARKIN et al., 2007). Then, the sequences were manually assessed and loaded into the Mega software version 7 (KUMAR; STECHER; TAMURA, 2016). Phylogenetic reconstruction was performed using the Neighbor-Joining algorithm (SAITOU; NEI, 1987) with bootstrap values calculated from 1000 replicates.

## 2.4. *Enzyme characterization*

The enzymatic assays were determined by evaluation of the enzymatic activity measuring the amount of reducing sugars released according to the DNS (3,5-dinitrosalicylic acid) method (MILLER, 1959), carried out at least in triplicate.

Initially, the substrate specificity was evaluated in different soluble polysaccharides 0.25% (m/ v) (Megazyme and Sigma-Aldrich) with reactions containing approximately 0.4 µg

of protein in 40 mM sodium acetate pH 5.3 at 45 °C for 60 minutes, followed by determination of the amount of reducing sugars released. In addition, the hydrolysis of 10% (w/ v) of Avicel PH-101 (Sigma Aldrich) and BMCC (bacterial microcrystalline cellulose) (CAMPOS et al., 2016; PINTO et al., 2015) were evaluated in reactions containing approximately 40 μg of enzyme in 40 mM of sodium acetate pH 5.3, incubated at 45 °C for 20 hours. The optimal temperature of enzymatic activity was evaluated in reactions containing 40 mM of sodium acetate pH 5.3 and 0.25% (m/ v) of oat β-glucan, incubated in a temperature range from 15 °C at 65 °C for 60 minutes, followed by determination of the amount of reducing sugars released. Subsequent, using 40 mM sodium acetate (pH 3.6 - 5.6) and 40 mM sodium phosphate (pH 5.8 - 8.0) (MOHAN, 2003), the optimal pH value was evaluated in reactions containing 0.25% (m/ v) of oat β-glucan, incubated at 45 °C for 60 minutes, followed by determination of the amount of reducing sugars released.

The thermal stability was assessed by incubating approximately 0.4 μg of the purified enzyme in 20 mM sodium phosphate and 50 mM NaCl pH 7.4 at temperatures of 40 ºC to 90 ºC, with the removal of aliquots at regular intervals of 0 to 360 minutes, followed by the evaluation of the residual enzymatic activity at 45 °C for 30 minutes in 0.25% (w/ v) oat β-glucan containing 40 mM sodium acetate pH 5.3.

To investigate the effect of metal ions on the enzymatic activity, two experiments were performed independently and simultaneously, considering that a small variation of ions in the purified enzymes could interfere in the assays. The enzymatic activity was determined in reactions containing 0.4 μg of protein mixed with 5 mM of different compounds (final concentration) (MO et al., 2010), 0.25% (w/ v) oat β-glucan and 40 mM sodium acetate pH 5.3, followed by incubation for 60 minutes at 45 ºC. For the analysis of the cation removal effect on the enzymatic activity, the chelating agent EDTA (ethylenediamine tetra-acetic acid - Calbiochem®) was used in the final concentration of 5 mM with the enzyme solution for 30 minutes at 4 °C, and then residual activity was verified at 45 °C for 60 minutes, without chelant removal. It was also evaluated the addition of 5 mM $CaCl_2$ in reactions containing EDTA-treated enzyme (without chelant removal) (LEE et al., 2014) to verify the ability of this metal salt to restore enzymatic activity. Moreover, the enzymatic activity was evaluated in the presence of 1% (v/ v) Triton ™ X-100 (Sigma-Aldrich) and 1 mM SDS (Sodium n-Dodecyl Sulfate - CALBIOCHEM®) in 0.25% (w/ v) of oat β-glucan containing 40 mM sodium acetate pH 5.3, followed incubation for 60 minutes at 45 °C. Likewise, reactions containing these agents were evaluated in the presence of 5 mM $CaCl_2$ and incubated at 45 °C for 60 minutes in

the optimum conditions. The relevance of the effect of these compounds on the enzymatic activity was verified through two-tailed Student's *t*-Test for two samples assuming equivalent variances in Microsoft Excel (2016), considering p-value <0.05.

## 2.5. *Determination of cleavage pattern*

The cleavage pattern was determined by capillary electrophoresis from the analysis of the hydrolysis product in 5 mM oligosaccharide (cellobiose-C2, cellotriosis-C3, cellotetraose-C4 and cellopentaose-C5, all from Megazyme), 0.25% (m/ v) of oat β-glucan and 10% (m/ v) Avicel PH-101 (Sigma Aldrich), using 0.4 µg and 40 µg of protein, respectively, after incubation at 45 °C for approximately 16 hours. Afterward the fractions were passed through vacuum drying and were derivatized with APTS fluorophore (8-Aminopyrene-1,3,6-Trisulfonic Acid, Trisodium Salt) by reductive amination (EVANGELISTA; LIU; CHEN, 1995). Following the conditions described in Alvarez *et al.* (2015), the separations were observed in the P / ACE MDQ system (Beckman Coulter) with laser induced fluorescence detector equipped with fused silica capillary (TSP050375, Polymicro Technologies).

## 2.6. *Binding assay*

The binding assays in insoluble polysaccharides were performed based on described previously (CAMPOS et al., 2016), with some modifications. An amount of 10 µg of $CelE2_{382-477}$ were incubated with 3% (m/ v) of Avicel and BMCC containing 25 mM ammonium acetate buffer pH 5.0 (200 µl) for 60 minutes at 4 °C under stirring. After incubation, the reactions were centrifuged at 12,000 xg for 5 minutes. The soluble fraction was recovered and concentrated. The insoluble fractions were washed three times with 25 mM ammonium acetate pH 5.0 containing 1 M NaCl. Both soluble and insoluble fractions were evaluated by SDS-PAGE.

The binding assays in soluble polysaccharides were also performed by affinity gel electrophoresis based on described previously (COCKBURN; WILKENS; SVENSSON, 2017; JAM et al., 2016), with some modifications. β-glucan from oat, xylan from beechwood and mannan were incorporate into the gel (6% bis-acrylamide) prior to polymerization at the final concentration of 0.25% (v/ v), using approximately 14 µg of $CelE2_{382-477}$. A control gel without polysaccharides was prepared and run simultaneously. As a negative control for polysaccharide

binding, approximately 7 μg of bovine serum albumin (BSA) was added to the gels as a comparison parameter. Electrophoresis was conducted in a vat containing all gels at 120 V for 165 minutes at 4 °C.

## 2.7. *Circular dichroism (CD) and thermal denaturation analysis*

Far-UV circular dichroism spectra (195-260 nm) and thermal denaturation were performed according to Mandelli *et al.* (2013) and Pimentel *et al.* (2017) using the JASCO 815 spectropolarimeter (JASCO Inc., Tokyo, Japan), equipped with a Peltier temperature control unit. The purified enzymes were diluted using 100 mM sodium acetate buffer pH 5.3 to the final concentration of 0.2 mg.mL$^{-1}$ in the presence and absence of 5 mM CaCl$_2$. Thermal denaturation analyzes were performed by monitoring the molar ellipticity for CelE2, CelE2$_{1-381}$ and CelE2$_{382-477}$, respectively at 222 nm, 221 nm and 215 nm, with a temperature increase of 20 to 100 °C at a rate of 1 °C.min$^{-1}$.

## 2.8. *Crystallization and processing of X-ray data*

The enzyme CelE2$_{1-381}$ eluted from size exclusion chromatography was concentrated to 10 mg.mL$^{-1}$ and submitted to initial crystallization experiments. The first screenings were set up with a Honey Bee 963 robot from ROBOLAB facility at Brazilian Biosciences National Laboratory (LNBio – Campinas, Brazil), using commercial kits as initial conditions (Hampton). The drops were made with 0.5 μl protein sample plus 0.5 μl reservoir solution in sitting drop vapor diffusion plates. The plates were incubated at 291 K. A second round of crystallization experiments were prepared in order to refine the conditions where the first crystals grown. This time, a hanging drop vapor diffusion technique was used with the drops containing 2 μl protein sample plus 2 μl reservoir solution.

Prior to data collection, crystals were soaked in a cryoprotectant solution (reservoir solution incremented with 20% ethylene glycol) and mounted directly on the goniometer at the beamline. The data sets were collected at MX-2 beamline at the Brazilian Synchrotron Light Laboratory (LNLS - Campinas, Brazil) equipped with a Pilatus 6M detector.

The collected data were indexed and integrated with XDS (KABSCH, 2010) and scaled with Aimless (EVANS, 2011). Initial phases were found by molecular replacement with Phaser (MCCOY et al., 2007), using the structure of an endoglucanase from *Acidothermus*

*cellulolyticus* (PDBid 1ECE – 54% identity) as the search model. After several rounds of manual adjustments of the model with COOT (EMSLEY et al., 2010) and structure refinements with Phenix (ADAMS et al., 2010) the final structure factors and model were validated with Molprobity (CHEN et al., 2010).

## 2.9. *SAXS analyzes*

SAXS data were acquired at the SAXS1 beamline of the Brazilian Synchrotron Light Laboratory (LNLS − CNPEM). The purified proteins were diluted using 100 mM sodium acetate buffer pH 5.3 at concentrations of 2 mg.mL$^{-1}$ in the presence of 5 mM $CaCl_2$, and 1.0 mg. mL$^{-1}$ and 3 mg. mL$^{-1}$ without the addition of $CaCl_2$, and the analysis were conducted at 20 °C. The scattering patterns were recorded using a Pilatus detector. The wavelength of the incident radiation was set at $\lambda = 1.54$ Å, with a detector-sample distance of 908 mm, resulting in a scattering vector q (defined by $q = 4\pi\ sin\ (\theta)\ /\ \lambda$, where $2\theta$ is the scattering angle) ranging from $0.012 <q <0.466$ Å-1. The data were corrected for detector response and staggered by beam intensity and samples uptake. The buffer scattering was subtracted from the scattering of the corresponding samples and integration of the SAXS standards was performed using Fit2D software (HAMMERSLEY, 1997).

Guinier analyzes were conducted using the Primus program interface (KONAREV et al., 2003) of the ATSAS package. The adjustment of the experimental curve and the evaluation of the distribution function of the distance pairs, *p (r)*, were conducted using the Gnom software (SVERGUN, 1992) of the ATSAS package. The low resolution dummy (DAM) model was generated using the *ab initio* routine implemented in the Dammif and Dammin programs (SVERGUN, 1999). The FoXS software (SCHNEIDMAN-DUHOVNY et al., 2016) was used to simulate the scattering pattern and to evaluate the parameters of the crystallographic structure. The low resolution model generated from the SAXS curve and the crystallographic model were overlaid using the SUPCOMB program (KOZIN; SVERGUN, 2001).

# 3. Results and discussion

## 3.1. *Sequences analyzes and production of recombinant proteins*

The full-length CelE2, recently characterized by Pimentel *et al.* (2017), has in its modular architecture two conserved domains predicted on NCBI, the N-terminal predicted as a GH5 and a C-terminal domain predicted as a Calx-β domain. The Calx-β domains are found in different organisms such as *Drosophila melanogaster, Physarum polycephalum* and *Amphimedon queenslandica* and, in some cases, are associated in calcium interaction and binding (GRICE et al., 2017; MAEKAWA et al., 2006; SCHWARZ; BENZER, 1997). These domains are also found to be associated to glycoside hydrolases (THE UNIPROT CONSORTIUM, 2017), but have no defined function in these cases.

Two constructs were designed for the catalytic domain (CelE2$_{1-381}$ and CelE2$_{1-363}$) and two for the Calx-β accessory domain (CelE2$_{364-477}$ and CelE2$_{382-477}$) (Fig. 1A). The sequence of each protein was analyzed by the ProtParam tool for determination of biochemical and biophysical parameters (Table 2), considering N-terminal 6-His tag and other fused amino acids. The recombinant proteins were produced and purified by affinity and size exclusion chromatography, and SDS-PAGE analysis indicated protein bands with molecular masses corresponding to the theoretical values (Fig. 1B).
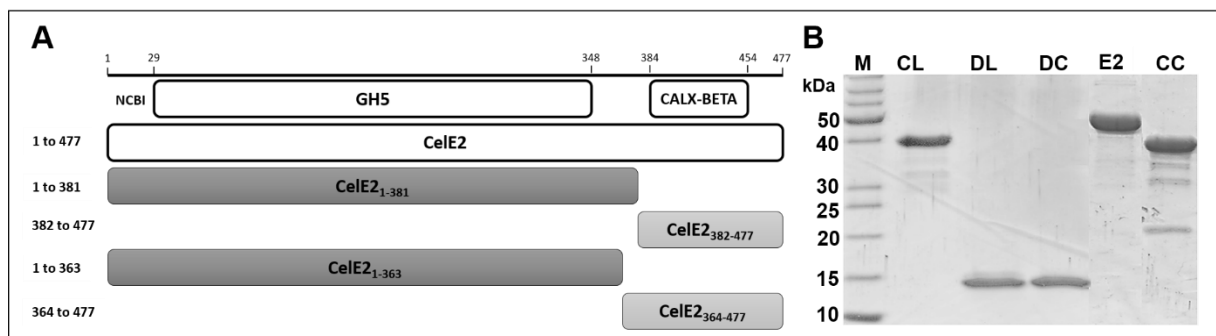


**Fig. 1 A:** The modular schematics of CelE2 predicted by NCBI and the constructs used in this study. Numbers indicates amino acid interval of each construct. **B:** SDS-PAGE of the purified proteins after affinity and size exclusion chromatography. M - molecular weight marker (PageRuler ™ Unstained Protein Ladder - 10-250 kDa, Thermo Scientific), CL - CelE2$_{1-381}$, DL - CelE2$_{364-477}$, DC - CelE2$_{382-477}$, CC - CelE2$_{1-363}$ and E2 - CelE2 Full-lenght.

**Table 2** Biochemical and biophysical parameters of each construct obtained by the Protparam tool.

| Proteins | Amino acid Number | Molecular Weight (kDa) | Isoelectric Point | Molar Extinction Coefficient |
|----------|-------------------|------------------------|-------------------|------------------------------|
| CelE2 | 500 | 53.8 | 4.95 | $123.300 \ m^{-1}cm^{-1}$ |
| CelE2$_{1-381}$ | 404 | 44.2 | 5.19 | $120.320 \ m^{-1}cm^{-1}$ |
| CelE2$_{1-363}$ | 386 | 42.2 | 5.02 | $117.340 \ m^{-1}cm^{-1}$ |
| CelE2$_{382-477}$ | 119 | 12.0 | 6.10 | $2.980 \ m^{-1}cm^{-1}$ |
| CelE2$_{364-477}$ | 137 | 14.0 | 6.63 | $5.960 \ m^{-1}cm^{-1}$ |

*3.2.  Phylogenetic analysis*

CelE2$_{1-381}$ has similarity to sequences belonging to subfamily 1 of the GH5 family (GH5_1) from the dbCAN database. Phylogenetically, the amino acid sequence of CelE2$_{1-381}$ is closely related to the GH5 catalytic domain of an endoglucanase from *Azorhizobium caulinodans* (accession number BAF87299.1 – 62% identity), and with the GH5 catalytic domain of an enzyme annotated as chitinase, as well as, with an endoglucanase from *Mycobacterium gilvum* (accession number ABP45478.1 – and ADT98979.1, respectively) (both with 70% identity) (Fig. 2). It is noteworthy that all these sequences have Calx-β domains in their architecture, according to the Pfam and Smart annotations. In addition, the higher boostrap value denoted a consistent phylogenetic relationship among the sequences.
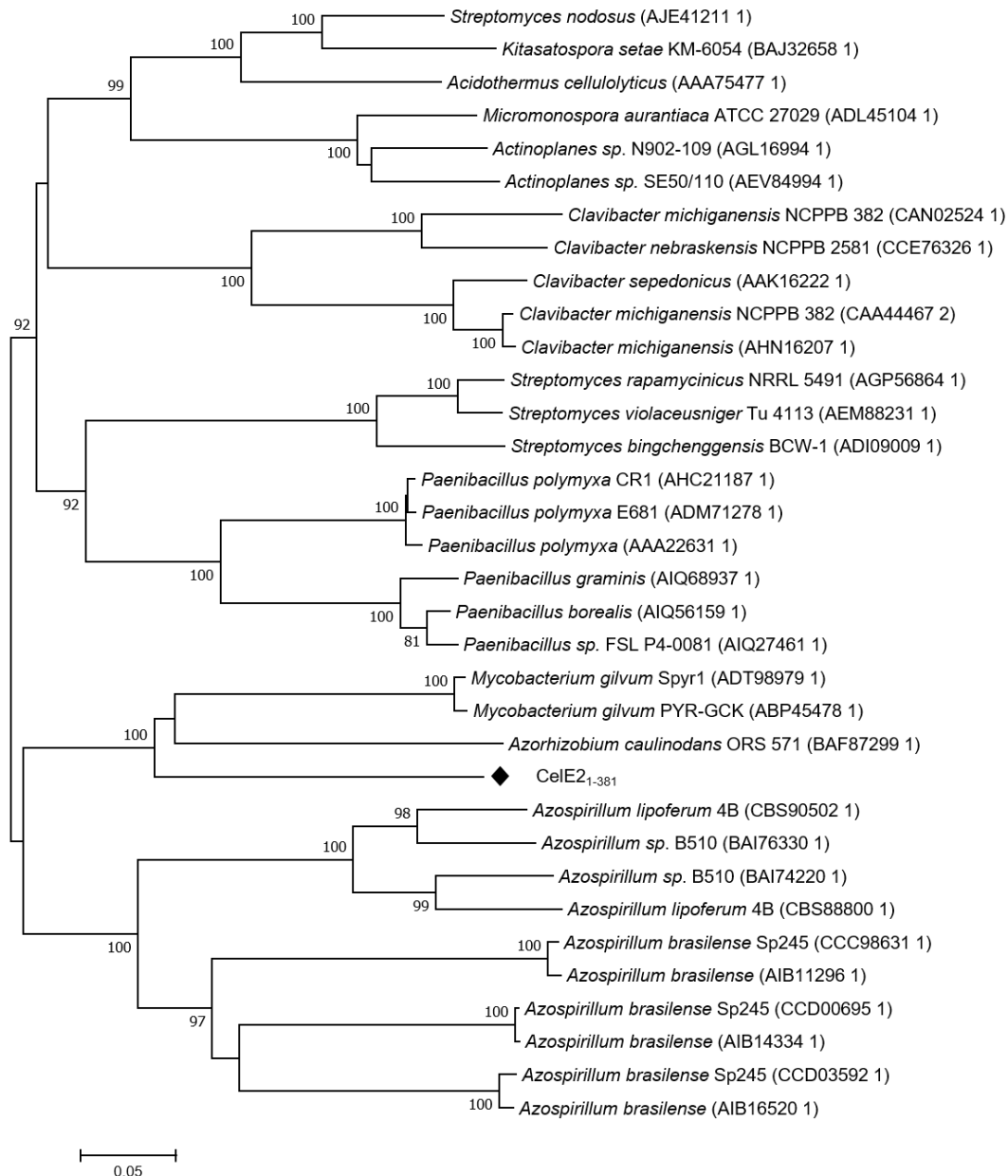
**Fig. 2.** Phylogenetic position of CelE2$_{1-381}$ relative to related sequences recovered from dbCAN based on amino acid sequences. The tree was construct using Neighbor-Joining method implemented in MEGA7 (KUMAR; STECHER; TAMURA, 2016; SAITOU; NEI, 1987) and the evolutionary distance was based on p-distance method (NEI; KUMAR, 2000). Genbank accession numbers are listed in parentheses. The bootstrap values obtained by repeating the analysis 1000 times are shown in percentages at the nodes.

The best blast hit of CelE2$_{382-477}$ (e-value $2.91^{e-27}$, score 97.8) was for the group ENOG4107R46 referring to category of transport and metabolism of carbohydrates, annotated as MAFp3 aggregation core factor, isoform C. The tree (Fig. 3) indicated a close association of CelE2$_{382-477}$ for a Na-Ca exchanger/integrin-beta4 *Nitrosomonas sp.* (accession number

WP_013646438.1) (49% identity). Unlike CelE2$_{1-381}$, the low boostrap value indicates phylogenetic divergence among CelE2$_{382-477}$ and the respective sequences.
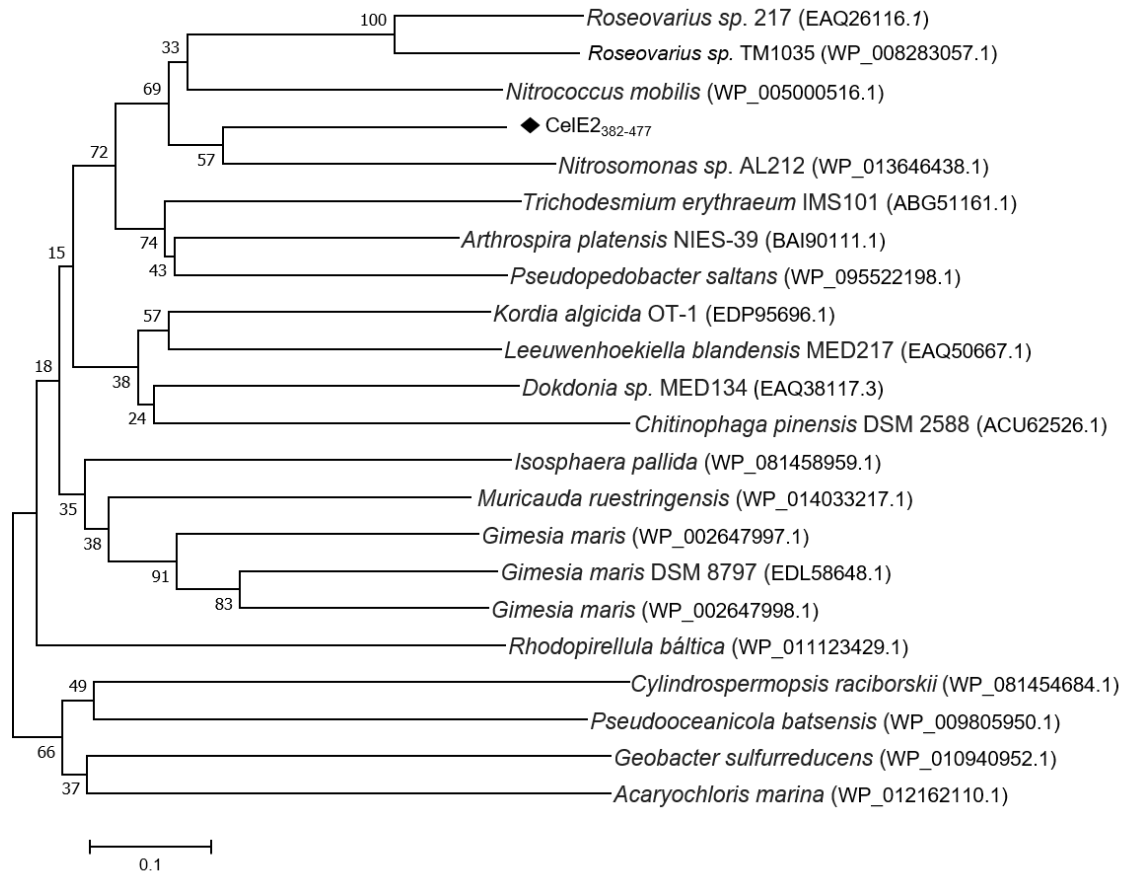


**Fig. 3.** Phylogenetic position of CelE2$_{382-477}$ relative to related sequences recovered from Eggnog database based on amino acid sequences. The tree was construct using Neighbor-Joining method implemented in MEGA7 (KUMAR; STECHER; TAMURA, 2016; SAITOU; NEI, 1987) and the evolutionary distance was based on p-distance method (NEI; KUMAR, 2000). Genbank accession numbers are listed in parentheses. The bootstrap values obtained by repeating the analysis 1000 times are shown in percentages at the nodes.

## 3.3. The Effect of Calx-β domain on the enzymatic properties

Among the evaluated constructs (CelE2$_{1-381}$, CelE2$_{1-363}$, CelE2$_{382-477}$ and CelE2$_{364-477}$), only CelE2$_{1-381}$ showed enzymatic activity on oat β-glucan (which was considered as relative activity of 100% ± 0%), lichenan (relative activity 84% ± 13%) and CMC (relative activity 21% ± 2%), which shared the same substrate specificity with full-length CelE2 (Fig. 4A). Therefore, this construct was chosen for further studies.

In the Avicel hydrolysis, an insoluble substrate containing approximately 40% amorphous region (HALL et al., 2010), CelE2$_{1\text{-}381}$ showed activity approximately 76% higher than the full-length CelE2 (specific activity of $1.7 \pm 0.0$ µmol/hour/mg and $0.4 \pm 0.0$ µmol/hour/mg, respectively) (Fig. 4B), suggesting that this result is directly related to the excision of the Calx-β accessory domain. Although a large number of enzymes are dependent on their accessory domains, such as BlCel5B (LIBERATO et al., 2016), studies suggest that increase on protein flexibility or of the catalytic residues may promote an increase in catalytic efficiency, as occurs with the removal of the Ig-like domain of the endoglucanase Cel9A (BHABHA et al., 2011; YOUNESI et al., 2016). Thus, it is possible that the removal of the Calx-β domain promotes higher flexibility for the catalytic domain on insoluble polysaccharides. However, further studies would be needed to evaluate the structural flexibility of CelE2 and CelE2$_{1\text{-}381}$ and its effect on the enzymatic activity, similar to what has been described in Younesi *et al.* (2016). In contrast, no enzymatic activity was observed for both constructs on the BMCC substrate (Fig. 4B), which compared to Avicel has a high degree of crystallinity (approximately 95% crystallinity (PARK et al., 2010)) and few amorphous regions.
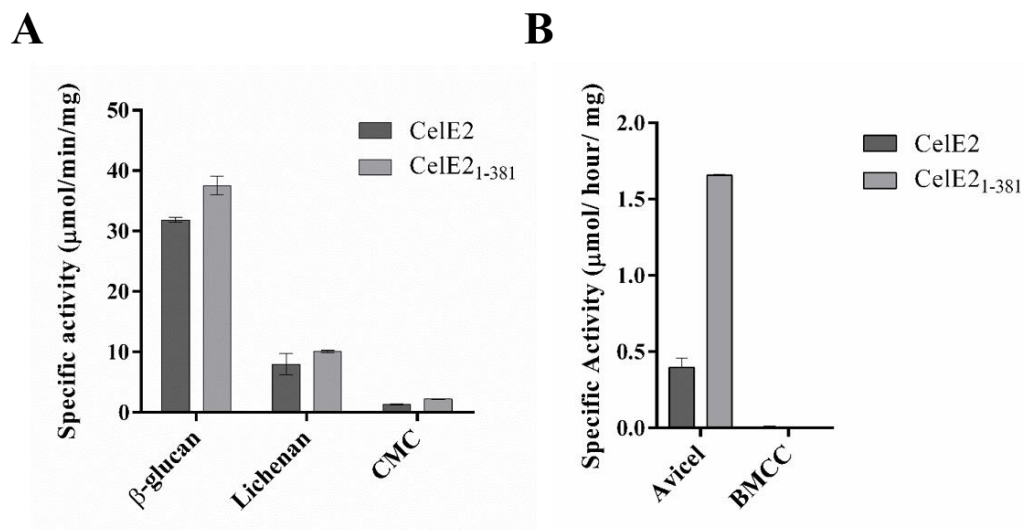


**Fig. 4. (A)** Specific activity of CelE2 and CelE2$_{1\text{-}381}$ on the preferred substrates β-glucan, lichenan and CMC (final concentration of 0.25% - m/ v) for 60 minutes and **(B)** in Avicel and BMCC (final concentration of 10% - m/ v) for 20 hours.

The biochemical properties such as optimal pH and temperature shown for CelE2 (PIMENTEL et al., 2017) were maintained for the truncated enzyme CelE2$_{1\text{-}381}$, with optimal

temperature at 45 ºC and pH of 5.3 (Fig. 5A and 5B). The same could be observed with respect to thermal stability (Fig. 5C), which CelE2$_{1-381}$ showed considerable thermal stability at 40 °C after 360 minutes of incubation, and at 50 °C and 60 °C was able to maintain, respectively, about 97% and 48% of the enzymatic activity, after that period. The removal of the Calx-β accessory domain also did not interfere with enzymatic activity at elevated temperatures, and after 360 minutes of incubation at 70 °C, CelE2$_{1-381}$ showed approximately 8% of residual activity, and about 52% and 20% after 30 minutes incubation at 80 °C and 90 °C, respectively. Frequently, the removal of an accessory module results in the change of the optimum temperature to a lower value and a decrease in thermal stability (ARAKI et al., 2004). However, CelE2$_{1-381}$ maintained biochemical characteristics similar to full-length CelE2. Unlike the endoglucanase rPglA from *Paenibacillus sp.*, which the exclusion of one or more accessory domains promoted change in the enzymatic characteristics (CHENG et al., 2013).
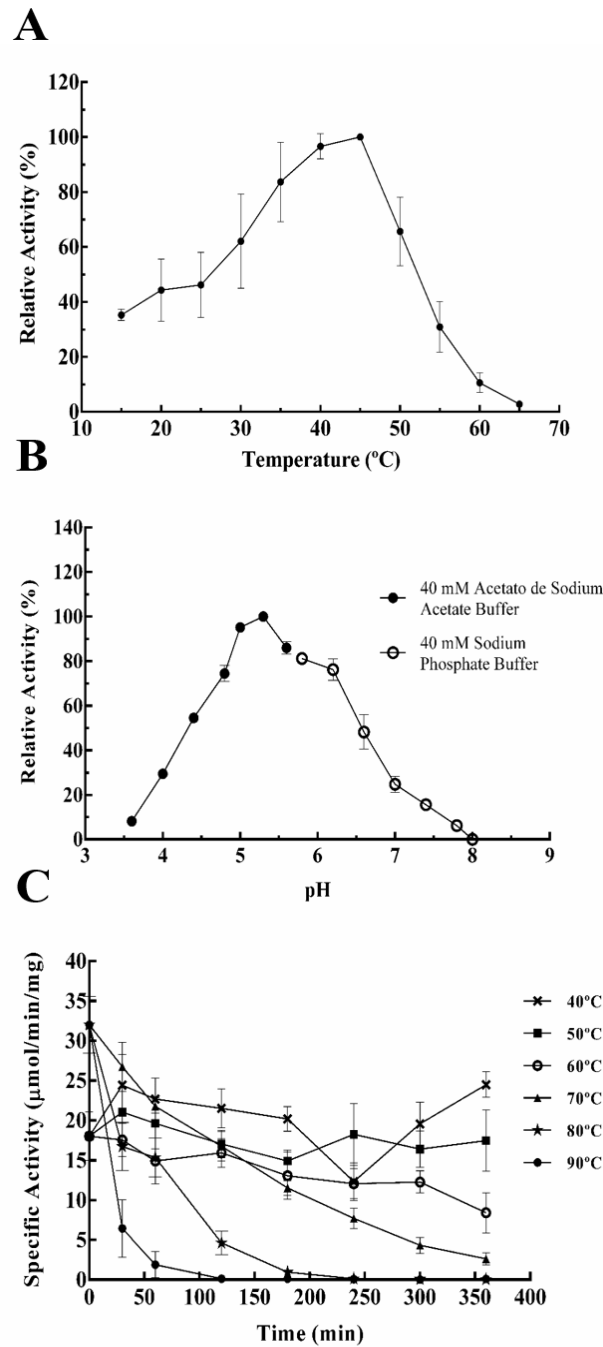
**Fig. 5.** The enzymatic characteristics of truncated CelE2$_{1-381}$ using β-glucan (0.25% m/ v). **A:** Influence of different temperatures on the enzymatic activity (15 ℃ – 65 ℃) with incubation of 60 minutes. **B:** Influence of different pHs using 40 mM sodium acetate (3.6 - 6.6) and 40 mM sodium phosphate (5.8 - 8.0) at optimal temperature (45 ℃) for 60 minutes. **C:** Thermal stability of purified CelE2$_{1-381}$ incubated from 0 to 360 minutes at temperatures of 40 ℃ to 90 ℃, followed by analysis of residual activity under optimal conditions for 30 minutes.

### 3.4. Effect of metal ions and others compounds

The presence of metal ions on the enzymatic activity of the truncated $CelE2_{1-381}$ showed similar effects to those observed for the full-length CelE2 (PIMENTEL et al., 2017). The addition of 5 mM of different metal ions on the $CelE2_{1-381}$ enzymatic activity, showed significant effects (p-value $<0.05$) of $CaCl_2$ and $CoCl_2$ activation, and inhibition by $CuCl_2$, $FeCl_3$, $FeSO_4$ and $ZnSO_4$ (Fig. 6A). Although the effect of metal ions on the enzymatic activity of endoglucanases is quite varied, such as $Zn^{2+}$, $Fe^{2+}$ and $Cu^{2+}$ that can be found as activators or inhibitors, $Ca^{2+}$ and $Co^{2+}$ are frequently associated to activation and $Fe^{3+}$ to inhibition of the enzymatic activity of these enzymes (PEREIRA et al., 2017). However, calcium activation for $CelE2_{1-381}$ suggests that the Calx-β domain is not directly involved in the activation of enzymatic activity by calcium ion. In integrins β4 of cell-surface adhesion receptors, the absence of cations in the three-dimensional structure of the Calx-β domain, in addition to biophysical data, suggested that this Calx-β domain does not bind to calcium ion (ALONSO-GARCÍA et al., 2009).

The calcium effect on the enzymatic activity of $CelE2_{1-381}$ was further evaluated through assays containing the enzyme treated with 5 mM EDTA and the 5 mM $CaCl_2$ supplementation after the same treatment. The treatment with EDTA reduced around 27% of the enzymatic activity. However, the addition of $CaCl_2$ containing the same treatment was able to completely restore the enzymatic activity, confirming that like CelE2, the enzymatic activity is also influenced by the presence of divalent cations in the purified enzyme (Fig. 6B).

Regarding the influence of Triton X-100 and SDS surfactants, $CelE2_{1-381}$ activity maintained similar effects to those presented for CelE2, with a significant reduction of activity by 1% of Triton X-100 (approximately 58% residual activity) and complete inactivation with 1 mM SDS. However, the addition of 5 mM $CaCl_2$ in the reactions with these surfactants promoted a positive effect only with Triton X-100, so that $CaCl_2$ was not able to restore or decrease the SDS effects (Fig. 6C).
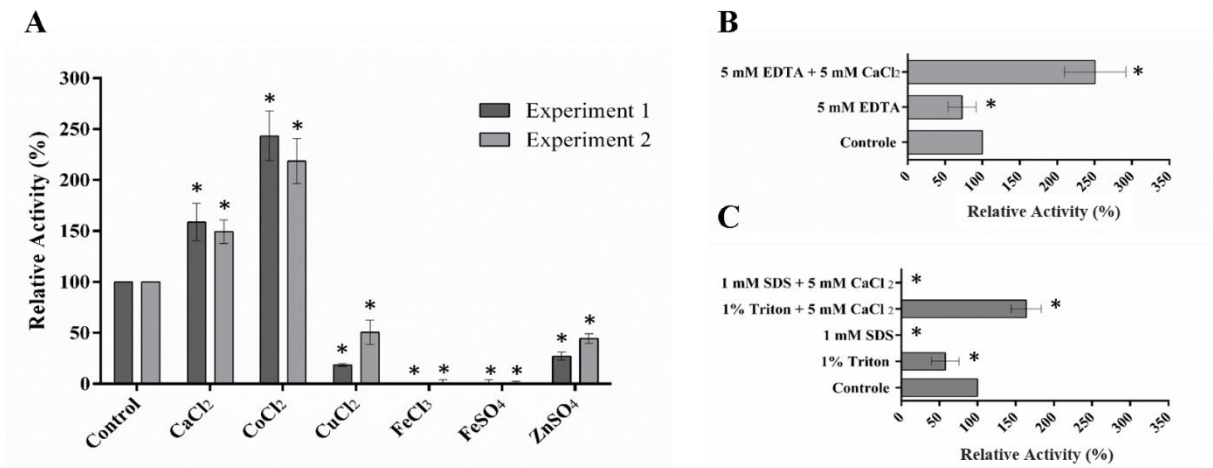
**Fig. 6. A:** Effect of metal ions on the enzymatic activity of CelE2$_{1-381}$ with the addition of 5 mM of different compounds, carried out under the optimal conditions for 60 minutes. **B:** Effect of the chelating agent EDTA in the presence e absence of 5 mM CaCl$_2$ with incubation for 30 minutes at 4 ºC, followed by incubation for 60 minutes under optimal conditions. **C:** Effect of 1 % Triton X-100 and 1mM SDS surfactants in the presence and absence of 5 mM CaCl$_2$, with incubation for 60 minutes under optimal conditions. * indicates statistically significant results (p-value <0.05) compared to respective controls using Student's *t*-Test.

### 3.5.  CelE2$_{1-381}$ hydrolysis and transglycosilation products

The capillary electrophoresis indicated that there is no change in the hydrolysis and transglycosylation products after removal of the Calx-β. Regarding to the degradation of the major soluble polysaccharide β-glucan, CelE2$_{1-381}$ was able to generate oligosaccharides of various sizes, whereas in Avicel it was able to generate only di- and trisaccharides (Fig. 7).

The evaluation of the cellooligosaccharide hydrolysis product showed that CelE2$_{1-381}$ was not able to cleave C2 and C3 (data not shown), only C4 and C5, as it was observed for CelE2. CelE2$_{1-381}$ maintained evidence of transglycosylation in the C4 hydrolysis, with the presence of di-, tri- and tetrasaccharides without resulting in glucose, and with the appearance of peaks with a higher degree of polymerization than the C4 substrate (Fig. 7). The hydrolysis of C5 generated mainly di- and trisaccharides, but without clear evidence of transglycosylation. Besides the cleavage profile of endoglucanases be characteristic for the random attack of the polysaccharides with the production of oligosaccharides of different sizes, the transglycosylation is also described for other GH5 (MORRILL et al., 2015; ZHOU et al., 2016).
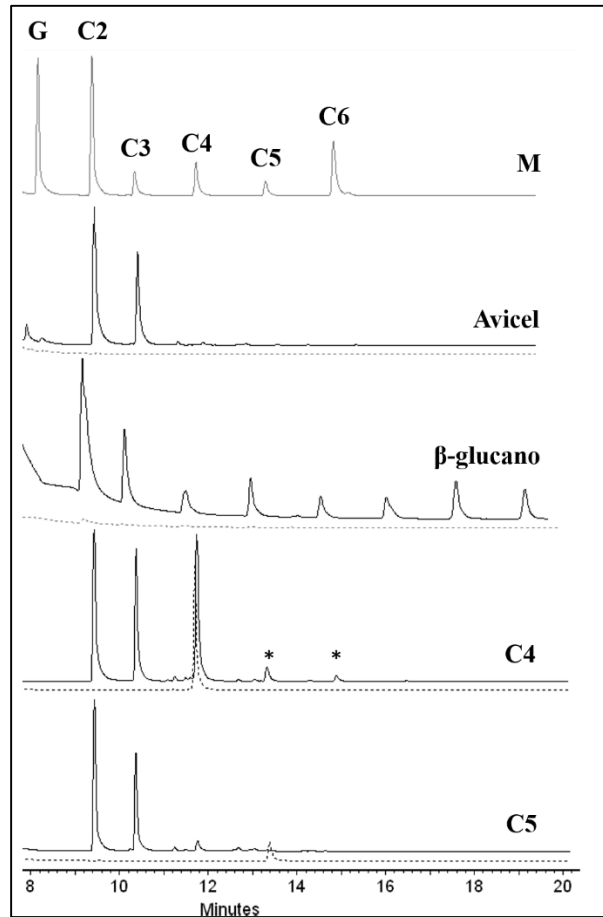
**Fig. 7.** Cleavage pattern of CelE2$_{1-381}$ by capillary electrophoresis evaluated through the hydrolysis on Avicel, β-glucan, C4 and C5 substrates for 16 hours in optimal conditions. The gray line represents the standard containing G (glucose), C2 (cellobiose), C3 (cellotriose), C4 (cellotetraose), C5 (cellopentaose) and C6 (cellohexaose). The black lines represent the hydrolysis profile, together with the respective control reactions represented by the dotted lines. * indicate peaks with degrees of polymerization higher than the substrate.

### 3.6. Binding of CelE2$_{382-477}$ in soluble and insoluble substrates

The interaction with insoluble substrates was evaluated (Fig. 8), and no evidence of specific binding was observed due to the major presence of CelE2$_{382-477}$ in the soluble fractions, which represent proteins not bound to the substrate. Furthermore, the assays performed with β-glucan, xylan and mannan by affinity gel electrophoresis also showed no interaction of CelE2$_{382-477}$ with those soluble substrates (data not shown).

The absence of tryptophan residues in the sequence of CelE2$_{382-477}$ theoretically corroborates its lack of interaction of Calx-β with polysaccharides, since tryptophan residues

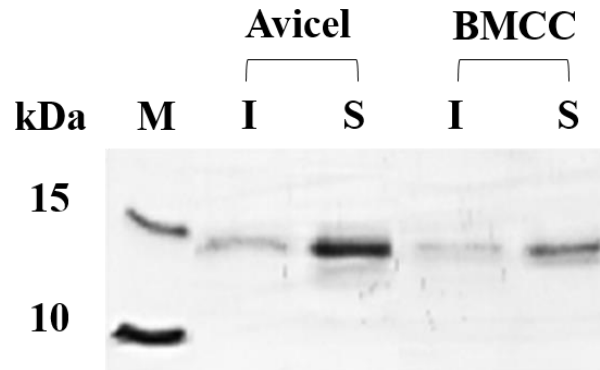generally plays a key role in the binding with polysaccharides (GUILLÉN; SÁNCHEZ; RODRÍGUEZ-SANOJA, 2010).



**Fig. 8.** SDS-PAGE gel from the supernatant fractions and precipitate obtained by centrifugation after incubation of CelE2$_{382-477}$ with 3% of insoluble polysaccharides, for 60 minutes ate 4 ℃. M: molecular weight marker (PageRuler ™ Unstained Protein Ladder - 10-250 kDa, Thermo Scientific); I: insoluble fractions and S: soluble fractions of the assay containing 10 μg of CelE2$_{382-477}$.

## 3.7. *Proteins CD spectra and thermal unfolding*

Full-length CelE2 CD spectra and thermal denaturation were evaluated using the 100 mM sodium acetate buffer pH 5.3 in the absence of 5 mM CaCl$_2$. As observed by Pimentel *et al.* (2017) in sodium phosphate buffer, CelE2 showed a similar profile with two negative bands around 206 and 222 nm and a positive band near 200 nm, maintaining the secondary structure folding profile of alpha-beta proteins (GREENFIELD, 2007) (Fig. 9A). Analysis of the temperature-induced changes at 206 and 222 nm revealed that the melting temperature (Tm) estimated at 222 nm through the sigmoidal fitting was 57 °C. However, the presence of 5 mM CaCl$_2$ increased the Tm value to 60 °C (Fig. 9B).

As regards the secondary structure of CelE2$_{1-381}$, a predominant α-helice profile was observed with two negative bands around 207 and 221 nm and a positive band near 195 nm (Fig. 9C) (CORRÊA; RAMOS, 2009). In addition, the thermal denaturation estimated at 221 nm from the sigmoidal adjustment showed that the presence of CaCl$_2$ promoted an increase in stability similar to that observed for CelE2, with a Tm of 52.4 °C in the absence of CaCl$_2$ and of 58.5 °C in the presence of CaCl$_2$ (Fig. 9D).

Regarding to CelE2$_{382-477}$, the CD spectra at 20 °C indicated a predominant profile of β-strands, with a negative band at 215 nm and a positive band around 200 nm (CORRÊA; RAMOS, 2009) (Fig. 10A). The Tm values obtained at 215 nm were smaller compared to the values of CelE2 and CelE2$_{1-381}$, and there was no increase in Tm in the presence of CaCl$_2$, being 49.5 ºC in the absence of CaCl$_2$ and 48.4 ºC in the presence of CaCl$_2$ (Fig. 10B).

In summary, these results revealed that the presence of CaCl$_2$ conferred higher thermal stability for CelE2 and CelE2$_{1-381}$ due to the increase in Tm values. The removal of the Calx-β accessory domain resulted in a Tm reduction of 4.6 ºC. In contrast, in the presence of CaCl$_2$, this difference was reduced to 1.5 ºC. In addition to changes in biochemical parameters, the removal of accessory modules may cause changes in Tm values as observed for the TpMan (GH5) and Amy703 (GH13) glycoside hydrolases in which removal of the C-terminal domain respectively promoted the decrease of 12 ºC and the increase of 7 ºC in the Tm value, respectively (LU et al., 2016; SANTOS et al., 2012).

Collectively, our data indicated that the presence of the accessory domain favors the thermal stability, as well the presence of CaCl$_2$ promotes additionally increase in thermal stability for both constructs.
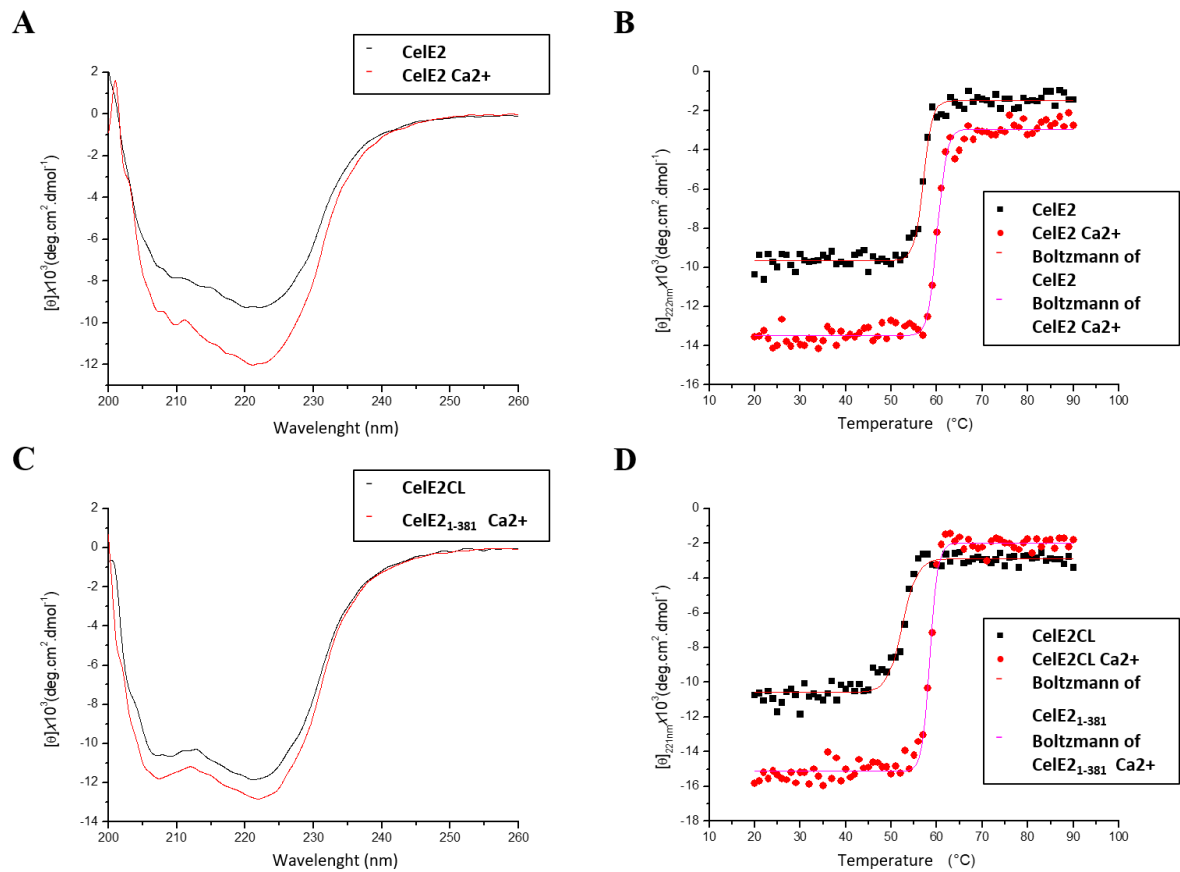
**Fig. 9.** Circular dichroism of CelE2 and CelE2$_{1\text{-}381}$. CD spectra in the far ultraviolet region, where the black line represents the CelE2 **(A)** and CelE2$_{1\text{-}381}$ **(B)** spectra at 20 °C in 100 mM sodium acetate buffer pH 5.3 and the red line represents the spectrum at 20 °C with the addition of 5 mM CaCl$_2$ in the same buffer. CelE2 molar ellipticity **(B)** and CelE2$_{1\text{-}381}$ **(D)** at 222 and 221 nm, respectively, as a function of temperature. Bold squares represent absence of CaCl$_2$ and red circles represent the presence of CaCl$_2$. The continuous red and pink lines represent the Sigmoidal adjustment of the data obtained from thermal denaturation.
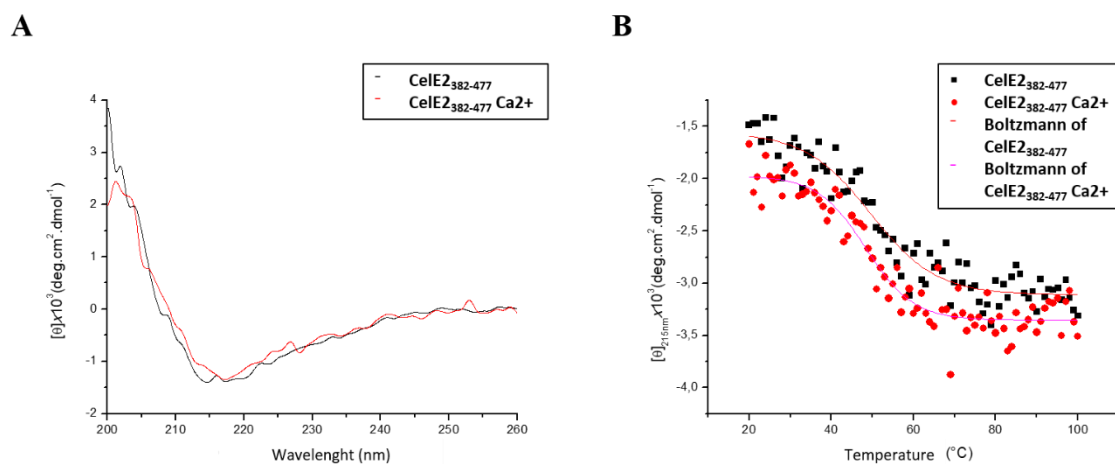
**Fig. 10. A:** CD spectra of CelE2$_{382-477}$. The black and red lines represent the spectrum in the absence and presence of 5 mM CaCl$_2$, respectively. **B:** CelE2$_{382-477}$ molar ellipticity at 215 nm as a function of temperature (bold squares for absence of CaCl$_2$ and red circles for presence of CaCl$_2$). The continuous red and pink lines represent the Sigmoidal adjustment of the data obtained from thermal denaturation.

### 3.8. The crystal structure of CelE2 catalytic domain

The crystal structure of CelE2$_{1-381}$ was obtained in condition containing 20% PEG3350, 0.2 M sodium bromide and 0.1 M BIS-TRIS propane, pH 6.5. The structure was solved to 2.1 Å resolution and the statistics from data collection and refinement are described in Table 3.

**Table 3** - Statistics from X-ray diffraction data collection and refinement.

| | CelE2$_{1\text{-}381}$ |
|---|---|
| **Data collection** | |
| Wavelength (Å) | 1.46 |
| Space group | P1 |
| Cell dimensions | |
|   *a, b, c* (Å) | 97.56, 97.82, 100.20 |
|   *α, β, γ*(°) | 66.0, 69.7, 70.0 |
| Resolution (Å) | 49.19–2.10 (2.14–2.10)* |
| Total reflections | 599255 |
| Unique reflections | 170614 |
| $R_{merge}$ | 0.13 (0.91) |
| $R_{pim}$ | 0.08 (0.61) |
| I/σI | 9.3 (1.4) |
| Completeness (%) | 94.7 (92.8) |
| Redundancy | 3.5 (3.2) |
| CC ½ (%) | 99.3 (43.3) |
| **Refinement** | |
| $R_{work}$ / $R_{free,}$ (%) | 19.17/22.30 |
| No. atoms | |
|   Protein | 22403 |
|   Solvent | 2264 |
|   Ions | 8 |
| *B*-factors | |
|   Protein | 33.8 |
|   Solvent | 36.4 |
|   Ions | 34.8 |
| R.m.s. deviations | |
|   Bond lengths (Å) | 0.003 |
|   Bond angles (°) | 0.54 |
| Ramachandran | |
|   Favored | 96.29 |
|   Allowed | 3.39 |
|   Outliers | 0.32 |

* Highest resolution shell is shown in parentheses.

The final model presented an octamer in the asymmetric unit and all monomers are practically identical. CelE2 has a typical folding of GH5 that is found in most of GHs, known as (α/β)8 barrel, where eight β-strands are surrounded by 8 α-helix. The loops that connect the secondary structures, at the carboxy-terminal side of the β-strands, form the catalytic cleft where the catalytic residues E179 and E297 can be found (Fig. 11). A bromide ion was modeled at the surface of the protein, coordinated by the asparagine N60 and lysine K62. However, this interaction is probably nonspecific. The bromide ion was present in the crystallization condition and it may be occupying a possible calcium binding site, since a modulation of the enzymatic

activity in the presence of this ion was observed. However, new experiments are needed to confirm this hypothesis.
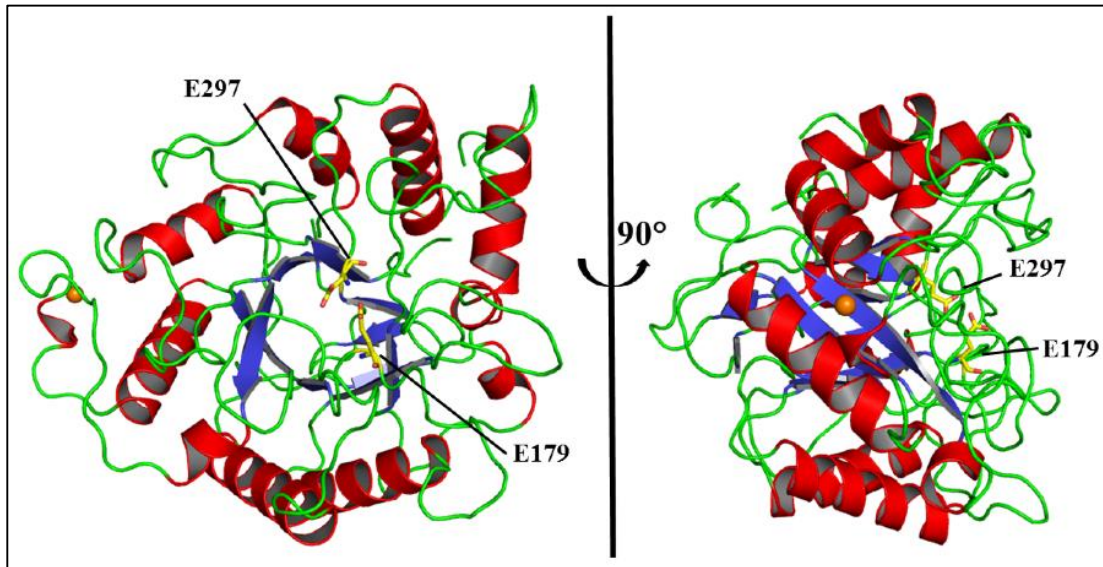


**Fig. 11.** Crystallographic structure of CelE2. The folding is composed of 8 β-strands (blue) surrounded by 8 α-helix (red), which are connected by loops (green). The catalytic residues (yellow) are located in the binding cleft. A bromide ion (orange sphere) was found non-specifically interacting with the protein at the surface of the structure.

According to the PDBefold server (KRISSINEL; HENRICK, 2004), CelE2$_{1-381}$ presents similar structural identity with representatives of GH5_1 subfamily (Fig. 12A). The most is an endocellulase from *Acidothermus cellulolyticus* (PDBid 1ECE) with a rmsd of 0.94 Å, followed by an endoglucanase from *Xanthomonas campestris* (PDBid 4TUF) and a hyperthermophilic endocellulase from *Pyrococcus horikoshii* (PDBid 3W6L), both with rmsd of 1.33 Å. According to a structural alignment of CelE2$_{1-381}$ with the 1ECE structure, complexed with a cellotetraose, it was observed that the active site of this group of enzymes is well conserved, presenting point mutations that in most cases do not alter the chemical character (Fig. 12B). The structural superposition reveals that cellotetraose, which extends from subsite -2 (non-reducing) to +2 (reductant), would be stabilized in CelE2$_{1-381}$ by hydrogen bonds with amino acids Q48, H130, T141, N178, E179, Q262, E297 and D349, in addition to CH-π hydrophobic interactions amino acids W228 and W260 (Fig. 13C). The only notable difference in relation to the other subfamily structures lies in the positioning of the arginine R131, whose side chain points to the solvent. The same residue in the three most similar structures forms hydrogen bonds with the substrate. The differential positioning in CelE2$_{1-381}$ is probably related

to steric hindrances caused by loop 3 (which connects the β3-strand with α-helix 3), and is more extensive in CelE2$_{1-381}$ than in the others (Fig. 12A).
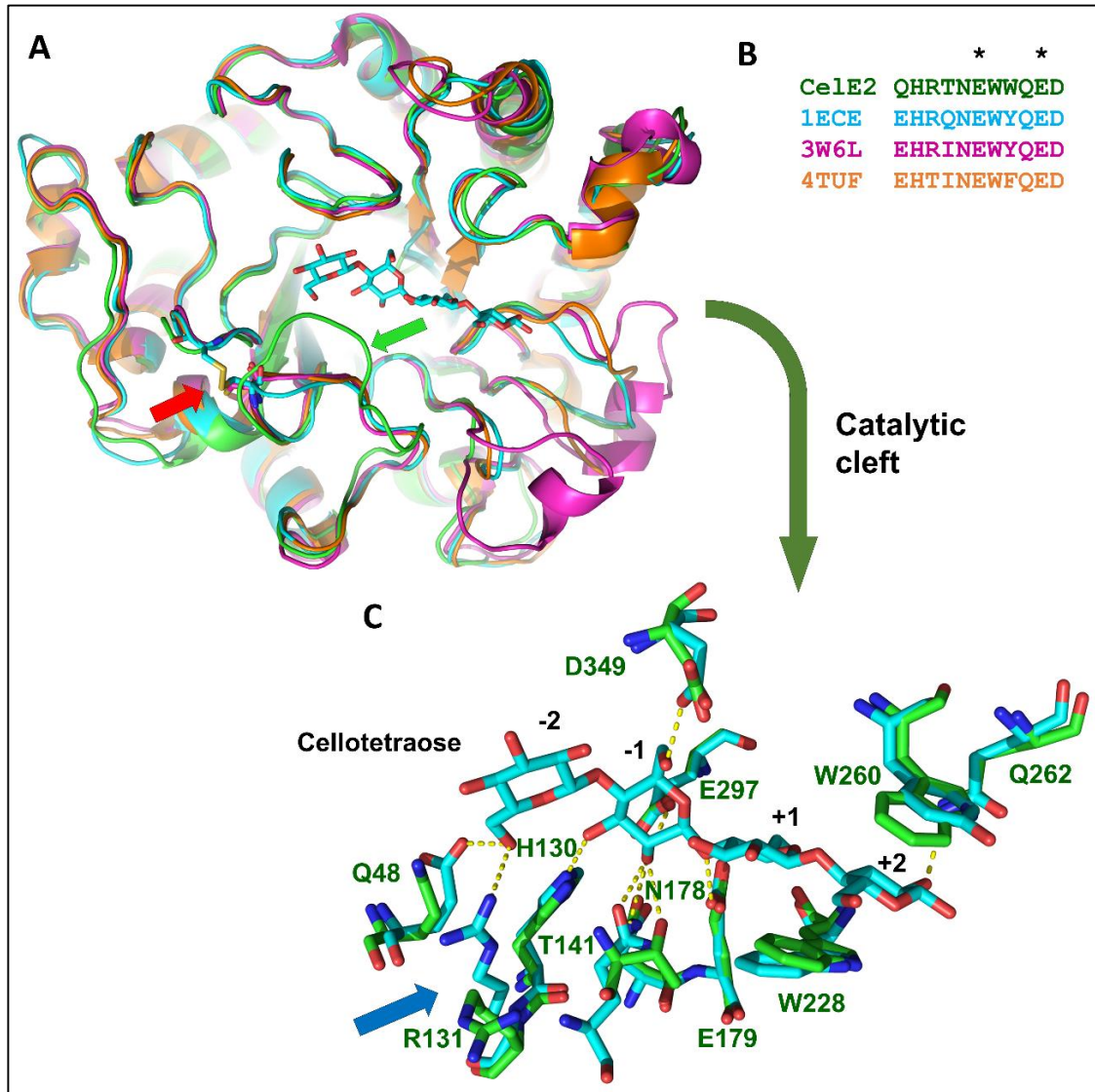


**Fig. 12.** Comparison of the CelE2$_{1-381}$ crystal structure with similar structures and inference of the active site. **(A)** Overlap of CelE2$_{1-381}$ structures (green), 1ECE (cyan), 3W6L (magenta) and 4TUF (orange). The arrows indicate the main points of difference. The position of loop 3 (green arrow) and the presence of disulfide bonds in the structures, with the exception of CelE2$_{1-381}$ (red arrow). **(B)** Alignment of the amino acids that make up the catalytic site of the compared structures, based on the 1ECE that is complexed with cellotetraose. **(C)** Composition of the active site of CelE2$_{1-381}$, superimposed on the 1ECE structure. The yellow dashes indicate hydrogen bonds between the protein and the substrate, which extends from the subsites -2 to +2. The blue arrow indicates difference in arginine R131 positioning of CelE2$_{1-381}$.

*3.9. SAXS data collection and analyzes*

SAXS analyzes showed results of $CelE2_{1\text{-}381}$ and $CelE2_{364\text{-}477}$ in the presence and absence of $CaCl_2$. A Guinier analysis of $CelE2_{1\text{-}381}$ within the limit q.Rg <1.3 (GUINIER, A. AND FOURNET, 1955) showed a significant decrease in the radius of gyration (Rg) in the presence of $CaCl_2$ (Table 4), suggesting that the calcium ion provides greater stability to the protein, as revealed by the decrease in slope of the line (Fig. 13A). Fig. 13B shows the distribution function of the distances pairs (*p (r)*), with the decrease in maximum diameter in the presence of calcium, evidencing a lower tendency of aggregation. The structural parameters of the experimental data obtained by SAXS and the simulated scattering of the crystal structure of $CelE2_{1\text{-}381}$ are shown in Table 4.

**Table 4.** Structural parameters obtained experimentally by SAXS and the simulated scattering of the crystallographic structure for $CelE2_{1\text{-}381}$ (Cat)

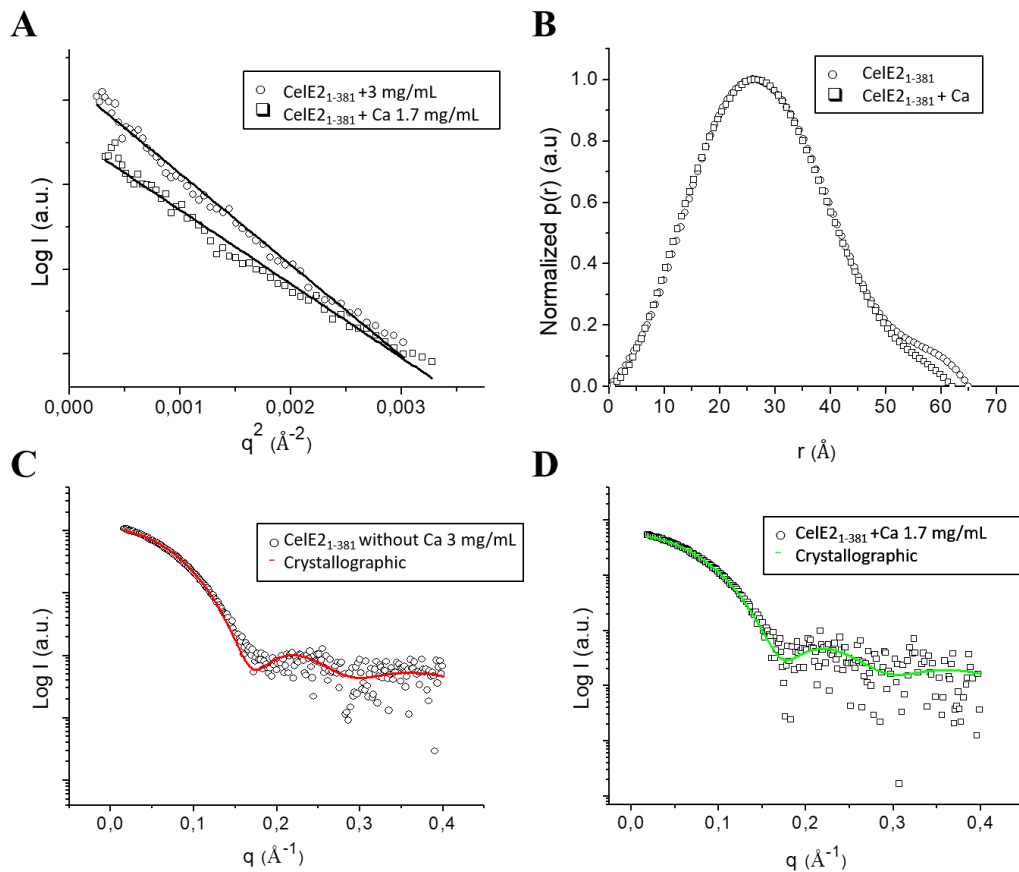| Parameters | Cat (3 mg.mL$^{-1}$) | Cat + Ca (1.7 mg.mL$^{-1}$) | DAM | Crystallographic |
|---|---|---|---|---|
| $R_g$ (Guinier) (Å) | 23,70±0.08 | 22,50±0,08 | - | - |
| $R_g$ (Å) | 22,11±0,02 | 21,46±0,03 | 20,84 | 19,03 |
| $D_{max}$ (Å) | 65,00 | 62,00 | 66,28 | 61,66 |

**Fig. 13.** (**A**) Guinier analysis and (**B**) distribution function of the distance pairs demonstrating the decrease in maximal diameter of CelE2$_{1-381}$ in the presence of 5 mM CaCl$_2$. Adjustments of the simulated scattering of the crystallographic model to the experimental curve in the absence (**C**) and presence of 5 mM CaCl$_2$ (**D**), revealing tendency of aggregation in the absence of CaCl$_2$.

The discrepancy ($\chi^2$) of the simulated scattering of the crystallographic model of CelE2$_{1-381}$ in relation to the experimental curve decreased from 3.09 to 2.07 in the presence of CaCl$_2$ (Fig. 13C and 13D). The difference in discrepancy is caused by a tendency of sample aggregation in the absence of calcium. From the experimental curve in the presence of CaCl$_2$, the *Dummy* atom model (DAM) was generated. Fig. 14 shows the overlap of the models.
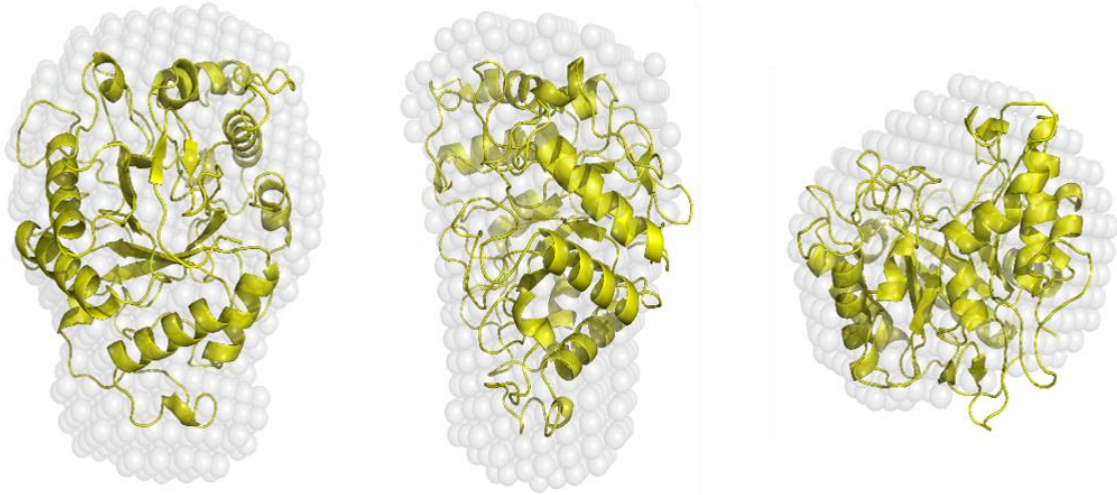
**Fig. 14.** Low-resolution *"dummy"* atom model superimposed on the crystallographic model of CelE2$_{1-381}$. The center and right models were rotated 90 ° about the y-axis and 90 ° about the x-axis of the orientation shown in the model on the left.

However, for CelE2$_{364-477}$ evidence of unfolding of the sample was verified, rendering the analyzes unfeasible. Similar results are shown in Rambo & Tainer (RAMBO; TAINER, 2011). Fig. 15 shows this unfolding tendency.

**Fig. 15.** Experimental curve of CelE2$_{364-477}$ **(A)**, Kratky plot **(B)**, Porod-Debye plot **(C).**

## 4    Concluding Remarks

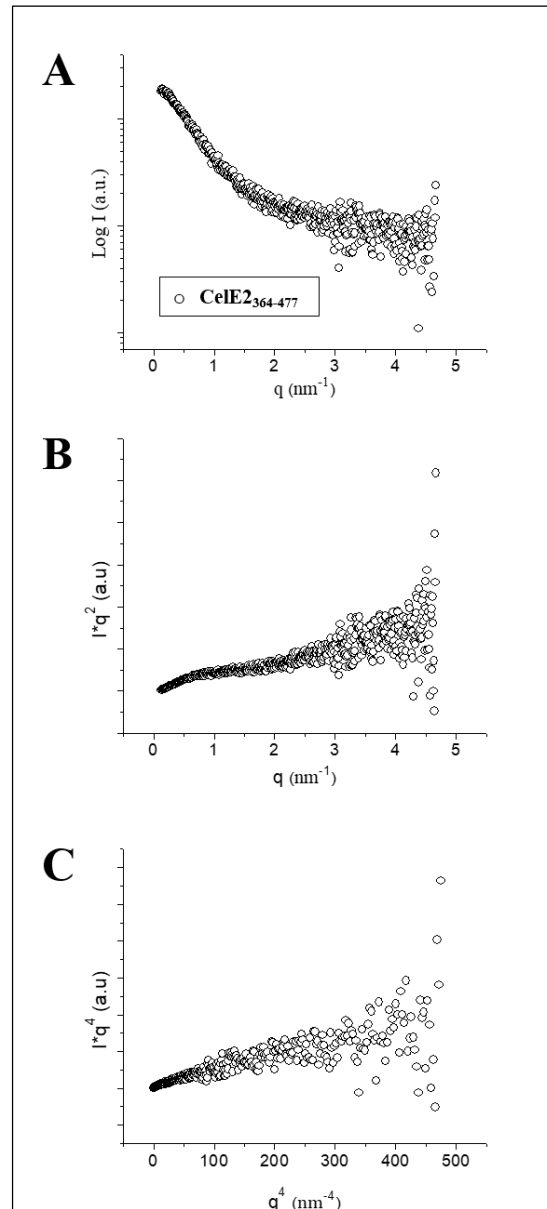The truncated domains of CelE2 were produced in *E. coli* and submitted to biochemical and structural studies. Our results showed that the removal of the Calx-β accessory domain (CelE2$_{382-477}$) did not promote significant functional changes, besides suggesting that the activation and effects of stability generated by calcium are not directly related to this domain. However, deleting of this domain increased more than 4-fold the enzymatic activity in the insoluble polysaccharide Avicel and reduction of 4.6 ºC of melting temperature. It was not possible to show, according to the results of polysaccharide binding analyzes, that CelE2$_{382-477}$

has affinity to the tested substrates. Finally, this work contributed with a new enzyme structure of the family 5 of glycoside hydrolases and presents new information about the individual CelE2 domains.

**Acknowledgments**

**SUPPLEMENTARY MATERIAL – DOCUMENT 2**

**Frame S1:** Conditions of PCR reactions for CelE2 amplification and their derivative constructs.

| Genes | Conditions of PCR reactions | | | | | |
|---|---|---|---|---|---|---|
| $cele2_{1-381}$ | Temperature Time | 98 ºC 1 min | 98 ºC 30 s | 62 ºC 50 s | 72 ºC 1.5 min | 72 ºC 10 min |
| | Nº cycles | 1 | 30 | | | 1 |
| $cele2_{1-363}$ | Temperature Time | 98 ºC 1 min | 98 ºC 30 s | 55 ºC 50 s | 72 ºC 1.5 min | 72 ºC 10 min |
| | Nº cycles | 1 | 30 | | | 1 |
| $cele2_{364-477}$ | Temperature Time | 98 ºC 1 min | 98 ºC 30 s | 62 ºC 30 s | 72 ºC 1.5 min | 72 ºC 10 min |
| | Nº cycles | 1 | 25 | | | 1 |
| $cele2_{382-477}$ | Temperature Time | 98 ºC 1 min | 98 ºC 30 s | 62 ºC 30 s | 72 ºC 1.5 min | 72 ºC 10 min |
| | Nº cycles | 1 | 25 | | | 1 |
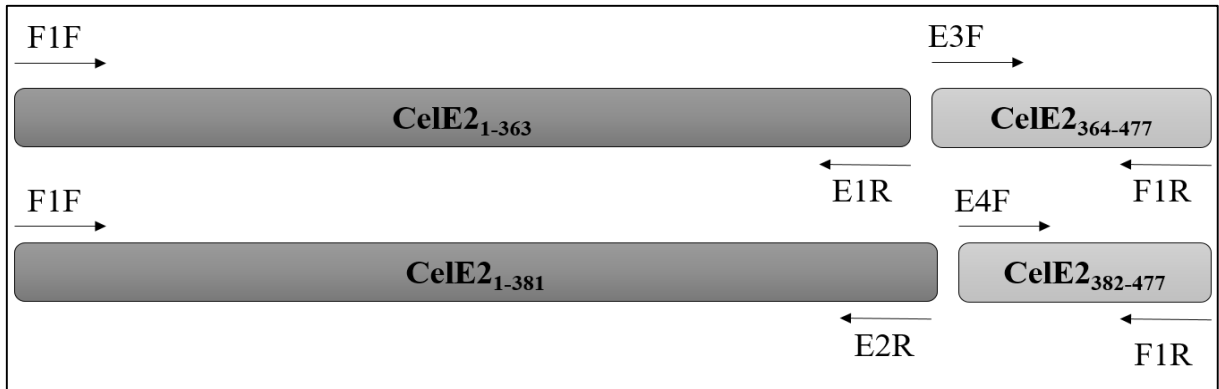
**Fig. S1.** Representation of the annealing of the primer oligonucleotides for each construct. F1F and F1R correspond to the forward and reverse primer oligonucleotides designed for *cele2* amplification. E1R and E2R correspond to the reverse primer oligonucleotides specific for the constructs *cele2*$_{1-363}$ and *cele2*$_{1-381}$ (catalytic domain), respectively. E3F and E4F correspond to the specific forward primer oligonucleotides for the constructs *cele2*$_{364-477}$ and *cele2*$_{382-477}$ (accessory domain), respectively.

## CAPÍTULO 3 – CONSIDERAÇÕES FINAIS

## DISCUSSÃO GERAL

O CAPÍTULO 1, como revisão bibliográfica, traça um panorama geral sobre a importância energética, composição e aplicações da biomassa lignocelulósica e de seus principais componentes, relacionando ainda a importância comercial de enzimas que atuam em carboidratos presentes na biomassa vegetal, como as celulases, que podem ser obtidas a partir de microrganismos não-cultiváveis.

Os resultados apresentados no CAPÍTULO 2 apresenta informações acerca da celulase multimodular CelE2, bem como a relação de seus domínios individuais para o seu funcionamento. No documento 1 foram apresentados resultados referentes às propriedades bioquímicas e biofísicas de CelE2, a qual possui um domínio não-convencional (Calx-β) para as GHs e pouco estudado. Além das principais características de atuação, CelE2 mostrou indícios de atividade de transglicosilação e a presença de $CaCl_2$ promoveu aumento da atividade enzimática e diminuição da tendência de agregação. Estes resultados levantaram questionamentos sobre o papel de cada domínio diante à uma arquitetura não-convencional.

O domínio Calx-β tem sido relacionado com a interação com $Ca^{2+}$, seguindo a definição do Interpro e os estudos relacionados com o domínio Calx-β de proteínas da membrana plasmática. No entanto, apesar de ser encontrado em algumas glicosil hidrolases, não há informações sobre a sua função definida para estes casos. No caso da β-glicosidase (GH3) BglM1 de *Physarum polycephalum*, sugeriram que este domínio poderia estar envolvido na manutenção da estrutura terciária da proteína.

Os domínios acessórios presentes em muitas enzimas podem desempenhar diferentes funções para a atividade enzimática. Com isso, foram construídas derivações truncadas de CelE2 para avaliar as propriedades bioquímicas e estruturais de cada domínio e sua influência para a atividade enzimática, como apresentado no documento 2. Os resultados mostraram que a deleção do domínio acessório Calx-β não provocou alterações nas características bioquímicas de CelE2, bem como o perfil de degradação e as evidências de transglicosilação. Além disso, aumento da atividade enzimática e a diminuição da tendência de agregação na presença de cálcio para o domínio catalítico, indicam que o domínio Calx-β não está relacionado aos efeitos do cálcio. No entanto, a remoção do domínio Calx-β resultou no aumento de 4 vezes da atividade enzimática em Avicel e um aumento de 4,6 ºC na temperatura

de desnaturação. Os ensaios de avaliação da capacidade de ligação à polissacarídeos indicaram que o domínio Calx-β não tem afinidade para os substratos testados. Contudo, a retirada do domínio Calx-β permitiu a cristalização e a resolução da estrutura tridimensional do domínio catalítico GH5, o qual não foi possível ser obtido para a construção completa.

A presença de domínios acessórios pode promover um efeito variado entre as enzimas, e esta questão tem impulsionado estudos de caracterização e influência de domínios individuais. A retirada destes módulos adjacentes pode não ter efeito sobre uma enzima modular ou pode resultar no aumento da eficiência catalítica, como para a endoglucanase Cel9A com a retirada do domínio Ig-like. Porém, existem enzimas em que a retirada de um mais domínios acessórios promovem alterações bioquímicas ou a completa inativação, como para as celulases rPglA e BlCel5B, respectivamente. Apesar dos domínios Calx-β estarem presentes em algumas GHs, até o momento não há informações sobre sua função para a atividade destas enzimas.

## CONCLUSÃO GERAL

Os ensaios de caracterização de CelE2 e seus domínios individuais exibidos no CAPÍTULO 2, mostraram que a retirada do domínio acessório Calx-β não alterou as principais características mostradas para a proteína completa no documento 1, e que o domínio Calx-β não deve estar relacionado com os efeitos promovidos pelo cálcio. No entanto, a remoção deste domínio resultou no aumento de atividade enzimática do domínio catalítico em Avicel e a diminuição da temperatura de desnaturação. Com isso, os resultados apresentados nesta dissertação contribuem com a primeira caracterização bioquímica e biofísica de uma endoglucanase GH5 que exibe esta organização modular não-convencional, e com a construção detalhada e caracterização dos domínios individuais da mesma. Além da estrutura tridimensional do domínio catalítico GH5, novos estudos poderão ser conduzidos para comparar os efeitos deste domínio Calx-β em outras representantes de GHs.

# REFERÊNCIAS

ABDUL KHALIL, H. P. S.; BHAT, A. H.; IREANA YUSRA, A. F. Green composites from sustainable cellulose nanofibrils: A review. **Carbohydrate Polymers**, v. 87, n. 2, p. 963–979, jan. 2012.

ABOT, A. et al. CAZyChip: dynamic assessment of exploration of glycoside hydrolases in microbial ecosystems. **BMC Genomics**, v. 17, n. 1, p. 671, 23 dez. 2016.

ADAMS, P. D. et al. PHENIX : a comprehensive Python-based system for macromolecular structure solution. **Acta Crystallographica Section D Biological Crystallography**, v. 66, n. 2, p. 213–221, 1 fev. 2010.

ADITIYA, H. B. et al. Second generation bioethanol production: A critical review. **Renewable and Sustainable Energy Reviews**, v. 66, p. 631–653, dez. 2016.

ALONSO-GARCÍA, N. et al. Structure of the Calx-beta domain of the integrin beta4 subunit: Insights into function and cation-independent stability. **Acta Crystallographica Section D: Biological Crystallography**, v. 65, n. 8, p. 858–871, 2009.

ALONSO, D. M.; BOND, J. Q.; DUMESIC, J. A. Catalytic conversion of biomass to biofuels. **Green Chem.**, v. 12, n. October, p. 1493–1513, 2010.

ALVAREZ, T. M. **Desenvolvimento de uma biblioteca de enzimas a partir de metagenoma de solo**. [s.l.] Universidade Estadual de Campinas, 2013.

ALVAREZ, T. M. et al. Structure and function of a novel cellulase 5 from sugarcane soil metagenome. **PLoS ONE**, v. 8, n. 12, p. 1–9, 2013a.

ALVAREZ, T. M. et al. Development and Biotechnological Application of a Novel Endoxylanase Family GH10 Identified from Sugarcane Soil Metagenome. **PLoS ONE**, v. 8, n. 7, 2013b.

ALVAREZ, T. M. et al. A Novel Member of GH16 Family Derived from Sugarcane Soil Metagenome. **Applied Biochemistry and Biotechnology**, v. 177, n. 2, p. 304–317, 2015.

ARAKI, R. et al. Essential role of the family-22 carbohydrate-binding modules for β-1,3-1,4-glucanase activity of Clostridium stercorarium Xyn10B. **FEBS Letters**, v. 561, n. 1–3, p. 155–158, 12 mar. 2004.

ARAUJO, G. J. F.; NAVARRO, L. F. S.; SANTOS, B. A. S. O Etanol De Segunda Geração E Sua Importância Estratégica Ante O Cenário Energetico Internacional Contemporâneo. **Fórum Ambiental de Alta Paulista**, v. 9, n. 5, p. 1–11, 2013.

ARO, N.; PAKULA, T.; PENTTILÄ, M. Transcriptional regulation of plant cell wall degradation by filamentous fungi. **FEMS Microbiology Reviews**, v. 29, n. 4, p. 719–739, set.

2005.

ASPEBORG, H. et al. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). **BMC Evolutionary Biology**, v. 12, n. 1, p. 186, 2012.

ASSAD, L. Aproveitamento de resíduos do setor sucroalcooleiro desafia empresas e pesquisadores. **Ciência e Cultura**, v. 69, n. 4, p. 13–16, out. 2017.

BAEYENS, J. et al. Challenges and opportunities in improving the production of bio-ethanol. **Progress in Energy and Combustion Science**, v. 47, p. 60–88, abr. 2015.

BAKER, J. O. et al. Catalytically enhanced endocellulase Cel5A from Acidothermus cellulolyticus. **Applied biochemistry and biotechnology**, v. 121–124, p. 129–148, 2005.

BASSO, L. C.; BASSO, T. O.; ROCHA, S. N. Ethanol Production in Brazil: The Industrial Process and Its Impact on Yeast Fermentation. In: **Biofuel Production-Recent Developments and Prospects**. [s.l.] InTech, 2011. v. 1530.

BÉRA-MAILLET, C. et al. Biochemical characterization of MI-ENG1 , a family 5 endoglucanase secreted by the root-knot nematode Meloidogyne incognita. **Eur. J. Biochem.**, v. 267, p. 3255–3263, 2000.

BERMAN, H. M. et al. The protein data bank. **Nucleic acids research**, v. 28, n. 1, p. 235–242, 2000.

BESSERER, G. M. et al. The second Ca2+-binding domain of the Na+ Ca2+ exchanger is essential for regulation: Crystal structures and mutational analysis. **Proceedings of the National Academy of Sciences**, v. 104, n. 47, p. 18467–18472, 20 nov. 2007.

BHABHA, G. et al. A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. **Science**, v. 332, n. 6026, p. 234–238, 2011.

BHARGAVA, A.; FUENTES, F. Biofuels: A Green Technology for the Future. In: BHARGAVA, A.; SRIVASTAVA, S. (Eds.). . **Biotechnology: recent trends and emerging dimensions**. [s.l: s.n.]. p. 39.

BHATTACHARYA, M. et al. The effect of renewable energy consumption on economic growth: Evidence from top 38 countries. **Applied Energy**, v. 162, n. January, p. 733–741, jan. 2016.

BIASINI, M. et al. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. **Nucleic Acids Research**, v. 42, n. W1, p. 1–7, 2014.

BIOENERGY AUSTRALIA, B. A. **Ethanol**. Disponível em: <http://biofuelsassociation.com.au/biofuels/ethanol/>. Acesso em: 23 jun. 2018.

BOLAM, D. N. et al. X4 Modules Represent a New Family of Carbohydrate-binding

Modules That Display Novel Properties. **Journal of Biological Chemistry**, v. 279, n. 22, p. 22953–22963, 28 maio 2004.

BORASTON, A. B. et al. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. **Biochemical Journal**, v. 382, n. 3, p. 769–781, 15 set. 2004.

CAMPOS, B. M. et al. A Novel Carbohydrate-binding Module from Sugar Cane Soil Metagenome Featuring Unique Structural and Carbohydrate Affinity Properties. **Journal of Biological Chemistry**, v. 291, n. 45, p. 23734–23743, 4 nov. 2016.

CANNELLA, D. et al. Light-driven oxidation of polysaccharides by photosynthetic pigments and a metalloenzyme. **Nature Communications**, v. 7, p. 11134, 4 abr. 2016.

CARDONA, C. A.; SÁNCHEZ, Ó. J. Fuel ethanol production: Process design trends and integration opportunities. **Bioresource Technology**, v. 98, n. 12, p. 2415–2457, set. 2007.

CHEN, V. B. et al. MolProbity : all-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D Biological Crystallography**, v. 66, n. 1, p. 12–21, 1 jan. 2010.

CHEN, Z. et al. Tracing determinants of dual substrate specificity in glycoside hydrolase family 5. **Journal of Biological Chemistry**, v. 287, n. 30, p. 25335–25343, 2012.

CHENG, R. et al. Recombinant production and characterization of full-length and truncated β-1,3-glucanase PglA from Paenibacillus sp. S09. **BMC Biotechnology**, v. 13, n. 1, p. 105, 2013.

COCKBURN, D.; WILKENS, C.; SVENSSON, B. Affinity Electrophoresis for Analysis of Catalytic Module-Carbohydrate Interactions. In: ABBOTT D., L. VAN B. A. (EDS) (Ed.). . **Protein-Carbohydrate Interactions. Methods in Molecular Biology**. New York, NY: Humana Press, 2017. v. 1588p. 119–127.

CONAB. **Acompanhamento da Safra Brasileira: Cana-de- açúcar - V.5 - Safra 2018/19 N.1 - Primeiro levantamento, Maio 2018.** [s.l: s.n.].

CORRÊA, D. H. A.; RAMOS, C. H. I. The use of circular dichroism spectroscopy to study protein folding , form and function. v. 3, n. 5, p. 164–173, 2009.

CORRÊA, T. L. R.; DOS SANTOS, L. V.; PEREIRA, G. A. G. AA9 and AA10: from enigmatic to essential enzymes. **Applied Microbiology and Biotechnology**, v. 100, n. 1, p. 9–16, 2016.

DE SOUZA LIMA, M. M.; BORSALI, R. Rodlike Cellulose Microcrystals: Structure, Properties, and Applications. **Macromolecular Rapid Communications**, v. 25, p. 771–787, abr. 2004.

DE SOUZA, W. R. Microbial Degradation of Lignocellulosic Biomass. In: PESEK, K. (Ed.). . **Sustainable Degradation of Lignocellulosic Biomass - Techniques, Applications and Commercialization**. [s.l.] InTech, 2013.

DEMAIN, A. L. Biosolutions to the energy problem. **Journal of Industrial Microbiology & Biotechnology**, v. 36, n. 3, p. 319–332, 10 mar. 2009.

DILOKPIMOL, A. et al. Recombinant production and characterisation of two related GH5 endo-β-1,4-mannanases from Aspergillus nidulans FGSC A4 showing distinctly different transglycosylation capacity. **Biochimica et Biophysica Acta - Proteins and Proteomics**, v. 1814, n. 12, p. 1720–1729, 2011.

DOS SANTOS, L. V. et al. Second-Generation Ethanol: The Need is Becoming a Reality. **Industrial Biotechnology**, v. 12, n. 1, p. 40–57, fev. 2016.

DWIVEDI, P.; ALAVALAPATI, J. R. R.; LAL, P. Cellulosic ethanol production in the United States: Conversion technologies, current production status, economics, and emerging developments. **Energy for Sustainable Development**, v. 13, n. 3, p. 174–182, 2009.

EMSLEY, P. et al. Features and development of Coot. **Acta Crystallographica Section D Biological Crystallography**, v. 66, n. 4, p. 486–501, 1 abr. 2010.

EVANGELISTA, R. A.; LIU, M.-S.; CHEN, F.-T. A. Characterization of 9-aminopyrene-1,4,6-trisulfonate derivatized sugars by capillary electrophoresis with laser-induced fluorescence detection. **Analytical Chemistry**, v. 67, p. 2239–2245, 1995.

EVANS, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. **Acta Crystallographica Section D Biological Crystallography**, v. 67, n. 4, p. 282–292, 1 abr. 2011.

FARHAD, S.; SAFFAR-AVVAL, M.; YOUNESSI-SINAKI, M. Efficient design of feedwater heaters network in steam power plants using pinch technology and exergy analysis. **International Journal of Energy Research**, v. 32, n. 1, p. 1–11, jan. 2008.

FINN, R. D. et al. The Pfam protein families database: towards a more sustainable future. **Nucleic acids research**, v. 44, n. D1, p. D279–D285, 2015.

FINN, R. D. et al. InterPro in 2017—beyond protein family and domain annotations. **Nucleic Acids Research**, v. 45, n. D1, p. D190–D199, 4 jan. 2017.

FISCHER, H. et al. Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. **Journal of Applied Crystallography**, v. 43, p. 101–109, 2010.

FRANCO CAIRO, J. P. L. et al. Deciphering the synergism of endogenous glycoside

hydrolase families 1 and 9 from Coptotermes gestroi. **Insect Biochemistry and Molecular Biology**, v. 43, n. 10, p. 970–981, 2013.

GASTEIGER, E. et al. Protein Identification and Analysis Tools on the ExPASy Server. In: WALKER, J. M. (Ed.). . **The Proteomics Protocols Handbook**. [s.l.] Humana Press Inc, 2005. p. 571–607.

GHATGE, S. S. et al. Characterization of modular bifunctional processive endoglucanase Cel5 from Hahella chejuensis KCTC 2396 Characterization of modular bifunctional processive endoglucanase Cel5 from Hahella chejuensis KCTC 2396. **Appl Microbiol Biotechnol**, 2013.

GILBERT, H. J.; KNOX, J. P.; BORASTON, A. B. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. **Current Opinion in Structural Biology**, v. 23, n. 5, p. 669–677, out. 2013.

GÍRIO, F. M. et al. Hemicelluloses for fuel ethanol: A review. **Bioresource Technology**, v. 101, n. 13, p. 4775–4800, jul. 2010.

GREENFIELD, N. Using circular dichroism spectra to estimate protein secondary structure. **Nat Protoc.**, v. 1, n. 6, p. 2876–2890, 2007.

GRICE, L. F. et al. Origin and evolution of the sponge aggregation factor gene family. **Molecular Biology and Evolution**, p. msx058, 19 jan. 2017.

GUERRIERO, G. et al. Lignocellulosic bioma ss : Biosynthesis, degradation, and industrial utilization. **Engineering in Life Sciences**, v. 16, n. 1, p. 1–16, jan. 2016.

GUILLÉN, D.; SÁNCHEZ, S.; RODRÍGUEZ-SANOJA, R. Carbohydrate-binding domains: multiplicity of biological roles. **Applied Microbiology and Biotechnology**, v. 85, n. 5, p. 1241–1249, 12 fev. 2010.

GUINIER, A. AND FOURNET, G. Small-angle scattering of X-ray. **New York: Wiley**, 1955.

GUO, M.; SONG, W.; BUHAIN, J. Bioenergy and biofuels: History, status, and perspective. **Renewable and Sustainable Energy Reviews**, v. 42, p. 712–725, fev. 2015.

GUPTA, A.; VERMA, J. P. Sustainable bio-ethanol production from agro-residues: A review. **Renewable and Sustainable Energy Reviews**, v. 41, p. 550–567, jan. 2015.

GUPTA, V. K. et al. Fungal Enzymes for Bio-Products from Sustainable and Waste Biomass. **Trends in Biochemical Sciences**, v. 41, n. 7, p. 633–645, jul. 2016.

HALL, M. et al. Cellulose crystallinity - a key predictor of the enzymatic hydrolysis rate. **FEBS Journal**, v. 277, n. 6, p. 1571–1582, mar. 2010.

HAMELINCK, C. N.; HOOIJDONK, G. VAN; FAAIJ, A. P. Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term. **Biomass and Bioenergy**, v. 28, n. 4, p. 384–410, abr. 2005.

HAMMERSLEY, A. P. FIT2D: An Introduction and Overview. **ESRF Internal Report.**, v. 68, 1997.

HANDELSMAN, J. Metagenomics: Application of Genomics to Uncultured Microorganisms. **Microbiology and Molecular Biology Reviews**, v. 68, n. 4, p. 669–685, 1 mar. 2004.

HARMSEN, P. et al. **Literature review of physical and chemical pretreatment processes for lignocellulosic biomass**. [s.l: s.n.]. Disponível em: <http://www.researchgate.net/publication/254853217_Literature_review_of_physical_and_chemical_pretreatment_processes_for_lignocellulosic_biomass>.

HAYASE, M. et al. Properties, intracellular localization, and stage-specific expression of membrane-bound β-glucosidase, BglM1, from Physarum polycephalum. **The International Journal of Biochemistry & Cell Biology**, v. 40, n. 10, p. 2141–2150, jan. 2008.

HENDRIKS, A. T. W. M.; ZEEMAN, G. Pretreatments to enhance the digestibility of lignocellulosic biomass. **Bioresource Technology**, v. 100, n. 1, p. 10–18, 2009.

HENRISSAT, B.; BAIROCH, A. Updating the sequence-based classification of glycosyl hydrolases. **Biochemical Journal**, v. 316, n. 2, p. 695–696, 1 jun. 1996.

HIMMEL, M. E. et al. Biomass recalcitrance: engineering plants and enzymes for biofuels production. **Science (New York, N.Y.)**, v. 315, n. 5813, p. 804–807, 2007.

HORN, S. J. et al. Novel enzymes for the degradation of cellulose. **Biotechnology for Biofuels**, v. 5, n. 1, p. 45, 2012.

HUA, M. et al. Direct detection, cloning and characterization of a glucoside hydrolase from forest soil. **Biotechnology Letters**, v. 37, n. 6, p. 1227–1232, 2015.

HUERTA-CEPAS, J. et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. **Nucleic Acids Research**, v. 44, n. D1, p. D286–D293, 4 jan. 2016.

HUGENHOLTZ, P.; TYSON, G. W. Metagenomics. **Pediatric Infectious Disease Journal**, v. 455, n. 25, 2008.

IEA, I. E. A. **World Energy Outlook 2017**. Disponível em: <http://www.iea.org/weo2017/>. Acesso em: 20 jun. 2018a.

IEA, I. E. A. **Climate Change**. Disponível em: <https://www.iea.org/topics/climatechange/>. Acesso em: 21 jun. 2018b.

IEA, I. E. A. **Key world energy statistics** (Intergovernmental Panel on Climate Change, Ed.)**IEA Publications**. Cambridge: [s.n.]. Disponível em: <http://www.enecho.meti.go.jp/statistics/total_energy/results.html>.

IEA, I. E. A. **Energy Technology Perspectives 2017 - Catalysing energy technology transformationsIEA Publications**. [s.l.] OECD, 7 jun. 2017d. Disponível em: <http://www.oecd.org/about/publishing/Corrigendum_EnergyTechnologyPerspectives2017.pdf>.

IEA, I. E. A. **CO2 Emissions from Fuel Combustion 2017International Energy Agency**: CO2 Emissions from Fuel Combustion. [s.l.] OECD Publishing, 31 out. 2017e. Disponível em: <https://www.iea.org/publications/freepublications/publication/CO2EmissionsfromFuelCombustionHighlights2017.pdf>.

JAM, M. et al. Unraveling the multivalent binding of a marine family 6 carbohydrate-binding module with its native laminarin ligand. **FEBS Journal**, v. 283, n. 10, p. 1863–1879, 2016.

JIMENEZ, D. J. et al. Characterization of three plant biomass-degrading microbial consortia by metagenomics- and metasecretomics-based approaches. **Applied Microbiology and Biotechnology**, v. 100, n. 24, p. 10463–10477, 2016.

JÚNIOR, A. D. N. F. et al. Applied Metagenomics for Biofuels Development and Environmental Sustainability. In: BUCKERIDGE, M. S.; DE SOUZA, A. P. (Eds.). . **Advances of Basic Science for Second Generation Bioethanol from Sugarcane**. Cham: Springer International Publishing, 2017. p. 107–129.

KABSCH, W. XDS. **Acta Crystallographica Section D Biological Crystallography**, v. 66, n. 2, p. 125–132, 1 fev. 2010.

KATAEVA, I. A. et al. The fibronectin type 3-like repeat from the Clostridium thermocellum cellobiohydrolase CbHa promotes hydrolysis of cellulose by modifying its surface. **Applied and Environmental Microbiology**, v. 68, n. 9, p. 4292–4300, 2002.

KEARSE, M., MOIR, R., WILSON, A., STONES-HAVAS, S., CHEUNG, M., STURROCK, S., BUXTON, S., COOPER, A., MARKOWITZ, S., DURAN, C., THIERER, T., ASHTON, B., MENTJIES, P., DRUMMOND, A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. **Bioinformatics**, v. 28, n. 12, p. 1647–1649, 2012.

KIM, S. J. et al. Characterization of a gene encoding cellulase from uncultured soil bacteria. **FEMS Microbiology Letters**, v. 282, n. 1, p. 44–51, 2008.

KIMBLE, M.; PASDELOUP, M.-V.; SPENCER, C. **Sustainable Bioenergy Development in UEMOA Member Countries**. [s.l: s.n.]. Disponível em: <http://www.globalproblems-globalsolutions-files.org/gpgs_files/pdf/UNF_Bioenergy/UNF_Bioenergy_full_report.pdf>.

KONAREV, P. V. et al. PRIMUS : a Windows PC-based system for small-angle scattering data analysis. **Journal of Applied Crystallography**, v. 36, n. 5, p. 1277–1282, 1 out. 2003.

KOZIN, M. B.; SVERGUN, D. I. Automated matching of high- and low-resolution structural models. **Journal of Applied Crystallography**, v. 34, n. 1, p. 33–41, 1 fev. 2001.

KRISSINEL, E.; HENRICK, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. **Acta Cryst.**, v. D60, p. 2256–2268, 2004.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Molecular Biology and Evolution**, v. 33, n. 7, p. 1870–1874, jul. 2016.

KUMAR, V. et al. Global Market Scenario of Industrial Enzymes. In: BENIWAL, V.; SHARMA, A. K. (Eds.). . **In: Industrial Enzymes: Trends, Scope and Relevance**. [s.l.] Nova Science Publishers, Inc. New York, 2014. p. 173–196.

LAEMMLI, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. **Nature**, v. 227, p. 680–685, 1970.

LARKIN, M. A. et al. Clustal W and Clustal X version 2.0. **Bioinformatics**, v. 23, n. 21, p. 2947–2948, 1 nov. 2007.

LAZARIDOU, A. et al. A comparative study on structure-function relations of mixed-linkage (1>3), (1>4) linear beta-D-glucans. **Food Hydrocolloids**, v. 18, n. 5, p. 837–855, 2004.

LEE, C.-M. et al. Screening and Characterization of a Novel Cellulase Gene from the Gut Microflora of Hermetia illucens Using Metagenomic Library. **Journal of Microbiology and Biotechnology**, v. 24, n. 9, p. 1196–1206, 28 set. 2014.

LEE, R. A.; LAVOIE, J.-M. From first- to third-generation biofuels: Challenges of producing a commodity from a biomass of increasing complexity. **Animal Frontiers**, v. 3, n. 2, p. 6–11, 2013.

LEE, T. K.; KIM, C. H. Molecular Cloning and Expression of an Endo- $\beta$ -1 , 4- D -glucanase I ( Avicelase I ) Gene from Bacillus cellulyticus K-12 and Characterization of the Recombinant Enzyme. **Applied Biochemistry and Biotechnology**, v. 80, 1999.

LEE, H. V.; HAMID, S. B. A.; ZAIN, S. K. Conversion of Lignocellulosic Biomass to Nanocellulose: Structure and Chemical Process. **The Scientific World Journal**, v. 2014, p. 1–20, 2014.

LETUNIC, I.; DOERKS, T.; BORK, P. SMART: Recent updates, new developments and status in 2015. **Nucleic Acids Research**, v. 43, n. D1, p. D257–D260, 2015.

LEVASSEUR, A. et al. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. **Biotechnology for Biofuels**, v. 6, n. 1, p. 41, 2013.

LIBERATO, M. V et al. Molecular characterization of a family 5 glycoside hydrolase suggests an induced-fit enzymatic mechanism. **Scientific Reports**, v. 6, n. 1, p. 23473, 1 set. 2016.

LIU, H. et al. Molecular simulations provide new insights into the role of the accessory immunoglobulin-like domain of Cel9A. **FEBS Letters**, v. 584, n. 15, p. 3431–3435, 2010.

LIU, J. et al. Cloning and functional characterization of a novel endo-beta-1,4- glucanase gene from a soil-derived metagenomic library. **Applied Microbiology and Biotechnology**, v. 89, n. 4, p. 1083–1092, 2011.

LOMBARD, V. et al. The carbohydrate-active enzymes database (CAZy) in 2013. **Nucleic Acids Research**, v. 42, n. D1, p. D490–D495, jan. 2014.

LOSORDO, Z. et al. Cost competitive second-generation ethanol production from hemicellulose in a Brazilian sugarcane biorefinery. **Biofuels, Bioproducts and Biorefining**, v. 10, n. 5, p. 589–602, set. 2016.

LU, Z. et al. Truncation of the unique N-terminal domain improved the thermos-stability and specific activity of alkaline α-amylase Amy703. **Scientific Reports**, v. 6, n. 1, p. 22465, 1 abr. 2016.

LYND, L. R. et al. Microbial Cellulose Utilization: Fundamentals and Biotechnology. **Bioresource Technology**, v. 66, n. 3, p. 506–577, 2002.

MAEKAWA, A. et al. A cDNA cloned from Physarum polycephalum encodes new type of family 3 β-glucosidase that is a fusion protein containing a calx-β motif. **The International Journal of Biochemistry & Cell Biology**, v. 38, n. 12, p. 2164–2172, jan. 2006.

MANDELLI, F. et al. The characterization of a thermostable and cambialistic superoxide dismutase from thermus filiformis. **Letters in Applied Microbiology**, v. 57, n. 1, p. 40–46, 2013.

MANISHA; YADAV, S. K. Technological advances and applications of hydrolytic enzymes for valorization of lignocellulosic biomass. **Bioresource Technology**, dez. 2017.

MARIN, F. R. et al. Prospects for Increasing Sugarcane and Bioethanol Production on Existing Crop Area in Brazil. **BioScience**, v. 66, n. 4, p. 307–316, 1 abr. 2016.

MARUTHAMUTHU, M. et al. A multi-substrate approach for functional metagenomics-

based screening for (hemi)cellulases in two wheat straw-degrading microbial consortia unveils novel thermoalkaliphilic enzymes. **BMC genomics**, v. 17, n. 1, p. 86, 2016.

MCCOY, A. J. et al. Phaser crystallographic software. **Journal of Applied Crystallography**, v. 40, n. 4, p. 658–674, 1 ago. 2007.

MCGAVIN, M.; FORSBERG, C. W. Isolation and characterization ofendoglucanases 1 and 2 from Bacteroides succinogenes S85. **J. Bacteriol.**, v. 170, n. 7, p. 2914–2922, 1988.

MEDIE, F. M. et al. Genome analyses highlight the different biological roles of cellulases. **Nature Reviews Microbiology**, v. 10, n. 3, p. 227–234, 2012.

MILLER, G. L. Use of dinitrosalicylic acid reagent for determination of reducing sugar. **Analytical Chemistry**, v. 31, p. 426–442, 1959.

MO, X. et al. Identification and characterization of a novel xylanase derived from a rice straw degrading enrichment culture. **Applied Microbiology and Biotechnology**, v. 87, n. 6, p. 2137–2146, 22 ago. 2010.

MOHAN, C. **Buffers**. [s.l.] EMD Biosciences, Inc, 2003.

MORRILL, J. et al. The GH5 1,4-β-mannanase from Bifidobacterium animalis subsp. lactis Bl-04 possesses a low-affinity mannan-binding module and highlights the diversity of mannanolytic enzymes. **BMC Biochemistry**, v. 16, n. 1, p. 26, 11 dez. 2015.

NANDA, S. et al. Pathways of lignocellulosic biomass conversion to renewable fuels. **Biomass Conversion and Biorefinery**, v. 4, n. 2, p. 157–191, 26 jun. 2014.

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION, U. S. N. L. OF M. **NCBI**. Disponível em: <https://www.ncbi.nlm.nih.gov/>.

NAZIR, A. Review on Metagenomics and its Applications. **Imperial Journal of Interdisciplinary Research**, v. 2, n. 3, p. 277–286, 2016.

NEI, M.; KUMAR, S. **Molecular evolution and phylogenetics.** Oxford Uni ed. New York: [s.n.].

NICOLL, D. A. et al. The Crystal Structure of the Primary Ca 2+ Sensor of the Na + /Ca 2+ Exchanger Reveals a Novel Ca 2+ Binding Motif. **Journal of Biological Chemistry**, v. 281, n. 31, p. 21577–21581, 4 ago. 2006.

PARK, S. et al. Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. **Biotechnology for Biofuels**, v. 3, n. 10, p. 1–10, 2010.

PAYNE, C. M. et al. Fungal Cellulases. **Chemical Reviews**, v. 115, n. 3, p. 1308–1448, 11 fev. 2015.

PEREIRA, J. D. C. et al. Effect of Metal Ions, Chemical Agents and Organic Compounds on

Lignocellulolytic Enzymes Activities. In: **Enzyme Inhibitors and Activators**. [s.l.] InTech, 2017.

PÉREZ, J. et al. Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. **International Microbiology**, v. 5, n. 2, p. 53–63, 27 jun. 2002.

PETERSEN, T. N. et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, v. 8, n. 10, p. 785–786, 2011.

PETOUKHOV, M. V; SVERGUN, D. I. Global rigid body modeling of macromolecular complexes against small-angle scattering data. **Biophysical journal**, v. 89, n. 2, p. 1237–50, 2005.

PHITSUWAN, P. et al. Present and potential applications of cellulases in agriculture, biotechnology, and bioenergy. **Folia Microbiologica**, v. 58, n. 2, p. 163–176, 26 mar. 2013.

PIMENTEL, A. C. et al. Biochemical and biophysical properties of a metagenome-derived GH5 endoglucanase displaying an unconventional domain architecture. **International Journal of Biological Macromolecules**, v. 99, p. 384–393, jun. 2017.

PINTO, E. R. P. et al. Transparent composites prepared from bacterial cellulose and castor oil based polyurethane as substrates for flexible OLEDs. **Journal of Materials Chemistry C**, v. 3, n. 44, p. 11581–11588, 2015.

QUIROZ-CASTAÑEDA, R. E.; FOLCH-MALLOL, J. L. Plant cell wall degrading and remodeling proteins: current perspectives. **Biotecnologia Aplicada**, v. 28, n. 4, p. 205–215, 2011.

RAMBO, R. P.; TAINER, J. A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. **Biopolymers**, v. 95, n. 8, p. 559–571, ago. 2011.

RFA, R. F. A. **2018 Ethanol Industry Outlook**. [s.l: s.n.]. Disponível em: <http://www.ethanolrfa.org/wp-content/uploads/2016/02/Ethanol-Industry-Outlook-2016.pdf%5Cnhttp://ethanolrfa.org/pages/monthly-fuel-ethanol-production-demand>.

RIESENFELD, C. S.; SCHLOSS, P. D.; HANDELSMAN, J. Metagenomics: Genomic Analysis of Microbial Communities. **Annual Review of Genetics**, v. 38, n. 1, p. 525–552, 2004.

RODRIGUES MOTA, T. et al. Plant cell wall composition and enzymatic deconstruction. **AIMS Bioengineering**, v. 5, n. 1, p. 63–77, 2018.

RUBIN, E. M. Genomics of cellulosic biofuels. **Nature**, v. 454, n. 7206, p. 841–845, 14 ago. 2008.

RUBINI, M. R. et al. Cloning, characterization and heterologous expression of the first Penicillium echinulatum cellulase gene. **Journal of Applied Microbiology**, v. 108, n. 4, p. 1187–1198, 2010.

SAITOU, N.; NEI, M. The neighgbor-joining method: a new method for reconstructing phylogenetic trees. **Mol. Biol. Evol.**, v. 4, n. 4, p. 406–425, 1987.

SÁNCHEZ, C. Lignocellulosic residues: Biodegradation and bioconversion by fungi. **Biotechnology Advances**, v. 27, n. 2, p. 185–194, mar. 2009.

SANDHU, S. K. et al. Cellulosic Biomass-Hydrolyzing Enzymes. In: SINGHANIA, R. R. ET AL (Ed.). . **Waste to Wealth, Energy, Enviroment, and Sustainability**. [s.l: s.n.]. p. 441–456.

SANTOS, C. R. DOS et al. Molecular insights into substrate specificity and thermal stability of a bacterial GH5-CBM27 endo-1,4-β-d-mannanase. **Journal of Structural Biology**, v. 177, n. 2, p. 469–476, fev. 2012.

SANTOS, W. D.; GÓMEZ, E. O.; BUCKERIDGE, M. S. Bioenergy and the Sustainable Revolution. In: BUCKERIDGE, M. S.; GOLDMAN, G. H. (Eds.). . **Routes to Cellulosic Ethanol**. New York, NY: Springer New York, 2011.

SCHELLER, H. V.; ULVSKOV, P. Hemicelluloses. **Annual Review of Plant Biology**, v. 61, n. 1, p. 263–289, 2 jun. 2010.

SCHNEIDMAN-DUHOVNY, D. et al. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. **Nucleic Acids Research**, v. 44, n. W1, p. W424–W429, 8 jul. 2016.

SCHRÖDER, R. et al. LeMAN4 endo-β-mannanase from ripe tomato fruit can act as a mannan transglycosylase or hydrolase. **Planta**, v. 224, n. 5, p. 1091–1102, 2006.

SCHULTZ, J. et al. SMART, a simple modular architecture research tool: identification of signaling domains. **Proceedings of the National Academy of Sciences of the United States of America**, v. 95, n. 11, p. 5857–64, 1998.

SCHWARZ, E.; BENZER, S. Calx beta trend 1999.pdf. **Trends in biochemical sciences**, v. 24, p. 260, 1999.

SCHWARZ, E. M.; BENZER, S. Calx, a Na-Ca exchanger gene of Drosophila melanogaster. **Proceedings of the National Academy of Sciences**, v. 94, n. 19, p. 10249–10254, 16 set. 1997.

SEGATO, F. et al. Genomics Review of Holocellulose Deconstruction by Aspergilli. **Microbiology and Molecular Biology Reviews**, v. 78, n. 4, p. 588–613, 2014.

SHAHZADI, T. et al. Advances in lignocellulosic biotechnology: A brief review on lignocellulosic biomass and cellulases. **Advances in Bioscience and Biotechnology**, v. 5, n. 3, p. 246–251, 2014.

SHARADA, R. et al. Applications of Cellulases – Review. **International Journal of Pharmaceutical, Chemical and Biological Sciences**, v. 4, n. 2, p. 424–437, 2014.

SHARMA, A. et al. Cellulases: Classification, Methods of Determination and Industrial Applications. **Applied Biochemistry and Biotechnology**, v. 179, n. 8, p. 1346–1380, 11 ago. 2016.

SHOKRI, J.; ADIBKI, K. Application of Cellulose and Cellulose Derivatives in Pharmaceutical Industries. In: **Cellulose - Medical, Pharmaceutical and Electronic Applications**. [s.l.] InTech, 2013. p. 47–66.

SILVA, R. N. **Mycology: Current and Future Developments Fungal Biotechnology for Biofuel Production**. Vol 1 ed. [s.l.] Bentham Science Publishers, 2016.

SILVA, V. C. T. et al. Effect of pH , Temperature , and Chemicals on the Endoglucanases and Beta-Glucosidases from the Thermophilic Fungus Myceliophthora heterothallica F . 2 . 1 . 4 . Obtained by Solid-State and Submerged Cultivation. **Biochemistry Research International**, v. 2016, p. 9, 2016.

SINDHU, R.; BINOD, P.; PANDEY, A. Biological pretreatment of lignocellulosic biomass – An overview. **Bioresource Technology**, v. 199, p. 76–82, jan. 2016.

SINGH, A.; OLSEN, S. I. A critical review of biochemical conversion, sustainability and life cycle assessment of algal biofuels. **Applied Energy**, v. 88, n. 10, p. 3548–3555, out. 2011.

SINGHANIA, R. R. et al. Cellulases. In: **Current Developments in Biotechnology and Bioengineering**. [s.l.] Elsevier, 2017. p. 73–101.

SINNOTT, M. L. Catalytic mechanism of enzymic glycosyl transfer. **Chemical Reviews**, v. 90, n. 7, p. 1171–1202, 1990.

SORRELL, S. Reducing energy demand: A review of issues, challenges and approaches. **Renewable and Sustainable Energy Reviews**, v. 47, p. 74–82, 2015.

SRIVASTAVA, N. et al. A Review on Fuel Ethanol Production From Lignocellulosic Biomass. **International Journal of Green Energy**, v. 12, n. 9, p. 949–960, 2 set. 2015.

SVERGUN, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. **Journal of Applied Crystallography**, v. 25, n. 4, p. 495–503, 1 ago. 1992.

SVERGUN, D. I. et al. CRYSOL - a Program to Evaluate X-ray Solution Scattering of

Biological Macromolecules from Atomic Coordinates. **Journal of Applied Crystallography**, v. 28, p. 768–773, 1995.

SVERGUN, D. I. Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. **Biophysical Journal**, v. 76, n. 6, p. 2879–2886, jun. 1999.

TAIZ, L.; ZEIGER, E. **Fisiologia Vegetal**. 3ª ed. [s.l.] Artmed, 2006.

TAMURA, K. et al. MEGA6: Molecular evolutionary genetics analysis version 6.0. **Molecular Biology and Evolution**, v. 30, n. 12, p. 2725–2729, 2013.

TANABE, T. et al. Novel chitosanase from Streptomyces griseus HUT 6037 with transglycosylation activity. **Bioscience, biotechnology, and biochemistry**, v. 67, n. 2, p. 354–364, 2003.

TEJIRIAN, A.; XU, F. Inhibition of Cellulase-Catalyzed Lignocellulosic Hydrolysis by Iron and Oxidative Metal Ions and Complexes. **APPLIED AND ENVIRONMENTAL MICROBIOLOGY**, v. 76, n. 23, p. 7673–7682, 2010.

THE UNIPROT CONSORTIUM. UniProt: a hub for protein information. **Nucleic Acids Research**, v. 43, n. Database issue, p. D204-12, 2014.

THE UNIPROT CONSORTIUM. UniProt: the Universal Protein knowledgebase. **Nucleic Acids Research**, v. 45, n. Database issue, p. D158–D169, 2017.

VAN WYK, N. et al. Current perspectives on the families of glycoside hydrolases of Mycobacterium tuberculosis : their importance and prospects for assigning function to unknowns. **Glycobiology**, v. 27, n. 2, p. 112–122, jan. 2017.

VÁRNAI, A. et al. Carbohydrate-Binding Modules of Fungal Cellulases. In: **Advances in Applied Microbiology**. [s.l: s.n.]. v. 88p. 103–165.

VILLARES, A. et al. Lytic polysaccharide monooxygenases disrupt the cellulose fibers structure. **Scientific Reports**, v. 7, n. 1, p. 40262, 10 dez. 2017.

VOLYNETS, B.; EIN-MOZAFFARI, F.; DAHMAN, Y. Biomass processing into ethanol: pretreatment, enzymatic hydrolysis, fermentation, rheology, and mixing. **Green Processing and Synthesis**, v. 6, n. 1, p. 1–22, 1 jan. 2017.

WALTER, A. et al. Sustainability assessment of bio-ethanol production in Brazil considering land use change, GHG emissions and socio-economic aspects. **Energy Policy**, v. 39, n. 10, p. 5703–5716, out. 2011.

WANG, M. et al. Life-cycle Energy Use and Greenhouse Gas Emission Implications of Brazilian Sugarcane Ethanol Simulated with the GREET Model. **International Sugar**

**Journal**, v. 110, n. 1317, 2008.

WANG, M. et al. Well-to-wheels energy use and greenhouse gas emissions of ethanol from corn, sugarcane and cellulosic biomass for US use. **Environmental Research Letters**, v. 7, n. 4, p. 045905, 1 dez. 2012.

WILSON, C.; GRUBLER, A.; WINTERFELDT, D. VON. Lessons from the history of technology and global change for the emerging clean technology cluster. **Natural Ressources Forums**, v. 35, n. January, p. 165–184, 2011.

WOOD, P. J.; WEISZ, J.; BLACKWELL, B. A. Structural Studies of (1-3)(1-4)- beta-D-Glucans by 13C-Nuclear Magnetic Resonance Spectroscopy and by Rapid Analysis of Cellulose-Like Regions Using High-Performance Anion-Exchange Chromatography of Oligosaccharides Released by Lichenase. **Cereal Chem.**, v. 71, p. 301–307, 1994.

XU, J. Metagenomics and Ecosystems Biology: Conceptual Frameworks, Tools and Methods. In: MARCO, D. (Ed.). . **Metagenomics: Theory, Methods and Applications**. [s.l.] Horizon Scientific Press, 2010.

YANG, B. et al. Enzymatic hydrolysis of cellulosic biomass. **Biofuels**, v. 2, n. 4, p. 421–449, 9 jul. 2011.

YENNAMALLI, R. M. et al. Endoglucanases: insights into thermostability for biofuel applications. **Biotechnology for Biofuels**, v. 6, n. 1, p. 136, 2013.

YIN, Y. et al. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. **Nucleic Acids Research**, v. 40, n. W1, p. W445–W451, 1 jul. 2012.

YOUNESI, F. S. et al. Deleting the Ig-Like Domain of Alicyclobacillus acidocaldarius Endoglucanase Cel9A Causes a Simultaneous Increase in the Activity and Stability. **Molecular Biotechnology**, v. 58, n. 1, p. 12–21, 4 jan. 2016.

ZHANG, K.; PEI, Z.; WANG, D. Organic solvent pretreatment of lignocellulosic biomass for biofuels and biochemicals: A review. **Bioresource Technology**, v. 199, p. 21–33, jan. 2016.

ZHENG, F.; DING, S. Processivity and enzymatic mode of a glycoside hydrolase family 5 endoglucanase from Volvariella volvacea. **Applied and Environmental Microbiology**, v. 79, n. 3, p. 989–996, 2013.

ZHOU, Y. et al. A novel efficient β-glucanase from a paddy soil microbial metagenome with versatile activities. **Biotechnology for Biofuels**, v. 9, n. 1, p. 36, 13 dez. 2016.

## ANEXO 1. DOCUMENTOS CIENTÍFICOS PUBLICADOS NO PERÍODO DE DESENVOLVIMENTO DA DISSERTAÇÃO

**Publicação 1**

# Chemical stability of a cold-active cellulase with high tolerance toward surfactants and chaotropic agent

Thaís V. Souza[a,1], Juscemácia N. Araujo[a,1], Viviam M. da Silva[a], Marcelo V. Liberato[b], Agnes C. Pimentel[b], Thabata M. Alvarez[b], Fabio M. Squina[b], Wanius Garcia[a,*]

[a] Centro de Ciências Naturais e Humanas, Universidade Federal do ABC (UFABC), Santo André, SP, Brazil
[b] Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, SP, Brazil

**ABSTRACT**

CelE1 is a cold-active endo-acting glucanase with high activity at a broad temperature range and under alkaline conditions. Here, we examined the effects of pH on the secondary and tertiary structures, net charge, and activity of CelE1. Although variation in pH showed a small effect in the enzyme structure, the activity was highly influenced at acidic conditions, while reached the optimum activity at pH 8. Furthermore, to estimate whether CelE1 could be used as detergent additives, CelE1 activity was evaluated in the presence of surfactants. Ionic and nonionic surfactants were not able to reduce CelE1 activity significantly. Therefore, CelE1 was found to be promising candidate for use as detergent additives. Finally, we reported a thermodynamic analysis based on the structural stability and the chemical unfolding/refolding process of CelE1. The results indicated that the chemical unfolding proceeds as a reversible two-state process. These data can be useful for biotechnological applications.
© 2015 The Author. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

## 1. Introduction

Cellulose is the most abundant polysaccharide found in the cell walls of plants, and provides the major renewable energy source on the planet [1,2]. Cellulases are important enzymes produced by various groups of microorganisms that can hydrolyze the $\beta$-1,4-linkage between the glucose units of the cellulose polymer. These enzymes are classified into glycoside hydrolase (GH) families and can be divided into three groups: endoglucanases (EC 3.2.1.4), cellobiohydrolases (EC 3.2.1.91), and $\beta$-glucosidases (EC 3.2.1.21) [3].

The complete enzymatic hydrolysis of cellulose polymers to glucose units requires at least one type of each enzyme described above working synergistically [4–6]. Briefly, endoglucanases randomly cleave the $\beta$-1,4-glycosidic linkages of cellulose, while cellobiohydrolases cleave off cellobiose units (dimer of glucose) from the reducing or nonreducing end of cellulose chain [4,5]. At last, $\beta$-glucosidases hydrolyze the released cellobiose to glucose [6,7].

A recent study reported a novel cellulase identified from a sugarcane soil metagomic library [8]. This cellulase named CelE1,

belonging to GH5 family, showed unusual catalytic properties. CelE1 is an endo-acting glucanase with high catalytic activity at a broad temperature range. Furthermore, CelE1 showed high catalytic activity at alkaline pH values, conditions that commonly cause enzyme inactivation of classical acidic cellulases [8]. Also, CelE1 is a cold-active enzyme that retains more than 70% of the activity at temperatures between 10 °C and 50 °C [8].

Cellulases are being used in several industrial applications, for example, in the pulp and paper, textile, and detergent industries, and more recently in the field of plant structural polysaccharides conversion into bioenergy [9–11]. Specifically, as essential criteria for using of cellulases as detergent additives, the enzyme should be active under low temperature, stable under alkaline conditions, and compatible with detergents [11]. Alkaline cellulases are added into detergents for laundry purposes. For example, cellulases remove cellulose microfibrils produced during manufacturing and washing of the cotton-based cloth [9]. The detergent industry has interest on cellulases for lower washing temperatures and also reduction in water consumption [12].

In this study, we examined the effects of pH on the secondary and tertiary structures, net charge, and also enzymatic activity of CelE1 using biophysical and biochemical techniques. Furthermore, to estimate whether CelE1 could be used as detergent additives, CelE1 activity was evaluated in the presence of surfactants (ionic and nonionic). Finally, we reported a thermodynamic analysis based on the structural stability and the chemical unfolding/

---

* *Corresponding author.
*E-mail address:* wanius.garcia@ufabc.edu.br (W. Garcia).
[1] These authors contributed equally to this work.

**Publicação 2**

# Biochemical and biophysical properties of a metagenome-derived GH5 endoglucanase displaying an unconventional domain architecture

CrossMark

Agnes C. Pimentel [a,b], Gabriela C.G. Ematsu [a], Marcelo V. Liberato [a], Douglas A.A. Paixão [a,b], João Paulo L. Franco Cairo [a,b], Fernanda Mandelli [a], Robson Tramontina [a,c], César A. Gandin [d], Mario de Oliveira Neto [d], Fabio M. Squina [a,**], Thabata M. Alvarez [a,*]

[a] Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Caixa Postal 6192, CEP 13083-970, Campinas, São Paulo, Brazil
[b] Departamento de Bioquímica, Instituto de Biologia (IB), Universidade Estadual de Campinas (UNICAMP), R. Monteiro Lobato, 255 – Cidade Universitária, Campinas, São Paulo, Brazil
[c] Programa de Pós Graduação em Biociências e Tecnologia de Produtos Bioativos (BTPB), Instituto de Biologia (IB) – CP 6109, Universidade Estadual de Campinas (UNICAMP), 13083-970, Campinas, SP, Brazil
[d] Departamento de Física e Biofísica, Instituto de Biociências de Botucatu, UNESP Univ Estadual Paulista, Distrito de Rubião Jr. s/n, Botucatu, SP, Brazil

## ARTICLE INFO

## ABSTRACT

Endoglucanases are key enzymes in the degradation of cellulose, the most abundant polymer on Earth. The aim of this work was to perform the biochemical and biophysical characterization of CelE2, a soil metagenome derived endoglucanase. CelE2 harbors a conserved domain from glycoside hydrolase family 5 (GH5) and a C-terminal domain with identity to Calx-beta domains. The recombinant CelE2 displayed preference for hydrolysis of oat beta-glucan, followed by lichenan and carboxymethyl cellulose. Optimum values of enzymatic activity were observed at 45 °C and pH 5.3, and CelE2 exhibited considerable thermal stability at 40 °C for up to 360 min. Regarding the cleavage pattern on polysaccharides, the release of oligosaccharides with a wide degree of polymerization indicated a characteristic of endoglucanase activity. Furthermore, the analysis of products generated from the cleavage of cellooligosaccharides suggested that CelE2 exhibited transglycosylation activity. Interestingly, the presence of $CaCl_2$ positively affect CelE2, including in the presence of surfactants. SAXS experiments provided key information on the effect of $CaCl_2$ on the stability of CelE2 and dummy atom and rigid-body models were generated. To the best of our knowledge this is the first biochemical and biophysical characterization of an endoglucanase from family GH5 displaying this unconventional modular organization.

© 2017 Elsevier B.V. All rights reserved.

## 1. Introduction

The main drivers for the development of renewable energy sources are the high dependence on nonrenewable energy sources coupled with their price fluctuation and environmental concerns regarding the emission of greenhouse gases from burning of fossil fuels [1]. In this context, lignocellulosic biomass represents a promising renewable source for the production of biofuels [2,3].

Lignocellulosic biomass is typically composed of 40–60% cellulose, 20–40% hemicellulose, 10–25% lignin and minor amounts of ash, acids and extractives [4]. The production of biofuel from lignocellulosic biomass involves deconstruction of the plant cell wall and depolymerization of cellulose and hemicellulose into monomeric sugars through a combination of physicochemical pre-treatment and enzymatic hydrolysis [4,5]. Glycoside hydrolase (GH) is a class

* Corresponding author. Present address: Universidade Positivo, Master in Industrial Biotechnology, R. Prof. Pedro Viriato Parigot de Souza, 5300, Cidade Industrial, 81280-330, Curitiba, Paraná, Brazil.
** Corresponding author. Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Caixa Postal 6192, CEP 13083-970, Campinas, São Paulo, Brazil.
E-mail addresses: agnescristinap@gmail.com (A.C. Pimentel), gabriela.ematsu@bioetanol.org.br (G.C.G. Ematsu), marcelovliberato@gmail.com (M.V. Liberato), douglas.paixao@bioetanol.org.br (D.A.A. Paixão), jpcairo@gmail.com (J.P.L. Franco Cairo), fernanda.mandelli@bioetanol.org.br (F. Mandelli), robson.tramontina@gmail.com (R. Tramontina), cesargandin@gmail.com (C.A. Gandin), mario.neto@ibb.unesp.br (M. de Oliveira Neto), fmsquina@gmail.com (F.M. Squina), thabata.alvarez@up.edu.br (T.M. Alvarez).

## ANEXO 2 – FORMULÁRIO DE APROVAÇÃO DA COMISSÃO INTERNA DE BIOSSEGURANÇA (CIBIO)

Uso exclusivo da CIBio:

| Número de projeto / processo: | 2016-26 |

Formulário de encaminhamento de projetos de pesquisa com OGMs para análise da CIBio - CNPEM

**1. Título do projeto:** INFLUÊNCIA DE UM DOMÍNIO ACESSÓRIO NAS CARACTERÍSTICAS BIOQUÍMICAS E ESTRUTURAIS DA CELULASE CelE2

**2. Pesquisador responsável:** Fabio Squina (orientador), Thabata M. Alvarez (coorientadora)

**3. Experimentador(es):**
Agnes Cristina Pimentel

Nível do treinamento do experimentador:   ||-Iniciação científica,   |X|-mestrado,   | |-doutorado,
| |-doutorado direto,   |  |-pós-doutorado,   | |-nível técnico,  | |-outro, especifique:

**4. Unidade operativa:** [   ] LNLS   [   ] LNNano   [X] CTBE   [   ] LNBio

**5. Maior Classe de risco de OGM deste projeto:** [ X ] Risco I    [ ] Risco II    [   ] Risco III    [   ] Risco IV

**6. O projeto é confidencial?** [ x ] não [   ] sim

**7. No caso de projeto confidencial, o título do projeto pode constar em lista aberta no CNPEM?** [   ] não [   ] sim

**8. Qual é o objetivo do projeto?** Elucidação da influência de um domínio proteico acessório nas propriedades bioquímicas e estruturais da celulase CelE2

**9. Informe um número e nome para cada OGM, organismo receptor, organismo doador, o transgene e classe de risco do OGM.** Ex: (1) Ecoli-CelE2_Cat_C, bactéria E Coli, origem metagenoma de solo, domínio catalítico de celulase, risco I. (2) Ecoli-CelE2_Cat_L, bactéria E.coli, origem metagenoma de solo, domínio catalítico de celulase, risco I. (3) Ecoli-CelE2_Dom_C, bactéria E Coli, origem metagenoma de solo, domínio proteico acessório de celulase, risco I. (4) Ecoli-CelE2_Dom_L, bactéria E Coli, origem metagenoma de solo, domínio proteico acessório de celulase, risco I.

**10. Descreva brevemente a função dos transgenes de cada OGM:** (1) e (2) domínio catalítico de celulase: apresenta atividade de endoglucanase. A diferença entre ambos é o tamanho da construção, de modo que (2) apresenta extensão na região C-terminal. (3) e (4) domínio proteico acessório de celulase: sem função definida e o objetivo deste projeto é buscar elucidar a função deste domínio. Estudos anteriores indicam possível influência na manutenção da estabilidade proteica. A diferença entre ambos é o tamanho da construção, de modo que (4) apresenta extensão na região N-terminal

**11. Algum OGM produz proteína tóxica, oncogênica ou pode gerar produtos deletérios para saúde humana, animal ou meio ambiente?** Não.
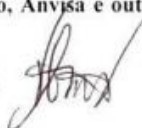
**12. Algum OGM é agente patogênico esporulante?** |  x  | Não   |   | Sim: _____

**13. Algum OGM é agente patogênico e pode se propagar pelo ar?** | x | Não   |   | Sim : _____

**14. Algum transgene confere infectividade ou patogenicidade para os OGMs? Descreva.** Todos os OGMs apresentam resistência ao antibiótico canamicina.

O pesquisador principal tem conhecimento de que os experimentadores envolvidos devem ser supervisionados na execução do projeto e que é o responsável pelo treinamento de biossegurança adequado às suas necessidades para a manipulação, armazenamento, descarte e transporte de OGMs, atendendo a legislação e normativas preconizadas pela CTNBio, Anvisa e outros órgãos e agências regulamentadoras e fiscalizadoras.

Assinatura eletrônica do pesquisador responsável:

Número de projeto / processo:    2016-26

Formulário de encaminhamento de projetos de pesquisa com OGMs para análise da CIBio - CNPEM

A CIBio analisou este projeto em reunião realizada no dia: _5/7/2016_ .

Parecer final: [X]-projeto aprovado, [ ]-projeto recusado, [ ]-projeto com deficiências.

comentários da CIBio:

Presidente da CIBio – CNPEM-LNBio
Marcio Chaim Bajgelman

Membro da CIBio da CNPEM –CTBE
Roberto Ruller

Membro da CIBio – CNPEM -LNBio
Celso Eduardo Benedetti
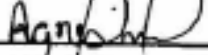
Membro da CIBio da CNPEM –CTBE
Sindélia Freitas Azzoni

Membro da CIBio – CNPEM -LNBio
Carolina Borsoi Moraes H. Freitas

Membro da CIBio da CNPEM –CTBE
Marina Camara Mattos Martins Soldi

**ANEXO 3 – TERMO DE DIREITOS AUTORAIS**

Declaração

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **"Características bioquímicas e estruturais do domínio catalítico da celulase CelE2 e sua relação com o domínio acessório Calx-β"**, não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 03 de agosto de 2018

Assinatura : _____

Nome do(a) autor(a): **Agnes Cristina Pimentel**

RG n.º 495330255

Assinatura : _____

Nome do(a) orientador(a): **Fabio Marcio Squina**

RG n.º 209812-77