UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

FRANCIELI COLUSSI

**Caracterização bioquímica, biofísica e estrutural da Celobiohidrolase I
de *Trichoderma harzianum* envolvida na hidrólise da biomassa
lignocelulósica**

São Carlos
2012

FRANCIELI COLUSSI

**Caracterização bioquímica, biofísica e estrutural da Celobiohidrolase I
de *Trichoderma harzianum* envolvida na hidrólise da biomassa
lignocelulósica**

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Doutor em Ciências.

Área de concentração: Física Aplicada
Opção: Física Biomolecular

Orientador: Prof. Dr. Igor Polikarpov

Versão Corrigida

(Versão original disponível na Unidade que aloja o Programa)

**São Carlos
2012**

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pelo Serviço de Biblioteca e Informação do IFSC,
com os dados fornecidos pelo(a) autor(a)

Colussi, Francieli

Caracterização bioquímica, biofísica e estrutural da Celobiohidrolase I de *Trichoderma harzianum* envolvida na hidrólise da biomassa lignocelulósica / Francieli Colussi; orientador Igor Polikarpov - versão corrigida -- São Carlos, 2012.
191 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Aplicada Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2012.

1. Celobiohidrolase I. 2. Hidrólise enzimática. 3. Etanol celulósico. 4. *Trichoderma harzianum*. I. Polikarpov, Igor, orient. II. Título.

À minha família, meu porto seguro.

AGRADECIMENTOS

Ao Prof. Dr. Igor Polikarpov, pela oportunidade de fazer parte do Grupo de Biotecnologia Molecular e pela orientação,

À Maria Auxiliadora, por ter sempre aquele sorriso sereno e calmante para nos acolher; Susana, pelas nossas boas conversas e planos; Bianca, Livia e Kelven, obrigada pela amizade e pelo apoio técnico.

Às minhas amigas Viviane Serpa, por me fazer aprender muito mais que o mundo das proteínas e a ser uma pessoa melhor a cada dia de convivência, Maria Luiza, meu anjo da guarda, sempre! Renata Alves, com seu sorriso sempre ali, independente da situação, Ana Cristina e Amanda Bernardes por estarem sempre prontas, tanto para o trabalho quanto para o 'happy hour', Marisa Lima, pela paciência, companheirismo e amizade enquanto dividimos casa, mesa e bancada por mais de 3 anos, Lis Schwartz e Andrei Nicoli, pelos jantares glamurosos, Vanessa e ao meu amigo Flavio Rodolfo, por me salvar de todos os problemas de informática que tive no doutorado, a todos vocês agradeço pelos bons momentos, pelas discussões e a ajuda de cada um! Vocês fizeram meus dias mais doces!

Às queridas Livia Faim, Daiana Evelin, Maria Amélia, Fernanda Costa, Simone Michelin, Michele Fagundes, Teresa, Fernanda Batista e aos Fernandes Bachega e Maluf, Jademilson, César, Tavin, Humberto, Caio Vinicius, Ivan, pela amizade e aprendizado com cada um.

Às amigas de longe, Paola Pissaia, Juliana Moraes e Barbara Fiorin.

Ao Bruno, meu ex-aluno de iniciação científica, que certamente trilhará uma vida brilhante na pesquisa.

Ao Dr. Wanius Garcia e Dra. Larissa Textor, pelo auxílio na interpretação de resultados.

Aos meus pais, pelo incentivo, apoio, confiança e orgulho que sentem de mim. Sem isso, não seria quem eu sou, nem teria chegado onde estou, minha mãe por ser essa mulher forte, companheira e tão amiga, e meu pai, por sempre nos incentivar a ir um pouquinho mais a frente... exemplo de energia e perseverança.

Ao meu irmão Jairo, pelo abraço apertado de “está tudo bem, logo você estará em casa de novo”, em cada despedida, e, com a Marina, por terem me dado a Júlia, minha sobrinha linda e fofucha,

Ao meu irmão Lucas, pelo otimismo e alegria de viver, sempre me fazendo achar as coisas menos difíceis, e pelo companheirismo, mesmo que de longe,

Á minha vó, Jandira, a vó mais legal que alguém pode ter! Sorrio só de lembrar as suas histórias!

Ao Paulo, meu namorado e amigo, por estar comigo por todo este caminho durante o doutorado, por aceitar as ausências, ter paciência, quando eu mesmo já não tinha mais, e nas

madrugadas de muito cansaço ir ao laboratório tirar ponto dos meus experimentos! Essas coisas são verdadeiras provas de amor!

Ao pessoal da secretaria, Silvio, Ricardo e Patricia;

Às meninas da biobiblioteca;

Ao IFSC, LNLS, LNNano, LNBio e CTBE pela disponibilidade do uso dos laboratórios para que esta pesquisa pudesse ser realizada.

À Capes pela concessão da bolsa de doutorado;

Enfim a todos que direta ou indiretamente auxiliaram para que este trabalho desse certo.

“É provável que a felicidade como coisa duradoura não exista, mas se ela pode ser encontrada em algum lugar, mesmo que de forma fugidía, é no presente. Para enxergá-la precisamos estar de olhos bem abertos, livres das sombras do passado e das luzes que cegam o futuro. Não é fácil, mas quem disse que seria simples?”

Ivan Martins

RESUMO

COLUSSI, F. **Caracterização bioquímica, biofísica e estrutural da Celobiohidrolase I de *Trichoderma harzianum* envolvida na hidrólise da biomassa lignocelulósica.** 2012. 191p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2012.

Devido à sua importante atividade celulolítica, o fungo *Trichoderma harzianum* possui um grande potencial de aplicação na hidrólise da biomassa. No entanto, as celulasas deste fungo filamentoso ainda não foram caracterizadas em profundidade. A celobiohidrolase I (CBHI) é a principal enzima celulolítica produzida por *Trichoderma* sp. e atualmente é uma das celulasas mais investigadas para aplicações de biocombustíveis. A CBHI hidrolisa celulose cristalina à unidades solúveis de celobiose, o que a torna uma enzima chave para a produção de açúcares fermentáveis a partir da biomassa. O objetivo deste trabalho foi purificar e caracterizar a CBHI de *Trichoderma harzianum* (ThCBHI) bioquímica, biofísica e estruturalmente. Primeiramente foi estabelecido um protocolo de purificação eficiente da proteína a partir da expressão homóloga no fungo. A caracterização bioquímica ThCBHI mostrou que a proteína possui uma massa molecular de 66 kDa, pI de 5,23 e o pH e a temperatura de atividade ótima foram 5,0 e 50 °C, respectivamente. A influência do pH e temperatura sobre as estruturas secundárias e terciárias e atividade enzimática da ThCBHI foram analisados por espectroscopia de CD, fluorescência e SAXS, e os resultados mostraram que as perturbações de pH e de temperatura afetam a estabilidade por dois mecanismos diferentes. As variações de pH podem modificar a protonação dos resíduos, afetando diretamente sua atividade, levando a desestabilização estrutural apenas em limites extremos de pH, como pH 9,0. A temperatura, por outro lado, tem uma influência direta sobre enovelamento e compactação da enzima, fazendo com que na temperatura em torno de 60 °C ocorra perda da estrutura secundária, e terciária. Quando as análises foram realizadas na presença do produto de reação e também inibidor competitivo, celobiose, a estabilidade térmica da ThCBHI aumentou significativamente de 61,5 para 65,9 °C. Os estudos estruturais e simulações de dinâmica molecular mostraram que a flexibilidade do resíduo Tyr260, em comparação com a Tyr247 do homólogo de *T. reesei* CBHI (TrCBHI), é aumentada devido às cadeias laterais curtas adjacentes de Val216 e Ala384 criando uma abertura adicional na face lateral do túnel catalítico. A ThCBHI também apresenta um *loop* encurtado na entrada do túnel de interação com a celulose, o que tem sido descrito como o responsável por interagir com o substrato de TrCBHI. Estas características estruturais podem explicar por que a ThCBHI apresenta maior valor de k_{cat} e menor inibição pelo produto em comparação com TrCBHI.

Palavras-chave: Celobiohidrolase I. Hidrólise enzimática. Biomassa. Etanol celulósico. *Trichoderma harzianum*.

ABSTRACT

COLUSSI, F. **Biochemistry, biophysics and structural characterization of cellobiohydrolase I from *Trichoderma harzianum* involved in the hydrolysis of lignocellulosic biomass.** 2012. 191p. Tese (Doutorado em Ciências), - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2012.

Trichoderma harzianum is a fungus that has a considerable potential in biomass hydrolysis application due to its elevated cellulolytic activity. Cellulases from *Trichoderma reesei* have been widely used as model in studies of cellulose breakdown. However, cellulases from *Trichoderma harzianum* are less-studied enzymes which have not been characterized biophysically and biochemically as yet. CBHI, a cellobiohydrolase I, is the major cellulolytic enzyme produced by *Trichoderma* sp. and is currently one of the most investigated cellulases for biofuel applications. CBHI hydrolyzes crystalline cellulose to soluble cellobiose units, which turns it into a key enzyme for producing fermentable sugars from biomass. The aim of this work was to purify and characterize the CBHI of *Trichoderma harzianum* (ThCBHI). We established an efficient purification protocol of ThCBHI, from the homologous expression. The biochemical characterization of ThCBHI showed that the protein has a molecular mass of 66 kDa, a pI of 5,23, and the optimum pH and temperature for its activity are 5,0 and 50 °C, respectively. The effect of pH and temperature on secondary and tertiary structure and enzymatic activity of ThCBHI were analyzed by CD and Fluorescence spectroscopy and showed that they affect protein stability by two distinct mechanisms. Variations of pH modify protonation of the residues, affecting directly its activity, leading to structural destabilization only at extreme pH values, such as pH 9, 0. On the other hand, temperature has direct influence on mobility, fold and compactness of the folding enzyme, at temperatures above 60 °C, there is loss of secondary and tertiary structure. When the assays were conducted in the presence of the cellobiose, a competitive inhibitor, thermal stability of ThCBHI was significantly increased to 61,5 to 65,9 °C. Structural studies and molecular dynamics simulations showed that the flexibility of Tyr260, in comparison to the Tyr247 from the homologous *T. reesei* CBHI, is enhanced due to the short side chains of adjacent Val216 and Ala384 residues and creates an additional gap at the side face of the catalytic tunnel. In addition, CBHI of *T. harzianum* has a shortened loop at the entrance of the cellulose-binding tunnel, which has been described to interact with the substrate in *T. reesei* CBHI. These structural features might explain why *T. harzianum* enzyme displays higher k_{cat} value and lower product inhibition on both glucosides and lactosides substrates in comparison to *T. reesei* CBHI.

Keywords: Cellobiohydrolase I. Enzymatic hydrolysis. Biomass. Cellulosic ethanol. *Trichoderma harzianum*.

LISTA DE FIGURAS

Figura 1 -	Estrutura da biomassa de origem vegetal.....	37
Figura 2 -	Excesso de bagaço de cana-de-açúcar nas usinas.....	39
Figura 3 -	Compostos químicos da hemicelulose.....	40
Figura 4 -	Estrutura geral da molécula de hemicelulose.....	40
Figura 5 -	Estrutura tridimensional da lignina.....	41
Figura 6 -	Estrutura da celobiose e ligação β -glicosídica.....	42
Figura 7 -	Estrutura da celulose.....	43
Figura 8 -	Ilustração da dinâmica da enzima CBHI durante o processo de desconstrução da celulose presente na biomassa lignocelulósica.....	48
Figura 9 -	Vista esquemática da biodegradação de celulose.....	49
Figura 10 -	Estrutura da enzima CBHI, demonstrando o linker altamente glicosilado.	50
Figura 11 -	Placa de meio PDA com o fungo <i>Trichoderma harzianum</i> esporulado, depois de 7 dias mantida em estufa a 30 °C.....	59
Figura 12 -	Câmara de Neubauer, com ênfase nos quadrantes utilizados para contagem dos esporos.....	59
Figura 13 -	Molécula de p-NP- β -D-cellobioside.....	66
Figura 15 -	Cinética enzimática durante os dias de cultivo de fungo <i>Trichoderma harzianum</i>	81
Figura 16 -	Cromatograma do extrato de proteínas totais de <i>Trichoderma harzianum</i>	82
Figura 17 -	SDS-PAGE da purificação da ThCBHI.....	84
Figura 18 -	Sequencia de ThCBHI depositada no banco de dados do NCBI (gi 50400675).....	85
Figura 19 -	Focalização isoeletrica de ThCBHI.....	86
Figura 20 -	Gel Nativo da ThCBHI e ThCCD.....	87
Figura 21 -	Gráfico de superfície de resposta de temperatura e pH em função da atividade da enzima CBHI com substrato pNPC.....	88
Figura 22 -	Estabilidade térmica celobiohidrolase mantida por 48h à	88

	50°C.....	
Figura 23 -	Atividades da ThCBHI intacta e seu CCD em diferentes substratos celulósicos.....	92
Figura 24 -	Efeito do pH na estrutura secundária da ThCBHI em 25°C monitorado por espectroscopia de Dicroísmo Circular.....	94
Figura 25 -	Efeito do pH na estrutura terciária da ThCBHI a 25 °C, monitorada por espectroscopia de fluorescência intrínseca em diferentes valores de pH.	95
Figura 26 -	Modelos gerados no programa I-Tasser para o domínio catalítico e módulo de ligação de celulose (CBM) mostrando a distribuição de triptofanos em uma molécula Th CBHI.....	96
Figura 27 -	Curva de espalhamento experimental em função do pH de ThCBHI.....	98
Figura 28 -	Desenovelamento térmico da ThCBHI.....	100
Figura 29 -	Análise utilizando a técnica de DLS para diferentes temperaturas para a ThCBHI.....	102
Figura 30 -	Curva de SAXS para calcular o Rg da Cellobiohidrolase I (lnI(q) vs q).....	103
Figura 31 -	Curva de espalhamento experimental em função da temperatura.....	104
Figura 32 -	Cristais de ThCBHI CCD em gota suspensa.....	106
Figura 33 -	Padrões de difração obtidos para ThCBHI.....	108
Figura 34 -	A). Organização estrutural da CBHI de <i>Trichoderma harzianum</i>	109
Figura 35 -	Resíduos laterais do túnel de ThCBHI e TrCBHI.....	109
Figura36 -	Estrutura baseada em alinhamento das celobiohidrolases.....	111
Figura 37 -	Superposição estrutural do Cel7A de <i>T. harzianum</i> (cinza) e <i>T. reesei</i> (azul, código PDB 8CEL).....	112
Figura 38 -	Representação superfície eletrostática potencial da CBHI de (A) <i>T. harzianum</i> (cinza), (B) <i>P. chrysosporium</i> (ciano) e (C) <i>T. reesei</i> (azul.....	113
Figura 39 -	Dinâmica Molecular de Celobiohidrolase I de <i>Trichoderma harzianum</i> e <i>Trichoderma reesei</i>	116
Figura 40 -	Microscopia Eletrônica de Varredura de Papel Filtro mostrando a ação da ThCBHI e uma endoglucanase de <i>T. harzianum</i> (ThEGI).....	117

Figura 41 -	Microscopia Eletrônica de Varredura de Papel Filtro mostrando a ação da ThCBHI e (ThEGI).....	120
-------------	---	-----

LISTA DE TABELAS

Tabela 1	Análise da Variância ANOVA	67
Tabela 2	Parâmetros cinéticos de ThCBHI.	69
Tabela 3	Passos da purificação da CBHI de <i>T. harzianum</i>	91
Tabela 4	Estatísticas de refinamentos para ThCBHI CCD	108

LISTA DE ABREVIATURAS E SIGLAS

CBH	celobiohidrolase
CBM	domínio de ligação a celulose
BG	beta-glicosidase
CD	dicroísmo circular
CMC	carboximetil celulose
D _{max}	distância máxima
DAM	modelo de átomos <i>dummy</i>
DC	domínio catalítico
DCCR	delineamento composto central rotacional
DNS	ácido dinitrosalicílico
DTT	ditioteritol
EDTA	ácido etilenodiamino tetra-acético
EG	endoglucanase
EX	exoglucanase
FPLC	fast performance liquid chromatography
GH	glicosil hidrolase
GOD	glicose oxidase
HPLC	high performance liquid chromatography
MS	espectrometria de massa
NR	extremidade não redutora
R	extremidade redutora
R _g	raio de giro
SAXS	espalhamento de raios X a baixo ângulo
SDS-PAGE	sodium dodecyl sulfate polycrylamide gel electrophoresis
SAXS	espalhamento de raios X a baixo ângulo
ThCBH	celobiohidrolase de <i>Trichoderma harzianum</i>
TrCBH	celobiohidrolase de <i>Trichoderma reesei</i>
UV	ultravioleta

LISTA DE SÍMBOLOS

Å	Angstrom
µm	Micrômetro
µM	Micromolar
° C	graus Celsius
CO ₂	dióxido de carbono
Da	Dalton
h	hora
L	Litro
M	Molar
min	minuto
mL	mililitro
nm	nanômetro
pH	potencial hidrogeniônico
rpm	rotação por minuto
λ	comprimento de onda

SUMÁRIO

1 INTRODUÇÃO.....	29
1.1 Biomassas Lignocelulósicas	36
1.2 Bagaço da Cana-de-Açúcar	38
1.3 Composição da Biomassa Lignocelulósica.....	39
1.3.1 Hemicelulose	39
1.3.2 Lignina.....	41
1.3.3 Celulose	42
1.4 Pré-tratamento.....	43
1.5 Enzimas Produzidas por Fungos	44
1.7 Hidrólise Enzimática da Biomassa Lignocelulose.....	46
1.7.1 Glicosil Hidrolases	46
1.7.2 Endoglucanases	47
1.7.3 Exoglucanases.....	47
1.7.4 β -glicosidase.....	48
1.9 As aplicações Industriais das Glicosil Hidrolases	52
2 JUSTIFICATIVA.....	53
3 MATERIAIS E MÉTODOS.....	55
3.1 Cultivo e estocagem do fungo <i>Trichoderma harzianum</i> IOC 3844	57
3.2 Meio de Cultura para pré-inóculo e produção das enzimas.....	58
3.3 Produção da enzimas	58
3.4 Purificação da CBHI.....	59
3.5 Separação proteolítica do domínio catalítico (CCD) e módulo de ligação à celulose (CBM).....	59
3.6 Análise de perfil proteico por eletroforese SDS-PAGE	60
3.7 Dosagem da concentração de proteína.....	60

3.8	Análise por eletroforese em gel não desnaturante (gel nativo)	60
3.9	Determinação do ponto isoeletrico	61
3.10	Estocagem das proteínas	62
3.11	Determinação da sequencia e identificação da proteína por espectrometria de massas	62
3.12	Ensaio de Atividade	62
3.12.1	Atividade em substrato solúvel, pNPCase	63
3.12.3	β -glicosidase	64
3.12.4	CMCase	64
3.12.5	Ensaio de atividade de FPase	65
3.12.7	Teste de conversão de celulose em diferentes substratos	65
3.13	Teste de estabilidade térmica da ThCBHI	66
3.14	Determinação das condições ótimas de atividade quanto ao pH e a temperatura por Delineamento Composto Central Rotacional (DCCR) para ThCBHI	66
3.15	Determinação de parâmetros cinéticos K_m e V_{max} para a enzima ThCBHI	67
3.16	Dicroísmo Circular	68
3.17	Modelo gerado por homologia	69
3.18	Determinação de Estabilidade por Dicroísmo Circular e Thermofluor	69
3.19	Espectroscopia de Emissão de Fluorescência	71
3.20	Espalhamento de Luz Dinâmico (DLS)	72
3.21	Estudos de espalhamento de raios-X a baixo ângulo (SAXS)	72
3.22	Predição dos sítios de glicosilação	73
3.2	Cristalização, coleta de dados e processamento	73
3.24	Microscopia Eletrônica de Varredura	75
4	RESULTADOS E DISCUSSÃO	77
4.1	Determinação do tempo de cultivo do <i>T. harzianum</i>	79
4.2	Purificação e identificação de ThCHBI	80
4.3	Espectrometria de Massa	82

4.4	Determinação do ponto isoelétrico (pI) da ThCBHI	83
4.5	Análise por eletroforese em gel não desnaturante (gel nativo).....	84
4.6	Estabilidade térmica da ThCBHI mantida por 48h à 50°C.....	86
4.7	Constantes cinéticas	87
4.8	Teste de Conversão da Celulose em diferentes substratos	88
4.10	Estudos biofísicos e estruturais da ThCBHI	90
4.10.1	Influência do pH sobre a estabilidade estrutural da ThCBHI	90
4.10.2	Análises da ThCBHI por Espalhamento de Raios X a Baixo Ângulo (SAXS) .	94
4.10.3	Influência da temperatura sobre a estabilidade e compactidade da ThCBHI.....	96
4.10.3	Influência da celobiose na estabilidade térmica da ThCBHI	98
4.10.3.1	Espalhamento dinâmico de luz (DLS)	99
4.10.3.2	SAXS.....	100
4.11	Estrutura cristalográfica do CCD de ThCBHI.....	102
4.11.1	Cristalização	103
4.11.2	Descrição da estrutura cristalográfica determinada	105
4.11.3	Comparação da estrutura de ThCBHI com outras celobiohidrolases I.....	108
4.12	Microscopia Eletrônica de Varredura	114
5	CONCLUSÕES	119
	REFERÊNCIAS	123
	ANEXO	139

1 INTRODUÇÃO

Atualmente, as mudanças climáticas que estão acontecendo, decorrentes do aquecimento global, têm se tornado um problema sério em todo o mundo e afetando drasticamente o clima em diferentes regiões da Terra, com consequências graves para a humanidade. Vários estudos têm associado o aquecimento global com crescentes níveis de CO₂ atmosférico gerado, inclusive, pela queima de combustíveis fósseis.¹ Assim, um amplo esforço científico tem sido empregado para a produção de bioenergia através da utilização e recuperação dos recursos naturais.² Energias alternativas como a eólica e solar são consideradas até o momento como as fontes mais populares de energias renováveis, no entanto, vários materiais biológicos, como resíduos e culturas agrícolas têm se mostrado como fontes potenciais para geração de bioenergia atualmente.

As mudanças climáticas e a elevação nos custos do petróleo aliadas às necessidades estratégicas de produção de energia têm motivado uma corrida sem precedentes à produção de combustíveis alternativos, preferencialmente de fontes renováveis. Isso teve início de forma efetiva na década de 70, devido à alta dos preços do petróleo e também, possibilidade de escassez devido a forte dependência em relação aos países produtores, politicamente instáveis. Nesse cenário, o Brasil despontou como o país com as tecnologias e políticas mais avançadas do mundo devido à pioneira utilização do etanol, obtido a partir da cana de açúcar como combustível. Em 1975, buscando uma solução para estes problemas, foi lançado no Brasil o Programa Nacional do Alcool – o Pró-álcool- primeiro programa de sucesso para produção de biocombustíveis em larga escala. Este Programa teve como objetivo a produção de etanol a partir da biomassa vegetal abundante e disponível, principalmente de sacarose e amido. O uso do bioetanol, além de possibilitar a produção de combustível a partir de matéria-prima não-fóssil e renovável, apresenta vantagens em relação à gasolina como menor emissão de poluentes, queima mais eficiente, e menor produção de precursores de ozônio e CO₂. Outra grande vantagem do bioetanol, é que o carbono liberado na sua queima é equivalente ao absorvido pela planta durante o seu crescimento, enquanto o dióxido de carbono gerado pela queima de gasolina se acumula na atmosfera.^{3; 4; 5}

Em relação ao petróleo refinado, a conversão de biomassa lignocelulósica oferece nova logística e desafios científicos, principalmente porque a energia desta biomassa por hectare é baixa.⁶ O transporte da biomassa, em longas distâncias até biorrefinarias lignocelulósicas, também ainda é um importante fator limitante.^{6; 7} A dificuldade de acesso

aos polissacarídeos da parede celular desses vegetais, denominada recalcitrância, é outra grande dificuldade para sua conversão, e consolidação da biomassa como fonte de energia.⁸

O Brasil e os EUA são os líderes mundiais na produção de bioetanol, a partir da fermentação de açúcares derivados de cana-de-açúcar e milho, respectivamente. Hoje, o Brasil desfruta de uma posição confortável em relação à tecnologia da produção de etanol. Nesse setor, o Brasil é líder mundial na produção de cana-de-açúcar, etanol e açúcar.⁹ A produção total de açúcar para a safra 2012/2013 está estimada em 38,85 milhões de toneladas, que equivalem a 777,0 milhões de sacas de 50 Kg. Para a produção de etanol serão esmagadas 302,2 milhões de toneladas de cana para a produção de 23,96 bilhões de litros de etanol, 4,81% maior que a produção da safra 2011/12.¹⁰

O Estado de São Paulo é o maior produtor de etanol no Brasil sendo, o segundo maior produtor depois dos EUA com ¼ da produção mundial de etanol.¹¹

Possivelmente, em breve, estas matérias-primas não serão suficientes para atender à crescente demanda mundial por esse combustível e alternativas precisam ser buscadas. Portanto a utilização de resíduos lignocelulolíticos pode contritubuir muito para a solução desse problema. Uma solução para a de escassez de energia são os biocombustíveis de segunda geração, que correspondem àqueles produzidos utilizando biomassas lignocelulósicas como matéria-prima, como por exemplo, o etanol.

1.1 Biomassas Lignocelulósicas

A biomassa lignocelulósica é uma fonte de matéria prima renovável que está sendo amplamente estudada. Produzir biocombustível a partir de biomassa lignocelulósica residual é uma alternativa ambientalmente favorável, uma vez que a emissão de CO₂ na atmosfera é compensada pela absorção deste gás durante o desenvolvimento vegetativo de novas biomassas.¹²

Vásquez et al., (2007)¹³ já reportaram a tendência por processos biotecnológicos que permitam a utilização de biomassas residuais de composição lignocelulósica. Essas matérias-primas são provenientes de resíduos agroindustriais como o sabugo e palha de milho, o bagaço da cana-de-açúcar, caule e cascas, resíduos sólidos urbanos (papel e papelão, lixo de jardim) e da indústria de papel e celulose.¹⁴

Por ano, a humanidade produz milhões de toneladas de resíduos agroflorestais. Em 2007, por exemplo, o Brasil produziu, incluindo os resíduos gerados na indústria de celulose, aproximadamente 400 milhões de toneladas de subprodutos agroflorestais.¹⁵

O material lignocelulósico é formado por estruturas organizadas e fibrosas, compostas principalmente por polissacarídeos chamados celulose e hemicelulose (cerca de 70% da massa seca), entremeados por outra macromolécula, a lignina (Figura 1), aos quais se encontram unidos por ligações covalentes e de hidrogênio. A função da lignina é de aumentar a resistência da estrutura da parede celular da planta a ataques químicos e enzimáticos.¹⁶ A partir da biomassa podem ser produzidos combustíveis sólidos, líquidos e gasosos. Assim, independentemente da forma e da fonte de energia utilizada, ela tem se mostrado, ao longo de décadas, um dos mais determinantes fatores de desenvolvimento econômico e social dos países industrializados e, por extensão, se tornando igualmente determinante em muitos aspectos da vida econômica e social no contexto atual.¹⁷

Figura 1 - Estrutura da biomassa de origem vegetal, demonstrando as principais componentes da parede celular vegetal: lignina, hemicelulose e celulose. (adaptado de www.genomics.energy.org)

Anualmente, são processadas $5,4 \times 10^8$ toneladas de cana-de-açúcar em todo o mundo, e, em média, 1 tonelada de cana-de-açúcar geram 280 Kg de bagaço.¹⁸ Quase a metade do total de bagaço produzido é utilizada para geração de energia elétrica nas usinas, o restante é armazenado. Isso justifica o interesse no desenvolvimento de métodos de produção biocombustíveis e produtos químicos a partir desse bagaço remanescente, uma vez que, oferece vantagens econômicas e ambientais.¹⁹ Esse material é produzido em grandes

quantidades pelas indústrias sucro-alcooleiras principalmente do Brasil, Índia, Cuba, China, México, Indonésia e Colômbia.¹⁹

Com o aumento da quantidade de excedente de bagaço e a existência de processos mais eficientes para a extração de glicídios, o bagaço pode ser destinado a etapas de pré-tratamento e posterior transformação em substâncias de maior interesse, aumentando a produtividade das indústrias, sem acréscimos nas áreas de plantio.²⁰ O bagaço e a palha da cana oferecem potencial para ampliar a fabricação nacional de etanol na atual capacidade instalada em mais de 35%, se apresentando como potencial solução para o déficit anual de 1 bilhão de litros. Tudo isso sem investimentos adicionais em terras ou competição com a produção de alimentos.¹³

1.2 *Bagaço da Cana-de-Açúcar*

O bagaço da cana-de-açúcar, até então considerado um resíduo agroindustrial, vem ganhando espaço na produção de energia de segunda geração. Com o desenfreado crescimento populacional e a demanda por energia, as indústrias investem em alternativas para continuar aumentando a produção do etanol e utilizar os subprodutos formados como fonte de energia através de recursos naturais. Dentre estes subprodutos, o bagaço de cana-de-açúcar certamente ocupa uma posição de destaque na geração de energia renovável. É composto basicamente por celulose (41-44%), hemicelulose (25-27%) e lignina (20-22%).²¹ Cerca de 2/3 da energia existente na cana-de-açúcar estão contidas no bagaço e na palha. Diversos processos já utilizam o bagaço excedente, principalmente na geração de energia. No entanto, como mostra a Figura 2, ainda sobra grande quantidade de bagaço e palha que ficam armazenados e podem ser utilizados para obtenção de etanol. Para tanto se faz necessária a hidrólise do bagaço e da palha, a fim de se obter açúcares que possam ser fermentados por leveduras industriais.

Figura 2 - Excesso de bagaço de cana-de-açúcar nas produzidos nas usinas de álcool e açúcar.

1.3 *Composição da Biomassa Lignocelulósica*

A biomassa lignocelulósica é formada por basicamente por três componentes principais:

1.3.1 *Hemicelulose*

A hemicelulose atua como elo químicos entre a celulose e a lignina.^{22; 23} É formada por uma estrutura complexa e ramificada de carboidratos poliméricos (p. ex. xilanos, arabinoxilanos, arabinanos, galactanos, arabinogalactanos, galactoglucomananos, xiloglucanos, glucanos) de ligações do tipo β -D- 1,3 e 1,4.²⁴ Xilano é o principal componente da fração hemicelulósica e o segundo polissacarídeo mais abundante na natureza.²⁵ A Figura 3 apresenta a estrutura dos principais monossacarídeos que compõem esses carboidratos.

Figura 3 – Principais compostos químicos da hemicelulose

A sua cadeia principal pode ser constituída de uma só unidade desses açúcares (homopolímeros) ou de duas ou mais unidades (heteropolímero) ligadas entre si por ligações do tipo glicosídicas β -1,4. As ramificações da hemicelulose são formadas a partir da cadeia principal e estabilizadas por pontes de hidrogênio com cadeias adjacentes. A Figura 4 apresenta a estrutura geral da hemicelulose.²⁶

Figura 4 - Estrutura geral da molécula de hemicelulose (adaptado de <http://www.ccruc.uga.edu>).

Devido à estrutura complexa da fração hemicelulósica, torna-se necessário uma variedade de enzimas para sua degradação ou modificação. As duas principais enzimas

capazes de despolimerizar a cadeia de hemicelulose são a endo-1,4- β -D-xilanase e a endo-1,4- β -D-mananase, ambas glicosil hidrolases.²⁷

1.3.2 Lignina

A lignina é um polímero abundante na natureza, que compõe até 20% da parede celular conferindo a planta suporte estrutural, impermeabilidade e resistência, assim como proteção contra ataques de patógenos ou pragas. As unidades básicas são interligadas por diferentes tipos de ligação, que incluem ligações C-C, aril-éter, aril-glicerol e β -ariléter como formas predominantes.²⁸ Essa barreira de lignina na composição lignocelulósica pode ser removida pela utilização de lignina peroxidases, deixando a estrutura mais susceptível à hidrólise dos outros polissacarídeos por ação de celulasas e hemicelulasas.²⁹ A Figura 5 apresenta a estrutura da lignina.

Figura 5 - Estrutura da lignina

1.3.3 Celulose

A celulose é o biopolímero renovável mais abundante na Terra,³⁰ o principal polímero constituinte da parede celular das plantas e, constantemente ressinetizado pela fotossíntese.³¹ Ela compõe cerca de 35-50% do peso seco dos tecidos vegetais, variando de acordo com o tipo de célula e a maturidade da planta.²⁷ É mais comumente encontrada associada a hemicelulose e lignina na parede celular, mas também pode existir na forma pura, como no algodão.²⁷ A celulose possui uma estrutura relativamente simples, sendo formada de monômeros de D-glicose unidos por ligações glicosídicas do tipo β -1,4. Cada resíduo de glicose tem uma orientação de 180° em relação ao resíduo adjacente, formando assim a celobiose, a subunidade repetitiva do polímero (Figura 6). Cada microfibrila é sintetizada independentemente e contém entre 2.000 e 20.000 resíduos de D-glicose. A extremidade da cadeia em que se encontra o resíduo de glicose cujo carbono anomérico está livre é chamada de “extremidade redutora”; a outra extremidade é chamada “não redutora”.³²

Figura 6 - Estrutura da celobiose e da ligação β -glicosídica na celulose (adaptado de Sandgren, 2005).

A estrutura cristalina da celulose está disposta em glicoses unidas por ligações de hidrogênio tanto intra quanto intercadeias de celulose (Figura 7). As camadas adjacentes de celulose são unidas por forças de *van der Waals*. Isso resulta em uma estrutura compacta, capaz de impedir o acesso de enzimas e até mesmo de moléculas pequenas, como a água. As regiões cristalinas da celulose são intermediadas por regiões menos compactas, chamadas amorfas, em que os polímeros de celulose apresentam maior distância uma das outras (Figura 7A). Além das regiões amorfas, as fibras de celulose apresentam irregularidades, falhas e

microporos que aumentam a superfície de contato da celulose, permitindo o acesso de moléculas de água e outras moléculas maiores, como enzimas celulolíticas.²⁷

Figura 7 - Estrutura da celulose. A) esquema mostrando organização de fibras de celulose, com regiões amorfas e cristalinas. B) Representação das ligações de hidrogênio entre cadeias (inter) e entre resíduos de glicose da mesma cadeia (intra) (Castro et. al. 2010)

1.4 Pré-tratamento

Matérias-primas baratas e tecnologias eficientes de conversão da biomassa são dois requisitos fundamentais para o sucesso na implementação de biorrefinarias de celulose.³³ As tecnologias de sacarificação enzimáticas são atualmente caras e ineficientes devido à recalcitrância das paredes celulares das plantas para produtos químicos e enzimas, assim, muitos estudos ainda buscam o desenvolvimento de estratégias eficientes de pré-tratamento.^{8;}

³³ Sendo que, o principal objetivo da maioria dos pré-tratamentos é superar essa barreira de lignina-hemicelulose que impedem a eficiente ação enzimática, e em alguns casos para reduzir a cristalinidade de celulose.

O rendimento da hidrólise pode variar em mais de 70% quando feito ou não o pré-tratamento da biomassa. Estes pré-tratamentos podem ser classificados como físicos, físico-químicos, químicos e biológicos, conforme o agente que atua na alteração estrutural.³⁴ Após os tratamentos para remoção da lignina e hemicelulose, a celulose é submetida a hidrólise enzimática para a obtenção de glicose.¹⁹

A hidrólise enzimática é catalisada por celulasas, enzimas que são secretadas por fungos e bactérias.

1.5 Enzimas Produzidas por Fungos

Na biotecnologia, os fungos destacam-se pela capacidade biodegradadora de resíduos naturais. Um grande número de fungos produzem importantes grupos de enzimas hemicelulolíticas, celulolíticas e lignolíticas que desempenham um papel fundamental no ciclo do carbono, pela decomposição de resíduos vegetais. Essas enzimas possuem diversas aplicações em processos de diferentes áreas industriais. Isto se deve, ao fato de enzimas atuarem como agentes que melhoram os processos industriais quanto à qualidade e eficiência do produto final, bem como aspectos de ordem econômica e ambiental.³⁵ Estes microrganismos, capazes de decompor hemicelulose, celulose e lignina, produzem uma série de enzimas com diferentes especificidades, que podem atuar em sinergia.³⁶

Entretanto, a exploração do potencial desses fungos como produtores de diversas enzimas extracelulares de aplicação industrial tais como, amilases, celulasas, xilanases, peroxidases, invertases, pectinases, queratinases, lipases e proteases, entre outras, somente se tornou possível com o avanço do conhecimento de sua fisiologia, bioquímica e genética.¹¹ A maior parte das enzimas hoje utilizadas é extraída de organismos mesófilos; portanto são bastante sensíveis a variação das condições ótimas de pH, temperaturas e presença de solventes orgânicos, por exemplo.³⁷

Apesar de vantagens, ainda existem vários empecilhos ao uso de enzimas em processos industriais e entre eles está relacionado ao alto custo destas. Atualmente as enzimas de origem microbiana são utilizadas no processamento de alimento, na manufatura de detergentes, nas indústrias têxtil e farmacêutica, terapia médica e na biologia molecular.³⁵

1.6 Fungos do Gênero *Trichoderma*

Fungos filamentosos são conhecidos por serem bons produtores de enzimas que degradam celulose e hemicelulose, e seu uso biotecnológico é um mercado crescente com promissoras aplicações no campo da bioenergia.² *Trichoderma* (Teleomorph: *Hypocrea*) é um gênero de fungos filamentosos amplamente distribuídos na natureza e que podem ser encontrados em ecossistemas que variam desde a tundra até o tropical.³⁸ Essa capacidade de sobrevivência em regiões tão diferentes pode ser atribuída a sua capacidade metabólica diversificada, que permite a degradação e utilização de carboidratos complexos, a sua agressividade natural competitiva, bem como a facilidade de proliferação incluindo ataque físico a outros fungos, características estas que despertam grande interesse biotecnológico e econômico.³⁹ Fungos do gênero *Trichoderma* possuem características que facilitam o seu reconhecimento, tais como crescimento rápido em cultura, produção de numerosos esporos (conídios- os quais são normalmente verdes) e micélio.⁴⁰

Trichoderma reesei (*T. reesei*) é o fungo celulolítico mais estudado atualmente, tendo várias estruturas cristalográficas resolvidas da maioria das enzimas do complexo. Estas enzimas por já terem suas características bem conhecidas, tornaram-se altamente eficientes e largamente utilizadas na indústria têxtil e degradação de celulose.⁴¹

Trichoderma harzianum (*T. harzianaum*), ainda pouco estudado, é um fungo filamentoso celulolítico com grande potencial biotecnológico. Estudos com algumas cepas de *T. harzianum* têm demonstrado que este fungo é capaz de produzir um complexo celulolítico com maior atividade β -glicosidase quando comparado ao *T. reesei*.²³ Além disso, *T. harzianum* secreta de um sistema enzimático eficiente e bem equilibrado de enzimas capazes de hidrolisar substratos celulósicos a glicose⁴². Isso representa uma vantagem considerável sobre *T. reesei*, uma vez que o extrato enzimático do fungo exibe um baixo nível de atividade de β -glicosidase, o que dificulta um processo de completa hidrólise e a fermentação da celulose em etanol.²⁷

1.7 Hidrólise Enzimática da Biomassa Lignocelulose

1.7.1 Glicosil Hidrolases

As celulases, envolvidas na hidrólise enzimática da biomassa lignocelulósica, são definidas pelo “Carbohydrate-Active enZYmes Database” (CAZy)⁴³ como Glicosil Hidrolases (GH). As GHs são classificadas com base na especificidade ao substrato com critérios estabelecidos pela União Internacional de Bioquímica e Biologia Molecular (IUBMB). De acordo com a IUBMB cada enzima recebe um código denominado ‘EC number’ (*Enzyme Commission number*), que são representadas da seguinte forma: EC 3.2.1.x, onde cada enzima dígitos caracteriza o tipo de reação catalisada: 1º dígito – classe de acordo com o tipo de reação química que catalisam, 2º dígito – subclasse que definem em que tipo de grupo as enzimas atuam, 3º dígito – subsubclasse continuando com o mesmo exemplo e o 4º dígito - indica o substrato, mas também pode representar em alguns casos o mecanismo molecular ou o tipo de ligação entre as moléculas.

A extensa variedade de carboidratos encontrados na parede celular das plantas requerem uma multiplicidade de enzimas para sua degradação. As glicosil hidrolases (GH) são enzimas-chave na hidrólise de carboidratos fazendo parte do grupo de enzimas que clivam ligações glicosídicas,⁴³ estas, denominadas celulases, reconhecem e clivam as ligações β -1,4 entre as moléculas de glicose.

As celulases são enzimas que agem em consórcio e podem ser separadas em três classes, basicamente, as endoglucanases, as exoglucanases e as β -glucosidases.²⁸⁻²⁹

As endoglucanases (EC: 3.2.1.4) são classificadas nas famílias GH 5, 6, 7, 8, 9, 12, 44, 45, 48, 51, 61; as exocelulases, celobiohidrolase I e II (EC: 3.2.1.91) pertencem as famílias GH 7 e 6, respectivamente enquanto a exocelulase glucanohidrolase (EC: 3.2.1.74) pertence as famílias GH 1 e 3; a terceira classe, as beta glicosidases (EC: 3.2.1.21), por sua vez, são classificadas nas famílias GH 1, 3, 30, 43 e 116.⁴⁴

1.7.2 Endoglucanases

As endoglucanases são também classificadas como celulasas, endo β -1,4glucanases e carboximetilcelulasas. Seu substrato natural é a celulose e xiloglicana, apresentando também especificidade sobre a carboximetilcelulose (CMC), β -glucana e xilana. Estas enzimas catalisam a hidrólise interna de ligações β -1,4-D-glucosídicas da celulose. Podendo, também, hidrolisar ligações β -1,4 em D-glucanas que contenham ligações β -1,3. As regiões de menor organização estrutural são mais facilmente atacadas, o que possibilita a ação da enzima ao longo da cadeia de celulose e reduz consideravelmente seu grau de polimerização,⁴⁵ pois possuem cadeias que não estão envolvidas em interações de hidrogênio intermoleculares tão fortes quanto as que ocorrem nas regiões cristalinas, levando, conseqüentemente, a uma maior exposição das ligações glicosídicas mais internas das cadeias de celulose.⁴⁶⁻⁴⁷

1.7.3 Exoglucanases

Conhecida também como β -1,4-celobiosidase, celobiohidrolase, β -1,4 celobiohidrolase ou avicelase, as exo-1,4- β -D-glucanases atuam nas extremidades redutoras e não redutoras da cadeia de celulose e celotetraose liberando majoritariamente celobiose, além de glicose e celotriose.⁴⁸ Estas enzimas atuam com bem menos eficiência sobre celulosas solúveis por haver um impedimento estereoquímico causado pelos grupos substituintes, seja carboximetílico (CMC) ou hidroxietílico (HEC). As exoglucanases atuam sobre Avicel, produzindo uma redução lenta e gradual do seu grau de polimerização. O sítio ativo das celobiohidrolases possui forma de túnel⁴⁹ através da qual a cadeia de celulose entra e sofre hidrólise de suas ligações glicosídicas terminais, liberando principalmente celobiose (Figura 8).^{50; 51; 52}

A celobiohidrolase I (CBHI), em particular, aparece como sendo a enzima mais importante para a indústria na conversão de celulose.²⁷ É considerada a enzima-chave no processo de hidrólise pois ela sozinha é capaz de hidrolisar celulose cristalina extensivamente.⁵³ Além disso, CBHI é a proteína predominantemente secretada pela maioria dos microorganismos produtores de celulasas. CBHs de espécies do gênero *Trichoderma* sp., quando

cultivadas sob condições induzidas para a produção de celulases, atingem até 60% do total de proteínas secretadas.⁵⁴

Figura 8 - Ilustração da dinâmica da enzima CBHI durante o processo de desconstrução da celulose presente na biomassa lignocelulósica. Figura adaptada de Beckham et al, 2010.⁵⁵

1.7.4 β -glicosidase

Estas enzimas são responsáveis pela hidrólise da celobiose liberada no meio reacional, pela ação das celobiohidrolases, em moléculas de glicose. Diminuindo, assim, o efeito de inibição das celobiohidrolases pelo produto, garantindo um processo de degradação eficiente.^{56; 57} Apresentam ampla especificidade por β -D-glicosídeos, podendo hidrolisar também β -D-galactosídeos, β -D-xilosídeos, α -L-arabinosídeos e β -D-fucosídeos.^{35; 58}

O processo de hidrólise começa com as endoglucanases que aleatoriamente hidrolisam ligações internas β -1,4-glicosídicas na região amorfa da celulose, reduzindo

significativamente o grau de polimerização do substrato e liberando novas extremidades na cadeia da celulose. As extremidades redutoras e não redutoras de celulose são atacadas pelas celobiohidrolases tipo I e tipo II, (CBHI e CBHII), respectivamente, liberando moléculas de celobiose (Figura 9). As unidades de celobiose liberados são, então, hidrolisadas em unidades de glicose pelas β -glicosidasas^{21; 27; 59; 60}.

Figura 9 - Vista esquemática da biodegradação de celulose. Celobiohidrolases atuam nos terminais redutores ou não redutores da fibras de celulose para processivamente libertar celobiose. Endo-1,4- β -glucanases aleatoriamente clivam cadeias de celulose e β -glicosidasas hidrolisam celobiose ou celooligômeros de glicose a partir das extremidades não redutoras. Figura Adaptada de Watanabe e Tokuda (2010).

1.8 Organização estrutural das celulases

A maioria das celulases exibem uma estrutura semelhante de organização, são constituídas por dois distintos domínios funcionais e estruturais: um domínio de ligação de

celulose (CBD) e um domínio catalítico (CCD) (Figura 10). Os domínios são unidos por uma sequência longa e altamente O-glicosilada.⁶¹

Figura 10 - Estrutura da enzima CBHI, demonstrando o *linker* altamente glicosilado. Figura adaptada a partir de Beckham, et al. (2010).

- Domínio Catalítico (CCD)

O domínio catalítico é a parte das celulases onde ocorre a hidrólise da cadeia de celulose. Este domínio é a maior parte da enzima, e o tamanho varia muito entre as diferentes celulases, por exemplo, os domínios catalíticos das celulases de *T. reesei* variam de 166 a 430 aminoácidos.⁶²

- Região do “Linker”

Os dois domínios da enzima são conectados por um *linker* flexível. O comprimento deste *linker* varia em tamanho, de menos de 20 para mais de 40 aminoácidos. Usualmente são

ricos em treoninas, serinas e prolinas, e são altamente glicosilados.⁶³ A função do *linker* é provavelmente manter os domínios independentes, e restringir os seus movimentos de um em relação ao outro, de modo que o domínio catalítico está sempre a uma curta distância do CBM, que se liga na a superfície da fibra de celulose. Provavelmente a glicosilação do *linker* faz com que ele tenha menos flexibilidade, e diminua a sensibilidade para enzimas proteolíticas que também são secretadas pelo microorganismo.⁶³⁻⁶⁴

- Módulo de Ligação à Celulose (CBM)

O CBM é o menor domínio das celulasas, consistindo de aproximadamente 35 aminoácidos. Ele pode ser ligado através do *linker* flexível ao módulo catalítico. A função do CBM é se ligar na fibra, e servir como uma "âncora" para a enzima, mantendo-o fortemente adsorvido à superfície da celulose.⁶⁵⁻⁶⁶ Isto reduz a necessidade de forte ligação do domínio catalítico com a celulose, e, assim, permite que a enzima tenha maior taxa de hidrólise.⁶⁷ Ainda não há evidência de que os CBMs de fungos podem penetrar na fibra de celulose e romper sua estrutura, ou que tenha alguma atividade catalítica.

- *Glicosilação*

Em geral, as celulasas secretadas por fungos filamentosos, são glicoproteínas que contém sítios O- ou N- glicosilados.⁶⁸ Estudos estruturais têm mostrado que a N-glicosilação da CBHI nas cepas de *T. reesei* QM9414 e ALKO2877 (derivado QM9414) possuem uma única modificação de N-acetilglicosamina em três ou quatro locais potenciais localizados na domínio catalítico da proteína.⁶³ A ocorrência de tais modificações depende de vários fatores, incluindo a estirpe e as condições de cultivo e gera várias isoformas da enzima com semelhantes propriedades catalíticas e adsorção.⁶⁹⁻⁷⁰

Um grande número de estruturas tridimensionais de celulasas de *T. reesei* foi determinado. As estruturas de CCD de CBHI (PDB ID 1Cel),⁷¹ CBHII (3CBH),⁷² EG1 (1EG1)⁵⁹ e EGIII (1H8V),⁷³ CBM da CBHI (1CBH)⁷⁴ e EGI (1AZJ)⁷⁵ possuem as estruturas

cristalográficas já determinadas. As proteínas intactas ou *full length* ainda não estão disponíveis porque a presença da região de ligação extremamente móvel e glicosilada dificulta a cristalização. Estruturas do domínio catalítico da *T. reesei* CBHI (TrCBHI) foram também determinados complexadas com uma série de oligossacarídeos e ligantes ⁵⁰. Geralmente, esse domínio é composto de um grande número de folhas- β anti paralelas formando uma estrutura terciária β -sanduíche. Um túnel alongado de ligação à celulose envolto por loops e pontes dissulfeto é formado ao longo do lado côncavo do β -sanduíche. O túnel inclui o sítio ativo em seu extremo, onde os resíduos catalíticos de aspartato e glutamatos estão localizados ⁷⁶. A presença deste túnel alongado mostra que a especificidade da enzima é obtida por um número substancial de interações entre as extremidades da cadeia de celulose e os resíduos localizados nessa região ⁵¹. A cadeia de polissacarídeo atravessa o túnel de ligação e cada segunda ligação glicosídica do substrato é corretamente reconhecida pelo aparato catalisador, o que explica a processividade destas enzimas que liberam unidades de dissacarídeo (celobiose) da cadeia de celulose ⁷⁷ (Figura 8).

1.9 As aplicações Industriais das Glicosil Hidrolases

As GHs, fúngicas ou bacterianas, são amplamente utilizadas em aplicações industriais. As celulasas têm sido utilizadas na indústria têxtil para enzimaticamente produzir o efeito de lavagem (efeito *stone-wash*) em jeans, ³⁵ e como agentes em detergentes para remover fibras curtas na superfície têxtil (*depilling*), para produzir tecidos mais finos. ⁷⁷ Nas indústrias de suco e vinho, pectinases e celulasas são utilizadas para maceração e clarificação das bebidas. Na alimentação animal, a adição de GHs, principalmente β -glucanases e xilanases, têm aumentado a digestibilidade dos alimentos. ⁷⁸ Nas indústrias de papel, xilanases e mananases são adicionados em processos de branqueamento. Enfim, existe uma gama de aplicações industriais para às celulasas, principalmente em indústrias de processamento de amido, fermentação do álcool de cereais, malte e cerveja, extração de sucos de frutas e vegetais, celulose, papel e têxtil. ^{48; 79} Um futuro promissor para aplicações destas celulasas, incluem a produção de etanol combustível, a partir da biomassa lignocelulósica abundantes, através da degradação enzimática da celulose, substituindo assim os combustíveis derivados do petróleo.

2 JUSTIFICATIVA

Nos últimos anos, a preocupação com a diminuição dos recursos naturais tem gerado uma busca por fontes de energias renováveis, e que não atinjam as áreas agriculturáveis, não comprometendo assim, a produção de alimentos. Com isso, a procura por biocombustíveis tem se tornado de extrema importância também para o meio ambiente. Insere-se, nesse contexto, com grande potencial, o uso do bioetanol como fonte renovável de energia. Contudo, a quantidade de álcool necessária para suprir a demanda, cada vez maior, cria a necessidade de pesquisas por novas formas de obtenção dos biocombustíveis. A conversão enzimática de resíduos agroindustriais, materiais lignocelulósicos, como cascas de cereais, bagaço de cana e resíduos da indústria madeireira tem ganhado destaque nessa busca por soluções aos problemas gerados pelos combustíveis fósseis. Os materiais lignocelulósicos são considerados as mais abundantes fontes renováveis na Terra e parece claro que essa biomassa (primária e residual) será, num futuro próximo, a principal fonte de recursos para a obtenção de energia. A hidrólise enzimática, que converte biomassa lignocelulósica em açúcares fermentáveis pode ser o passo mais complexo no processo de produção de etanol de segunda geração.

Com isso, o entendimento e conseqüente melhoramento da atividade de enzimas capazes de degradar a celulose e a hemicelulose à moléculas de açúcares fermentáveis a etanol tem assumido papel essencial nas pesquisas relacionadas aos biocombustíveis.

O fungo *T. harzianum* apresenta potencial atividade celulolítica, contudo, pouco se sabe sobre a composição do complexo celulolítico deste, tão pouco suas características moleculares. Tendo em vista a alta atividade celulolítica deste fungo filamentoso, já estudada por Castro et. al., 2010, é de suma importância a compreensão dos mecanismos moleculares envolvidos, bem como a caracterização estrutural das enzimas desse complexo celulolítico.

A fim de investigar as enzimas que fazem parte deste complexo celulolítico, os objetivos deste trabalho foram purificar a celobiohidrolase I produzida por expressão homóloga, fazer a caracterização enzimática, bioquímica e estrutural da mesma. Desta forma, pretende-se contribuir para a otimização da produção de etanol a partir do bagaço da cana-de-açúcar pelo processo de hidrólise enzimática.

3 MATERIAIS E MÉTODOS

3.1 Cultivo e estocagem do fungo *Trichoderma harzianum* IOC 3844

Esporos de *Trichoderma harzianum* IOC 3844, mantidos em freezer -80 °C com 20% (v/v) glicerol, foram usados para inocular placas de Petri com meio PDA (potato dextrose agar). Após inóculo, as placas foram mantidas em estufa a 28 °C, por sete dias (Figura 11).

Figura 11 - Placa de meio PDA com o fungo *Trichoderma harzianum* esporulado, depois de 7 dias mantida em estufa a 28 °C.

Estes esporos foram ressuspensos em solução salina 0,9% (p/v) NaCl e a contagem de esporos foi realizada em câmara de Neubauer (Figura 12).

Figura 12 - Câmara de Neubauer, com ênfase nos quadrantes utilizados para contagem dos esporos.

Depois de feita a média dos esporos nos 5 quadrantes escolhidos (em destaque vermelho na Figura 12), foram realizados os cálculos necessários para se obter os números de esporos por mL a partir da equação:

$$N = \frac{Ne \times F}{Vq} \times 1000$$

Onde:

N= número de esporos por mL de suspensão (esporos/mL)

ne= Média dos 5 quadrantes contados na câmara

F= fator de diluição

Vq= volume do quadrículo : 0.004mm³

1000 = transformação de µL para mL

3.2 Meio de Cultura para pré-inóculo e produção das enzimas

Os meios utilizados para inóculo e pré-inóculo foram adaptados de Mandels,⁸¹ os quais contêm: 0.3 g Uréia /L; 2.0 g KH₂PO₄ /L; 1.4 g (NH₄)₂SO₄ /L; 0.4 g CaCl₂ · 2H₂O /L; 0.3 g MgSO₄ 7H₂O /L; 0.6 g Extrato de levedura /L; 5.0 mg FeSO₄7H₂O /L; 1.6 mg MnSO₄ ·4H₂O /L; 1.4 mg ZnSO₄ ·7H₂O /L; 2.0 mg CoCl₂ 6H₂O /L; 10 g glicose /L. Para a expressão das celulasas o meio de cultura foi o mesmo meio do pré-inóculo, mas a fonte de carbono utilizada foi Avicel 10g/L, substituindo a glicose.

3.3 Produção da enzimas

Erlenmeyers de 500 ml contendo 100 ml de meio de cultura foram inoculados em uma concentração final de 10⁶ esporos/mL e incubados por 48 horas a 28 °C em shaker a 250 rotações min⁻¹. Fermentador Bioflo 115 (New Brunswick Scientific Co., EUA) 7,5 L, equipado com controle automático de temperatura, pH, agitação, aeração e adição de espuma. O volume de meio para a produção da enzima foi de 4 L, inoculado com 10% (v/v) da cultura e temperatura mantida a 28°C, agitação de 200 rotações min⁻¹. A taxa de aeração foi ajustada de modo que o nível de O₂ dissolvido nunca fosse abaixo de 60% de saturação de ar no meio de cultura. O pH foi de 5,0 ± 0,2 e controlado automaticamente por meio da adição de ácido fosfórico ou NaOH 20% (p/v). A formação de espuma foi suprimida pela adição do antiespumante Antifoam (Sigma). Após 5 dias, o cultivo foi encerrado, o micélio e o Avicel remanescente foram removidos por ultrafiltração de membrana PVDF de 0,22µm. O sobrenadante foi concentrado 20 vezes, por ultrafiltração tangencial em sistema *hollowfiber*, com cartucho de 10 kDa de corte (UFP-10-E-35) (GE) e dialisado contra tampão Tris-HCl 50 mM, pH 7,0. A preparação da celulasas foi, então, armazenada a 4 °C até a purificação

3.4 Purificação da CBHI

A celobiohidrolase I de *Trichoderma harzianum* ThCBHI foi purificada por cromatografia líquida em equipamento FPLC (*Fast Performance Liquid Chromatography* – Amersham Bioscience).

O extrato bruto dialisado este foi submetida à coluna DEAE Sepharose. A proteína foi separada em tampão Tris 50 mM pH 7,0 e eluída em gradiente de 0 a 100% (em 5 volumes de coluna) do mesmo tampão de equilíbrio, adicionado de NaCl 1M, o fluxo utilizado para a eluição foi de 3mL/min. A fração de proteínas que não teve interação com esta coluna foi preparada para outros passos de purificação das outras celulases. A fração que teve interação, era composta basicamente por ThCBHI, sem a necessidade de um segundo passo de purificação. Todas as frações foram analisadas quanto ao teor de proteína e sua atividade de avicelase.

3.5 Separação proteolítica do domínio catalítico (CCD) e módulo de ligação à celulose (CBM)

A ThCBHI purificada, foi então digerida com papaína (Sigma) em tampão fosfato 40 mM, pH 6,0, L-cisteína 5 mM, EDTA 2 mM nas proporções de 1:5, 1:25, 1:50 e 1:100, (mg/mL de papaína: mg/mL celobiohidrolase, respectivamente). A reação foi mantida a 25 °C e alíquotas de 0,01 ml foram removidas em intervalos de tempo de 10, 20 e 30 minutos e analisados por SDS-PAGE.

O isolamento do CCD e do CBM foi feito em uma coluna Superdex 75 10/30 (GE) e eluídos com tampão Tris 50 mM, pH 7,0 adicionado de NaCl 200 mM. Os eluatos foram monitorados por absorbância no UV a 280nm e coletados em frações contendo as proteínas purificadas, que posteriormente foram concentradas em concentradores VivaSpin (GE).

3.6 *Análise de perfil proteico por eletroforese SDS-PAGE*

Para a análise do perfil de proteínas totais foi realizado um gel de poliacrilamida a 15% (SDS-PAGE), contendo gel de empacotamento a 3%, de acordo com o sistema descrito por LAEMMLI (1970). As amostras foram adicionadas de 5% de 2-β-mercaptoetanol e aquecidas, em banho-maria, por aproximadamente 5 min., a 95 °C. Os géis foram carregados com 1 mg de proteína referentes à amostra e a eletroforese processada a 100 V, a 25°C. As proteínas foram reveladas através da coloração dos géis com 0,2% de Coomassie Brilliant Blue R-250 em metanol:ácido acético:água (50:10:40 v/v/v) .⁸² Como marcadores de massa molecular foram utilizadas as seguintes proteínas (kit Bio-Rad): soro albumina (66,2 kDa); ovoalbumina (45kDa); anidrase carbônica (31 kDa); inibidor de tripsina (21,5 kDa); lisozima (14,4 kDa); aprotinina (6,5 kDa). As massas moleculares aparentes das proteínas e proteases foram determinadas por comparação com essas proteínas padrões de massas moleculares conhecidas aplicadas ao gel de poliacrilamida.

3.7 *Dosagem da concentração de proteína*

A quantificação de proteínas totais (mg/mL) foi realizada pelo método de Bradford (1976). A ligação do corante azul de Coomassie G250 (molécula negativamente carregada) à proteína gera um produto colorido cuja absorbância a 595 nm utilizando uma curva padrão de albumina de soro bovino (BSA). A concentração final de ThCBHI purificada, foi determinada por absorbância UV a 280 nm utilizando coeficiente de extinção teórico baseado na composição sequencial de aminoácidos.

3.8 *Análise por eletroforese em gel não desnaturante (gel nativo)*

Este gel foi usado para a verificação de possíveis monômeros, dímeros ou multímeros das enzimas em questão. Essa técnica permite a separação das proteínas baseada na diferença de cargas, massa molecular e conformação, enquanto mantém sua integridade conformacional

e sua atividade biológica. A eletroforese foi feita com o sistema *Phast* (Amersham Biosciences) em gel não desnaturante com gradiente de 8-25% de acrilamida. Como marcadores moleculares foram utilizadas as seguintes proteínas de massa molecular conhecidas: tiroglobulina (669 kDa), ferritina (440 kDa), catalase (232 kDa) e aldolase (140 kDa), o gel foi corado com a solução 0,1% PhastGel Blue R-350 (GE) e descorado com 30% de metanol e solução de ácido acético a 10% em água destilada (3:1:6).

3.9 Determinação do ponto isoelétrico

As proteínas são macromoléculas anfóteras, ou seja, possuem cargas líquidas positivas ou negativas dependendo do pH. A rede de cargas nas proteínas é uma soma algébrica de todas as cargas das cadeias laterais dos aminoácidos que as compõe. Cada proteína possui um pH específico na qual a somatória das cargas é nula ou zero e cujo valor é denominado ponto isoelétrico (pI).⁸³ A focalização isoelétrica (IEF, isoelectric focusing) é uma técnica eletroferética de alta resolução que separa proteínas em uma mistura de acordo com seus pontos isoelétricos. Ela é realizada em condições não desnaturantes e têm como princípio a movimentação destas macromoléculas carregadas em um gradiente de pH quando um campo elétrico é aplicado. Nesse gradiente, as proteínas se movimentam até atingirem seus pIs. Em valores de pH acima do pI da proteína, estas são negativamente carregadas e em valores abaixo do seu pI e são positivamente carregadas.

A focalização isoelétrica foi realizada em *PhastSystem* (GE), utilizando IEF PhastGel abrangendo os intervalos de pH de 3-9. A separação levou aproximadamente 30 minutos após o qual o gel foi corado com 0,02% PhastGel Blue R-350 da solução (GE) contendo metanol 30% e 10% de ácido acético em água destilada e 0,1% (w/v) de CuSO₄, fixado em 20% ácido tricloroacético e descorado com uma solução contendo metanol 30% e 10% de ácido acético em água destilada.

3.10 *Estocagem das proteínas*

Após a purificação e diálise da proteína de interesse, essa foi dividida em alíquotas de 20-100 μL , adicionado 5% de glicerol, congeladas rapidamente em nitrogênio líquido e mantidas em freezer a -80°C .

3.11 *Determinação da sequência e identificação da proteína por espectrometria de massas*

A espectrometria de massas é uma técnica analítica utilizada para identificar compostos, elucidar algumas propriedades estruturais e químicas de moléculas, e determinar sequências aminoacídicas de peptídeos que serão comparadas com massas e sequências depositadas em bancos de dados.

Para confirmação da identidade da proteína, a banda da suposta ThCBHI, foi recortada do gel de poliacrilamida e clivada com tripsina (Sigma) por digestão em gel, de acordo com Tarentino e Plummer (1994).⁸⁴ Os peptídeos derivados foram analisados em um espectrômetro de massas do tipo Electrospray Tandem MS (Q-TOF), equipado com interfase nanospray (Micromass), no LNNano, Campinas, Brasil. O programa MASCOT (www.matrixscience.com) foi usado na identificação de proteínas. A lista de massas dos peptídeos foi pesquisada usando a base de dados de proteína NCBItr (base de dados não redundante do National Center for Biotechnology information).

3.12 *Ensaio de Atividade*

3.12.1 *Avicelase*

A atividade de exoglucanase contra o substrato insolúvel foi determinada usando como substrato Avicel 1% dissolvido em tampão citrato de sódio 50 mM, pH 5,0. As amostras e os controles foram incubados a 50°C por 120 minutos, com agitação manual a cada 20 min. A quantidade de açúcares redutores foi determinado pelo reagente de DNS

conforme descrito por Miller (1988) ⁶¹ e a absorbância medida em espectrofotômetro a 540nm. A curva padrão foi construída com glicose para determinar a quantidade (mM) do açúcar redutor liberado. O cálculo de atividade foi feito de acordo com a seguinte equação :

$$Atividade(UI / mL) = \frac{[AçúcaresLiberados (g/L) \times 10^6]}{180} \times \frac{1}{1000} \times \frac{1}{t (\text{ min })} \times \frac{Vt (mL)}{ve (mL)},$$

Onde:

t = tempo de reação, Vt = volume total de meio reacional, Ve = volume de enzima empregado no ensaio.

Todas as atividades foram expressas em unidades internacionais (IU), onde uma unidade de atividade corresponde a 1 μ mol de açúcar redutor (glicose equivalente) liberado por minuto de reação.

3.12.1 Atividade em substrato solúvel, *pNPCase*

Atividade com *p*-NP- β -D-cellobioside (*pNPC*, Sigma, EUA) um substrato sintético, (Figura 13), foi determinada em pH 5,0. Alíquota de 25 mM de solução estoque de substrato preparado em tampão citrato 50 mM (0,05 ml) foi misturada com 0,01 ml de enzima diluída no mesmo tampão. A reação foi aquecida a 50 °C por 30 min e 0,1 ml de uma solução de carbonato de sódio 1 M foi adicionado para paralisar e conferir cor à reação. A medida da absorbância foi feita em $\lambda = 410$ nm. ⁸⁵ A curva padrão foi construída com *p*-nitrofenol para determinar a concentração (mM) do açúcar liberado. Após a reação da enzima com o substrato, em temperatura e pH adequados, ocorre a clivagem do substrato, sendo liberado celobiose e *p*-nitrofenil (*pNP*). Quando adicionado carbonato de sódio na reação, o *pNP* reage conferindo cor ao experimento. Assim, quanto mais *pNPC* clivado, mais forte a cor revelada.

Figura 13 - Molécula de p-NP-β-D-cellobioside após a ação da enzima CBHI, liberando celobiose e pNP, o qual, quando reage com carbonato de sódio confere cor ao experimento.

3.12. 3 β-glicosidase

O extrato enzimático foi incubado com solução de celobiose 2% em tampão citrato de sódio 50 mM pH 5.0. A reação ocorreu a 50°C por 15 min., seguido da adição de GOD, um kit comercial de glicose oxidase para quantificação de glicose. As soluções tiveram suas absorbâncias medidas a 505nm. Cada análise foi realizada em triplicata.⁵⁸

3.12. 4 CMCase

A determinação da atividade de endoglucanase foi realizada incubando-se o extrato enzimático com solução de carboximetilcelulose (CMC) média viscosidade 1% em tampão citrato de sódio 50 mM pH 5,0 a 50 °C por 15 min. A quantificação dos açúcares redutores totais foi obtida pelo método DNS descrito acima no qual as amostras tiveram a absorbância medida a 540 nm. Cada análise foi realizada em triplicata.⁵⁸

3.12.5 Ensaio de atividade de FPase

A atividade de FPase foi realizada incubando-se o extrato enzimático com 1 esfera de papel filtro Whatman nº1 de 7 mm de diâmetro e tampão citrato de sódio 50 mM pH 5,0 em placa de 96 poços. A reação enzima substrato deu-se a 50 °C por 60 min. A quantificação de açúcares redutores totais foi feita pelo método de DNS. Os testes foram realizados em triplicata.⁵⁸

3.12.7 Teste de conversão de celulose em diferentes substratos

Para este experimento foram utilizados 5 substratos celulósicos diferentes:

- celulose microcristalina bacteriana (BMCC) onde o filme ⁸⁶ foi gentilmente cedida pelos pesquisadores Fábio Squina e Hernane Barud, tratada e preparada para os ensaios como descrito por Valjamae e colaboradores,⁸⁷
- papel filtro (FP) (Whatman nº 1),
- Avicel PH 101,
- Sigmacell 20 e
- carboximetilcelulose (CMC) de média viscosidade (Sigma). Todos os substratos foram utilizados na concentração final de 10 g/L.

A celulose microcristalina, é considerada um bom substrato para ensaios de atividade de exoglucanases, porque tem relativamente baixa acessibilidade e grau de polimerização (DP). BMCC tem um reduzido grau de polimerização e maior cristalinidade do que avicel. Esta celulose é altamente cristalina, com uma morfologia simples e tem sido utilizado como um substrato modelo para estudar o mecanismo das enzimas sobre a fibra. O índice de cristalinidade medido para BMCC por difração de raios-X é de 95,2.⁸⁸ Embora, Avicel, o principal substrato utilizado para testes com esta enzima, seja principalmente composta de celulose cristalina, ainda contém uma fração significativa de celulose amorfa e impurezas como manose, xilose, galactose e arabinose perfazendo 5% dos seus polissacarídeos, o que pode dificultar a clivagem do polímero, o mesmo acontece com o Sigmacell. A principal diferença entre Avicel e Sigmacell 20, utilizados neste teste, é a cristalinidade - 91,7, e 84,8%, respectivamente ⁸⁸, sendo que Sigmacell 20 tem uma porção maior de celulose amorfa. O CMC é um derivado da celulose com grupos carboximetil ligados ao seu esqueleto carbônico

de forma randômica. Desta forma, a CBHI pode atacar o substrato pela sua extremidade redutora, hidrolisando os resíduos não substituídos até que o progresso ao longo do polímero fosse bloqueado por resíduos. Tais resíduos substituídos não se acomodam no sítio ativo de exocelulases, tornando o substrato resistente ao seu ataque.^{46; 89}

A concentração da ThCBHI utilizada foi 15 μ M, (tanto para ela intacta, domínio catalítico e de ligação à celulose, quanto para CCD) e a reação foi conduzida em tampão citrato de sódio, 50mM pH5, por 24h, em banho termostático com temperatura constante de 50°C. Para evitar a inibição da reação pela celobiose, foi adicionado 10uM de β -glicosidase (Novozymes SP 188) à reação.⁹⁰ Os pontos foram tirados periodicamente, variando 1 hora até 24 horas. As reações foram interrompidas pela a adição de ácido dinitrossalicílico⁹¹, aquecidos a 95 °C por 5 minutos, resfriados até a temperatura ambiente e as absorbâncias determinadas em espectrofotômetro a 540 nm. A quantidade de glicose produzida pela enzima foi estimada a partir de uma curva de calibração de glicose. Para confirmar que a enzima estava ativa durante todo o experimento à 50°C, paralelamente foi feito acompanhamento de estabilidade térmica da CBHI.

3.13 *Teste de estabilidade térmica da ThCBHI*

Para verificação da estabilidade da proteína a mesma foi incubada a 50°C, em pH 5, e alíquotas foram retiradas a cada hora, por 48 horas, para análise de sua atividade com avicel 1% (p/v) como substrato.

3.14 *Determinação das condições ótimas de atividade quanto ao pH e a temperatura por Delineamento Composto Central Rotacional (DCCR) para ThCBHI*

O Delineamento Composto Central Rotacional (CCRD), contendo cinco níveis para cada variável independente (codificados como -1.414, -1, 0, 1, 1.414) foi empregada para determinar as faixas ideais e avaliar os efeitos de dois parametros principais da atividade ótima da ThCBHI: temperatura e pH. Essas variáveis ($k = 2$) e seus níveis foram predeterminados baseado na literatura sobre CBHI de *Trichoderma reesei*, onde, para a temperatura, a variação ficou entre 21,80 e 78,20 °C e, para o pH, variou entre 2,18 e 7,82. Os valores codificados e reais das variáveis independentes estão apresentados na Tabela 1.

Tabela 1: Variáveis independentes (valores codificados e reais)

Variáveis	Símbolo	Níveis				
		-1.4142	-1	0	+1	+1.4142
Temperatura (°C)	A	21,80	30	50	70	78,20
pH	B	2,18	3	5	7	7,82

Um total de onze experimentos foram realizados, sendo 2^2 experimentos para o delineamento fatorial completo e 3 pontos centrais para avaliar o erro puro. Com a resposta obtida (atividade da ThCBHI para pNPC) foi montado um modelo de segunda ordem polinomial. Depois disso, os dados experimentais foram analisados através do software Statistica 7.0 (StatSoft, Inc.) para obter a significância estatística dos coeficientes de regressão para as variáveis dependentes, variância de ajuste do modelo (ANOVA) e a plotagem de superfície tridimensional resposta (RSM) (Tabela 2).

Tabela 2: Análise da Variância ANOVA

Fonte	Soma dos quadrados	df	Mean Squares	F value	Prob>F
Modelo	17042.70	5	3408.54	161.08	0.000016
Erro	108.38	5	21.16		
Total	17151.08	10			

R^2 = coeficiente de determinação = 0.994

Prob>F menor que 0.05 é significativo

df = graus de liberdade

3.15 Determinação de parâmetros cinéticos K_m e $V_{máx}$ para a enzima ThCBHI

O ensaio de atividade de pNPCase foi descrito no item 3.12.2. A taxa da reação (V_o) foi determinada mantendo-se fixa a concentração de CBHI em 700nM e variando a concentração do substrato pNPC em tampão citrato de sódio 50 mM, pH 5,0 a 50 °C. A atividade de CBHI foi monitorada nas concentrações de 0,78; 1,56; 3,125; 6,25; 12,5; 25; 50 e 100 mM. Os dados obtidos da leitura de absorbância a 410 nm, pelo tempo de 5 minutos, para cada concentração foram representados graficamente e a velocidade de reação (V_o) foi calculada por concentração de produto liberado (mM) dividido pelo tempo de reação. Assim

foi possível calcular o efeito do pNPC e também de 2-chloro-4-nitrophenyl-P-lactoside (CNPL) na velocidade inicial da reação e, a partir disso, determinar a velocidade máxima ($V_{\text{máx}}$) da reação e obter a constante de Michaelis-Menten (K_m) para o substrato. Com os valores de velocidade de reação (V_o) e da concentração de substrato foram calculados os parâmetros cinéticos com o software SigmaPlot 10.0. A partir destes dados foram calculados também K_{cat} e a eficiência catalítica da enzima. Para constante de inibição (K_i), foi testado celobiose como inibidor nas concentrações de 0-100mM, para pNPC como substrato. A obtenção dos valores de K_i para celobiose deu-se através da representação gráfica dos valores de K_m aparentes/ $V_{\text{máx}}$ para pNPC na presença e ausência do inibidor, podendo também demonstrar o tipo de inibição da celobiose com pNPC. Para efeito de comparação dos resultados nas mesmas condições, os testes foram realizados também com a CBHI de *Trichoderma reesei* (TrCBHI), produzida e purificada nas mesmas condições da ThCBHI.

3.16 Dicroísmo Circular

As medidas de CD foram realizadas a temperatura controlada de 25°C utilizando o espectropolarímetro JASCO J-720 (JASCO Corporation, Tóquio, Japão) na faixa de 190 a 260 nm em cubetas cilíndricas de quartzo de caminho óptico de 1 nm. Os espectros de CD foram tipicamente recuperados empregando-se médias de 8 varreduras. As contribuições dos tampões obtidos sob condições idênticas foram subtraídas e todos os espectros de CD foram corrigidos a fim de eliminar qualquer efeito de ruído. Os espectros originais foram filtrados com Transformadas de Fourier, preservando as bandas típicas de cada espectro. Em todos os casos, eles foram obtidos em elipticidade (θ) e, convenientemente transformados em elipticidade molar ($[\theta]$). Estes experimentos de dicroísmo circular foram realizados no Laboratório de Biofísica Molecular (IFSC-USP).

A solução proteica proveniente da purificação, foi utilizada para analisar a integridade das estruturas secundárias nas condições de extração e monitorar as possíveis mudanças conformacionais induzidas por influência do pH através das medidas de CD. A concentração da proteína foi de 1,5 μM em tampões k_2PO_4 /ácido cítrico e Glicina/NaOH 20mM. A proteína foi incubada por 12 horas nas soluções tamponadas a pH variando 2,6 a 10,0.

As contribuições de estrutura secundária pelo espectro de CD da ThCBHI foram calculadas utilizando o pacote de desconvolução *CDPro* software package, que inclui três programas SELCON3, CONTINLL e CDSSTR, providos de um grupo de proteínas de referência contendo 43 espectros de CD.⁹² A partir dos resultados fornecidos pelos três programas, foi escolhido o resultado que mais se aproximava com as estruturas de celobiohidrolases disponíveis na literatura, para encontrar as frações de estrutura secundária dos espectros de CD.

3.17 *Modelo gerado por homologia*

O modelo gerado por homologia da ThCBHI foi construído com base na estrutura primária da enzima pelo programa I-Tasser.^{93; 94} O conteúdo de estrutura secundária do modelo foi analisada por Porter.⁹⁵ Para gerar o modelo utilizado, o programa não baseou-se unicamente na sequência de aminoácidos da proteína alvo, ele considera também informações sobre estruturas de proteínas experimentalmente conhecidas, através do uso de recursos como bibliotecas de fragmentos, preditores de estrutura secundária, preditores de tipos de enovelamento, dentre outros.⁹⁶⁻⁹⁷

3.18 *Determinação de Estabilidade por Dicroísmo Circular e Thermofluor*

- *Dicroísmo:*

Os estudos de desnaturação térmica da ThCBHI foram caracterizados por dicroísmo circular somente no pH 5,0, onde a proteína possui sua maior atividade, na presença e ausência de inibidor celobiose, medindo-se as mudanças de elipticidade a 212 nm induzidas pelo aumento de temperatura de 20 a 80°C com taxa de aquecimento de 1°C por minuto.

- *Thermofluor*:

Com o intuito de monitorar a estabilidade da proteína frente a diferentes valores pH e de temperatura foi realizado experimento de deslocamento térmico. Essa técnica biofísica, baseada no fenômeno de fluorescência, conta com um marcador fluorescente, nesse caso o Sypro® Orange (Invitrogen), sensível às condições do ambiente em que está inserido. O ensaio funciona com o monitoramento das mudanças de sinal emitidas pelo marcador enquanto a proteína vai se desenovelando e variando o ambiente ao qual o marcador está ligado. O marcador se liga às regiões hidrofóbicas da proteína e quando esta se encontra perfeitamente enovelada, sem regiões hidrofóbicas expostas, o marcador fica exposto ao solvente, sem sinal de fluorescência. Com aumento de temperatura ou pH desfavorável, a proteína começa perder estabilidade e desenovelar, com a consequente interação da sonda com as regiões hidrofóbicas e a emissão de sinal fluorescente (Figura 14).

Foi avaliada a estabilidade da ThCBHI em diferentes pHs (2,0; 5,0 e 10,0) e na presença e ausência do ligante. Os experimentos foram realizados em um MiniOpticon Real-Time PCR (Bio-Rad), em triplicata e com comprimento de onda de excitação de 470-505 nm e emissão de 540-700 nm. Os dados foram coletados com uma variação de temperatura de 20°C a 80°C, com um aumento de 0,5°C/min. Foram preparados 200 µL de uma solução-mãe, contendo 1 µL da sonda, em tampão 50 mM de acetato de sódio pH 5,0, de modo que a concentração final na reação fosse de 1,5 mg/mL. O volume final de reação foi de 20 µL (2 µL da amostra + 18 µL da condição testada). Como ligante foi utilizado celobiose (solução 2% , p/v), em tampão de acetato de sódio 50 mM, pH 5,0.

Figura 14 - Princípio Thermofluor. Fluorescência de Sypro Orange é dramaticamente aumentada quando ele se liga às partes hidrofóbicas expostas das proteínas após desnaturação. A estabilidade da proteína (por exemplo, T_m) pode ser estimada por meio da análise da dependência da temperatura da intensidade de fluorescência (Quiagen News, 2010).

3.19 Espectroscopia de Emissão de Fluorescência

As medidas de emissão de fluorescência da ThCBHI foram realizadas utilizando-se o espectrofluorímetro K2 (ISS Inc. Champaign, IL, EUA). Os espectros de fluorescência foram obtidos a temperatura controlada de 25°C. Amostras foram excitadas a 295nm e a emissão da fluorescência foi monitorada no intervalo de 300 a 450 nm. Cubetas de quartzo de 1cm de caminho ótico foram utilizadas nas medidas. A fim de minimizar o efeito do espalhamento de luz, os espectros de emissão de fluorescência dos tampões foram subtraídos dos espectros da amostra. Foram analisadas as possíveis mudanças conformacionais da ThCBHI induzidas por influência do pH nas mesmas condições das medidas de CD.

3.20 *Espalhamento de Luz Dinâmico (DLS)*

A metodologia de espalhamento de luz dinâmico permite a determinação da massa molecular e o raio de giro de macromoléculas em solução a partir das flutuações de luz espalhada pelas amostras. O espalhamento de raios X a baixo ângulo é um processo de espalhamento elástico que ocorre quando o feixe de raios X atravessa a amostra e interage com os elétrons do material. As frequências das flutuações dependem do quão rapidamente as moléculas estão se movendo por difusão, essa técnica permite relacionar esse fato com o tamanho da molécula estudada. Essa técnica foi empregada para avaliarmos o estado de oligomerização em solução da proteína ThCBHI.⁹⁸

Para este ensaio foi utilizado 0,5 mg /mL de ThCBHI a pH 5,0 foi centrifugado a 4,000x g durante 15 min à temperatura ambiente e imediatamente antes da medições colocada em uma cubeta de quartzo (caminho ótico de 1 cm). As medidas de DLS foram realizadas utilizando o instrumento Zetasizer (Malvern Instrument, Worcestershire, Reino Unido) no Laboratório de Biofísica Molecular (IFSC-USP). A temperatura foi elevada de 25-80° C com 5 °C intervalos a cada 120 s para o equilíbrio das temperaturas. O efeito do tampão foi subtraído de cada medida para eliminar o ruído produzido por ele. O raio hidrodinâmico foi calculado usando a Zetasizer Software IV.

3.21 *Estudos de espalhamento de raios-X a baixo ângulo (SAXS)*

Os estudos de espalhamento de raios-X a baixo ângulo são utilizados para análises estruturais de baixa resolução de macromoléculas em solução e nos dão informações sobre o estado de oligomerização da proteína e dados complementares às obtidas pela cristalografia de proteínas. As medidas de SAXS foram realizadas na linha SAXS 2 no Laboratório Nacional de Luz Síncrotron (LNLS), em Campinas-SP, em colaboração com o Dr. Mario de Oliveira Neto do Instituto de Física da USP/São Carlos.

A primeira parte dos estudos de SAXS com a ThCBHI, foram realizados pela Dra. Viviane Serpa, onde foi definido o tamanho da proteína, a massa molecular, e um modelo tridimensional, gerado por modelagem *ab initio*.

Em uma segunda parte foram realizados estudos de estabilidade para observar o efeito de pH e temperatura na estrutura e na compactação da ThCBHI.

Os dados de (SAXS) para a ThCBHI, utilizando comprimento de onda $\lambda=1.48 \text{ \AA}$, detector bidimensional (MarCCD) e distância amostra-detector de 1026.96 mm, cobrindo um intervalo de transferência de momento de $0.015 \text{ \AA}^{-1} < q < 0.33 \text{ \AA}^{-1}$. As curvas de espalhamento da solução da ThCBHI e do tampão foram coletadas em 1 etapa de 300 segundos para monitoramento dos danos causados pela radiação à amostra, além da estabilidade do feixe. Análise sistemática foi realizada subtraindo os dados pelo ruído gerado pelo CCD no mesmo tempo de exposição, em seguida, normalizando-os pela intensidade do feixe incidente e multiplicando-os pela absorção da amostra. O espalhamento da solução tampão foi subtraído da curva de espalhamento da solução de proteína nas diferentes concentrações.

A integração dos padrões bidimensionais de SAXS foi realizada usando software Fit2D⁹⁹, e as curvas foram escalonadas pela concentração de proteína. A concentração de ThCBHI para este experimento de rampa térmica foi 0,5 mg/mL. Amostras de ThCBHI foram incubadas em pH variando de 2 a 10 e os tampões utilizados foram 20 mM de fosfato / ácido cítrico ou Gly / NaOH ajustados para os valores de pH diferentes. As incubações foram por 12h a 25 °C. A desnaturação térmica de ThCBHI em pH 5,0 também foi caracterizado por SAXS numa variação de temperatura de 20 a 70 °C, com dados coletados a cada 5 °C.

3.22 *Predição dos sítios de glicosilação*

Para predição dos sítios de glicosilação foi utilizado o programa CBS Prediction Servers para Glicosilação n-ligada – NetNGlyc (<http://www.cbs.dtu.dk/services/NetNGlyc/>) e Glicosilação o-ligada – NetOGlyc (<http://www.cbs.dtu.dk/services/NetOGlyc/>).

3.2 *Cristalização, coleta de dados e processamento*

Visto que, o número de variáveis que afetam a cristalização, tais como, concentração, temperatura, pH, força iônica, aditivos específicos e precipitantes, é grande e combinatória, o

domínio catalítico purificado de ThCBHI com 46 kDa (aminoácidos 18-449), na concentração de 8 mg/mL foi submetido a uma busca de condições de cristalização utilizando a técnica de difusão de vapor em gota suspensa (*hanging-drop*). Gotas de 2 µL de volume final (relação 1:1 de proteína e de solução precipitante) foram feitas automaticamente com o robô de cristalização Honeybee 931 (Genomic Solutions Inc), usando uma variedade de soluções precipitantes disponíveis comercialmente (nesta etapa foram testadas 650 condições diferentes) e as placas foram mantidas à temperatura de 18 °C. Pequenos cristais em forma de agulhas, que foram posteriormente utilizadas como cristais-sementes, cresceram na solução contendo 35% de Polietilenoglicol (PEG) 4000 como precipitante. Técnicas como cristalização em gota imersa em óleo e *seeding* foram utilizadas para otimização e obtenção de cristais tridimensionais. Para a técnica de cristalização em óleo foram utilizadas placas de 72 poços (72-well MicroBatch plates, Hampton) com uma mistura 1:1 de óleo silicone e parafina para permitir a difusão da água pelo óleo e consequentemente aumentar a concentração da gota de cristalização. *Streakseeding*, método que permite a introdução de cristais-sementes em uma gota de cristalização, foi implementado tanto em gotas suspensas como em gotas sentadas em placas de 24 poços (24-well Linbro plates). Esses cristais-sementes, obtidos anteriormente, foram inseridos na gota pré-preparada utilizando-se um pelo de cavalo.

Um aglomerado de cristais do CCD de ThCBHI na forma de agulhas foi manipulado com uma agulha de acupuntura para separação dos mesmos e uma única agulha cristalina foi transferida para uma solução crioprotetora contendo uma mistura de 35% de PEG 4000 e 20% de PEG 400. O cristal foi então montado em um *loop* e congelado diretamente no fluxo de vapor nitrogênio, antes da análise dos dados de raios-X. O conjunto de dado de difração do cristal foi coletado na linha MX2 do Laboratório Nacional de Luz Síncrotron LNLS (Campinas, Brasil), utilizando-se um detector marCCD¹⁰⁰ e comprimento de onda (λ) de 1,46 nm. O cristal apresentou alto poder de difração com resolução máxima de 1,7Å. Esse conjunto de dados foi coletado aplicando-se o método de oscilação com cobertura de um total de 180° do cristal. A indexação, integração e o escalonamento dos dados foram realizados com o programa HKL2000.¹⁰¹ A estrutura cristalográfica do CCD de ThCBHI foi determinada pelo método de substituição molecular com o programa Phaser¹⁰² utilizando-se as coordenadas atômicas da estrutura cristalográfica do CCD de TrCBHI (código PDB 8CEL) como modelo. O mapa $2Fo - Fc$ calculado da solução única resultante das funções de rotação e translação, aplicadas para a obtenção de fases do modelo, revelou uma densidade eletrônica contínua para a proteína e também foi possível observar densidades eletrônicas no interior do sítio catalítico

para moléculas que não foram incluídas originalmente no modelo de busca. Modelagem e refinamento da estrutura foram realizadas com o programa Coot¹⁰³ e Refmac¹⁰². Um ciclo de refinamento “*simulated-annealing*” também foi realizado com o programa Phenix¹⁰⁴. Moléculas de água foram adicionadas ao modelo com o programa Arp/wArp¹⁰⁵, e a qualidade estereoquímica do modelo final foram validadas com o programa MolProbity¹⁰⁶.

As simulações de dinâmica molecular foram realizadas pelo grupo do Prof. Dr. Munir Skaf do Instituto de Química da Unicamp. As coordenadas iniciais para simulações de MD do CCD de *T. harzianum* e *T. reesei* foram tomadas a partir da estrutura cristalográfica obtida e do Protein Data Bank (PDB ID 8CEL).

3.24 Microscopia Eletrônica de Varredura

A microscopia eletrônica de varredura (MEV ou SEM) é uma técnica bem estabelecida para a obtenção de imagens superficiais de sólidos orgânicos e inorgânicos utilizada em diversas áreas, sobretudo em biologia, geologia e ciência dos materiais. As imagens obtidas são formadas a partir da interação de um feixe de elétrons focalizado sobre a amostra que varre a sua superfície ponto a ponto. A morfologia do papel filtro, antes e depois de ser submetido ao ataque da enzima ThCBHI, foi analisada utilizando um microscópio eletrônico de varredura, equipado com canhão de emissão de campo (FEI Quanta 650). Este experimento foi realizado no LNNano (Laboratório Nacional de Nanotecnologia), Campinas, SP. A hidrólise ocorreu durante 24h a 50 °C, colocando-se um disco de 0,7 cm de papel de filtro Wathman nº 1 em cada tubo de eppendorf.com 1mg/mL de ThCBHI em tampão citrato (pH 5,0).

As diferentes amostras de papel hidrolisado foram secas em estufa a 60 °C e fixadas em suportes de amostras de MEV, usando um pequeno pedaço de fita de carbono. Em seguida, foram revestidas com ouro em um metalizador SCD Balzers 050 por pulverização catódica (corrente 40 mA, tempo de 60 s). Um grande número de imagens foi obtido em diferentes amostras para garantir a reprodutibilidade dos resultados.

4 RESULTADOS E DISCUSSÃO

4.1 Determinação do tempo de cultivo do *T. harzianum*

Para a determinação do tempo de cultivo do *T. harzianum* em fermentador, foram retiradas amostras a cada 24h e realizado, além da quantificação das proteínas secretadas, também os testes de atividades CMCase; β -glicosidase e FPase. A Figura 15, a seguir mostra o comportamento do cultivo deste fungo durante sete dias.

Figura 15 – Perfil do processo de produção de enzimas do fungo *Trichoderma harzianum*. CMCase em vermelho; β -glicosidase em verde; FPase em azul; e em roxo a dosagem da concentração de proteínas.

Observando o decaimento da expressão das proteínas de interesse, com base na cinética de atividade enzimática, com diferentes substratos, o cultivo foi interrompido entre o

quarto e quinto dia de fermentação (Figura 15). A concentração das proteínas totais continuou aumentando no quinto dia, mas as atividades, para todos os substratos testados, estavam diminuindo. Isso pode ocorrer, pois o fungo pode começar a expressar proteínas proteolíticas, prejudicando assim o rendimento das celulasas de interesse.

O extrato total bruto foi preparado para iniciar a separação das proteínas de interesse, eliminando, como primeiro passo o avicel, que não foi consumido, e o micélio.

4.2 Purificação e identificação de ThCHBI

Após cinco dias de crescimento do *T. harzianum* em fermentador, iniciou-se o processo de purificação das celulasas. A concentração de proteínas e testes de atividade de celulasas foram medidos em cada passo da purificação (Tabela 3). Um único passo de cromatografia DEAE-Sepharose foi suficiente para eluir a ThCBHI na presença de NaCl 100mM, com uma fração com maior atividade avicelase (Figura 16) .

Figura 16 – Cromatograma do extrato de proteínas totais de *Trichoderma harzianum* em coluna DEAE-Sepharose, equilibrada com tampão Tris 50mM, pH 7,0, e eluída com NaCl 1M. A fração contendo a CBHI está representada no pico 4.

O rendimento desta purificação pode ser considerado bom, pois obtivemos aproximadamente 33% de ThCBHI purificada do total de proteínas do extrato bruto, como pode ser observado na Tabela 3.

Tabela 3 - Passos da purificação da CBHI de *T. harzianum* em coluna DEAE Sepharose A50, atividades para Avicel como substrato.

Passos da purificação	Proteína (mg)	Atividade Avicelase (U/mg)	“Purification fold”	Rendimento (%)
Extrato bruto concentrado	380 mg	0.082±0,007	1,0	100
DEAE-Sephadex A50	121,2 mg	0.91±0,09	11,41	31,9

A fim de se avaliar a pureza da enzima estudada, o perfil das bandas proteicas desta enzima foram obtidas através de SDS-PAGE, que mostra a fração purificada continha uma banda de massa molecular de aproximadamente 66 kDa, Figura 17. A massa molecular teórica de ThCBHI, calculado pelo programa ProtParam, foi de 53 kDa, porém como já esperávamos, no gel aparece um valor experimental maior devido à glicosilação extensa como já relatado.⁶⁸ Para comparação, o peso teórico molecular de *T. reesei* CBHI (TrCBHI), calculado com base em sua sequência de aminoácidos é de 52 kDa, enquanto que o valor observado experimentalmente é próximo a 66 kDa.¹⁰⁷

Depois de purificada, ThCBHI foi submetida à clivagem parcial pela protease papaína, para separação do domínio catalítico (CCD) e domínio de ligação à celulose (CBM), com diferentes proporções de papaína e tempos de reação. O melhor resultado foi obtido na proporção 1:50 (papaína:ThCBHI) e tempo de 20 minutos, liberando dois grupos de massa molecular de cerca de 50 e 16 kDa. Os fragmentos digeridos foram separados por gel filtração em coluna Superdex 75 10/30. O domínio catalítico pode ser observado no gel SDS-PAGE da Figura 17. A ThCBHI intacta migrou no gel como uma proteína de 66 KDa, enquanto que a ThCBHI digerida com papaína mostrou ter maior mobilidade nas mesmas condições, sendo uma banda na altura de 50kDa.

Figura 17 - SDS-PAGE da purificação da ThCBHI. Coluna 1 Marcador de Massa Molecular Coluna 2, extrato bruto obtido após 5 dias de crescimento em fermentador na presença de Avicel como indutor. Coluna 3, ThCBHI eluída com 100mM de NaCl em tampão Tris-HCl 50 mM, pH 7,0., Coluna 4, CCD de ThCBHI após a digestão da enzima com papaína (Sigma).

A combinação do domínio catalítico e o domínio de ligação à celulose da ThCBHI mostrou um comportamento sinérgico, revelando que a taxa de hidrólise da proteína clivada quando comparada com a ThCBHI intacta teve quase 50% da sua atividade diminuída para o substrato Avicel (0,89 e 0,47 U/mg, respectivamente), que é um substrato insolúvel. Resultado parecido obteve-se para TrCBHI.¹⁰⁸ Estes estudos confirmam a visão atual de que uma organização da estrutura tridimensional de uma enzima intacta é necessária para o seu pleno funcionamento catalítico.

4.3 Espectrometria de Massa

Para identificação da enzima purificada, a mesma foi submetida a análise por espectrometria de massa. Os resultados da pesquisa mostraram que os perfis peptídicos foram consistentes com a sequência de aminoácidos de ThCBHI depositada, demonstrando que a proteína purificada de 66 kDa foi efetivamente a ThCBHI. A Figura 18 apresenta a sequência depositada para CBHI de *T. harzianum* (gi|50400675) em banco de dados do NCBI¹⁰⁹ com os peptídeos identificados por espectrometria de massa em destaque. A cobertura de peptídeos foi de 14%, um número relativamente pequeno, mas isso pode ser explicado pela dificuldade para a tripsina digerir essa proteína, pois ela possui 10 pontes dissulfeto e é altamente

glicosilada, e estas características podem interferir negativamente no momento da clivagem.

110

1	MYRKLAVISA FLAARAQOV CTQQAETHPP LTWQKCTASG CTPQQGSVVL
51	DANWRWTHDT KSTTNCYDGN TWSSTLCRDD ATCAKNCCLD GANYSGTYGV
101	TTSGDALTLQ FVTASNVSRR LYLMANDSTY QEFTLSGNEF SFDVDVSQLP
151	CGLNGALYFV SMDADGGQSK <u>YPGNAAGAKY</u> <u>GTGYCDSQCP</u> RDLKFINGQA
201	NVEGWEPSSN NANTGVGGHG SCCSEMDIWE ANSISEALTP HPCETVQTM
251	CSGDSCGGTY SDR <u>YGGTCD</u> <u>PDGCDWNPYR</u> LGNTSFYGGP SSFALDTTKK
301	LTVVTQFATD GSISR <u>YYVQN</u> <u>GVEFQQPNAQ</u> VGSYSNTIN TDYCAAEQTA
351	FGGTSFTDKG GLAQINKAFQ GGMVLVMSLW DDYAVNMLWL DSTYPTNATA
401	STPGAKRGSC <u>STSSGVPAQV</u> <u>EAQSPNSKVI</u> <u>YSNIRFGPIG</u> STGGNTGSNP
451	PGTSTTRAPP SSTGSSPTAT QTHYGQCGGT GWTGPTRCAS GYTCQVLNPF
501	YSQCL

Figura 18 - Sequência de ThCBHI depositada no banco de dados do NCBI (gi | 50400675). Em negrito e sublinhado, os peptídeos derivados da digestão com tripsina,. A mistura de peptídeos resultantes foram analisados por espectrometria de massa e confirmaram a identidade da celobiohidrolase de *T. harzianum*.

4.4 Determinação do ponto isoelétrico (pI) da ThCBHI

O pI das proteínas pode ser determinado experimentalmente, e é definido como o pH no qual a molécula não migra quando submetida ao campo elétrico. Cada proteína possui seu pI característico, que reflete a proporção entre aminoácidos ácidos e básicos presentes na sua estrutura.

Para melhor caracterizar esta enzima, medimos o seu ponto isoelétrico. O pI de ThCBHI determinado experimentalmente, foi calculado usando a equação de regressão linear ($pI = -0.1508x + 9,4565$) obtida após a separação da amostra sobre o IEF PhastGel (Figura 19A). O pI determinado experimentalmente de ThCBHI foi de aproximadamente 5,23 (Figura 19B), revelando que essa é uma enzima ácida. Este valor de pI é ligeiramente superior ao valor previsto teoricamente para a ThCBHI (gi|50400675),¹¹¹ usando ferramenta ProtParam que gerou um $pI = 4,8$. A diferença pode ser atribuída à superfície de distribuição de carga influenciado pela glicosilação da proteína. O $pI < 5,5$ é comum para as principais celulasas de *T. reesei*, fornecendo propriedades ácidas para estas glicoproteínas, e esta característica auxilia na purificação utilizando cromatografia de troca iônica.¹¹²⁻¹¹³

Figura 19 - Focalização isoeétrica de ThCBHI. (A) Curva padrão de pI construída a partir do Kit Pharmacia de calibração PI: Amilglicosidase, pI 3,50; inibidor de tripsina, pI 4,55; A β -lactoglobulina, pI 5,20; anidrase bovina carbônica B, pI 5,85; anidrase carbônica humana B, pI 6,55; mioglobina, faixa ácida, pI 6,85; mioglobina, banda de base, pI 7,35; lectina de lentilha, ácidos, pI 8,15; lectina de lentilha, médio, pI 8,45; lectina de lentilha, básico, pI 8,65. (B) IEF PhastGel 3-9. Linha 1, Th CBHI purificada. Linha 2, Padrões Pharmacia PI (GE).

4.5 Análise por eletroforese em gel não desnaturante (gel nativo)

A proteína CBHI inteira e seu domínio catalítico isolado foram caracterizados quanto ao seu estado nativo pela aplicação de um gel não desnaturante de 8-25% de acrilamida (Figura 20). No gel de eletroforese em condições nativas podemos verificar que a proteína é monomérica e apresenta uma única banda para a concentração medida.

Este resultado reflete o fato de termos CBHI em fração pura e monodispersa em solução.

Figura 20 - Gel Nativo da ThCBHI intacta e CCD. Linha M, os padrões de massa tiroglobulina (669 kDa), ferritina (440 kDa), catalase (232 kDa) e aldolase (140 kDa); Coluna, ThCBHI purificada e clivada com de papaína. Coluna 2: ThCBHI.

4.6 Perfis de temperatura e pH para atividade enzimática

A atividade de uma enzima é dependente do pH e da temperatura. E esses são fatores que causam mudanças na enzima, no seu enovelamento, conformação, e protonação. Como um primeiro passo do nosso estudo, a atividade exoglucanase CBHI foi estudada em função do pH. Estes resultados mostraram que a 50 °C, a atividade exoglucanase aumenta à medida que o pH aumenta dentro da faixa de pH de 3,0 a 5,0 e, em seguida, diminui dentro da faixa de pH de 5,0 a 7,0 (Figura 21). A enzima atinge a sua atividade ótima em pH 5,0 e 50 °C, condições semelhantes às encontradas para outras celobiohidrolases.^{114; 115}

Em seguida, a atividade exoglucanase foi estudada como uma função da temperatura, em pH 5,0. Dentro da temperatura na faixa de 20 a 50 °C, a atividade aumentou com a elevação da temperatura atingindo o seu máximo em T = 50 °C e diminuindo de forma constante até zero em temperaturas acima de 50°C.

Figura 21 - Gráfico de superfície de resposta de temperatura e pH em função da atividade da enzima CBHI com substrato pNPC.

4.6 Estabilidade térmica da ThCBHI mantida por 48h à 50°C

Para confirmar que a enzima estava ativa durante os experimentos de longa duração à 50°C, foi feito acompanhamento de estabilidade térmica da ThCBHI. Observamos que ela manteve atividade inalterada até 26 h (Figura 22), quando mantidas à 50°C em pH 5,0..

Figura 22 - Estabilidade térmica celobiohidrolase mantida por 48h à 50°C, medida através do acompanhamento da atividade da ThCBHI com 1% de avicel como substrato.

4.7 Constantes cinéticas

Assumindo cinética de Michaelis-Menten da ThCBHI intacta, o K_m determinado experimentalmente e os valores de k_{cat} são, respectivamente, 3,4 mM e 23,2 mM min⁻¹ para pNPC e 3,7 e 44,6 min⁻¹ para cNPL (Tabela 4). Em comparação com os parâmetros cinéticos relatados para TrCBHI, que varia de 0,4 até 1,23 mM^{116, 52, 2; 117, 118}, os nossos resultados mostram que ThCBHI exibe afinidade significativamente mais baixa para estes substratos, enquanto que o k_{cat} é aproximadamente 3,5 vezes maior do que a de TrCBHI sobre o substrato lactoside (~ 12 min⁻¹).^{116, 52} Além disso, os ensaios de inibição mostram que ThCBHI é competitivamente inibida por celobiose assim como já descrita para TrCBHI¹¹⁹, embora com uma constante inibição (k_i) muito mais elevada de $7,2 \pm 1,2$ mM em comparação com o k_i de 0,02 mM para TrCBHI¹¹⁹. Isso pode ser importante para potenciais aplicações biotecnológicas desta enzima.

Tabela 4 - Os parâmetros cinéticos de ThCBHI. As medições de atividade com a enzima na concentração de 700nm em 50 mM de tampão citrato de sódio pH 5,0, a 50 °C.

	pNPC			cNPL			Cellobios
	K_m	V_{max}	K_{cat}	K_m	V_{max}	K_{cat}	K_i
	(mM)	(mM.min ⁻¹)	(min ⁻¹)	(mM)	(mM.min ⁻¹)	(min ⁻¹)	(mM)
ThCBH	3.4±0.15	0.016±0.001	23.2 ± 1,7	3.7±1,2	0.031±0.004	44.6 ± 4.7	7.2

4.8 Teste de Conversao da Celulose em diferentes substratos

Diversos substratos foram utilizados para monitorar a conversão da celulose a açúcar redutor, pela ThCBHI.

Atividades específicas e ensaios de hidrólise do substrato da ThCBHI intacta e seu CCD são mostrados na Figura 23. Estes resultados demonstraram que a enzima está ativa e é eficiente na degradação da celulose microcristalina bacteriana (BMCC), atingindo quase 90% de conversão do substrato em 24 horas sob as condições do experimento. Como esperado, a CCD é significativamente menos eficiente na hidrólise de celulose cristalina (Avicel PH101, Sigmacell 20 e BMCC) em comparação com a enzima intacta, contendo os dois domínios (CCD e CBM) (Figura 23 A e B). Além disso, ThCBHI é capaz clivar carboximetilcelulose (CMC) e, neste caso, a eficiência enzimática é baixa e não há diferença significativa na atividade específica entre CCD e da enzima intacta (Figura 23 A e B). Também, nem enzima intacta ou o CCD foram capazes de hidrolisar mais de 10% do substrato, mesmo após 24 horas de reação. A conversão de CMC rapidamente atingiu um platô e permaneceu assim até o final do experimento (Figura 23 C e D). Ambas as atividades específicas e os rendimentos de conversão da ThCBHI para Avicel 101 e Sigmacell 20 estão nos níveis intermédios entre os resultados observados para BMCC e CMC, mas há uma diferença considerável da atividade e os rendimentos de hidrólise entre a enzima intacta e seu CCD. Curiosamente, as atividades específicas de ThCBHI são mais elevadas do que a de TrCBHI. De acordo com os dados relatados, as atividades específicas de TrCBHI contra Avicel ou Sigmacell estão entre de 0,014 - 0,3 U / mg, e 0,04 - até 0,07 U / mg para CMC sob condições experimentais semelhantes.^{61; 120} Estes resultados mostram que a exoglucanase CBHI de *T. harzianum* não

só é capaz de hidrolisar substratos celulósicos, mas também é muito eficiente na despolimerização de BMCC, Avicel e Sigmacell.

Figura 23 - Atividades da ThCBHI intacta e seu CCD em diferentes substratos celulósicos. Os ensaios foram realizados usando 50 mM de citrato de sódio tampão pH 5,0, a 50 ° C. (A) Atividades específicas de ThCBHI. (B) % de conversão de substrato celulósico à glicose em 24 horas. (C) Conversão de substrato por ThCBHI e seu CCD (D) em 24 horas de reação.

Estes resultados mostram que a atividade celobiohidrolase de *T. harzianum* é ativa contra substratos celulósicos com diferentes graus de cristalinidade e de polimerização. O perfil de hidrólise específica obtido de ThCBHI em geral se correlaciona bem com os dados publicados para TrCBHI.¹²⁰ ThCBHI se mostrou bastante eficiente na hidrólise de BMCC, que é um substrato altamente cristalino e solúvel com mais de 80% na conversão desse

substrato a açúcar. Para Avicel e Sigmacell, ambos cristalinos, mas insolúveis essa conversão foi inferior (entre 50-60%), e para CMC a ThCBHI se mostrou pouco eficaz na hidrólise para 24h. Rendimentos menores na conversão de Avicel e Sigmacell podem estar relacionados ao menor acesso destes substratos e presença da hemicelulose (~ 5%) em sua composição.^{89; 121} Xilose e xilooligossacarídeos foram descritos como sendo fortes inibidores da hidrólise da celulose por celulasas e pode ter um efeito prejudicial sobre a conversão dos últimos substratos.¹²² Para todos os três substratos cristalinos, a presença de CBM foi benéfica e conduziu a uma maior atividade específica e melhores rendimentos na conversão. ThCBHI, como uma enzima prolessiva, é ineficaz na hidrólise de CMC, o que também foi visto para TrCBHI, que é conhecido por ter atividade CMCase baixa.¹²³ De acordo com a nossa análise cinética, as taxas catalíticas de ThCBHI, são consistentemente mais elevadas do que as atividades mencionadas para TrCBHI^{61, 120} utilizando tanto substratos glicosídeo/lactosídeo. A enzima também mostra a inibição pelo produto final da reação, a celobiose. Embora, os mecanismos moleculares que poderiam explicar a baixa inibição de ThCBHI por celobiose ainda não foram completamente elucidados, esta característica da enzima pode ser importante em ambientes industriais.

4.10 Estudos biofísicos e estruturais da ThCBHI

O estudo de aglomerados, mudanças conformacionais ou enovelamento de macromoléculas biológicas são parte de um amplo campo de pesquisas em biofísica. Uma melhor compreensão dos processos de mudança conformacional da CBHI é importante para estudar a relação entre a estrutura da enzima, estabilidade e sua atividade. Assim, para correlacionar os dados de atividade enzimática com a estrutura da enzima, foram conduzidos uma série de estudos detalhados sobre o enovelamento e a estabilidade da CBHI em função da temperatura e pH, empregando uma série de estudos biofísicos e estruturais.

4.10.1 Influência do pH sobre a estabilidade estrutural da ThCBHI

A influência do pH sobre a estabilidade estrutural da ThCBHI foi estudada usando espectroscopia de dicroísmo circular (CD), espectroscopia de fluorescência intrínseca, e espalhamento de raios X a baixo ângulo (SAXS).

O CD é uma poderosa ferramenta para investigar a conformação das proteínas e, portanto foi usado para analisar a estrutura secundária da ThCBHI e avaliar seu enovelamento. A forma geral do espectro da proteína em pH 5,0 (pH ótimo) e a 25 °C (Figura 24), com máximo positivo próximo de 195 ± 1 nm e máximo negativo na faixa entre de 205-215 ± 2 nm, esses dados são característicos da classe das β -proteínas, comparável com a estrutura secundária de outras celobiohidrolases.^{71; 124; 125}

Além disso, a ausência de uma banda negativa em torno de 222 ± 1 nm é indicativo de uma proteína com baixo conteúdo de α -hélices na estrutura secundária. A Figura 24 mostra a influência, do pH sobre a estrutura secundária da ThCBHI a 25 °C, em diferentes valores de pH (2,0; 5,0; 6,0 e 10,0). Estes resultados mostram que não existem alterações estruturais significativas da ThCBHI dentro do intervalo de pH 2,0-8,0. No entanto, quando o valor de pH foi maior que 8,0, o espectro foi progressivamente alterado com uma perda do formato característico, indicando um forte efeito do pH sobre a estrutura secundária.

Figura 24 - Efeito do pH na estrutura secundária da ThCBHI em 25°C monitorado por espectroscopia de Dicroísmo Circular. As medidas foram feitas varrendo o comprimento de onda de 195 a 250nm.

Em seguida, a estrutura terciária da ThCBHI foi investigada por espectroscopia de fluorescência intrínseca e SAXS. O espectro de emissão de resíduos triptofanos expostos à água possui um máximo deslocado ao redor de 345 - 350 nm, enquanto que os resíduos de triptofanos enterrados em regiões hidrofóbicas das proteínas têm um máximo de emissão deslocados para 310 - 320 nm,¹²⁶ sabendo disso, observamos o comportamento da ThCBHI em diferentes pHs.

A Figura 25 mostra os espectros de emissão de fluorescência de ThCBHI após excitação a 295 nm nos diferentes valores de pH. O espectro da amostra em pH 5,0 (linha vermelha, pH ótimo) é caracterizado por uma emissão máxima a 347 ± 1 nm típico de resíduos de triptofano parcialmente exposta ao solvente, indicando que a maioria dos resíduos de triptofano presentes na sequência ThCBHI estão parcialmente expostos ao solvente. Resultados semelhantes foram observados quando a proteína foi incubada em pH 2,0 e 6,0, indicando que não há mudanças conformacionais apreciáveis induzidas na proteína e nas posições dos resíduos de triptofano. No entanto, o espectro em pH 10,0 é caracterizado por um máximo emissão a 354 ± 1 nm e um desvio do máximo de emissão a direita, indicando que os resíduos de triptofano estão mais expostos ao solvente, permitindo-nos considerar que a proteína está perdendo a estrutura terciária.

Figura 25 - Efeito do pH na estrutura terciária da ThCBHI a 25 °C, monitorada por espectroscopia de fluorescência intrínseca em diferentes valores de pH. O comprimento de onda de excitação utilizado

foi de 295 nm e a emissão medida de 300 – 400 nm. A concentração de proteína empregada foi de 1,5 μ M nos tampões K_2PO_4 /ácido cítrico e Glicina/NaOH 20m, de pH 2 a 10. No inserto do gráfico, a figura está ilustrado a variação do comprimento de onda máximo de emissão para cada pH.

Para se certificar que esses dados de fluorescência estavam coerentes, foi medida a área da acessibilidade ao solvente de cada resíduo de triptofano presentes na molécula. Para isso, o programa Areaimol (CCP4-Packages\ccp4-6.1.3\html\areaimol.html) (Collaborative Computational Project, Number 4 (CCP4), 1994) e o modelo I-Tasser derivado da estrutura primária de ThCBHI, foram usados.

O modelo gerado mostra nove resíduos de triptofano localizado no domínio catalítico nas posições W33, W54, W56, W72, W205, W229, W276, W380, W389 e um resíduo de triptofano no CBM da ThCBHI W482 (Figura 26). O cálculo da área de acesso ao solvente demonstrou que cinco resíduos de triptofano W54 (50,20 A^2), W56 (119,00 A^2), W380 (50,10 A^2), W389 (44,60 A^2) e W482 (82,10 A^2) estão expostos ao solvente, considerando que outros cinco, W33 (5,90 A^2), W72 (14,40 A^2), W205 (29,90 A^2), W229 (1,80 A^2), W276 (0,00 A^2) estão total ou parcialmente enterrado no interior do núcleo da proteína, onde o solvente é inacessível. Portanto, a acessibilidade do triptofano é consistente com o máximo de emissão de 347 nm detectado pela fluorescência intrínseca da ThCBHI nativa.

Figura 26 - Modelos gerados no programa I-Tasser para o domínio catalítico e módulo de ligação de celulose (CBM) mostrando a distribuição de triptofanos em uma molécula Th CBHI

4.10.2 Análises da ThCBHI por Espalhamento de Raios X a Baixo Ângulo (SAXS)

O mais importante aspecto da técnica de SAXS é que ela permite o estudo de proteínas em seu estado nativo em soluções quase fisiológicas, e analisar mudanças estruturais em resposta a variações das condições externas. Desse modo, ela fornece informações valiosas sobre a relação entre essas mudanças estruturais e o funcionamento desses sistemas. SAXS também é aplicável para a análise de formas globais, dimensões e estados oligoméricos de macromoléculas. A Figura 27A mostra as curvas de espalhamento de raios-X da ThCBHI para os diferentes valores de pH testados (2,0; 5,0; 8,0 e 10,0). Neste caso, os dados de espalhamento mostram claramente que não houve alterações estruturais da molécula a pHs abaixo de 8,0. No entanto, o espalhamento foi visivelmente alterado quando o valor de pH foi aumentado acima de 8, indicando que existem alterações estruturais na também estrutura terciária da ThCBHI em valores de pH alcalinos.

Para uma análise mais detalhada dos efeitos do pH sobre a compactação global da ThCBHI, os dados de SAXS foram analisados em termos dos gráficos de Kratky. Para um sistema com uma forma compacta, este gráfico mostra uma curva bem definida com uma porção inicial aumentada seguido por uma descida acentuada formando assim um pico máximo definido ^{127; 128; 129}. No entanto, a curva esperada para um polímero com forma estendida ou conformação helicoidal aleatória não possui máximo bem definido ¹³⁰. Como pode ser visto na Figura 27B todas as curvas correspondentes a medições feitas em pH abaixo de 8 mostram os máximos bem definidos, como esperado para estruturas compactas. No entanto, no caso dos valores de pH acima de 8,0 o gráfico de Kratky indicou uma alteração na intensidade do pico e a sua posição mostrando que as alterações para a compactidade da estrutura pode ter ocorrido. Este comportamento é tipicamente exibido por cadeias alongadas ¹³¹, e que em pHs alcalinos a ThCBHI está desenovelada, em contraste com o que foi observado para valores de pH abaixo de 8,0. Tomados em conjunto, resultados de CD, fluorescência intrínseca e SAXS nos permite inferir que não houve alterações significativas na estrutura secundária e terciária da ThCBHI a valores de pH abaixo de 8,0, indicando que a proteína é resistente a pHs ácidos.

Alterações, tanto em níveis estruturais secundários, como em terciários foram observadas para valores de pH alcalinos. Como no intervalo de pH 2,0 – 8,0 não houve apreciáveis alterações conformacionais na estrutura da proteína, a variação da atividade observada foi provavelmente devido à variação na ionização das cadeias laterais dos aminoácidos no túnel hidrofóbico de ligação a celulose, o qual tinha a sua atividade ótima em pH = 5. No entanto, a diminuição da atividade observada para os valores de pH 8,0 – 10,0 é mais provavelmente mediada pelas alterações na estrutura secundária e terciária da enzima, com a formação de um glóbulo fundido –causando o seu desenovelamento, provocada pela variação pH.

Figura 27 - Curva de espalhamento experimental em função do pH de ThCBHI. (A) Curva de SAXS mostra como algoritmo da intensidade de espalhamento ($\log I$) versus vetor transferência de momento, q . (B) Gráfico de Kratky da ThCBHI para os pHs 2,0 – 10,0.

4.10.3 Influência da temperatura sobre a estabilidade e compacidade da ThCBHI

A natureza da transição térmica no enovelamento da ThCBHI foi estudada em pH 5 por espectroscopia de CD, Thermofluor, e de SAXS variando a temperatura de 10-90 °C. O

espectro de CD da ThCBHI a 25 ° C é típica de estrutura secundária de proteínas contendo elementos de folhas- β . O espectro permanece relativamente constante, abaixo 55 °C, e isto pode ser facilmente seguido acompanhando a elipticidade em 212 nm (Figura 27 A). O perfil do espectro foi progressivamente alterado, no entanto, quando a temperatura ficou acima de 55 °C, com uma perda da estrutura secundária. Além disso, a enzima ficou completamente desenovelada em temperaturas acima 80 °C.

A alteração estrutural observada pelo método de Thermofluor foi acompanhada por um aumento no nível de ligação do Sypro Orange pela proteína (Figura. 27B). Sypro Orange é uma sonda fluorescente extrínseca, possui baixo rendimento quântico em solução aquosa, mas é altamente fluorescente em ambientes não polares, tais como sítios hidrofóbicos em proteínas¹³²⁻¹³³. Quando uma proteína se desenovela ou começa perder estrutura, a sonda liga-se a regiões hidrofóbicas expostas da proteína, resultando em um aumento significativo na intensidade de emissão da fluorescência. A intensidade de fluorescência atinge um máximo e, em seguida, começa a diminuir, provavelmente devido à precipitação do complexo da sonda fluorescente e proteína desestruturada. Os resultados obtidos para a ThCBHI são mostrados na Figura 27 B, onde é indicada a mudança conformacional que envolve a exposição dos referidos resíduos, mostrando que o Sypro Orange mostra baixa afinidade para a proteína a temperaturas abaixo de 55 °C, mostrando a presença de poucas regiões hidrofóbicas expostas na estrutura nativa. Defendemos que esta baixa afinidade para a ThCBHI nativa, em temperaturas abaixo de 55 °C é provavelmente devido à interação entre o corante e o túnel hidrofóbico de ligação à celulose. Quando a temperatura foi aumentada para valores acima de 55 °C, o aumento na ligação de Sypro Orange e proteína foi observada, o que implica a exposição de regiões hidrofóbicas. A razão para a diminuição na atividade da ThCBHI em temperaturas entre 55 e 65 °C, é provavelmente, a formação de um estado como uma estrutura mais compactada. Enquanto entre 55 e 65 °C atividade da enzima diminuiu, pois a estrutura expõe superfícies hidrofóbicas para o ambiente externo e tem uma tendência para formar agregados. No entanto, quando a temperatura era valores acima de 65 °C, o sinal da fluorescência mostrou como resultado uma redução na afinidade do corante com a proteína, presumivelmente devido à oclusão da superfície hidrofóbica e agregação e/ou desnaturação proteica.

Figura 28 - Desenovelamento térmico da ThCBHI. (A) Desnaturação térmica de ThCBHI em pH 5,0, na ausência (círculos) ou presença (triângulos) de celobiose, foi monitorada por medição da elipticidade a 212 nm em função da temperatura. (B) Emissão da intensidade de fluorescência de Sypro Orange, em função da temperatura, da ThCBHI em pH 5,0, na ausência (triângulos) e na presença de celobiose. Os comprimentos de onda de excitação e de emissão foram 490 e 575 nm, respectivamente.

4.10.3 Influência da celobiose na estabilidade térmica da ThCBHI

Durante a ação da celobiohidrolase sobre a celulose, a cadeia do substrato polissacarídico desliza através do túnel de ligação, e a cada segunda ligação glicosídica apresentada corretamente aos resíduos catalíticos, liberam progressivamente, unidades dissacarídicas (celobiose), a partir da cadeia de celulose.⁸⁵ Após clivagem da ligação glicosídica β 1-4, a ThCBHI pode permanecer ligado ao produto da reação, celobiose, que neste caso funciona como um inibidor competitivo. (Assim, realizamos estudos térmicos monitorando o desenovelamento por espectroscopia de CD far-UV e thermofluor), na ausência e presença de celobiose, a fim de investigar a influência deste inibidor sobre a estabilidade térmica de ThCBHI. O valor de temperatura de fusão (T_m) para a ThCBHI, é igual a $61,5 \pm 0,2$ °C (Figura 27 A). Por outro lado, quando a ThCBHI foi incubada com 1 mM de celobiose a estabilidade térmica do complexo aumenta, acompanhada por uma mudança no T_m para $65,9 \pm 0,2$ °C. Além disso, a intensidade de fluorescência do Sypro Orange foi monitorada em função da temperatura na ausência e na presença de celobiose. Na figura 27 B, a sonda mostra uma baixa afinidade para a ThCBHI a temperaturas abaixo de 55 °C, indicando a presença de regiões hidrofóbicas, provavelmente devido à interação entre a sonda e o túnel hidrofóbico de ligação à celulose. No entanto, a adição celobiose resultou numa diminuição da intensidade de emissão Sypro Orange, em temperaturas inferiores a 50 °C, provavelmente agora, devido à interação entre a celobiose e a celulose no túnel de ligação. Por outro lado, a intensidade da fluorescência aumentou apenas para valores de temperatura acima de 60 °C, indicando um aumento na estabilidade térmica da proteína na presença de celobiose. Quando a temperatura era aumentada para valores acima de 60 °C, o nível de ligação da ThCBHI ao Sypro Orange foi aumentada, indicativo da exposição das regiões hidrofóbicas. No entanto, quando a temperatura foi superior a 73 °C, uma diminuição no sinal de fluorescência foi detectada. Tomados em conjunto, os resultados de CD e thermofluor, confirmam que a celobiose aumentou significativamente a estabilidade térmica de ThCBHI.

4.10.3.1 Espalhamento dinâmico de luz (DLS)

Para investigar agregados da ThCBHI em altas temperaturas, foi utilizado espalhamento dinâmico de luz (DLS), que distingue diferentes estados de oligomerização e/ou agregação da proteína. Quando ThCBHI foi analisado por DLS em 25 e 50 °C, os perfis observados foram característicos de proteína monodispersa em solução. Os raios

hidrodinâmicos determinados nestas temperaturas foram 3,11 e 3,19 nm, respectivamente, coerente, em ambos os casos, com uma macromolécula monomérica em solução. No entanto, quando a ThCBHI foi incubada a 80 °C, o perfil foi alterado e mostrou um raio hidrodinâmico de 36,78 nm, o que corresponde a agregados de elevada massa molecular (Figura 29).

Figura 29 - Análise utilizando a técnica de DLS para diferentes temperaturas para a ThCBHI. Demonstrando que a proteína se mantém intacta até 50°C (vermelho), e mostrando um estado de agregação quando medida a 80°C.

4.10.3.2 SAXS

Experimento de SAXS para a ThCBHI, mostrou um raio de giro (R_g) de 2,7 nm, esse resultado corrobora com os resultados obtidos para DLS, e também para CBHI de *Melanocarpus albomyces* que mostrou um raio de 3,2 nm.¹³⁴ O R_g foi obtido a partir da curva de SAXS usando aproximação Guinier (Figura 30-inserção).

Figura 30 – Curva de SAXS para calcular o R_g da Cellobiohydrolase I ($\ln I(q)$ vs q). Insert: Guinier plot ($\ln I(q)$ vs q^2) para calcular o R_g .

Para obter informações mais detalhadas sobre a estrutura terciária da enzima, também foram obtidos os dados de SAXS em função da temperatura. A Figura 31 mostra as curvas de dispersão de raios X da ThCBHI para diferentes temperaturas ($T = 20, 50, 55, 60$ e 65 °C). Numa primeira análise, os dados mostram que não há alterações estruturais na estrutura de baixa resolução de ThCBHI a temperaturas entre 20 e 55 °C. No entanto, a dispersão foi alterada quando a temperatura foi aumentada acima de 55 °C, indicando que ThCBHI começou a perder estrutura terciária. Para uma mais análise detalhada dos efeitos da temperatura sobre a estrutura terciária de ThCBHI, os dados de SAXS foram também analisados em termos de gráficos de Kratky. Como pode ser visto na figura 31 B, todas as curvas correspondentes aos experimentos conduzidos entre as temperaturas de 20 e 55 °C se mostram bem, como prevista, a máxima compactação para a proteína.

No entanto, para temperaturas acima de 55 °C, o gráfico de Kratky mostra claramente uma alteração na intensidade de ambos os picos, indicando que ocorreram fortes alterações para a compactação da estrutura da proteína. Em conjunto, os resultados de CD e fluorescência, corroboram com os resultados de SAXS que demonstraram que a estrutura secundária e terciária de ThCBHI e sua compacidade são fortemente dependentes da temperatura. Em conclusão, o comportamento térmico para o desenovelamento da ThCBHI é um processo que pode ser dividido em duas etapas, em pH 5,0. O primeiro passo levou a proteína à formação de uma estrutura mais globular, entre 55 e 65 °C, seguido por um segundo passo que era devido ao desenovelamento do estado globular a temperaturas acima de 70 °C. Este estado globular apresenta as seguintes características: elas reduziram a atividade da enzima, pois tiveram um conteúdo de estrutura secundária reduzido, perdendo assim o perfil da estrutura mais compacta, e expondo, mais a superfície hidrofóbica ao solvente.

Figura 31 - Curva de espalhamento experimental em função da temperatura. (A) Curvas de SAXS da Th Cel7A mostrado como logaritmo da intensidade de espalhamento ($\log I$) versus vetor transferência de momento, q (B) Gráfico de Kratky para Th Cel7A. Os valores de temperaturas testadas foram de 20-65 °C

4.11 Estrutura cristalográfica do CCD de ThCBHI

4.11.1 Cristalização

O domínio catalítico (CCD) da ThCBHI foi cristalizado usando robô cristalização e vários conjuntos comerciais de condições de cristalização que cobrem mais de 650 soluções diferentes. Formas cristalinas vistas como pequenas agulhas foram observadas somente em uma condição, que continha 35% de PEG 4000 como precipitante. Aglomerados de pequenas agulhas apareceram muito rapidamente nesta condição de cristalização, de modo estavam completamente formados após 10 h de incubação (Figura 32 A). Várias tentativas foram realizadas para tentar reduzir o número de núcleos formados e para favorecer o crescimento de cristais maiores e único. Variações no pH, concentração de proteína e precipitante e testes de diferentes aditivos foram ineficazes, o que continuou resultando exclusivamente em um amontoado de microcristais ou conjuntos de agulhas. Essas agulhas foram testadas para confirmar o seu conteúdo proteico e para estimar o poder de difração de raios-X das mesmas. Surpreendentemente, um único cristal em formato agulha com dimensões aproximadas de 10x5x5 μm mostrou um padrão de difração, que se estendeu a uma resolução de cerca de 2,9 Å (Figura 32 A). Com isso, foram iniciados experimentos de *seeding* para obtermos uma condição de cristalização com cristais melhores, onde os cristais aglomerados eram divididos, “pescados” com um pêlo e colocados num poço novo, com a mesma condição de cristalização anterior. Nesta etapa, conseguimos cristais um pouco melhores, mas ainda assim eram muito pequenos e em formato de placas (Figura 32 B).

Figura 32 - Cristais de ThCBHI CCD em gota suspensa. A) 35% (w / v) de PEG 4000, cristais utilizados para a coleta do primeiro conjunto de dados, B) Cristais provenientes da técnica de *Seeding*.

Esses cristais em formato de placas foram submetidos à coleta de dados após serem protegidos com soluções crio, e apresentaram padrões de difração com pontos únicos e resolução máxima de $1,7\text{\AA}$ (Figura 33 B), resolução essa superior ao conjunto coletado anteriormente, o que mostra a eficiência da metodologia aplicada para a otimização das condições de cristalização. Os dois conjuntos de dados foram processados, integrados e escalonados e as estatísticas de coleta estão mostrados na Tabela 5.

Figura 33 - Padrões de difração obtidos para ThCBHI. A) Padrão obtido na primeira coleta, com $2,89\text{\AA}$ e B) Padrão obtido na segunda coleta, após experimentos de *seeding*, com $1,7\text{\AA}$.

A tabela abaixo mostra as estatísticas de refinamento das estruturas de ThCBHI.

Tabela 5 - Estatística de refinamentos das duas coletas feitas para ThCBHI CCD

	Dataset 1	Dataset 2
Dados de difração		
Grupo Espacial	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Parâmetros da célula unitária (Å)	a=56,9 b= 69,3 c=90,7	a=56,7 b= 69,2 c=90,6
Intervalo de resolução (Å)	55,1-2,9 (3,0-2,9)	45,30-1,67 (1,71-1,67)
Número de reflexões únicas	8471 (1154)	42070 (4077)
Multiplicidade	7,0 (6,7)	6,1 (5,5)
Completeza (%)	99,6 (99,6)	99,8 (98,8)
R_{merge} (%)^b	11,9 (29,5)	9,1 (32,9)
<I>/<sigma (I)>^b	6,5 (2,8)	24,4 (4,7)
Refinamento		
R_{factor}/R_{free} (%)	22,2/26,2	14,6/17,3
Número de átomos		
Proteína	3169	3193
N-acetyl-D-glucosamine	42	42
PEG	10	23
Água	77	562
B-fator (Å²)		
Proteína	20,8	10,3
N-acetyl-D-glucosamina	47,5	29,7
PEG	30,9	30,4
Água	19,1	24,2
R.m.s.d.^c		
Comprimento das ligações (Å)	0,003	0,011
Ângulo das ligações (°)	1,035	1,331
Ramachandran	97,7/2,3	98,1/1,9

^a Os valores entre parênteses são para a faixa de mais alta resolução.

^b $R_{\text{merge}} = \sum |I_j - \langle I \rangle| / \sum \langle I_j \rangle$, onde I_j é a intensidade observada de uma reflexão individual e $\langle I \rangle$ é a intensidade média dessa reflexão.

^c desvio quadratico médio.

4.11.2 Descrição da estrutura cristalográfica determinada

O modelo refinado da estrutura do ThCBHI compreende os resíduos 18 a 445. O domínio catalítico da enzima assemelha-se ao enovelamento característico das celobiohidrolases da família 7A (Figura 34) ^{135; 136} que compreende várias folhas- β antiparalelas formando uma estrutura terciária β -sanduíche que em conjunto com algumas regiões helicoidais e *loops*, formam um túnel alongado de ligação à celulose ao longo do lado côncavo do β -sanduíche. O túnel inclui o sítio ativo na sua porção terminal, onde os resíduos catalíticos de aspartato e glutamatos estão localizados.

Os dois modelos obtidos são quase idênticos (rms desvio de 0,3 Å para todos os pares de resíduos alinhados C α), apenas poucas diferenças conformacionais são vistas e em sua maioria em cadeias laterais dos resíduos expostos ao solvente. A ThCBHI CCD contém um total de 20 resíduos de cisteína, todos envolvidos em pontes dissulfeto, esses resíduos são conhecidos por estabilizar os *loops* ao redor da fenda de ligação de celulose, como comumente encontrados em enzimas secretadas por fungos ¹³⁵⁻¹³⁶, e uma glutamina N-terminal ciclizada por um resíduo pyroglutamato.

Figura 34—A). Organização estrutural da CBHI de *Trichoderma harzianum*. B). Diagrama de ribbon representado a estrutura cristalográfica de ThCBHI destacando as moléculas de PEG no túnel de ligação à celulose e os loops numerados de L1 a L5 que compõem esse túnel.

Além desses resíduos proteicos, a densidade eletrônica da estrutura cristalográfica revelou claramente a presença de sítios de N-glicosilação em 3 dos 4 resíduos preditos de acordo com a sequência N-X-S/T (onde X representa qualquer resíduo exceto prolina). Assim, moléculas N-acetilglucosamina (NAG) foram modeladas nos resíduos Asn126, Asn283 e Asn397 e todos os resíduos NAG fazem contato com as moléculas de proteína simetricamente relacionadas.

Praticamente todos os resíduos localizados no túnel de ligação a celulose exibem as conformações laterais parecidas nas Estruturas 1 e 2, exceto para Tyr260. Este resíduo está localizado no “loop” 5, o que parece ser uma região de proteína flexível em ambas as estruturas como revelado pelos fatores de temperatura mais altas, em comparação com outras partes da proteína. Na estrutura 1, a cadeia lateral do Tyr260 está apontando para o lado

interior do túnel, enquanto na Estrutura 2, a densidade eletrônica da cadeia lateral de Tyr260 é fraca e foi modelada em uma conformação mais exposta, interagindo hidrofobicamente com Ala384, como mostrado na Figura 35.

Figura 35 - Resíduos laterais do túnel de ThCBHI e TrCBHI. (A) sobreposição estrutural de ThCBHI do conjunto de dados 1 (cinza claro) e conjunto de dados 2 (cinza escuro) destacando as diferentes conformações assumidas pela Tyr260. (B) Sobreposição estrutural de TrCBHI 1CEL (azul) e 7CEL TrCBHI (roxa), destacando a dependência em conformações assumidas pelo Tyr247 e Tyr371.

4.11.3 Comparação da estrutura de ThCBHI com outras celobiohidrolases I

Consistente com a alta identidade sequencial entre CBHI de *T. reesei* e *T. harzianum*, a sobreposição estrutural mostra que ambas as moléculas têm um enovelamento muito semelhante, com um desvio rms de 0,5 Å entre os 426 pares de resíduos alinhados. Como esperado, os maiores desvios estruturais estão associadas com as regiões dos *loops* que exibem fatores de temperatura alta e, portanto, são regiões móveis ou estruturalmente desordenadas. No entanto, uma primeira diferença estrutural significativa entre ThCBHI e TrCBHI encontra-se no *loop* 3 (L3) na entrada do túnel (resíduos Ser99 a Gln101 na sequência TrCBHI). A localização do L3 é crucial para a conformação geral dessa abertura e alguns resíduos que compõem esse *loop* em TrCBHI estão ausentes em ThCBHI. Isso faz com que a fenda de acesso ao túnel seja maior em ThCBHI do que em TrCBHI (Figuras 36 e

37). O equivalente L3 em TrCBHI engloba o resíduo Gln101 que é responsável pelo reconhecimento por pontes de hidrogênio das unidades glicosídicas da celulose nos sítios -6 e -5⁵⁰, e essas interações são, portanto, indisponíveis em ThCBHI.

T_harzianum	18	EQVCTQQAE TH PP ^{LT} WQKCTA.S.GCTPQQGSVVLDANWRW TH DTKST ^{TN}
T_emersonii	1	EQAGTATAEN ^{HP} PL ^{TW} QE ^{CT} A.PGSCTTQNGAVVLDANWRW ^{VH} DVNGY ^{TN}
P_chrysosporium	1	EQAGTNTAEN ^{HP} QLQSQQCTT.SGGCKPLSTKVVLD ^{SN} WRW ^{VH} STSGY ^{TN}
T_reesei	1	ESACTLQSE TH PP ^{LT} WQKCSSGG.TCTQQTGSVVIDANWRW TH ATNSS ^{TN}
L3		
T_harzianum	66	CYDGN ^{TW} SS ^{TL} CPDDATCAK ^{NC} CLDGANYS ^{GT} YGV ^{TT} SGDAL ^{TL} QFV ^{TA} .
T_emersonii	50	CYTGN ^{TW} DP ^{TY} CPDDE ^{TC} AQ ^{NC} ALDGADYEG ^{TY} GV ^{TS} SGSSLK ^{LN} EFV ^{TG} .
P_chrysosporium	50	CYTGN ^{EW} DTSLCPDGK ^{TC} AAN ^{CA} ALDGADYS ^{GT} YGI ST GTAL ^{TL} KFV ^{TG} .
T_reesei	50	CYDGN ^{TW} SS ^{TL} CPDNETCAK ^{NC} CLDGAAYAS ^{TY} GV ^{TT} SGNSLSIDFV ^{TQS}
L3		
T_harzianum	115	..SNVGS ^{RL} YL ^{MA} NDSTY ^{QE} FTLSGNEFS ^{FD} VDV ^{SQ} LP ^{CG} LN ^{GA} LYFV SM
T_emersonii	99	..SNVGS ^{RL} YL ^{LQ} DDSTY ^{QI} FKLLNREFS ^{FD} VDV ^{SN} LP ^{CG} LN ^{GA} LYFV ^{AM}
P_chrysosporium	99	..SNVGS ^{RV} YL ^{MA} DDTHY ^{QL} LKLLNQE ^{FT} FDVDM ^{SN} LP ^{CG} LN ^{GA} LYLS ^{AM}
T_reesei	100	AQK ^{NV} GAR ^{LY} LMASD ^{TY} QE ^{FT} LLGNEFS ^{FD} VDV ^{SQ} LP ^{CG} LN ^{GA} LYFV SM
L5		
T_harzianum	163	DADGGQ ^{SK} YPGNAAGAKY ^{GT} GY ^{CD} SQCP RD LK ^{FI} NGQANVEGW ^{EP} SSN ^{NA}
T_emersonii	147	DADGGV ^{SK} YPNNKAGAKY ^{GT} GY ^{CD} SQCP RD LK ^{FI} DGEANVEGW ^{QP} PS.snn
P_chrysosporium	147	DADGGM ^{SK} YPGNKAGAKY ^{GT} GY ^{CD} SQCP ^{KD} IK ^{FI} NGEANVGNWTETGS ^N
T_reesei	150	DADGGV ^{SK} YPTNTAGAKY ^{GT} GY ^{CD} SQCP RD LK ^{FI} NGQANVEGW ^{EP} SSN
L5		
T_harzianum	213	N.TGVGGH ^{GS} CCSEMDIWEANSISEAL ^{TP} HP ^{PC} ETV ^{GQ} TMCSG ^{DS} CGG ^{TYS}
T_emersonii	196	anTGIGDH ^{GS} CCAEMDVWEANSISNAV ^{TP} HP ^{PC} DTPG ^Q TMCSG ^{DD} CGG ^{TYS}
P_chrysosporium	196	..TGTGSYGTCCSEMDIWEANNDAAAF ^{TP} HP ^{PC} TTT ^{GQ} TRCSG ^{DD} CA....
T_reesei	200	N.TGIGGH ^{GS} CCSEMDIWQANSISEAL ^{TP} HP ^{PC} TTV ^{GQ} EICEGDGCGG ^{TYS}
L5		
T_harzianum	262	ND ^{RY} GGTCD ^{PD} GD ^{CD} WN ^{PF} RLGNTS ^{FY} GP ^{GS} SFALD ^{TT} KK ^{KL} TVV ^{TQ} FAT..
T_emersonii	246	ND ^{RY} AGTCD ^{PD} GD ^{CD} FN ^{PF} RMGNTS ^{FY} GP..KIID ^{TT} KK ^{PF} TVV ^{TQ} FL ^{TDD}
P_chrysosporium	240	..RNTGLCD ^{GD} GD ^{CD} FN ^{SF} FMGDK ^{TF} L ^{GK} G..MTVD ^{TS} KK ^{PF} TVV ^{TQ} FL ^{TND}
T_reesei	249	DN ^{RY} GGTCD ^{PD} GD ^{CD} WN ^{PF} RLGNTS ^{FY} GP ^{GS} SFTLD ^{TT} KK ^{KL} TVV ^{TQ} FET..
A384		
T_harzianum	310	DGSIS ^{RY} YVQNGVKF ^{QQ} PNAQVGSYS.GNTINTDYCAAEQ ^{TA} FG
T_emersonii	294	GTDTGTLSEIK ^{RF} YIQNSNVI ^{PQ} PN ^{SD} ISGVT.GNSITTE ^{FC} TAQKQ ^{AF} FG
P_chrysosporium	286	NTSTGTLSEIR ^{RI} YIQNGKVI ^{QNS} VANIPGVDPVNSITDN ^{FC} CAQQK ^{TA} FG
T_reesei	297	SGAIN ^{RY} YVQNGVTF ^{QQ} PNAELGSYS.GNELND ^{DY} CTAEAE ^{EF} FG
A384		
T_harzianum	353	.GTSF ^{TD} KGGLAQINKAFQGG ^{ML} VLM ^{SL} WDDY ^{AN} NMLWLD ST YPTNATAS
T_emersonii	343	DTDDFSQH ^{GL} LAKMGAAMQQG ^{ML} VLM ^{SL} WDDY ^{AA} QMLWLD ^{SD} YPTDADPT
P_chrysosporium	336	DTNWF ^{AQ} KGGLKQMG ^{EAL} GN ^{GM} VLALSIWDDH ^{AA} NMLWLD ^{SD} YPTDKDPS
T_reesei	340	.GSSF ^{SD} KGGLTQFKKATSGG ^{ML} VLM ^{SL} WDDY ^{AN} NMLWLD ST YPTNETSS
A384		
T_harzianum	402	TPGAK ^{RG} SC ^{TS} SGVFAQVEAQSPNSKVIY ^{SN} IRFG ^{PI} GSTGGntgsn
T_emersonii	393	TPGIARG ^{TC} PTDSGV ^{PS} SDVESQSPNSYV ^{TY} SNIK ^{FG} PINSTFTas
P_chrysosporium	386	APGVARG ^{TC} ATTSGV ^{PS} SDVESQSPNSQV ^{VF} SNIK ^{FG} DIGSTFSGTS
T_reesei	389	TPGAV ^{RG} SC ^{TS} SGVFAOVESOSPNAKV ^{TF} SNIK ^{FG} PIGSTGNPSG

Figura 36 - Estrutura baseada em alinhamento das celobiohidrolases I de *T. harzianum* (UniProt Q9P8P3), *T. emersonii* (UniProt Q8TFL9, PDB 1Q9H), *P. chrysosporium* (UniProt Q7LIJ0, PDB 1GPI) e *T. reesei* (UniProt P07981.1, APO 1CEL). O alinhamento foi realizado no site Strap (<http://3d-alignment.eu/>). Numeração de sequências de proteínas está de acordo com as estruturas depositadas no PDB. Posição da Ala384 resíduos e lacunas em loops 3 e 5 estão retratados em caixas retangulares

Figura 37 - Superposição estrutural da CBHI de *T. harzianum* (cinza) e *T. reesei* (azul, código PDB 8CEL). Em foco a ausência do loop em *T. harzianum* destacando Gln101 e os resíduos de glicose a de TrCBHI (azul) na entrada do túnel de celulose de ligação. As ligações de H são mostradas com linhas tracejadas.

Um encurtamento semelhante do L3 também foi descrito para CBHIs de *T. emersonii* (TeCBHI) e *P. chrysosporium*. (PcCBHI). Os domínios catalíticos de TeCBHI e PcCBHI (códigos PDB 1Q9H e 1GPI, respectivamente) exibem uma identidade de sequência de 70% e 60% com o CCD de ThCBHI e um desvio estrutural rms de 0,8 Å para 417 Cα resíduos alinhados de TeCBHI e 0,9 Å para 415 Cα de PcCBHI, respectivamente. O equivalente L3 em ambas as estruturas é semelhante ao da ThCBHI e, conseqüentemente, eles também exibem uma entrada de túnel mais aberta quando comparado com o TrCBHI.¹³⁵⁻¹³⁶

Uma segunda característica importante na estrutura ThCBHI CCD está relacionada com os loops que formam as paredes laterais do túnel de ligação ao substrato (*loops* de 4 a 6), nos quais a conformação e composição da sequência são cruciais para a determinação do formato das faces laterais do túnel. Loops 5 e 6 de ThCBHI incluem uma parte da superfície do túnel lateral e criam uma abertura que dá acesso ao túnel de ligação ao local do substrato (Figura 38 A). Considerando que o *loop* 5 de PcCBHI seja significativamente menor, isso resulta em um túnel totalmente exposto (Figura 38B), a face lateral do canal de celulose de TrCBHI é quase completamente coberta (Figura 38C). A entrada para o túnel catalítico de

ThCBHI é também mais exposta, em comparação com TrCBHI (embora não tanto como em PcCBHI) os resíduos Ile203 e Tyr371 de ThCBHI são substituídos pelos resíduos de cadeias laterais mais curtas Val216 e Ala3840 (Figura 38 A).

Figura 38 - Representação superfície eletrostática potencial da CBHI de (A) *T. harzianum* (cinzento), (B) *P. chrysosporium* (ciano) e (C) *T. reesei* (azul). A face lateral do túnel está representada dentro de uma caixa retangular. Resíduos e loops que compõem a face lateral do túnel estão realçados.

Simulações de dinâmica molecular demonstraram que a presença do resíduo Ala384 no *loop* 6 de ThCBHI no lugar de Tyr371 em TrCBHI é responsável pelos efeitos notáveis sobre a dinâmica dos *loops* que englobam o sítio de ligação ao substrato. Nossas estruturas cristalográficas revelam que Tyr260 podem adotar conformações diferentes no domínio

catalítico de ThCBHI. Conformações da cadeia lateral da Tyr260 são mais expostas ao solvente e pode obstruir parcialmente o maior acesso ao túnel de ligação à celulose da ThCBHI. Em relação aos resíduos catalíticos de ligação à celulose, a única diferença conformacional observada está associada com a Glu217, que é uma ligação de H na unidade de glicose no sítio do produto +1 em TrCBHI, enquanto que o Glu230 equivalente de ThCBHI assume uma conformação alternativa formando uma ligação de H para Asp154.

Essa caracterização estrutural do *T. harzianum* CBHI CCD revela que esta enzima tem o enovelamento clássico da família de glicosil hidrolase 7, compreende várias folhas- β antiparalelas formando uma estrutura terciária β -sanduíche que em conjunto com algumas regiões helicoidais e *loops*, formam um túnel alongado de ligação à celulose ao longo do lado côncavo do β -sanduíche, típicos das celobiohidrolases, ao invés de um sulco muito mais aberto como encontrado em endoglucanases.^{50; 71} Apesar da elevada identidade da sequência de enzima CBHI a partir de *T. reesei*, a CBHI, no ponto de vista estrutural, é um intermediário entre a TrCel7A e PcCBHI, com similaridades significantes com a TeCBHI em termos de entrada de túnel e o formato (Figuras 36 e 38).

A fenda identificada ao lado do túnel, a qual dá acesso direto ao sítio ativo da enzima, expõe os glutamatos catalíticos de ThCBHI ao solvente, e é diretamente dependente da conformação da cadeia lateral Tyr260. A fenda está parcialmente obstruída quando este resíduo é exposto ao solvente. A flexibilidade observada de ThCBHI Tyr260 na ausência de ligante não é inesperado. Efeito semelhante foi previamente relatado para o Tyr247 equivalente em TrCBHI.^{50; 52; 71}

No entanto, as simulações MD, feitas em colaboração com o Prof. Dr. Munir Skaf, mostram que na ausência de substrato a presença da cadeia lateral curta de Ala384 permite um movimento praticamente livre de Tyr260 para as conformações internas e expostas ao solvente em ThCBHI. Os movimentos da cadeia lateral da Tyr247 em TrCBHI são restritos e são realizados simultaneamente com as mudanças conformacionais de Tyr371 (Ala384 em ThCBHI), que aponta para a parte interna do túnel quando Tyr247 está na conformação exposta ao solvente e vice-versa. Apesar da relativa liberdade conformacional da Tyr247, nossos resultados de MD indicam que este resíduo faz contatos persistentes com resíduo Tyr371 no “loop” 6 de TrCBHI. Assim, a abertura lateral existente quando Tyr260 é deslocado para o túnel em ThCBHI é na sua maior parte, coberta em TrCBHI quer por Tyr247 ou Tyr371 (Figuras 38 e 39).

A análise da sequência de celobiohidrolases I mostra que a posição ocupada pelo Ala384, embora moderadamente conservada, exhibe tipos predominantemente de certos

resíduos, entre os quais alanina ou tirosina são mais frequentes. Curiosamente, a maioria das enzimas CBHI têm uma abertura no *loop* 3 e, assim, apresentam uma maior entrada de túnel, também contém um resíduo de alanina, o que em *T. harzianum* é observado para garantir um acesso adicional para o túnel de ligação à celulose. Duas destas proteínas, PcCBHI e TeCBHI, têm as suas estruturas de raios-X determinada.^{135; 136}

A abertura lateral no túnel de ligação do substrato, mencionado acima, é estruturalmente novo. Embora a estrutura cristalográfica de TeCBHI careça de parte da face lateral do túnel devido a um *loop* flexível¹³⁶, a estrutura PcCBHI mostra um túnel completamente aberto.¹³⁵ Isto é devido a um *loop* 5 mais curto, que expõe completamente o túnel e pode estar associado a taxas mais elevadas de hidrólise, processabilidade reduzida e atividade de endoglucanase considerável de PcCBHI.^{87; 135}

Considerando-se a conservação do tamanho, forma e resíduos do túnel, é esperado que os contatos de açúcar descritos para TrCBHI⁵⁰ sejam mantidas em ThCBHI. Exceto para Gln101, todos os resíduos que interagem com as unidades de glicose em *T. reesei* são conservados em *T. harzianum*, incluindo os quatro resíduos de triptofano que formam os sítios -7, -4, -2, e +1⁵⁰ que, juntamente com resíduos de tirosina, interagem com as faces hidrofóbicas das unidades de açúcar.¹³⁷ Apesar das semelhanças descritas, uma vez que o resíduo Q101 correspondente TrCBHI não está presente na ThCBHI, a ausência de ligações específicas de H em -6 e -5 reduz a afinidade de ligação de ThCBHI para o substrato, mas o deslizamento da cadeia de celulose no sítio ativo pode ser favorecida. Além disso, a ligação específica descrita para a Tyr247 TrCBHI com o substrato no sítio -2 pode ser mais fraca em ThCBHI devido ao movimento menos restrito da Tyr260 equivalente e, assim, também contribuir para o deslizamento das unidades de açúcar e da catálise, como observado nos nossos estudos enzimáticos da enzima.

Figura 39 - Dinâmica Molecular de Celobiohidrolase I de *Trichoderma harzianum* e *Trichoderma reesei*, comparando uma região. Superfícies moleculares estão representadas em vermelho. O Tyr260 resíduo é mais móvel que Tyr247 e as mudanças conformacionais de Tyr260 em ThCBHI criam um local de ligação completamente exposto. Embora o resíduo Tyr247 (TrCBHI) também possa sofrer alterações conformacionais significativas, estes não são suficientes para criar um sítio de ligação aberto, devido à presença de resíduos de Tyr371 no *loop* 6, em oposição a Ala384 no *loop* 6 em ThCBHI.

4.12 Microscopia Eletrônica de Varredura

As técnicas de microscopia eletrônica de varredura têm promovido um grande impacto em ciência dos materiais devido à possibilidade de obtenção de imagens com alta resolução, com aparência tridimensional e por exigirem uma etapa relativamente simples de preparo de amostras. Nosso objetivo nesta parte do trabalho foi caracterizar as alterações morfológicas que ocorrem nas fibras de celulose do papel filtro como resultado da ação da enzima ThCBHI. A Figura 40 mostra o aspecto geral das fibras de papel filtro antes (A), e depois (B) da ação da ThCBHI, além do efeito de uma endoglucanase de *T. harzianum* (ThEGI) (C), e da ação sinérgica destas duas enzimas (D).

Figura 40 - Microscopia eletrônica de varredura de papel filtro mostrando a ação da ThCBHI e de uma endoglucanase de *T. harzianum* (ThEGI). Em A) controle, em B) papel depois de 24h sob a ação da CBHI, em C) papel depois de 24h sob a ação da ThEGI, em D) Ação conjunta da ThEGI e ThCBHI sobre as fibras. Barra de escala de 200 µm.

A análise da morfologia do papel filtro como modelo de substrato celulósico, claramente demonstra o efeito hidrolítico das celulasas. Antes da ação das enzimas (Figura 40A), o papel de filtro apresenta uma estrutura bastante fechada, com fibras mais espessas emaranhadas a

fibrilas mais finas. O efeito da ThEGI atuando sozinha sobre o substrato é suave (Figura 40C), mas a ação da CBHI, por outro lado, é bastante notável, como pode ser observado na Figura 40B. Pode-se observar que a estrutura do tecido em geral aparece bem mais aberta, com vazios maiores e com as fibras espessas do papel, que devem ser as mais resistentes às enzimas, bem mais independentes do que as que são originalmente observadas no papel de controle.

A morfologia mostrada na Figura 40D é efeito da ação conjunta das duas enzimas e é bastante similar à imagem da Figura 40B, com o papel apresentando uma estrutura aberta, de fibras independentes. Observando-se cuidadosamente a imagem em (D), é possível notar que o papel parece ligeiramente mais hidrolisado que em (B), mostrando fibras espessas ainda mais independentes, o que pode ser atribuído ao efeito sinérgico das duas enzimas.

As fibrilas que fecham a estrutura, interligando as fibras mais grossas, foram digeridas. Essas fibrilas mais finas desaparecem completamente em (B) e (D), enquanto as fibras mais espessas mantêm a mesma espessura. Esse resultado indica que aplicar algum método de desfibrilamento da estrutura do papel, como moagem a seco ou moagem das amostras úmidas (usando um liquidificador, por exemplo) pode facilitar muito a ação enzimática. De uma forma geral, considera-se que o processo de moagem (antes e depois dos pré-tratamentos da biomassa) melhora a ação enzima sobre o substrato devido à diminuição no tamanho de partículas.¹³⁸

Os resultados obtidos por MEV mostram, no entanto, que não é apenas uma questão de diminuir o tamanho das partículas em geral, mas principalmente de desfibrilar as fibras mais espessas, para que elas tornem-se também disponíveis para a ação da CBH I. Apesar de incluir uma etapa extra de pré-tratamento, isso poderia ser compensado pela remoção desse importante entrave à ação das enzimas.

Na Figura 41, o efeito de que a superfície está ficando mais lisa sob a ação das duas enzimas pode ser visto com mais detalhes em aumentos maiores (barra de escala de 10 μm). A superfície da fibra com a ação da ThEGI é bem semelhante a do papel controle (Figura 43 A e C), mas a do papel tratado com CBHI (B) ou com as duas enzimas juntas apresenta um aspecto muito mais liso (D), indicando que as fibrilas superficiais de mais fácil acesso foram consumidas pela enzima.

No geral, pode-se observar que cada uma das fibras espessas de celulose que aparecem na Figura 43, apresentam várias camadas de fibras menores (fibrilas) e têm um aspecto

rugoso, devido às fibrilas de celulose que se sobrepõem e se entrelaçam nessas camadas. As fibrilas da superfície são as mais expostas e depois da ação das enzimas, são as primeiras a serem removidas, indicando que as enzimas estavam ativas e agiram nas extremidades livres da cadeia, fazendo uma “limpeza” na estrutura do complexo lignocelulósico do papel filtro.

Estas alterações observadas pela ação da ThCBHI podem constituir a primeira visualização direta apoiando a ideia de que as exocelulases seletivamente hidrolisam extremidades livres da celulose. A acessibilidade limitada destas extremidades pode contribuir de forma significativa para a lentidão limitante da velocidade de hidrólise da celulose.¹³⁹

Figura 41 - Microscopia Eletrônica de Varredura de Papel Filtro mostrando a ação da ThCBHI e (ThEGI). Em A) controle, e em B) Papel depois de 24h sob a ação da CBHI, C) Papel depois de 24h sob a ação da ThEGI, D) Ação conjunta das enzimas ThCBHI e ThEGI. Barra de escala de 10 µm.

5 CONCLUSÕES

Os resultados apresentados e discutidos nessa tese, mostram que os objetivos delineados no início do doutorado foram realizados com sucesso. O primeiro passo do projeto foi estabelecer protocolos eficientes de expressão e purificação para a enzima CBHI do fungo filamentoso *Trichoderma harzianum*. Como esperado, muitas informações obtidas das análises biofísicas realizadas com a ThCBHI como rendimento da produção protéica, temperatura ótima de atividade próxima de 50°C e ponto isoelétrico ácido foram similares as obtidas com a CBHI do fungo *T. reesei*. Resultados de espectroscopia de dicroísmo circular (CD) mostraram que a ThCBHI possui uma estrutura secundária composta por uma grande porcentagem de estrutura regular do tipo folhas- β , o que não difere de outras celobiohidrolases descritas na literatura científica.

Outras análises da estrutura secundária em função da variação do pH, empregando CD, mostraram que o espectro da ThCBHI sofre modificações não apreciáveis em pHs ácidos, porém em pH 10 o espectro de CD mostrou modificações drásticas sugerindo perda parcial ou total de sua estrutura secundária regular. Resultados similares foram obtidos quando a estrutura terciária da ThCBHI foi analisada empregando a técnica de espectroscopia de fluorescência intrínseca. Em pHs ácidos nenhuma modificação apreciável no máximo de emissão dos triptofanos foi verificada. Porém, em pHs básicos extremos, um deslocamento apreciável do máximo de emissão e do centro de massa da curva para o vermelho foi observado, indicando modificações na estrutura terciária da proteína. As análises de estabilidade térmica empregando dicroísmo circular ou o método de termofluor mostraram que a ThCBHI tem um T_m em torno de aproximadamente 62 °C em pH 5,0, contudo, em pHs extremos ela diminui drasticamente este valor. Nossos resultados para ensaios de espalhamento de raios - X a baixo angulo mostram que o pH e perturbações de temperatura afetam a estabilidade da ThCBHI por dois mecanismos diferentes. Variações na protonação na modificação do pH dos resíduos de enzimas, afetando diretamente a sua atividade, levando à desestabilização estrutural apenas a pH extremo. Temperatura, por outro lado, tem uma influência direta sobre a mobilidade e a compacidade da enzima, causando desdobramento de ThCBHI apenas acima da temperatura ótima. Finalmente, foi demonstrado que a incubação com celobiose, o produto da reação e um inibidor competitivo, aumentaram significativamente a estabilidade térmica da ThCBHI. Estes estudos podem fornecer percepções sobre compreensão, a nível molecular sobre a interação entre estrutura e atividade das ThCBHI em diferentes pH e condições de temperatura.

Estudos cinéticos com a ThCBHI, quando comparada nas mesmas condições de purificação e ensaio com a TrCBHI, possui maior eficiência, o que torna esta enzima de

grande interesse biotecnológico para hidrólise de glicosídios. A estrutura cristalográfica determinada pode explicar os resultados cinéticos obtidos pois diferenças estruturais, como uma maior abertura na entrada do túnel e a existência de uma fenda lateral, podem facilitar a dinâmica tanto de entrada do substrato como saída do produto, aumentando assim sua eficiência catalítica quando comparada a TrCBHI. Todos os resultados apresentados nessa tese indicam que a enzima ThCBHI pode ser de grande interesse para o futuro do etanol de segunda geração provindos de hidrólise enzimática de resíduos lignocelulolíticos.

REFERÊNCIAS

- 1 CAO, S. Y.; XIE, G. D.; ZHEN, L. Total embodied energy requirements and its decomposition in China's agricultural sector. **Ecological Economics**. v. 69, n. 7, p. 1396-1404, 2010. ISSN 0921-8009.

- 2 VON OSSOWSKI, I. et al. Engineering the exo-loop of *Trichoderma reesei* cellobiohydrolase, Cel7A. A comparison with *Phanerochaete chrysosporium* Cel7D. **Journal of Molecular Biology**, v. 333, n. 4, p. 817-829, 2003. ISSN 0022-2836.

- 3 ZAPATA, C.; NIEUWENHUIS, P. Exploring innovation in the automotive industry: new technologies for cleaner cars. **Journal of Cleaner Production**. v. 18, n. 1, p. 14-20, Jan. 2010. ISSN 0959-6526.

- 4 DEMAINE, A. L.; NEWCOMB, M.; WU, J. H. D. Cellulase, clostridia, and ethanol. **Microbiology and Molecular Biology Reviews**. v. 69, n. 1, p. 124, Mar. 2005. ISSN 1092-2172.

- 5 LYND, L. R. et al. Fuel ethanol from cellulosic biomass. **Science**. v. 251, n. 4999, p. 1318-1323, Mar. ISSN 0036-8075.

- 6 RICHARD, T. L. Challenges in scaling up biofuels infrastructure. **Science**. v. 329, n. 5993, p. 793-796, Aug. 2010. ISSN 0036-8075.

- 7 ADEN, A.; FOUST, T. Technoeconomic analysis of the dilute sulfuric acid and enzymatic hydrolysis process for the conversion of corn stover to ethanol. **Cellulose**. v. 16, n. 4, p. 535-545, Aug. 2009. ISSN 0969-0239.

- 8 HIMMEL, M. E. Biomass recalcitrance: engineering plants and enzymes for biofuels production. **Science**. v. 316, n. 5827, p. 982-982, 2007. ISSN 0036-8075.

- 9 CORTEZ, L. A. B. **Bioetanol de cana-de-açúcar: P&D para produtividade e sustentabilidade**. São Paulo: Blucher, 2010. ISBN 9788521205319.

- 10 COMPANHIA NACIONAL DE ABASTECIMENTO. (CONAB) **Acompanhamento de safra brasileira:** cana-de-açúcar, Safra 2012-2013, primeiro levantamento. Acompanhamento da Safra Brasileira. Brasília: abril/2012.
- 11 BUCKERIDGE, M. S.; SANTOS, W. D.; SOUZA, A. P. As rotas para o etanol celulósico no Brasil. In: BLUCHER (Ed.). **Bioetanol de cana-de-açúcar: P&D para produtividade e sustentabilidade.** São Paulo: Cortez, 2010. ISBN 9788521205319.
- 12 BRAGATTO, J. **Avaliação do potencial da casca de *Eucaliptus* spp. para a produção de bioetanol.** 2010. 154p. (Doutorado) – Departamento de Fisiologia e Bioquímica de Plantas, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2010.
- 13 VASQUEZ, M. P. et al. Enzymatic hydrolysis optimization to ethanol production by simultaneous saccharification and fermentation. **Applied Biochemistry and Biotechnology**. v. 137, p. 141-153, Sept. 2007. ISSN 0273-2289.
- 14 BASTOS, V. D. Etanol, alcoolquímica e biorrefinarias. **BNDS Setorial**, p.5-38, 2007
- 15 PEREIRA JR., N.; COUTO, M.A.P.G.; SANTA ANNA, L.M.M. Biomass of lignocellulosic composition for fuel ethanol production and the context of biorefinery. **Series on Biotechnology**, v.2, n. 3, p.45 , 2008
- 16 BANSAL, P. et al. Modeling cellulase kinetics on lignocellulosic substrates. **Biotechnology Advances**. v. 27, n. 6, p. 833-848, Nov.- Dec. 2009. ISSN 0734-9750.
- 17 COUTO, S. R. et al. Stainless steel sponge: a novel carrier for the immobilisation of the white-rot fungus *Trametes hirsuta* for decolourization of textile dyes. **Bioresource Technology**, v. 95, n. 1, Oct. 2004. ISSN 0960-8524.
- 18 CERQUEIRA, D. A.; RODRIGUES FILHO, G.; MEIRELES, C. D. S. Optimization of sugarcane bagasse cellulose acetylation. **Carbohydrate Polymers**, v. 69, n. 3, p. 579-582, June 25 2007. ISSN 0144-8617.

19 CARDONA, C. A.; QUINTERO, J. A.; PAZ, I. C. Production of bioethanol from sugarcane bagasse: status and perspectives. **Bioresource Technology**. v. 101, n. 13, p. 4754-4766, July 2010. ISSN 0960-8524.

20 FERREIRA, V. **Produção de β -glucosidase em *Saccharomyces cerevisiae* recombinante e avaliação de seu emprego no processo de hidrólise enzimática simultânea à fermentação para a produção de etanol de segunda geração** 2010. 132p. Tese (Doutorado) - Escola de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

21 YANG, B. et al. Enzymatic hydrolysis of cellulosic biomass. **Biofuels**. v. 2, n. 4, p. 29, 2011.

22 REGULAMENTADA a lei do tratamento do Lixo. Brasília, 2011. Ministério da Agricultura. Disponível em:
<<http://www.agricultura.gov.br/vegetal/noticias/2011/01/regulamentada-a-lei-do-tratamento-do-lixo>>. Acesso em: 05/07/2012.

23 NEUREITER, M. et al. Dilute-acid hydrolysis of sugarcane bagasse at varying conditions. **Applied Biochemistry and Biotechnology**. v. 98, p. 49-58, 2002. ISSN 0273-2289.

24 CASTRO, A. M. **Produção e propriedades de celulases de fungos filamentosos obtidas a partir de celulignina de bagaço de cana-de-açúcar (*Saccharum spp.*)**. 2006. 156p. Dissertação (Mestrado) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

25 SHALLOM, D.; SHOHAM, Y. Microbial hemicellulases. **Current Opinion in Microbiology**. v. 6, n. 3, p. 219-228, June 2003. ISSN 1369-5274.

26 COLLINS, T.; GERDAY, C.; FELLER, G. Xylanases, xylanase families and extremophilic xylanases. **Fems Microbiology Reviews**. v. 29, n. 1, Jan. 2005. ISSN 0168-6445.

27 ALVES, R. E. **Caracterização de fibras lignocelulósicas pré-tratadas por meio de técnicas espectroscópicas e microscópicas ópticas de alta resolução**. 2011. 117p. Dissertação (Mestrado) - Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, 2011.

- 28 LYND, L. R. et al. Microbial cellulose utilization: fundamentals and biotechnology. **Microbiology and Molecular Biology Reviews**. v. 66, n. 3, p. 506, Sept. 2002. ISSN 1092-2172. 29.
- 29 PEREZ, J. et al. Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. **International Microbiology**. v.5, n.2, p. 53-63, June, 2002.
- 30 ARISTIDOU, A.; PENTTILA, M. Metabolic engineering applications to renewable resource utilization. **Current Opinion in Biotechnology**. v. 11, n. 2, p. 187-198, Apr. 2000. ISSN 0958-1669.
- 31 BHAT, M.; BHAT, S. Cellulose degrading enzymes and their potential industrial applications. **Biotechnology Advances**. v. 15, n. 3-4, p. 583-620, 1997. ISSN 0734-9750.
- 32 SUN, J. et al. Isolation and characterization of cellulose from sugarcane bagasse. **Polymer Degradation and Stability**. v. 84, n. 2, p. 331-339, May 2004. ISSN 0141-3910.
- 33 SANDGREN, M.; STAHLBERG, J.; MITCHINSON, C. Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. **Progress in Biophysics & Molecular Biology**. v. 89, n. 3, p. 246-291, 2005. ISSN 0079-6107.
- 34 CHUNDAWAT, S. P. S. et al. Multi-scale visualization and characterization of lignocellulosic plant cell wall deconstruction during thermochemical pretreatment. **Energy & Environmental Science**, v. 4, n. 3, p. 973-984, Mar. 2011. ISSN 1754-5692.
- 36 HAMELINCK, C. N.; VAN HOOIJDONK, G.; FAAIJ, A. P. C. Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term. **Biomass & Bioenergy**. v. 28, n. 4, p. 384-410, 2005. ISSN 0961-9534.
- 36 BON, E. P. S.; FERRARA, M. A.; CORVO, M. L. **Enzimas em biotecnologia: produção, aplicações e mercado**. Rio de Janeiro: Ed. Interciência, 2008. p. 506.

37 BEGUIN, P.; AUBERT, J. P. The biological degradation of cellulose. **Fems Microbiology Reviews**. v. 13, n. 1, p. 25-58, Jan. 1994. ISSN 0168-6445.

38 CARREA, G.; COLOMBO, G. Coupling high enzyme activity and stability: a challenging target. **Trends in Biotechnology**. v. 18, n. 10, p. 401-402, Oct. 2000. ISSN 0167-7799.

39 SAMUELS, G. J. Trichoderma: a review of biology and systematics of the genus. **Mycological Research**. v. 100, p. 923-935, 1996. ISSN 0953-7562.

40 HARMAN, G. E. et al. Trichoderma species - opportunistic, avirulent plant symbionts. **Nature Reviews Microbiology**. v. 2, n. 1, p. 43-56, 2004. ISSN 1740-1526.

41 BISSETT, J. A revision of the genus Trichoderma. **Canadian Journal of Botany- Revue Canadienne De Botanique**. v. 69, n. 11, p. 2357-2372, 1991. ISSN 0008-4026.

42 JEOH, T. et al. Implications of cellobiohydrolase glycosylation for use in biomass conversion. **Biotechnology for Biofuels**, v.1, n.10, 2008. ISSN 1754-6834. DOI:10.1186/1754-6834-1-10

43 DE CASTRO, A. M. et al. High-yield endoglucanase production by Trichoderma harzianum IOC-3844 cultivated in pretreated sugarcane mill byproduct. **Enzyme Research**. p. 8, 2010. DOI: 10.4061/2010/854526.

44 HENRISSAT, B. A classification of glycosyl hydrolases based on amino-acid-sequence similarities. **Biochemical Journal**. v. 280, p. 309-316, 1991. ISSN 0264-6021.

45 TOMME, P. et al. Studies of the cellulolytic of *Trichoderma reesei* QM-9414 - analysis of domain function in 2 cellobiohydrolases by limited proteolysis. **European Journal of Biochemistry**. v. 170, n. 3, p. 575-581, 1988. ISSN 0014-2956.

46 COUTINHO, P. M.; HENRISSAT, B. Carbohydrate-active enzymes: an integrated database approach. In: GILBERT, H. J.; DAVIES, G. J. et al (Ed.). **Recent advances in carbohydrate bioengineering**. London: The Royal Society of Chemistry, 1999.p. 3-12. ISBN 0260-62910-85404-774-3.

47 CANTAREL, B. L. et al. The Carbohydrate-Active EnZymes Database (CAZy): an expert resource for Glycogenomics. **Nucleic Acids Research**, v. 37, p. D233-D238, 2009. ISSN 0305-1048.

48 BISCHOFF, K. M. et al. Purification and characterization of a family 5 endoglucanase from a moderately thermophilic strain of *Bacillus licheniformis*. **Biotechnology Letters**. v. 28, n. 21, p. 1761-1765, Nov. 2006. ISSN 0141-5492.

49 GILKES, N. R. et al. Attack of carboxymethylcellulose at opposite ends by two cellobiohydrolases from *Cellulomonas fimi*. **Journal of Biotechnology**. v. 57, n. 1-3, p. 83-90, Sept. 1997. ISSN 0168-1656.

50 PALONEN, H. et al. Adsorption of *Trichoderma reesei* CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin. **Journal of Biotechnology**. v. 107, n. 1, p. 65-72, Jan. 2004. ISSN 0168-1656.

51 GAO, J. M. et al. Production and characterization of cellulolytic enzymes from the thermoacidophilic fungal *Aspergillus terreus* M11 under solid-state cultivation of corn stover. **Bioresource Technology**. v. 99, n. 16, p. 7623-7629, 2008. ISSN 0960-8524.

52 DIVNE, C. et al. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. **Science**. v. 265, n. 5171, p. 524-528, July 1994. ISSN 0036-8075.

53 _____. High-resolution crystal structures reveal how a cellulose chain is bound in the 50 angstrom long tunnel of cellobiohydrolase I from *Trichoderma reesei*. **Journal of Molecular Biology**. v. 275, n. 2, p. 309-325, 1998. ISSN 0022-2836.

54 BARR, B. K. et al. Identification of two functionally different classes of exocellulases. **Biochemistry**. v. 35, n. 2, p. 586-592, 1996. ISSN 0006-2960.

55 STAHLBERG, J. et al. Activity studies and crystal structures of catalytically deficient mutants of cellobiohydrolase I from *Trichoderma reesei*. **Journal of Molecular Biology**. v. 264, n. 2, p. 337-349, 1996. ISSN 0022-2836.

56 FAGERSTAM, L. G.; PETTERSSON, L. G. The 1.4-Beta-glucan cellobiohydrolases of *Trichoderma reesei* QM-9414 - a new type of cellulolytic synergism. **Febs Letters**. v. 119, n. 1, p. 97-100, 1980. ISSN 0014-5793.

57 NUMMI, M. et al. Cellobiohydrolase from *Trichoderma reesei*. **Biochemical Journal**. v. 215, n. 3, p. 677-683, 1983. ISSN 0264-6021.

58 BECKHAM, G. T., BOMBLE Y. J., BAYER E. A.; HIMMEL, M. E., Crowley M. F.. Applications of computational science for understanding enzymatic deconstruction of Cellulose. **Current Opinion in Biotechnology**. v. 22, p. 1–8, 2010.

59 KOROTKOVA, O. G. et al. Isolation and properties of fungal beta-glucosidases.). **Biochemistry (Mosc)**, v.74, n. 5, p.569-77,May 2009. ISBN 1608-3040

60 LIU, L. et al. Study on binding modes between cellobiose and beta-glucosidases from glycoside hydrolase family 1. **Bioorganic & Medicinal Chemistry Letters**. v. 22, n. 2, p. 837-843, Jan.15 2012. ISSN 0960-894X.

61 GHOSE, T. K. Measurement of cellulase activities. **Pure and Applied Chemistry**. v. 59, n. 2, p. 257-268, 1987. ISSN 0033-4545.

62 KLEYWEGT, G. J. et al. The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 angstrom resolution, and a comparison with related enzymes. **Journal of Molecular Biology**. v. 272, n. 3, p. 383-397, 1997. ISSN 0022-2836.

63 WATANABE, H.; TOKUDA, G. Cellulolytic systems in insects. **Annual Review of Entomology**. v. 55, 2010. ISSN 0066-4170. DOI:10.1146/annurev-ento-112408-085319

64 SANDGREN, M. et al. Comparison of family 12 glycoside hydrolases and recruited substitutions important for thermal stability. **Protein Science**. v. 12, n. 4, p. 848-860, 2003. ISSN 0961-8368.

65 HARRISON, M. J. et al. Modified glycosylation of cellobiohydrolase I from a high cellulase-producing mutant strain of *Trichoderma reesei*. **European Journal of Biochemistry**. v. 256, n. 1, p. 119-127, 1998. ISSN 0014-2956.

66 SRISODSUK, M. et al. Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose. **Journal of Biological Chemistry**. v. 268, n. 28, p. 20756-20761, 1993. ISSN 0021-9258.

67 REINIKAINEN, T. et al. Investigations of the function of mutated cellulose binding domains of *Trichoderma reesei* cellobiohydrolase I. **Proteins-Structure Function and Genetics**. v. 14, n. 4, p. 475-482, Dec.1992. ISSN 0887-3585.

68 LINDER, M.; TEERI, T. T. The cellulose-binding domain of the major cellobiohydrolase of *Trichoderma reesei* exhibits true reversibility and a high exchange rate on crystalline cellulose. **Proceedings of the National Academy of Sciences of the United States of America**. v. 93, n. 22, p. 12251-12255, Oct. 1996. ISSN 0027-8424.

69 TEERI, T. et al. Domain Function in *Trichoderma reesei* cellobiohydrolases. **Journal of Biotechnology**. v. 24, n. 2, p. 169-176, June 1992. ISSN 0168-1656.

70 STAHLBERG, J.; JOHANSSON, G.; PETTERSSON, G. A new model for enzymatic hydrolysis of cellulose based on the 2-domains structure of cellobiohydrolase-I. **Bio-Technology**. v. 9, n. 3, p. 286-290, 1991. ISSN 0733-222X.

71 STALS, I. et al. Factors influencing glycosylation of *Trichoderma reesei* cellulases. I: postsecretorial changes of the O- and N-glycosylation pattern of Cel7A. **Glycobiology**. v. 14, n. 8, p. 713-724, 2004. ISSN 0959-6658.

72 MEDVE, J. et al. Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II from *Trichoderma reesei*: adsorption, sugar production pattern, and synergism of the enzymes. **Biotechnology and Bioengineering**. v. 59, n. 5, p. 621-634, 1998. ISSN 0006-3592.

73 TEERI, T. T. et al. Homologous domains in *Trichoderma reesei* cellulytic enzymes - gene sequence and expression of cellobiohydrolase II. **Gene**. v. 51, n. 1, p. 43-52, 1987. ISSN 0378-1119.

74 DIVNE, C. et al. The 3-dimensional crystal-structure of the catalytic core of cellobiohydrolase-I from *Trichoderma reesei*. **Science**. v. 265, n. 5171, p. 524-528, July 1994. ISSN 0036-8075.

75 ROUVINEN, J. et al. 3-Dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. **Science**. v. 249, n. 4967, p. 380-386, 1990. ISSN 0036-8075.

76 SANDGREN, M. et al. The X-ray crystal structure of the *Trichoderma reesei* family 12 endoglucanase 3, Cel12A, at 1.9 angstrom resolution. **Journal of Molecular Biology**. v. 308, n. 2, p. 295-310, 2001. ISSN 0022-2836.

77 KRAULIS, P. J. et al. Determination of the 3-Dimensional structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei* - a study using nuclear magnetic resonance and hybrid distance geometry dynamical simulated annealing. **Biochemistry**. v. 28, n. 18, p. 7241-7257, 1989. ISSN 0006-2960.

78 MATTINEN, M. L. et al. Three-dimensional structures of three engineered cellulose-binding domains of cellobiohydrolase I from *Trichoderma reesei*. **Protein Science**, v. 6, n. 2, p. 294-303, Feb. 1997. ISSN 0961-8368.

79 DIVNE, C. et al. The 3-Dimensional crystal structure of the catalytic core of cellobiohydrolase from *Trichoderma reesei*. **Science**. v. 265, n. 5171, p. 524-528, July 1994. ISSN 0036-8075.

80 VASCONCELOS, A.; CAVACO-PAULO, A. Enzymatic removal of cellulose from cotton/polyester fabric blends. **Cellulose**. v. 13, n. 5, p. 611-618, Oct. 2006. ISSN 0969-0239.

81 BHAT, M. K. Cellulases and related enzymes in biotechnology. **Biotechnology Advances**. v. 18, n. 5, p. 355-383, 2000. ISSN 0734-9750.

82 ZHOU, J. et al. Identification and purification of the main components of cellulases from a mutant strain of *Trichoderma viride* T 100-14. **Bioresource Technology**. v. 99, n. 15, p. 6826-6833, Oct. 2008. ISSN 0960-8524.

83 AHAMED, A.; VERMETTE, P. Culture-based strategies to enhance cellulase enzyme production from *Trichoderma reesei* RUT-C30 in bioreactor culture conditions. **Biochemical Engineering Journal**. v. 40, n. 3, p. 399-407, 2008. ISSN 1369-703X.

84 MANDELS, M.; WEBER, J. Production of cellulases. **Advances in Chemistry Series**. n. 95, p. 391-414, 1969. ISSN 0065-2393.

85 LAEMMLI, U. K. Cleavage of structural proteins during assembly of head of bacteriophage-T4. **Nature**. v. 227, n. 5259, p. 680-685, 1970. ISSN 0028-0836.

86 STRYER, L. Molecular-basis of visual excitation. **Cold Spring Harbor Symposia on Quantitative Biology**. v. 53, p. 283-294, 1988. ISSN 0091-7451.

87 TARENTINO, A. L.; PLUMMER, T. H. Oligosaccharide chain-cleaving enzymes for deglycosylation of asparagine-linked glycans. **Journal of Cellular Biochemistry**. p. 256-256, 1994. ISSN 0730-2312.

88 WOOD, T. M.; BHAT, K. M. Methods for measuring cellulase Activities. In: WOOD, W.A.; KELLOG, S.T. **Methods in enzymology**. San Diego: Academic Press, 1988. v. 160, cap.9, p. 87-112. ISSN 0076-6879.

89 VALJAMAE, P. et al. Acid hydrolysis of bacterial cellulose reveals different modes of synergistic action between cellobiohydrolase I and endoglucanase I. **European Journal of Biochemistry**. v. 266, n. 2, p. 327-334, Dec. 1999. ISSN 0014-2956.

90 PARK, S. et al. Measuring the crystallinity index of cellulose by solid state C-13 nuclear magnetic resonance. **Cellulose**. v. 16, n. 4, p. 641-647, Aug. 2009. ISSN 0969-0239.

91 ZHANG, Y. H. P.; HIMMEL, M. E.; MIELENZ, J. R. Outlook for cellulase improvement: screening and selection strategies. **Biotechnology Advances**. v. 24, n. 5, p. 452-481, 2006. ISSN 0734-9750.

92 JALAK, J.; VAELJAMAE, P. Mechanism of initial rapid rate retardation in cellobiohydrolase catalyzed cellulose hydrolysis. **Biotechnology and Bioengineering**. v. 106, n. 6, p. 871-883, Aug. 15 2010. ISSN 0006-3592.

93 MILLER, G. L. Use of dinitrosalicylic acid reagent for determination of reducing sugar. **Analytical Chemistry**. v. 31, n. 3, p. 426-428, 1959. ISSN 0003-2700.

94 SREERAMA, N.; WOODY, R. W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. **Analytical Biochemistry**. v. 287, n. 2, p. 252-260, 2000. ISSN 0003-2697.

95 WU, S. T.; SKOLNICK, J.; ZHANG, Y. Ab initio modeling of small proteins by iterative TASSER simulations. **Bmc Biology**. v. 5, 2007. ISSN 1741-7007. DOI: 10.1186/1741-7007-5-17

96 WU, S. T.; ZHANG, Y. MUSTER: improving protein sequence profile-profile alignments by using multiple sources of structure information. **Proteins-Structure Function and Bioinformatics**. v. 72, n. 2, p. 547-556, 2008. ISSN 0887-3585.

97 POLLASTRI, G.; MCLYSAGHT, A. Porter: a new, accurate server for protein secondary structure prediction. **Bioinformatics**. v. 21, n. 8, p. 1719-1720, 2005. ISSN 1367-4803.

98 HELLES, G. A comparative study of the reported performance of ab initio protein structure prediction algorithms. **Journal of the Royal Society Interface**. v. 5, n. 21, Apr. 2008. ISSN 1742-5689.

99 MOULT, J. et al. Critical assessment of methods of protein structure prediction-Round VIII. **Proteins-Structure Function and Bioinformatics**. v. 77, 2009. ISSN 0887-3585. DOI: 10.1002/prot.22589.

100 CREIGHTON, T. E. **The physical and chemical basis of molecular biology**. Omaha: Helvetian Press, 2010. 688p. ISBN-10: 0956478107.

101 HAMMERSLEY, A. P. **FIT2D**: an introduction and overview. ESRF Internal Report. 1997. Disponível em:<http://www.csb.yale.edu/userguides/datamanip/fit2d/fit2d_ref.pdf>. Acesso em: 05 maio 2012.

102 GUIMARAES, B. G. et al. The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS. **Journal of Synchrotron Radiation**. v. 16, p. 69-75, Jan. 2009. ISSN 0909-0495.

103 OTWINOWSKI, Z.; MINOR, W. Processing of X-ray diffraction data collected in oscillation mode. **Macromolecular Crystallography. Pt A**, v. 276, p. 307-326, 1997. ISSN 0076-6879.

104 STORONI, L. C.; MCCOY, A. J.; READ, R. J. Likelihood-enhanced fast rotation functions. **Acta Crystallographica Section D-Biological Crystallography**. v. 60, p. 432-438, 2004. ISSN 0907-4449.

105 EMSLEY, P.; COWTAN, K. Coot: model-building tools for molecular graphics. **Acta Crystallographica Section D-Biological Crystallography**. v. 60, p.2126-2132, 2004. ISSN 0907-4449.

106 ADAMS, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. **Acta Crystallographica Section D-Biological Crystallography**. v. 66, p. 213-221, 2010. ISSN 0907-4449.

107 PERRAKIS, A. et al. ARP/wARP and molecular replacement. **Acta Crystallographica Section D-Biological Crystallography**. v. 57, p. 1445-1450, 2001. ISSN 0907-4449.

108 DAVIS, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. **Nucleic Acids Research**. v. 35, p. W375-W383, 2007. ISSN 0305-1048.

109 BOER, H.; TEERI, T. T.; KOIVULA, A. Characterization of Trichoderma reesei cellobiohydrolase Cel7A secreted from Pichia pastoris using two different promoters. **Biotechnology and Bioengineering**. v. 69, n. 5, p. 486-494, Sept. 2000. ISSN 0006-3592.

110 WANG, L. S. et al. Comparison of domains function between cellobiohydrolase I and endoglucanase I from *Trichoderma pseudokoningii* S-38 by limited proteolysis. **Journal of Molecular Catalysis B-Enzymatic**. v. 24-5, p. 27-38, 2003. ISSN 1381-1177.

111 GUILFOILE, P. et al. Cloning and sequencing of a cellobiohydrolase gene from *Trichoderma harzianum* FP108. **Journal of the Minnesota Academy of Science**. v. 64, n. 1, p. 18-22, 1999.

112 ZHANG, H. M. et al. Simultaneous reduction and digestion of proteins with disulfide bonds for hydrogen/deuterium exchange monitored by mass spectrometry. **Analytical Chemistry**. v. 82, n. 4, Feb. 15 2010. ISSN 0003-2700.

113 GASTEIGER E., H. C. et al. Protein identification and analysis tools on the *ExPASy* serve. In: WALKER, J. M. (Ed.). **Proteomics protocols handbook**. Totowa: Humana Press, 2005. p.571-607.

114 MEDVE, J.; LEE, D.; TJERNELD, F. Ion-exchange chromatographic purification and quantitative analysis of *Trichoderma reesei* cellulases cellobiohydrolase I, II and endoglucanase II by fast protein liquid chromatography. **Journal of Chromatography A**. v. 808, n. 1-2, p. 153-165, May 1998. ISSN 0021-9673.

115 MANDELS, M.; REESE, E. T. Induction of cellulase in *Trichoderma viride* induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals. **Journal of Bacteriology**. v. 73, n. 2, p. 269-278, 1957. ISSN 0021-9193.

116 BOMMARIUS, A. S. et al. Cellulase kinetics as a function of cellulose pretreatment. **Metabolic Engineering**. v. 10, n. 6, p. 370-381, 2008. ISSN 1096-7176.

117 HUI, J. P. M.; WHITE, T. C.; THIBAUT, P. Identification of glycan structure and glycosylation sites in cellobiohydrolase II and endoglucanases I and II from *Trichoderma reesei*. **Glycobiology**. v. 12, n. 12, p. 837-849, Dec. 2002. ISSN 0959-6658.

118 CLAEYSSSENS, M. et al. Fungal cellulase systems - comparison of the specificities of cellobiohydrolases isolated from *Penicillium ponophilum* and *Trichoderma reesei*. **Biochemical Journal**. v. 261, n. 3, p. 819-825, Aug. 1989. ISSN 0264-6021.

119 DESHPANDE, M. V.; ERIKSSON, K. E.; PETTERSSON, L. G. An assay for selective determination of exo-1,4,-beta-glucanase in a mixture of cellulolytic enzymes. **Analytical Biochemistry**. v. 138, n. 2, p. 481-487, 1984. ISSN 0003-2697.

120 VOUTILAINEN, S. P. et al. Cloning, expression, and characterization of novel thermostable family 7 cellobiohydrolases. **Biotechnology and Bioengineering**. v. 101, n. 3, p. 515-528, Oct. 2008. ISSN 0006-3592.

121 VONHOFF, S. et al. Inhibition of cellobiohydrolases from *Trichoderma reesei*. synthesis and evaluation of some glucose-, cellobiose-, and cellotriose-derived hydroximolactams and imidazoles. **Helvetica Chimica Acta**, v. 82, n. 7, p. 963-980, 1999. ISSN 0018-019X.

122 JAGER, G. et al. Practical screening of purified cellobiohydrolases and endoglucanases with alpha-cellulose and specification of hydrodynamics. **Biotechnology for Biofuels**. v. 3, 2010. DOI: 10.1186/1754-6834-3-18.

123 VARNAI, A.; SIIKA-AHO, M.; VIKARI, L. Restriction of the enzymatic hydrolysis of steam-pretreated spruce by lignin and hemicellulose. **Enzyme and Microbial Technology**. v. 46, n. 3-4, p.185-193, Mar. 2010. ISSN 0141-0229.

124 QING, Q.; YANG, B.; WYMAN, C. E. Xylooligomers are strong inhibitors of cellulose hydrolysis by enzymes. **Bioresource Technology**. v. 101, n. 24, p. 9624-9630, Dec. 2010. ISSN 0960-8524.

125 IRWIN, D. C. et al. Activity studies of 8 purified cellulases - specificity, synergism, and binding domain effects. **Biotechnology and Bioengineering**. v. 42, n. 8, p. 1002-1013, 1993. ISSN 0006-3592.

126 BOER, H.; KOIVULA, A. The relationship between thermal stability and pH optimum studied with wild-type and mutant *Trichoderma reesei* cellobiohydrolase Cel7A. **European Journal of Biochemistry**. v. 270, n. 5, p. 841-848, 2003. ISSN 0014-2956.

127 VOUTILAINEN, S. P. et al. Improving the thermostability and activity of *melanocarpus albomyces* cellobiohydrolase Cel7B. **Applied Microbiology and Biotechnology**. v. 83, n. 2, p. 261-272, May 2009. ISSN 0175-7598.

128 EFTINK, M. R. Use of fluorescence spectroscopy as thermodynamics tool. **Energetics of Biological Macromolecules Pt C.** v. 323, p. 459-473, 2000. ISSN 0076-6879.

129 DONIACH, S. et al. Partially folded states of proteins - characterization by x ray scattering. **Journal of Molecular Biology.** v. 254, n. 5, p. 960-967, 1995. ISSN 0022-2836.

130 GARCIA, W. et al. Probing conformational changes in orphan nuclear receptor: the NGFI-B intermediate is a partially unfolded dimer. **Biophysical Chemistry.** v. 137, n. 2-3, p. 81-87, Oct. 2008. ISSN 0301-4622.

131_____. The stability and aggregation properties of the GTPase domain from human SEPT4. **Biochimica et Biophysica Acta - Proteins and Proteomics.** v. 1784, n.11, p. 1720-1727, Nov. 2008. ISSN 1570-9639.

132 DARIMONT, B. D. et al. Structure and specificity of nuclear receptor-coactivator interactions. **Genes & Development.** v. 12, n. 21, p. 3343-3356, 1998. ISSN 0890-9369.

133 SEMISOTNOV, G. V. et al. Protein globularization during folding. a study by synchrotron small-angle X-ray scattering. **Journal of Molecular Biology.** v. 262, n. 4, p. 559-574, 1996. ISSN 0022-2836.

134 ERICSSON, U. B. et al. Thermofluor-based high-throughput stability optimization of proteins for structural studies. **Analytical Biochemistry.** v. 357, n. 2, p. 289-298, 2006. ISSN 0003-2697.

135 GARCIA, W. et al. An intermediate structure in the thermal unfolding of the GTPase domain of human septin 4 (SEPT4/Bradeion-beta) forms amyloid-like filaments in vitro. **Biochemistry.** v. 46, n. 39, p. 11101-11109, Oct. 2007. ISSN 0006-2960.

136 PARKKINEN, T. et al. Preliminary X-ray analysis of cellobiohydrolase Cel7B from *Melanocarpus albomyces*. **Acta Crystallographica Section F - Structural Biology and Crystallization Communications.** v. 63, p. 754-757, 2007. ISSN 1744-3091.

137 MUNOZ, I. G. et al. Family 7 cellobiohydrolases from *Phanerochaete chrysosporium*: crystal structure of the catalytic module of Cel7D (CBH58) at 1.32 angstrom resolution and homology models of the isozymes. **Journal of Molecular Biology**. v. 314, n. 5, p. 1097-1111, 2001. ISSN 0022-2836.

138 GRASSICK, A. et al. Three-dimensional structure of a thermostable native cellobiohydrolase, CBHIB, and molecular characterization of the cel7 gene from the filamentous fungus, *Talaromyces emersonii*. **European Journal of Biochemistry**. v. 271, n. 22, p. 4495-4506, 2004. ISSN 0014-2956.

139 LIENEMANN, M. et al. Toward understanding of carbohydrate binding and substrate specificity of a glycosyl hydrolase 18 family (GH-18) chitinase from *Trichoderma harzianum*. **Glycobiology**. v. 19, n. 10, Oct. 2009. ISSN 0959-6658.

140 ZHU, J. Y.; PAN, X. J. Woody biomass pretreatment for cellulosic ethanol production: technology and energy consumption evaluation. **Bioresource Technology**. v. 101, n. 13, July 2010. ISSN 0960-8524.

141 LIU, Y. S. et al. Cellobiohydrolase hydrolyzes crystalline cellulose on hydrophobic faces. **Journal of Biological Chemistry**. v. 286, n. 13, p.11195-11201, Apr. 2011. ISSN 0021-9258.

ANEXO

Francieli Colussi,^a Larissa C. Textor,^a Viviane Serpa,^a Roberto Nobuyuki Maeda,^b Nei Pereira Jr^b and Igor Polikarpov^{a*}

^aInstituto de Física de São Carlos, Universidade de São Paulo, Avenida Trabalhador Saocarlene 400, 13560-970 São Carlos-SP, Brazil, and
^bUniversidade Federal do Rio de Janeiro, Centro de Tecnologia, Escola de Química, Laboratório de Desenvolvimento de Bioprocessos (LaDeBio), 21949-900 Rio de Janeiro-RJ, Brazil

Correspondence e-mail: ipolikarpov@ifsc.usp.br

Received 28 April 2010

Accepted 7 July 2010

Purification, crystallization and preliminary crystallographic analysis of the catalytic domain of the extracellular cellulase CBHI from *Trichoderma harzianum*

The filamentous fungus *Trichoderma harzianum* has a considerable cellulolytic activity that is mediated by a complex of enzymes which are essential for the hydrolysis of microcrystalline cellulose. These enzymes were produced by the induction of *T. harzianum* with microcrystalline cellulose (Avicel) under submerged fermentation in a bioreactor. The catalytic core domain (CCD) of cellobiohydrolase I (CBHI) was purified from the extracellular extracts and submitted to robotic crystallization. Diffraction-quality CBHI CCD crystals were grown and an X-ray diffraction data set was collected under cryogenic conditions using a synchrotron-radiation source.

1. Introduction

A variety of microorganisms degrade cellulose in nature. The enzymatic process that is involved in the hydrolysis of cellulosic materials has been studied in depth for several filamentous fungi, such as, for example, *Trichoderma reesei*. It is widely accepted that this hydrolytic process is accomplished by several cellulases through a series of reactions that are mediated by exoglucanases (cellobiohydrolases I and II), endoglucanases and β -glucosidases (Kubicek *et al.*, 1993). These enzymes are of considerable biotechnological interest, while the reaction conditions and the production costs of these enzymes significantly influence their range of application (Zhou *et al.*, 2008). Cellulases have been used in waste recycling and in the processing of cellulose-rich raw materials in the food, detergent, paper and textile industries (Bhat, 2000). Recently, the application of cellulases that have potential in the production of biofuels has gained significant importance.

Of the cellulases secreted by *T. reesei*, cellobiohydrolases (CBHs) are the most abundant enzymes, comprising roughly 60% of the cellulases produced by this fungus. These are key enzymes for the degradation of crystalline cellulose and the absence of CBHs reduces the rate of cellulose hydrolysis by as much as 70% (Lynd *et al.*, 2002). In general, the isoforms CBHI and CBHII act on the reducing and nonreducing ends of the cellulose chain, respectively, cleaving the internal glycoside bonds to release the disaccharide cellobiose (Hui *et al.*, 2002). These enzymes are known to be glycosylated proteins with acidic pIs (Medve *et al.*, 1998). Structurally, CBHI is a two-domain enzyme with a tertiary structure arrangement that is characteristic of glycosyl hydrolases (GHs) of family 7 (Coutinho & Henrissat, 1999). CBHs consist of a carbohydrate-binding module (CBM) and a catalytic core domain (CCD) separated by a heavily glycosylated peptide linker that is rich in prolines, serines and threonines (Gruno *et al.*, 2004). Cellulose-binding modules (CBMs) are found in most cellulases (Linder & Teeri, 1996) and play an important role in the interaction of GHs with the insoluble cellulose (Palonen *et al.*, 1999), which is crucial for the initiation and processivity of exoglucanases (Teeri *et al.*, 1998). The activity of the enzyme towards insoluble polymeric substrate in the absence of the CBM is frequently considerably lower than that of the full-length protein (Mattinen *et al.*, 1997). Structural studies have demonstrated that the active site of the CBHI from *T. reesei* is centrally located and is formed mainly by a

© 2010 International Union of Crystallography
All rights reserved

large antiparallel β -sandwich followed by a 50 Å long cellulose-binding tunnel composed of four surface loops. As the cellulose chain passes through the tunnel the hydrolytic product cellobiose is formed (Divne *et al.*, 1994).

T. harzianum is a filamentous fungus that secretes a cellulolytic complex that is able to efficiently hydrolyze a range of substrates to their reducing sugars (Saddler *et al.*, 1985). With respect to the three major groups of cellulases, *T. harzianum* displays well balanced β -glucosidase, endoglucanase and exoglucanase activities, efficiently hydrolyzing cellulolytic substrates into monomeric glucose (de Castro *et al.*, 2010). Sequencing of the gene for *T. harzianum* exoglucanase type I (CBHI) has shown that the enzyme is closely related to *T. reesei* CBHI (81% sequence identity; Guilfoile *et al.*, 1999). To date, little is known about the three-dimensional structure of CBHI from *T. harzianum*. In order to structurally characterize orthologues of *T. reesei* CBHI and to better understand possible differences in the mechanisms of action of cellobiohydrolases, we have purified and crystallized the catalytic domain of CBHI from *T. harzianum* for X-ray diffraction analysis.

2. Experimental methods

2.1. Protein production

The enzymes of the cellulolytic complex were produced by submerged fermentation of *T. harzianum* in a 7.5 l BioFlo 110 bioreactor (New Brunswick Scientific) with an operating volume of 5 l. The culture was maintained under 50% oxygen saturation by aeration and agitation at a temperature of 298 K for 120 h. The fermentation medium was prepared as described previously (Mandels *et al.*, 1976), with microcrystalline cellulose (Avicel) as the carbon source to induce the expression of cellulases.

2.2. Protein isolation and purification

Full-length CBHI (amino acids 1–505, with a theoretical molecular weight of 53 kDa) was purified from a 5 l culture of *T. harzianum*. On the fifth day the culture was harvested for 20 min at 8000g and 277 K and the clarified extract was submitted to filtration and concentration using hollow-fibre technology. The extract was then applied onto a Q-Sepharose (GE Healthcare) column equilibrated with 50 mM Tris buffer pH 7.0 and eluted with a linear gradient of 0–1 M sodium chloride in the same buffer. The CBHI-containing fractions were concentrated with a Vivaspin concentrator and further purified by hydrophobic interaction chromatography on a Phenyl Sepharose column (GE Healthcare) equilibrated with 50 mM Tris buffer pH 7.0 and eluted with a linear gradient of 0–1 M ammonium sulfate in the same buffer. The protein fractions were monitored for exoglucanase activity using 1% (w/v) Avicel in 50 mM sodium citrate pH 5.0 as a substrate (Ghose, 1987) and the 3,5-dinitrosalicylic acid (DNS) method for quantification of the produced reducing sugars (Tomme *et al.*, 1988). Finally, the fractions containing CBHI activity were pooled and concentrated. The Avicelase activity of pure CBHI was determined according to the procedures of the IUPAC Commission on Biotechnology (Ghose, 1987) using cellobiose as a standard curve.

As the next step, the CBHI was submitted to photolytic cleavage by the addition of papain at a protein:enzyme ratio of 50:1 (w:w). The reaction was carried out in 40 mM potassium phosphate buffer pH 6.0 at room temperature for 30 min. Size-exclusion chromatography was applied for further purification of the catalytic core domain using a Superdex 75 10/30 column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 7.0 and 300 mM sodium chloride. The eluted protein was dialyzed overnight against 50 mM Tris-HCl pH 7.0 and con-

centrated prior to crystallization. Protein quantification was carried out by the method of Bradford (1976) using bovine serum albumin (BSA) as a standard.

2.3. Crystallization

The purified catalytic core domain of CBHI (46 kDa; amino acids 18–449) at a concentration of 8 mg ml⁻¹ was submitted to crystallization screening using the sitting-drop vapour-diffusion technique. Drops of 2 µl final volume (1:1 ratio of protein and screen solutions) were set up automatically with a Honeybee 931 crystallization robot (Genomic Solutions Inc.) using a variety of commercially available screens (Qiagen) and maintained at a temperature of 291 K. Needles grew using 35% PEG 4000 as a precipitant. Crystal-growth optimization trials were carried out in both hanging-drop and sitting-drop plates (24-well Linbro plates).

2.4. Data collection and processing

The cluster of needle-like CBHI CCD crystals was manipulated using an acupuncture needle and transferred to a cryosolution containing a mixture of 35% PEG 4000 and 20% PEG 400. One single needle-like protein crystal was mounted in a cryoloop and directly flash-frozen in a nitrogen stream prior to X-ray data analysis. The diffraction data were collected on beamline MX2 at the Synchrotron Light Source Laboratory (LNLS), Campinas, Brazil using a MAR charge-coupled device detector (Guimarães *et al.*, 2009). Data covering 180° were collected using the oscillation method at a wavelength of 1.46 Å. Data integration was carried out using the program XDS (Kabsch, 2010) and data were scaled with the program SCALA (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

3.1. Enzyme purification and crystallization

As expected, CBHI comprised the majority of the protein that was produced in the extracellular extract of *T. harzianum* (Fig. 1, lane 1). Highly pure full-length CBHI could be obtained after the second purification step *via* hydrophobic interaction chromatography (Fig. 1, lane 3). Its enzymatic activity was determined to be 0.3 U mg⁻¹ on Avicel substrate at pH 5.0 and 323 K. This value is at the upper limit or just slightly above the reported Avicelase activity values obtained for *T. reesei* CBHI, which range from 0.014 to 0.26 U mg⁻¹ depending

Figure 1
SDS-PAGE gel. Lane 1, supernatant of *T. harzianum* culture broth; lane 2, full-length CBHI eluted from the anion-exchange column; lane 3, full-length CBHI eluted from the hydrophobic interaction column; lane 4, CBHI CCD after cleavage with papain and size-exclusion chromatography purification. Lane M contains protein markers (labelled in kDa).

Table 1

Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

Beamline	MX2, LNLS
Wavelength (Å)	1.46
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 56.9, b = 69.3, c = 90.7$
Resolution range (Å)	55.1–2.9 (3.0–2.9)
No. of unique reflections	8493 (1176)
Mosaicity (°)	0.5
Redundancy	7.0 (6.7)
Completeness (%)	99.6 (99.6)
$R_{\text{merge}}^{\dagger}$ (%)	11.9 (29.5)
$\langle I/\sigma(I) \rangle$	6.5 (2.8)

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity of an individual reflection and $\langle I(hkl) \rangle$ is the average intensity of that reflection.

on the initial substrate concentration and the experimental conditions (Tomme *et al.*, 1988; Elgogary *et al.*, 1989; Vehmaanpera *et al.*, 2009). Following protein cleavage with papain, the CCD of CBHI was successfully separated from the CBM by size-exclusion chromatography (Fig. 1, lane 4). According to the size-exclusion chromatography elution volume and SDS–PAGE, the CBHI CCD exhibits a molecular weight of about 50 kDa, which is close to the sequence-based theoretical molecular weight of the CBHI CCD (46 kDa).

The CBHI CCD was crystallized using a crystallization robot and several commercial crystallizations kits covering more than 650 different conditions. Crystals only grew in only one condition, which contained 35% PEG 4000 as a precipitant. Small needles appeared very rapidly using this crystallization condition, so that the needle cluster was completely formed after 10 h of incubation time (Fig. 2). Various attempts were carried out to attempt to reduce the number of nuclei being formed and to favour single-crystal growth. Variations in the pH and the protein and precipitant concentrations and tests of different additives were ineffective, resulting exclusively in a shower of microcrystals or clusters of needles. Needles grown in clusters were tested for X-ray diffraction quality. Surprisingly, a small single needle-like crystal of approximate dimensions $10 \times 5 \times 5 \mu\text{m}$ showed a diffraction pattern that extended to a resolution of about 2.9 Å (Fig. 3). A complete native data set was collected on the dedicated wiggler beamline MX2 (LNLS, Brazil; Guimarães *et al.*, 2009). The data set was reduced in the orthorhombic space group $P222$, with unit-cell parameters $a = 56.9, b = 69.3, c = 90.7$ Å. Analysis of the systematic absences revealed the presence of twofold screw symmetry along all three unit-cell axes. According to the Matthews

Figure 2

Crystals of CBHI CCD grew as clusters in hanging drops using 35% (w/v) PEG 4000 as precipitant. The black bar corresponds to 10 μm .

coefficient of $1.77 \text{ Å}^3 \text{ Da}^{-1}$ (Matthews, 1968), there is only one molecule in the asymmetric unit and the solvent content of the crystal is 31%. Data-collection and processing statistics are shown in Table 1.

4. Conclusions

The cellulase CBHI was successfully purified from the soluble extracellular fraction of *T. harzianum* cultured in a bioreactor. The catalytic core domain of the enzyme was submitted to crystallization and diffraction-quality crystals were obtained. An X-ray diffraction data set from a single needle-like CBHI catalytic core domain crystal was collected using a synchrotron source.

(a)

(b)

Figure 3

(a) Diffraction pattern collected from one single needle-shaped crystal extending to a resolution of 2.9 Å. (b) Close-up view of the diffraction image with enhanced contrast.

We are grateful to the MX2 staff members for beamline support and to the Brazilian Synchrotron Light Source (LNLS). This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) via research grant Nos. 08/56255-9, 2007/08706-9 and 2009/05349-6, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) via INCT do Bioetanol and grant No. 471834/2009-2.

References

- Bhat, M. K. (2000). *Biotechnol. Adv.* **18**, 355–383.
- Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
- Castro, A. M. de, Pedro, K. C., Cruz, J. C., Ferreira, M. C., Leite, S. G. & Pereira, N. Jr (2010). *Appl. Biochem. Biotechnol.* doi:10.1007/s12010-010-8986-0.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Coutinho, P. M. & Henrissat, B. (1999). *Recent Advances in Carbohydrate Bioengineering*, edited by H. J. Gilbert, G. J. Davies, B. Henrissat & B. Svensson, pp. 3–12. Cambridge: Royal Society of Chemistry.
- Divne, C., Stahlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K. C., Teeri, T. T. & Jones, T. A. (1994). *Science*, **265**, 524–528.
- Elgogary, S., Leite, A., Crivellaro, O., Eveleigh, D. E. & Eldorri, H. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 6138–6141.
- Ghose, T. K. (1987). *Pure Appl. Chem.* **59**, 257–268.
- Gruno, M., Valjamae, P., Pettersson, G. & Johansson, G. (2004). *Biotechnol. Bioeng.* **86**, 503–511.
- Guilfoile, P., Burns, R. G., Zu-Yi, Amundson, M. & Chang, F.-H. (1999). *J. Minn. Acad. Sci.* **64**, 18–22.
- Guimarães, B. G., Sanfelici, L., Neuenschwander, R. T., Rodrigues, F., Grizolli, W. C., Raulik, M. A., Piton, J. R., Meyer, B. C., Nascimento, A. S. & Polikarpov, I. (2009). *J. Synchrotron Rad.* **16**, 69–75.
- Hui, J. P. M., White, T. C. & Thibault, P. (2002). *Glycobiology*, **12**, 837–849.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kubicek, C. P., Messner, R., Gruber, F., Mach, R. L. & Kubicekpranz, E. M. (1993). *Enzyme Microb. Technol.* **15**, 90–99.
- Linder, M. & Teeri, T. T. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 12251–12255.
- Lynd, L. R., Weimer, P. J., van Zyl, W. H. & Pretorius, I. S. (2002). *Microbiol. Mol. Biol. Rev.* **66**, 506–577.
- Mandels, M., Andreotti, R. & Roche, C. (1976). *Biotechnol. Bioeng. Symp.* **6**, 21–33.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Mattinen, M. L., Kontteli, M., Kerovuori, J., Linder, M., Annala, A., Lindeberg, G., Reinikainen, T. & Drakenberg, T. (1997). *Protein Sci.* **6**, 294–303.
- Medve, J., Lee, D. & Tjerneld, F. (1998). *J. Chromatogr. A*, **808**, 153–165.
- Palonen, H., Tenkanen, M. & Linder, M. (1999). *Appl. Environ. Microbiol.* **65**, 5229–5233.
- Saddler, J. N., Hogan, C. M. & Louisseize, G. (1985). *Appl. Environ. Microbiol.* **22**, 139–145.
- Teeri, T. T., Koivula, A., Linder, M., Wohlfahrt, G., Divne, C. & Jones, T. A. (1998). *Biochem. Soc. Trans.* **26**, 173–178.
- Tomme, P., Vantilbeurgh, H., Pettersson, G., Vandamme, J., Vandekerckhove, J., Knowles, J., Teeri, T. & Claeyssens, M. (1988). *Eur. J. Biochem.* **170**, 575–581.
- Vehmaanpera, J., Alapuranen, M., Puranen, T., Siika-aho, M., Kallio, J., Hooman, S., Voutilainen, S., Halonen, T. & Viikari, L. (2009). US Patent 20090042266.
- Zhou, J., Wang, Y.-H., Chu, J., Zhuang, Y.-P., Zhang, S.-L. & Yin, P. (2008). *Bioresour. Technol.* **99**, 6826–6833.

Purification and Characterization of a Thermostable Cellobiohydrolase from *Fomitopsis pinicola*

Shin, Keum¹, Yoon Hee Kim¹, Marimuthu Jeya², Jung-Kul Lee^{2*}, and Yeong-Suk Kim^{1*}

¹Department of Forest Products, Kookmin University, Seoul 136-702, Korea

²Department of Chemical Engineering, Konkuk University, Seoul 143-701, Korea

Received: August 9, 2010 / Revised: September 12, 2010 / Accepted: September 13, 2010

A screening for cellobiohydrolase (CBH) activity was performed and *Fomitopsis pinicola* KMJ812 was selected for further characterization as it produced a high level of CBH activity. An extracellular CBH was purified to homogeneity by sequential chromatography of *F. pinicola* culture supernatants. The molecular mass of the *F. pinicola* CBH was determined to be 64 kDa by SDS–PAGE and by size-exclusion chromatography, indicating that the enzyme is a monomer. The *F. pinicola* CBH showed a $t_{1/2}$ value of 42 h at 70°C and catalytic efficiency of $15.8 \text{ mM}^{-1} \text{ s}^{-1}$ (k_{cat}/K_m) for *p*-nitrophenyl- β -D-cellobioside, one of the highest levels seen for CBH-producing microorganisms. Its internal amino acid sequences showed a significant homology with hydrolases from glycoside hydrolase family 7. Although CBHs have been purified and characterized from other sources, the *F. pinicola* CBH is distinguished from other CBHs by its high catalytic efficiency and thermostability.

Keywords: Cellobiohydrolase, enzyme production, *Fomitopsis pinicola*, glycoside hydrolase, purification, stability

Cellulose, a main component of plant cell wall, represents the most abundant renewable biomass available on Earth. It is a linear biopolymer composed of 100–10,000 β -D-glucopyranosyl units linked by β -1,4-D-glucosidic bonds. Microbial cellulases catalyzing the hydrolysis of plant polysaccharides are industrially important enzymes used to saccharify industrial and agricultural cellulose-containing residues, treat cellulose pulp wastes in the paper industry, enhance the extraction of fermentable substances in the beer brewing and alcohol fermentation industries, etc. [33]. The cellulose degradation requires a multienzymatic system

composed of three activities: endo-1,4- β -glucanase (EG, E.C. 3.2.1.4), cellobiohydrolase (CBH, E.C. 3.2.1.91), and β -glucosidase (BGL, E.C. 3.2.1.21) [3]. EG and CBH act cooperatively and synergistically in depolymerizing cellulose to cellobiose and short oligosaccharides, which are converted by BGL to glucose [6]. CBHs are most efficient on highly ordered crystalline cellulose and cleave mainly cellobiose from the opposite ends of the glucose chains, whereas EGs act randomly in the middle of the chains, probably in the more amorphous regions of cellulose [37].

In the process of cellulose hydrolysis, enzyme production is still the most crucial and costly step. Filamentous fungi are the major source of cellulases and hemicellulases. Cellobiohydrolases (CBHs) are important components in the multienzyme cellulase complexes [36] and they display an exo-type of attack on polymeric substrates, and the major product of their action on cellulose is cellobiose. They are classified into three glycoside hydrolase families (GH6, GH7, and GH48) according to amino acid sequence similarity [14, 16–18, 24]. Of these families, only GH7 is thought to be exclusively of fungal origin, and this family contains CBH I cellobiohydrolases and EG I endoglucanases from both ascomycete and basidiomycete fungi [10]. Potent cellulolytic fungi generally produce two different CBHs, CBH I and CBH II. These two types of enzymes, which are classified based on sequence identity, have extended tunnel-shaped active sites and can achieve, although slowly, complete solubilization of cellulose crystals even without help of endoglucanases [36].

Brown-rot fungi primarily utilize the cellulose and hemicellulose components of wood biomass and then rapidly depolymerize the cellulose without removing the surrounding lignin that normally prevents microbial attack [19]. In the present study, a potent CBH-producing fungus, *Fomitopsis pinicola*, was screened. Under optimized conditions, *F. pinicola* produced CBH with a 4.3 U/mg-protein specific activity. Here, we purified a novel CBH from the isolated *F. pinicola* KMJ812 (KACC 93064P) to

*Corresponding author

J.-K.L.

Phone: +82-2-450-3505; Fax: +82-2-458-3504;

E-mail: jkrhee@konkuk.ac.kr

Y.-S.K.

Phone: +82-2-910-4822; Fax: +82-2-919-5901;

E-mail: yskim@kookmin.ac.kr

homogeneity and characterized its physiological and kinetic parameters. The properties of the enzyme, including its substrate specificity, molecular form, inhibition by product cellobiose, and partial amino acid sequence revealed that this enzyme is a CBH, a member of GH7.

MATERIALS AND METHODS

Screening of CBH-Producing Strain

All fungal strains were obtained from the Korean Agricultural Culture Collection (KACC). Initial screening of CBH-producing fungi was carried out on agar plates containing 10 mM 4-methylumbelliferyl cellobioside. Based on the fluorescence observed, 15 strains were inoculated into 3 ml of the growth medium containing (g/l) peptone (8), yeast extract (2), KH_2PO_4 (5), K_2HPO_4 (5), $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ (3), Thiamine-HCl (0.005), and microcrystalline cellulose (Sigma, MO, USA), and cultivated at 28°C with agitation at 200 rpm for 7 days. The CBH activity of the culture broth was analyzed using *p*-nitrophenyl- β -D-cellobioside (pNPC; Sigma, MO, USA) as described previously [8]. After analyses, the strain with the highest CBH activity was selected.

Culture Conditions

For flask culture, the mycelia of *F. pinicola* KMJ812 were inoculated into 100 ml of potato dextrose broth. Precultures (5 ml) were inoculated into 200 ml of cellulolytic medium in a 1-l fermenter. The medium was the same as the growth medium in the screening step. The effect of carbon or nitrogen source on CBH production was investigated after 17 days of cultivation in flasks containing medium composed of 50 g/l of carbon sources and various nitrogen sources. The concentration of nitrogen source was adjusted to the same content of nitrogen using the Kjeldahl method [5]. The medium used for CBH production was the same as the growth medium except that rice straw was used instead of crystalline cellulose.

Enzyme Assay

The CBH activity was assayed using pNPC (Sigma, MO, USA) as substrate. The enzymatic reaction mixtures (1 ml) containing 100 μ l of enzyme solution and 10 mM pNPC (final concentration) in 100 mM sodium acetate buffer (pH 5.0) were incubated for 15 min at 50°C [8]. The amount of *p*-nitrophenol released was measured at 405 nm after addition of 2 M Na_2CO_3 to the reaction mixtures [8]. One unit of pNPC-hydrolyzing activity was defined as the amount of enzyme equivalent to release the of 1 mmol of *p*-nitrophenol per minute.

Purification of Cellobiohydrolase

All procedures were performed at 4°C, and 50 mM potassium phosphate buffer (pH 7.0) containing 1 mM dithiothreitol (DTT) was used in the purification procedures unless otherwise stated. Protein estimations were performed by the Bradford method [4] using bovine serum albumin as a standard. Proteins in the column effluents were monitored by measuring the absorbance at 280 nm. All chromatographic separations were performed using an ÄKTA FPLC system (ÄKTA, Sweden).

Step 1: Preparation of crude enzyme. Cells from the culture broth were harvested by centrifugation at 10,000 $\times g$ for 30 min. After

washing with 20 mM sodium acetate buffer (pH 5.0), the washes and supernatants were combined, concentrated, and desalted by ultrafiltration through a polyether sulfone membrane (30 kDa cutoff) in a stirred cell (Amicon, Beverly, MA, USA).

Step 2: DEAE Sepharose chromatography. The dialyzed enzyme solution was loaded on a DEAE Sepharose Fast Flow column (1.6 $\times$ 10 cm; Amersham Biosciences, Uppsala, Sweden) equilibrated with 20 mM sodium acetate buffer (pH 5.0), and proteins were eluted with a 180-min linear gradient of 0–0.5 M NaCl in the same buffer at a flow rate of 1.0 ml/min. Fractions of 1 ml each were collected and assayed for CBH activity. Active fractions were pooled, dialyzed against the same buffer, and concentrated by ultrafiltration for further purification.

Step 3: Sephacryl gel filtration chromatography. The concentrated enzyme solution was loaded on a HiPrep 16/60 Sephacryl S-300 HR column (1.0 cm $\times$ 120 cm; Amersham Biosciences, Uppsala, Sweden) equilibrated with 20 mM sodium acetate buffer containing 100 mM NaCl at pH 5.0, and proteins were eluted with the same buffer at a flow rate of 0.5 ml/min. Active fractions were pooled, dialyzed against the same buffer, and concentrated by ultrafiltration.

Step 4: MonoQ ion-exchange chromatography. The enzyme was further purified with a MonoQ ion-exchange column 5/50 GL (1.0 $\times$ 10 cm; Amersham Biosciences, Uppsala, Sweden) previously equilibrated with 20 mM sodium acetate buffer (pH 5.0). The enzyme was eluted with a 180-min linear gradient of 0–0.5 M NaCl in the same buffer at a flow rate of 0.5 ml/min. The combined active fractions were pooled, concentrated, and dialyzed against the same buffer and concentrated with a Centricon (Millipore, Bedford, MA, USA) ultrafiltration device with a molecular mass cutoff of 30 kDa, and then used as a purified enzyme in the following experiments.

Determination of pH and Temperature Optima

The optimal pH of CBH activity was determined by incubating the purified enzyme at 50°C for 15 min in different buffers: citrate (100 mM, pH 3–4.5), sodium acetate (100 mM, pH 4.5–5.5), and phosphate (100 mM, pH 5.5–8). To determine the optimal temperature, the enzyme was incubated in sodium acetate buffer (100 mM, pH 5) for 15 min at different temperatures (30–80°C). To determine the thermostability of CBH activity, the purified enzyme was incubated at different temperatures (30°C–70°C) in the absence of substrate. After keeping them for certain periods of time (0–25 days), the residual CBH activity was determined as described above.

PAGE and Molecular Mass Determination

For the determination of subunit molecular mass, sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS–PAGE) was performed as described previously [22], with 10% gels. Protein bands were visualized by staining with Coomassie brilliant blue R-250 (Sigma, MO, USA). The molecular mass of the purified enzyme was determined by size-exclusion chromatography using a Superose 12 (Amersham Pharmacia Biotech, Uppsala, Sweden) column attached to an ÄKTA FPLC system (GE Health Care). The enzyme was eluted with 20 mM sodium acetate (pH 5.0) at a flow rate of 0.5 ml/min.

Substrate Specificity

The substrate specificity of CBH was determined by using pNPC, *p*-nitrophenyl- β -D-glucopyranoside (pNPG), *p*NP- β -galactopyranoside (pNPgal), *p*NP- β -D-mannopyranoside (pNPM), *p*NP- β -D-xylopyranoside (pNPX), *p*NP- β -D-lactopyranoside (pNPL), cellobiose, cellotriose,

cellotetraose, and cellopentaose as substrates at 10 mM concentration. CBH activities on polysaccharides carboxymethylcellulose (CMC), xylan, lichenan, laminarin, and avicel (1%) were also tested. The *p*-nitrophenol released was determined under standard enzyme assay conditions. The activities on oligosaccharides were estimated by assaying the amount of released glucose using the GOD-POD method [27]. The enzyme activity on polysaccharides was estimated by measuring released reducing sugars by the DNS method [29] using glucose or xylose as standards.

Determination of Kinetic Parameters and Inhibition Constants

The values of the Michaelis constant (K_m) and the maximum velocity (V_{max}) were determined for CBH by incubating in 100 mM sodium acetate buffer (pH 5) at 50°C with pNPC at concentrations ranging from 0.5 to 50 mM. Inhibition of CBH by cellobiose was determined in the presence of pNPC as the substrate. Values for K_m , V_{max} , and K_i were determined from Lineweaver–Burk plots using standard linear regression techniques.

Effects of Metals and Reagents

The effects of various metal ions ($BaCl_2$, $MgCl_2$, $MnCl_2$, $ZnCl_2$, $CaCl_2$, $NiCl_2$, or $CoCl_2$) and reagents (2-mercaptoethanol and cysteine) at 1 mM on CBH activity were determined by preincubating the enzyme with the individual reagents in 20 mM sodium acetate buffer (pH 5.0) at 30°C for 30 min. Activities were then measured at 50°C for 15 min in the presence of the metal ions or reagents. The activity assayed in the absence of metal ions or reagents was recorded as 100%.

Internal Amino Acid Sequence of CBH

Protein cleavage for peptide mapping was carried out at 37°C for 4 h with 100 ng of endoproteinase Asp-N or endoproteinase Lys-C or trypsin (Promega, Madison, WI, USA) to digest 20 µg of purified enzyme in 50 µl of 100 mM $(NH_4)_2CO_3$ (pH 8.5). The resulting peptide fragments were separated by SDS–PAGE (15% polyacrylamide), and the separated peptides were transferred to a polyvinylidene fluoride (PVDF) membrane (Bio-Rad). Peptide bands were visualized by 0.1% Coomassie brilliant blue R-250 staining in 40% methanol. The partial amino acid sequence was determined by Edman degradation with an automatic protein sequencer (model 491A; Applied Biosystems, Division of Perkin-Elmer) at The National Instrumentation Center for Environmental Management (Seoul, Korea). The partial amino acid sequence was used to identify analogous proteins through a BLAST search of the nonredundant protein database from the National Center for Biotechnology Information. A protein domain similarity search was performed using the CDD (Conserved Domain Database) at <http://www.ncbi.nlm.nih.gov/Structure/cdd>.

RESULTS

Screening and Selection of CBH-Producing Strain

Among 112 fungal strains screened for CBH activity, 15 strains were selected based on the fluorescence observed when the agar plates containing 10 mM of MUC were exposed under UV. Out of 15 strains, an efficient CBH-producing *Fomitopsis pinicola* KMJ812 was selected for further study.

Optimization of Carbon and Nitrogen Sources for CBH Production

To select a suitable carbon source for CBH production, *F. pinicola* KMJ812 was cultivated in media with yeast extract (10 g/l) and varied carbon sources (cellulose, rice straw, wheat bran, xylan, Avicel, CMC, cellobiose, glucose, maltose, lactose, sucrose, or wood fiber). Among the carbon sources tested, rice straw was found to be the best carbon source for CBH production, leading to a CBH-specific activity of 4.1 U/mg-protein in a flask culture (Table 1).

Since the mechanisms that govern the formation of extracellular enzymes are influenced by the availability of precursors for protein synthesis, the effect of inorganic and organic nitrogen sources on CBH synthesis was also studied (Table 1). Rice straw (20 g/l) was used as a carbon source. Among the various nitrogen sources (peptone, corn steep powder, yeast extract, urea, ammonium sulfate, potassium nitrate, and sodium nitrate), a combination of yeast extract (5 g/l) and peptone (5 g/l) favored maximum CBH production (4.6 U/mg-protein), followed by yeast extract, whereas the other nitrogen sources were poor sources of nitrogen in the absence of pH control. The time

Table 1. Effects of various carbon and nitrogen sources on CBH production.

Carbon source (20 g/l)	CBH activity (U/mg-protein)	Total protein (mg)	pH
Cellulose	4.0±0.27	4.0±0.22	3.92
Glucose	0.4±0.04	0.9±0.11	2.57
Lactose	2.8±0.13	2.3±0.23	2.95
Maltose	1.4±0.13	3.3±0.14	2.72
Cellobiose	0.8±0.11	3.1±0.32	2.63
CMC	2.2±0.21	2.7±0.18	3.90
Xylan	1.3±0.07	4.8±0.11	3.99
Rice straw	4.1±0.32	5.3±0.32	3.71
Avicel	2.9±0.10	3.4±0.12	3.83
Sucrose	3.7±0.28	3.5±0.21	3.36
Softwood powder	1.2±0.04	3.03±0.31	5.67
Hardwood powder	3.9±0.21	4.55±0.12	3.77
Nitrogen sources (10 g/l)	CBH activity (U/mg-protein)	Total protein (mg)	pH
Yeast extract+peptone	4.3±0.34	5.5±0.42	3.92
Yeast extract	4.1±0.32	5.3±0.32	3.71
Peptone	1.8±0.10	3.7±0.15	3.95
Corn steep powder	0.9±0.06	2.4±0.17	4.01
Urea	1.7±0.11	2.9±0.24	7.48
$(NH_4)_2SO_4$	1.5±0.08	1.1±0.04	5.05
KNO_3	0.1±0.01	0.6±0.02	4.86
$NaNO_3$	0.3±0.02	0.9±0.07	3.37

All flasks were incubated at 25°C and 150 rpm for 15 days. Each value represents the mean of triplicate measurements and varied from the mean by not more than 15%.

Fig. 1. Time course of cellobiohydrolase production by *F. pinicola* KMJ812 on rice straw. Filled circles, specific activity; empty circles, pH; filled triangles, residual reducing sugar level.

course of CBH production by *F. pinicola* grown on rice straw, yeast extract, and peptone was studied (Fig. 1). CBH activity increased up to 17 days, after which it decreased gradually.

Purification of a Cellobiohydrolase

CBH was purified as described in the Materials and Methods section, and the results are summarized in Table 2. Fractionation by ultrafiltration increased the specific activity about 2-fold, with 32% recovery of CBH activity. The active fractions were applied to a DEAE Sepharose column, and CBH was eluted with approximately 0.1 M NaCl. A subsequent gel filtration step produced three protein peaks of which the second peak showed CBH activity. FPLC elution of a MonoQ ion-exchange chromatography column with 0.5 M NaCl produced an active CBH protein peak. These chromatography methods resulted in a 7-fold purification of CBH with a recovery of 0.1%. Analysis of the purified enzyme by gel electrophoresis in the presence of SDS (Fig. 2A, lane 4) revealed one band with an M_r of $64,000 \pm 1,000$. Size-exclusion chromatography on a Sephacryl S-300 high resolution column resulted in the elution of the enzyme activity as a symmetrical peak corresponding to an M_r of

Fig. 2. PAGE and determination of molecular mass of CBH purified from the *F. pinicola*.

A. PAGE of CBH purified from *F. pinicola*. Lane 1, molecular marker; lane 2, cell extract; lane 3, DEAE ion-exchange fraction; lane 4, SDS-PAGE of MonoQ ion-exchange fraction. **B.** Determination of native molecular mass of *F. pinicola* CBH by gel filtration chromatography on a Sephacryl S-300 high-resolution column. The column was calibrated with standard molecular mass proteins such as aldolase (168 kDa), ovalbumin (43 kDa), chymotrypsinogen (25 kDa), and ribonuclease A (13.7 kDa).

approximately 64,000 (Fig. 2B). These results indicate that the enzyme migrates as a monomer in gel filtration under the mild conditions used, and thus may also be present active as a monomer in solution.

Identification of the Partial Peptide Fragment

The pure enzyme (1.5 μ g) was separated by 10% SDS-PAGE and blotted onto a PVDF membrane. Automated

Table 2. Purification of CBH from the culture broth of *F. pinicola* KMJ812.

Procedure	Total protein (mg)	Total activity (U)	Specific activity (U/mg-protein)	Yield (%)	Purification fold
Cell extract	1,200	4,590	4	100	1
Ultrafiltration (PES30)	383	2,240	6	32	1.6
Salting out	195	1,760	9	16	2.3
DEAE ion exchange chromatography	76	988	13	6	3.4
Gel filtration-chromatography	10	149	15	1	3.9
MonoQ ion-exchange chromatography	1.3	34	26	0.1	6.8

Edman degradation of the enzyme protein was unsuccessful, implying that the N-terminal of the enzyme was blocked. The CBH was partially digested with trypsin, endoproteinase Asp-N, and endoproteinase Lys-C, and then separated by 12.5% SDS-PAGE and blotted onto a PVDF membrane. Three fragments, a Lys-C fragment (LYS), an Asp-N fragment (ASP), and a trypsin fragment (TRY), were sequenced on an automatic protein sequencer. The LYS fragment contained a RFLEQYDK segment. The TRY and ASP fragments contained GGLEAMGESLDR and TYLMQD, respectively.

Optimum pH and Temperature

The optimum pH for the CBH was 5.0, with 90% and 93% of the maximum activity appearing at pH 4.5 and 5.5,

Fig. 3. Effects of pH and temperature on the activity of *F. pinicola* CBH.

A. Effect of pH on the activity of purified CBH from *F. pinicola* KMJ812. The enzyme activity was assayed by the standard assay method by changing the buffer to obtain the desired pHs. The buffers used were citrate (pH 3.0 to 4.5), sodium acetate (pH 4.5 to 5.5), and phosphate (pH 5.5 to 8.0). **B.** Effect of temperature on the activity of purified CBH from *F. pinicola* KMJ812. The enzyme activity was assayed at various temperatures by the standard assay method. Each value represents the mean of triplicate measurements and varied from the mean by not more than 15%.

respectively (Fig. 3A). An acidic pH optimum and maximal activity at about pH 4.5 are common features of similar CBH enzymes isolated from diverse microbial systems [25]. The isoelectric point, pI, was determined to be 5.1, which is typical for extracellular CBHs [1]. The optimum temperature for the hydrolysis reaction was 50°C with 80% and 95% of the maximum activity at 45 and 55°C, respectively (Fig. 3B).

Thermostability of *F. pinicola* CBH

The stability of purified CBH was studied at various temperatures from 40°C to 80°C. According to the thermostability assay, the purified *F. pinicola* CBH was highly stable and maintained ~90% activity when incubated at 60°C for 24 h. About 90% activity was maintained after 8 h of incubation at 70°C. At temperatures over 80°C, the CBH activity sharply decreased depending on the incubation time. The enzyme showed $t_{1/2}$ values of 552 h, 144 h, 96 h, and 42 h at 40°C, 50°C, 60°C, and 70°C, respectively (Fig. 4).

Substrate Specificity and Kinetic Parameters of *F. pinicola* CBH

The activities of *F. pinicola* CBH with various substrates are shown in Table 3. Whereas pNPL, Avicel, and lichenan showed an activity of 42.1%, 29.6%, and 33.8%, respectively, the other compounds did not serve as substrates. *F. pinicola* CBH was almost completely inactive with polymeric cellulose and xylan, but had a high preference for pNPC.

Initial velocities were determined in the standard assay mixture at pH 5. All the substrates tested had hyperbolic saturation curves, and the corresponding double-reciprocal plots were linear. The concentration of pNPC varied from

Fig. 4. Thermal inactivation of *F. pinicola* CBH.

The enzymes were incubated at 40°C (filled circles), 50°C (empty circles), 60°C (filled triangles), and 70°C (empty triangles) for varying periods of time. Samples were withdrawn at each time interval and relative activities were determined.

Table 3. Substrate specificity of CBH purified from *F. pinicola* KMJ812.

Substrate (10 mM)	CBH activity (U/mg-protein)	Relative activity (%)
<i>p</i> -Nitrophenyl- β -D-glucopyranoside	0.03	0.11
<i>p</i> -Nitrophenyl- β -D-galactopyranoside	0.07	0.26
<i>p</i> -Nitrophenyl- β -D-cellobioside	26.0	100
<i>p</i> -Nitrophenyl- β -D-mannopyranoside	0.17	0.64
<i>p</i> -Nitrophenyl- β -D-xylopyranoside	0.17	0.64
<i>p</i> -Nitrophenyl- β -D-lactopyranoside	11.2	42.1
Cellobiose	0.03	0.11
Cellobiose	0	0
Cellotetraose	0	0
1% Xylan	0.3	1.13
1% Avicel	7.89	29.6
1% CMC	0.07	0.26
1% Lichenan	9.01	33.8
1% Laminarin	0.4	1.5

The purified enzyme was assayed in the standard assay condition with various compounds. Each value represents the mean of triplicate measurements and varied from the mean by not more than 10%.

0 to 50 mM. Fig. 5 shows the typical Michaelis–Menten-type kinetics for CBH activity with increasing pNPC concentrations. The Lineweaver–Burk plot (inset in Fig. 5) obtained for the conversion of pNPC under standard assay conditions showed a K_m of 2.1 mM and a V_{max} of 31.1 mmol/min/mg. Product inhibition studies under nonsaturating conditions showed that cellobiose inhibited CBH competitively, with a K_i value of 14.8 mM.

Effects of Metal Ions and Various Compounds

CBH activity was measured in the presence of metal ions or various other compounds (data not shown). The CBH activity was not stimulated by $BaCl_2$, $MgCl_2$, $MnCl_2$, $ZnCl_2$, $CaCl_2$, $NiCl_2$, or $CoCl_2$ (each at 1 mM concentration), and it was neither inhibited nor activated by EDTA at concentrations ranging from 1 to 10 mM. In contrast, Hg^{2+} ions caused significant inhibition (60%) that appeared to be competitive with respect to the substrate pNPC. *N*-Ethylmaleimide (NEM) is often used to inactivate enzymes, presumably by reacting with the thiol group of cysteine residues. Incubation of purified CBH with 10 mM NEM resulted in an approximately 30% inhibition of CBH activity, suggesting that cysteine is at least partly responsible for CBH activity. The time required to reach half-maximal inhibition ($t_{1/2}$) with NEM was approximately 6.5 min.

DISCUSSION

Although the purification and properties of CBH from several strains have been reported, this is the first report on

Fig. 5. Effects of substrate concentration on the activities of *F. pinicola* CBH.

The CBH activity of the enzyme (1 unit) was measured in the presence of the indicated concentrations of pNPC, at pH 5.0. The inset shows a Lineweaver–Burk plot of initial velocity versus various fixed pNPC concentrations. Each value represents the mean of triplicate measurements and varied from the mean by not more than 15%.

the purification and characterization of CBH from *F. pinicola*, a basidiomycete. The present study reveals a unique *F. pinicola* CBH showing one of the highest CBH activities ever reported. Table 4 shows a comparison of the properties of various CBHs from a number of different sources [2, 7, 13, 15, 21, 26, 31, 38]. In comparison, the K_m values of purified CBHs for pNPC from other fungi range from 0.58 to 6.8. Thus, the K_m value (2.1 mM) of *F. pinicola* CBH was in agreement with those recorded for CBHs from other fungi. However, the catalytic efficiency value ($k_{cat}/K_m = 15.8 \text{ mM}^{-1}\text{s}^{-1}$) of *F. pinicola* CBH was significantly higher than other CBHs from *Trichoderma reesei* ($0.045 \text{ mM}^{-1}\text{s}^{-1}$), *Talaromyces emersonii* ($2.52 \text{ mM}^{-1}\text{s}^{-1}$), *Chrysosporium lucknowense* ($0.17 \text{ mM}^{-1}\text{s}^{-1}$), *Cellulomonas fimi* ($8.0 \text{ mM}^{-1}\text{s}^{-1}$), etc. The extracellular CBH purified from *F. pinicola* was a monomer with a molecular mass of 64 kDa. The molecular mass of this CBH is in agreement with those of many extracellular CBHs characterized from other fungal sources.

CBH 1A of *Talaromyces emersonii*, a moderately thermophilic fungus, showed a $t_{1/2}$ value of 0.57 h at 80°C [38]. *Trichoderma reesei* CBH, despite showing an optimum temperature of 70°C, had a $t_{1/2}$ at 65°C of less than 1 h [28]. The $t_{1/2}$ value (42 h at 70°C) of *F. pinicola* CBH was higher than other CBHs from *Chaetomium thermophilus* CT2 (1 h at 70°C), *Penicillium janthinellum* (0.67 h at 65 °C), *Trametes versicolor* (0.5 h at 60 °C), etc. Thus, *F. pinicola* CBH shows one of the highest thermostabilities among the CBHs isolated thus far from fungi, which suggests that it has potential industrial uses.

Table 4. Properties of CBHs from various sources.

Microorganism	M _r (kDa)	Quaternary structure	K _m (mM)	Opt. pH	Opt. temp (°C)	Half-life (t _{1/2})	Specific activity (U/mg)	Reference
<i>Dichomitus squalens</i> Ex-1	39	Monomer	NR	5.0	60	NR	13.6	[34]
<i>Irpex lacteus</i> Ex-1	53	NR	NR	5.0	50	NR	33.2	[15]
<i>Irpex lacteus</i> Ex-2	56	NR	NR	5.0	50	NR	34.0	[15]
<i>Chrysosporium lucknowense</i> CBH IA	65	Monomer	NR	4.5	NR	5 h at 50°C (>90% retained)	0.02	[13]
<i>Chrysosporium lucknowense</i> CBH IB	60	Monomer	NR	NR	NR	5 h at 50°C (>90% retained)	0.02	[12]
<i>Chrysosporium lucknowense</i> CBH IIA	43	Monomer	NR	5.5	65	5 h at 50°C (>90% retained)	0.08 ^b	[6]
<i>Trametes versicolor</i>	55	Monomer	0.58	5.0	40	0.5 h at 60°C	1.0	[23]
<i>Penicillium occitanis</i> CBH I	60	NR	1 ^a	4.0–5.0	60	0.5 h at 60°C	1.09	[26]
<i>Penicillium occitanis</i> CBHII	55	NR	5 ^a	4.0–5.0	65	0.5 h at 60°C	0.03	[26]
<i>Talaromyces emersonii</i> CBHIA	66	Monomer	2.1	3.6	78	0.57 h at 80°C	7.7	[38]
<i>Chaetomium thermophilus</i> CT2	67	Monomer	0.95	5.0	65	1 h at 70°C	NR	[28]
<i>Penicillium janthinellum</i>	57	Monomer	0.82	5.0	50	0.67 h at 65°C	0.07	[20]
<i>Coniophora puteana</i> CBH I	52	NR	6.8	5.0	40	NR	0.46	[35]
<i>Coniophora puteana</i> CBH II	50	NR	4.3	5.0	40	NR	0.4	[35]
<i>Fomitopsis pinicola</i>	64	Monomer	2.1	5.0	50	42 h at 70°C	26.0	This work

NR, not reported.

^aKinetic parameters of CBHs are shown for pNPC.^bAvicel as a substrate.

Based on amino acid sequence similarities, GHs have been classified into several families, with most CBHs belonging to either family 6, family 7, or family 48 [14, 16–18, 24]. Of these families, only GH7 is thought to be exclusively fungal, and this family contains CBH I cellobiohydrolases and EG I endoglucanases from both ascomycete and basidiomycete fungi [10]. The three-dimensional structures of GH7 CBH catalytic modules from mesophilic fungi *Trichoderma reesei* [9] and *Phanerochaete chrysosporium* [30], from a thermophilic fungus *Talaromyces emersonii* [11], and *Melanocarpus albomyces* [32] have been solved. The overall fold is a β -sandwich, where loops extending from the β -sandwich form an enclosed cellulose-binding tunnel. CBHs of *P. chrysosporium* have a GGLKQMGE sequence, similar to the *F. pinicola* fragment GGLEAMGE. The peptide fragments of *F. pinicola* CBH, RFLEQYDK and GGLEAMGESLDR, are identical to those of the *celA* (*Cellulomonas flavigena* DSM 20109) and *cel3A* (*Schizophyllum commune*) CBHs belonging to GH7, respectively.

In conclusion, a potent CBH-producing *F. pinicola* strain was identified. *F. pinicola* CBH possesses the highest thermostability and catalytic efficiency among the CBHs reported. The successful purification and characterization of CBH produced by *F. pinicola* allows us to characterize a novel CBH and now sets the stage for more detailed investigations of this enzyme, such as cloning of the full-length gene followed by protein engineering studies.

Acknowledgments

This study was supported by a grant (S210709L010110) from the Korea Forest Service. This study was also supported by a grant (PJ007449201006) from BioGreen 21 Program, Rural Development Administration, Republic of Korea.

REFERENCES

- Baldeian, P. and V. Valaskova. 2008. Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiol. Rev.* **32**: 501–521.
- Baldrian, V. P. 2006. Degradation of cellulose and hemicelluloses by the brown rot fungus *Piptoporus betulinus* – production of extracellular enzymes and characterization of the major cellulases. *Microbiology* **152**: 3613–3622.
- Bayer, E. A., H. Chanzy, R. Lamed, and Y. Shoham. 1998. Cellulose, cellulases and cellosomes. *Curr. Opin. Struct. Biol.* **8**: 548–557.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**: 248–254.
- Bradstreet, R. B. 1954. Kjeldahl method for organic nitrogen. *Anal. Chem.* **26**: 185–187.
- Bukhtojarov, F. E., B. B. Ustinov, T. N. Salanovich, A. L. Antonov, A. V. Gusakov, O. N. Okunev, and A. P. Sinitsyn. 2004. Cellulose complex of the fungus *Chrysosporium lucknowense*: Isolation and characterization of endoglucanases and cellobiohydrolases. *Biochemistry* **69**: 542–551.

7. Cao, W. G. and D. L. Crawford. 1993. Purification and some properties of β -glucosidase from the ectomycorrhizal fungus *Piptoporus betulinus* strain Smf. *Can. J. Microbiol.* **39**: 125–129.
8. Deshpande, M. V., K. E. Eriksson, and L. G. Pettersson. 1984. An assay for selective determination of exo-1,4- β -glucanases in a mixture of cellulolytic enzymes. *Anal. Biochem.* **138**: 481–487.
9. Divne, C., J. Ståhlberg, T. Reinikainen, L. Ruohonen, G. Pettersson, J. K. Knowles, T. T. Teeri, and T. A. Jones. 1994. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science* **265**: 524–528.
10. Edwards, I. P., R. A. Upchurch, and R. Z. Donald. 2008. Isolation of fungal cellobiohydrolase I genes from sporocarps and forest soils by a PCR. *Appl. Environ. Microbiol.* **74**: 3481–3489.
11. Grassick, A., P. G. Murray, R. Thompson, C. M. Collins, L. Byrnes, G. Birrane, T. M. Higgins, and M. G. Tuohy. 2004. Three-dimensional structure of a thermostable native cellobiohydrolase, CBH IB, and molecular characterization of the *cel7* gene from the filamentous fungus, *Talaromyces emersonii*. *Eur. J. Biochem.* **271**: 4495–4506.
12. Gusakov, A. V., T. N. Salanovich, A. I. Antonov, B. B. Ustinov, O. N. Okunev, R. Burlingame, M. Emalfarb, M. Baez, and P. Arkady. 2007. Design of highly efficient cellulase mixtures for enzymatic hydrolysis of cellulose. *Biotechnol. Bioeng.* **97**: 1028–1038.
13. Gusakov, A. V., A. P. Sinitsyn, T. N. Salanovich, F. E. Bukhtoyarov, A. V. Markov, B. B. Ustinov, C. van Zeijl, P. Punt, and R. Burlingame. 2005. Purification, cloning and characterization of two forms of thermostable and highly active cellobiohydrolase I (Cel7A) produced by the industrial strain of *Chrysosporium lucknowense*. *Enz. Microbial Technol.* **36**: 57–69.
14. Haakana, H., A. Miettinen-Oinonen, V. Joutsjoki, A. Mantyla, P. Suominen, and J. Vehmaanpera. 2004. Cloning of cellulose genes from *Melanocarpus albomyces* and their efficient expression in *Trichoderma reesei*. *Enz. Microbial Technol.* **34**: 159–167.
15. Hamada, N., K. Ishikawa, N. Fuse, R. Kodaira, M. Shimozaka, Y. Amano, T. Kanda, and M. Okazaki. 1999. Purification, characterization and gene analysis of exo-cellulase II (Ex-2) from the white rot basidiomycete *Irpex lacteus*. *J. Biosci. Bioeng.* **87**: 442–451.
16. Henrissat, B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarity. *Biochem. J.* **280**: 309–316.
17. Hong, J., H. Tamaki, K. Yamamoto, and H. Kumagai. 2003. Cloning of a gene encoding thermostable cellobiohydrolase from *Thermoascus aurantiacus* and its expression in yeast. *Appl. Microbiol. Biotechnol.* **63**: 42–50.
18. Jia, J., P. S. Dyer, and J. F. Buswell. 1999. Cloning of the CBHI and CBHII genes involved in cellulose utilization by the straw mushroom *Volvariella volvacea*. *Mol. Gen. Genet.* **261**: 985–993.
19. Kerem, Z., K. Jensen, and K. Hammel. 1999. Biodegradative mechanism of the brown rot basidiomycete *Gloeophyllum trabeum*: Evidence for an extracellular hydroquinone-driven Fenton reaction. *FEBS Lett.* **446**: 49–54.
20. Koch, A., C. T. Weigel, and G. Schulz. 1993. Cloning, sequencing, and heterologous expression of a cellulase-encoding cDNA (*cbh1*) from *Penicillium janthinellum*. *Gene* **124**: 57–65.
21. Konstantinidis, A. K., I. Marsden, and M. L. Sinnott. 1993. Hydrolyses of alpha- and beta-cellobiosyl fluorides by cellobiohydrolases of *Trichoderma reesei*. *J. Biochem.* **291**: 883–888.
22. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680–685.
23. Lahjouji, K., R. Storms, Z. Xiao, K. B. Joung, Y. Zheng, J. Powlowski, A. Tsang, and L. Varin. 2007. Biochemical and molecular characterization of a cellobiohydrolase from *Trametes versicolor*. *Appl. Microbiol. Biotechnol.* **75**: 337–346.
24. Lee, C. C., D. W. Wong, and G. H. Robertson. 2001. Cloning and characterization of two cellulose genes from *Lentinula edodes*. *FEMS Microbiol. Lett.* **205**: 355–360.
25. Liete, R. S. R., E. Gomes, and R. Da-Silva. 2007. Characterization of β -glucosidases from a mesophilic *Aureobasidium pullulana* and thermophilic *Thermoascus aurantiacus*. *Process Biochem.* **42**: 1101–1106.
26. Limam, F., S. E. Chaabouni, R. Ghrir, and N. Marzouki. 1995. Two cellobiohydrolases of *Penicillium occitanis* mutant Pol 6: Purification and properties. *Enz. Microbial Technol.* **17**: 340–346.
27. Lin, J., B. Pillay, and S. Singh. 1999. Purification and biochemical characterization of β -glucosidase from a thermophilic fungus, *Thermomyces lanuginosus* - SSBP. *Biotechnol. Appl. Biochem.* **30**: 81–87.
28. Li, Y. L., D. C. Li, and F. C. Teng. 2006. Purification and characterization of a cellobiohydrolase from the thermophilic fungus *Chaetomium thermophilus* CT2. *Wei Sheng Wu Xue Bao* **46**: 143–146.
29. Miller, G. L. 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* **31**: 426–428.
30. Muñoz, I. G., W. Ubhayaseker, H. Henriksson, I. Szabó, G. Pettersson, G. Johansson, S. L. Mowbray, and J. Ståhlberg. 2001. Family 7 cellobiohydrolases from *Phanerochaete chrysosporium*: Crystal structure of the catalytic module of Cel7D (CBH58) at 1.32 Å resolution and homology models of the isozymes. *J. Mol. Biol.* **314**: 1097–1111.
31. Nikolova, P. V., A. L. Creagh, S. J. B. Duff, and C. A. Haynes. 1997. Thermostability and irreversible activity loss of exoglucanase/xylanase Cex from *Cellulomonas fimi*. *J. Biochem.* **36**: 1381–1388.
32. Parkkinen, T., A. Koivula, J. Vehmaanperä, and J. Rouvinen. 2008. Crystal structures of *Melanocarpus albomyces* cellobiohydrolase Cel7B in complex with cello-oligomers show high flexibility in the substrate binding. *Protein Sci.* **17**: 1383–1394.
33. Percival Zhang, Y. H., M. E. Himmel, and J. R. Mielenz. 2006. Outlook for cellulose improvement: Screening and selection strategies. *Biotechnol. Adv.* **24**: 452–481.
34. Rouau, X. and E. Odier. 1986. Purification and properties of 2 enzymes from *Duchomitus squalens* which exhibit both cellobiohydrolase and xylanase activity. *Carbohydr. Res.* **145**: 279–292.
35. Schmidhalter, D. R. and G. Canevascini. 1993. Purification and characterization of two exo-cellobiohydrolases from the brown-rot fungus *Coniophora puteana* (Schum ex Fr) Karst. *Arch. Biochem. Biophys.* **300**: 551–558.
36. Teeri, T. T. 1997. Crystalline cellulose degradation: New insight into the function of cello-biohydrolases. *Trends Biotechnol.* **15**: 160–167.
37. Teeri, T. T., P. Lehtovaara, S. Kauppinen, I. Salovuori, and J. Knowles. 1987. Homologous domains in *Trichoderma reesei* cellulolytic enzymes: Gene sequence and expression of cellobiohydrolase II. *J. Gene* **51**: 43–52.
38. Tuohy, M. G., D. J. Walsh, P. G. Murray, M. Claeysens, M. Cuffe, A. V. Savage, and M. P. Coughlan. 2002. Kinetic parameters and mode of action of the cellobiohydrolases produced by *Talaromyces emersonii*. *Biochim. Biophys. Acta* **1596**: 366–380.

Effect of pH and temperature on the global compactness, structure, and activity of cellobiohydrolase Cel7A from *Trichoderma harzianum*

Francieli Colussi · Wanius Garcia · Flávio Rodolfo Rosseto ·
Bruno Luan Soares de Mello · Mário de Oliveira Neto ·
Igor Polikarpov

Received: 21 July 2011 / Revised: 3 October 2011 / Accepted: 11 October 2011 / Published online: 3 November 2011
© European Biophysical Societies' Association 2011

Abstract Due to its elevated cellulolytic activity, the filamentous fungus *Trichoderma harzianum* (*T. harzianum*) has considerable potential in biomass hydrolysis application. Cellulases from *Trichoderma reesei* have been widely used in studies of cellulose breakdown. However, cellulases from *T. harzianum* are less-studied enzymes that have not been characterized biophysically and biochemically as yet. Here, we examined the effects of pH and temperature on the secondary and tertiary structures, compactness, and enzymatic activity of cellobiohydrolase Cel7A from *T. harzianum* (*Th* Cel7A) using a number of biophysical and biochemical techniques. Our results show that pH and temperature perturbations affect *Th* Cel7A stability by two different mechanisms. Variations in pH modify protonation of the enzyme residues, directly affecting its activity, while leading to structural destabilization only at extreme pH limits. Temperature, on the other hand, has direct influence on mobility, fold, and compactness of the enzyme, causing unfolding of *Th* Cel7A just above the optimum temperature limit. Finally, we demonstrated that incubation with cellobiose, the product of the reaction and a competitive inhibitor, significantly increased the thermal stability of *Th* Cel7A. Our studies might provide insights into

understanding, at a molecular level, the interplay between structure and activity of *Th* Cel7A at different pH and temperature conditions.

Keywords *Trichoderma harzianum* · Cellobiohydrolase · Cel7A · Cellulose · Circular dichroism

Introduction

Cellulases are important enzymes that can hydrolyze the β -1,4-linkage between the glucose units of the cellulose chain. These enzymes are classified into glycoside hydrolase (GH) families (<http://www.cazy.org/Glycoside-Hydrolases.html>) (Henrissat and Davies 1997) and can be divided into three types: endo-1,4- β -D-glucanase (EC 3.2.1.4), 1,4- β -D-cellobiohydrolase (EC 3.2.1.91), and β -glucosidase (EC 3.2.1.21). Cellulases are being used in several industrial applications, for example, in the pulp and paper, textile, and detergent industries, and more recently in industrial conversion of cellulose-containing biomass to fermentable sugars as an alternative to fossil fuels (Gray et al. 2006; Kumar et al. 2008).

A large part of currently produced bioethanol is obtained from starch, corn, or sucrose, thus competing directly or indirectly with food. Therefore, biomass utilization offers an alternative carbon source for second-generation biofuel production (Bommarius et al. 2008; Kumar et al. 2008). Cellulose, a polysaccharide composed of β -D-glucose units joined by 1,4-glycosidic bonds, is the major component of the cell wall of plants and the most abundant polymer in nature (Igarashi et al. 2006; Kumar et al. 2008). The degree of crystallinity differs with the source of the cellulose, and crystalline regions are specifically resistant to enzymatic hydrolysis. Thus, degradation of cellulose is one of the

Francieli Colussi and Wanius Garcia contributed equally to this work.

F. Colussi · F. R. Rosseto · B. L. S. de Mello ·
M. de Oliveira Neto · I. Polikarpov (✉)
Grupo de Cristalografia, Instituto de Física de São Carlos (IFSC),
Universidade de São Paulo (USP), Av. Trabalhador SãoCarlense
400, P.O. Box 369, São Carlos, SP 13560-970, Brazil
e-mail: ipolikarpov@ifsc.usp.br

W. Garcia
Centro de Ciências Naturais e Humanas (CCNH), Universidade
Federal do ABC (UFABC), Santa Adélia, 148, Santo André,
SP, Brazil

most important processes in the carbon cycle, normally involving a combined action of several synergistically functioning enzymes.

The cellulolytic systems of filamentous fungi *Trichoderma reesei* (*Hypocrea jecorina*) and *Humicola insolens* are among the most thoroughly functionally, biochemically, and structurally studied (von Ossowski et al. 2003; Kumar et al. 2008). These enzymatic mixtures consist of at least of two cellobiohydrolases, Cel7A and Cel6A (formerly CBHI and CBHII), that act, respectively, upon reducing and nonreducing ends of crystalline cellulose-releasing cellobiose units, and several endoglucanases that hydrolase amorphous regions of cellulose, thus decreasing the degree of polymerization of the substrate and generating new cellulose chain termini. Enzymatic action of β -glucosidases completes depolymerization by hydrolysis of cellobioses in glucose monomers (Kubicek et al. 1993; Kleywegt et al. 1997). The cellobiohydrolase Cel7A is the major component of total extract produced by *Trichoderma* and is the key enzyme for efficient degradation of crystalline cellulose. For example, removal of the Cel7A gene reduces overall activity of *Trichoderma reesei* (*T. reesei*) on crystalline cellulose by 70%, and inactivation of Cel7A during cellulose hydrolysis is the major factor responsible for the decrease in the hydrolysis rate (Davies and Henrissat 1995; Teeri 1997).

Structural studies show that Cel7A from *T. reesei* is composed of a large catalytic core domain joined by an O-glycosylated linker peptide to a small cellulose-binding domain that is responsible for anchoring the enzyme on solid substrate surfaces (Davies and Henrissat 1995; Hui et al. 2002; Boraston et al. 2004). Glycosylation plays an important role in protecting the linker peptide against proteolysis. The linker seems to be important for enhanced synergy between the catalytic core domain and the cellulose-binding domain on its natural substrate, since the shortening or the deletion of the linker drastically reduces enzymatic activity on crystalline cellulose (Shen et al. 1991; Srisodsuk et al. 1993). Crystallographic studies reveal that the catalytic core domain has a large enclosed tunnel structure formed by a β -sandwich fold. The tunnel is 50 Å long and contains the active site to mediate the hydrolysis of cellulose from the chain ends (Divne et al. 1994; Davies and Henrissat 1995; Munoz et al. 2001; Grassick et al. 2004; Gray et al. 2006; Parkkinen et al. 2007).

Trichoderma harzianum (*T. harzianum* or *Hypocrea lixii*) is a filamentous fungi that secretes a well balanced cellulolytic complex able to efficiently hydrolyze cellulolytic substrates into monomeric glucose. Due to its elevated cellulolytic activity, the fungus *T. harzianum* has considerable potential in biomass hydrolysis applications (Castro et al. 2010). The mixture of cellulases produced by

T. reesei has been widely used as a model in studies of cellulose breakdown. However, cellulases from *T. harzianum* are less-studied enzymes that have not been characterized biophysically and biochemically as yet. Cel7A from *T. harzianum* (NCBI accession number AAF36391.1) is 89% homologous to *T. reesei* Cel7A and presumably also has a bimodular molecular organization composed of a large catalytic domain linked to a small carbohydrate-binding domain by a flexible linker peptide.

Industrial conditions frequently impose restrictions on the pH range applicable at a particular step of the processes. The use of high temperatures generally accelerates rates of enzymatic reactions, reduces bacterial contamination, and in most cases, increases the solubility of substrates and thus is important for biotechnological applications and in bioreactors. High temperature and inadequate pH conditions, on the other hand, will eventually inactivate and/or denature enzymes. To comprehend cellulase inactivation at a molecular level, we studied the effects of pH and temperature on enzymatic activity and secondary and tertiary structures of cellobiohydrolase Cel7A from fungus *T. harzianum* by employing several biophysical and biochemical tools. This work is the first study that examined the effect of pH and temperature on the structural stability and enzymatic activity of Cel7A from *T. harzianum*.

Materials and methods

Materials

Protein standards used as sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) markers, Avicel[®] PH101, and cellobiose were purchased from Sigma Aldrich. DEAE-Sephadex A-50 (5 × 40 cm) column was purchased from GE Healthcare. Sypro Orange dye was purchased from Molecular Probes. All chemicals and reagents used in this study were of the highest purity analytical grade. To cover a pH range of 2–6 and 7–10, phosphate/citric acid and Gly/NaOH were chosen as buffers, respectively, and prepared at a final concentration of 20 mM.

Expression and purification of cellobiohydrolase Cel7A from *T. harzianum*

The expression and purification of wild-type cellobiohydrolase Cel7A enzyme from *T. harzianum* (*Th* Cel7A) was carried out as described previously (Colussi et al. 2010, 2011). The resulting *Th* Cel7A after purification was exhaustively dialyzed in 50 mM Tris-HCl buffer, pH 7.0 for elimination of NaCl. The purity was checked by

SDS-PAGE, and the final concentration of *Th* Cel7A was determined by UV absorbance at 280 nm using a theoretical extinction coefficient based on the amino acid composition (Gill and von Hippel 1989). The final product *Th* Cel7A was then frozen and stored at -80°C and was melted on ice before use.

Enzymatic activity measurements

Exoglucanase activity against insoluble substrate was determined using as substrate 1% Avicel dissolved in a buffer. Activity of *Th* Cel7A was measured in the temperature range between 20 and 70°C and pH values from 3.0 to 7.0 (20 mM phosphate/citric acid or Gly/NaOH buffers adjusted to the different pH values). The final protein concentration for each experiment was 1.5 μM . For measurement of chemical stability, samples and controls were incubated at 50°C for 120 min and agitated every 20 min. The amount of reducing sugars was determined using DNS reagent as described by Wood and Bhat (1988). One unit of the enzyme was defined as 1 μmol of glucose equivalent released per minute under the assay conditions.

Far-UV circular dichroism (CD) spectroscopy

Far-UV circular dichroism (CD) measurements were performed using a Jasco J-815 spectropolarimeter (JASCO, Japan) equipped with a Peltier temperature control unit. All data were measured using quartz cuvettes of 1 mm path length, scanning speed of 20 nm min^{-1} , a spectral bandwidth of 1 nm, and a response time of 1 s. The spectra were recorded over the wavelength range 195–240 nm and were determined as an average of 16 accumulations. The buffer contributions were subtracted in all the analyses to eliminate background effects, and all experiments were carried out in triplicate. The *Th* Cel7A concentration used was approximately 5 μM in all the experiments. *Th* Cel7A samples (300 μl) were incubated at pH values ranging from 2 to 10 in 20 mM phosphate/citric acid or Gly/NaOH buffers and adjusted to the different pH values. The incubations took 12 h at 25°C prior to analysis. Thermal denaturation of *Th* Cel7A, in the presence or absence of cellobiose at pH 5, was also characterized by measuring the ellipticity changes at 212 nm induced by a temperature increase to 90°C from 20°C at a heating rate of 10°C/h . The samples were incubated in each condition, in the presence or absence of cellobiose, for 12 h before the analysis. The fraction of denatured protein (α) was calculated from the relationship: $\alpha = (\theta_n - \theta_{\text{obs}})/(\theta_n - \theta_d)$ and $\alpha + \beta = 1$, in which θ_{obs} is the ellipticity obtained at a particular temperature and θ_d and θ_n are the values of the ellipticity characteristic of the denatured and native states, respectively.

Intrinsic fluorescence spectroscopy

Intrinsic fluorescence spectroscopy measurements were performed on a steady-state spectrofluorimeter model K2 ISS, equipped with a refrigerated circulator. The *Th* Cel7A concentration was 2 μM in all of the experiments. *Th* Cel7A samples (400 μl) were incubated at pH values ranging from 2 to 10 in 20 mM phosphate/citric acid or Gly/NaOH buffers and adjusted to the different pH values. The samples were incubated at 25°C for 12 h before analysis. *Th* Cel7A has 10 tryptophans, which were excited at 295 nm, and fluorescence emission spectra were recorded from 300 to 400 nm from samples loaded into a quartz cuvette of 1 cm path length. The buffer contributions were subtracted in all analyses, and all experiments were carried out in triplicate.

Thermofluor method

Solutions of 2 μl of $300\times$ Sypro Orange were added to 18 μl of *Th* Cel7A at 5 μM , pH 5, in the absence or presence of cellobiose. Subsequently, the mixtures were added to the wells of a 96-well thin-wall PCR plate (Bio-Rad). Water was measured as a control. The plates were sealed with Optical-Quality Sealing Tape (Bio-Rad) and heated in an iCycler iQ Real Time PCR Detection System (Bio-Rad) from 10 to 90°C in increments of 0.5°C . Fluorescence changes in the wells of the plate were monitored simultaneously with a charge-coupled device camera (CCD). The wavelengths for excitation were 470–505 nm and emission 540–700 nm.

Dynamic light scattering (DLS)

Th Cel7A at 0.5 mg/ml and pH 5 was centrifuged at $14,000\times g$ for 15 min at room temperature and immediately loaded into a quartz cuvette (path length of 1 cm) prior to measurements. DLS was performed using a Zetasizer instrument (Malvern Instrument, Worcestershire, UK). The temperature was raised from 25 to 80°C at 5°C intervals every 120 s for equilibration of temperatures. The effect of buffer was subtracted from each recording to eliminate the noise produced by it. The hydrodynamic radius was calculated using the Zetasizer μV Software.

Small-angle X-ray scattering (SAXS)

All data were collected on the D02A-SAXS2 beamline of the Synchrotron Light National Laboratory (LNLS, Campinas, Brazil) (Kellermann et al. 1997). The wavelength of incident monochromatic X-ray beam λ was set at 1.488 Å. The X-ray patterns were recorded using a two-dimensional CCD detector (MarResearch, USA). The sample-detector

distance was set at 1,022.5 mm, resulting in a scattering vector (q) range of $0.013\text{--}0.33\text{ \AA}^{-1}$, where $q = 4 \times \pi \times \sin(\theta)/\lambda$ and θ is half the scattering angle. The samples were centrifuged for 15 min at $13,000\times g$ (at 4°C) prior to measurement to remove any aggregates or particles and then placed on ice. The samples were loaded into cells made of two thin parallel mica windows. Two successive frames of 300 s each were recorded for each sample to monitor radiation damage and beam stability. Buffer scattering was recorded before each sample scattering. The SAXS patterns were individually corrected for the detector response and scaled by the incident beam intensity and the sample's absorption. The parasitic scattering from air and beamline windows was subtracted from the total measured intensities. The integration of SAXS patterns was performed using Fit2D software, and the curves were scaled by protein concentration. The *Th* Cel7A concentrations used for small-angle X-ray scattering measurements were approximately 4 and 2 mg/ml. *Th* Cel7A samples (200 μl) were incubated at pH values ranging from 2 to 10 in 20 mM phosphate/citric acid or Gly/NaOH buffers and adjusted to the different pH values. The incubations were prolonged for a further 12 h at 25°C before the analysis. Thermal denaturation of *Th* Cel7A at pH 5 was also characterized by SAXS induced by a temperature increase from 20 to 70°C .

Results and discussion

Temperature and pH profiles of *Th* Cel7A enzymatic activity

The activity of an enzyme is dependent on pH and temperature, conditions that invoke changes in enzyme folding, conformation, and protonation. As a first step of our studies, the exoglucanase activity of *Th* Cel7A was studied as a function of pH. Our results showed that, under the conditions of our study (at 50°C), the exoglucanase activity of *Th* Cel7A increases as pH increases within the pH range of 3–5 and then decreases within the pH range of 5–7 (Fig. 1a). The enzyme reaches its optimum activity at pH 5 and 50°C , conditions similar to those found for other cellobiohydrolases (Bommarius et al. 2008; Hui et al. 2002). Interestingly, even at pH 3, it was found that the remaining activity of the enzyme was 76% and at pH levels above 7 the remaining activity was virtually zero (data not shown). Next, the exoglucanase activity of *Th* Cel7A was studied as a function of temperature at pH 5. Within the temperature range of $20\text{--}50^\circ\text{C}$ (pH 5), the exoglucanase activity of *Th* Cel7A increased with the increase in temperature reaching its maximum at $T = 50^\circ\text{C}$ (Fig. 1b). The exoglucanase activity of *Th* Cel7A steadily decreased to zero at

temperatures above 50°C (Fig. 1b). A better understanding of the biochemistry and biophysics of the processes of conformational change of *Th* Cel7A is very important to study the relationship of enzyme structure, stability, and activity. Thus, to correlate the enzymatic activity data with the structure of the enzyme, we conducted a series of detailed studies of the unfolding and stability of *Th* Cel7A as a function of temperature and pH, employing a series of biophysical studies.

Influence of pH on the structural stability and compactness of *Th* Cel7A

Influence of pH on the structural stability of *Th* Cel7A was studied using far-UV circular dichroism (CD) spectroscopy, intrinsic fluorescence spectroscopy, and small-angle X-ray scattering (SAXS). CD spectroscopy is a powerful

Fig. 1 Effect of pH (a) and temperature (b) on the enzymatic activity of *Th* Cel7A

tool to investigate the conformation of proteins; therefore CD spectroscopy was used to probe the secondary structure of the purified *Th* Cel7A to assess its folding in terms of secondary structure. The general shape of the spectrum at pH 5 (optimum pH) and 25°C (Fig. 2, solid line), with positive maxima near 195 ± 1 nm and a broad, negative trough at $205\text{--}215 \pm 1$ nm is characteristic of a β -class protein, comparable to the secondary structure of other cellobiohydrolases (Divne et al. 1994; Boer and Koivula 2003; Voutilainen et al. 2009). In addition, the absence of a negative band around 222 ± 1 nm is indicative of a protein with low content of α -helical elements in the secondary structure. Figure 2 shows the influence of pH on the secondary structure of *Th* Cel7A at 25°C under different pH values (pH = 2, 5, 6, and 10). Our results show that there are no appreciable structural changes in *Th* Cel7A within the pH range of 2–8 at 25°C. However, when the pH value was increased above 8 (Fig. 2, highly alkaline pH), the spectral profile was progressively altered with a loss of the minimum characteristic at ~ 210 nm, indicating a strong effect of pH on the secondary structure.

As a next step, the tertiary structure of *Th* Cel7A was investigated by intrinsic fluorescence spectroscopy and SAXS. Figure 3 shows the emission spectra of *Th* Cel7A after excitation at 295 nm at several pH levels. The spectrum at pH 5 (solid line, optimum pH) is characterized by a maximum emission at 336 ± 1 nm typical of tryptophan residues partially exposed to water environment, indicating that the majority of tryptophan residues present in the *Th* Cel7A sequence are somewhat exposed to the solvent. Similar results were observed when the protein was incubated at pH 2 and 6 (Fig. 3), indicating that there are no

appreciable conformational changes induced in the protein and its tryptophan residue positions. However, the spectrum at pH 10 (dash line) is characterized by a maximum emission at 340 ± 1 nm and a red shift in the spectral center of mass indicating that the tryptophan residues are more exposed to the buffer environment, allowing us to consider the protein unfolded.

SAXS is a technique applicable to the analysis of the global shapes, dimensions, and oligomeric states of macromolecules in solution. Figure 4a shows the X-ray scattering curves for *Th* Cel7A at different pH values (pH = 2, 5, 8, and 10). The raw scattering data clearly show that there are no structural alterations of the molecule at pH below 8. However, the raw scattering was altered when the pH value was increased above 8 indicating that there are structural alterations in the tertiary structure of *Th* Cel7A at highly alkaline pH values. For a more detailed analysis of the effects of pH on the global compactness of *Th* Cel7A, the SAXS data were analyzed in terms of the Kratky plot (Fig. 4b) (Doniach et al. 1995; Garcia et al. 2008a, b). The Kratky plot, in Fig. 4b, aids in the structural interpretation of such alterations. For a system with a compact shape, this plot shows a well defined curve with an initial rising portion followed by a marked descent thereby forming a well defined maximum (Semisotnov et al. 1996). Nevertheless, the expected curve for a polymer in an extended or random coil conformation rises to a characteristic plateau, with no well defined maximum (Darimont et al. 1998). As can be seen in Fig. 4b all the curves corresponding to measurements made in pH below 8 shows the well defined maxima expected for compact structures. However, in the case of pH values above 8 the Kratky plot indicates an alteration in

Fig. 2 Effect of pH on the secondary structure of *Th* Cel7A at 25°C monitored by far-UV CD spectroscopy. Far-UV CD spectra as a function of pH. The pH values were 2 (dot), 5 (solid line), 6 (dash dot), and 10 (dash). The experimental conditions are described in “Materials and methods”

Fig. 3 Effect of pH on the tertiary structure of *Th* Cel7A at 25°C monitored by intrinsic fluorescence spectroscopy. The pH values are 2 (dot), 5 (solid line), 6 (dash dot), and 10 (dash). The excitation wavelength was 295 nm, and emission was collected from 300 to 400 nm. The experimental conditions are described in “Materials and methods”

Fig. 4a, b Experimental solution scattering curves as a function of pH. **a** SAXS curves from *Th* Cel7A shown as the logarithm of the scattering intensity ($\log I$) versus the momentum transfer q scale. **b** Kratky plots for *Th* Cel7A. The pH values are 2 (black filled spheres), 5 (gray filled spheres), 8 (half filled squares), and 10 (triangles). The experimental conditions are given in “Materials and methods” section

the peak intensity and its position indicating that alterations to the compactness of the structure have occurred (Fig. 4b). This behavior is typically exhibited by elongated chains (Semisotnov et al. 1996), indicating that at highly alkaline pH values *Th* Cel7A is unfolded, in contrast to that observed for pH values below 8. Taken together, CD, intrinsic fluorescence, and SAXS results permit us to infer that there were no significant changes in the secondary and tertiary structure of *Th* Cel7A at pH values below 8, indicating that the protein is a pH acid resistant. However, alterations at both secondary and tertiary structural levels were observed at highly alkaline pH values.

Because in the pH range 2–8 there were no appreciable conformational changes in the protein structure that could

be detected by the biophysical techniques used in this work, the variation in activity observed was probably due to the variation in ionization of the side chains of the amino acids in the hydrophobic cellulose-binding tunnel, which each had their optimum at $\text{pH} = 5$. However, the observed decrease in activity observed for the pH range 8–10 is most probably mediated by the changes in secondary and tertiary structure of the enzyme, with the formation of molten globule-like state, and its unfolding, provoked by pH variation.

Influence of temperature on *Th* Cel7A stability and compactness

The nature of the thermal unfolding transition of *Th* Cel7A at pH 5 was studied by CD spectroscopy, thermofluor method, and SAXS in the range 10–90°C. The CD spectrum of purified *Th* Cel7A at 25°C was typical of proteins containing elements of β -sheet secondary structure (Fig. 2). The spectrum remains relatively constant below 55°C (data not shown), and this can be readily followed by tracking the ellipticity at 212 nm (Fig. 5a). The spectral profile progressively altered, however, when the temperature increased above 55°C, with a loss of the regular secondary structure (Fig. 5a). In addition, the enzyme could be completely thermally unfolded at temperatures above 80°C.

The structural alteration observed by CD spectroscopy was accompanied by an increase in the level of binding of Sypro Orange dye to the protein (Fig. 5b). Sypro Orange is an extrinsic fluorescent probe with low quantum yield in aqueous solution but is highly fluorescent in nonpolar environments with low dielectric constants such as hydrophobic sites in proteins (Ericsson et al. 2006; Garcia et al. 2007). When a protein starts to unfold or melt, the dye binds to exposed hydrophobic regions of the protein, resulting in a significant increase in fluorescence intensity emission. The fluorescence intensity reaches a maximum and then starts to decrease, probably due to precipitation of the complex of the fluorescent probe and the unfolded protein. Sypro Orange is known to bind to hydrophobic surfaces, and the results shown in Fig. 5b therefore indicate that the observed conformational change involves the exposure of such residues.

As can be seen in Fig. 5b, Sypro Orange shows a finite affinity for the protein at temperatures below 55°C, indicating the presence of low exposed hydrophobic regions on the native structure. The catalytic domain structures of Cel7A from *T. reesei*, for example, reveal tunnel-shaped structures formed by loops (Divne et al. 1994; Davies and Henriessat 1995; Munoz et al. 2001; Grassick et al. 2004; Gray et al. 2006; Parkkinen et al. 2007). A striking feature of the Cel7A structure is an extensive 50-Å-long cellulose-

Fig. 5a, b Thermal unfolding of *Th* Cel7A. **a** Thermal denaturation of *Th* Cel7A at pH 5 in the absence (circles) or presence (triangles) of cellobiose was monitored by measuring the ellipticity at 212 nm as a function of temperature. **b** Fluorescence intensity emission from Sypro Orange in the presence of *Th* Cel7A as a function of temperature. Circles *Th* Cel7A at pH 5 in the absence of cellobiose, triangles *Th* Cel7A at pH 5 in the presence of cellobiose. The wavelengths for excitation and emission were 490 and 575 nm, respectively

binding tunnel formed by four surface loops, which runs along the concave face of the sandwich and contains the active site at its far end. We advocate that the finite affinity for the native *Th* Cel7A at temperatures below 55°C is probably due to the interaction between the dye and the hydrophobic cellulose-binding tunnel. When the temperature was increased to values above 55°C, an improved binding of Sypro Orange to the protein was observed (Fig. 5b), implying exposure of hydrophobic regions. The reason for the decrease in the activity of *Th* Cel7A at temperatures between 55 and 65°C (Fig. 1b) is most likely the formation of a molten globule-like state. While between 55 and 65°C the enzyme activity decreased, the molten

globule-like state created exposes hydrophobic surfaces to the external environment and has a tendency to form aggregates. However, when the temperature was further increased to values above 65°C, the fluorescence signal decayed as a result of a reduction in the affinity of the dye to the protein, presumably due to occlusion of the hydrophobic surfaces by aggregation and/or protein unfolding.

To investigate whether *Th* Cel7A aggregates at high temperatures, we employed dynamic light scattering (DLS) to distinguish different oligomeric protein states and/or amorphous aggregation. When *Th* Cel7A was analyzed by DLS at 25 and 50°C, the observed profiles were characteristic of monodisperse solutions (data not shown). The hydrodynamic radii determined at 25 and 50°C were 3.11 and 3.19 nm, respectively, consistent, in both cases, with a monomeric macromolecule in solution. However, when *Th* Cel7A was incubated at 80°C, the profile was altered and showed a hydrodynamic radius of 36.78 nm, which corresponds to high molecular weight aggregates.

To obtain more detailed information about the global compactness of the enzyme, we collected SAXS data as a function of temperature. Figure 6a shows the X-ray scattering curves for *Th* Cel7A at different temperatures ($T = 20, 50, 55, 60$, and 65°C). The raw scattering (Fig. 6a) data show that there are no structural alterations in the low resolution structure of *Th* Cel7A at temperatures between 20 and 55°C. However, the raw scattering was altered when the temperature was increased above 55°C, indicating that *Th* Cel7A had started to unfold. For a more detailed analysis of the effects of temperature on the global compactness of *Th* Cel7A, the SAXS data were also analyzed in terms of the Kratky plot (Fig. 6b). As can be seen in Fig. 6b all the curves corresponding to experiments conducted at the temperatures between 20 and 55°C show the well defined maxima expected for compact structures. However for temperatures above 55°C, the Kratky plot clearly shows an alteration in both peak intensity and position indicating that strong alterations to the compactness of the protein structure occurred (Fig. 6b). Taken together, the CD and fluorescence results matched the SAXS experiments and jointly confirmed that the secondary and tertiary structure of *Th* Cel7A and its compactness are strongly dependent on temperature. In conclusion, the thermal unfolding behavior of *Th* Cel7A was found to be a two-step process at pH 5. The first step led to the formation of molten globule-like states between 55 and 65°C, followed by a second step that was due to the unfolding of the molten globule-like state at temperatures above 70°C. These molten globule-like states present the following characteristics: they have reduced enzyme activity, they have reduced secondary structure content, they have largely lost their compact core, and they exhibit more hydrophobic surface exposed to the solvent.

Fig. 6a, b Experimental solution scattering curves as a function of temperature. **a** SAXS curves from *Th* Cel7A shown as the logarithm of the scattering intensity ($\log I$) versus the momentum transfer q scale. **b** Kratky plots for *Th* Cel7A. The temperature values are 20°C (black filled circles), 50°C (half filled squares), 55°C (triangles), 60°C (gray filled circles), and 65°C (circles).

Cellobiose-binding influence on *Th* Cel7A thermal stability

In the course of cellobiohydrolase action upon cellulose, the polysaccharide chain slides through the substrate binding tunnel, and every second glycosidic bond is correctly presented to the catalytic apparatus, explaining why these enzymes progressively liberate disaccharide units (cellobiose) from the cellulose chain (Wood and Bhat 1988). After cleavage of the $\beta(1 \rightarrow 4)$ glycosidic bond, *Th* Cel7A might remain bound to the reaction product, cellobiose, which in this case functions as a competitive inhibitor. Thus, we performed thermal unfolding studies monitored by far-UV CD spectroscopy and thermofluor method (Fig. 5) in the absence (apo-*Th* Cel7A) and presence (cellobiose-*Th* Cel7A) of cellobiose in order to investigate the influence of this inhibitor on thermal stability of *Th* Cel7A. The melting temperature value for apo-

Th Cel7A, T_m , is equal to $61.5 \pm 0.2^\circ\text{C}$ (Fig. 5a). On the other hand, when the *Th* Cel7A was incubated with 1 mM cellobiose, thermal stability of the complex increased, accompanied by a T_m shift to $65.9 \pm 0.2^\circ\text{C}$. In addition, Sypro Orange fluorescence intensity was monitored as a function of temperature in the absence and presence of cellobiose. In Fig. 5b, Sypro Orange shows a finite affinity for the apo-*Th* Cel7A at temperatures below 55°C indicating the presence of hydrophobic regions, probably due to the interaction between the dye and the hydrophobic cellulose-binding tunnel. However, the cellobiose addition resulted in a decrease in Sypro Orange intensity emission at temperatures below 50°C , probably now due to the interaction between the cellobiose and the cellulose-binding tunnel. On the other hand, the fluorescence intensity increased only to temperature values above 60°C , indicating an increase in the protein thermal stability in the presence of cellobiose. When the temperature was further increased to values above 60°C , the level of Sypro Orange binding to *Th* Cel7A further increased (Fig. 5b), indicative of the exposure of hydrophobic regions. However, when the temperature increased above 73°C , a decrease in the fluorescence signal was detected, indicating a reduction in the affinity of the dye, presumably due to the enzyme aggregation and/or protein unfolding. Taken together, the CD and thermofluor method results confirm that the cellobiose significantly increased the thermal stability of *Th* Cel7A.

Conclusions

Due to its elevated cellulolytic activity, the fungus *T. harzianum* has considerable potential in several biotechnological applications, such as biomass hydrolysis. In this study, we examined the effects of pH and temperature on the compactness and structure of *Th* Cel7A and the correlations with enzymatic activity. Variations of pH modify the protonation of the enzyme residues, directly affecting its activity, while leading to structural destabilization only at extreme pH limits. Temperature, however, has a direct influence on mobility, fold, and compactness of the enzyme, causing unfolding just above the optimum temperature limit. Our studies might provide insights into understanding, at the molecular level, the interplay between structure and activity of *Th* Cel7A at different pH and temperature conditions, which could be useful for biotechnological applications.

Acknowledgments This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) via grant numbers 2008/56255-9, 2009/54035-4, 2010/08370-3, 2010/16542-9; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) via grant number 482166/2010-0; and Coordenação de

Aperfeiçoamento de Pessoal de Nível Superior (CAPES). We also thank Laboratório Nacional de Luz Sincrotron (LNLS) for beam time and help with X-ray data collection.

References

- Boer B, Koivula A (2003) The relationship between thermal stability and pH optimum studied with wild-type and mutant *Trichoderma reesei* cellobiohydrolase Cel7A. *Eur J Biochem* 270:841–848
- Bommarius AS, Katona A, Cheben SE, Patel AS, Ragauskas AJ, Knudson K, Pu Y (2008) Cellulase kinetics as a function of cellulose pretreatment. *Metab Eng* 10(6):370–381
- Boraston AB, Bolam DN, Gilbert HJ, Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* 382:769–781
- Castro AM, Pedro KC, Cruz JC, Ferreira MC, Leite SG, Pereira N (2010) *Trichoderma harzianum* IOC-4038: a promising strain for the production of a cellulolytic complex with significant β -glucosidase activity from sugarcane bagasse cellulignin. *Appl Biochem Biotechnol* 162(7):2111–2122
- Colussi F, Textor LC, Serpa V, Maeda RN, Pereira N, Polikarpov I (2010) Purification, crystallization and preliminary crystallographic analysis of the catalytic domain of the extracellular cellulase CBHI from *Trichoderma harzianum*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 66:1041–1044
- Colussi F, Serpa V, Delabona PS, Manzone LR, Voltatódio ML, Alves R, Mello BL, Pereira JRN, Farinas CS, Golubev AM, Santos MAM, Polikarpov I (2011) Purification, biochemical and biophysical characterization of cellobiohydrolase I from *Trichoderma harzianum* IOC 3844. *J Microbiol Biotech* 21:808–817
- Darimont BD, Wagner RL, Apriletti JW, Stallcup MR, Kushner PJ, Baxter JD, Fletterick RJ, Yamamoto KR (1998) Structure and specificity of nuclear receptor-coactivator interactions. *Genes Dev* 12(21):3343–3356
- Davies G, Henrissat B (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* 3:853–859
- Divne C, Stahlberg J, Reinikainen T, Ruohonen L, Pettersson G, Knowles JKC, Teeri TT, Jones TA (1994) The 3-dimensional crystal-structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science* 265(5171):524–528
- Doniach S, Basile J, Garel T, Orland H (1995) Partially folded states of proteins—characterization by X-ray-scattering. *J Mol Biol* 254(5):960–967
- Ericsson UB, Hallberg BM, DeTitta GT, Dekker N, Nordlund P (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* 357(2):289–298
- Garcia W, Araújo APU, Lara F, Foguel D, Tanaka M, Tanaka T, Garratt RC (2007) An intermediate structure in the thermal unfolding of the GTPase domain of human Sept4 (SEPT4/Bradeion- β) forms amyloid-like filaments in vitro. *Biochemistry* 46:11101–11109
- Garcia W, Figueira ACM, Oliveira Neto M, Guzzi CA, Buzzá HH, Portugal RV, Calgaro MR, Polikarpov I (2008a) Probing conformational changes in orphan nuclear receptor: the NGFI-B intermediate is a partially unfolded dimer. *Biophys Chem* 137:81–87
- Garcia W, Rodrigues NC, Oliveira Neto M, Araújo APU, Polikarpov I, Tanaka M, Tanaka T, Garratt RC (2008b) The stability and aggregation properties of the GTPase domain from human SEPT4. *Biochim Biophys Acta* 1784:1720–1727
- Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182:319–326
- Grassick A, Murray PG, Thompson R, Collins CM, Byrnes L, Birrane G, Higgins TM, Tuohy MG (2004) Three-dimensional structure of a thermostable native cellobiohydrolase, CBH IB, and molecular characterization of the cel7 gene from the filamentous fungus, *Talaromyces emersonii*. *Eur J Biochem* 271(22):4495–4506
- Gray KA, Zhao LS, Emptage M (2006) *Bioethanol. Curr Opin Chem Biol* 10(2):141–146
- Henrissat B, Davies G (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Chem Biol* 7(5):637–644
- Hui JPM, White TC, Thibault P (2002) Identification of glycan structure and glycosylation sites in cellobiohydrolase II and endoglucanases I and II from *Trichoderma reesei*. *Glycobiology* 12(12):837–849
- Igarashi K, Wada M, Hori R, Samejima M (2006) Surface density of cellobiohydrolase on crystalline celluloses—a critical parameter to evaluate enzymatic kinetics at a solid-liquid interface. *FEBS J* 273(13):2869–2878
- Kellermann G, Vicentin F, Tamura E, Rocha M, Tolentino H, Barbosa A, Craievich A, Torriani I (1997) The small-angle X-ray scattering beamline of the Brazilian Synchrotron Light Laboratory. *J Appl Crystallogr* 30(2):880–883
- Kleywegt GJ, Zou JY, Divne C, Davies GJ, Sinning I, Stahlberg J, Reinikainen T, Srisodsuk M, Teeri TT, Jones TA (1997) The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 angstrom resolution, and a comparison with related enzymes. *J Mol Biol* 272(3):383–397
- Kubicek CP, Messner R, Gruber F, Mach RL, Kubicekpranz EM (1993) The *Trichoderma* cellulase regulatory puzzle—from the interior life of a secretory fungus. *Enzyme Micro Technol* 15(2):90–99
- Kumar R, Singh S, Singh OV (2008) Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *J Ind Microbiol Biotechnol* 35:377–391
- Munoz IG, Ubhayasekera W, Henriksson H, Szabo I, Pettersson G, Johansson G, Mowbray SL, Stahlberg J (2001) Family 7 cellobiohydrolases from *Phanerochaete chrysosporium*: crystal structure of the catalytic module of Cel7D (CBH58) at 1.32 angstrom resolution and homology models of the isozymes. *J Mol Biol* 314(5):1097–1111
- Parkinen T, Koivula A, Vehmaanpera J, Rouvinen J (2007) Preliminary X-ray analysis of cellobiohydrolase Cel7B from *Melanocarpus albomyces*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 63:754–757
- Semisotnov GV, Kihara H, Kotova NV, Kimura K, Amemiya Y, Wakabayashi K, Serdyuk IN, Timchenko AA, Chiba K, Nikaido K et al (1996) Protein globularization during folding. A study by synchrotron small-angle X-ray scattering. *J Mol Biol* 262(4):559–574
- Shen H, Schmuck M, Pilz I, Gilkes NR, Kilburn DG, Miller RC, Warren RAJ (1991) Deletion of the linker connecting the catalytic and cellulose-binding domains of endoglucanase-A (CENA) of *Cellulomonas fimi* alters its conformation and catalytic activity. *J Biol Chem* 266(17):11335–11340
- Srisodsuk M, Reinikainen T, Penttilä M, Teeri TT (1993) Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with the crystalline cellulose. *J Biol Chem* 268(28):20756–20761
- Teeri TT (1997) Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. *Trends Biotechnol* 15:160–167
- von Ossowski I, Stahlberg J, Koivula A, Piens K, Becker D, Boer H, Harle R, Harris M, Divne C, Mahdi S et al (2003) Engineering the exo-loop of *Trichoderma reesei* cellobiohydrolase, Cel7A. A

- comparison with *Phanerochaete chrysosporium* Cel7D. J Mol Biol 333(4):817–829
- Voutilainen SP, Boer H, Alapuranen M, Jänis J, Vehmaanperä J, Koivula A (2009) Improving the thermostability and activity of *Melanocarpus albomyces* cellobiohydrolase Cel7B. Appl Microbiol Biotechnol 83:261–272
- Wood TM, Bhat KM (1988) Methods for measuring cellulase activities. Methods Enzymol 160:87–112

Joint x-ray crystallographic and molecular dynamics study of cellobiohydrolase i from *trichoderma harzianum*: deciphering the structural features of cellobiohydrolase catalytic activity

Larissa C. Textor^{1*}, Francieli Colussi^{1*}, Rodrigo L. Silveira^{2*}, Viviane Serpa¹, Bruno Luan de Mello¹, João Renato C. Muniz¹, Fabio M. Squina³, Nei Pereira Jr⁴, Munir S. Skaf² and Igor Polikarpov¹

* Equally contributing authors

¹Instituto de Física de São Carlos, Universidade de São Paulo, Av. Trabalhador Saocarlense 400, São Carlos, SP, 13560-970, Brazil, emails: larissa@consani-textor.com; fcolussi@ursa.ifsc.usp.br; jrcmuniz@gmail.com; ipolikarpov@ifsc.usp.br

²Institute of Chemistry, State University of Campinas – UNICAMP, Cx. P. 6154, Campinas, SP 13084-862, Brazil, emails: rodrigolsilveira@gmail.com; skaf@iqm.unicamp.br

³Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Rua Giuseppe Máximo Scolfaro, 10000, Campinas, SP, 13083-970, Brazil, email: fabio.squina@bioetanol.org.br

⁴Centro de Tecnologia, Escola de Química, Laboratório de Desenvolvimento de Bioprocessos (LaDeBio), Universidade Federal do Rio de Janeiro, RJ, 21949-900, Brazil, email: nei@eq.ufrj.br

Article type : Original Article

Corresponding author Igor Polikarpov, Instituto de Física de São Carlos, Universidade de São Paulo, Av. Trabalhador Saocarlense 400, São Carlos, SP, 13560-970, Brazil
phone +55 16 3373-8088, fax +55 16 3373-9881, e-mail: ipolikarpov@ifsc.usp.br

Running title: Crystal structure of Cel7A from *T. harzianum*

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/febs.12049

© 2012 The Authors Journal compilation © 2012 FEBS

Keywords X-ray structure; cellobiohydrolase; enzymatic hydrolysis; biomass; cellulosic ethanol; molecular dynamics

Abbreviations GH, glycosyl hydrolase; CBHI, cellobiohydrolase I; CCD, catalytic core domain; CBD, cellulose-binding domain; pNPC, p-nitrophenyl-D-cellobioside; cNPL, 2-chloro-4-nitrophenyl - β -lactoside; CMC, carboxy-methyl cellulose; rmsd, root-mean-square deviation; PDB, Protein Data Bank

Abstract

Aiming to contribute towards the characterization of new, biotechnologically relevant cellulolytic enzymes, we report here the first crystal structure of the catalytic core domain of Cel7A (CBHI) from the filamentous fungus *Trichoderma harzianum* IOC 3844. Our structural studies and molecular dynamics simulations show that the flexibility of Tyr260, in comparison to the Tyr247 from the homologous *T. reesei* Cel7A, is enhanced due to the short side chains of adjacent Val216 and Ala384 residues and creates an additional gap at the side face of the catalytic tunnel. *T. harzianum* CBHI also has a shortened loop at the entrance of the cellulose-binding tunnel, which has been described to interact with the substrate in *T. reesei* Cel7A. These structural features might explain why *T. harzianum* enzyme displays higher k_{cat} and K_m values and lower product inhibition on both glucosides and lactosides substrates as compared to *T. reesei* Cel7A.

Introduction

It is now generally accepted that environmental changes became a serious problem worldwide and might drastically affect the Earth's climate with severe consequences for humankind if not addressed properly. Several studies have associated global warming with increasing levels

of atmospheric CO₂ generated, among other sources, by burning of fossil fuels [1-3]. Hence, extensive scientific efforts have been invested toward utilization and recovery of natural renewable resources for bioenergy production [4]. Environment-friendly alternatives such as wind, solar power, and biomass represent important sources of renewable energy. Since lignocellulosic material comprises most of the biomass on Earth, biodegradation of cellulose and hemicellulose, has become one of the focus of research on renewable liquid biofuels. Hydrolysis of cellulose represents considerable challenge since the substrate is insoluble and recalcitrant against depolymerization and enzymatic action.

The biotechnological use of filamentous fungi is a growing field with promising practical applications in bioenergy and biotechnology since these microorganisms are known to express several hemicellulose- and cellulose-degrading enzymes at high yields [4]. There is a wide range of industrial technologies using cellulases, mainly in starch processing, grain alcohol fermentation, malting and brewing, extraction of fruit and vegetable juices, pulp, paper, and textile [5, 6]. Future promising applications for cellulases include the production of fuel ethanol from the abundant lignocellulosic biomass through the enzymatic breakdown of cellulose [7].

The well-studied filamentous fungus *Trichoderma reesei* expresses a number of cellulases that can be broadly divided into three types: endoglucanases (endo-1, 4- β -D-glucanases), cellobiohydrolases (CBHs) or exoglucanases (exo-1, 4- β -D-glucanases), and β -glucosidases (1, 4- β -D-glucosidases) [5, 8]. The classical view of cellulose hydrolysis is described as a simultaneous and synergistic action of these three types of enzymes. The hydrolysis process starts with endoglucanases that randomly hydrolyze internal β -1, 4-glycosidic linkages in the amorphous region of cellulose microfibrils, reducing significantly the degree of substrate polymerization, creating new terminal ends. Simultaneously, the accessible reducing and non-reducing ends of cellulose chains are attacked by the types I and

II cellobiohydrolases, respectively. The released cellobiose units are then hydrolyzed into glucose units by β -glucosidases [8].

Cellulases can be classified into glycosyl hydrolase families (GH) based on their amino acid sequence and folding similarities [9]. Together with some endoglucanases and chitosanases, type I cellobiohydrolases belong to the glycosyl hydrolases family 7, GH7 [10]. The exo-glucanase Cel7A, is the most important single enzyme component for cellulose depolymerization and conversion into cellobiose in fungi [11]. It is considered to be the key enzyme in the hydrolysis process since it is able to extensively hydrolyze crystalline cellulose, albeit at a slow rate [12]. Moreover, Cel7A is the predominant enzyme secreted by the fungus *Trichoderma sp.* when grown under conditions induced for the production of cellulases, reaching up to 60% of the total secreted protein [13].

In general, the cellulases produced by *T. reesei* are glycoproteins frequently containing both *O*- and *N*-linked glycosylation sites [14]. Most cellulases display similar structural organization consisting of two distinct functional and structural domains (Fig. 1A): a cellulose-binding domain (CBD) and an N-glycosylated catalytic core domain (CCD) [15]. The domains are joined together by a long and highly O-glycosylated linker region. Structural studies have shown that N-glycosylation of Cel7A from *T. reesei* strains QM9414 and ALKO2877 (derived from QM9414) is present as single N-acetylglucosamine modification on three (Asn45, Asn270, and Asn384) out of four potential sites located in the protein catalytic domain [16]. The occurrence of such modifications depends on several factors, including strain and culturing conditions, and generates several enzyme isoforms with similar catalytic and adsorption properties [17, 18].

X-ray structures of four GH family 7 cellobiohydrolases have been published to date: two of them correspond to the CCDs of *Trichoderma reesei* Cel7A (PDB entry **1CEL**) [19]

and of *Phanerochaete chrysosporium* Cel7D (PDB entry **1GPI**) [20] whereas two other structures represent single-domain thermostable filamentous fungi cellobiohydrolases Cel7A from *Talaromyces emersonii* (PDB id code **1Q9H**) [21] and Cel7B from *Melanocarpus albomyces* (PDB id code **2RFW**) [22]. In addition, crystallographic structure of cellobiohydrolase Cel7A from *Heterobasidion annosum* is available at the PDB under the code **2XSP**, but has not yet been published. Furthermore, solution structures of several CBMs have been determined, including the one of *Trichoderma reesei* Cel7A (PDB entry code **1CBH**, [23])

Crystallographic structures of the full-length proteins have not been determined yet, presumably due to the presence of the highly mobile linker region, which hampers crystallization. At the same time, CCD structures of *T. reesei* CBHI (Tr_Cel7A) have been determined in complex with a number of oligosaccharides and ligands [24]. The CCD fold consists of two large anti-parallel β -sheets forming a β -sandwich tertiary structure. An elongated cellulose-binding tunnel surrounded by disulfide-bridged loops is formed along the concave side of the β -sandwich. The tunnel encloses the active site at its far end where the catalytic residues aspartate and both glutamates are located [19]. The presence of the elongated tunnel mediates the enzyme specificity that is achieved by a substantial number of interactions between the cellulose chain ends and the residues of the tunnel [25]. The polysaccharide chain slides through the substrate binding tunnel and every second glycosidic bond is correctly presented to the catalytic apparatus, which explains why these enzymes progressively liberate disaccharide units (cellobiose) from the reducing end of a cellulose chain [19]. After enzymatic cleavage, the cellobiose is found bound to sites +1 and +2 near the exit of the tunnel [19]. The catalytic triad of carboxylate groups is strictly conserved among the enzymes of GH family 7 [26].

Trichoderma harzianum is a filamentous fungus that exhibit both cellulolytic and mycoparasitic activities, and unlike *T. reesei*, its cellulolytic activity and biotechnological potential are still largely unexplored. Enzymatic studies of several strains of *T. harzianum* have shown that these fungi are able to produce cellulolytic complexes with higher β -glucosidase activity as compared to *T. reesei*. Moreover, *T. harzianum* secretes an effective and well balanced enzymatic system able to completely hydrolyze cellulosic substrates into monomeric glucose [27]. The hydrolysis efficiency of *T. harzianum* represents a considerable advantage over *T. reesei* due to the higher levels of xylanase and β -glucosidase activities and, consequently, potentially more efficient biomass hydrolysis.

In order to further investigate the enzymes that compose the cellulolytic complex secreted by *T. harzianum*, we studied the enzymatic kinetics and inhibition by cellobiose of Cel7A from *T. harzianum* and compared these data with enzymatic parameters of the homologue *T. reesei* Cel7A (81% amino acid sequence identity with *T. harzianum* Cel7A). We also determined the crystallographic structure of *T. harzianum* cellobiohydrolase I (Th_Cel7A), and performed molecular dynamics (MD) computer simulations on these structures to get better information about structural features of the enzymes and their dynamics. Our analyses provide insights into the molecular basis of the observed differences in the enzymatic activity and kinetics between Tr_Cel7A and Th_Cel7A.

Results

Kinetic constants and enzyme specificity for different soluble and insoluble substrates

Assuming Michaelis–Menten kinetics of the full-length *Trichoderma harzianum* Cel7A, the experimentally determined K_m and k_{cat} values are respectively 3.4 mM and 23.2 min⁻¹ for pNPC and 3.7 mM and 44.6 min⁻¹ for cNPL (Table 1). In comparison to the reported kinetic

parameters of Tr_Cel7A, which ranges from 0.4 up to 1.23 mM [4, 26, 28, 29, 30], our results show that Th_Cel7A exhibits significantly lower affinity for these substrates, whereas the turnover number k_{cat} is roughly 3.5 times larger than that of Tr_Cel7A on the lactoside substrate ($\sim 12 \text{ min}^{-1}$) [26, 28]. Moreover, our inhibition assays show that Th_Cel7A is competitively inhibited by cellobiose, though with a much higher inhibition constant of $7.2 \pm 1.2 \text{ mM}$ as compared to the K_i of 0.02 mM for Tr_Cel7A [31]. Low inhibition by the end-product and competitive inhibitor, cellobiose, might be potentially important for biotechnological applications of the enzyme.

Specific activity and substrate hydrolysis assays of the full-length Th_Cel7A and its CCD are shown in Fig. 2. Our experiments reveal that Th_Cel7A is active and very efficient in degrading of bacterial microcrystalline cellulose (BMCC), reaching almost 90% conversion of the later substrate in 24 hours under the conditions of our experiment. As expected the CCD is significantly less efficient in the hydrolysis of crystalline cellulose (Avicel-101, Sigmacell 20 and BMCC) as compared to the full-length enzyme containing both CCD and CBM domains (Fig. 2A&B). Furthermore, Th_Cel7A is capable of enzymatical cleavage of amorphous carboxy-methyl cellulose (CMC) however, in this case, the enzymatic efficiency is low and there is no significant difference in specific activity between CCD and the full-length enzyme (Fig. 2 A&B). Furthermore, neither full-length enzyme nor CCD was able to hydrolyze more than 10% of this substrate after 24 hours of reaction. Conversion of CMC rapidly reaches a plateau and reaction would not proceed any further (Fig. 2 C&D). Both specific activities and hydrolytic yields of Th_Cel7A Avicel 101 and Sigmacell 20 are at the intermediate levels in between the ones observed for BMCC and CMC and there is a considerable difference in activity and hydrolysis yields between complete enzyme and its CCD (Fig. 2). The measured specific activities of Th_Cel7A are comparable to that of Tr_Cel7A (Fig. 2A). According to the reported data, specific activities of Tr_Cel7A against

Avicel or Sigmacell are in the range of 0.014 U/mg to 0.3 U/mg, and 0.04 U/mg up to 0.07 U/mg for CMC under similar experimental conditions [15, 32]. These results show that the exoglucanase Cel7A from *T. harzianum* is efficient in hydrolyzing crystalline cellulosic substrates such as BMCC, Avicel and Sigmacell, but is not effective in hydrolysis of substituted substrates, such as CMC.

Overall structure of *Trichoderma harzianum* Cel7A CCD

The refined Cel7A CCD model (Table 2), comprise residues 18-445 and assume the characteristic fold of Cel7A enzymes [19-22, 24, 26], with the β -strands of the β -sandwich structure connected by the loops (L1 to L6) that enclose the long cellulose-binding tunnel (Fig. 1B).

The CCD contains a total of 20 cysteine residues, all involved in disulphide bridges known to stabilize the loops around the cellulose-binding cleft. As commonly found in enzymes secreted by fungi [19-22, 24, 26], the N-terminal glutamine is cyclized into a pyroglutamate residue. Three of the four glycosylation sites predicted according to the sequence motif N-X-S/T (where X is any amino acid residue, except Proline) could be identified in the electron density maps and single *N*-acetyl glucosamine (NAG) residues were modeled covalently bound to Asn126, Asn283, and Asn397. All NAG residues make contacts with symmetry-related protein molecules.

Overall comparison of the *Trichoderma harzianum* Cel7A CCD with other Cel7A CCD structures

Consistent with the high sequence homology between the Cel7A from *T. harzianum* and *T. reesei*, structural superposition shows that both molecules have a very similar global fold, with an rms deviation of 0.5 Å for 426 C α aligned residue pairs. As expected, the largest

structural deviations are associated with the loop regions that display high temperature factors and therefore are mobile or structurally disordered. The first important structural difference between Th_Cel7A and Tr_Cel7A is associated with the loop 3 at the entrance of the tunnel (residues Ser99 to Gln101 in the Tr_Cel7A sequence). Loop 3 is critical in defining the overall opening profile of the ligand-binding site. Some of the residues that compose this loop in Tr_Cel7A are absent in Th_Cel7A and consequently the enzyme has a more open entrance to the binding tunnel (Fig. 3 and 4). Gln101 from Tr_Cel7A loop 3 is a key residue in the substrate recognition at subsites -6 and -5 [24] and such interactions are thus not available in Th_Cel7A.

A similar shortening of the loop 3 was also described for GH7 cellobiohydrolases from *T. emersonii* (Te_Cel7A) and *P. chrysosporium* (Pc_Cel7D). The catalytic core domains of Te_Cel7A and Pc_Cel7D display a sequence similarity of 70% and 60% with the CCD of Th_Cel7A and a structural rms deviation of 0.8 Å for 417 Cα Te_Cel7A aligned residues and 0.9 Å for 415 Cα Pc_Cel7D aligned residues, respectively. The equivalent loop 3 in both structures is similar to that of Th_Cel7A and consequently they also exhibit a more open tunnel entrance when compared to the Tr_Cel7A [20, 24].

A second important feature of the Th_Cel7A CCD structure is related to the loops that form the sidewalls of the ligand-binding tunnel (loops 4 to 6), in which loops conformation and sequence composition are crucial for determining the shape of the channel side faces. Loops 5 and 6 of Th_Cel7A enclose part of the side tunnel surface and create a gap that gives access to the binding tunnel at the substrate site (Fig. 5A). Whereas the equivalent loop 5 in Pc_Cel7D is significantly shorter resulting in a totally exposed tunnel (Fig. 5B), the side face of the cellulose channel from Tr_Cel7A is almost completely covered (Fig. 5C). The entrance to the catalytic tunnel of Th_Cel7A is also more exposed as compared to Th_Cel7A (although not as much as in Pc_Cel7D) since Th_Cel7A residues Ile203 and Tyr371 are substituted by

Accepted Article

shorter side chains of Val216 and Ala384 in Th_Cel7A (Fig. 5A). As discussed below, molecular dynamics simulations demonstrate that the presence of the residue Ala384 in loop 6 of Th_Cel7A in place of Tr_Cel7A Tyr371 is responsible for notable effects on the dynamics of the loops enclosing the substrate binding site. Our crystal structure and molecular dynamics simulations reveal that Tyr260 can adopt different conformations in the Th_Cel7A CCD. Conformations in which Tyr260 side chain is more exposed to the solvent can partially obstruct the enhanced access to the cellulose binding channel in Th_Cel7A. Regarding the catalytic residues at the cellulose binding cleft, the only observed conformational difference is associated with Glu217, which is H-bonded to the glucosyl unit at the product +1 site in Tr_Cel7A, whereas the equivalent Glu230 of Th_Cel7A assumes an alternative conformation forming an H-bond to Asp154.

Molecular dynamics simulations of *T. harzianum* and *T. reesei* Cel7A CCDs

To gain further insight into the mobilities of Th_Cel7A and Tr_Cel7A CCD, we performed molecular dynamics (MD) simulations of these enzymes starting from their *apo* structures. According to the results of our simulations, the presence of the residue Ala384 in the Th_Cel7A loop 6 in place of Tyr371 in Tr_Cel7A is responsible for considerable differences in the dynamics of the loops enclosing the substrate binding site. In particular, residue Tyr260 (loop 5) experiences significant conformational changes in Th_Cel7A, consistent with the higher values of B-factors obtained from the crystal structures. Even more interestingly, the MD simulations show that the side chain conformation of Tyr260 plays an important role in modulating the degree of exposure of the substrate binding site in the unliganded Th_Cel7A CCD. Fig. 6A displays the time evolution of the rmsd values of the residues Tyr247 (loop 5, Tr_Cel7A) and Tyr260 (loop 5, Th_Cel7A), where we have observed two states. State 1 denotes protein conformations whose rmsd values relative to the average structure are below

3.0 Å, whereas state 2 refers to those whose rmsd are above 3.0 Å. Fig. 6B shows representative snapshots from states 1 and 2 in the region of the loops 5 and 6. The results show that the side chain of Th_Cel7A Tyr260 is substantially more mobile than the corresponding Tyr247 in Tr_Cel7A structure. In addition, Tyr260 interactions with the residues of the loop 6, mainly Ala384, are frequently broken. Whenever such interactions are broken, the binding site becomes completely exposed and the residue Tyr260 adopts a conformation that very closely resembles the one in our 2.9 Å resolution structure (Structure 1). By the contrast, the side chains of residues Tyr247 (loop 5) and Tyr371 (loop 6) in Tr_Cel7A remain constantly in contact and the loops enclosing the binding site are never found open. Even though residue Tyr247 in Tr_Cel7A is also capable of undergoing conformational changes, as shown by the peak in the Tyr247 rmsd near 60 ns, the fluctuations are not sufficient to break the contact between loops 5 and 6, thus preventing the binding site from being completely exposed (opened). Thus, the MD simulations suggest that side chain conformations of Tyr260 in loop 5 may break contacts with Ala384 in loop 6 and expose the binding site in Th_Cel7A, whereas its counterpart residues Tyr247 (loop 5) and Tyr371 (loop6) in Tr_Cel7A are never seen opened apart in the course of our simulations, thus preventing the exposure of the binding site.

Analysis of root mean square fluctuations (rmsf) of the α -carbons depicted in Figure 7 also provides valuable information on the dynamics of these enzymes. The data show that the flexibilities of most loops of Th_Cel7A are greater than those of Tr_Cel7A. In particular, the data show that loop 4 is substantially more mobile in Th_Cel7A. This is interesting, since loop 4 is the main loop that covers the tunnel and is located at the entrance of the binding site. A larger flexibility in this region might facilitate the entry of the substrate into the tight binding tunnel of the enzyme.

Discussion

Our structural characterization of the *T. harzianum* Cel7A CCD reveals that this secreted enzyme has the canonical fold of the glycosyl hydrolase family 7, with the loops of the β -sandwich domain forming typical cellobiohydrolases cellulose-binding tunnel instead of the far more open groove found in endoglucanases [19, 24]. Despite the enzyme high sequence homology with the Cel7A from *T. reesei*, Th_Cel7A, from the structural point of view, is an intermediate between Tr_Cel7A and Pc_Cel7D, with significant similarities to the structure of Te_Cel7A in terms of tunnel entrance and shape (Figs. 3 and 5).

The identified tunnel side cleft, which gives direct access to the enzyme's active site and exposes the catalytic glutamates of Th_Cel7A to the solvent, is directly dependent on the conformation of the Tyr260 side chain. The cleft is partially obstructed when this residue is solvent-exposed. The observed flexibility of Th_Cel7A Tyr260 in the absence of ligands is not unexpected. Similar effect has been previously reported for the equivalent Tyr247 in unliganded Tr_Cel7A [19, 26]. Nevertheless, our MD simulations in the absence of substrate indicate that the presence of the short side chain from Ala384 allows a nearly free motion of Tyr260 towards the inner and solvent-exposed conformations in Th_Cel7A. The movements of Tyr247 side chain in Tr_Cel7A are restrained by the interactions with Tyr371 (Ala384 in Th_Cel7A) and are accomplished simultaneously with Tyr371 conformational changes, which points towards the inner part of the tunnel when Tyr247 is in the solvent-exposed conformation and vice-versa. Despite the relative conformational freedom of Tyr247, our MD results indicate that this residue makes persistent contacts with residue Tyr371 at loop 6 of Tr_Cel7A. Thus, the existing side opening in Th_Cel7A structure when Tyr260 is shifted outwards the tunnel, in Tr_Cel7A is mostly covered either by Tyr247 or Tyr371 (Figs. 3 and 6).

Sequence analysis of cellobiohydrolases I shows that the position occupied by Ala384, although moderately conserved, displays predominantly certain types of residues, amongst which alanine or tyrosine are most frequent. Interestingly, that most of Cel7A enzymes which have a gap at loop 3, and thus exhibit a wider tunnel entrance, also contain an alanine residue, which in *T. harzianum* is observed to guarantee an additional access to the cellulose-binding tunnel. The dynamically-enhanced side opening to the substrate-binding tunnel, alluded to above, is structurally novel. While the crystallographic structure of Te_Cel7A lacks part of the side tunnel face due to a flexible loop [21], the Pc_Cel7D structure shows a completely opened tunnel [20]. This is due to a shorter loop 5 that completely exposes the tunnel and may be associated to higher rates of hydrolysis, reduced processivity, and considerable endoglucanase activity of Pc_Cel7D [20, 21].

Considering the overall tunnel shape, length and residues conservation, it is expected that the sugar contacts described for Tr_Cel7A [24] are maintained in Th_Cel7A. Except for Gln101, all the residues interacting with the glucosyl units in *T. reesei* are conserved in *T. harzianum*, including the four tryptophan residues that account for the formation of sites -7, -4, -2, and +1 in Tr_Cel7A, which along with tyrosine residues interact with the hydrophobic glucosyl faces of the sugar units and presumably act as lubricating agents facilitating processive action of the enzyme [24]. It is important to mention that the corresponding Tr_Cel7A residue Q101 is not present in Th_Cel7A, the absence of specific H-bonds at sites -6 and -5 lowers the binding affinity of Th_Cel7A to the substrate, which is consistent with the higher K_M experimentally observed for Th_Cel7A (as compared to Tr_Cel7A). However, the same amino acid substitution might be favorable for the sliding of the cellulose chain into the active site. Moreover, due to the less restricted conformational mobility of Th_Cel7A Tyr260 which is equivalent to more conformationally-constrained Tr_Cel7A Tyr247, the specific H-bond interactions with the substrate at site -2 might be weaker in Th_Cel7A, contributing to the

sliding of the sugar units within the Th_Cel7A catalytic tunnel. This is consistent with the higher k_{cat} value, as observed in our enzymatic studies of the enzyme.

Our activity studies show that the cellobiohydrolase from *T. harzianum* is active against cellulosic substrates with different degrees of crystallinity and polymerization. Th_Cel7A was efficient in depolymerization of highly crystalline BMCC, less so for crystalline and insoluble Avicel and Sigmacell and poorly efficient in CMC hydrolysis (Fig. 2). The obtained specific hydrolysis profile of Th_Cel7A in general correlates well with the published data for Tr_Cel7A [32]. Smaller conversion yields of Avicel and Sigmacell might be related to several factors, such as lower accessibility of these substrates as compared to BMCC; their dehydration/rehydration prior to experimental studies; significantly different degree of polymerization; different ratio between $I\alpha$ and $I\beta$ -types of crystalline cellulose between these substrates; higher amorphous content and a presence of the minor hemicellulose fraction in their composition [33; 34]. BMCC was used in our experiments in a never-dried form, whereas both Avicel and Sigmacell have been purchased in a dried form and re-hydrated prior to their use as substrates for enzymatic hydrolysis. It is known that cycles of dehydration and re-hydration of cellulosic materials have significant effect on their porosity and accessibility, which both tend to quite strongly decrease after dehydration/rehydration [34, 35]. This renders substrates that are less susceptible to enzymatic cleavage [36, 37]. Furthermore, aggregation of cellulose microcrystals, entangled structures, non-cellulosic components attached to the cellulose as well as amorphous regions of the cellulose were recently shown to represent obstacles for processive cellobiohydrolase action and to significantly slow down Tr_Cel7A cellulose hydrolysis [38, 39]. It is known that native crystalline cellulose is composed of two different crystalline components, cellulose $I\alpha$ and cellulose $I\beta$ [40]. Bacterial type celluloses are mostly composed of the $I\alpha$ (triclinic) form, whereas $I\beta$ (monoclinic) form of cellulose is dominant in the cotton/wood type of biomass [40-42]. Less thermodynamically

stable I α form of cellulose was shown to be more susceptible to enzymatic degradation by *Trichoderma* cellulases than I β cellulose and in the mixed substrates I α form of cellulose has been degraded first by cellulases [42]. Finally, xylose and xylooligosaccharides were found to be strong inhibitors of cellulose hydrolysis by cellulases and might also have some detrimental effect on the Th_Cel7A conversion of Avicel and Sigmacell [43]. The presence of hemicellulose was demonstrated to be detrimental for biomass hydrolysis using commercial enzymatic preparations, a problem which can be alleviated by addition of xylanase activity [44, 45].

For hydrolysis of all three crystalline substrates (BMCC, Avicel and Sigmacell-20) the presence of CBM is beneficial and leads to the higher specific activity and conversion yields. Th_Cel7A, as a processive enzyme, is ineffective in CMC hydrolysis, a functional feature shared by Tr_Cel7A which is known to have low CMCase activity [46, 47]. The presence of CBM makes no difference for hydrolysis of CMC (Fig. 2) and the tunnel-like shape of processive Th_Cel7A makes it not very effective for hydrolysis of this soluble and substituted non-crystalline substrate.

In agreement with our kinetic analysis, the Th_Cel7A catalytic rates are consistently higher than the reported activities for Tr_Cel7A [15, 32] using both glucoside and/or lactoside substrates. This seems to be consistent with the more open catalytic active cleft of the enzyme. Te_Cel7A and Pc_Cel7D, both presenting shorter active site loops and thus more open and accessible binding site, have higher k_{cat} values as compared to Tr_Cel7A, particularly for short chain oligosaccharides [20, 21]. More open substrate binding site also seems to be positively correlated with the decreased processivity and increased endo-glucanase activity of the enzyme [20, 48]. This allows us to speculate that Th_Cel7A will probably have lesser processivity but higher additional endo-glucanase activity as compared

to Tr_Cel7A. However its endo-glucanase activity probably will be smaller than that of Pc_Cel7D [48].

To date, *Melanocarpus albomyces* Cel7B (Ma_Cel7B) is the only enzyme which was structurally characterized in its native form in the complex with oligosaccharides [22]. (The Tr_Cel7A-cellooligosaccharide complexes were determined using catalytically deficient mutants [19]). The Ma_Cel7B is strongly inhibition by the product of enzymatic reaction, cellobiose ($K_d = 6$ mM), which probably was an important factor for success in obtaining of the enzyme crystal structures in the complex with oligosaccharides [22]. The Th_Cel7A enzyme shows very weak inhibition by the end product of reaction, cellobiose. Unfortunately, all our attempts to solve its native X-ray structure in complex with oligosaccharides failed and production of the catalytically deficient mutants was impeded by the necessity to produce recombinant enzymes expressed in its soluble form using an appropriate expression system. Although the molecular mechanism underlying the low inhibition of Th_Cel7A by cellobiose still remains elusive, this feature of the enzyme might be important for its applications under industrial settings.

Materials and methods

Expression and purification of Th_Cel7A

The expression and purification of the wild-type Cel7A from *T. harzianum* IOC 3844 was carried out following the protocol described elsewhere [49]. In short, the cells were cultivated for 48 h at 25 °C and used to inoculate a 7.5 L vessel BioFlo 115 Fermenter (New Brunswick Scientific Co, USA) cultured for 5 days at 28 °C, pH 5.0 and 200 rpm. The aeration rate was adjusted to keep the dissolved O₂ levels below 60 % in the culture media. The enzyme was purified using the columns DEAE-Sephadex (A-50, Sigma) and Phenyl Sepharose (GE). pH

and temperature optimum for Th_ Cel7A enzymatic activity are reached at pH 5.0 and 50 °C, respectively [49]. Possible endoglucanase contamination in the protein preparations was analyzed using activity assays against hydroxyethyl cellulose (HEC, Sigma) [26], with no significant activity detected. The purified protein was dialyzed overnight against 50 mM Tris-HCl pH 7.0 and concentrated to 8 mg/ml prior crystallization. Protein concentration was monitored by measuring the absorbance at 280 nm.

Enzyme activity assays

1% suspensions (w/v) of substrates Avicel® PH101 (Sigma), Sigmacell 20 (Sigma) and a solution of carboxymethyl cellulose (CMC, Sigma) were prepared according to the manufacturer's instructions and used for the analysis of the enzyme activity on insoluble and soluble substrates, respectively. Bacterial microcrystalline cellulose (BMCC) was prepared from commercially available *Acetobacter xylinum* cellulose, as described previously [50].

Avicel® and Sigmacel 20 have degree of polymerization (DP) between 150 and 500, are mostly composed of crystalline cellulose (85-95%), with minor fractions of amorphous cellulose and up to 5% of xylose, mannose, galactose and arabinose. The main difference between these two substrates is their crystallinity, which is approximately 92% for Avicel and 85% for Sigmacel 20, respectively [33, 51].

Soluble ionic-substituted non-crystalline carboxymethyl cellulose (CMC) is often used for determining activity of endoglucanases, which randomly cleave intramolecular β -1,4-glucosidic bonds, resulting in a dramatic reduction in the degree of polymerization (DP) and specific viscosity of CMC [33]. CMC is not a very good substrate for cellobiohydrolases since due to the steric confinement imposed by the active site tunnel only non-substituted glucose units are accessible for enzymatic hydrolysis [19, 33, 47]. Therefore, an exoglucanase will trim off the unsubstituted residues until further progress along the CMC is prevented by the

Accepted Article

carboxymethyl groups. Finally, BMCC has a low DP (100 to 150), higher crystallinity, accessibility and purity as compared to Avicel and Sigmacel [33]. This highly crystalline cellulose with a simple morphology has been used as a model substrate to study the mechanism of crystal erosion [42] and cellobiohydrolase processivity [48]. The crystallinity index of BMCC measured by X-ray diffraction was reported to be 95.2% [51].

Activity assays for both full-length Th_Cel7A and its CCD were carried out in 50 mM citrate buffer at the optimum pH and temperature of the enzyme (pH 5.0 at 50 °C) [49]. The reaction mixture of 15uM either of full-length or CCD was incubated for 1h for the soluble and the insoluble substrates. The conversion yields were measured for 24 hours under the same conditions as described above. No agitation was applied and to avoid inhibition of Th_Cel7A by the cellobiose over time, 10uM of *Aspergillus niger* β -glucosidase (Novozymes SP188; Novo Nordisk A/S, Bagsvaerd, Denmark) was added to the reaction mixtures [38]. To ensure that the Th_Cel7A was totally active during our experiments, we monitored its activity during 24 hours under the same experimental conditions. No changes in the specific activity were observed (data not shown). The quantification of released reducing sugars was done with the DNS reagent [15]. All enzyme activities were expressed in International Units, i.e., one unit of activity corresponding to the quantity of enzyme hydrolyzing 1 μ mol of substrate or releasing 1 μ mol of reducing sugars (in glucose equivalents) per minute.

The kinetic assays were carried out in 50 mM of sodium acetate buffer (pH 5.0) using 700 nM of enzyme with pNPC and cNPL (Sigma) as soluble substrates in 12 different final concentrations (from 0 to 15 mM). The reaction started by the addition of pNPC or cNPL at 50°C for 10 min and it was stopped with 1.0 M sodium carbonate. The amount of p-or c-nitrophenol released was measured at λ = 405nm (spectrophotometer U-2001 Hitachi) and quantified according to the respective standard curves. All experiments were carried out in triplicate.

The effects of Cel7A inhibition were analyzed as described above using pNPC as substrate and different concentrations of cellobiose (0 to 40 mM) at pH 5.0, 50°C. The kinetic parameters were obtained from a Michaelis-Menten plot V_0 versus pNPC or cNPL by a rectangular hyperbola curve fitting. Both linear and non-linear fits were carried out using GraphPad Prism version 5.04 for Windows, GraphPad Software, La Jolla California, USA (www.graphpad.com).

X-ray structure determination

The X-ray data collection was carried out on beamline MX2 at the Synchrotron Light National Laboratory (LNLS, Campinas, Brazil) [52]. The first X-ray diffraction dataset was collected at 2.9 Å resolution from a needle-shaped crystal obtained by the hanging-drop vapor diffusion technique in the presence of 25% PEG 4000 grown using 35% PEG 4000 as a precipitant at pH=7.0 [53] (data not shown). The second dataset at 1.67 Å resolution was collected from optimized three-dimensional crystals grown in 25% PEG 4000 and 5 mM CaCl_2 . The diffraction data were processed using the program HKL2000 [54]. The crystal structures of Cel7A catalytic core domain (residues 18-449) were determined by molecular replacement with the program Phaser [55] using the CCD of Cel7A from *T. reesei* (1CEL) as template (82% sequence identity over 494 amino acid of the CCD). The $2Fo - Fc$ map calculated from the unique solution of the top rotation and translation searches showed a clear and contiguous electron density for the protein and active site molecules not included in the search model. Model building and refinement was carried out using the program Coot [56]. Cycles of restrained refinement with the program Refmac5 [57] were initially carried out using overall temperature factors and later on isotropic temperature factors. One cycle of simulated-annealing was applied with the program Phenix [58]. Water molecules were added using the program Arp/wArp [59] and the stereochemical quality of the protein complex was

validated with the program MolProbity [60]. Statistics of the refined structures are presented in Table 2. The visualization and all representations of the structures were carried out using the program PyMol (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC).

Molecular dynamics simulations

The initial coordinates for MD simulations of the CCD of *T. harzianum* and *T. reesei* were taken from the obtained crystallographic structure and from the Protein Data Bank (PDB entry **8CEL**), respectively. Hydrogen atoms were added and special attention was paid to the ionizable residues, whose protonation states were estimated using the server H++ at pH=6 [61, 62]. All substrate molecules were removed as well as the glucosamine groups. The structures were solvated by water molecules within a rectangular box such that the solvent layer was at least 16 Å thick using Packmol [63]. Sodium and chloride ions were added to a total concentration of 0.10 mol/L in order to render the systems electrically neutral. The constructed systems consisted of approximately 60,000 atoms.

The simulations were performed in the ensemble NpT at 1 atm and 310 K using the Langevin/Nosé-Hoover barostat and the Langevin thermostat for pressure and temperature control, respectively. The CHARMM force field [64] was used for the protein and the TIP3P model for water molecules [65]. The van der Waals interactions were truncated at a cut off radius of 12 Å and electrostatics were evaluated using the particle mesh Ewald method [66]. All bonds involving apolar hydrogen atoms were constrained at their equilibrium lengths using the SHAKE algorithm [67]. The time step for the numerical integration of the equations of motion was 2.0 fs.

The following steps were performed to prepare the systems for the production runs. (1) 1000 steps of conjugated gradient (CG) energy minimization followed by a 200 ps MD run

with all protein atoms harmonically restrained; (2) 1000 steps of CG followed by 200 ps MD with only the protein alpha carbons harmonically restrained; (3) 600 ps MD without any restraints. After these initial steps, 60 ns production trajectories were generated and considered for analyses. For each system, three independent simulations were performed. All simulations were performed with NAMD [68].

Accession numbers

The coordinates and structure factors of Cel7A CCD from *T. harzianum* have been deposited at the Protein Data Bank (PDB codes **2YOK** and **R2YOKSF**, respectively).

Acknowledgements

We thank the support from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) via grants 2008/56255-9, 2009/54035-4 and 2010/08680-2; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) via grants 490022/2009-0, 471834/2009-2, 550985/2010-7 and 151951/2008-0 and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

References

1. Cao, L., Bala, G., Caldeira, K., Nemani, R. & Ban-Weiss, G. (2010) Importance of carbon dioxide physiological forcing to future climate change, *Proceedings of the National Academy of Sciences of the United States of America*. **107**, 9513-9518.
2. Idso, S. B. (1989) *Carbon dioxide and global change : earth in transition*, IBR Press, Tempe, Ariz, U.S.A. (631 E. Laguna Dr., Tempe 85282).
3. Koerner, C. & Bazzaz, F. A. (1996) *Carbon dioxide, populations, and communities*, Academic Press, San Diego.
4. von Ossowski, I., Stahlberg, J., Koivula, A., Piens, K., Becker, D., Boer, H., Harle, R., Harris, M., Divne, C., Mahdi, S., Zhao, Y. X., Driguez, H., Claeyssens, M., Sinnott, M. L. & Teeri, T. T. (2003) Engineering the exo-loop of *Trichoderma reesei* cellobiohydrolase, Cel7A. A comparison with *Phanerochaete chrysosporium* Cel7D, *Journal of Molecular Biology*. **333**, 817-829.
5. Gao, J. M., Weng, H. B., Zhu, D. H., Yuan, M. X., Guan, F. X. & Xi, Y. (2008) Production and characterization of cellulolytic enzymes from the thermoacidophilic fungal *Aspergillus*

- terreus M11 under solid-state cultivation of corn stover, *Bioresource Technology*. **99**, 7623-7629.
6. Zhou, J., Wang, Y. H., Chu, J., Zhuang, Y. P., Zhang, S. L. & Yin, P. (2008) Identification and purification of the main components of cellulases from a mutant strain of *Trichoderma viride* T 100-14, *Bioresource Technology*. **99**, 6826-6833.
7. Ahamed, A. & Vermette, P. (2008) Culture-based strategies to enhance cellulase enzyme production from *Trichoderma reesei* RUT-C30 in bioreactor culture conditions, *Biochemical Engineering Journal*. **40**, 399-407.
8. Kleywegt, G. J., Zou, J. Y., Divne, C., Davies, G. J., Sinning, I., Stahlberg, J., Reinikainen, T., Srisodsuk, M., Teeri, T. T. & Jones, T. A. (1997) The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 angstrom resolution, and a comparison with related enzymes, *Journal of Molecular Biology*. **272**, 383-397.
9. Coutinho, P. M. & Henrissat, B. (1999) Carbohydrate-active enzymes: An integrated database approach in *Recent Advances in Carbohydrate Bioengineering* (Gilbert, H. J., Davies, G. J., Henrissat, B. & Svensson, B., eds) pp. 3-12.
10. Davies, G. J., Wilson, K. S. & Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases, *Biochemical Journal*. **321**, 557-559.
11. Lynd, L. R., Weimer, P. J., van Zyl, W. H. & Pretorius, I. S. (2002) Microbial cellulose utilization: Fundamentals and biotechnology (vol 66, pg 506, 2002), *Microbiology and Molecular Biology Reviews*. **66**, 739-739.
12. Fagerstam, L. G. & Pettersson, L. G. (1980) The 1.4-Beta-Glucan Cellobiohydrolases of *Trichoderma-Reesei* Qm-9414 - a New Type of Cellulolytic Synergism, *Febs Letters*. **119**, 97-100.
13. Nummi, M., Niku-Paavola, M. L., Lappalainen, A., Enari, T. M. & Raunio, V. (1983) Cellobiohydrolase from *Trichoderma reesei*, *The Biochemical journal*. **215**, 677-83.
14. Stals, I., Sandra, K., Geysens, S., Contreras, R., Van Beeumen, J. & Claeyssens, M. (2004) Factors influencing glycosylation of *Trichoderma reesei* cellulases. I: Postsecretorial changes of the O- and N-glycosylation pattern of Cel7A, *Glycobiology*. **14**, 713-724.
15. Tomme, P., Vantilbeurgh, H., Pettersson, G., Vandamme, J., Vandekerckhove, J., Knowles, J., Teeri, T. & Claeyssens, M. (1988) Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis, *European Journal of Biochemistry*. **170**, 575-581.
16. Harrison, M. J., Nouwens, A. S., Jardine, D. R., Zachara, N. E., Gooley, A. A., Nevalainen, H. & Packer, N. H. (1998) Modified glycosylation of cellobiohydrolase I from a high cellulase-producing mutant strain of *Trichoderma reesei*, *European Journal of Biochemistry*. **256**, 119-127.
17. Medve, J., Lee, D. & Tjerneld, F. (1998) Ion-exchange chromatographic purification and quantitative analysis of *Trichoderma reesei* cellulases cellobiohydrolase I, II and endoglucanase II by fast protein liquid chromatography, *Journal of Chromatography A*. **808**, 153-165.
18. Teeri, T. T., Lehtovaara, P., Kauppinen, S., Salovuori, I. & Knowles, J. (1987) Homologous domains in *Trichoderma reesei* cellulolytic enzymes: gene sequence and expression of cellobiohydrolase I, *Gene*. **51**, 43-52.
19. Divne, C., Stahlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K. C., Teeri, T. T. & Jones, T. A. (1994) The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*, *Science*. **265**, 524-528.
20. Munoz, I.G., Ubhayasekera, W., Henriksson, H., Szabo, I., Pettersson, G., Johansson, G., Mowbray, S.L. & Stahlberg, J. (2001) Family 7 cellobiohydrolases from *Phanerochaete chrysosporium*: crystal structure of the catalytic module of Cel7D (CBH58) at 1.32 Å

resolution and homology models of the isozymes, *Journal of Molecular Biology*. **314**, 1097-1111.

21. Grassick, A., Murray, P.G., Thompson, R., Collins, C.M., Byrnes, L., Birrane, G., Higgins, T.M. & Tuohy, M.G. (2004) Three-dimensional structure of a thermostable native cellobiohydrolase, CBH IB, and molecular characterization of the cel7 gene from the filamentous fungus, *Talaromyces emersonii* *European Journal of Biochemistry*. **271**, 4495–450622.
22. Parkkinen, T., Koivula, A. & Rouvinen, J. (2008) Crystal structures of *Melanocarpus albomyces* cellobiohydrolase Cel7B in complex with cello-oligomers show high flexibility in the substrate binding. *Protein Science*. **17**, 1383-1394
23. Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J. & Gronenborn, A. M. (1989) Determination of the 3-dimensional solution structure of the C-terminal domain of cellobiohydrolase-I from *Trichoderma reesei* - a study using nuclear magnetic-resonance and hybrid distance geometry dynamical simulated annealing, *Biochemistry*. **28**, 7241-7257.
24. Divne, C., Stahlberg, J., Teeri, T. T. & Jones, T. A. (1998) High-resolution crystal structures reveal how a cellulose chain is bound in the 50 angstrom long tunnel of cellobiohydrolase I from *Trichoderma reesei*, *Journal of Molecular Biology*. **275**, 309-325.
25. Barr, B. K., Hsieh, Y. L., Ganem, B. & Wilson, D. B. (1996) Identification of two functionally different classes of exocellulases, *Biochemistry*. **35**, 586-592.
26. Stahlberg, J., Divne, C., Koivula, A., Piens, K., Claeysens, M., Teeri, T. T. & Jones, T. A. (1996) Activity studies and crystal structures of catalytically deficient mutants of cellobiohydrolase I from *Trichoderma reesei*, *Journal of Molecular Biology*. **264**, 337-349.
27. Castro, A. M., Pedro, K. C., Cruz, J. C., Ferreira, M. C., Leite, S. G. & Pereira, N., Jr. (2010) *Trichoderma harzianum* IOC-4038: A promising strain for the production of a cellulolytic complex with significant beta-glucosidase activity from sugarcane bagasse cellulignin, *Appl Biochem Biotechnol*. **162**, 2111-2122
28. Claeysens, M., Van Tilbeurgh, H., Tomme, P., Wood, T. M. & McRae, S. I. (1989) Fungal cellulase systems. Comparison of the specificities of the cellobiohydrolases isolated from *Penicillium pinophilum* and *Trichoderma reesei*, *The Biochemical journal*. **261**, 819-25.
29. Voutilainen, S. P., Puranen, T., Siika-Aho, M., Lappalainen, A., Alapuranen, M., Kallio, J., Hooman, S., Viikari, L., Vehmaanpera, J. & Koivula, A. (2008) Cloning, expression, and characterization of novel thermostable family 7 cellobiohydrolases, *Biotechnology and Bioengineering*. **101**, 515-28.
30. Deshpande, M. V., Eriksson, K.-E. & Gran Pettersson, L. (1984) An assay for selective determination of exo-1,4,- β -glucanases in a mixture of cellulolytic enzymes *Analytical Biochemistry*. **138**, 481-487.
31. Vonhoff, S., Piens, K., Pipelier, M., Braet, C., Claeysens, M. & Vasella, A. (1999) Inhibition of cellobiohydrolases from *Trichoderma reesei*. Synthesis and evaluation of some glucose-, cellobiose-, and cellotriose-derived hydroxymolactams and imidazoles, *Helvetica Chimica Acta*. **82**, 963–980.
32. Jager, G., Wu, Z., Garschhammer, K., Engel, P., Klement, T., Rinaldi, R., Spiess, A. C. & Buchs, J. (2010) Practical screening of purified cellobiohydrolases and endoglucanases with alpha-cellulose and specification of hydrodynamics, *Biotechnol Biofuels*. **3**, 3-18.
33. Zhang, Y. H. P., Himmel, M. E. & Mielenz, J. R. (2006) Outlook for cellulase improvement: Screening and selection strategies, *Biotechnology Advances*. **24**, 452-481.
34. Bragatto, J., Segato, F., Cota, J., Mello, D. B. Oliveira, M.M., Buckeridge, M.S., Squina, F.M. & Driemeier, C. (2012) Insights on how the activity of an endoglucanase is affected by physical properties of insoluble celluloses, *J. Phys. Chem. B* **116**, 6128–6136.

35. Esteghlalian, A.R., Bilodeau, M., Mansfield S.D. & Saddler, J.N. (2001) Do enzymatic hydrolyzability and simons' stain reflect the changes in the accessibility of lignocellulosic substrates to cellulase enzymes? *Biotechnology Progress* **17**, 1049–1054.
36. Arantes, V. & Saddler, J.N. (2010) Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnology for Biofuels* **3**, 4.
37. Arantes, V. & Saddler, J.N. (2011) Cellulose accessibility limits the effectiveness of minimum cellulase loading on the efficient hydrolysis of pretreated lignocellulosic substrates. *Biotechnology for Biofuels* **4**, 3.
38. Jalak, J. & Vaeljamae, P. (2010) Mechanism of initial rapid rate retardation in cellobiohydrolase catalyzed cellulose hydrolysis, *Biotechnology and Bioengineering*. **106**, 871-883.
39. Jürgen Jalak, Mihhail Kurašin, Hele Teugjas and Priit Väljamäe (2012) Endo-exo synergism in cellulose hydrolysis revisited *J. Biol. Chem.* Epub ahead of print, doi:10.1074/jbc.M112.381624.
40. Atalla, R.H. and VanderHart, D.L. (1984) Native cellulose: a composite of two distinct crystalline forms, *Science*, **223**, 283-285.
41. Wiley, J.H. and Atalla, R.H. (1987) Band assignments in the Raman spectra of celluloses, *Carbohydrate Research*. **160**, 113-129.
42. Hayashi, N., Sugiyama, J., Okano, T. & Ishihara, M. (1998) The enzymatic susceptibility of cellulose microfibrils of the algal-bacterial type and the cotton-ramie type, *Carbohydrate Research* **305**, 261-269.
43. Qing, Q., Yang, B. & Wyman, C. E. (2010) Xylooligomers are strong inhibitors of cellulose hydrolysis by enzymes, *Bioresource Technology*. **101**, 9624-9630.
44. Berlin, A., Gilkes, N., Kilburn, D., Bura, R., Markov, A., Skomarovsky, A., Okunev, O., Gusakov, A., Maximenko, V., Gregg, D., Sinitsyn, A. & Saddler, J. (2005) Evaluation of novel fungal cellulase preparations for ability to hydrolyze softwood substrates - evidence for the role of accessory enzymes, *Enzyme and Microbial Technology*. **37**, 175-184.
45. Varnai, A., Siika-aho, M. & Viikari, L. (2010) Restriction of the enzymatic hydrolysis of steam-pretreated spruce by lignin and hemicellulose, *Enzyme and Microbial Technology*. **46**, 185-193.
46. Irwin, D. C., Spezio, M., Walker, L. P. & Wilson, D. B. (1993) Activity studies of 8 purified cellulases - specificity, synergism, and binding domain effects, *Biotechnology and Bioengineering*. **42**, 1002-1013.
47. Penttilä, M. E., Andre, L., Lehtovaara, P., Bailey, M. J., Teeri, T. T. & Knowles, J. K. C. (1988). Efficient secretion of two fungal cellobiohydrolases by *Saccharomyces cerevisiae*. *Gene*, **63**, 103-112
48. Kurasin, M. & Valjamae, P. Processivity of cellobiohydrolases is limited by the substrate (2011) *Journal of Biological Chemistry*. **286**, 169–177
49. Colussi, F., Serpa, V., Delabona, P. S., Manzone, L. R., Voltatodio, M. L., Alves, R., Mello, B. L., Pereira, N., Farinas, C. S., Golubev, A. M., Santos, M. A. M. & Polikarpov, I. (2011) Purification, and Biochemical and Biophysical Characterization of Cellobiohydrolase I from *Trichoderma harzianum* IOC 3844, *Journal of Microbiology and Biotechnology*. **21**, 808-817.
50. Valjamae, P., Sild, V., Nutt, A., Pettersson, G. & Johansson, G. (1999) Acid hydrolysis of bacterial cellulose reveals different modes of synergistic action between cellobiohydrolase I and endoglucanase I, *European Journal of Biochemistry*. **266**, 327-334.
51. Park, S., Johnson, D. K., Ishizawa, C. I., Parilla, P. A. & Davis, M. F. (2009) Measuring the crystallinity index of cellulose by solid state C-13 nuclear magnetic resonance, *Cellulose*. **16**, 641-647.

52. Guimaraes, B. G., Sanfelici, L., Neuenschwander, R. T., Rodrigues, F., Grizolli, W. C., Raulik, M. A., Piton, J. R., Meyer, B. C., Nascimento, A. S. & Polikarpov, I. (2009) The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS, *Journal of Synchrotron Radiation*. **16**, 69-75.
53. Colussi, F., Textor, L. C., Serpa, V., Maeda, R. N., Pereira, N. & Polikarpov, I. (2010) Purification, crystallization and preliminary crystallographic analysis of the catalytic domain of the extracellular cellulase CBHI from *Trichoderma harzianum*, *Acta Crystallographica Section F-Structural Biology and Crystallization Communications*. **66**, 1041-1044.
54. Otwinowski, Z. & Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Macromolecular Crystallography, Pt A*. **276**, 307-326.
55. Storoni, L. C., McCoy, A. J. & Read, R. J. (2004) Likelihood-enhanced fast rotation functions, *Acta Crystallographica Section D-Biological Crystallography*. **60**, 432-438.
56. Emsley, P. & Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta crystallographica Section D, Biological crystallography*. **60**, 2126-32.
57. Skubak, P., Murshudov, G.N. & Pannu, N.S. (2004) Direct incorporation of experimental phase information in model refinement *Acta Cryst. Sec. D - Biol. Cryst.***60**, 2196-2201
58. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta crystallographica Section D, Biological crystallography*. **66**, 213-21.
59. Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. (2001) ARP/wARP and molecular replacement, *Acta Crystallographica Section D-Biological Crystallography*. **57**, 1445-1450.
60. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., Snoeyink, J., Richardson, J. S. & Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids, *Nucleic Acids Research*. **35**, W375-W383.
61. Gordon, J. C., Myers, J. B., Folta, T., Shoja, V., Heath, L. S. & Onufriev, A. (2005) H⁺⁺: a server for estimating pK_as and adding missing hydrogens to macromolecules, *Nucleic Acids Res.* **33**, W368-371.
62. Myers, J., Grothaus, G., Narayanan, S. & Onufriev, A. (2006) A simple clustering algorithm can be accurate enough for use in calculations of pK_as in macromolecules, *Proteins*. **63**, 928-938.
63. Martinez, L., Andrade, R., Birgin, E. G. & Martinez, J. M. (2009) PACKMOL: a package for building initial configurations for molecular dynamics simulations, *Journal of Computational Chemistry*. **30**, 2157-2164.
64. MacKerell, A. D., Bashford, D., Bellott, Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wirkiewicz-Kuczera, J., Yin, D. & Karplus, M. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins, *The Journal of Physical Chemistry B*. **102**, 3586-3616.
65. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water, *The Journal of Chemical Physics*. **79**, 926-935.
66. Darden, T., York, D. & Pedersen, L. (1993) Particle mesh Ewald: An N . log(N) method for Ewald sums in large systems, *The Journal of Chemical Physics*. **98**, 10089-10092

67. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1997) Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes, *Journal of Computational Chemistry*. **23**, 327-341.
68. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L. & Schulten, K. (2005) Scalable molecular dynamics with NAMD, *Journal of Computational Chemistry*. **26**, 1781-1802.

Tables

Table 1. Kinetic parameters of Th_Cel7A.

Activity measurements have been done using 700nM enzyme in 50 mM sodium acetate buffer pH 5.0 at 50 °C.

	pNPC			cNPL			Cellobiose
	K_m (mM)	V_{max} (mM.min ⁻¹)	K_{cat} (min ⁻¹)	K_m (mM)	V_{max} (mM.min ⁻¹)	K_{cat} (min ⁻¹)	K_i (mM)
Th_Cel7A	3.4±0.5	0.016±0.001	23.2 ± 1.1	3.7±1.2	0.031±0.004	44.6 ± 4.7	7.2

Table 2. X-ray refinement statistics

<i>Diffraction data</i>	
Space group	P2 ₁ 2 ₁ 2 ₁
Unit cell parameters (Å)	a=56.7 b= 69.2 c=90.6
Resolution limits (Å) ^a	45.3-1.67 (1.71-1.67)
Number of unique reflections	42070 (4077)
Redundancy	6.1 (5.5)
Completeness (%)	99.8 (98.8)
Rmerge (%) ^b	9.1 (32.9)
I/sigI (I) ^b	24.4 (4.7)
<i>Refinement</i>	
Rfactor/Rfree (%)	13.1/16.2
Number of atoms	
Protein	3281
N-acetyl-D-glucosamine	42
PEG	24
Water	694
B factors (Å ²)	
Protein	10.5
N-acetyl-D-glucosamine	28.3
PEG	33.0
Water	27.2
R.m.s.d. ^c	
Bond length (Å)	0.011
Bond angle (°)	1.494
Ramachandran (preferred/allowed)	98.1/1.9

^a Values in parentheses are for the highest resolution shell.

^b $R_{\text{merge}} = \sum |I_j - \langle I \rangle| / \sum \langle I_j \rangle$, where I_j is the observed intensity of an individual reflection and $\langle I \rangle$ is the average intensity of that reflection.

^c Root-mean-square-deviation.

Fig. 1. Structure of CBHI from *Trichoderma harzianum*. (A) Structural organization.

Predicted N-glycosylation sites are identified and O-glycosylated linker is shown. (B) Ribbon diagram of the crystallographic structure of the CCD from Th_Cel7A highlighting the loops composing the tunnel. Observed glycosylation sites at N126, N283 and N397 are displayed.

Fig. 2. Activities of purified Th_CBHI and its CCD on different cellulosic substrates. Assays were performed using 50 mM sodium citrate buffer pH 5.0 at 50 °C. (A) Specific Activities of Th_CBHI, (B) % of substrate conversion in 24 hours. (C) Time course of substrate conversion by Th_CBHI and (D) by its CCD. (Please, note the change in scale between (C) and (D)).

T_harzianum	18	EQVCTQQAETHPPLTWQKCTA.S.GCTPQQGSVVLDA ^{L3} NWRWTHD ^{L3} TKSTTN
T_emersonii	1	EQAGTATAENHPPLTWQE ^{L3} CTA.PGSC ^{L3} TTQNGAVVLDA ^{L3} NWRWVHDVNGYTN
P_chrysosporium	1	EQAGTNTAENHPQLQSQQCTT.SGGCKPLSTKVVLD ^{L3} SNWRVHSTSGGYTN
T_reesei	1	ESACTLQSETHPPLTWQKCSSGG.TCTQQTGSVVIDA ^{L3} NWRWTHATNSS ^{L3} TN
T_harzianum	66	CYDGN ^{L3} TWSS ^{L3} TLCPDDATCAKNCCLDGANYS ^{L3} SGTYGV ^{L3} TTSGDALTLQFV ^{L3} TA.
T_emersonii	50	CYTGNTWDPT ^{L3} YCPDDE ^{L3} TCAQNCALDGADYEGTYGV ^{L3} TS ^{L3} SGSSLKLN ^{L3} FV ^{L3} TG.
P_chrysosporium	50	CYTGN ^{L3} EWDTSLCPDGK ^{L3} TCAANCALDGADYS ^{L3} SGTYGI ^{L3} STGTALT ^{L3} LK ^{L3} FV ^{L3} TG.
T_reesei	50	CYDGN ^{L3} TWSS ^{L3} TLCPDNE ^{L3} TCAKNCCLDGAA ^{L3} YASTYGV ^{L3} TS ^{L3} SGNSLSID ^{L3} FV ^{L3} TQS
T_harzianum	115	..SNVGSRLYL ^{L3} MANDSTYQEFTLSGNEF ^{L3} SFDVDV ^{L3} SQLP ^{L3} CGLN ^{L3} GALYFV ^{L3} SM
T_emersonii	99	..SNVGSRLYL ^{L3} LQDDSTYQIFKLLNREF ^{L3} SFDVDV ^{L3} SNL ^{L3} PCGLN ^{L3} GALYFV ^{L3} AM
P_chrysosporium	99	..SNVGSRLYL ^{L3} ADDTHYQLLKL ^{L3} NQEFTFDVDM ^{L3} SNL ^{L3} PCGLN ^{L3} GALYLS ^{L3} AM
T_reesei	100	AQKNVGARLYL ^{L3} MASDTYQEFTLLGNEF ^{L3} SFDVDV ^{L3} SQLP ^{L3} CGLN ^{L3} GALYFV ^{L3} SM
T_harzianum	163	DADGGQSKYPGNAAGAKYGTGYCDSQCP ^{L3} RD ^{L3} LKFINGQANVEG ^{L3} WEPSS ^{L3} NNA
T_emersonii	147	DADGGVSKYPNNKAGAKYGTGYCDSQCP ^{L3} RD ^{L3} LKFIDGEANVEG ^{L3} WQPS.snn
P_chrysosporium	147	DADGGMSKYPGNKAGAKYGTGYCDSQCP ^{L3} KDI ^{L3} KFINGEANVGNWTETGS ^{L3} N.
T_reesei	150	DADGGVSKYPTNTAGAKYGTGYCDSQCP ^{L3} RD ^{L3} LKFINGQANVEG ^{L3} WEPSS ^{L3} N
T_harzianum	213	N.TGVGGHSGCCSEMDIWEANSISEALTPHP ^{L5} CTTVGQ ^{L5} TMCSGDS ^{L5} CGGTYS
T_emersonii	196	anTGIGDHSGCCAE ^{L5} MDVWEANSISNAVTPHP ^{L5} CDTPGQ ^{L5} TMCSGDD ^{L5} CGGTYS
P_chrysosporium	196	..TGTSYGTCCSEMDIWEANNDAAFTPH ^{L5} PCTTGQ ^{L5} TRCSGDDCA....
T_reesei	200	N.TGIGGHSGCCSEMDIWQANSISEALTPHP ^{L5} CTTVGQEICEG ^{L5} DGCGGTYS
T_harzianum	262	NDRYGGTCDPDGCDWNPYRLGNTSFY ^{L5} FGSGSFALD ^{L5} TTK ^{L5} KLTV ^{L5} TVTQFAT..
T_emersonii	246	NDRYAGTCDPDGCDFN ^{L5} YRMGNTSFY ^{L5} FG..KIIDTTK ^{L5} PFTV ^{L5} TVTQFLTDD
P_chrysosporium	240	..RNTGLCDGDGCDFNSFRMGDKTFLGK ^{L5} ..MTVDTSK ^{L5} PF ^{L5} TV ^{L5} TVTQFLTND
T_reesei	249	DNRYGGTCDPDGCDWNPYRLGNTSFY ^{L5} FGSGSFTLD ^{L5} TTK ^{L5} KLTV ^{L5} TVTQFET..
T_harzianum	310	DGSISRY ^{L5} YVQNGVKFQQPNAQVGSYS.GNTINTDYCAA ^{L5} EQTA ^{L5} AFG
T_emersonii	294	GTDTGTLSEIKR ^{L5} FIQNSNVI ^{L5} PPNSDISGVT.GNSITTE ^{L5} FTAQ ^{L5} QQA ^{L5} AFG
P_chrysosporium	286	NTSTGTLSEIR ^{L5} FIQNGKVIQNSVANI ^{L5} PGVD ^{L5} PNSITDN ^{L5} FC ^{L5} AAQ ^{L5} QKTA ^{L5} AFG
T_reesei	297	SGAINRY ^{L5} YVQNGVTFQQPNAELGSYS.GNELNDDYCTAEEAE ^{L5} FG
T_harzianum	353	.GTSFTDKGGLAQINKAFQGGV ^{L5} VLVMSLWDDY ^{L5} ANMLWLDSTYPTNATAS
T_emersonii	343	DTDDFSQHGGLAKMGAA ^{L5} MQGGV ^{L5} VLVMSLWDDY ^{L5} AAQMLWLDSDYPTDADPT
P_chrysosporium	336	DTNWFQAQKGLKQMG ^{L5} EALNGV ^{L5} VLALSIWDDH ^{L5} AAANMLWLDSDYPTDKDPS
T_reesei	340	.GSSFSDKGGLTQFKKATSGGV ^{L5} VLVMSLWDDY ^{L5} ANMLWLDSTYPTNETSS
T_harzianum	402	TFGAKRGSCSTSSGVPAQVEAQSPNSKVIYSNIR ^{L5} FGPIG ^{L5} STGGntgsn
T_emersonii	393	TFGIARGTCPTDSGVPSDVESQSPNSYV ^{L5} TYSNIR ^{L5} FGPINSTFTas
P_chrysosporium	386	APGVARGTCATTSGVPSDVESQV ^{L5} PNSQVVF ^{L5} SNIR ^{L5} FGDIGSTFSGTS
T_reesei	389	TFGAVRGSCSTSSGVPAOVESOSPNAKVTF ^{L5} SNIR ^{L5} FGPIG ^{L5} STGNPSG

Fig. 3. Structure-based alignment of the cellobiohydrolases I from *T. harzianum* (Uniprot Q9P8P3), *T. emersonii* (Uniprot Q8TFL9, PDB 1Q9H), *P. chrysosporium* (Uniprot Q7LIJ0, PDB 1GPI) and *T. reesei* (Uniprot P07981.1, PDB 1CEL). The alignment was carried out with Strap (<http://3d-alignment.eu/>). Numbering of protein sequences is according to the deposited PDB structures. Position of residue Ala384 and gaps in loops 3 and 5 are depicted in rectangular boxes.

Fig. 4. Structural superposition of the Cel7A from *T. harzianum* (gray) and *T. reesei* (blue, PDB code **8CEL**). Closer view of the loop absent in *T. harzianum* highlighting Gln101 and the glucosyl residues from Tr_Cel7A (blue) at the entrance of the cellulose-binding tunnel. H-bonds are shown as dashed lines.

Fig. 5. Molecular surface representation of the Cel7A from (A) *T. harzianum* (gray), (B) *P. chrysosporium* (cyan) and (C) *T. reesei* (slate). The side face of the tunnel is depicted in a rectangular box. Residues and loops composing the tunnel side face are highlighted. Regions in red and dark blue represent negative and positive charged residues respectively.

Fig. 6. (A) Time evolution of the rmsd for residues Tyr260 (Th_Cel7A) and Tyr247 (Tr_Cel7A), showing the states 1 (rmsd below 3.0 Å) and 2 (rmsd above 3.0 Å). (B) Snapshots of the region between the loops 5 and 6 representative of the states 1 and 2. Molecular surfaces are depicted light red. The residue Tyr260 is more mobile than Tyr247 and the conformational changes of Tyr260 in Th_Cel7A create a completely exposed binding site. Although the residue Tyr247 (Tr_Cel7A) can also undergo significant conformational changes, these are not sufficient to create an opened binding site due to the presence of residue Tyr371 at loop 6, as opposed to Ala384 at loop 6 in Th_Cel7A.

Fig. 7. Root mean squared fluctuation (rmsf) of the α -carbons of Th_Cel7A and Tr_Cel7A.

The upper part shows the rmsf as a function of the residue numbers and the lower part shows superposed snapshots of the MD simulations colored according to the rmsf of each α -carbon (see the color scale). Highlighted is the loop 4, which covers the tunnel at the entrance of the binding site and is substantially more mobile in Th_Cel7A.