JUNIO COTA SILVA

"ESTUDOS FUNCIONAIS E ESTRUTURAIS DE HIDROLASES GLICOLÍTICAS BACTERIANAS VISANDO APLICAÇÕES EM BIOPROCESSOS"

"FUNCTIONAL AND STRUCTURAL STUDIES OF BACTERIAL GLYCOSYL HYDROLASES AIMING APPLICATIONS IN BIOPROCESSES"

CAMPINAS

2013

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA DE ALIMENTOS

JUNIO COTA SILVA

"ESTUDOS FUNCIONAIS E ESTRUTURAIS DE HIDROLASES GLICOLÍTICAS BACTERIANAS VISANDO APLICAÇÕES EM BIOPROCESSOS"

Orientador(a): Prof[a]. Dr[a]. Glaucia Maria Pastore
Co-orientador: Prof. Dr. Fabio Marcio Squina

"FUNCTIONAL AND STRUCTURAL STUDIES OF BACTERIAL GLYCOSIL HYDROLASES AIMING APPLICATIONS IN BIOPROCESSES"

Tese de Doutorado apresentada ao Programa de Pós-graduação em Ciência de Alimentos da Faculdade de Engenharia de Alimentos da Universidade Estadual de Campinas (FEA/UNICAMP) para a obtenção do título de Doutor em Ciência de Alimentos.

Doctorate Thesis presented to the Food Science Post-graduate Program of the Faculty of Food Engineering of the University of Campinas (FEA/UNICAMP) to obtain the Ph.D grade in Food Science.

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELO ALUNO JUNIO COTA SILVA E ORIENTADO PELA PROFA. DRA. GLAUCIA M. PASTORE

Assinatura do Orientador(a)

_____

CAMPINAS
2013

FICHA CATALOGRÁFICA ELABORADA POR
CLAUDIA AP. ROMANO DE SOUZA – CRB8/5816 - BIBLIOTECA DA FACULDADE DE ENGENHARIA DE ALIMENTOS – UNICAMP

Silva, Júnio Cota, 1985-
Si38e Estudos funcionais e estruturais de hidrolases glicolíticas bacterianas visando aplicações em bioprocessos / Júnio Cota Silva. -- Campinas, SP: [s.n.], 2013.

Orientador: Gláucia Maria Pastore.
Coorientador: Fabio Marcio Squina.
Tese (doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia de Alimentos.

1. Glicosil hidrolases. 2. Enzimas hipertermofílicas. 3. Engenharia de proteínas. 4. β-Glicosidases. 5. Caracterização. I. Pastore, Gláucia Maria. II. Squina, Fabio Marcio. III. Universidade Estadual de Campinas. Faculdade de Engenharia de Alimentos. IV. Título.

Informações para Biblioteca Digital

Título em Inglês: Functional and structural studies of bacterial glycosyl hydrolases aiming applications in bioprocesses
Palavras-chave em Inglês:
Glycosyl hydrolases
Hyperthermophilic enzymes
Protein engineering
β-Glucosidases
Characterization
Área de concentração: Ciência de Alimentos
Titulação: Doutor em Ciência de Alimentos
Banca examinadora:
Gláucia Maria Pastore [Orientador]
Hélia Harumi Sato
Mário Roberto Maróstica Júnior
George Jackson de Moraes Rocha
Juliano Lemos Bicas
Data da defesa: 27-03-2013
Programa de Pós Graduação: Ciência de Alimentos

BANCA EXAMINADORA

_____

Prof[a]. Dr[a]. Glaucia Maria Patore (FEA/UNICAMP) – Orientador(a)

_____

Prof. Dr. George Jackson de Moraes Rocha (CTBE/CNPEM) – Membro Titular

_____

Prof. Dr. Juliano Lemos Bicas (UFSJ) – Membro Titular

_____

Prof[a]. Dr[a]. Hélia Harumi Sato (FEA/UNICAMP) – Membro Titular

_____

Prof. Dr. Mario Roberto Maróstica Júnior (FEA/UNICAMP) – Membro Titular

_____

Dr. Ednildo de Alcântara Machado (UFRJ) – Membro Suplente

_____

Prof. Dr. Antônio José de Almeida Meirelles (FEA/UNICAMP) – Membro Suplente

_____

Prof[a]. Dr[a]. Gabriela Alves Macedo (FEA/UNICAMP) – Membro Suplente

"A investigação científica consiste em ver o que todos já viram, mas pensar o que ninguém ainda pensou. "

(Albert Szent-Györgyi)

Dedico aos meus pais Geraldo e Helena, que não mediram esforços para que eu até aqui chegasse!

Agradecimentos

A DEUS por me dar vida, saúde e vigor para enfrentar os desafios e porque, sem Ele, não teria chegado até aqui.

À minha esposa Valéria, pelo grande amor e carinho ao longo de mais de 6 anos de caminhada juntos, e por tornar meus dias mais agradáveis. Aos meus pais Geraldo e Helena pelo amor, pelo constante apoio, e por não medirem esforços para que todos os filhos estudassem. À Viviane e ao Renato, meus queridos irmãos, por sempre me apoiarem nessa caminhada, pelo carinho, e pelas conversas sempre motivadoras. À minha sogra e aos meus cunhados pelo carinho e pela torcida!

À Prof^{a}. Dr^{a}. Glaucia Maria Pastore pela oportunidade desde o mestrado, pela orientação, incentivo e por acreditar no meu trabalho.

Ao Dr. Fabio Squina pela oportunidade única de desenvolver o trabalho no CTBE, pela co-orientação ao longo de todo o trabalho, pela coordenação das pesquisas no CTBE, e também por confiar no meu trabalho.

Ao Dr. Mario Murakami e ao Dr. Roberto Ruller pelas valiosas discussões e contribuições nos trabalhos.

Aos professores membros da banca por avaliarem e darem sugestões preciosas para os trabalhos: Prof^{a}. Dr^{a}. Hélia Sato, Prof. Dr. Mário Maróstica, Prof. Dr. Juliano Bicas, Prof. Dr. George Jackson, Prof. Dr. Antônio Meirelles, Prof. Dr. Ednildo Machado e Prof^{a}. Dr^{a}. Gabriela Macedo.

Aos amigos do CTBE/CNPEM pela amizade e tornarem o ambiente de trabalho mais agradável. Agradecimento especial à Zaira, pela grande amizade e pelos vários trabalhos desenvolvidos em parceria, ao amigo "Capitão" Douglas por me ensinar um pouco de biologia molecular, aos amigos Leandro, João Paulo, Damásio, Segato e Bragatto pela excelente convivência e pelas coaborações científicas, à Daniela, ao Rodrigo, à Ana Paula, à Rebeca por organizarem os labs do CTBE e nos dar condições de trabalho, e a todos os colegas de lab pelo bom convívio e troca de experiências.

Aos amigos do laboratório de Bioaromas pela amizade e boa convivência, especialmente aos amigos Dr. Fábio Bichão, Prof. Dr. Biscatê, Dr^{a}. Xispita, MSc Molineno e MSc Cris, que apesar de termos nos espalhado a amizade continua!

Aos amigos da BIO UFV 2003 pelo companheirismo, especialmente ao Rômulo, grande amigo desde os tempos remotos de Viçosa.

Aos funcionários da Pós-garaduação Cosme e Marcão pela paciência e por estarem sempre prontos a nos atender e tirar nossas dúvidas.

Aos amigos da Igreja Presbiteriana de Barão Geraldo pelo companheirismo e apoio ao longo da jornada, e também pelos bons momentos de diversão.

Ao CTBE/CNPEM por fornecer toda a estrutura necessária ao desenvolvimento dos trabalhos.

À Universidade Federal de Viçosa pela exccelente formação ao longo de quatro anos de graduação.

À Universidade Estadual de Campinas pela oportunidade de cursar a pós-graduação em um centro de excelência em pesquisa, tanto no mestrado como agora no doutorado.

À FAPESP e ao CNPq pelos recursos destinados à realização das pesquisa.

Ao CNPq pela concessão da bolsa de doutorado.

SUMÁRIO

LISTA DE TABELAS

LISTA DE FIGURAS

Figure 2.1. Response surface (A) and contour plot (B) for the influence of temperature and pH on TpLam activity. (C) Curve of decay of time of half life ($t_{1/2}$) using four different temperatures (70, 80, 90 and 95 °C). (D) Thermal stability of laminarinase analyzed by circular dichroism. CD spectra were taken at 20°C (solid line), at 90°C (dashed line) and after 18 hours at 90°C.

Figure 2.2. Capillary zone electrophoresis of APTS-labeled oligosaccharides. (A) Incomplete and complete hydrolysis of APTS-reducing-end-labeled-laminarihexaose. (B) Incomplete and complete hydrolysis of laminarin (APTS-labeled after the enzymatic reaction). G1,G2, G3, G4, G5, G6, G7, G8 and G9 indicate the degree of polymerization of glucose oligomers.

Figure 2.3. Laminarinase analysis by SAXS. (A) Experimental scattering curve (open circles) and fit produced by GNOM (solid line). (B) Distance distribution function computed from the experimental data. (C) Laminarinase model with CBM1 (magenta), CD2 (green) and CBM3 (blue) domains fitted into the envelope obtained from SAXS data is shown in different orientations.

Figure 3.1. Characterization of XylLich using structure-based models (SB) through molecular dynamics (MD) simulations. **(A)** The figure on the top shows the free energy profile as a function of the radius of gyration (R_g) and the distance between the centers of mass of the XynA and BglS domains (CM). Free energy is given in units of kT and presented as a color scale. The figure was obtained using extensive simulation (100 ns) and no intra-domain interactions (see methods). **(B)** Average solvent accessible surface area (SAS) per residue, calculated from structure-based model simulations and **(C)** root mean square fluctuation (RMSF) per residue are presented. Figures **A** and **B** show the chimera (XylLich) in black, XynA only in blue, BglS only in green and the residues related to the catalytic site as red dots.

Figure 3.2. The chimera construction and experimental validation by SAXS. **(A)** Flow chart showing the process of fusion PCR for the XylLich construction. **(B)** SDS-PAGE of the chimera and wild-type proteins, indicating that the fused enzyme Xylanase-Lichenase had the predicted molecular weight. The protein molecular weight marker is shown in the first lane (M), and the values are displayed in kDa. **C)** The small angle X-ray scattering profile for the experimental and theoretical evaluation of XylLich, which was taken from the free energy basin with χ^2=2.80, R_g=26.0Å and *CM* distance=40.7Å. The scattering intensity is shown on a logarithm scale as a function of the momentum transfer (q). **(D)** The XylLich model comprising XynA (purple), the linker (red) and BglS (orange). CRYSOL was employed to generate the theoretical curve and VMD [55] for the denoted cartoon.

Figure 3.3. The effects of pH and temperature on the XylLich catalytic activity. pH (A and B) and temperature (C and D) curves for the chimera (■) and the parental enzymes (▲). The influence of pH on the enzymatic activity of XylLich compared to that of XynA (A) and BglS (B) using beechwood xylan (A) and lichenan (B) as substrates. The effect of temperature on the enzymatic activity of the chimeric enzyme compared to that of xylanase (C) and lichenase (D) using beechwood xylan and lichenan as substrates, respectively, is represented.

Figure 3.4. Capillary zone electrophoresis analysis of the breakdown products released by XylLich, XynA and BglS. The products after enzymatic hydrolysis of APTS-labeled xylohexaose (A), lichenan (B) and xylohexaose plus lichenan together (C) are presented. X2, X3, X4, X6 and G4 indicate the degree of polymerization of the produced xylose and glucose oligomers. The APTS-labeled xylohexaose used in the assays is indicated in the upper right boxes (A and C).

Figure 3.5. Secondary structure evaluation by circular dichroism. **(A)** Far-UV CD spectra of XylLich and individual enzymes at pH 7.4 and 20°C and **(B)** thermal denaturation curve at pH 7.4. The circular dichroism spectrum represents an average of eight scans. The thermal denaturation curve was obtained by monitoring at 220 nm for xylanase, 218 nm for lichenase and 218.5 nm for the chimera.

Figure 3.6. Biotechnological appealing for producing the chimeric enzyme**. (A)** Analysis of the conversion efficiency of a composite consisting of beechwood xylan and lichenan by the parental enzymes, the mixture of the parental enzymes (XynA + BglS) and XylLich. **(B)** Assessment of the activity yield (AY) using the crude *E. coli* cell extract as an enzyme source.

Figure SI-3.1. Analysis of the candidates to describe the chimera arrangement in solution. A) Superposition of some conformations extracted from the free energy basin and the B) Structural alignment at the XynA domain. C) Theoretical scattering of the conformations and experimental curves. The scattering curves shown in Figure C share the same color as the derived protein model presented in A and B.

Figure SI-3.2. Scattering curves of conformations extracted from the regions outside the free energy basin. The evaluated conformations did not show a satisfactory agreement with the experimental result.

Figure 4.1. Influence of pH and temperature on the pNPGase activity. Response surface plot and contour plot for the effects of pH and temperature on β-glucosidase activity of PfBgl1 (A, B), TpBgl1 (C, D), and TpBgl3 (E, F). The statistic models which predict the enzyme activities have a coefficient of variation of 0.95, 0.98, 0.94 respectively, and were obtained by using experimental data.

Figure 4.2. Effects of pH on pNPGase activity and stability. (A) Curve of pH for enzymatic activities of PfBgl1 (●), TpBgl1 (■), and TpBgl3 (▲) at 70 °C. (B) The bar graph show the pH stability for PfBgl1 (small grid bar), TpBgl1(large grid bar), and TpBgl3 (horizontal lines bar). The bars present the residual β-glucosidase activities after 1 week incubated at 8 °C in McIlvaine-glycine added buffer in several pH levels. The percentages where accessed by comparing the activities to initial activity before incubation.

Figure 4.3. Influence of several ligands on pNPGase activity. Inhibition/activation effects of 6 monosaccharides at 12 concentrations (up to 1 M) on β-glucosidase activity of PfBgl1 (A), TpBgl1

(B), and TpBgl3 (C). The monosaccharides assayed were glucose (● black), galactose (■ dark yellow), arabinose (▲ orange), mannose (▼ green), fructose (♦ blue), and xylose (○ red).

Figure 4.4. Characterization of BGL by circular dichroism. Far-UV CD spectra and thermal denaturation in different pH levels for PfBgl1 (A, B), TpBgl1 (C, D), and TpBgl3 (E, F). The circular dichroism spectrum represents an average of eight scans. The thermal denaturation was obtained by monitoring at 222 nm.

Figure SI-4.1. Capillary zone electrophoresis of APTS-labeled oligosaccharides. (A) Incomplete and complete hydrolysis of APTS-reducing-end-labeled-cellohexaose and cellobiose APTS-labeled after the enzymatic reaction (left box). (B) Incomplete and complete hydrolysis of APTS-reducing-end-labeled-laminarihexaose. G1, G2, G3, G4, G5, G6, L1, L2, L3, L4, and L6 indicate the degree of polymerization of glucose oligomers. This is a typical pattern found for the three hyperthermophilic BGLs.

LISTA DE SÍMBOLOS E ABREVIATURAS

ANOVA análise de variância

AY rendimento de atividade enzimática (*activity yield*)

BamHI enzima de restrição

BGL β-glicosidase

BglS liquenase de *Bacillus subtilis*

CAZy banco de dados de enzimas (*Carbohydrate-Active EnZymes*)

CBH celobiohidrolase

CBM domínio de ligação a carboidrato (*carbohydrate binding domain*)

CCRD delineamento composto central rotacional (*central composite rotatable design*)

CD dicroísmo circular

CD domínio catalítico (*catalytic domain*)

CM centro de massa (*center mass*)

CMC carboximetilcelulose

CRYSOL *software* para análise de SAXS

CZE eletroforese capilar de zona (*capillary zone eletrophoresis*)

DAMIN *software* para análise de SAXS

D_{max} diâmetro máximo (Å)

DNA ácido desoxirribonucleico (*deoxyribonucleic acid*)

EC sistema de classificação da comissão de enzimas (*enzyme commission*)

EcoRI enzima de restrição

GH sistema de classificação em famílias (*glycosyl hydrolase*)

GNOM *software* para análise de SAXS

IPTG	isopropil-β-D-1-tiogalactopiranosídeo
k_{cat}	constante catalítica (número de *turnover*)
k_{cat}/K_m	constante de especificidade
K_i	constante de inibição (mM)
K_m	constante de Michaelis-Menten (mg/mL ou mM)
LB	meio de cultura (Luria Bertani)
NdeI	enzima de restrição
pBAD/Myc	vetor de clonagem para expressão de proteínas
PCR	reação em cadeia da polimerase (*polymerase reaction chain*)
PDB	banco de dados de proteínas (*protein data bank*)
pET28a(+)	vetor de clonagem para expressão de proteínas
PfBgl1	β-glicosidase de *Pyrococcusfuriosus* da família GH1
PMSF	*phenylmethylsulfonyl fluoride*
pNPG	4-nitrofenil-β-D-glicopiranosídeo (*4-nitrophenyl- β-D-glucopyranoside*)
R_g	raio de giro de Guinier (Å)
RMSF	raiz quadrada média das flutuações (*root mean square fluctuations*)
RSM	metodologia de superfície de resposta (*response surface methodology*)
SAS	área de superfície acessível ao solvente (*solvent accessible surface*)
SAXS	espalhamento de raios-X a baixos ângulos (*small angle X-ray scattering*)
SB	modelos baseados em estrutura (*structure based models*)
SDS-PAGE	eletroforese em gel desnaturante para proteínas (*sodium dodecyl sulfate polyacrylamide gel electrophoresis*)
Tm	temperatura média de desnaturação
TpBgl1	β-glicosidase de *Thermotoga petrophila* da família GH1

TpBgl3	β-glicosidase de *Thermotoga petrophila* da família GH3
TpLam	laminarase de Thermotoga petrophila
V_{max}	velocidade máxima (IU/mg de proteína ou IU/nmol de proteína)
XylLich	enzima quimérica formada por domínios de xilanase e liquenase
XynA	xilanase de *Bacillus subtilis*
APTS	*8-aminopyreno-1,3,6-trisulfonic acid*
$t_{1/2}$	tempo de meia vida da atividade enzimática (*time of half-life*) (min)
DNS	ácido dinitrosalicílico (*3,5-dinitrosalicylic acid*)

RESUMO GERAL

Atualmente há uma crescente demanda para o desenvolvimento de combustíveis não-fósseis alternativos. Assim, como a biomassa lignocelulósica é uma das fontes de energia mais abundantes na natureza, pode ser estabelecida uma economia verde e sustentável, com o objetivo de processar a grande quantidade de energia estocada nessas matérias-primas. O etanol de cana-de-açúcar é uma das melhores opções em biocombustíveis e sua produção pode mais que dobrar, se os açúcares constituintes da parede celular vegetal forem utililizados. No entanto, o alto custo de produção das enzimas para hidrolisar e processar os materiais lignocelulósicos é um fator altamente limitante para o uso de tecnologias verdes. Este trabalho se propôs a avaliar novos biocatalisadores e construir uma enzima quimérica na tentativa de obter glicosidases com melhor desempenho que as já relatadas. Enzimas despolimerizadoras de β-1,3-glucanos têm consideráveis aplicações biotecnológicas, incluindo produção de biocombustíveis, insumos químicos e farmacêuticos. No segundo capítulo, mostramos a caracterização funcional e a estrutura de baixa resolução da laminarase hipertermofílica de *Thermotoga petrophila* (TpLam), além de seu modo de operação por eletroforese capilar de zona, mostrando que ela cliva especificamente ligações β-1,3-glucosídicas internas. O dicroismo circular (CD) UV-distante demonstrou que TpLam é formada principalmente por elementos estruturais do tipo beta, e a estrutura secundária é preservada após incubação por 16 horas a 90 º C. A forma determinada pelo pequeno espalhamento de raios-X a baixo ângulo revelou uma arquitetura de multi-domínio da enzima, com um arranjo de envelope em forma de V, no qual os dois módulos de ligação de carboidrato estão ligados ao domínio catalítico. A engenharia de enzimas multifuncionais pode melhorar coquetéis enzimáticos para tecnologias emergentes de biocombustíveis. Dinâmica molecular através de modelos baseados em estrutura (SB) é uma ferramenta eficaz para avaliar a disposição tridimensional das enzimas quiméricas, bem como para inferir a viabilidade funcional antes da validação experimental. No terceiro capítulo, descrevemos a montagem computacional de uma quimera bifuncional xilanase-liquenase (XylLich), usando os genes *xynA* e *bglS* de *Bacillus subtilis*. As análises *in silico* da área de superfície acessível ao solvente (SAS) e da raiz quadrada média das flutuações (RMSF) previram uma quimera completamente funcional, ou seja, uma enzima cujo substrato tem acesso ao seu sítio catalítico com pequenas flutuações e variações ao longo das cadeias polipeptídicas. A quimera preservou as características bioquímicas das enzimas parentais, com exceção de uma pequena variação na temperatura de operação e na eficiência catalítica (k_{cat} / K_m). Também foi verificado ausência de

mudanças significativas no modo de operação catalítico. Além disso, a produção de enzimas quiméricas pode ser mais rentável do que a produção de uma única enzima separadamente, comparando-se o rendimento da produção de proteína recombinante e a atividade hidrolítica da enzima quimérica com as enzimas parentais. β-Glicosidases (BGLs) são enzimas muito úteis e com grande potencial para serem empregadas em diversos processos industriais. Entretanto, algumas características são essenciais para tornar viáveis as aplicações, como por exemplo estabilidade à temperatura e ao pH, bem como baixa inibição por íons e outros compostos químicos. Assim, no quarto capítulo buscamos estudar três BGLs dos organismos extremófilos *Pyrococcus furiosus* e *Thermotoga petrophila*. Os genes *PfBgl1*, *TpBgl1* and *TpBgl3* foram clonados no vetor pET28a e as proteínas expressadas em *Escherichia coli* e posteriormente purificadas em duas etapas cromatográficas. As enzimas purificadas foram avaliadas quanto ao pH e temperatura de atividade, sendo que as BGLs da família GH1 (PfBgl1 e TpBgl1) apresentaram faixas mais largas de pH e temperatura de operação do que a família GH3 (TpBgl3). As BGLs mostraram grande estabilidade ao pH e o maior tempo de meia-vida (a 99 ° C) foi verificado no pH 6, e além disso, não foram significativamente afetadas pela presença de EDTA ou de íons, exceto a TpBgl1 que foi inibida por Hg^{2+} e Fe^{2+}. As atividades específicas para um conjunto de diferentes substratos sugeriram que TpBgl3 é mais específica que as BGLs GH1. O k_{cat} e k_{cat} / K_m em 4-nitrofenol-β-D-glicopiranosídeo (pNPG) indicam que TpBgl3 é a mais eficiente para hidrólise do substrato, embora seja a enzima que foi inibida com a menor concentração de glicose (30.1 mM). Além disso, as BGLs foram analisadas quanto à influência de seis monossacarídeos na catálise, e demonstraram serem fracamente inibidas pela maioria dos açúcares testados. Os ensaios de CD UV-distante revelaram que a estrutura secundária das BGLs não é afetada pelas variações de pH, e os estudos de desnaturação térmica evidenciaram que as BGLs são proteínas hipertermofílicas.

Palavras-chave: *glicosil hidrolases, enzimas hipertermofílicas, engenharia de proteínas, β-glicosidases, caracterização.*

ABSTRACT

There is an increasing demand for the development of alternative non-fossil fuels. Thus, since the lignocellulosic biomass is the most abundant source in nature, it may be settled a green and sustainable economy, aiming to process the great amount of energy stocked in these raw materials. The ethanol from sugarcane is one of the best options concerning biofuels and its productivity could be raised more than double if the use of sugars constituents of plant cell wall is considered. However the high production cost of the enzymes to hydrolyze and process lignocellulose is a great limiting factor for green technologies. In this way, this work proposed to evaluate new enzymes and engineer a chimeric enzyme in the attempt to prospect glycosyl hydrolases with better performance than those reported up to date. 1,3-β-Glucan depolymerizing enzymes have considerable biotechnological applications including the production of biofuels, feedstock-chemicals and pharmaceuticals. In the first chapter we showed the comprehensive functional characterization and low-resolution structure of hyperthermophilic laminarase from *Thermotoga petrophila* (TpLam), besides its mode of operation through capillary zone electrophoresis, which specifically cleaves internal β-1,3-glucosidic bonds. Far-UV circular dichroism demonstrated that LamA is formed mainly by beta structural elements, and the secondary structure is maintained after incubation up to 16 hours at 90ºC. The structure determined by small angle X-ray scattering revealed a multi-domain structural architecture of the enzyme with a V-shape envelope arrangement of the two carbohydrate binding modules in relation to the catalytic domain. Multifunctional enzyme engineering can improve enzyme cocktails for emerging biofuel technology. Molecular dynamics through structure-based models (SB) is an effective tool for assessing the tridimensional disposal of chimeric enzymes as well as for inferring the functional practicability before experimental validation. In the second chapter we describe the computational design of a bifunctional xylanase-lichenase chimera (XylLich) using the *xynA* and *bglS* genes from *Bacillus subtilis*. In silico analysis of the average surface accessible area (SAS) and the root mean square fluctuation (RMSF) predicted a fully functional chimera, i.e. the substrate has access to the catalytic pocket with minor fluctuations and variations along the polypeptide chains. The chimera preserved the biochemical characteristics of the parental enzymes, with the exception of a slight variation in the temperature of operation and the catalytic efficiency (k_{cat}/K_m). The absence of substantial shifts in the catalytic mode of operation was also verified. Furthermore, the production of chimeric enzymes could be more profitable than producing a single enzyme separately, based on comparing the recombinant protein production

yield and the hydrolytic activity achieved for XylLich with that of the parental enzymes. β-Glucosidases (BGLs) are very useful enzymes with a great potential to be employed in several industrial processes. However, some features are required to become viable the enzyme applications, such as temperature and pH stability as well, low ions and chemicals inhibition. Thus this work aimed to study three BGLs from the extremophiles organisms *Pyrococcus furiosus* and *Thermotoga petrophila*. The genes *PfBgl1*, *TpBgl1* and *TpBgl3* were cloned into pET28a vector and the proteins were expressed in *Escherichia coli* and further purified in two chromatographic steps. The purified enzymes were evaluated for pH and temperature of activity, which showed that BGLs from the glycosyl hydrolases family 1 (PfBgl1, TpBgl1) presented a wider range of pH and temperature operation than BGL from family 3 (TpBgl3). The BGLs showed great stability to a range of pH (4-10) and the highest time of half-life (at 99 °C) was at pH 6, besides they were not significantly affected by the presence of EDTA or ions, except TpBgl1 that was inhibited by Hg^{2+} and Fe^{2+}. The specific activities in a set of different substrates suggested that TpBgl3 is more specific than GH1 BGLs. The k_{cat} and k_{cat}/K_m in pNPG indicate that TpBgl3 is the most efficient among BGLs characterized herein, although this enzyme is inhibited with the lowest glucose concentration (30.1 mM). Furthermore, the BGLs were assayed for influence of six monosaccharides in catalysis, which the results suggested a weak inhibition by the most of those carbohydrates tested. The CD experiments revealed that the secondary structure of BGLs is not affected by the pH variations and the denaturation studies evidenced that the BGLs are indeed hyperthermophilic.

Key words: *glycosyl hydrolases, hyperthermophilic enzymes, protein engineering, β-glucosidases, characterization.*

CAPÍTULO 1

APRESENTAÇÃO DO TRABALHO E REVISÃO BIBLIOGRÁFICA

1.1 INTRODUÇÃO GERAL

Um dos maiores desafios para as nações no século 21 é suprir a crescente demanda por energia em sistemas de transporte, de aquecimento e em processos industriais, e ainda fornecer de forma sustentável matéria-prima para a indústria (Hahn-Hägerdal et al., 2006). Assim, o atual panorama político-econômico mundial concentra esforços no sentido de buscar alternativas energéticas ao uso irracional do petróleo e seus derivados. Há uma tendência global em direção ao desenvolvimento econômico sustentável, que atenda à crescente demanda energética e ao mesmo tempo contemple a redução dos impactos ambientais, evidenciando a importância da substituição de combustíveis fósseis por fontes renováveis de energia (Brasil - MME, 2007). A oscilação de preços e as perspectivas de esgotamento de reservas de petróleo, aliadas ao risco relativo à dependência da importação deste combustível de países politicamente instáveis, também impulsionam os países na busca por novas matrizes energéticas e sustentáveis (Bastos, 2007).

Desde o início dos anos 90, o foco no tratamento de resíduos está sendo substituído pelo princípio da exploração dos resíduos que são inevitavelmente produzidos em larga escala (Anastas et al., 1998; Tuck et al., 2012). Ninguém poderia imagirar que em 2011 existiria um empreendimento viável envolvendo remessa de milhares de toneladas de resíduos domésticos de um país para outro. E esse sistema já existe e responde pelo transporte de aproximadamente 200.000 toneladas/ano de resíduos domésticos da Itália para Rotterdam, para uso como matéria-prima na geração de energia elétrica em plantas industriais holandesas (Truck et al., 2012). Pouco tempo depois, em 2012, a Suécia também importou 800.000 toneladas de lixo de países europeus para produzir energia (Smith, 2012). Estima-se que aproximadamente 10^8 t de resíduos sejam gerados anualmente em todo o mundo (Truck et al., 2012). Assim, uma bio-economia deve ser estabelecida, fornecendo meios de aproveitamento, processos e demanda para essa enorme quantidade de energia estocada na forma de biomassa.

A biomassa lignocelulósica é a fonte de energia renovável mais abundantemente encontrada na natureza (Castro e Pereira Jr., 2010); compreendida pelos resíduos agrícolas, agroindustriais e florestais, além de materiais desperdiçados, denominados biomassas residuais, dentre os quais o bagaço e o palhiço da cana, o sabugo e a palha de milho, as palhas de trigo e arroz, os restos de madeira processada e os resíduos municipais baseados em papel (Bastos, 2007). No Brasil, merecem destaque os sub-produtos da cultura de cana-de-açúcar, uma vez que 50% dessa biomassa se encontram disponíveis nas instalações industriais e, em 2005, somaram 118 milhões de toneladas

em base seca. O palhiço mais o bagaço representam, aproximadamente, dois terços do conteúdo energético da cana-de-açúcar. Enquanto o palhiço é deixado no solo, como adubo para a cultura, uma parte do potencial de uso do bagaço é aproveitada para a geração de bioeletricidade, tornando as usinas de açúcar e álcool auto-suficientes no consumo energético (Brasil - MME, 2007). Os excedentes de bagaço e palhiço têm o potencial para serem transformados em etanol de segunda geração. Nesta nova condição, o aproveitamento da cana será integral (Cerqueira-Leite, 2009).

Vários países têm incrementado o uso de bioetanol em sua matriz energética visando a uma redução da dependência de combustíveis fósseis. O etanol de segunda geração de cana-de-açúcar é atualmente a opção de biomassa energética de maior produtividade por unidade de área cultivada e de melhor balanço energético. O Brasil é o maior produtor mundial de cana e de açúcar e o segundo maior produtor de bioetanol (Cerqueira-Leite, 2009). Apesar dos seus altos e baixos desde a criação do Programa Nacional do Álcool (Proálcool) por meio do Decreto 76.593/75, a trajetória bem-sucedida do bioetanol combustível, aliada aos planos nacionais que contemplam a manutenção do alto percentual de energia renovável na matriz energética, reúne vantagens ao Brasil frente a outros países para destacar-se mundialmente na produção e uso de biomassa como recurso energético (Bastos, 2007). O País reúne as condições necessárias, tanto do ponto de vista dos recursos naturais considerados essenciais, como também de competência técnica, infraestrutura e recursos financeiros. O diferencial brasileiro estaria no clima e qualidade do solo (Cerqueira-Leite, 2009) e na integração existente dentro da própria usina, que permite atender ao requisito de substancial balanço energético positivo (Bastos, 2007).

Entretanto, a demanda por maior quantidade de matérias-primas para produção de bioetanol, aliada à racionalização do uso da terra, tem estimulado cientistas de diversos países a direcionarem suas pesquisas para o desenvolvimento de tecnologias que proporcionem uma maior utilização de resíduos lignocelulósicos (Bastos, 2007). O termo geral biomassa vegetal compreende a matéria vegetal gerada pela fotossíntese e seus diversos produtos e subprodutos derivados. A energia contida nestes materiais é aproveitada mediante processos de conversão dos mais diversos (Brasil - MME, 2007). É interessante salientar que cada resíduo lignocelulósico tem seu valor de mercado regulado pela aplicação a ele dada, podendo o valor (em dólares) de uma tonelada variar de $ 60 – 150 para geração de eletricidade, $ 70 – 200 para alimentação animal, $ 200 – 400 para produção de combustível, até $ 1.000 para produção de químicos (Bals & Dale, 2012; Sanders et al., 2007).

Um bioprocesso eficiente e econômico é a base da indústria biotecnológica em um mercado competitivo. Em todo o mundo, indústrias de alimentos, agroindústrias e indústrias de madeira

produzem, anualmente, milhões de toneladas de resíduos, incluindo bagaços, cascas, tortas e até resíduos líquidos, muitos dos quais representam sérios problemas referentes ao tratamento dos mesmos. A coleta e o tratamento desses resíduos são geralmente bastante onerosos para as indústrias, e por isso, há uma grande necessidade de um melhor gerenciamento desses resíduos, baseado no conceito: reduzir, reutilizar e reciclar (Makkar & Cameotra, 2002). Nesse sentido, nos últimos anos a indústria e a ciência tem se aliado para tentar reduzir a quantidade desses resíduos, propondo usos alternativos para eles. Devido a sua composição rica em açúcares, nitrogênio e minerais, os resíduos agroindustriais podem ser facilmente utilizáveis por micro-organismos e modificados por enzimas microbianas, sendo, portanto, considerados matérias-primas adequadas à produção de compostos industrialmente relevantes e de maior valor agregado (Orzua *et al*., 2009).

É importante salientar que as aplicações da alcoolquímica não estão restritas ao uso como combustível, havendo também interesse no etanol grau químico, uma matéria-prima utilizada em diversos setores da indústria de transformação. O modelo de produção norte-americano baseado no petróleo barato tornou-se hegemônico, sendo que, segundo dados de 2007, 90% da matéria-prima para síntese de moléculas orgânicas era proveniente do petróleo (Bastos, 2007). Neste contexto, surge o termo biorrefinaria, que se refere ao uso de matérias-primas renováveis e de seus resíduos para produção de uma variedade de substâncias e energia de modo a maximizar o valor da biomassa (Bastos, 2007, Truck et al., 2012).

1.2 OBJETIVOS

1.2.1 Objetivo Geral

O objetivo desse trabalho foi clonar, expressar e caracterizar seis hidrolases glicolíticas bacterianas utilizando ferramentas bioquímicas e biofísicas, visando entender melhor o funcionamento das mesmas, para futuramente aplicá-las na bioconversão de biomassa lignocelulósica.

1.2.2 Objetivos Específicos

- Clonar em pET28a(+) os genes que codificam as seis enzimas, expressá-las de forma heteróloga em *E. coli* BL 21, e posteriormente purificá-las por cromatografia de afinidade;
- Caracterizar a laminarase (endo-1,3-β-glucanase) de *Thermotoga petrophila* quanto aos aspectos bioquímicos (temperatura e pH de atividade, modo de operação, especificidade pelos substratos, estabilidade térmica, parâmetros cinéticos) e biofísicos (composição de estrutura secundária, desnaturação térmica, forma em solução);
- Simular computacionalmente e construir uma enzima híbrida a partir de uma endo-1,4-β-xilanase e uma 1,3-1,4-β-glucanase, ambas de *Bacillus subtilis*. Caracterizar as proteínas nativas e a quimera quanto aos aspectos bioquímicos (temperatura e pH de atividade, modo de operação, especificidade pelos substratos, parâmetros cinéticos) e biofísicos (composição de estrutura secundária, desnaturação térmica, disposição dos domínios em solução);
- Comparar três β-glicosidases hipertermofílicas de *Thermotoga petrophila* e *Pirococcus furiosus*, quanto às propriedades bioquímicas (temperatura e pH de atividade, modo de operação, especificidade pelos substratos, estabilidade térmica e em função do pH, parâmetros cinéticos) e biofísicos (composição de estrutura secundária, desnaturação térmica e em função do pH, forma em solução em função do pH).

1.3 REVISÃO BIBLIOGRÁFICA

1.3.1 Biomassa vegetal como fonte de energia

Os resíduos agroindustriais gerados são, em sua maior parte, biomassa vegetal de constituição ligninocelulósica. Na lignocelulose há três componentes principais: celulose (38-50%), hemicelulose (23-32%) e lignina (15-30%) (Sierra et. Al., 2008). O maior componente da parede celular é a celulose, que consiste em um homopolissacarídeo linear de resíduos de D-glicopiranose unidos por ligações β-1,4 glicosídicas. Esse polissacarídeo é sintetizado na membrana plasmática e posteriormente depositado na parede celular, sendo sua principal função conferir rigidez à parede celular (Minic & Jouanin, 2006). Caracteriza-se por possuir regiões de alta cristalinidade e regiões amorfas (não-cristalina), sendo que as ligações de hidrogênio entre cadeias de celulose adjacentes contribuem para o alto grau de cristalinidade, insolubilidade em água e resistência da celulose à degradação/hidrólise (Carvalho et. Al., 2009).

A fração hemicelulósica é a segunda mais abundante na parede celular vegetal, apresentando polissacarídeos principalmente na forma de xilano, além de xiloglucano e quantidades menores de diversos polissacarídeos contendo manose, os quais são sintetizados no Complexo de Golgi e direcionados para a parede celular. O componente majoritário da hemicelulose, o xilano, é formado por uma cadeia principal de resíduos de β-1,4-D-xilopiranose substituídos com L-arabinofuranose, ácido glicurônico, ácido 4-O-metilglicurônico, e cadeias laterais de acetil. Essas substituições são encontradas em diferentes frequências dependendo da origem da biomassa vegetal. O xiloglucano tem como cadeia principal resíduos de β-1,4-D-glicose substituído com xilose, galactose e L-fucose. A estrutura hemicelulósica, diferentemente da celulose, não apresenta cristalinidade (Madson et al., 2003; Ogeda e Petri, 2010).

A lignina é uma macromolécula aromática heterogêneo sintetizado a partir de três alcoóis p-hidróxi-cinamílicos precursores: p-cumarílico, coniferílico e sinapílico; é formada por ligações éter que aumentam a resistência da estrutura a ataques químicos e enzimáticos (Castro e Pereira Jr., 2010). Cadeias de celulose são circundadas por hemicelulose e lignina, associadas entre si por meio de ligações covalentes e interações fracas, formando estruturas que medeiam a estabilidade estrutural da parede celular vegetal (Carvalho et. Al., 2009) (Figura 1.1). Assim, a natureza cristalina da celulose e as complexas interações entre celulose, hemicelulose e lignina tornam a lignocelulose altamente resistente a processos hidrolíticos (Ogeda e Petri, 2010).

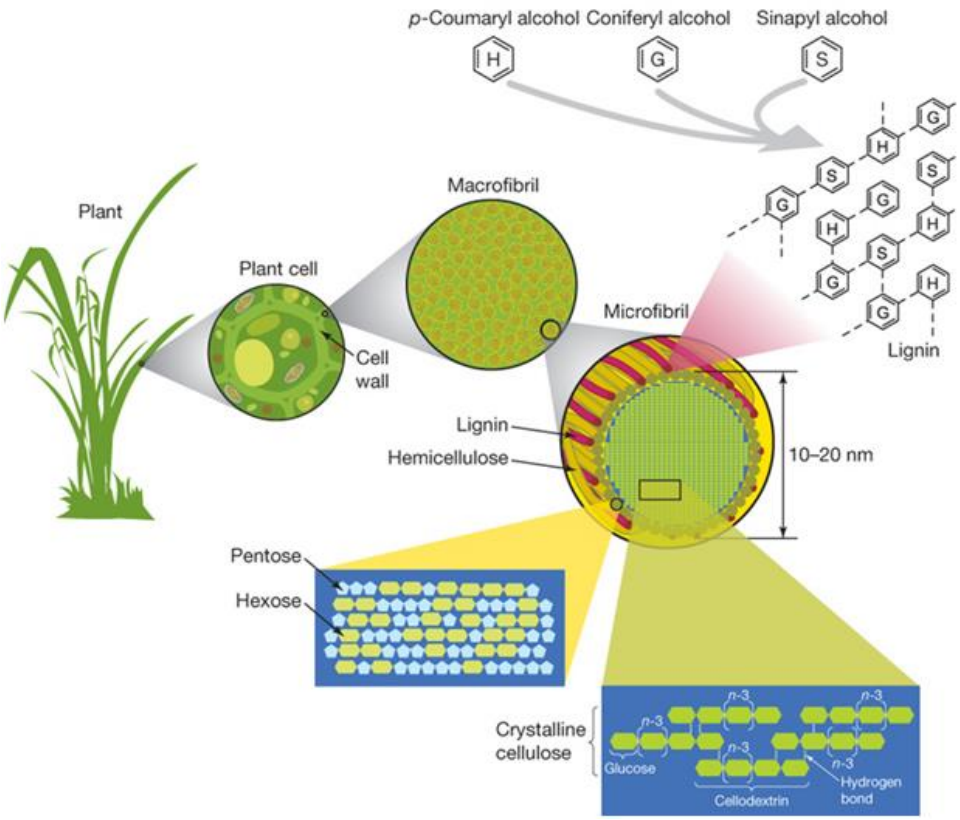


Figura 1.1. Esquema representativo da estrutura micro e nanométrica da lignocelulose. (Adaptada de Potters *et. al.*, 2010).

No que diz respeito à cana-de-açúcar, uma das maiores culturas brasileiras, a composição média do bagaço é, em porcentagem, 43,1 ± 1,4 de celulose, 25,2 ± 1,9 de hemicelulose, 22,9 ± 1,1 de lignina, 2,8 ± 1,4 de cinzas e 4,3 ± 1,6 de extrativos (Rocha *et al.*, 2012). Assim, levando-se em consideração os milhões de toneladas de "resíduos" gerados pela indústria sucroalcooleira, é evidente o grande potencial de emprego desta biomassa lignocelulósica como fonte de açúcares fermentescíveis. A glicose é o produto de hidrólise de maior interesse industrial, tendo em vista que sua tecnologia de conversão em etanol está consolidada (Castro e Pereira Jr., 2010).

O etanol de primeira geração, obtido a partir da fermentação da sacarose, apresenta um rendimento médio de 88 litros por tonelada de cana, em condições otimizadas (Cerqueira-Leite, 2009). Na cana-de açúcar, o bagaço, a palha e o caldo respondem cada um por um terço da fitomassa (Cerqueira-Leite, 2009). A celulose e a hemicelulose do bagaço tem potencial para a produção de 281,6 litros e 170,3 litros de etanol por tonelada de matéria seca, respectivamente (www1.eere.energy.gov/biomass/ethanol_yield_calculator.html). Por consequência, uma tonelada de bagaço seco tem o potencial de produção de 451,9 litros de etanol. Dessa forma, a produção de etanol poderia ser ampliada de 88 litros (primeira geração) podendo chegar a 235 litros por tonelada

de cana, utilizado-se o processo de hidrólise do bagaço, isso sem considerar o potencial gerador de energia da palha. Portanto, o bioetanol celulósico é referido como uma segunda geração de biocombustíveis cujo processamento é uma das mais promissoras tecnologias em fase de desenvolvimento (Bastos, 2007).

1.3.2 Hidrólise enzimática de biomassa

A hidrólise consiste na conversão da biomassa celulósica em açúcares fermentescíveis, seja principalmente via processos químicos (hidrólise ácida) ou biológicos (hidrólise enzimática). O processo de hidrólise quebra as cadeias de celulose e hemicelulose, liberando hexoses e pentoses, respectivamente. A fermentação das hexoses e das pentoses permite a produção de etanol (Neto *et al*., 2008). As enzimas glicosídicas são catalisadores biológicos, de natureza protéica, e que apresentam alta atividade catalítica e seletividade específica sobre polímeros de carboidratos, podendo transformar, desestabilisar ou reorganizar essas cadeias orgânicas (Fry, 2004). Microrganismos celulolíticos são capazes de produzir várias enzimas com diferentes especificidades, atuando simultânea e sinergicamente para liberar açúcares fermentescíveis. Além dos fungos decompositores e bactérias, organismos que se alimentam de matéria lignocelulósica são capazes de produzir essas enzimas, como cracas, baratas e cupins (Watanabe & Tokuda, 2010).

A conversão de biomassas lignocelulósicas a bioetanol pode ser de forma integrada ou em etapas independentes (Figura 1.2). No processo SHF (*Separate Hydrolysis Fermentation*) o bioprocesso ocorre em quatro passos separadamente, como indicado na figura 1.2 (Wright et al., 1988). No processo SSF (*Simultaneous Saccharification Fermentation*), a produção de enzimas hidrolíticas e a hidrólise de hemiceluloses ocorrem em dois passos distintos, enquanto que a sacarificação ocorre simultâneamente à fermentação (Takagi et al., 1977), reduzindo a inibição das celulases pelos produtos de hidrólise, pois a glicose liberada é concomitantemente fermentada. A condução da cofermentação de pentoses e hexoses simultaneamente à hidrólise das frações celulósica e hemicelulósica compreende o processo de SSCF (*Simultaneous Saccharification and Co-Fermentation*). Por fim, na forma mais integrada das etapas de conversão, nomeada Bioprocesso Consolidado, CBP (*Consolidated BioProcessing*), a produção enzimática, a hidrólise de celulose e hemicelulose e a fermentação acontecem em um mesmo reator, utilizando-se um único micro-organismo com a habilidade de produzir enzimas, bem como de fermentar os produtos oriundos da hidrólise (Wang et al., 1983; Lynd et al., 1991 e 2005).

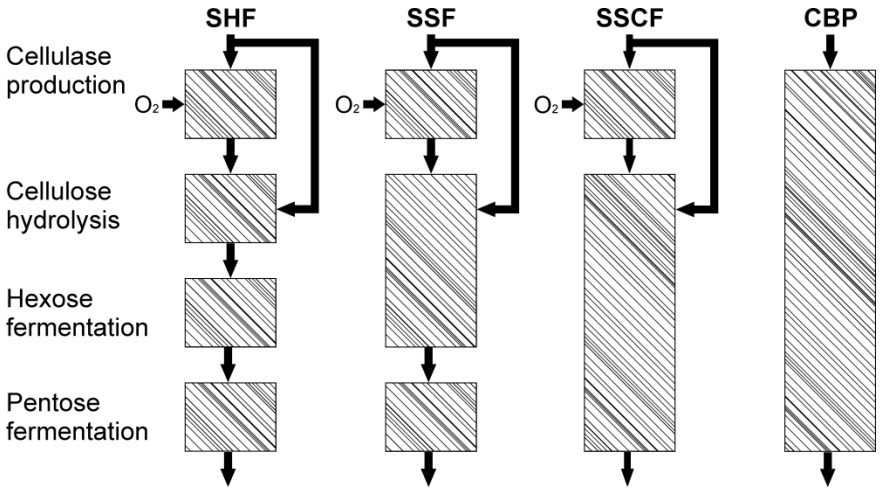


Figura 1.2. Fluxograma mostrando as estratégias de processamento de biomassa por hidrólise enzimática. Cada caixa representa um bioreator (não em escala). (Adaptada de Lynd *et al.*, 2002).

1.3.3 Enzimas ativas em carboidratos

De acordo com a nomenclatura da Comissão de Enzimas (*EC-Enzyme Commission*) da IUBMB (*International Union of Biochemistry and Molecular Biology*) hidrolases glicolíticas recebem a identificação EC 3.2.1.x, sendo que: o primeiro dígito indica que as enzimas catalisam reações de hidrólise, o segundo dígito indica que são glicosilases, o terceiro dígito indica que são glicosidases e que hidrolisam ligações *O*-glicosídicas e *S*-glicosídicas, e o último dígito remete aos tipos de substratos reconhecidos por essas enzimas.

Henrissat (1991) foi o primeiro a propor uma classificação abrangente para as hidrolases glicolíticas, baseando-se nas sequências de aminoácidos para agrupá-las em famílias. Um total de 291 sequências de aminoácidos, correspondentes a 39 grupos de classificação EC, foram classificadas em 35 famílias, sendo apenas 10 sequências não alocadas em nenhuma família. Em relação à especificidade pelo substrato, 18 famílias foram identificadas como mono específicas e 17 como poli específicas. Posteriormente essa classificação foi atualizada, aumentando o número de sequências para 482 e o de famílias para 45, sendo 22 poli específicas (Henrissat & Bairoch, 1993). Esse sistema de nomenclatura foi sendo continuamente atualizado (Henrissat & Davies, 1997, Cantarel *et al*., 2009). Atualmente já existem pelo menos 113 famílias descritas em um banco de dados, o CAZy (*Carbohydrate-Active EnZymes*), disponível em http://www.cazy.org/ (Cantarel *et al*., 2009).

1.3.3.1 Hemicelulases

Hemicelulases são frequentemente classificadas de acordo com sua ação em diferentes substratos. Xilana é o principal polissacarídeo encontrado na hemicelulose e sua completa degradação requer a ação conjunta de várias enzimas hidrolíticas. As hemiceluloses são biodegradadas em monossacarídeos e ácido acético. As enzimas endo-β-1,4-xilanase e β-1,4-xilosidase são funcionalmente distintas: a primeira produz oligossacarídeos a partir da quebra da xilana, enquanto que a segunda hidrolisa os oligossacarídeos produzindo xilose (Lynd, 2002; Pérez et al., 2002).

Tendo em vista que arabinoxilana e β-1,3-1,4-glucano são os polissacarídeos dominantes na hemicelulose da cana-de-açúcar (Lima et al., 2001), a hidrólise dessa fração requer a ação de duas enzimas principais: a β-1,4-xilanase e a β-1,3-1,4-glucanase que fragmentarão, respectivamente, a arabinoxilana e o β-glucano em oligossacarídeos menores. Arabinoxilanas e β-1,3-1,4-glucanos também tem sido encontrados em outras plantas como trigo, centeio, cevada, aveia, arroz e sorgo (Planas et al., 2000; Polizeli et al., 2005). As xilanases (EC 3.2.1.8) são enzimas específicas produzidas principalmente por microrganismos (Damásio et al., 2011; Polizeli et al., 2005), e que participam da degradação da parede celular vegetal, juntamente com outras hidrolases glicolíticas. Além disso, atuam na hidrólise da xilana durante a germinação de algumas sementes (Polizeli et al., 2005), e são classificadas normalmente nas famílias GH (*glycosyl hydrolase*) 10 e GH11 no CAZy.

β-1,3-Glucanos são polímeros importantes encontrados principalmente em leveduras e fungos filamentosos, e são produzidos por plantas (calose) em resposta a danos nos tecidos. A laminarina, um tipo de β-1,3-Glucano, é o principal polissacarídeo de reserva da macro alga marrom do gênero *Laminaria*. Esses polímeros são também produzidos por algumas espécies de fungos e bactérias como exo-polissacarídeos insolúveis (Hrmova & Fincher, 1993; Storseth *et al.*, 2005; Vasur *et al.*, 2007; Zverlov *et al.*, 1997). A hidrólise enzimática dos β-glucanos é catalizada por hidrolases glicolíticas endógenas, como β-1,4-D-glucanase (EC 3.2.1.4), β-1,3-D-glucanase ou laminarase (EC 3.2.1.39) e β-1,3-1,4-D-glucanase ou liqueninase (3.2.1.73), sendo essas enzimas geralmente alocadas na família GH16 do CAZy (Cota et al., 2011, Planas et al., 2000). As β-1,3-D-glucanases são amplamente encontradas em bactérias, fungos e plantas, sendo geralmente classificadas em dois tipos, exo-β-1,3-glucanases (EC 3.2.1.58) e endo-β-1,3-glucanases (EC 3.2.1.6 and 3.2.1.39), que hidrolizam ligações glicosídicas terminais ou internas, respectivamente (Cantarel *et al.*, 2009). Essas enzimas são também classificadas nas famílias GH 16, 17, 55, 64 e 81,

do sistema CAZy, que é baseado principalmente em similaridades estruturais de sequências de aminoácidos (Cantarel *et al.*, 2009).

A biodegradação completa de hemicelulose requer a ação de outras enzimas adicionais, como xilana esterases (EC 3.1.1.72), feruloil esterases (EC 3.1.1.73), endo-β-1,4-mananases (EC 3.2.1.78), β-1,4-manosidases (EC 3.2.1.100), α-1,5-L-arabinanases (EC 3.2.1.99), α-L-arabinofuranosidases (EC 3.2.1.55), α-4-O-metil glucuronidases (EC 3.2.1.139), xiloglucanases (EC 3.2.1.151), entre outras. Essas devem atuar sinergicamente para uma hidrólise eficiente da fração hemicelulósica da parede celular (Lynd et al., 2002; Pérez et al., 2002). A Figura 1.3b-c apresenta a ação conjunta dessas enzimas na biodegradação da hemicelulose. Além de todo o arsenal de enzimas já mensionado, existem enzimas adicionais específicas que atuam na degradação da lignina ligada à hemicelulose, como por exemplo as lacases (EC 1.10.3.2), as lignina peroxidases (EC 1.11.1.14) e as manganês peroxidases (EC 1.11.1.13).

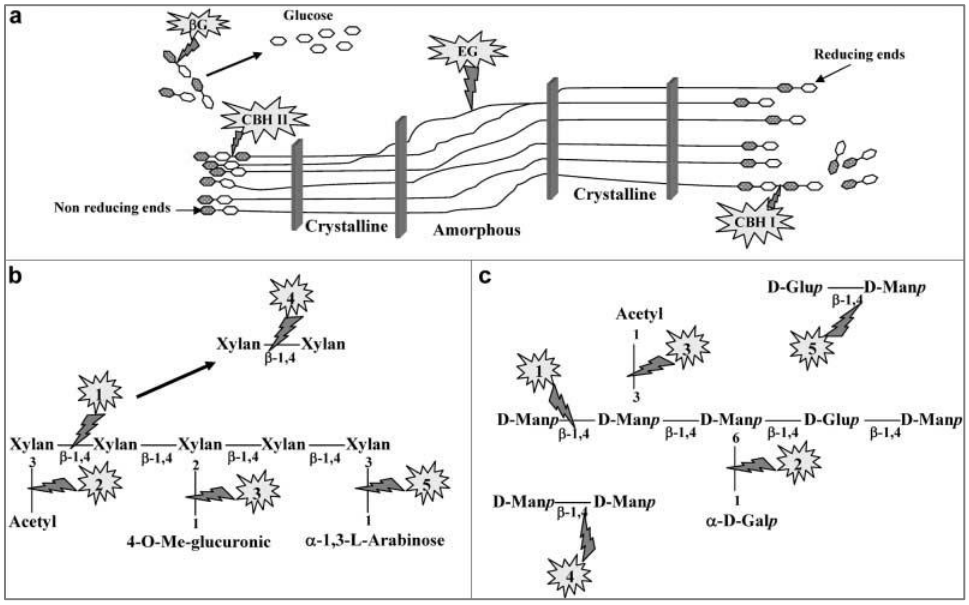


Figura 1.3. (**a**) Degradação enzimática da celulose a glicose. *CBH I* Celobiohidrolase I age nas extremidades redutoras; *CBH II* Celobiohidrolase II age nas extremidades não-redutoras; *EG* endoglucanases hidrolisam ligações internas; *β-G* β-Glicosidases clivam o dissacarídeo celobiose liberando glicose. (**b**) Degradação enzimática de glucuronoxilanas. *1* Endoxilanase, *2* acetilxilana-esterase, *3* α-glucuronidase, *4* β-xilosidase, *5* α-arabinanase. (**c**) Degradação enzimática de glucomanana. *1* Endo-mananase, *2* α-galactosidase, *3* acetilglucomanana-esterase, *4* β-manosidase, *5* β-Glicosidase. (Pérez *et al*., 2002).

1.3.3.2 Celulases

As celulases hidrolisam as ligações glicosídicas do tipo β-1,4 da celulose. Tradicionalmente, elas são divididas em três classes maiores: (i) as endoglucanases (EC 3.2.1.4), que podem hidrolisar ligações internas, preferencialmente nas regiões amorfas da celulose, liberando novas extremidades; (ii) as exoglucanases (EC 3.2.1.74) e as celobiohidrolases (EC 3.2.1.91), que atuam nas extremidades das cadeias geradas pelas endoglucanases, produzindo moléculas de celobiose e outros oligossacarídeos; (iii) as β-glicosidases (EC 3.2.1.21) enzimas cruciais que hidrolisam a celobiose, produzindo duas moléculas de glicose. As celulases podem degradar regiões amorfas de celulose, porém, excepcionalmente as celobiohidrolases são as únicas enzimas que podem agir eficientemente na região cristalina (Lynd *et al*., 2002; Pérez *et al*., 2002). Por hidrolisarem a celobiose, as β-glicosidases desempenham um papel crucial na sacarificação da biomassa. A Figura 1.3a representa a atuação conjunta dessas enzimas na celulose. Além das hidrolases já mencionadas que atuam na degradação da celulose, existem proteínas acessórias sem atividade catalítica que atuam rompendo as ligações de hidrogênio entre as cadeias paralelas de celulose, facilitando o acesso às enzimas celulolíticas (McQueen-Mason & Cosgrove, 1994).

1.3.3.3 β-Glicosidases

β-glicosidases (BGLs) ou celobiases (EC 3.2.1.21) constituem um grupo de enzimas biologicamente importantes e bem caracterizadas, e que hidrolisam uma grande variedade de glicosídeos, incluindo aril e alquil-β-D-glicosídeos, ou oligossacarídeos, dos quais essas hidrolases liberam resíduos de glicose a partir da extremidade não redutora (Bhatia et al., 2002). As funções fisiológicas das BGLs podem variar em função de sua origem (*Archea*, bactérias, fungos, plantas ou animais) e da especificidade pelo substrato. A expressão de β-glicosidases e de outras hidrolases como xilanases e endoglucanases em bactérias vem desempenhando um papel fundamental na conversão de biomassa (Li et al., 2004), mais precisamente na hidrólise de oligossacarídeos de cadeia pequena e celobiose, resultantes da ação sinérgica de outras hidrolases (Bhatia et al., 2002).

As celulases e as BGLs agem sequencialmente e sinergicamente para degradar a celulose à glicose. A principal função da β-glicosidase, comumente chamada de celobiase, é a hidrólise da celobiose e outras celodextrinas à glicose, reduzindo o efeito inibidor da celobiose sobre as enzimas

endo e exoglucanases e celobiohidrolases (Coughlan, 1985; Bhatia, 2002). A β-glicosidase é, portanto, um fator limitante na velocidade global do processo de hidrólise enzimática da celulose. No entanto, as β-glicosidases são em geral inibidas competitivamente pela glicose, o que limita sua atividade (Gueguen et al., 1995). A inibição por glicose e a baixa estabilidade térmica das β-glicosidases constituem os dois maiores desafios ao desenvolvimento de processos comerciais de hidrólise enzimática de celulose a glicose para a subsequente produção de etanol combustível (Woodward & Wiseman, 1982).

Os polissacarídeos complexos de parede celular de plantas possuem um elevado potencial para processos fermentativos em indústrias de alimentos. Entretanto, a liberação dos monossacarídeos necessita da atuação conjunta de várias enzimas hidrolíticas, dentre elas as β-glicosidases, enzimas chave no processo de hidrólise. As BGLs podem ser empregadas em vários processos biotecnológicos em indústrias de alimentos para obtenção de produtos com maior valor agregado. Essas enzimas podem hidrolisar glicosídeos de flavonóides e isoflavonóides que ocorrem naturalmente em frutas, vinho tinto, soja, vegetais e chá, produzindo uma porção aglicona que pode ter efeitos terapêuticos e até atividade antitumoral (Bhatia, 2002). A hidrólise de daidzina e genistina pela β-glicosidase libera, além de glicose, daidzeína e genisteína, respectivamente, o que contribui para a redução dos glicosídeos amargos e adstringentes provenientes da soja (Matsuda et al., 1994). Na indústria de alimentos, o uso de goma gelana, um exopolissacarídeo produzido por *Sphingomonas paucimobilis*, é bem limitado devido a sua alta viscosidade e baixa solubilidade. Nesse caso, β-glicosidases podem ser aplicadas nesses alimentos a fim de reduzir a viscosidade. Outra importante aplicação para β-glicosidases é a produção de compostos voláteis de aroma (terpenóis) por meio da hidrólise de glicosídeos naturais precursores, muito comum em processos de fabricação de vinhos e sucos de frutas (Bhatia, 2002).

De acordo com a nomenclatura da IUBMB as β-glicosidases estão alocadas no grupo EC 3.2.1.21. Baseando-se no sistema de classificação de famílias do CAZy, as β-glicosidases estão inseridas, em geral, nas famílias 1 ou 3 das hidrolases glicolíticas. A família 1 compreende BGLs de diversas origens, que podem apresentar significativa atividade de β-galactosidase além da atividade principal de celobiase. A família 3 consiste de β-glicosidases de bactérias, fungos filamentosos e levedura, que são constituídas por 5 regiões distintas: os resíduos N-terminais, o domínio catalítico N-terminal, uma região não homóloga, um domínio C-terminal de função desconhecida e os resíduos C-terminais. Existem ainda outros possíveis sistemas de classificação para as β-glicosidases, como os baseados na especificidade pelo substrato ou na identidade da sequência de

nucleotídeos (Henrissat & Bairoch, 1996). Baseando-se na especificidade pelo substrato as β-glicosidases podem ser agrupadas em três classes distintas: (1) aril β-glicosidases, que agem em aril-glicosídeos; (2) celobiases verdadeiras, que hidrolisam celobiose e liberam glicose; e (3) enzimas com especificidade por diversos substratos, que atuam em um largo espectro de substratos. Essa última classe engloba muitas das β-glicosidases caracterizadas até o momento (Bhatia *et al*., 2002).

As β-glicosidases catalisam a hidrólise de ligações glicosídicas formadas entre o grupo hemiacetal-OH de uma aldose ou glicose cíclica e o grupo –OH de outros compostos, como açúcar, amino-álcool, aril-álcool ou alcoóis primários, secundários ou terciários. Essa reação acontece em três passos e seu mecanismo está esquematizado na figura 1.4 (Withers & Street, 1988; Bhatia, 2002).

A identificação de resíduos catalíticos pode ser feita por diversas técnicas, como alinhamento de sequências, modificações químicas, mutagênese sítio dirigida, entre outras. Essas informações podem auxiliar no entendimento do mecanismo catalítico que fundamentam as atividades hidrolíticas e sintéticas das β-glicosidases. Resíduos hidrofóbicos como triptofano e tirosina podem desempenhar um papel crucial na forma de ligação ao substrato e orientação ou diretamente na atividade enzimática. Há evidências de que aminoácidos com o grupo –COOH, como os ácidos glutâmico e aspártico, participem como nucleófilo na catálise das β-glicosidases. Resíduos de histidina podem estar diretamente envolvidos na atividade catalítica, ou ainda como potenciais doadores de prótons. Há ainda indícios de que resíduos de cisteína sejam essenciais para a atividade enzimática, possivelmente na formação do intermediário enzima-tio-glicosil (Bhatia, 2002).

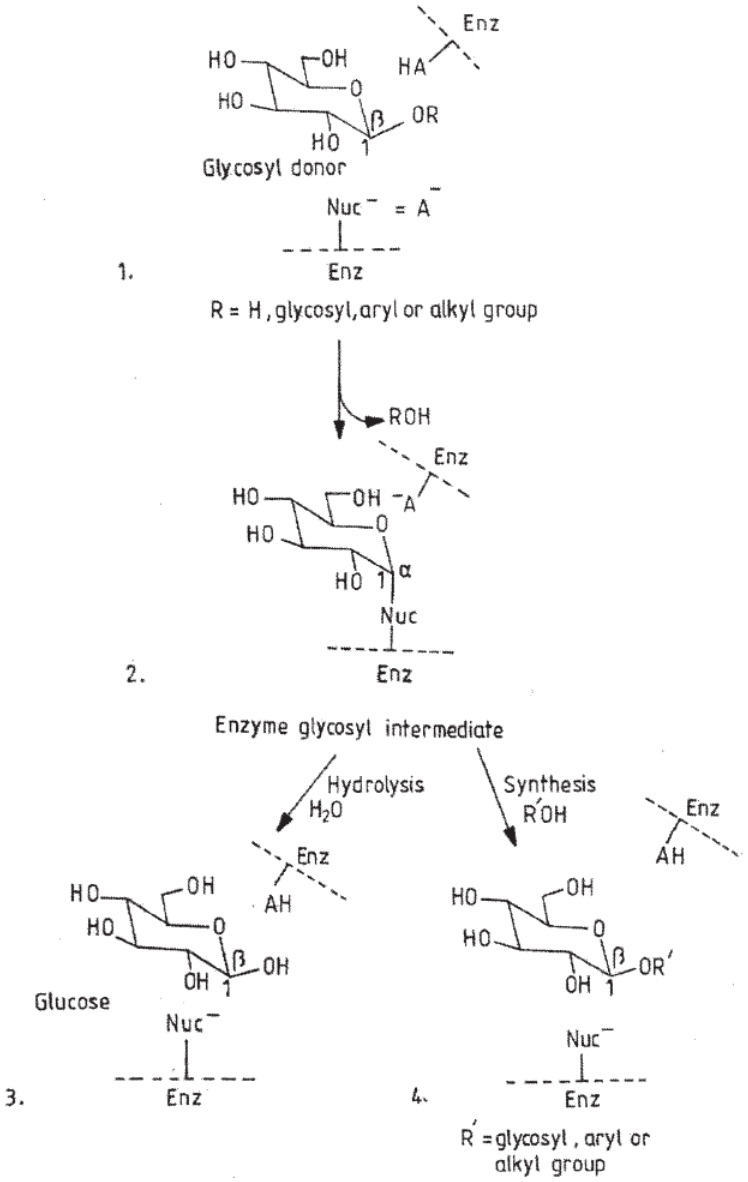


Figura 1.4. Mecanismo da reação de β-glucosidase. (1) o nucleófilo presente no sítio ativo ataca o centro anomérico do substrato, resultando na formação de um intermediário α-glicosil-enzima, ligados covalentemente por meio de um estado de transição do tipo *oxocarbenium ion*; (2) outro resíduo ativo do sítio catalítico funciona como um catalisador ácido base e doa um próton (H^+) para o oxigênio glicosídico, auxiliando assim na excisão do grupo aglicona, ou outro glicídio, como em dissacarídeos; (3) o intermediário glicosil-enzima é então hidrolisado por meio de um ataque geralmente catalisado por base, no qual a água ataca o centro anomérico, liberando β-glicose como produto. (Withers & Street, 1988; Bhatia, 2002).

1.3.4 Engenharia de enzimas

A Engenharia de Enzimas é atualmente uma das mais importantes ferramentas em biotecnologia, e tem sido amplamente utilizada para se obter ganhos em relação às propriedades nativas das proteínas. Várias estratégias de biologia molecular têm sido empregadas no melhoramento de enzimas, sejam elas racionais ou irracionais. A evolução dirigida de proteínas é um processo que gera variabilidade genética e visa a selecionar aquelas que apresentarem melhor desempenho para uma determinada característica. Há atualmente uma variedade de técnicas disponíveis para a geração de bibliotecas de diversidade gênica, principalmente mutagênese aleatória por PCR (*error-prone* PCR; do inglês *Polymerase Chain Reaction*) e recombinação gênica *in vitro* por embaralhamento de DNA (DNA *shuffle*). Essas são consideradas metodologias "irracionais", pois geram variabilidade ao acaso, e não mutações sítio específicas. A técnica de *error-prone* PCR tem sido bastante utilizada na evolução dirigida de enzimas (Leung et al.,1989; Cadwell & Joyce, 1994), contrariando hipóteses referentes à escolha dos melhores sítios para alteração, que embasam metodologias "racionais" (Wittrup, 2002). Muitas bibliotecas de mutantes têm sido construídas para estudos de evolução dirigida de enzimas, como as hidrolases glicolíticas. Os interesses que movem a evolução *in vitro* de proteínas podem ser os mais variados, tais como diminuição da afinidade pelo inibidor, em β-D-xilosidase (Fan et al., 2010); melhorar a catálise em temperaturas mais baixas, em β-glicosidase hipertermofílica (Lebbink et al., 2000); aumento da atividade enzimática, em β-glicosidase (Hardiman et al., 2010).

Por outro lado, uma outra abordagem racional amplamente utilizada é a construção de híbridos de proteínas, seja por recombinação *in vivo* (Weber & Weissmann, 1983), por meio de enzimas de restrição e ligação dos fragmentos *in vitro* (Ay et al., 1998; Ribeiro et al., 2011), ou por fusão de fragmentos por PCR *in vitro* (Horton et al., 1989). Na literatura há vários trabalhos que fazem uso dessa técnica de PCR de fusão, envolvendo desde o aprimoramento do método (Heckman & Pease, 2007) até a construção de quimeras de enzimas glicolíticas por simples justaposição dos genes inteiros separados por um linker (Fan et al., 2009; Levasseur et al., 2006; Lu & Feng, 2008) (Figura 1.5). A produção de proteínas híbridas possui um grande apelo econômico, uma vez que torna-se possível a produção de duas enzimas com diferentes atividades catalíticas em um único polipeptídeo. Além do mais, muitas vezes é possível obter ganhos significativos no perfil catalítico da proteína fusionada em relação às enzimas parentais (Levasseur et al., 2006; Ribeiro et al., 2011). Assim, considerando-se que um dos maiores entraves à viabilidade do etanol celulósico é

o alto custo de produção das enzimas hídrolíticas, o desenvolvimento de enzimas quiméricas tem grande potencial e aplicabilidade para as indústrias de biotransformação de biomassa lignocelulósica.

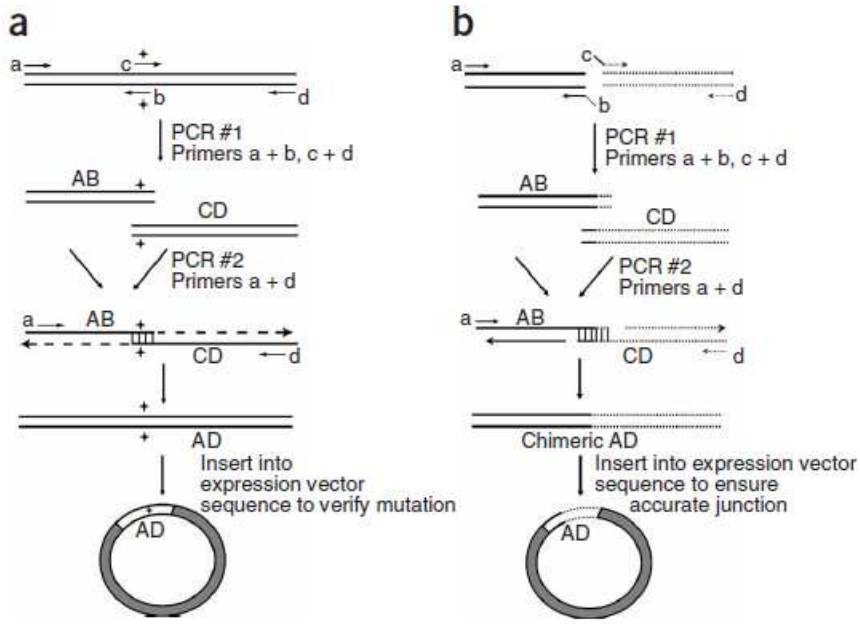


Figura 1.5 A extensão por sobreposição mediada por PCR pode criar mutações em nucleotídeos específicos ou gerar produtos como genes quiméricos (**a**) A mutagênese sítio dirigida é realizada utilizando *primers* mutagênicos (b e c) e *primers* externos (a e d) para gerar os produtos de PCR intermediários AB e CD, que são os fragmentos sobrepostos do produto inteiro AD. Os produtos AB e CD são desnaturados quando utilizado como DNA molde para a segunda PCR; fitas de cada produto hibridizam em suas regiões complementares que se sobrepõem, que também contêm a mutação desejada (indicado por +). A amplificação do produto AD na PCR #2 é conduzida pelos *primers* a e d. O produto final AD pode ser inserido em um vetor de expressão (círculo cinza) para gerar grandes quantidades de DNA, que também deve ser sequenciado para assegurar a presença da mutação desejada. (**b**) Genes quiméricos podem ser gerados por duas PCRs, como em **a**, exceto pelos *primers* internos b e c não serem mutagênicos. Em vez disso, por ser o objetivo unir dois segmentos de genes diferentes por justaposição, os *primers* b e c geram as sequências que se sobrepõem, incluindo os nucleotídeos que se estendem à junção dos segmentos AB (linha sólida) e CD (linha tracejada). A segunda PCR gera o gene híbrido AD, que pode ser inserido num vetor (círculo cinza) para a produção em maior escala e verificação da correta junção dos segmentos AB e CD. (Heckman & Pease, 2007).

1.3.5 Cinética enzimática

Os parâmetros cinéticos de um grande número de enzimas podem ser estimados utilizando o modelo matemático de Michaelis & Menten. Este modelo é adequado para predizer reações onde há saturação com o substrato, isto é, não há incremento na velocidade de reação a partir de um determinado valor de oferta de substrato (Dixon & Webb, 1979; Doran, 2002). A reação enzimática mais simples envolve a transformação de um único substrato em um único produto, e pode ser representada pela sequência a seguir:

$$E + S \overset{k_1}{\Leftrightarrow} ES \overset{k_2}{\Leftrightarrow} EP \overset{k_3}{\Leftrightarrow} E + P$$

onde E representa a enzima, S o substrato, e P o produto, sendo ES e EP os complexos centrais (Segel, 1979).

A equação de velocidade para esse sistema unireacional simples pode ser representada pela equação de Michaelis-Menten, que fornece a velocidade inicial para uma dada concentração de substrato:

$$V_0 = \frac{V_{max}\,[S]}{K_m + [S]}$$

onde a velocidade inicial V_0 aumenta com o acréscimo na concentração de substrato [*S*], e se aproxima de sua máxima taxa V_{max} de modo assintótico quando todas todas as moléculas de enzima estão ligadas ao substrato (V_{max} é extrapolada da curva, V_0 nunca alcança V_{max}), e K_m representa a concentração de substrato onde o valor de *v* é metade de V_{max} (Nelson & Cox, 2005; Segel, 1979).

Na maioria das situações, a etapa limitante da reação é EP → E + P, e neste caso, muitas das enzimas estão na forma EP durante a saturação, e $V_{max} = k_3[E_t]$. Assim, pode-se definir uma constante mais genérica k_{cat} (também chamada de número de *turnover*), que é uma constante de velocidade de primeira ordem e, portanto, sua unidade é de tempo recíproco. A melhor forma de se comparar as eficiências catalíticas de diferentes enzimas ou o *turnover* de diferentes substratos pela mesma enzima é utilizando a relação k_{cat}/K_m (constante de especificidade) para as duas situações distintas. O parâmetro k_{cat}/K_m é uma constante de velocidade de segunda ordem e sua unidade é dada em $M^{-1}s^{-1}$ (Nelson & Cox, 2005).

Os parâmetros cinéticos podem ser estimados através dos dados experimentais utilizando métodos de linearização, como o de Lineweaver-Burk o de Eadie-Hofstee, que são rearranjos da equação de Michelis-Menten. Cada método apresenta suas limitações, sendo o primeiro mais

indicado para o cálculo de V_{max} e o segundo para o cálculo de K_m (Dixon & Webb, 1979; Nelson & Cox, 2005).

1.3.6 Tendências futuras para o uso de biomassa vegetal

A biomassa vegetal representa, neste momento, uma fonte de energia renovável de custo relativamente baixo e de muita abundância. Entretanto, as tecnologias para biotransformação de materiais lignocelulósicos são frequentemente questionadas quanto ao seu maior gargalo: o custo de produção das mais variadas enzimas. Muitas pesquisas têm sido realizadas no sentido da busca por novos organismos produtores de coquetéis enzimáticos mais eficientes. O Brasil detém uma das maiores diversidades biológicas do planeta. Milhões de espécies já foram catalogadas e muitos outros milhões ainda são desconhecidas (Moreira *et al*., 2006). Seus solos e seus ecossistemas apresentam enorme diversidade de microrganismos, o que garante ao Brasil uma grande vantagem frente aos demais países do mundo na prospecção de novas e mais eficientes enzimas hidrolíticas para sacarificação de biomassa vegetal. Dessa forma, o Brasil tem todas as condições para alavancar ainda mais a produção de bioetanol, e dar passos maiores no desenvolvimento de tecnologias para produção do etanol celulósico.

De modo paralelo, no caso de fungos filamentosos celulolíticos/hemicelulolíticos, o melhoramento genético pode, potencialmente proporcionar, ganhos significativos na secreção de enzimas, o que é muito relevante em um processo *scale up*. Além disso, pode-se aumentar significativamente, por engenharia genética, a eficiência catalítica de enzimas recombinantes, bem como alterar sua especificidade por substratos em função da demanda de cada aplicação. Assim, as ferramentas de biologia molecular têm alavancado o desenvolvimento de enzimas sob medida, sendo essa uma forte tendência para os mais variados tipos de bioprocessos.

Atualmente, muito se discute sobre a temática de biocombustíveis de segunda geração, isto é, aqueles produzidos a partir de biomassa vegetal. Não obstante, o conceito de química verde tem se mostrado um excelente caminho para a criação de biorefinarias integradas, onde outras moléculas de maior valor agregado que o etanol podem ser extraídas ou sintetizadas a partir dos componentes da parede celular vegetal. Entre os químicos que podem ser obtidos, pode-se destacar: butanol, isobutanol, propanol, ácido butírico, xilooligossacarídeos, ácido ferúlico e outros compostos aromáticos (Carvalho *et al.*, 2013; Damásio *et al.*, 2012; Dellomonaco *et al.*, 2010, Gonçalves *et al.*, 2012; Wei *et al.*, 2013). Dessa forma, há um grande potencial no aproveitamento de resíduos

agroindustriais e outras fontes de biomassa vegetal para produção de, não apenas etanol de segunda geração, mas também outros combustíveis e moléculas de maior valor agregado, estabelecendo-se então um desenvolvimento sustentável.

REFERÊNCIAS BIBLIOGRÁFICAS

Anastas, P. T.; Warner, J. C. Green Chemistry: Theory and Practice. Oxford University Press, Oxford, 1998.

Aÿ, J.; Götz, F.; Borriss, R.; Heinemann, U. Structure and function of the *Bacillus* hybrid enzyme GluXyn-1: native-like jellyroll fold preserved after insertion of autonomous globular domain, P. Natl. Acad. Sci. USA, v. 95, p. 6613-6618, 1998.

Bals, B. D.; Dale, B. E. Developing a model for assessing biomass processing technologies within a local biomass processing depot. Bioresour. Technol., v. 106, p. 161-169, 2012.

Bastos V.D. Etanol, alcoolquímica e biorrefinarias. BNDES Setorial, Rio de Janeiro, n. 25, p. 5-38, 2007.

Brasil. Ministério de Minas e Energia; colaboração Empresa de Pesquisa Energética. Matriz Energética Nacional 2030. Brasília, 254p., 2007.

Bhatia, Y.; Mishra, S.; Bisaria, V. S. Microbial β-glicosidases: cloning, properties and applications. Crit. Rev. Biotechnol., v. 22, p. 375-407, 2002.

Cadwell, R. C.; Joyce, G. F. Mutagenic PCR. PCR Meth. Appl., v. 3, p. S136-140, 1994.

Cantarel, B. L.; Coutinho, P. M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res., v. 37, p. 233-238, 2009.

Carvalho, A. F. A.; Neto, P. O.; Silva, D. F. *et al*. Xylo-oligosaccharides from lignocellulosic materials: chemical structure, health, benefits and production by chemical and enzymatic hydrolysis. Food Res. Int., v. 51, p. 75-85, 2013.

Carvalho, W; Canilha, L.; Ferraz, A.; Milagres, A. M. F. Uma visão sobre a estrutura, função e biodegradação da madeira. Quim. Nova, v. 32, n. 8, p. 2191-2195, 2009.

Castro, A. M.; Pereira Jr., N. Produção, propriedades e aplicação de celulases na hidrólise de resíduos agroindustriais. Quim. Nova, v. 33, n. 1, p. 181-188, 2010.

Cerqueira-Leite, R. C. (Org). Bioetanol combustível: uma oportunidade para o Brasil. Brasília: Centro de Gestão e Estudos Estratégicos, 536 p., 2009.

Cota, J.; Alvarez, T. M.; Citadini, A. P.; Santos, C. R.; Oliveira-Neto, M.; Oliveira, R. R.; Pastore, G. M.; Ruller, R.; Prade, R. A.; Murakami, M. T.; Squina, F. M. Mode of operation and low-resolution structure of a multi-domain and hyperthermophilic endo-β-1,3-glucanase from *Thermotoga petrophila*, Biochem. Bioph. Res. Co., v. 406, p. 590-594, 2011.

Coughlan, M. P. The properties of fungal and bacterial cellulases with comment on their production and application. Biotech. Gen. Eng. Rev., v. 3, p. 39-109, 1985.

Damásio, A. D. R. L.; Braga, C. M. P.; Brenelli, L. B. *et al*. Biomass-to-bio-products application of feruloyl esterase from *Aspergillus clavatus*. Appl. Microbiol. Bioechnol., 2012.

Damásio, A. D. R. L.; Silva, T. M.; Almeida, F. B. D. R. *et al*. Heterologous expression of an *Aspergillus niveus* xylanase GH11 in *Aspergillus nidulans* and its characterization and application. Process Biochem., v. 46, p. 1236-1242, 2011.

Dellomonaco, C.; Fava, F.; Gonzalez, R. The path to next generation biofuels: successes and challenges in the era of synthetic biology. Microbial Cell Factories, v. 9, n. 3, p. 1-15, 2010.

Dixon, M.; Webb, E. C.; Enzymes, 3rd Ed., Academic Press, New York, 1116 p., 1979.

Doran, P. M.; Bioprocess Engineering Principles, Academic Press, New York, 439 p., 2002.

Fan, Z.; Wagschal, K.; Lee, C. C.; Kong, Q.; Shen, K. A.; Maiti, I. B.; Yuan, L. The construction and characterization of two xylan-degrading chimeric enzymes, Biotechnol. Bioeng., v. 102, p. 684-692, 2009.

Fan, Z.; Yuan, L.; Jordan, D.B.; Wagschal, K.; Heng, C.; Braker, J.D. Engineering lower inhibitor affinities in β-D-xylosidase. Appl. Microbiol. Biot., v. 86, p. 1099-1113, 2010.

Gonçalves, T. A.; Damásio, A. R. L.; Segato, F. *et al*. Functional characterization and synergic action of fungal xylanase and arabinofuranosidase for production of xylooligosaccharides. Biores. Tech., v. 119, p. 293-299, 2012.

Gueguen, Y.; Chemardin, P.; Arnaud, A.; Galzy, P. Purification and characterization of an intracellular β-glicosidase from *Botrytis cinerea*. Enzyme Microb. Tech., v. 17, p. 900-906, 1995.

Hahn-Hägerdal, B.; Galbe, M.; Gorwa-Grauslund, M. F.; Lidén, G.; Zacchi, G. Bio-ethanol – the fuel of tomorrow from the residues of today. Trends Biotechnol., v. 24, n. 12, p. 549-556, 2006.

Hardiman, E.; Gibbs, M.; Reeves, R.; Bergquist, P. Directed evolution of a thermophilic β-glicosidase for cellulosic bioethanol production. Appl. Biochem. Biotech., v. 161, p. 301-312, 2010.

Heckman, K. L.; Pease, L. R. Gene splicing and mutagenesis by PCR-driven overlap extension, Nat. Protoc., v. 2, p. 924-932, 2007.

Henrissat, B. A classification of glycosyl-hydrolases based on amino acid sequence similarities. Biochem. J., v. 280, p. 309-316, 1991.

Henrissat, B.; Bairoch, A. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J., v. 293, p. 781-788, 1993.

Henrissat, B.; Bairoch, A. Updating sequence-based classification of glycosyl hydrolases. Biochem. J., v. 316, p. 695-696, 1996.

Henrissat, B.; Davies, G.J. Structural and sequence based classification of glycosyl hydrolases. Curr. Opin. Struc. Biol., v. 7, p. 637-644, 1997.

Horton, R. M.; Hunt, H. D.; Ho, S. N.; Pullen, J. K.; Pease, L. R: Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene, v. 77, p. 61-68, 1989.

Hrmova, M.; Fincher, G.B. Purification and properties of three (1,3)-β-D-glucanase isoenzymes from young leaves of barley (*Hordeum vulgare*). Biochem. J., v. 289, p. 453–461, 1993.

Lebbink, J.H.G.; Kaper, T.; Bron, P.; van der Oost, J.; de Vos, W.M. Improving low-temperature catalysis in the hyperthermostable *Pyrococcus furiosus* β-glicosidase CelB by directed evolution. Biochemistry, v. 39, p.3656-3665, 2000.

Leung, D.W.; Chen, E.; Goeddel, D.V. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. Technique, v. 1, p. 11-15, 1989.

Levasseur, A.; Saloheimo, M.; Navarro, D.; Andberg, M.; Monot, F.; Nakari-Setälä, T.; Asther, M.; Record, E. Production of a chimeric enzyme tool associating the *Trichoderma reesei* swollenin with the *Aspergillus niger* feruloyl esterase A for release of ferulic acid. Appl. Microbiol. Biotechnol., v. 73, p. 872-880, 2006.

Li, X.; Ljungdahl, L.G.; Ximenes, E.A.; Chen, H.; Felix, C.R.; Cotta, M.A.; Dies, B.S. Properties of a recombinant β-glicosidase from the polycentric anaerobic fungus *Orpinomyces* PC-2 and its application for cellulose hydrolysis. Appl. Biochem. Biotechnol., v. 113-116, p. 233-250, 2004.

Lima, D. U.; Santos, H. P.; Tiné, M. A.; Molle, F. R. D.; Buckeridge, M. S. Patterns of expression of cell wall related genes in sugarcane. Genet. Mol. Biol., v. 24, p. 191-198, 2001.

Lynd, L. R.; Cushman, J. H.; Nichols, R. J.; Wyman, C. E. Fuel ethanol from cellulosic biomass. Science, v. 251, p. 1318-1323, 1991.

Lynd, L. R.; Weimer, P. J.; van Zyl, W. H.; Pretorius, I. S. Microbial Cellulose Utilization: Fundamentals and Biotechnology. Microbiol. Mol. Biol. R., v. 66, p. 506-577, 2002.

Lynd, L. R. et. al. Consolidated bioprocessing of cellulosic biomass: an update. Curr. Opin. Biotechnol. v. 16, p. 577-583, 2005.

Lu, P.; Feng, M. G. Bifunctional enhancement of a β-glucanase-xylanase fusion enzyme by optimization of peptide linkers. Appl. Microbiol. Biotechnol., v. 79, p. 579-587, 2008.

Madson, M.; Dunand, C.; Li, X.; Verma, R.; Vanzin, G.F.; Caplan, J.; Shoue, D.A.; Carpita, N.C.; Reiter, W.D. The MUR3 gene of *Arabidopsis* encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. Plant Cell, v. 15, p. 1662-1670, 2003.

Makkar, R. S.; Cameotra, S. S. An update on the use of unconventional substrates for biosurfactant production and their new applications. Appl. Microbiol. Biotechnol., v. 58, p. 428-434, 2002.

Marana, S. R.; Terra, W. R.; Ferreira, C. Purification and properties of a β-glycosidase purified from midgut cells of *Spodoptera frugiperda* (Lepidoptera) larvae. Insect Biochem. Molec., v. 30, p. 1139-1146, 2000.

Matsuda, S.; Norimoto, F.; Matsumoto, Y.; Ohba, R.; Teramoto, Y.; Ohta, N.; Veda, S. Solubilization of a novel isoflavone glycoside hydrolysing β-glicosidase from *Lactobacillus casei* subsp. *rhamnosus*. J. Ferment. Bioeng., v. 77, p. 439-441, 1994.

McQueen-Mason, S.; Cosgrove, D.J. Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. Proc. Natl Acad. Sci. USA, v. 91, p. 6574-6578, 1994.

Minic, Z.; Jouanin, L. Plant glycoside hydrolases involved in cell wall polysaccharide degradation. Plant Physiol. Bioch., v. 44, p. 435-449, 2006.

Moreira, F. M. S.; Siqueira, J. O.; Brussaard, L. Soil Biodiversity in Amazonian and Other Brazilian Ecosystems. Oxfordshire: CABI Publishing , 304p, 2006.

Nelson, D.L.; Cox, M.M. Lehninger principles of biochemistry, 4th edition, W.H. Freeman, New York, 1119 p., 2005.

Neto, A. Q. M. (Org). Álcool Combustível. Série Indústria em Perspectiva. Brasília: Instituto Euvaldo Lodi, 163p., 2008.

Nogueira, L.A.H. (Org). Bioetanol de Cana-de-Açúcar: Energia para o Desenvolvimento Sustentável. Rio de Janeiro: BNDES/CGEE, 316p, 2008.

Ogeda, T. L.; Petri, D. F. S. Hidrólise enzimática de biomassa. Quim. Nova, v. 33, n. 7, p. 1549-1558, 2010.

Orzua, M. C.; Mussatto, S. I.; Contreras-Esquivel, J. C.; Rodriguez, R.; de la Garza, H.; Teixeira, J. A.; Aguilar, C. N. Exploitation of agro industrial wastes as immobilization carrier for solid-state fermentation. Ind. Crop. Prod., v. 30, p. 24-27, 2009.

Pérez, J.; Munõz-Dorado, J.; de La Rubia, T.; Martínez, J. Biodegradation and biological treatments of cellulose, hemicelluloses and lignin: an overview. Int. Microbiol., v. 5, p. 53–63, 2002.

Planas, A. Bacterial 1,3-1,4-beta-glucanases: structure, function and protein engineering, Biochim. Biophys. Acta, v. 1543, p. 361-382, 2000.

Polizeli, M. L. T. M.; Rizzatti, A. C. S.; Monti, R.; Terenzi, H. F.; Jorge, J. A.; Amorim, D. S. Xylanases from fungi: properties and industrial applications, Appl. Microbiol. Biotechnol., v. 67, p. 577-591, 2005.

Potters, G.; Van Goethem, D.; Schutte, F. Promising biofuel resources: lignocellulose e algae. Nat. Educ., v. 3, n. 9, p. 14, 2010.

Rodrigues, M. I.; Iemma, A. F. Planejamento de experimentos e otimização de processos: uma estratégia seqüencial de planejamentos. Campinas: Editora Casa do Pão, 326p, 2005.

Ribeiro, L. F.; Furtado, G. P.; Lourenzoni, M. R. *et al*. Engineering bifunctional laccase-xylanase chimerae for improved catalytic performance. J. Biol. Chem., v. 286, p. 40026-40038, 2011.

Ruller, R.; Deliberto, L. A.; Ferreira, T. L.; Ward, R. J. Thermostable variants of the recombinant xylanase A from *Bacillus subtilis* produced by directed evolution show reduced heat capacity changes. Proteins, v. 70, p. 1280-1293, 2008.

Ruller, R.; Rosa, J. C.; Faca, V. M.; Greene, L. J.; Ward, R. J. Efficient constitutive expression of *Bacillus subtilis* Xylanase A in *Escherichia coli* DH5alpha under the control of the *Bacillus* BsXA promotor. Biotechnol. Appl. Biochem., v. 28, p. 9-15, 2006.

Sanders, J.; Scott, E.; Weusthuis, R.; Mooibroek, H. Bio-refinery as the bio-inspired process to bulk chemicals. Macromol. Biosci., v. 7, p. 105-117, 2007.

Segel, I. H. Bioquímica: teoria e problemas; 2nd Ed.; Livros Técnicos e Científicos Editora; Rio de Janeiro; 529 pp.; 1979.

Sierra, R.; Smith, A.; Granda, C.; Holtzapple, M. T. Producing fuels and chemicals from lignocellulosic biomass. SBE Special Section - Biofuels, p. S10-S18, 2008.

Smith B. Sweden: Making Money And Energy Off Of Euro Trash. RedOrbit, 29/10/12. Disponível em: http://www.redorbit.com/news/science/1112722160/sweden-waste-to-energy-102912/. Acesso em 01.nov.2012.

Storseth, T. R.; Hansen, K.; Reitan, K. I.; Skjermo, J. Structural characterization of β-D-(1,3)-glucans from different growth phases of the marine diatoms *Chaetoceros mülleri* and *Thalassiosira weissflogii*. Carbohydr. Res., v. 340, p. 1159–1164, 2005.

Takagi, M.; Abe, S.; Suzuki, S.; Emert, G.H.; Yata, N. A method for production of ethanol directly from cellulose using cellulase and yeast. In: Ghose, T.K. (Ed.), Proceedings of Bioconversion Symposium. IIT, Delhi, p. 551-571, 1977.

Truck, C. O.; Pérez, E.; Horváth, I. T.; Sheldon, R. A.; Poliakoff, M. Valorization of biomass: deriving more value from waste. Science, v. 337, p. 695-699, 2012.

Vasur, J.; Kawai, R.; Andersson, E. et al. X-ray crystal structures of *Phanerochaete chrysosporium* Laminarinase 16A in complex with products from lichenin and laminarin hydrolysis. FEBS J., v. 276, p. 3858–3869, 2007.

Wang, D. I. C.; Avgerinos, G. C.; Biocic, I.; Wang, S. D.; Fang, H. Y. Ethanol from cellulosic biomass. Philos. Trans. R. Soc. Lond. Ser. Bot., v. 300, p. 323-333, 1983.

Weber, H.; Weissmann, C. Formation of genes coding for hybrid proteins by recombination between related, cloned genes in *Escherichia coli*. Nucleic Acids Res., v. 11, p. 5661-5669, 1983.

Wei, D.; Liu, X.; Yang, S. T. Butyric acid production from sugarcane bagasse hydrolysate by Clostridium tyrobutyricum immobilized in a fibrous-bed bioreactor. Biores. Tech., v. 129, p. 553-560, 2013.

Withers, S.G.; Street, I.P. Identification of a covalent α-D-glucopyranosyl enzyme intermediate formed on a β-glicosidase. J. Am. Chem. Soc., v. 110, p. 88551-8553, 1988.

Wittrup, K. D. Directed Evolution of Binding Proteins by Cell Surface Display: Analysis of the Screening Process. In: Brakmann, S.; Johnsson, K. (Org). Directed Molecular Evolution of Proteins. Weinheim: Wiley-VCH, 357p, 2002.

Woodward, J.; Wiseman, A. Fungal and other β-D-glycosidases – their properties and applications. Enzyme Microb. Tech., v. 4, p. 73-79, 1982.

Wright, J. D. Ethanol from biomass by enzymatic hydrolysis. Chem. Eng. Prog., v. 84, p. 62-74, 1988.

Zverlov, V. V.; Volkov, I. Y.; Velikodvorskaya, T. V.; Schwarz, W.H. Highly thermostable endo-1, 3-beta-glucanase (laminarinase) LamA from *Thermotoga neapolitana*: nucleotide sequence of the gene and characterization of the recombinant gene product. Microbiology, v. 143, p. 1701–1708, 1997.

CAPÍTULO 2

HYDROLYTIC PATTERN CHARACTERIZATION AND LOW-RESOLUTION STRUCTURE OF HYPERTHERMOPHILIC ENDO-β-1,3-GLUCANASE FROM THERMOTOGA PETROPHILA

ABSTRACT

1,3-β-Glucan depolymerizing enzymes have considerable biotechnological applications including biofuel production, feedstock-chemicals and pharmaceuticals. Herein we show the comprehensive functional characterization and low-resolution structure of hyperthermophilic laminarase from *Thermotoga petrophila* (TpLam). We determined the TpLam enzymatic mode of operation through capillary zone electrophoresis, which specifically cleaves internal β-1,3-glucosidic bonds. The enzyme most frequently attach the 3rd and 4th bond from the non-reducing end residue on oligosaccharides (laminohexaose), whereas produces glucose, as well laminaribiose and laminaritriose as end as major products from polysaccharides. Far-UV circular dichroism demonstrated that LamA is formed mainly by beta structural elements, and the secondary structure is maintained after incubation up to 16 hours at 90ºC. The structure determined by small angle X-ray scattering revealed a multi-domain structural architecture of the enzyme with a V-shape envelope arrangement of the two carbohydrate binding modules in relation to the catalytic domain.

Keywords: endo-β-1,3-glucanase; glycoside hydrolase family 16; hyperthermostable enzyme; capilary zone Electrophoresis, small angle X-ray scattering, hyperthermostable enzyme, *Thermotoga sp.*

2.1 INTRODUCTION

Renewable structural polysaccharides are abundant resource in the biosphere and represent a valuable industrial substrate for many industrial processes, such as biofuels production, feedstock-chemicals and pharmaceuticals [1]**.** β-1,3-glucans are important polymers found mainly in yeast and filamentous fungi and are produced by plants (as callose) in response to tissue damage. This polymer is a major structural storage polysaccharide (laminarin) of the marine brown macro-algae of the genus *Laminaria*. β-1,3-glucan is also produced as insoluble exopolisaccharide by some fungi and bacteria species [2-5]**.**

The β-1,3-glucanases are widely distributed among bacteria, fungi, and plants and are generally classified into two types, exo-β-1,3-glucanases (EC 3.2.1.58) and endo-β-1,3-glucanases (EC 3.2.1.6 and 3.2.1.39), which hydrolyze terminal or internal linkages, respectively. This enzymes are further classified as been members into glycosyl hydrolase families (GH) 16, 17, 55, 64 and 81, by structure and similarity of the amino acid sequence. However, these GHs display a wide range of enzymatic activity, and little information regarding biochemical and biophysical properties are available.

To date, crystal structures from family GH 16 endo-β-1,3-glucanases (3.2.1.39) have been determined for *Pyrococcus furiosus* [6], *Rhodothermus marinus* [6], *Streptomyces sioyaensis* [7] and *Nocardiopsis* sp [8].The active site region of family GH16 bears three acidic residues of bacterial GH16, two glutamic acid and one aspartic acid in the active site, nucleophile, and acid and base components, and two tryptophan which promote the substrate binding and transition-state stabilization, are conserved in this group of enzymes [9]. The overall fold of these enzymes consists of a classical sandwich-like β-jelly-roll motif formed by the face-to-face packing of two antiparallel sheets containing seven and eight strands with a deep cavity [6-8,10].

In this work, we report a comprehensive characterization of hyperthermostable endo-β-1,3-glucanases from *Thermotoga petrophila* (TpLam). Here, we show a detailed description of the hydrolytic mode of operation and hyperthermophilicity of the recombinant enzyme along with structural insights of TpLam.

2.2 MATERIAL AND METHODS

2.2.1 Cloning and purification of TpLam

The full-length coding TpLam gene (Tpet_0899), cloned into pBAD/Myc-His B vector (Invitrogen), was used as template for standard PCR method for cloning the mature enzyme without signal peptide. The primer set used for amplification was 5´-GCTAGCCAAAACATCCTTGGCAACGC-3´ and 5´-GGATCCTCATTGAGGACTCACCGAAA-3´, and the gene segment was cloned into the pET28a (Novagen) vector using NdeI and BamHI restriction sites. The protein expression and purification steps, included a Ni^{2+}-chelating affinity and size-exclusion chromatography, were performed following Squina et al., 2010 [11]. The purified TpLam was further analyzed by SDS-PAGE. Protein concentration was determined by absorbance at 280 nm.

2.2.2 Enzyme characterization

The standard enzymatic assays for TpLam evaluation were performed following previous work [29]. The determination of the optimum pH and temperature profiles, thermostability evaluation and 8-aminopyreno-1,3,6-trisulfonic acid (APTS) labeling was performed as described previously [28]. The mixture was incubated at 90 ^{o}C for 10 min in standard assays with laminarin or 60 min aiming at to determine the specific activity in a set of polyssacharides (purchased from the best source possible, Sigma Aldrich and Megazyme). The enzymatic activity and substrate specificity were determined from the amount of reducing sugar liberated from different polysaccharide substrates by the DNS method [12]. One unit of enzyme was defined as the quantity of enzyme that released reducing sugars at rate of 1 μmol/min. The kinetic parameters for TpLam activity were estimated from initial rates at twelve substrates concentration in the 1.6–65 mg/mL range in standard assay by using Michaelis-Menten approach.

A Response Surface Methodology was performed to optimize the reaction conditions for TpLAM. The variables analyzed here were the pH together with temperature, whereas a Central Composite Design (k = 2) with four central points totalizing 12 experiments (Table 2.1) was considered for optimization of these variables. Details concerning the statistical approaches for these experiments can be found in Myers and Montgomery [13] and all β-1,3-glucanase activity

assays were carried out following our previous work using laminarin as substrate (stock 0,5% in water) [14, 15].

Capillary electrophoresis of oligosaccharides and Far-UV circular dichroism (CD) measurements (190–260 nm) and prediction of secondary structure from the dichroism circular spectrum was conducted as described as previously described and [14 - 16].

2.2.3 Small angle X-ray scattering

Small Angle X-ray Scattering (SAXS) data for TpLam at the concentrations 4 and 8 mg/mL were collected on the SAXS2 beamline at the Brazilian Synchrotron Light Laboratory. The radiation wavelength was set to 1.48 Å and a 165 mm MarCCD detector was used to record the scattering patterns. The sample-to-detector distance was set to 1022.5 mm to give a scattering vector-range from 0.013 to 0.33 nm^{-1}, where q is the magnitude of the q-vector defined by $q= 4\pi/ \lambda \sin \theta$ (2θ is the scattering angle). Protein samples were prepared in 20mM phosphate 50 mM NaCl buffer pH 6.0. The integration of SAXS patterns were performed using Fit2D software [17] and the curves were scaled by protein concentration. The radius of gyration, R_g was approximated using two independent procedures, by Guinier equation [18] and by indirect Fourier transform method using GNOM program [18]. The distance distribution functions p® also was evaluated by GNOM and the maximum diameter, D_{max} was obtained. Molecular weight was calculated for all constructions and complexes using the procedure implemented on webtool SAXSmoW [19]. Dummy atom models (DAMs) were calculated from the experimental SAXS data using ab initio procedure implemented in Dammin program [20]. Several runs of ab initio shape determination with different starting conditions led to consistent results as judged by the structural similarity of the output models, yielding nearly identical scattering patterns and fitting statistics in a stable and self-consistent process. CRYSOL program calculated the simulated scattering curve from DAM and homology model [21]. The evaluation of R_g, D_{max}, and discrepancy parameters (χ) were performed with the same program.

2.2.4 Homology molecular modeling

Homology models for each domain of TpLam were constructed using the HHPred server [22]. These models were renumbered according to the wild protein and spatially disposed in this order: 1,

2, and 3, in a manner to avoid clashes, with MASSHA [23]. CRYSOL program calculated the theoretical scattering curve for each model and it was input for the next approach [24]. BUNCH software, which employs rigid body model (RBM) and simulated annealing routines, was used to find the best relative positions of the high resolution models [25]. This solution was superposed in the SAXS envelope with SUPCOMB [25].

2.3 RESULTS AND DISCUSSION

2.3.1 Functional and Biophysical Characterization of TpLam

TpLam activity was investigated at different pH values and temperatures. The enzyme exhibited at least 80% of its optimal activity over a rather broad pH range, from 5 to 7 (data not shown), which resemble other hyperthermophilic laminarases [2, 26]. Purified TpLam showed optimum activity at 87°C with a relative activity ranging from 82% to 86% at temperatures of 78°C and 95°C, respectively (data not shown). A CCRD (central composite rotatable design) with two variables was performed to optimize reaction conditions. The model (Figure 2.1A and 2.1B) was validated by an analysis of variance (ANOVA) showing that the model was significant at high confidence level (95%), with $R^2 = 0.96$. The best condition of enzyme activity was reached pH 6.2 and 91 ºC. To validate the model it was chosen pH 6 and temperature of 91ºC. The predicted laminarinase activity would be 250.0 IU/mg protein and the value found in the experiment was 260.8 IU/mg protein.

TpLam was fully heat-stable at 70°C and maintain 60% of residual activity after 16 hours at 80°C. The enzyme showed a time of half-life ($t_{1/2}$) of 577 and 126 minutes at 90 and 95°C, respectively (Figure 2.1C). The optimum temperature reported for the TpLam is comparable to other thermophilic bacterial endo-β-1,3-glucanases from *Thermotoga neapolitana* and *Pyrococcus furiosus*, which display temperatures 95°C and 100°C, respectively [2,26].

Table 2.1. CCRD matrix (2^2) and the response of TpLam activity after 10 minutes of incubation.

Run no.	Coded levels (X_1 = pH; X_2 = *T*)		Actual levels (X_1 = pH; X_2 = *T*)		Laminarinase activity (IU/mg protein)
	X_1	X_2	X_1	X_2	
1	-1	-1	4.6	74.2	139,8
2	1	-1	7.4	74.2	170,9
3	-1	1	4.6	94.8	175,2
4	1	1	7.4	94.8	204,2
5	-1.414	0	4.0	84.5	68,6
6	1.414	0	8.0	84.5	154,6
7	0	-1.414	6.0	70.0	159,7
8	0	1.414	6.0	99.0	238,9
9	0	0	6.0	84.5	234,6
10	0	0	6.0	84.5	241,3
11	0	0	6.0	84.5	247,8
12	0	0	6.0	84.5	247,1

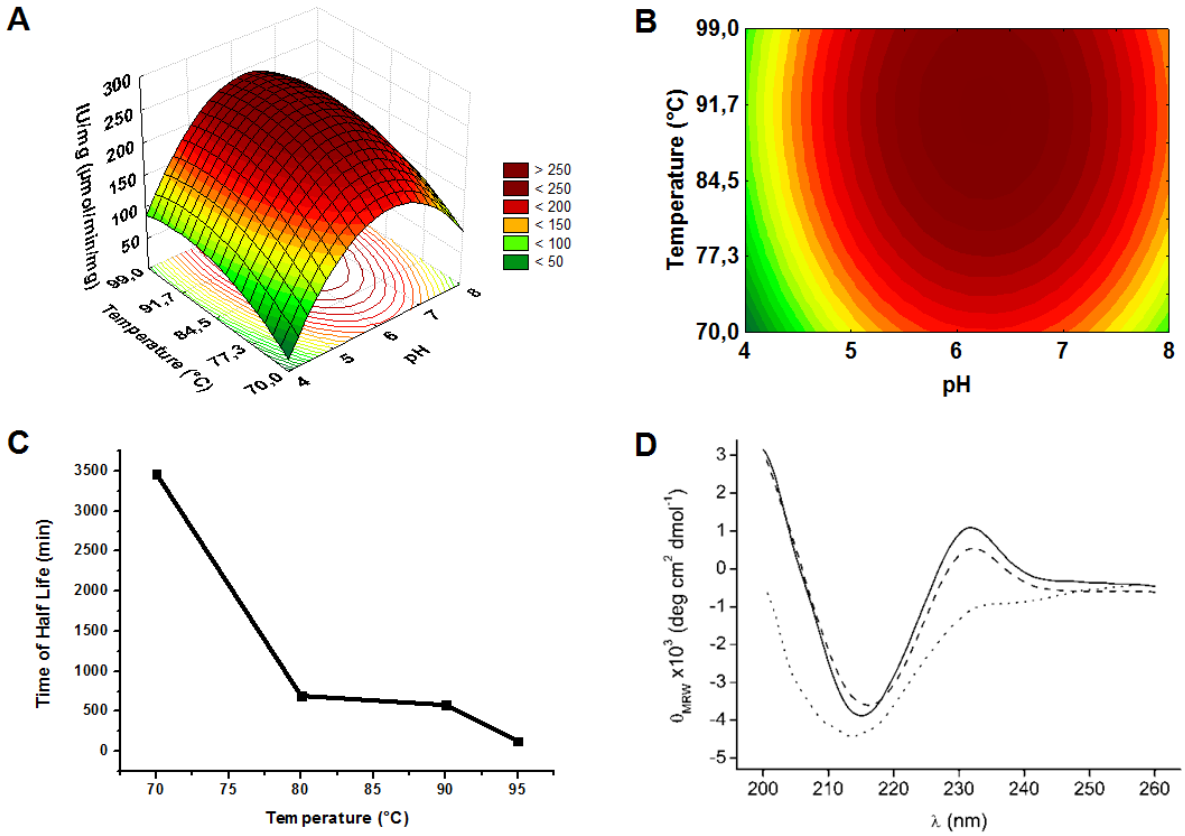


Figure 2.1. Response surface (A) and contour plot (B) for the influence of temperature and pH on TpLam activity. (C) Curve of decay of time of half life ($t_{1/2}$) using four different temperatures (70, 80, 90 and 95 °C). (D) Thermal stability of laminarinase analyzed by circular dichroism. CD spectra were taken at 20°C (solid line), at 90°C (dashed line) and after 18 hours at 90°C.

The recombinant enzyme was highly specific on β-1,3 contain polysaccharides, such as laminarin (48.07 IU/mg) and (β-glucan 41.55 IU/mg) from barley and lichenan in a lower scale (21.37 IU/mg). No β-1,4 glucanase activity was detected, based on enzymatic assays using avicel, carboxymethylcellulose and cellopentose, as well, neither activity was found in xylans, arabinan and manan containing polysaccharides. The constants values were determined experimentally from the initial rates of laminarin hydrolysis at each substrate concentration and standard assay conditions. A V_{max} of 558.2 μmol/min/mg and a K_m of 3.3 mg/mL were found for purified TpLam. Two other kinetic constants were determined from the experimental data and the values for catalytic constant (K_{cat}) and K_{cat}/K_m were 680.3 s^{-1} and 206.2, respectively.

The CD spectrum of native laminarinase presents a negative peak at 215 nm and a positive peak at 232 nm (Figure 2.1D). A positive peak would also be observed around 195 nm if we could take data at smaller wavelengths. The 215 nm peak is characteristic of β-structure and the 232 nm peak is explained by the aromatic residues contribution. All the enzyme domains have a high content of beta-sheet, with few contributions of α-helices to the structure. Glycoside hydrolases are rich in aromatic residues, which are notably important to form the active site pocket and to interact with the substrate. Those side chains can contribute to the CD spectrum, especially in protein of low helix content, such as the laminarinase, giving positive bands in the 220-230 region [27]. At 90ºC little change is observed in the CD spectrum of laminarinase, showing conservation of the secondary structure content (Figure 2.1D). After incubation at 90ºC for 18 hours and a CD spectrum has been taken. The disappearing of the 232 nm peak would indicate the protein unfolding, but the 215 nm peak increasing indicates that β-sheet secondary structure was maintened (Figure 2.1D).

Capillary zone electrophoresis analysis of the corresponding hydrolysis products was performed in an attempt to define the mode of action for TpLam. During the course of hydrolysis of APTS-labeled laminarihexaose (non-reducing end available) the enzyme most likely attacks the internal glycosidic linkages releasing fluorescent dimers, trimers and tetramers (Figure 2.2A). Total hydrolysis resulted in only dimers and trimers, suggesting that they cannot be degraded further (Figure 2.2A). Nevertheless through APTS labeling after the enzymatic hydrolys of laminarihexaose was possible to see the formation of glucose (data not shown). Using natural laminarin (APTS-labeled after the reaction) the degradation pattern was characteristic for endoglucanases, production of intermediates with different degree of polimerization, yielding as final product glucose, laminaribiose and laminaritriose (Figure 2.2B).

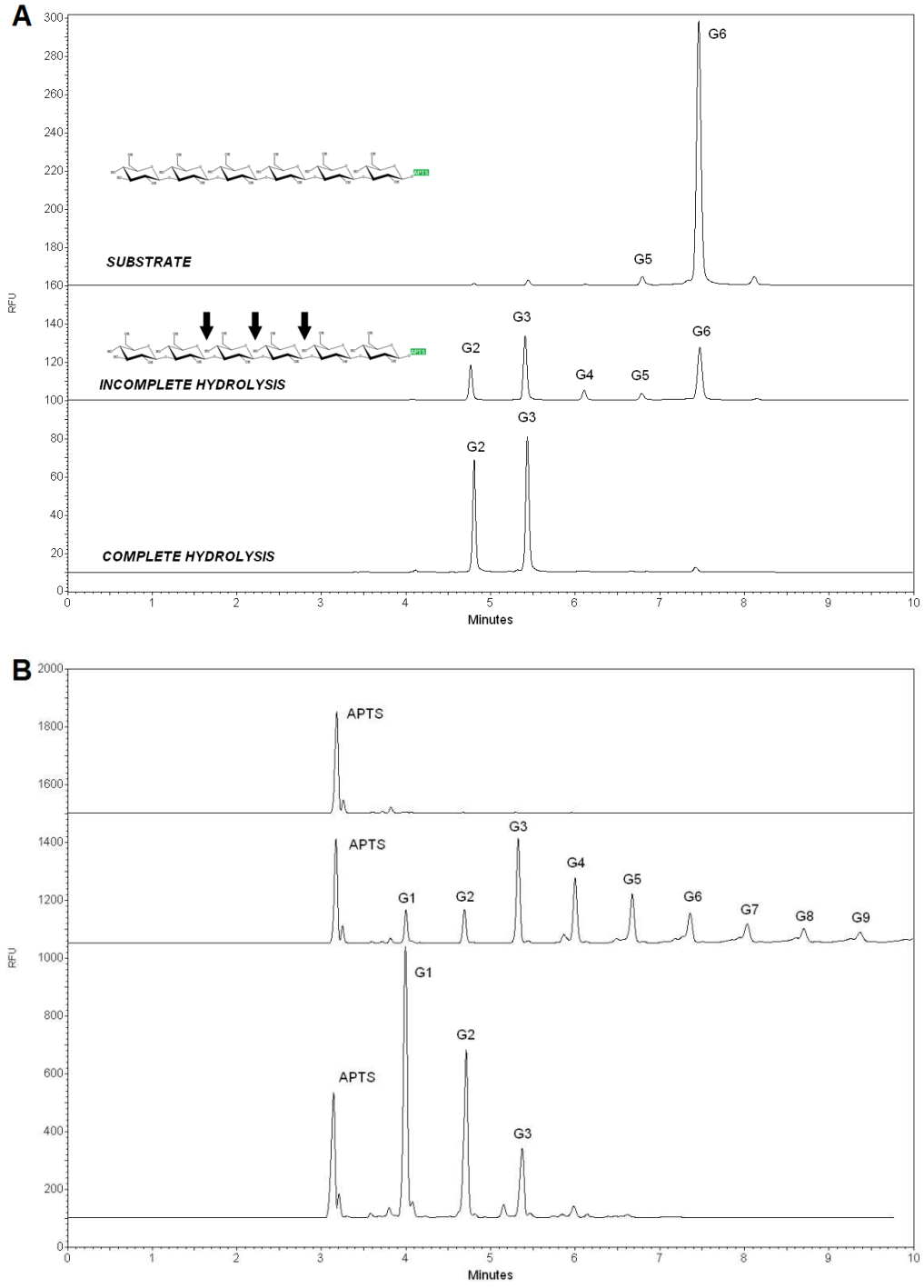


Figure 2.2. Capillary zone electrophoresis of APTS-labeled oligosaccharides. (A) Incomplete and complete hydrolysis of APTS-reducing-end-labeled-laminarihexaose. (B) Incomplete and complete hydrolysis of laminarin (APTS-labeled after the enzymatic reaction). G1,G2, G3, G4, G5, G6, G7, G8 and G9 indicate the degree of polymerization of glucose oligomers.

2.3.2 Low-resolution structure of laminarinase

The structural parameters derived from experimental curve, DAM and homology model are shown in Table 2.2. The values obtained from these different calculations are very similar, indicating a good agreement between the SAXS data, the SAXS-derived envelope and the high resolution model. Moreover, these data indicate that the maximum intraparticle distance (D_{max}) and radius of gyration (R_g) of laminarinase are 130 and 42 Å, respectively, meaning that the Dmax is 3 times greater than the Rg, and the particle should have an elongated form.

Table 2.2. Structural parameters derived from SAXS data for TpLam.

Parameters/Sample	exp[a]	DAM[b]	RBM[c]
D_{max} (Å)	130.0 ± 5.0	130.8	135.3
R_g (Å)	40.10 (Guinier) 42.0 ± 0.1 (GNOM)	41.99	42.20
Resolution (Å)	19.0	19.0	-
MW_{SAXS} (kDa)	67.9	-	-

Resolution: $2\pi/q$.

[a]Calculated from the experimental data at 4 mg/mL.

[b]Calculated from dummy atom model.

[c]Calculated from rigid body model (RBM), BUNCH program.

The scattering curve, the distance distribution function and the low resolution structure are shown in Figure 2.3A. As expected from calculations, the DAM is elongated, exhibiting a V-shape with two asymmetric arms. The three domain high resolution model was well fitted into the DAM; the catalytic domain (CD) occupying the middle and the CBMs forming the arms. The asymmetry of the molecule is due to fact that the linker between CD2 and CBM3 (residues 459 to 486) contains more residues than the linker between domains CBM1 and CD2 (residues 187 to 199). This feature permitted us unambiguously to distinguish the two CBM domains. Moreover, the linkers could act as hinges during enzymatic process, moving the CBMs in relation to the CD and conferring flexibility to the protein. This mobility is particularly evident in the CBM3 containing arm of the SAXS model, where the average of different positions resulted in a bigger volume than expected for high resolution model.

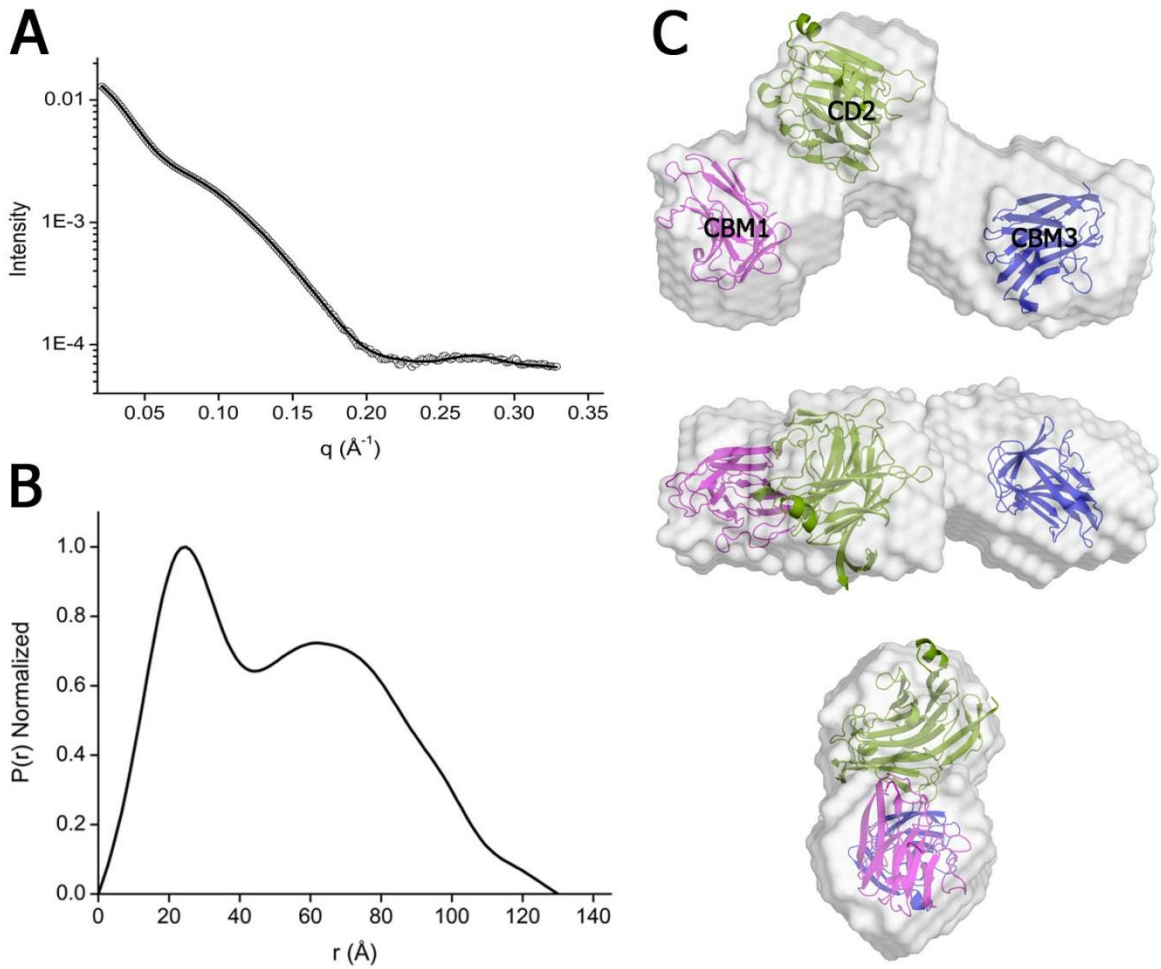


Figure 2.3. Laminarinase analysis by SAXS. (A) Experimental scattering curve (open circles) and fit produced by GNOM (solid line). (B) Distance distribution function computed from the experimental data. (C) Laminarinase model with CBM1 (magenta), CD2 (green) and CBM3 (blue) domains fitted into the envelope obtained from SAXS data is shown in different orientations.

In conclusion, we evaluated the TpLam an specific endo β-1-3 acting glucanase with remarkable thermostable properties. The study herein depicted enzymatic mode of attack, performed a comprehensive spectrometric analysis of hyperthermophilicity and determined the low-resolution structure of a multi-domain V-shape enzyme bearing two carbohydrate binding modules. Our findings in this study provide biochemical and structural basis for further studies of endo glucanases, which are important components of enzymatic repertory in polysaccharide degradation.

ACKNOWLEDGEMENTS

This research was supported by grants from FAPESP (08/58037-9) to FMS and CNPq (478059/2009-4) to MTM and scholarship from CNPq (140420/2009-6) to JC.

REFERENCES

[1] M.E. Himmel, S.Y. Ding, D.K. Johnson, et al., Biomass recalcitrance. Engineering plants and enzymes for biofuels production, Science 315 (2007) 804–807.

[2] V.V. Zverlov, I.Y. Volkov, T.V. Velikodvorskaya, W.H. Schwarz, Highly thermostable endo-1, 3-beta-glucanase (laminarinase) LamA from *Thermotoga neapolitana*: nucleotide sequence of the gene and characterization of the recombinant gene product, Microbiology 143 (1997) 1701–1708.

[3] J. Vasur, R. Kawai, E. Andersson, et al., X-ray crystal structures of *Phanerochaete chrysosporium* Laminarinase 16A in complex with products from lichenin and laminarin hydrolysis, FEBS J. 276 (2007) 3858–3869.

[4] T.R. Storseth, K. Hansen, K.I. Reitan, J. Skjermo, Carbohydr. Res. 340 (2005) 1159–1164.

[5] M. Hrmova, G.B. Fincher, Purification and properties of three (1-3)-beta-D-glucanase isoenzymes from young leaves of barley (*Hordeum vulgare*), Biochem. J. (1993) 453–461.

[6] A. Ilari, A. Fiorillo, S. Angelaccio, et al., Crystal structure of a family 16 endoglucanase from the hyperthermophile *Pyrococcus furiosus*–structural basis of substrate recognition, FEBS J. 276 (2009) 1048–1058.

[7] T.Y. Hong, Y.T. Hsiao, M. Meng, T.T. Li, The 1.5 A structure of endo-1, 3-betaglucanase from *Streptomyces sioyaensis*: evolution of the active-site structure for 1, 3-beta-glucan-binding specificity and hydrolysis Acta Crystallogr, D Biol. Crystallogr. 64 (2008) 964–970.

[8] G. Fibriansah, S. Masuda, N. Koizumi, et al., The 1.3 A crystal structure of a novel endo-beta-1, 3-glucanase of glycoside hydrolase family 16 from alkaliphilic *Nocardiopsis* sp. strain F96, Proteins 69 (2007) 683–690.

[9] O.J. Gaiser, K. Piotukh, M.N. Ponnuswamy, et al., Structural basis for the substrate specificity of a *Bacillus* 1, 3–1, 4-beta-glucanase, J. Mol. Biol. 357 (2006) 1211–1225.

[10] M. Krah, R. Misselwitz, O. Politz, et al., The laminarinase from thermophilic eubacterium *Rhodothermus marinus* conformation, Stability, and identification of active site carboxylic residues by site-directed mutagenesis, Eur. J. Biochem. 257 (1998) 101–111.

[11] F.M. Squina, C.R. Santos, D.A. Ribeiro, et al., Substrate cleavage pattern, biophysical characterization and low-resolution structure of a novel hyperthermostable arabinanase from *Thermotoga petrophila*, Biochem. Biophys. Res. Com. 399 (2010) 505–511.

[12] G.L. Miller, Use of dinitrosalicilic acid reagent for determination of reducing sugar, Anal. Chem. 31 (1959) 426–428.

[13] R.H. Myers, D.C. Montgomery, Response Surface Methodology, second ed., John Wiley & Sons, New York, 2001.

[14] C.R. Santos, F.M. Squina, A.M. Navarro, Et al., Functional and biophysical characterization of a hyperthermostable GH51 alpha-L-arabinofuranosidase from *Thermotoga petrophila*, Biotech. Lett. 33 (2010) 131–137.

[15] C.R. Santos, A.N. Meza, Z.B. Hoffmam, et al., Thermal-induced conformational changes in the product release area drive the enzymatic activity of xylanases 10B: Crystal structure, conformational stability and functional characterization of the xylanase 10B from *Thermotoga petrophila* RKU-1, Biochem. Biophys. Res. Com. 403 (2010) 214–219.

[16] A. Lobley, L. Whitmore, B.A. Wallace, DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra, Bioinformatics 18 (2002) 211–212.

[17] A. P. Hammersley, FIT2D: An Introduction and Overview. In ESRF Internal Report, (1997).

[18] A. Guinier, G. Fournet, Small-Angle Scattering of X-Rays, John Wiley and Sons, New York, 1955.

[19] H. Fischer, M. Oliveira Neto, H.B. Napolitano, et al., Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale, J. Appl. Cryst. 43 (2010) 101–109.

[20] D.I. Svergun, Determination of the regularization parameter in indirect transform methods using perceptual criteria, J. Appl. Cryst. 25 (1992) 495–503.

[21] D.I. Svergun, Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing, Biophys. J. 76 (1999) 2879–2886.

[22] J. Söding, A. Biegert, A.N. Lupas, The HHpred interactive server, Nucleic Acids Res. 33 (2005) W244–W248.

[23] P.V. Konarev, M.V. Petoukhov, D.I. Svergun, MASSHA - a graphics system for rigid-body modeling of macromolecular complexes against solution scattering data, J. Appl. Crystallogr. 34 (2001) 527–532.

[24] D.I. Svergun, C. Barberato, M.H.J. Koch, CRYSOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, J. Appl. Crystallogr. 28 (1995) 768–773.

[25] M.V. Petoukhov, D.I. Svergun, Global rigid body modeling of macromolecular complexes against small-angle scattering data, Biophys. J. 89 (2005) 1237–1250.

[26] Y. Gueguen, W.G.B. Voorhorst, J. van der Oost, W.M. de Vos, Molecular and biochemical characterization of an endo- b-1, 3-glucanase of the Hyperthermophilic archaeon *Pyrococcus furiosus*, J. Biol. Chem. 272 (1997) 31258–31264.

[27] R.W. Woody, Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins, Eur. Biophys. J. 23 (1994) 253–262.

[28] B.L. Cantarel, P.M. Coutinho, C. Rancurel, et al., The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics, Nucleic Acids Res. 37 (2009) D233–D238.

[29] F.M. Squina, A.J. Mort, S.R. Decker, et al., Xylan decomposition by *Aspergillus clavatus* endo-xylanase, Protein Expression Purif. 68 (2009) 65–71.

CAPÍTULO 3

ASSEMBLING A XYLANASE-LICHENASE CHIMERA THROUGH ALL-ATOM MOLECULAR DYNAMICS SIMULATIONS

ABSTRACT

Multifunctional enzyme engineering can improve enzyme cocktails for emerging biofuel technology. Molecular dynamics through structure-based models (SB) is an effective tool for assessing the tridimensional arrangement of chimeric enzymes as well as for inferring the functional practicability before experimental validation. This study describes the computational design of a bifunctional xylanase-lichenase chimera (XylLich) using the xynA and bglS genes from *Bacillus subtilis*. *In silico* analysis of the average solvent accessible surface area (SAS) and the root mean square fluctuation (RMSF) predicted a fully functional chimera, with minor fluctuations and variations along the polypeptide chains. Afterward, the chimeric enzyme was built by fusing the *xynA* and *bglS* genes. XylLich was evaluated through small-angle X-ray scattering (SAXS) experiments, resulting in scattering curves with a very accurate fit to the theoretical protein model. The chimera preserved the biochemical characteristics of the parental enzymes, with the exception of a slight variation in the temperature of operation and the catalytic efficiency (k_{cat}/K_m). The absence of substantial shifts in the catalytic mode of operation was also verified. Furthermore, the production of chimeric enzymes could be more profitable than producing a single enzyme separately, based on comparing the recombinant protein production yield and the hydrolytic activity achieved for XylLich with that of the parental enzymes.

Keywords: multifunctional enzymes; small-angle X-ray scattering; molecular dynamics; computational characterization; experimental validation.

3.1 INTRODUCTION

Plant biomass saccharification and biofuel production have been described as promising renewable alternatives to petroleum and natural gas. However, the polysaccharide network in plant cell walls is one of the most complex structures in nature, which jeopardizes the production of biofuels from plant biomass. First, biomass feedstock must go through a recalcitrance-reducing step (pretreatment). Then, enzymatic cocktails are used to precisely breakdown the polysaccharides into simple sugars suitable for several bioprocesses, such as fermentation to ethanol [1].

The enzymatic cocktail for plant biomass saccharification and biofuel production must include cellulolytic hydrolases, such as cellobiohydrolases, endo-glucanases and β-glucosidases. For hemicellulose degradation, synergistic action by hydrolytic enzymes is required at the polysaccharide backbone, side chains and decorating units. For instance, the hydrolysis of feedstock containing arabinoxylan requires several hydrolytic enzymes, such as endo-xylanases, β-xylosidases, arabinofuranosidases, ferulic acid esterases, glucuronidases and other enzymes [2]. β-glucans, polysaccharides with β-1,3 and β-1,4 glucosidic linkages, are also abundant in many plant cell walls, especially in sugarcane [3]. The enzymatic depolymerization of β-1,3-1,4-glucans is catalyzed by β-1,4-D-glucan 4-glucanohydrolase (EC 3.2.1.4), β-1,3-D-glucan 3-glucanohydrolase (EC 3.2.1.39) and β-1,3-1,4-D-4-glucanohydrolase or lichenase (EC 3.2.1.73) [4].

The engineering of multifunctional proteins with a synergistic catalytic capacity has the potential to streamline biomass conversion strategies [5]. However, the unsupervised construction of enzyme fusions can result in nonfunctional chimeras because of misfolding and catalytic restriction [6]. Catalytic modules connected through a linker peptide are widely used because this process allows inter-domain flexibility and usually retains the original wild-type functionality [7]. Numerous linkers have been described for protein fusion, including AAA [8], GGGG [9, 10], HHHHHH [11] and (GGGGS)4 [12].

An important issue regarding the construction of a protein chimera is how to determine the structural organization of the domains in a simple and reliable manner. Currently, small-angle X-ray scattering (SAXS) has become a central tool in structural biology for characterizing proteins in solution. Models using flexible regions and studied with normal mode analysis [13], molecular dynamics (MD) [14-16] or Monte Carlo simulations [17] have provided not only successful data validation but also accurate fitting of the scattering profile because of the potential to explore the protein conformation in space. Motivated by low computational costs, high control of the energetic

parameters and good agreement with experiments, the models based on the energy landscape theory [18] have been extensively employed in several molecular systems, including protein folding, conformational changes and dynamic molecular machines [18-24]. The combination of simulations and SAXS can also provide information about the dynamic equilibrium of proteins in solution [25], for instance, conformational changes of *holo* and *apo* protein states and their correlation to regulatory mechanisms [26].

The development of computational approaches for predicting chimeric behavior, such as the stability and arrangement of the domains in solution, can contribute to the development of automated searching pipelines for optimal linkers and enzyme modules. In this paper, we described a computational approach based on energy landscape theory for designing a bifunctional enzyme containing the endo-xylanase and lichenase catalytic modules from *Bacillus subtilis:* XynA and BglS, respectively. The properties related to the structure of the chimera, substrate accessibility to the active site and dynamical behavior in solution were first calculated through computational tools. Then, we designed the multifunctional enzyme (XylLich), which was comprehensively evaluated and validated both biochemically and structurally (SAXS). The effectiveness of the chimera in hydrolyzing beechwood xylan and lichenan polysaccharide composites was also evaluated. Furthermore, based on the recombinant protein production yield and hydrolytic activity achieved for XylLich compared with for the parental enzymes, producing chimeric enzymes could be more advantageous than producing single enzymes separately.

3.2 MATERIAL AND METHODS

3.2.1 Chimera construction for simulations

The chimera was constructed using the endo-xylanase (XynA; PDB id: 1xxn) and endo-β-1,3-1,4-glucanase (BglS; PDB id: 3o5s) domains connected by a 4-glycine linker. The chosen linker, which has been studied previously [9], is simple and guarantees reasonable separation between the monomers. The linker was modeled using the MOLMOL program [27], where the angles ϕ (Phi) and ψ (Psi) were defined as -100 degrees and 120 degrees, respectively, which is an allowed region of the Ramachandran plot. The obtained structure was solvated (water TIP-3), and 5000 steps of energy minimization using the default conjugated gradient were carried out using

NAMD 2.9 [28] with the CHARMM 2.2 force field [29]. The conformation employed for further simulations was obtained after removing the water.

3.2.2 Molecular dynamics simulations of flexible models

The SMOG web server [30] was used to generate the force field for an all-atoms model. The total energy (V) of the system for a given conformation Γ is calculated relative to the native state Γ_0 by the following equation:

$$V(\Gamma,\Gamma_0) = \sum_{bond} \varepsilon_r(r-r_0)^2 + \sum_{angle} \varepsilon_\theta(\theta-\theta_0)^2 + \sum_{planar/improper} \varepsilon_v(v-v_0)^2 + \sum_{backbone} \varepsilon_{BB}\left\{[1-\cos(\phi-\phi_0)] + \frac{1}{2}[1-\cos(3(\phi-\phi_0))]\right\} + \sum_{sidechain} \varepsilon_{SC}\left\{[1-\cos(\phi-\phi_0)] + \frac{1}{2}[1-\cos(3(\phi-\phi_0))]\right\} + \sum_{contact} \varepsilon_c\left[\left(\frac{\delta_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{\delta_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{non-contact} \varepsilon_{NC}\left(\frac{\delta_{NC}}{r_{ij}}\right)^{12}$$

The values for r_0, θ_0, v_0, ϕ_0 and δ_{ij} were taken from the conformation obtained in the previous step (*see chimera construction for simulations*) and corresponded to the distance between the covalent bonds, the angle between three consecutive connected atoms, the improper/planar dihedral angles, the dihedral angles and the pair distance between the pair of atoms i and j, respectively, for all cases in the initial state. r_{ij} is the pair distance in the conformation Γ. The energetic terms were given as a function of the contact energy ε_c=1 kT, where $\varepsilon_r=100\varepsilon_c/Å^2$, $\varepsilon_\theta=20\varepsilon_c/rad^2$, $\varepsilon_v=10\varepsilon_c/rad^2$ and $\varepsilon_{NC}=0.01\varepsilon_c$. The map of interactions between the atoms was calculated using the "Shadow algorithm" [31]. The excluded volume term was defined by non-contact pairs with δ_{NC} defined as *1.25 Å*. ε_c, ε_{BB} and ε_{SC} were adjusted as described by Whitford and collaborators [32]. When increasing the conformational sampling, no inter-domain interactions were included, indicating that the simulations were mostly entropically driven.

The simulation steps were integrated using the GROMACS software package 4 [33] at a temperature slightly lower than the folding temperature and coupled by a thermal bath controlled by Langevin dynamics for a total of *2,000,000,000* steps using a time step=*0.0005* for a total of *100 ns*.

The free energy profile was calculated using the simple histogram method [34] for the following reaction coordinates: radius of gyration (R_g) and distance from the center of mass of the XynA to the center of mass of the BglS in the chimera (CM distance Xyl-Lich), both of which were calculated using GROMACS analysis tools [33]. The conformations with lower free energy values were considered the computational suggestions for describing the arrangement of the domains in solution.

3.2.3 Assessment of the enzymatic functionality using simulations

All-atom SB were employed to simulate $5x10^8$ steps using the energetic parameters previously defined and without modifications to the topological files generated by the SMOG web server (the contact map included inter-domain interactions). To evaluate possible unexpected features, the root mean square fluctuation (RMSD) and average solvent accessible surface area (SAS) parameters per residue were calculated using GROMACS tools [33].

3.2.4 Assembly and protein expression of the XynA-BglS chimera (XylLich)

Genomic DNA from *Bacillus subtilis* 168 was used as the template for PCR amplification of the *xynA* (NCBI-GI: 16078944) and *bglS* (NCBI-GI: 16080958) genes. The forward and reverse primers were 5'-tatatagctagcagcacagactactggcaaaa-3' and 5'-tatataggatcccacactgttacgttagaac-3', respectively, for xynA, and the forward and reverse primers were 5'-tatatagctagccaaacaggtggatcgttttt-3' and 5'-tatataggatccttatttttttgtatagcgca-3', respectively, for *bglS*. The restriction sites are underlined in the primer sequences. Two different primers were tailor designed to fuse the genes (*xynA* reverse 5'-accaccaccaccccacactgttacgttagaac-3' and *bglS* forward 5'-ggtggtggtggtcaaacaggtggatcgttttt-3') and included a four-glycine residues linker [9] (underlined nucleotides). The amplified fragments were used as templates for the overlap extension PCR technique [35] to fuse the two genes in a unique ORF. The chimeric fragment was digested with the restriction enzymes *Nhe*I and *BamH*I and cloned in pET28a(+) (Novagen).

All plasmid constructions, pET-XynA, pET-BglS and pET-XylLich, were produced in BL21 DE3 cells. The protein expression was induced with a final concentration of 0.5 mM IPTG in 0.5 L LB medium for 5 hours with shaking at 200 rpm and 37 °C. The culture was harvested at the end of fermentation and resuspended in cell lysis buffer consisting of 20 mM phosphate buffer at pH 7.4

with 5 mM imidazole, 1 mM PMSF and 0.5 mg/mL lysozyme. Then, the cells were disrupted by sonication. Two protein purification steps, including Ni^{2+}-chelating affinity and size-exclusion chromatography, were performed according to Squina et al. [36]. The purified proteins were further analyzed using SDS-PAGE, and the protein concentration was determined from the absorbance at 280 nm.

3.2.5 SAXS data collection and validation of the predicted conformations in solution

Small angle X-ray scattering (SAXS) data for XylLich were collected on the SAXS2 beam line at the Brazilian Synchrotron Light Laboratory at the concentrations of 1 and 2 mg/mL. The radiation wavelength was set to 1.48 Å, and a 165 mm MarCCD detector was used to record the scattering patterns. The sample-to-detector distance was set to 1022.5 mm to give a range of the scattering vector q from *0.013* to *0.33* Å-1, where q is the magnitude of the q-vector defined by $q = 4\pi \sin\theta / \lambda$ (2θ is the scattering angle). Protein samples were prepared in a buffer with 20 mM phosphate and 50 mM NaCl at pH 7.4. The SAXS patterns were integrated using Fit2D software [37], and the curves were scaled by the protein concentration. The molecular weight was calculated using the procedure implemented in the web tool SAXSmoW [38].

CRYSOL 2.7 [39] was used to generate theoretical scattering curves and to compare the experimental and theoretical data. The parameter for goodness of fitting, X^2, was defined as follows:

$$\chi^2 = \frac{1}{N_q}\left[\frac{I^{experimental}(q) - I^{theoretical}(q)}{\varsigma(q)}\right]^2$$

where N_q is the total number of experimental points, $I^{experimental}(q)$ *and* $I^{theoretical}(q)$ are the experimental and theoretical intensities of the scattering vector q, respectively, and $\varsigma(q)$ is the standard deviation of the experimental intensity values.

3.2.6 Assessing the enzymatic properties

All the assays were carried out with an automated pipetting system: epMotion® 5075 (Eppendorf). Reducing sugars were determined using the 3,5-dinitrosalicylic acid (DNS) method and monitored colorimetrically at 540 nm [40] using an Infinite® 200 PRO microplate reader (TECAN). One unit of enzyme was defined as the quantity of enzyme necessary to release reducing

sugars at rate of 1 µmol per minute under standard conditions. The standard assay was conducted for 10 minutes in 40 mM McIlvaine's buffer with glycine at pH 6, 50 °C. The final concentrations were 2.5 mg/mL for the substrates and 1 µM for the enzymes.

The optimal pH and temperature were determined for XynA and BglS using beechwood xylan and lichenan as substrates, respectively, under standard conditions. The enzymatic activity rates of the chimera and independent modules were evaluated with other substrates, including sugar beet, debranched arabinan, linear arabinan, rye arabinoxylan, larch arabinogalactan, galactomannan, xyloglucan, oat spelt xylan, wheat arabinoxylan and xylan beechwood. The polysaccharides were purchased from Sigma Aldrich or Megazyme.

Xylohexaose and the oligosaccharides derived from lichenan (Megazyme) were derivatized with 8-aminopyrene-1,3,6-trisulfonic acid (APTS) by reductive amination [41]. Thus, the enzymatic hydrolysis of these labeled substrates was performed at 50 °C. Capillary zone electrophoresis (CZE) was performed in a P/ACE MQD instrument (Beckman Coulter) equipped with a laser-induced fluorescence detector, as previously described [42]. A fused-silica capillary (TSP 050375, Polymicro Technologies) with an internal diameter of 50 µm and total length of 31 cm was used as a separation column for the oligosaccharides. The electrophoresis conditions were 15 kV/70–100 µA at a controlled temperature of 20 °C using sodium phosphate buffer (40 mM, pH 2.5).

The kinetic parameters were estimated for all enzymes from the initial rates using twelve substrates concentrations in the range of 1–30 mg/mL. The assays were carried out under standard conditions, using beechwood xylan and lichenan as substrates to assess the V_{max}, K_m and k_{cat} of the activity of xylanase and lichenase, respectively, using the Hill method.

Far-UV circular dichroism (CD) spectra were taken on a JASCO J-810 spectropolarimeter (Jasco Inc., Tokyo, Japan.) equipped with a Peltier temperature control unit using a wavelength range of 195-250 nm, a 0.1 cm path quartz cuvette. The solvent spectra were subtracted in all experiments to eliminate background effects. The CD spectra were the average of 8 accumulations using a scanning speed of 100 nm min^{-1}, spectral bandwidth of 1 nm, and response time of 0.5 s. The protein concentration was 0,2 mg/mL in 50 mM sodium phosphate buffer at pH 7.4. The thermal denaturation of the xylanase-lichenase chimera was characterized by measuring the ellipticity changes at 218,5 nm induced by a temperature increase from 20 °C to 100 °C at a heating rate of 1 °C min^{-1} [43].

The biotechnological potential of the chimeric enzyme was evaluated under standard conditions using a polysaccharide composite prepared with beechwood xylan and lichenan in a 1:1

ratio (10 mg/mL). The conversion rates were assessed for XynA, BglS, XynA plus BglS mixture and XylLich at the same molar concentration (1 μM). The amount of reducing sugar was measured as previously described.

To compare the hydrolytic activity and recombinant protein production yield of the chimera and the parental enzymes, the enzymes were produced at the same scale as previously described. The total produced cell mass was weighed. Afterward, the crude enzyme extract was prepared through cell lysis, as previous described. The enzymatic activities were measured under standard conditions using the crude preparations of XynA, BglS and XylLich. The total protein content was also determined for all the samples using a commercial Bradford kit from Bio-Rad®. Finally, to correlate the overall activity and the total recombinant protein production yield, a parameter dubbed activity yield (AY) was generated by dividing the total enzyme units (U) by the total protein content (mg) and cell mass that was produced (g).

3.3 RESULTS

3.3.1 Simulations suggested a unique ensemble of structures

Molecular dynamics simulations were performed to characterize the most probable conformation of the chimera in solution. To assess probable arrangements of the chimera in solution, 2,000,000 conformations were generated through simulations using SB models and evaluated. Inter-region contacts (among XynA, BglS and the linker) were removed to allow an extensive search of the configuration space. Figure 3.1A illustrates the universe of conformational possibilities derived from these simulations. The free energy profile was calculated as a function of the following reaction coordinates: CM distance, R_g and total energy of the system. An ensemble of candidate structures covered the lowest free-energy region, within which an accurate chimera arrangement was also expected. As shown in Figure 1, a large number of possible configurational states were sampled within a distinctive basin of low-energy conformations, which was restricted to the R_g region near 26 Å with a CM close to 42 Å, which represented the accommodated chimera in solution. These results were obtained without any experimental support.

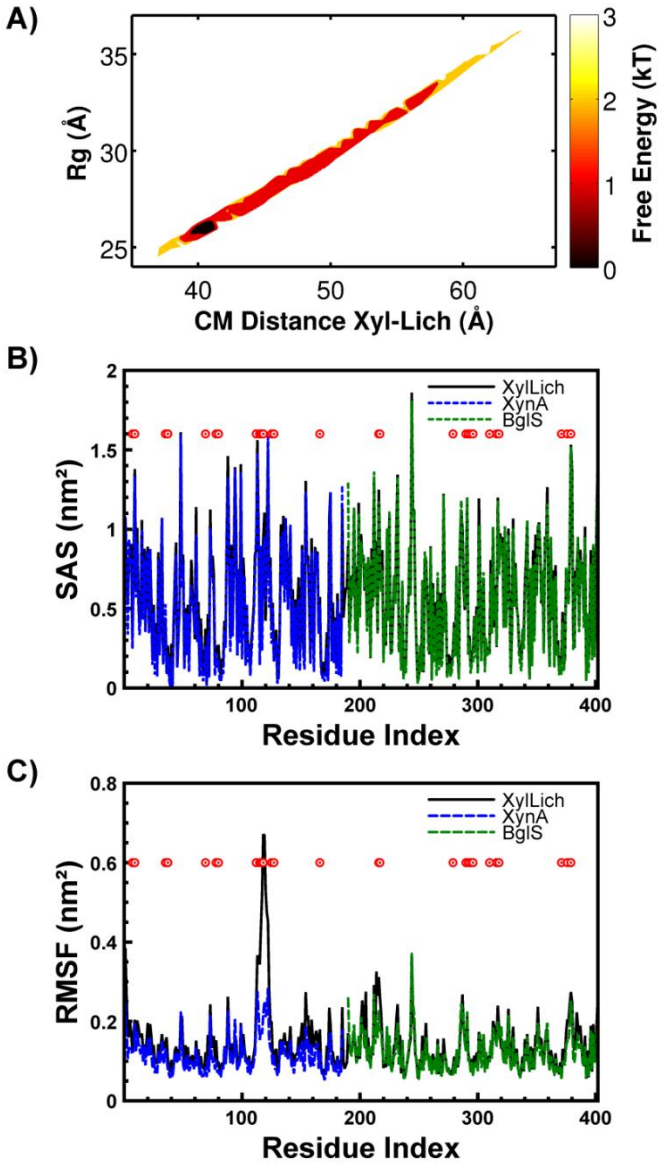


Figure 3.1. Characterization of XylLich using structure-based models (SB) through molecular dynamics (MD) simulations. **(A)** The figure on the top shows the free energy profile as a function of the radius of gyration (R_g) and the distance between the centers of mass of the XynA and BglS domains (CM). Free energy is given in units of kT and presented as a color scale. The figure was obtained using extensive simulation (100 ns) and no intra-domain interactions (see methods). **(B)** Average solvent accessible surface area (SAS) per residue, calculated from structure-based model simulations and **(C)** root mean square fluctuation (RMSF) per residue are presented. Figures **A** and **B** show the chimera (XylLich) in black, XynA only in blue, BglS only in green and the residues related to the catalytic site as red dots.

Conformations outside the free-energy basin revealed disagreements (high χ^2 values), which were caused by incorrect protein arrangements (Fig SI-3.2). The compacted conformations showed

deformations because of the high repulsive force between the residues. The extended conformations presented local unfolding and distortions near the linker. In both cases, the protein accommodation in solution was not quenched.

Simulations using XynA, BglS and the computationally suggested chimera were performed to predict possible modifications in their functional behavior. One conformation from the basin was chosen for a new simulation (R_g=26.7Å, *CM* distance=42.4Å). Simulations using the chimera and the enzyme domains separately (XynA and BglS) displayed minor variations in the SAS (Fig. 3.1B), suggesting that the accessibility of the substrate to the pocket was not affected in the designed chimera. RMSF analysis (Fig. 3.1C) also showed minor variations, except for the XynA residues SER:117 and ILE:118, which were located into the "thumb" region of the protein [44, 45].

3.3.2 Validation of the computationally predicted conformations through SAXS experiments

The genes *xynA* and *bglS* were fused by PCR-mediated overlap extension, and four glycine residues were included as a linker (Fig. 3.2A). The resulting amplicon, which was cloned into a pET28a vector, was 1209 bp long (Fig. 3.2A). Restriction analysis and DNA sequencing confirmed the molecular cloning. The chimera and its parental enzymes, XynA (22.8 kDa) and BglS (26.7 kDa), were over expressed in *E. coli* and purified through two chromatographic steps: first, nickel-affinity and then size-exclusion chromatography. The apparent molecular weight of the xylanase-lichenase chimera was 47.2 kDa including the 6-His tag (Fig. 3.2B).

Because the candidate conformations were obtained without previous experimental information, SAXS was employed to validate the computationally predicted structures. The superimposed theoretical and experimental scattering curves of the selected conformation and the XylLich model are presented in Figure 3.2C. Theoretical scattering curves (*I(q)* versus *q*) within the broad basin of low-energy conformations were compared to the experimental scattering curve using CRYSOL. Good agreements were obtained for these conformations, i.e., low χ^2 values, which were equal to 2.9 on average (some examples are presented in Fig. SI-3.1A and Fig. SI-3.1B), validating our simulation strategy. The obtained conformations displayed no obstructions or deformations at the catalytic region, indicating that the chimera would be fully functional *in vitro.*

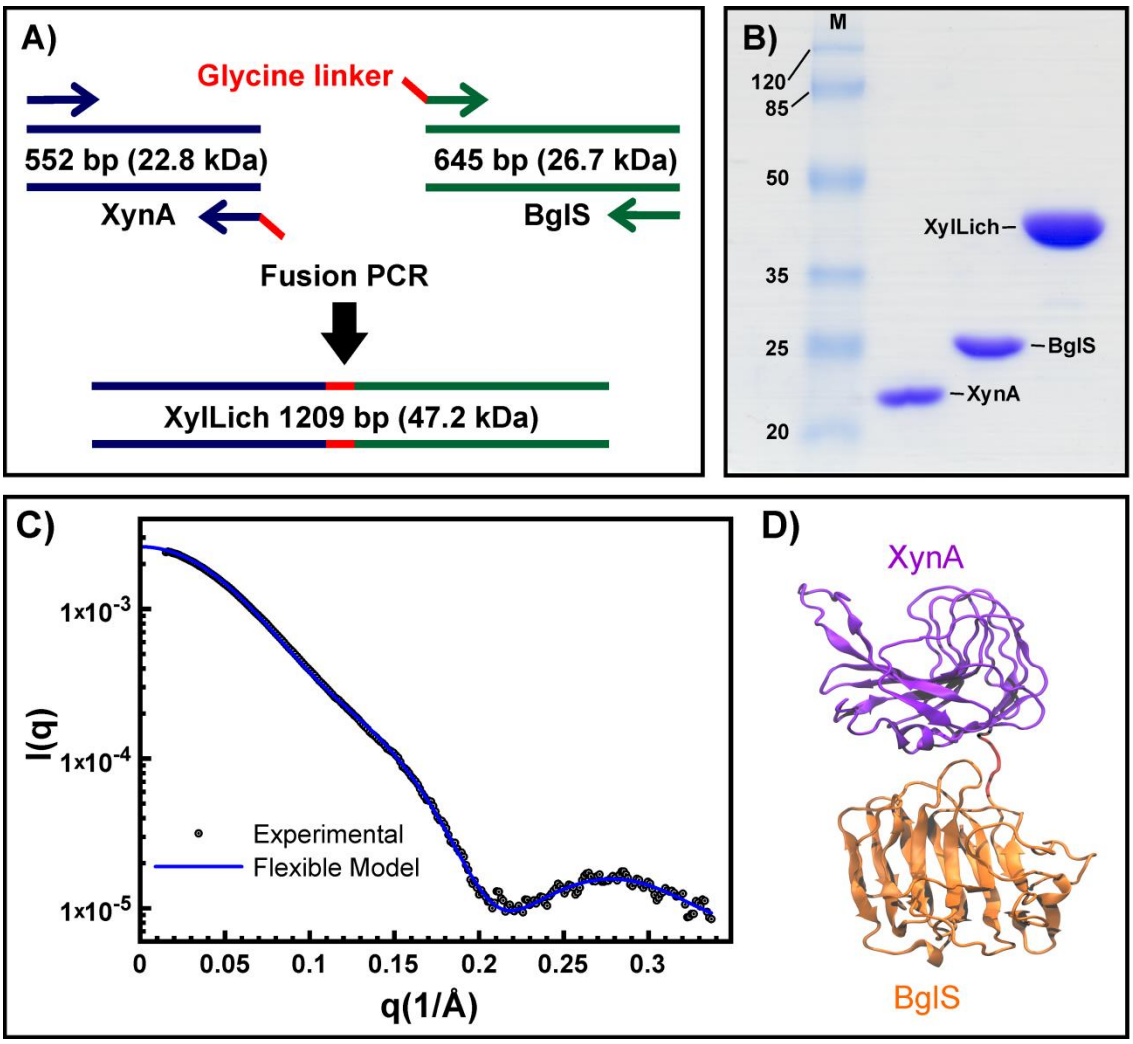


Figure 3.2. The chimera construction and experimental validation by SAXS. **(A)** Flow chart showing the process of fusion PCR for the XylLich construction. **(B)** SDS-PAGE of the chimera and wild-type proteins, indicating that the fused enzyme Xylanase-Lichenase had the predicted molecular weight. The protein molecular weight marker is shown in the first lane (M), and the values are displayed in kDa. **C)** The small angle X-ray scattering profile for the experimental and theoretical evaluation of XylLich, which was taken from the free energy basin with χ^2=2.80, R_g=26.0Å and *CM* distance=40.7Å. The scattering intensity is shown on a logarithm scale as a function of the momentum transfer (q). **(D)** The XylLich model comprising XynA (purple), the linker (red) and BglS (orange). CRYSOL was employed to generate the theoretical curve and VMD [55] for the denoted cartoon.

3.3.3 The chimera maintained the functional characteristics of the parental enzymes

The optimal temperature and pH of operation did not change drastically between the chimeric enzyme and the parental enzymes. The optimal pH for XynA and BglS was just 0.5 units lower than that of XylLich (Fig. 3.3A and 3.3B). There was no statistically significant shift in the xylanase temperature-dependent activity for the chimera, whereas the optimum temperature for lichenan hydrolysis was slightly greater for XylLich than for the parental enzymes (Fig. 3.3C and 3.3D).

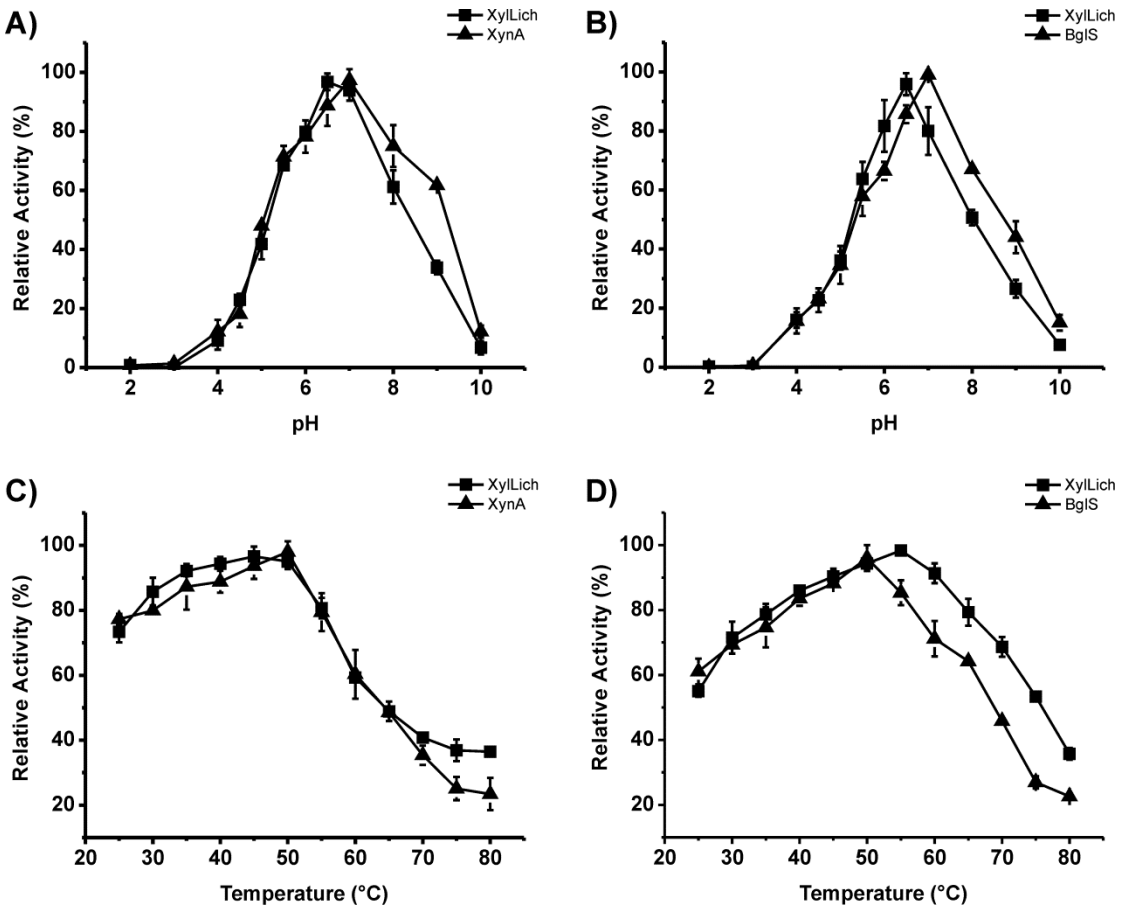


Figure 3.3. The effects of pH and temperature on the XylLich catalytic activity. pH (A and B) and temperature (C and D) curves for the chimera (■) and the parental enzymes (▲). The influence of pH on the enzymatic activity of XylLich compared to that of XynA (A) and BglS (B) using beechwood xylan (A) and lichenan (B) as substrates. The effect of temperature on the enzymatic activity of the chimeric enzyme compared to that of xylanase (C) and lichenase (D) using beechwood xylan and lichenan as substrates, respectively, is represented.

The enzymes were biochemically assessed using a set of 10 natural polysaccharides. The specific activities of the chimera and the parental enzymes on these polysaccharides are summarized in Table 3.1. XynA and XylLich hydrolyzed birchwood xylan, rye arabinoxylan and beechwood xylan more efficiently than other polysaccharides. Taking into account the standard deviation, there was no statistically significant difference between XynA and XylLich in hydrolyzing birchwood xylan and rye arabinoxylan (see Tables SI-3.1 and SI-3.2 in supplementary information section). XylA and XylLich presented the lowest xylan-degrading activity (29-36%) on wheat arabinoxylan, which was the most complex and insoluble substrate tested. XylLich and BglS hydrolyzed lichenan and β-glucan equally. The relative activities of XylLich and BglS on lichenan were 72 and 75% lower, respectively, on β-glucan.

Table 3.1. Specific activities of the parental and chimeric enzymes on different types of substrates.

	Specific activity (U/nmol)		
Substrate	**XynA**	**BglS**	**XylLich**
Birchwood Xylan	3.73 ± 0.29	ND	2.71 ± 0.13
Beechwood Xylan	3.17 ± 0.07	ND	2.87 ± 0.08
Rye Arabinoxylan	3.73 ± 0.14	ND	3.03 ± 0.15
Wheat Arabinoxylan	1.36 ± 0.12	ND	0.88 ± 0.07
Oat Spelt Xylan	3.28 ± 0.27	ND	2.15 ± 0.06
Lichenan	ND	3.65 ± 0.29	3.85 ± 0.16
β-Glucan	ND	5.03 ± 0.20	5.11 ± 0.07

ND means not determined. There was no activity on laminarin, xyloglucan and glucomannan (konjac) for all enzymes. Values are given by the mean ± S.D. of three independent assays.

Figure 3.4 describes our attempt to evaluate the mode of action of XylLich, XynA and BglS through CZE analysis of APTS-labeled oligosaccharides. Both parental and chimeric xylanase produced the same degradation pattern for APTS-labeled xylohexaose (X6), which included xylotetraose (X4), xylotriose (X3) and xylobiose (X2) (Fig. 3.4A). BglS was not able to hydrolyze X6. After hydrolysis of lichenan, XylLich and BglS produced glucotetraose (G4), and XynA was not able to break down lichenan (Fig. 3.4B). As shown in Figure 3.4C, the enzymes were also assayed against the two substrates (X6 and lichenan) simultaneously. The results confirmed our

previous findings from using the substrates separately (Fig. 3.4A and 3.4B), highlighting that the mode of operation of the chimera was exactly similar to that of the parental enzymes.

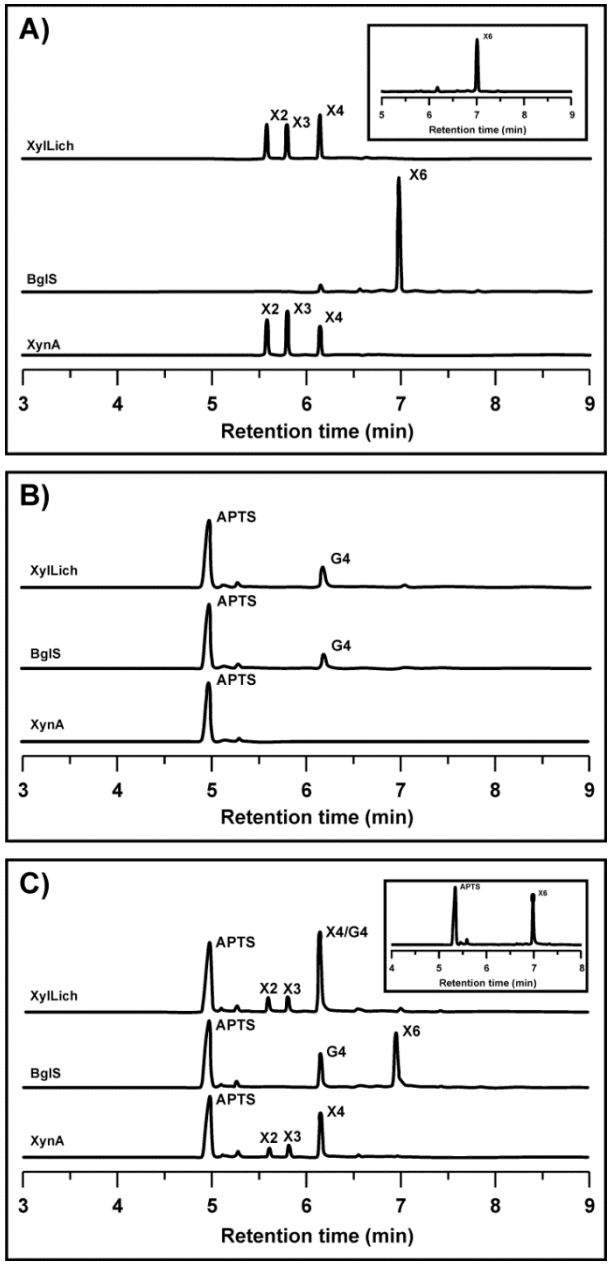


Figure 3.4. Capillary zone electrophoresis analysis of the breakdown products released by XylLich, XynA and BglS. The products after enzymatic hydrolysis of APTS-labeled xylohexaose (A), lichenan (B) and xylohexaose plus lichenan together (C) are presented. X2, X3, X4, X6 and G4

indicate the degree of polymerization of the produced xylose and glucose oligomers. The APTS-labeled xylohexaose used in the assays is indicated in the upper right boxes (A and C).

Saturation assays were performed using beechwood xylan and lichenan. The kinetic constants of the wild-type and chimeric enzymes are compared in Table 3.2. The maximum xylan degradation rate (V_{max}) of the chimera was 30% lower than that of the parental enzyme (XynA), but for lichenan hydrolysis, the V_{max} was 33% greater for XylLich than for BglS. Using xylan as a substrate, XylLich had a lower K_m (41%) and turnover number (30%) than XynA. In contrast, the K_m of the chimeric enzyme was approximately 55% higher than that of BglS. In addition, the k_{cat} increased by up to 28% for XylLich using lichenan as substrate. Moreover, the catalytic efficiency (k_{cat}/K_m) of the chimera was greater (18%) for xylan degradation and lower (18%) for lichenan hydrolysis than for the parental enzymes.

Table 3.2. Kinetic parameters of the chimeric and parental enzymes.

Substrate	***Beechwood Xylan***				***Lichenan***			
	V_{max} *µmol/min/µmol*	K_m *mg/mL*	k_{cat} s^{-1}	k_{cat}/K_m $mL.mg^{-1}.s^{-1}$	V_{max} *µmol/min/µmol*	K_m *mg/mL*	k_{cat} s^{-1}	k_{cat}/K_m $mL.mg^{-1}.s^{-1}$
XynA	10420 ± 497	7.87 ± 0.71	173.7 ± 8.3	22.1 ± 1.1	ND	ND	ND	ND
BglS	ND	ND	ND	ND	11430 ± 407	3.39 ± 0.26	197.2 ± 8.5	58.1 ± 2.5
XylLich	7300 ± 212	4.67 ± 0.26	121.7 ± 3.5	26.1 ± 0.8	15190 ± 1987	5.27 ± 1.13	253.2 ± 33.1	48.1 ± 6.3

ND means not determined. Values are given by the mean ± S.D. of three independent assays. The *Vmax* unit was defined as µmol of reducing sugars released per minute per µmol of enzyme.

The far-UV CD spectra of XylLich showed a negative peak near 218 nm, suggesting that XylLich consists primarily of β-sheets (Fig. 3.5A). This profile was expected because the wild-type enzymes have β-sheet predominance [46-48]. Another band was detected in the 220-230 nm region of the BglS spectrum, indicating the contribution of side-chain aromatic residues [47]. Thermal denaturation experiments were performed to compare the stability of the chimera and wild-type enzymes. The XylLich melting temperature (Tm) was 43.6°C, whereas the Tm values of the parental enzymes were 53.3°C and 42.5°C for XynA and BglS, respectively (Fig. 3.5B).

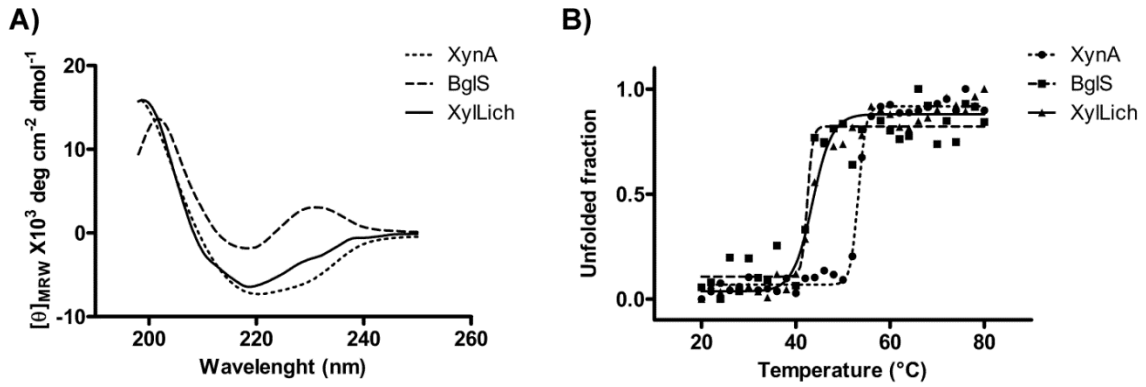


Figure 3.5. Secondary structure evaluation by circular dichroism. **(A)** Far-UV CD spectra of XylLich and individual enzymes at pH 7.4 and 20°C and **(B)** thermal denaturation curve at pH 7.4. The circular dichroism spectrum represents an average of eight scans. The thermal denaturation curve was obtained by monitoring at 220 nm for xylanase, 218 nm for lichenase and 218.5 nm for the chimera.

3.3.4 Biotechnological appeal for producing the chimeric enzyme

To evaluate the biotechnological potential of using the chimeric enzyme in biomass to bio-products applications, hydrolysis assays were performed using a polysaccharide composite comprising beechwood xylan and lichenan in a 1:1 ratio. The conversion efficiencies of *XynA*, *BglS*, an equimolar mixture of the parental enzymes and the chimeric enzyme on this polysaccharide composite are described in Figure 3.6A. The rate of substrate conversion for XylLich was statistically equal to that of the parental enzymes combined (Fig. 3.6A).

The enzymatic performance was also evaluated using the crude *E. coli* cell lysate as an enzyme source. Based on the total protein content in the cell lysate and the overall mass produced by *E. coli* cells, the amount of enzyme required to release 1 µmol of reducing sugar per minute per protein milligram per gram of cell (U/mg/g) was calculated, which was defined in AY units (see methods). The chimera AY (0.35 U/mg/g) was approximately 6 times higher than XynA AY (0.06 U/mg/g) using xylan as a substrate. Likewise, the AY was 20% greater for XylLich than for BglS (Fig. 3.6B) using lichenan as a substrate.

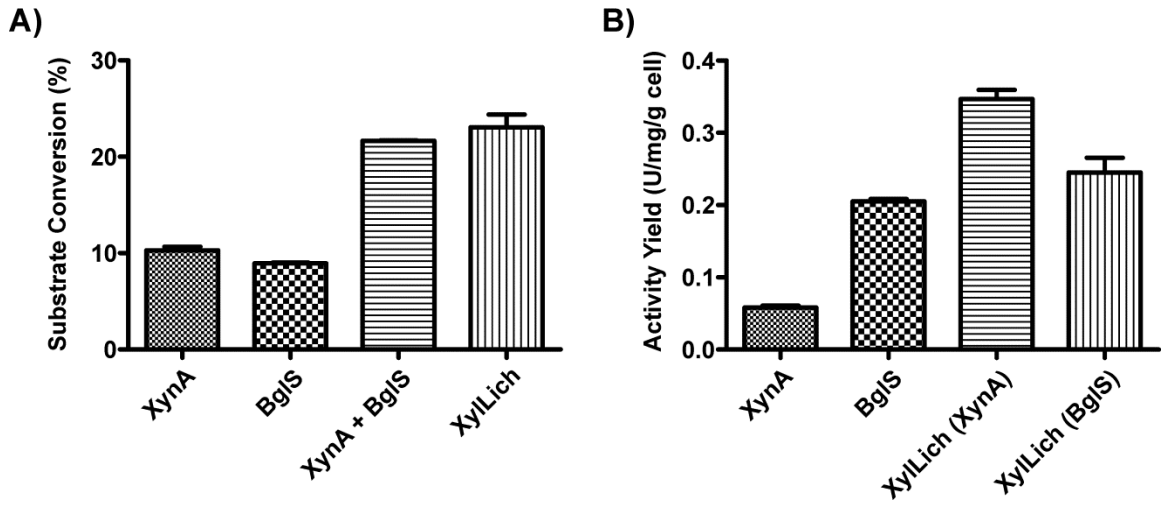


Figure 3.6. Biotechnological appealing for producing the chimeric enzyme**. (A)** Analysis of the conversion efficiency of a composite consisting of beechwood xylan and lichenan by the parental enzymes, the mixture of the parental enzymes (XynA + BglS) and XylLich. **(B)** Assessment of the activity yield (AY) using the crude *E. coli* cell extract as an enzyme source.

3.4 DISCUSSION

This study is an initial step toward the development of automated screening pipelines for optimal linkers and modules and functional chimeras before experimental validation. XynA and BglS from *B. subtilis* were chosen because these enzymes have been fully characterized and crystallographic structures are available [44-46, 48-50]. Accordingly, the small and simple linker comprising four glycine residues was very well described previously [9]. However, based on the number of conformations available to XylLich (Fig. 3.1A, SI-3.1 and SI-3.2), its predicted organization in solution was not obvious without computational tools or SAXS.

Simulations employing SB models require significantly less computational time than traditional molecular dynamics using explicit water or complex force fields. This key benefit allows the straightforward investigation of several constructions and larger systems. Despite being minimalistic approaches, SB models take into account the atomic restrictions to which real proteins are susceptible and thus corroborate experimental results [22]. The solvent properties implicitly included in these models support the calculation of an SAS, which predicts the change of the substrate accessibility to the enzyme [51] and estimates alterations in the enzymatic functionality. Simulations employing the chimeric model ($\chi 2$=2.8) and the parental enzymes (XynA or BglS) presented minor changes in the SAS (Fig. 3.1B). These results suggested that variations in the

operation mode would not be expected because the binding site was not obstructed by deformations or steric effects between the domains. Supporting our hypothesis, the RMSF analysis (Fig. 3.1C) showed a similar profile for all cases. The only exception was the residue ILE:118 of the XynA thumb region located near the catalytic binding site, which could influence the enzymatic catalysis [44, 45].

To validate the molecular dynamics simulations, comprehensive biochemical characterization was performed with the parental enzymes and the chimera. Collectively, the biochemical analysis confirmed the simulations. The main exception was the slight variation of the XylLich temperature of operation and catalytic efficiency (k_{cat}/K_m). There were no substantial shifts in either the substrate specificities or the mode of operation. In addition, circular dichroism analysis revealed a greater similarity to the thermal denaturation properties of XylLich to BglS than those of XynA.

The minor shifts in the optimal pH and temperature of operation can be explained by an alteration in the microenvironment of the fused enzyme; one enzyme module preferentially catches protons and acidifies the local pH [52]. In addition, protein fusion can disturb the tertiary structure of a chimeric enzyme [52], which can cause displacements in the pH and temperature of operation, phenomena that are very well described in the literature [5, 50, 52].

The engineering of fused proteins has long been considered attractive for industrial processes because these proteins are more efficient (based on costs and catalytic efficiency) than wild-type proteins [5, 49, 50, 53, 54]. Along with the lack of substantial changes in the catalytic performance of the chimera, the parameter AY suggested an advantage in producing the fused protein rather than the separate wild-type ones. Certainly more studies on XylLich properties are required before use in a scaled-up process.

In conclusion, this work presented a novel approach for predicting the arrangement of chimeric domains in solution before experimental validation. The computational strategy proposed herein is fast and robust for characterizing a large number of enzymes in a few weeks, as well as for identifying possible binding site obstructions or large dynamical variations differing from the single domains. The methodology can be extended to multi-domain chimeras and, by the same token, to a wide variety of biological systems. Finally, we expect that our findings will increase the pace of finding novel and cost-effective approaches for converting plant biomass into bio-products. One of the great challenges in biomass saccharification is decreasing the cost of enzyme production. Therefore, the use of multifunctional hydrolases, which act synergistically in plant polysaccharide

degradation, is a promising venue for the improvement of enzyme cocktails for second-generation biofuels.

ACKNOWLEDGEMENTS

This research was supported by grants from FAPESP (2008/58037-9) and CNPq (475022/2011-4 and 310177/2011-1). JC received a scholarship from CNPq (140420/2009-6), and LCO, TMA and ARLD received scholarships from FAPESP (2011/13424-7, 2010/11499-1 and 2011/02169-7, respectively). X-ray scattering data were collected at the Brazilian Synchrotron Light Laboratory (LNLS) that integrates the Center of Research in Energy and Material (CNPEM). The authors would like to thank the LNLS for support. The computational analyses were supported by resources supplied by the Center for Scientific Computing (NCC/Grid UNESP) of São Paulo State University (UNESP) and CENAPAD-SP (Centro Nacional de Processamento de Alto Desempenho em São Paulo), project UNICAMP/FINEP-MCT**.**

REFERENCES

[1] M.E. Himmel, S.-Y. Ding, D.K. Johnson, W.S. Adney, M.R. Nimlos, J.W. Brady, T.D. Foust, Biomass recalcitrance: engineering plants and enzymes for biofuels production, Science (New York, N.Y.), 315 (2007) 804-807.

[2] T. Collins, C. Gerday, G. Feller, Xylanases, xylanase families and extremophilic xylanases, FEMS microbiology reviews, 29 (2005) 3-23.

[3] D.U. Lima, H.P. Santos, M.a. Tiné, F.R.D. Molle, M.S. Buckeridge, Patterns of expression of cell wall related genes in sugarcane, Genetics and Molecular Biology, 24 (2001) 191-198.

[4] A. Planas, Bacterial 1,3-1,4-beta-glucanases: structure, function and protein engineering, Biochimica et Biophysica Acta, 1543 (2000) 361-382.

[5] Z. Fan, K. Wagschal, C.C. Lee, Q. Kong, K.a. Shen, I.B. Maiti, L. Yuan, The construction and characterization of two xylan-degrading chimeric enzymes, Biotechnology and Bioengineering, 102 (2009) 684-692.

[6] S.Y. Hong, J.S. Lee, K.M. Cho, R.K. Math, Y.H. Kim, S.J. Hong, Y.U. Cho, S.J. Cho, H. Kim, H.D. Yun, Construction of the bifunctional enzyme cellulase-beta-glucosidase from the hyperthermophilic bacterium *Thermotoga maritima*, Biotechnology Letters, 29 (2007) 931-936.

[7] C.J. Crasto, J.A. Feng, LINKER: a program to generate linker sequences for fusion proteins, Protein Engineering, 13 (2000) 309-312.

[8] J.L. Casey, A.M. Coley, L.M. Tilley, M. Foley, Green fluorescent antibodies: novel in vitro tools, Protein Engineering, 13 (2000) 445-452.

[9] R. Arai, H. Ueda, a. Kitayama, N. Kamiya, T. Nagamune, Design of the linkers which effectively separate domains of a bifunctional fusion protein, Protein Engineering, 14 (2001) 529-532.

[10] R.A. Griep, C. van Twisk, J.M. van der Wolf, A. Schots, Fluobodies: green fluorescent single-chain Fv fusion proteins, Journal of Immunological Methods, 230 (1999) 121-130.

[11] K. Morino, H. Katsumi, Y. Akahori, Y. Iba, M. Shinohara, Y. Ukai, Y. Kohara, Y. Kurosawa, Antibody fusions with fluorescent proteins: a versatile reagent for profiling protein expression, Journal of Immunological Methods, 257 (2001) 175-184.

[12] M. Peipp, D. Saul, K. Barbin, J. Bruenke, S.J. Zunino, M. Niederweis, G.H. Fey, Efficient eukaryotic expression of fluorescent scFv fusion proteins directed against CD antigens for FACS applications, Journal of Immunological Methods, 285 (2004) 265-280.

[13] O. Miyashita, C. Gorba, F. Tama, Structure modeling from small angle X-ray scattering data with elastic network normal mode analysis, J. Struct. Biol., 173 (2011) 451-460.

[14] M. Pelikan, G.L. Hura, M. Hammel, Structure and flexibility within proteins as identified through small angle X-ray scattering, General Physiology and Biophysics, 28 (2009) 174-189.

[15] S. Yang, L. Blachowicz, L. Makowski, B. Roux, Multidomain assembled states of Hck tyrosine kinase in solution, Proceedings of the National Academy of Sciences of the United States of America, 107 (2010) 15757-15762.

[16] S. Yang, S. Park, L. Makowski, B. Roux, A rapid coarse residue-based computational method for x-ray solution scattering characterization of protein folds and multiple conformational states of large protein complexes, Biophysical Journal, 96 (2009) 4449-4463.

[17] S.J. Kim, C. Dumont, M. Gruebele, Simulation-based fitting of protein-protein interaction potentials to SAXS experiments, Biophysical Journal, 94 (2008) 4924-4931.

[18] J.N. Onuchic, P.G. Wolynes, Theory of protein folding, Current Opinion in Structural Biology, 14 (2004) 70-75.

[19] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Funnels, pathways, and the energy landscape of protein folding: a synthesis, Proteins, 21 (1995) 167-195.
[20] L.L. Chavez, J.N. Onuchic, C. Clementi, Quantifying the roughness on the free energy landscape: entropic bottlenecks and protein folding rates, Journal of American Chemical Society, 126 (2004) 8426-8432.
[21] C. Clementi, Coarse-grained models of protein folding: toy models or predictive tools?, Curr. Opin. Struct. Biol., 18 (2008) 10-15.
[22] C. Clementi, P.A. Jennings, J.N. Onuchic, Prediction of folding mechanism for circular-permuted proteins, Journal of Molecular Biology, 311 (2001) 879-890.
[23] J.G. Lyubovitsky, H.B. Gray, J.R. Winkler, Mapping the cytochrome C folding landscape, Journal of American Chemical Society, 124 (2002) 5481-5485.
[24] L.C. Oliveira, A. Schug, J.N. Onuchic, Geometrical features of the protein folding mechanism are a robust property of the energy landscape: a detailed investigation of several reduced models, Journal of Physical Chemistry B, 112 (2008) 6131-6136.
[25] M.a. Jamros, L.C. Oliveria, P.C. Whitford, J.N. Onuchic, J.a. Adams, D.K. Blumenthal, P.a. Jennings, Proteins at work: A combined SAXS and theoretical determination of the multiple structures involved on the protein kinase functional landscape, Journal of Biological Chemistry, (2010).
[26] M.A. Jamros, L.C. Oliveira, P.C. Whitford, J.N. Onuchic, J.A. Adams, P.A. Jennings, Substrate-specific Reorganization of the Conformational Ensemble of CSK Implicates Novel Modes of Kinase Function, Plos Computational Biology, 8 (2012).
[27] R. Koradi, M. Billeter, K. Wüthrich, MOLMOL: a program for display and analysis of macromolecular structures, Journal of Molecular Graphics, 14 (1996) 51-55, 29-32.
[28] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD, Journal of Computational Chemistry, 26 (2005) 1781-1802.
[29] A.D. MacKerell, D. Bashford, Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins{\dag}, Journal of Physical Chemistry B, 102 (1998) 3586-3616.

[30] J.K. Noel, P.C. Whitford, K.Y. Sanbonmatsu, J.N. Onuchic, SMOG@ctbp: simplified deployment of structure-based models in GROMACS, Nucleic Acids Research, 38 Suppl (2010) W657-661.

[31] J.K. Noel, P.C. Whitford, J.N. Onuchic, The Shadow Map: A general contact definition for capturing the dynamics of biomolecular folding and function, The Journal of Physical Chemistry B, 116 (2012) 8692-8702.

[32] P.C. Whitford, J.K. Noel, S. Gosavi, A. Schug, K.Y. Sanbonmatsu, J.N. Onuchic, An all-atom structure-based potential for proteins: bridging minimal models with all-atom empirical forcefields, Proteins, 75 (2009) 430-441.

[33] B. Hess, C. Kutzner, D.V.D. Spoel, E. Lindahl, GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation, Journal of Chemical Theory Computation, 4 (2008) 435-447.

[34] A.M. Ferrenberg, R.H. Swendsen, New Monte Carlo technique for studying phase transitions, Physical Review Letters, 61 (1988) 2635-2638.

[35] K.L. Heckman, L.R. Pease, Gene splicing and mutagenesis by PCR-driven overlap extension, Nature Protocols, 2 (2007) 924-932.

[36] F.M. Squina, C.R. Santos, D.a. Ribeiro, J. Cota, R.R. de Oliveira, R. Ruller, A. Mort, M.T. Murakami, R.a. Prade, Substrate cleavage pattern, biophysical characterization and low-resolution structure of a novel hyperthermostable arabinanase from *Thermotoga petrophila*, Biochemical and Biophysical Research Communications, 399 (2010) 505-511.

[37] A.P. Hammersley, FIT2D: an introduction and overview, in, Grenoble, France, 1997, pp. 1-33.

[38] H. Fischer, M. de Oliveira Neto, H.B. Napolitano, I. Polikarpov, a.F. Craievich, Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale, Journal of Applied Crystallography, 43 (2009) 101-109.

[39] D. Svergun, C. Barberato, M.H.J. Koch, CRYSOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, Journal of Applied Crystallography, 28 (1995) 768-773.

[40] G.L. Miller, Use of dinitrosalicylic acid reagent for determination of reducing sugar, Analytical Chemistry, 31 (1959) 426-428.

[41] R. Naran, M.L. Pierce, A.J. Mort, Detection and identification of rhamnogalacturonan lyase activity in intercellular spaces of expanding cotton cotyledons, The Plant Journal, 50 (2007) 95-107.

[42] C.R. Santos, F.M. Squina, A.M. Navarro, D.P. Oldiges, A.F. Leme, R. Ruller, A.J. Mort, R. Prade, M.T. Murakami, Functional and biophysical characterization of a hyperthermostable GH51 alpha-L-arabinofuranosidase from *Thermotoga petrophila*, Biotechnol Lett, 33 (2011) 131-137.

[43] J. Cota, T.M. Alvarez, A.P. Citadini, C.R. Santos, M. de Oliveira Neto, R.R. Oliveira, G.M. Pastore, R. Ruller, R.a. Prade, M.T. Murakami, F.M. Squina, Mode of operation and low-resolution structure of a multi-domain and hyperthermophilic endo-β-1,3-glucanase from *Thermotoga petrophila*, Biochemical and Biophysical Research Communications, 406 (2011) 590-594.

[44] M.T. Murakami, R.K. Arni, D.S. Vieira, L. Degrève, R. Ruller, R.J. Ward, Correlation of temperature induced conformation change with optimum catalytic activity in the recombinant G/11 xylanase A from *Bacillus subtilis* strain 168 (1A1), FEBS Letters, 579 (2005) 6505-6510.

[45] D.S. Vieira, R.J. Ward, Conformation analysis of a surface loop that controls active site access in the GH11 xylanase A from *Bacillus subtilis*, Journal of Molecular Modeling, 18 (2012) 1473-1479.

[46] G.P. Furtado, L.F. Ribeiro, C.R. Santos, C.C. Tonoli, A.R. Souza, R.R. Oliveira, M.T. Murakami, R.J. Ward, Biochemical and structural characterization of a β-1,3–1,4-glucanase from *Bacillus subtilis* 168, Process Biochemistry, 46 (2011) 1202-1206.

[47] I.B. Grishina, R.W. Woody, Contributions of tryptophan side chains to the circular dichroism of globular proteins: exciton couplets and coupled oscillators, Faraday Discussions, 99 (1994) 245-262.

[48] R. Ruller, L. Deliberto, T.L. Ferreira, R.J. Ward, Thermostable variants of the recombinant xylanase A from *Bacillus subtilis* produced by directed evolution show reduced heat capacity changes, Proteins, 70 (2008) 1280-1293.

[49] J. Aÿ, F. Götz, R. Borriss, U. Heinemann, Structure and function of the *Bacillus* hybrid enzyme GluXyn-1: native-like jellyroll fold preserved after insertion of autonomous globular domain, Proceedings of the National Academy of Sciences of the United States of America, 95 (1998) 6613-6618.

[50] L.F. Ribeiro, G.P. Furtado, M.R. Lourenzoni, A.J. Costa-Filho, C.R. Santos, S.C. Peixoto Nogueira, J.a. Betini, M.D.L.T.M. Polizeli, M.T. Murakami, R.J. Ward, Engineering bifunctional laccase-xylanase chimerae for improved catalytic performance, The Journal of Biological Chemistry, 286 (2011) 40026-40038.

[51] M.L. Connolly, SCIENCE : Solvent-accessible surfaces of proteins and nucleic acids, Science, 221 (1983) 709-713.

[52] L. Bülow, Characterization of an artificial bifunctional enzyme, beta-galactosidase/galactokinase, prepared by gene fusion, European Journal of Biochemistry, 163 (1987) 443-448.
[53] L. Chang, M. Ding, L. Bao, Y. Chen, J. Zhou, H. Lu, Characterization of a bifunctional xylanase/endoglucanase from yak rumen microorganisms, Applied Microbiology and Biotechnology, 90 (2011) 1933-1942.
[54] R. Khandeparker, M.T. Numan, Bifunctional xylanases and their potential use in biotechnology, Journal of Industrial Microbiology & Biotechnology, 35 (2008) 635-644.
[55] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, J. Mol. Graph., 14 (1996) 33-38.

SUPPLEMENTARY INFORMATION

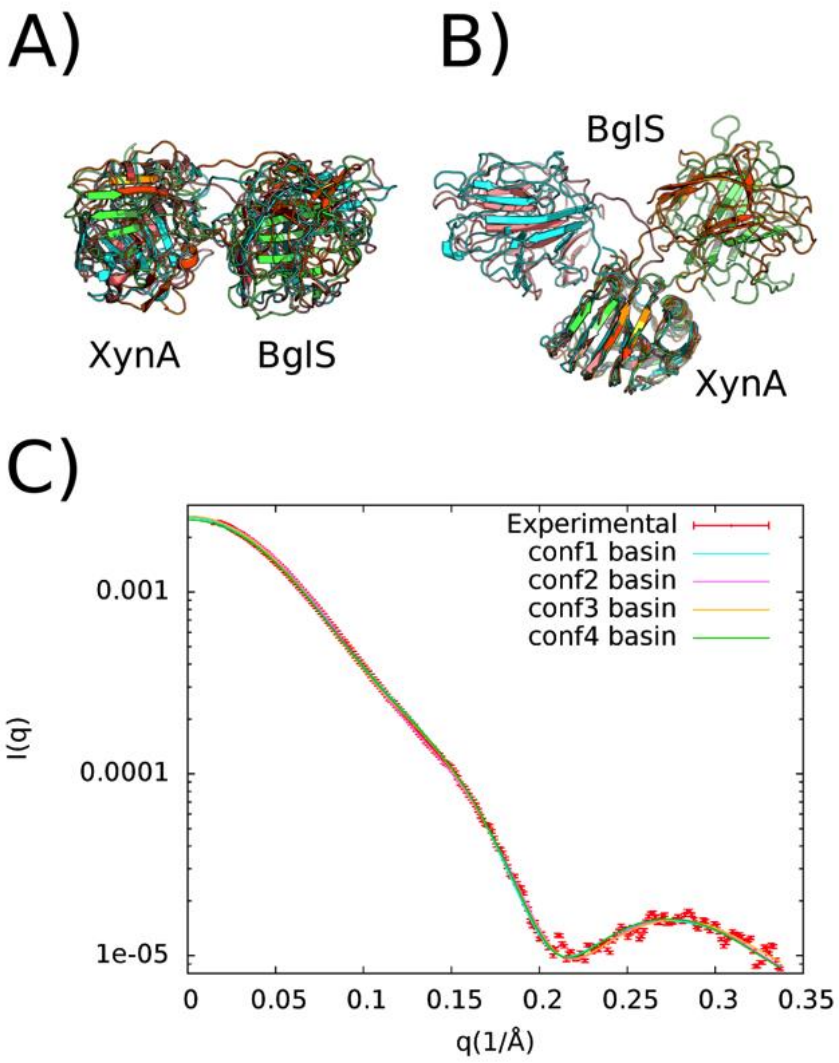


Figure SI-3.1. Analysis of the candidates to describe the chimera arrangement in solution. A) Superposition of some conformations extracted from the free energy basin and the B) Structural alignment at the XynA domain. C) Theoretical scattering of the conformations and experimental curves. The scattering curves shown in Figure C share the same color as the derived protein model presented in A and B.

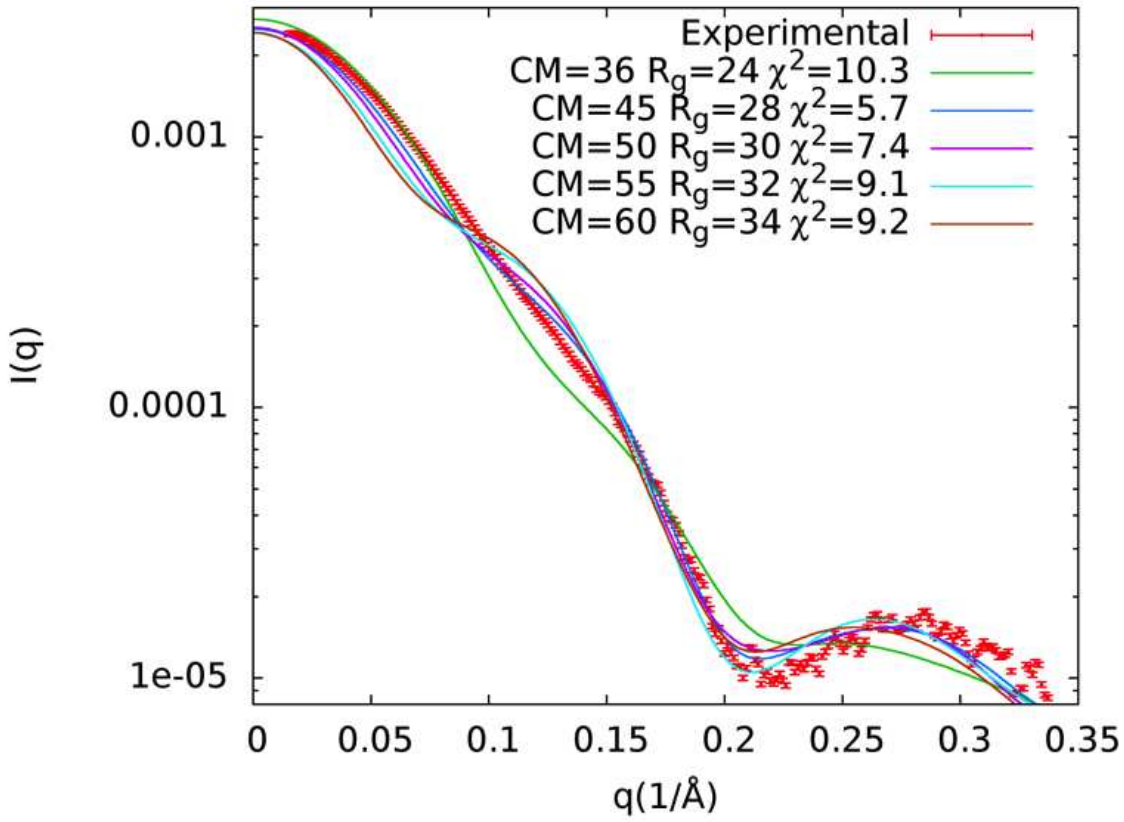


Figure SI-3.2. Scattering curves of conformations extracted from the regions outside the free energy basin. The evaluated conformations did not show a satisfactory agreement with the experimental result.

Tukey test for the xylanase activity of XynA

Table SI-3.1. ANOVA table for the specific xylanase activities of XynA.

ANOVA		Design	Casualised		
Source of variation	DF	SS	MS	F	Prob{>F}
Treatment	4	11.87	2.97	76.29	0.0000**
Residual	10	0.39	0.04		
Total	14	12.26			

Significance level: **: 1%; *: 5%.

Mean	3.11
Standard deviation	0.20
HSD*	0.53
Coefficient of variation %	6.35

*HSD means the honestly significant difference. (x) More than one value because of many different observations between the treatments.

Tukey test (5%)

Treatments	Mean	Significance
Rye arabinoxylan	3.81	a
Beechwood xylan	3.79	a
Birchwood xylan	3.36	ab
Oat spelt xylan	3.17	b
Wheat arabinoxylan	1.40	c

The same letter indicates that there is no difference between the means at 5% of significance. Thus, the specific activities of XynA in birchwood xylan, beechwood xylan and rye arabinoxylan are statistically the same (a). However the activity in birchwood xylan seems to be equal to oat spelt xylan (b), and the wheat arabinoxylan (c) reached the lowest value of activity.

Tukey test for the xylanase activity of XylLich

Table SI-3.2. ANOVA table for the specific xylanase activities of XylLich.

ANOVA	Design	Casualised			
Source of variation	DF	SS	MS	F	Prob{>F}
Treatment	4	6.65	1.66	132.27	0.0000**
Residual	8	0.10	0.01		
Total	12	6.75			

Significance level: **: 1%; *: 5%.

Mean	2.45
Standard deviation	0.11
HSD*	(x)
Coefficient of variation %	4.59

*HSD means the honestly significant difference. (x) More than one value because of many different observations between the treatments.

Tukey test (5%)

Treatments	Mean	Significance
Rye arabinoxylan	2.95	a
Beechwood xylan	2.88	a
Birchwood xylan	2.75	a
Oat spelt xylan	2.15	b
Wheat arabinoxylan	0.88	c

The same letter indicates that there is no difference between the means at 5% of significance. Thus, the specific activities of XylLich in birchwood xylan, beechwood xylan and rye arabinoxylan are statistically the same (a). Furthermore the activity in birchwood xylan (b) seems to be different, as well as the wheat arabinoxylan (c), which reached the lowest value of activity.

CAPÍTULO 4

STRUCTURAL AND BIOCHEMICAL pH-DEPENDENT INSIGHTS FOR HYPERTHERMOPHILIC β-GLUCOSIDASES

ABSTRACT

β-Glucosidases (BGLs) are very useful enzymes with a great potential to be employed in several industrial processes. However, some features are required to become viable the enzyme applications, such as temperature and pH stability as well, low ions and chemicals inhibition. Thus this work aimed to study three BGLs from the extremophiles organisms *Pyrococcus furiosus* and *Thermotoga petrophila*. The genes *PfBgl1*, *TpBgl1* and *TpBgl3* were cloned into pET28a vector and the proteins were expressed in *Escherichia coli* and further purified in two chromatographic steps. The purified enzymes were evaluated for pH and temperature of activity, which showed that BGLs from the glycosyl hydrolases family 1 (PfBgl1, TpBgl1) presented a wider range of pH and temperature operation than BGL from family 3 (TpBgl3). The BGLs showed great stability to a range of pH (4-10) and the highest time of half-life (at 99 °C) was at pH 6, besides they were not significantly affected by the presence of EDTA or ions, except TpBgl1 that was inhibited by Hg^{2+} and Fe^{2+}. The specific activities in a set of different substrates suggested that TpBgl3 is more specific than GH1 BGLs. The k_{cat} and k_{cat}/K_m in pNPG indicate that TpBgl3 is the most efficient among BGLs characterized herein, although this enzyme is inhibited with the lowest glucose concentration (30.1 mM). Furthermore, the BGLs were assayed for influence of six monosaccharides in catalysis, which the results suggested a weak inhibition by the most of those carbohydrates tested. The CD experiments revealed that the secondary structure of BGLs is not affected by the pH variations and the denaturation studies evidenced that the BGLs are indeed hyperthermophilic.

4.1 INTRODUCTION

An effective bioprocess is the basis for the biotechnological industry in a competitive market [1]. However, to develop a profitable enzymatic process it is necessary to overcome a great bottle-neck: the cost-effectiveness ratio of producing and employing enzymes, which should be a function of the products values. β-Glucosidases (BGLs) are very useful enzymes that can be employed in a large number of processes and industrial approaches, such the production of anti-cancer compounds [2], the reduction of the toxicity of animal feed [3], the improvement of the aromatic potential of wine [4], and the hydrolysis of oligosaccharides from biomass for biofuels production [5].

A minimal cellulolytic system requires at least three classes of enzymes, like endo-1,4-β-glucanase (E.C. 3.2.1.4), cellobiohydrolase (E.C. 3.2.1.91), and β-glucosidase [6]. BGLs are exo-type enzymes that operate at the non-reducing ends, degrading cellobiose and other cellooligosaccharides produced by depolymerizing enzymes [7]. They are classified as E.C. 3.2.1.21 according to their function and as glycosyl hydrolases families 1 or 3 depending on their structures and molecular mechanisms [8, 9]. Several BGLs have been identified and characterized in different microorganisms, including insects [10, 11], filamentous fungi [12, 13], yeast [14], bacteria [15, 16], and from metagenomic approaches [17].

The screening of new industrial enzymes should consider several parameters, which the most important are thermostability, wide pH range of action, high catalytic efficiency, and low product inhibition. Thus, because of many emerging pretreatments of biomass like chemical [18] or steam explosion [19], it become necessary to find out new enzymes with an enhanced pH and thermal resistance aiming to the development of co-hydrolysis process by adding enzymes in the pretreated biomass slurry [20, 21]. In addition, an enzyme with this profile has a great potential to be applied in some high temperature processes of food industry [7, 22, 23], such as those in dairy and brewing industries.

During the last three decades it is known that the great challenge regarding the BGL usages may be placed in two main points: the product (glucose) inhibition and its low thermal stability [24]. In this way, the objective of this study was to characterize three hyperthermostable bacterial β-glucosidases from the organisms *Thermotoga petrophila* and *Pyrococcus furiosus* in order to identify BGLs with high potential for industrial use. This work comes up with several comparisons between β-glucosidases from two distinct families of glycosyl hydrolases (GH1 and GH3), looking

up the differences towards the glucose inhibition, pH range of activity and stability, half-life in high temperature, the substrates specificities, and reaction mechanisms.

4.2 EXPERIMENTAL PROCEDURES

4.2.1 Cloning, expression and purification of β-glucosidases genes

The full-length coding genes [*PfBgl1* (PF0073) from *P. furiosus*, *TpBgl1* (Tpet_TpBgl1) and *TpBlg3* (Tpet_TpBgl3) from *T. petrophila*] cloned into pBAD/Myc-His plasmid (Invitrogen, CA), were used as template for standard PCR method for cloning the mature enzyme without signal peptide. The primer sets used for amplification were: 5'-TATATAGCTAGCATGAAGTTCCCAAAAAACTT-3' and 5'-TATATAGGATCCCTACTTTCTTGTAACAAATT-3' for *PfBgl1*, 5'-TATATAGCTAGCATGAACGTGAAAAAGTTCCC-3' and 5'-TATATAGAATTCTCAATCTTCCAGACTGTTGC-3' for *TpBgl1*, 5'-TATATAGCTAGCATGGGAAAGATCGATGAAAT-3' and 5'-TATATAGGATCCTCATGGTTTGAATCTCTTCT-3' for *TpBgl3*. The amplicons were cloned into the pET28a (Novagen) vector using NdeI and BamHI/EcoRI restriction sites. All the enzymes were produced in *E. coli* BL21(DE3) pRARE2. The proteins expression and purification steps, included a Ni^{2+}-chelating affinity and size-exclusion chromatography, were performed according to Squina et al., 2010 [25]. The purified enzymes were further analyzed by SDS-PAGE and the protein concentrations were determined by absorbance at 280 nm.

4.2.2 Standard assay and enzyme characterization

The standard enzymatic assay for β-glucosidase evaluation was performed at optimal pH for each enzyme in appropriate buffer (phosphate or acetate) and temperature of 70 °C using 0.5 mM 4-nitrophenyl-β-D-glucopyranoside (pNPG) as substrate. The reaction was stopped by addition of 1 volume of 1 M Na_2CO_3 and the formation of 4-nitrophenol was monitored colorimetrically at 400 nm [4] using Infinite® 200 PRO microplate reader (TECAN). All the assays were carried out employing the automated pipetting system ep*Motion*® 5075 (Eppendorf). One unit of enzyme was defined as the quantity of enzyme that released 4-nitrophenol or reducing sugars (in case of

polysaccharides or cellulosic substrates) at rate of 1 µmol per minute under standard conditions. Reducing sugars were determined using the 3,5-dinitrosalicylic acid (DNS) method and monitored colorimetrically at 540 nm [26].

A Response Surface Methodology (RSM) was employed to optimize the reaction conditions for β-glucosidase activity. A Central Composite Rotatable Design (CCRD) totalizing 12 experiments with four central points (Table 4.1) was performed for optimizing the variables pH and temperature (k = 2). Details concerning the statistical approaches for those experiments can be found in Myers and Montgomery [27]. In addition a simple curve of pH variation was carried out for the enzymes under standard conditions.

In order to evaluate thermal stability of hyperthermophilic BGLs, those enzymes were incubated (1 µM) at 99 °C in several pH levels (Table 4.3) and aliquots were taken at different incubation times and kept on ice until the evaluation of remaining enzymatic activity. The residual activities were performed in standard assay and the time of half-life ($t_{1/2}$) of β-glucosidase activity was estimated using a mathematical approach previously described [28]. The pH stabilities of BGL were evaluated by measuring the residual pNPGase activity after incubation in McIlvaine-glycine added buffer in different pH levels at 8 °C for 3 hours and 1 week.

The substrate specificities of β-glucosidases were evaluated using standard assays in three sets of substrates (Table 4.5). The nitrophenol derived substrates were assayed in a 10 minutes reaction. Polyssacharides (purchased from the best source possible, Sigma Aldrich and Megazyme) were assayed in a 60 minutes reaction and cellulosic substrates were assayed in a 12 hours reaction. The enzymatic activity and substrate specificity were determined from the amount of p-nitrophenol released from pNPG by the Na_2CO_3 method. In case of polysaccharides and cellulosic substrates it was employed DNS method [26] to measure the amount of reducing sugars released, or Glucose Oxidase/Peroxidase commercial kit method for the cellobiose and lactose substrates.

The products of the BGLs hydrolysis were assessed by means of capillary zone electrophoresis (CZE) of oligosaccharides and the 8-aminopyreno-1,3,6-trisulfonic acid (APTS) labeling and were performed according to previous work [29, 30]. The effects of several ions and other chemical compounds were evaluated under standard assay at the final concentration of 10 mM.

The kinetic parameters for β-glucosidase activities were estimated from initial rates at twelve substrate concentrations in the ranges of 0.1 – 30 mM of pNPG, and 1 – 100 mM of

cellobiose in standard assays. The kinetic constants were assessed using the Michaelis-Menten approach.

4.2.3 Circular dichroism

Far-UV circular dichroism (CD) measurements (190–260 nm) were carried out using a JASCO 810 spectropolarimeter (JASCO Inc., Tokyo, Japan) equipped with a Peltier module for temperature control. The measurements parameters were the same previously described [29]. Blank corrections were made in all spectra. All spectroscopic measurements were carried out with a sample prepared in 20 mM McIlvaine-glycine added buffer at the final concentration of 0.06 mg/ml. Results were expressed as mean residual molar ellipticity (q) in 10^3 deg cm^2 $dmol^{-1}$.

4.3 RESULTS

4.3.1 Optimization of pH and temperature of catalysis

A Central Composite Design with two variables (pH and temperature) was used to optimize BGL reaction conditions and all runs and results are presented in Table 4.1. The statistical significance of the models was checked by the F test (ANOVA) and the results are shown in Table 4.2. The experimental data were adjusted with very accuracy, showing variation coefficients (R^2) of 0.94, 0.98, 0.92 for PfBgl1, TpBgl1, and TpBgl3 respectively, besides the highly significant F test values (24.93, 52.04, 17.81) for the regression, which were at least 4.1 times higher than the critical value. Considering the inherent variability of bioprocesses, the model could be considered to be predictive and was therefore used to generate a response surface and contour plot (Fig. 4.1) for β-glucosidase activity. In addition, the fit of the statistical models to the experimental data was validated at the optimal region for activity, resulting in low deviations of the predict value in comparison to the experimental value (6.4, 1.5, and 5.6 %) (data not shown).

The results indicate that the enzymes are hyperthermophilic, and work better at pH range of 4.0 – 6.0 (PfBgl1; Fig. 4.1A, 4.1B), 5.5 – 7.5 (TpBgl1; Fig. 4.1C, 4.1D), and 3.5 – 4.5 (TpBgl3; Fig. 4.1E, 4.1F) and at a wide temperature range of 90 – 99 °C (PfBgl1; Fig. 4.1A, 4.1B), 80 – 99 °C (TpBgl1; Fig. 4.1C, 4.1D), and 90 – 99 °C (TpBgl3; Fig. 4.1E, 4.1F). Similar profiles were found in the experiment varying only pH, showing the optimum activity at pH 4.5 (PfBgl1), 6.0 (TpBgl1),

and 4.0 (TpBgl3), with a relative activity of at least 70% ranging from 3.0 to 6.0, 5.0 to 7.5, and 3.0 to 4.5 (Fig. 4.2A).

Table 4.1. Central composite rotatable design (CCRD) for pNPGase activity of β-glucosidases using pH (X1) and temperature (X2). Specific activity is given by the ratio between the standard unit (see methods) and the quantity of enzyme used in the assay (nmol).

Run	pH		Temperature (°C)		Specific Activity (IU/nmol)[a]		
	Coded	Real	Coded	Real	PfBgl1	TpBgl1	TpBgl3
1	-1	4.2	-1	70.0	36.21	4.21	7.57
2	1	7.8	-1	70.0	11.67	5.03	0.68
3	-1	4.2	1	94.0	82.37	4.50	9.23
4	1	7.8	1	94.0	41.25	8.45	1.97
5	-1.414	3.5	0	82.0	36.51	1.48	7.87
6	1.414	8.5	0	82.0	18.06	4.93	0.75
7	0	6.0	-1.414	65.0	24.59	4.49	0.89
8	0	6.0	1.414	99.0	99.87	9.59	5.83
9	0	6.0	0	82.0	56.53	9.23	2.88
10	0	6.0	0	82.0	61.09	9.25	2.94
11	0	6.0	0	82.0	57.89	9.66	2.74
12	0	6.0	0	82.0	55.71	9.34	2.74

[a]Average of triplicate measurements

Table 4.2. ANOVA of the models generated to predict the pNPGase activity of BGLs.

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F-value*
PfBgl1				
Model	7013.52	5	1402.70	24.93
Residual	337.57	6	56.26	
Total	7351.09	11		
TpBgl1				
Model	87.10	5	17.42	52.04
Residual	2.01	6	0.33	
Total	89.10	11		
TpBgl3				
Model	92.70	5	18.54	17.81
Residual	6.25	6	1.04	
Total	98.95	11		

Coefficient of determination (R^2): 0.94 for PfBgl1, 0.98 for TpBgl1, and 0.92 for TpBgl3

*$F_{0.05;\,5;\,6} = 4.39$

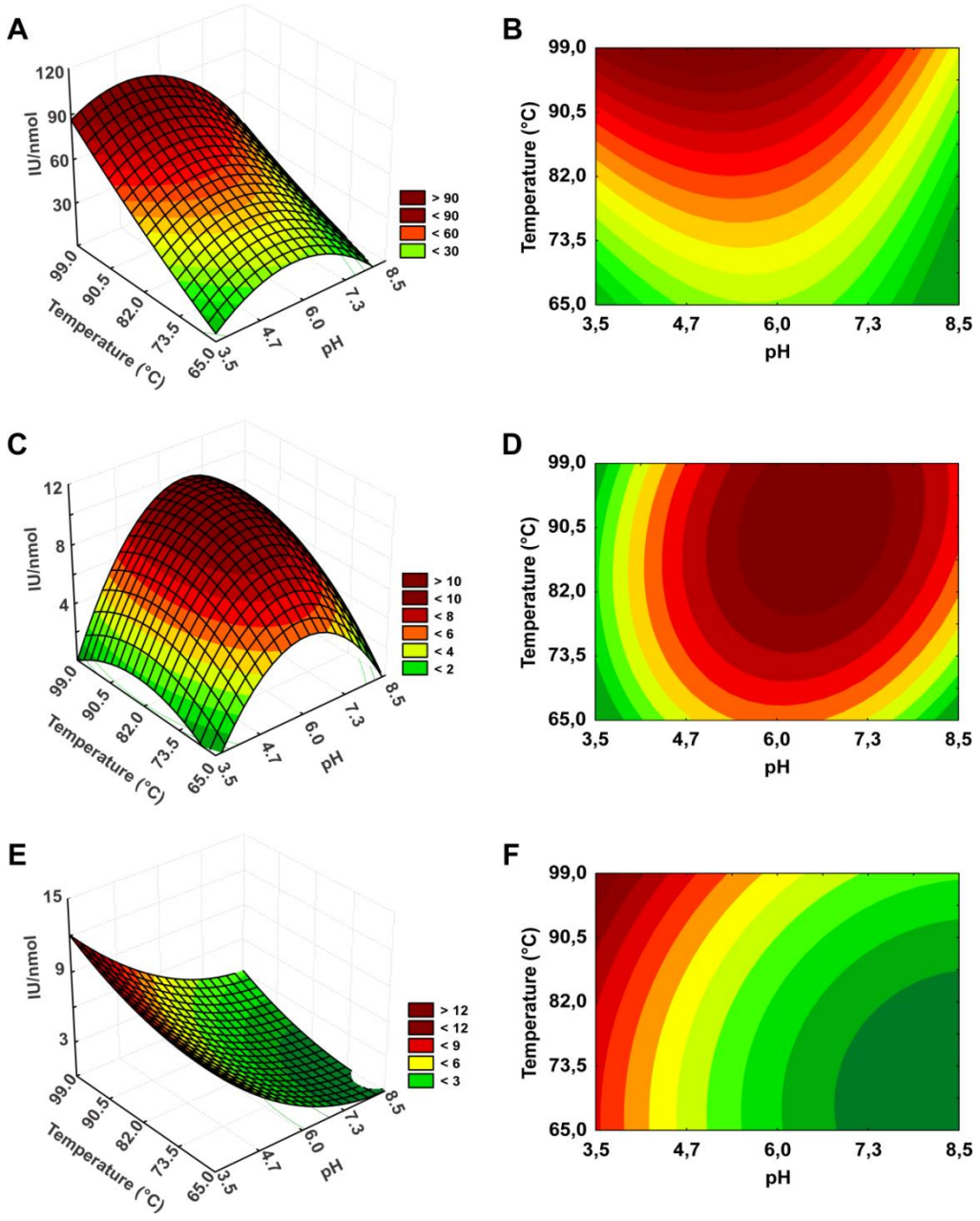


Figure 4.1. Influence of pH and temperature on the pNPGase activity. Response surface plot and contour plot for the effects of pH and temperature on β-glucosidase activity of PfBgl1 (A, B), TpBgl1 (C, D), and TpBgl3 (E, F). The statistic models which predict the enzyme activities have a coefficient of variation of 0.95, 0.98, 0.94 respectively, and were obtained by using experimental data.

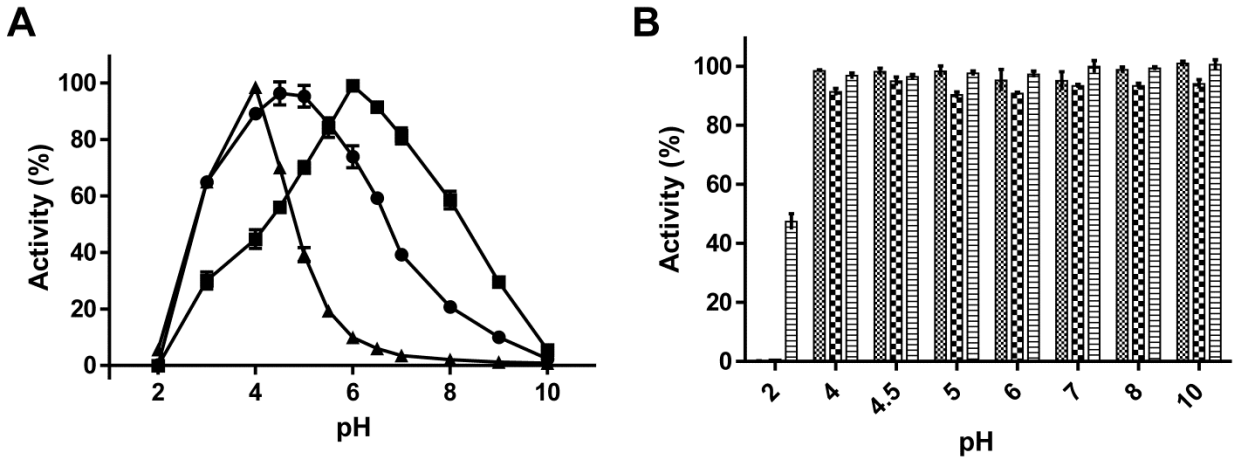


Figure 4.2. Effects of pH on pNPGase activity and stability. (A) Curve of pH for enzymatic activities of PfBgl1 (●), TpBgl1 (■), and TpBgl3 (▲) at 70 °C. (B) The bar graph show the pH stability for PfBgl1 (small grid bar), TpBgl1(large grid bar), and TpBgl3 (horizontal lines bar). The bars present the residual β-glucosidase activities after 1 week incubated at 8 °C in McIlvaine-glycine added buffer in several pH levels. The percentages where accessed by comparing the activities to initial activity before incubation.

4.3.2 Stability to pH and temperature

The enzymatic stability to pH was accessed by incubating the purified proteins in buffer ranging from pH 2.0 to 10.0 using 8 pH values (Fig. 4.2B). The residual pNPGase activity was measured after 3 hours (data not shown) and 7 days incubations in standard assays. In both conditions the same profile was reached for the three enzymes, in which there were no significant differences in β-glucosidase stability with pH ranging from 4.0 to 10.0. Only GH3 enzyme (TpBgl3) partially held out at pH 2.0 preserving 47% of activity, while the GH1 enzymes (PfBgl1 and TpBgl1) lost 100% of their enzymatic activities. The percentages of remaining activity were calculated from the initial enzymatic activity before incubation.

Three temperatures (70, 80, and 90 °C) were firstly tested in the thermal stability experiments, but there were no significant changes in enzymatic activities after 16h of incubation at 70 or 80 °C, and at 90 °C the decay of β-glucosidase activity was very slow. Thus, aiming to evaluate the enzymatic thermal stability, the assays were carried out at 99 °C in seven different pH

(4.0, 4.5, 5.0, 6.0, 7.0, 8.0 and 10.0) using standard assays. All the estimated times of half-life of the enzymatic activity are presented in Table 4.3. The highest thermostability at 99 °C was achieved when the protein was incubated at pH 6.0 (115.5 h, PfBgl1; 10.5 h, TpBgl1; 7.7 h, TpBgl3), and the lowest was at pH 4.0 (76 min, PfBgl1; < 5 min, TpBgl1; < 5 min, TpBgl3). The thermostability was reduced in an average of 30% (PfBgl1), 70% (TpBgl1) and 85% (TpBgl3) by incrementing or decreasing 1 pH unit regarding the maximum (pH 6.0), while PfBgl1 retained about 33% and 55% of the highest $t_{1/2}$ (pH 6.0) at pH 4.5 and 8.0, respectively.

Table 4.3. Thermostability of β-glucosidases at 99 °C in different pHs.

	Time of Half-Life (min)		
pH	**PfBgl1**	**TpBgl1**	**TpBgl3**
4.0	76	< 5	< 5
4.5	2311	-	-
5.0	4332	217	102
6.0	6932	630	462
7.0	5776	165	112
8.0	3851	58	26
10.0	495	11	16

4.3.3 Effect of EDTA and ions in enzymatic catalysis

Aiming to evaluate the effects of several ions and EDTA on the β-glucosidase activity a set of experiments was performed. A control reaction was carried out under standard assay and it was taken as 100%. Table 4.4 points the influence of all ions in enzymatic activity. The addition of EDTA did not affect the β-glucosidase activity of BGLs. Only the salts containing Fe^{2+} and Hg^{2+} caused substantial losses in pNPGase activity of BGLs. The catalytic activity was reduced to 81.7, 16.7 and 62.2% in presence of Fe^{2+} and 37.5, 0 and 69.0% when Hg^{2+} was added, for respectively PfBgl1, TpBgl1 and TpBgl3. All other ions did not have effect on the activity of BGLs, except Ni^{+}, Co^{2+}, Mn^{2+} and Cd^{2+} that reduced the TpBgl1 activity to 50 to 70%.

Table 4.4. Effects of ions and EDTA on the catalytic activity of BGLs.

	Relative Activity (%)		
	PfBgl1	**TpBgl1**	**TpBgl3**
Control	100.0	100.0	100.0
EDTA	108.0	105.4	108.8
LiCl	108.0	98.2	93.3
KCl	109.9	89.6	105.9
$CaCl_2$	96.8	104.2	107.7
$NiCl_2$	82.6	49.4	100.7
$CoCl_2$	87.4	71.2	108.2
$MgCl_2$	108.8	88.7	105.9
$MnCl_2$	101.7	64.9	104.3
$CdCl_2$	92.7	58.9	107.7
$FeCl_2$	81.7	16.7	62.2
$HgCl_2$	37.5	0.0	69.0

4.3.4 Substrate specificities of β-glucosidases

Our findings into substrate specificity show that the three BGLs hydrolyse β-(1,4) and β-(1,3) linkages. Table 4.5 summarizes the specific activities in 15 distinct substrates. The enzymes were not able by themselves to release reducing sugars from six cellulosic substrates assayed neither from nine cell-wall polysaccharides substrates tested (data not presented). When evaluating the hydrolysis efficiency in several polysaccharides we found that those enzymes only work in β-(1,4) or β-(1,3) linkages glucose containing substrates, like laminarin, lichenan and β-glucan. In this way the highest specific activity in polysaccharides group was found to be 100% in laminarin for PfBgl1 and TpBgl3 and in lichenan for TpBgl1. The ratio of the activities in laminarin:lichenan:β-glucan was very diverse among the enzymes, as 6:2:1 for PfBgl1, 1:1:1 for TpBgl1, and 12:1:2 for TpBgl3 (Table 4.5).

When assayed to hydrolyze disaccharides, like cellobiose and lactose, all the β-glucosidases worked in cellobiose as expected, but only the GH1 enzymes (PfBgl1 and TpBgl1) were able to hydrolyze lactose, which was in a rate of 55 – 60% of specific activity in cellobiose. Regarding the group of pNP derived substrates, the highest specific activities were reached, and there were many differences in the activities on those substrates among the enzymes (Table 4.5). The best substrates to β-glucosidase activity were 4-NP-β-D-Glucopyranoside for TpBgl1 and TpBgl3, and 4-NP-β-D-

Fucopyranoside for PfBgl1. Notably the GH1 enzymes were more promiscuous than GH3 TpBgl3, because they hydrolyzed many substrates, while the TpBgl3 was specific to cellobiose and pNPGlu.

The product of enzymatic hydrolysis of cellohexaose and laminarihexaose were analyzed by CZE. In both cases the major reaction product was glucose, indicating an exo-acting pattern of cleavage for all the BGLs (see Figure SI-4.1 in supplementary information section).

Table 4.5. Specific activities of β-glucosidases on different types of substrates.

	PfBgl1		TpBgl1		TpBgl3	
Substrate	**Specific Activity**	**Relative Activity**	**Specific Activity**	**Relative Activity**	**Specific Activity**	**Relative Activity**
Noncellulosic substrates	***μmol/min/nmol***	***%***	***μmol/min/nmol***	***%***	***μmol/min/nmol***	***%***
Laminarin	0.06 ± 0.00	100.0	0.13 ± 0.01	97.5	0.06 ± 0.00	100.0
Lichenan	0.02 ± 0.00	28.1	0.14 ± 0.01	100.0	0.005 ± 0.00	8.5
β-Glucan (Barley)	0.01 ± 0.00	17.3	0.13 ± 0.01	97.9	0.01 ± 0.00	16.9
Disaccharides substrates	***μmol/min/nmol***	***%***	***μmol/min/nmol***	***%***	***μmol/min/nmol***	***%***
Cellobiose	2.24 + 0.03	100.0	2.03 ± 0.05	100.0	0.56 ± 0.04	100.0
Lactose	1.25 ± 0.07	55.5	1.24 ± 0.01	61.1	0.01 ± 0.00	2.2
4-nitrophenyl substrates	***μmol/min/nmol***	***%***	***μmol/min/nmol***	***%***	***μmol/min/nmol***	***%***
4-NP-β-D-Cellobioside	1.70 ± 0.22	28.8	3.28 ± 0.28	35.0	0.30 ± 0.03	3.3
4-NP-β-D-Glucopyranoside	5.89 ± 0.17	100.0	9.36 ± 0.28	100.0	9.03 ± 0.90	100.0
4-NP-α D-Glucopyranoside	0.08 ± 0.01	1.3	0.04 ± 0.00	0.4	0.03 ± 0.01	0.4
2-NP-β-D-Glucopyranoside	0.99 ± 0.05	16.7	1.50 ± 0.06	16.1	1.34 ± 0.04	14.8
4-NP-β-D-Galactopyranoside	2.00 ± 0.20	34.0	3.33 ± 0.20	35.6	0.05 ± 0.01	0.6
4-NP-β-D-Xylobioside	0.38 ± 0.04	6.5	0.17 ± 0.00	1.8	0.85 ± 0.05	9.4
4-NP-β-D-Mannopyranoside	0.12 ± 0.02	2.1	0.03 ± 0.00	0.3	0.04 ± 0.02	0.5
4-NP-α-L-Arabinofuranoside	0.10 ± 0.04	1.7	0.02 ± 0.00	0.2	0.06 ± 0.01	0.7
4-NP-α-L-Arabinopyranoside	0.43 ± 0.07	7.3	0.42 ± 0.02	4.5	0.04 ± 0.01	0.4
4-NP-β-D-Fucopyranoside	8.36 ± 0.51	142.0	8.76 ± 0.29	93.6	0.10 ± 0.01	1.1

Values are given by the mean ± S.D. of three independent assays. The specific activity was defined as μmol of reducing sugars released per minute per nmol of enzyme. The enzymes were not able by themselves to release reducing sugars from six cellulosic substrates assayed (CMC, avicel, cellulose microcrystalline, sigmacell, sugarcane bagasses *in natura* and pretreated) neither from nine cell-wall polysaccharides tested (sugar beet, glucomannan, potato starch, beechwood xylan, xyloglucan, rye arabinoxylan, linear arabinan, wheat arabinoxylan, mannan).

4.3.5 Assessment of the kinetic parameters and monosaccharides inhibition

The kinetic parameters were estimated for BGLs using the GraphPad Prism® 5.0 and they are displayed in Table 4.6. It was found in pNPG a V_{max} value at least about 70% greater than in

cellobiose. Looking at the parameters V_{max}, K_m and k_{cat} it is possible to affirm that BGLs are more processive in pNPG than in cellobiose. Concerning to catalytic efficiency, the extremely high ratio k_{cat}/K_m found in pNPG is several times higher than those found in cellobiose. This can be explained by the great difference between K_m values in pNPG and cellobiose, which are notably lower in pNPG.

Table 4.6. Kinetic parameters of hyperthermophilic BGLs.

	pNPG			Cellobiose		
Parameter	PfBgl1	TpBgl1	TpBgl3	PfBgl1	TpBgl1	TpBgl3
V_{max} (IU/nmol)	19.06 ± 0.75	16.56 ± 0.21	34.25 ± 0.64	11.36 ± 0.37	7.11 ± 0.16	1.95 ± 0.07
K_m (mM)	0.24 ± 0.04	0.28 ± 0.02	0.38 ± 0.03	6.38 ± 0.71	18.24 ± 1.18	39.32 ± 3.18
K_i (mM)	207.00	1100.00	30.10	N.D.	N.D.	N.D.
k_{cat} (s^{-1})	317.67 ± 12.48	276.00 ± 3.53	570.83 ± 10.63	189.33 ± 6.24	118.47 ± 2.72	32.57 ± 1.22
k_{cat}/K_m	1348.90 ± 323.57	999.28 ± 166.59	1490.43 ± 348.53	29.66 ± 8.82	6.49 ± 2.31	0.83 ± 0.38

N.D. means not determined. Values are given by the mean ± S.D. of three independent assays. The V_{max} unit was defined as µmol of reducing sugars released per minute per nmol of enzyme.

Six monosaccharides were assayed by increasing its concentration up to 1 M in the pNPGase reaction (Fig. 4.3). Fructose, xylose and arabinose did not have significant effect on catalysis for PfBgl1 and TpBgl3. Galactose did not affect TpBgl3 either, but caused reduction of activity at higher concentrations for PfBgl1. Glucose and mannose also inhibited those two enzymes. TpBgl1 showed an unusual response to the saccharides, and only fructose did not have influence on its activity. Glucose, mannose, galactose, and arabinose stimulated in different levels the catalysis of TpBgl1 and provide decrease at higher concentrations. The highest stimulation of the TpBgl1 activity was by xylose, but in the studied range (up to 1 M) there was no decay, only stabilization of the curve.

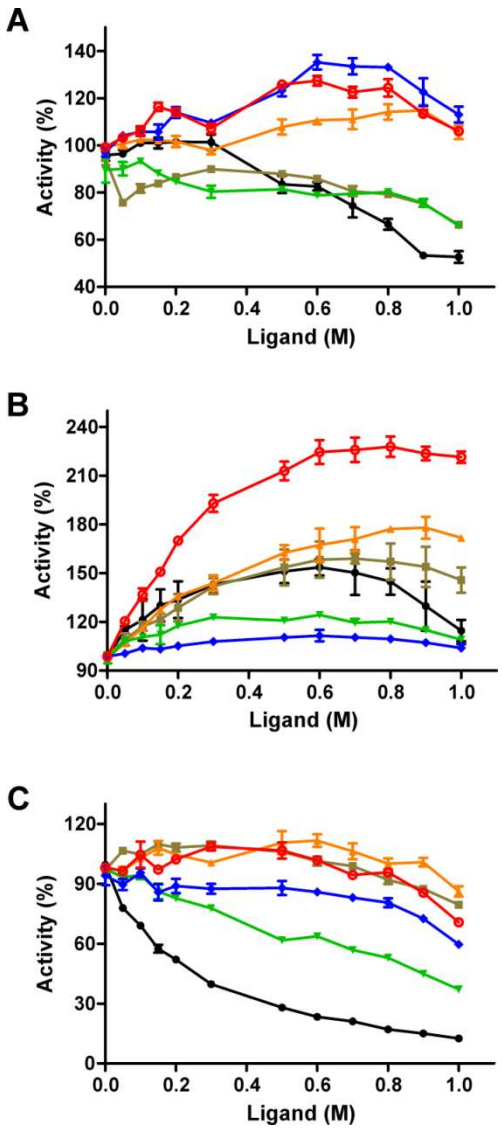


Figure 4.3. Influence of several ligands on pNPGase activity. Inhibition/activation effects of 6 monosaccharides at 12 concentrations (up to 1 M) on β-glucosidase activity of PfBgl1 (A), TpBgl1 (B), and TpBgl3 (C). The monosaccharides assayed were glucose (● black), galactose (■ dark yellow), arabinose (▲ orange), mannose (▼ green), fructose (♦ blue), and xylose (○ red).

4.3.6 Evaluation of stability to pH and temperature by circular dichroism

Far-UV CD spectrum of PfBgl1, TpBgl1 and TpBgl3 indicate that those enzymes present distinct susceptibilities to change in pH. Although all enzymes remain folded in the tested pH (Fig. 4.4A, 4.4C, 4.4E), low level of denaturation was found for PfBgl1 and TpBgl3 (Fig. 4.4A, 4.4E).

However, the transfer from acidic (pH 2) to a basic environment (pH 10) did not result in substantial CD spectral changes.

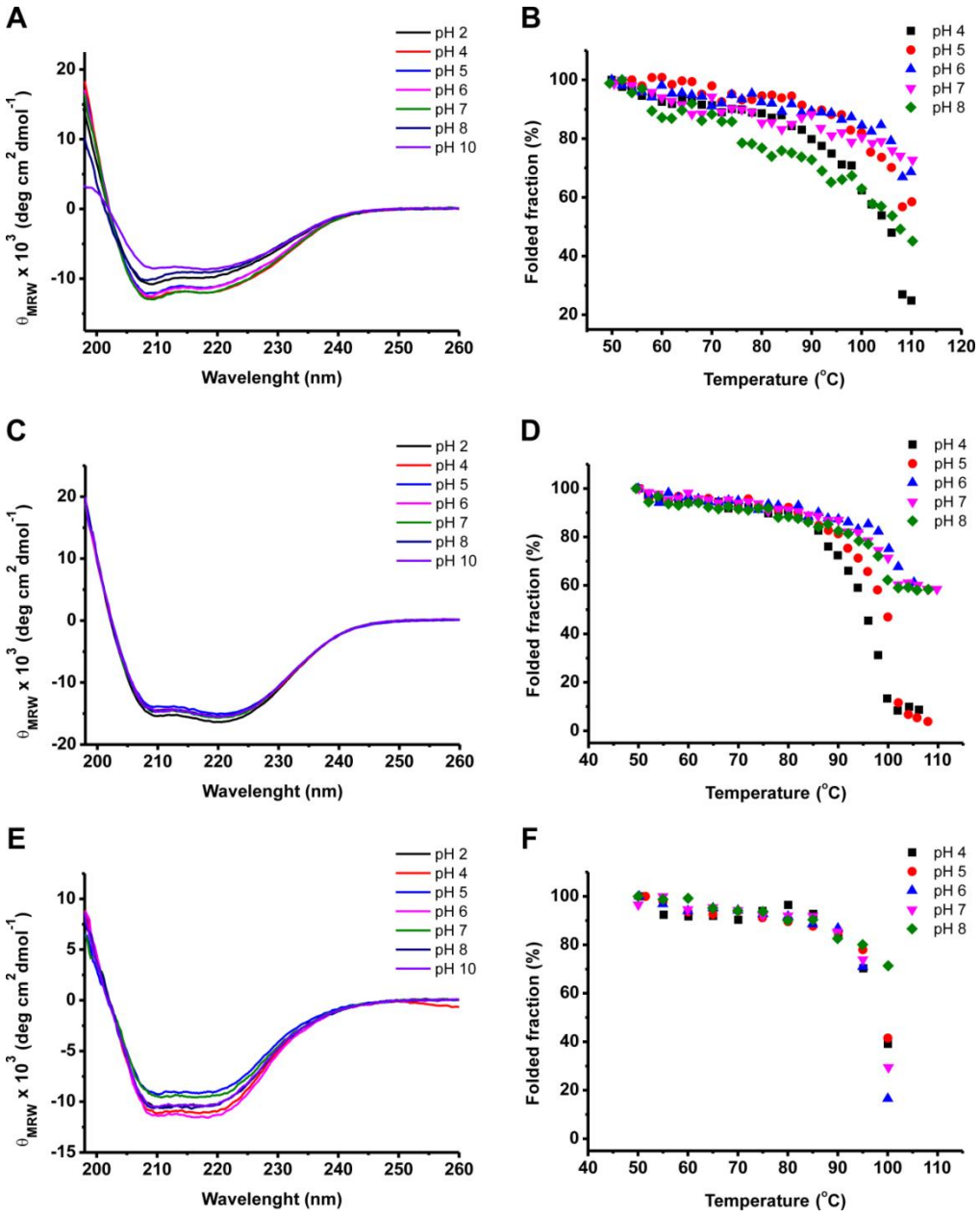


Figure 4.4. Characterization of BGL by circular dichroism. Far-UV CD spectra and thermal denaturation in different pH levels for PfBgl1 (A, B), TpBgl1 (C, D), and TpBgl3 (E, F). The circular dichroism spectrum represents an average of eight scans. The thermal denaturation was obtained by monitoring at 222 nm.

The profile of thermal denaturation curve demonstrates that PfBgl1 is very stable in higher temperatures. At pH 4 and pH 8, the enzyme started to lose the secondary structure content in lower temperatures that when in other pHs. Specifically, PfBgl1 was very sensible in pH 8. In the other tested pH, the secondary structure is preserved at optimum temperature (Fig. 4.4B). TpBgl1 was not as stable at higher temperatures as PfBgl1 (Fig. 4.4B, 4.4D), the protein was sensible to temperatures higher than 95 °C. Moreover, acidic environment greatly affected the protein stability (Fig. 4.4D). At pH 4 and pH 5, TpBgl1 drastically lose structure after 85 °C with 100% of denaturation at 100 °C. TpBgl3 presented high susceptibility to pH at higher temperatures (Fig 4.4F). At optimum activity temperature, the protein was only structured at pH 8. At the other pH values, low secondary structure content was observed at 100 °C.

4.4 DISCUSSION

BGLs are an important class of glycosyl hydrolases that can work in a wide and diverse range of substrates, and they can be found in all domains of living organisms [31]. Because of their diverse functions, they can be applied in several processes and are assigned in different E.C. numbers. Regarding their amino acid sequence and structural similarity classification system (CAZY) they are allocated in two main families, GH1 and GH3 [8, 32, 33]. Here, we have studied two β-glucosidases from GH1 family and one from GH3. Only PfBgl1 (PDB ID: 3APG) has its tridimensional structure solved [34] and placed into the Protein Data Bank (PDB), however there are homologous proteins with high identity in PDB for TpBgl1 (PDB ID: 1OD0) [35] and TpBgl3 (PDB ID: 2X40) [36] with 99 and 88% of identity, respectively.

The target enzymes of this work are hyperthermophilic, showing the highest activities at temperature ranges above 80 °C (Fig. 4.1). However, the GH1 BGLs seem to operate at a wider range of pH than GH3 (Fig. 4.1, 4.2A). The outcomes from CCRD experiments are in agreement to the pH and temperature curves reported for homologous enzymes, like non-recombinant PfBgl1 [37]. It has been demonstrated that the optimal pH for catalysis of recombinant TpBgl1 is around 8.0 [38], however our results from two distinct methodologies have supported that the best pH is 6.0, and the residual activity on pH 8.0 is near to 60%. The BGL with the highest identity to TpBgl3 in PDB is from *Thermotoga neapolitana* [39] and it presented similar temperature profile, although the optimal pH differed more than in one unit.

The assessment of enzymatic activities of BGLs after one week of incubation in buffers with different pHs indicate a great stability to pH variations, since there were no significant shifts in the pH range from 4 to 10 (Fig. 4.2B). The extreme acid treatment (pH 2) lead to a full catalytic function loss in GH1 enzymes, but TpBgl3 still retrieved 47% of initial activity which means that its catalytic pocket was not completely disorganized. Besides pH-stability, another fundamental feature to allow enzymatic technical applications is thermostability. The highest time of half-life (carried out at 99 °C) was found at pH 6.0 for the three β-glucosidases (Table 4.3), suggesting that at pH 6.0 there is the best charge equilibrium for the tertiary structure stabilization, which may differs from the charge balance for the catalysis. PfBgl1 presented a $t_{1/2}$ at least 10 times higher than BGLs from *T. petrophila*, probably because of its tetrameric organization where there is a very strong interface between monomers A and C stable even in SDS-PAGE gel [34].

Furthermore, the analysis on the effects of ions in pNPGase activity showed that the BGLs are plenty resistant to ions interference, although TpBgl1 is more susceptible than PfBgl1 and TpBgl3 (Table 4.4). The addition of EDTA did not change significantly the catalytic activity of BGLs, and it supports that the enzymes studied do not require any divalent ion to be functional. It is expected since BGLs are known to not be ion dependent [31, 40, 41]. Several metal ions have been described to not inhibit β-glucosidases, although Ag^{+}, Hg^{2+}, Cu^{2+}, Fe^{2+} and Fe^{3+} has been reported to be potent inhibitors [15, 31, 42]. Nevertheless the PfBgl1 and TpBgl3 seem to be very robust enzymes since they were Fe^{2+} and Hg^{2+} resistant. Another bacterial GH1 BGL from *Thermoanaerobacterium thermosaccharolyticum* has been identified to be Fe^{2+} tolerant [41].

Our findings regarding the specific activities towards the polysaccharides indicated that the BGLs have different substrates specificities. As assigned by the manufacturer Megazyme®, laminarim, lichenan, and β-glucan are polymers composed of glucose, but have distinct linkages, like β-(1,3) in backbone and β-(1,6) in branches for laminarin, and β-(1,3) and β-(1,4) alternating in backbone for lichenan and β-glucan, although not in the same frequency. The diversity in bonds specificity among the enzymes may be due to differences in subsites affinity for substrates, as it has been reported [31, 40, 43]. Reviewing the catalytic activities in cellobiose, lactose and the pNP derived substrates, it becomes evident that TpBgl3 is an extremely specific β-glucosidase as it is known for this family GH3 [44], while the PfBgl1 and TpBgl1 are more promiscuous enzymes, which has been described for GH1 family [15, 40, 45]. Interestingly the specific activity in pNP-fucopyranoside may be higher than in pNPG for some GH1 BGLs like we found for PfBgl1 [40].

Bathia et al. [7] has grouped the BGLs in three main classes based on substrate specificity: (i) aryl β-glucosidases, which act on aryl-glucosides, (ii) true cellobiases, which hydrolyze cellobiose to release glucose, and (iii) broad substrate specificity β-glucosidases, which act on a wide spectrum of substrates. Regarding the pNPGase activity, the TpBgl3 was the most efficient among the β-glucosidases, exhibiting the highest k_{cat} and k_{cat}/K_m (Table 4.6), which is in agreement to the great substrate specificity of that enzyme. The K_m and k_{cat}/K_m values of the hyperthermophilic BGLs in pNPG and cellobiose were lower than those reported in the literature [15, 40, 46, 47]. The inhibition constant of glucose accessed for TpBgl1 was the highest among the BGLs studied or even those reported up to date [15, 45, 48], while the K_i for Pf Bgl1 is similar to the most of those disclosed. Although TpBgl3 had a K_i of 30.10 mM, this is a value greater than those published for GH3 BGLs [49]. The curves for the effects of monosaccharides in reaction suggest that the enzymes support high concentrations (up to 1 M) of those reducing sugars. TpBgl1 exhibited a different behavior like an allosteric regulation, and there was an apparent stimulation by some saccharides. Adding xylose up to 1M in the reaction, TpBgl1 reached a maximum increase of about 130% in catalytic activity.

The Far-UV CD spectra of BGLs support that there is no significant shifts on secondary structure in extremely acid environment (pH 2), despite of functional experiments had showed the BGLs suffer some tridimensional deformation, since they lost partially or even fully catalytic activity after that. Among studied β-glucosidases, TpBgl3 presented the highest susceptibility to pH at higher temperatures (Fig. 4.4), although that was the unique enzyme that supported pH ranging from 2 to 10 and without fully losing its activity. Based on time of half-life (Table 4.3) and CD (Fig. 4.4), the PfBgl1 seems to be the most stable BGL, indeed exhibiting thermostability many times higher than other two BGLs from *T. petrophila*.

In conclusion, the β-glucosidases studied seem to have a great potential for industrial uses, especially those from GH1 family because of its broad substrate specificity and weakly inhibition by glucose and other saccharides. Indeed it is necessary further studies on scale up to employ hyperthermophilic BGLs in industrial processes. Nevertheless, this work opens the horizons for future experiments, in order to improve any specific feature of BGLs according to the application demand.

ACKNOWLEDGEMENTS

This research was supported by grants from FAPESP (2008/58037-9) and CNPq (475022/2011-4 and 310177/2011-1). JC received a scholarship from CNPq (140420/2009-6. All data were collected at the Brazilian Bioethanol Science and Technology (CTBE) that integrates the Center of Research in Energy and Material (CNPEM). The authors would like to thank the CNPEM for support.

REFERENCES

[1] R.S. Makkar, S.S. Cameotra, An update on the use of unconventional substrates for biosurfactant production and their new applications, Applied Microbiology and Biotechnology, 58 (2002) 428-434.

[2] L.H. Quan, J.W. Min, D.U. Yang, Y.J. Kim, D.C. Yang, Enzymatic biotransformation of ginsenoside Rb1 to 20(S)-Rg3 by recombinant beta-glucosidase from *Microbacterium esteraromaticum*, Applied Microbiology and Biotechnology, 94 (2012) 377-384.

[3] V. Gopalan, A. Pastuszyn, W.R.J. Galey, R.H. Glew, Exolytic hydrolysis of toxic plant glucosides by guinea pig liver cytosolic glucosidases, The Journal of Biological Chemistry, 267 (1992) 14027-14032.

[4] A. Gallifuoco, L. D'Ercole, F. Alfani, M. Cantarella, G. Spagna, P.G. Pifferi, On the use of chitosan-immobilized beta-glucosidase in wine-making: kinetics and enzyme inhibition, Process Biochem, 33 (1998) 163-168.

[5] D. Sternberg, P. Vijayakumar, E.T. Reese, beta-Glucosidase: microbial production and effect on enzymatic hydrolysis of cellulose, Can J Microbiol, 23 (1977) 139-147.

[6] L.R. Lynd, P.J. Weimer, W.H. van Zyl, I.S. Pretorius, Microbial Cellulose Utilization: Fundamentals and Biotechnology, Microbiology and Molecular Biology Reviews, 66 (2002) 506-577.

[7] Y. Bhatia, S. Mishra, V.S. Bisaria, Microbial β-glucosidases: cloning, properties, and applications, Critical Reviews in Biotechnology, 22 (2002) 375-407.

[8] B. Henrissat, A classification of glycosyl hydrolases based on amino acid sequence similarities, Biochemical Journal, 280 (1991) 309-316.

[9] B. Henrissat, A. Bairoch, Updating the sequence-based classification of glycosyl hydrolases, Biochemical Journal, 316 (1996) 695-696.

[10] C. Ferreira, S.R. Marana, C. Silva, W.R. Terra, Properties of digestive glycosidases and peptidases and the permeability of the peritrophic membranes of *Abracris flavolineata* (Orthoptera Acrididae), Comparative Biochemistry and Physiology Part B, 123 (1999) 241-250.

[11] S.R. Marana, W.R. Terra, C. Ferreira, Purification and properties of a β-glycosidase purified from midgut cells of *Spodoptera frugiperda* (Lepidoptera) larvae, Insect Biochemistry and Molecular Biology, 30 (2000) 1139-1146.

[12] J. Hong, H. Tamaki, H. Kumagai, Cloning and functional expression of thermostable beta-glucosidase gene from *Thermoascus aurantiacus*, Applied Microbiology and Biotechnology, 73 (2007) 1331-1339.

[13] R.L. Mach, B. Seiboth, A. Myasnikov, R. Gonzalez, J. Strauss, A.M. Harkki, C.P. Kubicek, The bgl1 gene of *Trichoderma reesei* QM 9414 encodes an extracellular, cellulose-inducible β-glucosidase involved in cellulase induction by sophorose, Molecular Microbiology, 16 (1995) 687-697.

[14] B.C. Saha, R.J. Bothast, Glucose tolerant and thermophilic β-glucosidases from yeasts, Biotechnology Letters, 18 (1996) 155-158.

[15] N. Ait, N. Creuzet, J. Cattaneo, Properties of β-glucosidase purified from *Clostridium thermocellum*, Journal of General Microbiology, 128 (1982) 569-577.

[16] M.W. Bauer, R.M. Kelly, The family 1 β-glucosidases from *Pyrococcus furiosus* and *Agrobacterium faecalis* share a common catalytic mechanism, Biochemistry, 37 (1998) 17170-17178.

[17] Z. Fang, Cloning and characterization of a β-glucosidase from marine microbial metagenome with excellent glucose tolerance, Journal of Microbiology and Biotechnology, 20 (2010) 1351-1358.

[18] T. Robinson, B. Chandran, P. Nigam, Effect of pretreatments of three waste residues, wheat straw, corncobs and barley husks on dye adsorption, Bioresource Technology, 85 (2002) 119-124.

[19] M. Laser, D. Schulman, S.G. Allen, J. Lichwa, M.J. Antral Jr., L.R. Lynd, A comparison of liquid hot water and steam pretreatments of sugar cane bagasse for bioconversion to ethanol, Bioresource Technology, 81 (2002) 33-44.

[20] M.H. Studer, S. Brethauer, J.D. Demartini, H.L. McKenzie, C.E. Wyman, Co-hydrolysis of hydrothermal and dilute acid pretreated populus slurries to support development of a high-throughput pretreatment system, Biotechnology for biofuels, 4 (2011) 19.

[21] H. Wang, F.M. Squina, F. Segato, A. Mort, D. Lee, K. Pappan, R.A. Prade, High temperature enzymatic cellulose breakdown, Applied and Environmental Microbiology, 77 (2011) 5199-5206.

[22] I. Petzelbauer, B. Splechtna, B. Nidetzky, Development of an ultrahigh-temperature process for the enzymatic hydrolysis of lactose. III. Utilization of two thermostable beta-glycosidases in a continuous ultrafiltration membrane reactor and galacto-oligosaccharide formation under steady-state conditions, Biotechnol Bioeng, 77 (2002) 394-404.

[23] I. Petzelbauer, R. Zeleny, A. Reiter, K.D. Kulbe, B. Nidetzky, Development of an ultra-high-temperature process for the enzymatic hydrolysis of lactose: II. Oligosaccharide formation by two thermostable beta-glycosidases, Biotechnol Bioeng, 69 (2000) 140-149.

[24] J. Woodward, A. Wiseman, Fungal and other β-D-glucosidases — Their properties and applications, Enzyme and Microbial Technology, 4 (1982) 73-79.

[25] F.M. Squina, R.A. Prade, H. Wang, M.T. Murakami, Expression, purification, crystallization and preliminary crystallographic analysis of an endo-1,5-alpha-L-arabinanase from hyperthermophilic *Thermotoga petrophila*, Acta crystallographica. Section F, Structural biology and crystallization communications, 65 (2009) 902-905.

[26] G.L. Miller, Use of dinitrosalicylic acid reagent for determination of reducing sugar, Analytical Chemistry, 31 (1959) 426-428.

[27] R.H. Myers, D.C. Montgomery, C.M. Anderson-Cook, Response surface methodology: process and product optimization using designed experiments, Wiley, 2009.

[28] E. Aguiar-Oliveira, F. Maugeri, Thermal stability of the immobilized fructosyltransferase from *Rhodotorula* sp., Braz J Chem Eng, 28 (2011) 363-372.

[29] J. Cota, T.M. Alvarez, A.P. Citadini, C.R. Santos, M. de Oliveira Neto, R.R. Oliveira, G.M. Pastore, R. Ruller, R.A. Prade, M.T. Murakami, F.M. Squina, Mode of operation and low-resolution structure of a multi-domain and hyperthermophilic endo-beta-1,3-glucanase from *Thermotoga petrophila*, Biochemical and biophysical research communications, 406 (2011) 590-594.

[30] R. Naran, M.L. Pierce, A.J. Mort, Detection and identification of rhamnogalacturonan lyase activity in intercellular spaces of expanding cotton cotyledons, The Plant Journal: For Cell and Molecular Biology, 50 (2007) 95-107.

[31] J.R. Ketudat Cairns, A. Esen, beta-Glucosidases, Cellular and molecular life sciences : CMLS, 67 (2010) 3389-3405.

[32] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics, Nucleic Acids Research, 37 (2009) D233-238.

[33] B. Henrissat, G. Davies, Structural and sequence-based classification of glycoside hydrolases, Current Opinion in Structural Biology, 7 (1997) 637-644.

[34] Y. Kado, T. Inoue, K. Ishikawa, Structure of hyperthermophilic beta-glucosidase from *Pyrococcus furiosus*, Acta crystallographica. Section F, Structural biology and crystallization communications, 67 (2011) 1473-1479.

[35] D.L. Zechel, A.B. Boraston, T. Gloster, C.M. Boraston, J.M. Macdonald, M.G. Tilbrook, R.V. Stick, G.J. Davies, Iminosugar glycosidase inhibitors: structural and thermodynamic dissection of the binding of isofagomine and 1-deoxynojirimycin to β-glucosidases, Journal of American Chemical Society, 125 (2003) 14313-14323.

[36] T. Pozzo, J.L. Pasten, E.N. Karlsson, D.T. Logan, Structural and functional analyses of beta-glucosidase 3B from *Thermotoga neapolitana*: a thermostable three-domain representative of glycoside hydrolase 3, Journal of Molecular Biology, 397 (2010) 724-739.

[37] S.W.M. Kengen, E.J. Luesink, A.J.M. Stams, A.J.B. Zehnder, Purification and characterization of an extremely thermostable β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*, European Journal of Biochemistry, 213 (1993) 305-312.

[38] I.U. Haq, M.A. Khan, B. Muneer, Z. Hussain, S. Afzal, S. Majeed, N. Rashid, M.M. Javed, I. Ahmad, Cloning, characterization and molecular docking of a highly thermostable beta-1,4-glucosidase from *Thermotoga petrophila*, Biotechnol Lett, 34 (2012) 1703-1709.

[39] P. Turner, D. Svensson, P. Adlercreutz, E.N. Karlsson, A novel variant of Thermotoga neapolitana beta-glucosidase B is an efficient catalyst for the synthesis of alkyl glucosides by transglycosylation, Journal of Biotechnology, 130 (2007) 67-74.

[40] M.R. Hong, Y.S. Kim, C.S. Park, J.K. Lee, Y.S. Kim, D.K. Oh, Characterization of a recombinant beta-glucosidase from the thermophilic bacterium *Caldicellulosiruptor saccharolyticus*, Journal of Bioscience and Bioengineering, 108 (2009) 36-40.

[41] J. Pei, Q. Pang, L. Zhao, S. Fan, H. Shi, *Thermoanaerobacterium thermosaccharolyticum* beta-glucosidase: a glucose-tolerant enzyme with high specific activity for cellobiose, Biotechnology for Biofuels, 5 (2012) 31.

[42] E.M. Bowers, L.O. Ragland, L.D. Byers, Salt effects on beta-glucosidase: pH-profile narrowing, Biochimica et Biophysica Acta, 1774 (2007) 1500-1507.
[43] S. Badieyan, D.R. Bevan, C. Zhang, Probing the active site chemistry of beta-glucosidases along the hydrolysis reaction pathway, Biochemistry, 51 (2012) 8907-8918.
[44] Y.K. Li, J. Chir, F.Y. Chen, Catalytic mechanism of a family 3 β-glucosidase and mutagenesis study on residue Asp-247, Biochemical Journal, 355 (2001) 835-840.
[45] R. Breves, K. Bronnenmeier, N. Wild, F. Lottspeich, W.L. Staudenbauer, J. Hofemeister, Genes encoding two different β-glucosidases of *Thermoanaerobacter brockii* are clustered in a common operon, Applied and Environmental Microbiology, 63 (1997) 3902-3910.
[46] Y. Xue, X. Song, J. Yu, Overexpression of β-glucosidase from *Thermotoga maritima* for the production of highly purified aglycone isoflavones from soy flour, World Journal of Microbiology and Biotechnology, 25 (2009) 2165-2172.
[47] V.V. Zverlov, I.Y. Volkov, T.V. Velikodvorskaya, W.H. Schwarz, *Thermotoga neapolitana* bgIB gene, upstream of lamA, encodes a highly thermostable β-glucosidase that is a laminaribiase, Microbiology, 143 (1997) 3537-3542.
[48] D. Jabbour, B. Klippel, G. Antranikian, A novel thermostable and glucose-tolerant beta-glucosidase from *Fervidobacterium islandicum*, Applied microbiology and biotechnology, 93 (2012) 1947-1956.
[49] C. Bohlin, S.N. Olsen, M.D. Morant, S. Patkar, K. Borch, P. Westh, A comparative study of activity and apparent inhibition of fungal beta-glucosidases, Biotechnol Bioeng, 107 (2010) 943-952.

SUPPLEMENTARY INFORMATION

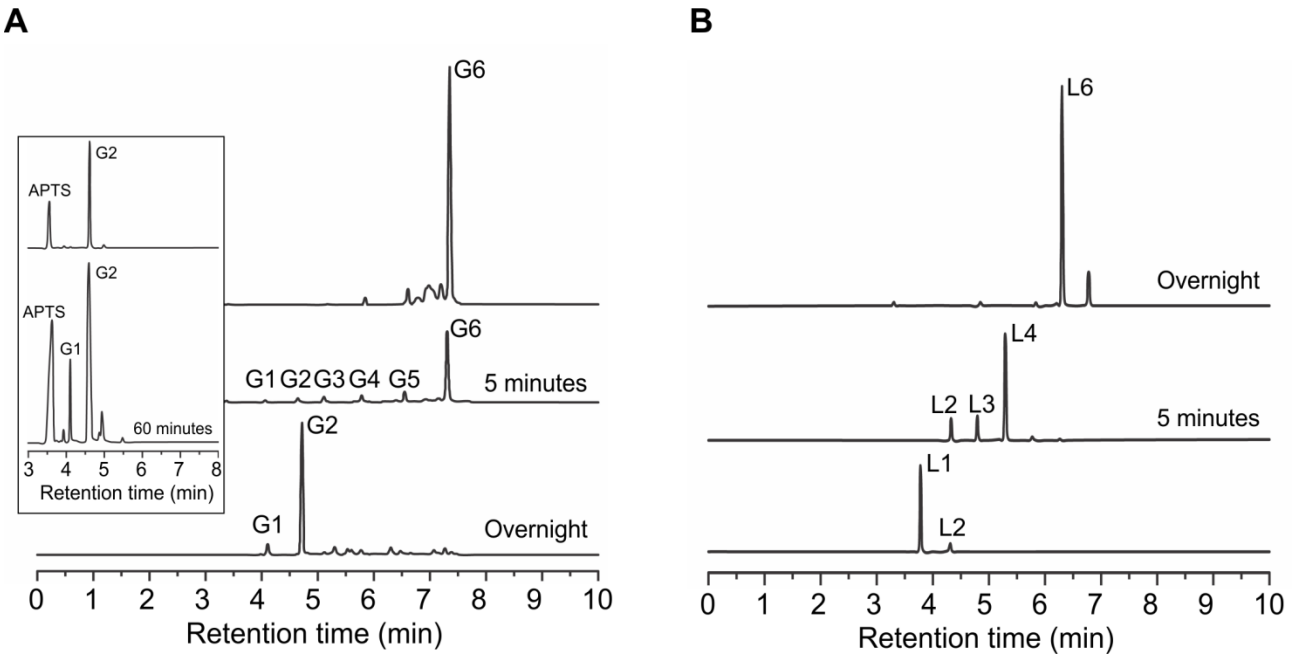


Figure SI-4.1. Capillary zone electrophoresis of APTS-labeled oligosaccharides. (A) Incomplete and complete hydrolysis of APTS-reducing-end-labeled-cellohexaose and cellobiose APTS-labeled after the enzymatic reaction (left box). (B) Incomplete and complete hydrolysis of APTS-reducing-end-labeled-laminarihexaose. G1, G2, G3, G4, G5, G6, L1, L2, L3, L4, and L6 indicate the degree of polymerization of glucose oligomers. This is a typical pattern found for the three hyperthermophilic BGLs.

CAPÍTULO 5

CONCLUSÕES GERAIS E PERSPECTIVAS

Conclusões gerais e sugestões de trabalhos futuros

A TpLam é uma enzima do tipo endo-β-1,3-glucanase com propriedades termoestáveis marcantes. Este estudo apresentou o modo de ataque enzimático, realizou uma abrangente análise de espectrometria de hipertermofilicidade e determinou a estrutura de baixa resolução de um multi-domínio do tipo *V-shape*, em que a enzima apresenta dois domínios de ligação a carboidratos. Nossos resultados fornecem bases bioquímicas e estruturais para estudos posteriores de endo-β-1 ,3-glucanases, que são componentes importantes do repertório enzimático na degradação de polissacarídeos, podendo ter muitas aplicações biotecnológicas. Esse trabalho fornece as bases para que sejam realizados futuramente ensaios de aplicação, nos quais coquetéis comerciais podem ser suplementados com a enzima TpLam visando a aumentar o rendimento na sacarificação de biomassa rica em β-glucanos, como por exemplo o bagaço de cana-de-açúcar. Além disso, a TpLam é uma enzima que, devido a sua grande estabilidade térmica, possui um enorme potencial para ser aplicada em vários processos na indústria de alimentos, como por exemplo na produção de oligossacarídeos a partir de β-glucanos.

A engenharia de proteínas fusionadas tem sido considerada interessante para processos industriais, porque estas proteínas são mais eficientes (com base nos custos e na eficiência catalítica) que as proteínas nativas. Juntamente com a falta de alterações substanciais no desempenho catalítico da quimera catalítica, o parâmetro de rendimento AY sugeriu uma vantagem em produzir a proteína quimérica, em vez de produzir as proteínas isoladamente. Certamente mais estudos sobre as propriedades da quimera são necessários antes de se fazer um escalonamento de processo. Além do mais, trabalhos futuros com bagaço de cana-de-açúcar e outras biomassas vegetais devem ser realizados utilizando as proteínas isoladas e quimérica, buscando entender se há sinergia na ação dessas duas enzimas e se houve algum ganho nesse sentido com o uso da proteína fusionada.

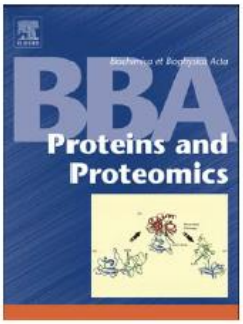
O trabalho de engenharia enzimática apresentou uma nova abordagem para a predição da disposição de domínios quiméricos em solução, antes mesmo da validação experimental. A estratégia computacional aqui proposta é rápida e robusta para a caracterização de um grande número de enzimas dentro de algumas semanas, bem como para a identificação de possíveis obstruções do local de ligação ou grandes variações dinâmicas que diferem das proteínas nativas. Esses resultados devem ajudar na busca por novas abordagens de baixo custo para a conversão de

biomassa vegetal em produtos biotecnológicos. Um dos grandes desafios na sacarificação de biomassa é diminuir o custo da produção da enzima. Portanto, o uso das hidrolases multifuncionais, que atuem em sinergia na degradação dos polissacarídeos de planta, é uma estratégia promissora para a melhoria de coquetéis enzimáticos para biocombustíveis de segunda geração. A metodologia utilizada nesse trabalho pode ser estendida a várias quimeras multi-domínios e, da mesma forma, para uma grande variedade de sistemas biológicos. Assim, em trabalhos futuros, essa ferramenta pode ser consolidada como uma estratégia de *screening* de possíveis e melhores proteínas quiméricas, resultando em economia de tempo e de recursos.

As β-glucosidases estudadas parecem ter um grande potencial para usos industriais, especialmente aquelas da família GH1, por causa de sua ampla especificidade de substratos e por causa de serem fracamente inibidas por glicose e outros sacarídeos. Além disso, as BGLs se mostraram enzimas que possuem alta estabilidade à temperatura e pH, e também a diversos íons. De fato, são necessários estudos aprofundados sobre o escalonamento e aplicação dessas BGLs hipertermofílicas em processos industriais. No entanto, este trabalho abre os horizontes para futuros experimentos de engenharia enzimática, de modo a melhorar alguma característica específica das BGLs de acordo com a demanda do processo. Além disso, devido às características acima reportadas, as BGLs podem ser futuramente estudadas em escala industrial visando aplicações tanto na sacarificação de biomassa para produção de combustíveis de segunda geração, bem como em diversos processos na indústria de alimentos, como por exemplo na vinificação.

ANEXO

TRABALHOS PUBLICADOS E COLABORAÇÕES



Assembling a xylanase-lichenase chimera through all-atom molecular dynamics simulations

Junio Cota, Leandro C. Oliveira, André R.L. Damásio, Ana P. Citadini, Zaira B. Hoffmam, Thabata M. Alvarez, Carla A. Codima, Vitor B.P. Leite, Glaucia Pastore, Mario de Oliveira-Neto, Mario T. Murakami, Roberto Ruller, Fabio M. Squina

PII: S1570-9639(13)00093-9
DOI: doi: 10.1016/j.bbapap.2013.02.030
Reference: BBAPAP 39043

To appear in: *BBA - Proteins and Proteomics*

Received date: 8 October 2012
Revised date: 23 January 2013
Accepted date: 20 February 2013

Please cite this article as: Junio Cota, Leandro C. Oliveira, André R.L. Damásio, Ana P. Citadini, Zaira B. Hoffmam, Thabata M. Alvarez, Carla A. Codima, Vitor B.P. Leite, Glaucia Pastore, Mario de Oliveira-Neto, Mario T. Murakami, Roberto Ruller, Fabio M. Squina, Assembling a xylanase-lichenase chimera through all-atom molecular dynamics simulations, *BBA - Proteins and Proteomics* (2013), doi: 10.1016/j.bbapap.2013.02.030

Biochemical and Biophysical Research Communications 406 (2011) 590–594

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Mode of operation and low-resolution structure of a multi-domain and hyperthermophilic endo-β-1,3-glucanase from *Thermotoga petrophila*

Junio Cota [a,d], Thabata M. Alvarez [a], Ana P. Citadini [a], Camila Ramos Santos [b], Mario de Oliveira Neto [c], Renata R. Oliveira [b], Glaucia M. Pastore [d], Roberto Ruller [a], Rolf A. Prade [e], Mario T. Murakami [b], Fabio M. Squina [a,*]

[a] Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), do Centro Nacional de Pesquisa em Energia e Materiais, Campinas SP, Brazil
[b] e Laboratório Nacional de Biociências (LNBio), do Centro Nacional de Pesquisa em Energia e Materiais, Campinas SP, Brazil
[c] Instituto de Física de São Carlos (IFSC), Universidade de São Paulo, São Carlos SP, Brazil
[d] Faculdade de Engenharia de Alimentos (FEA), Universidade Estadual de Campinas, Campinas SP, Brazil
[e] Department of Microbiology and Molecular Genetics, Oklahoma State University (OSU), Stillwater, OK, USA

ARTICLE INFO

Article history:
Received 16 February 2011
Available online 23 February 2011

Keywords:
Endo-β-1,3-glucanase
Glycoside hydrolase family 16
Hyperthermostable enzyme
Capillary zone electrophoresis
Small angle X-ray scattering

ABSTRACT

1,3-β-Glucan depolymerizing enzymes have considerable biotechnological applications including biofuel production, feedstock-chemicals and pharmaceuticals. Here we describe a comprehensive functional characterization and low-resolution structure of a hyperthermophilic laminarinase from *Thermotoga petrophila* (TpLam). We determine TpLam enzymatic mode of operation, which specifically cleaves internal β-1,3-glucosidic bonds. The enzyme most frequently attacks the bond between the 3rd and 4th residue from the non-reducing end, producing glucose, laminaribiose and laminaritriose as major products. Far-UV circular dichroism demonstrates that TpLam is formed mainly by beta structural elements, and the secondary structure is maintained after incubation at 90 °C. The structure resolved by small angle X-ray scattering, reveals a multi-domain structural architecture of a V-shape envelope with a catalytic domain flanked by two carbohydrate-binding modules.

1. Introduction

Renewable structural polysaccharides are abundant resources in the biosphere and represent a valuable industrial substrate for many industrial processes, such as biofuels production, feedstock-chemicals and pharmaceuticals [1]. β-1,3-Glucans are important polymers found mainly in yeast and filamentous fungi and are produced by plants (as callose) in response to tissue damage. This polymer is a major structural storage polysaccharide (laminarin) of the marine brown macro-algae of the genus *Laminaria*. β-1,3-glucan is also produced as insoluble exopolysaccharide by some fungal and bacterial species [2–5].

The β-1,3-glucanases are widely distributed among bacteria, fungi, and plants and are generally classified into two types, exo-β-1,3-glucanases (EC 3.2.1.58) and endo-β-1,3-glucanases (EC 3.2.1.6 and 3.2.1.39), which hydrolyze terminal or internal glycoside linkages, respectively [28]. These enzymes are further classified as family members (GH) 16, 17, 55, 64, and 81 of the CaZy system, which is based on structural similarities of the amino acid sequence [28]. However, CaZy GH families display a dispersed range of enzymatic activity, and updating biochemical and biophysical information strengthens the Cazy pupose.

To date, crystal structures from family GH 16 endo-β-1,3-glucanases (3.2.1.39) have been determined for *Pyrococcus furiosus* [6], *Rhodothermus marinus* [6], *Streptomyces sioyaensis* [7], and *Nocardiopsis sp* [8]. The active site region of family GH16 consists of three acidic residues, two glutamic and one aspartic acid in the active site, and two conserved tryptophan residues, which promote substrate binding and transition-state stabilization [9]. The overall fold of these enzymes consists of a classical sandwich-like β-jelly-roll motif formed by the face-to-face packing of two ant parallel sheets containing seven and eight strands with a deep cavity [6–8,10].

In this work is reported a comprehensive functional characterization of an endo-β-1,3-glucanase from *Thermotoga petrophila* (TpLam), including a detailed description of the hydrolytic mode of operation, along with structural insights of this hyperthermostable enzyme.

* Corresponding author. Address: Rua Giuseppe Máximo Scolfaro, 10.000 Polo II de Alta Tecnologia, Caixa Postal 6170, Campinas, SP 13083-970, Brazil. Fax: +55 19 35183104.
E-mail address: fabio.squina@bioetanol.org.br (F.M. Squina).

0006-291X/$ - see front matter Crown Copyright © 2011 Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2011.02.098

Bioresource Technology 131 (2013) 500–507

Contents lists available at SciVerse ScienceDirect

Bioresource Technology

journal homepage: www.elsevier.com/locate/biortech

Understanding the cellulolytic system of *Trichoderma harzianum* P49P11 and enhancing saccharification of pretreated sugarcane bagasse by supplementation with pectinase and α-L-arabinofuranosidase

Priscila da Silva Delabona [a,*], Júnio Cota [a], Zaira Bruna Hoffmam [a], Douglas Antonio Alvaredo Paixão [a], Cristiane Sanchez Farinas [b], João Paulo Lourenço Franco Cairo [a], Deise Juliana Lima [a], Fábio Marcio Squina [a], Roberto Ruller [a], José Geraldo da Cruz Pradella [a]

[a] *Brazilian Bioethanol Science and Technology Laboratory – CTBE, Rua Giuseppe Maximo Scolfaro 10000, Pólo II de Alta Tecnologia, Caixa Postal 6192, CEP 13083-970, Campinas, São Paulo, Brazil*

[b] *Embrapa Instrumentation, Rua XV de Novembro 1452, CEP 13560-970, São Carlos, São Paulo, Brazil*

HIGHLIGHTS

- The Amazon rainforest *Trichoderma harzianum* is a potential candidate for saccharification of plant biomass.
- Important roles for pectinase and α-L-arabinofuranosidase in the enzymatic synergies.
- Sugarcane bagasse hydrolysis is more effective and competitive for second generation ethanol.

ARTICLE INFO

Article history:
Received 30 July 2012
Received in revised form 27 November 2012
Accepted 14 December 2012
Available online 22 December 2012

Keywords:
Accessory enzymes
Trichoderma harzianum
Pectinase
α-L-arabinofuranosidase
Enzymatic hydrolysis

ABSTRACT

Supplementation of cellulase cocktails with accessory enzymes can contribute to a higher hydrolytic capacity in releasing fermentable sugars from plant biomass. This study investigated which enzymes were complementary to the enzyme set of *Trichoderma harzianum* in the degradation of sugarcane bagasse. Specific activities of *T. harzianum* extract on different substrates were compared with the extracts of *Penicillium echinulatum* and *Trichoderma reesei*, and two commercial cellulase preparations. Complementary analysis of the secretome of *T. harzianum* was also used to identify which enzymes were produced during growth on pretreated sugarcane bagasse. These analyses enabled the selection of the enzymes pectinase and α-L-arabinofuranosidase (AF) to be further investigated as supplements to the *T. harzianum* extract. The effect of enzyme supplementation on the efficiency of sugarcane bagasse saccharification was evaluated using response surface methodology. The supplementation of *T. harzianum* enzymatic extract with pectinase and AF increased the efficiency of hydrolysis by up to 116%.

1. Introduction

The conversion of lignocellulosic biomass into fuels and other chemicals can be achieved using a multi-enzyme system acting in synergy. An enzymatic cocktail containing different enzymes involved in the hydrolysis of each part of the lignocellulosic structure is crucial to increase enzymatic hydrolysis (EH) yields (Alvira et al., 2011; Berlin et al., 2007; Gao et al., 2011). In addition, the viable bioconversion of biomass requires not only that the cost of the enzymes be reduced, but also that improvements be made in the effectiveness of the enzymatic extracts used in the hydrolysis mixture. These aspects are highly dependent on both the raw material and the pretreatment employed (Delabona et al., 2012a; Jorgensen and Olsson, 2006; Sorensen et al., 2011; Sukumaran et al., 2009). On-site production of enzymes can help to reduce enzyme costs since there is less need to stabilize the enzyme preparations, avoiding costs associated with transport and storage for long periods. Furthermore, the use of the same lignocellulosic biomass for enzyme production and hydrolysis could reduce the production costs of bioethanol, since both process could be co-located and share infrastructure and utilities (Delabona et al., 2012a). The on-site enzyme production using a split stream from the bioethanol process as part of the fermentation medium could be an attractive alternative (Kovacs et al., 2009; Sorensen et al., 2011). Moreover, the production of enzymes using as carbon source the same lignocellulosic material that will be used in the hydrolysis has shown that these enzyme preparations can present better performance (Delabona et al., 2012a; Jorgensen and Olsson, 2006).

* Corresponding author. Tel.: +55 (19) 3518 3154; fax: +55 (19) 3518 3104.
E-mail address: priscila.delabona@bioetanol.org.br (P.S. Delabona).

0960-8524/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.biortech.2012.12.105

OPEN ACCESS Freely available online

PLOS ONE

The *Penicillium echinulatum* Secretome on Sugar Cane Bagasse

Daniela A. Ribeiro[1], Júnio Cota[1], Thabata M. Alvarez[1], Fernanda Brüchli[1], Juliano Bragato[1], Beatriz M. P. Pereira[1], Bianca A. Pauletti[1], George Jackson[1], Maria T. B. Pimenta[1], Mario T. Murakami[2], Marli Camassola[3], Roberto Ruller[1], Aldo J. P. Dillon[3], Jose G. C. Pradella[1], Adriana F. Paes Leme[1], Fabio M. Squina[1]*

1 Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, (CNPEM), Campinas, São Paulo, Brazil, **2** Laboratório de Espectrometria de Massas, Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, (CNPEM), Campinas, São Paulo, Brazil, **3** Instituto de Biotecnologia, Universidade de Caxias do Sul (UCS), Caxias do Sul, Rio Grande do Sol, Brazil

Abstract

Plant feedstocks are at the leading front of the biofuel industry based on the potential to promote economical, social and environmental development worldwide through sustainable scenarios related to energy production. *Penicillium echinulatum* is a promising strain for the bioethanol industry based on its capacity to produce large amounts of cellulases at low cost. The secretome profile of *P. echinulatum* after grown on integral sugarcane bagasse, microcrystalline cellulose and three types of pretreated sugarcane bagasse was evaluated using shotgun proteomics. The comprehensive chemical characterization of the biomass used as the source of fungal nutrition, as well as biochemical activity assays using a collection of natural polysaccharides, were also performed. Our study revealed that the enzymatic repertoire of *P. echinulatum* is geared mainly toward producing enzymes from the cellulose complex (endogluganases, cellobiohydrolases and β-glucosidases). Glycoside hydrolase (GH) family members, important to biomass-to-biofuels conversion strategies, were identified, including endoglucanases GH5, 7, 6, 12, 17 and 61, β-glycosidase GH3, xylanases GH10 and GH11, as well as debranching hemicellulases from GH43, GH62 and CE2 and pectinanes from GH28. Collectively, the approach conducted in this study gave new insights on the better comprehension of the composition and degradation capability of an industrial cellulolytic strain, from which a number of applied technologies, such as biofuel production, can be generated.

Citation: Ribeiro DA, Cota J, Alvarez TM, Brüchli F, Bragato J, et al. (2012) The *Penicillium echinulatum* Secretome on Sugar Cane Bagasse. PLoS ONE 7(12): e50571. doi:10.1371/journal.pone.0050571

Editor: Gustavo Henrique Goldman, Universidade de Sao Paulo, Brazil

Received July 27, 2012; **Accepted** October 23, 2012; **Published** December 5, 2012

Funding: This research was supported by grants from FAPESP (2008/58037-9) and CNPq (475022/2011-4 e 310177/2011-1). TMA received a fellowship from FAPESP (2010/11499-1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: fabio.squina@bioetanol.org.br

Introduction

Plant structural polysaccharides are the most abundant and renewable biomass in the biosphere. Plant feedstocks are at the leading front of the biofuel industry based on the potential to promote economical, social and environmental development worldwide through sustainable scenarios related to energy production [1]. However, saccharification and bioproduct manufacturing from lignocellulose biomass are complex and lengthy processes. The current schemes for the biotechnological conversion of plant cell wall polysaccharides rely on first reducing biomass recalcitrance through a pretreatment step, and afterward, enzymatic cocktails are needed to breakdown biomass into more simple, fermentable saccharides, which could be fed into several bioprocesses, such as bioethanol production.

Despite the advantages of enzyme-catalyzed processes, *i.e.*, speed, specificity and mildness, the high cost of enzyme production and low catalytic efficiency are still major hindrances for cellulosic bioethanol. Thus, relevant biotechnological challenges in this field include the improvement of the catalytic efficiency of enzymes, the economic benefit and the synergy between the type of pretreatment and enzymatic load, and reduction of the cost of enzyme production by filamentous fungi [2,3].

The enzymatic degradation of cellulosic materials by fungal enzyme systems, especially those produced by *Tricoderma reesei* and *Aspergillus* spp., has been extensively studied due to its effectiveness for the liberation of fermentable sugars for bioethanol production [4,5]. Several efforts, such as optimization of fermentation processes [6, 7, and 8] and genetic modifications of the microorganisms [9, 10, 11, and 12], are being targeted for improvement of fungal enzymatic systems. More recently, *Penicillium echinulatum* has been the focus of attention due to its potential to produce large amounts of cellulases at low costs; it has also been considered a promising strain for the bioethanol industry [7, 8, 9, 10, 13 and 14].

Advances in proteomics analysis have pushed forward secretome studies on filamentous fungi [15, 16, 17 and 18], mainly because they can highlight pathways as well as target genes for deletions or co-expression, to improve strains for biotechnological purposes [19]. In addition, secretome studies are appealing for basic research regarding the role of filamentous fungi not only as ubiquitous saprophytes in nature but also as cell factories to secrete

Journal of Structural Biology 177 (2012) 469–476

Contents lists available at SciVerse ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Molecular insights into substrate specificity and thermal stability of a bacterial GH5-CBM27 endo-1,4-β-D-mannanase

Camila Ramos dos Santos [a], Joice Helena Paiva [a], Andreia Navarro Meza [a], Junio Cota [b], Thabata Maria Alvarez [b], Roberto Ruller [b], Rolf Alexander Prade [c], Fabio Marcio Squina [b], Mario Tyago Murakami [a,*]

[a] Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, SP, Brazil
[b] Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, SP, Brazil
[c] Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, USA

ARTICLE INFO

Article history:
Received 9 July 2011
Received in revised form 4 November 2011
Accepted 18 November 2011
Available online 3 December 2011

Keywords:
Mannan endo-1,4-β-mannosidase
Glycoside hydrolase family 5
Carbohydrate binding module 27
Thermotoga petrophila RKU-1
Crystal structure
Substrate recognition

ABSTRACT

The breakdown of β-1,4-mannoside linkages in a variety of mannan-containing polysaccharides is of great importance in industrial processes such as kraft pulp delignification, food processing and production of second-generation biofuels, which puts a premium on studies regarding the prospection and engineering of β-mannanases. In this work, a two-domain β-mannanase from *Thermotoga petrophila* that encompasses a GH5 catalytic domain with a C-terminal CBM27 accessory domain, was functionally and structurally characterized. Kinetic and thermal denaturation experiments showed that the CBM27 domain provided thermo-protection to the catalytic domain, while no contribution on enzymatic activity was observed. The structure of the catalytic domain determined by SIRAS revealed a canonical $(\alpha/\beta)_8$-barrel scaffold surrounded by loops and short helices that form the catalytic interface. Several structurally related ligand molecules interacting with TpMan were solved at high-resolution and resulted in a wide-range representation of the subsites forming the active-site cleft with residues W134, E198, R200, E235, H283 and W284 directly involved in glucose binding.

1. Introduction

Mannan endo-1,4-β-D-mannosidase or 1,4-β-D-mannan mannohydrolase (EC 3.2.1.78), commonly referred as β-mannanase, catalyzes the hydrolysis of β-1,4-mannoside linkages in various mannan-containing polysaccharides, such as glucomannans and galactomannans (Stålbrand et al., 1993; de Vries and Visser, 2001). Degradation of these polysaccharides represents a key step for a number of industrial applications including delignification of kraft pulps (Tenkanen et al., 1997; Montiel et al., 2002), food processing (Sachslehner et al., 2000; Dhavan and Kaur, 2007) and production of second-generation biofuels (Dhavan and Kaur, 2007). In general, these biotechnological processes such as biomass pre-treatments, are performed under extreme environmental conditions regarding pH, osmolarity and temperature. Thus, β-mannanases being stable and functional at high temperatures offer substantial techno-economical advantages.

In addition to their biotechnological relevance, mannan-degrading enzymes also participate in a number of biological processes such as fruit ripening (Pressey, 1989), seed germination (Black, 1996) and remodeling of plant cell walls (reviewed in Schröder et al., 2009). These enzymes have also been used in structural characterization of polysaccharides having β-mannosidic linkages and sequencing of heteropolysaccharides and carbohydrates attached to glycoproteins (Dhavan and Kaur, 2007).

Thermotoga petrophila strain RKU-1 (T) is a hyperthermophilic bacterium isolated from the Kubiki oil reservoir in Niigata (Japan) that grows optimally at 80 °C (Takahata et al., 2001). Some hyperthermostable enzymes produced by this microorganism have demonstrated great potential for industrial applications and served as models for investigating structure–function-stability relationships in multidomain glycosyl hydrolases (Santos et al., 2010, 2011; Squina et al., 2010; Cota et al., 2011).

The β-mannanase from *T. petrophila* RKU-1, TpMan, consists of a CaZy GH5 catalytic core connected to a CBM27 accessory domain by an 100-residue-long linker. To date, only six structures of GH5 endo-β-1,4-mannanases have been solved: *Thermobifida fusca*

Abbreviations: Mannanases: TpMan, from *Thermotoga petrophila*; TfMan, from *Thermobifida fusca*; TrMan, from *Trichoderma reesei*; LeMan, from *Lycopersicon esculentum*; CmMan, from *Cellvibio mixtus*; CjMan, from *Cellvibrio japonicus*; BaMan, from *Bacillus agaradhaerens*; BsMan, from *Bacillus subtilis*; VsMan, from *Vibrio* sp. strain MA-138.

* Corresponding author. Address: Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Rua Giuseppe Maximo Scolfaro, 10000 Campinas, 13083-970 SP, Brazil. Fax: +55 19 3512 1004.
E-mail address: mario.murakami@lnbio.org.br (M.T. Murakami).

1047-8477/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2011.11.021

BJ www.biochemj.org

Biochem. J. (2012) **441**, 95–104 (Printed in Great Britain) doi:10.1042/BJ20110869

online data

Dissecting structure–function–stability relationships of a thermostable GH5-CBM3 cellulase from *Bacillus subtilis* 168

Camila R. SANTOS*[1], Joice H. PAIVA*[1], Maurício L. SFORÇA*, Jorge L. NEVES*, Rodrigo Z. NAVARRO*, Júnio COTA†, Patrícia K. AKAO*, Zaira B. HOFFMAM†, Andréia N. MEZA*, Juliana H. SMETANA*, Maria L. NOGUEIRA*, Igor POLIKARPOV‡, José XAVIER-NETO*, Fábio M. SQUINA†, Richard J. WARD§, Roberto RULLER†, Ana C. ZERI* and Mário T. MURAKAMI*[2]

*Laboratório Nacional de Biociências, Centro Nacional de Pesquisa em Energia e Materiais, Campinas SP, Brazil, †Laboratório Nacional de Ciência e Tecnologia do Bioetanol, Centro Nacional de Pesquisa em Energia e Materiais, Campinas SP, Brazil, ‡Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, Brazil, and §Departamento de Química, Universidade de São Paulo, Ribeirão Preto, Brazil

Cellulases participate in a number of biological events, such as plant cell wall remodelling, nematode parasitism and microbial carbon uptake. Their ability to depolymerize crystalline cellulose is of great biotechnological interest for environmentally compatible production of fuels from lignocellulosic biomass. However, industrial use of cellulases is somewhat limited by both their low catalytic efficiency and stability. In the present study, we conducted a detailed functional and structural characterization of the thermostable BsCel5A (*Bacillus subtilis* cellulase 5A), which consists of a GH5 (glycoside hydrolase 5) catalytic domain fused to a CBM3 (family 3 carbohydrate-binding module). NMR structural analysis revealed that the *Bacillus* CBM3 represents a new subfamily, which lacks the classical calcium-binding motif, and variations in NMR frequencies in the presence of cellopentaose showed the importance of polar residues in the carbohydrate interaction. Together with the catalytic domain, the CBM3 forms a large planar surface for cellulose recognition, which conducts the substrate in a proper conformation to the active site and increases enzymatic efficiency. Notably, the manganese ion was demonstrated to have a hyper-stabilizing effect on BsCel5A, and by using deletion constructs and X-ray crystallography we determined that this effect maps to a negatively charged motif located at the opposite face of the catalytic site.

Key words: accessory domain, cellulase 5A, carbohydrate-binding module, kinetics, structure, thermal stability.

INTRODUCTION

The production of ethanol from lignocellulosic biomass is well-placed among other possibilities to produce energy, owing to its potential sustainability and agro-economic benefits [1]. Different chemical and enzymatic strategies have been proposed for the saccharification of lignocellulosic biomass [2,3]. The latter has been implemented at different organization levels, from point mutations of cellulolytic enzymes to the engineering of whole metabolic pathways in micro-organisms [4]. Furthermore, these two approaches have been employed synergistically to increase the yields of fermentable sugars from lignocellulosic biomass. Although enzymatic hydrolysis has been successfully implemented as a biomass-to-bioenergy technology, it is widely regarded as an expensive and wasteful link in the whole chain of bioethanol production, mainly because of the large amounts of enzymes required to compensate for their low catalytic efficiency and stability [5,6]. Driven by this need, a number of enzymes with biological and commercial value have been systematically modified through structure-based rational approaches, or by intelligent use of serendipity [7–9]. In spite of these efforts, the molecular basis of protein stability and its intricate correlation with catalysis is still elusive, putting a premium on studies that address structure–function–stability relationships and the discovery of new alternative pathways to improve thermal stability of the enzymes.

Endo-β-1,4-glucanases (EC 3.2.1.4), also referred to as endoglucanases, are the major enzymes responsible for the breakdown of internal glycosidic bonds of cellulose chains. Several endoglucanases from different *Bacillus subtilis* strains (BsEgls) have been cloned and characterized, aiming at their potential applications in the biofuels industry [10–17]. BsEgls encompass a catalytic domain belonging to family 5A, containing a CBM3 (family 3 carbohydrate-binding module) appended to their C-terminus. BsEgls show maximum activity at approximately pH 6 and 60 °C [10,14] and are capable of hydrolysing CMC (carboxymethylcellulose) and lichenan, but not xylan, chitosan or laminarin [10,18]. Typically, BsEgls are thermostable enzymes, retaining 90 % of activity after incubation for 2 h at 65 °C, 70 % after 30 min at 75 °C, and 12 % after 10 min at 80 °C [11,13,14].

It has been proposed that the non-catalytic domain of BsEgls is important to bind insoluble substrates [13], increasing the catalytic efficiency by a mechanism that involves the disruption of the compact cellulose structure and delivering of the substrate to the CC (catalytic core) in a favoured configuration for catalysis [19–21]. As it is often the case with GHs (glycoside hydrolases), deletion of the accessory domain decreases the thermal stability

Abbreviations used: BsCel5A, *Bacillus subtilis* cellulase 5A; CBM, carbohydrate-binding module; BsEgl, *Bacillus subtilis* endoglucanase; CC, catalytic core; CMC, carboxymethylcellulose; CtCipACBM, *Clostridium thermocellum* CBM3; GH, glycoside hydrolase; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhancement spectroscopy; PEG, poly(ethylene glycol); SAXS, small-angle X-ray scattering; T_m, melting temperature; TmCel5A, *Thermotoga maritima* Cel5A; WT, wild-type.

[1] These authors contributed equally to this work.

[2] To whom correspondence should be addressed (email mario.murakami@lnbio.org.br).

Structural factors and atomic co-ordinates of the *Bacillus subtilis* cellulase 5A catalytic core have been deposited in the Protein Data Bank under accession codes 3PZV (Form I), 3PZU (Form II) and 3PZT (Form II*). NMR data of the *Bacillus subtilis* cellulase 5A CBM3 were deposited in the Biological Magnetic Resonance Bank and Protein Data Bank under codes 17399 and 2L8A respectively.

Appl Microbiol Biotechnol
DOI 10.1007/s00253-012-4548-4

Biomass-to-bio-products application of feruloyl esterase from *Aspergillus clavatus*

André R. L. Damásio · Cleiton Márcio Pinto Braga · Lívia B. Brenelli · Ana Paula Citadini · Fernanda Mandelli · Junio Cota · Rodrigo Ferreira de Almeida · Victor Hugo Salvador · Douglas Antonio Alvaredo Paixao · Fernando Segato · Adriana Zerlotti Mercadante · Mario de Oliveira Neto · Wanderley Dantas do Santos · Fabio M. Squina

Received: 15 September 2012 / Revised: 22 October 2012 / Accepted: 23 October 2012

Abstract The structural polysaccharides contained in plant cell walls have been pointed to as a promising renewable alternative to petroleum and natural gas. Ferulic acid is a ubiquitous component of plant polysaccharides, which is found in either monomeric or dimeric forms and is covalently linked to arabinosyl residues. Ferulic acid has several commercial applications in food and pharmaceutical industries. The study herein introduces a novel feruloyl esterase from *Aspergillus clavatus* (AcFAE). Along with a comprehensive functional and biophysical characterization, the low-resolution structure of this enzyme was also determined by small-angle X-ray scattering. In addition, we described the production of phenolic compounds with antioxidant capacity from wheat arabinoxylan and sugarcane bagasse using AcFAE. The ability to specifically cleave ester linkages in hemicellulose is useful in several biotechnological applications, including improved accessibility to lignocellulosic enzymes for biofuel production.

Keywords *Aspergillus clavatus* · Feruloyl esterase · Sugarcane bagasse · Wheat arabinoxylan · Phenolic compounds · Antioxidant capacity

André R. L. Damásio, Cleiton Márcio Pinto Braga, and Lívia Brenelli contributed equally to this work.

A. R. L. Damásio · C. M. P. Braga · L. B. Brenelli · A. P. Citadini · F. Mandelli · J. Cota · R. F. de Almeida · D. A. A. Paixao · F. Segato · F. M. Squina (✉)
Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM),
Rua Giuseppe Máximo Scolfaro, no. 10.000, Caixa, Postal 6170 13083-970, Campinas, São Paulo, Brazil
e-mail: fabio.squina@bioetanol.org.br

F. Mandelli · J. Cota · A. Z. Mercadante
Departamento de Ciência de Alimentos, Faculdade de Engenharia de Alimentos, UNICAMP,
Campinas, São Paulo, Brazil

M. de Oliveira Neto
Departamento de Física e Biofísica, Instituto de Biociências, UNESP,
Botucatu, São Paulo, Brazil

V. H. Salvador · W. D. do Santos
Departamento de Bioquímica, Centro do Ciências Biológicas, UEM,
Maringá, Brazil

Introduction

The climate change due to the greenhouse gas emissions and other human activities has encouraged the quest for renewable means of fuel and chemical production (Chundawat et al. 2011). Lignocellulosic resources have the most favorable net energy ratio to maintain environmental sustainability. Therefore, the structural polysaccharides derived from plant cell walls have been pointed to as a promising renewable alternative to petroleum and natural gas (Crepin et al. 2003b).

Plant biomass saccharification and biofuel production are not a trivial process. The polysaccharide network in plant cell walls is one of the most complex structures in nature. The plant cell wall recalcitrance is highly influenced by ester cross-linked hydroxycinnamic acids, such as ferulic

Published online: 16 November 2012

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

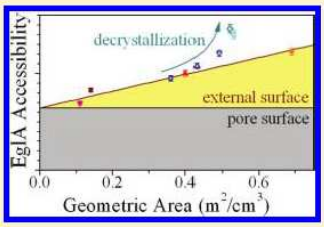
Insights on How the Activity of an Endoglucanase Is Affected by Physical Properties of Insoluble Celluloses

Juliano Bragatto, Fernando Segato, Junio Cota, Danilo B. Mello, Marcelo M. Oliveira, Marcos S. Buckeridge, Fabio M. Squina, and Carlos Driemeier*

Laboratório Nacional de Ciência e Tecnologia do Bioetanol, CTBE, Caixa Postal 6170, 13083-970 Campinas, São Paulo, Brazil

Supporting Information

ABSTRACT: Cellulose physical properties like crystallinity, porosity, and particle size are known to influence cellulase activity, but knowledge is still insufficient for activity prediction from such measurable substrate characteristics. With the aim of illuminating enzyme–substrate relationships, this work evaluates a purified hyperthermophilic endo-1,4-beta-glucanase (from *Pyrococcus furiosus*) acting on 13 celluloses characterized for crystallinity and crystal width (by X-ray diffraction), wet porosity (by thermoporometry), and particle size (by light scattering). Activities are analyzed by the Michaelis–Menten kinetic equation, which is justified by low enzyme–substrate affinity. Michaelis–Menten coefficients K_m and k_{cat} are reinterpreted in the context of heterogeneous cellulose hydrolysis. For a set of as-received and milled microcrystalline celluloses, activity is successfully described as a function of accessible substrate concentration, with accessibility proportional to K_m^{-1}. Accessibility contribution from external particle areas, pore areas, and crystalline packing are discriminated to have comparable magnitudes, implying that activity prediction demands all these substrate properties to be considered. Results additionally suggest that looser crystalline packing increases the lengths of released cello-oligomers as well as the maximum endoglucanase specific activity (k_{cat}).



1. INTRODUCTION

Cellulosic biomass will substitute a significant share of fossil resources currently used for materials, chemicals, and fuels. Enzymatic saccharification is one of the most studied strategies to convert the recalcitrant and insoluble cellulosic feedstocks into simple sugars that are reactive soluble intermediates.[1,2] A major use for glucose obtained from cellulose hydrolysis is its fermentation to ethanol, which is already a widespread transportation fuel.[3]

Three different types of cellulases are involved in cellulose hydrolysis: endoglucanases (EC 3.2.1.4), cellobiohydrolases (EC 3.2.1.91, EC 3.2.1.176), and β-glucosidases (EC 3.2.1.21). Current paradigm states that endoglucanases randomly hydrolyzes β-1,4 bonds within cellulose amorphous regions. At the chain ends created by endoglucanases, cellobiohydrolases can start the processive deconstruction of crystalline cellulose, releasing soluble cellobiose that β-glucosidases hydrolyze to glucose.[4−6]

As reviewed extensively by many authors,[5−9] enzymatic hydrolysis of cellulosic substrates is complex. There are synergies between the different types of enzymes. Reactions are inhibited by accumulation of hydrolysis products. Realistic cellulosic resources have, in addition to cellulose, hemicellulose and lignin; these other components compete for enzyme adsorption, hinder access to cellulose, and ultimately require other enzymes such as xylanases and ligninases. Finally, cellulosic substrates are insoluble in water, implying heterogeneous reactions that depend on substrates' tridimensional organization.

The complexity from substrate structure is partly summarized in the concept of accessibility.[10,11] This concept highlights that only a fraction of the substrate mass is actually reachable by cellulases, whose movements are restricted by their sizes of about 5 nm. Measurable substrate properties like surface areas and porosity[12−14] have been proposed to modulate accessibility and, in this way, affect hydrolysis rate.

There is also the issue of cellulose crystallinity.[7,15,16] Understanding its effect on hydrolysis rates faces two fundamental challenges. First, cellulose crystallinity measurements are somewhat ambiguous,[16] and unfortunately, careless crystallinity measurements are common. For instance, X-ray diffraction (XRD) factors significant for cellulose crystallinity measurements[17] are often neglected. Second, there are cellulases specialized in crystalline as well as in noncrystalline cellulose.[5] Hence, relationships between substrate crystallinity and hydrolysis rate are a function of the cellulase system.

Against this background of complexity, this article investigates enzymatic hydrolysis in rather simple systems with the aim of illuminating mechanisms in complex realistic systems. Hydrolysis was performed with a single purified cellulase, the hyperthermophilic endo-1,4-beta-glucanase from *Pyrococcus furiosus* (EglA).[18,19] EglA activities were compared on 13 insoluble cellulosic substrates with extensive characterization of physical properties (crystallinity, crystal width, wet porosity, and particle size). Activities were analyzed by Michaelis–

Received: March 6, 2012
Revised: May 2, 2012
Published: May 11, 2012

Microbiology (2013), **159**, 112–125 DOI 10.1099/mic.0.062901-0

Boto, a class II transposon in *Moniliophthora perniciosa*, is the first representative of the *PIF/Harbinger* superfamily in a phytopathogenic fungus

Jorge Fernando Pereira,[1]† Ana Paula Morais Martins Almeida,[1]‡ Júnio Cota,[1]§ João Alencar Pamphile,[2] Gilvan Ferreira da Silva,[1]‖ Elza Fernandes de Araújo,[1] Karina Peres Gramacho,[3] Sérgio Hermínio Brommonschenkel,[4] Gonçalo Amarante Guimarães Pereira[5] and Marisa Vieira de Queiroz[1]

Correspondence
Marisa V. de Queiroz
mvqueiro@ufv.br

[1]Universidade Federal de Viçosa, Departamento de Microbiologia, CEP 36571-000, Viçosa, MG, Brazil

[2]Universidade Estadual de Maringá, Departamento de Biologia Celular e Genética, CEP 87020-900, Maringá, PR, Brazil

[3]CEPLAC, Centro de Pesquisa do Cacau, CEP 45600-000, Itabuna, BA, Brazil

[4]Universidade Federal de Viçosa, Departamento de Fitopatologia, CEP 36571-000, Viçosa, MG, Brazil

[5]Universidade Estadual de Campinas, Departamento de Genética e Evolução, CEP 13083-970, Campinas, SP, Brazil

Boto, a class II transposable element, was characterized in the *Moniliophthora perniciosa* genome. The *Boto* transposase is highly similar to plant *PIF*-like transposases that belong to the newest class II superfamily known as *PIF/Harbinger*. Although *Boto* shares characteristics with *PIF*-like elements, other characteristics, such as the transposase intron position, the position and direction of the second ORF, and the footprint, indicate that *Boto* belongs to a novel family of the *PIF/Harbinger* superfamily. Southern blot analyses detected 6–12 copies of *Boto* in C-biotype isolates and a ubiquitous presence among the C- and S-biotypes, as well as a separation in the C-biotype isolates from Bahia State in Brazil in at least two genotypic groups, and a new insertion in the genome of a C-biotype isolate maintained in the laboratory for 6 years. In addition to PCR amplification from a specific insertion site, changes in the *Boto* hybridization profile after the *M. perniciosa* sexual cycle and detection of *Boto* transcripts gave further evidence of *Boto* activity. As an active family in the genome of *M. perniciosa*, *Boto* elements may contribute to genetic variability in this homothallic fungus. This is the first report of a *PIF/Harbinger* transposon in the genome of a phytopathogenic fungus.

Received 15 August 2012
Revised 22 October 2012
Accepted 24 October 2012

†**Present address:** Embrapa Trigo, Rodovia BR 285 Km 294, CEP 99001-970, Passo Fundo, RS, Brazil.

‡**Present address:** Universidade Federal de Minas Gerais, Departamento de Microbiologia, Av. Antônio Carlos 6627, CEP 31270-901, Belo Horizonte, MG, Brazil.

§**Present address:** Laboratório Nacional de Ciência e Tecnologia do Bioetanol, Rua Giuseppe Máximo Scolfaro 10000, CEP 13083-970, Campinas, SP, Brazil.

‖**Present address:** Embrapa Amazônia Ocidental, Rodovia AM-010 Km 29, CEP 69010-970, Manaus, AM, Brazil.

Abbreviations: IS, insertion sequence; MITE, miniature inverted-repeated transposable element; TIR, terminal inverted repeat; TSD, target site duplication.

The GenBank/EMBL/DDBJ accession number for the *Boto* sequence reported in this paper is EU218539.

INTRODUCTION

Eukaryotic transposable elements are divided into two main categories according to their transposition mechanism: the class I elements that transpose by an intermediate RNA and are further divided into the five orders LTR, DIRS, Penelope-like, LINEs and SINEs (Wicker *et al.*, 2007); and the class II elements that transpose directly at the DNA level, not requiring an RNA transposition intermediate. Class II elements can be further divided into subclasses, superfamilies and families by the transposition mechanisms and structural features of the terminal inverted repeats (TIRs), the transposase and the target site duplication (TSD) (Daboussi & Capy, 2003; Wicker *et al.*, 2007). Class II elements belonging to the

Biochemical and Biophysical Research Communications 403 (2010) 214–219

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Thermal-induced conformational changes in the product release area drive the enzymatic activity of xylanases 10B: Crystal structure, conformational stability and functional characterization of the xylanase 10B from *Thermotoga petrophila* RKU-1

Camila Ramos Santos [a], Andreia Navarro Meza [a], Zaira Bruna Hoffmam [b], Junio Cota Silva [b], Thabata Maria Alvarez [b], Roberto Ruller [b], Guilherme Menegon Giesel [c], Hugo Verli [c], Fabio Marcio Squina [b], Rolf Alexander Prade [d], Mario Tyago Murakami [a,*]

[a] Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, SP, Brazil
[b] Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais, Campinas, SP, Brazil
[c] Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
[d] Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, USA

ARTICLE INFO

Article history:
Received 27 October 2010
Available online 9 November 2010

Keywords:
Xylanase
Glycoside hydrolase family 10
Crystal structure
Thermostability
Thermotoga petrophila RKU-1

ABSTRACT

Endo-xylanases play a key role in the depolymerization of xylan and recently, they have attracted much attention owing to their potential applications on biofuels and paper industries. In this work, we have investigated the molecular basis for the action mode of xylanases 10B at high temperatures using biochemical, biophysical and crystallographic methods. The crystal structure of xylanase 10B from hyperthermophilic bacterium *Thermotoga petrophila* RKU-1 (TpXyl10B) has been solved in the native state and in complex with xylobiose. The complex crystal structure showed a classical binding mode shared among other xylanases, which encompasses the −1 and −2 subsites. Interestingly, TpXyl10B displayed a temperature-dependent action mode producing xylobiose and xylotriose at 20 °C, and exclusively xylobiose at 90 °C as assessed by capillary zone electrophoresis. Moreover, circular dichroism spectroscopy suggested a coupling effect of temperature-induced structural changes with this particular enzymatic behavior. Molecular dynamics simulations supported the CD analysis suggesting that an open conformational state adopted by the catalytic loop (Trp297-Lys326) provokes significant modifications in the product release area (+1,+2 and +3 subsites), which drives the enzymatic activity to the specific release of xylobiose at high temperatures.

1. Introduction

Xylanases (β-1,4-xylan xylanohydrolase, E.C. 3.2.1.8) are responsible for breaking down xylan, the major hemicellulosic component of plant cell walls, into short xylooligosaccharides by a general acid–base mechanism involving two glutamic acid residues [1,2]. Typically, these enzymes can be classified into glycoside hydrolase (GH) families 10 and 11 based on amino-acid sequence similarities [3]. However, xylanases belonging to families 5, 8 and 43 have also been identified [4,5]. The biotechnological applications of xylanases include food, feed, textile and paper industries [6], and recently they have received much attention owing to their use in degradation of lignocellulosic biomass for biofuels production [7,8]. In this context, enzymes exhibiting high enzymatic efficiency combined with hyper thermal and chemical tolerance are extremely desired. Although countless extremophilic enzymes have been characterized, the molecular basis for their mode of action at high temperatures is still poorly understood. Several structures of family 10 xylanases have been solved, from mesophilic to hyperthermophilic organisms enabling an in-depth comparative study in terms of thermophilicity, thermostability and action mechanism. Thus, in order to gain insights into the structural determinants of their mode of action at high temperatures we have performed a functional and structural characterization of a hyperthermostable xylanase 10B from the bacterium *Thermotoga petrophila* RKU-1 (TpXyl10B). Our studies addressed important

Abbreviations: TpXyl10B, xylanase 10B from *Thermotoga petrophila* RKU-1; TmXyl10B, xylanase 10B from *Thermotoga maritima* MSB8; GH, glycoside hydrolase; CD, circular dichroism; CZE, capillary zone electrophoresis; U, unit; MD, molecular dynamics; XTAL, X-ray crystallography.

* Corresponding author. Address: Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Rua Giuseppe Maximo Scolfaro, 10000, Campinas, 13083-970 SP, Brazil. Fax: +55 19 3512 1004.
E-mail address: mario.murakami@lnbio.org.br (M.T. Murakami).

0006-291X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2010.11.010

Biochemical and Biophysical Research Communications 399 (2010) 505–511

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Substrate cleavage pattern, biophysical characterization and low-resolution structure of a novel hyperthermostable arabinanase from *Thermotoga petrophila*

Fabio M. Squina [a,*], Camila R. Santos [b], Daniela A. Ribeiro [a], Júnio Cota [a,c], Renata R. de Oliveira [b], Roberto Ruller [a], Andrew Mort [d], Mario T. Murakami [b], Rolf A. Prade [e,**]

[a] Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, SP, Brazil
[b] Laboratório Nacional de Biociências (LNBio), do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, SP, Brazil
[c] Departamento de Ciências de Alimentos da Universidade Estadual de Campinas, SP, Brazil
[d] Department of Biochemistry, Oklahoma State University, Stillwater, OK 74078, USA
[e] Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK 74078, USA

ARTICLE INFO

Article history:
Received 25 June 2010
Available online 1 August 2010

Keywords:
Arabinanase
Capillary zone electrophoresis
Small angle X-ray scattering (SAXS)
Hyperthermostable enzyme
Arabinan

ABSTRACT

Arabinan is a plant structural polysaccharide degraded by two enzymes; α-L-arabinofuranosidase and endo-1,5-α-L-arabinanase. These enzymes are highly diversified in nature, however, little is known about their biochemical and biophysical properties. We have characterized a novel arabinanase (AbnA) isolated from *Thermotoga petrophila* with unique thermostable properties such as the insignificant decrease of residual activity after incubation up to 90 °C. We determined the AbnA mode of operation through capillary zone electrophoresis, which accumulates arabinotriose and arabinobiose as end products after hydrolysis of arabinan-containing polysaccharides. Spectroscopic analyses by Far-UV circular dichroism and intrinsic tryptophan fluorescence emission demonstrated that AbnA is folded and formed mainly by β-sheet structural elements. *In silico* molecular modeling showed that the AbnA structure encompasses a five-bladed β-propeller catalytic core juxtaposed by distorted up-and-down β-barrel domain. The low-resolution structure determined by small angle X-ray scattering indicated that AbnA is monomeric in solution and its molecular shape is in full agreement with the model.

1. Introduction

Plant structural polysaccharides are the most abundant source of renewable carbon in the biosphere and represent a valuable industrial substrate in several applications, such as bio-energy production, pulp and paper, food technology, detergent, textile, nutritional or medical research and organic synthesis [7]. Aiming at the development of sugar-platform based chemicals; plant biomass hydrolytic enzymes are the key technological component for efficient use of renewable feedstocks through an environmental friendly bioconversion route.

Arabinan is composed of an α-1–5 linked L-arabinofuranosyl residue backbone, substituted with α-1,3 and α-1,2 linked chains of L-arabinofuranosyl units. They occur as either homoglycans, arabinans, or as heteroglycans such as arabinoxylans and arabinogalactans [25]. Two major enzyme families are involved in arabinan hydrolysis, the α-L-arabinofuranosidases (ABFs) (EC 3.2.1.55), which remove arabinose side chains, allowing the access of endo-1,5-α-L-arabinanase (ABNs) (EC 3.2.1.99) that releases arabino-oligosaccharides and L-arabinose as hydrolysis products. Arabinases and arabinofuranosidases are classified by CAZy [5,6] as being members of glycoside hydrolase families GH3, GH43, GH51, GH54 and GH62, which display a wide range of enzymatic activities. Only little information regarding biochemical and biophysical properties is available.

To date, crystal structures from bacterial family GH51 arabinofuranosidases (EC 3.2.1.55) have been published from *Geobacillus stearothermophilus* [9], *Clostridium thermocellum* [23] and *Thermobacillus xylanilyticus* [15], and crystal structures from arabinases of the related family GH43 (EC 3.2.1.99) have been determined from *Cellvibrio japonicus* [13], *Bacillus subtillis* [17] and *G. stearothermophilus* [1].

Members from CAZy GH43 family shows three conserved acidic residues, two aspartic acids and one glutamic acid, located at the active site providing the general acid and base components for glycosidic bond hydrolysis, with inversion of the anomeric configuration [16]. These catalytic residues are close in space in the five-bladed beta-propeller fold, which was structurally determined for ABN from *C. japonicus* [13] and *B. subtilis* [17]. A long V-shaped surface groove, partially enclosed at one end, forms a single extended

* Corresponding author. Address: Rua Giuseppe Máximo Scolfaro, 10.000 Polo II de Alta Tecnologia, Caixa Postal 6170 – Campinas-SP, CEP: 13083-970 Campinas, SP, Brazil. Tel.: +55 19 3518 3104.
** Corresponding author.
E-mail address: fabio.squina@bioetanol.org.br (F.M. Squina).

0006-291X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2010.07.097

Revisão

Biotechnological production of bioflavors and functional sugars

Produção biotecnológica de bioaromas e açúcares funcionais

Juliano Lemos BICAS[1], Júnio Cota SILVA[2], Ana Paula DIONÍSIO[2], Gláucia Maria PASTORE[2*]

Abstract

Bioflavors and oligosaccharides are two classes of substances that may be produced biotechnologically through microbial bioprocesses. These compounds have attracted the interest of pharmaceutical and food industries not only due to their technological properties (sweetening/fiber or flavoring, respectively), but also as a consequence of other functional properties such as, for example, health promoting benefits. The use of agro-industrial residues as substrates in biotechnological processes seems to be a valuable alternative in helping to overcome the high manufacturing costs of industrial fermentations. This manuscript reviews the most important advances in biotechnological production of bioflavors and oligosaccharides. The use of some agro-industrial residues in such processes is also cited and discussed, showing that this is a rising trend in biotechnology.

Keywords: *bioflavors; biotransformation; de novo synthesis; functional food; oligosaccharides.*

Resumo

Bioaromas e oligossacarídeos são duas classes de substâncias que podem ser produzidas biotecnológicamente por meio de bioprocessos microbiológicos. Estes compostos têm atraído o interesse das indústrias farmacêutica e de alimentos não só devido às suas propriedades tecnológicas (adoçantes/fibras ou aromatizantes, respectivamente), mas também como consequência de outras propriedades funcionais como, por exemplo, benefícios na promoção da saúde. O uso de resíduos agroindustriais como substrato em processos biotecnológicos parece ser uma alternativa valiosa para superar os altos custos de manufatura envolvidos nas fermentações industriais. Este manuscrito faz uma revisão dos mais importantes avanços na produção biotecnológica de bioaromas e oligossacarídeos. O uso de alguns resíduos agroindustriais nestes processos também são citados e discutidos, mostrando que esta é uma tendência crescente na biotecnologia.

Palavras-chave: *bioaromas; biotransformação; síntese de novo; alimentos funcionais; oligossacarídeos.*

1 Introduction

In recent years, it has been considered that food additives should supply not only basic nutritional and technological attributes to the final product. Their desirable properties are natural origin, different technological functions in food and favorable contribution to the promotion of consumer's health and well-being. These characteristics, typical of the so-called functional foods, may be found in bioflavors and some oligosaccharides. Bioflavors, for example, have shown biological activity in vitro and in vivo against certain types of tumor (CROWELL, 1999). Some of them can also be used to induce detoxifying enzymes, for the inhibition of potato sprouting, as an anti-microbial agent, insect repellent, among other uses (DE CARVALHO; DA FONSECA, 2006). Fructooligosaccharide and oligosaccharides, on the other hand, act as dietary fiber since they are non-digestible oligosaccharides (NDO). These NDO are currently considered prebiotics because they are fermented mainly by bifidobacteria and lactic acid bacteria in the colon, reporting beneficial effects for health by increasing the amount of desirable bacteria in the human gut (TOMOMATSU, 1994; ROBERFROID; SLAVIN, 2000; DELZENNE, 2003). Therefore, they have been a substitute for sucrose in some food products. The importance of functional foods is reflected by their global market, which represents tens of billion dollars and is growing annually at a rate of 7.5% (SANGEETHA; RAMESH; PRAPULLA, 2005; SLOAN, 2002).

Biotechnology might be broadly defined as any technique that uses living organisms (or parts of them) to make or modify products, to improve plants or animals, or to develop microorganisms for specific uses. Therefore, it is the application of science and engineering, using living organisms or their derived substances, to generate products or to perform functions that can benefit the human condition. In this sense, biotechnological bioprocesses is a subdivision of biotechnology that is responsible for translating the discoveries of life science into practical products, processes, or systems that can serve the needs of society (CBE, 1992). The biotechnological production of food products has been used by human beings since pre-history for the preparation of fermented beverages, breads etc. However, it was only in the last decades of 20th century that the use of this technique was extended to the manufacturing of ingredients, e.g. flavors and oligosaccharides. Since then, the

Recebido para publicação em 1/2/2010
Aceito para publicação em 8/3/2010 (0001RV)
[1] *Campus Alto Paraopeba, Universidade Federal de São João del-Rei, Rod. MG-443, Km 7, C.P. 131, CEP 36420-000, Ouro Branco – MG, Brazil*
[2] *Departamenrto de Ciência de Alimentos, Laboratório de Bioaromas, Universidade Estadual de Campinas, Rua Monteiro Lobato, 80, CEP 13083-862, Campinas – SP, Brazil, E-mail: glaupast@fea.unicamp.br*
**A quem a correspondência deve ser enviada*