

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS**

Leonardo Rodrigues de Almeida

**Estudo estrutural e funcional de hidrolases de glicosídeos
de organismos termofílicos**

São Carlos

2022

Leonardo Rodrigues de Almeida

**Estudo estrutural e funcional de hidrolases de glicosídeos
de organismos termofílicos**

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

Área de concentração: Física Aplicada
Opção: Física Biomolecular

Orientador: Prof. Dr. João Renato Carvalho
Muniz

Versão corrigida

(Versão original disponível na Unidade que aloja o Programa)

São Carlos

2022

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

Ficha catalográfica revisada pelo Serviço de Biblioteca e Informação Prof. Bernhard Gross, com os dados fornecidos pelo(a) autor(a)

Almeida, Leonardo Rodrigues de

Estudo estrutural e funcional de hidrolases de glicosídeos de organismos termofílicos / Leonardo Rodrigues de Almeida ; orientador João Renato Carvalho Muniz – versão corrigida – São Carlos, 2022.

184 p. : il. (algumas color.) ; 30 cm.

Tese (Doutorado - Programa de Pós-Graduação em Física Aplicada) – Instituto de Física de São Carlos, Universidade de São Paulo, 2022.

1. Hidrolases de glicosídeo. 2. Biomassa lignocelulósica. 3. *Clostridium thermocellum*. 4. β -glicosidases. 5. *Bacillus licheniformis*. 6. Cristalografia de raios X. 7. Surfactantes não iônicos. 8. Bioetanol. 9. Enzimas. I. Muniz, João Renato Carvalho, orient. II. Título.

Este trabalho é dedicado à minha família, por cada momento que renunciei de estarmos juntos nos últimos anos.

AGRADECIMENTOS

Os meus agradecimentos a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo apoio financeiro, por meio da bolsa de estudos do Projeto 16/09152-6, o que permitiu minha dedicação para a realização dos trabalhos aqui apresentados além da minha formação profissional.

Agradeço ao Prof. Dr. João R. C. Muniz, por aceitar ser meu orientador neste projeto, pelos ensinamentos, pela paciência quando estive perdido, pela compreensão e pelas oportunidades.

Ao Prof. Dr. Wanius Garcia da UFABC, por me apoiar no projeto, e ser um grande parceiro do nosso grupo, principalmente pelas análises biofísicas de SAXS e espalhamentos de luz, discussões e enriquecedoras colaborações nos trabalhos.

À Dra. Amanda Bernardes, por toda ajuda nas etapas iniciais do trabalho, pelas valiosas dicas no laboratório, pelos ensinamentos dos experimentos, e por todo apoio oferecido ao nosso grupo.

À minha família, em especial minha mãe, Bertilha, meu pai, Luiz Carlos, minha avó, Joana D'arc e minha tia, Lilian, por todo apoio em momentos de dificuldade, pelo apoio emocional, pelo amor e carinho, tudo em prol da minha busca de realização desse sonho de fazer o doutorado aqui na USP de São Carlos.

Agradeço imensamente aos amigos, de longa data, e também aos novos que essa fase trouxe para minha vida. Agradecerei pessoalmente, quando possível, um por um!

Aos amigos e colegas de casa, da Rep onde vivi o início do doutorado, pela recepção, dicas, apoio, descontrações e aprendizados.

Aos amigos e vizinhos da Vila, onde vivi a maioria da fase de doutorado, incluindo a sobrevivência em meio à terrível pandemia que atingiu todo o planeta. Sou profundamente grato por todo apoio, amizade, compreensão e aprendizado.

Os meus agradecimentos a alguns dos professores que foram meus grandes incentivadores para perseverar nessa caminhada acadêmica, em especial aos professores do Ens. médio: Sebastião; do curso Técnico: Kenia, Marcos, Cleusa e Taciano; e da Graduação: Hamilton, Valter, Geórgia, José Rubens e Lauro.

Aos colaboradores com quem pude trabalhar em conjunto neste período, e juntos produzirmos alguns dos trabalhos aqui apresentados, e outros que estão por vir.

Aos companheiros de laboratório (Biofísica e Biologia Estrutural), e dos vários dias de lutas, que certamente cometerei o equívoco de esquecer algum nome neste momento,

portanto, não me arriscarei em nomear todos.

Os meus agradecimentos ao corpo de docentes do IFSC-USP em especial aos professores Drs, Richard Garratt, André Ambrósio, Otavio Thiemann, Marcos Navarro, Javier Ellena, *in memoriam* Ricardo De Marco, por todas as aulas das disciplinas, por todas as eventuais discussões científicas, direcionamentos e contribuições para minha formação.

Aos técnicos dos Laboratórios que compõem o Grupo de Biofísica e Biologia Estrutural "Sérgio Mascarenhas", em especial Susana Sculaccio e Humberto Pereira.

À Universidade de São Paulo e o Instituto de Física de São Carlos, por toda infraestrutura oferecida e apoio à minha pesquisa.

Aos envolvidos nos serviços da secretaria do Programa de Pós-graduação do IFSC-USP por toda assistência prestada ao longo desses anos, em especial, Ricardo, Silvio e Lhaís.

Aos técnicos do ScInfor, e os técnicos da gráfica do IFSC-USP por todos os serviços prestados, além de toda disposição em ajudar.

Os meus agradecimentos às bibliotecárias e demais colaboradores da biblioteca, pelos excelentes serviços prestados, aos integrantes do projeto abnT_EX₂, ao Grupo desenvolvedor do Pacote USPSC, e a todos que contribuíram para que a produção de trabalhos acadêmicos consoante as normas ABNT com L^AT_EX fosse possível.

À todas as comunidades GNU/Linux, também aos criadores e desenvolvedores de programas *Open source*, pois muito neste trabalho foi beneficiado pelo acesso às ferramentas disponíveis para fins acadêmicos.

Finalizo os agradecimentos, com a necessidade de ser grato a mim mesmo, pela resiliência de chegar até aqui, sabendo o quão difícil foi tudo o que passei e sobrevivi até a chegada deste momento de conclusão.

“Uma vez que todos nós ansiamos por ser felizes, e que ficou manifesto que nos tomamos tais graças ao uso, e ao uso correto, das coisas, e que é a ciência que produz a correção e a boa fortuna, é preciso então, segundo parece, que todo homem, sem exceção, tome providências, de todas as maneiras, para que se tome o mais sábio possível.”

Platão, Eutidemo, 282a

RESUMO

ALMEIDA, L. R. **Estudo estrutural e funcional de hidrolases de glicosídeos de organismos termofílicos**. 2022. 184p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2022.

O esgotamento das reservas de petróleo, a crescente demanda energética de países emergentes e a necessidade de redução da emissão de dióxido de carbono sinalizam a importância pela busca de novas fontes de energias renováveis. Contudo, para que tais fontes tenham um caráter de efetiva substituição e não apenas de complementação, se fazem obrigatórias pesquisas por novas formas de se obtê-la. Assim, a produção do bioetanol mediante a hidrólise de biomassas lignocelulósicas, como do bagaço de cana-de-açúcar, por celulases tem recebido grande destaque. Devido ao seu potencial de evolução e redução de custos, a hidrólise enzimática da celulose pode ser uma peça fundamental para a produção de açúcares, etanol 2G e ácidos carboxílicos, com um custo competitivo a longo prazo. Alguns organismos termófilos, como as bactérias *Clostridium thermocellum* e *Lactobacillus gasseri*, o fungo *Thermothielavioides terrestris* e a bactéria mesófila *Bacillus licheniformis*, são exemplos conhecidos por sua capacidade em secretar celulases. *C. thermocellum* produz um complexo proteico multienzimático extracelular chamado celulosomo, que demonstra elevada capacidade de realizar uma eficiente degradação de biomassa celulósica, em especial na porção cristalina da celulose. A enzima CtBgl3B, β -glicosidase (BGL) celulosomal de *C. thermocellum* da família de hidrolases de glicosídeos GH3, foi expressa em *E. coli* (BL21), purificada por cromatografia de afinidade e de separação por massa molecular. Apresentou características termofílicas, mostrando alta atividade em temperatura acima de 60 °C, variando o pH de 4,5—7,0, com maior atividade em pH 5,5 a 70 °C. T_m em ≈ 70 °C, determinada por fluorimetria diferencial de varredura (DSF) e dicroísmo circular (CD). Além disso, teve avaliada sua atividade enzimática e parâmetros cinéticos avaliados com *p*NPG: At. Esp. = 124 U mg⁻¹, $K_M = 0,37$ (mM), constante catalítica $k_{cat} = 173$ (s⁻¹), e eficiência catalítica $k_{cat}/K_M = 469$ (mM⁻¹ s⁻¹). A enzima também demonstrou capacidade em atuar contra os substratos sintéticos como *p*NPX ($\approx 40\%$) e *Cp*NPG2 ($\approx 25\%$). Enquanto no teste de aditivos, teve suas taxas de conversões aumentadas em média 70% na presença de Triton X-100, e 60% com Tween20. Por outro lado, apresentou queda significativa nas reações com a presença de íons divalentes, DMSO e SDS respectivamente abaixo de 25-75% da atividade relativa ao *p*NPG. A estrutura cristalográfica da CtBGL3 foi determinada por meio da difração de raios-X usando substituição molecular, com os limites de resolução anisotrópicos em Å nas respectivas direções 3,47 (a*), 2,22 (b*) e 1,78 (c*). Demonstrando um arranjo dimérico, e com indicadores da qualidade do refinamento R_{work} e R_{free} , respectivamente 0,20 e 0,23. Experimentos de espalhamentos de luz em múltiplos ângulos (MALS) e espalhamento de raios-X a baixo ângulos (SAXS)

corroboram que em solução *CtBGL3* também adota arranjo dimérico. A bactéria mesofílica *B. licheniformis* produz uma proteína hipotética *BlYlmD*, que foi expressa em *E. coli* (Rosetta DE3) e teve sua estrutura determinada a 2,66 Å de resolução. Esse domínio proteico, cujo alinhamento indicava a presença de um domínio Cu-oxidase_4 (PF02578), também encontrado em lacases, e amplamente distribuído entre procariotos e eucarioto, ou de uma CNF1/YfiH-like cisteína hidrolase putativa de *B. licheniformis* *BlYlmD*, chamou a atenção devido suas semelhanças sequenciais com domínios de lacases, bem conhecidos, e se tornou o objeto de estudo nesta pesquisa. Sua estrutura é classificada como uma proteína α/β que forma quatro camadas de $\alpha/\beta/\beta/\alpha$ com folhas beta antiparalelas mistas e se reúne em um dímero na unidade assimétrica. Não teve sua função caracterizada para as funções sugeridas. Recentemente, foi comparada a uma purina nucleosídeo fosforilase PNP FAMIL multifuncional. Porém, também sem detecção de nenhuma atividade enzimática, sendo o principal indício estrutural da falta de atividade centrado na conformação da His47, e as diferenças entre os resíduos Tyr106 (Phe) e Glu77 (Asp), que podem afetar o volume da cavidade catalítica impedindo a entrada dos substratos.

Palavras-chave: Hidrolases de glicosídeo. Biomassa lignocelulósica. *Clostridium thermocellum*. β -glicosidases. *Bacillus licheniformis*. Cristalografia de raios X. Surfactantes não iônicos. Bioetanol. Enzimas.

ABSTRACT

ALMEIDA, L. R. **Structural and functional studies of glycoside hydrolases from thermophilic microorganisms**. 2022. 184p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2022.

The depletion of petroleum reserves, the growing energy demand from emerging countries, and the need to reduce the emission of carbon dioxide signal the importance of searching for new sources of renewable energy. However, in order for such sources to have a character of effective replacement and not only complementation, it is also mandatory to search for new strategies to obtain it. Thus, the production of bioethanol through the hydrolysis of lignocellulosic biomass, such as sugarcane bagasse, by cellulases has received great attention. Due to its potential for evolution and cost reduction, the enzymatic hydrolysis of cellulose can be a fundamental part for the production of sugars, 2G ethanol and carboxylic acids, with a competitive cost in the long term. Some thermophilic organisms, such as the bacteria *Clostridium thermocellum* and *Lactobacillus gasseri*, the fungus *Thermothielavioides terrestris* and the mesophilic bacterium *Bacillus licheniformis*, are examples known for their ability to secrete cellulases. *C. thermocellum* produces an extracellular multienzymatic protein complex called cellulosome, which demonstrates a high capacity to perform an efficient degradation of cellulosic biomass, especially in the crystalline portion of cellulose. The enzyme CtBgl3B, β -glycosidases (BGls) cellulosomal from *C. thermocellum* from the GH3 family of glycoside hydrolases, was expressed in *E. coli* (BL21), purified by affinity and size-exclusion chromatography. It showed thermophilic characteristics, showing high activity at temperatures above 60 °C, ranging from pH 4.5 to 7.0, with greater activity at pH 5.5 to 70 °C. T_m in ≈ 70 °C, determined by differential scanning fluorimetry (DSF) and circular dichroism (CD). In addition, its enzymatic activity and kinetic parameters were evaluated with *p*NPG: Sp. act. = 124 U mg⁻¹, $K_M = 0.37$ (mM), catalytic constant $k_{cat} = 173$ (s⁻¹), and catalytic efficiency $k_{cat}/K_M = 469$ (mM⁻¹ s⁻¹). The enzyme also demonstrated the ability to act against synthetic substrates such as *p*NPX ($\approx 40\%$) and *Cp*NPG2 ($\approx 25\%$). While in the additive test, it had its conversion rates increased 70% in the presence of Triton X-100, and 60% with Tween20. On the other hand, it showed a significant decrease in the reactions with the presence of divalent ions, DMSO and SDS, respectively, below 25-75% its relative activity to *p*NPG. The crystallographic structure of CtBGl3 was determined employing X-ray diffraction using molecular replacement, with anisotropic resolution limits in Å along their respective directions 3.47 (a*), 2.22 (b*) e 1.78 (c*), demonstrating a dimer arrangement, indicators of refinement quality R_{work} and R_{free} , respectively 0.20 and 0.23. Multi-angle light scattering (MALS) and small-angle X-ray scattering (SAXS) experiments corroborate that in solution, CtBGl3 also adopts dimer arrangement. The mesophilic bacterium *B.*

licheniformis produces a hypothetical protein *BIYlmD*, which was expressed in *E. coli* (Rosetta DE3) and had its structure determined at 2.66 Å resolution. This protein domain, whose alignment indicated the presence of a Cu-oxidase_4 domain (PF02578), also found in laccases, and widely distributed between prokaryotes and eukaryotes, or a putative CNF1/YfiH-like cysteine hydrolase from *B. licheniformis* *BIYlmD*, drew attention due to its sequential similarities with well-known laccase domains, and became the secondary object of study in this research. Its structure is classified as a α/β protein that forms four layers of $\alpha/\beta/\beta/\alpha$ with mixed antiparallel beta sheets and assembles into a dimer in the asymmetric unit. It did not have its function characterized for the suggested functions. It has recently been compared to a multifunctional purine nucleoside phosphorylase PNP FAMIN. However, also without detection of any enzymatic activity, the main structural indication for its lack of activity being centred on the conformation of His47, and the differences between the residues Tyr106 (Phe) and Glu77 (Asp), which can affect the volume of the catalytic cavity, preventing the entry of substrates.

Keywords: Glycoside hydrolases. Lignocellulosic biomass. *Clostridium thermocellum*. β -glucosidases. *Bacillus licheniformis*. X-ray crystallography. Non-ionic surfactants. Bioethanol. Enzymes.

LISTA DE FIGURAS

Figura 1 – Esquema da disposição das moléculas de celulose e hemicelulose.	31
Figura 2 – Esquema de hidrólise de substratos celulósicos.	33
Figura 3 – Esquema do mecanismo da reação de hidrólise realizada via β -glicosidasas.	36
Figura 4 – Esquema ilustrativo do celulosomo e seus componentes.	40
Figura 5 – Representação da rede recíproca para <i>CtBgl3B</i> obtida na difração de raios X.	59
Figura 6 – Etapas para obtenção da estrutura cristalográfica.	61
Figura 7 – Curva de calibração <i>pNP</i>	64
Figura 8 – Purificação da <i>CtBgl3</i>	68
Figura 9 – Curvas de DSF avaliando o efeito de diferentes tampões nos perfis de desnaturação térmica da <i>CtBgl3B</i>	69
Figura 10 – Espectro de dicroísmo circular da <i>CtBgl3</i>	71
Figura 11 – Análise de SEC-MALS da <i>CtBgl3B</i>	71
Figura 12 – Espalhamento de raios X a baixo ângulo da <i>CtBgl3B</i>	73
Figura 13 – Envelope de SAXS da <i>CtBgl3B</i>	74
Figura 14 – Representação esquemática do arranjo estrutural da <i>CtBgl3</i>	75
Figura 15 – Fatores de temperatura (representação cilíndrica - <i>sausage</i>) da <i>CtBgl3B</i>	77
Figura 16 – Visão do mapa de densidade eletrônica (FEM) na localização das serinas fosforiladas (SEP735)	78
Figura 17 – Visão detalhada do sítio ativo e da região de interface DUF — cavidade catalítica	79
Figura 18 – Efeito do pH e temperatura na velocidade das reações catalisadas pela <i>CtBgl3</i>	80
Figura 19 – Efeito do tempo de incubação e temperatura nas reações catalisadas pela <i>CtBgl3</i>	81
Figura 20 – Efeito de aditivos químicos nas reações catalisadas pela <i>CtBgl3B</i>	82
Figura 21 – Efeito dos aditivos: Triton, Tween20 e EDTA nas reações catalisadas pela <i>CtBgl3B</i>	83
Figura 22 – Especificidade da <i>CtBgl3B</i> por substratos sintéticos, efeito dos aditivos surfactantes (Triton e Tween20), e esquema das reações.	85
Figura 23 – Efeito da concentração de substrato na velocidade das reações catalisadas pela <i>CtBgl3</i>	87
Figura 24 – Reações enzimáticas pelo domínio DUF152 da proteína FAMIN.	94
Figura 25 – Purificação da <i>BIYImD</i>	95
Figura 26 – T_m obtidas por DSF via curvas de desenovelamento proteico da <i>BIYImD</i> para variadas condições de tampão e pHs.	96

Figura 27 – Representação <i>cartoon</i> da unidade assimétrica da estrutura tridimensional da <i>BIYlmD</i>	97
Figura 28 – Alinhamento de sequência baseado em estrutura de 6DZD para os homólogos de proteína YfiH.	99
Figura 29 – Descrição dos íons ligantes da <i>BIYlmD</i>	101
Figura 30 – Perfis SEC-MALS da <i>BIYlmd</i> mostrando monômero em solução.	102
Figura 31 – Perfis SEC-MALS da <i>BIYlmd</i> mostrando monômero em solução.	104
Figura 32 – Comparação com estruturas similares e proposição do sítio ativo para <i>BIYlmD</i>	105
Figura 33 – Sobreposição das estruturas de YlmD evidenciando a região do sítio de ligação da inosina observado na estrutura 6T1B.	107
Figura 34 – A análise com STRING v11 de <i>BIYlmD</i> mostra a rede entre as proteínas identificadas.	108
Figura 35 – Resultado da IMAC para <i>CtBgl1A</i>	149
Figura 36 – Resultado da cromatografia de troca iônica para <i>CtBgl1A</i>	151
Figura 37 – Resultado da cromatografia de exclusão por massa molecular para <i>CtBgl1A</i>	152

LISTA DE TABELAS

Tabela 1 – Composição das condições de tamponamento testadas aplicando na análise de DSF.	54
Tabela 2 – Parâmetros de coleta e análises de SAXS da <i>CtBgl3B</i>	74
Tabela 3 – Estatísticas da coleta de dados e refinamento da estrutura da <i>CtBgl3B</i>	76
Tabela 4 – Parâmetros cinéticos determinados para <i>CtBgl3B</i> * e outras BGLs. . . .	90
Tabela 5 – Estatísticas da coleta de dados e refinamento da estrutura da <i>BlYlmD</i>	98
Tabela 6 – Resultados de correspondência de estrutura secundária usando ProFunc e PDBeFold com <i>BlYlmD</i>	100
Tabela 7 – Parâmetros das interfaces analisadas de estruturas YfiH diméricas. . .	103

LISTA DE QUADROS

- Quadro 1 – Sumário geral de informações das famílias CAZymes estudadas neste trabalho e em nosso grupo de pesquisa. 35
- Quadro 2 – Avaliação do efeito de diferentes tampões nos valores de T_m da *CtBgl3B*. 70

LISTA DE ABREVIATURAS E SIGLAS

AA	Atividades auxiliares
Abf	α -L-arabinofuranosidase
Água Milli-Q®	Água ultrapura do tipo 1
BGl	β -glucosidase
BLAST	Ferramenta de pesquisa e alinhamento local básico (no inglês <i>Basic Local Alignment Search Tool</i>)
BME	β -mercaptoetanol
CAZy	Repositório de sequências de enzimas ativas em carboidratos (no inglês <i>Carbohydrate-Active enZYmes Database</i>)
CBM	Módulo de ligação a carboidratos
CD	Dicroísmo circular (no inglês <i>Circular Dichroism</i>)
CE	Esterases de carboidratos
C _p NPG2	2-Cloro- <i>p</i> -NP- β -celobiosídeo
Da	Dalton
D_{max}	Diâmetro máximo
DMSO	Dimetilsulfóxido
D.O _{600nm}	Densidade óptica a 600 nm
DSF	<i>Differential Scanning fluorimetry</i>
DTT	1,4-dithiothreitol
EC	Classificação numérica para enzimas (no inglês <i>Enzyme Commission number</i>)
EDTA	Ácido etilenodiamino tetra-acético
FEM	<i>Feature-enhanced map</i>
GHs	Hidrolases de glicosídeos (no inglês <i>Glycoside Hydrolases</i>)
GT	Glicosiltransferase

HPLC	Cromatografia líquida de alta eficiência
IMAC	Cromatografia de afinidade por íons metálicos imobilizados (no inglês <i>Immobilized Metal Ion Affinity Chromatography</i>)
IPTG	Isopropil β -D-tiogalactopiranosídeo
K_{cat}	Constante catalítica
kDa	Kilo Dalton
K_M	Constante de afinidade ou constate de <i>Michaelis-Menten</i>
L	Litro
Lcc	Lacase
LIC	Clonagem independente de ligação (no inglês <i>Ligation Independent Cloning</i>)
LNLS	Laboratório Nacional de Luz Síncrotron
LP	Liases de polissacarídeos
LPMO	Mono-oxigenases líticas de olissacarídeos
μ M	Micromolar
M	Molar
MALS	Espalhamento de luz a múltiplos ângulos (no inglês <i>Multiple-Angle Light Scattering</i>)
mM	Milimolar
M_r	Referência de massa molecular
MR	Substituição molecular (no inglês <i>Molecular Replacement</i>)
MW	Massa molecular
NCBI	Centro Nacional de Informações sobre Biotecnologia (no inglês <i>National Center for Biotechnology Information</i>)
nm	Nanômetro
nM	Nanomolar
PEG	Polietilenoglicol

PgeF	Fator de edição de peptidoglicanos <i>Peptidoglycan editing Factor</i>
PMSF	Fluoreto de fenilmetilsulfonil
PNP	Purina nucleosídeo fosforilase
<i>p</i> -NP	<i>p</i> -nitrofenol ou 4-nitrofenol
<i>p</i> NPG	<i>p</i> -NP- β -D-glucopiranosídeo
<i>p</i> NPAF	<i>p</i> -NP- α -L-arabinofuranosídeo
<i>p</i> NPBGAL	<i>p</i> -NP- α -D-galactopiranosídeo
<i>p</i> NPBM	<i>p</i> -NP- β -D-manopiranosídeo
<i>p</i> NPX	<i>p</i> -NP- β -D-xilopiranosídeo
R_g	Raio de giro
RMSD	Desvio quadrático médio (no inglês <i>Root Mean Square Deviation</i>)
s	Segundos
SAD	<i>Single-wavelength Anomalous Dispersion</i>
SAXS	Espalhamento de raios-X a baixo ângulo (no inglês <i>Small Angle X-ray Scattering</i>)
SDS-PAGE	<i>Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis</i>
SEC	Cromatografia de exclusão por tamanho (no inglês <i>Size Exclusion Chromatography</i>)
SEC-MALS	Cromatografia de exclusão por tamanho acoplado com detector de espalhamento de luz a múltiplos ângulos (no inglês <i>Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering</i>)
TEV	<i>Tobacco Etch Virus</i> protease
TLS	<i>Translation liberation screw</i>
Tris	2-Amino-2-hidroximetil-propano-1,3-diol
Triton X-100 ou Triton	2-[4-(2,4,4-trimetilpentan-2-il)fenoxi]etanol
Trx	Tioredoxina
Tween 20 ou Tween	Polioxietileno glicol sorbitano monolaurato

UV	Espectroscopia no ultravioleta
UV-vis	Espectroscopia no ultravioleta visível
V_m	Volume molecular
V_{max}	Velocidade máxima

LISTA DE SÍMBOLOS

Å	Angstrom
α	Letra grega minúscula alfa
β	Letra grega minúscula beta
γ	Letra grega minúscula gama
δ	Letra grega minúscula delta
λ	Letra grega minúscula lambda
σ	Letra grega minúscula sigma
θ	Letra grega minúscula theta
ϵ	Letra grega minúscula epsilon
η	Letra grega minúscula eta
Δ	Letra grega maiúscula delta
\approx	Sinal de aproximadamente
\pm	Sinal de mais menos

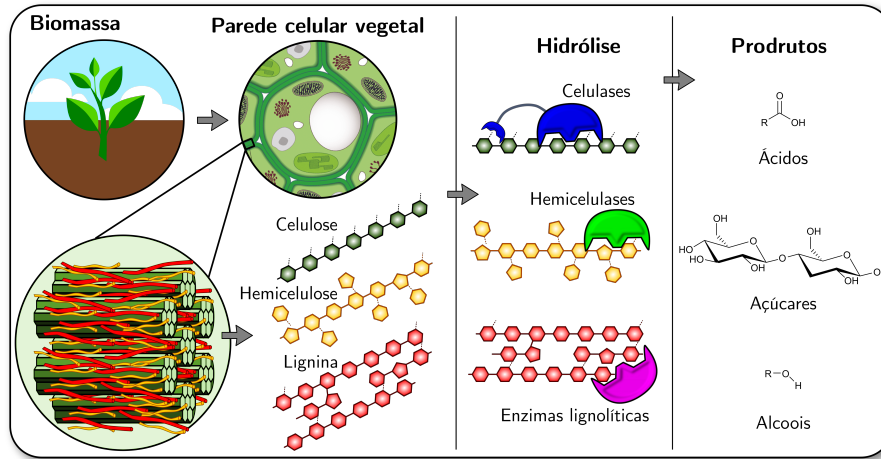
SUMÁRIO

1	INTRODUÇÃO	29
1.1	Contextualização	29
1.2	CAZYmes	32
1.2.1	Hidrolases de Glicosídeos	32
1.2.2	Estrutura e função das Hidrolases de Glicosídeos	34
1.3	Microorganismos termófilos e mesófilos	38
1.3.1	<i>Clostridium thermocellum</i>	38
1.3.1.1	BGI de <i>C. thermocellum</i>	39
1.3.2	<i>Lactobacillus gasseri</i>	41
1.3.2.1	BFfase de <i>L. gasseri</i>	41
1.3.3	<i>Thermotielavioides terrestris</i>	41
1.3.3.1	Abf62 de <i>T. terrestris</i>	42
1.3.4	<i>Bacillus licheniformis</i>	42
1.3.4.1	BIYImD de <i>B. licheniformis</i>	43
2	OBJETIVOS	45
2.1	GH1 e GH3	45
2.2	GH62 de <i>Thermotielavioides terrestris</i>	46
2.3	GH32 de <i>Lactobacillus gasseri</i>	46
2.4	BIYImD YfiH-like de <i>Bacillus licheniformis</i>	46
3	MATERIAIS E MÉTODOS	49
3.1	Busca das sequências e bioinformática	49
3.2	Produção das proteínas recombinantes	50
3.2.1	Transformação em <i>E coli</i> quimio competentes	50
3.2.2	Testes de Expressão	51
3.2.3	Expressão e purificação das proteínas recombinantes	51
3.3	Análises de estabilidade térmica	52
3.3.1	Espectroscopia de dicroísmo circular	52
3.3.2	Fluorimetria diferencial de varredura	53
3.4	Espalhamento de luz (DLS) e SEC-MALS	53
3.5	SAXS	56
3.6	Metodologia cristalográfica	56
3.6.1	Cristalização	56
3.6.2	Coleta de dados, processamento e determinação estrutural	57
3.6.3	Tratamento da Anisotropia	57

3.6.4	Faseamento e refinamento	61
3.7	Ensaio bioquímico Bgl	62
3.7.1	Cinéticas	63
3.7.1.1	Inibição por substrato	63
3.8	Ensaio bioquímico YlmD	64
3.9	Gráficos e Figuras	65
4	<i>β</i>-GLUCOSIDASE GH3 - CTBGL3B	67
4.1	Expressão e purificação	67
4.2	Análises de estabilidade térmica	68
4.2.1	DSF	68
4.2.2	CD	69
4.3	Análises estruturais	70
4.3.1	SEC-MALS	70
4.3.2	SAXS	72
4.3.3	Estrutura cristalográfica	73
4.4	Ensaio bioquímico	80
4.4.1	Efeito de aditivos químicos nas reações enzimáticas	81
4.4.2	Especificidade por substratos	84
4.4.3	Cinética enzimática	86
5	ESTRUTURA DA HIPOTÉTICA PURINA NUCLEOSÍDEO FOSFORILASE BLYLMD YFIH-LIKE DE BACILLUS LICHENIFORMIS	93
5.1	Introdução	93
5.2	Resultados	94
5.2.1	Expressão e purificação	94
5.2.2	Estabilidade térmica	95
5.2.3	Estrutura	97
5.2.4	Estado oligomérico	102
5.2.5	Atividade e sítio de ligação proposto	104
5.2.6	Rede PPI	107
6	CONSIDERAÇÕES FINAIS E PERSPECTIVAS	111
	REFERÊNCIAS	115
	APÊNDICES	141
	APÊNDICE A – INDICADORES DE QUALIDADE CRI- STALOGRAFICA	143

	APÊNDICE B – OS PARÂMETROS DOS ENSAIOS BIOQUÍMICOS	147
	APÊNDICE C – TESTES DE PURIFICAÇÃO CTBGL1A	149
C.1	Protocolos que não funcionaram neste caso	149
	ANEXOS	153
	ANEXO A – ARTIGOS PUBLICADOS	155
A.1	Sucrose-6-fosfato hidrolase GH32 de <i>Lactobacillus gasseri</i>	155
A.2	α-L-arabinofuranosidase GH62 de <i>Thermothielavioides terrestris</i>	168
A.3	Benzenesulfonyl incorporated chalcones: Synthesis, structural and optical properties	181
A.4	Protonation of inorganic 5-Fluorocytosine salts	183

1 INTRODUÇÃO



Neste capítulo é apresentada a contextualização do trabalho, junto da revisão de literatura, a problematização e alguns dos alvos abordados.

1.1 Contextualização

Vários benefícios para a sociedade estão subjacentes entre os esforços de pesquisas das universidades e departamentos de energia de diversos países, entre eles o Brasil e os Estados Unidos da América, para apoiar uma indústria viável e sustentável de biocombustíveis e bioprodutos derivados de biomassa vegetal lignocelulósica. Mais especificamente, esses benefícios incluem garantir a segurança energética futura, ratificar a emissão dos gases de efeito estufa para mitigar os impactos climáticos, conforme a convenção sobre as mudanças climáticas das Nações Unidas (COP21 na França em 2015), também expandir a diversidade de produtos de base biológica disponíveis, gerar menos produtos químicos tóxicos e resíduos, criando empregos nas áreas rurais e consequentemente melhorando a balança comercial.

A partir da biomassa, até a produção de biocombustíveis ou bioprodutos avançados, um viés de sustentabilidade pretende o direcionamento dessa produção com o mínimo de impactos ambientais. Para isso, tecnologias de desenvolvimento de biomassa com parede celular otimizada para facilitar sua desconstrução e conversão nos bioprodutos (1,2). A descoberta e melhoria de enzimas ou microorganismos, e novos métodos que aumentam a quebra da biomassa em açúcares, álcoois e ácidos carboxílicos, levando a produção de biocombustíveis e síntese de bioprodutos (1).

A preocupação com o meio ambiente, paralelamente com o esgotamento das fontes de combustíveis fósseis, tem gerado grande interesse por parte de vários setores da sociedade

em tornar mais eficientes os modos de produção de biocombustíveis (3–5). O Brasil, graças à sua imensa biodiversidade, permite ampla produção do setor agroindustrial que, por consequência, gera uma abundante variedade de resíduos agrícolas e agroindustriais, cujo bioprocessamento seria de notável interesse econômico e social. Somando todos os resíduos disponíveis, entre eles o bagaço da cana-de-açúcar (6), obtém-se aproximadamente 350 milhões de toneladas, resultando em uma capacidade de produção de 147 milhões de toneladas de açúcares (sacarose e glicose), isso reflete em um potencial de 92 bilhões de litros na produção de etanol. Contudo, se considerarmos as eficiências atuais de hidrólise e fermentação das biomassas, a capacidade de produção de etanol é de cerca de 60 bilhões de litros (7). Em trabalhos mais recentes, sobre o potencial para biocombustíveis, foi reportado que 117–236 L de etanol poderiam ser produzidos, a partir de cada tonelada de biomassa lignocelulósica de cana-de-açúcar fresca (8,9), e 506–617 L de etanol para cada tonelada de biomassa seca (8).

A lignocelulose é o material biológico mais abundante na Terra. Na maioria das vezes, contido nas paredes celulares das plantas é composto por longas cadeias de açúcares (polissacarídeos), que podem ser convertidos em biocombustíveis e bioprodutos por microorganismos que os metabolizam. No entanto, a celulose nas paredes celulares é resistente à degradação porque está incorporada em uma matriz de outros polímeros, incluindo hemicelulose e também a macromolécula lignina. Essa resistência, ou a recalcitrância, com a falta de métodos eficientes para converter lignocelulose em produtos úteis são os principais impedimentos para a produção econômica de biocombustíveis e bioprodutos a partir de biomassa vegetal (10). A inovação decorrente de pesquisas baseadas em biotecnologia avançada é fundamental para acelerar as melhorias necessárias na produção sustentável de biomassa lignocelulósica, sua desconstrução em açúcares e lignina, além da produção de biocombustíveis e bioprodutos (10).

O conceito de biocombustíveis de segunda geração corresponde à produção de etanol a partir de biomassa lignocelulósica. Essas matérias-primas são provenientes de resíduos agroindustriais como o sabugo e palha de milho, o bagaço da cana-de-açúcar (11), caule e cascas, resíduos sólidos urbanos (papel e papelão, lixo de jardim) e da indústria de papel e celulose (12). A celulose (Figura 1) é o principal componente desses materiais, mas o substrato lignocelulósico também é formado por outros dois polímeros, sendo a hemicelulose e a lignina (3, 12, 13). A composição geral da parede da cana-de-açúcar (folhas e colmo) em termos apenas de carboidratos, com base em estimativas derivadas da análise de monossacarídeos, baseado em perfil de glicoma e oligossacarídeos liberados enzimaticamente, é que a celulose representa $\approx 30\%$ da parede em ambos os órgãos, as hemiceluloses compreendem até $\approx 50\%$ e pectinas $\approx 10\%$ (14). Os intervalos de porcentagem para biomassa seca em estufa, variam de 38% a 43% de glucano, 25% a 32% de hemicelulose, 17% a 24% de lignina e 1,6% a 7,5% de extrativo (15).

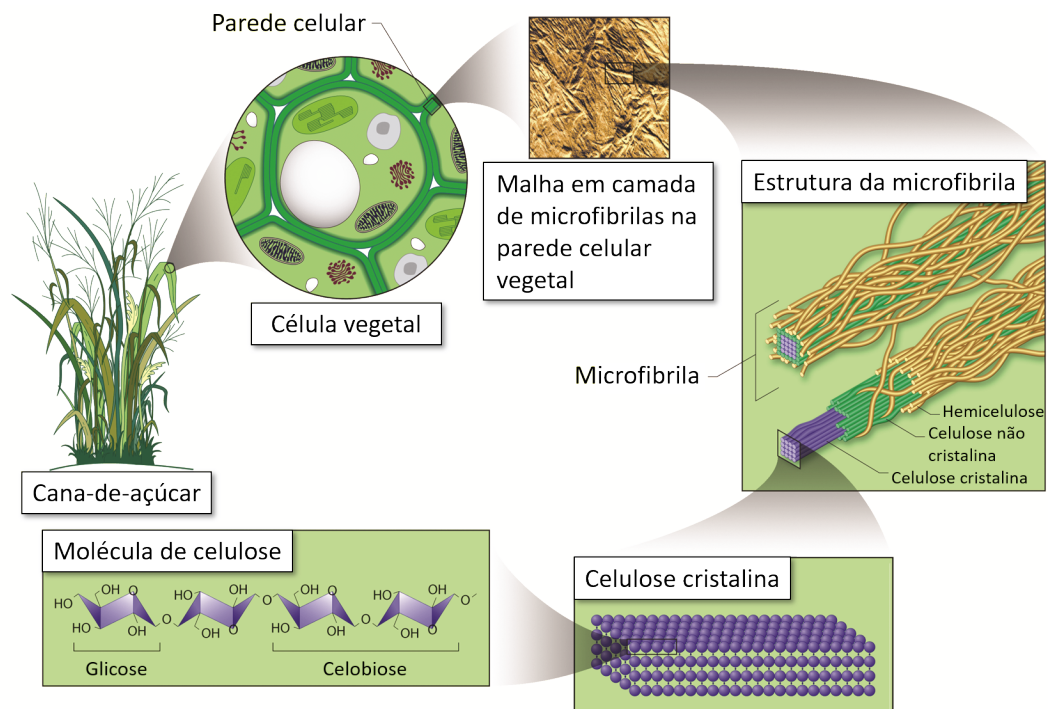


Figura 1 – Representação esquemática da disposição das moléculas de celulose e hemicelulose na parede celular das plantas e apresentação da estrutura da celulose.

Fonte: Disponível em: <http://www.genomicscience.energy.gov>. Acesso em: abr. 2018.

A celulose ($C_6H_{10}O_5$)_n tem sido utilizada pelo homem há séculos (16). Este homopolímero linear é formado por unidades de D-glicose que são unidas por meio de ligações glicosídicas β -1,4. Existem diferenças estereoquímicas entre resíduos de glicose adjacentes de forma que a unidade de repetição na celulose é representada pelo dissacarídeo celobiose que é o principal produto da hidrólise enzimática da celulose (17, 18). A celulose pode ser encontrada nas formas amorfa e cristalina sendo esta última livre de lignina e hemicelulose.

A bioconversão de celulose, matéria-prima renovável, para a produção do bioetanol representa uma ótima alternativa como fonte de energia. Contudo, para ser possível a utilização da biomassa é necessária à degradação das moléculas constituintes da parede celular a açúcares fermentáveis (14). Uma das formas de degradação consiste na hidrólise enzimática por complexos celulolíticos (CAZYmes), sendo esta bastante vantajosa em relação às outras técnicas, como, por exemplo, a hidrólise ácida, onde a corrosão de equipamentos e geração de subprodutos indesejados são alguns dos problemas encontrados. O uso dessas enzimas celulolíticas e hemicelulolíticas, porém, ainda se apresenta com custo elevado o que tem desencadeado, cada vez mais, estudos que buscam torná-las mais viáveis (10, 19, 20).

1.2 CAZYmes

A classificação das enzimas ativas de carboidratos (CAZymes), que incluem principalmente as celulases entre outras, e a anotação dos módulos que compõem as sequências de aminoácidos (estruturas primárias) são usadas principalmente para analisar, e comparar a capacidade de um organismo em quebrar e de montar carboidratos complexos. As sequências de proteínas do [GenBank](#) (21) são analisadas e os módulos de sequência correspondentes às famílias existentes Hidrolases de Glicosídeos (GH), Glicosiltransferases (GT), Liasas de Polissacarídeos (PL), Esterases de carboidratos (CE), Atividades Auxiliares (AA) — enzimas oxidativas — incluindo Mono-oxigenases Líticas de Polissacarídeos (LPMO) e também os não catalíticos módulos de ligação a carboidratos (CBM) são sistematicamente anotados e atribuídos às suas respectivas superfamílias, famílias, e clãs. Esta classificação está disponível no banco de dados [CAZy](#) (22), onde atualmente são disponibilizados mais de 23600 genomas distintos, com distribuição das famílias de proteínas (CAZymes), e divididos entre bactéria (22396), vírus (457), arquea (436) e eucariota (362), (acesso em Fev, 2022).

Todos os organismos que degradam celulose amorfa ou cristalina secretam sistemas proteicos de celulases. Tais sistemas são formados por uma variedade de enzimas com especificidades e modos de ação distintos, as quais agem em sinergia hidrolisando a celulose e hemicelulose (18,23,24). Alguns destes organismos, principalmente bactérias, leveduras e fungos (25), podem secretar um conjunto diversificado de CAZYmes independentes entre si (26), complexos multi enzimáticos chamados celulosomo (27,28), ou uma mistura de CAZYmes com variados módulos catalíticos agregados (29).

1.2.1 Hidrolases de Glicosídeos

As GHs, são enzimas que utilizam o mesmo mecanismo de catálise ácido-base, com inversão ou retenção da configuração anomérica da glicose (30). Há dois tipos comuns de sítios ativos de celulase: (i) as GHs com sítios ativos abertos (sulco, fenda), tipicamente, exibem atividade endocelulolítica (endocelulases), se ligando em qualquer região ao longo da extensão da molécula de celulose e hidrolisando a ligação β -1,4 glicosídica; (ii) GHs com sítios ativos semelhantes a um túnel, apresentam atividade exocelulolítica (celobiohidrolases ou exocelulases), se ligando nas extremidades da molécula de celulose e produzindo oligosacarídeos de comprimento unitário (30).

São três os principais tipos de enzimas celulolíticas, endoglucanases (EC 3.2.1.4), exoglucanases (EC 3.2.1.91) e β -glicosidases (EC 3.2.1.21) que separadamente, não hidrolisam de forma tão eficiente o substrato celulósico, sendo necessária uma ação sinérgica (31), como visto na [Figura 2](#), para um melhor rendimento (32–35). Cada uma das enzimas desempenha um papel complementar e importante na despolimerização da biomassa como segue:

- (i) as **endoglucanases** (endo-1,4- β -D-glucanase) (E.C. 3.2.1.4) são enzimas responsáveis pelo início da hidrólise clivando as ligações β -1,4-glicosídicas na cadeia de celulose, de forma randômica, nas regiões amorfas e liberando oligossacarídeos de vários graus de polimerização com terminais redutores e não redutores (36, 37);
- (ii) as **exoglucanases** (celobiohidrolases) (EC 3.2.1.91) são enzimas que atacam as fibras da celulose, atuam nos terminais redutores das cadeias de celulose, produzindo cadeias com extremidades livres, clivando e liberando celobiose (dímero de glicose) (36, 38–40);
- (iii) e as **β -glucosidases** (EC 3.2.1.21) hidrolisam a celobiose, liberando duas moléculas de glicose (41).

No modelo de ação das CAZymes, como visto na [Figura 2](#), a hidrólise das cadeias de celulose é desempenhada pelas principais celulases, seguindo uma distribuição de ações conforme destacado anteriormente pelas características de cada categoria de enzima. Tipicamente, exocelulases são enzimas processivas, isto é, estão ligadas à cadeia de celulose até que esteja completamente hidrolisada, enquanto endocelulases podem ser ambas processivas e não-processivas (30). Entretanto, apenas a ação conjunta destas enzimas não hidrolisa por completo a celulose. Para isso, as enzimas oxidativas LPMOs e LCCs (Lacases) contribuem para o aumento das taxas hidrolíticas (42, 43), são conhecidas por realizar a clivagem oxidativa de ligações glicosídicas em quitina e celulose, aumentando assim a atividade de enzimas despolimerizantes hidrolíticas bem conhecidas (44, 45).

1.2.2 Estrutura e função das Hidrolases de Glicosídeos

A extensa variedade de estereoquímica dos carboidratos é acompanhada por uma grande multiplicidade de enzimas envolvidas em seu metabolismo. As GHs (EC 3.2.1.x) são enzimas-chave para o metabolismo de carboidratos. Existe uma relação direta entre sequência e tipos de enovelamento (46), diversos autores (47–51) fizeram uma comparação sistemática das sequências de amino ácidos de GHs e encontraram (i) famílias contendo várias entradas de EC e (ii) enzimas com especificidades de substrato semelhantes que pertencem a famílias diferentes. Assim, propuseram uma classificação que refletisse melhor a similaridade de sequência (portanto estrutural) e apresentaram tal classificação para GHs. Algumas famílias de GHs foram alvos dos estudos executados neste trabalho e em nosso grupo de pesquisa, no [Quadro 1](#), onde se nota que diferentes famílias partilham de similares atividades conhecidas (GH1 e GH3), tanto como, o tipo de enovelamento do *core* dessas enzimas apesar de ser similar pode resultar em diferentes funções (GH32 e GH62). Estas classificações das famílias de GHs tornou-se possível devido ao crescente no número de sequências disponíveis e disponibilidade de novos métodos de comparação de sequências que podem lidar com baixa identidade de sequência (22, 47, 52).

Quadro 1 – Sumário geral de informações das famílias CAZymes estudadas neste trabalho e em nosso grupo de pesquisa. Dados obtidos no CAZy 2021.

Família	Nº de membros ^a	Nº de enzimas caracterizadas	Nº de estruturas ^b	Tipo de enovelamento
GH1	48637	329	77 (287)	barril (β/α) ₈
Atividades conhecidas				
β -glucosidase (EC 3.2.1.21); β -galactosidase (EC 3.2.1.23); β -mannosidase (EC 3.2.1.25); β -glucuronidase (EC 3.2.1.31); β -xylosidase (EC 3.2.1.37); β -D-fucosidase (EC 3.2.1.38); phlorizin hydrolase (EC 3.2.1.62); exo- β -1,4-glucanase (EC 3.2.1.74); 6-phospho- β -galactosidase (EC 3.2.1.85); 6-phospho- β -glucosidase (EC 3.2.1.86); strictosidine β -glucosidase (EC 3.2.1.105); lactase (EC 3.2.1.108); amygdalin β -glucosidase (EC 3.2.1.117); prunasin β -glucosidase (EC 3.2.1.118); vicianin hydrolase (EC 3.2.1.119); raucassicine β -glucosidase (EC 3.2.1.125); thioglucosidase (EC 3.2.1.147); β -primeverosidase (EC 3.2.1.149); isoflavonoid 7-O- β -apiosyl- β -glucosidase (EC 3.2.1.161); ABA-specific β -glucosidase (EC 3.2.1.175); DIMBOA β -glucosidase (EC 3.2.1.182); β -glycosidase (EC 3.2.1.-); hydroxyisourate hydrolase (EC 3.-.-.); β -rutosidase / α -L-rhamnose-(1,6)- β -D-glucosidase (EC 3.2.1.-); protodioscin 26-O- \hat{I}^2 -D-glucosidase (EC 3.2.1.186).				
GH3	42265	317	51 (173)	barril (β/α) ₈
Atividades conhecidas				
β -glucosidase (EC 3.2.1.21); xylan 1,4- β -xylosidase (EC 3.2.1.37); β -glucosylceramidase (EC 3.2.1.45); β -N-acetylhexosaminidase (EC 3.2.1.52); α -L-arabinofuranosidase (EC 3.2.1.55); glucan 1,4- β -glucosidase (EC 3.2.1.74); isoprimeverose-producing oligoxyloglucan hydrolase (EC 3.2.1.120); coniferin β -glucosidase (EC 3.2.1.126); exo-1,3-1,4-glucanase (EC 3.2.1.-); β -N-acetylglucosaminide phosphorylases (EC 2.4.1.-); β -1,2-glucosidase (EC 3.2.1.-); β -1,3-glucosidase (EC 3.2.1.-); xyloglucan-specific exo- β -1,4-glucanase / exoxyloglucanase (EC 3.2.1.155); stevioside- β -1,2-glucosidase (EC 3.2.1.-); lichenase / endo- β -1,3-1,4-glucanase (EC 3.2.1.73); protodioscin 26-O- \hat{I}^2 -D-glucosidase (EC 3.2.1.186).				
GH32	17627	237	21 (72)	5 α β -propeller
Atividades conhecidas				
invertase (EC 3.2.1.26); endo-inulinase (EC 3.2.1.7); β -2,6-fructan 6-levanbiohydrolase (EC 3.2.1.64); endo-levanase (EC 3.2.1.65); exo-inulinase (EC 3.2.1.80); fructan β -(2,1)-fructosidase/1-exohydrolase (EC 3.2.1.153); fructan β -(2,6)-fructosidase/6-exohydrolase (EC 3.2.1.154); sucrose:sucrose 1-fructosyltransferase (EC 2.4.1.99); fructan:fructan 1-fructosyltransferase (EC 2.4.1.100); sucrose:fructan 6-fructosyltransferase (EC 2.4.1.10); fructan:fructan 6G-fructosyltransferase (EC 2.4.1.243); levan fructosyltransferase (EC 2.4.1.-); [retaining] sucrose:sucrose 6-fructosyltransferase (6-SST) (EC 2.4.1.-); cycloinulo-oligosaccharide fructanotransferase (EC 2.4.1.-).				
GH62	637	25	10 (19)	5 α β -propeller
Atividades conhecidas				
α -L-arabinofuranosidase (EC 3.2.1.55).				

^a Estruturas primárias disponíveis no GenBank;

^b Estruturas disponíveis no PDB, valores entre () incluem eventuais estruturas de complexos com diferentes ligantes para uma mesma proteína.

Fonte: Elaborada pelo autor.

Bgls: as β -glicosidases (EC 3.2.1.21) são enzimas que hidrolisam as moléculas de celobiose e pequenos oligossacarídeos solúveis liberando moléculas de glicose [Figura 3 \(39, 53\)](#). As Bgls foram classificadas nas famílias de GHs: GH1, GH3, GH5, GH9 e GH30, com base em suas sequências de aminoácidos. As Bgls GH1, GH5 e GH30 compõem o GH Clan A, que consiste em proteínas com estruturas de enovelamento em barril (β/α)₈. Em contraste, o sítio ativo das enzimas GH3 compreende dois domínios, enquanto as enzimas GH9 possuem estruturas em barril (β/α)₆ ([54](#)).

Elas são cruciais para a hidrólise eficiente da celulose. Estas enzimas são responsáveis pela hidrólise da celobiose que é liberada no meio reacional, no primeiro passo de conversão da celulose em glicose [Figura 2](#), tendo em vista que em excesso, a celobiose inibe a ação das endoglucanases e exoglucanases, impedindo um processo de degradação eficiente ([55, 56](#)). A inibição das enzimas celulolíticas desempenha um papel importante na taxa de hidrólise e rendimento na formação de glicose. Glicose e celobiose reduzem significativamente a

taxa reacional, bem como o rendimento de formação destes produtos, onde ambas são consideradas inibidores das enzimas celulolíticas (57). A Celobiose, produto da ação das celobiohidrolases e em parte das endoglucanases, é inibidor direto dessas enzimas. Na prática, a glicose é um inibidor indireto, pois a ação sobre a β -glicosidase gera um acúmulo de celobiose, vindo a intensificar a inibição sobre as celobiohidrolases e as endoglucanases (58).

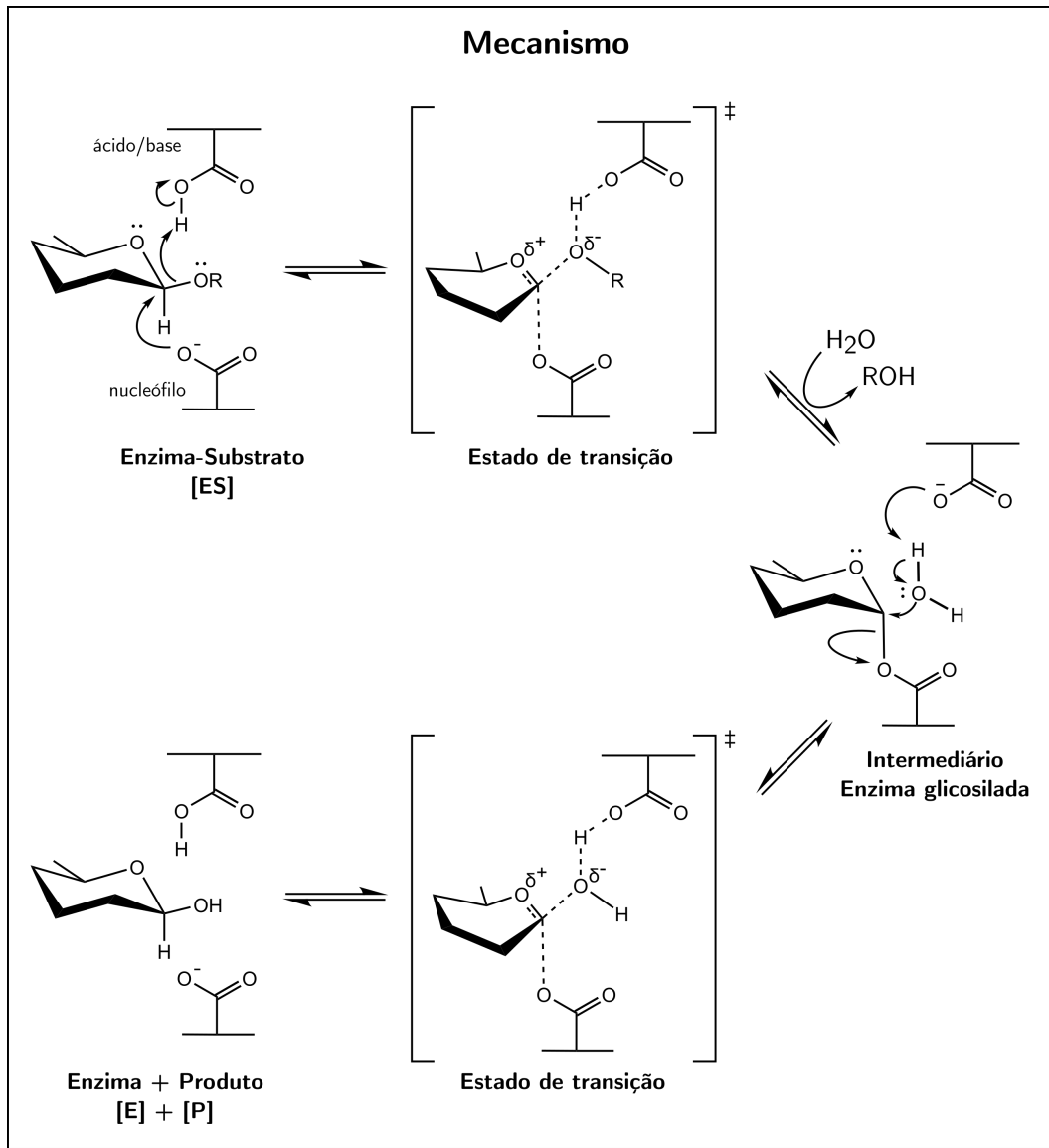


Figura 3 – Esquema do mecanismo da reação de hidrólise realizada pelas β -glicosidases, mostrando as duas etapas de glicosilação e desglicosilação, passando pela formação de um intermediário covalentemente ligado ao substrato.

Fonte: Elaborada pelo autor.

Por esse motivo, é essencial estudar a atividade de hidrólise das β -glicosidases a fim de resolver este problema, uma vez que estas enzimas são especializadas na quebra dos dímeros de glicose (celobiose). Pesquisas feitas atualmente dão mais foco à degradação da

celulose cristalina à celobiose, negligenciando estudos mais aprofundados da hidrólise da celobiose (23,59), o que tem sido uma adversidade enfrentada.

Abfs: as α -L-arabinofuranosidasas (EC 3.2.1.55) desempenham um papel fundamental na o processo de liberação de L-arabinose de polissacarídeos ou arabinooligossacarídeos atuando em resíduos de L-arabinofuranosil ligados a α -(1→2)-, α -(1→3)- ou α -(1→5)-da extremidade não redutora de poli e oligossacarídeos (60), aumentando a hidrólise das ligações glicosídicas de polissacarídeos por outras enzimas (61). As Abfs são agrupados em seis famílias distintas de glicosídeo hidrolases: GH2, GH3, GH5, GH39, GH43, GH51, GH54 e GH62 (52,62,63). Com base em suas sequências de aminoácidos, os membros da família GH62 foram divididos nas subfamílias GH62_1 e GH62_2 (64). Além disso, as famílias Abfs são ainda classificadas em três tipos com base na divergência entre si quanto à especificidade dos substratos (65,66).

Essas enzimas podem liberar arabinose de muitos substratos e fazem parte do sistema xilanolítico, sendo essenciais para romper as ligações arabinoxilanas (60,67). Portanto, Abfs são notavelmente importância biotecnológica (60), que se estende desde o uso em indústrias voltadas para a agricultura (68), alimentícia (69), e também em coquetéis para digestão de ração animal (68), valorização do sabor dos vinhos (68,70), melhoria na clarificação de sucos de frutas (68,71), e para a deslignificação da celulose na produção de biocombustíveis (72,73).

BFfases: as β -D-frutofuranosidasas (EC 3.2.1.26), pertencentes à família GH32, ou invertases (EC 3.2.1.26), são originalmente descritas como enzimas hidrolisadoras de sacarose, também possuem capacidade de transfrutossilagem em condições ótimas. Às vezes, eles são confundidas com “frutosiltransferases” (ou seja, sacarose 1-frutosil transferase ou inulosucrase; EC 2.4.1.9), as transglicosidasas produtoras de frutooligossacarídeos. Ambas as enzimas pertencem à família GH68 e processam a sacarose como substrato; entretanto, as frutosiltransferases realizam apenas a transferência da porção frutosil com mínima ou nenhuma hidrólise de sacarose, produzindo oligo- e polímeros de frutose. É certo que a fronteira entre esses dois grupos de enzimas não é nítida, pois a razão hidrólise/síntese difere fortemente de acordo com as condições de reação e a fonte particular.

As famílias GH32, GH43, GH62 e GH68 em suas estruturas, apesar de no geral apresentarem baixa homologia de sequência (<15% de identidade), elas compartilham o enovelamento em comum *5-fold β -propeller*, portanto, são combinados no clã GH-J (63,74,75). Além de três resíduos ácidos estritamente conservados (Asp, Asp e Glu) identificados em três blocos segundo os autores (75), operando com os mecanismos canônicos de reações, de inversão (GH43) ou retenção (GH32 e GH68) das configurações do carbono anomérico. Onde o primeiro Asp é reconhecido como a base nucleofílica e o Glu como o doador de próton (75).

A β -D-frutofuranosidase e suas habilidades de transfrutossilagem foram descobertas

observando-se a formação de outros produtos além de glicose e frutose, como frutanos com grau e polimerização DP2 e DP6 (76, 77), durante a hidrólise da sacarose catalisada por essas enzimas. Importantes principalmente, relacionadas a obtenção de bioprodutos como os frutanos ou fruto-oligossacarídeos (FOS), que também podem ser obtidos por meio de reações catalisadas por essas enzimas encontradas em fungos, bactérias ou plantas (78, 79).

1.3 Microorganismos termófilos e mesófilos

Os microrganismos termófilos são de grande interesse científico, principalmente em relação ao seu potencial biotecnológico, como a produção de enzimas termoestáveis (80, 81). A termofilia requer adaptações fisiológicas especializadas, pois as proteínas e ácidos nucleicos não podem ser desnaturados e a membrana deve manter-se funcional em altas temperaturas. As proteínas possuem um papel extremamente importante na adaptação dos termofílicos. O aumento da estabilidade destas enzimas é um resultado de várias mudanças nos aminoácidos, com efeitos cumulativos na estabilidade em relação à estrutura de proteínas homólogas de organismos mesofílicos (82, 83).

Algumas bactérias termofílicas apresentam crescimento ótimo em temperaturas bastante elevadas ($\approx 60 - 100\text{ }^{\circ}\text{C}$), e estão recebendo atenção especial como fontes de enzimas altamente termoestáveis (84, 85). A estabilidade intrínseca de proteínas de microrganismos termofílicos geralmente é maior que sua equivalente mesofílica, mesmo mantendo estrutura básica particular da família a qual pertencem (86). Organismos termofílicos podem ser encontrados em fontes geotérmicas, fontes vulcânicas (que expelem vapores e compostos sulfurados), fontes termais marinhas, onde erupções vulcânicas elevam a temperatura para mais de $100\text{ }^{\circ}\text{C}$.

Para aplicações industriais, as enzimas termoestáveis de maneira geral apresentam vantagens nos processos biotecnológicos conduzidos em temperaturas elevadas, diminuindo assim o risco de contaminação por microrganismos mesofílicos, já que estes constituem maior número em um ambiente industrial. Essa temperatura mais elevada ainda favorece a solubilidade dos substratos e produtos, aumentando as taxas de reação por redução da viscosidade e coeficiente de difusão dos substratos.

As enzimas termoestáveis possuem a característica de maior resistência às proteases, pois quanto menor o grau de maleabilidade da molécula, menos seus sítios de proteólise ficam expostos. Além disso, estudos mostraram que solventes orgânicos também tem ação limitada à desnaturação das proteínas termoestáveis (86).

1.3.1 *Clostridium thermocellum*

A bactéria *Clostridium thermocellum* (*C. thermocellum*), extraída do algodão, é flagelada, anaeróbica, gram-positiva e que produz esporos em forma de bastonetes. É um organismo termofílico com uma das maiores taxas de hidrólise de substratos cristalinos

(87–89). Esse organismo, possui uma eficiente forma de degradação da celulose, com a utilização de um complexo multienzimático extracelular chamado celulosomo (87, 90–92). O celulosomo demonstra vantagens na hidrólise enzimática de substratos celulósicos, pois atua predominantemente sobre a celulose cristalina (17, 93).

Recentes estudos identificaram os mecanismos moleculares de como o celulosomo é formado e o modo com que o complexo celulosomal é apresentado sobre a superfície da bactéria hospedeira (92–94). A arquitetura do celulosomo (Figura 4a) consiste no recrutamento de diversas enzimas hidrolíticas em uma subunidade não-catalítica chamada de escafoldina (91, 94–98). A integração e reconhecimento do complexo são feitos por meio da interação entre dois domínios proteicos existentes nas subunidades enzimáticas (doquerina) e na escafoldina (coesina) (94, 98, 99). O complexo multienzimático celulosomal da bactéria anaeróbica *C. thermocellum* Figura 4b consiste em: doquerinas individuais contendo enzimas integradas a 9 coesinas de uma escafoldina primária. Um conjunto de escafoldinas ancoradas possuem 1, 2 e 7 coesinas, respectivamente, para a fixação dos números correspondentes de escafoldina CipA primária e seu complemento de enzima a superfície da célula (27).

As escafoldinas, geralmente, possuem um módulo de ligação à celulose, do inglês *Carbohydrate-Binding Module* (CBM) não-catalítico, responsável por ancorar o complexo enzimático e fazer o reconhecimento do substrato polissacarídico, caracterizando então o celulosomo como uma maquinaria altamente eficiente na degradação de celulose (100). As celulasas expressas em *C. thermocellum* estão divididas em diversas famílias de hidrolases de glicosídeos, como as GH1, GH3 (alvos do estudo).

1.3.1.1 BGI de *C. thermocellum*

Através de um levantamento sobre estruturas de β -glicosidases depositadas no banco de dados de proteínas, PDB (<http://www.rcsb.org/pdb>), observamos que, quando a busca foi realizada com a expressão “*beta-glicosidases*”, para organismos em geral, observamos um total de 673 estruturas depositadas. Dessas, 314 são pertencentes as classes taxonômicas do domínio Eucarioto, 318 ao domínio Bactéria, 19 ao domínio Arquea, 21 a outros prospectados de sequências não classificadas e 2 para outras sequências (dados levantados em fevereiro, 2022). Em junho de 2018, o número total era de 348 estruturas.

Estudos mostraram que a maquinaria do celulosomo para degradação enzimática de celulose de *C. thermocellum* é de 50 a 100 vezes mais eficiente do que as enzimas de fungos filamentosos (complexos celulolíticos) que são explorados comercialmente para a hidrólise da celulose (27, 32, 101), dessa maneira, considerando os pontos abordados, fica evidente que as enzimas da bactéria *C. thermocellum* (em especial as β -glicosidases GH1 e GH3 alvos deste trabalho), são extremamente interessantes e que resultados, não apenas estruturais, mas também biofísicos e bioquímicos, podem ser relevantes na literatura

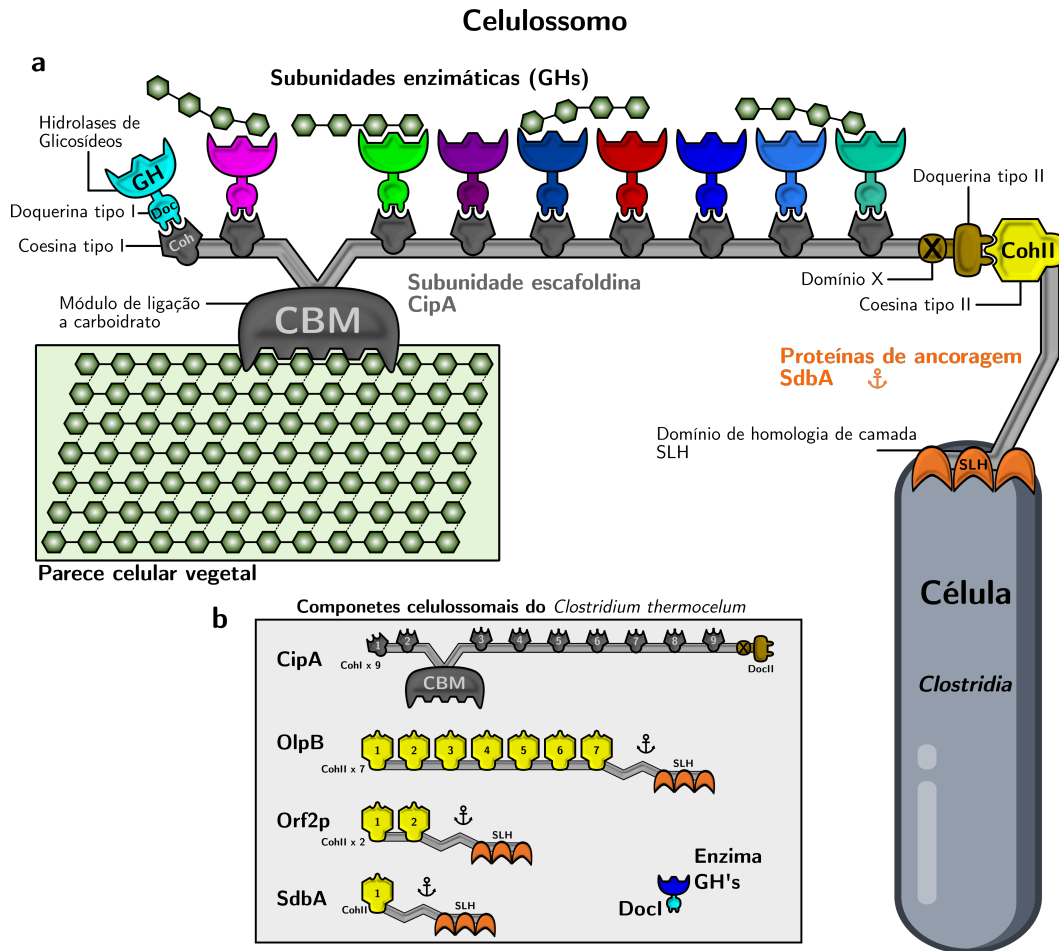


Figura 4 – (a) Esquema ilustrativo de um complexo enzimático celulosomal bacteriológico e seus componentes. (b) Esquema ilustrativo detalhando os componentes do celulosomo da bactéria anaeróbica *Clostridium thermocellum*.

Fonte: Adaptada de Bayer *et al.*(27).

científica.

Como existem poucas informações a respeito de Bgls de *C. thermocellum*, o objetivo principal deste projeto de pesquisa é tentar identificar e compreender as possíveis diferenças entre Bgl de *C. thermocellum* (constituente de celulosomo) em suas diferentes famílias, tendo como alvos a *CtBgl1A* da família GH1 e a *CtBgl3B* da família GH3. Neste trabalho não houve sucesso na produção da *CtBgl1A*, pois esta não foi purificada, apesar de ter sido transformada e expressa na fração solúvel, mesmo testando diferentes métodos de purificação não foi possível separar por cromatografia de afinidade, troca iônica, separação por exclusão de massa, nem por método desnaturante com uréia. Em contrapartida, para a *CtBgl3B* desde a determinação estrutural, caracterização biofísica, e comparação de atividades bioquímicas mostra em detalhes seu arranjo 3D, organização de domínios, interfaces, sugerindo modos de interação, e permitindo elucidar diferenças entre as enzimas que esclareçam as eficiências enzimáticas. Especialmente neste caso, observou-se um

aumento na atividade enzimática quando as reações ocorreram na presença dos surfactantes Triton X-100 e Tween20. A compreensão da diferença de funcionamento dessas enzimas pode levar ao desenvolvimento de processos economicamente viáveis, isso por meio do aumento da eficiência das enzimas hidrolíticas, como por exemplo, o seu emprego na otimização dos processos de produção do bioetanol.

1.3.2 *Lactobacillus gasseri*

Lactobacillus gasseri é um microrganismo homofermentativo que habita o trato gastrointestinal de humanos e exerce função probiótica, principalmente devido à síntese de frutanos prebióticos (polímeros de frutose) derivados da sacarose e seus hidrolisados (102, 103). Até o momento, apenas alguns genes potenciais pertencentes a este organismo relacionados com a produção de FOS foram caracterizados (104).

Várias espécies de *Lactobacillus* têm seu habitat natural no trato gastrointestinal, e muitos estudos associam as frutossiltransferases ao *pool* de enzimas produzidas pelo microrganismo que gera ou hidrolisa os frutanos (105–107). A importância desses achados é inegável, e o presente trabalho tem como objetivo contribuir para o aumento das informações relacionadas a essas enzimas, pois as propriedades tecnológicas e probióticas ainda são limitadas (103, 108). Estudos de enzimas isoladas podem levar a uma melhoria e melhor controle de condições ao invés de trabalhar com um *pool* de enzimas.

1.3.2.1 BFfase de *L. gasseri*

No artigo publicado por nosso grupo (79), seção A.1, tive a oportunidade de dividir a primeira autoria devido aos esforços nele dedicados. Nele, relatamos a clonagem, expressão, purificação, cristalização e determinação da estrutura cristalina de uma sucrose-6-fosfato hidrolase (BFfase) de *L. gasseri* (DSM 20243/ATCC) pertencente à família GH32 — *LgAs32*. Estruturas Apo-*LgAs32* e *LgAs32* em complexo com β -D-frutose foram determinados nas resoluções de 1,94 e 1,84 Å, respectivamente. A produção de FOS, frutanos, 1-kestose e nistose pela *LgAs32* recombinante, utilizando sacarose como substrato, mostrada neste estudo é bastante promissora, principalmente pela dupla capacidade da enzima tanto de hidrólise da sacarose como a produção de 1-kestose nas reações. Quando comparada à sua enzima homóloga de *Lactobacillus reuteri*, a produção de 1-kestose por *LgAs32* é aumentada; assim, *LgAs32* pode ser considerado uma alternativa na produção de frutanos e outras aplicações industriais.

1.3.3 *Thermothielavioides terrestris*

Fungos termofílicos e termotolerantes também são interessantes, pois suas enzimas geralmente podem suportar condições adversas usadas na indústria, como altas temperaturas, pHs extremos e solventes químicos (109). O fungo filamentoso *Thermothielavioides*

terrestris, pertence ao filo Ascomycota e está entre um número limitado de espécies eucarióticas classificadas como termófilas mostrando crescimento ótimo acima de 45 °C (110). Muitas enzimas caracterizadas de *T. terrestris*, como celulases e hemicelulases, têm sido atrativas para aplicações biotecnológicas envolvendo hidrólise de biomassa devido aos seus perfis catalíticos e de estabilidade favoráveis (110–113).

Num trabalho recente de Tölgo *et al.*(114), foram reportadas análises do genoma e do transcriptoma da cepa LPH172 de *T. terrestris* e revelaram em detalhes a maquinaria enzimática usada pelo fungo para quebrar a biomassa lignocelulósica. Usando dados de transcriptoma de crescimento em glicose, Avicel, palha de arroz e xilano de madeira de faia, os autores mostraram que o fungo depende de uma estratégia centrada em LPMO quando cultivado em substratos celulósicos e mais em hemicelulases quando cultivado em xilano. Suas características bioquímicas, incluem a capacidade de liberar xilo-oligossacarídeos e celo-oligossacarídeos, por exemplo, durante o pré-tratamento enzimático de biomassa vegetal para a produção de bioetanol, podendo ser vantajoso para aplicações (111).

1.3.3.1 Abf62 de *T. terrestris*

No artigo publicado por nosso grupo (115), seção A.2, tive a oportunidade de dividir a primeira autoria devido aos esforços nele dedicados. Visando contribuir para uma melhor compreensão das enzimas da família GH62, neste estudo descrevemos uma Abf (subfamília GH62_2; aqui denominada *TtAbf62*) do fungo termofílico *T. terrestris*. As características bioquímicas do *TtAbf62* purificada foram determinadas usando arabino de beterraba (SBA), arabinoxilano de trigo (WAX) e *p*NP-Araf como substratos. Além disso, os resultados permitiram inferir sobre a hidrólise de SBA, WAX e *p*NP-Araf por *T. terrestris*. Os resultados biofísicos e bioquímicos visam promover um melhor entendimento de seu mecanismo catalítico na degradação de biomassa renovável, tornando-os potenciais candidatos para aplicações biotecnológicas nas indústrias de biocombustíveis e alimentos. Além disso, análises estruturais revelaram a presença de uma troca de domínio, provavelmente um artefato de cristalização. No entanto, se trata de uma arquitetura que preservou sítio ativo da enzima onde a cavidade catalítica foi estruturalmente conservada entre a troca do domínio de uma das partes do β -*propeller*, conferindo estabilidade à proteína.

1.3.4 *Bacillus licheniformis*

Bacillus licheniformis é uma bactéria gram-positiva, comumente encontrada no solo e difundida no meio ambiente, que pode crescer em várias temperaturas variando de 37 °C a 55 °C, portanto, considerada mesófila (116, 117) ou termotolerante (118). Pode existir na forma de esporos aerobicamente formados para resistir a ambientes hostis, Sadiq *et al.*(119) também relataram a capacidade de seus esporos de sobreviver ao tratamento térmico de

115 °C por 30 min. *B. licheniformis* é conhecido por sua deterioração de alimentos, mas também seu potencial de intoxicação alimentar (118).

Muitos isolados de *B. licheniformis* foram relatados como potencialmente causadores de deterioração do leite na indústria de laticínios (119), e amplamente utilizados para a produção industrial em larga escala de proteínas (120), apresentando a capacidade de produzir proteases (121, 122), lipases (123, 124) e β -galactosidasas (125, 126). Esta bactéria também foi relatada devido à sua capacidade de aderir e formar biofilme no aço inoxidável durante a fabricação de laticínios em várias superfícies incluindo dutos, trocadores de calor por placas e evaporadores, destacando alguns dos problemas para a indústria de laticínios (127, 128).

1.3.4.1 *BLYmD* de *B. licheniformis*

A bactéria mesofílica *Bacillus licheniformis* tem demonstrado grande potencial para produção de enzimas industriais, devido a sua alta capacidade de secretar proteases. Com base na análise proteômica e de sequências, uma hipotética proteína CNF1/YfiH-like cisteína hidrolase putativa de *B. licheniformis* *BLYmD* chamou a atenção devido suas semelhanças sequenciais com domínios de lacases, bem conhecidos, e se tornou o objeto de estudo nesta pesquisa. *BLYmD* foi expressa em *E. coli* (Rosetta DE3) e purificada, tivemos sua estrutura cristalina determinada na resolução de 2,6 Å. A estrutura da *BLYmD* é classificada como uma proteína α/β que forma quatro camadas de $\alpha/\beta/\beta/\alpha$ com folhas beta antiparalelas mistas e se reúne em um dímero na unidade assimétrica.

Foram aplicadas ferramentas de bioinformática para identificar sete proteínas em diferentes organismos com razoável identidade sequencial e homologia estrutural com a superfamília do tipo CNF1/YfiH. Coletivamente, essas sequências são atribuídas ao domínio da lacase polifenol oxidoreductase multi-cooper (PF02578 Cu-oxidase_4) e ao domínio de função desconhecida (DUF152). Contudo, foi recentemente atribuída a função de purina nucleosídeo fosforilase (129), onde os autores observaram que a porção C-terminal de uma proteína humana relacionada com o ciclo de nucleotídeos de purinas, seria equivalente ao domínio DUF152 sendo este responsável por multivariadas atividades enzimáticas.

Uma análise comparativa entre *BLYmD* e suas estruturas homólogas, mostrou que elas compartilham a díade catalítica hipotética, proposta para a superfamília das cisteína hidrolases putativas como CNF1/YfiH, os resíduos Cys132 e His149 centralmente posicionados na cavidade aberta revelada pelo programa *Caver 3.0*, uma similaridade estrutural que pode ser observada nessas proteínas hipotéticas. Assim, a acessibilidade da via e a díade cisteína-histidina característica que representam o domínio catalítico hipotético de várias proteínas, como, de fator necrotizante citotóxico bacteriano (CNF), de toxina dermonecrótica (DNT) ou a de fator de edição de peptidoglicano (PgeF). Os íons zinco estão frequentemente presentes nas estruturas desta superfamília, como no caso do

BIYlmD, onde dois íons zinco são observados adotando uma coordenação tetraédrica para os resíduos de cisteína. Os resultados dessas comparações são apresentados no [Capítulo 5](#).

2 OBJETIVOS



Este capítulo reporta os objetivos do presente trabalho e a justificativa da realização do mesmo como projeto para Tese de doutorado.

2.1 GH1 e GH3

- Expressar em *E. coli* e purificar as β -glicosidases da bactéria termófila *C. thermocellum*;
- Realizar análises bioquímicas da atividade enzimática das enzimas em estudo;
- Estudar a organização e estabilidade das estruturas secundária e terciária das enzimas em diferentes condições físico-químicas, empregando espectroscopia de dicroísmo circular (CD) e espectroscopia de fluorescência (intrínseca e extrínseca), respectivamente, fluorimetria diferencial de varredura (DSF), espalhamento de luz (DLS);
- Conduzir estudos de análises da estrutura e flexibilidade das enzimas (em solução) empregando espalhamento de raios X a baixo ângulo (SAXS);
- Realizar ensaios de cristalização para obter cristais das proteínas, em estudo, que difratem a alta resolução;
- Determinar a estrutura tridimensional destas enzimas empregando cristalografia de macromoléculas utilizando difração de raios X;
- Comparar os dados obtidos das β -glicosidases de *C. thermocellum* com os dados obtidos paralelamente para as famílias GH1 e GH3 visando identificar e compreender possíveis diferenças entre elas já que as mesmas atuam em condições diferenciadas (como constituinte de um celulosomo);

- Utilizar todos os resultados obtidos para compreender melhor a relação entre a estabilidade, flexibilidade e atividade enzimática das enzimas em estudo, tentando assim obter uma condição que permita a otimização da bioconversão de biomassa lignocelulósica em bioetanol através de rotas enzimáticas.

2.2 GH62 de *Thermotielavioides terrestris*

- Expressar em *E. coli* e purificar a enzima α -L-arabinofuranosidase da família GH62;
- Realizar discussões correlacionando estruturas de GH62 com as análises bioquímicas;
- Realizar novos ensaios de cristalização para obter co-cristais da α -L-arabinofuranosidase com os possíveis ligantes, que difratassem a alta resolução;
- Realizar análise estrutural descritiva comparando com as estruturas similares encontradas no PDB;
- Escrever, editar, e finalizar o respectivo artigo.

2.3 GH32 de *Lactobacillus gasseri*

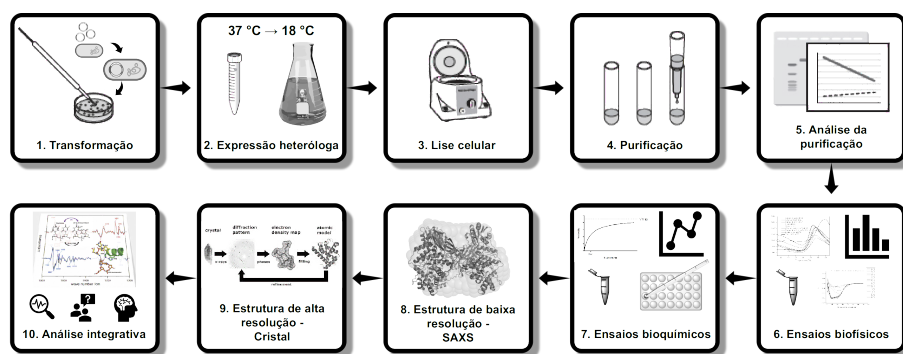
- Realizar a análise estrutural descritiva da enzima sucrose-6-fosfato hidrolase da família GH32, desde os dados de cristalografia e SAXS correlacionando com os aspectos bioquímicos previamente obtidos;
- Realizar discussões correlacionando estruturas de GH32, com as análises bioquímicas, indicando tanto as atividades como fructosiltransferase e hidrolítica;
- Utilizar análises de bioinformática que auxiliem na anotação de suas características.
- Escrever, editar, e finalizar o respectivo artigo.

2.4 *BIYlmD YfiH-like* de *Bacillus licheniformis*

- Expressar em *E. coli* e purificar a proteína hipotética YlmD YfiH-like;
- Realizar análises bioquímicas e identificar a atividade enzimática;
- Estudar a organização e estabilidade das estruturas secundária e terciária das enzimas em diferentes condições físico-químicas, empregando fluorimetria diferencial de varredura (DSF), e espalhamento de luz com ângulos variados acoplado com cromatografia de exclusão (SEC-MALS);
- Realizar novos ensaios de cristalização para obter co-cristais da *BIYlmD* com os possíveis ligantes, que difratem a alta resolução;

- Realizar análise estrutural descritiva comparando com as estruturas similares encontradas no PDB;
- Utilizar análises de bioinformática que auxiliem na anotação das características da proteína por meio da correlação estrutura-atividade.

3 MATERIAIS E MÉTODOS



Este capítulo é onde estão descritas todas as metodologias empregadas no trabalho.

3.1 Busca das sequências e bioinformática

As proteínas localizadas nos bancos de dados NCBI e CAZy (52, 62) com a ferramenta BLAST (130, 131) usada para a busca e alinhamento das sequências, alvos dos trabalhos, aqui apresentadas. Sendo elas: BGIs (*CtBgl1A*) da família GH1, e (*CtBgl3B*) da família GH3 da bactéria termofílica celulosossomal *C. thermocellum* disponíveis pelos códigos de acesso no NCBI: [ALX07828.1](#) e [ALX08008.1](#) respectivamente; proteína hipotética PNP (*BLYImD*) *YfiH-like* da bactéria *Bacillus licheniformis* (NCBI: [AEO64662.1](#)).

As medidas de distância dos pares de sequências foram obtidas usando o programa MEGA (132), e podem ser calculadas a partir de sequências de aminoácidos ou das regiões de codificação de sequências de nucleotídeos. Esta distância é a proporção (p) de sítio de aminoácidos onde às duas sequências a serem comparadas são diferentes, obtida por meio da divisão do número de diferenças de aminoácidos pelo número total de sítios comparados, sem fazer nenhuma correção para múltiplas substituições no mesmo sítio ou diferenças nas taxas evolutivas entre os sítios. Assim, temos $p = \frac{n_d}{n}$, enquanto sua variância é dada por $V(p) = \frac{p(1-p)}{n}$. Onde n_d e n são o número de diferenças de aminoácidos e o número total de aminoácidos comparados, respectivamente. A distância (p) é aproximadamente igual ao número de substituições de nucleotídeos por sítio (d) somente quando é pequena, ex.: $p < 0,1$ (133).

As análises estruturais usando as ferramentas de bioinformática foram realizadas utilizando o PDBeFOLD (134, 135) e Dali server (136) para estatísticas em alinhamento com estruturas semelhantes. ENDscript (137) para geração de imagens dos alinhamentos das sequências de aminoácidos e detalhamento das características da estrutura secundária. ProFunc (138) para correlacionar a função enzimática com base na estrutura tridimensional

de *BlyImD*. A categorização dos domínios estruturais das proteínas foi feita usando CATH (139), PFAM (140) e SCOP (141).

O servidor PDBePISA (142) foi aplicado para verificar estados oligoméricos avaliando interfaces e sua força de interação entre monômeros vizinhos gerados pela aplicação de simetria às coordenadas do arquivo de entrada no conjunto macromolecular, consequentemente, prevendo a forma biológica mais provável a partir das estruturas. O kit de ferramentas MOLE (143) e o *plugin* Caver 3.0 (144) foram aplicados para estimar cavidades das proteínas *CtBgl3B*, *BlyImD*, *TtAbf62* (115), *LgAs32* (79) e as possíveis vias de entrada de substratos. Todas as imagens estruturais foram elaboradas usando o programa PyMOL *Molecular Graphics System* (145).

As sequências obtidas e selecionadas dos bancos de dados NCBI e CAZy (52, 62) foram então traduzidas utilizando as ferramentas de bioinformática do servidor ExPASy (146, 147). Os genes que codificam as proteínas *CtBgl1A* e *CtBgl3B* já clonados das 2 enzimas escolhidas, para o trabalho principal, foram adquiridos com dinheiro da reserva técnica (Projeto FAPESP 2016/09152-6) junto a empresa FastBio. Conforme pedido, em vetor de expressão pET28a(+), conferindo resistência ao antibiótico canamicina à célula bacteriana. Além disso, o pET28a(+) permite a expressão da proteína recombinante com uma cauda no N-terminal (*His_(6x)-tag*), contendo seis histidinas, e um sítio de reconhecimento para a protease trombina.

Enquanto, os genes que codificam as proteínas *BlyImD* (AAU23292.2) foi clonado usando o método LIC (148, 149) em vetor pETTrx-1-a/LIC (Novagen). Desta forma, o uso de enzimas de restrição ou ligase não são necessários, possuindo um domínio de Trx fusionado a uma cauda contendo seis histidinas (*His_(6x)-Trx-tag*) adicionada as proteínas alvos, e um sítio de reconhecimento a protease TEV utilizada para clivagem deste domínio fusionado. Seguindo a mesma metodologia dos artigos seção A.1 (79) e seção A.2 (115).

A presença da cauda com 6xHis confere as proteínas recombinantes alta afinidade para interação com as resinas cromatográficas contendo íons metálicos, como por exemplo Ni^{2+} ou Co^{2+} . Assim, com os genes das proteínas inseridos, foi possível realizar a transformação em linhagens de bactérias competentes *E. coli* DH5 α , BL21(DE3) e Rosetta(DE3).

3.2 Produção das proteínas recombinantes

3.2.1 Transformação em *E. coli* quimio competentes

Para inserção dos vetores em célula bacteriana empregou-se o seguinte protocolo: em 10 μL de solução contendo os plasmídeos (0,1 μg), foram adicionados 40 μL de tampão de transformação (1 mL KCM 10x – 1 M de KCl, 0,3 M CaCl_2 , 0,5 M MgCl_2 ; 1,5 mL PEG 4000 10% m/v; 7,5 mL água *Milli-Q*® autoclavada) e 50 μL de células competentes.

A mistura foi incubada no gelo por 30 minutos e posteriormente à temperatura ambiente por 10 minutos para o choque térmico. Meio LB (*Lysogeny Broth* – 10 g L⁻¹ de Triptona; 5 g L⁻¹ de Extrato de levedura e 10 g L⁻¹ de NaCl) foi adicionado à mesma (1 mL) e incubada por 1 hora a 100 rpm, 37 °C. Meio ágar LB foi preparado para o isolamento dos transformantes adicionando-se os antibióticos nas seguintes concentrações cloranfenicol 35 µg/mL e canamicina 50 µg/mL (*E. coli* DH5α e BL21(DE3) apenas canamicina). 200 µL de amostra foram espalhados na superfície de 25 mL de meio com o auxílio de uma alça de Drigalski e incubado 16h (*overnight*) à 37 °C.

As células de *E. coli* DH5α servem como células de propagação gênica, onde por meio da utilização do Kit *MiniPrep*® foram recuperados os genes, para caso sejam necessárias futuras clonagens ou transformações isso possa ser feito partindo-se destes. Enquanto as células de *E. coli* BL21(DE3) e Rosetta(DE3) tem como objetivo sua utilização para a expressão heteróloga em larga escala das proteínas de interesse.

3.2.2 Testes de Expressão

Partindo de colônias isoladas das bactérias transformadas, foram inoculadas em meio LB líquido (5 mL, pré-inóculos) com canamicina (50 µL) e incubadas por 12 horas a 37 °C e 200 rpm, destas foram separadas alíquotas bacterianas condicionadas em frações glicerinadas (25%) para congelamento instantâneo com nitrogênio líquido e armazenadas a -80 °C. Partindo das células congeladas, foram preparados os pré-inóculos para os testes de expressão.

Os testes de expressão foram realizados como primeiro passo para estabelecer um protocolo de expressão heteróloga para *CtBgl1A* e *CtBgl3B*, para isso as seguintes variáveis foram consideradas: **1) Célula de expressão:** optando-se por *E. coli* BL21(DE3) e Rosetta(DE3). **2) Meio/Indução:** meio LB (*Lysogeny Broth* (150) – 10 g L⁻¹ g L⁻¹ de Triptona; 5 g L⁻¹ de Extrato de levedura e 10 g L⁻¹ de NaCl) expressão induzida com 1 mM de IPTG após atingir D.O_{600nm} entre 0,8-1,0 e meio auto-indutor (151) (3,3 g L⁻¹ de (NH₄)₂SO₄; 6,8 g L⁻¹ de KH₂PO₄; 7,1 g L⁻¹ de Na₂HPO₄; 0,15 g L⁻¹ de MgSO₄; 5 g L⁻¹ de glicerol; 0,5 g L⁻¹ de glicose; 2 g L⁻¹ de lactose; 10 g L⁻¹ de triptona; 5 g L⁻¹ de extrato de levedura); **3) Temperatura:** após atingir a D.O_{600nm} 0,8-1,0, foram testadas expressão à 18 e 37 °C respectivamente para ambos os meios de cultivo, sendo que para o meio LB a mudança para 18 °C foi feita logo após adição de IPTG.

3.2.3 Expressão e purificação das proteínas recombinantes

Expressão heteróloga: o protocolo estabelecido para ambas proteínas recombinantes, *CtBgl1A*, *CtBgl3B* e *BlyImD*, observando a sua obtenção na fração solúvel após a lise celular, respectivamente utilizando: **1) Célula de expressão:** *E. coli* BL21(DE3) e Rosetta(DE3); **2) Meio/Indução:** 1 L do meio auto-indutor (contendo canamicina 50

$\mu\text{g mL}^{-1}$), ou seja, não sendo necessária indução com IPTG; 3) Temperatura: após atingir a $D.O_{600nm}$ 0,6-0,8 mudança para 18 °C e mantido por 16 h.

Lise celular: as células foram colhidas por centrifugação a 4500 rpm durante 1 h a 4 °C, ressuspensas em tampão 50 mM fosfato de sódio pH 7,0, 250 mM de NaCl, e 10% de glicerol (*CtBgl1A* e *CtBgl3B*) e no tampão 50 mM Tris-HCl pH 8,0, 300 mM NaCl, 10% glicerol (*BIYImD*), ambos suplementados com 5 mM imidazol, 1 mM DTT, 1 mM de fenilmetilsulfonil fluoreto (PMSF) e 0,1 g L^{-1} de lisozima. As células foram rompidas por sonicação e centrifugadas por 45 min a 14000 rpm para remover detritos celulares. Os resultados das expressões foram analisados por eletroforese em gel desnaturante de poliacrilamida (SDS-PAGE) corado com azul de coomassie (152).

Purificação: para a *CtBgl3B* e *BIYImD* foram realizadas em três etapas: (1) IMAC Ni^{2+} , usando gradiente de concentração de imidazol no tampões para limpeza da resina (30, 500 mM) e para eluição (250 mM); (2), clivagem com trombina (para *CtBgl3B*) ou TEV protease (para *BIYImD*) durante 16h a 18 °C seguida de outra IMAC Ni^{2+} também em tampão com imidazol nas quantidades anteriores; (3), SEC usando uma coluna HiLoad 16/60 Superdex™ 200 e Superdex 75 16/60 (GE Healthcare Lifescience), respectivamente. o grau de pureza, ao nível qualitativo, das proteínas recombinantes foi verificado por gel 10% SDS-PAGE. Realizando-se a troca de tampão para 50 mM de sulfato de sódio em pH 7, 250 mM NaCl, e 10% de glicerol (*CtBgl3B*) e 50 mM Tris pH 7,5, 300 mM NaCl, 1 mM DTT e 5% de glicerol (*BIYImD*) durante a SEC. As proteínas purificadas foram então concentradas usando filtro concentrador de centrífuga Amicon™ Ultra-15 com membrana 30k e 15k (Merck Millipore). As concentrações das soluções de proteínas foram determinadas por meio da absorção de luz no comprimento de onda de 280 nm e com os coeficientes de extinção molar (ϵ), obtidos usando as sequências de aminoácidos de cada proteína com a ferramenta *ExpASy - ProtParam tool* (146), em espectrofotômetro modelo NanoDrop™ 2000 (Thermo Scientific).

3.3 Análises de estabilidade térmica

3.3.1 Espectroscopia de dicroísmo circular

A técnica de dicroísmo circular permite a obtenção de informações sobre a estrutura secundária de proteínas em solução, permitindo inferir sobre sua estabilidade. Os espectros de CD, foram medidos em um equipamento Jasco J-815 equipado com controle de temperatura, com amostras preparadas no tampão fosfato de sódio 50 mM pH 7,0, 150 mM de NaCl e glicerol 10% (*CtBgl3B*). Todos os dados foram coletados em uma cubeta de quartzo de 1 mm de caminho óptico, e os espectros medidos no intervalo de comprimento de onda de 200–270 nm. Foram medidos uma média de 8 acumulações para cada espectro com abertura de fenda de 0,2 nm. A contribuição do tampão foi subtraída em todos os experimentos realizados. Mudanças na estrutura secundária como função

da variação da temperatura (20–90 °C), foram monitoradas com incrementos de 5 °C. A estabilidade térmica foi caracterizada pela mudança da elipticidade ($\Delta\epsilon$) nos pontos de mínimo em 208 e 222 nm, permitindo obter valor da temperatura de transição de desnaturação (*denaturation midpoint* – T_m , °C) aplicando ajuste sigmoidal [Equação 3.1](#) pela função de *Boltzmann*:

$$y = A_2 + \frac{A_1 - A_2}{1 + \exp\left(\frac{X - X_0}{d_x}\right)} \quad (3.1)$$

Onde, A_1 corresponde ao sinal na região de pré-transição, A_2 descreve o sinal na região pós-transição, X_0 à T_m , d_x descreve a inclinação na região de transição, X é a temperatura e y é o sinal de elipticidade molar experimental.

3.3.2 Fluorimetria diferencial de varredura

Após a obtenção das proteínas *CtBgl3B* e *BlYlmD* solúveis e purificadas, buscou-se determinar as melhores condições de estabilidade por meio da escolha do tampão que fosse capaz de conferir tal característica, detalhe importante, capaz de influenciar no sucesso do processo de cristalização ([153](#)). O experimento de DSF foi realizado com este propósito, uma vez, que a técnica consiste na medida do aumento da fluorescência de uma sonda conforme ocorre a desnaturação de uma proteína quando submetida a uma variação de temperatura, permitindo o cálculo da temperatura do ponto médio de transição da curva de desnaturação (T_m) ([153](#)).

As medidas foram feitas em um equipamento *Real-time System* (Bio-Rad), utilizando a proteína purificada em concentração final de 1,5 mg mL⁻¹ incubada com os diferentes tampões [Tabela 1](#) para análise numa placa de 96 poços. Logo após a adição da sonda *SYPRO Orange* (Invitrogen), cuja diluição final foi de 1:2000, foram realizadas as medidas de fluorescência onde a excitação ocorre a 473 nm e emissão a 570 nm. As amostras nas suas respectivas condições foram aquecidas num intervalo de 25–90 °C, com todas as medidas feitas em triplicata. As análises das curvas de desnaturação e obtenção dos valores de T_m foram realizadas sob comparação utilizando o programa do próprio equipamento.

3.4 Espalhamento de luz (DLS) e SEC-MALS

Espalhamento de luz dinâmico (DLS) é uma técnica auxiliar de grande importância, para o trabalho com cristalografia de proteínas, aplicada na determinação de medidas hidrodinâmicas, polidispersividade e agregação das moléculas, parâmetros essenciais para a propensão de cristalização das biomoléculas. Nesse experimento, o objetivo era a avaliação do estado oligomérico das amostras de proteínas em solução para monitorar a condição da amostra a ser utilizada no preparo dos experimentos de cristalização. Para as medidas

Tabela 1 – Composição das condições de tamponamento testadas aplicando na análise de DSF.

Condição	pH	Tampão (100 mM)	Aditivo (300 mM)
A1	-	água	
A2	1,2	HCl	
A3	2,0	HCl	
A4	3,0	ácido cítrico	
A5	4,0	citrato de sódio	
A6	4,5	citrato de sódio	
A7	4,7	citrato de sódio	
A8	5,0	citrato de sódio	
A9	5,0	fosfato de potássio	
A10	5,5	citrato de sódio	
A11	5,5	fosfato de sódio	
A12	5,8	MES	
B1	6,0	fosfato de potássio	
B2	6,0	BIS-TRIS	
B3	6,2	MES	
B4	6,5	fosfato de sódio	
B5	6,5	cacodilato de sódio	
B6	6,5	BIS-TRIS	
B7	6,5	MES	
B8	6,7	BIS-TRIS	
B9	7,0	fosfato de potássio	
B10	7,0	HEPES	
B11	7,0	BIS-TRIS	
B12	7,3	acetato de amônio	
C1	7,5	fosfato de sódio	
C2	7,5	TRIS	
C3	8,0	imidazol	
C4	8,0	HEPES	
C5	8,0	TRIS	
C6	8,0	bicina	
C7	8,5	TRIS	
C8	9,0	Bicine	
C9	9,5	carbonato de sódio	
C10	10,0	carbonato de sódio	
C11	3,0	ácido cítrico	NaCl
C12	4,5	acetato de sódio	NaCl
D1	5,0	fosfato de potássio	NaCl
D2	5,5	BIS-TRIS	NaCl
D3	5,5	citrato de sódio	NaCl
D4	5,8	MES	NaCl
D5	6,0	BIS-TRIS	NaCl
D6	6,5	cacodilato de sódio	NaCl
D7	6,5	fosfato de sódio	NaCl
D8	7,0	BIS-TRIS	NaCl
D9	7,0	HEPES	NaCl
D10	7,3	acetato de amônio	NaCl
D11	7,5	TRIS	NaCl
D12	8,0	HEPES	NaCl

Fonte: Elaborada pelo autor.

foi utilizado o espectrofotômetro Zetasizer UV (Malvern), em concentrações de 1,2 e 5 mg mL⁻¹ no tampão de purificação. Foram registrados 16 scans em triplicata durante 8 minutos a 25° C, em cubeta de quartzo (caminho óptico 1,5 mm para volume de 12 µL) antes do início das medidas.

Os resultados preliminares obtidos no DLS, haja visto algumas de suas limitações, motivaram a realização de uma segunda técnica, mais versátil e completa, a cromatografia de exclusão por massa combinada com a detecção do espalhamento de luz de multi-ângulo (do inglês, *Size exclusion chromatography with Multiangle light scattering* — SEC-MALS), e o índice refrativo diferencial (dRI) sendo utilizado para medição de massa molecular absoluta e estados oligoméricos de proteínas, independentemente do tempo de retenção (154–158). Onde, um detector de dRI determina a concentração com base na mudança no índice de refração da solução devido à presença do analito. Um detector de MALS mede a proporção de luz espalhada por um analito em vários ângulos em relação ao feixe de laser incidente. Portanto, esta instrumentação determina a Mw independentemente do tempo de eluição, visto que a Mw pode ser calculada diretamente a partir da [Equação 3.2](#).

$$M = \frac{R(0)}{Kc\left(\frac{dn}{dc}\right)^2} \quad (3.2)$$

onde M é a massa molecular do analito, $R(0)$ é a razão de *Rayleigh* reduzida (ou seja, a quantidade de luz espalhada pelo analito em relação à intensidade do laser) determinada pelo detector MALS e extrapolada para o ângulo zero, c é a concentração em massa determinada pelo detector UV ou dRI, dn/dc o incremento do índice de refração do analito (essencialmente a diferença entre o índice de refração do analito e do tampão) e K uma constante óptica que depende das propriedades do sistema, como comprimento de onda e índice de refração do solvente (154). A coluna SEC é usada exclusivamente para separar as várias espécies em solução, de modo que elas entrem no MALS e nas células do detector de concentração individualmente. O tempo de retenção real não tem significado para a análise, exceto na medida em que resolve as espécies de proteína. Pela análise da variação angular da intensidade de espalhamento, o MALS também permite determinar o tamanho (raiz quadrática média do raio, R_g) das macromoléculas com raio geométrico maior que cerca de 12,5 nm (154).

SEC-MALS foi empregado com objetivo de observar os padrões de separação e as espécies oligoméricas encontradas em solução para as proteínas CtBgl3B e BlyImD realizando análises de MALS em tempo real. Para as medidas foi utilizado sistema SEC-MALS (WYATT) com coluna de SEC Superdex 200 HR 10/30, proteínas em concentração de 2 mg mL⁻¹ no tampão Tris-HCl 50 mM pH 7,0, 250 mM de NaCl e glicerol 10% para CtBgl3B. No caso de BlyImD avaliou-se a variação da concentração de proteína de 0,5 a 10 mg mL⁻¹, e também variando a concentração de sal NaCl, variando de 35 a 800 mM.

3.5 SAXS

Estudos da estabilidade e flexibilidade da *CtBgl3B* foi realizado por meio da técnica de espalhamento de raios X a baixo ângulo (SAXS). Os dados foram coletados na linha de SAXS do Laboratório Nacional de Luz Sincrotron (LNLS, Campinas). O comprimento de onda dos raios X incidentes é de 1,488 Å. Os raios X espalhados são medidos utilizando-se um detector CCD bidimensional (MarResearch, USA). A distância entre a amostra e detector fixada entre 800 e 1000 mm. As amostras foram centrifugadas durante 15 minutos a 14.000xg (4 °C) antes de cada medida para remover agregados.

As amostras serão colocadas em uma célula com paredes finas de mica. Dois frames sucessivos de 300 segundos foram medidos para monitorar danos na amostra biológica e também a estabilidade do feixe. A integração dos dados de SAXS foi realizada com o programa *Fit2D* (159). As concentrações proteicas foram de 1, 5 e 10 mg mL⁻¹. As análises foram feitas por meio da análise de *Guinier* utilizando os programas *PRIMUS* (160) e *GNOM* (161) do pacote *ATSAS v2.8* (162).

3.6 Metodologia cristalográfica

3.6.1 Cristalização

Ensaio de cristalização foram realizados com auxílio de robô (Phoenix robot — Art Robbins Instruments) utilizando os kits comerciais *PACT premier*TM e *JCSG-plus*TM(163), *SG1*TM(164), *BCS* (165), e *MORPHEUS*[®](166). Nas condições deste último kit de cristalização foram obtidos os melhores cristais para os experimentos de difração, para isso, a proteína *CtBgl3B* foi concentrada até 25 mg mL⁻¹ com filtro *Amicon Ultra MWCO* 30 kDa (Millipore, EUA) (tampão 10 mM fosfato de sódio, pH 7,0, com 250 mM NaCl e 10% glicerol) foi cristalizada com método da gota sentada (do inglês, *sitting drop*) por difusão de vapor em gotas com as proporções das gotas µL 1:1, 2:1 e 1:2 de solução de proteína e solução da condição, neste caso os melhores cristais, sendo 0,1 M Tris (base); BICINA pH 8,5, com 0,12 M de mistura de álcoois como aditivos (0,2 M 1,6-Hexanodiol; 0,2 M 1-Butanol; 0,2 M 1,2-Propanodiol; 0,2 M 2-Propanol; 0,2 M 1,4-Butanodiol; 0,2 M 1,3-Propanodiol), e 37,5% v/v de mistura de precipitantes (25% v/v MPD; 25% PEG 1000; 25% w/v PEG 3350). Cristais em forma de placas foram observados após 5 dias a 18 °C. Para a coleta de dados, os cristais foram resfriados instantaneamente com nitrogênio líquido diretamente a partir das gotas de cristalização.

A proteína *BlymD* foi inicialmente concentrada a 10 mg mL⁻¹ com filtro *Amicon Ultra MWCO* 10 kDa (Millipore, EUA) e os cristais foram obtidos a 18 °C pelo método de difusão de vapor em gota sentada utilizando uma solução precipitante contendo 200 mM de cloreto de cálcio, 100 mM Tris-HCl pH 8,5 e 20% PEG 4000. O volume e a proporção da gota foi de 2 µL (1:1), o volume do reservatório de 65 µL. O aglomerado de placas

foi separado com agulha de acupuntura e um único cristal transferido para uma solução crioprotetora com composição semelhante ao reservatório, suplementada com etilenoglicol a 15%. O cristal de proteína foi então colhido e resfriado em uma corrente de vapor de nitrogênio antes da coleta de dados.

3.6.2 Coleta de dados, processamento e determinação estrutural

Experimentos de difração de raios X para os cristais selecionados da *CtBgl3B* foram realizados a 100 K, na linha I24: *microfocus* MX, das instalações de luz síncrotron do *Diamond Light Source (UK's national synchrotron science facility, Harwell Science and Innovation Campus* em Oxfordshire). Os conjuntos de dados da difração foram obtidos usando detector Pilatus3 6M (DECTRIS). Processamento das imagens foi realizado usando o conjunto de programas Xia2 (167), indexadas e integradas com o programa XDS (168, 169), escalonadas com o programa Aimless (170, 171) ambos integrados ao pacote de programas CCP4 (172). O problema de anisotropia foi corrigido partindo do arquivo de reflexões não-mergeadas por meio do servidor STARANISO (173).

Os dados de difração da *BIYlmD* foram coletados num difratômetro APEX DUO (Bruker AXS) usando o comprimento de onda do Cobre ($\text{CuK}\alpha$ 1,5418 Å). As imagens de difração foram processadas usando o pacote de programas PROTEUM2 (174).

3.6.3 Tratamento da Anisotropia

Os conjuntos de dados de intensidade [$I_{(hkl)}$] do cristal da enzima *CtBgl3B* exibiram severa anisotropia, isto é, a queda de intensidade com resolução irá variar com a direção, hkl , e pode ser muito mais alta ao longo de um eixo do cristal do que ao longo de outro, como representado na Figura 5. Portanto, trata-se da propriedade da difração em ser direcionalmente dependente (anisotropia). Isso afeta diretamente duas propriedades anisotrópicas dos dados de difração de raios X medidos: o limite de difração, ou 'limite de resolução', e a atenuação da intensidade difratada devido ao movimento térmico e/ou desordem.

Se a estrutura deve ser refinada com um modelo atômico isotrópico, seja porque não há dados suficientes ou os programas usados não podem lidar com parâmetros anisotrópicos, então a queda dos valores de fatores de estrutura calculados (F_{cal}^2) serão, necessariamente, também isotrópicos. Nesta situação, uma melhora entre os valores de fatores de estrutura observados (F_{obs}^2) e F_{cal}^2 pode ser obtida usando um escalonamento anisotrópico durante a redução de dados para a distribuição de *Wilson* esperada das intensidades, ou incluindo um máximo de seis parâmetros anisotrópicos gerais durante o refinamento. Isso resultará em um conjunto isotrópico de valores para os F_{hkl}^2 . Para cristais com um alto grau de anisotropia nos dados experimentais, isso pode levar a uma queda substancial nos valores de R e R_{free} (175, 176).

Essa ambiguidade desaparece efetivamente com o uso de um modelo atômico anisotrópico. Os ADPs individuais, incluindo contribuições de desordem estática e térmica, assumem deslocamentos individuais, mas também a anisotropia geral dos valores F_{hkl}^2 experimentais. A importância da anisotropia pode representar imperfeição assimétrica do cristal ou deslocamento anisotrópico de moléculas na rede cristalina relacionado aos parâmetros TLS. O refinamento dos parâmetros TLS foram realizados utilizando *TLSMD web server* (177,178) e *PHENIX.REFINE* (179), que remove a contribuição geral do cristal para os ADPs.

O limite de difração: além do limite de difração da amostra, as intensidades mescladas (*merged*) de reflexões únicas caem, em média, abaixo do nível de significância estatística, portanto, devem ser consideradas não informativas. A significância estatística dos dados de intensidade após a mescla é comumente expressa como o valor da razão sinal/ruído médio $I/\sigma(I)$, ou o coeficiente de correlação de meio conjunto de dados $CC_{1/2}$ (180), calculado em camadas esféricas de espaço recíproco. A significância estatística, seja qual for a forma escolhida para expressá-la, é usada como critério para aplicar um corte de difração aos dados de reflexão que devem corresponder ao limite de difração o mais próximo possível.

Com o limite de difração anisotrópico, no STARANISO o valor médio local de $I/\sigma(I)$, em vez de sua média esférica, é usado para definir um corte de difração anisotrópico — isto é, um corte variando com a direção no espaço recíproco. A fronteira entre dados estatisticamente significativos e insignificantes define, portanto, uma superfície, a superfície de corte de difração anisotrópica.

A atenuação da intensidade difratada: o fator de atenuação é considerado relativo a um modelo de média esférica de um conjunto de estruturas semelhantes (181), após correção para aumento de simetria em certas zonas (182). Isso é comumente expresso como o fator *Debye-Waller* $T(\mathbf{s})$, este efeito pode ser parcialmente devido ao movimento térmico, podendo haver também uma grande contribuição decorrente da desordem:

$$T(\mathbf{s}) = \exp(-2\pi^2 \mathbf{s}^T \mathbf{U} \mathbf{s}) \quad (3.3)$$

onde \mathbf{s} é o vetor de espalhamento e \mathbf{U} é o tensor de deslocamento anisotrópico geral, normalmente expresso como um tensor de segunda ordem (ou seja, contendo apenas termos quadráticos nos deslocamentos) (183,184). Contornos de equi-valor do fator *Debye-Waller* são, portanto, elipsóides que obedecem à simetria da classe *Laue*. Desta forma, não há distorção dos contornos elipsoidais por variações na redundância de reflexão, pois esta última não tem efeito sistemático em I , apenas em $\sigma(I)$. O fator B está diretamente relacionado ao deslocamento isotrópico quadrado médio dos átomos:

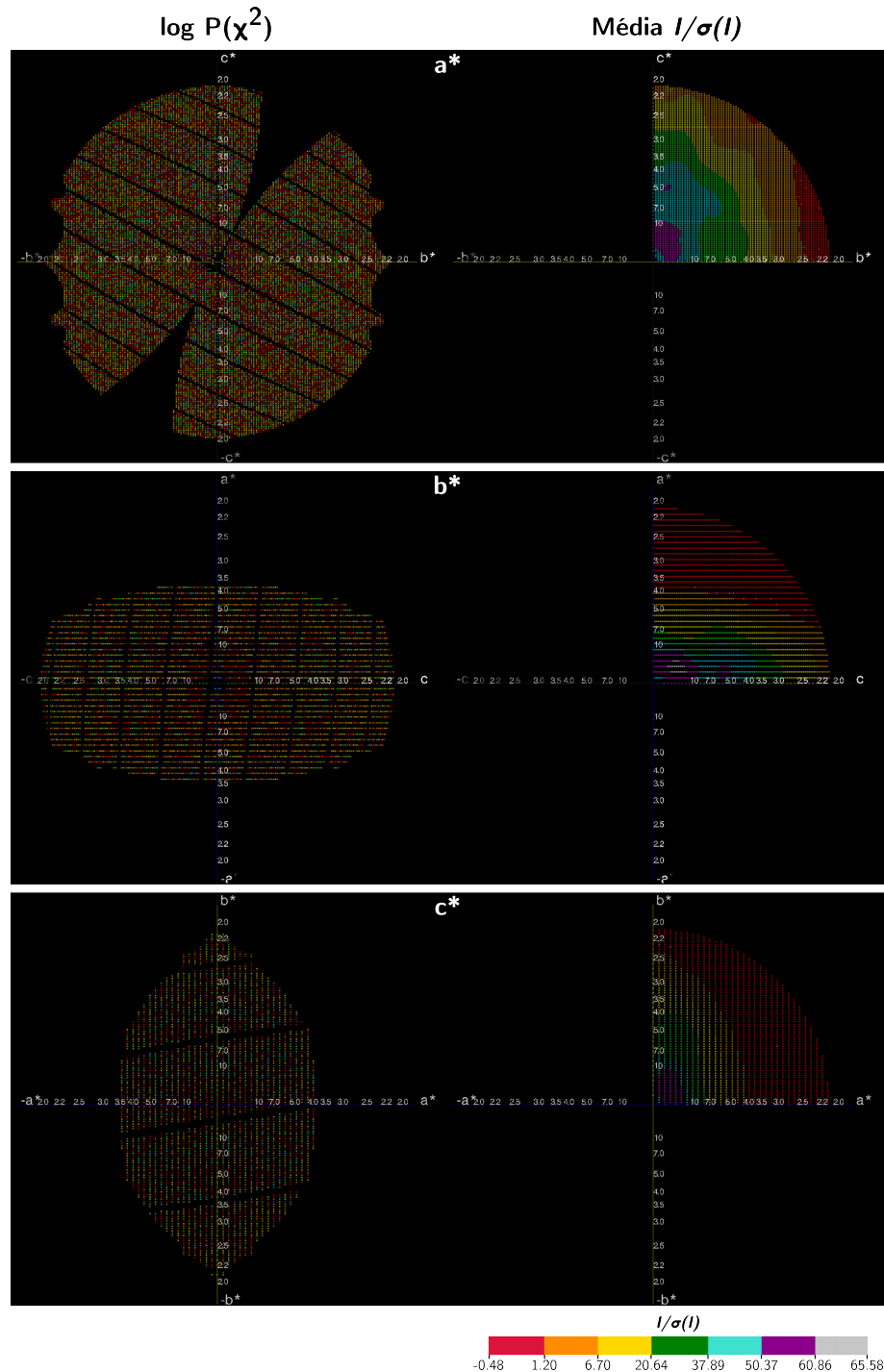


Figura 5 – Representação da rede recíproca para *CtBgl3B* obtida na difração de raios X. A representação do $\log P(\chi^2)$ se refere ao perfil do mapa de correlação para os índices de reflexões fornecidos, após o processamento com programa XDS utilizando dados não mesclados (*unmerged*). Na imagem é mostrada a seção transversal da rede recíproca ao longo de cada uma das respectivas direções (a^* , b^* e c^*). Além dos valores médios de $I/\sigma(I)$.

Fonte: Elaborada pelo autor.

$$B = 8\pi^2 \langle u^2 \rangle \quad (3.4)$$

A imagem de difração [Figura 5](#) mostra os contornos elipsoidais dos fatores *Debye-Waller*. Os pontos rosa, azul e vermelho têm o mesmo significado, enquanto as outras cores representam valores crescentes do fator *Debye-Waller* em direção ao centro do padrão de difração.

O inverso da raiz quadrada do fator *Debye-Waller* é aplicado como um fator de correção anisotrópico para as amplitudes após estimativa bayesiana das intensidades medidas e suas incertezas padrão (173). O método de [French e Wilson\(185\)](#), exceto que utiliza uma distribuição a priori anisotrópica da intensidade esperada, calculada a partir do fator *Debye-Waller* para cada reflexão, ao invés da isotrópica originalmente. O fator de correção anisotrópico é ajustado subtraindo o menor autovalor do tensor de deslocamento anisotrópico geral \mathbf{U} [Equação 3.5](#), dos elementos diagonais de \mathbf{U} .

$$\mathbf{U} = \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ U_{12} & U_{22} & U_{23} \\ U_{13} & U_{23} & U_{33} \end{pmatrix} \quad (3.5)$$

Um elemento U_{ij} de \mathbf{U} tem como dimensão (*comprimento*)². Os componentes do fator de escala anisotrópico são obtidos a partir de um ajuste de mínimos quadrados dos elementos de um tensor anisotrópico para F_{obs} :

$$\sum (|F_{obs}| - k(s)|F_{cal}|)^2 \rightarrow \min \quad (3.6)$$

onde o fator de escala $k = k_0 e^{-h^T U^* h}$. U^* é o tensor anisotrópico simétrico do espaço recíproco. Usando [Equação 3.5](#) k pode ser reescrito como:

$$k = e^{-(U_{11}a^{*2}h^2 + 2U_{12}a^*b^*hk + 2U_{12}a^*c^*hl + U_{22}b^{*2}k^2 + 2U_{23}b^*c^*kl + U_{33}c^{*2}l^2)} \quad (3.7)$$

O valor de k muda na forma de camadas elípticas concêntricas a partir do centro da rede recíproca. Os parâmetros U_{11} , U_{22} e U_{33} correspondem aos eixos principais da elipse. Um fator de escala anisotrópico deve ser aplicado para ambos $|F_{obs}|$ ou $|F_{cal}|$ para torná-los comparáveis, dado que $|F_{obs}|$ tem dependência direcional, portanto, a concordância entre eles seria muito baixa.

Assim, nenhuma correção é aplicada na direção do autovetor correspondente ao menor autovalor de \mathbf{U} (ou seja, a direção do menor deslocamento quadrático médio), e um fator de correção multiplicativo > 1 é aplicado em todas as outras direções. As amplitudes corrigidas, portanto, decaem isotropicamente em altos valores de d^* , na mesma taxa que aquelas na direção da difração mais forte. Em outras palavras, os dados nas direções de

difração mais fraca são aguçados de modo a decair na mesma taxa que aqueles na direção de difração mais forte, com a última permanecendo inalterada.

3.6.4 Faseamento e refinamento

O faseamento, dos conjuntos de dados da *CtBgl3B*, foi realizado pelo método de MR com o programa *Phaser* (186) usando como modelo a estrutura da Bgl de *Thermotoga neapolitana TnBgl3B* (PDB 2X40, 41% de identidade) determinada por Pozzo *et al.*(187). Devido à baixa identidade sequencial com modelos disponíveis, o faseamento apresentou problemas para obtenção de modelo mais preciso e completo, por exemplo, erro de registro, desta forma foi necessário a utilização de construções automáticas de modelo usando os programas *AutoBuild wizard* (188) do pacote de programas *PHENIX* (189), além do *Buccaneer* (190) do *CCP4*, resultando em modelos parcialmente completos (40 – 60%). A Figura 6 traz o esquema aplicado nas ações para obtenção da estrutura cristalográfica.

O faseamento, do conjunto de dados da *BlymD*, foi realizado também pelo método de MR com o programa *Phaser* (186) usando como modelo a estrutura da proteína não caracterizada de *Geobacillus stearothermophilus* (PDB 1T8H, 55% de identidade).

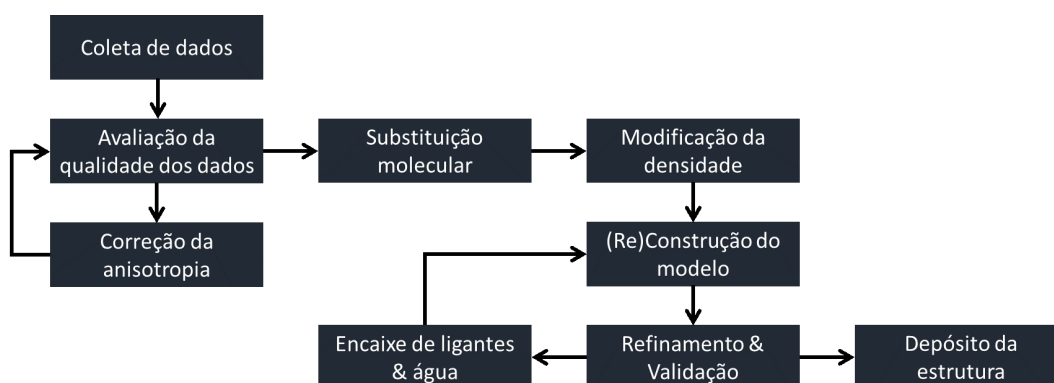


Figura 6 – Fluxograma de etapas para obtenção da estrutura partindo da coleta de dados.

Fonte: Elaborada pelo autor.

Para as etapas de interpretação dos mapas de densidades eletrônica, construção, refinamento em espaço real foi utilizado o programa *Coot* (191) alternadamente com ciclos de refinamento no espaço recíproco usando os programas *REFMAC5* (192) e *Phenix refine* (179). Moléculas de água foram manualmente adicionadas ou deletadas com programa *Coot* conforme a altura do pico ($3,0\sigma$ no mapa ($F_o - F_c$) e distância para um potencial doador ou aceitador de interações de hidrogênio ($<3,5 \text{ \AA}$). Refinamento de grupos de TLS (177, 178) e NCS foram usados, principalmente, nos últimos ciclos de refinamento. Para a validação do modelo utilizou-se o programa *MolProribity* (193). Os modelos finais obtidos foram depositados no banco de dados PDB (www.rcsb.org) (194). Todos os resultados

estatísticos de coleta de dados, faseamento e refinamento estão resumidos nas tabelas correspondentes.

As definições de alguns dos parâmetros de qualidade usados para estruturas cristalográficas estão apresentadas no [Apêndice A](#).

3.7 Ensaios bioquímicos Bgl

A atividade da *CtBgl3B* foi verificada, adaptando método previamente descrito (195–197), que utiliza *p*-nitrofenol- β -D-glucopiranosídeo (*p*NPG) como substrato, em função da variação de temperatura variando num intervalo de 20–90 °C (com auxílio de termociclador) e também do pH variando de 2–10, utilizando-se o tampão *McIlvaine* (198), que em sua composição com diferentes proporções de fosfato dissódico Na₂HPO₄ [0,2 M] e ácido cítrico C₆H₈O₇ [0,1 M], permite amplo espectro de variação do pH, com a adição de 150 mM NaCl e glicerol 10%. As reações de hidrólise foram realizadas utilizando-se 0,06 $\mu\text{g mL}^{-1}$ (7,14 nM) de proteína e 0,4 mM de *p*NPG, para um volume total em cada reação de 100 μL , durante o tempo de 5 minutos e as reações foram interrompidas pela adição de 100 μL de Na₂CO₃ 2%. A quantidade de *p*-nitrofenol (*p*-NP) hidrolisado foi medida por meio de leituras de absorvância a 410 nm em espectrofotômetro de placas (SpectraMax), e convertidos para valores de atividade relativa (%) gerando um gráfico de superfície de resposta.

Efeito de aditivos: para *CtBgl3B* o efeito de SDS, Triton X-100, Tween-20 e DMSO a (1%), enquanto EDTA, NiSO₄, LiCl₂, CaCl₂, KCl, ZnSO₄, MgCl₂, CuSO₄, MnSO₄ e CoCl₂ a 5 mM foi avaliado na atividade da enzima adicionando-se 10 μL em cada reação (volume final 100 μL). Para avaliar o efeito dos principais aditivos, que demonstraram afetar positivamente nas reações enzimáticas, foi aplicada análise de variância ANOVA (199) e comparação de valores médios individuais pelo teste de Tukey (200).

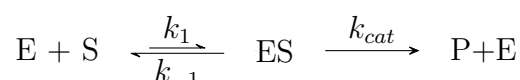
Atividade residual, com efeito, do tempo e temperatura: amostras de *CtBgl3B* (30 μM) foram submetidas à incubação nas temperaturas de 60, 65, 70 e 75 °C por 1, 2, e 3 h. Posteriormente, estas amostras foram conduzidas à reação de hidrólise do *p*NPG em tampão *McIlvaine* (pH 5,5) por 5 min em volume reacional de 100 μL nas respectivas temperaturas de incubação acima. Desta forma as atividades foram determinadas novamente e relacionadas à atividade previamente estabelecida como controle.

Especificidade de substrato: a atividade também foi testada para os diferentes substratos sintéticos além do *p*NPG, são eles: *p*NPAF (*p*-NP-alfa-L-arabinofuranosídeo), *p*NPB GAL (*p*-NP-alfa-D-galactopiranosídeo), *p*NPB M (*p*-NP-beta-D-Manopiranosídeo), *p*NPX (*p*-NP-beta-D-xylopiranosídeo) e *Cp*NPG2 (2-Cloro-*p*-NP-beta-celobiosídeo). Os ensaios foram realizados utilizando-se condições ótimas de pH e temperatura pré-estabelecidas.

3.7.1 Cinéticas

Conhecendo as condições de pH e temperatura ideais para a atividade enzimática, a cinética da hidrólise do *p*NPG pela *CtBgl3B* foi medida a 65 °C em tampão *McIlvaine* 50 mM pH 5,5 contendo 7,14 nM de proteína e variando desta vez a concentração do *p*NPG por meio de diluições seriadas partindo de 12,8 até 0,0125 mM, diminuindo a metade da concentração a cada diluição. A curva de calibração do *p*-nitrofenol ($R^2 = 0,9926$, dados não mostrados) foi utilizada para a obtenção dos valores de concentração do produto formado em cada reação. Aplicando a análise de *Michaelis-Menten* (201) para cinética enzimática foi possível calcular os parâmetros V_{max} , K_M , e K_{cat} . Todas as medidas foram realizadas em quadruplicata, respectivamente.

Os ensaios de cinética para atividade enzimática, foram realizados para avaliar o efeito da concentração do substrato nas reações catalisadas pelas enzimas, baseando-se no seguinte modelo de reação:



Onde no modelo acima, E = enzima; S = substrato; ES = complexo enzima-substrato; P = produto, k_1 = constante de reação de formação do complexo enzima-substrato, k_{-1} = constante de dissociação do complexo enzima-substrato e k_{cat} = constante da reação de formação do produto. A cinética foi analisada adaptando modelo de *Michaelis-Menten*, tal como:

$$v = \frac{d[P]}{dt} = V_{max} \frac{[S]}{K_M + [S]} = k_{cat}[E_t] \frac{[S]}{K_M + [S]} \quad (3.8)$$

Onde, velocidade v é dada pela taxa de formação de produto variando num intervalo de tempo $d[P]/dt$; V_{max} = velocidade máxima da reação; K_M = constante de *Michaelis-Menten* definida como a concentração de substrato que fornece uma velocidade de reação que é metade da velocidade máxima obtida sob condições de saturação do substrato; $[E_t]$ = concentração total de enzima. Assumindo que a concentração de substrato $[S]$ é muito maior que de enzima total $[E_t]$, permitindo reajustar a Equação 3.8 de tal maneira que: $V_{max} = k_{cat}[E_t]$. Assim, os resultados apresentados na seção 4.4 foram obtidos a partir de regressão não-linear nos pontos medidos após 5 minutos de reação entre $[E_t] + [S]$, e convertidos em valores de velocidade de formação do produto (mM s^{-1}) utilizando a curva de calibração conduzida com o produto *p*NP Figura 7.

3.7.1.1 Inibição por substrato

Em altas concentrações, alguns substratos também inibem a atividade enzimática. A inibição de substrato ocorre com cerca de 20% de todas as enzimas conhecidas. Isso

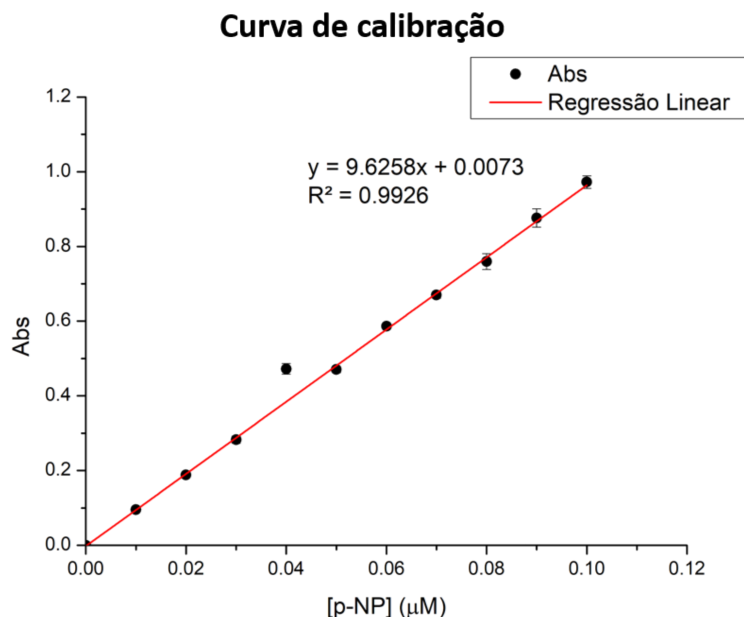


Figura 7 – Curva de calibração do *p*NP para referência das medidas de produto formado nas reações enzimáticas.

Fonte: Elaborada pelo autor.

acontece quando duas moléculas de substrato podem se ligar à enzima e, assim, bloquear a atividade. Conforme o modelo:

$$Y = \frac{V_{max} \cdot X}{(K_M + X \cdot (\frac{1+X}{K_i}))} \quad (3.9)$$

onde V_{max} é a velocidade máxima da reação enzimática, se o substrato também não inibiu a atividade enzimática, expressa nas mesmas unidades que Y (velocidade); K_i é a constante de dissociação para a ligação do substrato de tal forma que dois substratos podem se ligar a uma enzima.

3.8 Ensaios bioquímicos YImD

A proteína *BIYImD*, como mencionada anteriormente, é uma proteína hipotética, cujo alinhamento detecta domínios de lacase (EC 1.10.3.2) — polifenol oxidoreductase multi-cooper (PF02578 Cu-oxidase_4), também não caracterizados. Portanto, após a obtenção da proteína purificada e sem a cauda (6xHis-Trx-tag), foram realizados ensaios de atividade enzimática, para tentar verificar a atividade oxidativa por meio da coloração com o substrato ABTS (sal diamônio 2,2'-azino-bis(3-etilbenzotiazolina-6-sulfonato) (202, 203). Com a intenção de determinar parâmetros de pH (variando de 2 a 10) e temperatura ótimos, e de cinética. Todos os ensaios enzimáticos envolvendo o ABTS foram medidos em 420 nm, pois à medida que o ABTS é oxidado sua absorbância em 420 nm aumenta

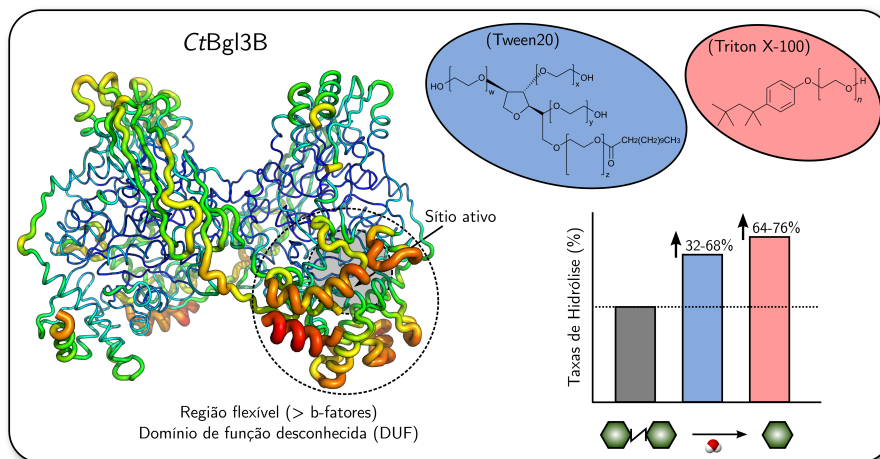
liberando a coloração azul. ABTS (ϵ_{420}) = 38000 M⁻¹cm⁻¹. Os ensaios, também foram realizados mediante adição de variadas concentrações (0,1 a 5 mM) de CuSO₄, ZnSO₄ e EDTA respectivamente.

Com base na comparação da sequência primária e estrutura da *BIYImD* com uma proteína reportada em janeiro de 2020 (129), foi levantada a hipótese de que a proteína pudesse ter similares atividades enzimáticas, tal como os autores demonstraram. As atividades reportadas, são variadas, como: purina nucleosídeo fosforilase (PNP) (EC 2.4.2.1), adenosina deaminase (ADEA) (EC 3.5.4.4) e metiltioadenosina fosforilase (MTAP) (EC 2.4.2.4). Desta forma, foram conduzidos experimentos para tentar verificar alguma dessas atividades enzimáticas, seguindo protocolo das reações conforme as metodologias reportadas (129, 204).

3.9 Gráficos e Figuras

Os gráficos e análises deste trabalho foram feitos usando os programas Origin (OriginLab), GraphPad Prism e STATISTICA. Todas as figuras de estruturas cristalográficas foram elaboradas usando o PyMOL (145). Figuras, esquemas e modelos foram elaborados e organizados extensivamente usando o programa Inkscape (205). Esquemas de reações químicas foram elaborados no programa ChemBioDraw[®]. As tabelas foram montadas com auxílio da ferramenta *online Tables Generator*. Este trabalho foi escrito e editado usando o modelo USPSC-modelo-IFSCp v3.0 (abnTeX2) nos editores L^AT_EX Overleaf e TeXstudio.

4 β -GLUCOSIDASE GH3 - CTBGL3B



Neste capítulo são apresentados os resultados obtidos com a proteína *CtBgl3B*, e suas discussões junto de comparações com dados da literatura.

4.1 Expressão e purificação

A proteína recombinante *CtBgl3B* foi expressa em *E. coli* BL21(DE3), usando meio de auto-indução, seguida pela lise celular, e então as proteínas foram submetidas às etapas purificação, utilizando-se IMAC contendo íon níquel Ni^{2+} . Bandas compatíveis com a massa da enzima *CtBgl3B* (84 kDa) podem ser observadas: na fração insolúvel (FI); solúvel (FS); *Flow-through* (FT) e na W30, contudo a eluição propriamente pode ser observada na fração de 250 mM de imidazol (E250) (Figura 8b). Após a primeira etapa de purificação, a fração E250 contendo a enzima *CtBgl3B* foi submetida a clivagem com trombina e novamente foi realizada purificação usando IMAC, separando a fração clivada (FT2), conforme observado na Figura 8c.

As frações contendo a enzima de interesse, clivada (FT2), foram concentradas ($\sim 1\text{-}2 \text{ mg mL}^{-1}$) e submetidas a terceira etapa de purificação, por cromatografia de exclusão por massa molecular (Superdex-200 16/60; tampão 50 mM fosfato de sódio pH 7,0; 250 mM NaCl; 10% glicerol). O cromatograma para a purificação por exclusão de massa molecular é apresentado na Figura 8a. As frações destacadas em amarelo 63-69 mL (Figura 8a e d) não foram utilizadas para ensaios por apresentarem baixa homogeneidade de estado oligomérico, sendo observados dímeros ($\approx 85\%$) e agregados desconhecidos ($\approx 15\%$) por meio de DLS, gel nativo e SEC-MALS (resultados não mostrados). Optando-se pela seleção das frações destacadas em verde 73-79 mL, que compreendem a parte central do pico de eluição e com alta homogeneidade para o estado dimérico da proteína *CtBgl3B*.

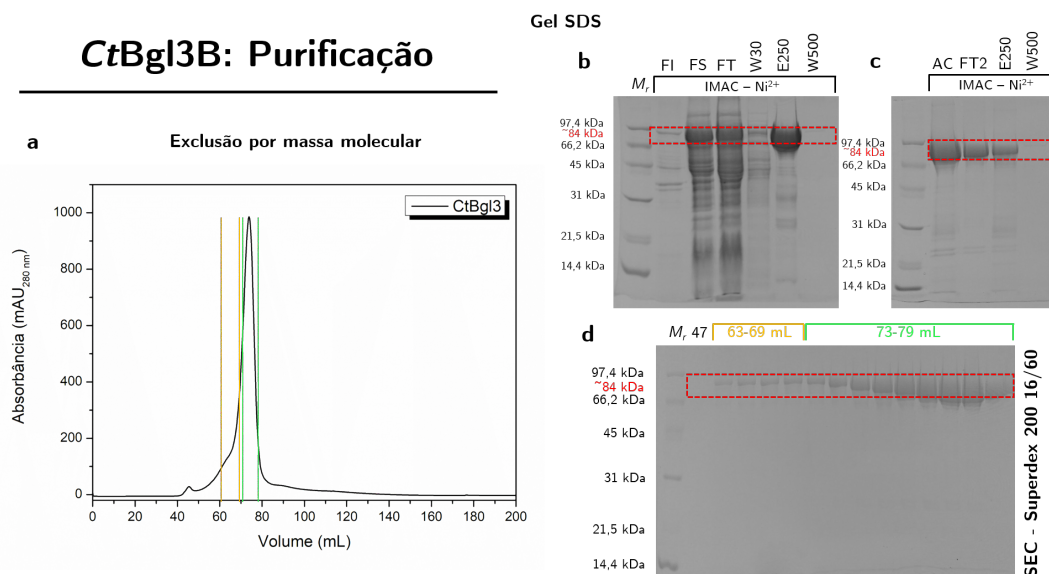


Figura 8 – **Purificação da CtBgl3B.** **a.** Cromatograma de purificação em coluna de exclusão por massa molecular (Superdex 200 16/60; V = 120 mL). **b.** Gel SDS-page 10% mostrando todas as etapas da purificação por coluna de afinidade com metal imobilizado (IMAC - Ni²⁺), onde: Mr - marcador; FI - fração insolúvel; FS - fração solúvel; FT - *flow through*; W30 - limpeza (30 mM imidazol); E250 - eluição (250 mM imidazol); W500 - (500 mM imidazol). **c.** Gel SDS-page 10% mostrando todas as etapas da segunda purificação por IMAC (Ni²⁺), onde: Mr - marcador; AC - antes da clivagem; FT2 - *flow through* proteínas clivadas; E250 - eluição de proteínas não clivadas (250 mM imidazol); W500 - (500 mM imidazol). **d)** Frações coletadas da eluição na coluna de exclusão por massa molecular (Superdex 200 16/60).

Fonte: Elaborada pelo autor.

Finalmente, a proteína CtBgl3B foi obtida purificada com rendimentos variando de 25-40 mg L⁻¹ de cultura, concentrada a 12,5 mg mL⁻¹ e armazenadas no freezer -80 °C, e a partir de alíquotas de proteína congelada foram realizados os ensaios bioquímicos aqui apresentados até o momento. Os ensaios biofísicos e de cristalização foram realizados com proteína fresca.

4.2 Análises de estabilidade térmica

4.2.1 DSF

O *screening* de condições usando fluorimetria diferencial de varredura permite a determinação das melhores condições tamponantes, de estabilidade, para as proteínas em solução, por meio da escolha do tampão que confira tal característica, visto que é de grande importância que as amostras proteicas estejam em condições favoráveis a manutenção do seu estado de enovelamento, fator diretamente correlacionado com a capacidade em desempenhar sua função, ou mesmo interferir na formação e obtenção de cristais (153). O

ponto de inflexão, das curvas de transição T_m , foram calculados usando o mesmo simples modelo da [Equação 3.1](#).

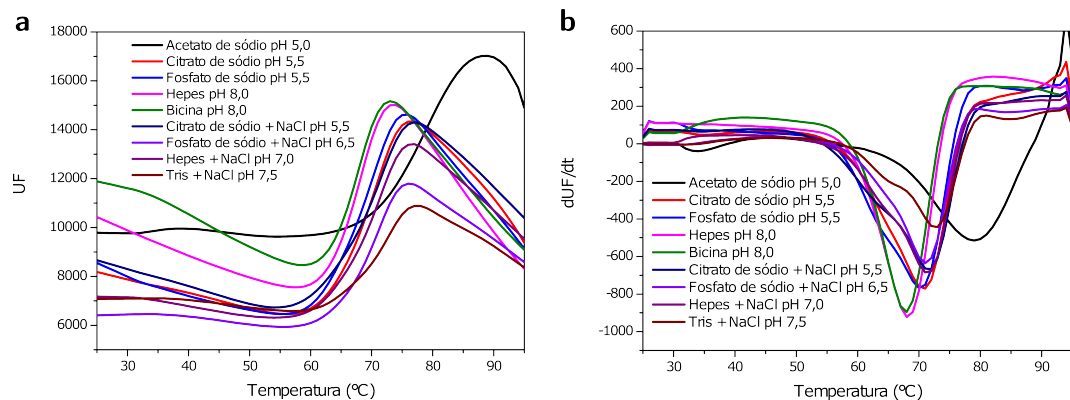


Figura 9 – **Curvas de DSF avaliando o efeito de diferentes tampões nos perfis de desnaturação térmica da *CtBgl3B*.** **a.** Curvas do perfil de deslocamento térmico de ligação da proteína com o corante fluorescente. **b.** As temperaturas da média de cada pico, nos máximos da primeira derivada do sinal de fluorescência, foram tomadas como a T_m das amostras de *CtBgl3B*.

Fonte: Elaborada pelo autor.

O [Quadro 2](#) apresenta os valores de T_m para cada condição avaliada no experimento, demonstrando o diferente comportamento em cada condição de pH e composição do tampão, bem como concentrações de NaCl. Por meio destes resultados e com a avaliação de cada curva sigmóide de fluorescência, tomando como critério o perfil de cada sigmóide, apesar de apresentar maior valor de T_m foram eliminados respectivamente os tampões de acetato de sódio pH 5,0 ($T_m = 79$ °C), acetato de amônio pH 7,3, NaCl 300 mM ($T_m = 73$ °C), Bis-Tris + NaCl pH 5,5 e 6,0, NaCl 300 mM ($T_m = 73$ °C) optando-se pela escolha do tampão de trabalho fosfato de sódio pH 6,5, NaCl 300 mM, cujo $T_m = 71$ °C.

4.2.2 CD

Com objetivo de identificar a desnaturação térmica para a enzima *CtBgl3B*, as medidas utilizando espectropolarimetria de dicroísmo circular foram realizadas em função da variação de temperatura, cujo intervalo foi de 20 a 90 °C ([Figura 10a](#)).

As temperaturas de desnaturação T_m foram estimadas usando função sigmoidal de *Boltzmann* ([Equação 3.1](#)), a partir da interpolação dos valores de $\Delta\epsilon$, cujos mínimos referentes aos comprimentos de onda 208 nm ($T_m = 68,08 \pm 0,5$ °C) e 222 nm ($T_m = 71,88 \pm 0,9$ °C), ([Figura 10b](#)), estes pontos podem ser associados a presença de fitas- β e α -hélices respectivamente, indicando o ponto médio de transição da curva de desnaturação, média da fase enovelada para desenovelada. Os resultados são compatíveis com o que foi observado no experimento de DSF onde $T_m = 71$ °C, e sugerem uma temperatura característica para enzimas termófilas, como era esperado. Fazendo contraste interessante,

Quadro 2 – Avaliação do efeito de diferentes tampões nos valores de T_m da *CtBgl3B* medidos via DSF. A graduação de cores reporta os diferentes graus de estabilidade ao longo da avaliação no experimento, indo do menos estável (vermelho) para o mais estável (verde-escuro), tracejados não apresentaram valores consideráveis. []f = 50 mM é a concentração dos sais que constituem cada tampão.

pH	Tampão ([]f = 50 mM)	[NaCl] (mM)	T_m	pH	Tampão ([]f = 50 mM)	[NaCl] (mM)	T_m
7,0	Água	0	49,00	7,3	Acetato de Amônio	0	67,00
1,2	HCl	0	-	7,5	Fosfato de Sódio	0	72,00
2,0	HCl	0	-	7,5	Tris	0	67,00
3,0	Ácido Cítrico	0	32,00	8,0	Imidazol	0	68,00
4,0	Citrato de Sódio	0	54,00	8,0	Hepes	0	71,00
4,5	Acetato de Sódio	0	60,00	8,0	Tris	0	68,00
4,7	Citrato de Sódio	0	65,00	8,0	Bicina	0	71,00
5,0	Acetato de Sódio	0	79,00	8,5	Tris	0	66,00
5,0	Fosfato de Potássio	0	69,00	9,0	Bicina	0	54,00
5,5	Citrato de Sódio	0	71,00	9,5	Carbonato de Sódio	0	46,00
5,5	Fosfato de Sódio	0	70,00	10,0	Carbonato de Sódio	0	-
5,8	Mes	0	69,00	3,0	Ácido Cítrico + NaCl	300	53,00
6,0	Fosfato de Potássio	0	70,00	4,5	Acetato de Sódio + NaCl	300	68,00
6,0	Bis-Tris	0	69,00	5,0	Fosfato de Potássio + NaCl	300	71,00
6,2	Mes	0	70,00	5,5	Bis-Tris + NaCl	300	72,00
6,5	Fosfato de Sódio	0	70,00	5,5	Citrato de Sódio + NaCl	300	68,00
6,5	Cacodilato de Sódio	0	71,00	5,8	Mes + NaCl	300	70,00
6,5	Bis-Tris	0	71,00	6,0	Bis-Tris + NaCl	300	72,00
6,5	Mes	0	71,00	6,5	Cacodilato de Sódio + NaCl	300	71,00
6,7	Bis-Tris	0	71,00	6,5	Fosfato de Sódio + NaCl	300	71,00
7,0	Fosfato de Potássio	0	69,00	7,0	Bis-Tris + NaCl	300	71,00
7,0	Hepes	0	70,00	7,0	Hepes + NaCl	300	71,00
7,0	Bis-Tris	0	71,00	7,3	Acetato de Amônio + NaCl	300	73,00
7,3	Acetato de Amônio	0	70,00	7,5	Tris + NaCl	300	70,00

Fonte: Elaborada pelo autor.

com os melhores valores observados nos ensaios de varredura para determinação de pH e temperatura ótima da sua atividade enzimática mostrados mais a frente na [Figura 18](#).

4.3 Análises estruturais

4.3.1 SEC-MALS

Os experimentos de espalhamentos de luz de multi-ângulos foram realizados para monitorar a qualidade da amostra, principalmente no que se refere ao estado oligomérico, preparada para os ensaios bioquímicos e de cristalização. Os resultados dos experimentos, para *CtBgl3B*, são mostrados na [Figura 11a](#), onde na análise de SEC-MALS nota-se a presença majoritária de espécies diméricas no Pico 1 com fração de massa igual a 96% cuja massa molecular corresponde a ≈ 158 kDa, enquanto no pico 2 com fração de apenas

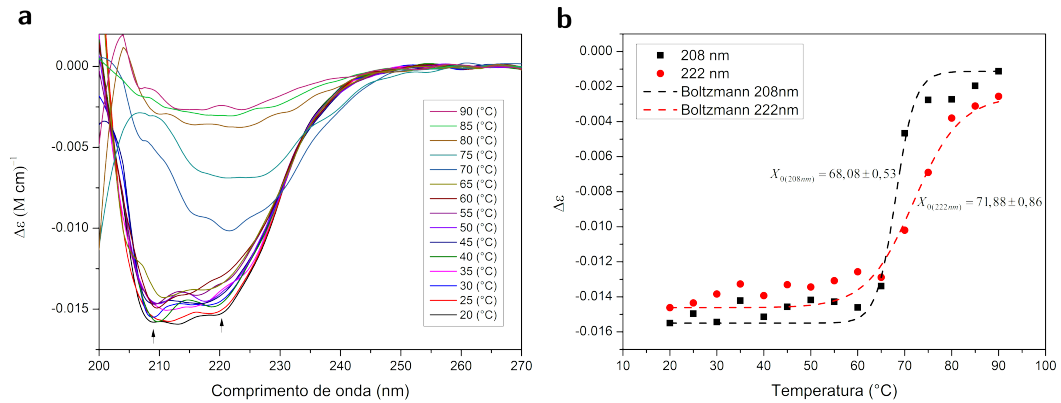


Figura 10 – **Espectro de dicroísmo circular da CtBgl3B.** **a.** Curvas de dicroísmo circular em função da variação de temperatura (20-90 °C). **b.** Ajuste de Boltzmann para os valores correspondentes a 208 nm (preto) e 222 nm (vermelho) respectivamente, onde X_0 (208 nm = 68,08 e 222 nm = 71,88) representam as T_m obtidas.

Fonte: Elaborada pelo autor.

4% a massa equivale a \approx 211 kDa sugere a presença sutil de agregados desconhecidos.

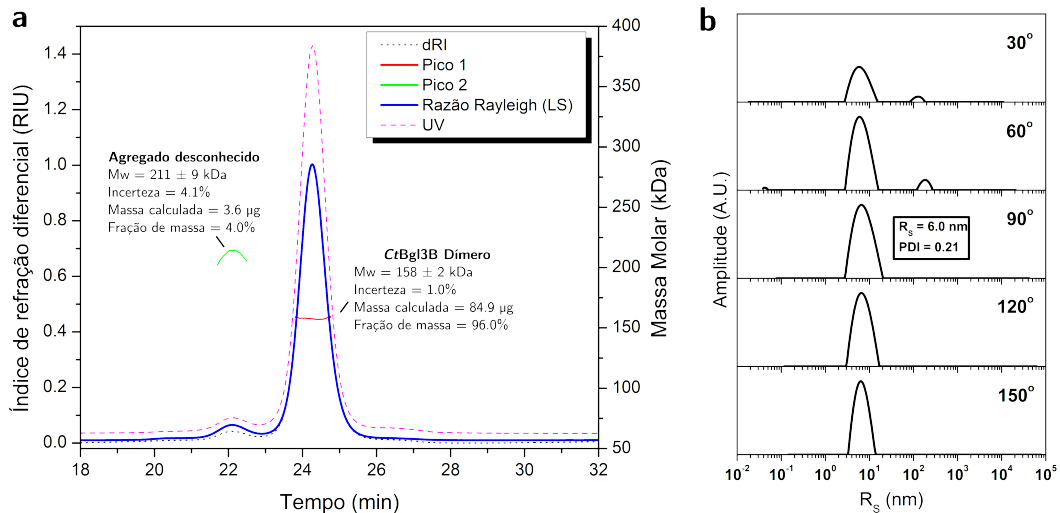


Figura 11 – **Análise de SEC-MALS da CtBgl3B.** **a.** Cromatograma de gel filtração analítica (Superdex 200 HR 10/30) com análise do espalhamento de luz com ângulo multivariado e índice de refração. Resultado da determinação do percentual de massa e estados oligoméricos observados para CtBgl3B, onde o pico 1 com 96% da massa representando o dímero (158 kDa). **b.** Estimativa do R_s pelas distribuições da amplitude do espalhamento de luz em diferentes ângulos.

Fonte: Elaborada pelo autor.

A análise usando apenas o espalhamento de luz multiângulos também foi realizada, indicando majoritariamente a presença de dímeros com raio hidrodinâmico R_s de 6.0 nm

(60 Å), observando homogeneidade nos picos referentes aos ângulos mais altos como de 120° e 150°, onde caso houvessem partículas maiores seria observado sinal no espalhamento de luz [Figura 11b](#). Contudo, acerca da informação sobre o estado oligomérico, estes resultados são especialmente significantes quando comparados aos dados apresentados a seguir obtidos por meio das análises de SAXS, esta que permite a obtenção de modelos de baixa resolução para o envelope molecular das partículas da enzima em solução.

4.3.2 SAXS

O experimento de SAXS fora realizado no intuito de se construir um modelo para o envelope molecular da proteína em solução. Seus parâmetros de coleta e estruturais são mostrados na [Tabela 2](#). Na [Figura 12a](#) é apresentada a curva de espalhamento sobreposta nas curvas calculadas para modelos teóricos de dímero e monômero da referente estrutura cristalográfica *CtBgl3B*, gerados pelo programa Crysol (206). Enquanto que, na [Figura 12d](#) observa-se a curva de SAXS sobreposta nas curvas de espalhamento calculadas para modelos de monômero e dímero da estrutura de uma GH3 (em complexo com glicerol) *ThBgl3B* (PDB **2X40**) (187) com 39% de identidade para *CtBgl3B*.

Usando aproximação de *Guinier* observa-se que a região de ajuste, localizada no início da curva [Figura 12b](#), exibe o perfil linear indicando não haver interferência de heteropartículas, ou mesmo contribuição causada por danos de radiação. Desta forma, o R_g foi determinado como o valor médio de 36 ± 2 Å para as diferentes concentrações das amostras analisadas [Tabela 2](#).

A partir das curvas de espalhamento, [Figura 12a](#), a massa molecular (Mw) foi calculada com valor médio de 161,7 kDa, conforme (207). No gráfico da [Figura 12c](#) encontra-se o resultado da função de distribuição de distâncias, $P(r)$, com perfil característico de proteínas globulares e D_{max} de 110,7 Å, e onde o R_g foi de 36,6 Å, assemelhando-se aos valores estimados por meio da aproximação de *Guinier*.

Os modelos *ab initio* foram construídos usando o programa DAMMIN. O envelope da *CtBgl3B* [Figura 13](#) reconstruído a partir dos dados experimentais de SAXS teve superposição com os modelos cristalográficos da própria enzima com $\chi^2 = 2,2$, apresentando ótimo encaixe. Portanto, os perfis das curvas calculadas para os modelos cristalográficos da *CtBgl3B* e **2X40** (187), respectivamente reproduziram as curvas compatíveis com perfil de um arranjo dímérico.

Os resultados corroboram, para formação arranjo dimérico da proteína em solução. Haja vista, as condições de homogeneidade da amostra (96%), de forma que os cristais formados da enzima foram suscetíveis aos experimentos de difração realizados e permitiram a elucidação da sua estrutura, cujo modelo do envelope molecular é mostrado na [Figura 13](#).

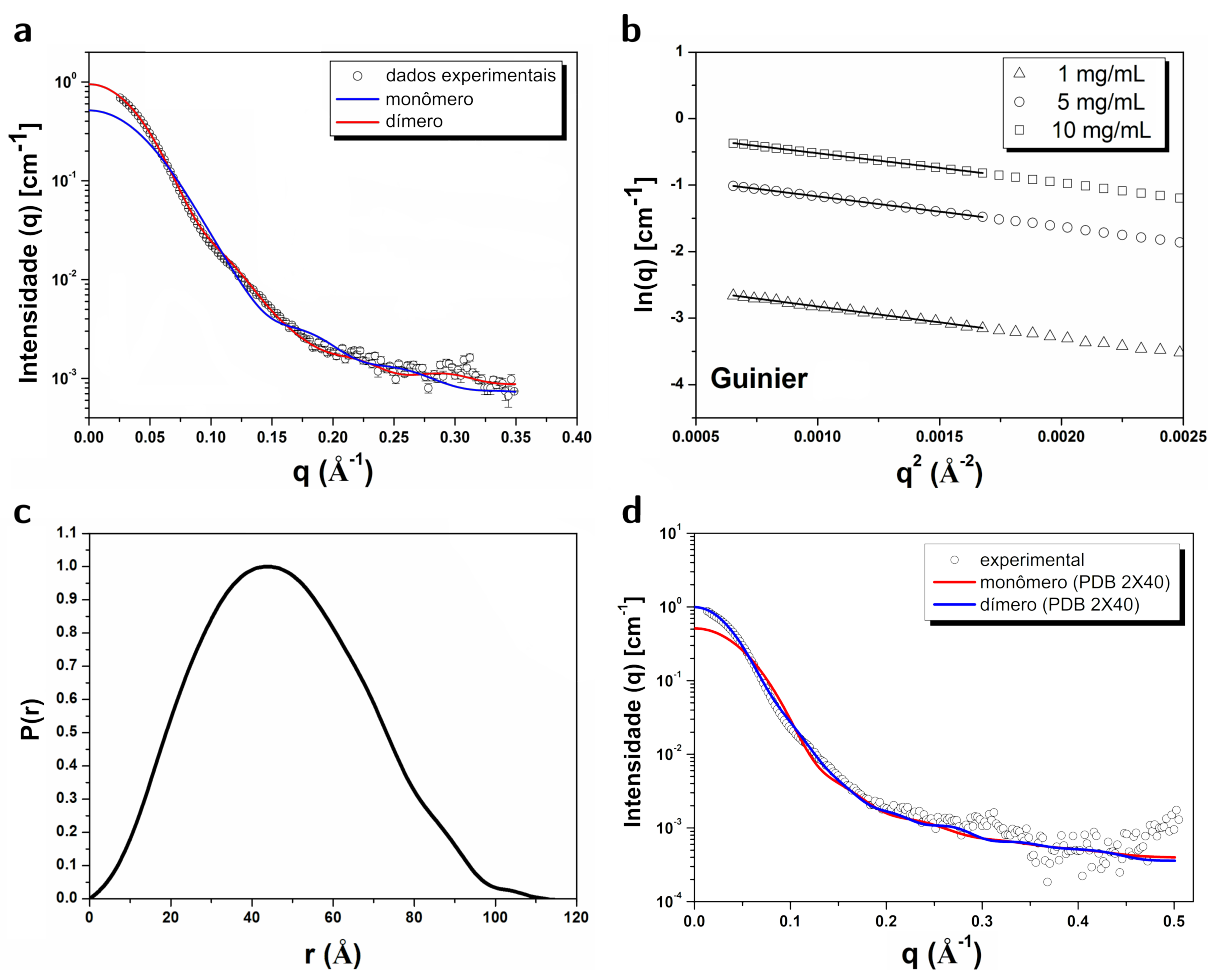


Figura 12 – Resultados do espalhamento de raios X a baixo ângulo da *CtBgl3B*. **a.** Curva experimental de SAXS para *CtBgl3B* sobreposta com curvas calculadas para modelos de monômero e dímero de estrutura depositada no PDB (**2X40**). **b.** Gráfico da análise de Guinier para amostras com diferentes concentrações, 1, 5 e 10 mg mL⁻¹ respectivamente.

Fonte: Elaborada pelo autor.

4.3.3 Estrutura cristalográfica

A estrutura cristalina de *CtBgl3B* foi determinada, com os limites de resolução anisotrópica distintos ao longo de cada direção, sendo de 3,47 em a^* , 2,22 em b^* e 1,78 em c^* (Å), por meio do método de substituição molecular usando como modelo de busca (PDB **2X40**, 42,43% de identidade). Grupo espacial $P2_12_12_1$ com parâmetros de cela unitária: $a=63,84$; $b=148,06$; $c=198,64$ (Å). A enzima apresenta empacotamento no estado oligomérico de dímero, sendo composta por 744 e 743 resíduos de amino ácidos, respectivamente nas cadeias A e B, arranjados em três domínios conectados por regiões desordenadas ou de baixa complexidade (*linkers*) (resíduos 307–320 e 537–599). A Tabela 3 resume os resultados estatísticos da coleta de dados e refinamento. *CtBgl3B* é uma proteína de multi-domínios,

Tabela 2 – Parâmetros de coleta e análises de SAXS da *CtBgl3B*.

Parâmetros de coleta	
Instrumento	LNLS D02A-SAXS2
Comprimento de onda λ (Å)	1,54
q-intervalo (Å ⁻¹) usado no modelo	0,017 a 0,034
Concentração de proteína (mg mL ⁻¹)	1, 5 e 10
Temperatura (°C)	20
Parâmetros estruturais	
	<i>CtBgl3B</i>
R_g (Å) Guinier	37,85 ± 1,2 (1 mg mL ⁻¹)
R_g (Å) Guinier	36,39 ± 0,3 (5 mg mL ⁻¹)
R_g (Å) Guinier	36,12 ± 0,4 (10 mg mL ⁻¹)
R_g (Å) $P(r)$	36,6 ± 0,5
D_{max} (Å)	110,7 ± 2
Determinação da massa molecular	
Mw (kDa) SAXSMow2.0	171,2 (1 mg mL ⁻¹)
Mw (kDa) SAXSMow2.0	157,9 (5 mg mL ⁻¹)
Mw (kDa) SAXSMow2.0	156,1 (1 mg mL ⁻¹)
Mw (kDa) seq AA	84,9

Fonte: Elaborada pelo autor.

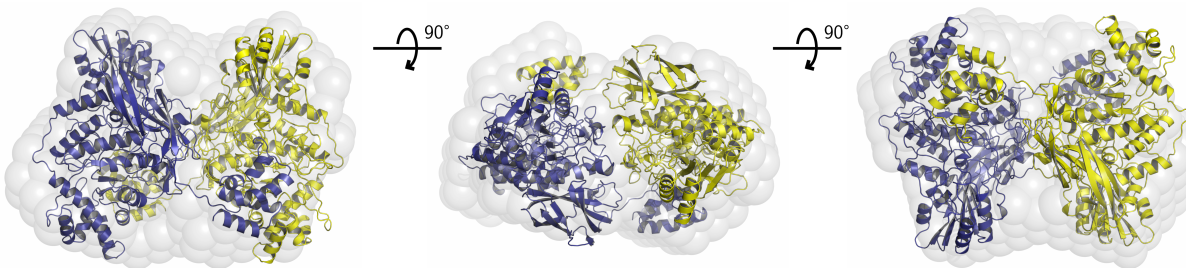


Figura 13 – **Envelope de SAXS da *CtBgl3B***. Sobreposição do modelo cristalográfico, da *CtBgl3B* (alta resolução), com envelope obtido em solução dos resultados de SAXS, evidenciando o estado dimérico da proteína em solução (baixa resolução).

Fonte: Elaborada pelo autor.

conforme visto na [Figura 14a](#), Domínio 1 (Glyco_hydro_3; resíduos 13–285) consiste em um $(\alpha/\beta)_8$ barril-TIM (PF00933), seguido pelo domínio 2 (Glyco_hydro_3_C; residues 316–550), que é um $(\alpha/\beta)_6$ sanduíche com uma folha β de seis fitas (PF01915). Estes dois domínios possuem papel direto na arquitetura do sítio ativo previsto. Enquanto o domínio 3 (resíduos 587–657) fibronectina tipo III (Fn3-like; PF00933), comumente referido como domínios de adesão a célula bacteriana, também similares a uma região com enovelamento

do tipo imunoglobulina (*Ig-like*) ou a família de domínios CARDB (PF07705).

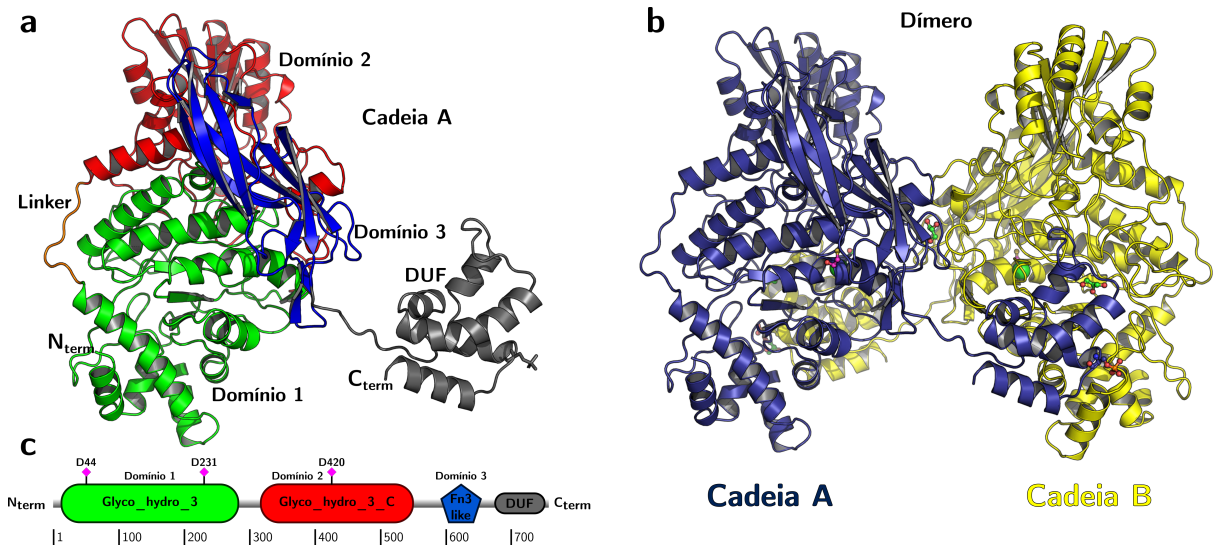


Figura 14 – **Representação esquemática do arranjo estrutural da *CtBgl3***. **a.** Identificação por cores dos domínios que compõem a estrutura, domínio 1 - N-terminal Glyco_hydro_3 com enovelamento do tipo $(\alpha/\beta)_8$ barril-TIM (verde), um *linker* entre os domínios 1–2 (laranja), domínio 2 - Glyco_hydro_3_C com enovelamento do tipo $\alpha/\beta/\alpha$ sanduíche com uma folha de 6 fitas (vermelho), domínio 3 - fibronectina tipo III com enovelamento tipo imunoglobulina formando um β -sanduíche de 7 fitas (azul) e o domínio de função desconhecida - DUF na região C-terminal composto por hélices e *loops* (cinza). **b.** Representação da estrutura quaternária dimérica (cadeias A e B) com DUF sendo acomodado numa interface próxima à localização da cavidade catalítica. Moléculas de ligantes e a serina fosforilada são representados no estilo bastão e bola, GOL (verde e vermelho), ligante BME (magenta, vermelho e amarelo), íons de Cl (esferas verdes). **c.** Representação da arquitetura para *CtBgl3* com esquemas de cores dos respectivos domínios, seguindo a numeração da sequência de amino ácidos, identificando os resíduos catalíticos.

Fonte: Elaborada pelo autor.

Por fim, na região C-terminal é observado um fragmento de cadeia polipeptídica (DUF; resíduos 658–755) (Figura 14a), que em parte pode ser alinhada com a estrutura de um domínio com função desconhecida DUF1810 (PF08837) observado na proteína hipotética de *Mycobacterium tuberculosis* (PDB **2JEK**) (208), possivelmente com papel importante para estabilidade estrutural da proteína (209), como também na formação de uma interface de contato com a cadeia parceira na dimerização (Figura 14b), mas talvez não seja essencial biologicamente já que sua função permanece desconhecida.

Os estados oligoméricos das proteínas desta família variam entre monômero, homodímero e homotetrâmero. A unidade assimétrica na célula unitária da *CtBgl3B* (7MS2), foi estabelecida como um conjunto dimérico. Usando o PISA, os valores de energia foram

Tabela 3 – **Estatísticas da coleta de dados e refinamento da estrutura da CtBgl3B.**
Valores estatísticos para camada de mais alta resolução estão mostrados entre parênteses.

Parâmetros	Valores
PDB ID	7MS2
Coleta de dados e processamento	
Grupo espacial	$P2_12_12_1$
Cela unitária	
a, b, c (Å)	63,84, 148,06, 198,64
α, β, γ (°)	90
Detector	PILATUS3 6M
Fonte de raios X	Síncrotron, DLS I24:microfocus
Comprimento de onda (Å)	0,9686
Resolução (Å)	69,37 – 2,14 (2,04 – 2,04)
Resolução anisotrópica (Å)	3,47 (a*), 2,21 (b*), 1,78 (c*)
Multiplicidade	25,8 (23,0)
$I/\sigma(I)$ médio	13,6 (1,9)
Wilson B-fator (Å ²)	25,19
R_{merge}	0,183 (2,100)
R_{meas}	0,187 (2,148)
$R_{p.i.m}$	0,037 (0,444)
CC (1/2)	0,999 (0,827)
Completeza (esférica) (%)	57,4 (21,5)
Completeza (elipsoidal) (%)	94,4 (97,7)
Reflexões totais	1786265 (79736)
Reflexões únicas	69276 (3465)
Parâmetros do refinamento	
Reflexões usadas no refinamento	69151 (2486)
Reflexões usadas para R_{free}	1999 (72)
R_{work} (%)	20,18 (28,92)
R_{free} (%)	23,01 (30,57)
N.º de átomos não hidrogênio	11772
N.º de átomos da proteína	11327
N.º de átomos dos ligantes	40
N.º de átomos do solvente	405
Resíduos da proteína	1495
Ramachandran	
Favorecidos (%)	98,25
Permitidos (%)	1,75
Outliers (%)	0,0
Clashscore(193)	3,82
B-fator médio (Å ²)	42,61
Proteína	42,98
Ligantes	42,34
Solvente	32,35
N.º de grupos TLS	40

DLS, *Diamond light source*.

Fonte: Elaborada pelo autor.

obtidos para comparar a estabilidade da interface (Figura 14b), possuindo área de 2394,1 Å², 39 ligações de hidrogênio, e 12 pontes salinas. Neste contexto, o valor de ΔG_{int} (-31,9 kcal mol⁻¹) é calculado como a diferença nas energias de solvatação totais, de estruturas isoladas e montadas, não incluindo o efeito das ligações de hidrogênio e pontes salinas nas interfaces. A energia livre de dissociação, ΔG_{diss} , corresponde à diferença de energia livre entre o estado dissociado e o associado, esse arranjo dimérico da *CtBgl3B* se dissocia em dois monômeros, mais os íons e ligantes, com $\Delta G_{diss} = 10,1$ kcal mol⁻¹ e seu valor positivo indica que uma força motriz externa é necessária para dissociar o conjunto. O valor de $T\Delta S_{diss}$ (16,1 kcal mol⁻¹) indica a mudança de entropia de corpo rígido nesse padrão de dissociação.

As regiões de núcleo dos domínios 1 e 2 da proteína estão bem ordenadas, exceções são destacadas pela representação dos B-fatores (Figura 15) nas regiões: do *loop* que conecta os domínios 1 com 2, 3 com 4; partes do domínio 1 nas hélices mais externas e no domínio DUF da região C-terminal. Estas regiões sugerem apresentar alta flexibilidade, sendo mais acentuada no C-terminal.

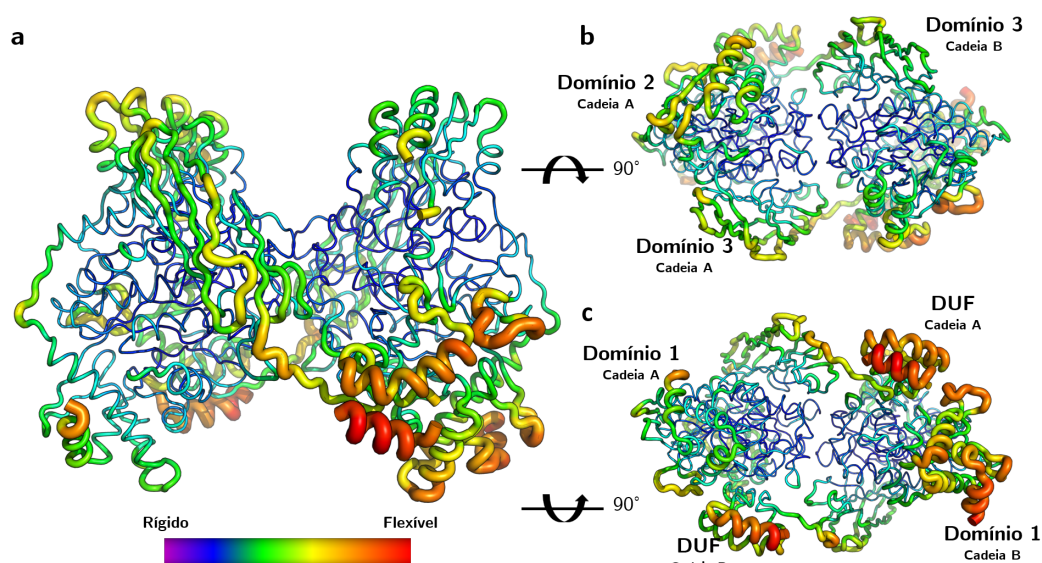


Figura 15 – Fatores de temperatura (representação cilíndrica - sausage) da *CtBgl3B*. **a**. Visão geral da estrutura, onde a espessura é proporcional aos valores de B-fator, destacando o nível de flexibilidade em cada região. Cores vermelhas indicam altos valores enquanto cores azuis menores valores para B-fator. Setas indicam a posição de rotação da estrutura. **b** e **c** são visões rotacionadas em +90° e -90°, respectivamente, em relação a **a**.

Fonte: Elaborada pelo autor.

O domínio 1 contém o resíduo nucleofílico Asp231 e pode ser sobreposto com os domínios correspondentes em outras estruturas nas posições dos resíduos Asp44, Asp231 e Glu420 que interagem diretamente com a molécula ligante o glicerol (GOL). Molécula

possivelmente oriunda do próprio tampão de purificação da proteína. Na estrutura da *CtBgl3*, outras moléculas de GOL, além das que se localizam no sítio catalítico aparecem na estrutura, sendo uma delas na região da interface entre o dímero estabelecendo ligações de hidrogênio com os resíduos Lys536 e Tyr532 de ambas as cadeias A e B (dados não mostrados). A outra molécula GOL(5) interagindo com uma serina fosforilada (SEP735, cadeia B) [Figura 16](#), está compondo o domínio DUF, na região próxima a interface estabelecida entre este domínio e a cavidade catalítica da molécula parceira. As serinas fosforiladas foram avaliadas pelos mapas de densidade eletrônica ($2F_o - F_c$) e FEM ([210](#)) durante o refinamento. Duas moléculas de β -mercaptoetanol (BME), oriundas do tampão usado na etapa de purificação da proteína, interagindo por ligação de hidrogênio com Ala181, e rodeada pelos resíduos Tyr129, Glu184, Asn185, e Arg630 (dados não mostrados).

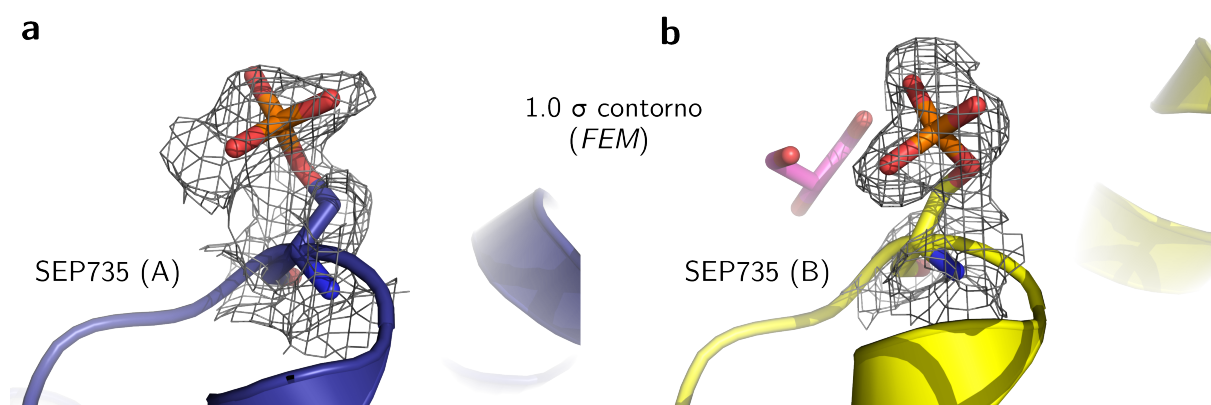


Figura 16 – Visão do mapa de densidade eletrônica *feature-enhanced map* (FEM) na localização das serinas fosforiladas (SEP735). Representação em bastão dos resíduos de serinas fosforiladas (SEP735) observados na estrutura da *CtBgl3B*, cadeias A e B respectivamente, e do mapa FEM na cor cinza com nível de contorno 1.0σ .

Fonte: Elaborada pelo autor.

A [Figura 17b](#) e [d](#) mostra em detalhes a região do sítio ativo da *CtBgl3B*, onde é possível observar que os resíduos estão organizados em estruturas secundárias de *loops*, sugerindo flexibilidade na acomodação dos substratos. A superfície acessível a solvente na interface de interação do dímero, cadeias A e B, foi calculada pelo servidor *PISA* ([142, 211](#)), sendo 2328 \AA^2 , composta por 41 interações de hidrogênio e 12 pontes salinas. São os indicativos, por meio da estrutura cristalográfica, de que o estado dimérico é o mais estável [Figura 17a](#), tal como observado na própria unidade assimétrica. Os experimentos em solução, SEC-MALS e SAXS também corroboram no sentido de que este é o arranjo mais estável. Todavia, nesta interface de contatos entre as cadeias, é possível notar que o domínio DUF da região C-terminal está localizado próximo à fenda catalítica da cadeia parceira, as interações estão destacadas na [Figura 17c](#).

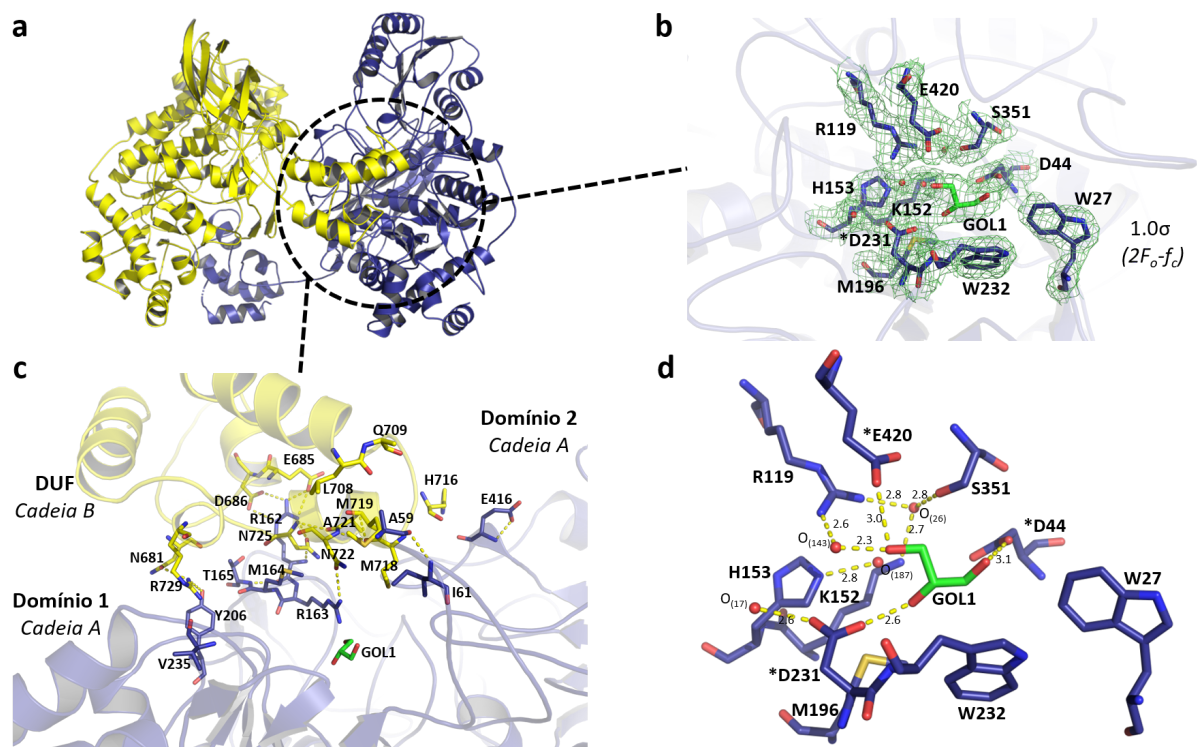


Figura 17 – Visão detalhada do sítio ativo e da região de interface entre DUF e cavidade catalítica. **a.** Representação *cartoon* do dímero de *CtBgl3B* com destaque para a região de interface entre o domínio de função desconhecida (DUF) com a região da cavidade catalítica. **b.** Mapa de densidade eletrônica para os resíduos do sítio catalítico e molécula de glicerol (GOL) como ligante mostrado verde-claro. Mapa ($2F_o - F_c$) com nível de contorno 1.0σ . **c.** Ampliação da região referente à interface de contato entre o domínio (DUF, em amarelo), da cadeia vizinha, com a região de hélices e *loops* em torno da cavidade catalítica (domínios 1 e 2, em azul). Resíduos que fazem contatos via interações de hidrogênio ou hidrofóbicas destacados por representação de bastão. **d.** Sítio catalítico destacando as interações com a molécula ligante (glicerol - GOL) e também com moléculas de água (esferas em vermelho). (*) Resíduos que interagem diretamente com o ligante. Representação de bastão para os resíduos, codificação de uma letra e coloração conforme a cadeia. Interações (linhas tracejadas em amarelo) e comprimento em angstrom (Å).

Fonte: Elaborada pelo autor.

Na Figura 17d, podem ser observados com detalhes os contatos diretos entre os resíduos Asp44, Asp231 e Glu420 com a molécula de GOL, os resíduos His153, Arg119, Ser351, Met196 e os Trp27 e Trp232 também compõem o sítio catalítico, e são comumente observados nas hidrolases de glicosídeos, mesmo em diferentes famílias. É comum, atribuir as características de seletividade para substratos dessas enzimas baseando-se nas posições destes resíduos no sítio ativo. E no caso desta enzima, *CtBgl3B*, também é razoável considerar que o domínio DUF pode ter alguma relação com a especificidade por substratos

ou influenciar diretamente nas taxas das reações por ela catalisadas.

4.4 Ensaios bioquímicos

Na primeira etapa dos ensaios de atividade foi necessário conhecer os parâmetros de pH e temperatura ($^{\circ}\text{C}$) ótimas para o trabalho com a enzima *CtBgl3B*. Onde por uma varredura dessas duas condições, apresentados na [Figura 18](#), observou-se melhor atividade catalítica para pH ácido de 5 e 5,5, sendo o último com maior percentual de atividade (100%). Todavia em pH 7, a enzima ainda apresenta cerca de 50% de sua atividade. Enquanto para temperatura, os melhores valores estão situados entre 60–80 $^{\circ}\text{C}$, com o maior percentual de atividade para 70 $^{\circ}\text{C}$, temperatura característica para enzimas termófilas.

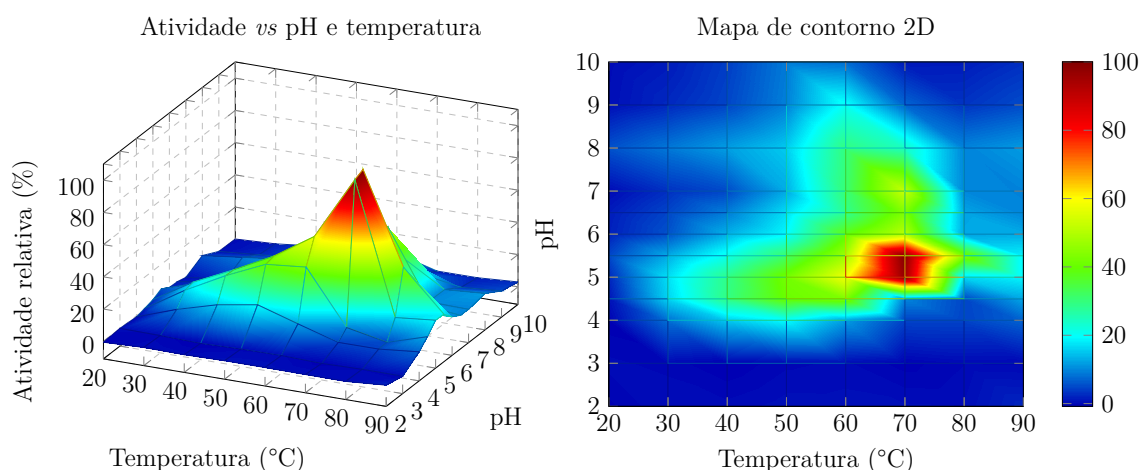


Figura 18 – **Efeito do pH e temperatura na velocidade das reações catalisadas pela *CtBgl3***. Análise de superfície de resposta do ensaio de varredura das condições de pH e temperatura ($^{\circ}\text{C}$) em função da Atividade relativa (%) da enzima *CtBgl3*. Onde, os valores para variável de resposta variam da cor azul (menores) ao vermelho (maiores). Todas as medidas foram realizadas em quadruplicata.

Fonte: Elaborada pelo autor.

Um detalhe a se destacar é que os valores de T_m observados, tanto pelas técnicas DSF (71 $^{\circ}\text{C}$) como CD (68 e 71 $^{\circ}\text{C}$), são compatíveis com a mais alta temperatura de melhor atividade enzimática, no entanto, conforme os ensaios térmicos 70 $^{\circ}\text{C}$ estaria no limite para desnaturação da enzima.

Para avaliar esses limites de temperatura, foi realizada a avaliação da atividade residual pelo efeito do tempo de incubação e temperatura, para estimar a estabilidade térmica da enzima. A [Figura 19](#) apresenta estes resultados, onde no gráfico inserido a proteína *CtBgl3B* demonstrou drástica perda na sua capacidade catalítica quando mantida nas temperaturas acima de 65 $^{\circ}\text{C}$, chegando a praticamente nula após o decorrer de 1 h.

Enquanto, a 60 °C, sofreu perdas de metade da sua capacidade catalítica após o mesmo tempo de experimento, observando baixa atividade ($\approx 20\%$) até depois de 3 h.

Os perfis das curvas de atividade residual sugerem que a enzima não resiste a prolongados períodos em altas temperaturas (acima de 65), ainda que tenha demonstrado anteriormente alta atividade em 70 °C. Apesar disso, abaixo do possível limite de 60 °C, a proteína quando incubada nas temperaturas de 35 – 50 °C manteve sua capacidade enzimática oscilando com valores acima de 85% pelo longo período de até 72 h. Indicando que, a 50 °C pode ser a mais alta temperatura onde a *CtBgl3B* mantém melhores índices de reatividade frente os substratos.

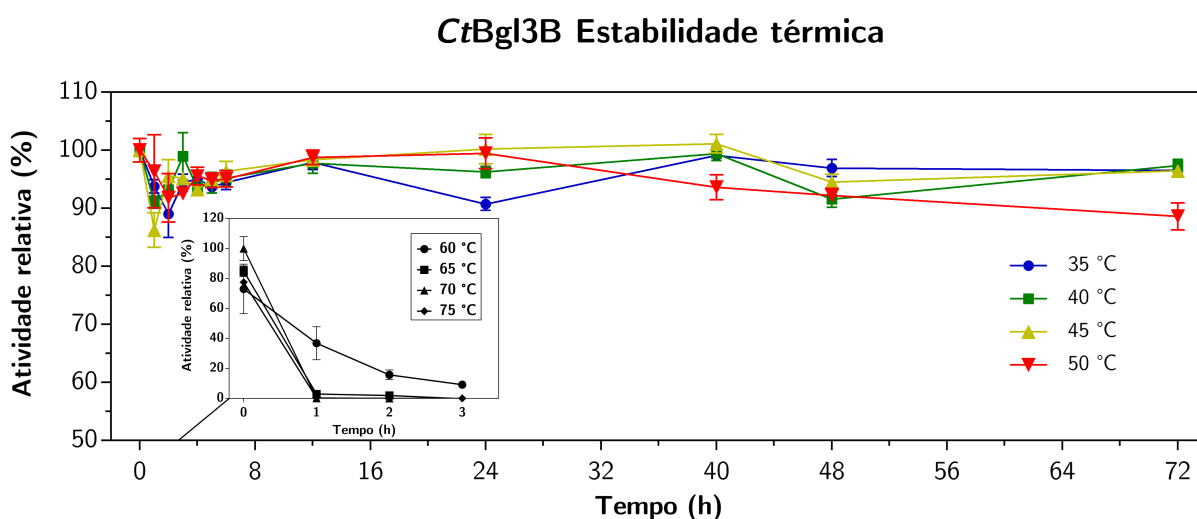


Figura 19 – Efeito da temperatura na reação catalisada por *CtBgl3B*. a. Curvas de atividade enzimática residual com proteína incubada em diferentes temperaturas. Experimento foi realizado com tempo máximo de 3 horas.

Fonte: Elaborada pelo autor.

4.4.1 Efeito de aditivos químicos nas reações enzimáticas

Aditivos químicos, presentes no meio reacional, podem interferir potencializando ou inibindo a atividade enzimática. Na Figura 20, tomando como base a reação controle (condições pré-estabelecidas), foi realizado experimento testando painel de aditivos variados entre íons salinos, solvente (DMSO), sal quelante (EDTA) e detergentes. Aqui, há que se destacar o aumento da atividade enzimática da *CtBgl3B* da ordem de $\approx 70\%$ mais ativa na condição com adição de Triton-X100, de $\approx 60\%$ com Tween20 e ligeiro aumento de $\approx 5\%$ com a presença de EDTA. Estimativas desconsiderando as margens de erro, tomando valores médios. Os destaques negativos ficam para as demais condições avaliadas, onde atividade esteve na ordem de apenas 25–75% de sua capacidade total, e as piores condições com adição de CuSO_4 e SDS onde a atividade foi abaixo dos 25% de sua capacidade.

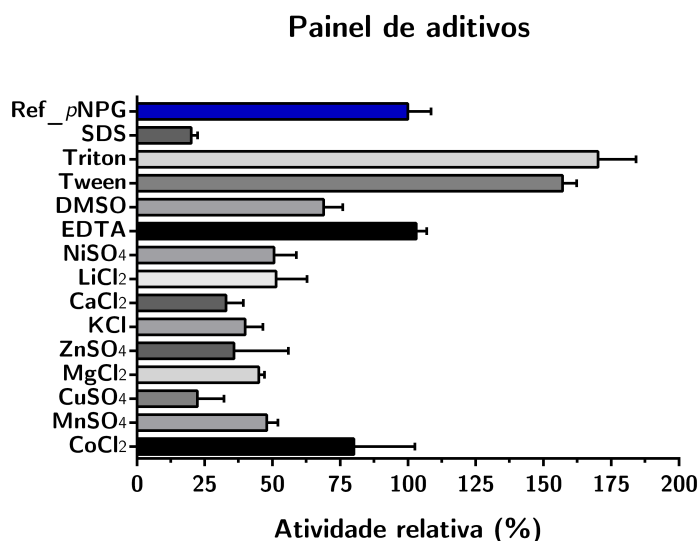


Figura 20 – **Efeito de aditivos químicos nas reações catalisadas pela *CtBgl3*.** Reações com base na concentração do produto formado na reação *p*-NP.

Fonte: Elaborada pelo autor.

Os detergentes Triton-X100 e Tween20 [Figura 22e](#), produziram expressivo aumento nas taxas de conversão das reações enzimáticas da *CtBgl3B*, além do sal quelante EDTA que apresentou valores próximos ao de referência com ligeiro aumento, conforme a [Figura 21a](#). Assim, estes aditivos foram os selecionados para reações variando suas concentrações, os detergentes 0,1%, 0,05% e 0,02%, e o EDTA 0,5, 0,25 e 0,1 mM. Seguido das análises de variância ANOVA e comparações múltiplas pelo teste de Tukey. Na [Figura 21d](#), está o resumo dos parâmetros da análise, e os resultados com P-valor < 0,0001 indicando diferenças significativas entre os valores comparados. O R^2 foi de 0,9559 para gráfico de resíduos ([Figura 21c](#)), neste que, são observados apenas dois pontos mais discrepantes para as condições de Triton 0,1% e 0,05%. Na [Figura 21b](#), está a representação gráfica dos intervalos de confiança (95%) dos resultados, observando que as condições com adição de EDTA não apresentaram diferenças significativas entre si, nas diferentes concentrações aqui testadas. Em contraste, devido ao aumento significativo dos valores, as condições de maior aprimoramento foram constatadas para as adições de Triton 0,1%, e Tween20 0,05%, respectivamente. Considerando para isso, a menor quantidade de aditivo com a maior influência positiva na variável resposta, atividade relativa (%) ou taxa de conversão em produto.

Os surfactantes são frequentemente usados durante a produção de bioetanol ([212](#)), fornecendo um meio para aumentar a disponibilidade superficial da biomassa (vs pré-tratamento ácido, explosão de vapor, etc.), mas, em simultâneo, podem minimizar a interação das enzimas com a lignina que leva à rápida desativação durante a hidrólise ([213, 214](#)).

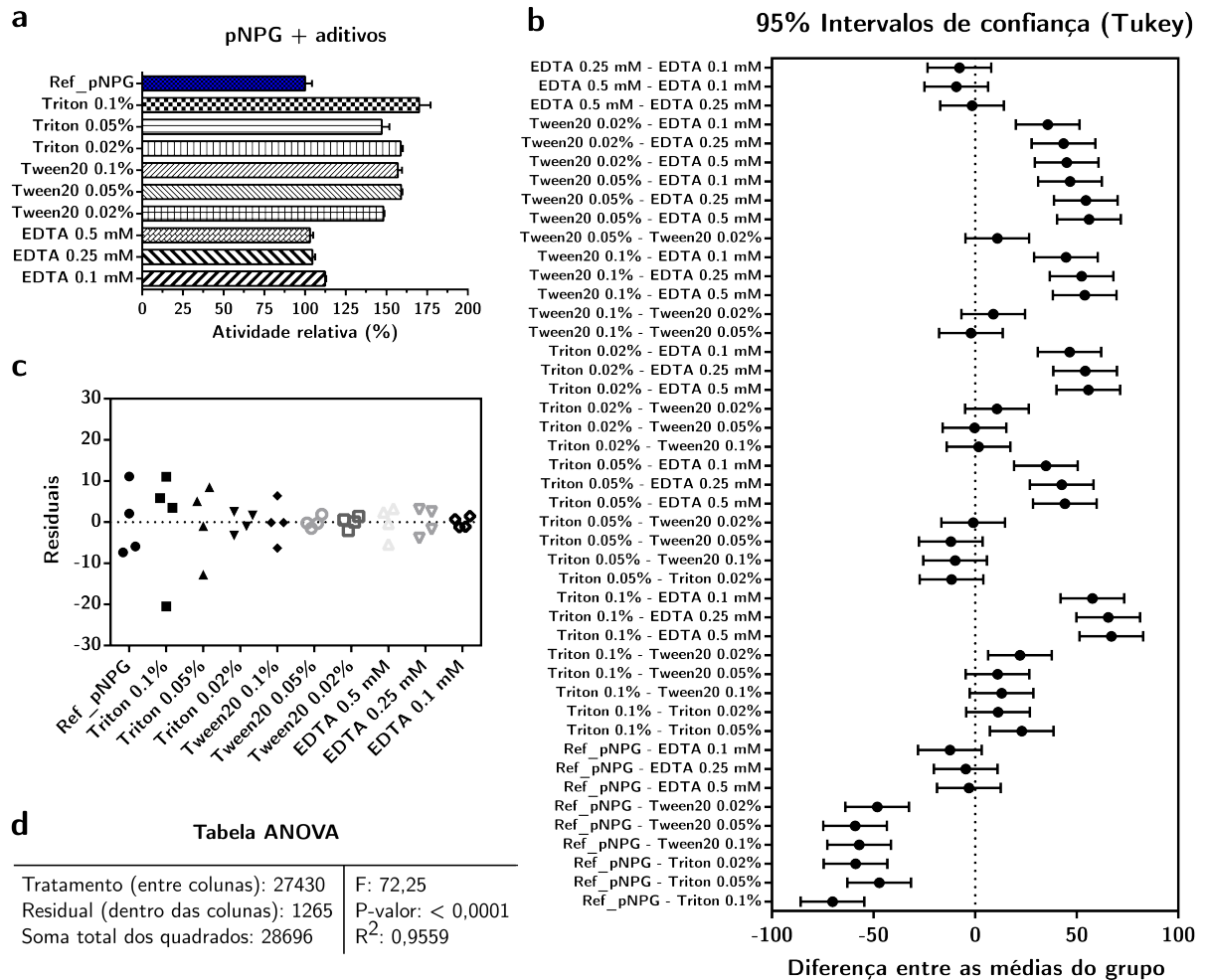


Figura 21 – Efeito dos aditivos: Triton, Tween20 e EDTA nas reações catalisadas pela *CtBgl3B*. **a**. Efeito nos valores de atividade relativa, sendo *p*NPG a referência, com os detergentes Triton-X100 e Tween20, nas concentrações 0,1%, 0,05% e 0,02%. E do EDTA, nas concentrações 0,5 mM, 0,25 mM e 0,1 mM. **b**. Gráfico dos intervalos de confiança (95%) das medidas no teste de Tukey. **c**. Gráfico de residuais da estatística. **d**. Tabela de parâmetros da estatística ANOVA, P-valor < 0,0001 indicando diferenças significativas entre os valores nas comparações múltiplas da análise com R^2 do gráfico de resíduos igual a 0,9559.

Fonte: Elaborada pelo autor.

Além disso, vários surfactantes, como os não iônicos Triton X-100 e Tween (20 e 80) (215–217), têm sido empregados para aumentar diretamente a eficiência enzimática das celulasas (215, 218, 219), embora com resultados modestos, $\approx 20\%$ de melhoria. Estes aumentos na eficiência também são dependentes das condições, do meio reacional, já que em situações com a celulose pura o aumento não ocorre de forma tão consistente (218).

O aumento da eficiência enzimática de celulase β -glucosidase frente a adição de tensoativos já foi reportada na literatura, por exemplo, no trabalho publicado por Seidel e Lee(220). Segundo os autores, o uso de um tensoativo fotoresponsivo (UV — 350 nm e Vis — 434 nm) baseado em compostos de azobenzeno (4-etil-4'(trimetilamino-butoxi)-azobenzeno brometo - azoTAB) promoveu uma transição progressiva no estado oligomérico para formação de monômeros, conforme o aumento na concentração do azoTAB. De forma que esta transição foi acompanhada de um aumento em 60% da atividade catalítica. Em contraste, a enzima em questão foi inativada na presença de tensoativos baseados em hidrocarbonetos de cadeia linear.

4.4.2 Especificidade por substratos

Também foram realizados ensaios para testar atividade da *CtBgl3B* contra outros substratos sintéticos, formados por diferentes grupos ligados com o *p*-NP, mimetizando diferentes substratos naturais que eventualmente a enzima seria capaz de hidrolisar. Novamente tomando como base os valores de hidrólise do *p*NPG como referência (100% da atividade), observou-se na Figura 22a que a enzima apenas demonstrou capacidade de hidrolisar os substratos *p*NPX ($\approx 20\%$ da atividade relativa) e *Cp*NPG2 ($\approx 50\%$ da atividade relativa), todos formados por ligação β -1,4 (Figura 22d). Indicando alta especificidade para um pequeno repertório de substratos e ligações glicosídicas, Figura 22f. Destacando que, frente aos substratos *p*NPAF, *p*NPB GAL e *p*NPBM não foi detectada atividade catalítica da enzima *CtBgl3B*.

Frente aos substratos *p*NPX e *Cp*NPG2 também foram avaliados os efeitos dos aditivos selecionados, Triton 0,1% e Tween20 0,05%, nas reações enzimáticas com a *CtBgl3B*. Similarmente aos valores observados contra o *p*NPG (referência), houve também aumento significativo nas taxas de conversões em produto. A adição de Triton (0,1%) significou para hidrólise de *p*NPX aumento de 23%, e para o *Cp*NPG2 aumento de 37%, Figura 22b. A adição de Tween20 (0,05%) significou aumentos das hidrólises do *p*NPX de 10%, e para o *Cp*NPG2 de 39%, Figura 22c. Estes valores foram estimados com base no valor de atividade relativa da hidrólise do *p*NPG sem adição de aditivos como referência, indicado pela linha tracejada nos gráficos da Figura 22.

De outra forma, agora considerando as taxas de hidrólise individuais de *p*NPX e *Cp*NPG2 sem a adição de aditivos como seus próprios valores de referência, e sendo assim o valor normalizado de 100% das suas atividades relativas. Temos para o *p*NPX que a

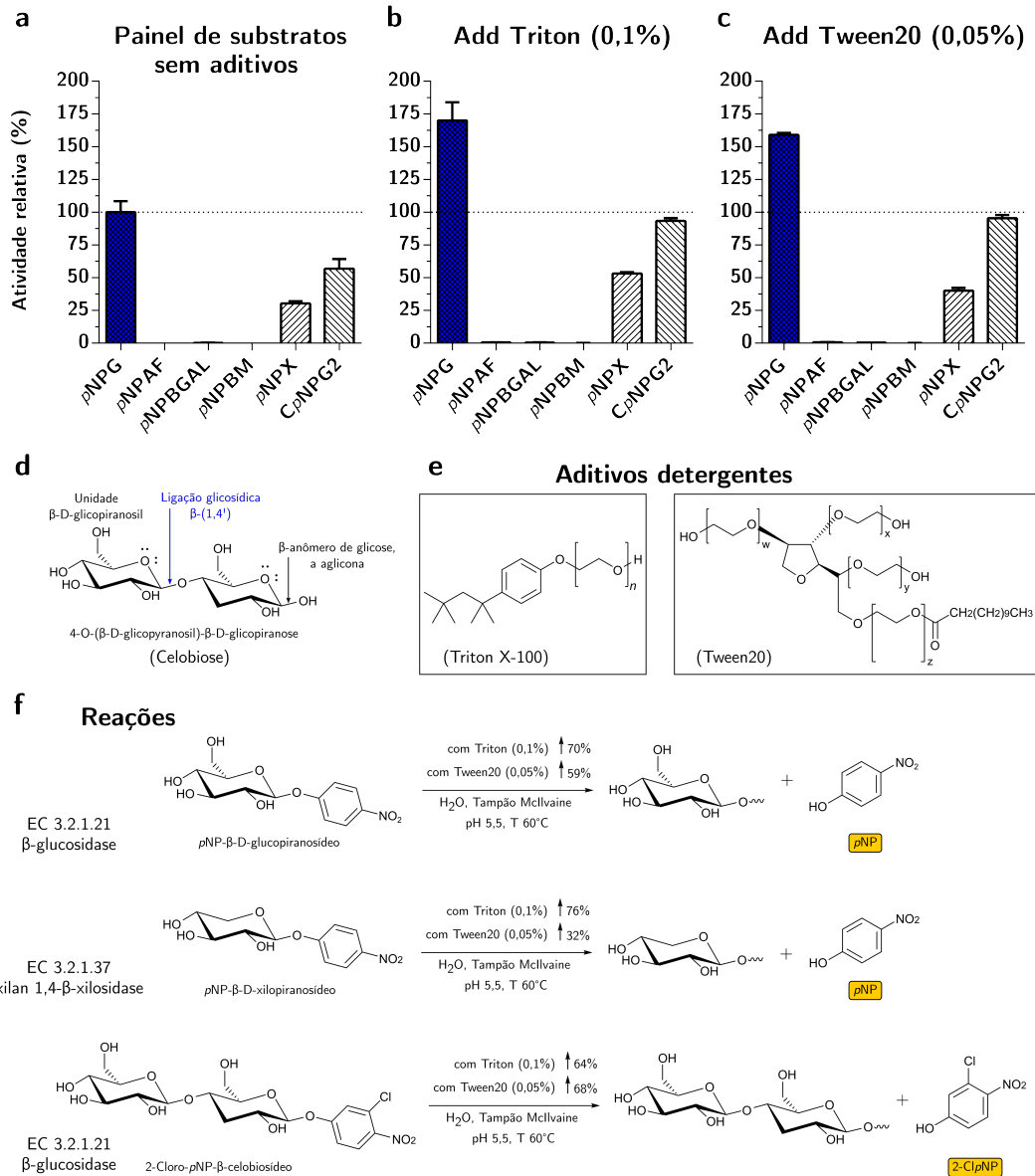


Figura 22 – **Painel de substratos.** **a-c.** Especificidade da *CtBgl3B* frente a diferentes substratos sintéticos, efeito dos aditivos surfactantes (Triton e Tween20). **d.** esquema da molécula de celobiose e da ligação β-(1,4'). **e.** Representação estrutural dos aditivos surfactantes Triton X-100 e Tween20. **f.** Esquema das reações catalisadas pela *CtBgl3B* e suas condições. Rendimentos das reações com base na concentração do produto (*pNP*) formado.

Fonte: Elaborada pelo autor.

adição de Triton (0,1%) e Tween20 (0,05%) causaram um aumento total de 76% e 32%, respectivamente. Enquanto para o *CpNPG2* esses aditivos promoveram um aumento total de 64% e de 68%, respectivamente nas taxas de conversões dos produtos das reações.

A atividade específica, definida como U mg^{-1} , é a quantidade de proteína que libera 1 μmol de *p*-nitrofenol por minuto por miligrama ($\mu\text{mol}/\text{min}/\text{mg}$) de substrato nas condições do ensaio. Os valores foram obtidos a partir da V_{max} medida nos experimentos de cinéticas, como mostrados na [Figura 23](#) e na [Tabela 4](#). A enzima *CtBgl3B* tem a atividade específica frente ao substrato *pNPG* (124 U mg^{-1}), com os aditivos Triton (153 U mg^{-1}) e Tween20 (136 U mg^{-1}). Frente ao *CpNPG2* (14 U mg^{-1}), com os aditivos Triton (55 U mg^{-1}) e Tween20 (65 U mg^{-1}). Estes são substratos para aferir a atividade enzimática de β -glicosidases (EC. 3.2.1.21), hidrólise de resíduos terminais de β -D-glicosil não redutores com liberação de β -D-glicose.

O substrato *pNPX*, serve para aferir a atividade enzimática de xilan 1,4- β -xilosidases (EC. 3.2.1.37), hidrólise de (1 \rightarrow 4) β -D-xilanos, para remover sucessivos resíduos de D-xilose dos terminais não redutores. Neste caso, a atividade específica foi de 32 U mg^{-1} sem os aditivos, com o Triton 105 U mg^{-1} e com o Tween20 69 U mg^{-1} . Os esquemas ilustrativos destas reações está representado na [Figura 22f](#). Conforme discutido anteriormente, os aditivos Triton (0,1%) e Tween20 (0,05%), [Figura 22e](#), conseguiram promover o aumento da taxa de conversão dos substratos, desta forma, eles têm influência também na especificidade por substratos da enzima aumentando assim a capacidade de conversão em produtos.

Quando comparados os valores de atividade catalítica com outras enzimas, da literatura, ou comerciais (Megazyme[®]), [Tabela 4](#), fica evidente que *CtBgl3B* é uma BGL específica, tal como observado para sua família GH3 ([41](#)). Possui atividade específica, contra o *CpNPG* (124 , 153 e 136 U mg^{-1}), inferior à *CrBgl* 419 U mg^{-1} ([221](#)), *AoBgl3* 294 U mg^{-1} ([222](#)), *PcBgl3* que a $60 \text{ }^\circ\text{C}$ é 300 U mg^{-1} (E-BGOSPC - Lot 121101f) e *ThBgl3B* $11.86 \times 10^4 \text{ U mg}^{-1}$ ([187](#)), sendo esta última a mais destacada dentre as comparadas. Porém, demonstra ser superior em relação as demais, com os valores maiores ou em similar escala em suas variadas condições (aditivadas com Triton e Tween20 respectivamente).

A enzima *CtBgl3B* quando comparada com duas D-xilosidases comerciais (Megazyme[®]), da família GH43, em relação à hidrólise do substrato *pNPX*, tem atividade específica de 32 , 105 e 69 U mg^{-1} , superior à *BpBxl43* 18 U mg^{-1} (E-BXSEBP - Lot 120201b). Enquanto, na condição com adição de Triton está em mesma ordem de grandeza com a *SrBxl43* 115 U mg^{-1} (E-BXSR-1KU - Lot 100301f).

4.4.3 Cinética enzimática

Para aprofundar nas características enzimáticas da *CtBgl3B*, os parâmetros cinéticos com modelo *Michaelis-Menten* foram determinados utilizando os substratos onde ela foi hidroliticamente ativa *pNPG*, *pNPX* e *CpNPG2*. Na [Figura 23](#) são apresentados os gráficos

dos ensaios cinéticos, enquanto os resultados estão sumarizados na [Tabela 4](#). Os ensaios foram realizados nas condições ótimas de pH e temperatura, previamente estabelecidas. Desta maneira, foi observado que a *CtBgl3B* hidrolisou *p*NPG com os valores de K_M sendo 0,371 mM (sem aditivos), e a afinidade pelo substrato foi aumentada com a presença do Tween20 para 0,405 mM e com Triton para 0,415 mM. Conforme esperado, da mesma forma ocorreu frente aos demais substratos, para *p*NPX foi de 0,736 mM (sem aditivos), aumentou na presença de Tween20 para 1,234 mM e de Triton para 2,606 mM. No caso do *Cp*NPG2 foi 0,308 mM (sem aditivos), com Tween20 0,348 mM e com Triton 0,359 mM. Os valores de K_M são usualmente uma medida relativa da afinidade de ligação ao substrato, assim, a presença dos aditivos afetou diretamente a variação dessa afinidade, principalmente frente ao substrato *p*NPX onde o aumento foi mais significativo.

Os valores de K_M da *CtBgl3B* são menores que os da *LinBgl* (48 mM) ([223](#)), *BtBgl* (1,3 mM) ([224](#)), *PaBgl* (14 mM) ([225](#)) e ligeiramente menor que da JMB19063 (0,79 mM) ([226](#)). É maior ou em escala similar as demais, exceto para *KmBglII* de *Kluyveromyces marxianus* que são, respectivamente, os menores valores 9.9×10^{-2} mM (*p*NPG) e 2.2×10^{-1} mM (*p*NPX) ([227](#)).

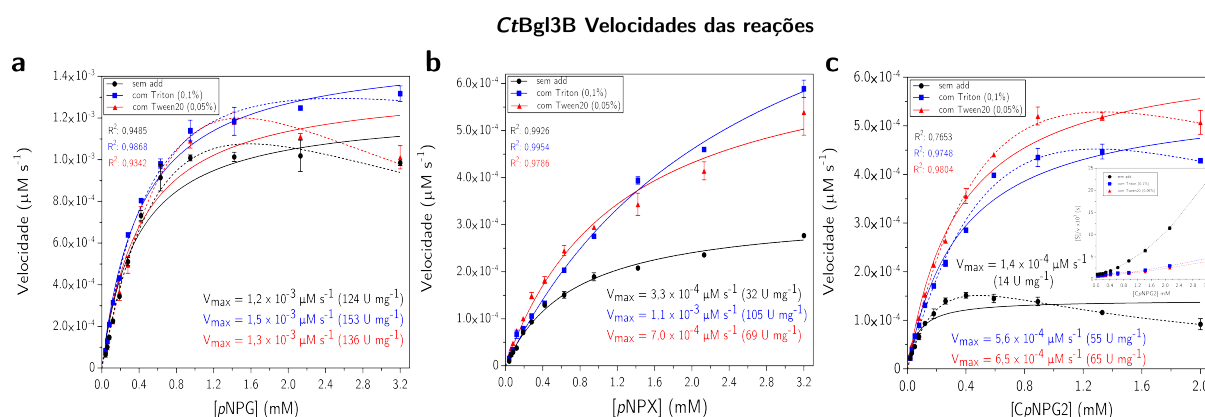


Figura 23 – **Análise do efeito da concentração de substrato na velocidade da reação catalisada por *CtBgl3***. Linhas sólidas são as Curvas de cinéticas enzimáticas usando regressão não linear ajustadas ao modelo de *Michaelis-Menten*. Linhas tracejadas representam ajustes feitos a equação modificada para contabilizar a inibição por substrato. Dados obtidos utilizando variação da concentração dos substratos **a.** *p*NPG, **b.** *p*NPX e **c.** *Cp*NPG2. Velocidades calculadas com base na concentração do produto formado *p*-NP nas reações.

Fonte: Elaborada pelo autor.

Os valores da constante catalítica (k_{cat}) das reações da *CtBgl3B*, que representa o número de renovação ou número de conversões que a enzima pode fazer, foram determinados para o *p*NPG sendo 173 s^{-1} sem aditivos, 191 s^{-1} com Tween20 e 214 s^{-1} com Triton. Para o *p*NPX, na cinética sem aditivos o valor foi de 46 s^{-1} , com Tween20 97 s^{-1} e com

Triton 148 s^{-1} . E para o *CpNPG2*, na condição sem aditivos não foi possível obter um valor confiável diretamente pelo modelo *Michaelis-Menten*, devido à falta de ajuste. Na condição com adição de Tween20 foi de 91 s^{-1} . Diferente das demais reações, frente os outros substratos onde adição de Triton apresentou maior aumento nas taxas de conversão, neste caso a condição com Triton apresentou sutil diferença menor em relação ao Tween20, com k_{cat} de 78 s^{-1} .

É claro que, a curva da cinética de reação referente ao *CpNPG2* na [Figura 23c](#), sem aditivos, não tem o perfil de uma cinética *Michaelis-Menten* convencional com o baixo ajuste R^2 0,7653. Isso é devido ao comportamento não linear observado no gráfico inserido na [Figura 23c](#), onde é representado pela não linearização ao se aplicar o modelo de *Hanes-Woolf* ($[S]/v$ vs $[S]$) ([228](#)), que uma linha reta seria um comportamento típico de uma cinética *Michaelis-Menten*.

Os dados indicam inibição por substrato, e esta foi responsável pela redução nas velocidades de reação em concentrações elevadas de substrato. De fato os dados são mais consistentes com a equação modificada [Equação B.3](#), onde K_i é a constante de dissociação para a ligação do substrato de tal forma que dois substratos podem se ligar a uma enzima, por exemplo, a molécula de celobiose formada na sua hidrólise, conforme visto no esquema da [Figura 22f](#). Considerando assim a formação de um complexo ES_2 cataliticamente inativo, ou seja, com via competidora ($ES_2 \leftrightarrow S + ES \rightarrow E + P$), justificando o baixo valor de eficiência catalítica tal como o falta de ajuste ao modelo inicialmente aplicado na análise. Usando o modelo de inibição por substrato, diante das condições de ensaio com a mesma concentração de enzima (7,14 nM), foram estimados os valores de K_i 0,7742 mM, enquanto o K_M usando este modelo foi de 0,3083 mM, valor corrigido maior em relação ao de *Michaelis-Menten* 0,06 mM.

A *CtBgl3B* tem os valores de k_{cat} inferiores apenas em relação a *AoBgl3* de *Aspergillus oryzae* (370 s^{-1}) ([222](#)), *TpBgl3* ($570,83 \text{ s}^{-1}$) ([229](#)) e *KmBgl1* que é $1.5 \times 10^2 \text{ s}^{-1}$ para *pNPG* e 1.7×10^1 para *pNPX* ([227](#)).

A eficiência catalítica de uma enzima é geralmente definida pela razão k_{cat}/K_M , uma medida dos efeitos nas ligações com o substrato e estabelece uma taxa constante de segunda ordem. Este parâmetro pode ser usado, tanto para comparar diferentes enzimas, como também diferentes substratos para uma enzima em particular. Para a enzima *CtBgl3B*, no caso do substrato *pNPG*, as condições com Triton e Tween20 representaram um aumento na eficiência catalítica, onde sem aditivos foi $469 \text{ mM}^{-1} \text{ s}^{-1}$, com Tween20 sutil diferença para $472 \text{ mM}^{-1} \text{ s}^{-1}$ e com Triton o aumento para $517 \text{ mM}^{-1} \text{ s}^{-1}$.

Para o substrato *pNPX* na condição sem aditivos à eficiência foi de $62 \text{ mM}^{-1} \text{ s}^{-1}$, apesar de o aditivo Triton ter aumentado tanto afinidade (K_M) como a constante catalítica k_{cat} , a eficiência apresentou ligeira diminuição indo para $56 \text{ mM}^{-1} \text{ s}^{-1}$. Em contrapartida, neste caso a presença de Tween20 aumentou a eficiência catalítica para $79 \text{ mM}^{-1} \text{ s}^{-1}$,

mesmo com os parâmetros de atividade específica, K_M e k_{cat} inferiores em relação à adição de Triton.

Similar aos substratos anteriores, com $CpNPG2$ nas condições com os aditivos, o aumento de todos os parâmetros previamente discutidos desde a atividade específica, a afinidade pelo substrato K_M , a constante catalítica k_{cat} , e a eficiência catalítica (k_{cat}/K_M), também teve mudanças significativas. Inclusive, na condição sem os aditivos, k_{cat}/K_M , apresentando o valor de $64 \text{ mM}^{-1} \text{ s}^{-1}$, contra $218 \text{ mM}^{-1} \text{ s}^{-1}$ com Triton e $262 \text{ mM}^{-1} \text{ s}^{-1}$ com Tween20.

As enzimas $TnBgl3B$ com $2002 \text{ mM}^{-1} \text{ s}^{-1}$, $TpBgl3$ $1490 \text{ mM}^{-1} \text{ s}^{-1}$, $KmBglI$ $1.5 \times 10^3 \text{ mM}^{-1} \text{ s}^{-1}$ e $7.6 \times 10^1 \text{ mM}^{-1} \text{ s}^{-1}$, $AoBgl3$ 1.3×10^3 , e por último a $HjCel3A$ de *Hypocrea jecorina* com $3.2 \times 10^5 \text{ mM}^{-1} \text{ s}^{-1}$ (230), são as que apresentam os melhores valores de eficiência catalítica, k_{cat}/K_M , para os respectivos substratos. Isso se deve principalmente, pelo fato de esta razão depender do valor de K_M , que no caso destas enzimas é substancialmente baixo, fazendo com que suas eficiências sejam tão mais altas. Particularmente nas condições com e sem aditivos, a $CtBgl3B$ apresentou maior k_{cat}/K_M que as demais enzimas.

Quando se compara a enzima $CtBgl3B$, com as da Tabela 4 em relação ao pH ótimo, a faixa de atividade não é tão discrepante e com a tendência para valores ácidos abaixo de pH 6,0, e variando do menor pH 4,0 para $TpBgl3$ de *Thermotoga petrophila* (229), até o mais alto pH 7,5 para a enzima comercial $BpBxl43$ de *Bacillus pumilus* (E-BXSEBP, Lot 120201b). Ao comparar a temperatura ótima, a $CtBgl3B$ fica na média em $65 \text{ }^\circ\text{C}$, onde a mínima é de $20 \text{ }^\circ\text{C}$ para a $LinBgl$ de *Listeria innocua* (223), e a máxima é de $90 \text{ }^\circ\text{C}$ para a $TnBgl3B$ de *Thermotoga neapolitana* (187). Esta última enzima que também apresentou o valor mais alto de atividade específica ($11.86 \times 10^4 \text{ U mg}^{-1}$) em relação as demais.

A Bgl $CtBgl3B$ se mostrou bastante promissora para aplicabilidade em processos de conversão da biomassa lignocelulósica em açúcar, especialmente devido a sua capacidade em clivar as ligações do tipo β -(1,4') de substratos glucopiranosídicos como o $pNPG$, $CpNPG2$, e xilopiranosídico como o $pNPX$. De forma que na presença dos surfactantes não iônicos, Triton X-100 e Tween20, ainda demonstrou aumento nas taxas de conversão, e aumento na atividade específica frente aos substratos. A $CtBgl3B$ resistiu por até 72h incubada nas temperaturas de 50, 40, 35 e $30 \text{ }^\circ\text{C}$, mantendo sua atividade catalítica residual acima de 80% em relação ao estado inicial, com poucas variações nessas taxas e nunca abaixo dos 80%. Considerando um processo industrial isso é mais uma vantagem.

A estrutura da $CtBgl3B$ revelou uma possível explicação para sua alta estabilidade por seu arranjo dimérico e disposição da interface de contato, onde, um domínio de função desconhecida DUF (resíduos 658–755) da molécula parceira na arquitetura interage com flexibilidade sobre a região da cavidade catalítica, associado a mobilidade dos *loops* com resíduos dispostos no sítio catalítico, e também os resíduos aromáticos da região

Tabela 4 – Parâmetros cinéticos determinados para *CtBgl3B** e outras BGLs.

Organismo, (Enzima ID), n.º acesso	Condição, substrato	At, Esp, U mg ⁻¹	K_M (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_M mM ⁻¹ s ⁻¹	Ref
<i>Clostridium thermocellum</i> , (CtBgl3B), ALX08008.1	pH 5,5, 65 °C					Este trabalho
	pNPG	124,10 a	0,371 ± 0,04a	173,8 ± 6a	469,7 ± 6 a	
	pNPG + Trt	153,40	0,415 ± 0,02	214,7 ± 4	517,3 ± 4	
	pNPG + Twn	136,70	0,405 ± 0,05	191,3 ± 8	472,3 ± 8	
	pNPX	32,91	0,736 ± 0,03	46,1 ± 1	62,6 ± 1	
	pNPX + Trt	105,90	2,606 ± 0,14	148,3 ± 5	56,9 ± 5	
	pNPX + Twn	69,74	1,234 ± 0,10	97,6 ± 4	79,1 ± 4	
	CpNPG2	14,14	0,067 ± 0,01	19,8 ± 1	295,5 ± 1	
	CpNPG2 + Trt	55,98	0,359 ± 0,03	78,3 ± 2	218,1 ± 2	
	CpNPG2 + Twn	65,31	0,348 ± 0,02	91,4 ± 2	262,6 ± 2	
Metagenomica, JMB19063	pH 6,5, 50 °C					
	pNPG	ND	0,79	150	198	(226)
<i>Thermotoga neapolitana</i> , (TnBgl3B), WP_038066943.1	pH 5,6, 90 °C					
	pNPG	ND	0,06	129	2002	(187)
<i>Thermotoga petrophila</i> , (TpBgl3)	pH 4,0, 70 °C					
	pNPG	ND	0,38	570,83	1490,43	(229)
<i>Bifidobacterium adolescentis</i> , (BaBgl3B), A1A2P2	pH 6,0, 30 °C					
	pNPG	ND	0,32	88	275	(231)
<i>Chaetomella raphigera</i> , (CrBgl), A0A068FT77	pH 5, 55 °C					
		419	0,12	ND	ND	(221)
<i>Pseudomonas aeruginosa</i> , (PaBgl), Q9I311	pH 7, 25 °C					
	pNPG	ND	14	0,057	4,1	(225)
<i>Bacteroides thetaiotaomicron</i> , (BtBgl), Q8A1U1	pH 5,5, 40 °C					
	pNPG	ND	1,3	9,8	7,3	(224)
<i>Listeria innocua</i> , (LinBgl), Q92AS9	pH 6,0, 20 °C					
	pNPG	ND	48	3,1	15	(223)
<i>Kluyveromyces marxianus</i> , (KmBglI), D1GCC6	pH 5,5, 30 °C					
	pNPG	ND	$9,9 \times 10^{-2}$	$1,5 \times 10^2$	$1,5 \times 10^3$	(227)
	pNPX		$2,2 \times 10^{-1}$	$1,7 \times 10^1$	$7,6 \times 10^1$	
<i>Aspergillus aculeatus</i> , (AoBgl3), P48825	pH 5,0, 37 °C					
	pNPG	294	0,42	ND	ND	(222)
<i>Aspergillus oryzae</i> , (AoBgl3)	pH 5,0, 50 °C					
	pNPG	ND	0,29	370	$1,3 \times 10^3$	(222)
<i>Hypocrea jecorina</i> , (HjCel3A), Q12715	pH 5,7, 37 °C					
	CpNPG2	ND	0,087	28,0	$3,2 \times 10^5$	(230)
Comercial (Megazyme®)						
<i>Aspergillus niger</i> , (AnBgl3), E-BGGLUC	pH 4,0, 40 °C					
	pNPG	80	ND	ND	ND	Lot 150102c
<i>Agrobacterium sp.</i> , (AsBgl1), E-BGOSAG	pH 6,5, 40 °C					
	pNPG	120	ND	ND	ND	Lot 100201d
<i>Phanerochaete chrysosporium</i> , (PcBgl3), E-BGOSPC	pH 5,0					
	pNPG, 40 °C	100	ND	ND	ND	Lot 121101f
	pNPG, 60 °C	300	ND	ND	ND	Lot 121101f
<i>Thermotoga maritima</i> , (TmBgl1), E-BGOSTM	pH 6,5, 40 °C					
	pNPG	70	ND	ND	ND	Lot 151102a
<i>Selenomonas ruminantium</i> , (SrBxl43), E-BXSR-1KU	pH 5,3, 40 °C					
	pNPX	115	ND	ND	ND	Lot 100301f
<i>Bacillus pumilus</i> , (BpBxl43), E-BXSEBP	pH 7,5, 35 °C					
	pNPX	18	ND	ND	ND	Lot 120201b

ND: não disponível; At. Esp.: Atividade específica; U: Unidade é definida como a quantidade de proteína que libera 1 μmol de *p*-nitrofenol por minuto por miligrama ($\mu\text{mol}/\text{min}/\text{mg}$) de proteína nas condições do ensaio;

^a Reações foram incubadas por 5 min com *CtBgl3B* na concentração de 7,14 nM em 100 μL (volume de reação);

^b Dados representam as médias \pm SD ($n = 4$ medidas independentes);

^c Trt: Triton X-100 (0,1%); Twn: Tween20 (0,05%).

Fonte: Elaborada pelo autor.

catalítica. A flexibilidade do DUF, talvez seja maior quando as reações têm a presença dos surfactantes Triton X-100 e Tween20 facilitando a entrada dos substratos, algo que explicaria o aumento da atividade catalítica nestas condições.

5 ESTRUTURA DA HIPOTÉTICA PURINA NUCLEOSÍDEO FOSFORILASE *BLYLMD* YfIH-LIKE DE *BACILLUS LICHENIFORMIS*

5.1 Introdução

A proteína hipotética *BLYmD* tem um domínio pertencente à família de proteínas contendo a díade catalítica Cys-His (Cisteína-Histidina), que é encontrada em três homólogos distantes: CNF1 (PF05785), CheD (PF03975) e YfiH (Acesso: cd16832) (232). A função YfiH permaneceu desconhecida por um longo tempo até que fragmentos de evidências sugeriram que ela poderia estar relacionada com o controle de qualidade de muopeptídeos e, portanto, implicando na manutenção da composição de peptídeo glicano e integridade estrutural em *Escherichia coli* (233). Após um longo período na ausência de qualquer informação funcional para este domínio de *Bacillus licheniformis*, observou-se que há ampla conservação entre gêneros bacterianos, e informações de que a complementação heteróloga de *E. coli* YfiH mutante com os ortólogos clonados de *Bacillus subtilis* e *Mycobacterium tuberculosis* fornecem evidências em favor da existência de um mecanismo conservado de fator contribuinte para a manutenção da composição bacteriana de peptidoglicanos (233).

A sequência de aminoácidos da *BLYmD* indicava a presença de um domínio Cu-oxidase_4 (PF02578), também encontrado em lacases, e amplamente distribuído entre procariotos e eucariotos. Para uma lacase típica, vários estudos estruturais e comparativos identificaram muitas regiões importantes e conservadas onde os resíduos de histidina coordenam os átomos de cobre (234–237). Existem três tipos principais envolvidos na atividade catalítica (202): Tipo-1 (T1) ou centro de cobre azul, Tipo-2 (T2) ou cobre padrão e Tipo-3 (T3) ou centros de cobre binuclear acoplados, cada um com sua maneira única de coordenar íons de cobre. Em T1, o cobre mostra coordenação com duas histidinas, uma cisteína e uma metionina. Em T2, o cobre é coordenado por duas histidinas e uma molécula de água, e em T3 o cobre é coordenado por três histidinas e uma ponte hidroxila (238).

Apesar do extenso estudo focado em oxidases multicobre (MCO), particularmente na família das lacases (239–241), pouco foi alcançado com a *YlmD* para uma atividade compatível com a família das lacases, e algumas dessas proteínas ainda estão sendo depositadas no PDB como proteínas de função desconhecida, surgindo a necessidade de novas investigações desta nova classe de proteína. Neste estudo, é apresentado desde a purificação da proteína até a determinação da estrutura de uma hipotética YfiH de *B. licheniformis* *BLYmD* que foi resolvida com a resolução de 2,66 Å. O modelo foi comparado com lacases típicas e atípicas, dados não mostrados devido a baixa similaridade. Também foi realizado ensaio de fluorimetria diferencial de varredura (DSF) para detectar as condições tampão que promovem a estabilidade da proteína.

Recentemente, Cader *et al.*(129), demonstraram que a proteína Ylmd de *Geobacillus stearothermophilus* (PBD: 6T1B; 6T0Y, 55,38% de identidade) consegue clivar fosforoliticamente a adenosina em adenina e ribose-1-fosfato. Esta proteína, cuja estrutura foi obtida em complexo com uma molécula de inosina como ligante na região ativa, é referente a porção C-terminal ou domínio DUF152 da purina nucleosídeo fosforilase (gene LACC1) de *Homo sapiens*, também conhecida como FAMIN (*Fatty Acid Metabolism-Immunity Nexus*), e é correlacionada com doenças humanas. Esta porção truncada, chamada FAMIN Δ 176 pelos autores, partilha homologia com outros ortólogos bacterianos e contém todas as atividades catalíticas mencionadas na Figura 24, funcionando como: adenosina deaminase (EC.3.5.4.4), purina nucleosídeo fosforilase (EC 2.4.2.1) e metiltioadenosina fosforilase (EC 2.4.2.8).

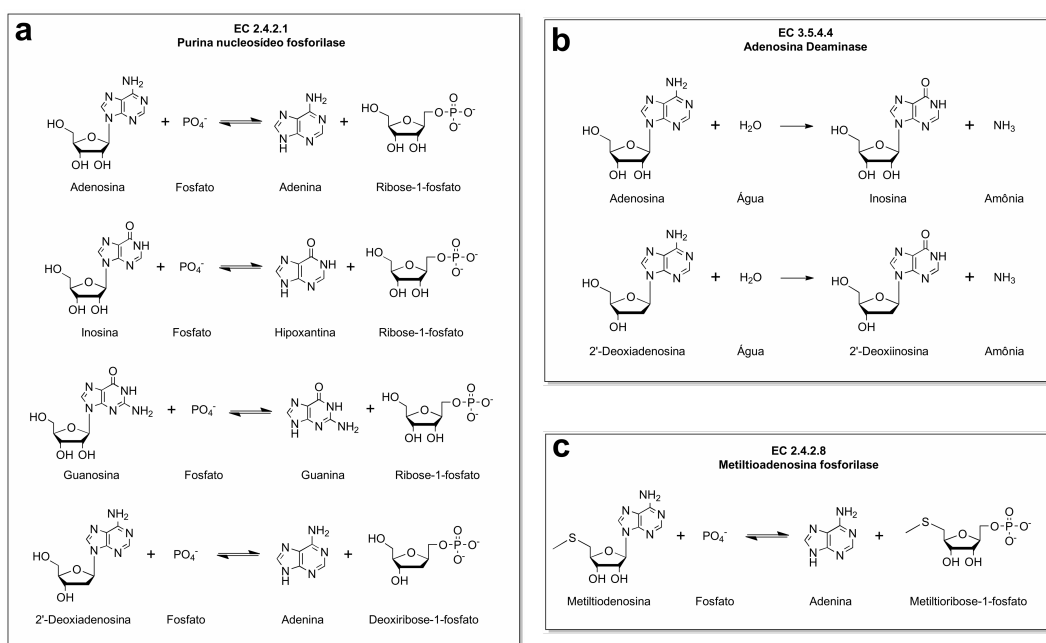


Figura 24 – Reações enzimáticas catalisadas pelo domínio DUF152 da proteína FAMIN, potencialmente reproduzidas pela *BIYlmd*.

Fonte: Elaborada pelo autor.

5.2 Resultados

5.2.1 Expressão e purificação

A proteína recombinante *BIYlmd* foi expressa em *E. coli* Rosetta(DE3), usando meio de auto-indução, seguida pela lise celular, e então a proteína foi submetida às etapas purificação, utilizando-se IMAC contendo íon níquel Ni^{2+} . Bandas compatíveis com a massa da proteína *BIYlmd* (31 kDa) podem ser observadas na Figura 25a: onde a linha da proteína eluída da coluna com 300 mM de imidazol (E300); amostra concentrada antes da clivagem (Con); e proteína clivada mais a TEV protease (Cli). As frações contendo a

enzima de interesse, clivada (FT2), foram concentradas ($\approx 1\text{--}2\text{ mg mL}^{-1}$) e submetidas a terceira etapa de purificação, por cromatografia de exclusão por massa molecular (Superdex 200 16/60; tampão 20 mM Tris-HCl pH 7,5; 150 mM NaCl; 5% glicerol). O cromatograma para a purificação por exclusão de massa molecular é apresentado na [Figura 25a](#). As frações coletadas referentes ao Pico 1 estão juntas na linha P1 ([Figura 25b](#)).

Finalmente, a proteína *BIYImD* foi obtida purificada com rendimentos variando de $40\text{--}54\text{ mg L}^{-1}$ de cultura, concentrada a 3 mg mL^{-1} , e armazenadas no freezer $-80\text{ }^{\circ}\text{C}$, e a partir de alíquotas de proteína congelada foram realizados os testes de atividades bioquímicas. Os ensaios biofísicos e de cristalização foram realizados com proteína fresca.

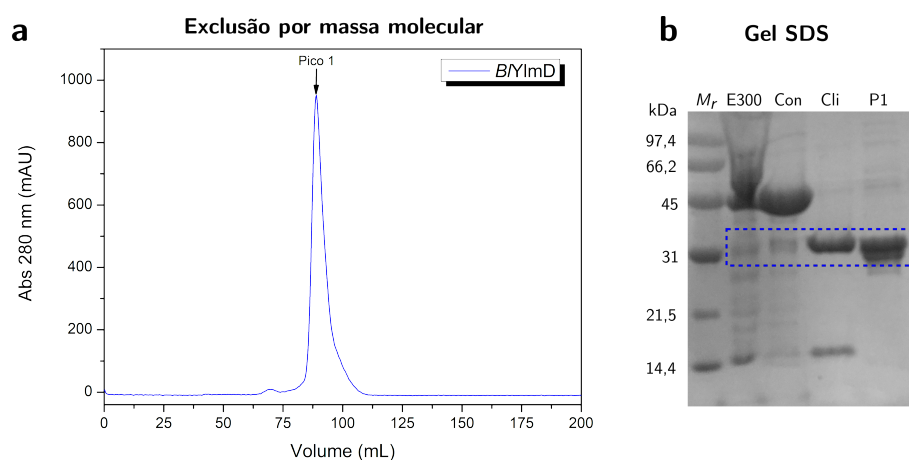


Figura 25 – **Purificação da *BIYImD***. **a**. Cromatograma de purificação em coluna de exclusão por massa molecular (Superdex 200 16/60; $V = 120\text{ mL}$). **b**. Gel SDS-page 10% resumindo as amostras das etapas da purificação, onde: M_r - marcador; E300 - eluição (300 mM imidazol, IMAC - Ni^{2+}); Con - solução da proteína concentrada antes da clivagem, com a cauda (45,5 kDa); Cli - solução da proteína após a clivagem com TEV protease; P1 - amostra do Pico 1 da SEC (**a**), *BIYImD* purificada e sem a cauda (31 kDa).

Fonte: Elaborada pelo autor.

5.2.2 Estabilidade térmica

A estabilidade térmica das proteínas está associada a uma ordem estrutural elevada e menos flexibilidade da proteína, sendo frequentemente acompanhada por homogeneidade conformacional da amostra de proteína ([153,242–244](#)). Aqui, é apresentado um conjunto de medidas para investigar a influência que o pH dos tampões, e a concentração de sal podem ter na estabilidade térmica de proteínas.

Os resultados obtidos delinearam as transições de desnaturação, que permitiram inferir análises quantitativas da desnaturação térmica das proteínas através do cálculo de T_m . A triagem de tampões foi realizada por DSF para identificar a composição do tampão mais promissor; então, em um segundo experimento, a solução tampão de melhor

desempenho foi usada como base para a triagem de aditivos. A amostra de proteína mais estável foi obtida em fosfato de sódio 50 mM seguido de Tris-HCl, ambos pH 7,5, apresentando T_m de 61 e 60 °C, respectivamente em torno de 5 a 7 graus acima da média das condições avaliadas. Em pH extremo, abaixo de 5,0 e acima de 9,5, a proteína se desenovela, resultando em curvas não interpretáveis. Na faixa de pH variando de 5,0 a 9,5, a proteína preservou sua estrutura com pequena variação na T_m , demonstrando alta estabilidade em uma ampla faixa de pH [Figura 26](#).

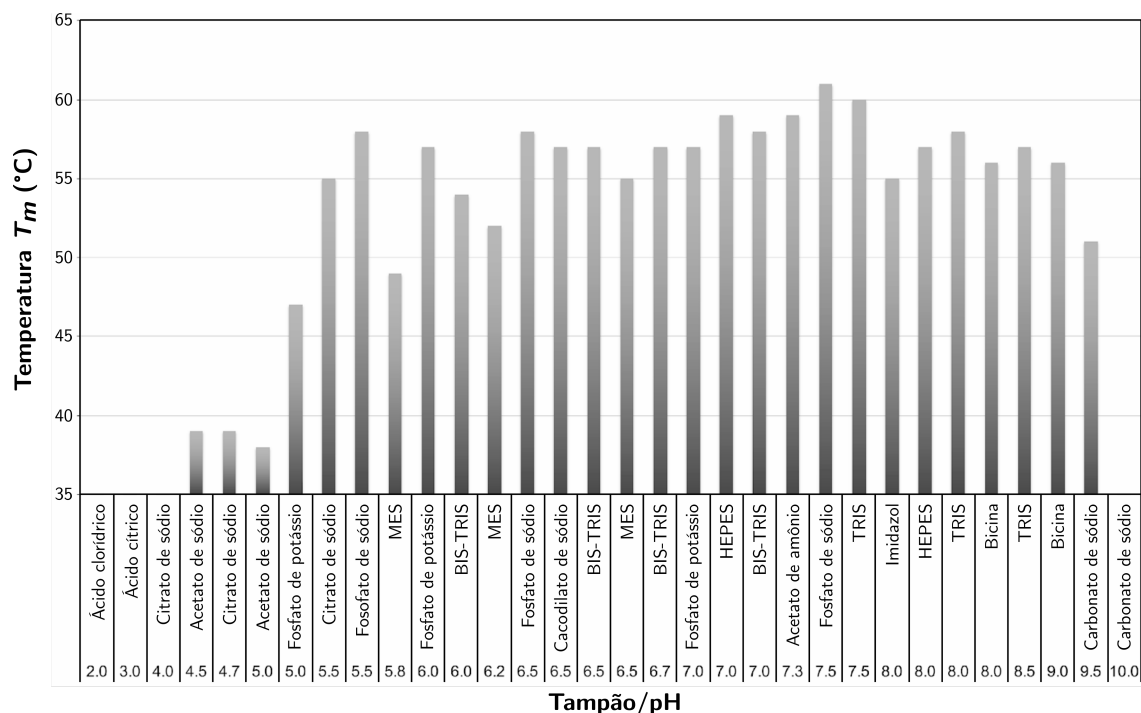


Figura 26 – T_m obtidas por DSF via curvas de desenovelamento proteico da *BIYImD* para variadas condições de tampão e pHs. Os valores de T_m foram determinados usando o modelo de Boltzmann (244). Concentração dos tampões usados nos experimentos foi 50 mM.

Fonte: Elaborada pelo autor.

O painel de aditivos da *Hampton Research* permitiu explorar 96 condições diferentes, incluindo sais de metal, ligantes, carboidratos, etc. A maioria deles não afetou a desnaturação térmica da proteína. Apenas os metais, que causaram uma instabilidade térmica da amostra, reduzindo a T_m em 4 °C. Da mesma forma, o perfil do cobre mostrou uma diminuição na estabilidade da proteína e esta característica também poderia explicar a ausência de atividade como lacase (dados não mostrados), que geralmente é dependente de íons de cobre. Em contrapartida, o íon zinco não promoveu nenhum efeito na estabilidade térmica da proteína, ao contrário do que se poderia esperar, dado que está presente na estrutura cristalina. Esse resultado pode ser explicado pelo fato de a proteína já estar carregada com esses íons metálicos desde o processo de expressão e purificação.

5.2.3 Estrutura

A estrutura cristalina de *BIYImD* foi determinada com resolução de 2,66 Å por substituição molecular (MR) no grupo espacial $P2_12_12_1$ com um homodímero por unidade assimétrica. Ambas as moléculas adotam conformação de *backbone* quase idêntica, com apenas algumas diferenças nas conformações de cadeias laterais. A estrutura da proteína adota o enovelamento do tipo $\alpha/\beta/\beta/\alpha$ de quatro camadas, da classe das proteínas alfa e beta (a + b) sendo principalmente fitas β antiparalelas, como relatado anteriormente (245), e nenhuma densidade eletrônica foi observada para os resíduos 1-3, 22-28 e 278-279 em ambas as cadeias. Os valores finais de R_{work} e R_{free} são 21,4% e 24,8%, respectivamente. Cada monômero de *BIYImD* é composto de 11 α -hélices e 15 fitas- β antiparalelas distribuídas em três camadas de folha- β , Figura 27. Onde as duas maiores camadas de folha- β (camada 1: $\beta 1$, $\beta 2$, $\beta 3$, $\beta 15$, $\beta 7$ e $\beta 4$; camada 2: $\beta 5$, $\beta 6$, $\beta 9$, $\beta 8$, $\beta 10$ e $\beta 14$) formam o sanduiche entre as α -hélices ($\alpha 1$) de um lado e as hélices α e 3_{10} ($\alpha 2$, $\alpha 3$, $\alpha 4$, $\alpha 5$, $\eta 4$ e $\eta 7$) por outro, enquanto a terceira folha- β ($\beta 11$, $\beta 13$ e $\beta 12$) está na interface do dímero também com $\eta 5$ e $\eta 6$. Alinhamentos de sequências de YfiH e elementos de estrutura secundária estão detalhados na Figura 28. A coleta do conjunto de dados e os detalhes do refinamento da estrutura estão resumidos na Tabela 5.

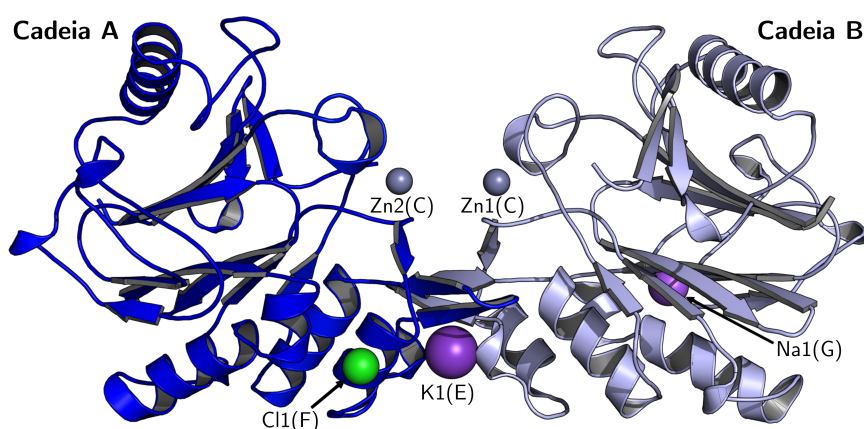


Figura 27 – Representação *cartoon* da unidade assimétrica da estrutura tridimensional da *BIYImD*. Cadeia A é mostrada em azul escuro; Cadeia B em azul claro; Todos os íons são representados por esferas de raio arbitrário: Zn^{2+} cinza; K^+ e Na^+ roxo; Cl^- verde. N-terminal e C-terminal são destacados por caixas amarelas, as estruturas secundárias estão identificadas ao longo da Cadeia A como α -hélices (α), hélices- 3_{10} (η) e fitas- β (β).

Fonte: Elaborada pelo autor.

Em uma busca por proteínas homólogas, os servidores ProFunc e PDBeFold encontraram sete estruturas com Q-score (142) superior a 0,6 e RMSD entre 0,9 a 1,6 Å, Tabela 6. A estrutura de *BIYImD* é semelhante àquelas previamente determinadas para a família putativa de cisteína hidrolases tipo CNF1/YfiH (SCOP) (246), incluindo

Tabela 5 – **Estatísticas da coleta de dados e refinamento da estrutura da *BIYImD*.**
Valores estatísticos para camada de mais alta resolução estão mostrados entre parênteses.

Parâmetros	Valores
PDB ID	6DZD
Coleta de dados	
Grupo espacial	$P2_12_12_1$
Cela unitária	
a, b, c (Å)	42,567, 69,464, 179,589
α, β, γ (°)	90
Detector	<i>APEX II</i> TM CCD
Fonte de raios X	Sealed tube; Bruker D8 quest
Comprimento de onda (Å)	1,5418
Resolução (Å)	54,49 — 2,79 (2,66 — 2,66)
Multiplicidade	5,9 (4,9)
$I/\sigma(I)$ médio	4,72 (1,57)
<i>Wilson</i> B-fator (Å ²)	20,31
R_{merge}	0,281 (0,678)
R_{meas}	0,308 (0,761)
$R_{p.i.m.}$	0,124 (0,341)
CC (1/2)	0,951 (0,705)
Completeza (%)	99,9 (100)
Reflexões totais	95332 (10116)
Reflexões únicas	16046 (2073)
Parâmetros do refinamento	
Reflexões usadas no refinamento	15975 (1556)
Reflexões usadas para R_{free}	824 (83)
R_{work} (%)	21,86 (27,25)
R_{free} (%)	22,73 (29,78)
N.º de átomos não hidrogênio	4357
N.º de átomos da proteína	4179
N.º de átomos dos ligantes	5
N.º de átomos do solvente	173
Resíduos da proteína	536
Ramachandran	
Favorecidos (%)	97,35
Permitidos (%)	2,27
Outliers (%)	0,38
<i>Clashscore</i> (193)	6,10
B-fator médio (Å ²)	24,13
Proteína	24,25
Ligantes	27,37
Solvente	20,92
N.º de grupos TLS	9

Fonte: Elaborada pelo autor.

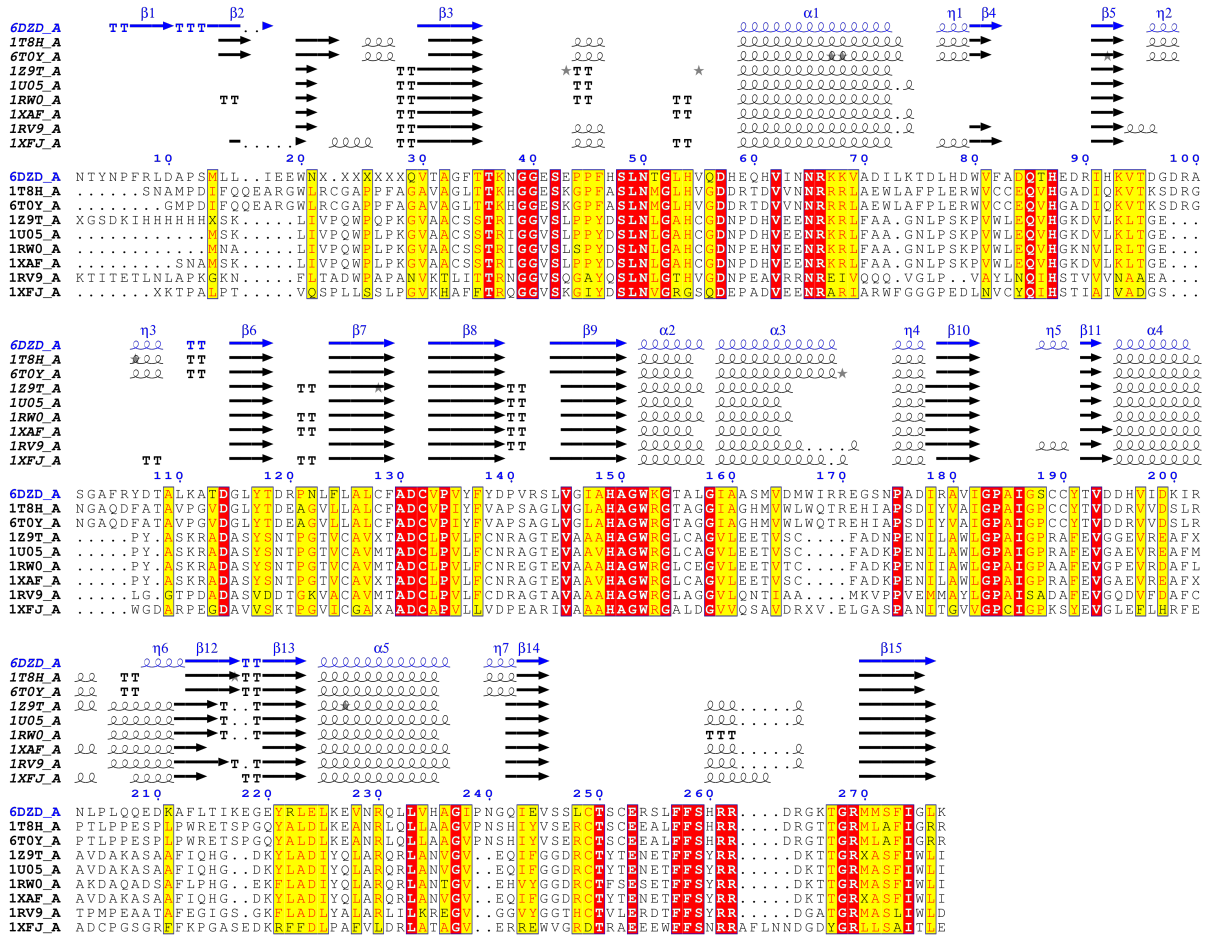


Figura 28 – Alinhamento de sequência baseado em estrutura de 6DZD para os homólogos de proteína YfiH de várias espécies, 1T8H (*Geobacillus stearothermophilus*), 1Z9T (*Escherichia coli*), 1U05 (*Shigella flexneri*), 1RW0 (*Salmonella typhi*), 1XAF (*Shigella flexneri*), 1RV9 (*Neisseria meningitidis*) e 1XFJ (*Caulobacter vibrioides*), com elementos de estrutura secundária correspondentes representados no topo. Os resíduos alternativos são destacados por estrelas cinzas. Resíduos idênticos e semelhantes são marcados em vermelho e amarelo, respectivamente. As hélices α e β_{10} foram marcadas respectivamente com α e η , fitas com β , β -turn com TT e α -turn com as letras TTT. A imagem foi feita usando ENDscript.

Fonte: Elaborada pelo autor.

a hipotética YfiH de *Shigella flexneri* 2a str. 2457T (PDB ID: 1XAF, Q-score 0,66669; id seq 34%) (245), e também as proteínas não caracterizadas de: *Shigella flexneri* (PDB ID: 1U05, Q-score 0,66669; id seq 34%), *Salmonella enterica* subsp. *enterica* sorovar *typhi* (PDB ID: 1RW0, Q 0,6987; id seq 33%), *Neisseria meningitidis* (PDB ID: 1RV9, Q 0,6998; is seq 31%), *Caulobacter vibrioides* (PDB ID: 1XFJ, Q 0,6411; id seq 28%), *Escherichia coli* (PDB ID: 1Z9T, Q 0,6897; id seq 34%), *Geobacillus stearothermophilus* (PDB ID: 1T8H, Q 0,8467; id seq 53%). A cadeia A de *BIYImD* foi sobreposta com

todas as estruturas disponíveis e acima mencionadas, ver na [Figura 32a](#), e elas são estruturalmente conservadas diferindo principalmente em termos de ligantes oriundos das condições da amostra. No servidor BLAST foram identificadas algumas YfiH de organismos diferentes, polifenol oxidoreductase multicobre YlmD de *Bacillus sonorensis* (82%), *Geobacillus thermoleovorans* (55%), *Bacillus stratosphericus* (54%), *Anoxybacillus flavithermus* (51%), *Geobacillus thermoglucosidas* (50%) e *Bacillus megaterium* (48%), todas não caracterizadas funcionalmente.

Tabela 6 – Resultados de correspondência de estrutura secundária usando ProFunc e PDBeFold com *BlYlmD* (6DZD) como consulta.

Q-score	RMSD (Å)	N _{align}	N _{gaps}	%Seq	Consulta	%sse	Alvo	
					N _{res}		N _{res}	PDB
0,8467	0,906	260	5	53,46	269	91	272	1T8H:A
0,6998	1,416	236	12	31,78	269	83	242	1RV9:A
0,6987	1,338	234	10	33,76	269	94	243	1RW0:A
0,6897	1,336	231	9	34,63	269	88	240	1Z9T:A
0,6762	1,295	229	11	34,06	269	88	243	1U05:A
0,6669	1,333	228	9	34,21	269	83	242	1XAF:A
0,6411	1,534	236	11	28,81	269	75	256	1XFJ:A

Parâmetros:

Q-score: Representa a função de qualidade do alinhamento de C α , maximizada pelo algoritmo de alinhamento SSM. Q-score chega a 1 apenas no caso de estruturas idênticas e diminui com o aumento do RMSD ou diminuição do comprimento do alinhamento.

RMSD: Calculado entre os átomos da C α de resíduos correspondentes na melhor sobreposição 3D da estrutura de consulta com os alvos.

N_{align}: comprimento do alinhamento, ou número de resíduos correspondentes.

N_{gaps}: o número de *gaps* é definido como uma seção de sequência consecutiva (contendo 1 ou mais resíduos) da cadeia de consulta que não pode ser correspondida a qualquer parte da cadeia alvo, encontrada entre duas seções de resíduos correspondidos.

%Seq: identidade sequencial é uma qualidade característica do alinhamento de C α .

N_{res}: o tamanho da cadeia de consulta, expresso no número de resíduos.

%sse: a fração de estrutura secundária da cadeia alvo identificada na cadeia de consulta.

Fonte: Elaborada pelo autor.

O modelo final também inclui dois Zn²⁺, um K⁺, um Cl⁻, um íon Na⁺ e 173 moléculas de água. Todos os íons foram avaliados por *CheckMyMetal (CMM): Metal Binding Site Validation Server* (247, 248) para confirmar sua razoabilidade em termos de coordenação. Em cada cadeia, o mesmo modo de coordenação para o íon zinco é observado: 4 átomos de enxofre (S γ) dos resíduos Cys 189, 190, 249 e 252 formando uma configuração tetraédrica clássica, mais detalhes na [Figura 29a](#). Modo de ligação semelhante para íon zinco foi observado para YlmD de *Geobacillus stearothermophilus* (dados não publicados, PDB 1T8H), e contrastando com YfiH de *Shigella flexneri* (245), o íon zinco não foi adicionado à solução de cristalização para *BlYlmD*. Analisando a estrutura cristalográfica,

é possível afirmar que a presença do zinco é importante, mas tal como nas amostras dos testes anteriores, sua adição não causou nenhuma intervenção positiva nos resultados do DSF. No entanto, sua remoção perturbou a estrutura da proteína, evidenciada por uma queda significativa da T_m em 6 °C, causada pela a adição de EDTA.

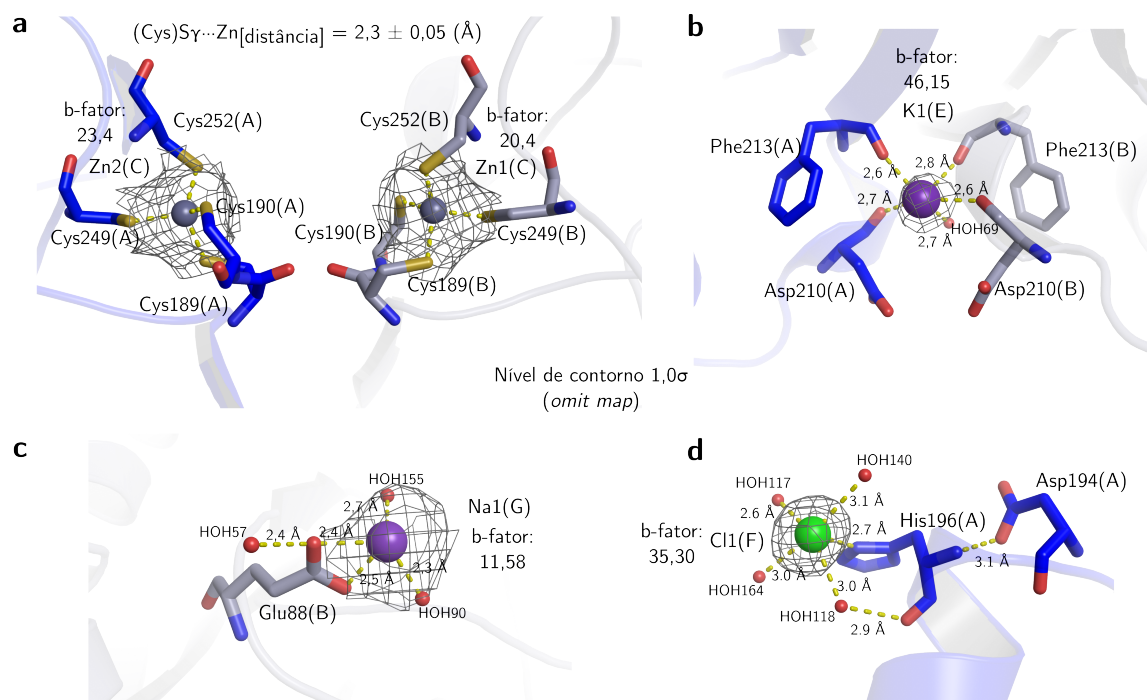


Figura 29 – **Descrição dos íons ligantes da BIYImD.** **a.** Os íons de zinco (Zn^{2+}) coordenados aos resíduos de Cys respectivamente em ambas as cadeias A e B (189, 190, 249 e 252) formando uma configuração tetraédrica. Próximo a interface de contato entre as cadeias. **b.** Íons de potássio (K^+) na interface entre as cadeias A e B, coordenado pela Phe213, Asp210 e H₂O-69 formando uma configuração octaédrica. **c.** Íon de sódio numa configuração bidentada com Glu88, e também as moléculas de H₂O 55 e 90. Íon de cloro (Cl^-) coordenado a His196, H₂O-117, 118, 140 e 164 formando uma configuração octaédrica distorcida. Todos íons foram verificados no *CheckMyMetal* (CMM). Todos os íons estão representados por esferas de raio arbitrário. As interações são representadas por linhas tracejadas amarelas e as distâncias são dadas em angstroms Å. O nível do mapa de contorno é de $1,0\sigma$ (omit map) na cor cinza.

Fonte: Elaborada pelo autor.

O íon potássio (K^+) encontrado na estrutura está localizado na interface de contatos entre as duas cadeias do dímero da proteína parceira, próximo aos resíduos Phe213 (A) e Asp210 (A), Phe213 (B) e Asp210 (B) e também por uma água (H₂O-69) (distâncias de interações $2,6 \pm 0,2$ Å) formando uma configuração octaédrica, conforme descrito na [Figura 29b](#). O íon sódio $Na^+(G)$ encontrado em BIYImD foi possivelmente derivado da condição de cristalização e está localizado em coordenação bidentada com Glu88 e também próximo às moléculas de água H₂O-55 e 90 (distâncias de interação $2,5 \pm 0,2$ Å),

na Figura 29c. Da mesma forma, o íon cloreto (Cl^-) é coordenado com His196 (2,7 Å), moléculas de água H_2O 117 (2,6 Å), 118 (3,0 Å), 140 (3,1 Å) e 164 (3,0 Å) num formato octaédrico distorcido (Figura 29d).

5.2.4 Estado oligomérico

O estado oligomérico das proteínas desta família varia entre monômero e homodímero. A unidade assimétrica na célula unitária para as proteínas *BIYlmD* (6DZD), 1U05, 1RW0 e 1XAF, foram estabelecidas como conjuntos diméricos Figura 30. Usando o PISA, os valores de energia foram obtidos para comparar a estabilidade de cada uma das interfaces (Tabela 7). Neste contexto, o valor ΔG_{int} é calculado como a diferença nas energias de solvatação totais de estruturas isoladas e montadas, não incluindo o efeito das ligações de hidrogênio e pontes de sal nas interfaces. A energia livre de dissociação (ΔG_{diss}) corresponde à diferença de energia livre entre o estado dissociado e o associado (142, 211).

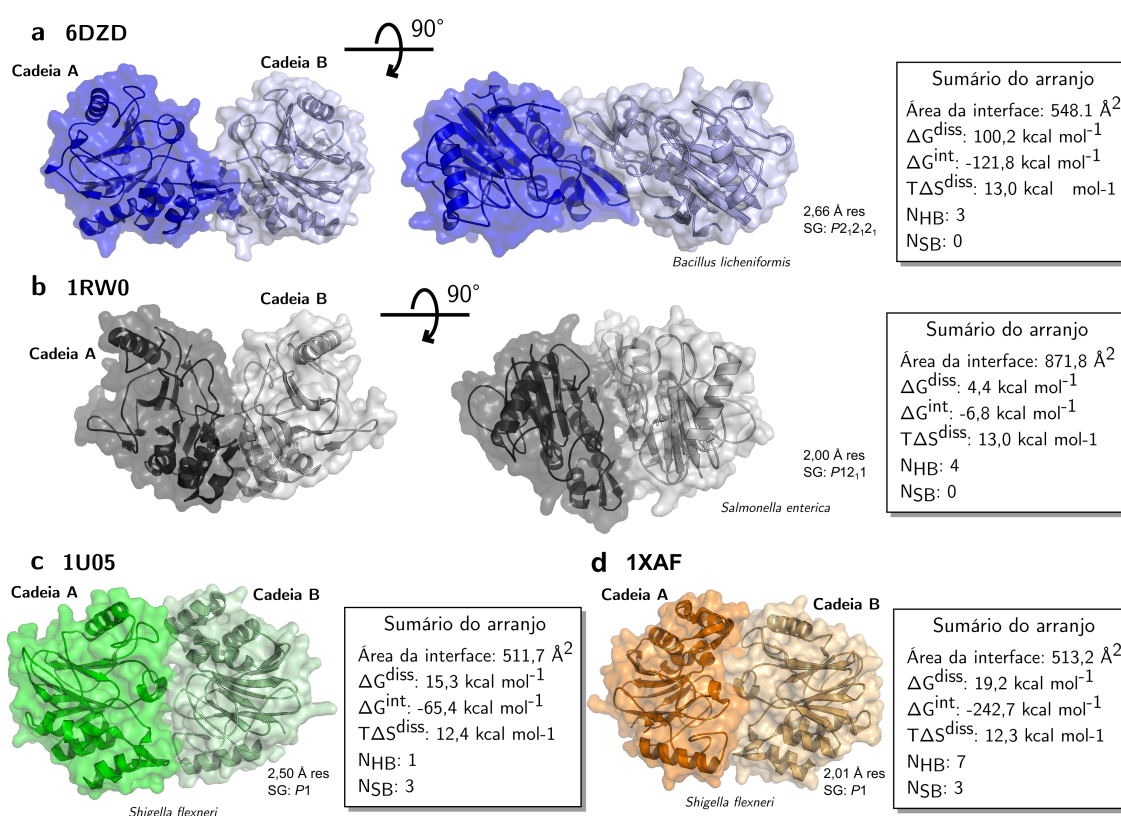


Figura 30 – Perfis SEC-MALS mostrando monômero em solução. **a.** Perfil de oligomerização em ampla variação das concentrações de *BIYlmD* (0,5 a 10 mg mL⁻¹). **b.** Efeito dos tampões com variação nas concentrações de NaCl (35 a 800 mM), alterando a força iônica nos estados de oligomerização de *BIYlmD*. Traços em preto representam a massa molecular calculada relativa ao pico.

Fonte: Elaborada pelo autor.

Tabela 7 – Parâmetros das interfaces analisadas de estruturas YfIH diméricas.

PDB, Organismo	Área da superfície (Å ²)	Área enterrada (Å ²)	ΔG_{int} (kcal mol ⁻¹)	ΔG_{diss} (kcal mol ⁻¹)	$T\Delta S_{diss}$ (kcal mol ⁻¹)
6DZD <i>Bacillus licheniformis</i>	20970,1	1632,9	-121,8	100,2	13,0
1U05 <i>Shigella flexneri</i>	20540,6	1220,9	-65,4	-15,3	12,4
1XAF <i>Shigella flexneri</i>	20183,1	3635,8	-242,7	19,2	12,3
1RW0 <i>Salmonella enterica</i>	19994,6	1743,6	-6,8	-4,4	13,0

Fonte: Elaborada pelo autor.

O arranjo dimérico de 6DZD (Figura 30a) se dissocia em dois monômeros mais os íons em $\Delta G_{diss} \approx 100$ kcal mol⁻¹. Esta alta barreira de dissociação pode ser principalmente devido à presença de íons Zn²⁺ perto da interface dos contatos, como mostrado na Figura 29a, pode se correlacionar razoavelmente com o aumento da energia livre de montagem (249), o que sugere o dímero como o conjunto mais estável. Os valores positivos de ΔG_{diss} indicam que uma força motriz externa é necessária para dissociar o conjunto.

A estrutura de 1XAF montada como um dímero na ASU (Figura 30d) e se dissocia em dois monômeros, seus respectivos ligantes e íons com $\Delta G_{diss} \approx 19$ kcal mol⁻¹. A presença de mais íons zinco não aumenta significativamente a barreira de energia, provavelmente devido ao modo de ligação e à posição desses íons zinco na estrutura. Por outro lado, a montagem da estrutura 1RW0 (Figura 30b) sugere estar próxima de uma associação instável, considerando o valor de $\Delta G_{diss} \approx -4$ kcal mol⁻¹. Como o erro de uma classificação incorreta do PISA é estimado em cerca de ± 5 kcal mol⁻¹, pode ser um monômero ou dímero em solução, uma vez que a diferença em ΔG_{diss} é muito pequena. Enquanto no caso de 1U05 (Figura 30c), a energia livre de dissociação $\Delta G_{diss} \approx -15$ kcal mol⁻¹ também sugere um arranjo instável, o que significa que também pode sofrer dissociação espontânea. Este valor também está muito próximo do erro de uma classificação incorreta.

Dada a incerteza na determinação do estado oligomérico usando apenas a estrutura cristalográfica, foram medidos os estados oligoméricos em solução. Os perfis SEC-MALS mostram um único pico cuja massa molecular de 30 kDa que é consistente com um monômero em solução. Este estado monomérico não é afetado por mudanças na concentração de proteína variando de 0,5 a 10 mg mL⁻¹ (Figura 31a) nem influenciado por mudanças na força iônica variando a concentração de NaCl de 35 mM a 800 mM (Figura 31b).

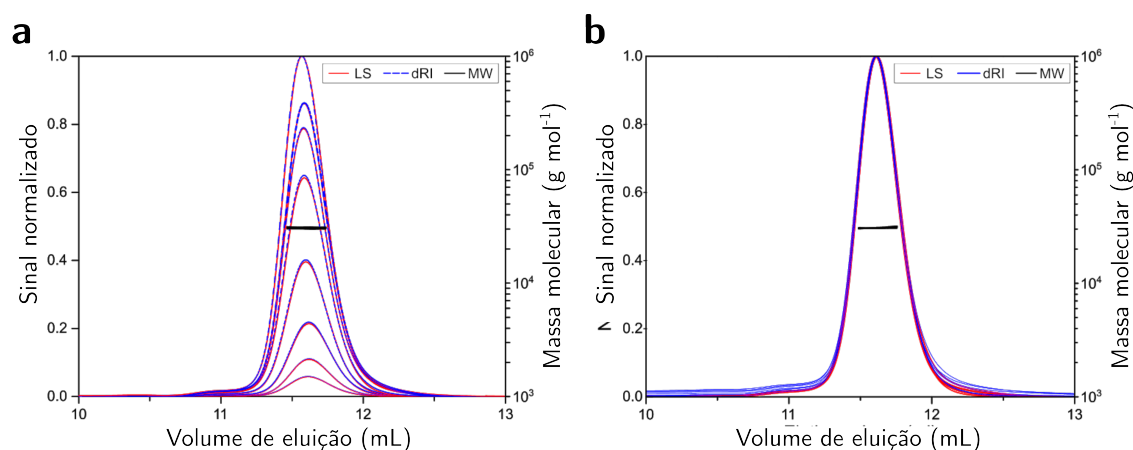


Figura 31 – Perfis SEC-MALS mostrando monômero em solução. **a.** Perfil de oligomerização em ampla variação das concentrações de *BLYmD* (0,5 a 10 mg mL⁻¹). **b.** Efeito dos tampões com variação nas concentrações de NaCl (35 a 800 mM), alterando a força iônica nos estados de oligomerização de *BLYmD*. Traços em preto representam a massa molecular calculada relativa ao pico.

Fonte: Elaborada pelo autor.

5.2.5 Atividade e sítio de ligação proposto

A sequência *BLYmD* contém a díade catalítica hipotética conservada Cys132 e His149, e esses resíduos estão localizados dentro da região funcional esperada da proteína. Para verificar se *BLYmD* poderia agir como uma lacase típica, experimentos usando siringaldazina e ABTS [2,2'-azino-bis (ácido 3-etilbenzotiazolina-6-sulfônico)] como substrato (250) foram realizados. No entanto, nem a siringaldazina nem o ABTS foram oxidados pela *BLYmD* na presença ou ausência de cobre, com a adição de substrato foi administrada de 0,5 a 10 mM de CuSO₄ (dados não mostrados). Tal como também não foi constatado significativa atividade com as proteínas *YlmD*, *YfiH* e *FAMIN* (129). Esta atividade era esperada conforme atribuída para outras proteínas DUF152 bacterianas por métodos espectroscópicos (202). Além disso, *BLYmD* não mostrou atividade na presença do substrato do bagaço de cana-de-açúcar contra um coquetel de enzimas comercial *Accelerase* (251).

Quando estruturas tridimensionais semelhantes da superfamília putativa de hidrolases de cisteína CNF1/*YfiH* são comparadas, Figura 32a, é importante notar que em torno da díade catalítica putativa Cys132 e His149 há uma fenda aberta que pode levar a um local de ligação para moléculas pequenas, Figura 32b. Dessa forma, selecionando todos os resíduos em torno de Cys132 e His149 como um ponto de partida, usando o programa CAVER 3.0 (144), dentro de um raio de 3 Å, uma visão abrangente deste sítio ativo proposto poderia ser obtida, Figura 32d. Um total de cinco caminhos (I a V) podem ser observados, como possíveis caminhos de entrada para acesso ao substrato. A sugestão é que a proposta fenda catalítica esteja localizada ao redor desses resíduos.

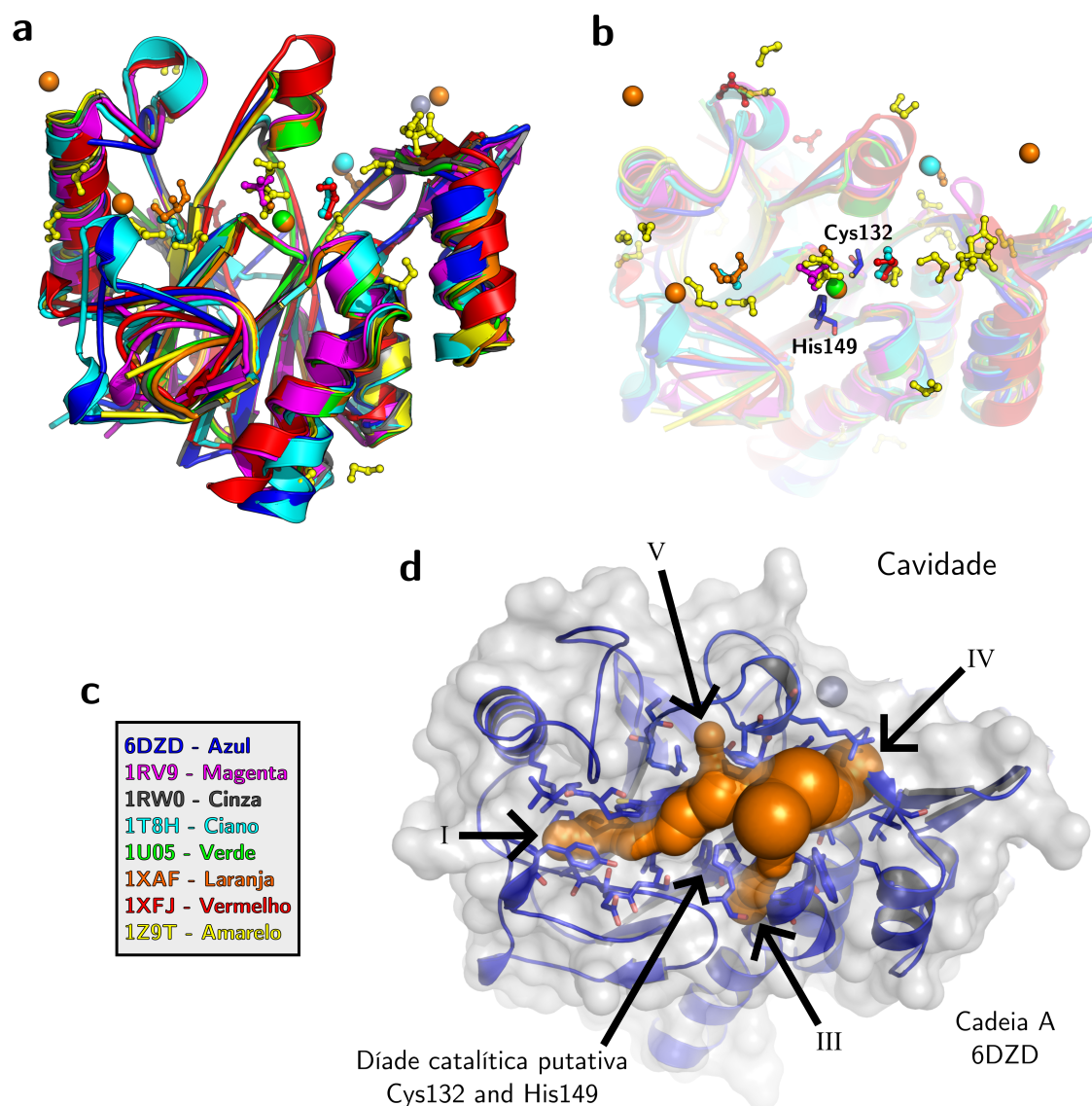


Figura 32 – Comparação com estruturas similares e proposição do sítio ativo. **a.** Sobreposição de estruturas similares da superfamília cisteína-hidrolases putativas CNF1/YfiH. **b.** Detalhamento dos ligantes em torno da cavidade, e a hipotética díade catalítica Cys132 e His149, que potencialmente permitem o acesso a caminhos de entradas para moléculas pequenas (ligantes ou substratos). **c.** Esquema de cores para representação das estruturas previamente reportadas, com os respectivos códigos do PDB. **d.** Possível caminho de acesso (mostrado como esferas laranjas) para o arranjo da *BIYImD*, calculados por meio do plugin do Pymol CAVER 3.0 (144), com raio de 3 Å. As estruturas de proteínas estão representadas na forma de cartoon, íons como esferas de raio arbitrário e as moléculas pequenas ligantes como bola e bastões, ambos coloridos conforme esquema de cores em (c).

Fonte: Elaborada pelo autor.

Parveen e Reddy(233), identificaram vários resíduos conservados que poderiam ser essenciais para a função hipotética de ortólogos YfiH como PgeF (fator de edição de peptidoglicano) de várias espécies bacterianas e de vertebrados, os equivalentes na estrutura aqui descrita 6DZD são representados entre parênteses, entre eles Asp-89 (Asp114), Asp-106 (Asp131), Cys-106 (Cys132), His-124 (His149) e finalmente o Asp-193 que é substituído por um glutamato Glu224, mostra-se conservado em *Bacillus subtilis*, *Danio rerio*, *Homo sapiens*, *E. coli*, *M. tuberculosis* e *M. leprae*. Os autores realizaram mutações para alanina, exceto Asp193, e verificaram a capacidade de complementar a atividade fisiológica in vivo. Observou-se que nenhum desses mutantes foi capaz de complementar um defeito de crescimento sugerindo a importância para a função fisiológica da YfiH, e também implicando que YfiH poderia ser uma enzima cataliticamente ativa, apesar de seu substrato fisiológico até então não identificado (233).

Para identificar o sítio de ligação, Cader *et al.*(129), realizaram o *soaking* dos cristais de YlmD com inosina. Estes difrataram até 1.2 Å de resolução, onde observaram a densidade eletrônica adicional (na 6T1B) não presente na estrutura nativa de YlmD (6T0Y), permitindo a identificação inequívoca do anel de purina da molécula de inosina ligada, conforme mostrado na Figura 33. A porção hipoxantina forma ligações de hidrogênio com a cadeia lateral da Arg59. Enquanto a porção da ribose é coordenada pelas cadeias laterais da Arg262, Cys121, His80, e His142, além de um íon de Zn^{2+} . Sendo a tríade Cys, His e His conservada nas YfiH e previamente preditas tanto por Kim *et al.*(245), como também explorada neste estudo, conforme apresentado anteriormente e indicando a díade Cys132 e His149 da *BIYlmD* (6DZD) Figura 33b.

Ensaio de reações para purina nucleosídeo fosforilases foram realizados exaustivamente, variando: concentrações dos substratos indicados na Figura 24; tampões com fosfato (10–100 mM) e sem usando o Tris (10–100 mM) com diferentes pHs (4–10); concentração de NaCl de 0–500 mM; adição de $CuSO_4$, $ZnSO_4$, e EDTA. Foram realizados experimentos de termoforese em microescala, para verificar se ocorria alguma interação entre a proteína e o substratos, também com resultados inconclusivos e divergentes entre si. Por fim, a última tentativa experimental foi a realização de espectrometrias de massas das amostras das misturas reacionais, e nenhum dos produtos de reação foi detectado, apenas picos referentes aos substratos (dados não mostrados).

Em contraste com a estrutura complexa com inosina (6T1B) (129), no cristal da estrutura apoYlmD, a cadeia lateral da His47 foi observada inserida na cavidade de ligação a base purina, consistente com as demais estruturas apo, previamente depositadas incluindo as deste estudo *BIYlmD* (6DZD). Assim, a cadeia lateral dessa His47 (His54 equivalente em 6DZD) pode atuar como uma controladora de entrada no sítio ativo que rotaciona para fora, causando uma abertura que permite o acesso e ligação do substrato. Estas estruturas possuem nesta região resíduos aromáticos, no *loop* correspondente, que

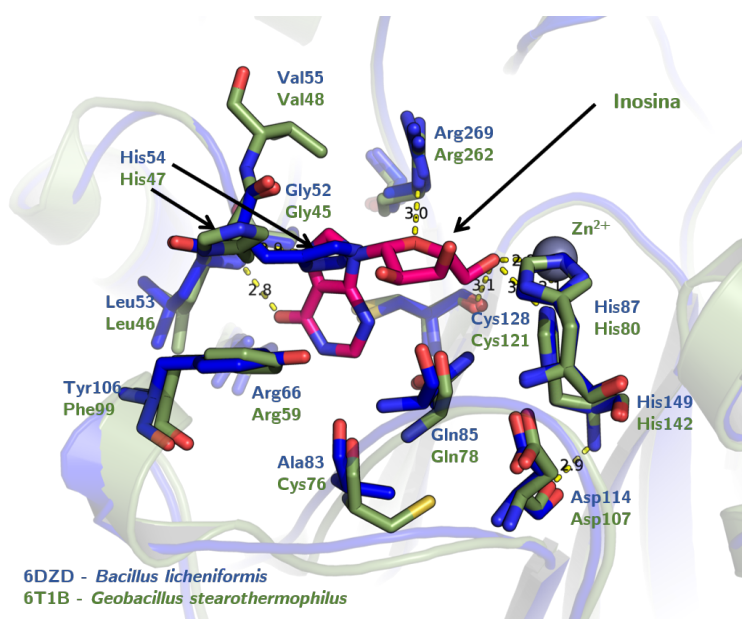


Figura 33 – Sobreposição das estruturas de YlmD evidenciando a região do sítio de ligação observado na estrutura 6T1B (verde), e a conservação do sítio e cadeias laterais equivalentes em 6DZD (azul), o ligante inosina (rosa) mostrando o perfeito encaixe com a mudança na posição da His47 (6T1B), de forma a estar conectada pelos resíduos Arg59 e Arg262, Cys121, His80 e His142.

Fonte: Elaborada pelo autor.

podem ter uma função análoga a His47, flexível para o acesso dos substratos. A posição da His54 em 6DZD, e uma combinação de microarranjos para os demais resíduos, ou condição fisiológicas, talvez sejam algumas das explicações para a não observação de atividade da *BlylmD*.

5.2.6 Rede PPI

A conservação de proteínas semelhantes entre diferentes organismos pode indicar importância fisiológica, e pode ser correlacionada com uma via de cooperatividade proteína-proteína para desempenhar sua função adequada. Analisando as associações diretas e indiretas das interações proteína-proteína usando o banco de dados STRING v11 (252), Figura 34, uma das associações sugere que *BlylmD* está relacionado às proteínas *ftsZ* e *sepF* (253–257), necessários para a divisão celular, pela distância intergênica no genoma de *Bacillus subtilis*. Tal vizinhança genômica pode indicar que contribuem em conjunto para a mesma função biológica em *B. licheniformis*, uma vez que *ftsZ* e *sepF* são proteínas essenciais no processo de divisão celular, formando a estrutura "Anel Z" no futuro sítio de divisão celular bacteriana. Por outro lado, YlmD pode estar relacionado ao processo de proliferação celular.

Após várias buscas de homologia de domínio e algumas poucas respostas, que

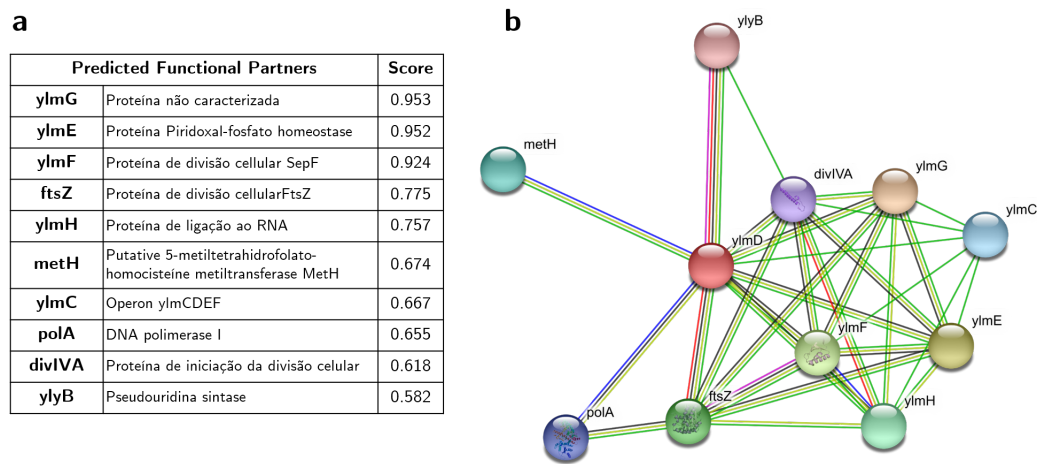


Figura 34 – **A análise com STRING v11 de *BIYlmD* mostra a rede entre as proteínas identificadas.** **a.** Os parceiros funcionais previstos e a pontuação para expressar confiança aproximada, em uma escala de zero a um. **b.** Rede de interação entre os genes visados. O tipo de evidência é ilustrado pela cor da borda: azul – co-ocorrência de genes, magenta – determinado experimentalmente, verde – vizinhança de genes, vermelho – fusões de genes, amarelo – mineração de texto e preto – coexpressão. A análise foi realizada com configuração de confiança média (0,400). Nós coloridos: proteínas de consulta e primeira camada de interatores. Nós vazios: proteínas de estrutura 3D desconhecida. Nós preenchidos: alguma estrutura 3D é conhecida ou prevista.

Fonte: Elaborada pelo autor.

direcionaram possíveis atividades, é difícil atribuir uma função para *BIYlmD*, dados os resultados negativos tanto como lacase ou purina nucleosídeo fosforilase. Portanto, apontar uma função metabólica ou processo biológico mais preciso em relação à YfiH, principalmente devido aos dados funcionais limitados dessa família de proteínas (245), pode ser uma tarefa difícil. Nenhuma atividade de YlmD foi detectada nesses substratos apesar do uso de várias preparações enzimáticas, condições de ensaios alteradas, como tampões, temperatura, pH ou mesmo após a adição de quantidades graduais de cofator metálico externo como cobre, fosfato, zinco e EDTA. A incapacidade de detectar qualquer atividade hidrolítica pode ser atribuída a várias razões, como a falta de eventuais cofatores necessários ou proteínas colaborativas no sistema de ensaio, necessidade de substrato desconhecido ou a falta de atividade hidrolítica. Comparada à PNP FAMIN multifuncional, o principal indício estrutural da falta de atividade permanece centrado na conformação da His47, associadas as variações nos resíduos Tyr106 (Phe), que podem afetar o volume da cavidade catalítica impedindo a entrada dos substratos caso seja por esta via.

Com a função desta proteína ainda não atribuída, ela se junta ao repertório de uma família PANTHER (PTHR30616) (258) com 52 genes que possuem uma estrutura tridimensional resolvida, mas sem atividade anotada. Nossos achados propõem mitigar

pelo menos a busca pelo sítio de ligação ativo mais provável, localizado na cavidade ao redor da hipotética díade catalítica de resíduos de Cys e His, com base em nossas comparações estruturais, os resultados negativos dos experimentos, a ligação de íons metálicos, similaridade no enovelamento com outras proteínas de funções conhecidas e as diferenças estruturais no possível sítio catalítico. Esses achados podem ajudar a esclarecer um pouco mais sobre essa difícil tarefa de atribuir uma função a estas proteínas.

6 CONSIDERAÇÕES FINAIS E PERSPECTIVAS

As crescentes demandas energéticas, demandas de consumos, crescimento populacional, desenvolvimento de processos industriais mais eficazes, mais sustentáveis e com responsabilidade ambiental, entre outros, sinalizam para a necessidade de reflexões quanto ao formato das fontes energéticas disponíveis. Atrelado à complexidade destas questões, algumas das melhores alternativas devem surgir, principalmente, graças ao desenvolvimento tecnológico.

Numa busca por desenvolver conhecimento sobre novas enzimas, que sejam, resistentes às altas temperaturas, a condições adversas como presença de inibidores, se mantenham ativas por longos períodos, com melhores taxas de recuperação, e versáteis para as aplicabilidades em processos industriais. Assim, se estabelecem os objetivos deste trabalho, e seus derivados, realizando os estudos estruturais e funcionais das hidrolases de glicosídeos *CtBgl1A*, *CtBgl3B*, *LgAs32*, *TtAbf62* e da proteína hipotética *BIYImD*.

Foi observado que a *CtBgl1A* não foi obtida purificada, apta para as aplicabilidades e ensaios aqui pretendidos, mesmo que diante de várias tentativas frustradas. Os problemas talvez estejam relacionados com questões básicas referentes a estratégia de clonagem ou metodologia de purificação. Considerando que, em dado momento os trabalhos para esta enzima deixaram de ser prioridade, fica em aberto a perspectiva futura de um melhor planejamento e mudanças nas abordagens para obtenção desta proteína.

Por outro lado, a Bgl *CtBgl3B* se mostrou bastante promissora para aplicabilidade em processos de conversão da biomassa lignocelulósica em açúcar, especialmente devido a sua capacidade em clivar as ligações do tipo β -(1,4') de substratos glucopiranosídicos como o *pNPG*, *CpNPG2*, e o xilopiranosídico *pNPX*. De forma que na presença dos surfactantes não iônicos, Triton X-100 e Tween20, ainda demonstrou aumento nas taxas de conversão, e aumento na atividade específica frente aos substratos. Foi observado que a *CtBgl3B* resistiu por até 72h incubada nas temperaturas de 50, 40, 35 e 30 °C, mantendo sua atividade catalítica residual acima de 80% em relação ao estado inicial, com poucas variações nessas taxas e nunca abaixo dos 80%. Considerando um processo industrial isso é mais uma grande vantagem.

A estrutura da *CtBgl3B* revelou uma possível explicação para sua alta estabilidade por seu arranjo dimérico e disposição da interface de contato, onde, um domínio de função desconhecida DUF (resíduos 658–755) da molécula parceira na arquitetura interage com flexibilidade sobre a região da cavidade catalítica, associado a mobilidade dos *loops* com resíduos dispostos no sítio catalítico, e também os resíduos aromáticos da região catalítica. A flexibilidade do DUF, talvez seja maior quando as reações têm a presença dos

surfactantes Triton X-100 e Tween20 facilitando a entrada dos substratos. O que explicaria o aumento da atividade catalítica nestas condições.

Nesta perspectiva de avaliar os efeitos destes aditivos, nas reações catalisadas pela *CtBgl3B*, faltará agora avaliar a estabilidade térmica da proteína nestas condições, para mensurar por quanto tempo isso se mantém e se seria viável. Além de investigar, com técnicas apropriadas, a influência destes surfactantes diretamente na estrutura. O que poderia confirmar ou ser um contraponto para a ideia de mobilidade do domínio, supostamente, regulador da entrada dos substratos.

Perspectiva futura: utilizar os dados dos ensaios enzimáticos da *CtBgl3B* variando pH e temperatura para estabelecer correlações com parâmetros cinéticos experimentais com dados teóricos, fornecendo valores cinéticos de constantes de reações (k_{cat}), e também valores termodinâmicos como energia livre de Gibbs (ΔG), [Equação B.7](#), e variação da energia livre de Gibbs ($\Delta\Delta G^\ddagger$, [Equação B.8](#), pelas relações entre estes parâmetros. Estas abordagens serão feitas com ajuda do Prof. Dr. Valter H. Carvalho, da Universidade Estadual de Goiás (UEG), um colaborador de longa data na minha trajetória que se dedica principalmente aos estudos de cinética química e fenômenos de estados de transição.

Considerações finais *BlYlmD*: após várias buscas de homologia de domínio e algumas poucas respostas, que direcionaram possíveis atividades, é difícil atribuir uma função para *BlYlmD*, dados os resultados negativos tanto como lacase ou purina nucleosídeo fosforilase. Portanto, apontar uma função metabólica ou processo biológico mais preciso em relação à *YfiH*, principalmente devido aos dados funcionais limitados dessa família de proteínas ([245](#)), pode ser uma tarefa difícil. Nenhuma atividade de *YlmD* foi detectada nesses substratos apesar do uso de várias preparações enzimáticas, condições de ensaios alteradas, como tampões, temperatura, pH ou mesmo após a adição de quantidades graduais de cofator metálico externo como cobre, fosfato, zinco e EDTA. A incapacidade de detectar qualquer atividade hidrolítica pode ser atribuída a várias razões, como a falta de eventuais cofatores necessários ou proteínas colaborativas no sistema de ensaio, necessidade de substrato desconhecido ou a falta de atividade hidrolítica. Comparada à PNP FAMILIN multifuncional, o principal indício estrutural da falta de atividade permanece centrado na conformação da His47, associadas as variações nos resíduos Tyr106 (Phe), que podem afetar o volume da cavidade catalítica impedindo a entrada dos substratos caso seja por esta via.

Com a função desta proteína ainda não atribuída, ela se junta ao repertório de uma família PANTHER (PTHR30616) ([258](#)) com 52 genes que possuem uma estrutura tridimensional resolvida, mas sem atividade anotada. Nossas comparações estruturais propõem mitigar pelo menos a busca pelo sítio de ligação, ou sítio ativo mais provável, localizado na cavidade ao redor da hipotética díade catalítica de resíduos de Cys e His, com base em nossas comparações estruturais, dada a similaridade no enovelamento com outras

proteínas de funções conhecidas e as diferenças estruturais no possível sítio catalítico. Contudo, diante dos resultados negativos nos experimentos de atividade para purina nucleosídeo fosforilases e lacases, principalmente, são fatos que retificam um pouco mais sobre a difícil tarefa de atribuir uma função as proteínas desta família. Tudo isso deixa margem para novos estudos sobre a evolução destas proteínas, investigando a perda de função, ou mesmo trabalhos feitos com os organismos nativos.

Em um recente trabalho publicado por [Meng-Sheng *et al.*\(259\)](#), os autores demonstraram que YfiH é uma hidrolase com a atividade anteriormente desconhecida específica para o mono-peptídeo UDP-MurNAc (uridina difosfato ácido N-acetilmurâmico), um dos precursores de nucleotídeos das etapas citoplasmáticas da via de biossíntese de peptidoglicanos. De tal forma, que a YfiH hidrolisa seletivamente UDP-MurNAc-Ser, um subproduto incorreto da reação de biossíntese, para garantir que apenas o precursor de peptidoglicanos correto, UDP-MurNAc-Ala, seja incorporado. Portanto, o trabalho de [Meng-Sheng *et al.*\(259\)](#) traz luz sobre reações da biossíntese de peptidoglicanos, podendo ser um novo ponto de partida para futuros estudos com a proteína *BYImD*.

REFERÊNCIAS

- 1 STICKLEN, M. Plant genetic engineering to improve biomass characteristics for biofuels. **Current Opinion in Biotechnology**, v. 17, n. 3, p. 315–319, 2006. ISSN 0958-1669. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0958166906000620>. Acesso em: 07 mar. 2022.
- 2 CHEN, F.; DIXON, R. A. Lignin modification improves fermentable sugar yields for biofuel production. **Nature Biotechnology**, v. 25, n. 7, p. 759–761, 2007. ISSN 1546-1696. Disponível em: <https://doi.org/10.1038/nbt1316>. Acesso em: 07 mar. 2022.
- 3 WANG, L. **Sustainable bioenergy production**. Boca Raton: CRC Press, 2014. ISBN 9781466505568.
- 4 SERRANO-RUIZ, J. C. **Advanced biofuels: using catalytic routes for the conversion of biomass platform molecules**. Oakville: Apple Academic Press, 2015. 330 p. ISBN 9781771882217.
- 5 LIU, J. *et al.* Systems integration for global sustainability. **Science**, v. 347, n. 6225, p. 963, 2 2015. Disponível em: <http://science.sciencemag.org/content/347/6225/1258832.abstract>. Acesso em: 21 mar. 2016.
- 6 RAMOS, L. P. **Aproveitamento integral de resíduos agrícolas e agroindustriais**. Centro de Pesquisa em Química Aplicada, Departamento de Química, Universidade Federal do Paraná, 2001. 11 p. Disponível em: http://web-resol.org/textos/artigo_pretratamento.pdf. Acesso em: 29 mar. 2016.
- 7 PEREIRA-JUNIOR, N. **Biomassas residuais de composição lignocelulósica para a produção de etanol e o contexto de refinaria**. Renewable Fuels Association, 2007. Disponível em: <http://www.ethanolrfa.org/pages/statistics>. Acesso em: 29 mar. 2016.
- 8 HOANG, N. V. *et al.* High-throughput profiling of the fiber and sugar composition of sugarcane biomass. **BioEnergy Research**, v. 10, n. 2, p. 400–416, 2017. ISSN 1939-1242. Disponível em: <https://doi.org/10.1007/s12155-016-9801-8>. Acesso em: 08 mar. 2022.
- 9 BENJAMIN, Y.; GARCÍA-APARICIO, M. P.; GÖRGENS, J. F. Impact of cultivar selection and process optimization on ethanol yield from different varieties of sugarcane. **Biotechnology for Biofuels**, v. 7, n. 1, p. 60, 2014. ISSN 1754-6834. Disponível em: <https://doi.org/10.1186/1754-6834-7-60>. Acesso em: 08 mar. 2022.
- 10 HIMMEL, M. E. *et al.* Biomass recalcitrance: engineering plants and enzymes for biofuels production. **Science**, v. 315, n. 5813, p. 804–807, 2 2007. Disponível em: <http://science.sciencemag.org/content/315/5813/804.abstract>. Acesso em: 21 mar. 2016.
- 11 BASTOS, V. D. Biorrefinarias, biocombustíveis e química renovável: revolução tecnológica e financiamento. **Revista do BNDES**, v. 38, p. 85–138, 2012. Disponível em: http://www.bndes.gov.br/SiteBNDES/export/sites/default/bndes_pt/Galerias/Arquivos/conhecimento/revista/rev3803.pdf. Acesso em: 29 mar. 2016.

12 CHUM, H. *et al.* Bioenergy. *In: EDENHOFER, O. et al.* (ed.). **IPCC special report on renewable energy sources and climate change mitigation**. Cambridge: Cambridge University Press, 2011.

13 SAHA, B. C. Lignocellulose biodegradation and applications in biotechnology. *In: Lignocellulose Biodegradation*. American Chemical Society, 2004, (ACS Symposium Series, v. 889). p. 2–34. ISBN 0-8412-3876-6. Disponível em: <http://dx.doi.org/10.1021/bk-2004-0889.ch001>. Acesso em: 29 mar. 2016.

14 SOUZA, A. P. de *et al.* Composition and structure of sugarcane cell wall polysaccharides: implications for second-generation bioethanol production. **BioEnergy Research**, v. 6, n. 2, p. 564–579, 6 2013. ISSN 1939-1242. Disponível em: <https://doi.org/10.1007/s12155-012-9268-1>. Acesso em: 31 ago. 2019.

15 MASARIN, F. *et al.* Chemical composition and enzymatic digestibility of sugarcane clones selected for varied lignin content. **Biotechnology for Biofuels**, v. 4, n. 1, p. 55, 2011. ISSN 1754-6834. Disponível em: <https://doi.org/10.1186/1754-6834-4-55>. Acesso em: 07 mar. 2022.

16 SUN, J. X. *et al.* Isolation and characterization of cellulose from sugarcane bagasse. **Polymer Degradation and Stability**, v. 84, n. 2, p. 331–339, 5 2004. ISSN 0141-3910. Disponível em: <http://www.sciencedirect.com/science/article/pii/S014139100400045X>. Acesso em: 29 mar. 2016.

17 BORNSCHEUER, U.; BUCHHOLZ, K.; SEIBEL, J. Enzymatic degradation of (ligno)cellulose. **Angewandte Chemie**, v. 53, n. 41, p. 10876–93, 2014. ISSN 1521-3773. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/25136976>. Acesso em: 01 dez. 2015.

18 BÉGUIN, P.; AUBERT, J.-P. The biological degradation of cellulose. **FEMS Microbiology Reviews**, Blackwell Publishing Ltd, v. 13, n. 1, p. 25–58, 1 1994. ISSN 1574-6976. Disponível em: <http://dx.doi.org/10.1111/j.1574-6976.1994.tb00033.x>. Acesso em: 02 dez. 2015.

19 LOSORDO, Z. *et al.* Cost competitive second-generation ethanol production from hemicellulose in a brazilian sugarcane biorefinery. **Biofuels, Bioproducts and Biorefining**, John Wiley & Sons, Ltd, v. 10, n. 5, p. 589–602, 2016. ISSN 1932-104X. Disponível em: <https://doi.org/10.1002/bbb.1663>. Acesso em: 07 mar. 2022.

20 CHANDEL, A. K. *et al.* Comparative analysis of key technologies for cellulosic ethanol production from brazilian sugarcane bagasse at a commercial scale. **Biofuels, Bioproducts and Biorefining**, John Wiley & Sons, Ltd, v. 13, n. 4, p. 994–1014, 2019. ISSN 1932-104X. Disponível em: <https://doi.org/10.1002/bbb.1990>. Acesso em: 07 mar. 2022.

21 SAYERS, E. W. *et al.* GenBank. **Nucleic Acids Research**, v. 49, n. D1, p. D92–D96, 2021. ISSN 0305-1048. Disponível em: <https://doi.org/10.1093/nar/gkaa1023>. Acesso em: 08 mar. 2022.

22 DRULA, E. *et al.* The carbohydrate-active enzyme database: functions and literature. **Nucleic Acids Research**, v. 50, n. D1, p. D571–D577, 2022. ISSN 0305-1048. Disponível em: <https://doi.org/10.1093/nar/gkab1045>. Acesso em: 08 mar. 2022.

-
- 23 KHAN, I.; AKHTAR, M. w. The biotechnological perspective of beta-glucosidases. **Nature Precedings**, Nature Publishing Group, 2010. ISSN 1756-0357. Disponível em: <http://precedings.nature.com/documents/4945/version/1>. Acesso em: 01 dez. 2015.
- 24 KUMAR, A.; GAUTAM, A.; DUTT, D. Biotechnological transformation of lignocellulosic biomass in to industrial products: an overview. **Advances in Bioscience and Biotechnology**, Scientific Research Publishing, v. 07, n. 03, p. 149–168, 2016. ISSN 2156-8456. Disponível em: <https://dx.doi.org/10.4236/abb.2016.73014>. Acesso em: 29 mar. 2016.
- 25 GARVEY, M. *et al.* Cellulases for biomass degradation: comparing recombinant cellulase expression platforms. **Trends in Biotechnology**, v. 31, n. 10, p. 581–593, 2013. ISSN 0167-7799. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0167779913001443>.
- 26 PAYNE, C. M. *et al.* Fungal cellulases. **Chemical Reviews**, American Chemical Society, v. 115, n. 3, p. 1308–1448, 2015. ISSN 0009-2665. Disponível em: <https://doi.org/10.1021/cr500351c>. Acesso em: 08 mar. 2022.
- 27 BAYER, E. A. *et al.* From cellulosomes to cellulomics. **The Chemical Record**, Wiley Subscription Services, Inc., A Wiley Company, v. 8, n. 6, p. 364–377, 2008. ISSN 1528-0691. Disponível em: <http://dx.doi.org/10.1002/tcr.20160>. Acesso em: 21 mar. 2016.
- 28 ARTZI, L.; BAYER, E. A.; MORAÏS, S. Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. **Nature Reviews Microbiology**, v. 15, n. 2, p. 83–95, 2017. ISSN 1740-1534. Disponível em: <https://doi.org/10.1038/nrmicro.2016.164>. Acesso em: 30 mar. 2021.
- 29 RAVACHOL, J. *et al.* Combining free and aggregated cellulolytic systems in the cellulosome-producing bacterium *Ruminiclostridium cellulolyticum*. **Biotechnology for Biofuels**, v. 8, n. 1, p. 114, 2015. ISSN 1754-6834. Disponível em: <https://doi.org/10.1186/s13068-015-0301-4>. Acesso em: 01 dez. 2015.
- 30 SUKHARNIKOV, L. O. *et al.* Cellulases: ambiguous nonhomologous enzymes in a genomic perspective. **Trends in Biotechnology**, v. 29, n. 10, p. 473–9, 2011. ISSN 1879-3096. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0167779911000746>. Acesso em: 25 abr. 2016.
- 31 BHATTACHARYA, A. S.; BHATTACHARYA, A.; PLETSCHKE, B. I. Synergism of fungal and bacterial cellulases and hemicellulases: a novel perspective for enhanced bio-ethanol production. **Biotechnology Letters**, v. 37, n. 6, p. 1117–1129, 2015. ISSN 1573-6776. Disponível em: <https://doi.org/10.1007/s10529-015-1779-3>. Acesso em: 08 mar. 2022.
- 32 LYND, L. R. *et al.* Microbial cellulose utilization: fundamentals and biotechnology. **Microbiology and Molecular Biology Reviews**, v. 66, n. 4, p. 506–577, 2002. ISSN 1092-2172. Disponível em: <https://doi.org/10.1128/MMBR.66.3.506-577.2002>. Acesso em: 09 ago. 2019.
- 33 ARANTES, V.; SADDLER, J. N. Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. **Biotechnology for Biofuels**, v. 3, n. 1, p. 1–11, 2010. ISSN 1754-6834. Disponível em: <http://dx.doi.org/10.1186/1754-6834-3-4>. Acesso em: 20 mar. 2016.

- 34 WOODWARD, J. Synergism in cellulase systems. **Bioresource Technology**, v. 36, n. 1, p. 67–75, 1991. ISSN 0960-8524. Disponível em: <http://www.sciencedirect.com/science/article/pii/096085249190100X>. Acesso em: 29 mar. 2016.
- 35 COLUSSI, F. *et al.* Oligomeric state and structural stability of two hyperthermophilic β -glucosidases from *Thermotoga petrophila*. **Amino Acids**, Springer Vienna, v. 47, n. 5, p. 937–948, 2015. ISSN 0939-4451. Disponível em: <http://dx.doi.org/10.1007/s00726-015-1923-3>. Acesso em: 04 dez. 2015.
- 36 RYU, D. D. Y.; MANDELS, M. Cellulases: biosynthesis and applications. **Enzyme and Microbial Technology**, v. 2, n. 2, p. 91–102, 1980. ISSN 0141-0229. Disponível em: <http://www.sciencedirect.com/science/article/pii/0141022980900630>. Acesso em: 29 mar. 2016.
- 37 YENNAMALLI, R. M. *et al.* Endoglucanases: insights into thermostability for biofuel applications. **Biotechnology for Biofuels**, BioMed Central, v. 6, p. 136, 2013. ISSN 1754-6834. Disponível em: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3856469/>.
- 38 SOARES-JÚNIOR, F. L. *et al.* Endo-and exoglucanase activities in bacteria from mangrove sediment. **Brazilian Journal of Microbiology**, SBM, v. 44, n. 3, p. 969–976, 2013. ISSN 1517-8382. Disponível em: <https://doi.org/10.1590/S1517-83822013000300048>. Acesso em: 29 mar. 2016.
- 39 CHAUVE, M. *et al.* Comparative kinetic analysis of two fungal beta-glucosidases. **Biotechnology for Biofuels**, v. 3, n. 1, p. 3, 2010. ISSN 1754-6834 (Electronic). Disponível em: <https://doi.org/10.1186/1754-6834-3-3>. Acesso em: 29 mar. 2016.
- 40 SHAHZADI, T. *et al.* Induced production of exoglucanase, and β -glucosidase from fungal co-culture of *T. viride* and *G. lucidum*. **Advances in Bioscience and Biotechnology**, Scientific Research Publishing, v. 05, n. 05, p. 426–433, 2014. ISSN 2156-8456. Disponível em: <http://www.scirp.org/journal/PaperInformation.aspx?PaperID=44482&#abstract>. Acesso em: 29 mar. 2016.
- 41 LI, Y.-K.; CHIR, J.; CHEN, F.-Y. Catalytic mechanism of a family 3 beta-glucosidase and mutagenesis study on residue Asp-247. **Biochemical Journal**, v. 355, n. Pt 3, p. 835–840, 2001. ISSN 0264-6021. Disponível em: <http://www.biochemj.org/content/355/3/835>. Acesso em: 24 ago. 2018.
- 42 SONG, B. *et al.* Real-time imaging reveals that lytic polysaccharide monoxygenase promotes cellulase activity by increasing cellulose accessibility. **Biotechnology for Biofuels**, v. 11, n. 1, p. 41, 2018. ISSN 1754-6834. Disponível em: <https://doi.org/10.1186/s13068-018-1023-1>. Acesso em: 08 mar. 2022.
- 43 GUSTAV, V.-K. *et al.* An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. **Science**, American Association for the Advancement of Science, v. 330, n. 6001, p. 219–222, 2010. Disponível em: <https://doi.org/10.1126/science.1192231>. Acesso em: 08 mar. 2022.
- 44 TANDRUP, T. *et al.* Recent insights into lytic polysaccharide monoxygenases (LPMOs). **Biochemical Society Transactions**, v. 46, n. 6, p. 1431–1447, 2018. ISSN 0300-5127. Disponível em: <https://doi.org/10.1042/BST20170549>. Acesso em: 08 mar. 2022.

-
- 45 AGGER, J. W. *et al.* Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. **Proceedings of the National Academy of Sciences**, Proceedings of the National Academy of Sciences, v. 111, n. 17, p. 6287–6292, 2014. Disponível em: <https://doi.org/10.1073/pnas.1323629111>. Acesso em: 08 mar. 2022.
- 46 CHOTHIA, C.; LESK, A. M. The relation between the divergence of sequence and structure in proteins. **EMBO Journal**, John Wiley & Sons, Ltd, v. 5, n. 4, p. 823–826, 1986. ISSN 0261-4189. Disponível em: <https://doi.org/10.1002/j.1460-2075.1986.tb04288.x>. Acesso em: 9 mar. 2022.
- 47 HENRISSAT, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochemical Journal**, v. 280, pt 2, p. 309–316, 1991. ISSN 0264-6021 (Print).
- 48 HENRISSAT, B.; DAVIES, G. Structural and sequence-based classification of glycoside hydrolases. **Current Opinion in Structural Biology**, v. 7, n. 5, p. 637–644, 1997. ISSN 0959-440X. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0959440X97800723>. Acesso em: 17 jul. 2017.
- 49 HENRISSAT, B.; DAVIES, G. J. Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. **Plant Physiology**, v. 124, n. 4, p. 1515–1519, 2000. ISSN 0032-0889 (Print).
- 50 MACGREGOR, E. A.; SVENSSON, B. A super-secondary structure predicted to be common to several α -1,4-D-glucan-cleaving enzymes. **Biochemical Journal**, Portland Press Ltd., v. 259, n. 1, p. 145–152, 1989. ISSN 0264-6021.
- 51 RAIMBAUD, E. *et al.* Hydrophobic cluster analysis of the primary sequences of α -amylases. **International Journal of Biological Macromolecules**, Elsevier, v. 11, n. 4, p. 217–225, 1989. ISSN 0141-8130.
- 52 CANTAREL, B. L. *et al.* The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. **Nucleic Acids Research**, Oxford Univ Press, v. 37, n. suppl_1, p. D233–D238, 2008. ISSN 0305-1048. Disponível em: <https://doi.org/10.1093/nar/gkn663>. Acesso em: 02 dez. 2015.
- 53 BHAT, M. K.; BHAT, S. Cellulose degrading enzymes and their potential industrial applications. **Biotechnology Advances**, v. 15, n. 3–4, p. 583–620, 1997. ISSN 0734-9750. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0734975097000062>. Acesso em: 29 mar. 2016.
- 54 CAIRNS, J. K.; ESEN, A. β -Glucosidases. **Cellular and Molecular Life Sciences**, SP Birkhäuser Verlag Basel, v. 67, n. 20, p. 3389–3405, 2010. ISSN 1420-682X. Disponível em: <http://dx.doi.org/10.1007/s00018-010-0399-2>. Acesso em: 01 dez. 2015.
- 55 LIU, L. *et al.* Study on binding modes between cellobiose and beta-glucosidases from glycoside hydrolase family 1. **Bioorganic & Medicinal Chemistry Letters**, v. 22, n. 2, p. 837–843, 2012. ISSN 1464-3405 (Electronic).
- 56 KOROTKOVA, O. G. *et al.* Isolation and properties of fungal beta-glucosidases. **Biochemistry. Biokhimiia**, v. 74, n. 5, p. 569–577, 2009. ISSN 1608-3040 (Electronic).

- 57 ZHANG, P. *et al.* A kinetics modeling study on the inhibition of glucose on cellulosome of *Clostridium thermocellum*. **Bioresource Technology**, v. 190, p. 36–43, 2015. ISSN 0960-8524. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0960852415005350>. Acesso em: 04 dez. 2015.
- 58 ANDRIĆ, P. *et al.* Reactor design for minimizing product inhibition during enzymatic lignocellulose hydrolysis: I. Significance and mechanism of cellobiose and glucose inhibition on cellulolytic enzymes. **Biotechnology Advances**, v. 28, n. 3, p. 308–324, 2010. ISSN 0734-9750. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0734975010000054>. Acesso em: 26 abr. 2016.
- 59 LEVINE, S. E. *et al.* A mechanistic model of the enzymatic hydrolysis of cellulose. **Biotechnology and Bioengineering**, Wiley Subscription Services, Inc., A Wiley Company, v. 107, n. 1, p. 37–51, 9 2010. ISSN 1097-0290. Disponível em: <http://dx.doi.org/10.1002/bit.22789>. Acesso em: 29 mar. 2016.
- 60 WILKENS, C. *et al.* GH62 arabinofuranosidases: structure, function and applications. **Biotechnology Advances**, 2017. ISSN 0734-9750. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0734975017300769>. Acesso em: 15 ago. 2017.
- 61 WANG, W. *et al.* Elucidation of the molecular basis for arabinoxylan-debranching activity of a thermostable family GH62 alpha-l-arabinofuranosidase from *Streptomyces thermoviolaceus*. **Applied and Environmental Microbiology**, v. 80, n. 17, p. 5317–5329, 2014. ISSN 1098-5336 (Electronic).
- 62 LOMBARD, V. *et al.* The carbohydrate-active enzymes database (CAZy) in 2013. **Nucleic Acids Research**, v. 42, n. D1, p. D490–D495, 2013. ISSN 0305-1048. Disponível em: <https://doi.org/10.1093/nar/gkt1178>.
- 63 NAUMOFF, D. G. Hierarchical classification of glycoside hydrolases. **Biochemistry**, v. 76, n. 6, p. 622–635, 2011. ISSN 1608-3040. Disponível em: <http://dx.doi.org/10.1134/S0006297911060022>. Acesso em: 17 jul. 2017.
- 64 SIGUIER, B. *et al.* First structural insights into alpha-L-arabinofuranosidases from the two GH62 glycoside hydrolase subfamilies. **Journal of Biological Chemistry**, v. 289, n. 8, p. 5261–5273, 2014. ISSN 1083-351X (Electronic).
- 65 BELDMAN, G. *et al.* Arabinans and arabinan degrading enzymes. *In*: STURGEON, R. J. (ed.). **Advances in macromolecular carbohydrate research**. London: Elsevier Science, 1997. v. 1, cap. 1, p. 1–64. ISBN 9780080552279.
- 66 LAERE, K. M. J. V. *et al.* Purification and mode of action of two different arabinoxylan arabinofuranohydrolases from *Bifidobacterium adolescentis* DSM 20083. **Applied Microbiology and Biotechnology**, v. 51, n. 5, p. 606–613, 1999. ISSN 1432-0614. Disponível em: <https://doi.org/10.1007/s002530051439>. Acesso em: 31 ago. 2019.
- 67 CONTESINI, F. J. *et al.* Structural and functional characterization of a highly secreted α -l-arabinofuranosidase (GH62) from *Aspergillus nidulans* grown on sugarcane bagasse. **Biochimica et Biophysica Acta (BBA) — Proteins and Proteomics**, v. 1865, n. 12, p. 1758–1769, 2017. ISSN 1570-9639. Disponível em:

<http://www.sciencedirect.com/science/article/pii/S1570963917302091>. Acesso em: 20 fev. 2018.

68 NUMAN, M. T.; BHOSLE, N. B. Alpha-L-arabinofuranosidases: the potential applications in biotechnology. **Journal of Industrial Microbiology & Biotechnology**, v. 33, n. 4, p. 247–260, 2006. ISSN 1367-5435 (Print).

69 KIM, P. Current studies on biological tagatose production using l-arabinose isomerase: a review and future perspective. **Applied Microbiology and Biotechnology**, v. 65, n. 3, p. 243–249, 2004. ISSN 1432-0614. Disponível em: <https://doi.org/10.1007/s00253-004-1665-8>. Acesso em: 03 jun. 2019.

70 SÁNCHEZ-TORRES, P.; GONZÁLEZ-CANDELAS, L.; RAMÓN, D. Expression in a wine yeast strain of the *Aspergillus niger* abfB gene. **FEMS Microbiology Letters**, v. 145, n. 2, p. 189–194, 1996. ISSN 0378-1097. Disponível em: <https://doi.org/10.1111/j.1574-6968.1996.tb08576.x>. Acesso em: 03 jun. 2019.

71 VORAGEN, A. G. J. *et al.* The degradation of arabinans by endo-arabinanase and arabinofuranosidases purified from *Aspergillus niger*. **Food Hydrocolloids**, v. 1, n. 5, p. 423–437, 1987. ISSN 0268-005X. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0268005X8780036X>. Acesso em: 03 jun. 2019.

72 SAHA, B. C. Alpha-L-arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. **Biotechnology Advances**, v. 18, n. 5, p. 403–423, 2000. ISSN 0734-9750 (Print).

73 SØRENSEN, H. R. *et al.* Efficiencies of designed enzyme combinations in releasing arabinose and xylose from wheat arabinoxylan in an industrial ethanol fermentation residue. **Enzyme and Microbial Technology**, v. 36, n. 5, p. 773–784, 2005. ISSN 0141-0229. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0141022905000141>. Acesso em: 03 jun. 2019.

74 LAMMENS, W. *et al.* Structural insights into glycoside hydrolase family 32 and 68 enzymes: functional implications. **Journal of Experimental Botany**, v. 60, n. 3, p. 727–740, 2009. ISSN 0022-0957. Disponível em: <https://doi.org/10.1093/jxb/ern333>. Acesso em: 16 fev. 2021.

75 PONS, T. *et al.* Three acidic residues are at the active site of a beta-propeller architecture in glycoside hydrolase families 32, 43, 62, and 68. **Proteins**, v. 54, n. 3, p. 424–432, 2004. ISSN 1097-0134 (Electronic).

76 CHUANKHAYAN, P. *et al.* Crystal structures of *Aspergillus japonicus* fructosyltransferase complex with donor/acceptor substrates reveal complete subsites in the active site for catalysis. **Journal of Biological Chemistry**, v. 285, n. 30, p. 23251–23264, 2010. ISSN 1083-351X (Electronic).

77 NASIR, A. *et al.* Production and characterization of a high molecular weight levan and fructooligosaccharides from a rhizospheric isolate of *Bacillus aryabhatai*. **Lebensmittel-Wissenschaft und-Technologie**, v. 123, p. 109093, 2020. ISSN 0023-6438. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0023643820300815>. Acesso em: 21 fev. 2021.

- 78 ZHANG, J. *et al.* Establishment of a rapid and effective plate chromogenic assay for screening of *Aspergillus* species with high β -fructofuranosidase activity for fructooligosaccharides production. **Journal of Microbiological Methods**, v. 166, p. 105740, 2019. ISSN 1872-8359 (Electronic).
- 79 LIMA, M. Z. T. de *et al.* Crystal Structure of a Sucrose-6-phosphate Hydrolase from *Lactobacillus gasseri* with Potential Applications in Fructan Production and the Food Industry. **Journal of Agricultural and Food Chemistry**, American Chemical Society, v. 69, n. 35, p. 10223–10234, 2021. ISSN 0021-8561. Disponível em: <https://doi.org/10.1021/acs.jafc.1c03901>. Acesso em: 09 set. 2021.
- 80 HAKI, G. D.; RAKSHIT, S. K. Developments in industrially important thermostable enzymes: a review. **Bioresource Technology**, v. 89, n. 1, p. 17–34, 2003. ISSN 0960-8524 (Print).
- 81 MARTINS, E. S. *et al.* Solid state production of thermostable pectinases from thermophilic *Thermoascus aurantiacus*. **Process Biochemistry**, v. 37, n. 9, p. 949–954, 2002. ISSN 1359-5113. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0032959201003004>. Acesso em: 29 mar. 2016.
- 82 ZAMOST, B. L.; NIELSEN, H. K.; STARNES, R. L. Thermostable enzymes for industrial applications. **Journal of Industrial Microbiology**, v. 8, n. 2, p. 71–81, 1991. ISSN 1476-5535. Disponível em: <http://dx.doi.org/10.1007/BF01578757>.
- 83 RABALHO, A. A. **Isolamento de linhagens microbianas termofílicas amilolíticas, produção, caracterização e aplicação das amilases na hidrólise do amido de mandioca**. 2002. 183 p. Dissertação (mestrado) — Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas, São José do Rio Preto, 2002. Disponível em: <http://hdl.handle.net/11449/88420>. Acesso em: 29 mar. 2016.
- 84 VIIKARI, L. *et al.* Thermostable enzymes in lignocellulose hydrolysis. **Advances in Biochemical Engineering/Biotechnology**, v. 108, p. 121–145, 2007. ISSN 0724-6145 (Print).
- 85 BISCHOFF, K. M. *et al.* Purification and characterization of a family 5 endoglucanase from a moderately thermophilic strain of *Bacillus licheniformis*. **Biotechnology Letters**, v. 28, n. 21, p. 1761–1765, 2006. ISSN 0141-5492 (Print).
- 86 GOMES, E. *et al.* Enzimas termoestáveis: fontes, produção e aplicação industrial. **Química Nova**, SBQ, v. 30, n. 1, p. 136–145, 2007. ISSN 0100-4042. Disponível em: <https://doi.org/10.1590/S0100-40422007000100025>. Acesso em: 29 mar. 2016.
- 87 DENG, L. *et al.* Cellulose hydrolysis ability of a *Clostridium thermocellum* cellulosome containing small-size scaffolding protein CipA. **Journal of Biotechnology**, v. 212, p. 144–152, 2015. ISSN 0168-1656. Disponível em: <http://www.sciencedirect.com/science/article/pii/S016816561530095X>. Acesso em: 04 dez. 2015.
- 88 XU, Q. *et al.* Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities. **Science Advances**, v. 2, n. 2, 2016. Disponível em: <https://doi.org/10.1126/sciadv.1501254>. Acesso em: 21 mar. 2016.

- 89 PRAWITWONG, P. *et al.* Direct glucose production from lignocellulose using *Clostridium thermocellum* cultures supplemented with a thermostable [beta]-glucosidase. **Biotechnology for Biofuels**, v. 6, n. 1, p. 184, 2013. ISSN 17546834. Disponível em: <https://doi.org/10.1186/1754-6834-6-184>. Acesso em: 04 dez. 2015.
- 90 BAYER, E. A. *et al.* The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. **Annual Review of Microbiology**, v. 58, p. 521–554, 2004. ISSN 0066-4227 (Print).
- 91 DOI, R. H.; KOSUGI, A. Cellulosomes: plant-cell-wall-degrading enzyme complexes. **Nature Reviews Microbiology**, v. 2, n. 7, p. 541–551, 2004. ISSN 1740-1526. Disponível em: <http://dx.doi.org/10.1038/nrmicro925>. Acesso em: 26 nov. 2015.
- 92 BÉGUIN, P.; ALZARI, P. M. The cellulosome of *Clostridium thermocellum*. **Biochemical Society Transactions**, v. 26, n. 2, p. 178–184, 1998. Disponível em: <http://www.biochemsoctrans.org/content/26/2/178.abstract>. Acesso em: 02 dez. 2015.
- 93 HAMANN, P. R. V. *et al.* Evaluation of plant cell wall degrading enzyme production by *Clostridium thermocellum* B8 in the presence of raw agricultural wastes. **International Biodeterioration & Biodegradation**, v. 105, p. 97–105, 11 2015. ISSN 0964-8305. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0964830515300688>. Acesso em: 04 dez. 2015.
- 94 SCHOELER, C. *et al.* Ultrastable cellulosome-adhesion complex tightens under load. **Nature Communications**, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., v. 5, n. 1, 2014. Disponível em: <http://dx.doi.org/10.1038/ncomms663510.1038/ncomms6635>. Acesso em: 05 dez. 2015.
- 95 BAYER, E. A.; MORAG, E.; LAMED, R. The cellulosome — A treasure-trove for biotechnology. **Trends in Biotechnology**, v. 12, n. 9, p. 379–386, 1994. ISSN 0167-7799. Disponível em: <http://www.sciencedirect.com/science/article/pii/0167779994900396>. Acesso em: 02 dez. 2015.
- 96 HAMBERG, Y. *et al.* Elaborate cellulosome architecture of *Acetivibrio cellulolyticus* revealed by selective screening of cohesin-dockerin interactions. **PeerJ**, v. 2, p. e636, 2014. ISSN 2167-8359 (Electronic).
- 97 SMITH, S. P.; BAYER, E. A. Insights into cellulosome assembly and dynamics: from dissection to reconstruction of the supramolecular enzyme complex. **Current Opinion in Structural Biology**, v. 23, n. 5, p. 686–94, 2013. ISSN 1879-033X. Disponível em: <https://doi.org/10.1016/j.sbi.2013.09.002>. Acesso em: 03 dez. 2015.
- 98 FONTES, C. M. G. A.; GILBERT, H. J. Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. **Annual Review of Biochemistry**, v. 79, p. 655–81, 2010. ISSN 1545-4509. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/20373916>. Acesso em: 26 nov. 2015.
- 99 SALAMITOU, S. *et al.* Recognition specificity of the duplicated segments present in *Clostridium thermocellum* endoglucanase CelD and in the cellulosome-integrating protein CipA. **Journal of Bacteriology**, v. 176, n. 10, p. 2822–2827, 1994. ISSN 0021-9193. Disponível em: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC205435/>. Acesso em: 29 mar. 2016.

- 100 POOLE, D. M. *et al.* Identification of the cellulose-binding domain of the cellulosome subunit S1 from *Clostridium thermocellum* YS. **FEMS Microbiology Letters**, v. 78, n. 2-3, p. 181–186, 1992. ISSN 0378-1097 (Print).
- 101 ZHANG, Y.-H. P.; LYND, L. R. Regulation of cellulase synthesis in batch and continuous cultures of *Clostridium thermocellum*. **Journal of Bacteriology**, v. 187, n. 1, p. 99–106, 2005. ISSN 0021-9193. Disponível em: <http://jb.asm.org/content/187/1/99.abstract>. Acesso em: 29 mar. 2016.
- 102 ANWAR, M. A. *et al.* Inulin and levan synthesis by probiotic *Lactobacillus gasseri* strains: characterization of three novel fructansucrase enzymes and their fructan products. **Microbiology**, v. 156, n. Pt 4, p. 1264–1274, 2010. ISSN 1465-2080 (Electronic).
- 103 CUNHA, J. S. *et al.* Synthesis and characterization of fructosyltransferase from *Aspergillus oryzae* IPT-301 for high fructooligosaccharides production. **Brazilian Journal of Chemical Engineering**, SciELO, v. 36, n. 2, p. 657–668, 2019. ISSN 0104-6632. Disponível em: <https://doi.org/10.1590/0104-6632.20190362s20180572>. Acesso em: 21 fev. 2021.
- 104 DÍEZ-MUNICIO, M. *et al.* Enzymatic synthesis and characterization of fructooligosaccharides and novel maltosylfructosides by inulosucrase from *Lactobacillus gasseri* DSM 20604. **Applied and Environmental Microbiology**, American Society for Microbiology, v. 79, n. 13, p. 4129–4140, 2013. ISSN 1098-5336. Disponível em: <https://doi.org/10.1128/AEM.00854-13>. Acesso em: 21 fev. 2021.
- 105 BUITRÓN, D. I. *et al.* Biotechnological approach for the production of prebiotics and search for new probiotics and their application in the food industry. **Applied Food Biotechnology**, v. 5, n. 4 SE - Review Article, p. 185–192, 2018. Disponível em: <https://journals.sbm.u.ac.ir/afb/article/view/20174>. Acesso em: 21 fev. 2021.
- 106 PICAZO, B. *et al.* Enzymes for Fructooligosaccharides Production: Achievements and Opportunities. In: KUDDUS, M. B. T. E. i. F. B. (ed.). **Enzymes in food biotechnology production, applications, and future prospects**. Academic Press, 2019. cap. 18, p. 303–320. ISBN 978-0-12-813280-7. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780128132807000189>. Acesso em: 21 fev. 2021.
- 107 SANDOVAL-GONZÁLEZ, R. S.; JIMÉNEZ-ISLAS, H.; NAVARRETE-BOLAÑOS, J. L. Design of a fermentation process for agave fructooligosaccharides production using endo-inulinases produced in situ by *Saccharomyces paradoxus*. **Carbohydrate Polymers**, v. 198, p. 94–100, 2018. ISSN 0144-8617. Disponível em: <https://www.sciencedirect.com/science/article/pii/S014486171830729X>. Acesso em: 21 fev. 2021.
- 108 GUNYAKTI, A.; ASAN-OZUSAGLAM, M. *Lactobacillus gasseri* from human milk with probiotic potential and some technological properties. **Lebensmittel-Wissenschaft und-Technologie**, v. 109, p. 261–269, 2019. ISSN 0023-6438. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0023643819303470>. Acesso em: 21 fev. 2021.
- 109 ATALAH, J. *et al.* Thermophiles and the applications of their enzymes as new biocatalysts. **Bioresource Technology**, Elsevier, v. 280, p. 478–488, 2019. ISSN 0960-8524.

- 110 BERKA, R. M. *et al.* Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. **Nature Biotechnology**, Nature Publishing Group, v. 29, n. 10, p. 922–927, 2011. ISSN 1087-0156. Disponível em: <http://dx.doi.org/10.1038/nbt.1976>. Acesso em: 15 ago. 2017.
- 111 GARCÍA-HUANTE, Y. *et al.* The thermophilic biomass-degrading fungus *Thielavia terrestris* Co3Bag1 produces a hyperthermophilic and thermostable β -1,4-xylanase with exo- and endo-activity. **Extremophiles**, v. 21, n. 1, p. 175–186, 2017. ISSN 1433-4909. Disponível em: <https://doi.org/10.1007/s00792-016-0893-z>. Acesso em: 15 ago. 2017.
- 112 LANGSTON, J. A. *et al.* Cloning, expression, and characterization of a cellobiose dehydrogenase from *Thielavia terrestris* induced under cellulose growth conditions. **Biochimica et Biophysica Acta (BBA) — Proteins and Proteomics**, v. 1824, n. 6, p. 802–812, 2012. ISSN 1570-9639. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1570963912000568>. Acesso em: 03 jun. 2019.
- 113 MARGARITIS, A.; MERCHANT, R. F. J.; YAGUCHI, M. Thermostable cellulases from thermophilic microorganisms. **Critical Reviews in Biotechnology**, Taylor & Francis, v. 4, n. 3, p. 327–367, 1986. ISSN 0738-8551. Disponível em: <https://doi.org/10.3109/07388558609150799>. Acesso em: 03 jun. 2019.
- 114 TÖLGO, M. *et al.* Genomic and transcriptomic analysis of the thermophilic lignocellulose-degrading fungus *Thielavia terrestris* LPH172. **Biotechnology for Biofuels**, v. 14, n. 1, p. 131, 2021. ISSN 1754-6834. Disponível em: <https://doi.org/10.1186/s13068-021-01975-1>. Acesso em: 11 mar. 2022.
- 115 CAMARGO, S. *et al.* Functional and structural characterization of an α -L-arabinofuranosidase from *Thermothielavioides terrestris* and its exquisite domain-swapped β -propeller fold crystal packing. **Biochimica et Biophysica Acta (BBA) — Proteins and Proteomics**, v. 1868, n. 12, p. 140533, 2020. ISSN 1570-9639. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1570963920301801>. Acesso em: 09 set. 2020.
- 116 SADIQ, F. A. *et al.* A RAPD based study revealing a previously unreported wide range of mesophilic and thermophilic spore formers associated with milk powders in China. **International Journal of Food Microbiology**, v. 217, p. 200–208, 2016. ISSN 0168-1605. Disponível em: <http://www.sciencedirect.com/science/article/pii/S016816051530163X>. Acesso em: 03 jun. 2020.
- 117 ZOU, M.; LIU, D. A systematic characterization of the distribution, biofilm-forming potential and the resistance of the biofilms to the CIP processes of the bacteria in a milk powder processing factory. **Food Research International**, v. 113, p. 316–326, 2018. ISSN 0963-9969. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0963996918305581>. Acesso em: 03 jun. 2020.
- 118 LÜCKING, G. *et al.* Characterization of aerobic spore-forming bacteria associated with industrial dairy processing environments and product spoilage. **International Journal of Food Microbiology**, v. 166, n. 2, p. 270–279, 2013. ISSN 0168-1605. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168160513003395>. Acesso em: 03 jun. 2020.

119 SADIQ, F. A. *et al.* The heat resistance and spoilage potential of aerobic mesophilic and thermophilic spore forming bacteria isolated from Chinese milk powders. **International Journal of Food Microbiology**, v. 238, p. 193–201, 2016. ISSN 0168-1605. Disponível em: <http://www.sciencedirect.com/science/article/pii/S016816051630469X>. Acesso em: 03 jun. 2020.

120 DEGERING, C. *et al.* Optimization of protease secretion in *Bacillus subtilis* and *Bacillus licheniformis* by screening of homologous and heterologous signal peptides. **Applied and Environmental Microbiology**, v. 76, n. 19, p. 6370 LP – 6376, 2010. Disponível em: <http://aem.asm.org/content/76/19/6370.abstract>. Acesso em: 03 jun. 2020.

121 JELLOULI, K. *et al.* Alkaline-protease from *Bacillus licheniformis* MP1: Purification, characterization and potential application as a detergent additive and for shrimp waste deproteinization. **Process Biochemistry**, v. 46, n. 6, p. 1248–1256, 2011. ISSN 1359-5113. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1359511311000699>. Acesso em: 03 jun. 2020.

122 SARKER, P. K. *et al.* Optimization and partial characterization of culture conditions for the production of alkaline protease from *Bacillus licheniformis* P003. **SpringerPlus**, v. 2, n. 1, p. 506, 2013. ISSN 2193-1801. Disponível em: <https://doi.org/10.1186/2193-1801-2-506>. Acesso em: 03 jun. 2020.

123 KHYAMI-HORANI, H. Thermotolerant strain of *Bacillus licheniformis* producing lipase. **World Journal of Microbiology and Biotechnology**, v. 12, n. 4, p. 399–401, 1996. ISSN 1573-0972. Disponível em: <https://doi.org/10.1007/BF00340219>. Acesso em: 03 jun. 2020.

124 NTHANGENI, M. B. *et al.* Over-expression and properties of a purified recombinant *Bacillus licheniformis* lipase: a comparative report on *Bacillus* lipases. **Enzyme and Microbial Technology**, v. 28, n. 7, p. 705–712, 2001. ISSN 0141-0229. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0141022901003167>. Acesso em: 03 jun. 2020.

125 JUAJUN, O. *et al.* Cloning, purification, and characterization of β -galactosidase from *Bacillus licheniformis* DSM 13. **Applied Microbiology and Biotechnology**, v. 89, n. 3, p. 645–654, 2011. ISSN 1432-0614. Disponível em: <https://doi.org/10.1007/s00253-010-2862-2>. Acesso em: 03 jun. 2020.

126 BEKLER, F. M. *et al.* Cloning, purification and characterization of a thermostable β -galactosidase from *Bacillus licheniformis* strain KG9. **Cellular and Molecular Biology**, v. 61, n. 3, p. 71–78, 2015. ISSN 1165-158X (Electronic).

127 SADIQ, F. A. *et al.* Propensity for biofilm formation by aerobic mesophilic and thermophilic spore forming bacteria isolated from chinese milk powders. **International Journal of Food Microbiology**, v. 262, p. 89–98, 2017. ISSN 0168-1605. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168160517304038>. Acesso em: 03 jun. 2020.

128 WANG, N. *et al.* Tandem mass tag-based quantitative proteomics reveals the regulators in biofilm formation and biofilm control of *Bacillus licheniformis*. **Food Control**, v. 110, p. 107029, 2020. ISSN 0956-7135. Disponível em:

<http://www.sciencedirect.com/science/article/pii/S0956713519306188>. Acesso em: 20 mai. 2020.

129 CADER, M. Z. *et al.* FAMIN is a multifunctional purine enzyme enabling the purine nucleotide cycle. **Cell**, Elsevier, v. 180, n. 2, p. 278–295, 2020. ISSN 0092-8674. Disponível em: <https://doi.org/10.1016/j.cell.2019.12.017>. Acesso em: 21 jul. 2020.

130 ALTSCHUL, S. F. *et al.* Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403–410, 1990. ISSN 0022-2836 (Print).

131 ALTSCHUL, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v. 25, n. 17, p. 3389–3402, 1997. ISSN 0305-1048 (Print).

132 KUMAR, S. *et al.* MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. **Molecular Biology and Evolution**, v. 35, n. 6, p. 1547–1549, 2018. ISSN 0737-4038. Disponível em: <https://doi.org/10.1093/molbev/msy096>.

133 KUMAR, M. N. S. *et al.* **Molecular evolution and phylogenetics**. New York: Oxford University Press, 2000. 333 p. ISBN 9780195135855.

134 KRISSINEL, E.; HENRICK, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. **Acta Crystallographica Section D — biological crystallography**, International Union of Crystallography (IUCr), v. 60, n. 12, p. 2256–2268, 2004. ISSN 0907-4449. Disponível em: <https://doi.org/10.1107/S0907444904026460>.

135 KRISSINEL, E.; HENRICK, K. Multiple Alignment of Protein Structures in Three Dimensions. *In*: BERTHOLD, M. R. *et al.* (ed.). **Computational life sciences**. Berlin, Heidelberg: Springer, 2005. p. 67–78. ISBN 978-3-540-31726-5.

136 HOLM, L.; ROSENSTRÖM, P. Dali server: conservation mapping in 3D. **Nucleic Acids Research**, v. 38, n. suppl_2, p. W545–W549, 2010. ISSN 0305-1048. Disponível em: <http://dx.doi.org/10.1093/nar/gkq366>. Acesso em: 15 fev. 2017.

137 ROBERT, X.; GOUET, P. Deciphering key features in protein structures with the new ENDscript server. **Nucleic Acids Research**, v. 42, n. W1, p. W320–W324, 2014. ISSN 0305-1048. Disponível em: <http://dx.doi.org/10.1093/nar/gku316>. Acesso em: 15 fev. 2017.

138 LASKOWSKI, R. A.; WATSON, J. D.; THORNTON, J. M. ProFunc: a server for predicting protein function from 3D structure. **Nucleic Acids Research**, v. 33, n. suppl_2, p. W89–W93, 2005. ISSN 0305-1048. Disponível em: <http://dx.doi.org/10.1093/nar/gki414>. Acesso em: 15 fev. 2017.

139 SILLITOE, I. *et al.* CATH: increased structural coverage of functional space. **Nucleic Acids Research**, v. 49, n. D1, p. D266–D273, 2021. ISSN 0305-1048. Disponível em: <https://doi.org/10.1093/nar/gkaa1079>. Acesso em: 09 fev. 2022.

140 MISTRY, J. *et al.* Pfam: The protein families database in 2021. **Nucleic Acids Research**, v. 49, n. D1, p. D412–D419, 2021. ISSN 0305-1048. Disponível em: <https://doi.org/10.1093/nar/gkaa913>. Acesso em: 09 fev. 2022.

- 141 ANDREEVA, A. *et al.* The SCOP database in 2020: expanded classification of representative family and superfamily domains of known protein structures. **Nucleic Acids Research**, v. 48, n. D1, p. D376–D382, 2020. ISSN 0305-1048. Disponível em: <https://doi.org/10.1093/nar/gkz1064>. Acesso em: 09 fev. 2022.
- 142 KRISSINEL, E.; HENRICK, K. Inference of Macromolecular Assemblies from Crystalline State. **Journal of Molecular Biology**, v. 372, n. 3, p. 774–797, 2007. ISSN 0022-2836. Disponível em: <file://www.sciencedirect.com/science/article/pii/S0022283607006420>. Acesso em: 15 fev. 2017.
- 143 BERKA, K. *et al.* MOLEonline 2.0: interactive web-based analysis of biomacromolecular channels. **Nucleic Acids Research**, v. 40, n. W1, p. W222–W227, 2012. ISSN 0305-1048. Disponível em: <https://doi.org/10.1093/nar/gks363>. Acesso em: 21 fev. 2021.
- 144 CHOVANCOVA, E. *et al.* CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. **PLOS Computational Biology**, Public Library of Science, v. 8, n. 10, p. e1002708, 2012. Disponível em: <https://doi.org/10.1371/journal.pcbi.1002708>. Acesso em: 09 jul. 2019.
- 145 SCHRÖDINGER, L.; DELANO, W. **The PyMOL molecular graphics system, version 2.4.1**. Schrödinger, LLC, 2020. Disponível em: <http://www.pymol.org/pymol>.
- 146 GASTEIGER, E. *et al.* ExPASy: the proteomics server for in-depth protein knowledge and analysis. **Nucleic Acids Research**, Oxford University Press, Oxford, UK, v. 31, n. 13, p. 3784–3788, 2003. ISSN 0305-1048. Disponível em: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC168970/>. Acesso em: 21 nov. 2017.
- 147 GASTEIGER, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. In: WALKER, J. M. (ed.). **The proteomics protocols handbook**. Totowa, NJ: Humana Press, 2005. p. 571–607. ISBN 978-1-59259-890-8. Disponível em: <https://doi.org/10.1385/1-59259-890-0:571>. Acesso em: 09 ago. 2019.
- 148 ESCHENFELDT, W. H. *et al.* A Family of LIC vectors for high-throughput cloning and purification of proteins. In: DOYLE, S. A. (ed.). **High throughput protein expression and purification: methods and protocols**. Totowa, NJ: Humana Press, 2009. p. 105–115. ISBN 978-1-59745-196-3. Disponível em: http://dx.doi.org/10.1007/978-1-59745-196-3_7. Acesso em: 15 fev. 2017.
- 149 DOYLE, S. A. High-throughput cloning for proteomics research. **Methods in Molecular Biology**, v. 310, p. 107–113, 2005. ISSN 1064-3745 (Print).
- 150 BERTANI, G. Studies on lysogenesis. **Journal of Bacteriology**, American Society for Microbiology, v. 62, n. 3, p. 293–300, 1951. Disponível em: <https://doi.org/10.1128/jb.62.3.293-300.1951>. Acesso em: 09 mar. 2022.
- 151 STUDIER, F. W. Protein production by auto-induction in high density shaking cultures. **Protein Expression and Purification**, v. 41, n. 1, p. 207–234, 2005. ISSN 1046-5928 (Print).
- 152 NEUHOFF, V.; STAMM, R.; EIBL, H. Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: A systematic analysis. **Electrophoresis**, v. 6, n. 9, p. 427–448, 1985. ISSN 15222683.

- 153 NIESEN, F. H.; BERGLUND, H.; VEDADI, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. **Nature Protocols**, Nature Publishing Group, v. 2, n. 9, p. 2212–2221, 2007. ISSN 1754-2189. Disponível em: <http://dx.doi.org/10.1038/nprot.2007.321>. Acesso em: 10 set. 2017.
- 154 WYATT, P. J. Light scattering and the absolute characterization of macromolecules. **Analytica Chimica Acta**, v. 272, n. 1, p. 1–40, 1993. ISSN 0003-2670. Disponível em: <https://www.sciencedirect.com/science/article/pii/000326709380373S>. Acesso em: 12 nov. 2021.
- 155 FOLTA-STOGNIEW, E. Oligomeric states of proteins determined by size-exclusion chromatography coupled with light scattering, absorbance, and refractive index detectors BT. *In: New and Emerging Proteomic Techniques*. Totowa, NJ: Humana Press, 2006. p. 97–112. ISBN 978-1-59745-026-3. Disponível em: <https://doi.org/10.1385/1-59745-026-X:97>. Acesso em: 04 set. 2019.
- 156 FOLTA-STOGNIEW, E.; WILLIAMS, K. R. Determination of molecular masses of proteins in solution: implementation of an HPLC size exclusion chromatography and laser light scattering service in a core laboratory. **Journal of Biomolecular Techniques**, Association of Biomolecular Resource Facilities, v. 10, n. 2, p. 51–63, 1999. ISSN 1524-0215. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/19499008https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2291588/>. Acesso em: 12 nov. 2021.
- 157 KENDRICK, B. S. *et al.* Online size-exclusion high-performance liquid chromatography light scattering and differential refractometry methods to determine degree of polymer conjugation to proteins and protein–protein or protein–ligand association states. **Analytical Biochemistry**, v. 299, n. 2, p. 136–146, 2001. ISSN 0003-2697. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0003269701954119>. Acesso em: 12 nov. 2021.
- 158 MOGRIDGE, J. Using light scattering to determine the stoichiometry of protein complexes. *In: Protein-protein interactions: methods and applications*. Totowa, NJ: Humana Press, 2004. p. 113–118. ISBN 978-1-59259-762-8. Disponível em: <https://doi.org/10.1385/1-59259-762-9:113>. Acesso em: 12 nov. 2021.
- 159 HAMMERSLEY, A. FIT2D: a multi-purpose data reduction, analysis and visualization program. **Journal of Applied Crystallography**, International Union of Crystallography (IUCr), v. 49, n. 2, p. 646–652, 2016. ISSN 1600-5767. Disponível em: <https://doi.org/10.1107/S1600576716000455>. Acesso em: 18 out. 2018.
- 160 KONAREV, P. V. *et al.* PRIMUS: a Windows PC-based system for small-angle scattering data analysis. **Journal of Applied Crystallography**, International Union of Crystallography (IUCr), v. 36, n. 5, p. 1277–1282, 2003. ISSN 0021-8898. Disponível em: <https://doi.org/10.1107/S0021889803012779>. Acesso em: 10 out. 2019.
- 161 SVERGUN, D. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. **Journal of Applied Crystallography**, International Union of Crystallography (IUCr), v. 25, n. 4, p. 495–503, 1992. ISSN 0021-8898. Disponível em: <https://doi.org/10.1107/S0021889892001663>. Acesso em: 10 out. 2019.
- 162 FRANKE, D. *et al.* ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. **Journal of Applied Crystallography**,

International Union of Crystallography, v. 50, n. 4, p. 1212–1225, 2017. ISSN 1600-5767. Disponível em: <https://doi.org/10.1107/S1600576717007786>. Acesso em: 18 out. 2018.

163 NEWMAN, J. *et al.* Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 61, n. 10, p. 1426–1431, 2005. Disponível em: <https://doi.org/10.1107/S0907444905024984>. Acesso em: 07 jul. 2021.

164 FAZIO, V. J.; PEAT, T. S.; NEWMAN, J. A drunken search in crystallization space. **Acta Crystallographica Section F** — structural biology communications, International Union of Crystallography (IUCr), v. 70, n. Pt 10, p. 1303–1311, 2014. ISSN 2053-230X. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25286930https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4188070/>. Acesso em: 07 jul. 2021.

165 CHAIKUAD, A.; KNAPP, S.; DELFT, F. von. Defined PEG smears as an alternative approach to enhance the search for crystallization conditions and crystal-quality improvement in reduced screens. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 71, n. 8, p. 1627–1639, 2015. Disponível em: <https://doi.org/10.1107/S1399004715007968>. Acesso em: 07 jul. 2021.

166 GORREC, F. The MORPHEUS protein crystallization screen. **Journal of Applied Crystallography**, International Union of Crystallography (IUCr), v. 42, n. Pt 6, p. 1035–1042, 2009. ISSN 0021-8898. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/22477774https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3246824/>. Acesso em: 07 jul. 2021.

167 WINTER, G. xia2: an expert system for macromolecular crystallography data reduction. **Journal of Applied Crystallography**, International Union of Crystallography (IUCr), v. 43, n. 1, p. 186–190, 2010. ISSN 0021-8898. Disponível em: <https://doi.org/10.1107/S0021889809045701>. Acesso em: 27 nov. 2017.

168 KABSCH, W. Integration, scaling, space-group assignment and post-refinement. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 66, n. Pt 2, p. 133–144, 2010. ISSN 1399-0047 (Electronic). Disponível em: <https://doi.org/10.1107/S0907444909047374>. Acesso em: 01 dez. 2015.

169 KABSCH, W. XDS. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 66, n. Pt 2, p. 125–132, 2010. ISSN 0907-4449. Disponível em: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2815665/>. Acesso em: 01 dez. 2015.

170 EVANS, P. Scaling and assessment of data quality. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 62, n. Pt 1, p. 72–82, 2006. ISSN 0907-4449 (Print). Disponível em: <https://doi.org/10.1107/S0907444905036693>. Acesso em: 15 ago. 2017.

171 EVANS, P. R.; MURSHUDOV, G. N. How good are my data and what is the resolution? **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 69, n. 7, p. 1204–1214, 2013. ISSN 0907-4449. Disponível em: <https://doi.org/10.1107/S0907444913000061>. Acesso em: 21 jun. 2018.

- 172 WINN, M. D. *et al.* Overview of the CCP4 suite and current developments. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 67, n. 4, p. 235–242, 2011. ISSN 0907-4449. Disponível em: <https://doi.org/10.1107/S0907444910045749>. Acesso em: 26 dez. 2018.
- 173 TICKLE, I. *et al.* **STARANISO**. Cambridge, United Kingdom: Global Phasing Ltd., 2018. Disponível em: <http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>. Acesso em: 26 dez. 2018.
- 174 BRUKER, A. Saint and SADABS. **Bruker AXS Inc.**, Madison, Wisconsin, USA, 2009.
- 175 SHERIFF, S.; HENDRICKSON, W. A. Description of overall anisotropy in diffraction from macromolecular crystals. **Acta Crystallographica Section A** — foundations and advances, International Union of Crystallography (IUCr), v. 43, n. 1, p. 118–121, 1987. ISSN 0108-7673. Disponível em: <https://doi.org/10.1107/S010876738709977X>. Acesso em: 18 nov. 2021.
- 176 MURSHUDOV, G. N. *et al.* **The effect of overall anisotropic scaling in macromolecular refinement**. [*S.l.: s.n.*]: Citeseer, 1998. 37–42 p.
- 177 PAINTER, J.; MERRITT, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 62, n. 4, p. 439–450, 2006. ISSN 1399-0047. Disponível em: <https://doi.org/10.1107/S0907444906005270>. Acesso em: 12 jul. 2019.
- 178 PAINTER, J.; MERRITT, E. A. TLSMD web server for the generation of multi-group TLS models. **Journal of Applied Crystallography**, International Union of Crystallography (IUCr), v. 39, n. 1, p. 109–111, 2006. ISSN 0021-8898. Disponível em: <https://doi.org/10.1107/S0021889805038987>. Acesso em: 10 mar. 2021.
- 179 AFONINE, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 68, n. Pt 4, p. 352–367, 2012. ISSN 0907-4449. Disponível em: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3322595/>. Acesso em: 05 mai. 2016.
- 180 KARPLUS, P. A.; DIEDERICHS, K. Linking crystallographic model and data quality. **Science**, v. 336, n. 6084, p. 1030–1033, 2012. ISSN 1095-9203. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/22628654https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3457925/>. Acesso em: 15 nov. 2021.
- 181 POPOV, A. N.; BOURENKOV, G. P. Choice of data-collection parameters based on statistic modelling. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 59, n. 7, p. 1145–1153, 2003. Disponível em: <https://doi.org/10.1107/S0907444903008163>. Acesso em: 11 nov. 2021.
- 182 WILSON, A. J. C. Treatment of enhanced zones and rows in normalizing intensities. **Acta Crystallographica Section A** — foundations and advances, International Union of Crystallography (IUCr), v. 43, n. 2, p. 250–252, 1987. Disponível em: <https://doi.org/10.1107/S0108767387099495>. Acesso em: 15 nov. 2021.

183 DEBYE, P. Interferenz von röntgenstrahlen und wärmebewegung. **Annalen der Physik**, John Wiley & Sons, Ltd, v. 348, n. 1, p. 49–92, 1913. ISSN 0003-3804. Disponível em: <https://doi.org/10.1002/andp.19133480105>. Acesso em: 15 nov. 2021.

184 WALLER, I. Zur frage der einwirkung der wärmebewegung auf die interferenz von röntgenstrahlen. **Zeitschrift für Physik**, v. 17, n. 1, p. 398–408, 1923. ISSN 0044-3328. Disponível em: <https://doi.org/10.1007/BF01328696>. Acesso em: 15 nov. 2021.

185 FRENCH, S.; WILSON, K. On the treatment of negative intensity observations. **Acta Crystallographica Section A** — foundations and advances, International Union of Crystallography (IUCr), v. 34, n. 4, p. 517–525, 1978. ISSN 0567-7394. Disponível em: <https://doi.org/10.1107/S0567739478001114>. Acesso em: 21 fev. 2021.

186 MCCOY, A. J. *et al.* Phaser crystallographic software. **Journal of Applied Crystallography**, International Union of Crystallography (IUCr), v. 40, n. 4, p. 658–674, 2007. ISSN 0021-8898. Disponível em: <http://dx.doi.org/10.1107/S0021889807021206>. Acesso em: 05 mai. 2016.

187 POZZO, T. *et al.* Structural and functional analyses of β -glucosidase 3B from *Thermotoga neapolitana*: a thermostable three-domain representative of glycoside hydrolase 3. **Journal of Molecular Biology**, v. 397, n. 3, p. 724–739, 2010. ISSN 0022-2836. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0022283610001270>.

188 TERWILLIGER, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 64, n. 1, p. 61–69, 2008. ISSN 0907-4449. Disponível em: <https://doi.org/10.1107/S090744490705024X>. Acesso em: 26 dez. 2018.

189 ADAMS, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 66, n. 2, p. 213–221, 2010. ISSN 0907-4449. Disponível em: <https://doi.org/10.1107/S0907444909052925>. Acesso em: 15 fev. 2017.

190 COWTAN, K. The Buccaneer software for automated model building. 1. Tracing protein chains. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 62, n. 9, p. 1002–1011, 2006. ISSN 0907-4449. Disponível em: <https://doi.org/10.1107/S0907444906022116>. Acesso em: 12 jul. 2019.

191 EMSLEY, P.; COWTAN, K. Coot: model-building tools for molecular graphics. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 60, n. 12, p. 2126–2132, 2004. ISSN 0907-4449. Disponível em: <http://dx.doi.org/10.1107/S0907444904019158>. Acesso em: 22 abr. 2016.

192 MURSHUDOV, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 67, n. 4, p. 355–367, 2011. ISSN 0907-4449. Disponível em: <http://dx.doi.org/10.1107/S0907444911001314>. Acesso em: 05 mai. 2016.

- 193 CHEN, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 66, n. 1, p. 12–21, 2010. ISSN 0907-4449. Disponível em: <https://doi.org/10.1107/S0907444909042073>. Acesso em: 15 fev. 2017.
- 194 BERMAN, H. M. *et al.* The protein data bank. **Nucleic Acids Research**, v. 28, n. 1, p. 235–242, 2000. ISSN 0305-1048. Disponível em: <https://doi.org/10.1093/nar/28.1.235>. Acesso em: 10 out. 2019.
- 195 DESHPANDE, M. V.; ERIKSSON, K.-E.; PETTERSSON, L. G. An assay for selective determination of exo-1,4,- β -glucanases in a mixture of cellulolytic enzymes. **Analytical Biochemistry**, v. 138, n. 2, p. 481–487, 1984. ISSN 0003-2697. Disponível em: <http://www.sciencedirect.com/science/article/pii/0003269784908431>. Acesso em: 20 jun. 2018.
- 196 GHOSE, T. K. Measurement of cellulase activities. **Pure and Applied Chemistry**, International Union of Pure and Applied Chemistry, v. 59, n. 2, p. 257–268, 1987. ISSN 1365-3075. Disponível em: <http://iupac.org/publications/pac/59/2/0257/>. Acesso em: 20 jun. 2018.
- 197 MARQUES, A. R. *et al.* *Sphingomonas paucimobilis* beta-glucosidase Bgl1: a member of a new bacterial subfamily in glycoside hydrolase family 1. **Biochemical Journal**, v. 370, n. 3, p. 793 LP – 804, 2003. Disponível em: <http://www.biochemj.org/content/370/3/793.abstract>.
- 198 MCILVAINE, T. C. A buffer solution for colorimetric comparison. **Journal of Biological Chemistry**, v. 49, n. 1, p. 183–186, 1921. Disponível em: <http://www.jbc.org/content/49/1/183.short>. Acesso em: 26 jun. 2018.
- 199 KAUFMANN, J.; SCHERING, A. G. **Analysis of variance ANOVA**. [S.l.: s.n.], 2014. ISBN 9781118445112. Disponível em: <https://doi.org/10.1002/9781118445112.stat06938>. Acesso em: 10 mar. 2022.
- 200 TUKEY, J. W. Comparing individual means in the analysis of variance. **Biometrics**, Wiley, International Biometric Society, v. 5, n. 2, p. 99–114, 1949. ISSN 0006341X, 15410420. Disponível em: <http://www.jstor.org/stable/3001913>. Acesso em: 10 mar. 2022.
- 201 MICHAELIS, L.; MENTEN, M. L. Die kinetik der invertinwirkung. **Biochemische Zeitschrift**, v. 49, n. 333-369, p. 352, 1913. Disponível em: https://www.chem.uwec.edu/Chem352_Resources/pages/readings/media/Michaelis_&_Menton_1913.pdf. Acesso em: 26 jun. 2018.
- 202 BELOQUI, A. *et al.* Novel polyphenol oxidase mined from a metagenome expression library of bovine rumen: biochemical properties, structural analysis, and phylogenetic relationships. **Journal of Biological Chemistry**, v. 281, n. 32, p. 22933–22942, 2006. Disponível em: <http://www.jbc.org/content/281/32/22933.abstract>. Acesso em: 15 fev. 2017.
- 203 COLLINS, P. J.; DOBSON, A. D. W.; FIELD, J. A. Reduction of the 2,2'-Azinobis(3-ethylbenzthiazoline-6-sulfonate) cation radical by physiological organic acids in the absence and presence of manganese. **Applied and Environmental Microbiology**, American Society for Microbiology, v. 64, n. 6, p. 2026–2031, 1998. ISSN 1098-5336.

Disponível em: <https://doi.org/10.1128/AEM.64.6.2026-2031.1998%0A>. Acesso em: 20 ago. 2021.

204 TORINI, J. R. *et al.* The molecular structure of *Schistosoma mansoni* PNP isoform 2 provides insights into the nucleoside selectivity of PNPs. **PLOS ONE**, Public Library of Science, v. 13, n. 9, p. e0203532, 2018. Disponível em: <https://doi.org/10.1371/journal.pone.0203532>. Acesso em: 27 ago. 2020.

205 INKSCAPE PROJECT. **Inkscape**. GPLv2, 2022. Disponível em: <https://inkscape.org>. Acesso em: 09 mar. 2022.

206 SVERGUN, D.; BARBERATO, C.; KOCH, M. H. J. CRY SOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. **Journal of Applied Crystallography**, International Union of Crystallography (IUCr), v. 28, n. 6, p. 768–773, 1995. ISSN 0021-8898. Disponível em: <https://doi.org/10.1107/S0021889895007047>. Acesso em: 21 fev. 2021.

207 PIIADOV, V. *et al.* SAXSMoW 2.0: Online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale. **Protein Science**, John Wiley & Sons, Inc., v. 28, n. 2, p. 454–463, 2019. ISSN 1469-896X. Disponível em: <https://doi.org/10.1002/pro.3528>. Acesso em: 11 mar. 2022.

208 GAREN, C. R. *et al.* The molecular structure of Rv1873, a conserved hypothetical protein from *Mycobacterium tuberculosis*, at 1.38 Å resolution. **Acta Crystallographica Section F** — structural biology communications, International Union of Crystallography (IUCr), v. 62, n. 12, p. 1201–1205, 2006. ISSN 1744-3091. Disponível em: <https://doi.org/10.1107/S1744309106046902>. Acesso em: 19 jul. 2019.

209 GOODACRE, N. F.; GERLOFF, D. L.; UETZ, P. Protein domains of unknown function are essential in bacteria. **mBio**, v. 5, n. 1, p. 00744–13, 2014. Disponível em: <http://mbio.asm.org/content/5/1/e00744-13.abstract>. Acesso em: 19 jul. 2019.

210 AFONINE, P. V. *et al.* FEM: feature-enhanced map. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 71, n. Pt 3, p. 646–666, 2015. ISSN 1399-0047. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/25760612https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4356370/>.

211 KRISSINEL, E. Crystal contacts as nature's docking solutions. **Journal of Computational Chemistry**, John Wiley & Sons, Ltd, v. 31, n. 1, p. 133–143, 2010. ISSN 0192-8651. Disponível em: <https://doi.org/10.1002/jcc.21303>. Acesso em: 11 set. 2019.

212 RASTEGARI, A. A. Identifying potential cationic surfactant – cellulase and computational approaches. *In*: GUPTA, V. K. B. T. (ed.). **New and future developments in microbial biotechnology and bioengineering**. Amsterdam: Elsevier, 2016. cap. 20, p. 223–235. ISBN 978-0-444-63507-5. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780444635075000204>. Acesso em: 12 mar. 2022.

213 ECKARD, A. D.; MUTHUKUMARAPPAN, K.; GIBBONS, W. A review of the role of amphiphiles in biomass to ethanol conversion. **Applied Sciences**, v. 3, n. 2, p. 396–419, 2013. ISSN 2076-3417. Disponível em: <https://www.mdpi.com/2076-3417/3/2/396>. Acesso em: 12 mar. 2022.

- 214 RASTEGARI, A. A.; BORDBAR, A.-K.; TAHERI-KAFRANI, A. Interaction of cellulase with cationic surfactants: Using surfactant membrane selective electrodes and fluorescence spectroscopy. **Colloids and Surfaces B — biointerfaces**, v. 73, n. 1, p. 132–139, 2009. ISSN 0927-7765. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0927776509002021>. Acesso em: 12 mar. 2022.
- 215 MENEGOL, D. *et al.* Increased release of fermentable sugars from elephant grass by enzymatic hydrolysis in the presence of surfactants. **Energy Conversion and Management**, v. 88, p. 1252–1256, 2014. ISSN 0196-8904. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0196890414001903>. Acesso em: 12 mar. 2022.
- 216 WU, J.; JU, L.-K. Enhancing enzymatic saccharification of waste newsprint by surfactant addition. **Biotechnology Progress**, American Chemical Society (ACS), v. 14, n. 4, p. 649–652, 1998. ISSN 8756-7938. Disponível em: <https://doi.org/10.1021/bp980040v>. Acesso em: 12 mar. 2022.
- 217 KAAR, W. E.; HOLTZAPPLE, M. T. Benefits from Tween during enzymic hydrolysis of corn stover. **Biotechnology and Bioengineering**, John Wiley & Sons, Ltd, v. 59, n. 4, p. 419–427, 1998. ISSN 0006-3592. Disponível em: [https://doi.org/10.1002/\(SICI\)1097-0290\(19980820\)59:4%3C419::AID-BIT4%3E3.0.COhttp://2-j](https://doi.org/10.1002/(SICI)1097-0290(19980820)59:4%3C419::AID-BIT4%3E3.0.COhttp://2-j). Acesso em: 12 mar. 2022.
- 218 ZHOU, Y. *et al.* Non-ionic surfactants do not consistently improve the enzymatic hydrolysis of pure cellulose. **Bioresource Technology**, v. 182, p. 136–143, 2015. ISSN 0960-8524. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0960852415001662>. Acesso em: 12 mar. 2022.
- 219 BÖRJESSON, J. *et al.* Effect of poly(ethylene glycol) on enzymatic hydrolysis and adsorption of cellulase enzymes to pretreated lignocellulose. **Enzyme and Microbial Technology**, v. 41, n. 1, p. 186–195, 2007. ISSN 0141-0229. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0141022907000191>. Acesso em: 12 mar. 2022.
- 220 SEIDEL, Z. P.; LEE, C. T. Enhanced activity of the cellulase enzyme β -glucosidase upon addition of an azobenzene-based surfactant. **ACS Sustainable Chemistry & Engineering**, American Chemical Society, v. 8, n. 4, p. 1751–1761, 2020. Disponível em: <https://doi.org/10.1021/acssuschemeng.9b05240>. Acesso em: 12 mar. 2022.
- 221 KAO, M.-R. *et al.* *Chaetomella raphigera* β -glucosidase D2-BGL has intriguing structural features and a high substrate affinity that renders it an efficient cellulase supplement for lignocellulosic biomass hydrolysis. **Biotechnology for Biofuels**, v. 12, n. 1, p. 258, 2019. ISSN 1754-6834. Disponível em: <https://doi.org/10.1186/s13068-019-1599-0>. Acesso em: 25 jun. 2021.
- 222 SUZUKI, K. *et al.* Crystal structures of glycoside hydrolase family 3 β -glucosidase 1 from *Aspergillus aculeatus*. **Biochemical Journal**, v. 452, n. 2, p. 211 LP – 221, 2013. Disponível em: <http://www.biochemj.org/content/452/2/211.abstract>. Acesso em: 26 jun. 2018.
- 223 NAKAJIMA, M. *et al.* Functional and structural analysis of a β -glucosidase involved in β -1,2-glucan metabolism in *Listeria innocua*. **PLOS ONE**, Public Library of Science,

v. 11, n. 2, p. e0148870, 2016. Disponível em: <https://doi.org/10.1371/journal.pone.0148870>. Acesso em: 25 jun. 2021.

224 ISHIGURO, R. *et al.* Function and structure relationships of a β -1,2-glucooligosaccharide-degrading β -glucosidase. **FEBS Letters**, John Wiley & Sons, Ltd, v. 591, n. 23, p. 3926–3936, 2017. ISSN 0014-5793. Disponível em: <https://doi.org/10.1002/1873-3468.12911>. Acesso em: 25 jun. 2021.

225 MAHASENAN, K. V. *et al.* Catalytic cycle of glycoside hydrolase BglX from *Pseudomonas aeruginosa* and its implications for biofilm formation. **ACS Chemical Biology**, American Chemical Society, v. 15, n. 1, p. 189–196, 2020. ISSN 1554-8929. Disponível em: <https://doi.org/10.1021/acscchembio.9b00754>. Acesso em: 25 jun. 2021.

226 MCANDREW, R. P. *et al.* From soil to structure, a novel dimeric β -glucosidase belonging to glycoside hydrolase family 3 isolated from compost using metagenomic analysis. **Journal of Biological Chemistry**, v. 288, n. 21, p. 14985–14992, 2013. Disponível em: <http://www.jbc.org/content/288/21/14985.abstract>. Acesso em: 26 jun. 2018.

227 YOSHIDA, E. *et al.* Role of a PA14 domain in determining substrate specificity of a glycoside hydrolase family 3 β -glucosidase from *Kluyveromyces marxianus*. **Biochemical Journal**, v. 431, n. 1, p. 39–49, 2010. ISSN 0264-6021. Disponível em: <https://doi.org/10.1042/BJ20100351>. Acesso em: 27 mai. 2021.

228 HANES, C. S. Studies on plant amylases: The effect of starch concentration upon the velocity of hydrolysis by the amylase of germinated barley. **Biochemical Journal**, v. 26, n. 5, p. 1406–1421, 1932. ISSN 0264-6021. Disponível em: <https://doi.org/10.1042/bj0261406>. Acesso em: 06 mar. 2022.

229 COTA, J. *et al.* Comparative analysis of three hyperthermophilic GH1 and GH3 family members with industrial potential. **New Biotechnology**, v. 32, n. 1, p. 13–20, 2015. ISSN 1871-6784. Disponível em: <http://www.sciencedirect.com/science/article/pii/S187167841402130X>. Acesso em: 19 jul. 2017.

230 KARKEHABADI, S. *et al.* Biochemical characterization and crystal structures of a fungal family 3 β -glucosidase, Cel3A from *Hypocrea jecorina*. **Journal of Biological Chemistry**, v. 289, n. 45, p. 31624–31637, 2014. Disponível em: <http://www.jbc.org/content/289/45/31624.abstract>. Acesso em: 29 jul. 2019.

231 FLORINDO, R. N. *et al.* Structural and biochemical characterization of a GH3 β -glucosidase from the probiotic bacteria *Bifidobacterium adolescentis*. **Biochimie**, v. 148, p. 107–115, 2018. ISSN 0300-9084. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0300908418300683>. Acesso em: 06 out. 2021.

232 BUETOW, L. *et al.* Structure of the Rho-activating domain of *Escherichia coli* cytotoxic necrotizing factor 1. **Nature Structural Biology**, v. 8, n. 7, p. 584–588, 2001. ISSN 1072-8368 (Print).

233 PARVEEN, S.; REDDY, M. Identification of YfiH (PgeF) as a factor contributing to the maintenance of bacterial peptidoglycan composition. **Molecular Microbiology**, Wiley/Blackwell, v. 105, n. 5, p. 705–720, 2017. ISSN 0950-382X. Disponível em: <https://doi.org/10.1111/mmi.13730>. Acesso em: 25 out. 2018.

- 234 THURSTON, C. F. The structure and function of fungal laccases. **Microbiology**, Microbiology Society, v. 140, n. 1, p. 19–26, 1994. ISSN 1350-0872. Disponível em: <http://mic.microbiologyresearch.org/content/journal/micro/10.1099/13500872-140-1-19>. Acesso em: 25 out. 2018.
- 235 KUNAMNENI, A. *et al.* Laccases and their applications: a patent review. **Recent Patents on Biotechnology**, v. 2, n. 1, p. 10–24, 2008. ISSN 18722083. Acesso em: 25 out. 2018.
- 236 KOMORI, H. *et al.* New insights into the catalytic active-site structure of multicopper oxidases. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 70, n. 3, p. 772–779, 2014. ISSN 1399-0047. Disponível em: <https://doi.org/10.1107/S1399004713033051>. Acesso em: 21 dez. 2018.
- 237 WHERLAND, S.; MIYAZAKI, K.; PECHT, I. Intramolecular electron transfer in the bacterial two-domain multicopper oxidase mgLAC. **Biochemistry**, American Chemical Society, v. 55, n. 21, p. 2960–2966, 2016. ISSN 0006-2960. Disponível em: <http://dx.doi.org/10.1021/acs.biochem.6b00158>. Acesso em: 10 jan. 2017.
- 238 DWIVEDI, U. N. *et al.* Structure–function relationship among bacterial, fungal and plant laccases. **Journal of Molecular Catalysis B** — enzymatic, v. 68, n. 2, p. 117–128, 2011. ISSN 1381-1177. Disponível em: <file://www.sciencedirect.com/science/article/pii/S1381117710002857>. Acesso em: 01 fev. 2017.
- 239 KALLIO, J. P. *et al.* Structure–function studies of a *Melanocarpus albomyces* laccase suggest a pathway for oxidation of phenolic compounds. **Journal of Molecular Biology**, v. 392, n. 4, p. 895–909, 2009. ISSN 0022-2836. Disponível em: <file://www.sciencedirect.com/science/article/pii/S0022283609007736>. Acesso em: 15 dez. 2017.
- 240 POLYAKOV, K. M. *et al.* Structure of native laccase from *Trametes hirsuta* at 1.8 Å resolution. **Acta Crystallographica Section D** — biological crystallography, International Union of Crystallography (IUCr), v. 65, n. 6, p. 611–617, 2009. ISSN 0907-4449. Disponível em: <https://doi.org/10.1107/S0907444909011950>. Acesso em: 10 jan. 2017.
- 241 CHEN, Z. *et al.* The role of Glu498 in the dioxygen reactivity of CotA-laccase from *Bacillus subtilis*. **Dalton Transactions**, The Royal Society of Chemistry, v. 39, n. 11, p. 2875–2882, 2010. ISSN 1477-9226. Disponível em: <http://dx.doi.org/10.1039/B922734B>. Acesso em: 15 fev. 2017.
- 242 PANTOLIANO, M. W. *et al.* High-density miniaturized thermal shift assays as a general strategy for drug discovery. **Journal of Biomolecular Screening**, SAGE Publications Inc STM, v. 6, n. 6, p. 429–440, 2001. ISSN 1087-0571. Disponível em: <http://journals.sagepub.com/doi/abs/10.1177/108705710100600609>. Acesso em: 25 out. 2018.
- 243 LO, M.-C. *et al.* Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. **Analytical Biochemistry**, v. 332, n. 1, p. 153–159, 2004. ISSN 0003-2697. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0003269704003756>. Acesso em: 25 out. 2018.

- 244 ERICSSON, U. B. *et al.* Thermofluor-based high-throughput stability optimization of proteins for structural studies. **Analytical Biochemistry**, v. 357, n. 2, p. 289–298, 2006. ISSN 00032697.
- 245 KIM, Y. *et al.* Crystal structure of hypothetical protein YfiH from *Shigella flexneri* at 2 Å resolution. **Proteins — structure, function, and bioinformatics**, Wiley Subscription Services, Inc., A Wiley Company, v. 63, n. 4, p. 1097–1101, 2006. ISSN 1097-0134. Disponível em: <http://dx.doi.org/10.1002/prot.20589>. Acesso em: 15 fev. 2017.
- 246 MURZIN, A. G. *et al.* SCOP: a structural classification of proteins database for the investigation of sequences and structures. **Journal of Molecular Biology**, v. 247, n. 4, p. 536–540, 1995. ISSN 0022-2836 (Print).
- 247 ZHENG, H. *et al.* Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. **Nature Protocols**, Nature Publishing Group, v. 9, p. 156, 12 2013. Disponível em: <https://doi.org/10.1038/nprot.2013.172>. Acesso em: 23 ago. 2019.
- 248 ZHENG, H. *et al.* CheckMyMetal: a macromolecular metal-binding validation tool. **Acta Crystallographica Section D — biological crystallography**, International Union of Crystallography (IUCr), v. 73, n. 3, p. 223–233, 2017. ISSN 2059-7983. Disponível em: <https://doi.org/10.1107/S2059798317001061>. Acesso em: 23 ago. 2019.
- 249 DITUSA, C. A. *et al.* Thermodynamics of metal ion binding. 1. Metal ion binding by wild-type carbonic anhydrase. **Biochemistry**, American Chemical Society, v. 40, n. 18, p. 5338–5344, 2001. ISSN 0006-2960. Disponível em: <https://doi.org/10.1021/bi001731e>. Acesso em: 08 out. 2019.
- 250 NIKU-PAAVOLA, M. L.; RAASKA, L.; ITÄVAARA, M. Detection of white-rot fungi by a non-toxic stain. **Mycological Research**, v. 94, n. 1, p. 27–31, 1990. ISSN 0953-7562. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0953756209812604>. Acesso em: 18 abr. 2017.
- 251 BELDMAN, G. *et al.* Synergism in cellulose hydrolysis by endoglucanases and exoglucanases purified from *Trichoderma viride*. **Biotechnology and Bioengineering**, Wiley Subscription Services, Inc., A Wiley Company, v. 31, n. 2, p. 173–178, 1988. ISSN 1097-0290. Disponível em: <http://dx.doi.org/10.1002/bit.260310211>. Acesso em: 18 abr. 2017.
- 252 SZKLARCZYK, D. *et al.* STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. **Nucleic Acids Research**, v. 47, n. D1, p. D607–D613, 2018. ISSN 0305-1048. Disponível em: <https://doi.org/10.1093/nar/gky1131>.
- 253 DONOVAN, C.; BRAMKAMP, M. Cell division in Corynebacterineae. **Frontiers in Microbiology**, v. 5, p. 132, 2014. Disponível em: <https://www.frontiersin.org/article/10.3389/fmicb.2014.00132>. Acesso em: 27 nov. 2019.
- 254 HAMOEN, L. W. *et al.* SepF, a novel FtsZ-interacting protein required for a late step in cell division. **Molecular Microbiology**, John Wiley & Sons, v. 59, n. 3, p. 989–999, 2006. ISSN 0950-382X. Disponível em: <https://doi.org/10.1111/j.1365-2958.2005.04987.x>. Acesso em: 27 nov. 2019.

-
- 255 HUANG, K.-H.; DURAND-HEREDIA, J.; JANAKIRAMAN, A. FtsZ Ring stability: of bundles, tubules, crosslinks, and curves. **Journal of Bacteriology**, v. 195, n. 9, p. 1859 LP – 1868, 2013. Disponível em: <http://jb.asm.org/content/195/9/1859.abstract>. Acesso em: 27 nov. 2019.
- 256 YOUNG, K. D. Bacterial shape: two-dimensional questions and possibilities. **Annual Review of Microbiology**, Annual Reviews, v. 64, n. 1, p. 223–240, 2010. ISSN 0066-4227. Disponível em: <https://doi.org/10.1146/annurev.micro.112408.134102>. Acesso em: 27 nov. 2019.
- 257 ZWEERS, J. C. *et al.* Towards the development of *Bacillus subtilis* as a cell factory for membrane proteins and protein complexes. **Microbial Cell Factories**, v. 7, n. 1, p. 10, 2008. ISSN 1475-2859. Disponível em: <https://doi.org/10.1186/1475-2859-7-10>. Acesso em: 27 nov. 2019.
- 258 THOMAS, P. D. *et al.* PANTHER: a library of protein families and subfamilies indexed by function. **Genome Research**, v. 13, n. 9, p. 2129–2141, 2003. ISSN 1088-9051 (Print).
- 259 MENG-SHENG, L. *et al.* Structural basis for the peptidoglycan-editing activity of YfiH. **mBio**, American Society for Microbiology, v. 13, n. 1, p. 03646–21, 2022. Disponível em: <https://doi.org/10.1128/mbio.03646-21>. Acesso em: 06 mar. 2022.
- 260 WILSON, A. The probability distribution of X-ray intensities. **Acta Crystallographica**, International Union of Crystallography (IUCr), v. 2, n. 5, p. 318–321, 1949. ISSN 0365-110X. Disponível em: <https://doi.org/10.1107/S0365110X49000813>.

APÊNDICES

APÊNDICE A – DEFINIÇÕES DOS INDICADORES DE QUALIDADE CRISTALOGRÁFICA MAIS COMUMENTE UTILIZADOS

- **Resolução** (d_{min})

A resolução de um conjunto de dados de difração descreve a extensão dos dados mensuráveis sendo calculada pela lei de Bragg [Equação A.1](#) com base no ângulo máximo 2θ incluído no conjunto de dados para um determinado comprimento de onda da coleta de dados λ .

$$d_{min} = \frac{\lambda}{2 \sin \theta} \quad (\text{A.1})$$

- **Completeza** (C)

A completeza C de um conjunto de dados de difração é definida como a fração das reflexões únicas em um determinado grupo espacial para uma determinada resolução d_{min} medida, pelo menos uma vez, durante a coleta de dados. C pode ser dado assumindo que a simetria de Friedel é aplicada ou não. No último caso, C também é conhecido como completeza anômala. Pode ser dada como a porcentagem (%) do número esperado de reflexões para os dados, grupo espacial e resolução, respectivamente.

- **Multiplicidade** ou **redundância** (N)

A multiplicidade ou redundância N de um conjunto de dados de difração define em média quantas vezes uma reflexão hkl foi observada durante o experimento de coleta de dados, incluindo pares de simetria e medições replicadas. N pode ser dado assumindo que a simetria de Friedel é aplicada ou não.

- **Coefficiente de correlação** (CC)

Os coeficientes de correlação são invariavelmente o coeficiente de correlação linear de Pearson, dos dados obtidos quando um mesmo conjunto de observações são feitos duas vezes (x_j, y_j) é dado por:

$$CC = \frac{\sum_j (x_j - \langle x \rangle)(y_j - \langle y \rangle)}{\sqrt{\sum_j (x_j - \langle x \rangle)^2 \sum_j (y_j - \langle y \rangle)^2}} \quad (\text{A.2})$$

com x_j e y_j sendo, as intensidades das reflexões únicas mescladas (*merged*) nas observações atribuídas randomicamente aos subconjuntos respectivamente, por exemplo, no caso de diferenças anômalas $[I(hkl) - I(\bar{h}\bar{k}\bar{l})]$ ou $[F(hkl) - F(\bar{h}\bar{k}\bar{l})]$ nos dois conjuntos de dados, $\langle x \rangle$ e $\langle y \rangle$ são suas médias, e os somatórios são sobre todas as

reflexões hkl para as quais existem observações em ambos os conjuntos de dados em todo intervalo de resolução ou dentro de uma camada de resolução em particular. CC_{anom} é um indicador confiável da força do sinal anômalo. Valores acima de 0,30 são considerados bons (180).

- **Wilson-plot B Fator (B_{Wilson})**

Um gráfico de Wilson (260) é um gráfico para uma série contígua de camadas de resolução do logaritmo da intensidade média em uma dada camada de resolução dividido pela soma dos quadrados fatores de forma atômica para todos os átomos na cela unitária avaliada na média dos limites de resolução da camada. A partir de um ajuste de mínimos quadrados de uma linha reta para a parte linear do gráfico de Wilson, o B fator B_{Wilson} pode ser derivado. Normalmente, os dados com resolução inferior a 4,5 Å são excluídos do ajuste. As determinações mais significativas de B_{Wilson} vêm de gráficos de Wilson que são lineares até a resolução d_{min} e minimizam a ocorrência de picos devido a anéis de gelo.

$$\ln \left(\frac{\langle I_{obs} \rangle}{\sum_i f_i^2} \right) = -\ln K_{wilson} - \frac{2B_{wilson}}{d^2} \quad (A.3)$$

onde $\langle I_{obs}(hkl) \rangle$ é a média sobre as intensidades de todas as reflexões hkl observadas em uma determinada camada de resolução. A soma \sum_i sobre todos os átomos na estrutura. O parâmetro d é o ponto médio da camada de resolução sobre a qual foi calculada a média de I_{obs} . K_{Wilson} é um fator de escala absoluto.

- **$B_{overall}$ ($Patterson$)**

B fator geral

$$\sqrt{\frac{8\pi^2\sigma_{Patt}}{(2)}} \quad (A.4)$$

onde σ_{Patt} é o desvio padrão da Gaussiana ajustada ao pico de origem de Patterson.

- **R_{merge}**

(R fator de mergeamento),

$$R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)} \quad (A.5)$$

$\langle I(hkl) \rangle$ é a média das medidas individuais $I_i(hkl)$ da intensidade da reflexão hkl .

- **R_{meas}** ou **$R_{r.i.m}$**

(R fator de mergeamento independente da redundância),

$$R_{meas} = \frac{\sum_{hkl} \sqrt{\frac{N(hkl)}{N(hkl)-1}} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)} \quad (\text{A.6})$$

$\langle I(hkl) \rangle$ é a média das $N(hkl)$ medidas individuais $I_i(hkl)$ da intensidade da reflexão hkl .

- **$R_{p.i.m}$**

(R fator de mergeamento indicando precisão),

$$R_{p.i.m} = \frac{\sum_{hkl} \sqrt{\frac{1}{N(hkl)-1}} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)} \quad (\text{A.7})$$

$\langle I(hkl) \rangle$ é a média das $N(hkl)$ medidas individuais $I_i(hkl)$ da intensidade da reflexão hkl .

- **R_{anom}**

(R fator anômalo)

$$R_{anom} = \frac{\sum_{hkl} |I(hkl) - I(\bar{h}\bar{k}\bar{l})|}{\sum_{hkl} \langle I(hkl) \rangle} \quad (\text{A.8})$$

$\langle I(hkl) \rangle$ é a intensidade média dos pares de Friedel da reflexão hkl , ou $\frac{1}{2}[I(hkl) + I(\bar{h}\bar{k}\bar{l})]$.

- **R** e **R_{free}**

(R fator cristalográfico e R_{free}),

$$R = \frac{\sum_{hkl} ||F_{obs}| - |F_{cal}||}{\sum_{hkl} |F_{obs}|} \quad (\text{A.9})$$

R_{free} é definido como o R fator cristalográfico para um subconjunto de reflexões excluídas do refinamento.

- **R.M.S.D's** dos parâmetros geométricos x

$$r.m.s.d(x) = \sqrt{\frac{\sum_i [x_i(ideal) - x_i(observado)]^2}{N}} \quad (\text{A.10})$$

x_i são os valores individuais, ideais ou observados, do parâmetro geométrico x e a soma sobre todos Nx_i observados. Os parâmetros geométricos x podem ser comprimentos de ligações, ângulos, ângulos diedrais, volumes quirais, desvios da planaridade, etc.

APÊNDICE B – OS PARÂMETROS DOS ENSAIOS BIOQUÍMICOS

- Modelo cinético de *Michaelis-Menten*

Esta é a expressão central para a cinética enzimática em estado estacionário:

$$v = \frac{V_{max}[S]}{K_M + [S]} = \frac{V_{max}}{1 + \frac{K_M}{S}} \quad (\text{B.1})$$

onde v é a velocidade da reação, o K_M é a concentração de substrato que fornece uma velocidade de reação que é metade da velocidade máxima V_{max} obtida sob condições de saturação do substrato.

- Constante catalítica k_{cat}

O número de renovação da enzima é determinado medindo-se a velocidade de reação sob condições de $[S] \gg K_M$ para que v se aproxime de V_{max} , conforme na equação:

$$k_{cat} = \frac{V_{max}}{[E]} \quad (\text{B.2})$$

- Inibição por substrato

Acima de uma concentração crítica de substrato, os dados se desviam significativamente do comportamento esperado. A ligação da segunda molécula inibidora do substrato pode ser explicada pela seguinte equação:

$$v = \frac{V_{max}[S]}{K_M + [S](1 + \frac{[S]}{K_i})} \quad (\text{B.3})$$

para simplificar, basta dividir a parte de cima e de baixo da equação pela concentração de substrato $[S]$, assim obtemos:

$$v = \frac{V_{max}}{1 + \frac{K_M}{[S]} + \frac{[S]}{K_i}} \quad (\text{B.4})$$

onde o termo K_i nas [Equação B.3](#) e [Equação B.4](#) representam a constante de dissociação para o complexo formado pela enzima e os dois substratos envolvidos (S ES). Os efeitos de inibição em concentrações de substrato muito altas podem ser detectados como a não linearidade nos gráficos de dados com os modelos de *Lineweaver—Burk* ou *Hanes-Woolf* [Equação B.5](#). Desta forma, se observa uma dramática curvatura dos dados perto da interceptação no eixo y .

- Linearização de Hanes-Woolf

Este tratamento leva a gráficos lineares quando $\frac{[S]}{v}$ é plotado em função de $[S]$:

$$\frac{[S]}{v} = [S] \left(\frac{1}{V_{max}} \right) + \frac{K_M}{V_{max}} \quad (\text{B.5})$$

Conhecido como gráfico de *Hanes-Woolf*. Neste gráfico, a inclinação é $\frac{1}{V_{max}}$, o intercepto y é $\frac{K_M}{V_{max}}$ e o intercepto x é $-K_M$.

- Reação no equilíbrio

$$\left(\frac{k_{cat}}{K_M} \right)_f [E][S] = \left(\frac{k_{cat}}{K_M} \right)_r [E][P] \quad (\text{B.6})$$

- Energia livre

$$\Delta G_{ES^\ddagger} = -RT \ln \left(\frac{k_{cat}}{K_M} \right) + RT \ln \left(\frac{k_B T}{h} \right) \quad (\text{B.7})$$

- Variação de energia livre

$$\Delta \Delta G_{ES^\ddagger} = -RT \ln \left[\frac{(k_{cat}/K_M)^1}{(k_{cat}/K_M)^2} \right] \quad (\text{B.8})$$

APÊNDICE C – TESTES DE PURIFICAÇÃO CTBGL1A

C.1 Protocolos que não funcionaram neste caso

Após a lise celular, conforme descrita no [seção 3.2](#), a amostra de *CtBgl1A* foi submetida aos testes de purificação, utilizando-se resina cromatográfica contendo íon níquel Ni Sepharose da GE Healthcare®, pelo seguinte protocolo: lavagem da coluna com 20 VC (volumes de coluna) de água *mili-Q*, equilíbrio da mesma com 10 VC (volume de coluna), em tampão Tris (50 mM, pH 7,5; 200 mM de NaCl; 10 mM de imidazol e 10% de glicerol), lavagens com 5 VC no mesmo tampão, com concentração crescente de imidazol, iniciando em 10; 20; 30; 60; 70; 80; 90; 100; 150 e 200 mM, respectivamente. A eluição da enzima alvo pode ser observada nas frações do *Flow-through* (FT) 10, 20, 30 e 60 mM de imidazol, o que significa que a proteína não interagiu suficientemente com a resina da coluna, o que possibilitaria sua purificação. A limpeza da coluna foi feita com 200 mM de imidazol. As alíquotas do teste de purificação foram analisadas utilizando-se gel de poliacrilamida 10% SDS-PAGE e corados com *Comassie Blue* para averiguação da migração das bandas, como apresentado na [Figura 35a](#).

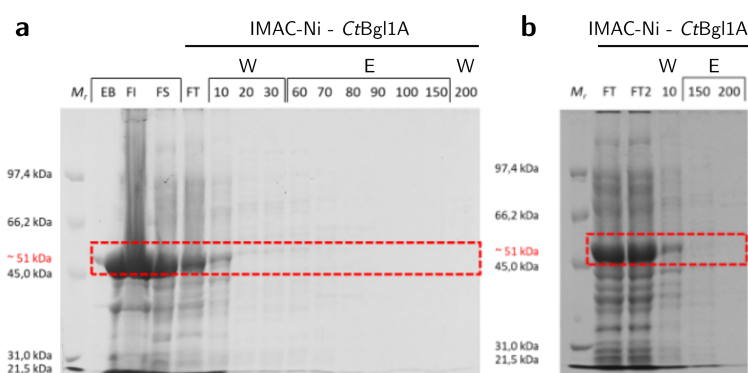


Figura 35 – **Resultado da IMAC para *CtBgl1A***. **a.** Teste de purificação por coluna de afinidade IMAC Ni-NTA da proteína *CtBgl1A*. **b.** Aplicação das amostras após coluna de afinidade Ni-NTA com tampão modificado pela adição de 500 mM de ureia. M_r - Marcador de referência; EB - Extrato bruto; FI - Fração insolúvel; FS - Fração solúvel; W (X) - Fração de limpeza com "X"mM de imidazol; FT - *Flow through*; FT2 - *Flow through* na condição de tampão com ureia; E - frações de eluição com "X"mM de imidazol. Destaque em vermelho para faixa referente à proteína de interesse (51 kDa).

Fonte: Elaborada pelo autor.

É possível observar levemente a presença de bandas nas frações de 10, 20 e 30 mM de imidazol, que seriam apenas para limpeza da coluna, constatando algum problema relacionado a *6xHis-tag* que possivelmente não esteve exposta nas condições de tampão

utilizadas. O que torna necessário a otimização do processo, provavelmente modificando-se o tampão. Considerando que os testes foram realizados utilizando o tampão mais recorrente para esta proteína e por métodos semelhantes aos já reportados com sucesso (35, 229).

Diante do problema nas primeiras tentativas de purificação, tentou-se refazer a passagem pela coluna de afinidade, desta vez modificando o tampão adicionando 500 mM de ureia para uma condição de desnaturação, em seguida deixando a proteína e resina incubadas, com agitação por 1 h à 4 °C, tentando favorecer a interação com a resina. Conforme na [Figura 35b](#).

Da mesma forma que anteriormente a proteína não interagiu suficientemente com a resina impossibilitando sua purificação. Provavelmente pelo mesmo motivo da não disponibilidade da *6xHis-tag* para interagir com os íons de Ni. Na sequência após as fracassadas tentativas de purificação, considerando que a proteína *CtBgl1A* não interagiu como desejado com a coluna de afinidade, foi feita a tentativa de purificação por cromatografia em coluna de troca iônica (Mono Q™ 5/50 GL) no HPLC Akta™. Onde, neste caso o *pellet* celular passou pela lise utilizando o tampão sem adição de NaCl, visto que o cloreto de sódio foi eluído em gradiente de concentração com a injeção da amostra. O resultado desta etapa é apresentado na [Figura 36a](#).

Observou-se um pico mais acentuado nas frações dos tubos de 3–14 (separando os 3, 6, 9, 12, 13 e 14 para aplicação no gel), outros picos surgiram no cromatograma e também tiveram suas frações separadas aplicadas em gel de poliacrilamida 10% SDS-PAGE e observou-se que a proteína de interesse foi eluída em praticamente todas as frações, significando também baixa afinidade de interação com esta coluna nas condições descritas, tal como observado na [Figura 36b–d](#).

Contudo, as frações dos tubos 6, 9, 13, e 14 apresentaram a melhor possibilidade de purificação, até o momento, para serem injetadas em coluna de exclusão por massa molecular (SEC), considerando que estas frações saíram com menores quantidades de outras proteínas, enquanto nas demais, boa parte delas foram eluídas ainda que bandas da proteína de interesse estejam presentes, o que pode significar também que houve dificuldade para interagir com a resina desta coluna.

Após juntar as frações dos tubos 6, 9, 13 e 14, estas foram concentradas para aplicação em coluna de exclusão por massa molecular (Superdex™ 75 10/300 GL), o cromatograma e géis das frações selecionadas resultantes desta etapa podem ser observados na [Figura 37a](#).

Observa-se o acentuado pico no cromatograma referente a proteína de interesse que teve eluição nas frações coletadas nos tubos de 7–10 e suas bandas podem ser vistas no gel ainda que discretas evidenciando sua presença. Até o momento esta foi a mais promissora forma de purificação pré-clivagem da *CtBgl1A*. Contudo, é possível notar a perda de

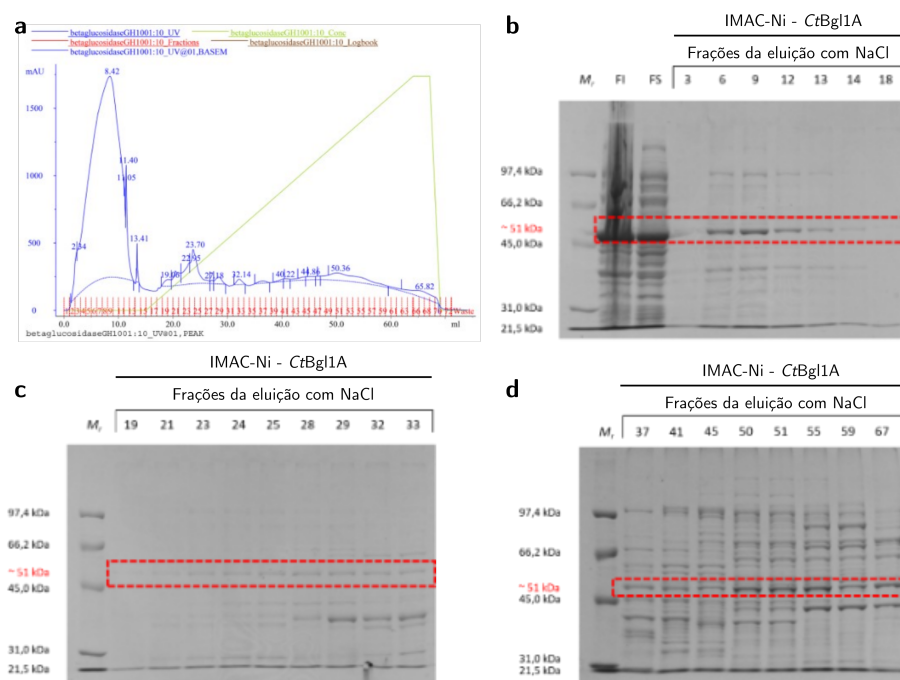


Figura 36 – **Resultado da cromatografia de troca iônica para *CtBgl1A*.** **a.** Cromatograma da proteína *CtBgl1A* em coluna de troca iônica (Mono Q™ 5/50 GL). **b–d.** Géis de poliacrilamida 10% SDS-PAGE com as amostras: M_r - Marcador de referência; FI - Fração insolúvel; FS - Fração solúvel; demais números são referentes aos tubos das frações coletadas após passar pela coluna de troca iônica. Destaque em vermelho para faixa referente à proteína de interesse (51 kDa).

Fonte: Elaborada pelo autor.

proteína de interesse devido ao número de etapas, sendo ainda necessária a realização da clivagem com trombina e conseqüentemente pelo menos mais uma etapa de purificação por exclusão de massa molecular.

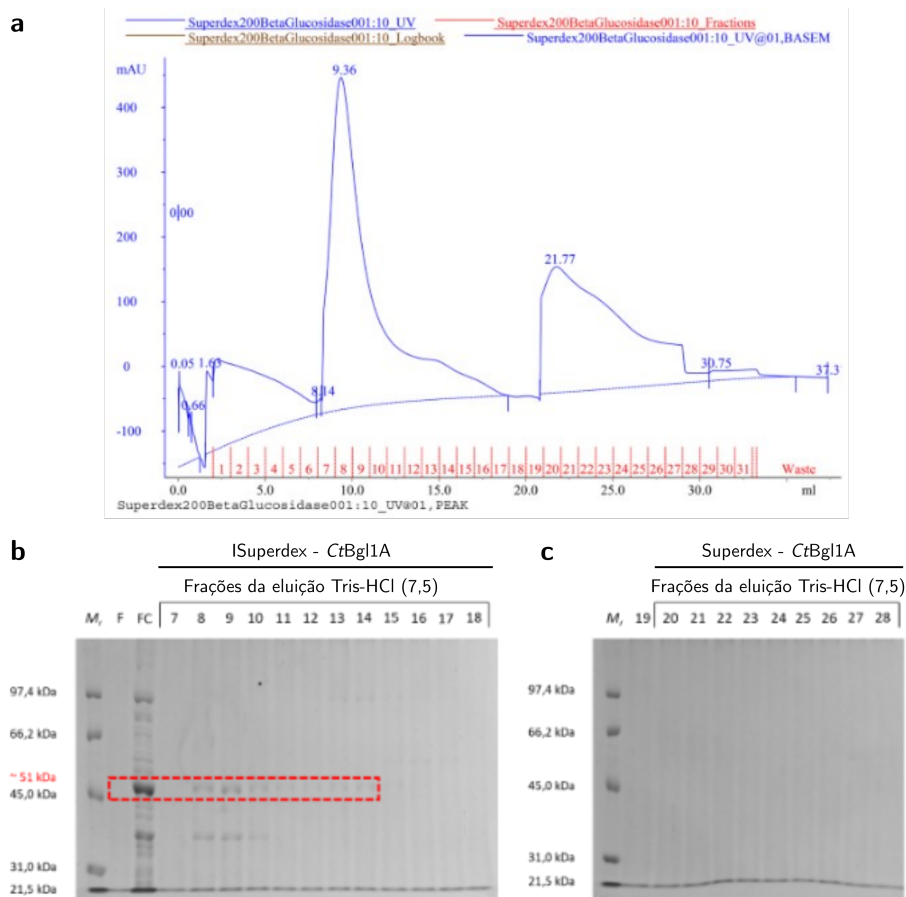


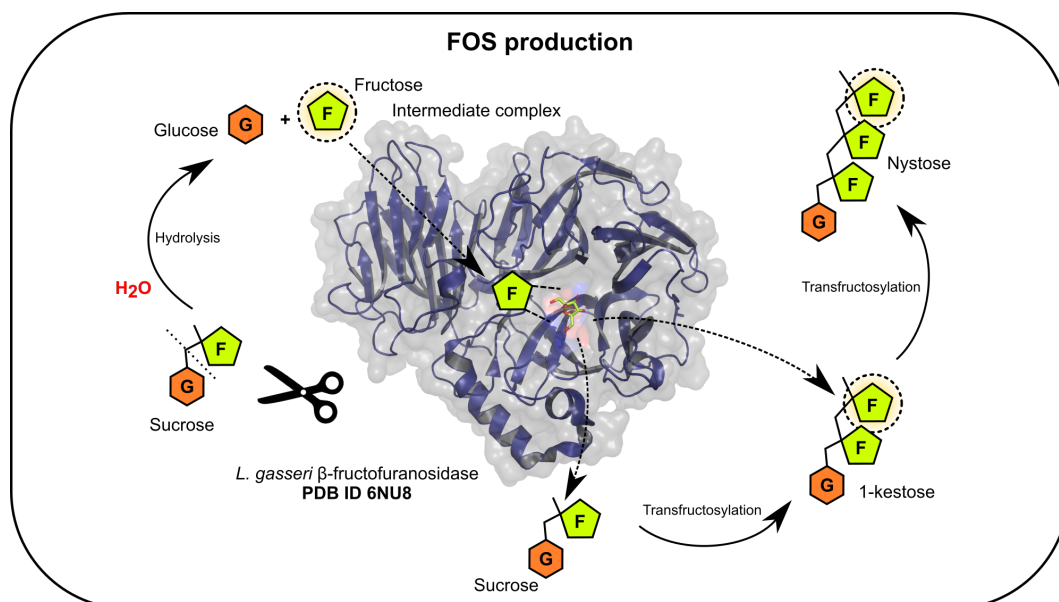
Figura 37 – **Resultado da cromatografia de exclusão por massa molecular para CtBgl1A.** **a.** Cromatograma da proteína CtBgl1A em coluna de exclusão por massa (Superdex™ 75 10/300 GL). **b–c.** Géis de poliacrilamida 10% SDS-PAGE de: F - Filtrado; FC - Fração concentrada 7–28 frações coletadas. Destaque em vermelho para faixa referente à proteína de interesse (51 kDa).

Fonte: Elaborada pelo autor.

ANEXOS

ANEXO A – ARTIGOS PUBLICADOS

A.1 Crystal Structure of a Sucrose-6-phosphate Hydrolase from *Lactobacillus gasseri* with Potential Applications in Fructan Production and the Food Industry



A partir da clonagem, expressão heteróloga, purificação, obtenção da estrutura de uma fructosilfrutofuranosidase do microorganismo homofermentativo *Lactobacillus gasseri* que habita o trato gastrointestinal de humanos e exerce uma função probiótica, especialmente por meio da síntese de de frutanos prebióticos derivados da sacarose e seus hidrolizados, nós realizamos sua caracterização bioquímica e estrutural para compreender seu mecanismo e condições ótimas de ação. Destacando-se, sua capacidade tanto para hidrólise da sacarose, como também a formação de fruto-oligosacarídeos como a 1-kestose e nistose. Associando-se, também as análises biofísicas de SAXS ao comparar o efeito do pH aos estados oligoméricos de dímeros, com diferentes interfaces de interação, observados dentro da família GH32.

Crystal Structure of a Sucrose-6-phosphate Hydrolase from *Lactobacillus gasseri* with Potential Applications in Fructan Production and the Food Industry

Mariana Z. T. de Lima,[#] Leonardo R. de Almeida,[#] Alain M. Mera, Amanda Bernardes, Wanius Garcia, and João R. C. Muniz*



Cite This: *J. Agric. Food Chem.* 2021, 69, 10223–10234



Read Online

ACCESS |



Metrics & More



Article Recommendations



Supporting Information

ABSTRACT: Fructooligosaccharides (FOSs) are polymers of fructose with a prebiotic activity because of their production and fermentation by bacteria that inhabit the gastrointestinal tract and are widely used in the industry and new functional foods. *Lactobacillus gasseri* stands out as an important homofermentative microorganism related to FOS production, and its potential applications in the industry are undeniable. In this study, we report the production and characterization of a sucrose-6-phosphate hydrolase from *L. gasseri* belonging to the GH32 family. Apo-LgAs32 and LgAs32 complexed with β -D-fructose structures were determined at a resolution of 1.94 and 1.84 Å, respectively. The production of FOS, fructans, 1-kestose, and nystose by the recombinant LgAs32, using sucrose as a substrate, shown in this study is very promising. When compared to its homologous enzyme from *Lactobacillus reuteri*, the production of 1-kestose by LgAs32 is increased; thus, LgAs32 can be considered as an alternative in fructan production and other industrial applications.

KEYWORDS: fructosyltransferases, fructooligosaccharides, *Lactobacillus gasseri*, LgAs32, fructan, food industry

1. INTRODUCTION

Lactobacillus gasseri is a homofermentative microorganism that inhabits the gastrointestinal tract of humans and exerts a probiotic function, especially due to the synthesis of prebiotic fructans (polymers of fructose) derived from sucrose and its hydrolyzates.^{1,2} To date, only a few potential genes belonging to this organism related to the production of FOS have been characterized.³ Fructans have prebiotic activities by stimulating the multiplication of beneficial bacteria in the human gastrointestinal tract related to several health benefits.^{4,5} The production of different fructans is associated with the physiological needs of each class of organism,⁶ and the most beneficial fructans to the human body have a degree of polymerization (DP) between DP2 and DP6, depending on the amount of fructose associated, attributable to their biomodulator characteristics.^{7,8} Fructans or fructooligosaccharides can also be obtained through reactions catalyzed by the enzymes found in fungi, bacteria, or plants named fructansucrases or fructosyltransferases, besides chemical synthesis.⁹ These enzymes belong to families 32 and 68 of glycoside hydrolases according to the CAZy database (www.cazy.org).

The advances in the development of functional foods are especially related to the progress in intestinal microbiota research.¹⁰ Bioactive oligosaccharides have attracted attention because of their many applications in the food and agro-industry, lowering their costs and expanding their production.^{3,9,11} As a result, prebiotics emerged as an alternative to stimulate the growth and/or activity of beneficial bacteria that play an important role in the human body. The use of

prebiotics as food resources for functional purposes has been severely hampered by the difficulty of maintaining their active conditions in the market for a long period given the requirements for temperature regulation and storage.

Several species of *Lactobacillus* have their natural habitat in the gastrointestinal tract, and many studies have associated fructosyltransferases with the pool of enzymes produced by the microorganism that generates or hydrolyzes fructans.^{12–14} The importance of these findings is undeniable, and the present work aims to contribute to the increase of information related to these enzymes, as technological and probiotic properties are still limited.^{2,15} Studies of isolated enzymes can lead to an improvement and better condition control rather than working with a pool of enzymes.

In this study, we report for the first time the cloning, expression, purification, crystallization, and crystal structure determination of a recombinant sucrose-6-phosphate hydrolase (β -fructofuranosidase, E.C. 3.2.1.26)¹⁶ from *L. gasseri* (DSM 20243/ATCC 33323/LGAS_1779 gene) belonging to the GH32 family (here named LgAs32). As previously reported by our research group, this class of enzymes can be found in many organisms.¹⁶ Crystal structures of apo-LgAs32 and LgAs32 in complex with β -D-fructose were determined at high resolution.

Received: June 29, 2021

Revised: August 13, 2021

Accepted: August 18, 2021

Published: August 27, 2021



LgAs32 showed hydrolytic activity in the presence of sucrose and the production of fructooligosaccharides nystose and 1-kestose, as shown by HPAEC/PAD and spectrometry assays by transfructosylation activity. Therefore, the present study reveals an enzyme with potential for fructan production and capable of improving biotechnological processes.^{1,5}

2. MATERIALS AND METHODS

2.1. Chemicals. Imidazole, kanamycin, chloramphenicol, 2-amino-2-(hydroxymethyl)propane-1,3-diol (Tris), HCl, NaCl, $(\text{NH}_4)_2\text{SO}_4$, KH_2PO_4 , Na_2HPO_4 , glycerol, glucose, lactose, tryptone, yeast extract, 1,4 dithiothreitol (DTT), phenylmethanesulfonyl fluoride (PMSF), lysozyme, and protein standards used as sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) markers were purchased from Sigma-Aldrich. Ni Sepharose 6 Fast Flow was purchased from GE Healthcare. All chemicals and reagents used in this study were of the highest purity analytical grade.

2.2. Cloning, Recombinant Expression, and Purification of LgAs32. The gene LGAS_1779 (UniProt: A0A133PGM1) was cloned using the Ligase Independent Cloning (LIC)^{17,18} in a modified pET-Trx vector, where the use of restriction or ligase enzymes is not required. A thioredoxin domain fused to a His-tag was also added to the target protein to enhance solubility and facilitate protein purification by metal affinity chromatography.¹⁹

For heterologous protein expression, the *E. coli* strain Rosetta was used for overexpression of the cloned gene. Cells containing the vector were aerobically cultured at 37 °C in an Auto Induction broth [3.3 g L⁻¹ $(\text{NH}_4)_2\text{SO}_4$, 6.8 g L⁻¹ KH_2PO_4 , 7.1 g L⁻¹ Na_2HPO_4 , 5 g L⁻¹ glycerol, 0.5 g L⁻¹ glucose, 2 g L⁻¹ lactose, 10 g L⁻¹ tryptone, and 5 g L⁻¹ yeast extract] supplemented with 50 $\mu\text{g mL}^{-1}$ kanamycin and 200 $\mu\text{g mL}^{-1}$ chloramphenicol (final concentration) and agitation at 150 rpm. After the absorbance at 600 nm reached values between 0.5 and 0.7, the temperature was reduced to 17 °C and the culture was kept under agitation at 150 rpm for 19 h. Cells were then harvested by centrifugation (7000g, 20 min, 4 °C) and resuspended in buffer A (50 mM Tris-HCl pH 7.5, 10% glycerol, 5 mM imidazole, 50 mM NaCl, 1 mM DTT, 0.1 mM PMSF, and 200 $\mu\text{g mL}^{-1}$ lysozyme). The suspension was subjected to sonication for cell disruption in an ice bath for 5 min, alternating between sonication and rest cycles of 30 s each, using a 550 Sonic Dismembrator Sonifier (Fisher Scientific) and then centrifuged at 40,000g for 20 min at 4 °C.

After sonication, the recombinant protein was found in the soluble fraction, which was subjected to immobilized metal affinity chromatography (IMAC) using a His-trap HP column (GE Healthcare) equilibrated with buffer B (50 mM Tris-HCl pH 7.5, 500 mM NaCl, and 10 mM imidazole). The same buffer was used to wash unbound proteins from the resin column. The protein was then eluted with an imidazole gradient (30–150 mM) applying a gradual increase of buffer B. The fraction containing the purified tagged protein was diluted four times in buffer comprising 50 mM Tris-HCl pH 7.5 and TEV (*tobacco etch virus*) protease²⁰ (1 mg of TEV for each 50 mg of protein)²¹ for the thioredoxin-tag cleavage at 4 °C overnight (16 h). The cleaved fusion protein was loaded again onto an equilibrated (50 mM Tris-HCl pH 7.5 and 20 mM imidazole) His-trap HP column (GE Healthcare) and eluted with the buffer without imidazole. All purification procedures were performed at 4 °C or in an ice bath. A final step of purification was performed employing the FPLC chromatographic system (GE Healthcare, USA) connected to a Superdex 200 Hiload 16/60 (GE Healthcare, USA) size-exclusion chromatography column with 50 mM Tris-HCl pH 7.5 containing 150 mM NaCl as the running buffer. The portion having the purified LgAs32 from *L. gasserii* was concentrated by centrifugation using an Amicon (Millipore, 30 kDa cutoff). All the eluted samples were analyzed at 15% (w/w) SDS-PAGE. After the purification steps, LgAs32 presented high purity and the expected molecular mass value under denaturing conditions.

Protein concentration was determined by measuring the absorbance at 280 nm using a NanoDrop spectrophotometer (Thermo Scientific) and the theoretical extinction coefficient based

on the amino acid composition ($\epsilon_{280\text{nm}} = 100,075 \text{ M}^{-1} \text{ cm}^{-1}$). The oligomeric state of purified LgAs32 was inspected using a native PhastGel with an 8–25% gradient, stained with Coomassie brilliant blue R-250 (Bio-Rad Laboratories), and discolored with 10% acetic acid (data not shown).

2.3. Dynamic Light Scattering (DLS). The dimensions of purified LgAs32 were examined using the dynamic light scattering method (Malvern Instruments Ltd., Malvern, UK).²² A protein solution (1 mg mL⁻¹) in a 25 mM acetate–borate–phosphate buffer adjusted at pH 6 containing 150 mM NaCl was loaded into a quartz cuvette before measurement. For thermostability analyses, the temperature was raised from 25 to 65 °C and the sample was sat for 5 min at each temperature prior to DLS measurements. Hydrodynamic radii (R_h) were determined from a second-order cumulant fit to the intensity autocorrelation function (size distribution by volume).

2.4. Crystallization, Data Collection, and Structure Determination. Crystals of *apo*-LgAs32 grew in 7 days in SWISSCI plates at 18 °C using the vapor diffusion method. Optimized crystals of a suitable size were achieved by mixing the protein at 26 mg mL⁻¹ with 25% (w/v) PEG 1500 and 0.1 M MMT buffer at pH 9.0 in a 2:1 ratio (protein/buffer). The diffraction data of the *apo*-LgAs32 crystals were collected at cryogenic temperatures (100 K) in a nitrogen vapor stream on the Bruker APEX DUO diffractometer equipment with an APEX II CCD detector using a wavelength of 1.5418 Å. Strategies to collect X-ray diffraction data and subsequent determinations of unit cell parameter (a , b , c , α , β , and γ) indexing and integration of the set of diffraction images were carried out using SIBAS and SAINT (Bruker, 2012, Bruker AXS Inc., Madison, Wisconsin, USA). The LgAs32 diffraction data in complex with β -D-fructose, a hydrolysis product obtained by soaking the *apo* crystal in sucrose (LgAs32:FRU), were collected under cryogenic nitrogen stream (100 K) at the National Synchrotron Light Laboratory (LNLS, Brazil) on the MX-2 beamline. Strategies for X-ray diffraction data collection, indexing, and integrating the set of diffraction were performed with iMOSFLM.²³ Data reduction and scaling were done with Truncate²⁴ and AIMLESS^{25,26} for both data sets (*apo*-LgAs32 and LgAs32:FRU). The number of molecules per asymmetric unit and the percentage of the solvent were estimated by applying Matthews' coeff.²⁷ The phases for LgAs32 were calculated by molecular replacement using PHASER,²⁸ with the three-dimensional structure of β -fructosidase (invertase) from *Thermotoga maritima* (PDB ID: 1UYP) sharing 30% sequential identity as a search model. The initial crystallographic model was built by ARP/wARP.²⁹ The phases for LgAs32:FRU were calculated by adopting the *apo* structure as a search model in PHASER. The structural model for *apo*-LgAs32 was refined using PHENIX_refine,³⁰ and the crystallographic model for LgAs32:FRU was refined with REFMACS.³¹ Intermittent manual rebuilding and refinement of individual B-factors after applying a TLS correction, alternating with inspection cycles and manual reconstruction of the real space through graphical visualization, were done using COOT.³² The qualities of the generated models were analyzed using MolProbity.³³ The summary of data collection, phasing, and refinement statistics is described in Table S2. The MOLE toolkit^{34,35} was applied to estimate cavities and substrate entry pathways. All structural images were drawn using the PyMOL Molecular Graphics System software.³⁶

2.5. Small-Angle X-ray Scattering (SAXS). SAXS measurements were performed at the SAXS beamline (LNLS, Brazil). LgAs32 was measured at different concentrations (at 25 °C) in the 25 mM acetate–borate–phosphate buffer containing 150 mM NaCl and adjusted at pH 6.0, 7.5, and 8.5. The wavelength of the X-ray beam was 1.48 Å, and the sample-to-detector distance was set at 0.95 m. X-ray patterns were recorded using a two-dimensional CCD detector (MarResearch, USA). The integration of SAXS patterns was done with the FIT2D program (www.esrf.eu/computing/scientific/FIT2D). The scattering of water measured on the same sample cell was employed to normalize the data on an absolute scale. The GNOM program was used for the determination of the radii of gyration (R_g)³⁷ (also calculated using Guinier approximation),

evaluation of the distance distribution function $P(r)$, and the maximum diameter (D_{\max}) calculation. Dummy atom models (DAMs) were calculated from the experimental SAXS using *ab initio* procedures implemented in the DAMMIN program.³⁸ The CRYSOLOG program³⁹ was applied to generate theoretical scattering curves from the three-dimensional crystallographic structures. The PISA program⁴⁰ was used to search for biological assemblies for the monomeric crystal structure.

2.6. Enzymatic Activity Measurements. The enzymatic activity of the purified LgAs32 was measured using 500 mM sucrose at 37 °C and 20 nM of the enzyme according to the previously reported method.⁴¹ A standard enzymatic assay for optimum pH and temperature determination was performed using sucrose as the substrate. Fifty microliters of the buffer solution adjusted at pH values from 2.0 to 10.0 (20 mM sodium phosphate/sodium citrate/glycine buffer), 40 μ L of 500 mM sucrose as substrate and enzyme at 20 nM. Following incubation at 37 °C for 15 min, the reaction was stopped by adding 100 μ L of DNS (3,5-dinitrosalicylic acid) and the total reaction volume (500 μ L) was boiled at 100 °C for 5 min. For the determination of the optimum temperature, the reaction was carried out within the buffer at the best pH previously established and a temperature ranging from 20 to 80 °C with 15 min of incubation. One hundred microliters of DNS was added, and the reaction was stopped. All experiments were done in triplicate, monitored colorimetrically at 540 nm,⁴² and the average values are reported.

For kinetic parameter determination, the end-point velocities of the enzyme were measured in 40 mM sodium citrate pH 5.5 at 50 °C and 20 nM of the enzyme using sucrose in concentrations ranging from 0.25 to 500 mM. The volume of all samples was adjusted to 100 μ L. For sucrose hydrolysis, the amount of glucose released was measured using the DNS method by measuring the absorbances at 540 nm. Kinetic parameters (K_M and k_{cat}) were calculated by fitting the Michaelis–Menten model using the nonlinear least-squares regression analysis with the Origin software (version 2018). All assays were carried out in triplicate, and the average values were reported.

2.7. Thermal Assays. Enzyme thermostability was analyzed in the presence and absence of sucrose. The experiment consisted of incubating for 5, 10, 30, and 90 min and 24 and 48 h at temperatures between 40 and 70 °C until it reaches its optimum pH and temperature. After incubation, 500 mM sucrose (final concentration) was added for each sample and the reaction was performed at 50 °C for 5 min. The reaction was also stopped by adding 100 μ L of DNS and monitored colorimetrically at 540 nm. All experiments were performed in triplicate, and the average values are reported.

Differential scanning fluorimetry (DSF) was evaluated as a method to detect a buffer composition that increases the structural integrity and homogeneity of the protein in the solution. The assay was performed based on previous reports,^{43,44} and 48 screen conditions were applied, which comprised 13 different buffers, covering a pH range from 1.2 to 10.0, to test the effects of the buffer and pH on protein stability. The composition of the screen is given in Table S1. DSF was carried out in a real-time thermocycler (CFX96 Real-Time System, Bio-Rad) with a 96-well thin-wall PCR plate and optical-quality sealing tape (Bio-Rad). The fluorescence emission intensity was initially measured at 25 °C, and then the temperature was increased to 90 °C with a stepwise increment of 1 °C min⁻¹. The fluorescence intensity was monitored with the wavelengths for excitation and emission set at 490 and 580 nm, respectively. The protein was added to each well at 7 μ M as a final concentration, 100 mM of each buffer, and the fluorescent dye SYPRO Orange (1:1000) (Molecular Probes). Data analysis and determination of the temperature derivative of the melting curve (T_m) were performed using the GraphPad Prism software (version 5.0).

Circular dichroism (CD) spectra were measured using a Jasco J-815 spectropolarimeter. The LgAs32 concentration was 0.25 mg mL⁻¹ in the 25 mM acetate–borate–phosphate buffer adjusted at pH 6 containing 150 mM NaCl. The spectra were collected at 25 °C using a quartz cuvette over the wavelength range of 195 to 260 nm. Eight measures were averaged to form each CD spectrum, using a scanning speed of 100 nm min⁻¹, a spectral bandwidth of 1 nm, and a response

time of 0.5 s, and obtained on a degree scale. The spectral contribution of the buffer solution was subtracted, and then the spectrum was transformed to molar ellipticity [θ] using the mean weight residue and protein concentration.

2.8. Carbohydrate Analysis Using High-Performance Anion-Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD). The protein was diluted against a stable buffer (from the DSF experiment) described as buffer C (sodium citrate pH 5.5). In buffer C, the protein was diluted to a final concentration of 20 nM and the experiment was held at 50 °C. Distinct times were selected for investigation (5 min, 8 h, 10 h, 12 h, and 24 h) for different sucrose concentrations (50, 75, 100, 125, 150, 200, 250, 300, 350, 400, 450, 500, 750, 1000, and 2000 mM). The flask was boiled in a hot-water bath for 10 min to stop the reaction. Each reaction was diluted (10 μ L of sample and 990 μ L of ultrapure water) and filtered using a nitrocellulose membrane (0.45 μ m) (Millipore). Samples were analyzed by HPAEC/PAD in a Dionex system ICS-3000 (USA) with a Carbo Pack PA 1 column (4 \times 250 mm). The elution was carried out with a 150 mM sodium hydroxide gradient (eluent A) and 500 mM sodium acetate (eluent B) as previously reported.^{42,45} The elution time for each diluted sample was 40 min at a rate of 0.250 mL min⁻¹. The peaks were identified using sugar standards (fructose, glucose, sucrose, 1-kestose, and nystose from Sigma-Aldrich and neokestose from Dr. Norio Shiomi, Japan) in distinct concentrations (25, 50, 100, and 200 μ g mL⁻¹), considering the concentrations and the sugar peak area.

2.9. Mass Spectrometry Analyses. After the best time of incubation and sucrose concentration were determined, the products of the reaction were analyzed using mass spectrometry. One microliter of the reactional product from LgAs32 (20 nM) and sucrose (300 and 2000 mM) incubated at different times (8 and 10 h) was mixed separately with 1 μ L of the DHB matrix (dihydroxybenzoic acid) at a concentration of 10 mg mL⁻¹ diluted in 30% acetonitrile to allow it to dry above the superficial plate.⁴⁶ Sucrose, 1-kestose, and nystose were used as standards in the same matrix. Samples were analyzed using a Microflex LT MALDI-TOF Mass Spectrometer (Bruker Daltonics, USA). Measures were performed using the positive-ion linear mode in a range of 300–720 m/z , with laser power of 61 and 600 shots in the same spot. The molecular mass average was processed using flexAnalysis (Bruker Daltonics, USA).

3. RESULTS AND DISCUSSION

3.1. Cloning, Recombinant Expression, and Purification of LgAs32. After cloning and expression in *E. coli* Rosetta, the recombinant sucrose-6-phosphate hydrolase from *L. gasseri* (LgAs32) fused to a thioredoxin-His₆ tag was found in the supernatant after lysis. The recombinant fusion protein (~69 kDa) was then purified by metal affinity chromatography (Figure S1, lane 4). Subsequently, the fusion protein was cleaved with the TEV protease (Figure S1, lane 5). Finally, LgAs32 (~55 kDa) and the thioredoxin-His₆-tag (~14 kDa) were separated by size-exclusion chromatography (Figure S1, lane 6). Upon SDS-PAGE under denaturing conditions, the purified LgAs32 migrated as a single homogeneous band with the expected molecular mass of ~55 kDa for a monomer (Figure S1, lane 6), and the final enzyme yield was typically 25 mg L⁻¹ of culture.

3.2. Structural Analysis of the Recombinant LgAs32. Circular dichroism (CD) spectroscopy was used to analyze the secondary structure of purified LgAs32 at pH 6 and 25 °C. The CD spectrum of the LgAs32 (Figure S2A) is characterized by a broad minimum centered about 215 nm, showing a high content of β -sheet secondary structure. The small positive band around 237 nm is characteristic of an aromatic-rich protein (LgAs32 structure has 10 tryptophan, 30 tyrosine, and 17 phenylalanine residues). When LgAs32 at pH 6 and 25 °C was analyzed by DLS in solution, the observed profile was

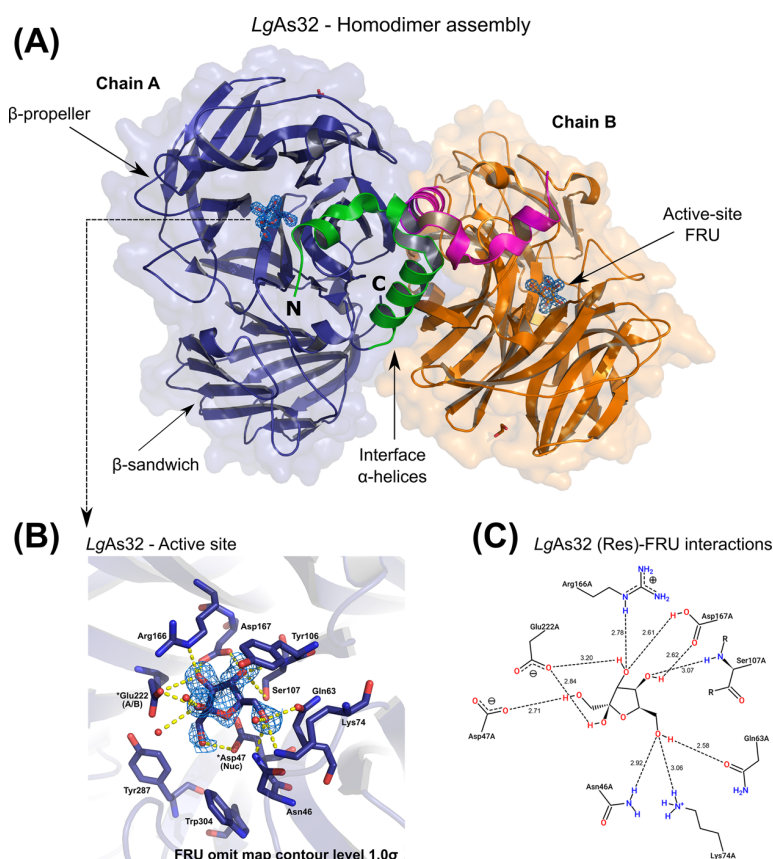


Figure 1. Structural organization and active site of β -fructofuranosidase from *Lactobacillus gasseri*. (A) Homodimer of *LgAs32* predicted by the PISA program and experimentally verified by SAXS (Section 3.4). The protruding α -helix responsible for maintaining the homodimeric oligomeric state is highlighted in green and magenta for chains A and B, respectively. (B) A detailed view of the electron density map for the fructose molecule bound in the active site is shown in both chains, with an omit map contour level of 1.0σ , occupancy = 1.00, and an average B-factor of 20.8 \AA^2 (PDB: 6NU8). (C) A two-dimensional diagram made with PoseView⁹³ showing the hydrogen bonds with the number of residues interacting with fructose.

typical of a monodisperse solution (inset of Figure S2B). The R_s value determined for *LgAs32* at pH 6 was 4.5 ± 0.3 nm, which corresponds to a molecular mass of 114 ± 18 kDa, consistent with the expected molecular mass of a homodimer in solution (~ 55 kDa for a monomer). The homodimer oligomerization state has been previously described in other members of this GH32 family and is strictly related to its activity and stability under a wide range of pH conditions.¹⁶

3.3. Crystal Structure of *LgAs32*. Despite sharing on average of only 30% of sequence identity with other enzymes of the same family, *LgAs32* has a conserved three-dimensional folding with the root-mean-square deviation (rmsd) ranging from 2.690 \AA to the closest structure from *Termotoga maritima* *TmINV*⁴⁷ (PDB ID 1UYP; 424 atoms aligned) to 6.061 \AA to the most distant from *Bifidobacterium adolescentis* *BaFfase* (PDB ID 6NOB; 288 atoms aligned), respectively, and considering only bacterial structures. Several bacterial structures for this family have already been described.^{47–50} In short, the enzyme shares a common fold within members of the GH32 family^{47,48,51–53} and also with the GH68 levansucrase family of proteins (both belonging to Clan-J of GHs),^{54–56} which comprises two domains: an N-terminal five-bladed β -propeller domain (residues 37–340) with each blade composed of four antiparallel β -strands that include the catalytic site creating a “funnel-like” shape channel and a β -

sandwich C-terminal domain (residues 336–486) composed of two six-stranded antiparallel β -sheets forming a β -sandwich architecture; see Figure 1A. Contrary to most representative structures available in the PDB belonging to the GH32 family, the *LgAs32* structure has a bent α -helix at its N-terminal (from Glu2 to Gln26), making a protuberance in blades I and II next to the structure dimerization interface (Figure 1A).

The catalytic site is at the bottom of the β -propeller cavity with a “funnel-like” opening, characteristic of enzymes that can hydrolyze fructose polymers with high degrees of polymerization, like inulin.⁵¹ The coordination at the catalytic site is formed by Asp47 and Glu222, on the first strand of blades I and IV (Figure 1B,C), which have a carbonyl oxygen distance of approximately 6.5 \AA that provides hydrolysis of the sucrose substrate with the retaining configuration and in an analogous location for the terminal fructosyl moiety at the -1 subsite.^{47,48,57–59} The other residues directly involved in polar character interactions with the fructofuranoside ring are Asn46, Gln63, Lys74, Ser107, Arg166, and Asp167 (Figure 1B,C). A Tris molecule from the buffer solution is found in a similar position in the *apo* form of the enzyme and coordinated by Asp47, Ser107, Arg166, Asp167, and Glu222 (data not shown). The C_α rmsd between the *apo* crystal structure and the fructose complex is 0.155 \AA (488 aligned atoms). The conformation of the active site residues in both structures is

practically unchanged after fructofuranose binding. Atomic coordinates and structure factors of the *apo* and the complex with fructose (*LgAs32:FRU*) have been deposited in the *Protein Data Bank*⁶⁰ with the PDB IDs of 6NU7 and 6NU8, respectively.

A search for structures similar to *LgAs32* was performed using BLASTp⁶¹ (data not shown), and the most representative structures were compared in terms of three-dimensional folding, amino acid sequence, ligands, and enzymatic reactions (Figure 2). Selected structures were superposed against *LgAs32* (6NU8; see details in Figure 2A–G): the β -fructofuranosidase from *Bifidobacterium longum* KN29.1 (*BIF*fase) that has an elongation of 35 residues at the N-terminus containing a five-turn α -helix (PDB ID 3PIJ); the inactive *AtcwINV1*-mutant *E203Q* in complex with sucrose as

an alternate binding mode (PDB ID 2OXB);⁶² the cell-wall invertase from the *Arabidopsis thaliana* *AtcwINV1*-mutant *D239A* in complex with sucrose (PDB ID 2QQU *D239A*);⁵⁹ the invertase mutant *E190D* from *Thermotoga maritima* (*TmINV-E190D*) in complex with the trisaccharide raffinose (PDB ID 1W2T);⁶³ the β -fructofuranosidase *BmSUC1* from *Bombyx mori* (PDB IDs 7BWB and 7BWC);⁶⁴ and the fructan exohydrolase 1-FEH IIa mutant *E201Q* from *Cichorium intybus* (*Ci1-FEHIIa E201Q*) in complex with 1-kestose (PDB ID 2AEZ).⁵⁸

Despite the similarity in the overall folding and the –1 subsite structure, the different conformations are observed for four loops near the catalytic site: Loop1 (residues 126–133), Loop2 (residues 157–164), Loop3 (residues 239–255), and Loop4 (residues 304–313) (Figure 2A). Loop1 in *BIF*fase (3PIJ) is structured as a short helix inserted above the catalytic pocket; Loop2 in *AtcwINV1* (2OXB and 2QQU), *BmSUC1*, and *Ci1-FEHIIa* (2AEZ) has an absence of a regular secondary structure, while in the other structures, it corresponds to a β -strand where Asp167 is located. Great divergence because of flexibility and insertions is found at Loop3, resulting in slightly longer loops for *LgAs32*, *BIF*fase, and *BmSUC1*; finally, Loop4 in *LgAs32*, *BmSUC1*, and *TmINV* is projected into the catalytic pocket, restricting the active site cavity. These flexible loops (Figure 2A) that form the catalytic pocket confer a wide substrate specificity, and *LgAs32*, which has sucrose as the main substrate, can also interact with greater fructooligosaccharides and fructans even with lower efficiency.

From the comparison between members of the GH32 family, the catalytic dyad aspartate (Asp47) stands out as a nucleophile,⁶⁵ whereas the glutamate (Glu222) acts as the general acid/base catalyst,⁶⁶ typical for GH32. Tyrosine (Tyr287) was suggested by Verhaest *et al.* (2005)⁵² to be involved in the pK_a modulation of the equivalent to Glu222 acid/base catalyst because of its proximity, while Yuan *et al.* (2012)⁶⁷ showed that the arginine equivalent to Arg166 could be a more suitable candidate based on theoretical pK_a calculations and molecular dynamics simulations.

The multiple sequence alignment of fructosyltransferases, β -fructofuranosidases, invertases, levanases, inulinases, fructan 1-exohydrolase IIa (1-FEH IIa), and sucrose-6-phosphate hydrolases revealed the conserved RDP motif (Arg166-Asp167-Pro168 in *LgAs32* structure), which is presumed to share a functional role.^{68,69} Residues Arg166 and Asp167 take part in substrate binding and are directly involved in hydrogen bond interactions with the hydroxyl groups from fructose, important for recognition of the sugar ring toward fructosyl moieties. The substitution of Asp by an Asn on the RDP motif affects the sucrose hydrolysis but maintains the enzyme specificity.⁶⁹

From a structural superposition, there is a notable open cleft around the canonical catalytic dyad Asp47 (nucleophile) and Glu222 (general acid/base catalyst) that leads to the binding of substrate molecules (see Figure S3G), suggesting a substrate pathway related to their biological function. Thus, all residues surrounding the fructofuranosyl moiety were selected as a starting point using the MOLE 2.0 program,^{34,35} and a list of tunnels close to the binding site was generated (see Figure S3G). Six tunnels (A to F) can be observed buried inside the β -propeller cavity as an entry pathway (Figure S3A–H and Table S3). The calculated tunnels are shallow, with the exception of tunnel A, which fills the internal path between the blades of the β -propeller domain (Figure S3A), with moderate

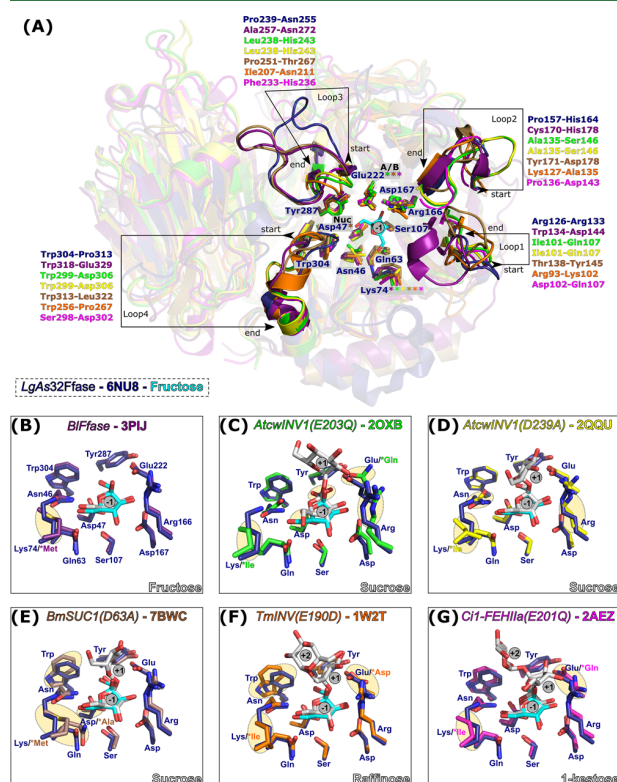


Figure 2. Structural superposition of *LgAs32* (6NU8) with representative fructosyltransferases of the GH32 family with different ligands for comparison. (A) Superposition of the fructosyltransferase structures detailing residues in the catalytic pocket and their respective conformation (*mutations in structures according to the color code and PDB ID). The distinguished difference between structures is observed in the N-terminal α -helices for *LgAs32* (6NU8 in blue), *BIF*fase (3PIJ in purple), and *BmSUC1* (7BWC in brown). Loop regions around the catalytic cleft that present insertions or flexibility are identified by their respective color scheme (*AtcwINV1*: 2OXB in green and 2QQU in yellow, *TmINV*: 1W2T in orange, and *Ci1-FEHIIa*: 2AEZ in magenta). (B–G) Catalytic pocket of selected GH32 enzymes highlighting the fructosyl moiety with retained configuration and also the different substrates that could be perfectly fitted to the same position considering the surrounding cleft, and also with an indication of the fructose, sucrose, raffinose, and 1-kestose subsites. The most relevant structural differences and mutations are shown by the light-yellow contour. Residues are represented as sticks and are labeled with a three-letter code.

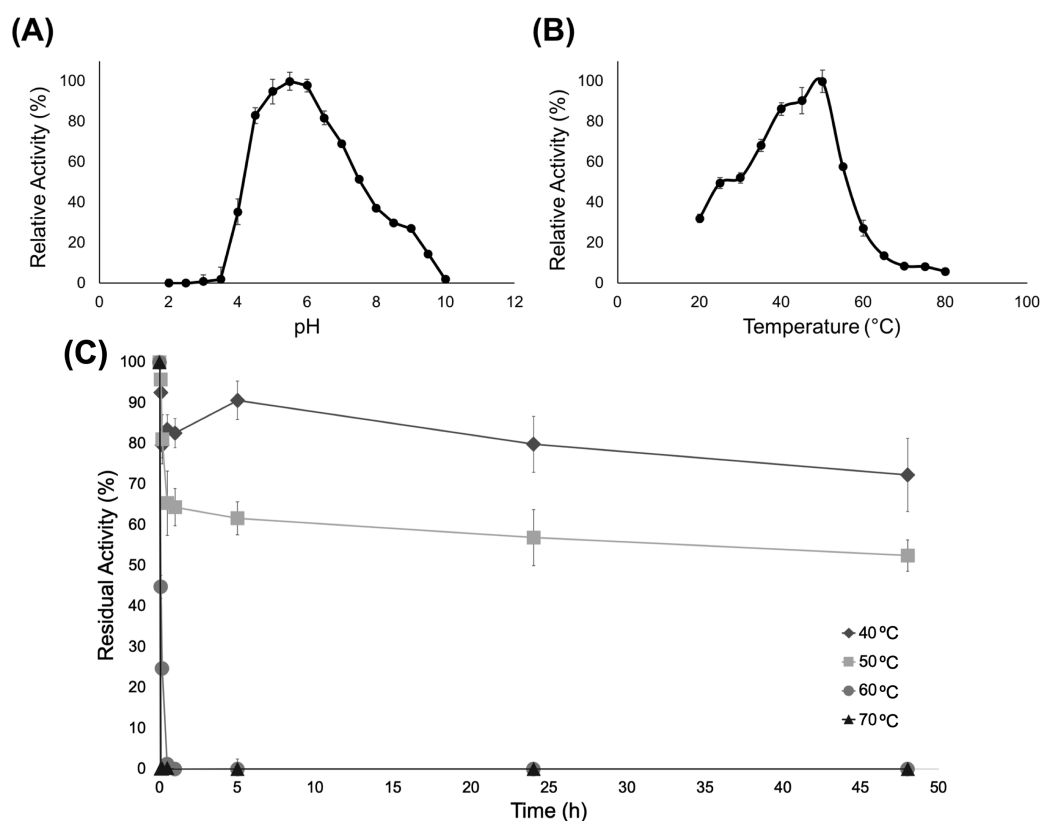


Figure 3. Effects of pH and temperature variation, and heat stability assay performed for LgAs32. (A) Effects of pH variation at 37 °C. (B) Effects of temperature variation at pH 5.5. (C) The remaining activity (%) was measured after the incubation of LgAs32 at pH 5.5 and different temperature values.

scores of hydrophilicity,⁷⁰ polarity,⁷¹ and higher relative mutability⁷² (see Table S3). The dimensions of tunnels B, C, D, and F are compatible to accommodate a molecule of sucrose as a substrate (Figure S3G,H), as observed in the LgAs32 structure (PDB ID 6NU8). Our calculations show that tunnels B, C, D, E, and F are mostly hydrophilic and can take in other molecules, e.g., sucrose, raffinose, and 1-kestose (Table S3), consistent with the literature.^{48,49,51–53,58,62,64,73–77} This allows us to consider that those tunnels could effectively represent the cleft in LgAs32, which performs more types of enzymatic reactions using sucrose as a substrate (e.g., Figure S12A,B).

3.4. LgAs32 Forms Stable Homodimers in Solution.

The SAXS analysis indicated that LgAs32 in solution is a molecule with dimensions ($R_g \sim 30$ Å and $D_{max} \sim 105$ Å) larger than the dimensions of the LgAs32 monomeric crystallographic structure ($R_g = 23.25$ Å and $D_{max} = 73.65$ Å) (Figure 1). Therefore, we searched for biological assemblies in the crystal structure of LgAs32 using the PISA program.⁴⁰ Only one stable homodimeric biological assembly was detected (Figure S4A), with the free energy barrier of assembly dissociation (ΔG^{diss}) of 5.9 kcal mol⁻¹ (*apo*-LgAs32) and 9.6 kcal mol⁻¹ (LgAs32:FRU), while the solvation free energy gains upon the formation of the assembly (ΔG_{int}) were -26.6 kcal mol⁻¹ (*apo*-LgAs32) and -4.7 kcal mol⁻¹ (LgAs32:FRU), respectively. The rigid-body entropy changes at the dissociation of dimeric arrangement ($T\Delta S^{diss}$) were estimated to be 14.7 kcal mol⁻¹ for both LgAs32 structures. The positive value determined for ΔG^{diss} shows that the LgAs32 homodimer is

thermodynamically stable. Twenty-seven amino acid residues participate in the formation of the homodimeric interface (Tyr12, Lys13, Trp15, Asp16, Ala17, Gln18, Leu20, Leu21, Gln24, Ala25, Ala28, Thr29, Pro31, His35, Phe67, Pro82, His88, Tyr89, Leu90, Pro92, Asp95, Asn144, Asn145, His146, Asn486, Ser487, and Leu489). Ten hydrogen bonds related to the stability of the dimeric arrangement are listed in Table S4. The surface is symmetrical, as the same residues of one monomer are interacting with those of the other monomer. The interface area was estimated to be 1450 Å² by the PISA program. When comparing with BIFase (PDB ID 3PIJ) from the Gram-positive bacterium *Bacillus longum* that also presents a dimeric arrangement (Figure S4B), the interface of oligomerization is different and less stable than LgAs32, with a ΔG^{diss} of 0.4 kcal mol⁻¹ and ΔG_{int} of -35.4 kcal mol⁻¹. Curiously, both structures present a $T\Delta S^{diss}$ of 14.7 kcal mol⁻¹ even with a slightly smaller interface area of 1368 Å² for BIFase and a substantially higher number of hydrogen bonds (N_{HB} : 36) and salt bridges (N_{SB} : 25).

When comparing the oligomeric states of enzymes of the GH32 family from yeast, the *Schwanniomyces occidentalis* and *Kluyveromyces marxianus* present a dimeric assembly in their crystal structure, with a larger interface area than LgAs32, e.g., an invertase SoInv (PDB ID 3KF3 and 3KF5),⁷⁸ a β -fructofuranosidase SoFfase (PDB ID 6S1T and 6S2B),⁷⁹ and an exoinulinase KmINU1 (PDB ID 6J0T) (data not published) with an interface area of 1991, 2025, and 2420 Å², respectively (see Figure S4C–E). SoFfase is shown to be more stable with a ΔG^{diss} of 45.8 kcal mol⁻¹ and a considerable number of

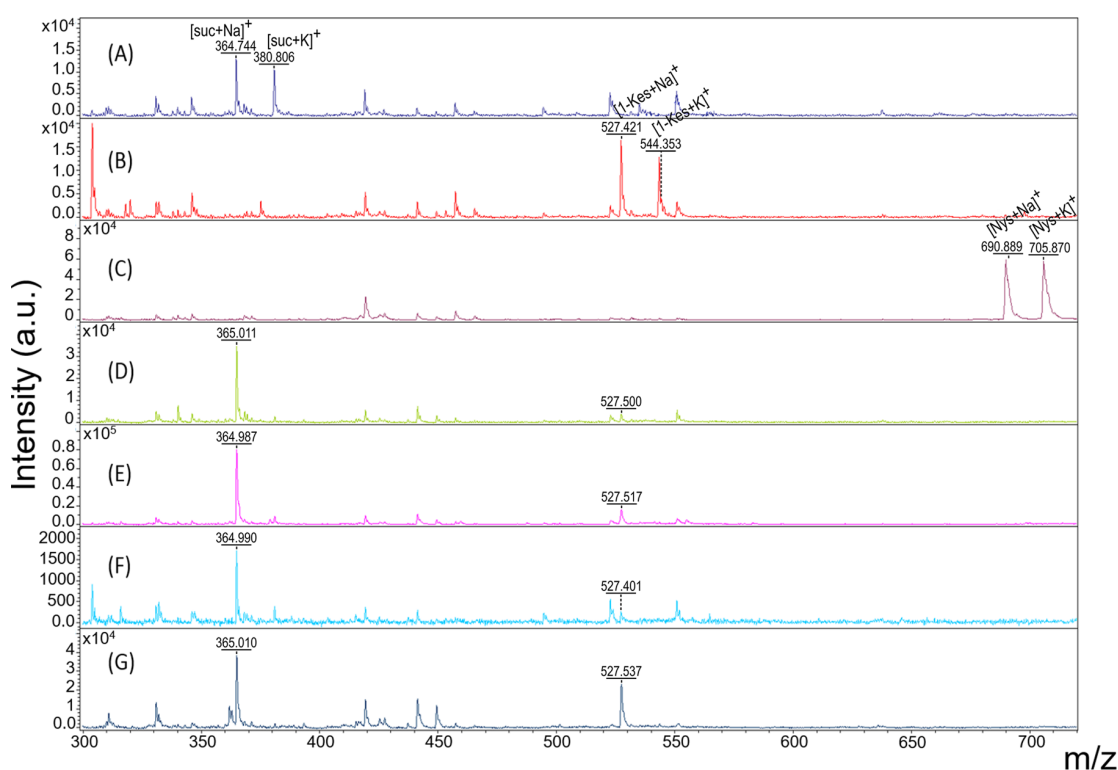


Figure 4. MALDI-TOF mass spectrometry with FOS 1-kestose identification. Since Na^+ and K^+ ions were present in the buffer, the patterns for sucrose (A), 1-kestose (B), and nystose (C) were displaced. The product of the reaction catalyzed by *LgAs32* after incubation with 300 mM of sucrose for 10 h (D) and after incubation with 2 M of sucrose for 10, 12, and 24 h, respectively (E–G).

hydrogen bonds (N_{HB} : 37) and also salt bridges (N_{SB} : 2). The *SoInv* and *KmINU1* present ΔG_{diss} values of 10.9 and 16.3 kcal mol^{-1} , respectively, slightly higher than *LgAs32*. In contrast, the positive ΔG_{int} for *SoInv* (23.3 kcal mol^{-1}) and *SoFfase* (18.9 kcal mol^{-1}) indicates a solvation free energy gain upon the formation of dimers, suggesting that their dimeric assembly in the crystallographic structure may be different than that in the biological environment. The *KmINU1* has a ΔG_{int} of -9.3 kcal mol^{-1} , similar to the other structures. Despite all the differences between the interfaces, the $T\Delta S_{\text{diss}}$ values do not differ significantly, being 15.1 kcal mol^{-1} (for *KmINU1*) and 10.0 kcal mol^{-1} (for *SoInv*), the highest and lowest value, respectively.

The variety of interfaces of oligomerization can be related to environmental conditions. The pH values of crystallization buffers may result in different crystallographic space groups (SGs) and dimer architectures for each structure, e.g., *LgAs32* (pH 9.0; SG: $P4_21_2$), *BIFfase* (pH 7.0; SG: $P3_21_1$), *SoInv* (pH 7.0; SG: $P12_11$), *SoFfase* (pH 7.0; SG: $P12_11$), and *KmINU1* (pH 6.0; SG: $I4_122$). Therefore, through the SAXS experiments, we evaluated three pH conditions.

Supporting Information Section S1.1 and Figure S5 show the experimental SAXS curves for *LgAs32*, measured at the three pH values, superimposed on the theoretical scattering curves based on the monomeric and homodimeric crystallographic structures (the latter detected by the PISA program). The theoretical scattering curves based on the homodimeric structure exhibited an excellent fit to the experimental data when compared to the monomeric structure. Therefore, the SAXS data for *LgAs32* are very consistent with a homodimeric molecule in solution, as also shown by the DLS analyzes. The

LgAs32 homodimeric crystal structure is a molecule with $R_g = 30.98$ Å and $D_{\text{max}} = 105.5$ Å (Figure 1). These values are in substantial agreement with the results obtained by the SAXS analyses (Table S5). The low-resolution model of *LgAs32* at pH 7.5 was built from X-ray scattering data using the DAMMIN program³⁸ (Figure S8). The low-resolution model of *LgAs32* (restored at 18 Å resolution) showed a molecule with D_{max} and R_g of 100.30 and 30.56 Å, respectively. Finally, the superposition of the low-resolution model and *LgAs32* homodimeric crystal structure showed good agreement (Figure S8). A summary of the main SAXS data is given in Table S5.

3.5. The Effects of pH, Temperature, and Thermal Stability on the Enzymatic Activity of *LgAs32*. Our results showed that the enzymatic activity of *LgAs32* at 37 °C increases between pH 4.5 and 6.5, reaching its maximum at pH 5.5; decays after pH 6; and becomes virtually zero at pH values below 4 and above 10, as shown in Figure 3A. For the temperature variation at pH 5.5, the activity raises from 20 to 50 °C with an exponential decay from 50 to 80 °C (Figure 3B). Taken together, our results show that the optimum activity for *LgAs32* occurs at pH 5.5 and 50 °C (Figure 3).

At physiological temperatures of the human body (around 37 °C), the optimum value for pH was 5.5, which is very similar to the gastrointestinal tract.⁸⁰ With an optimum temperature at 50 °C, *LgAs32* exhibits higher temperature stability, a significant advantage in industrial processes.^{81,82}

Heat stability assays were performed to evaluate the ability of *LgAs32* to keep its activity for long periods at different temperatures (40, 50, 60, and 70 °C) (Figure 3C). The results show that *LgAs32* was heat-stable at 40 °C and kept around 73% of its activity even after 48 h of incubation and

Table 1. Production of the Trisaccharide 1-Kestose by *LgAs32* and *LrFTF-Le*, and enzyme kinetic parameters of the *LgAs32* and GH32 enzymes.^a

(enzyme ID) source organism	substrate ^b	conditions			1-kestose production		reference
(<i>LgAs32</i>) <i>Lactobacillus gasseri</i>	1 M	12 h at 37 °C			23 mg mL ⁻¹		this study
(<i>LrFTF-lev</i>) <i>Lactobacillus reuteri</i>	0.26 M ^b	17 h at 37 °C			9.5 mg mL ^{-1,c}		83
	Assay conditions ^b	Sp. Act (U mg ⁻¹)	K _M (mM)	k _{cat} (s ⁻¹)	k _{cat} /K _M (mM ⁻¹ s ⁻¹)		
(<i>LgAs32</i>) <i>Lactobacillus gasseri</i>	pH 5.5, 50 °C	NA	200 ± 27 ^c	2.5 × 10 ³ ± 0.26 ^c	12.5 ^c		this study
(<i>LrFTF-Inu</i>) <i>Lactobacillus reuteri</i> 121	pH 5.4, 37 °C	97	NA	NA	NA		84
(<i>LrFTF-Inu</i>) <i>Lactobacillus reuteri</i> 121	pH 5.4, 50 °C	103.7	NA	NA	NA		83
(<i>BmSUC1</i>) <i>Bombyx mori</i>	pH 7.0, 30 °C	NA	4.89	97.3	19.8		64
(<i>TmBfA</i>) <i>Thermotoga maritima</i>	pH 5.5, 75 °C	NA	64	2.6 × 10 ³	40.6		94
(<i>BIFase</i>) <i>Bifidobacterium longum</i>	pH 6.2, 37 °C	NA	31.45	NA	NA		48
(<i>SoInv</i>) <i>Schwanniomyces occidentalis</i>	pH 5	NA	6.4	106	16.6		78
(<i>SoFase</i>) <i>Schwanniomyces occidentalis</i>	pH 5.6, 50 °C	NA	4.9	80.5	16.4		95
(<i>AtcwINV1</i>) <i>Arabidopsis thaliana</i>	pH 5.0, 30 °C	NA	0.35	59	168.6		96

^aSp. Act.: specific activity, NA: not available, U: unit defined as the release of 1 μmol of product per min per mg of sucrose under the specific assay conditions. ^bUsing sucrose as a substrate. ^b Value converted from percentage (%). ^c Value converted from g L⁻¹. ^c Data are mean ± SD (*n* = 3 independent measurements).

approximately 53% of activity after incubation at 50 °C for 48 h. However, when incubated at higher temperatures (60 and 70 °C), its activity was completely lost after 1 h of incubation.

The DSF technique was used to monitor the effect of buffer conditions on *LgAs32* stability as a function of melting temperature (*T_m*). Notably, we observed that *LgAs32* exhibited resistance against acidic pH, with higher stability between pH 4 and 6.5 and low variation against buffer salts in this same pH range (Figure S9A). However, the buffer screening approach identified conditions containing sodium citrate pH 5.5 and bistris pH 6.0 as optimal buffers for *LgAs32*, reaching a *T_m* of 62 °C. These data agree with the results for the enzymatic activity as detailed above, which confirm the optimum activity around pH 5.5. Therefore, it is possible to associate the *LgAs32* catalytic activity because of its structural stability. In contrast, a significant reduction in the activity at temperatures above 60 °C is caused by protein unfolding since the highest *T_m* value established was 62 °C. The inset of Figure S2B shows the variation of *R_S* as a function of temperature for *LgAs32* at pH 6. The *R_S* of *LgAs32* exhibited minimal temperature dependence in the range of 20 to 54 °C. For temperatures above 54 °C, the *R_S* rose steeply, suggesting amorphous aggregates because of the denaturation process. The unfolding process was irreversible under the conditions described in this study.

3.6. Analysis of Carbohydrates by Mass Spectrometry and HPAEC/PAD Identified the Presence of Fructooligosaccharides. Mass spectrometry analysis has shown the production of fructooligosaccharide 1-kestose by *LgAs32* with the incubation time and the sucrose concentration (Figure 4 and Figure S10A). Figure S10B indicates the specific activity of the enzyme for different sucrose concentrations (1 and 2 M) and incubation periods (8, 10, 12, and 24 h). Similarly, Figure S10C presents the production of nystose, and Figure S10D shows the specific activity of the enzyme.

All samples that revealed fructans in mass spectrometry were subjected to a quantitative analysis of their production by HPAEC/PAD. The chromatograms of the products are shown in Figure S11. Great amounts of glucose and fructose were observed from the reaction with sucrose, as well as small peaks with retention times between 5 and 10 min, suggesting not only hydrolytic activity but also fructooligosaccharide formation with a lower degree of polymerization (complex-ramified FOS): nystose and 1-kestose that later were

quantitatively measured (Figure S10). Such hydrolytic and transfructosylation activity is characteristic of this family of enzymes. However, we did not identify the production of other types of FOS.

Based on our experiments, the most efficient condition for the enzymatic production of 1-kestose by *LgAs32* was with 1 M sucrose after 12 h of incubation, yielding 23 mg mL⁻¹ (1.92 mg mL⁻¹/h) of the product with a specific activity of 27 U mg⁻¹ (Table 1 and Figure S10). After 24 h of incubation, the product concentration was around 20 mg mL⁻¹ (0.83 mg mL⁻¹/h) with a specific activity of 25 U mg⁻¹. When incubating *LgAs32* with 2 M of sucrose after 12 h, the production of 1-kestose drops to only 5 mg mL⁻¹ (0.42 mg mL⁻¹/h) with a specific activity of 6 U mg⁻¹. Lower concentrations of the product nystose (compared to 1-kestose) were also observed when refining the incubation time (8 and 10 h) with a reciprocal specific activity (Figure S10C,D). However, when compared in Table 1 with its homologous enzyme from *L. reuteri*, which is frequently used for fructan production,⁸³ although nystose production is smaller (0.26 mg mL⁻¹ versus 0.02 g L⁻¹ for *LgAs32*), 1-kestose production (9.5 mg mL⁻¹ for the *L. reuteri* recombinant enzyme) is significantly increased for recombinant *LgAs32* (23 mg mL⁻¹).

Anwar *et al.* (2012)⁸⁴ published an extensive work involving a wide variety of mutations in a fructosyltransferase enzyme inulosucrase of *Lactobacillus reuteri* (*LrINU*), reporting the effects of 15 single and 4 multiple mutations. Most mutants behaved similarly to the wild-type enzyme, except for the G416E mutation at the rim of the active pocket, increasing hydrolytic activity but causing no significant changes in transglycosylation activity. When the transglycosylation activity is compared in terms of specific activity, these (*LrINU*) mutants (ranging from 2.5 to 131 U mg⁻¹) and also the wild-type enzyme (44 U mg⁻¹) are similar to *LgAs32* (27 U mg⁻¹), which suggest that *LgAs32* could be further explored for FOS production. Once the targets for site-directed mutagenesis and the reaction pathway are decipherable, the results can be improved even further.

3.7. Kinetic Analysis. *LgAs32* efficiently hydrolyzed sucrose as described in Table 1, showing a higher *K_M* (200 mM) than other GH32 and close to 6-fold toward *BIFase* (31.45 mM). It suggests that *LgAs32* has a lower affinity for the substrate; in contrast, the turnover number (2.5 × 10³ s⁻¹)

is compatible with that observed for *TmBfrA* ($2.6 \times 10^3 \text{ s}^{-1}$), and both are significantly higher than those of other enzymes previously characterized. The k_{cat}/K_M of $12.5 \text{ mM}^{-1} \text{ s}^{-1}$ is the lowest value observed. It has been shown for fructosyltransferases that the sucrose concentrations directly affect hydrolysis, promoting it at low concentrations and being beneficial to transfructosylation at high concentrations.⁸⁵ With *L. reuteri*, the transfructosylation activity achieves 90% when the sucrose concentration is 1.7 M; this can also be increased with the enzyme concentration.⁸⁶ The high concentration of substrate will hinder the occupation of water molecules in the active site of the enzyme, increasing the transfructosylation activity and reducing the hydrolysis.⁷

Chambert *et al.* (1974)⁸⁷ proposed, from kinetic studies on initial velocity, a reaction with a "Ping-Pong BiBi" behavior, also called a double-displacement reaction, proposing the formation of a covalently bonded enzyme–fructose intermediary (Ping) and the release of glucose (Pong). This complex could interact with a water molecule (Ping), and the fructosyl moiety is then transferred, releasing the second product fructose (Pong) necessary for the branching of fructose chains, such as short-chain fructooligosaccharides GF_{*n*} (*n* = number of fructoses from 2 to 4).^{88,89} Figure S12 illustrates the reaction pathway performed by the *LgAs32*. Figure S12B presents a schematic sequence adapted from and first suggested by Jung *et al.* (1989)⁹⁰ for FOS production through transferase activity. The biochemical mechanism of retaining glycosidases is well described by Beine *et al.* (2008)⁹¹ and implies that the catalytic residues Asp47 (Nuc) and Glu222 (A/B) are in different protonation states, with a higher pKa for Glu222 acting as a general acid/base. The mechanism has two steps: (i) the sucrose (glucose–fructose) enters the active site and covalently binds (via fructose) to the enzyme (Asp47), forming an enzyme–fructofuranosyl intermediate,^{48,55} and (ii) a new sucrose molecule is recognized at the active site, and fructoses are transferred toward the fructooligosaccharide chain. Meanwhile, Asp167 or the enzyme–substrate hydrogen bonds stabilize the oxocarbenium-like transition state of the anomeric carbon between these two steps, promoting catalysis.⁵⁵

We report the production of fructans (1-kestose and nystose) from a recombinant enzyme. Compared to its homologous enzyme from *L. reuteri* (9.5 mg mL⁻¹ of 1-kestose production),⁸³ the production of 1-kestose by *LgAs32* can be increased. We tested only sucrose as a substrate, yet future studies could verify if *LgAs32* can produce varied FOS from different oligosaccharides once structural analysis shows that *LgAs32* owns a catalytic pocket very similar to other fructosyltransferases.

We observed that FOS production occurs expressively for the native enzyme *LgAs32* (this study) and the recombinant form from *A. niger*⁹² (see Table 1). In studies with mutants from *L. reuteri*, 1-kestose production was evaluated and may help to show the direction for future studies.^{83,84}

These results have the potential to improve FOS production by enzymes with considerable transferase activity, resulting in fructan production in a faster time and with a better turnover since its production currently occurs mainly through the use of bacteria or fungi, which can generate a greater number of variables. More information about enzymes can lead to the accurate engineering of FOS production and then contribute to industrial processes spending less on raw materials and allowing for enzymatic cycling.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jafc.1c03901>.

Experimental SAXS curves and dimerization (Section S1.1); composition of tested buffering conditions (Table S1); summary of data collection, phasing, and refinement statistics (Table S2); physicochemical properties of *LgAs32* (6NU8) tunnels (Table S3); list of hydrogen bonds in the homodimeric interface of *LgAs32* (6NU8) (Table S4); general SAXS results from *LgAs32* (Table S5); SDS-PAGE 15% under reducing conditions (Figure S1); CD spectroscopy and thermostability characterization of *LgAs32* (Figure S2); results of the MOLE 2.0 calculation for putative tunnels of the reaction (Figure S3); comparison between the interfaces of dimerization (Figure S4); SAXS data (Figure S5); Guinier plots and experimental distance distribution functions (Figure S6); dimensionless Kratky plots and experimental SAXS curves (Figure S7); a low-resolution model of *LgAs32* (Figure S8); results of the DSF assay (Figure S9); quantification of 1-kestose and nystose production by *LgAs32* (Figure S10); FOS production (Figure S11); and schematic view of the reaction pathway mediated by *LgAs32* Bfase (Figure S12) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

João R. C. Muniz – *Sao Carlos Institute of Physics (IFSC), University of Sao Paulo (USP), Sao Carlos, SP 13563-120, Brazil*; orcid.org/0000-0003-2731-3966; Phone: (+55)-16-3373-6668; Email: jrcmuniz@ifsc.usp.br

Authors

Mariana Z. T. de Lima – *Sao Carlos Institute of Physics (IFSC), University of Sao Paulo (USP), Sao Carlos, SP 13563-120, Brazil*; Present Address: M.Z.T.L.: Mackenzie Presbyterian University, Campinas, SP, 13073-148, Brazil

Leonardo R. de Almeida – *Sao Carlos Institute of Physics (IFSC), University of Sao Paulo (USP), Sao Carlos, SP 13563-120, Brazil*

Alain M. Mera – *Sao Carlos Institute of Physics (IFSC), University of Sao Paulo (USP), Sao Carlos, SP 13563-120, Brazil*; Present Address: A.M.M.: BioforLatam, Lambayeque, 14008, Peru.

Amanda Bernardes – *Sao Carlos Institute of Physics (IFSC), University of Sao Paulo (USP), Sao Carlos, SP 13563-120, Brazil*

Wanuis Garcia – *Centro de Ciências Naturais e Humanas (CCNH), Universidade Federal do ABC (UFABC), Santo André, SP 09210-580, Brazil*; orcid.org/0000-0003-3712-3488

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jafc.1c03901>

Author Contributions

[#]M.Z.T.L. and L.R.A. contributed equally to this paper

Author Contributions

M.Z.T.L., A.M., and J.R.C.M. designed the research. M.Z.T.L. and A.M. purified *LgAs32*. M.Z.T.L. characterized *LgAs32* enzymatically and performed HPAEC-PAD and mass spectroscopy assays. M.Z.T.L. and J.R.C.M. prepared cryo-crystal

samples, collected data sets, processed X-ray diffraction images, and refined the atomic coordinates. M.Z.T.L. and A.B. designed the differential scanning fluorimetry experiments. W.G. characterized LgAs32 biophysically. L.R.A. performed bioinformatics and structural analysis. M.Z.T.L., L.R.A., W.G., and J.R.C.M. wrote the paper with support from all other authors.

Notes

The authors declare no competing financial interest.

The atomic coordinates and structure factors (code 6NU7 and 6NU8) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

ACKNOWLEDGMENTS

We would like to thank the National Synchrotron Light Laboratory (LNLS, Brazil) and Central Experimental Multi-usuário da Universidade Federal do ABC (CEM/UFABC). We thank the São Paulo Research Foundation (FAPESP) for the financial support via grant nos. 2017/16291-5, 2014/50897-0 (J.R.C.M.), 2016/09152-6 (L.R.A.), and 2017/17275-3 (W.G.); São Carlos Institute of Physics/University of Sao Paulo; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the financial support via grant nos. 308865/2018-9 (J.R.C.M.), 305740/2017-2, 422132/2018-7, and 305816/2020-9 (W.G.); and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the fellowship to M.Z.T.L. and A.M.

REFERENCES

- (1) Anwar, M. A.; Kralj, S.; Piqué, A. V.; Leemhuis, H.; van der Maarel, M. J. E. C.; Dijkhuizen, L. Inulin and Levan Synthesis by Probiotic *Lactobacillus gasseri* Strains: Characterization of Three Novel Fructansucrase Enzymes and Their Fructan Products. *Microbiology* **2010**, *156*, 1264–1274.
- (2) Cunha, J. S.; Ottoni, C. A.; Morales, S. A. V.; Silva, E. S.; Maiorano, A. E.; Perna, R. F. Synthesis and Characterization of Fructosyltransferase from *Aspergillus oryzae* IPT-301 for High Fructooligosaccharides Production. *Braz. J. Chem. Eng.* **2019**, *657*–668.
- (3) Díez-Municio, M.; de las Rivas, B.; Jimeno, M. L.; Muñoz, R.; Moreno, F. J.; Herrero, M. Enzymatic Synthesis and Characterization of Fructooligosaccharides and Novel Maltosylfructosides by Inulosucrase from *Lactobacillus gasseri* DSM 20604. *Appl. Environ. Microbiol.* **2013**, *79*, 4129–4140.
- (4) Cruz, B. C. S.; Sarandy, M. M.; Messias, A. C.; Gonçalves, R. V.; Ferreira, C. L. L. F.; Peluzio, M. C. G. Preclinical and Clinical Relevance of Probiotics and Synbiotics in Colorectal Carcinogenesis: A Systematic Review. *Nutr. Rev.* **2020**, *78*, 667–687.
- (5) Granato, D.; Barba, F. J.; Bursać Kovačević, D.; Lorenzo, J. M.; Cruz, A. G.; Putnik, P. Functional Foods: Product Development, Technological Trends, Efficacy Testing, and Safety. *Annu. Rev. Food Sci. Technol.* **2020**, *11*, 93–118.
- (6) Ritsema, T.; Smeekens, S. Fructans: Beneficial for Plants and Humans. *Curr. Opin. Plant Biol.* **2003**, *6*, 223–230.
- (7) Chuankhayan, P.; Hsieh, C.-Y.; Huang, Y.-C.; Hsieh, Y.-Y.; Guan, H.-H.; Hsieh, Y.-C.; Tien, Y.-C.; Chen, C.-D.; Chiang, C.-M.; Chen, C.-J. Crystal Structures of *Aspergillus japonicus* Fructosyltransferase Complex with Donor/Acceptor Substrates Reveal Complete Subsites in the Active Site for Catalysis. *J. Biol. Chem.* **2010**, *285*, 23251–23264.
- (8) Nasir, A.; Sattar, F.; Ashfaq, I.; Lindemann, S. R.; Chen, M.-H.; Van den Ende, W.; Oner, E. T.; Kirtel, O.; Khaliq, S.; Ghauri, M. A.; Anwar, M. A. Production and Characterization of a High Molecular Weight Levan and Fructooligosaccharides from a Rhizospheric Isolate of *Bacillus aryabhattai*. *LWT* **2020**, *123*, 109093.
- (9) Zhang, J.; Wang, L.; Luan, C.; Liu, G.; Liu, J.; Zhong, Y. Establishment of a Rapid and Effective Plate Chromogenic Assay for Screening of *Aspergillus* Species with High β -Fructofuranosidase Activity for Fructooligosaccharides Production. *J. Microbiol. Methods* **2019**, *166*, 105740.
- (10) Mitsuoka, T. Development of Functional Foods. *Biosci. Microbiota, Food Health* **2014**, *33*, 117–128.
- (11) Saad, N.; Delattre, C.; Urdaci, M.; Schmitter, J. M.; Bressollier, P. An Overview of the Last Advances in Probiotic and Prebiotic Field. *LWT - Food Sci. Technol.* **2013**, *50*, 1–16.
- (12) Buitrón, D. I.; Sepúlveda, L.; Martínez, T. K. M.; Aguilar, C. N.; Medina, D. D.; Rodríguez-Herrera, R.; Flores-Gallegos, A. C. Biotechnological Approach for the Production of Prebiotics and Search for New Probiotics and Their Application in the Food Industry. *Appl. Food Biotechnol.* **2018**, *5*, 185–192.
- (13) Picazo, B.; Flores-Gallegos, A. C.; Muñoz-Márquez, D. B.; Flores-Maltos, A.; Michel-Michel, M. R.; de la Rosa, O.; Rodríguez-Jasso, R. M.; Rodríguez-Herrera, R.; Aguilar-González, C. N. Enzymes for Fructooligosaccharides Production: Achievements and Opportunities. In *Enzymes in Food Biotechnology Production, Applications, and Future Prospects*; Kuddus, M., Ed.; Academic Press, 2019; pp. 303–320. DOI: 10.1016/B978-0-12-813280-7.00018-9.
- (14) Sandoval-González, R. S.; Jiménez-Islas, H.; Navarrete-Bolaños, J. L. Design of a Fermentation Process for Agave Fructooligosaccharides Production Using Endo-Inulinases Produced in Situ by *Saccharomyces paradoxus*. *Carbohydr. Polym.* **2018**, *198*, 94–100.
- (15) Gunyakti, A.; Asan-Ozusaglam, M. *Lactobacillus gasseri* from Human Milk with Probiotic Potential and Some Technological Properties. *LWT* **2019**, *109*, 261–269.
- (16) Mera, A.; de Lima, M. Z. T.; Bernardes, A.; Garcia, W.; Muniz, J. R. C. Low-Resolution Structure, Oligomerization and Its Role on the Enzymatic Activity of a Sucrose-6-Phosphate Hydrolase from *Bacillus licheniformis*. *Amino Acids* **2019**, *51*, 599–610.
- (17) Eschenfeldt, W. H.; Lucy, S.; Millard, C. S.; Joachimiak, A.; Mark, I. D. A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. In *High Throughput Protein Expression and Purification: Methods and Protocols*; Doyle, S. A., Ed.; Humana Press: Totowa, NJ, 2009; pp 105–115, DOI: 10.1007/978-1-59745-196-3_7.
- (18) Doyle, S. A. High-Throughput Cloning for Proteomics Research. *Methods Mol. Biol.* **2005**, *310*, 107–113.
- (19) Paraskevopoulou, V.; Falcone, F. H. Polyionic Tags as Enhancers of Protein Solubility in Recombinant Protein Expression. *Microorganisms* **2018**, *6*, 47.
- (20) Kim, Y.; Dementieva, I.; Zhou, M.; Wu, R.; Lezondra, L.; Quartey, P.; Joachimiak, G.; Korolev, O.; Li, H.; Joachimiak, A. Automation of Protein Purification for Structural Genomics. *J. Struct. Funct. Genomics* **2004**, *5*, 111–118.
- (21) Studier, F. W. Protein Production by Auto-Induction in High Density Shaking Cultures. *Protein Expression Purif.* **2005**, *41*, 207–234.
- (22) Hall, M.; Rubin, J.; Behrens, S. H.; Bommarius, A. S. The Cellulose-Binding Domain of Cellobiohydrolase Cel7A from *Trichoderma reesei* Is Also a Thermostabilizing Domain. *J. Biotechnol.* **2011**, *155*, 370–376.
- (23) Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W. IMOSFLM: A New Graphical Interface for Diffraction-Image Processing with MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 271–281.
- (24) French, S.; Wilson, K. On the Treatment of Negative Intensity Observations. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1978**, *34*, 517–525.
- (25) Evans, P. Scaling and Assessment of Data Quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 72–82.
- (26) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S.

Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235–242.

(27) Matthews, B. W. Solvent Content of Protein Crystals. *J. Mol. Biol.* **1968**, *33*, 491–497.

(28) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

(29) Hattne, J.; Lamzin, V. S. Pattern-Recognition-Based Detection of Planar Objects in Three-Dimensional Electron-Density Maps. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2008**, *64*, 834–842.

(30) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 352–367.

(31) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355–367.

(32) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.

(33) Chen, V. B.; Arendall, W. B., III; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12–21.

(34) Berka, K.; Hanák, O.; Sehnal, D.; Banáš, P.; Navrátilová, V.; Jaiswal, D.; Ionescu, C.-M.; Svobodová Vařeková, R.; Koča, J.; Otyepka, M. MOLEonline 2.0: Interactive Web-Based Analysis of Biomacromolecular Channels. *Nucleic Acids Res.* **2012**, *40*, W222–W227.

(35) Sehnal, D.; Svobodová Vařeková, R.; Berka, K.; Pravda, L.; Navrátilová, V.; Banáš, P.; Ionescu, C.-M.; Otyepka, M.; Koča, J. MOLE 2.0: Advanced Approach for Analysis of Biomacromolecular Channels. *Aust. J. Chem.* **2013**, *5*, 39.

(36) Schrödinger, L. L. C. *The PyMOL Molecular Graphics System, Version 2.4.1*; 2015.

(37) Svergun, D. I. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *J. Appl. Crystallogr.* **1992**, *25*, 495–503.

(38) Svergun, D. I. Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. *Biophys. J.* **1999**, *76*, 2879–2886.

(39) Svergun, D.; Barberato, C.; Koch, M. H. J. CRY SOL - a Program to Evaluate X-Ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **1995**, *28*, 768–773.

(40) Krissinel, E.; Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **2007**, *372*, 774–797.

(41) Pijning, T.; Anwar, M. A.; Böger, M.; Dobruchowska, J. M.; Leemhuis, H.; Kralj, S.; Dijkhuizen, L.; Dijkstra, B. W. Crystal Structure of Inulosucrase from *Lactobacillus*: Insights into the Substrate Specificity and Product Specificity of GH68 Fructansucrases. *J. Mol. Biol.* **2011**, *412*, 80–93.

(42) Miller, G. L. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Anal. Chem.* **1959**, *31*, 426–428.

(43) Ericsson, U. B.; Hallberg, B. M.; DeTitta, G. T.; Dekker, N.; Nordlund, P. ThermoFluor-Based High-Throughput Stability Optimization of Proteins for Structural Studies. *Anal. Biochem.* **2006**, *357*, 289–298.

(44) Phillips, K.; de la Peña, A. H. The Combined Use of the ThermoFluor Assay and ThermoQ Analytical Software for the Determination of Protein Stability and Buffer Optimization as an Aid in Protein Crystallization. *Curr. Protoc. Mol. Biol.* **2011**, *94*, 10–28.

(45) Shiomi, N.; Onodera, S.; Chatterton, N. J.; Harrison, P. A. Separation of Fructooligosaccharide Isomers by Anion-Exchange Chromatography. *Agric. Biol. Chem.* **1991**, *55*, 1427–1428.

(46) Wise, C. S.; Dimler, R. J.; Davis, H. A.; Rist, C. E. Determination of Easily Hydrolyzable Fructose Units in Dextran Preparations. *Anal. Chem.* **1955**, *27*, 33–36.

(47) Alberto, F.; Bignon, C.; Sulzenbacher, G.; Henrissat, B.; Czjzek, M. The Three-Dimensional Structure of Invertase (β -Fructosidase) from *Thermotoga maritima* Reveals a Bimodular Arrangement and an Evolutionary Relationship between Retaining and Inverting Glycosidases. *J. Biol. Chem.* **2004**, *279*, 18903–18910.

(48) Bujacz, A.; Jedrzejczak-Krzepkowska, M.; Bielecki, S.; Redzyna, I.; Bujacz, G. Crystal Structures of the Apo Form of β -Fructofuranosidase from *Bifidobacterium longum* and Its Complex with Fructose. *FEBS J.* **2011**, *278*, 1728–1744.

(49) Ernits, K.; Eek, P.; Lukk, T.; Visnapuu, T.; Alamäe, T. First Crystal Structure of an Endo-Levanase – the BT1760 from a Human Gut Commensal *Bacteroides thetaiotaomicron*. *Sci. Rep.* **2019**, *9*, 8443.

(50) Park, J.; Kim, M.-I.; Park, Y.-D.; Shin, I.; Cha, J.; Kim, C. H.; Rhee, S. Structural and Functional Basis for Substrate Specificity and Catalysis of Levan Fructotransferase. *J. Biol. Chem.* **2012**, *287*, 31233–31241.

(51) Verhaest, M.; Lammens, W.; Le Roy, K.; De Coninck, B.; De Ranter, C. J.; Van Laere, A.; Van den Ende, W.; Rabijns, A. X-Ray Diffraction Structure of a Cell-Wall Invertase from *Arabidopsis thaliana*. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 1555–1563.

(52) Verhaest, M.; Ende, W. V. D.; Roy, K. L.; De Ranter, C. J.; Laere, A. V.; Rabijns, A. X-Ray Diffraction Structure of a Plant Glycosyl Hydrolase Family 32 Protein: Fructan 1-Exohydrolase IIa of *Cichorium intybus*. *Plant J.* **2005**, *41*, 400–411.

(53) Nagem, R. A. P.; Rojas, A. L.; Golubev, A. M.; Korneeva, O. S.; Eneyskaya, E. V.; Kulminkaya, A. A.; Neustroev, K. N.; Polikarpov, I. Crystal Structure of Exo-Inulinase from *Aspergillus awamori*: The Enzyme Fold and Structural Determinants of Substrate Recognition. *J. Mol. Biol.* **2004**, *344*, 471–480.

(54) Naumoff, D. G. β -Fructosidase Superfamily: Homology with Some α -L-Arabinases and β -D-Xylosidases. *Proteins: Struct., Funct., Bioinf.* **2001**, *42*, 66–76.

(55) Meng, G.; Fütterer, K. Structural Framework of Fructosyl Transfer in *Bacillus subtilis* Levansucrase. *Nat. Struct. Biol.* **2003**, *10*, 935–941.

(56) Martínez-Fleites, C.; Ortíz-Lombardía, M.; Pons, T.; Tarbouriech, N.; Taylor, E. J.; Arrieta, J. G.; Hernández, L.; Davies, G. J. Crystal Structure of Levansucrase from the Gram-Negative Bacterium *Gluconacetobacter diazotrophicus*. *Biochem. J.* **2005**, *390*, 19–27.

(57) Meng, G.; Fütterer, K. Donor Substrate Recognition in the Raffinose-Bound E342A Mutant of Fructosyltransferase *Bacillus subtilis* Levansucrase. *BMC Struct. Biol.* **2008**, *8*, 16.

(58) Verhaest, M.; Lammens, W.; Le Roy, K.; De Ranter, C. J.; Van Laere, A.; Rabijns, A.; Van den Ende, W. Insights into the Fine Architecture of the Active Site of Chicory Fructan 1-Exohydrolase: 1-Kestose as Substrate vs Sucrose as Inhibitor. *New Phytol.* **2007**, *174*, 90–100.

(59) Lammens, W.; Le Roy, K.; Van Laere, A.; Rabijns, A.; Van den Ende, W. Crystal Structures of *Arabidopsis thaliana* Cell-Wall Invertase Mutants in Complex with Sucrose. *J. Mol. Biol.* **2008**, *377*, 378–385.

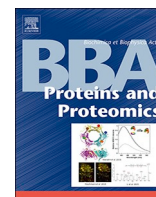
(60) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(61) Madden, T. L.; Tatusov, R. L.; Zhang, J. [9] Applications of network BLAST server. *Methods Enzymol.* **1996**, *266*, 131–141.

(62) Mátrai, J.; Lammens, W.; Jonckheer, A.; Le Roy, K.; Rabijns, A.; Van den Ende, W.; De Maeyer, M. An Alternate Sucrose Binding Mode in the E203Q *Arabidopsis* Invertase Mutant: An X-Ray Crystallography and Docking Study. *Proteins: Struct., Funct., Bioinf.* **2008**, *71*, 552–564.

(63) Alberto, F.; Jordi, E.; Henrissat, B.; Czjzek, M. Crystal Structure of Inactivated *Thermotoga maritima* Invertase in Complex with the Trisaccharide Substrate Raffinose. *Biochem. J.* **2006**, *395*, 457–462.

- (64) Miyazaki, T.; Oba, N.; Park, E. Y. Structural Insight into the Substrate Specificity of *Bombyx mori* β -Fructofuranosidase Belonging to the Glycoside Hydrolase Family 32. *Insect Biochem. Mol. Biol.* **2020**, *127*, 103494.
- (65) Reddy, V. A.; Maley, F. Identification of an Active-Site Residue in Yeast Invertase by Affinity Labeling and Site-Directed Mutagenesis. *J. Biol. Chem.* **1990**, *265*, 10817–10820.
- (66) Reddy, A.; Maley, F. Studies on Identifying the Catalytic Role of Glu-204 in the Active Site of Yeast Invertase. *J. Biol. Chem.* **1996**, *271*, 13953–13958.
- (67) Yuan, S.; Le Roy, K.; Venken, T.; Lammens, W.; Van den Ende, W.; De Maeyer, M. pKa Modulation of the Acid/Base Catalyst within GH32 and GH68: A Role in Substrate/Inhibitor Specificity? *PLoS One* **2012**, *7*, No. e37453.
- (68) Sprenger, N.; Bortlik, K.; Brandt, A.; Boller, T.; Wiemken, A. Purification, Cloning, and Functional Expression of Sucrose:Fructan 6-Fructosyltransferase, a Key Enzyme of Fructan Synthesis in Barley. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 11652–11656.
- (69) Batista, F. R.; Hernández, L.; Fernández, J. R.; Arrieta, J.; Menéndez, C.; Gómez, R.; Támbara, Y.; Pons, T. Substitution of Asp-309 by Asn in the Arg-Asp-Pro (RDP) Motif of *Acetobacter diazotrophicus* Levansucrase Affects Sucrose Hydrolysis, but Not Enzyme Specificity. *Biochem. J.* **1999**, *337*, 503–506.
- (70) Kyte, J.; Doolittle, R. F. A Simple Method for Displaying the Hydrophobic Character of a Protein. *J. Mol. Biol.* **1982**, *157*, 105–132.
- (71) Zimmerman, J. M.; Eliezer, N.; Simha, R. The Characterization of Amino Acid Sequences in Proteins by Statistical Methods. *J. Theor. Biol.* **1968**, *21*, 170–201.
- (72) Jones, D. T.; Taylor, W. R.; Thornton, J. M. The Rapid Generation of Mutation Data Matrices from Protein Sequences. *Bioinformatics* **1992**, *8*, 275–282.
- (73) Ma, J.; Li, T.; Tan, H.; Liu, W.; Yin, H. The Important Roles Played in Substrate Binding of Aromatic Amino Acids in Exo-Inulinase From *Kluyveromyces cicerisporus* CBS 4857. *Front. Mol. Biosci.* **2020**, *268*.
- (74) Lammens, W.; Le Roy, K.; Yuan, S.; Vergauwen, R.; Rabijns, A.; Van Laere, A.; Strelkov, S. V.; Van den Ende, W. Crystal Structure of 6-SST/6-SFT from *Pachysandra terminalis*, a Plant Fructan Biosynthesizing Enzyme in Complex with Its Acceptor Substrate 6-Kestose. *Plant J.* **2012**, *70*, 205–219.
- (75) Trujillo Toledo, L. E.; Martínez García, D.; Pérez Cruz, E.; Rivera Intriago, L. M.; Pérez, J. N.; Pais Chanfrau, J. M. *Chapter 26 - Fructosyltransferases and Invertases: Useful Enzymes in the Food and Feed Industries*; Kuddus, M. B. T.-E. in F. B., Ed.; Academic Press, 2019; pp. 451–469. DOI: 10.1016/B978-0-12-813280-7.00026-8.
- (76) Hothorn, M.; Van den Ende, W.; Lammens, W.; Rybin, V.; Scheffzek, K. Structural Insights into the pH-Controlled Targeting of Plant Cell-Wall Invertase by a Specific Inhibitor Protein. *Proc. Natl. Acad. Sci.* **2010**, *107*, 17427–17432.
- (77) Pouyez, J.; Mayard, A.; Vandamme, A.-M.; Roussel, G.; Perpète, E. A.; Wouters, J.; Housen, I.; Michaux, C. First Crystal Structure of an Endo-Inulinase, INU2, from *Aspergillus ficuum*: Discovery of an Extra-Pocket in the Catalytic Domain Responsible for Its Endo-Activity. *Biochimie* **2012**, *94*, 2423–2430.
- (78) Álvaro-Benito, M.; Polo, A.; González, B.; Fernández-Lobato, M.; Sanz-Aparicio, J. Structural and Kinetic Analysis of *Schwanniomyces occidentalis* Invertase Reveals a New Oligomerization Pattern and the Role of Its Supplementary Domain in Substrate Binding. *J. Biol. Chem.* **2010**, *13930*.
- (79) Rodrigo-Frutos, D.; Jiménez-Ortega, E.; Piedrabuena, D.; Ramírez-Escudero, M.; Míguez, N.; Plou, F. J.; Sanz-Aparicio, J.; Fernández-Lobato, M. New Insights into the Molecular Mechanism behind Mannitol and Erythritol Fructosylation by β -Fructofuranosidase from *Schwanniomyces occidentalis*. *Sci. Rep.* **2021**, *11*, 7158.
- (80) Fallingborg, J. Intraluminal pH of the Human Gastrointestinal Tract. *Dan. Med. Bull.* **1999**, *46*, 183–196.
- (81) Sangeetha, P. T.; Ramesh, M. N.; Prapulla, S. G. Recent Trends in the Microbial Production, Analysis and Application of Fructooligosaccharides. *Trends Food Sci. Technol.* **2005**, *16*, 442–457.
- (82) Vaughan, E. E.; Heilig, H. G. H. J.; Zoetendal, E. G.; Satokari, R.; Collins, J. K.; Akkermans, A. D. L.; de Vos, W. M. Molecular Approaches to Study Probiotic Bacteria. *Trends Food Sci. Technol.* **1999**, *10*, 400–404.
- (83) van Hijum, S. A. F. T.; van Geel-Schutten, G. H.; Rahaoui, H.; van der Maarel, M. J. E. C.; Dijkhuizen, L. Characterization of a Novel Fructosyltransferase from *Lactobacillus reuteri* That Synthesizes High-Molecular-Weight Inulin and Inulin Oligosaccharides. *Appl. Environ. Microbiol.* **2002**, *68*, 4390–4398.
- (84) Anwar, M. A.; Leemhuis, H.; Pijning, T.; Kralj, S.; Dijkstra, B. W.; Dijkhuizen, L. The Role of Conserved Inulosucrase Residues in the Reaction and Product Specificity of *Lactobacillus reuteri* Inulosucrase. *FEBS J.* **2012**, *279*, 3612–3621.
- (85) Kralj, S.; Leeflang, C.; Sierra, E. I.; Kempinski, B.; Alkan, V.; Kolkman, M. Synthesis of Fructooligosaccharides (FosA) and Inulin (InuO) by GH68 Fructosyltransferases from *Bacillus agaradhaerens* Strain WDG185. *Carbohydr. Polym.* **2018**, *179*, 350–359.
- (86) Ozimek, L. K.; Kralj, S.; Kaper, T.; van der Maarel, M. J. E. C.; Dijkhuizen, L. Single Amino Acid Residue Changes in Subsite – 1 of Inulosucrase from *Lactobacillus reuteri* 121 Strongly Influence the Size of Products Synthesized. *FEBS J.* **2006**, *273*, 4104–4113.
- (87) Chambert, R.; Treboul, G.; Dedonder, R. Kinetic Studies of Levansucrase of *Bacillus subtilis*. *Eur. J. Biochem.* **1974**, *41*, 285–300.
- (88) Vega, R.; Zuniga-Hansen, M. E. A New Mechanism and Kinetic Model for the Enzymatic Synthesis of Short-Chain Fructooligosaccharides from Sucrose. *Biochem. Eng. J.* **2014**, *82*, 158–165.
- (89) Singh, R. S.; Singh, R. P.; Kennedy, J. F. Recent Insights in Enzymatic Synthesis of Fructooligosaccharides from Inulin. *Int. J. Biol. Macromol.* **2016**, *85*, 565–572.
- (90) Jung, K. H.; Yun, J. W.; Kang, K. R.; Lim, J. Y.; Lee, J. H. Mathematical Model for Enzymatic Production of Fructo-Oligosaccharides from Sucrose. *Enzyme Microb. Technol.* **1989**, *11*, 491–494.
- (91) Beine, R.; Moraru, R.; Nimtz, M.; Na'amnieh, S.; Pawlowski, A.; Buchholz, K.; Seibel, J. Synthesis of Novel Fructooligosaccharides by Substrate and Enzyme Engineering. *J. Biotechnol.* **2008**, *138*, 33–41.
- (92) Hidaka, H.; Hirayama, M.; Sumi, N. A Fructooligosaccharide-Producing Enzyme from *Aspergillus niger* ATCC 20611. *Agric. Biol. Chem.* **1988**, *52*, 1181–1187.
- (93) Stierand, K.; Rarey, M. Drawing the PDB: Protein-Ligand Complexes in Two Dimensions. *ACS Med. Chem. Lett.* **2010**, *1*, 540–545.
- (94) Liebl, W.; Brem, D.; Gotschlich, A. Analysis of the Gene for β -Fructosidase (Invertase, Inulinase) of the Hyperthermophilic Bacterium *Thermotoga maritima*, and Characterisation of the Enzyme Expressed in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **1998**, *50*, 55–64.
- (95) Álvaro-Benito, M.; de Abreu, M.; Fernández-Arrojo, L.; Plou, F. J.; Jiménez-Barbero, J.; Ballesteros, A.; Polaina, J.; Fernández-Lobato, M. Characterization of a β -Fructofuranosidase from *Schwanniomyces occidentalis* with Transfructosylating Activity Yielding the Prebiotic 6-Kestose. *J. Biotechnol.* **2007**, *132*, 75–81.
- (96) Le Roy, K.; Lammens, W.; Verhaest, M.; De Coninck, B.; Rabijns, A.; Van Laere, A.; Van den Ende, W. Unraveling the Difference between Invertases and Fructan Exohydrolases: A Single Amino Acid (Asp-239) Substitution Transforms *Arabidopsis* Cell Wall Invertase1 into a Fructan 1-Exohydrolase. *Plant Physiol.* **2007**, *145*, 616–625.



Functional and structural characterization of an α -L-arabinofuranosidase from *Thermothielavioides terrestris* and its exquisite domain-swapped β -propeller fold crystal packing

Suelen Camargo^{a,1}, Evandro J. Mulinari^{a,1}, Leonardo R. de Almeida^{a,1}, Amanda Bernardes^a, Rolf A. Prade^b, Wanius Garcia^c, Fernando Segato^d, João R.C. Muniz^{a,*}

^a São Carlos Institute of Physics (IFSC), University of São Paulo (USP), São Carlos, SP, Brazil

^b Department of Microbiology and Molecular Genetics, Oklahoma State University (OSU), Stillwater, OK, USA

^c Centro de Ciências Naturais e Humanas, Universidade Federal do ABC (UFABC), Santo André, SP, Brazil

^d Department of Biotechnology, Lorena School of Engineering, University of São Paulo (USP), Lorena, SP, Brazil



ARTICLE INFO

Keywords:

Thermothielavioides terrestris
 α -L-arabinofuranosidase
 Domain-swapping
 GH62
 TtAbf62

ABSTRACT

The fungus *Thermothielavioides terrestris* plays an important role in the global carbon cycle with enzymes capable of degrading polysaccharides from biomass, therefore an attractive source of proteins to be investigated and understood. From cloning to a three-dimensional structure, we foster a deeper characterization of an α -L-arabinofuranosidase, a glycoside hydrolase from the family 62 (TtAbf62), responsible to release arabinofuranose from non-reducing ends of polysaccharides. TtAbf62 was tested with synthetic (pNP-Araf) and polymeric substrates (arabinan and arabinoxylan), showing optimal temperature and pH (for pNP-Araf) of 30 °C and 4.5–5.0, respectively. Kinetic parameters revealed different specific activity for the three substrates, with a higher affinity for pNP-Araf (K_M : 4 ± 1 mM). The hydrolyzing activity of TtAbf62 on sugarcane bagasse suggests high efficiency in the decomposition of arabinoxylan, abundant hemicellulose presented in the sugarcane cell wall. The crystal packing of TtAbf62 reveals an exquisite domain swapping, located at the supramolecular arrangement through a disulfide bond. All crystallographic behaviors go against its monomeric state in solution, indicating a crystal-induced artifact. Structural information will form the basis for further studies aiming the development of optimized enzymatic properties to be used in biotechnological applications.

1. Introduction

The α -L-arabinofuranosidases (Abfs; EC 3.2.1.55) play a key role in the releasing process of L-arabinose from polysaccharides or arabinooligosaccharides acting into α -(1 \rightarrow 2)-, α -(1 \rightarrow 3)- or α -(1 \rightarrow 5)- linked L-arabinofuranosyl residues from the non-reducing end of poly- and oligosaccharides [1], increasing the hydrolysis of glycosidic bonds of polysaccharides by other enzymes [2]. According to the Carbohydrate-Active Enzymes database (CAZy; www.cazy.org), Abfs are grouped into six distinct families of glycoside hydrolases: GH2, GH3, GH5, GH39, GH43, GH51, GH54, and GH62 [3–5]. Based on their amino acid sequences, members of GH62 family has been divided into GH62_1 and GH62_2 subfamilies [6]. Additionally, the Abf families are further classified into three types based on divergence among themselves concerning substrates specificity [7,8].

These enzymes can release arabinose from many substrates and they are part of the xylanolytic system, being essential to disrupt arabinoxylan bonds [1,9]. Therefore, Abfs are of remarkably biotechnological importance [1], which extends from the use in industries oriented to agriculture [10], food [11], and also in cocktails for digestion of animal feed [10], enhancement of wines flavor [10,12], improvement in the clarification of fruit juices [10,13], and for the delignification of cellulose on biofuels production [14,15].

The filamentous fungus *Thermothielavioides terrestris* (previously *Thielavia terrestris*), belongs to the Ascomycota phylum and is among a limited number of eukaryotic species classified as thermophiles showing optimal growth above 45 °C [16]. Many enzymes characterized from *T. terrestris*, such as cellulases and hemicellulases, have been shown attractive for biotechnological applications involving biomass hydrolysis due to its favorable catalytic and stability profiles [16–19].

* Corresponding author at: n 1100, Jardim Santa Angelina - CEP 13563-120, São Carlos, São Paulo, Brazil.

E-mail address: jrcmuniz@ifsc.usp.br (J.R.C. Muniz).

¹ Authors contributed equally to this manuscript.

<https://doi.org/10.1016/j.bbapap.2020.140533>

Received 19 April 2020; Received in revised form 25 July 2020; Accepted 12 August 2020

Available online 29 August 2020

1570-9639/ © 2020 Elsevier B.V. All rights reserved.

Aiming to contribute towards a better comprehension of these classes of enzymes, in this study we describe an α -L-arabinofuranosidase from the thermophilic fungus *T. terrestris* (subfamily GH62_2; named here *TtAbf62*). The biochemical characteristics of purified *TtAbf62* were determined using sugar beet arabinan (SBA), wheat arabinoxylan (WAX), and *p*NP-Araf as substrates. Moreover, the results presented here allowed inferences to be drawn on the hydrolysis of SBA, WAX, and *p*NP-Araf by *T. terrestris*. The biophysical and biochemical outcomes aim to promote a better understanding of its catalytic mechanism when degrading renewable biomass, making them potential candidates for biotechnological applications in the biofuel and food industries. Furthermore, structural analyzes revealed the presence of a domain-swapping, probably a crystallization artifact. However, this new and beautiful shape elegantly preserves the active site of the enzyme where the catalytic pocket was structurally conserved between the 5-bladed β -propeller fold and confers protein stability.

2. Methods

2.1. Fungus cultivation

The *T. terrestris* strain UAMH 3264 was purchased at Microfungus Collection of University of Alberta, Devonian Botanic Garden Edmonton. The fungus was maintained in Potato Dextrose Agar medium (Sigma-Aldrich), at 45 °C for 5 days and the resulting mycelium was used for further cultivation that was performed in 1 L of the minimal medium [20] containing: 50 mL of 20 \times Clutterbuck solution [21], 1 mL of a trace elements solution (100 mL MilliQ[®] water: 2.2 g of ZnSO₄, 1 g of H₃BO₃, 0.5 g of MnCl₂·4H₂O, 0.5 g of FeSO₄·7H₂O, 0.16 g of CoCl₂·5H₂O, 0.16 g of CuSO₄·5H₂O, 0.11 g of Na₂MoO₄·4H₂O and 5 g of Na₂EDTA at pH 6.5) and supplemented with 5 mM glutamine and 2.4 μ M thiamine at pH 5.0 [22] using a rotary shaker (150 rpm) at 45 °C for 48 h.

2.2. Cloning, heterologous expression, and purification of *TtAbf62*

A total of 10 mg of the mycelium of *T. terrestris* was frozen in liquid nitrogen and then triturated using a pestle, placed in a 1.5 mL tube where 600 μ L and the protocol for genomic DNA isolation was followed according to Damásio and collaborators [23]. To confirm the quality of the extracted DNA, the sample was analyzed on agarose gel (0.8%) by using the technique of electrophoresis.

The gene encoding the *TtAbf62* protein (GenBank: [AE064662.1](#)) was amplified from *T. terrestris* genomic DNA by Polymerase Chain Reaction (PCR) and cloned into pETTrx-1-a/LIC expression vector (Novagen) allowing the expression of recombinant protein with a His₆-Thioredoxin-tag (His₆-Trx-tag) fused at its N-terminal [24,25] using specific oligonucleotides (Supplemental SM1). The vector containing the gene fragment of interest was transformed into *E. coli* Rosetta-gami[™] 2 (DE3) (Novagen) expression system and the positive colonies were selected based on its resistance of growing in the presence of kanamycin and chloramphenicol. The recombinant *TtAbf62* was over-expressed in a 2 L Erlenmeyer containing 1 L of auto-induction medium (Supplemental SM2) added with kanamycin (0.05 mg·mL⁻¹) and chloramphenicol (0.036 mg·mL⁻¹), incubated in an orbital shaker with 150 rpm at 37 °C until an optical density of 0.6 at 600 nm; then the temperature was lowered and stabilized at 17 °C to induce protein expression for 17 h. The expression was monitored by SDS-PAGE. Cells were separated by centrifugation at 4500 \times g for 20 min at 4 °C, and the pellet stored at -80 °C.

The pellet was resuspended in buffer A (50 mM Tris-HCl at pH 7.5, 150 mM NaCl and 10% (w/v) of glycerol) for cell disruption. The treated pellet was sonicated on ice in 3 cycles (30 s sonication, 30 s rest), in a sonicator (Fisher Scientific[™] Model 505 Sonic Dismembrator) set with 45% amplitude. The lysate was then centrifuged at 14,000 \times g at 4 °C for 40 min. Immobilized metal affinity chromatography (IMAC)

with 2 mL Ni Sepharose 6 Fast Flow (GE Healthcare Life Sciences) was performed as the first purification step and after washing the column with buffer A, the recombinant protein was eluted with the same buffer with 150 mM imidazole added on it. Eluted solution was dialyzed overnight at 4 °C with buffer A. TEV protease (1:50 mg) was used for cleavage of the N-terminal His₆-Trx-tag. The second purification step repeats the IMAC purification step but elutes the protein with 20 mL of buffer A without imidazole. A third purification step was performed using size exclusion chromatography (SEC) on a HiLoad 16/60 Superdex 200 column (GE Healthcare Life Sciences) with buffer A. The purity of the recombinant protein was monitored by SDS-PAGE.

2.3. Biochemical characterization

Optimal pH for *TtAbf62* was determined by measuring the release of 4-nitrophenol from *p*-nitrophenyl- α -L-arabinofuranoside (*p*NP-Araf; 2 mM; Sigma-Aldrich) as the substrate in 100 μ L reactions containing 10 μ g of the purified enzyme, and 50 mM Mcllvaine buffer [26] with the pH ranging from 2.0 to 8.0 at 50 °C for 10 min. The optimal temperature was evaluated as described before, using the optimal pH varying the temperature from 10.0 to 60.0 °C, incubated for 15 min. After incubation, all reactions were quenched by the addition of 100 μ L of a stop reagent (saturated sodium tetraborate solution) and the absorbances of the released *p*-nitrophenol was measured at 410 nm. For kinetic parameters determination, the end-point velocities of the enzymes were measured in 50 mM Mcllvaine buffer at optimal pH and temperature, using *p*NP-Araf in concentrations ranging from 0.1 to 50.0 mM, sugar beet arabinan (SBA) and wheat arabinoxylan (WAX) (Megazyme) in concentrations ranging from 1 to 27 mg·mL⁻¹ for 25 min for polysaccharides and 7 min for *p*NP-Araf. All reactions were carried out in 100 μ L and using 2 μ g of the enzyme. The reducing sugars released in enzymatic reactions with polysaccharides was measured by the 3,5-dinitrosalicylic acid method (DNS) by measuring the absorbances at 540 nm [27]. The kinetics parameters (K_M and k_{cat}) were estimated by using direct weighting nonlinear least-squares regression analysis with Origin software (version 2018). The adjusted R-Square (Adj. R-Square) value was determined as a measure of the goodness of fit of experimental data. A calibration curve for the DNS method was plotted and used to calculate the amount of non-reducing end released. The 4-nitrophenol (*p*NP) was used to calculate the amount of arabinose released in *p*NP-Araf reactions. All assays were carried out in triplicate and the averaged values reported.

The activity of *TtAbf62* in the presence of sugarcane bagasse *in natura* was evaluated by incubating the reaction in the agitation of 150 rpm in a rotatory shaker at 50 °C for 20 h. The duplicate reaction with 1.25 mL contained 20 mg of sugarcane bagasse *in natura* in 1.25 mL reaction with 40 mM Mcllvaine buffer pH 4.5, 30 μ g *TtAbf62* and 0.02% sodium azide. In the enzyme-free reaction (control), the volume corresponding to the enzyme was replaced by 50 mM Tris-HCl at pH 8.0 (stock buffer). The reaction was centrifuged at 13,000 \times g, aliquoted in 100 μ L, and the amount of reducing sugars was measured by DNS [27]. The thermal stability of *TtAbf62* was evaluated by measuring the residual activity at optimal pH and incubating the enzyme in temperatures of 30 and 50 °C. Amounts corresponding to 10 μ g of the enzyme were aliquoted at time intervals and assays carried out using *p*NP-Araf (0.5 mM final concentration) as the substrate in 100 μ L reactions up to the total time of 36 h of incubation. All measurements were performed in triplicates and conducted in a 96-well microplate (Greiner) by using a Spectramax Plus 384 spectrophotometer (Molecular Device).

2.4. Thermal denaturation analysis

2.4.1. Differential scanning fluorimetry (DSF)

The DSF method was used to evaluate the structural integrity of *TtAbf62* in different conditions including a variety of buffer salts and

pH values. A CFX96 Real-Time System® (Bio-Rad) was used to detect the dye fluorescence allowing the determination of *TtAbf62* melting temperature T_m [28]. The assay was carried out applying 30 µg of purified *TtAbf62* to a 96 well plate containing different buffer solutions (Supplementary Table A.1), in a final concentration of 50 mM and 1:2000 of SYPRO Orange (Molecular Probes) and analysis were performed within the temperature range from 25 to 90 °C (with 1 °C increment of each 30 s). The fluorescence variation of SYPRO Orange was measured using excitation at 490 nm and emission at 530 nm. The construction and analysis of the curves, including the determination of the T_m of each sample, were carried out using Origin, version 2018 (OriginLab Corporation).

2.4.2. Circular dichroism (CD) spectroscopy

CD transition curve of *TtAbf62* was obtained using a JASCO J-815 CD spectropolarimeter in 50 mM McIlvaine, pH 4.5, and 50 mM Tris-HCl, 0.3 M NaCl, pH 7.5 buffers, using a 0.1 cm quartz cuvette, monitoring the ellipticity at 220 nm from 20 °C to 80 °C at a heating rate of 1 °C/min.

2.5. Protein crystallization, data collection, and processing

Crystals of *TtAbf62* were obtained using the purified protein in a concentration of 20 mg.mL⁻¹ in previously described buffer A. The protein concentration was determined by NanoDrop® 2000 spectrophotometer (Thermo Scientific). The initial screening of crystallization conditions was performed using a Crystal Gryphon LCP® crystallization robot (Art Robbins Instruments). Crystals were obtained after 30 days during exposure to 0.1 M sodium acetate (pH 4.6) and 2 M ammonium sulfate as reservoir solution at 18 °C.

The crystal diffraction and data collection were performed at the synchrotron radiation source at the Brazilian National Laboratory of Synchrotron Light from National Center for Energy and Materials (LNLS-CNPEN, Brazil). Diffraction datasets from the *TtAbf62* crystal were collected on beamline W01-MX2 using a PILATUS2M detector (Dectris). The data-collection strategy was determined using the iMosflm Strategy function [29], targeting > 95% completeness for data set collected to the given resolutions. The data were indexed and integrated using the programs XIA2 [30] and XDS [31] and scaled using Aimless [32,33] from the CCP4 suite [34]. The data completeness and $CC_{1/2}$ were used to determine the resolution limit.

2.6. Structure determination and refinement

Structure of *TtAbf62* was initially solved using molecular replacement with Phaser program [35]. The structure of a thermostable family GH62 α -L-arabinofuranosidase from *Streptomyces thermoviolaceus* (PDB id; 4O8N) [2] was used as a search model (79% of sequence identity). Crystallographic refinement was carried out using REFMAC5 [36] and manual model building was performed with Coot program [37] using weighted $2F_{obs} - F_{calc}$ and $F_{obs} - F_{calc}$ electron-density maps. Water molecules and additional fortuitous ligands such as sulfate ions and ethylene glycol molecules were manually placed appropriately by using Coot. The behavior of R_{work} and R_{free} was used as the main criteria to validate the refinement protocol and its convergence. The stereochemical quality of the model was evaluated with the MolProbity program [38]. The PISA server [39] was applied to verify oligomeric states by evaluating interfaces and their strength of the interaction between neighboring monomers generated by applying symmetry to the input file coordinates on the macromolecular assembly of *TtAbf62*, consequently, predict the most likely biological form of the structure established through domain swapping.

2.7. Size exclusion chromatography with multiangle light scattering (SEC-MALS)

Using Waters 600 HPLC coupled to a Wyatt DAWN EOS light scattering instrument, Wyatt Optilab refractometer and Wyatt dynamic light scattering module (Wyatt Technology Corp., Santa Barbara, CA), SEC-MALS experiments were employed for chromatographic separations to verify the oligomeric state of *TtAbf62* using a Superdex 75 10/300 GL (GE Healthcare Life Sciences) equilibrated with 20 mM Tris-HCl at pH 8.0 and 300 mM NaCl buffer, with an eluent flow rate of 0.5 mL.min⁻¹. Molecular mass calculations were performed by ASTRA software.

2.8. Glutaraldehyde cross-linking

Glutaraldehyde treatment with cross-linkers was performed in 20 mM HEPES at pH 8.5. Reaction mix containing 1 µg of protein (control and *TtAbf62*) in a total volume of 100 µL were treated with 5 and 10 µL of 2.3% freshly prepared solution of glutaraldehyde for 5 min at 37 °C. The reaction was stopped by the addition of 10 µL of 1 M Tris-HCl at pH 8.5. Cross-linked proteins were solubilized by the addition of an equal volume of Laemmli buffer [40], and verified by SDS-PAGE [41].

3. Results and discussion

3.1. Expression and purification of *TtAbf62*

The *TtAbf62* gene was localized in *T. terrestris* chromosome I between positions 9,454,228 and 9,455,223 at the complement strand. The gene sequence did not present introns and showed a region of 78 bp that correspond to the signal peptide (SP) in the *TtAbf62* protein. To oligonucleotides synthesis, the SP region was not considered and the gene showing 903 bp was amplified by PCR, cloned and expressed. The recombinant *TtAbf62* was expressed in *E. coli* system containing a tag of 6 histidine residues (His₆) and a thioredoxin (Trx) fused-protein. The total protein yield was approximately 2 mg.L⁻¹ of culture medium. Before TEV protease treatment, the SDS-PAGE analysis revealed a molecular mass bellow 45 kDa, which was consistent with the product of cloning on pETTrx-1-a/LIC (Fig. 1). After TEV cleavage the final molecular mass was around 31 kDa, which contains the residual site of TEV, plus one methionine, resulting in a molecular mass calculated from the mature amino acid sequence of 32.8 kDa (Fig. 1).

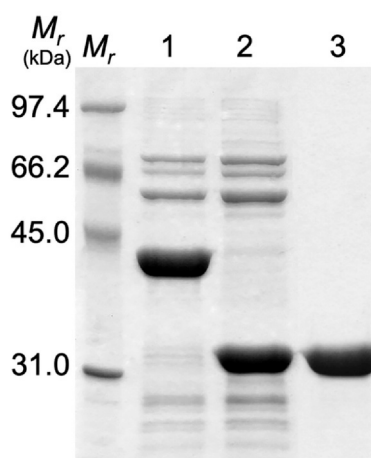


Fig. 1. Purification of *TtAbf62*. M_r – molecular mass; 1 – *TtAbf62* after an IMAC step of purification (43 kDa); 2 – The protein with approximately 33 kDa after cleavage with TEV protease; 3 – Purified protein after size exclusion chromatography.

3.2. Thermal denaturation analysis

To investigate the thermostability under miscellaneous buffer conditions, the DSF was performed for *TtAbf62* and its melting point temperatures (T_m), at different buffering conditions, were determined, as shown in the Supplementary Table A.1. An overview of the analyzed conditions indicates that the higher T_m was 64 °C in 50 mM Bis-Tris buffer at pH 6.7. These results also demonstrated a reasonable thermal tolerance at temperatures near to 60 °C, decreasing the values of T_m only when the buffer pH was extremely acidic or alkaline regardless of the composition (more details in the Supplementary Table A.1).

The DSF analyses established that Bis-Tris buffer at pH 6.7 was the best condition for protein stability with a T_m of 64 °C, followed by the same buffer at pH 6.0 and Tris-HCl buffer at pH 7.5, both with a T_m of 63 °C. The addition of NaCl did not affect the T_m . Thermal denaturation was also examined by CD as a complementary technique providing additional information on T_m . Both CD conditions resulted in T_m values of 59.7 ± 0.4 and 62.6 ± 0.2 °C for McIlvaine and Tris buffers, respectively (Supplementary Fig. A.1). These data suggest that the enzyme has great structural stability over a wide spectrum of pH values and buffers conditions.

Considering that increased thermal stability of the enzyme is correlated with enhanced structural arrangement often followed by conformational integrity of protein [42–44], the crystallization tests of *TtAbf62* were performed in Tris-HCl buffer at pH 7.5, one of the best conditions pointed out by the DSF experiment.

3.3. Biochemical characterization

3.3.1. Optimum pH and temperature for *TtAbf62* activity

Some enzymes of family GH62 can hydrolyze the substrate *pNP-Araf*, and also display arabinofuranosidase activity when assayed with polymeric substrates [2,45,46]. The activity of *TtAbf62* was verified on polymeric substrates (Supplementary Fig. A.2) and did not show activity in the synthetic substrates as *p*-nitrophenyl- β -D-mannopyranoside, *p*-nitrophenyl- β -D-glucopyranoside, *p*-nitrophenyl- β -D-galactopyranoside, *p*-nitrophenyl- β -D-xylopyranoside (data not shown) although only against *pNP-Araf* the activity was detected and applied to further analysis. The activity on arabinoxylan and arabinan reveals its capacity to release arabinose from the branches of water-soluble α -L-Araf-(1 \rightarrow 3) and α -L-Araf-(1 \rightarrow 2) bonds [47]. Beyond arabinoxylan, it is important to note that the enzymatic activity in branched arabinan, which contains a backbone of (1 \rightarrow 5)- α -L-Araf linkage, shows that *TtAbf62* may be acting in both polysaccharides from hemicellulose and pectin of sugarcane cell wall sources. The preference of the enzyme to hydrolyze substrates decorated with Araf- α -(1 \rightarrow 3) and Araf- α -(1 \rightarrow 2) bonds is clear since the significant activity is observed for WAX, indicating that the arabinosyl side chain is critical for the substrate recognition [48], making it plausible to classify the enzyme as arabinoxylan arabinofuranohydrolases [49].

TtAbf62 exhibits high activity for *pNP-Araf* over the pH range from 4.5 to 5.0 (Fig. 2a), with an abrupt decrease in its catalytic strength for pH values lower than 4.0 and over 5.0, suggesting an activity sensitive to pH. Since the DSF results pointed high structural stability for *TtAbf62* towards pH changes, the decrease in the enzyme activity should not be related with the loss in structural stability, but probably due to an impact on the overall charge of the amino acid residues involved in its activity.

When analyzed the optimal temperature of activity the highest values are in a range between 25 and 35 °C (Fig. 2b). Even being a thermophilic fungus, these values are compatible with the optimum pH and temperature previously observed for GH62 family towards *pNP-Araf* as a substrate, e.g.: *Podospora anserina* pH 5.0 and 37 °C [6]; *Ustilago maydis* pH 5.0 and 37 °C [6] (Table 2).

Furthermore, the enzyme stability was assessed by measuring the residual activity after prolonged periods (up to 36 h) of incubation at 30

and 50 °C (Supplementary Fig. A.3). At 50 °C, the enzyme maintained a residual activity of 30.6% after 36 h, and in the same period, the residual activity at 30 °C was 43.5%. Even at the higher temperature tested (50 °C), the enzymatic activity demonstrated that the *TtAbf62* tolerates this condition for extended periods, whereas the complete inactivation was not observed, also the complete inactivation at 30 °C was not observed as well.

3.3.2. Kinetic parameters

The kinetic parameters were determined on substrates with all Adj. R-square greater than 0.95 where *TtAbf62* had the uppermost activity based on the substrate panel: *pNP-Araf*, WAX and SBA (Table 1). As described in Table 1 equivalent values of k_{cat} was achieved for the substrates SBA (with a $K_M = 9 \pm 3$ mg.mL⁻¹) and *pNP-Araf* ($K_M = 4 \pm 1$ mM⁻¹), wherein the *TtAbf62* shows higher affinity by *pNP-Araf*, as well as other enzymes previously characterized. Comparing k_{cat} towards arabinan 5.8 ± 0.8 s⁻¹, and *pNP-Araf* 6.8 ± 0.3 s⁻¹, the turnover number is in the same order of magnitude, suggesting the similar conversions rate independently of the polymeric or synthetic substrate. Nevertheless, the enzymatic preference (k_{cat}/K_M) is higher for *pNP-Araf* when compared to SBA, whereas *SthAbf62A*, *AnAbf62A* and *ScAbf62A* (Table 1) prefer WAX. *TtAbf62* also presented the highest catalytic efficiency and enzymatic activity (U mg⁻¹) in *pNP-Araf* and SBA when compared to other enzymes as described in Table 1.

The activity towards *pNP-Araf* can be explained by the insights from the structural analysis of the *TtAbf62* (section 3.4), which has the catalytic triad composed of two aspartates (Asp58 and Asp165) and one glutamate (Glu217) coordinating the substrate. Also, the residues Asp58 and Asp165 are acting as catalytic residues, interacting with a *p*-nitrophenolate ring that stabilizes the intermolecular interactions.

3.3.3. Activity of *TtAbf62* in sugarcane bagasse

The sugarcane cell wall contains 40% arabinoxylan (AX), where arabinoside side chain may cause a steric hindrance for xylanases [55]. The hydrolysis with *TtAbf62* was qualitatively tested against natural biomass to estimate the potential of this enzyme to act on complex substrates. Evaluated by DNS method [27], the addition of *TtAbf62* resulted in a higher concentration of reducing terminals (4.5 μ mol glucose equivalent), compared with the enzymatic cocktail itself (0.2 μ mol glucose equivalent). This might be related to its ability to remove arabinose side chains from the AX, as well as arabinan, decreasing steric restraint in branched polysaccharides, making it more susceptible to hydrolysis by exo and endoxylanases.

3.4. Structural characterization

3.4.1. Overall structure

After the crystallization process, orthorhombic crystals of *TtAbf62* were obtained ($I222$; $a = 75.267$ Å, $b = 79.747$ Å, $c = 82.922$ Å) and diffracted up to 1.87 Å resolution. The structure was determined by molecular replacement with one molecule in the asymmetric unit. The protein chain was comprised of 303 amino acid residues (from Gly28 to Arg331), two ions: calcium (Ca²⁺), sulfate (SO₄²⁻), and one molecule of ethylene glycol (EDO). All amino acids and ligands built in the structure exhibited a well-defined electron density. Final crystallographic R_{work} and R_{free} values were 14.7% and 19.5%, respectively, with more than 96% of the residues located on the sterically allowed regions of the Ramachandran plot. Further refinement statistics are described in Table 2.

The *TtAbf62* asymmetric unit is shown in Fig. 3a. The N-terminus began in a loop from Gly28 to Thr54 residue, followed by five β -sheets (called blades I to V), usually composed of four antiparallel β -strands. Blades I, II, and III are composed of four antiparallel β -strands; blade IV by five mixed β -strands plus a η 4 helix; and blade V by three antiparallel β -strands, respectively, forming a set of 20 β -strand, from β 1 to β 20, linked by loops with variant lengths. Also, five helices (called η 1 to

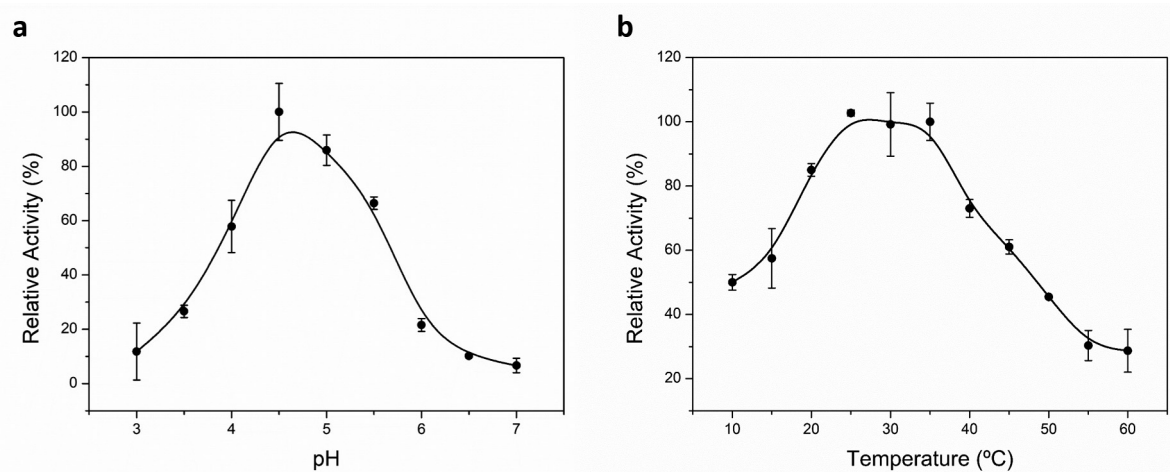


Fig. 2. Optimal pH and temperature of *TtAbf62* for *pNP-Araf*. In both cases (a) and (b), the solid line is just a guide for the eye.

Table 1

Kinetic parameters of *TtAbf62*. Substrate specificity and kinetic data of GH62 enzymes with known structure deposited on PDB. *pNP*: para-nitrophenol, *pNP-Araf*: para-nitrophenyl α -L-arabinofuranoside, *pNP-X*: para-nitrophenyl β -D-xylopyranoside, *pNP-Arap*: *pNP*- α -L-arabinopyranoside, *SBA*: sugar beet arabinan, *WAX*: wheat arabinoxylan, *HV*: high viscosity, *LV*: low viscosity. Sp. act.: Specific activity, U: unit defined as the release of 1 μ mol of product per min per mg of the substrate under the specific assay conditions. *values are normalized according to the unit.

Source, (protein ID), accession no.	Assay conditions, substrate	Activity	K_M	k_{cat}	k_{cat}/K_M	Reference
<i>Thermothielavioides terrestris</i> (<i>TtAbf62</i>), AEO64662.1	pH 4.5, 30 °C	Sp. Act (U mg ⁻¹)	mM	s ⁻¹	s ⁻¹ .mM ⁻¹	This study
	<i>pNP-Araf</i>	0.78	4 ± 1 mg.mL ⁻¹	6.8 ± 0.3 s ⁻¹	2 s ⁻¹ .mg ⁻¹ .mL	
	WAX	0.12				
<i>Streptomyces coelicolor</i> A3 (<i>ScAbf62A</i>), CAA16189.1	pH 5.5, 30 °C		mM	s ⁻¹ *	s ⁻¹ .mM ⁻¹ *	[46]
	<i>pNP-Araf</i>		1.9 mg.mL ⁻¹	0.02 s ⁻¹	0.01 s ⁻¹ .mg ⁻¹ .mL	
	WAX		7.3	0.30	0.04	
<i>Podospora anserina</i> S mat+ (<i>PaAbf62A</i>), CAP62336.1	pH 5.0, 37 °C	Sp. Act (U mg ⁻¹)	mM	s ⁻¹ *	s ⁻¹ .mM ⁻¹ *	a [6]; ^b [50];
	<i>pNP-Araf</i>	0.61 ^a 0.38 ^b	6.1 ^{a,b}	0.38 ^{a,b}	0.06 ^{a,b}	
	WAX-LV	1 ^b				
<i>Streptomyces thermoviolaceus</i> OPC-520 (<i>SthAbf62A-m2,3</i>), BAB84113.1	pH 7.0, 55 °C	Sp. Act (U mg ⁻¹)*	mM	s ⁻¹	s ⁻¹ mM	[2,51]
	<i>pNP-Araf</i>	0.003	3	1.3	0.5	
	<i>pNP-X</i>	7.8·10 ⁻⁶				
	<i>pNP-Arap</i>	5.2·10 ⁻⁷	8.2 mg.mL ⁻¹	0.01 s ⁻¹	0.001 s ⁻¹ .mg ⁻¹ .mL	
	WAX-HV	0.04	12	180	15	
<i>Ustilago maydis</i> 521 (<i>UmAbf62Am2,3</i>), EAK85571.1	pH 5.0, 37 °C ^a	Sp. Act (U mg ⁻¹)	mM	s ⁻¹ *	s ⁻¹ .mM ⁻¹ *	a [6]; ^b [52]
	pH 4.5, 40 °C ^b					
	<i>pNP-Araf</i>	0.06 ^a ; 0.012 ^b	7.5 ^a	0.04 ^a	0.005 ^a	
	WAX-LV	30 ^b				
<i>Aspergillus nidulans</i> FGSC A4 (<i>AnAbf62A-m2,3</i>), EAA59562.1	pH 5.5, 37 °C	Sp. Act (U mg ⁻¹)			s ⁻¹ .mM ⁻¹	[53]
	<i>pNP-Araf</i>	1.6			0.26 s ⁻¹ .mg ⁻¹ .mL	
	WAX-LV	67	4.9 mg.mL ⁻¹	s ⁻¹	36 s ⁻¹ .mg ⁻¹ .mL	
<i>Scytalidium thermophilum</i> CBS 625.91 (<i>StAbf62Cm2,3</i>), AHZ56660.1	pH 7.0, 50 °C	Sp. Act (U mg ⁻¹)				[47]
	<i>pNP-Araf</i>	0.02			s ⁻¹ .mg ⁻¹ .mL	
	WAX-HV		3.66 mg.mL ⁻¹	0.28 s ⁻¹	0.07 s ⁻¹ .mg ⁻¹ .mL	
<i>Coprinopsis cinerea</i> (CcAbf62A), BAK14423.1	pH 7.0, 40 °C	Sp. Act (U mg ⁻¹)*				[54]
	WAX	1.5 ± 0.4				

Table 2

Data collection and refinement statistics. Values in parentheses are for the highest resolution shell.

Name	TtAbf62 (PDB ID: 6CC7)
Data collection	
Beamline	W01A-MX2, LNLS
Wavelength (Å)	1.4586
Space group	I222
a, b, c (Å)	75.27, 79.75, 82.92
V _M ^a (Å ³ .Da ⁻¹)	1.79
Solvent content (%)	31.3
Resolution (Å)	33.05–1.87 (1.91–1.87)
Unique reflections	20,946 (1318)
Completeness (%)	99.8 (97.4)
Multiplicity	5.5 (2.8)
CC _{1/2} (%)	99.8 (83.8)
R _{merge} (%)	7.1 (38.6)
R _{meas} (%)	8.4 (52.2)
R _{p.i.m.} (%)	4.5 (34.9)
Mean I/σ(I)	11.9 (2.2)
Refinement	
Resolution (Å)	57.48–1.87
No. of reflections	19,767
Completeness (%)	99.65
R _{work} (%)	14.74
R _{free} ^b (%)	19.53
Average B factor (Å ²)	23.69
Protein	23.78
Waters	35.26
Metal	39.66
Inorganic	39.33
Organic	29.19
No. of protein atoms	2334
No. of solvent atoms	298
No. of molecules in the asymmetric unit	1
R.m.s.d. ^b , bonds (Å)	0.014
R.m.s.d., angles (°)	1.628
Ramachandran plot	
Favored (%)	96.36
Allowed (%)	3.31
Outliers (%)	0.33
Clashscore	0.44
MolProbity score [38]	0.90

^a Matthews coefficient.

^b The test set uses ~5% of the data.

η5) are distributed along the chain (Fig. 3b).

Analysis using PISA server [39] showed that the opened interface adopted by the structure of TtAbf62 allows contacts between chains related by a symmetry operation ($-x + 1, y, -z$) throughout 125 residues (41.3% of the whole structure). These interactions embrace a variety of hydrogen bonds, a salt bridge between His283A and Asp58B ($d_{ND1-OD1} = 3.10$ Å), and also a disulfide bond between Cys32A (from chain A) and Cys300B (from chain B) with a distance d_{S-S} of 2.03 Å. This interchain disulfide is connecting a loop (formed by residues 28 to 54) located at the N-terminus of chain A (Cys32A) and a η4 located at the C-terminus of chain B (Cys300) (Fig. 4).

The disulfide bond parameters are calculated according to the empirical formula described by Katz et al. [56] and Schmidt et al. [57] (<https://services.mbi.ucla.edu/disulfide/>). Interestingly, the presence of a disulfide bond summed with the opened interface adopted by the structure provided the possibility of an exquisite phenomenon of domain swapping, first observed in the GH62 family. This domain swapping generates an interface of contacts among neighbor protein chains involving blade I from chain B (blade IB) fitted onto chain A and blade I from chain A (blade IA) shuffled to chain B. Additionally, this phenomenon also relies on the so-called “molecular velcro”, joining the N-terminal β1 in one of the modular β-sheets (β20) via hydrogen bonds [58]. However, in this particular case, the molecular velcro is arranged by hydrogen bonds connecting the loop (28–54) from N-terminal upon β20 from C-terminal (details in Supplementary Fig. A.4). As a result, the

enzyme TtAbf62 adopts a homodimeric blade-swapped β-propeller fold (see scheme in Supplementary Fig. A.5), each part containing five blades (I to V) radially arranged in two pseudo-5-fold axes wherein the beta-sheets (β1 to β20) are repeated, as shown in Fig. 4a and Supplementary Fig. A.5, maintaining the catalytic residues on its appropriated position (Fig. 4b).

The domain swapping occurs when two or more molecules make a partial exchange of domains among their structures, usually to form oligomers, from their monomeric subunits [59–63]. It represents not only an isolated event of structural innovation but differently, these swaps enable the creation of new interfaces of interactions previously absent in the monomer or a closed arrangement [59]. Considered an important innovative aspect in the structure and function of a protein, even been commonly found, such a mechanism is still poorly understood structurally.

A search for similar structures of TtAbf62 was performed using the Dali server [64] (Supplementary Table A.2) and the results are displayed as aligned sequences with graphical enhancements performed by ESPript 3.0 [65] (Supplementary Fig. A.6). Comparisons to similar structures from *Streptomyces coelicolor* ScAbf62A (PDB ID 3WMY; 3WN0; 3WN1; 3WN2) [46], *S. thermoviolaceus* SthAbf62A (PDB ID 4O8N; 4O8O and 4O8P) [2], *Podospora anserina* PaAbf62A (PDB ID 4N4B and 4N2Z), *Ustilago maydis* UmAbf62A (PDB ID 4N1I and 4N2R) [6], *Aspergillus nidulans* AnAbf62A (PDB ID 5UBJ) [9], *Scytalidium thermophilum* SthAbf62C (PDB ID 4PVA and 4PVI) [47], *Coprinopsis cinerea* CcAbf62A (PDB ID 5B6S) [66] and *Talaromyces pinophilus* TpAbf62 (PDB ID 6F1J) [67].

3.4.2. Catalytic pocket description

In-depth analysis side-by-side of structurally characterized GH62 enzymes revealed that the catalytic triad [68]: an aspartate (acting as a general acid); glutamate (acting as a general base); and an aspartic acid supposedly modulating the pKa of the catalytic general acid, as observed in many crystallographic structures [2,6,46] and supported by site-directed mutagenesis and kinetic data [2,46], were also conserved in TtAbf62, named Asp58 (from β1), Asp165 (from β9), and Glu217 (from β13) (Fig. 4b, Supplementary Table A.3, and Fig. A.4). These residues interact with the ligands similarly as described in ScAbf62A (3WN0) [46], SthAbf62A (4O8O) [2] and UmAbf62A (4N2R) [6] for an arabinose residue; in SthAbf62A (4O8P) [2] for a xylotetraose residue; in ScAbf62A (3WN1) [46] and SthAbf62C (4PVI) [47] for xylotriose residues; in ScAbf62A (3WN2) [46] for a xylohexaose residue; and in PaAbf62A (4N2Z) [6] for a cellotriose residue. They are also conserved independently of their subclassification of the family, GH43, and GH62. For TtAbf62 (6CC7), a molecule of ethylene glycol (EDO, from the crystallization condition) is observed at the catalytic site, interacting via hydrogen bonds directly with Lys57 ($d_{NZ-O1} = 2.84$ Å) and Asp58 ($d_{OD2-O1} = 2.72$ Å).

Alignments of the three-dimensional structures (Fig. 5) and amino acid sequences (Supplementary Fig. A.6) revealed that in the catalytic pocket, the residues at the position equivalent to 57 and 283 in TtAbf62 (6CC7) are conserved in all structures, except in TpAbf62 (6F1J), since the Lys57 and His283 were mutated by an arginine (Arg) and serine (Ser), respectively. The key residues at position 58 (acting as a general acid), position 125 and 165 (acting as a pKa modulator), position 217 (acting as a general base), and positions 307 and 317 are fully conserved in all available structures.

The calcium ions (Ca²⁺) listed in Table A.3 are usually hydrogen-bonded to a histidine residue (Ca-His), equivalent to His283 (Supplementary Fig. A.6), with a few exceptions for TtAbf62 (6CC7) and SthAbf62C (4PVA and 4PVI). The presence of this histidine residue in the catalytic pocket has been proposed to be an asset in combination with the catalytic triad residues [47].

Some particular water molecules (H₂O) also appear to play a significant role in the catalytic site. In TtAbf62, the H₂O 533, 542, 574, and 594 are present, and their equivalent positions in other structures

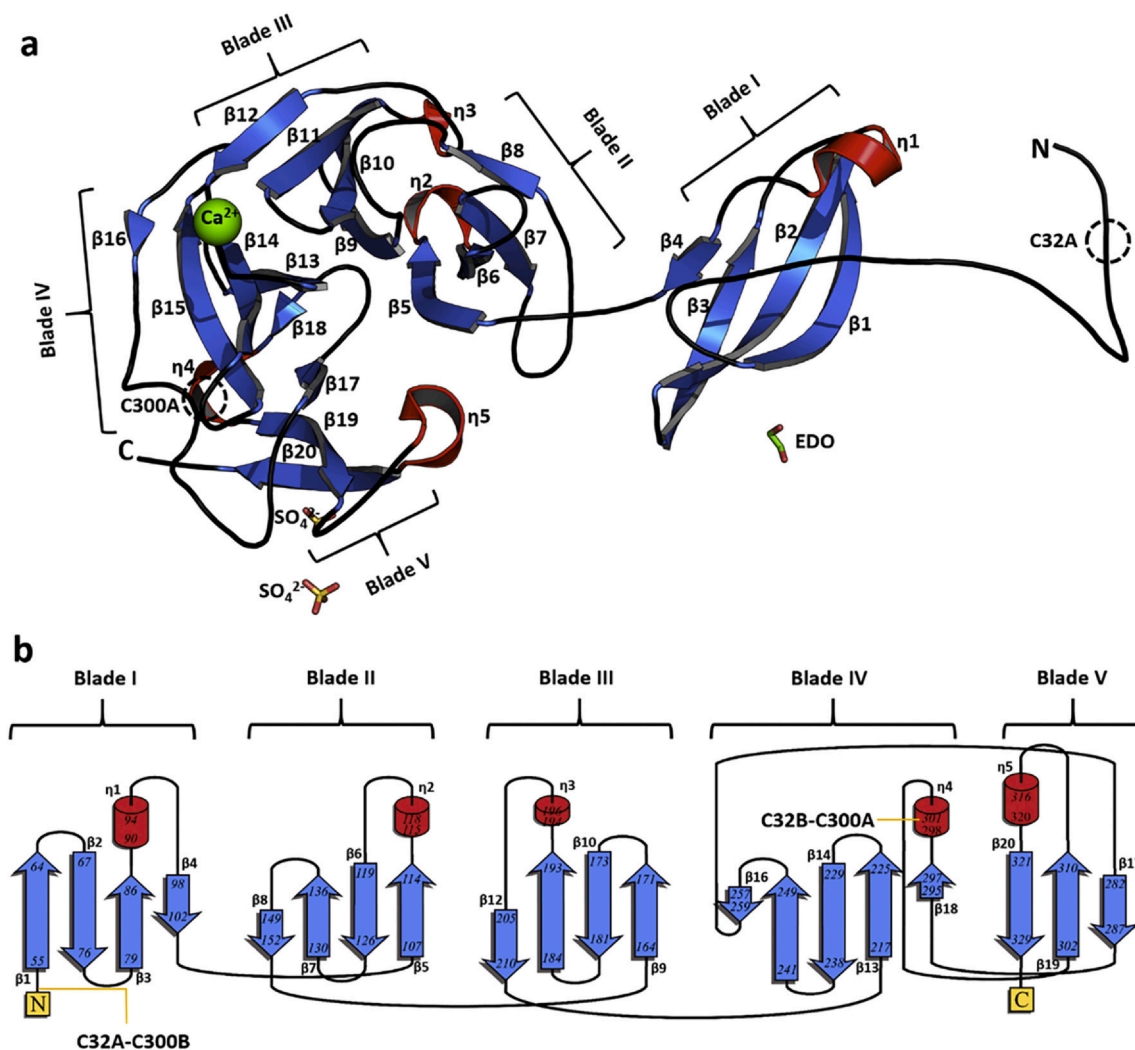


Fig. 3. Topology and top view of the asymmetric unit. a) The overall structure of *TtAbf62* revealing a five-bladed β -propeller fold, top view showing the numbering of 5 blades (blades I to V), a disulfide bond between Cys32A and Cys300B, an ethylene glycol (EDO) molecule located at partial catalytic cleft correspondent to channel of the fold, two sulfate ions (SO_4^{2-}) close to Blade V, and a calcium ion (Ca^{2+}) coordinated close to Blade IV. b) The overall topology of *TtAbf62* revealing a five-bladed β -propeller fold. Amino acid residue numbers are given at each start and end of the respective secondary structure elements, β -strands are indicated as arrows β 1– β 20 (shown in blue) and helices are indicated by cylinders η 1– η 5 (shown in red), respectively.

are often connected to the His283 (His-H₂O) as shown in 3WN1, 3WN2, 4PVI and 5B6S. The H₂O542 intermediates an interaction between residues Gln307 and Tyr317 (Gln-H₂O-Tyr), however, it was not observed in 3WN0 and 4N2R, where the water molecule is replaced by β -L-arabinofuranose, and for 5UBJ and 4PVA was replaced by a calcium and phosphate ion, respectively. The list of structural elements in the catalytic pocket for the Abfs GH62 family is presented in Supplementary Table A.3.

The calcium ion (Ca^{2+}) presented in the structure of *TtAbf62* is located in a completely different position when compared to other solved structures for the same family. In *TtAbf62* the Ca^{2+} is located at the interface of contacts between two dimeric arrangements, specifically by the residue Thr212 and (H₂O)730 both within 3.9 Å of distance as detailed in Supplementary Fig. A.7a. Furthermore, the chelation of Ca^{2+} by EDTA had no apparent effect on the enzymatic activity (data not shown). Fortuitous ligands from crystallization condition, the sulfate ions (SO_4^{2-}) 402 and 403 are coordinated on the surface of the structure by residues Trp91 and Asn312, respectively, with the addition of some water molecules (Supplementary Fig. A.7b-c).

3.4.3. Structural insights

According to CAZY, all the structures of characterized GH62 family

share a high degree of similarity, adopting the β -propeller fold with five blades (I-V), and their active site is composed of previously highlighted catalytic triad [68], that often interacts with the substrate.

The structure of *TtAbf62* was compared to other Abfs belonging to the family GH62 in complex with xylooligosaccharides and revealed that their orientations were similar (Supplementary Fig. A.8). According to Maehara et al. [46], the residues Trp270 and Tyr461 in *ScAbf62A* are interacting with xylotriose (PDB ID 3WN1) and xylohexaose (PDB ID 3WN2) ring at subsites +3NR and +1, respectively. The same positions were found in *TtAbf62*, corresponding to Trp126 and Tyr317, respectively. The residue Asn318 in *TtAbf62* (PDB ID 6CC7), also has the same orientation as described in the structures complexed with xylooligosaccharides: *ScAbf62A* (PDB IDs 3WN1 and 3WN2) [46] with an equivalent position of Asn462; Asn339 in *SthAbf62C* (PDB ID 4PVI) [47]; Asn321 in *PaAbf62A* (PDB ID 4N22) [6]; and Asp350 in *SthAbf62A* (PDB ID 4O8P) [2]. Amino acids at this position are usually involved in the hydrogen bond to subsite +2NR and reported as responsible for affecting the activity specifically in polysaccharides [46].

To demonstrate the most likely bio-assembly of *TtAbf62* in solution, SEC-MALS and cross-linking experiments were carried out. Differently of the crystal structure, assays with *TtAbf62* in solution demonstrate

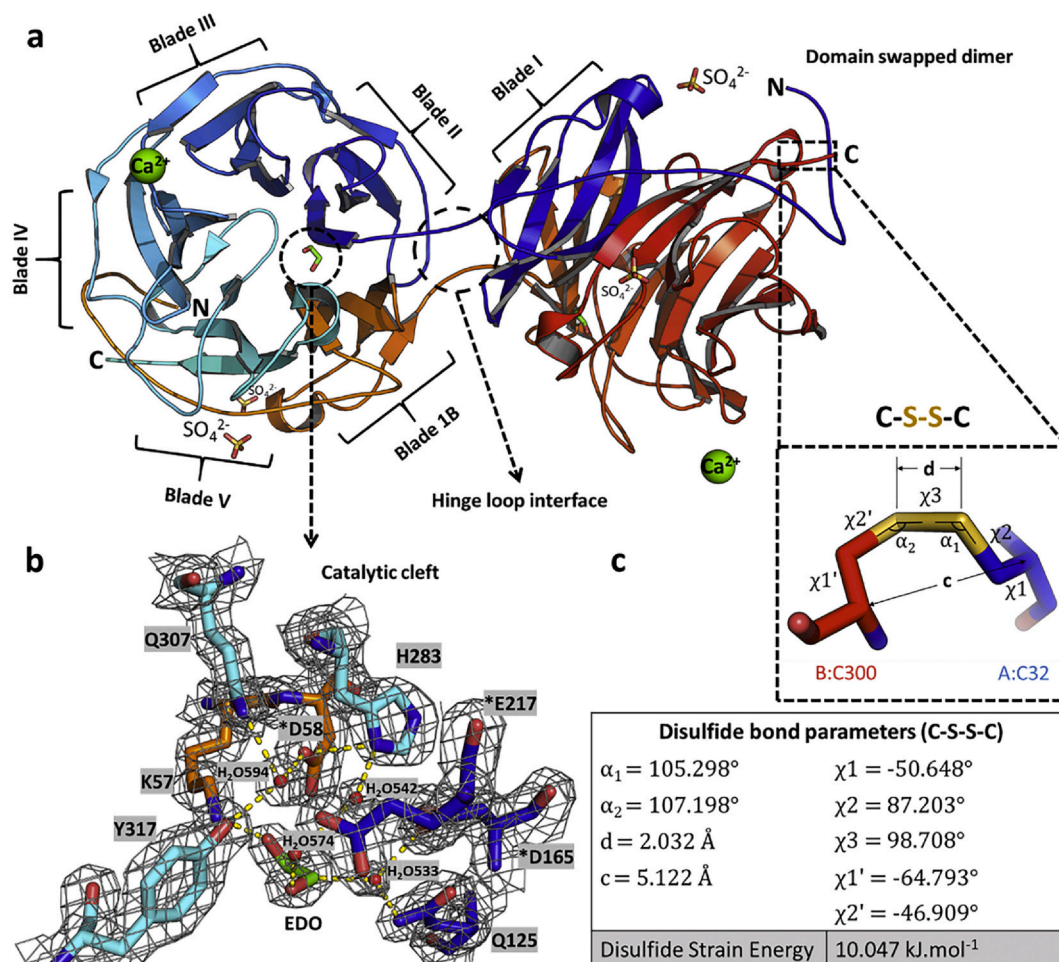


Fig. 4. Detailed view of the swapped domain from TtAbf62, its catalytic pocket, and disulfide bond. a) TtAbf62 represented as a homodimer through a domain swap. In detail, the N- (dark blue) and C- (dark red) terminus from their monomer chains respectively, which are connected via a disulfide bond. Five blades of β -strands can be observed assembling the dimeric system required for this structure. Calcium ions (Ca^{2+}) are represented as spheres colored in green; sulfate ions (SO_4^{2-}) and ethylene glycol (EDO) ligands can also be observed. b) Detailed view of the catalytic pocket with hydrogen bonds between water molecules mediating interactions to EDO and also a salt bridge evolving H283 and *D58 residues. Isomesh level 1.0. c) Detailed view of the disulfide bond between A:C32 (belonging to the N terminus) and B:C300 (belonging to the C terminus) with its respective parameters.

evidence of a monomeric arrangement, suggesting that a closed interface regarding all blades in a classical β -propeller fold is most reliable.

In general, the domain swapping phenomena has no clear explanations, and one possible reason could be due to specific conditions that the protein had been exposed [60]. These conditions can be justified through some processes such as pH and temperature changes, mutation, and denaturing agents [63,69]. Also, the physiological conditions of *T. terrestris* are at acidic pH and temperature of 45 °C [16], providing the basis to believe that the new molecular arrangement (dimeric) of the enzyme with domain swapping and the complement of catalytic pocket by surrounding molecule to be an artifact arising from molecular alterations caused by the environmental and physiological conditions during crystallization.

3.5. Size exclusion chromatography with multiangle light scattering (SEC-MALS) and glutaraldehyde cross-linking

When analyzed by SEC-MALS, TtAbf62 exhibited a symmetrical peak with an average molecular mass of (32.1 ± 0.2) kDa (Fig. 6b). The expected theoretical molecular mass under denaturing conditions for TtAbf62 is 32.8 kDa, which represents only 2% of error for the estimative mass. This result indicates that TtAbf62 is likely a monomer in solution; despite the crystal structure has shown a possible homodimer

or even a homotetramer (if considering the dimer-dimer interface for Ca^{2+}). According to Yang et al. [70], domain swapping happens at high protein concentration where a monomer becomes less favorable entropically and proteins tend to form homodimers. The exchanged blade generates an oligomerization in such matter to afford to rise in protein stability by providing additional contacts between chains [71]. Therefore, measurements of the molecular mass for a sample with higher protein concentration, close to the condition of the drop in the condition of crystallization were carried out. Thus, the graphs of Fig. 6b-c show two more concentrations: 2 mg.mL^{-1} in buffer A and 54 mg.mL^{-1} in 1:1 buffer A and crystallization solution, respectively, with both confirming the protein as a monomer in solution. Although there are small peaks with a mass of approximately (51 ± 1) and (52 ± 2) kDa, they do not correspond to another oligomeric state of TtAbf62, but rather a contaminant not eliminated during the purification processes, as can be seen in the SDS-PAGE gel of Fig. 6a in two concentrations: 2 and 54 mg.mL^{-1} .

Moreover, TtAbf62 was also analyzed by SDS-PAGE after glutaraldehyde cross-linking reaction, and again a single band corresponding a monomer predominates, indicating that TtAbf62 does not form another oligomeric state in solution (Fig. 7). To conclude, TtAbf62 is a monomer in solution, adopting a more compact structure like all other proteins from the GH62 family.

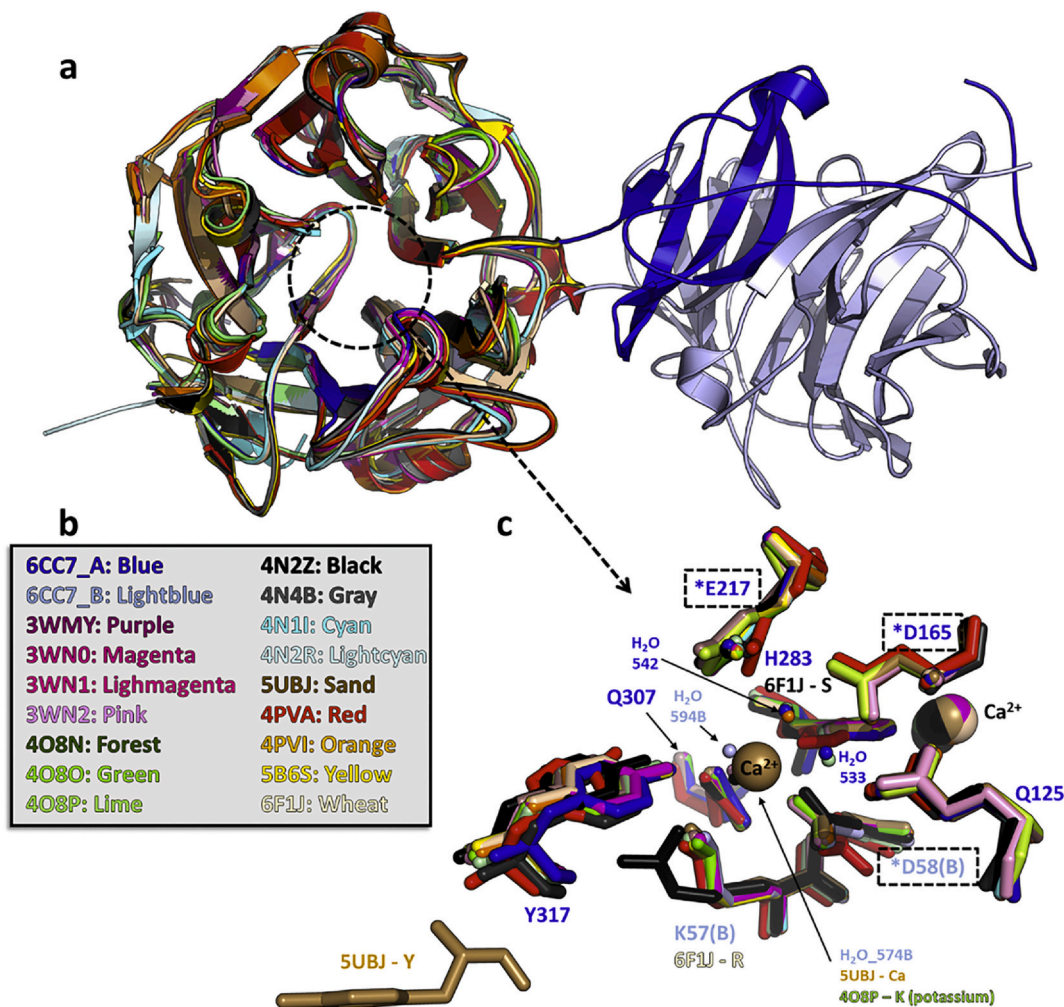


Fig. 5. Structural superposition of *TtAbf62* (6CC7) with representative Abf from the GH62 family with similar three-dimensional structures. a) Superimposition of representative Abf structures; b) Color scheme adopted for identification of the structures; c) Residues in the active site of similar Abf with their equivalent in *TtAbf62*, represented by one letter code (colored as described in b). The catalytic residues *D165 and *E217 belong to chain A of *TtAbf62* and residue *D58 belongs to chain B of *TtAbf62*, forming its catalytic site. Water molecules with structurally conserved positions are represented as small spheres (colored by chains and structure as described in b). Calcium ions (Ca²⁺) are represented as spheres with arbitrary radii colored by chains. Residues are labeled by one-letter code. Additional structural comparisons of *TtAbf62A* are summarized in Supplemental Table A.3.

4. Conclusion

The enzyme *TtAbf62* catalyzes the hydrolysis of arabinose residues of the arabinan branches, arabinoxyylan as well as acts in sugarcane bagasse *in natura*, whereas, suggesting a great versatility of the fungal enzyme against the natural substrates, emerging its biotechnological potential. The structural differences assisted in the clarification of the reasons for these observed functional differences. Sequence alignments of glycoside hydrolase GH62 show conserved blocks. The catalytic pocket was structurally conserved between the 5-bladed-propeller fold. The β -propeller architecture followed a pattern, with the junction of C and N-terminus blades forming a pseudo-5-fold symmetry, which confers protein stability. This stability is also strengthened by disulfide bonds and the “non-classical molecular velcro”. The structural outcome for *TtAbf62* is unexpected, considering the predominant monomeric nature of the homologous proteins adopting classical β -propeller folds. Therefore, the possibility of evolution or mere crystallization artifact on protein oligomerization for this family of enzymes was shown. Such things may lead to its thermostability to be related to the folding. All analyses presented conclude that the quaternary structure of the enzyme *TtAbf62*, in crystal, is in a dimeric arrangement, once the interdependence between the catalytic residues and domain-swapped (blade

I) are necessary for complementarity, and it is only possible through this molecular architecture. Therefore, the domain swapping reveals a new structure within the GH62 family proteins and suggests a novel structural arrangement. The nature of this behavior may be related to the favorable entropic as a form of storage or compacting of the enzyme in high concentrations.

Author contributions

S.C., E.J.M., L.R.A., and J.R.C.M. designed the research. S.C. and E.J.M. purified *TtAbf62*. L.R.A. and J.R.C.M. prepared cryo-crystals samples and collected datasets. S.C., E.J.M., A.B., and F.S. characterized *TtAbf62* enzymatically. W.G. characterized *TtAbf62* biophysically. L.R.A. and J.R.C.M. processed X-ray diffraction images and refined the atomic coordinates. S.C., E.J.M., L.R.A., R.P., and J.R.C.M. wrote the paper with support from all other authors.

Footnote

This article contains supplementary data: SM1, SM2, Table A.1, Table A.2, Table A.3, Fig. A.1, Fig. A.2, Fig. A.3, Fig. A.4, Fig. A.5, Fig. A.6, Fig. A.7, and Fig. A.8.

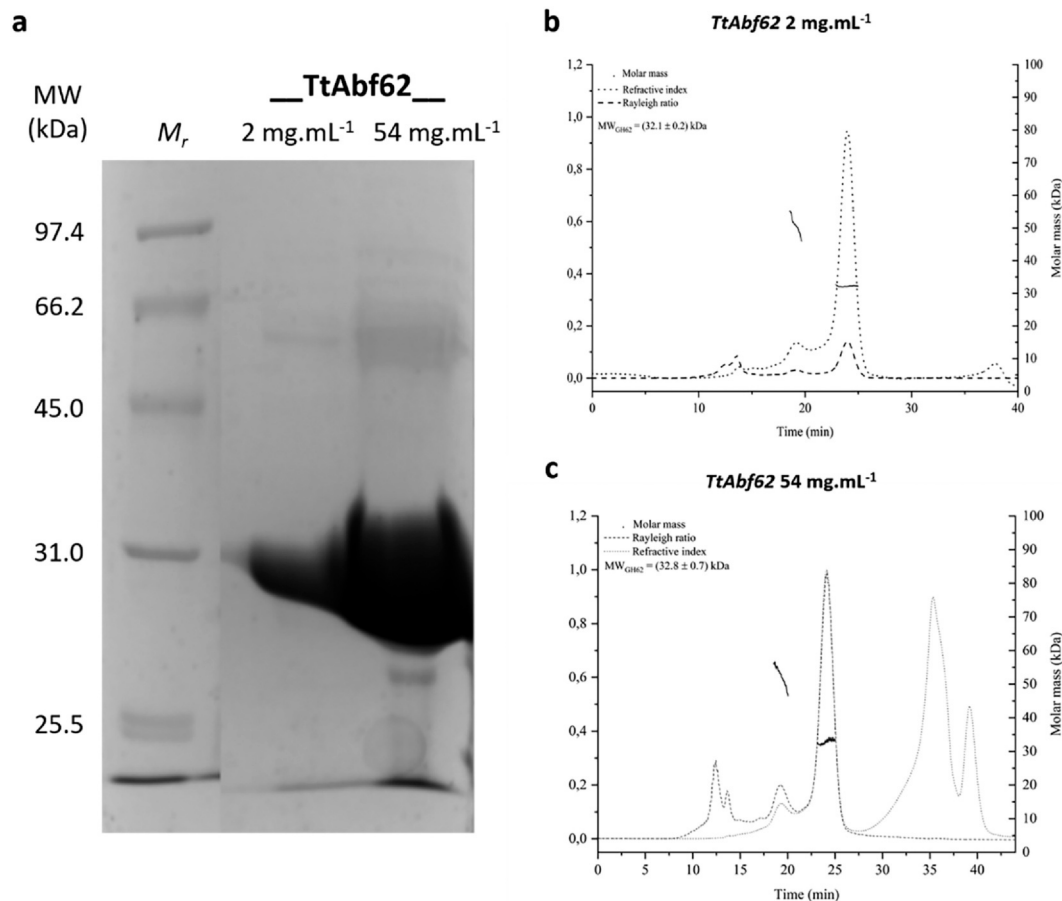


Fig. 6. Multiangle Light Scattering (SEC-MALS) analysis for different concentrations of *TtAbf62* in solution. a) 12% SDS-PAGE after SEC. M_r – molecular mass; lanes 2 and 3 show *TtAbf62* in two different concentrations: 2 and 54 mg.mL⁻¹, respectively. SEC-MALS of purified *TtAbf62* at 2 mg.mL⁻¹ (b) and 54 mg.mL⁻¹ (c). For all samples, the small peak represents a possible contaminant confirmed by SDS-PAGE (a) with the highest peak representing *TtAbf62* in the monomeric state in solution in agreement with the theoretical molar mass of 32.8 kDa. MALS calculations are dependent on a good estimate of the refractive index increment (dn/dc) of the sample, near to 0.185 for proteins [72]. The peaks in (b) correspond to (51 ± 1) and (32.1 ± 0.2) kDa, respectively; peaks in (c) correspond to (52 ± 2) and (32.8 ± 0.7) kDa, respectively.

The atomic coordinates and structure factors (code 6CC7) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

Declaration of Competing Interest

All authors of this work declare that have no potential conflict of interest and that there is no financial, consultant, institutional or other relationships that might lead to bias or conflicts of interest in this research. Financial grants, infrastructure and fellowships supporting this

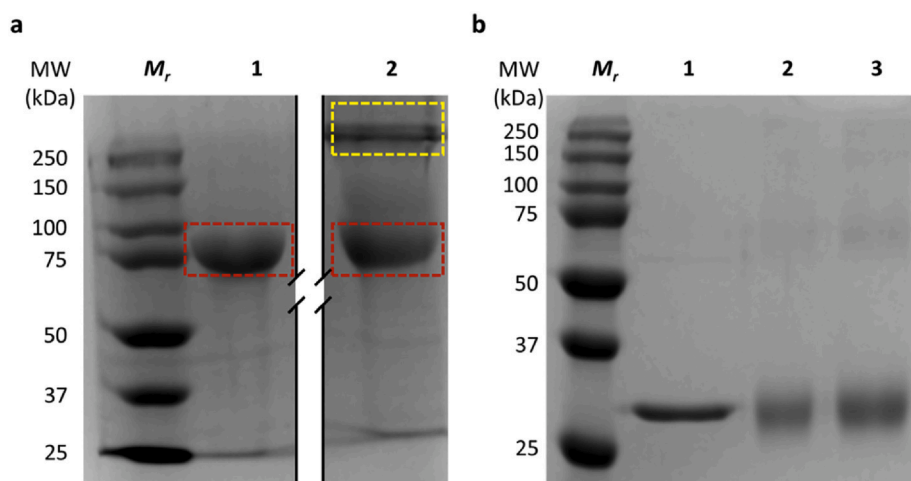


Fig. 7. Glutaraldehyde Cross-linking reaction. a) 10% SDS-PAGE of the reaction (positive control) using a known sample of a homodimer in solution; M_r – molecular mass reference; 1 – protein control of 84.9 kDa (monomeric, red box); 2 – after crosslink control reaction with 5 μL of glutaraldehyde mixture (dimers - yellow box and monomers - red box). b) 15% SDS-PAGE after reaction with *TtAbf62*. M_r molecular mass; 1 – only the purified protein. The *TtAbf62* after crosslink reaction with 5 μL and 10 μL of glutaraldehyde mixture: lanes 2 and 3, respectively.

work are described below.

Acknowledgments

We would like to thank the National Synchrotron Light Laboratory (LNLS, Brazil); CNPEM – National Center for Research in Energy and Materials; São Paulo Research Foundation (FAPESP) for the financial support via grants # 2017/16291-5, 2014/50897-0 (JRCM), 2014/18714-2, 2017/50025-0, 2019/22284-7 (FS) and 2016/09152-6 (LRA); São Carlos Institute of Physics/University of Sao Paulo; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the financial support via grant # 308865/2018-9 (JRCM) and 302627/2018-9 (FS); and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for fellowship to SC and EJM.

Appendix A. Supplementary data

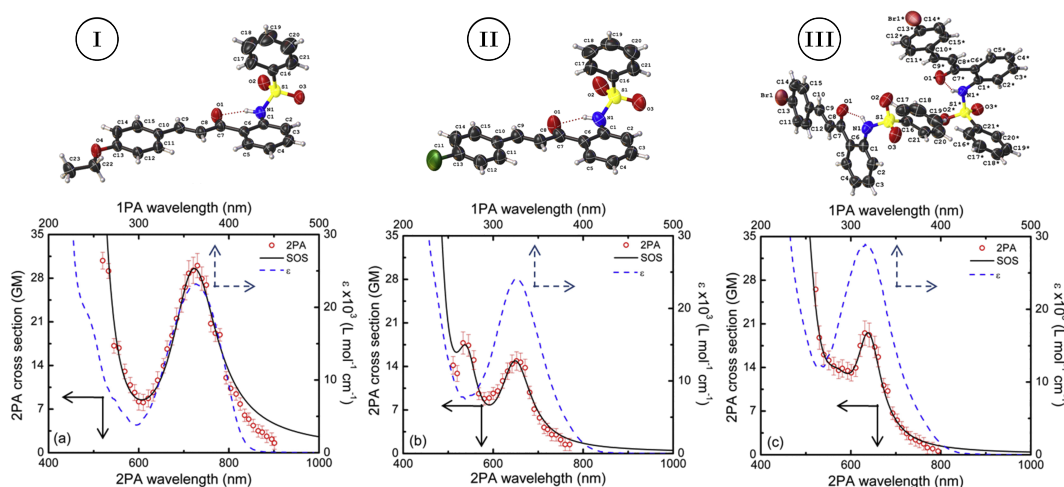
Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bbapap.2020.140533>.

References

- [1] C. Wilkens, S. Andersen, C. Dumon, J.-G. Berrin, B. Svensson, GH62 arabinofuranosidases: structure, function and applications, *Biotechnol. Adv.* 35 (6) (2017) 792–804, <https://doi.org/10.1016/j.biotechadv.2017.06.005>.
- [2] W. Wang, G. Mai-Gisondi, P.J. Stogios, A. Kaur, X. Xu, H. Cui, O. Turunen, A. Savchenko, E.R. Master, Elucidation of the molecular basis for arabinoxylan-debranching activity of a thermostable family GH62 alpha-L-arabinofuranosidase from *Streptomyces thermoviolaceus*, *Appl. Environ. Microbiol.* 80 (2014) 5317–5329.
- [3] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, The carbohydrate-active EnZymes database (CAZy): an expert resource for glyco-genomics, *Nucleic Acids Res.* 37 (2009) D233–D238.
- [4] V. Lombard, H.G. Ramulu, E. Drula, P.M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, *Nucleic Acids Res.* 42 (2014) D490–D495.
- [5] D.G. Naumoff, Hierarchical classification of glycoside hydrolases, *Biochem.* 76 (2011) 622–635.
- [6] B. Siguier, M. Haon, V. Nahoum, M. Marcellin, O. Burlet-Schiltz, P.M. Coutinho, B. Henrissat, L. Mourey, M.J. O'Donohue, J.-G. Berrin, S. Tranier, C. Dumon, First structural insights into alpha-L-arabinofuranosidases from the two GH62 glycoside hydrolase subfamilies, *J. Biol. Chem.* 289 (2014) 5261–5273.
- [7] G. Beldman, H.A. Schols, S.M. Pitson, M.J.F. Searle-van Leeuwen, A.G.J. Voragen, Arabinans and arabinan degrading enzymes, in: R.J. Sturgeon (Ed.), *Adv. Macromol. Carbohydr. Res.*, vol. 1, 1997, pp. 1–64 146.
- [8] K.M.J. Van Laere, C.H.L. Voragen, T. Kroef, L.A.M. den Broek, G. Beldman, A.G.J. Voragen, Purification and mode of action of two different arabinoxylan arabinofuranohydrolases from *Bifidobacterium adolescentis* DSM 20083, *Appl. Microbiol. Biotechnol.* 51 (1999) 606–613.
- [9] F.J. Contesini, M.V. Liberato, M.V. Rubio, F. Calzado, M.P. Zubieta, D.M. Riaño-Pachón, F.M. Squina, F. Bracht, M.S. Skaf, A.R. Damasio, Structural and functional characterization of a highly secreted alpha-L-arabinofuranosidase (GH62) from *Aspergillus nidulans* grown on sugarcane bagasse, *Biochim. Biophys. Acta - Proteins Proteomics* 1865 (2017) 1758–1769.
- [10] M.T. Numan, N.B. Bhosle, Alpha-L-arabinofuranosidases: the potential applications in biotechnology, *J. Ind. Microbiol. Biotechnol.* 33 (2006) 247–260.
- [11] P. Kim, Current studies on biological tagatose production using L-arabinose isomerase: a review and future perspective, *Appl. Microbiol. Biotechnol.* 65 (2004) 243–249.
- [12] P. Sánchez-Torres, L. González-candelas, D. Ramón, Expression in a wine yeast strain of the *Aspergillus niger* abfB gene, *FEMS Microbiol. Lett.* 145 (1996) 189–194.
- [13] A.G.J. Voragen, F.M. Rombouts, M.F. Searle-van Leeuwen, H.A. Schols, W. Pilnik, The degradation of arabinans by endo-arabinanase and arabinofuranosidases purified from *Aspergillus niger*, *Food Hydrocoll.* 1 (1987) 423–437.
- [14] B.C. Saha, Alpha-L-arabinofuranosidases: biochemistry, molecular biology and application in biotechnology, *Biotechnol. Adv.* 18 (2000) 403–423.
- [15] H.R. Sørensen, S. Pedersen, A. Viksø-Nielsen, A.S. Meyer, Efficiencies of designed enzyme combinations in releasing arabinose and xylose from wheat arabinoxylan in an industrial ethanol fermentation residue, *Enzym. Microb. Technol.* 36 (2005) 773–784.
- [16] R.M. Berka, I.V. Grigoriev, R. Otillar, A. Salamov, J. Grimwood, I. Reid, N. Ishmael, T. John, C. Darmond, M.-C. Moisan, B. Henrissat, P.M. Coutinho, V. Lombard, D.O. Natvig, E. Lindquist, J. Schmutz, S. Lucas, P. Harris, J. Powlowski, A. Bellemare, et al., Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*, *Nat. Biotechnol.*, 29 (2011) 922–927.
- [17] Y. García-Huante, M. Cayetano-Cruz, A. Santiago-Hernández, C. Cano-Ramírez, R. Marsch-Moreno, J.E. Campos, G. Aguilar-Osorio, C.G. Benitez-Cardoza, S. Trejo-Estrada, M.E. Hidalgo-Lara, The thermophilic biomass-degrading fungus *Thielavia terrestris* Co3Bag1 produces a hyperthermophilic and thermostable beta-1,4-xylanase with exo- and endo-activity, *Extremophiles* 21 (2017) 175–186.
- [18] J.A. Langston, K. Brown, F. Xu, K. Borch, A. Garner, M.D. Sweeney, Cloning, expression, and characterization of a cellobiose dehydrogenase from *Thielavia terrestris* induced under cellulose growth conditions, *Biochim. Biophys. Acta - Proteins Proteomics* 1824 (2012) 802–812.
- [19] A. Margaritis, R.F.J. Merchant, M. Yaguchi, Thermostable Cellulases from thermophilic microorganisms, *Crit. Rev. Biotechnol.* 4 (1986) 327–367.
- [20] F. Segato, A.R.L. Damásio, T.A. Gonçalves, R.C. de Lucas, F.M. Squina, S.R. Decker, R.A. Prade, High-yield secretion of multiple client proteins in *Aspergillus*, *Enzym. Microb. Technol.* 51 (2012) 100–106.
- [21] A.J. Clutterbuck, Sexual and parasexual genetics of *Aspergillus* species, *Biotechnology* 23 (1992) 3–18.
- [22] B.P. Roy, F. Archibald, Effects of Kraft pulp and lignin on *Trametes versicolor* carbon metabolism, *Appl. Environ. Microbiol.* 59 (1993) 1855–1863.
- [23] André R. de Lima Damásio, Tony M. Silva, Fausto B. dos Reis Almeida, Fábio M. Squina, Daniela A. Ribeiro, Adriana F. Paes Leme, Fernando Segato, Rolf A. Prade, João A. Jorge, Hector F. Terenzi, Maria de Lourdes T.M. Polizeli, Heterologous expression of an *Aspergillus nidulans* xylanase GH11 in *Aspergillus nidulans* and its characterization and application, *Process Biochem.* 46 (2011) 1236–1242.
- [24] C. Aslanidis, P.J. de Jong, Ligation-independent cloning of PCR products (LIC-PCR), *Nucleic Acids Res.* 18 (1990) 6069–6074.
- [25] W.H. Eschenfeldt, S. Lucy, C.S. Millard, A. Joachimiak, I.D. Mark, A family of LIC vectors for high-throughput cloning and purification of proteins, in: S.A. Doyle (Ed.), *High Throughput Protein Expr. Purif. Methods Protoc.*, Humana Press, Totowa, NJ, 2009, pp. 105–115.
- [26] T.C. McIlvaine, A buffer solution for colorimetric comparison, *J. Biol. Chem.* 49 (1921) 183–186.
- [27] G.L. Miller, Use of Dinitrosalicylic acid reagent for determination of reducing sugar, *Anal. Chem.* 31 (1959) 426–428.
- [28] U.B. Ericsson, B.M. Hallberg, G.T. DeTitta, N. Dekker, P. Nordlund, Thermofluor-based high-throughput stability optimization of proteins for structural studies, *Anal. Biochem.* 357 (2006) 289–298.
- [29] T.G.G. Battye, L. Kontogiannis, O. Johnson, H.R. Powell, A.G.W. Leslie, iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 67 (2011) 271–281.
- [30] G. Winter, xia2: an expert system for macromolecular crystallography data reduction, *J. Appl. Crystallogr.*, 43 (2010) 186–190.
- [31] W. Kabsch, Integration, scaling, space-group assignment and post-refinement, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66 (2010) 133–144.
- [32] P. Evans, Scaling and assessment of data quality, *Acta Crystallogr. D Biol. Crystallogr.* 62 (2006) 72–82.
- [33] P.R. Evans, G.N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. Sect. D* 69 (2013) 1204–1214.
- [34] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, *Acta Crystallogr. Sect. D* 67 (2011) 235–242.
- [35] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, *J. Appl. Crystallogr.* 40 (2007) 658–674.
- [36] G.N. Murshudov, P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallogr. Sect. D* 67 (2011) 355–367.
- [37] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr. Sect. D* 60 (2004) 2126–2132.
- [38] V.B. Chen, W.B. Arendall III, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. Sect. D* 66 (2010) 12–21.
- [39] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, *J. Mol. Biol.* 372 (2007) 774–797.
- [40] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227 (1970) 680–685.
- [41] I. Migneault, C. Dartiguenave, M.J. Bertrand, K.C. Waldron, Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking, *Biotechniques* 37 (2004) 790–802.
- [42] T.W. Geders, G. Gustafson, B.C. Finzel, Use of differential scanning fluorimetry to optimize the purification and crystallization of PLP-dependent enzymes, *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* 68 (2012) 596–600.
- [43] F.H. Nielsen, H. Berglund, M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, *Nat. Protoc.* 2 (2007) 2212–2221.
- [44] M. Vedadi, F.H. Nielsen, A. Allali-Hassani, O.Y. Fedorov, P.J. Finerty, G.A. Wasney, R. Yeung, C. Arrowsmith, L.J. Ball, H. Berglund, R. Hui, B.D. Marsden, P. Nordlund, M. Sundstrom, J. Weigelt, A.M. Edwards, Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination, *Proc. Natl. Acad. Sci.* 103 (2006) 15835 LP – 15840.
- [45] L.E. Kellelt, D.M. Poole, L.M. Ferreira, A.J. Durrant, G.P. Hazlewood, H.J. Gilbert, Xylanase B and an arabinofuranosidase from *Pseudomonas fluorescens subsp. cellulosa* contain identical cellulose-binding domains and are encoded by adjacent genes, *Biochem. J.* 272 (1990) 369 LP – 376.
- [46] T. Maehara, Z. Fujimoto, H. Ichinose, M. Michikawa, K. Harazono, S. Kaneko, Crystal structure and characterization of the glycoside hydrolase family 62 alpha-L-Arabinofuranosidase from *Streptomyces coelicolor*, *J. Biol. Chem.* 289 (2014) 7962–7972.

- [47] A.P. Kaur, B.P. Nocek, X. Xu, M.J. Lowden, J.F. Leyva, P.J. Stogios, H. Cui, R. Di Leo, J. Powlowski, A. Tsang, A. Savchenko, Functional and structural diversity in GH62 alpha-L-arabinofuranosidases from the thermophilic fungus *Scytalidium thermophilum*, *Microb. Biotechnol.* 8 (2015) 419–433.
- [48] M. Mechelke, D.E. Koeck, J. Broeker, B. Roessler, F. Krabichler, W.H. Schwarz, V.V. Zverlov, W. Liebl, Characterization of the arabinoxylan-degrading machinery of the thermophilic bacterium *Herbinix hemicellulosilytica*-Six new xylanases, three arabinofuranosidases and one xylosidase, *J. Biotechnol.* 257 (2017) 122–130, <https://doi.org/10.1016/j.jbiotec.2017.04.023>.
- [49] R. Pérez, J. Eyzaguirre, *Aspergillus fumigatus* produces two Arabinofuranosidases from Glycosyl hydrolase family 62: comparative properties of the recombinant enzymes, *Appl. Biochem. Biotechnol.* 179 (2016) 143–154.
- [50] M. Couturier, M. Haon, P.M. Coutinho, B. Henrissat, L. Lesage-Meessen, J.-G. Berrin, *Podospora anserina* hemicellulases potentiate the *Trichoderma reesei* secretome for saccharification of lignocellulosic biomass, *Appl. Environ. Microbiol.* 77 (2011) 237 LP – 246.
- [51] H. Tsujibo, C. Takada, Y. Wakamatsu, M. Kosaka, A. Tsuji, K. Miyamoto, Y. Inamori, Cloning and expression of an alpha-L-Arabinofuranosidase gene (stxIV) from *Streptomyces thermoviolaceus* OPC-520, and characterization of the enzyme, *Biosci. Biotechnol. Biochem.* 66 (2002) 434–438.
- [52] B.V. McCleary, V.A. McKie, A. Draga, E. Rooney, D. Mangan, J. Larkin, Hydrolysis of wheat flour arabinoxylan, acid-debranched wheat flour arabinoxylan and arabinoxylo-oligosaccharides by beta-xylanase, alpha-L-arabinofuranosidase and beta-xylosidase, *Carbohydr. Res.*, 407 (2015) 79–96.
- [53] C. Wilkens, S. Andersen, B.O. Petersen, A. Li, M. Busse-Wicher, J. Birch, D. Cockburn, H. Nakai, H.E.M. Christensen, B.B. Kragelund, P. Dupree, B. McCleary, O. Hindsgaul, M.A. Hachem, B. Svensson, An efficient arabinoxylan-debranching alpha-L-arabinofuranosidase of family GH62 from *Aspergillus nidulans* contains a secondary carbohydrate binding site, *Appl. Microbiol. Biotechnol.* 100 (2016) 6265–6277.
- [54] K. Hashimoto, M. Yoshida, K. Hasumi, Isolation and characterization of CcAbf62A, a GH62 alpha-L-Arabinofuranosidase, from the basidiomycete *Coprinopsis cinerea*, *Biosci. Biotechnol. Biochem.* 75 (2011) 342–345.
- [55] A.P. de Souza, D.C.C. Leite, S. Pattathil, M.G. Hahn, M.S. Buckeridge, Composition and structure of sugarcane Cell Wall polysaccharides: implications for second-generation bioethanol production, *BioEnergy Res.* 6 (2013) 564–579.
- [56] B.A. Katz, A. Kossiakoff, The crystallographically determined structures of atypical strained disulfides engineered into subtilisin, *J. Biol. Chem.* 261 (1986) 15480–15485.
- [57] B. Schmidt, L. Ho, P.J. Hogg, Allosteric disulfide bonds, *Biochemistry* 45 (2006) 7429–7433.
- [58] M. Paoli, An elusive propeller-like fold, *Nat. Struct. Biol.* 8 (2001) 744.
- [59] M.J. Bennett, S. Choe, D. Eisenberg, Domain swapping: entangling alliances between proteins, *Proc. Natl. Acad. Sci. U. S. A.* 91 (1994) 3127–3131.
- [60] M.J. Bennett, M.P. Schlunegger, D. Eisenberg, 3D domain swapping: a mechanism for oligomer assembly, *Protein Sci.* 4 (1995) 2455–2468.
- [61] A.M. Gronenborn, Protein acrobatics in pairs – dimerization via domain swapping, *Curr. Opin. Struct. Biol.* 19 (2009) 39–49.
- [62] Y. Liu, D. Eisenberg, 3D domain swapping: as domains continue to swap, *Protein Sci.* 11 (2002) 1285–1299.
- [63] F. Rousseau, J.W.H. Schymkowitz, L.S. Itzhaki, The unfolding story of three-dimensional domain swapping, *Structure* 11 (2003) 243–251.
- [64] L. Holm, P. Rosenström, Dali server: conservation mapping in 3D, *Nucleic Acids Res.* 38 (2010) W545–W549.
- [65] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Res.* 42 (2014) W320–W324.
- [66] T. Tonozuka, Y. Tanaka, S. Okuyama, T. Miyazaki, A. Nishikawa, M. Yoshida, Structure of the catalytic domain of alpha-L-Arabinofuranosidase from *Coprinopsis cinerea*, CcAbf62A, provides insights into structure-function relationships in glycoside hydrolase family 62, *Appl. Biochem. Biotechnol.* 181 (2017) 511–525.
- [67] O.V. Moroz, L.F. Sobala, E. Blagova, T. Coyle, W. Peng, K.B.R. Morkeberg Krogh, K.A. Stubbs, K.S. Wilson, G.J. Davies, Structure of a *Talaromyces pinophilus* GH62 arabinofuranosidase in complex with AraDNJ at 125 Å resolution, *Acta Crystallogr. Sect. F* 74 (2018) 490–495.
- [68] T. Pons, D.G. Naumoff, C. Martinez-Fleites, L. Hernandez, Three acidic residues are at the active site of a beta-propeller architecture in glycoside hydrolase families 32, 43, 62, and 68, *Proteins* 54 (2004) 424–432.
- [69] J.W. Schymkowitz, F. Rousseau, H.R. Wilkinson, A. Friedler, L.S. Itzhaki, Observation of signal transduction in three-dimensional domain swapping, *Nat. Struct. Biol.* 8 (2001) 888–892.
- [70] S. Yang, H. Levine, J.N. Onuchic, Protein oligomerization through domain swapping: role of inter-molecular interactions and protein concentration, *J. Mol. Biol.* 352 (2005) 202–211.
- [71] S. Bershtein, W. Mu, E.I. Shakhnovich, Soluble oligomerization provides a beneficial fitness effect on destabilizing mutations, *Proc. Natl. Acad. Sci.* 109 (2012) 4857 LP – 4862.
- [72] E. Foltá-Stogniew, Oligomeric States of Proteins Determined by Size-Exclusion Chromatography Coupled with Light Scattering, Absorbance, and Refractive Index Detectors, in: *New Emerg. Proteomic Tech.*, Humana Press, Totowa, NJ, 2006, pp. 97–112.

A.3 Benzenesulfonyl incorporated chalcones: Synthesis, structural and optical properties



Sintetizamos três análogos de sulfonamida-chalcona alterando o substituinte ligado ao anel sulfonamida (I = etoxi; II = Cl e III = Br). Esses compostos foram caracterizados por métodos espectroscópicos (RMN, IR e HRMS) e métodos térmicos (HSM e DSC/TGA). Suas estruturas cristalinas foram determinadas por difração de raios-X (SCXRD) e suas estruturas moleculares foram comparadas. Os resultados interessantes observados nas medidas de óptica linear e não-linear mostraram que diferentes grupos ligados ao mesmo *backbone* principal não modificaram significativamente suas propriedades ópticas, embora afetem fortemente suas estruturas cristalinas. Os espectros σ 2PA revelam que o primeiro estado excitado é permitido por 1PA e 2PA, o que indica assimetria na distribuição de carga ao longo da estrutura da molécula π -conjugada.



Benzenesulfonyl incorporated chalcones: Synthesis, structural and optical properties

Jean M.F. Custodio^{a,*}, Fernando Gotardo^b, Wesley F. Vaz^c, Giulio D.C. D'Oliveira^a, Leonardo R. de Almeida^b, Ruben D. Fonseca^{b,d}, Leandro H.Z. Cocca^b, Caridad N. Perez^a, Allen G. Oliver^e, Leonardo de Boni^b, Hamilton B. Napolitano^{c,**}

^a Instituto de Química, Universidade Federal de Goiás, Goiânia, Goiás, Brazil

^b Instituto de Física de São Carlos, Universidade de São Paulo, São Paulo, São Paulo, Brazil

^c Ciências Exatas e Tecnológicas, Universidade Estadual de Goiás, Anápolis, Goiás, Brazil

^d Campus Universitario, Universidad Popular del Cesar, Barrio Sabana, Valledupar, Colombia

^e Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN, 46556, USA

ARTICLE INFO

Article history:

Received 12 December 2019

Received in revised form

30 January 2020

Accepted 3 February 2020

Available online 14 February 2020

Keywords:

Sulfonamide chalcone

Nonlinear optical properties

Hybrid compounds

ABSTRACT

Although the optical properties and applicability of chalcones as potential nonlinear optical (NLO) materials is well known, hybrids having both chalcone and sulfonamide portions are relatively scarce. In this sense, we are seeking to show if combining both into a double functionalized compound will still have similar or better NLO responses. For this, we have synthesized three sulfonamide-chalcone analogues by changing the substituent bonded to the sulfonamide ring (I = ethoxy; II = Cl and III = Br). These compounds were characterized by spectroscopic methods (NMR, IR and HRMS) and thermal methods (HSM and DSC/TGA). Their crystal structures were determined by Single Crystal X-ray Diffraction (SCXRD) and their molecular structures were compared. Crystallographic results showed that one compound crystallizes in a triclinic system whereas the others crystallize in a monoclinic crystal system. Moreover, their crystal packing is dominated by C–H...O interactions. In addition to this study and to first characterize these compounds, linear and nonlinear optical (NLO) properties were performed on all three compounds dissolved in dimethyl sulfoxide. One-photon and two-photon absorption (2PA) spectra and incoherent second harmonic generation were obtained by employing different spectroscopic techniques. The interesting results observed in the linear and NLO measurements showed that different groups bounded to the same main backbone did not significantly modify their optical properties, although they strongly affect their crystal structures. σ 2PA spectra reveal that the first excited state is allowed by both 1PA and 2PA, which indicates asymmetry in the charge distribution along the π -conjugated molecule structure. NLO properties of compounds I-III agree with other chalcones previously studied, which indicates that adding the benzenesulfonyl group does not influence these properties. Considering the similar experimental values for different substituents, the results further motivate investigation on substituent-based optical properties for other sulfonamide-chalcone hybrids.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Chalcones and their derivatives have been extensively studied because of their wide range of applicability. The presence of the reactive keto-ethylenic group and two aryl rings linked by an enone

group is responsible for most of the biological properties observed for these compounds [1,2]. Many biological potentials, such as anticancer [3,4], antituberculosis [5], antileishmanial [6] and therapeutic [7] have been attributed to chalcones, and their use as potential optical devices has also been studied. Chalcones have delocalized electronic charge distribution and π -orbitals; this leads to a high electron density mobility making them ideal candidates for nonlinear optical (NLO) materials [8].

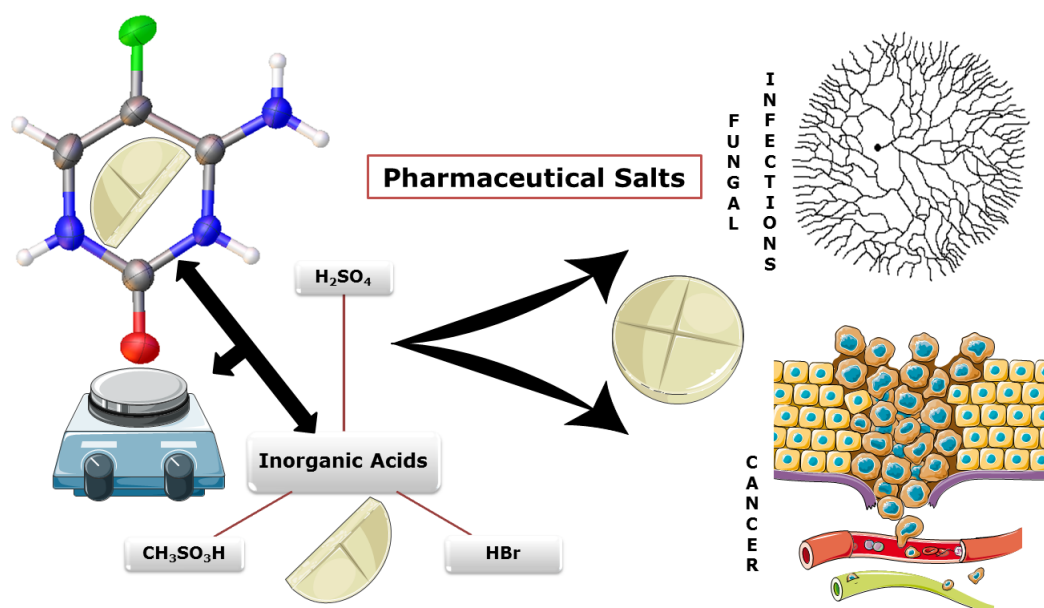
Not only the chalcone backbone, but also the substitution pattern around their aromatic rings is responsible for observed NLO

* Corresponding author.

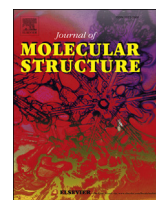
** Corresponding author.

E-mail addresses: jeanmfcustodio@gmail.com (J.M.F. Custodio), hbnapolitano@gmail.com (H.B. Napolitano).

A.4 Protonation of inorganic 5-Fluorocytosine salts



Foram investigados quatro sais da droga antifúngica/antineoplásica 5-Fluorocitosina (5-FC) com os ácidos inorgânicos farmacologicamente aceitos H_2SO_4 , HBr e $\text{CH}_3\text{SO}_3\text{H}$. As reações dão origem a duas formas com ácido sulfúrico, uma com ácido metanossulfônico e outra com ácido bromídrico. A difração de raios-X (SCXRD) foi usada para elucidar a estrutura cristalina.



Protonation of inorganic 5-Fluorocytosine salts

Matheus S. Souza, Cecília C.P. da Silva, Leonardo R. Almeida, Luan F. Diniz, Marcelo B. Andrade, Javier Ellena*

Instituto de Física de São Carlos, Universidade de São Paulo, CP 369, 13.560-970, São Carlos, SP, Brazil

ARTICLE INFO

Article history:

Received 1 December 2017

Received in revised form

15 February 2018

Accepted 16 February 2018

Available online 19 February 2018

Keywords:

5-Fluorocytosine

Active pharmaceutical ingredient

Inorganic acids

Pharmaceutical salts

X-ray diffraction

Crystal structure

ABSTRACT

5-Fluorocytosine (5-FC) has been widely used for the treatment of fungal infections and recently was found to exert an extraordinary antineoplastic activity in gene directed prodrug therapy. However, despite of its intense use, 5-FC exhibits tabletability issues due its physical instability in humid environments, leading to transition from the anhydrous to monohydrate phase. By considering that salt formation is an interesting strategy to overcome this problem, in this paper crystal engineering approach was applied to the supramolecular synthesis of new 5-FC salts with sulfuric, hydrobromic and methanesulfonic inorganic acids. A total of four structures were obtained, namely 5-FC sulfate monohydrate (1:1:1), 5-FC hydrogen sulfate (1:1), 5-FC mesylate (1:1) and 5-FC hydrobromide (1:1), the last one being a polymorphic form of a structure already reported in the literature. These novel salts were structurally characterized by single crystal X-ray diffraction and its supramolecular organization were analysed by Hirshfeld surface analysis. The vibrational behavior was evaluated by Raman spectroscopy and it was found to be consistent with the crystal structures.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Anhydrous form of the cytosine derivative [1] 5-Fluorocytosine (4-amino-5-fluoro-1,2-dihydropyrimidin-2-one, 5-FC, *Scheme 1*) is the most frequently used active pharmaceutical ingredient (API) in the treatment of fungal infections [2]. Recently, this drug has been used for cancer treatment by gene-directed enzyme prodrug therapy (GDEPT), since is converted into the antineoplastic drug 5-Fluorouracil (5-FU) in the presence of the cytosine deaminase enzyme (CD) [3–10]. Despite its intense use, 5-FC is sensitive to relative humidity (RH) variations. More specifically, in environments with high moisture content, 5-FC anhydrous converts into the monohydrate form, leading to serious issues during the manufacturing process and storage. As such, sensitivity is a crucial challenge when trying to develop stable 5-FC solid dosage forms [2].

One of the strategies used to settle manufacturing issues in an API formulation is the development of multicomponent crystal forms, e.g. salts [11–14]. In the pharmaceutical industry, this is a straightforward way to generate safe and effective solid medicines [15]. The rational design and the supramolecular synthesis of salts

may modulate physicochemical properties of the currently commercialized APIs which exhibit pharmacokinetic limitations, such as variable bioavailability [16]. Thus, the advantages of salt formation in the context of drug development and clinical performance may be linked, for example, to the solution of issues related to thermal stability, hygroscopicity and solubility [12,17–27].

To investigate the solid-state stability and mechanical properties of 5-FC in tablets, recently, Perumalla et al. reported two conjugated acid-base cocrystals involving the neutral API and the ion pair previously formed between protonated 5-FC and deprotonated cofomers (in this case, HCl and acesulfame, both listed as pharmaceutically accepted) [2]. Besides that, in the last few years an extensive amount of new 5-FC solid forms has been reported [2,11–13,28–30]. Despite that, it was observed a lack 5-FC structures with inorganic cofomers [11,31–36]. In a survey in the Cambridge Structural Database (CSD) [33], only 6 salts with inorganic acids were identified, representing just the 11.7% of all 5-FC deposited crystalline structures [11,31,32]. Within this framework, here we report the supramolecular synthesis and the vibrational behavior of four novel salts obtained from the reaction of 5-FC with three pharmaceutically accepted inorganic acids: sulfuric, methanesulfonic and hydrobromic.

* Corresponding author.

E-mail address: javiere@ifsc.usp.br (J. Ellena).