

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS JABOTICABAL**

**CARACTERIZAÇÃO ESTRUTURAL DE ENZIMAS
LIGNOCELULOLÍTICAS COM POTENCIAL DE APLICAÇÃO
EM BIORREFINARIA**

Pâmela Aparecida Maldaner Pereira

Bióloga

2020

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS JABOTICABAL**

**CARACTERIZAÇÃO ESTRUTURAL DE ENZIMAS
LIGNOCELULOLÍTICAS COM POTENCIAL DE APLICAÇÃO
EM BIORREFINARIA**

Discente: Pâmela Aparecida Maldaner Pereira

Orientadora: Profa. Dra. Eliana Gertrudes de Macedo Lemos

Coorientadora: Dra. Elisângela Soares Gomes Pepe

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutor em Microbiologia Agropecuária.

2020

FICHA CATALOGRÁFICA

P436c	Pereira, Pâmela Aparecida Maldaner Caracterização estrutural de enzimas lignocelulolíticas com potencial de aplicação em biorrefinaria / Pâmela Aparecida Maldaner Pereira. -- Jaboticabal, 2020 157 f. : tabs., fotos Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal Orientadora: Eliana Gertrudes de Macedo Lemos Coorientadora: Elisângela Soares Gomes-Pepe 1. Biotecnologia. 2. Enzimas. 3. Cristalografia de raio X. I. Título.
-------	---

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

**TÍTULO DA TESE: CARACTERIZAÇÃO ESTRUTURAL DE ENZIMAS LIGNOCELULOLÍTICAS COM
POTENCIAL DE APLICAÇÃO EM BIORREFINARIA**

AUTORA: PÂMELA APARECIDA MALDANER PEREIRA

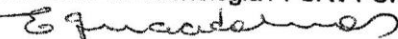
ORIENTADORA: ELIANA GERTRUDES DE MACEDO LEMOS

COORDINADORA: ELISÂNGELA SOARES GOMES PEPE

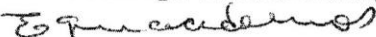
Aprovada como parte das exigências para obtenção do Título de Doutora em MICROBIOLOGIA
AGROPECUÁRIA, pela Comissão Examinadora:



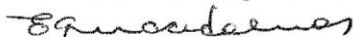
Profa. Dra. ELIANA GERTRUDES DE MACEDO LEMOS
Departamento de Tecnologia / FCAV / UNESP - Jaboticabal



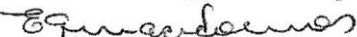
Dr. PEDRO MANUEL HENRIQUES MARQUES MATIAS (VIDEOCONFERÊNCIA)
Instituto de Tecnologia Química e Biológica-ITQB/UNL / Oeiras/Portugal



Dra. MARIANA RANGEL PEREIRA (VIDEOCONFERÊNCIA)
Universidade de Cambridge / Cambridge/Inglaterra



Dra. THAÍS CARVALHO MAESTER CASANOVA (VIDEOCONFERÊNCIA)
Ecobiotech Biotecnologia e Automação / Ribeirão Preto/SP



Profa. Dra. LUCIA MARIA CARARETO ALVES (VIDEOCONFERÊNCIA)
Departamento de Tecnologia / FCAV / UNESP - Jaboticabal



Jaboticabal, 04 de setembro de 2020

DADOS CURRICULARES DO AUTOR

Pâmela Aparecida Maldaner Pereira, nascida em 14 de outubro de 1990 na cidade de Rosana, São Paulo. Filha de João Bosco Pereira e Neiva Inês Maldaner Pereira (*in memoriam*) cresceu na cidade de São João, Paraná. Em 2008 iniciou sua graduação em Ciências Biológicas na Universidade Paranaense – UNIPAR, Francisco Beltrão, Paraná, obtendo o título de Bióloga no ano de 2011. Durante a graduação participou de projetos de iniciação científica nas áreas de citogenética de peixes e reaproveitamento de resíduos de cana-de-açúcar para cultivo axênico de cogumelos comestíveis. Em março de 2013 ingressou no curso de mestrado em Conservação e Manejo de Recursos Naturais na Universidade Estadual do Oeste do Paraná – UNIOESTE, câmpus Cascavel, sob orientação do Prof. Dr. Sílvio César Sampaio e coorientação da Dra. Danielle Medina Rosa. Durante o mestrado atuou na área de reuso de águas residuárias da suinocultura como alternativa de fertilizante em culturas anuais e obteve título de mestre no ano de 2015. Em agosto de 2016 ingressou no curso de doutorado em Microbiologia Agropecuária da Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP/FCAV, Jaboticabal, São Paulo, sob orientação da Profa. Dra. Eliana G. de Macedo Lemos e coorientação da Dra. Elisângela Soares Gomes Pepe. Em 2018 foi contemplada com uma bolsa de doutorado sanduíche através do Programa de Doutorado-sanduíche no exterior (PDSE – CAPES) para realizar parte de sua pesquisa no Laboratório de Cristalografia Aplicada à Indústria e a Medicina, do Instituto de Tecnologia Química e Biológica (ITQB) da Universidade Nova de Lisboa (UNL) em Oeiras, Portugal, sob orientação do Dr. Pedro Matias. Durante estágio sanduíche realizou, principalmente, estudos de cristalografia de raios-X das proteínas aqui descritas. Desenvolveu sua Tese de Doutorado no departamento de Tecnologia da UNESP/FCAV e, até o presente momento, dedicou-se às áreas de Bioquímica de Microrganismos, Biologia Molecular e Reaproveitamento de Resíduos Agroindustriais.

“Quando acordei hoje de manhã, eu sabia quem eu era, mas acho que já mudei
muitas vezes desde então.”

Alice no País das maravilhas (Lewis Carroll)

Dedico esta tese com todo meu amor e carinho aos meus pais, João e Neiva (*in memoriam*), pelo incentivo e amor incondicional; ao meu filho Vitor Otávio, por dar sentido à minha existência e ao meu esposo Vandrei, pelo apoio e paciência.

AGRADECIMENTOS

À Deus, pela dádiva da vida, por guiar meus passos, iluminar minhas escolhas e me conceder forças para nunca desistir.

À minha família, por todo amor e carinho, apoio e incentivo constante em minha vida, especialmente aos meus pais, João e Neiva (*in memoriam*), que sempre batalharam pela educação e felicidade dos filhos; e ao meu irmão, João Gabriel, pelo companheirismo e amizade. Vocês são meu alicerce e exemplo de força e dedicação! Amo vocês!

Ao meu filho Vitor Otávio, razão do meu viver, pelas alegrias e felicidades diárias, pelo seu sorriso e seu jeito todo carinhoso de ser. Ao meu esposo Vandrei, por me acompanhar em mais esta jornada, por todo amor, compreensão e paciência.

À Profa. Dra. Eliana G. de Macedo Lemos, pela orientação, oportunidade, confiança, paciência e conhecimentos compartilhados.

À minha coorientadora e amiga Dra. Elisângela Soares Gomes Pepe, pelo imenso auxílio e conhecimentos compartilhados, pela paciência ao me ensinar, pelo apoio e incentivo nos momentos difíceis.

Ao Dr. Pedro Matias, da Unidade de Cristalografia Macromolecular do ITQB-UNL, por me receber gentilmente em seu laboratório, por compartilhar seus conhecimentos sobre cristalografia e contribuir para o desenvolvimento deste trabalho; e aos colegas, Sara e Bruno, pelos auxílio sempre que necessário.

Às amigas que ficaram na minha terra amada, o Paraná, mas que me acompanharam no coração e continuaram fazendo parte da minha vida me incentivando e torcendo pelo meu sucesso!

Às amigas que fiz no laboratório e que levarei para a vida: Bárbara, Natália, Michelli, Milena, Bruna, Gabriela, Joana, Luís (Lutcho) e Farinacio. Grata pelo apoio, incentivo e divertimento!

Ao Dr. João Carlos Campanharo, pelo auxílio e conhecimento compartilhado durante os experimentos.

Aos pesquisadores e professores do Laboratório de Bioquímica de Microrganismos e Plantas (LBMP).

Aos funcionários e aos amigos do Departamento de Tecnologia UNESP, Câmpus de Jaboticabal.

A todos aqueles que de alguma maneira também contribuíram para a realização desta tese e não foram aqui mencionados.

À Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências Agrárias e Veterinárias (UNESP/FCAV) e ao Programa de Pós-Graduação em Microbiologia Agropecuária, câmpus Jaboticabal.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudos concedida. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES - Código de Financiamento 001 e PDSE-CAPES processo 88881.189918/2018-01).

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo recurso concedido ao Laboratório de Bioquímica de Microrganismos e Plantas, coordenado pela Professora Dra. Eliana Lemos, onde a tese foi desenvolvida.

Muito Obrigada!

SUMÁRIO

RESUMO.....	iv
ABSTRACT	vi
LISTA DE ABREVIATURAS.....	viii
CAPÍTULO 1 – Considerações gerais.....	1
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	4
2.1. Composição da biomassa lignocelulósica.....	4
2.2. Enzimas envolvidas na sacarificação da biomassa lignocelulósica	7
2.2.1. Celulases.....	8
2.2.2. Hemicelulases	9
2.2.3. Enzimas modificadoras de lignina	12
2.3. A biomassa lignocelulósica e a produção de biocombustíveis.....	15
2.4. Bioprodutos derivados da biomassa	16
2.5. Enzimas microbianas	17
2.6. Determinação e predição de estruturas	18
2.6.1. Predição computacional de estruturas	18
2.6.2. Cristalografia de raios-X.....	20
3. REFERÊNCIAS	23
CAPÍTULO 2 – Estrutura cristalográfica e estudos de ancoragem molecular da β -glicosidase Bg10 estimulada pela glicose	32
1. INTRODUÇÃO	33
2. MATERIAL E MÉTODOS	34
2.1. Origem da amostra e características enzimáticas.....	34
2.2. Expressão e purificação da proteína recombinante	34
2.3. Cristalização	35
2.4. Coleta e processamento de dados	37
2.5. Ancoragem molecular da Bg10.....	38
3. RESULTADOS E DISCUSSÃO.....	39
3.1. Expressão e Purificação da proteína recombinante.....	39
3.2. Cristalização e otimização dos cristais.....	40
3.3. Estrutura geral da Bg10	40
3.4. Sítio ativo da Bg10.....	44
3.5. Bases moleculares da tolerância/estimulação pelo produto	46
3.6. Simulações de ancoragem molecular	53
4. CONCLUSÕES	54
5. REFERÊNCIAS	55
CAPÍTULO 3 – Caracterização estrutural <i>in silico</i> e estudos de ancoragem molecular da β -xilosidase bacteriana BXil	59
1. INTRODUÇÃO	60
2. MATERIAL E MÉTODOS	61
2.1. Origem da ORF CB_612456.....	61
2.2. Análise da estrutura primária	62
2.3. Análise da estrutura secundária.....	62
2.4. Modelagem por homologia e caracterização dos domínios	62
2.5. Estudos de ancoragem molecular.....	63

3. RESULTADOS E DISCUSSÃO.....	64
3.1. Origem e descrição da ORF CB_612456.....	64
3.2. Análise da estrutura primária.....	65
3.3. Análise da estrutura secundária.....	67
3.4. Modelagem por homologia e caracterização dos domínios.....	68
3.5. Ancoragem molecular e caracterização do sítio ativo da BXil.....	74
4. CONCLUSÕES.....	82
5. REFERÊNCIAS.....	82
CAPÍTULO 4 – Expressão, extração e ensaios preliminares de cristalização da β -xilosidase recombinante (BXil) expressa em corpos de inclusão.....	87
1. INTRODUÇÃO.....	88
2. MATERIAL E MÉTODOS.....	90
2.1. Origem da ORF CB_612456 codificadora de BXil.....	90
2.2. Clonagem molecular da ORF CB_612456.....	90
2.3. Expressão da proteína recombinante.....	92
2.4. Extração nativa.....	93
2.5. Extração desnaturante.....	94
2.5.1. Lise celular.....	94
2.5.2. Lavagem celular e solubilização dos corpos de inclusão.....	95
2.6. Renaturação <i>in vitro</i>	95
2.7. Purificação proteica.....	96
2.8. Determinação da concentração da proteína.....	96
2.9. Teste de detecção de atividade enzimática.....	97
2.10. Análise de Espalhamento dinâmico de luz (DLS - Dynamic Light Scattering).....	97
2.11. “Western blotting”.....	97
2.12. Ensaios de cristalização.....	98
3. RESULTADOS E DISCUSSÃO.....	98
3.1. Clonagem molecular da ORF CB_612456.....	99
3.2. Expressão e extração nativa.....	100
3.3. Extração desnaturante – corpos de inclusão.....	110
3.4. Renaturação <i>in vitro</i>	114
3.5. Purificação proteica.....	115
3.6. Análise de Espalhamento dinâmico de luz e detecção da atividade enzimática.....	117
3.7. Ensaios de cristalização da BXil através do método de difusão de vapor em gota assente.....	118
4. CONCLUSÕES E PERSPECTIVAS.....	118
5. REFERÊNCIAS.....	119
CAPÍTULO 5 – Ensaios preliminares de cristalização da lacase metagenômica (LacMeta) prospectada em solo sob serapilheira de eucalipto.....	122
1. INTRODUÇÃO.....	123
2. MATERIAL E MÉTODOS.....	126
2.1. Origem da proteína.....	126
2.2. Expressão e extração.....	126
2.3. Purificação proteica.....	126
2.4. Experimentos de cristalização.....	127
2.5. Coleta de dados de difração.....	129

3. RESULTADOS E DISCUSSÃO.....	129
3.1. Expressão e purificação da proteína recombinante	129
3.2. Cristalização e otimização dos cristais.....	130
3.3. Coleta e processamento de dados	131
4. CONCLUSÕES E PERSPECTIVAS.....	132
5. REFERÊNCIAS	133
APÊNDICES.....	136
Apêndice A. Sequências de aminoácidos provenientes do banco de dados NCBI.....	137
Apêndice B. Métodos de expressão e extração da proteína recombinante.	138

CARACTERIZAÇÃO ESTRUTURAL DE ENZIMAS LIGNOCELULOLÍTICAS COM POTENCIAL DE APLICAÇÃO EM BIORREFINARIA

RESUMO – O cenário atual do uso de biocatalisadores em processos industriais movimenta a busca por novas enzimas, principalmente as que são capazes de degradar material lignocelulósico potencialmente utilizado na produção de biocombustíveis e outros produtos de alto valor agregado. O etanol de segunda geração é considerado um substituto sustentável aos combustíveis fósseis, devido ao processo ser desenvolvido a partir de resíduos agroindustriais. No entanto, a produção deste biocombustível detém diversos interferentes próprios do processo de produção, que contribuem para impulsionar a busca e o aperfeiçoamento de enzimas a fim de torná-las mais estáveis e eficientes. As enzimas β -glicosidase, β -xilosidase e lacase apresentam um amplo potencial biotecnológico, principalmente no que diz respeito à sacarificação da biomassa lignocelulósica. A resolução da estrutura tridimensional destas enzimas é de grande relevância, pois permite compreender alguns dos inúmeros mecanismos biológicos destas macromoléculas, como a especificidade da enzima pelo substrato, relação estrutura-função, arquitetura proteica e centros catalíticos. Deste modo, o presente estudo teve como objetivo expressar, purificar e caracterizar estruturalmente as enzimas β -glicosidase, β -xilosidase e lacase. A enzima β -glicosidase tolerante e estimulada pela glicose (Bg10) foi expressa e purificada em alta concentração e submetida a ensaios de cristalografia e difração de raios-X que forneceram dados de difração a uma resolução de 2,3 Å e permitiram determinar sua estrutura tridimensional. A estrutura proteica apresentou características estruturais conservadas entre as β -glicosidases GH1 tolerantes e estimuladas pela glicose. Além disso, seus dados estruturais forneceram percepções importantes sobre alguns dos mecanismos biológicos envolvidos na tolerância e estímulo pela glicose desta molécula, tais como a configuração do canal do sítio ativo e a presença de resíduos “gatekeepers”, os quais regulam a passagem e a orientação da glicose. A enzima β -xilosidase (BXil), por sua vez, teve sua caracterização estrutural realizada através de abordagens *in silico* e estudos de ancoragem molecular, os quais forneceram informações importantes sobre esta molécula, tais como a arquitetura de domínios composta pela repetição de dois domínios típicos de GH43 conservados e até hoje não caracterizados; a disposição do seu sítio catalítico e os principais resíduos envolvidos em interações com diferentes substratos. A expressão da enzima BXil ocorreu em corpos de inclusão e, assim, sua extração foi realizada através de métodos desnaturantes com otimização de protocolo e submetida a ensaios de cristalografia, no entanto, não ocorreu a formação de cristais. Sendo assim, novas abordagens para obtenção da proteína solúvel estão sendo realizadas, a fim de sustentar a caracterização *in silico* e desvendar a função e a interação entre seus domínios a partir de caracterização cinética e estrutural. Por último, a enzima lacase (LacMeta), também foi submetida a ensaios de cristalografia e difração de raios-X com dados coletados a uma resolução de 1,8 Å. Apesar disso, não foi possível determinar a estrutura da LacMeta através do método de substituição molecular devido à falta de um modelo de fase adequado no Banco de Dados de Proteínas. Diante disso, será necessária a realização de novos ensaios para obtenção de cristais derivados com selenometionina para obtenção das fases e determinação da estrutura. Este trabalho de caracterização permitiu elucidar algumas características estruturais

destas macromoléculas, conhecimento este que poderá ser útil para potencializar suas propriedades catalíticas de acordo com propósitos específicos para aplicação biotecnológica em biorrefinaria.

Palavras-chave: biotecnologia, cristalografia de raios-X, enzimas

STRUCTURAL CHARACTERIZATION OF LIGNOCELLULOLYTIC ENZYMES WITH POTENTIAL APPLICATION IN BIOREFINERY

ABSTRACT – The current scenario of the use of biocatalysts in industrial processes moves the research in the search for new enzymes, especially those that are capable of degrading lignocellulosic material potentially used in the production of biofuels and other products with high added value. Second generation ethanol is considered a sustainable substitute for fossil fuels, due to the process being developed from agro-industrial waste. However, the production of this biofuel has several interferents hat part of the production process, which contribute to boost the search and the improvement of enzymes in order to make them more stable and efficient. The enzymes β -glycosidase, β -xylosidase and laccase have a wide biotechnological potential, mainly with regard to the saccharification of lignocellulosic biomass. The resolution of the three-dimensional structure of these enzymes is of great relevance, as it will allow us to understand some of the innumerable biological mechanisms of these macromolecules, such as the substrate specificity, structure-function relationship, protein architecture and catalytic centers. Thus, the present study aimed to express, purify and structurally characterize the enzymes β -glycosidase, β -xylosidase and laccase. The glucose-stimulated and tolerant β -glucosidase enzyme (Bg10) was expressed and purified in high concentration and subjected to crystallography and X-ray diffraction tests that provided diffraction data for a resolution of 2.3 Å and allowed to determine its three-dimensional structure. The protein structure showed structural characteristics conserved among tolerant and glucose-stimulated GH1 β -glycosidases. In addition, its structural data provided important insights into some of the biological mechanisms involved in glucose tolerance and stimulation of this molecule, such as the configuration of the active site channel and the presence of gatekeeper's residues, which regulate the passage and orientation glucose. The β -xylosidase enzyme (BXil), in turn, had its structural characterization carried out through *in silico* approaches and molecular docking studies, which provided important information about this molecule, such as the domain architecture composed by the repetition of two typical domains of conserved and not yet characterized GH43; the disposition of its catalytic site and the main residues involved in interactions with different substrates. The expression of the BXil enzyme occurred in inclusion bodies and, therefore, its extraction was performed through denaturing methods with protocol optimization and subjected to crystallography tests, however, crystal formation did not occur. The factors that may explain the absence of crystals may be related to the characteristics of the sample obtained from denaturing extractions, being homogeneity, concentration and even its intrinsic characteristics. Therefore, new approaches to obtain the soluble protein are being carried out, in order to support the characterization *in silico* and unveil the function and the interaction between its domains based on kinetic and structural characterization. Finally, the laccase enzyme (LacMeta) was also subjected to crystallography and X-ray diffraction tests with data collected to a resolution of 1.8 Å. Despite this, it was not possible to determine the structure of LacMeta using the molecular substitution method, due to the lack of a suitable phasing model in the Protein Data Bank. Therefore, it will be necessary to carry out new tests to obtain crystals derivatized with selenomethionine to obtain the phases and determine the structure. This characterization work allowed to elucidate some structural characteristics of these macromolecules, a knowledge that could be

useful to enhance their catalytic properties according to specific purposes for biotechnological application in biorefinery.

Keywords: biotechnology, X-ray crystallography, enzymes

LISTA DE ABREVIATURAS

Å: Angström
ABTS: 2,2-Azino-bis-(3-etilbenzotiazolina-6-sulfonato)
Bis-tris: 2,2-Bis(hidroximetil)-2,2',2''-nitrilotrietanol
CBM: "Carbohydrate-Binding Module"
CO₂: gás carbônico
CTAB: brometo de cetiltrimetilamônio de sódio
DNA: ácido desoxirribonucleico
DO: densidade óptica
DTT: ditionitrito
EC: "Enzyme Commission"
GSH: glutathiona reduzida
GSSG: glutathiona oxidada
IMAC: cromatografia de afinidade de íons metálicos imobilizados
IPTG: isopropil-β-D-tiogalactopiranosídeo
kDa: kilodaltons
LB: Luria Bertani
M: molar
mg: miligrama
mM: milimolar
NCBI: "National Center for Biotechnology Information"
ORF: Open Reading Frame
p/v: peso por volume
pb: pares de bases
PCR: reação em cadeia da polimerase
PDB: Protein Data Bank
PEG: polietilenoglicol
pH: potencial hidrogeniônico - log [H⁺]
pI: ponto isoelétrico
pK_a: constante de ionização
pNP: *p*-nitrofenol
PVDF: fluoreto de polivinilideno
QMEAN: Qualitative Model Energy Analysis
RMN: Ressonância Magnética Nuclear
RMSD: Root Mean Square deviation
RNA: ácido ribonucleico
rpm: rotações por minuto
SDS: dodecil sulfato de sódio
SDS-Page: eletroforese em gel de poliacrilamida na presença de dodecil sulfato
SEC: cromatografia por exclusão de tamanho
SLAC: "small laccase"
Tris: tris[hidroximetil]aminometano
μg: micrograma
μL: microlitro
v/v: volume por volume

Abreviatura dos aminoácidos

A	Ala	Alanina
C	Cys	Cisteína
D	Asp	Ácido aspártico
E	Glu	Ácido glutâmico
F	Phe	Fenilalanina
G	Gly	Glicina
H	His	Histidina
I	Ile	Isoleucina
K	Lis	Lisina
L	Leu	Leucina
M	Met	Metionina
N	Asn	Asparagina
P	Pro	Prolina
Q	Gln	Glutamina
R	Arg	Arginina
S	Ser	Serina
T	Thr	Treonina
V	Val	Valina
W	Trp	Triptofano
Y	Tyr	Tirosina

CAPÍTULO 1 – Considerações gerais

1. INTRODUÇÃO

O elevado poder biotecnológico das enzimas tem atraído atenção de pesquisadores na busca por variedades mais eficientes e capazes de reduzir o custo de produção nos mais diferentes setores.

As enzimas são aplicadas nos processos de produção de biocombustíveis (Robak e Balcerek, 2018); no tratamento de águas residuárias (Cammarota e Freire, 2006; Unuofin et al., 2019); indústrias têxteis (Madhavan et al., 2017); indústrias de papel e celulose para o biobranqueamento e melhoria da qualidade da polpa (Maijala et al., 2008; Choi et al., 2015); em indústrias alimentícias para produção de prebióticos, adoçantes de baixa caloria e açúcares raros (Akoh et al., 2008); indústrias de bebidas para melhorar o rendimento e sua qualidade (Uzuner e Cekmecelioglu, 2018); indústrias farmacêuticas (Choi et al., 2015); na química fina para produção de compostos ativos enantiopuros (Nestl et al., 2011); indústrias de ração animal para aumentar a digestibilidade e absorção de nutrientes (Ojha et al., 2018), dentre outros.

O uso de enzimas na biocatálise industrial é vantajoso e sustentável uma vez que operam em condições de reação mais leves, operam com eficiência e especificidade, possuem rotas de produção mais simples, reduzem a geração de resíduos tóxicos e poluentes e reduzem o tempo dos processos (Akoh et al., 2008; Wohlgemuth, 2010; Choi et al., 2015).

Apesar das inúmeras vantagens, a digestão enzimática se configura em um desafio financeiro para sua utilização em escala industrial, considerando que o custo das enzimas ainda é considerado oneroso, especialmente devido à necessidade de importação e às constantes necessidades de aprimoramentos das formulações comerciais sob finalidade de torná-las mais efetivas e estáveis perante condições adversas do próprio processo em que participam (Dos Santos et al., 2016; Johnson, 2016; Branco et al., 2019). Estes são os principais fatores que têm impulsionado a busca por novas enzimas e o avanço da engenharia enzimática, a fim de obter enzimas mais robustas, estáveis, efetivas e de baixo custo de produção.

A biologia estrutural é a área responsável por decifrar e compreender a estrutura molecular e a dinâmica de macromoléculas biológicas e como sua função pode ser afetada por alterações em suas estruturas. As estruturas de proteínas podem ser obtidas a partir de abordagens computacionais, com a predição de estruturas, ou experimentalmente, através da cristalografia de raios-X, espectroscopia de ressonância magnética nuclear (RMN) ou microscopia crioelétrica (cryo-EM).

As informações estruturais servem como base para inúmeros campos de pesquisa, destacando-se aqueles relacionados à biotecnologia, que auxiliam na descoberta e no desenho de novas drogas e auxiliam no entendimento de interações e relações estrutura-função de proteínas a nível molecular (Nitta et al., 2018). A engenharia de proteínas também utiliza estas informações a fim de aprimorar propriedades específicas de enzimas de interesse biotecnológico, como por exemplo termoestabilidade e maior atividade catalítica.

O entendimento dos mecanismos moleculares de enzimas que atuam na desconstrução da biomassa lignocelulósica é de grande importância para as indústrias de biocombustíveis e bioprocessos, pois contribuem para a inovação tecnológica e favorecem a produção de catalisadores mais eficientes e de baixo custo.

Neste sentido, o presente estudo teve como objetivo expressar, purificar e caracterizar estruturalmente as enzimas β -glicosidase, β -xilosidase e lacase. Em um primeiro estudo, uma β -glicosidase tolerante e estimulada pela glicose (Bg10), proveniente de DNA metagenômico de solo e anteriormente caracterizada através de estudos cinéticos, foi submetida à ensaios de cristalografia e difração de raios-X e estudos de ancoragem molecular (Capítulo 2). Os melhores dados de difração foram coletados a uma resolução de 2,3 Å e permitiram determinar sua estrutura tridimensional. A Bg10 apresentou características estruturais conservadas entre as β -glicosidasas GH1, como por exemplo, a presença de enovelamento clássico do tipo TIM barril (β/α)₈. Além disso, foi possível observar características importantes sobre a conformação do seu canal catalítico e a presença de resíduos envolvidos na tolerância à glicose. Os estudos de ancoragem molecular permitiram observar o modo de ligação da glicose e outros substratos no sítio ativo, bem como identificar os resíduos de aminoácidos envolvidos em possíveis interações proteína-ligante. As propriedades estruturais observadas na Bg10 são características de GH1 tolerantes e estimuladas

pela glicose, o que corrobora com estudos de cinética enzimática realizados anteriormente. Estes resultados forneceram percepções importantes sobre as características estruturais desta molécula envolvidos na tolerância e estímulo pela glicose, as quais podem ser modificados através de engenharia de proteínas a fim de melhorar a atividade catalítica e a tolerância à glicose.

Em outro trabalho, uma β -xilosidase microbiana (BXil), anotada e prospectada no genoma de *Chitinophaga* sp., foi caracterizada estruturalmente através de abordagens *in silico* e estudos de ancoragem molecular (Capítulo 3). A enzima, denominada BXil, apresentou uma arquitetura de domínios não caracterizada até o momento constituído pela repetição de um domínio hélice beta característico de uma Glicosil Hidrolase da família 43. Além disso, foi possível prever a disposição de seu sítio ativo e os principais resíduos envolvidos em interações com diferentes substratos. Estes resultados forneceram informações importantes sobre a estrutura geral da enzima BXil, no entanto, a função de seu domínio N-terminal não pôde ser esclarecida, sendo necessários estudos adicionais de cristalografia e cinética enzimática dos domínios individuais e em conjunto.

Com intuito de compreender a função desta arquitetura de domínios e a interação entre eles, a BXil foi submetida à ensaios de expressão heteróloga a fim de caracterizá-la através de estudos cinéticos e estruturais (Capítulo 4). A enzima BXil foi expressa em corpos de inclusão e inúmeras alternativas para melhorar a solubilidade foram realizadas, no entanto, todas sem sucesso. Diante disso, sua extração foi realizada através de diferentes métodos desnaturantes e de renaturação proteica a fim de estabelecer um protocolo adequado. Em seguida, a enzima foi submetida à purificação e ensaios de cristalografia. Embora o protocolo de extração desnaturante e renaturação proteica tenha resultado na obtenção da proteína na forma ativa, não ocorreu a formação de cristais durante o período de análises e, portanto, não foi possível realizar os ensaios de difração de raios-X. Os fatores que podem explicar a ausência de cristais podem estar relacionados às características da amostra proteica, como homogeneidade, concentração e até mesmo características intrínsecas da enzima, como a presença de um “linker” demasiado grande entre os domínios, impedindo-os de assumirem uma orientação relativa uniforme e, consequentemente, impedir a cristalização. Sendo assim, novas abordagens para

obtenção da proteína solúvel estão sendo realizadas com os domínios separados através da técnica do truncamento incremental aleatório da sequência do gene alvo (Hart e Tarendeau, 2006). Esta técnica gera milhares de construções possíveis utilizadas para rastrear uma construção solúvel. A partir destes resultados espera-se poder desvendar a função e a interação entre seus domínios a partir de caracterização cinética e estrutural.

Finalizando os esforços empreendidos em elucidar aspectos estruturais de enzimas de interesse biotecnológico, uma “small” lacase proveniente de DNA metagenômico de solo (LacMeta), anteriormente caracterizada através de estudos cinéticos, foi submetida à ensaios de cristalografia e difração de raios-X (Capítulo 5). A formação dos cristais ocorreu com sucesso e os melhores dados de difração foram a uma resolução de 1,8 Å. Algumas tentativas de determinação da estrutura foram realizadas, no entanto, o método de substituição molecular não foi eficiente, principalmente devido à falta de um modelo de fase adequado no Banco de Dados de Proteínas. A baixa similaridade de sequência é uma característica comum entre “Multicopper oxidases” que costumam compartilhar conservação elevada restrita a domínios catalíticos, além disso, estes resultados corroboram com os estudos anteriores que indicaram que a LacMeta se trata de uma enzima ainda não descrita. Assim, é necessária a realização de novos ensaios através de outra técnica, onde cristais derivados com selenometiona são utilizados para obtenção das fases e determinação da estrutura (Hendrickson et al., 1990).

De modo geral, este estudo de caracterização elucidou algumas características estruturais das enzimas β -glicosidase, β -xilosidase e lacase. Este conhecimento poderá ser útil para potencializar as propriedades catalíticas destas macromoléculas de acordo com propósitos específicos para aplicação biotecnológica.

2. REVISÃO DE LITERATURA

2.1. Composição da biomassa lignocelulósica

Diversos tipos de resíduos lignocelulósicos estão disponíveis no Brasil, como resíduos florestais e agrícolas, sendo o bagaço e a palha da cana-de-açúcar oriundos

das indústrias de açúcar e etanol os mais abundantes. Isto deve-se ao fato que o Brasil produz quase 747 milhões de toneladas de cana-de-açúcar por ano (Food and Agriculture Organization FAO, 2018). A biomassa lignocelulósica é uma fonte abundante de açúcares renováveis e é composta principalmente por celulose (30-50%), hemicelulose (15-35%) e lignina (10-30%) (Zheng et al., 2017) (Figura 1).

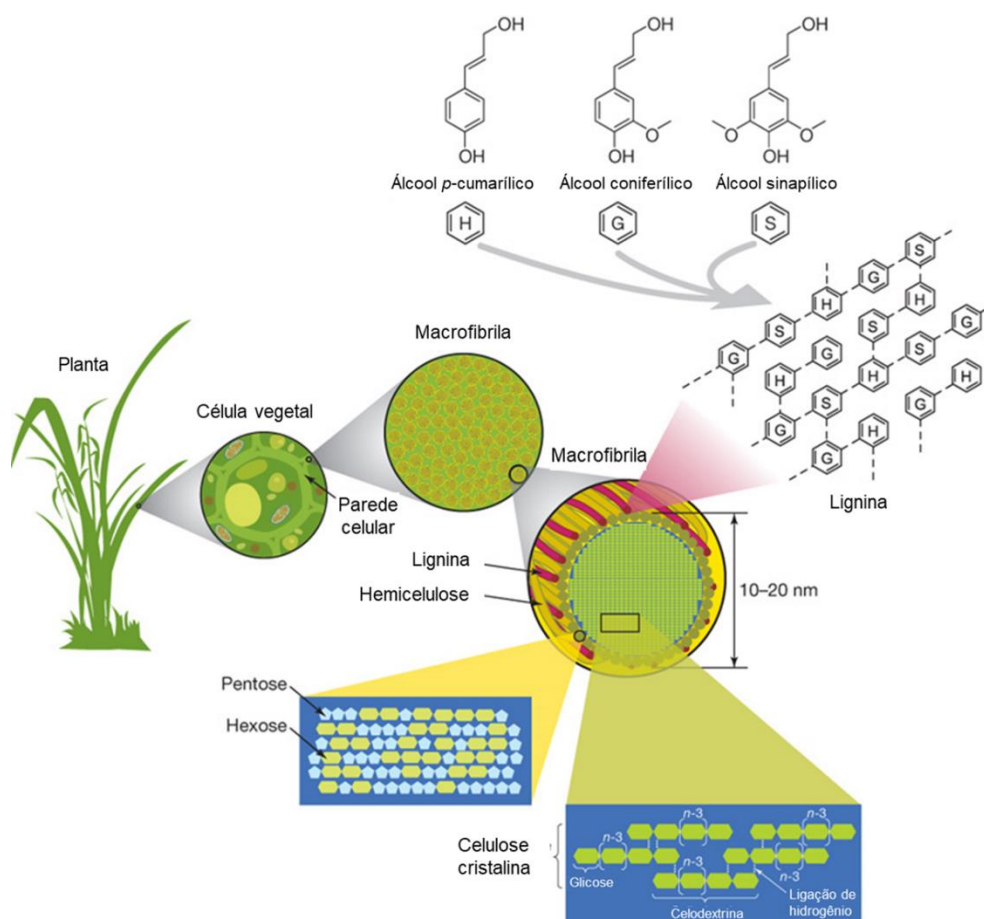


Figura 1. Composição geral da biomassa lignocelulósica representando as frações de celulose, hemicelulose e lignina. Adaptado de (Rubin, 2008).

A celulose é composta por um homopolissacarídeo linear formado por subunidades de D-glicose unidos entre si através de ligações glicosídicas do tipo β -1,4 (Pérez et al., 2002). Apresenta regiões altamente ordenadas e cristalinas que são insolúveis e resistentes à degradação química e enzimática, além de regiões amorfas com estrutura desorganizada (Horn et al., 2012).

A hemicelulose é o segundo maior componente da parede celular vegetal e a xilana é seu principal constituinte. A estrutura de sua cadeia principal é composta por

resíduos de β -xilopiranosse unidos por ligação glicosídica do tipo β (1-4), podendo apresentar substituições nas cadeias laterais entre diferentes espécies de plantas (Wang et al., 2014). Estas substituições ocorrem em diversas posições por diferentes grupos e cada tipo de substituição categoriza a xilana entre homoxilana, arabinoxilana, glucuronoxilana e glucuronoarabinoxilana (Wang et al., 2014; Chakdar et al., 2016). Além disso, as moléculas da xilana formam ligações covalentes com os resíduos fenólicos da lignina (Kamble e Jadhav, 2012). A heterogeneidade da xilana é resultado de diferentes tipos de ligações e ramificações que, do mesmo modo que a celulose cristalina, forma uma barreira física que interfere na sua solubilidade e dificulta a degradação enzimática (Ebringerová e Heinze, 2000).

A lignina, por sua vez, é um polímero aromático muito abundante na natureza e pode apresentar diferentes unidades monoméricas e tipos de ligação em diferentes espécies de plantas (Pollegioni et al., 2015). É formada por unidades de fenilpropano, incluindo álcool *p*-cumarílico, álcool sinapílico e álcool coniferílico, conectadas por ligações C-C e C-O-C (Zheng et al., 2017) (Figura 2). Sua estrutura heterogênea complexa é altamente recalcitrante, no entanto, organismos como os fungos de podridão branca (como *Pleurotus* sp.), actinomicetos (como *Streptomyces* sp. e *Rhodococcus* sp.), e proteobactérias (*Pseudomonas* e *Enterobacter*) produzem enzimas oxidativas eficientes para sua degradação (Wong, 2009; Lee et al., 2019).

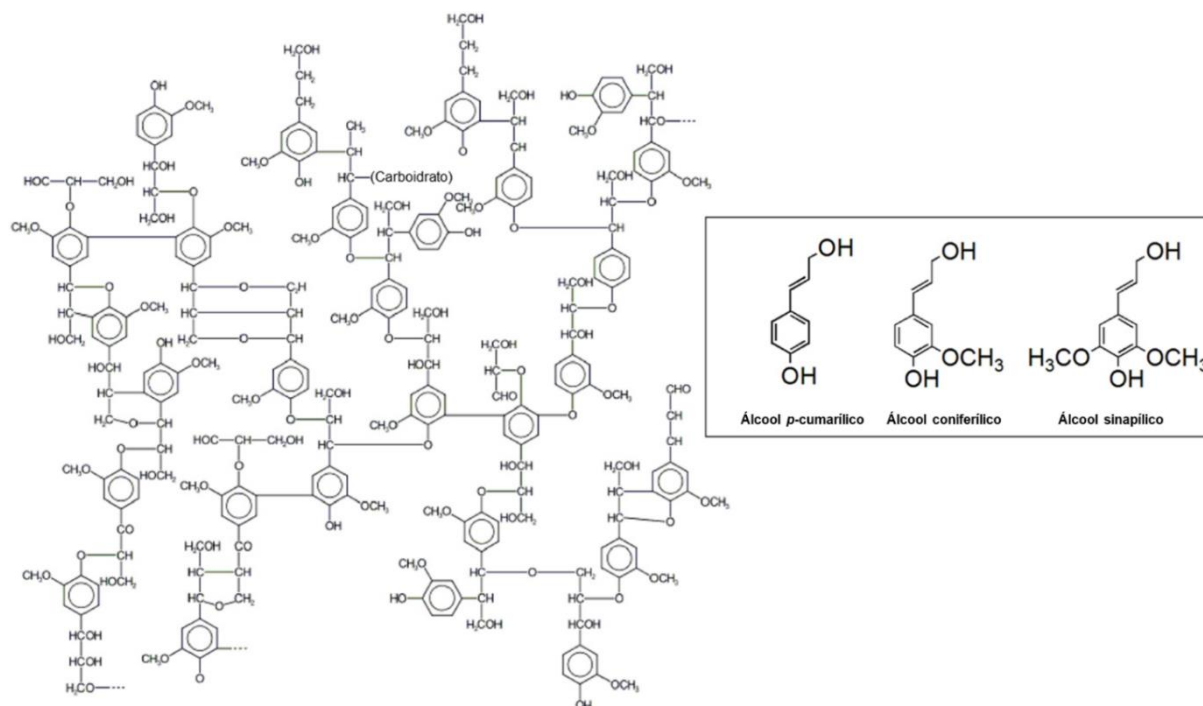


Figura 2. Representação esquemática da estrutura da lignina e das unidades de fenilpropano. Adaptado de (Christopher et al., 2014).

2.2. Enzimas envolvidas na sacarificação da biomassa lignocelulósica

As enzimas são classificadas em 6 grandes classes de acordo com o tipo de reação que catalisam (Tabela 1). Cada uma das classes recebe uma codificação numérica definida pela “Enzyme Commission”, que corresponde ao número E.C.

Tabela 1. Classificação enzimática de acordo com a “Enzyme Commission”.

CLASSES	REAÇÃO
EC 1. Oxidoredutases	Catalisam reações de oxirredução
EC 2. Transferases	Catalisam a transferência de grupos entre moléculas
EC 3. Hidrolases	Catalisam a reação de hidrólise de várias ligações
EC 4. Liases	Catalisam a clivagem de ligações covalentes
EC 5. Isomerases	Catalisam a transferência de grupos em uma mesma molécula.
EC 6. Ligases	Ligam as moléculas por meio de ligações covalentes

A complexidade estrutural da biomassa lignocelulósica requer a ação sinérgica de várias enzimas atuando em diferentes pontos da sua estrutura. As principais enzimas envolvidas na sacarificação da biomassa são as hidrolases (EC 3), que

atuam na quebra de ligações glicosídicas e são produzidas principalmente por fungos e bactérias. Enzimas pertencentes à outras classes também atuam na desconstrução da biomassa, tais como enzimas de atividade auxiliar (2.2.3).

2.2.1. Celulases

As enzimas que atuam em pontos específicos da celulose para sua conversão em glicose são as endo-1,4- β -glicanases (EC 3.2.1.4); exo-1,4- β -glicanases não redutoras (EC 3.2.1.91), exo-1,4- β -glicanases redutoras (EC 3.2.1.176) e β -glicosidases (EC 3.2.1.21) (Figura 3). Estas enzimas estão distribuídas em diferentes famílias de Glicosil Hidrolases (GHs) de acordo com o banco de dados CAZy ("Carbohydrate-Active enZymes Database", <http://www.cazy.org/>), que classifica as enzimas em diferentes famílias com base na sequência, especificidade e estrutura 3D de enzimas ativas em carboidratos.

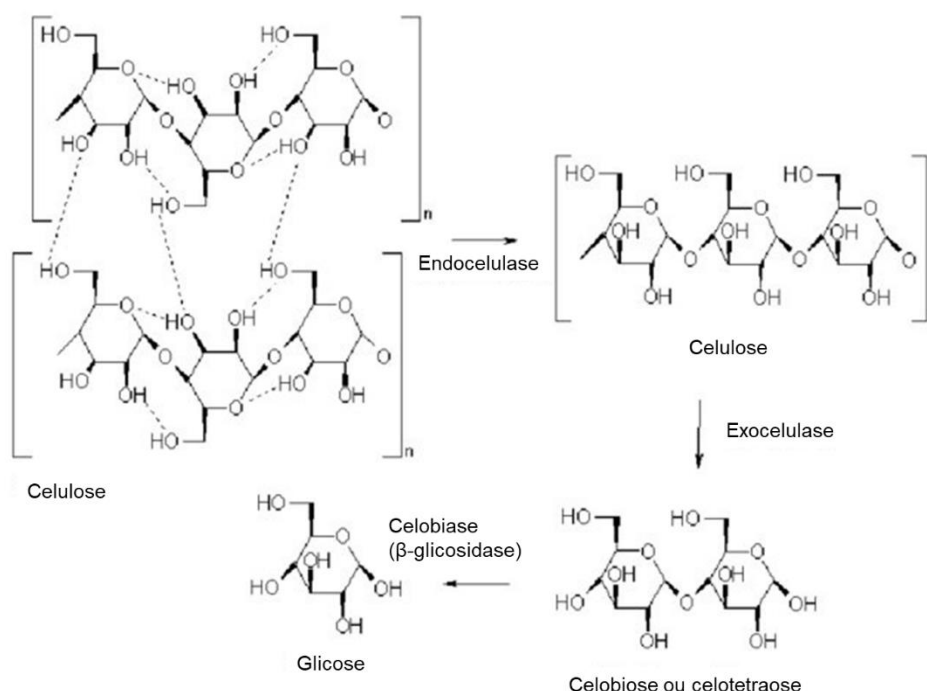


Figura 3. Representação esquemática dos pontos de atuação das celulases. Fonte: Bacovsky et al. (2010).

As endo-1,4- β -glicanases (EC 3.2.1.4), também chamadas de celulases, possuem função primordial durante a hidrólise, sendo responsável pela quebra

randômica das ligações internas do polímero de celulose, expondo as extremidades redutoras e não redutoras para degradação subsequente. Até o momento, estas enzimas estão distribuídas em 15 famílias de GHs (5, 6, 7, 8, 9, 10, 12, 26, 44, 45, 48, 51, 74, 124 e 148). Por sua vez, as exo-1,4- β -glicanases, ou celobiohidrolases (CBH), atuam nas extremidades não redutoras (EC 3.2.1.91) e redutoras (EC 3.2.1.176) da cadeia liberando unidades de celobiose (Sharma et al., 2018) e estão distribuídas entre as famílias de GHs 5, 6 e 9 e nas famílias 7, 9 e 48, respectivamente.

As β -glicosidases, ou celobiases (EC 3.2.1.21), atuam na hidrólise de resíduos não redutores de celobiose e celobio-oligossacarídeos liberando unidades de glicose e pertencem a diversas famílias de GHs (1, 2, 3, 5, 9, 16, 30, 39 e 116). As β -glicosidases pertencentes à família 1 são alvos de inúmeras pesquisas acadêmicas, pois possuem características de tolerância à glicose e ao etanol. Esta característica resulta na sua exploração para a sacarificação de fibra de celulose nas indústrias de etanol de segunda geração, no processamento de bebidas alcóolicas, como a maturação de vinhos para acelerar o envelhecimento através da quebra de compostos glicosídicos precursores de sabor (Gomes-Pepe et al., 2016). Deste modo, β -glicosidases tolerantes ao etanol e estimuladas pela glicose são de grande interesse biotecnológico, pois podem ser mais eficientes durante o processo de degradação do substrato e reduzir os custos de produção.

2.2.2. Hemicelulases

Devido à heterogeneidade de sua estrutura, a hidrólise completa da xilana requer um grupo de enzimas que atuam sobre sua estrutura principal, além de enzimas acessórias que atuam sinergicamente nos pontos de ramificação (Figura 4).

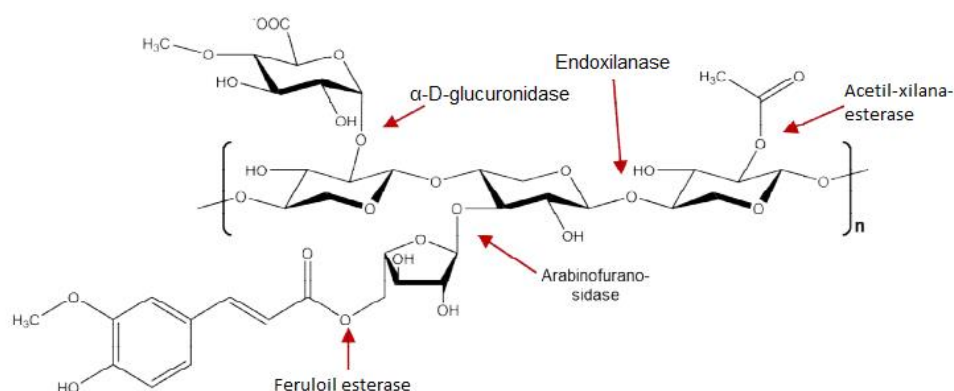


Figura 4. Representação esquemática de parte da estrutura química da xilana e pontos de atuação das enzimas (adaptado de KALIM et al., 2015).

As endo-1,4- β -xilanases (ou comumente descritas xilanases) e as β -D-xilosidases atuam em sinergia na hidrólise da cadeia principal de xilana liberando unidades simples de xilose e xilo-oligossacarídeos de diferentes tamanhos.

As xilanases (E.C. 3.2.1.8) estão envolvidas na hidrólise randômica de ligações glicosídicas β -1,4 entre resíduos de xilose na cadeia principal da xilana (Figura 3). Estas enzimas são agrupadas nas famílias 3, 5, 6, 8, 9, 10, 11, 16, 26, 30, 43, 44, 51, 62, 98 e 141 de GHs (<http://www.cazy.org/>). As xilanases mais estudadas pertencem às famílias GH10 e GH11 com mais de 300 sequências gênicas conhecidas (Shallom e Shoham, 2003). Os membros da família GH10 possuem massa molecular de >30 kDa, ponto isoelétrico (pI) ácido, enquanto que os membros da família GH11 possuem massa molecular de <30 kDa e pI básico (Polizeli et al., 2005; Kocabas et al., 2015).

As β -xilosidases (E.C. 3.2.1.37), ou somente xilosidases, também atuam sobre a cadeia principal da xilana e estão envolvidas na hidrólise de ligações β -1,4 de xilo-oligossacarídeos e xilobiose liberando unidades simples de xilose (Knob et al., 2010; Juturu e Wu, 2012) (Figura 4). Estas enzimas pertencem à diversas famílias de GHs (1, 2, 3, 5, 10, 30, 39, 43, 51, 52, 54, 116 e 120, <http://www.cazy.org/>), das quais os membros pertencentes à família 43 catalisam a hidrólise da ligação glicosídica invertendo a configuração do carbono anomérico do substrato (Till et al., 2014).

Algumas xilosidases da família GH43 apresentam atividade bifuncional para arabinofuranosidase e β -xilosidase, principalmente em substratos de aril-glicosídeo

(Falck et al., 2016). Isso pode ser explicado pelas semelhanças espaciais das ligações glicosídicas e grupos hidroxilas entre os resíduos de D-xilopirranose e L-arabinofuranose (Shallom e Shoham, 2003; Huy et al., 2013).

A ação sinérgica das xilanases e β -xilosidases aumenta a eficiência da hidrólise da xilana e, conseqüentemente, eleva a produção de açúcares fermentescíveis. Isto é possível graças à ação catalítica das β -xilosidases que hidrolisa os xilo-oligossacarídeos que inibem as endoxilanases e limitam a hidrólise da xilana (Polizeli et al., 2005; Knob et al., 2010; Canilha et al., 2012; Juturu e Wu, 2012). Em vista disso, as xilosidases são consideradas essenciais para o complexo xilanolítico e têm-se tornado alvo de pesquisas acadêmicas e industriais, principalmente para o desenvolvimento de bioprocessos e produtos, como na indústria de celulose e papel, na indústria de alimentos e bebidas e ração animal.

As demais enzimas envolvidas na hidrólise completa da xilana atuam sobre os vários substitutos das cadeias laterais em sua estrutura. O grupo das enzimas acessórias inclui α -L-arabinofuranosidases (E.C. 3.2.1.55), α -D-glucuronidases (E.C.3.2.1.139), acetil-xilana-esterases (E.C. 3.1.1.72) e feruloil-esterases (EC. 3.1.1.73) (Collins; Gerday; Feller, 2005; Moreira; Filho, 2016; Wang, Guozeng et al., 2014).

As enzimas acessórias α -L-arabinofuranosidases pertencem à classe das GHs e atuam na hidrólise de resíduos terminais não redutores de α -L-arabinofuranose em arabinoxilanas (Saha, 2000). As α -glucuronidases, por sua vez, são responsáveis pela remoção dos grupos laterais de ácido D-glucurônico através da hidrólise de ligações glicosídicas do tipo α -1,2 (Septiningruma et al., 2015; Yan et al., 2017). Acetil-xilana-esterases catalisam a hidrólise de ligações ésteres liberando ácido acético de acetil xilano (Adesioye et al., 2016).

O ácido ferúlico é o ácido hidroxicinâmico mais abundante. Consiste em uma ramificação lateral ligada covalentemente à xilana por meio de ligações éster (Wong, 2006; Fazary e Ju, 2007). As enzimas feruloil esterases (FAEs) são responsáveis pela hidrólise da ligação éster entre polissacarídeos (principalmente xilana e pectina) e ácido ferúlico liberando ácidos fenólicos (como ácidos ferúlicos ou outros ácidos cinâmicos) (Fazary e Ju, 2007; Topakas et al., 2007; Li et al., 2011). As FAEs são enzimas acessórias essenciais na remoção das cadeias laterais de ácido ferúlico e

suas reticulações, aumentando a acessibilidade às fibras de hemicelulose por outras enzimas e, conseqüentemente, favorecendo a degradação da biomassa lignocelulósica (Wong, 2006).

2.2.3. Enzimas modificadoras de lignina

As enzimas modificadoras da lignina (LMEs) são oxidoredutases e não atuam diretamente sobre os carboidratos, no entanto, atuam na lignina que está associada a eles, auxiliando as demais enzimas. Portanto, as LMEs foram agrupadas em uma classe denominada Atividade Auxiliares (AAs) no banco de dados CAZy. Atualmente, a classe de AAs agrupa 9 famílias de enzimas lignolíticas e 6 famílias de monooxigenases de polissacarídeos líticos (LPMOs) (<http://www.cazy.org/>).

As enzimas LMEs mais importantes são divididas em lignina peroxidases (EC 1.11.1.14), manganês peroxidases (EC 1.11.1.13) e lacases (EC 1.10.3.2), além de enzimas acessórias (Falade et al., 2017).

As lignina peroxidases e as manganês peroxidases, relatadas em fungos e bactérias, fazem parte da classe heme-peroxidases do tipo II, que realizam a oxidação de substratos orgânicos e inorgânicos na presença de peróxido de hidrogênio (como a própria nomenclatura sugere) e contêm um grupo heme (Falade et al., 2017). As lignina peroxidases atuam em compostos não fenólicos da lignina para liberar radicais aril-cátion, enquanto que as manganês peroxidases catalisam a oxidação do íon Mn^{2+} para formar Mn^{3+} e, então oxidar as frações fenólicas da lignina para produzir radicais livres (Vrsanska et al., 2016; Falade et al., 2017).

Ao contrário das peroxidases, as lacases não necessitam de co-substratos, como o peróxido de hidrogênio (Roth e Spiess, 2015). Estas enzimas são amplamente distribuídas em fungos, plantas e bactérias (Balcázar-López et al., 2016; Baldrian, 2006). Também chamadas oxidoredutases, elas atuam numa ampla variedade de compostos como polifenólicos, fenólicos, não fenólicos e aminas aromáticas (Balcázar-López et al., 2016; Singh e Gupta, 2020).

Juntamente com outras enzimas, as lacases compõem a superfamília de “Multicopper oxidases” (MCOs), caracterizadas por possuírem em cada um dos seus monômeros, geralmente, quatro átomos de cobre em seu centro catalítico

coordenados por um resíduo de cisteína e 10 resíduos de histidina altamente conservados (Balcázar-López et al., 2016; Baldrian, 2006).

O centro cupredoxina que acomoda estes átomos de cobre são divididos em três. O centro T1 que acomoda um átomo de cobre, o centro T2 que acomoda outro cobre e o T3 com 2 átomos de cobre. Os centros T2 e T3 formam o cluster trinuclear de cobre (TNC). Cada um destes centros possui função específica e a presença dos átomos de cobre é essencial para a correta atividade destas enzimas (Solomon et al., 1996; Komori et al., 2009; Giardina et al., 2010; Sirim et al., 2011; Gunne et al., 2014) (Figura 5).

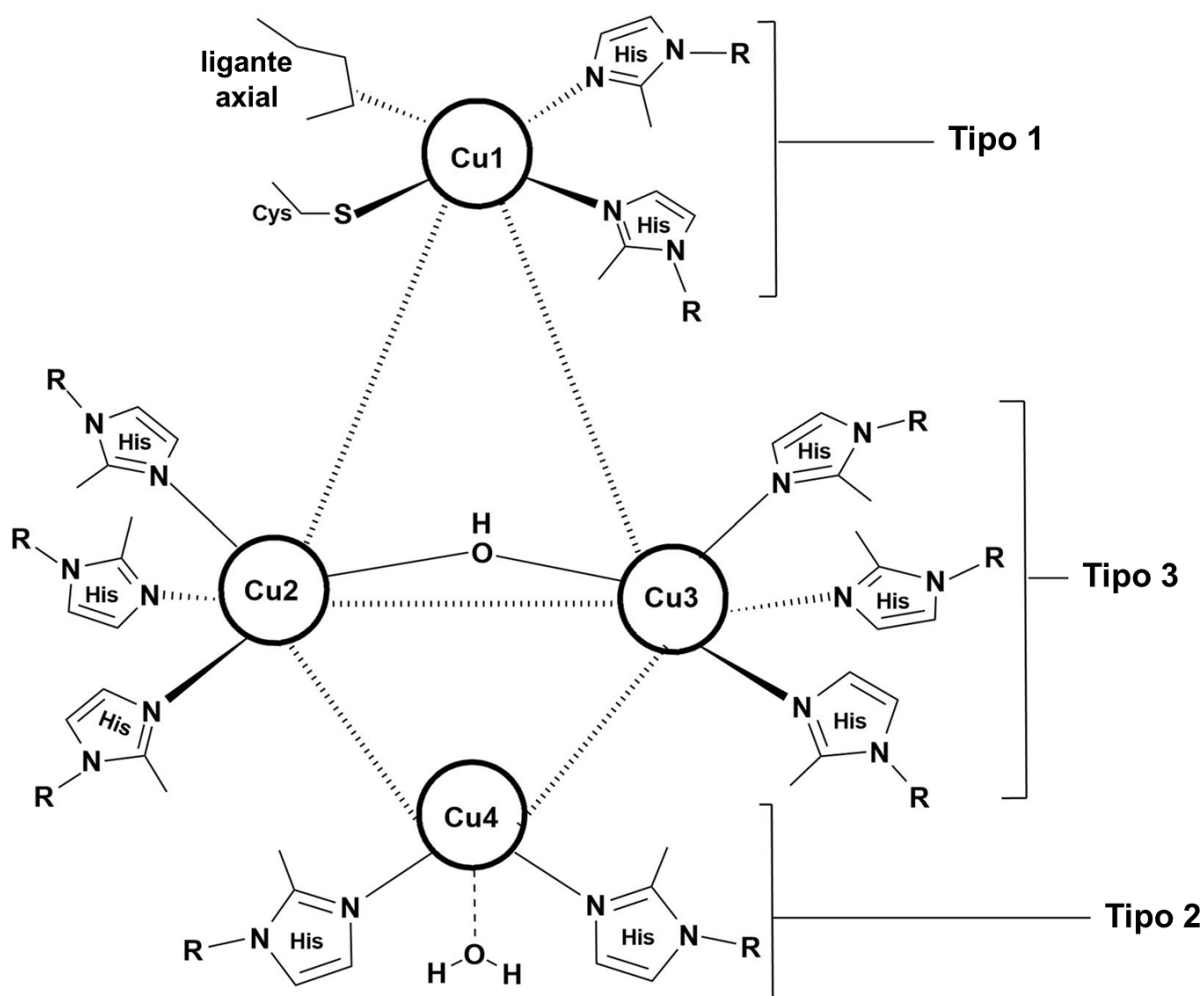


Figura 5. Representação esquemática do arranjo dos centros de ligação ao cobre encontrado nas MCOs. Adaptado de (Agrawal et al., 2018).

O centro T1, ou centro azul, é o responsável pela coloração azul destas enzimas e também por receber 4 elétrons provenientes da redução do substrato e

transferi-los para o TNC (Figura 5) que, por sua vez, é responsável por reduzir uma molécula de oxigênio à água (Bento et al., 2005; Skálová et al., 2009).

As lacases típicas possuem 3 domínios de ligação ao cobre por monômero, os domínios Cu-oxidase 1 e 3 são responsáveis por acomodar os centros de ligação ao cobre enquanto o domínio Cu-oxidase 2 auxilia na formação da fenda de ligação ao substrato e no posicionamento correto dos outros domínios para formar o TNC (Machczynski et al., 2004; Gunne et al., 2014). Lacases de dois domínios vêm sendo identificadas, sendo que nas quais, geralmente, o domínio Cu-oxidase 1 encontra-se ausente. A primeira lacase de dois domínios, também chamada de “small laccase” (SLAC), foi relatada no genoma de *Streptomyces coelicolor* (Tishchenko et al., 2015).

As enzimas SLACs formam estruturas homotriméricas com o TNC dos monômeros adjacentes a fim de reestabelecer sua atividade (Machczynski et al., 2004; Skálová et al., 2009).

A versatilidade de atuação das lacases em diferentes substratos deve-se à baixa especificidade destas enzimas, o que é visto como vantagem para sua utilização em diversos processos industriais (Machczynski et al., 2004).

Além da aplicação das lacases na biorrefinaria para o tratamento da biomassa, despolimerização e modificação da lignina e detoxificação da biomassa hidrolisada (Roth e Spiess, 2015); outros setores também demonstram interesse nessas enzimas. Nas indústrias alimentícias as lacases são utilizadas como aditivos no processamento de alimentos e bebidas, atuando na remoção de polifenóis durante processo de produção do vinho, cerveja e sucos; na panificação melhorando a qualidade da farinha; na melhoria dos fatores sensoriais de alimentos, como odor e sabor; como agentes gelificantes; dentre outros (Minussi et al., 2002; Osma et al., 2010).

Em indústrias de processamento de papel e celulose, as lacases são aplicadas na deslignificação e branqueamento da polpa removendo a lignina residual (Virk et al., 2012).

Além da aplicação das lacases durante a realização de processos industriais, seu uso também ganhou destaque no tratamento de efluentes das fábricas de cerveja e de destilaria, na remoção de corantes de efluentes de indústrias têxteis e na biorremediação de solos e águas contaminados com hidrocarbonetos aromáticos policíclicos (Osma et al., 2010; Singh e Gupta, 2020).

As indústrias têxteis, de cosméticos, nanotecnologia, e síntese orgânica também fazem uso de lacases para os mais diversos fins devido à sua ampla estabilidade e ampla especificidade de substrato (Riva, 2006; Rodríguez Couto e Toca Herrera, 2006; Sirim et al., 2011; Zheng et al., 2016; Singh e Gupta, 2020).

2.3. A biomassa lignocelulósica e a produção de biocombustíveis

O bagaço e a palha de cana-de-açúcar são os resíduos lignocelulósicos mais abundantes no país, considerando que o Brasil é um dos maiores produtores de cana-de-açúcar do mundo, com quase 747 milhões de toneladas produzidas em 2018 (Food and Agriculture Organization FAO, 2018), o que representa aproximadamente 39,2% da produção mundial.

O etanol de segunda geração (2G) é um dos produtos biotecnológicos mais promissores e seu processo de produção baseia-se no reaproveitamento de diferentes tipos de resíduos lignocelulósicos agroindustriais disponíveis localmente e em grande quantidade. O reaproveitamento desse material resulta no aumento da produção de biocombustíveis sem necessidade de expansão das áreas de cultivo, além de reduzir a emissão de CO₂ na atmosfera. Por estas e outras razões, é considerado um substituto sustentável para os combustíveis fósseis.

No entanto, a produção deste biocombustível apresenta alguns desafios a serem superados. A recalcitrância da biomassa e o alto custo de produção quando comparado ao etanol de primeira geração são as principais razões, sobretudo devido a necessidade de importação e a baixa atividade catalítica das enzimas utilizadas na etapa de hidrólise (Dos Santos et al., 2016; Robak e Balcerek, 2018; Branco et al., 2019; Malico et al., 2019).

A fração cristalina da celulose e a barreira de lignina existente ao redor das fibras celulósicas aumentam a resistência estrutural dificultando o acesso enzimático aos açúcares fermentescíveis configurando um fator limitante à produção de biocombustíveis (Canilha et al., 2012; Castro, 2010; Lucy; Petri, 2010). Em vista disso, é necessária uma grande quantidade de enzimas hidrolíticas para degradá-la, o que aumenta consideravelmente os custos de produção. Como alternativa mais econômica, uma etapa de pré-tratamento é realizada para reduzir a cristalinidade da

celulose e remover a lignina de maneira a aumentar a susceptibilidade da celulose e hemicelulose à hidrólise enzimática (Canilha; Chandel; et al., 2012; Castro, 2010; Lucy; Petri, 2010).

Embora a degradação da lignina seja realizada por métodos físico-químicos, o uso de enzimas é considerado ecológico e mais politicamente correto, pois evita o uso de grandes quantidades de água e reduz a produção de resíduos tóxicos liberados pelo pré-tratamento ácido que interferem negativamente no processo de fermentação (Avanthi; Banerjee, 2016; Boruah et al., 2016).

A etapa seguinte ao pré-tratamento e detoxificação envolve reações multienzimáticas onde celulasas, hemicelulasas e outras enzimas acessórias atuam sinergicamente para liberar os açúcares fermentescíveis que serão transformados em etanol por leveduras.

Além da complexidade estrutural da biomassa e o difícil acesso aos açúcares fermentescíveis, algumas enzimas podem apresentar suas próprias limitações, como baixa especificidade de substrato, baixa tolerância à variações de pH e temperatura, inibição por radicais livres e íons, inibição alostérica pelo produto, dentre outros (Amorim et al., 2011; Canilha et al., 2012; Klein-Marcuschamer et al., 2012).

Diante disso, a busca pelo aprimoramento das enzimas já descobertas ou por novas enzimas com características promissoras e de baixo custo de produção direciona grande parte da pesquisa biotecnológica mundial.

2.4. Bioprodutos derivados da biomassa

Além da produção de biocombustíveis, a transformação da biomassa lignocelulósica em produtos de alto valor agregado tem crescido nos últimos anos. As biorrefinarias lignocelulósicas (LCB) surgiram com o propósito de reduzir o impacto ambiental causado pelo uso de matérias-primas não renováveis, como o petróleo, e substituí-las por fontes renováveis e prontamente disponíveis. Estas instalações integram a produção de energia, de produtos químicos e biocombustíveis através do processamento da biomassa (Takkellapati et al., 2018). Dentre os principais bioprodutos, podemos citar os biofármacos, insumos químicos, alimentos e ração animal.

Diferentes tipos de biomassa podem ser utilizados como matéria-prima nestas biorrefinarias, como lascas de madeira, cana-de-açúcar, palha de trigo e milho, casca de arroz e coco verde, dentre outros, cada um possuindo uma composição química diferente do outro (De Bhowmick et al., 2018; Takkellapati et al., 2018). A cana-de-açúcar é o principal tipo de biomassa encontrado no Brasil e sua composição química é de aproximadamente 42% de celulose, 25% de hemicelulose e 20% de lignina (Kim e Day, 2011).

Os principais produtos derivados da biomassa incluem o próprio etanol, butanol, 1,4-diácidos, ácido aspártico, ácido glutâmico, carvão ativado, lactato de etila, ácido láctico, ácido succínico, xilitol, xilo-oligossacarídeos, furfural, glicerol, sorbitol, dentre outros (Werpy e Petersen, 2004; Irmak, 2017).

A lignina é um composto que pode ser aproveitado para obtenção de vários bioprodutos, no entanto, sua utilização é limitada devido à alta complexidade de sua estrutura (Lee et al., 2019). Além disso, algumas estratégias vêm sendo relatadas a fim de agregar valor a este subproduto, como por exemplo, métodos para recuperar a lignina após o pré-tratamento (Lee et al., 2019).

2.5. Enzimas microbianas

A intensa demanda por biocatalisadores, associada à exploração da biodiversidade global, tem proporcionado a constante descoberta de genes funcionais, principalmente por meio da abordagem metagenômica, que permite o acesso à diversidade de organismos e genes provenientes de diversos ambientes. A partir da metagenômica é possível estudar o potencial metabólico de diferentes microrganismos não cultiváveis e identificar novos genes codificadores de enzimas de interesse biotecnológico.

As comunidades microbianas frequentemente são capazes de degradar o material lignocelulósico, uma vez que sinergicamente sintetizam inúmeras enzimas durante seu desenvolvimento nos mais diversos tipos de substratos. As enzimas microbianas são as mais utilizadas para fins industriais, uma vez que são mais viáveis economicamente, são de fácil disponibilidade, não sofrem interferências sazonais, apresentam rápida taxa de crescimento, alto rendimento, estabilidade e maior

atividade catalítica (Gurung et al., 2013). Estas particularidades possibilitam que as enzimas disponham de múltiplas aplicações biotecnológicas nos mais diversos setores, conforme discutido anteriormente.

Os principais microrganismos produtores de enzimas envolvidas em inúmeros produtos e processos biotecnológicos nos mais diversos segmentos como laticínios, panificação, bebidas, polpa e papel, ração animal e detergente, pertencem aos gêneros *Aspergillus*, *Bacillus*, *Streptomyces*, *Trichoderma* e *Lactobacillus* (Singh et al., 2016b).

2.6. Determinação e predição de estruturas

Diante da demanda por enzimas mais eficientes e de baixo custo, a caracterização estrutural nos permite compreender algumas características, como a especificidade pelo substrato, correlação estrutura-função, interação proteína-ligante e arquitetura de centros catalíticos. A partir desse conhecimento, é possível melhorar propriedades catalíticas de acordo com finalidades específicas, como solubilidade, termoestabilidade e especificidade pelo substrato. Assim, as enzimas podem ter sua eficiência catalítica aprimorada e a inibição reduzida durante os processos de produção de etanol de segunda geração e outras aplicações biotecnológicas.

A estrutura tridimensional de uma proteína pode ser predita a partir de abordagens computacionais, utilizando ferramentas de bioinformática para prever sua organização estrutural, ou obtida através de abordagens experimentais, como a cristalografia de raios-X e a espectroscopia de ressonância magnética nuclear (RMN). As informações estruturais obtidas a partir de vários métodos de determinação de estrutura, modelagem e bioinformática fazem parte da biologia estrutural integrativa, que busca reunir e interpretar tais informações a fim de ampliar a compreensão sobre sistemas biológicos de macromoléculas (Gutmanas et al., 2013; Rout e Sali, 2019).

2.6.1. Predição computacional de estruturas

Esta abordagem permite prever a estrutura de proteínas a partir de ferramentas de bioinformática, empregando as técnicas de modelagem comparativa, “threading”

ou reconhecimento de dobras, recombinação baseada em fragmentos e modelagem *ab initio* baseada em física (Deng et al., 2018). Cada uma destas técnicas possui suas vantagens e limitações, que devem ser consideradas durante sua escolha e de acordo com os dados e recursos disponíveis.

Na modelagem por homologia utiliza-se uma estrutura molde PDB (“Protein Data Bank”) com similaridade de sequência para prever a estrutura de uma proteína alvo. Atualmente, este método é o mais utilizado e o mais preciso; ele baseia-se na premissa que sequências semelhantes relacionadas evolutivamente possuem estruturas semelhantes (Deng et al., 2018). A precisão da modelagem por homologia depende do grau de similaridade entre a sequência alvo e a sequência molde, bem como da qualidade da resolução da estrutura molde depositada no PDB.

As etapas da modelagem por homologia consistem em seleção do modelo adequado, alinhamento da sequência alvo com a sequência da estrutura molde, modelagem estrutural a partir das coordenadas do modelo e avaliação do modelo (Deng et al., 2018; Kuhlman e Bradley, 2019).

Quando não há uma estrutura PDB com similaridade razoável, utiliza-se a técnica de reconhecimento de “motifs” ou dobras, ou “threading”, que consiste na busca por proteínas com similaridade estrutural à proteína alvo (Jones et al., 1992; Jones, 1999). Inicialmente, neste método realiza-se um alinhamento múltiplo de sequências relacionadas à proteína alvo, em seguida, as características estruturais locais são geradas com base na sequência da proteína alvo e seus homólogos e são utilizadas para construção do modelo tridimensional. Os modelos gerados são refinados e comparados entre si a fim de gerar a estrutura final (Kuhlman e Bradley, 2019).

A modelagem “*ab initio*” por não utilizar um molde, gera um modelo baseando-se nas leis da física e é utilizada principalmente para modelagem de pequenas moléculas (<100 resíduos) (Wu et al., 2007; Lee et al., 2009).

O método de recombinação baseada em fragmentos utiliza fragmentos curtos de estruturas conhecidas para construção de modelos. Existem diferentes ferramentas de predição para este método, tais como 3D-Shotgun e Frankenstein’s Monster, cada um utiliza abordagens diferentes para obtenção, análise e montagem dos fragmentos (Bujnicki, 2006).

2.6.2. Cristalografia de raios-X

As abordagens experimentais para obtenção de dados estruturais mais utilizados são a cristalografia de raios-X, a espectroscopia de ressonância magnética nuclear (RMN) e, atualmente em expansão, a microscopia crioelétrica (cryo-EM).

A RMN é uma técnica utilizada para determinação de estruturas e tem como vantagem a possibilidade de avaliar a interação e a movimentação dos alvos, no entanto, é limitada a pequenas moléculas (<20 kDa). O cryo-EM, por sua vez, é uma técnica relativamente nova que fornece dados tão detalhados quanto a cristalografia de raios-X e também é realizada em amostras vitrificadas em temperaturas criogênicas (Nitta et al., 2018).

No presente estudo foi utilizada a técnica de difração de raios-X, por este motivo as outras duas técnicas não são descritas detalhadamente, mas uma revisão breve pode ser encontrada em: (Wuthrich, 1990; Markwick et al., 2008; Nogales, 2015; Nitta et al., 2018; Bhella, 2019).

A cristalografia de difração de raios-X é a principal técnica utilizada para obtenção da estrutura tridimensional a partir de cristais que devem apresentar tamanho e qualidade suficientes para produzir imagens de difração (McPherson e Gavira, 2014), sendo que cristais de boa qualidade difratam com alta resolução (McPherson, 2017).

O método inicia com a purificação da proteína alvo, seja a partir de fontes naturais ou a partir de técnicas de DNA recombinante (McPherson, 2017). Quando necessário, a proteína deve ser concentrada até atingir uma concentração ideal para a formação de cristais, que geralmente ocorre a partir de 10 mg/mL⁻¹. A proteína purificada é submetida a diversas condições de cristalização, que por sua vez são constituídas por diferentes tampões, agentes precipitantes, sais e aditivos (Smyth e Martin, 2000).

O sucesso na formação dos cristais depende de algumas variáveis, como por exemplo as características da proteína, a concentração e homogeneidade da amostra, os agentes precipitantes, a presença de ligantes, pH e temperatura (McPherson e Gavira, 2014). Atualmente existem kits comerciais com diferentes condições precipitantes utilizados para “screening” inicial, permitindo que uma ampla quantidade

de condições sejam analisadas de uma única vez até que uma condição ideal seja encontrada (Smyth e Martin, 2000). Identificar as condições iniciais que formam algum material cristalino não é tão simples, uma vez que algumas proteínas resistem à formação de cristais e as causas disso podem ser desconhecidas; ao passo que a otimização da condição inicial é mais exigente e demanda mais tempo (McPherson e Gavira, 2014).

A cristalização de macromoléculas é um processo de precipitação lento e controlado que requer uma quantidade da proteína excedente ao limite de solubilidade para criar uma condição de supersaturação (McPherson e Gavira, 2014). Quando as condições corretas são encontradas e o equilíbrio da saturação é reestabelecido, os cristais são formados (McPherson e Gavira, 2014). No entanto, a obtenção de cristais de alta qualidade ainda é um fator limitante para a cristalografia de raios-X (Saridakis e Chayen, 2009).

A etapa de cristalização pode ser realizada por diálise ou pelo método de difusão de vapor em gota assente, da gota suspensa ou pelo método sanduíche, realizados em recipientes isolados (Figura 6) (Krauss et al., 2013; McPherson e Gavira, 2014).

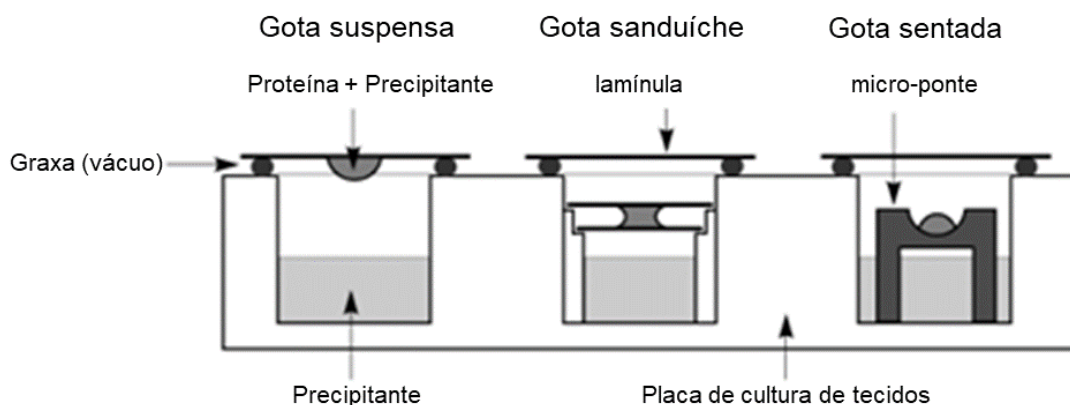


Figura 6. Diferentes estratégias utilizadas na cristalografia por difusão de vapor. Adaptado de (Wu et al., 2012).

Nos métodos de difusão de vapor a proteína purificada é misturada à solução de cristalização para formação da gota e equilibrada com a mesma solução precipitante armazenada no reservatório. A diferença entre eles está na posição da gota. No método da gota sentada a proteína fica disposta em um suporte no centro do

poço, enquanto que no método da gota suspensa a proteína encontra-se em uma lamínula suspensa e no método sanduíche ela permanece entre duas superfícies (Sun et al., 2010).

Uma vez produzidos e coletados, os cristais são imersos em uma solução crioprotetora e congelados em nitrogênio líquido a fim de evitar danos proporcionados pela radiação e submetidos a feixes de raios-X, os quais visam a formação de padrões de difração quando as condições são favoráveis (McPherson e Gavira, 2014). A partir dos avanços da Luz Síncrotron e das tecnologias dos detectores, a criocristalografia e o desenvolvimento de softwares mais sofisticados, a coleta de dados não é mais considerada um fator limitante (Ludwig, 2004; Taylor, 2010).

Resumidamente, a difração revela a densidade eletrônica de uma molécula e o padrão observado na difração é utilizado para determinar a intensidade de cada reflexão. O processamento destes dados é realizado inúmeras vezes com intuito de verificar a precisão do modelo e evitar possíveis erros (Smyth e Martin, 2000).

Em seguida, inicia-se a fase de determinação da estrutura e refinamento. A determinação da estrutura pode ser realizada pelo método de Substituição Molecular (MR), relativamente mais rápido e o mais utilizado atualmente. No método MR utilizam-se as coordenadas atômicas de uma proteína de estrutura similar à proteína alvo para resolver o problema de fases (Evans e McCoy, 2007; Taylor, 2010).

Na substituição molecular as fases de uma proteína molde estruturalmente similar à proteína alvo são utilizadas para estimar as fases da nova estrutura, calcular o mapa inicial e construir o modelo (Evans e McCoy, 2007). A escolha do modelo adequado influencia na precisão da nova estrutura e, normalmente, utiliza-se um modelo com similaridade de sequência >25%, embora possam existir algumas exceções (Evans e McCoy, 2007; Taylor, 2010).

Outros métodos são utilizados para obtenção das fases, como a difração anômala de comprimento de onda único (SAD) e de comprimento de onda múltiplo (MAD), onde átomos pesados são introduzidos na estrutura da proteína durante a etapa de expressão proteica e suas posições são utilizadas para resolver o problema de fases (Taylor, 2010). Em alguns casos, como em metaloproteínas, os átomos pesados ocorrem naturalmente na sua estrutura.

A etapa de refinamento é um processo iterativo não linear que tem como objetivo minimizar uma função baseada nas diferenças entre as amplitudes observadas experimentalmente e as calculadas a partir do modelo. Feito isto, a estrutura tridimensional pode ser depositada em bancos de dados de proteínas, como o PDB (<https://www.rcsb.org/>), sob um número de identificação.

3. REFERÊNCIAS

Adesioye, F.A., Makhalanyane, T.P., Biely, P. e Cowan, D.A. (2016). Phylogeny , classification and metagenomic bioprospecting of microbial acetyl xylan esterases. **Enzyme and Microbial Technology** 93–94: 79–91.

Agrawal, K., Chaturvedi, V. e Verma, P. (2018). Fungal laccase discovered but yet undiscovered. **Bioresources and Bioprocessing** 5: 1–12.

Akoh, C.C., Chang, S.W., Lee, G.C. e Shaw, J.F. (2008). Biocatalysis for the production of industrial products and functional foods from rice and other agricultural produce. **Journal of Agricultural and Food Chemistry** 56: 10445–10451.

McPherson, A. (2017). Protein Crystallization. In: Wlodawer, A., Jaskolski, M., & Dauter, Z. (Eds.), **Protein Crystallography: Methods and Protocols**. New York, NY: Springer Nature, p. 17–50.

Amorim, H. V., Lopes, M.L., de Castro Oliveira, J.V., Buckeridge, M.S. e Goldman, G.H. (2011). Scientific challenges of bioethanol production in Brazil. **Applied microbiology and biotechnology** 91: 1267–75.

Bacovsky, D., Dallos, M. e Wörgetter, M. (2010). **Status of 2nd Generation Biofuels Demonstration Facilities in June 2010**, IEA Bioenergy Task 39. In Anais do IEA Bioenergy Task 39: Commercializing 1st and 2nd generation liquid biofuels from biomass.

Balcázar-López, E., Méndez-Lorenzo, L.H., Batista-García, R.A., Esquivel-Naranjo, U., Ayala, M., Kumar, V.V., Savary, O., Cabana, H., et al. (2016). Xenobiotic compounds degradation by heterologous expression of a trametes sanguineus laccase in trichoderma atroviride. **PLoS ONE** 11: e0147997.

Bento, I., Martins, L.O., Lopes, G.G., Carrondo, M.A. e Lindley, P.F. (2005). Dioxygen reduction by multi-copper oxidases; a structural perspective. **Dalton Transactions** 4: 3507–3513.

Bhella, D. (2019). Cryo-electron microscopy: an introduction to the technique, and considerations when working to establish a national facility. **Biophysical Reviews** 515–519.

Branco, R.H.R., Serafim, L.S. e Xavier, A.M.R.B. (2019). Second generation bioethanol production: On the use of pulp and paper industry wastes as feedstock. **Fermentation** 5: 1–30.

Bujnicki, J.M. (2006). Protein-structure prediction by recombination of fragments. **ChemBioChem** 7: 19–27.

Cammarota, M.C. e Freire, D.M.G. (2006). A review on hydrolytic enzymes in the treatment of wastewater with high oil and grease content. **Bioresource Technology** 97: 2195–2210.

Canilha, L., Kumar Chandel, A., dos Santos Milessi, T.S., Fernandes Antunes, F.A., da Costa Freitas, W.L., das Graças Almeida Felipe, M. e da Silva, S.S. (2012). Bioconversion of sugarcane biomass into ethanol: an overview about composition, pretreatment methods, detoxification of hydrolysates, enzymatic saccharification, and ethanol fermentation. **Journal of biomedicine & biotechnology** 2012: 1–15.

Chakdar, H., Kumar, M., Pandiyan, K., Singh, A., Nanjappan, K., Kashyap, P.L. e Srivastava, A.K. (2016). Bacterial xylanases: biology to biotechnology. **3 Biotech** 6: 150.

Choi, J.M., Han, S.S. e Kim, H.S. (2015). Industrial applications of enzyme biocatalysis: Current status and future aspects. **Biotechnology Advances** 33: 1443–1454.

Christopher, L.P., Yao, B. e Ji, Y. (2014). Lignin biodegradation with laccase-mediator systems. **Frontiers in Energy Research** 2: 1–13.

Cotana, F., Cavalaglio, G., Nicolini, A., Gelosia, M., Coccia, V., Petrozzi, A. e Brinchi, L. (2014). Lignin as co-product of second generation bioethanol production from lignocellulosic biomass. **Energy Procedia** Elsevier B.V. 45: 52–60.

De Bhowmick, G., Sarmah, A.K. e Sen, R. (2018). Lignocellulosic biorefinery as a model for sustainable development of biofuels and value added products. **Bioresource Technology** 247: 1144–1154.

Deng, H., Jia, Y. e Zhang, Y. (2018). Protein structure prediction. **International Journal of Modern Physics B** 32: 1–18.

Dos Santos, L.V., De Barros Grassi, M.C., Gallardo, J.C.M., Pirolla, R.A.S., Calderón, L.L., De Carvalho-Netto, O.V., Parreiras, L.S., Camargo, E.L.O., et al. (2016). Second-Generation Ethanol: The Need is Becoming a Reality. **Industrial Biotechnology** 12: 40–57.

Ebringerová, A., & Heinze, T. (2000). Xylan and xylan derivatives–biopolymers with valuable properties, 1. Naturally occurring xylans structures, isolation procedures and properties. **Macromolecular rapid communications** 21: 542-556.

Evans, P. e McCoy, A. (2007). An introduction to molecular replacement. **Acta Crystallographica Section D: Biological Crystallography** International Union of Crystallography 64: 1–10.

Falade, A.O., Nwodo, U.U., Iweriebor, B.C., Green, E., Mabinya, L. V. e Okoh, A.I. (2017). Lignin peroxidase functionalities and prospective applications. **MicrobiologyOpen** 6: e00394.

Falck, P., Linares-pastén, J.A., Adlercreutz, P. e Karlsson, E.N. (2016). Characterization of a family 43 β -xylosidase from the xylooligosaccharide utilizing putative probiotic *Weissella* sp. strain 92. **Microbial Biology** 26: 193–202.

Fazary, A.E. e Ju, Y.H. (2007). Feruloyl esterases as biotechnological tools: Current and future perspectives. **Acta Biochimica et Biophysica Sinica** 39: 811–828.

Food and Agriculture Organization FAO. (2018). **FAOSTAT. Sugar cane production.**

Giardina, P., Faraco, V., Pezzella, C., Piscitelli, A., Vanhulle, S. e Sannia, G. (2010). Laccases: A never-ending story. **Cellular and Molecular Life Sciences** 67: 369–385.

Gomes-Pepe, E.S., Sierra, E.G.M., Pereira, M.R., Castellane, T.C.L. e De Lemos, E.G.M. (2016). Bg10: A novel metagenomics alcohol-tolerant and glucose-stimulated gh1 β -glucosidase suitable for lactose-free milk preparation. **PLoS ONE** 11: 1–25.

Gunne, M., Höppner, A., Hagedoorn, P.L. e Urlacher, V.B. (2014). Structural and redox properties of the small laccase Ssl1 from *Streptomyces svaceus*. **FEBS Journal** 281: 4307–4318.

Gurung, N., Ray, S., Bose, S., Rai, V. e K, W.F. (2013). A Broader View : Microbial Enzymes and Their Relevance in Industries , Medicine , and Beyond enzyme and its use were well known to the mankind but. 2013: 1-19.

Gutmanas, A., Oldfield, T.J., Patwardhan, A., Sen, S., Velankar, S. e Kleywegt, G.J. (2013). The role of structural bioinformatics resources in the era of integrative structural biology. **Acta Crystallographica Section D: Biological Crystallography** 69: 710–721.

Hart, D.J. e Tarendeau, F. (2006). Combinatorial library approaches for improving soluble protein expression in *Escherichia coli*. **Acta Crystallographica Section D: Biological Crystallography** 62: 19–26.

Hendrickson, W. A., Horton, J. R., & LeMaster, D. M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. **The EMBO journal** 9: 1665-1672.

Horn, S. J., Vaaje-Kolstad, G., Westereng, B., & Eijsink, V. (2012). Novel enzymes for the degradation of cellulose. **Biotechnology for biofuels** 5: 45.

Huy, N.D., Thayumanavan, P., Kwon, T.H. e Park, S.M. (2013). Characterization of a recombinant bifunctional xylosidase/arabinofuranosidase from *Phanerochaete chrysosporium*. **Journal of Bioscience and Bioengineering - Elsevier** 116: 152–159.

Irmak, S. (2017). Biomass as Raw Material for Production of High-Value Products. In: Jaya Shankar Tumuluru (Eds.), **Biomass Volume Estimation and Valorization for Energy**. London: IntechOpen, p. 201–225.

Johnson, E. (2016). Integrated enzyme production lowers the cost of cellulosic ethanol. **Biofuels, Bioproducts and Biorefining** 10: 164–174.

Jones, D.T. (1999). GenTHREADER: An efficient and reliable protein fold recognition method for genomic sequences. **Journal of Molecular Biology** 287: 797–815.

Jones, D.T., Taylor, W.R. e Thornton, J.M. (1992). A new approach to protein fold recognition. **Nature** 358: 86–89.

Juturu, V. e Wu, J.C. (2012). Microbial xylanases: Engineering, production and industrial applications. **Biotechnology Advances** Elsevier Inc. 30: 1219–1227.

Kalim, B., Böhringer, N., Ali, N. e Schäberle, T. (2015). Xylanases—from Microbial Origin to Industrial Application. **British Biotechnology Journal** 7: 1–20.

Kamble, R.D. e Jadhav, A.R. (2012). Isolation, purification, and characterization of xylanase produced by a new species of bacillus in solid state fermentation. **International Journal of Microbiology** 2012: 1–8.

Kim, M. e Day, D.F. (2011). Composition of sugar cane, energy cane, and sweet sorghum suitable for ethanol production at Louisiana sugar mills. **Journal of Industrial Microbiology and Biotechnology** 38: 803–807.

Klein-Marcuschamer, D., Oleskowicz-Popiel, P., Simmons, B. a. e Blanch, H.W. (2012). The challenge of enzyme cost in the production of lignocellulosic biofuels. **Biotechnology and Bioengineering** 109: 1083–1087.

Knob, A., Terrasan, C.R.F. e Carmona, E.C. (2010). β -Xylosidases from filamentous fungi: An overview. **World Journal of Microbiology and Biotechnology** 26: 389–407.

Kocabas, D.S., Güder, S. e Özben, N. (2015). Purification strategies and properties of a low-molecular weight xylanase and its application in agricultural waste biomass hydrolysis. **Journal of Molecular Catalysis B: Enzymatic** 115: 66–75.

Komori, H., Miyazaki, K. e Higuchi, Y. (2009). X-ray structure of a two-domain type laccase: A missing link in the evolution of multi-copper proteins. **FEBS Letters** 583: 1189–1195.

Krauss, I.R., Merlino, A., Vergara, A. e Sica, F. (2013). An overview of biological macromolecule crystallization. **International Journal of Molecular Sciences** 14: 11643–11691.

Kuhlman, B. e Bradley, P. (2019). Advances in protein structure prediction and design. **Nature Reviews Molecular Cell Biology** Springer US 20: 681–697.

Lee, J., Wu, S. e Zhang, Y. (2009). Ab initio protein structure prediction. **From Protein Structure to Function with Bioinformatics** 3–25.

Lee, S., Kang, M., Bae, J.H., Sohn, J.H. e Sung, B.H. (2019). Bacterial Valorization of Lignin: Strains, Enzymes, Conversion Pathways, Biosensors, and Perspectives. **Frontiers in Bioengineering and Biotechnology** 7: 1–18.

Li, J., Cai, S., Luo, Y. e Dong, X. (2011). Three feruloyl esterases in *Cellulosilyticum ruminicola* H1 act synergistically to hydrolyze esterified polysaccharides. **Applied and Environmental Microbiology** 77: 6141–6147.

Ludwig, M.L. (2004). X-Ray Determination of 3-D Structure in Proteins. In: J.Lennarz, W. & M. Daniel Lane (Eds.), **Encyclopedia of Biological Chemistry**, p. 422–428.

Machczynski, M.C., Vijgenboom, E., Samyn, B. e Canters, G.W. (2004). Characterization of SLAC: A small laccase from *Streptomyces coelicolor* with unprecedented activity. **Protein Science** 13: 2388–2397.

Madhavan, A., Sindhu, R., Binod, P., Sukumaran, R.K. e Pandey, A. (2017). Strategies for design of improved biocatalysts for industrial applications. **Bioresource Technology** 245: 1304–1313.

Maijala, P., Kleen, M., Westin, C., Poppius-Levlin, K., Herranen, K., Lehto, J.H., Reponen, P., Mäentausta, O., et al. (2008). Biomechanical pulping of softwood with enzymes and white-rot fungus *Physisporinus rivulosus*. **Enzyme and Microbial Technology** 43: 169–177.

Malico, I., Nepomuceno Pereira, R., Gonçalves, A.C. e Sousa, A.M.O. (2019). Current status and future perspectives for energy production from solid biomass in the European industry. **Renewable and Sustainable Energy Reviews** 112: 960–977.

Markwick, P.R.L., Malliavin, T. e Nilges, M. (2008). Structural biology by NMR: Structure, dynamics, and interactions. **PLoS Computational Biology** 4: e1000168.

McPherson, A. e Gavira, J.A. (2014). Introduction to protein crystallization. **Acta Crystallographica Section F: Structural Biology Communications** International Union of Crystallography 70: 2–20.

Minussi, R.C., Pastore, G.M. e Durán, N. (2002). Potential applications of laccase in the food industry. **Trends in Food Science and Technology** 13: 205–216.

Nestl, B.M., Nebel, B.A. e Hauer, B. (2011). Recent progress in industrial biocatalysis. **Current Opinion in Chemical Biology** 15: 187–193.

Nitta, R., Imasaki, T. e Nitta, E. (2018). Recent progress in structural biology: Lessons from our research history. **Microscopy** 67: 187–195.

Nogales, E. (2015). The development of cryo-EM into a mainstream structural biology technique. **Nature Methods** 13: 24–27.

Ojha, B.K., Singh, P.K. e Shrivastava, N. (2018). Enzymes in the animal feed industry. In: Kuddus, M. (Eds.), **Enzymes in Food Biotechnology: Production, Applications, and Future Prospects** Academic Press, p. 93–109.

Osma, J.F., Toca-Herrera, J.L. e Rodríguez-Couto, S. (2010). Uses of laccases in the food industry. **Enzyme Research** 918761.

Pérez, J., Muñoz-Dorado, J., De La Rubia, T. e Martínez, J. (2002). Biodegradation and biological treatments of cellulose, hemicellulose and lignin: An overview. **International Microbiology** 5: 53–63.

Polizeli, M.L.T.M., Rizzatti, A.C.S., Monti, R., Terenzi, H.F., Jorge, J.A. e Amorim, D.S. (2005). Xylanases from fungi: Properties and industrial applications. **Applied Microbiology and Biotechnology** 67: 577–591.

Pollegioni, L., Tonin, F. e Rosini, E. (2015). Lignin-degrading enzymes. **FEBS Journal** 282: 1190–1213.

Riva, S. (2006). Laccases: blue enzymes for green chemistry. **Trends in Biotechnology** 24: 219–226.

Robak, K. e Balcerek, M. (2018). Review of second generation bioethanol production from residual biomass. **Food Technology and Biotechnology** 56: 174–187.

Rodríguez Couto, S. e Toca Herrera, J.L. (2006). Industrial and biotechnological applications of laccases: A review. **Biotechnology Advances** 24: 500–513.

Roth, S. e Spiess, A.C. (2015). Laccases for biorefinery applications: a critical review on challenges and perspectives. **Bioprocess and Biosystems Engineering** Springer Berlin Heidelberg 38: 2285–2313.

Rout, M.P. e Sali, A. (2019). Principles for Integrative Structural Biology Studies. **Cell** Elsevier Inc. 177: 1384–1403.

Rubin, E.M. (2008). Genomics of cellulosic biofuels. **Nature** 454: 841–845.

Saha, B.C. (2000). α -L-Arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. **Biotechnology Advances** 18: 403–423.

Saridakis, E. e Chayen, N.E. (2009). Towards a “universal” nucleant for protein crystallization. **Trends in Biotechnology** 27: 99–106.

Septiningruma, K., Ohi, H., Waeonukul, R., Pason, P., Tachaapaikoon, C., Ratanakhanokchai, K., Sermsathanaswadi, J., Deng, L., et al. (2015). Enzyme and Microbial Technology The GH67 α -glucuronidase of *Paenibacillus curdlanolyticus* B-6 removes hexenuronic acid groups and facilitates biodegradation of the model xylooligosaccharide hexenuronosyl xylotriose. **Enzyme and Microbial Technology** Elsevier Inc. 71: 28–35.

Shallom, D. e Shoham, Y. (2003). Microbial hemicellulases. **Current Opinion in Microbiology** 6: 219–228.

Sharma, H.K., Qin, W. e Xu, C.C. (2018). Cellobiohydrolase (cbh) activity assays. **Methods in Molecular Biology** 1796: 185–199.

Singh, D. e Gupta, N. (2020). Microbial Laccase: a robust enzyme and its industrial applications. **Biologia** Biologia 1: 1–11.

Singh, R., Kumar, M., Mittal, A. e Mehta, P.K. (2016). Microbial enzymes: industrial progress in 21st century. **3 Biotech** Springer Berlin Heidelberg 6: 1-15.

Sirim, D., Wagner, F., Wang, L., Schmid, R.D. e Pleiss, J. (2011). The Laccase Engineering Database: A classification and analysis system for laccases and related Multicopper oxidases. **Database** 2011: 1–7.

Skálová, T., Dohnálek, J., Østergaard, L.H., Østergaard, P.R., Kolenko, P., Dušková, J., Štěpánková, A. e Hašek, J. (2009). The Structure of the Small Laccase from *Streptomyces coelicolor* Reveals a Link between Laccases and Nitrite Reductases. **Journal of Molecular Biology** 385: 1165–1178.

Smyth, M.S. e Martin, J.H.J. (2000). x Ray crystallography. **Molecular Pathology** 53: 8–14.

Solomon, E.I., Sundaram, U.M. e Machonkin, T.E. (1996). Multicopper oxidases and oxygenases. **Chemical Reviews** 96: 2563–2606.

Sun, L., Li, J., Xu, C., Yu, F., Zhou, H., Tang, L. e He, J. (2010). The sandwich method for protein crystallization and its effect on crystal growth. **Acta Biochimica et Biophysica Sinica** 42: 332–336.

Takkellapati, S., Li, T. e Gonzalez, M.A. (2018). An overview of biorefinery-derived platform chemicals from a cellulose and hemicellulose biorefinery. **Clean Technologies and Environmental Policy** 20: 1615–1630.

Taylor, G.L. (2010). Introduction to phasing. **Acta Crystallographica Section D: Biological Crystallography** International Union of Crystallography 66: 325–338.

Till, M., Goldstone, D., Card, G., Attwood, G. T., Moon, C. D., & Arcus, V. L. (2014). Structural analysis of the GH43 enzyme Xsa43E from *Butyrivibrio proteoclasticus*. **Acta crystallographica. Section F, Structural biology communications**, 70: 1193–1198.

Tishchenko, S., Gabdulkhakov, A., Trubitsina, L., Lisov, A., Zakharova, M. e Leontievsky, A. (2015). Crystallization and X-ray diffraction studies of a two-domain laccase from *Streptomyces griseoflavus*. **Acta Crystallographica Section:F Structural Biology Communications** 71: 1200–1204.

Topakas, E., Vafiadi, C. e Christakopoulos, P. (2007). Microbial production, characterization and applications of feruloyl esterases. **Process Biochemistry** 42: 497–509.

Unuofin, J.O., Okoh, A.I. e Nwodo, U.U. (2019). Aptitude of oxidative enzymes for treatment of wastewater pollutants: A laccase perspective. **Molecules** 24: 2064.

Uzuner, S. e Cekmecelioglu, D. (2018). Enzymes in the beverage industry. **Enzymes in Food Biotechnology: Production, Applications, and Future Prospects**, p. 29–43.

Virk, A.P., Sharma, P. e Capalash, N. (2012). Use of laccase in pulp and paper industry. **Biotechnology Progress** 28: 21–32.

Vrsanska, M., Voberkova, S., Langer, V., Palovcikova, D., Moulick, A., Adam, V. e Kopel, P. (2016). Induction of laccase, lignin peroxidase and manganese peroxidase activities in white-rot fungi using copper complexes. **Molecules** 21: 1553.

Wang, G., Huang, X., Ng, T.B., Lin, J. e Ye, X.Y. (2014). High phylogenetic diversity of glycosyl hydrolase family 10 and 11 xylanases in the sediment of Lake Dabusu in China. **PLoS ONE** 9: e112798.

Werpy, T. e Petersen, G. (2004). **Top value added chemicals from biomass volume I — Results of screening for potential candidates from sugars and synthesis gas.**

Wohlgemuth, R. (2010). Biocatalysis-key to sustainable industrial chemistry. **Current Opinion in Biotechnology** 21: 713–724.

Wong, D.W.S. (2006). Feruloyl esterase: a key enzyme in biomass degradation. **Applied Biochemistry and Biotechnology** 133: 87–112.

Wong, D.W.S. (2009). Structure and action mechanism of ligninolytic enzymes. **Applied Biochemistry and Biotechnology** 157: 174–209.

Wu, Q., Sibanda, L., Ochi, T., Bolanos-Garcia, V.M., Blundell, T.L. e Chirgadze, D.Y. (2012). Spatial and temporal organisation of multiprotein systems of cell regulation and signalling: what can we learn from NHEJ system of double-strand break repair?. In: Carrondo MA & P, S. (Eds.), **Macromolecular crystallography**. Springer Netherlands, p. 1–31.

Wu, S., Skolnick, J. e Zhang, Y. (2007). Ab initio modeling of small proteins by iterative TASSER simulations. **BMC Biology** 5: 1–10.

Wuthrich, K. (1990). Protein structure determination in solution by NMR spectroscopy. **Journal of Biological Chemistry** 243: 45–50.

Yan, R., Vuong, T. V., Wang, W. e Master, E.R. (2017). Action of a GH115 α -glucuronidase from *Amphibacillus xylanus* at alkaline condition promotes release of 4-O -methylglucopyranosyluronic acid from glucuronoxylan and arabinoglucuronoxylan. **Enzyme and Microbial Technology** 104: 22–28.

Zeng, Y., Himmel, M. E., & Ding, S. Y. (2017). Visualizing chemical functionality in plant cell walls. **Biotechnology for biofuels** 10: 1-16.

Zheng, F., Cui, B.K., Wu, X.J., Meng, G., Liu, H.X. e Si, J. (2016). Immobilization of laccase onto chitosan beads to enhance its capability to degrade synthetic dyes. **International Biodeterioration and Biodegradation** 110: 69–78.

Zheng, Y., Shi, J., Tu, M. e Cheng, Y.-S. (2017). Principles and Development of Lignocellulosic Biomass Pretreatment for Biofuels. In: Li, Y. & Ge, X. (Eds.), **Advances in Bioenergy**. Amsterdã: Elsevier B.V., p. 1–68.

CAPÍTULO 2 – Estrutura cristalográfica e estudos de ancoragem molecular da β -glicosidase Bg10 estimulada pela glicose

RESUMO – As β -glicosidases atuam sobre a cadeia principal da celulose e pertencem, principalmente, às famílias GH1 e GH3. β -glicosidases GH1 são amplamente investigadas por apresentarem características de tolerância e/ou estímulo pela glicose. Tais características favorecem muitos processos biotecnológicos uma vez que as enzimas não apresentam inibição alostérica pelo próprio produto (glicose), proporcionando maior rendimento quando utilizada em baixas concentrações, o que reduz consideravelmente os custos de produção. O presente estudo teve como objetivo caracterizar estruturalmente uma β -glicosidase tolerante e estimulada pela glicose (Bg10) obtida de uma biblioteca metagenômica desenvolvida a partir de DNA de comunidade microbiana de solo sob serapilheira de Eucalipto. A enzima Bg10 foi expressa, purificada e submetida a ensaios de cristalografia por difração de raios-X para determinação da estrutura cristalográfica e a ensaios de ancoragem molecular com glicose, celobiose, celotriose, celotetraose e *p*NP-G. Os melhores dados de difração coletados foram a uma resolução de 2,3 Å. A estrutura tridimensional e os resíduos do sítio ativo da Bg10 apresentaram conformação típica às outras glicosidases GH1. A entrada do canal catalítico é mais neutra em comparação ao sítio ativo, e a Bg10 possui um canal profundo e estreito com a presença de resíduos “gatekeepers” neutros que regulam a passagem e a orientação da glicose. Além disso, foi possível observar o modo de ligação da glicose e outros substratos no sítio ativo, bem como identificar os resíduos de aminoácidos envolvidos em possíveis interações proteína-ligante. As propriedades observadas na estrutura da Bg10 são características de GH1 tolerantes e estimuladas pela glicose, o que corrobora com estudos de cinética enzimática realizados em estudo anterior. Estes resultados permitiram compreender alguns de seus principais mecanismos de tolerância e estímulo que, futuramente aliados a técnicas de mutagênese sítio dirigida, poderão contribuir consideravelmente para a melhoria de β -glicosidases GH1, úteis no processo de sacarificação da biomassa lignocelulósica, dentre outras aplicações biotecnológicas.

Palavras-chave: ancoragem molecular, β -glicosidase, difração de raios-X, substituição molecular

1. INTRODUÇÃO

A aplicação de β -glicosidases (E.C. 3.2.1.21) se desenvolveu notavelmente nos últimos anos, principalmente, quanto ao seu uso na exploração na sacarificação da fibra de celulose nas indústrias de etanol de segunda geração, em aplicações químicas e processos industriais de alimentos e bebidas (Singh et al., 2016a; Srivastava et al., 2019; Zhang et al., 2019).

As β -glicosidases são celulasas pertencentes principalmente às famílias GH1 e GH3 (Glicosil Hidrolases) e atuam sinergicamente com endoglicanases e celobiohidrolases na hidrólise de ligações β -glicosídicas da celulose, liberando unidades de glicose.

A maioria das β -glicosidases caracterizadas são sensíveis ao produto formado durante a hidrólise - a glicose - e acabam sendo inibidas por ela (Zhao et al., 2013). Contudo, algumas β -glicosidases GH1 apresentam diferentes graus de tolerância à glicose, variando entre 10 a 1000 vezes mais tolerância do que glicosidases GH3 (De Giuseppe et al., 2014). Além desta característica, um pequeno número de GH1 são ainda estimuladas pela glicose, o que as tornam mais vantajosas do ponto de vista industrial, uma vez que a inibição enzimática enfrentada por glicosidases não tolerantes implica em menor rendimento do produto de interesse, sendo necessário o uso excedente de enzima, elevando os custos de produção, ao passo que uma enzima tolerante e eventualmente estimulada apresenta maior eficiência catalítica (De Giuseppe et al., 2014).

A fim de explorar completamente as características de β -glicosidases GH1 é necessário conhecer sua estrutura tridimensional para complementar as informações sobre as relações estrutura-função e, assim, compreender os mecanismos biológicos envolvidos na tolerância e estímulo pela glicose nestas macromoléculas. Ademais, as informações obtidas a partir de dados estruturais podem ser essenciais para aprimorar suas propriedades de acordo com propósitos específicos a partir da engenharia de enzimas.

No laboratório LBMP foi previamente relatada a caracterização bioquímica de uma β -glicosidase GH1 que apresentou tolerância e estímulo à glicose (até 350 e 100 mM, respectivamente) e tolerância a álcoois revelando excelentes propriedades

biotecnológicas para aplicação na indústria de biocombustível (Gomes-Pepe et al., 2016).

Diante disso, este estudo teve como objetivo realizar ensaios de cristalografia por difração de raios-X e determinação da estrutura cristalográfica da β -glicosidase metagenômica tolerante e estimulada pela glicose, bem como realizar ensaios de ancoragem molecular a fim de avaliar as propriedades estruturais gerais e do seu centro catalítico e prever possíveis interações proteína-ligante, características que podem estar diretamente relacionadas ao mecanismo de tolerância e estímulo pela glicose.

2. MATERIAL E MÉTODOS

2.1. Origem da amostra e características enzimáticas

No laboratório LBMP foi previamente relatada a clonagem e caracterização de uma β -glicosidase GH1, denominada Bg10 (número de acesso NCBI KX364386), obtida de uma biblioteca metagenômica desenvolvida a partir de DNA de comunidade microbiana de solo sob serapilheira de Eucalipto (Gomes et al., 2013; Gomes-Pepe et al., 2016). A enzima Bg10 foi predita como sendo um homodímero de aproximadamente 116.66 kDa, com pH e temperatura ótimos de 7.0 e 37 °C, respectivamente.

A atividade meia-vida da Bg10 foi retida a 350 mM de glicose, sendo que 90% da atividade foi observada em 150 mM, e a 500 mM a atividade residual manteve-se a 30%. Além disso, a atividade foi estimulada em 36.8% e 22% com 50 e 100 mM de glicose, respectivamente. Outra característica relevante para o processo de produção de biocombustível que esta enzima apresentou foi a tolerância ao etanol e ao metanol até concentrações de 500 mM e 400 mM, respectivamente (Gomes-Pepe et al., 2016).

2.2. Expressão e purificação da proteína recombinante

O gene Bg10 foi previamente clonado em vetor pET-28a(+) (Novagen, Madison, Wisconsin, Estados Unidos) e a proteína recombinante expressa em *E. coli* BL21 (DE3) (Novagen) (Gomes-Pepe et al., 2016). Para os experimentos realizados

neste estudo, as condições variaram parcialmente. Inicialmente, as células se desenvolveram em meio líquido LB suplementado com canamicina ($50 \mu\text{g/mL}^{-1}$) durante 16 horas a 37°C e 200 rpm em agitador orbital. Em seguida, 10% (v/v) do pré-inóculo foi transferido para 1000 mL de LB e incubado nas mesmas condições até atingir a DO_{600} entre 0,4 – 0,6, antes da cultura ser induzida por 0,1 mM de IPTG. Após 6 horas de agitação, a cultura foi centrifugada ($11000 \times g$ a 4°C durante 20 minutos), o sobrenadante descartado e o “pellet” ressuspendido em tampão de lise (20 mM Tris-HCl pH 8, 200 mM NaCl, 20 mM Imidazol, 5% (v/v) de glicerol, $20 \mu\text{g mL}^{-1}$ lisozima, $20 \mu\text{g mL}^{-1}$ DNase I, 1 pastilha Pierce Protease Inhibitor (Thermo Fisher Scientific, Waltham, Massachusetts, EUA), lisado em homogeneizador de alta pressão (APV Model 2000 Homogenizer, Delavan, Wisconsin, EUA) a 11600 psi e centrifugado a $27.000 \times g$ a 4°C durante 30 minutos. A proteína recombinante presente na fração solúvel foi purificada por cromatografia de afinidade por íons metálicos imobilizados (IMAC) utilizando uma coluna Sepharose 5 mL (HisTrap HP – GE Healthcare, Little Chalfont, Reino Unido), seguindo um gradiente crescente de imidazol, sendo eluída, principalmente, na concentração de 500 mM de imidazol. As frações eluídas foram reunidas e concentradas através de filtros Amicon® Ultra-15 (Merck MilliPore, Burlington, Massachusetts, EUA) com capacidade de retenção de moléculas com massa molecular superior a 30 kDa e, posteriormente, purificadas por cromatografia líquida por exclusão molecular (SEC) utilizando uma coluna Superdex 200 Increase 10/300 GL (GE Healthcare) conectada em um aparelho de automatização cromatográfica ÄKTA FPLC System (GE Healthcare) e equilibrada em tampão Tris HCl pH 9, 200 mM NaCl e 10% (v/v) de glicerol.

A concentração da proteína purificada foi calculada a partir da medida de absorção a 280 nm utilizando um espectrofotômetro NanoDrop 1000 (Thermo Fisher Scientific) com base nos valores do coeficiente de extinção molar ($71.515 \text{ M}^{-1} \text{ cm}^{-1}$) e na massa molecular do monômero (58,33 kDa), conforme parâmetros preditos com base na sequência de aminoácidos com auxílio do software ProtParam (“ProtParam tool - ExPASy”, <https://web.expasy.org/protparam/>).

2.3. Cristalização

Os primeiros ensaios de cristalização e o “screening” inicial foram realizados no Laboratório Nacional de Biociências (LNBio), Campinas – SP, através de proposta de pesquisa enviada pela pesquisadora Dra. Elisângela Soares Gomes Pepe. Os cristais obtidos nas condições avaliadas eram frágeis e quebradiços, o que impediu sua coleta e posterior análise por difração de raios-X. Diante disso, a amostra da enzima Bg10 foi gentilmente cedida pela pesquisadora para a presente pesquisa.

Os ensaios descritos a seguir foram realizados durante estágio doutoral no Laboratório de Cristalografia Aplicada à Indústria e a Medicina, do Instituto de Tecnologia Química e Biológica (ITQB) da Universidade Nova de Lisboa (UNL) em Oeiras, Portugal, sob orientação do pesquisador Dr. Pedro Matias.

Os novos experimentos de cristalização foram realizados manualmente em placas de 24 poços utilizando o método de difusão de vapor em gota assente, com um volume de gota de 1 μL (variando a proporção da proteína e do precipitante) e volume de reservatório de 500 μL . As placas de cristalização foram seladas e mantidas em ambientes sob temperatura controlada (4 °C, 20 °C e 30 °C). Os ensaios iniciais foram realizados utilizando kits comerciais JCSG+ Screen, Structure I & II, Pact Premier, The BCS Screen, Morpheus HT-96 (Molecular Dimensions, Sheffield, Inglaterra) e os kits The Pact Suíte e The AmSO₄ Suíte (Qiagen, Hilden, Alemanha). Neste caso, cada condição (gota) foi realizada em triplicata.

As gotas foram examinadas sob uma lupa estereoscópica diariamente e, por conseguinte, as condições iniciais foram otimizadas em placas de 96 poços preparadas pelo robô de pipetagem Mosquito LCP (SPT Labtech, Royston, Reino Unido) (Figura 1). Nesta condição, a gota era formada por 250 nL da proteína (50% v/v), 200 nL da solução de cristalização (40% v/v) dos kits comerciais e 50 nL de aditivos (10% v/v). A condição ideal estabelecida foi avaliada com os kits de aditivos Additive Screen HT (Hampton Research, Aliso Viejo, California, EUA) e Morpheus Additive Screen (Molecular Dimensions). As otimizações foram necessárias a fim de se obter cristais com a melhor morfologia e melhor tamanho possível.



Figura 1. Imagem ilustrativa do robô de pipetagem Mosquito LCP (SPT Labtech) com câmara de controle de umidade que impede a evaporação da amostra.

As placas foram visualizadas em lupa estereoscópica e os cristais únicos, com faces e bordas lisas e com tamanho a partir de 200 μm foram coletados, imersos em crioprotetor (15% [v/v] de glicerol incorporado à sua respectiva solução de cristalização), rapidamente congelados em nitrogênio líquido e usados para coleta de dados.

2.4. Coleta e processamento de dados

Os cristais da Bg10 obtidos de diferentes condições foram submetidos às análises na Linha de Luz I04 no sincrotron Diamond Light Source em Oxfordshire, Reino Unido, usando um detector Eiger2 XE 16M. Os melhores dados medidos para os cristais Bg10 foram usados para resolver a estrutura da proteína. Os dados de difração foram processados com XDS (Kabsch et al., 2010) através da suíte autoPROC (Vonrhein et al., 2011), e posteriormente convertidos para utilização na suíte de programas CCP4 (Winn et al., 2011). Uma busca por estruturas de proteínas similares foi realizada no banco de dados PDBe ("Protein Data Bank in Europe"). A estrutura da proteína Bg10 foi resolvida usando a técnica de substituição molecular com o programa PHASER (McCoy et al., 2007) integrado ao CCP4 (Winn et al., 2011), utilizando como modelo a estrutura de uma β -glicosidase de *Streptomyces* sp. (PDB 1gnx com identidade de sequência de 69%). O modelo foi corrigido numa estação gráfica 3D com COOT (Emsley et al., 2010) e o refinamento da estrutura realizado com a suíte de programas PHENIX (Adams et al., 2010). O modelo final foi validado através da ferramenta Molprobit (Chen et al., 2010).

As imagens da estrutura foram obtidas e visualizadas no software PyMol (The PyMOL Molecular Graphics System, Version 2.3.0 Schrödinger, LLC.).

2.5. Ancoragem molecular da Bg10

O modo de ligação dos ligantes glicose (BGC), celobiose (CBI), celotriose (CTR), celotetraose (CTT) e *p*NP-G (*p*-nitrofenil β -D-glicopiranosídeo) com o sítio ativo da Bg10 foi avaliado por meio de simulação de ancoragem molecular. A estrutura da Bg10, denominada receptor, foi preparada removendo as águas e adicionando átomos de hidrogênios polares, e em seguida, estabelecidas as coordenadas da grade de busca usando o AutoDock Tools 1.5.6 (Morris et al., 2010). O tamanho da grade foi estabelecido para $x=20$, $y=20$ e $z=20$ com espaçamento de 1Å, o centro da grade nas dimensões $x=39,760$, $y=14,426$, $z=27,444$ e o parâmetro de exaustividade igual a 10.

A ancoragem molecular foi realizada com o AutoDock Vina 1.1.2 (Trott e Olson, 2010) utilizando o receptor e cada um dos ligantes juntamente com os parâmetros da grade definidos anteriormente, considerando o receptor como rígido e os ligantes como flexíveis. O melhor modo de ligação para cada ligante foi considerado como

sendo o modo com menor energia livre de ligação, o qual foi alinhado com a estrutura do receptor e escolhido para a análise seguinte. A visualização da estrutura proteína-ligante e a preparação de figuras foram realizadas com o software PyMol (The PyMOL Molecular Graphics System, Version 2.3.0 Schrödinger, LLC.).

As interações entre proteína-ligante foram avaliadas por meio da ferramenta LigPlot+ (Laskowski e Swindells, 2011) utilizando o complexo proteína-ligante (BGC, CBI, CTR, CTT e *p*NP-G) a fim de conhecer os resíduos envolvidos e o tipo de interação estabelecida com os ligantes.

3. RESULTADOS E DISCUSSÃO

3.1. Expressão e Purificação da proteína recombinante

As frações eluídas da coluna durante a purificação por SEC, referentes a principal fração eluente visualizada no cromatograma, foram analisadas por eletroforese em gel SDS-Page (Davis, 1964) (Figura 2).

O rendimento total dessa enzima após a purificação por cromatografia por exclusão de tamanho foi de aproximadamente 33 mg/mL^{-1} por litro de cultura. Quando necessário, as frações eluídas foram reunidas e submetidas à concentração até ser possível recuperar a proteína alvo em uma concentração $\geq 10 \text{ mg/mL}^{-1}$.

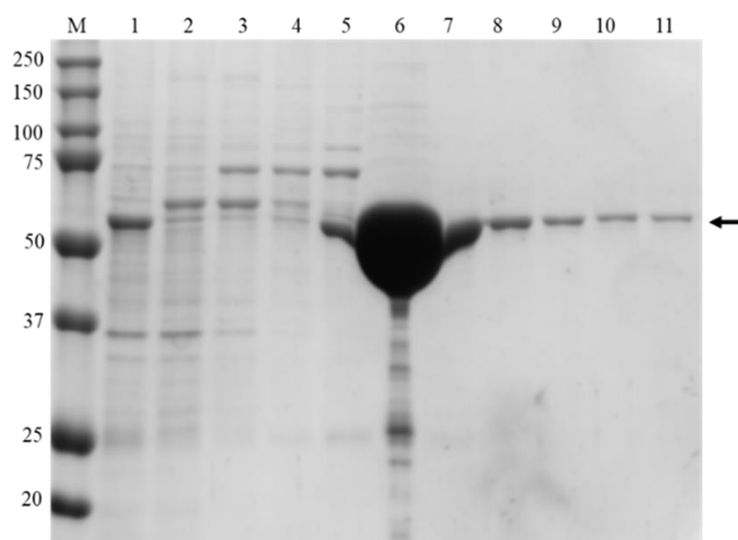


Figura 2. Eletroforese por SDS-Page após purificação da Bg10 por SEC. M: Marcador de massa molecular (“Precision Plus Protein Dual Color Standards” – BioRad,

Hercules, Califórnia, EUA); Linhas 1-11: frações eluídas. A seta indica a Bg10 com massa molecular estimada em aproximadamente 58,33 kDa.

3.2. Cristalização e otimização dos cristais

Os cristais obtidos a partir das placas preparadas pelo robô de pipetagem Mosquito LCP (SPT Labtech) foram formados em uma condição considerada ideal contendo 12% (v/v) PEG 3550, 250 mM citrato de potássio e 10% (v/v) de 1,5-pentanediol 40% (v/v) a uma temperatura constante de 20 °C (Figura 3). Os dados de maior difração para os cristais Bg10 foram a uma resolução de 2.3 Å.

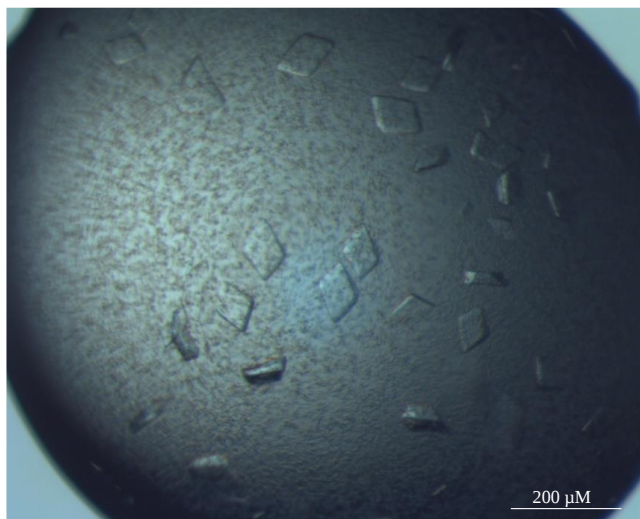


Figura 3. Cristais obtidos para Bg10 sob a condição contendo 12% (v/v) PEG 3550, 250 mM citrato de potássio e 10% (v/v) de 1,5-pentanediol 40% (v/v) crescido sob temperatura constante de 20 °C.

3.3. Estrutura geral da Bg10

Apesar do grande número de sequências de GH1 serem identificadas através da inferência de homologia por similaridade de sequências em bancos genômicos e metagenômicos, como por exemplo NCBI (“National Center for Biotechnology Information”), CAZy (“Carbohydrate-Active Enzymes Database”) e (“SwissProt/Uniprot”), poucas enzimas pertencentes à família GH1 foram efetivamente estudadas do ponto de vista estrutural e um número ainda menor de β -glicosidasas (EC 3.2.1.21) desta família possuem estrutura tridimensional depositada no PDB.

No presente trabalho, uma β -glicosidase metagenômica (Bg10) pertencente à família GH1 foi previamente caracterizada bioquimicamente e mostrou-se estimulada e tolerante a glicose, bem como apresentou outras características interessantes para uso industrial como a tolerância aos álcoois etanol e metanol além de ampla faixa de temperatura de atuação (Gomes-Pepe et al., 2016). Tais características tornam esta enzima uma boa candidata ao processo de produção de etanol 2G ou mesmo ao processamento de bebidas alcólicas como a maturação de vinhos para acelerar o envelhecimento através da quebra de compostos glicosídicos precursores de sabor (Gomes-Pepe et al., 2016).

A estrutura da Bg10 a uma resolução de 2.3 Å foi determinada através de substituição molecular, usando como molde uma β -glicosidase de *Streptomyces* sp. (69% de similaridade de sequência, PDB 1gnx). A Tabela 1 mostra a estatística e processamento de dados da Bg10. O modelo foi refinado até os valores finais de R_{work} 0.232 e R_{free} 0.264, além disso, o diagrama de Ramachandran mostrou que 95,42% dos resíduos da Bg10 se situaram em regiões favoráveis.

Tabela 1. Coleta de dados, processamento e refinamento da estrutura da Bg10.

Coleta de dados	
Linha de Luz	104
Detector	Eiger2 XE 16M
Comprimento de onda (Å)	0.97950
Grupo espacial	C2
<i>Parâmetros da malha unitária</i>	
a, b, c (Å)	92,43; 138,48; 114,84
β (°)	102,47
Processamento dos dados (AutoPROC/STARANISO)	
Limites de resolução do elipsoide montado na superfície de corte da resolução (Å)	3,43; 2,60; 2,29
Limites esféricos de resolução (Å)	75.6 – 2.31 (2.60 – 2.31)
Nº. total de observações	114356 (4334)
Nº total de reflexões únicas	33702 (1686)
Multiplicidade	3,4 (2,6)
Completude esférica (%)	54,4 (8,9)
Completude elipsoidal (%)	88,9 (59,3)
R-merge ^a	0,193 (0,643)
R-p.i.m. ^b	0,124 (0,486)
$\langle I/\sigma(I) \rangle$	5,6 (1,5)
CC (1/2)	0,982 (0,584)

Factor B de Wilson (Å ²)	55,9
Estatísticas de refinamento	
Faixa de resolução (Å)	112,13 – 2,30 (2,37 – 2,30)
R _{work} (%) ^c	23,2 (45,2)
R _{free} (%)	26,4 (50,1)
<i>Composição e completude do modelo</i>	
Regiões omitidas	1-16; 485-493 em ambas as cadeias A e B
Nº de átomos de proteína não-hidrogênios	7099
Moléculas de solvente	21
<i>Valores médios do fator B (Å²)</i>	
Proteína	29,1
Moléculas de solvente	22,7
<i>Desvios da geometria do modelo em relação aos valores ideais</i>	
RMS ligações (Å)	0,002
RMS ângulos (°)	0,537
Centros quirais (Å ³)	0,040
Grupos planares (Å)	0,003
<i>Validação do modelo ^d</i>	
Ramachandran favorecidos (%)	95,42
Ramachandran outliers (%)	0
Rotâmeros outliers (%)	0,69
C ^β outliers (%)	0
Clash score	3,10

Os valores entre parênteses são para a resolução “Outershell”. ^a: R-merge = $(\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|) / (\sum_{hkl} \sum_i I_i(hkl)) \times 100 \%$; ^b: R-p.i.m. = $\sum_{hkl} [1/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / (\sum_{hkl} \sum_i I_i(hkl)) \times 100 \%$ (Diederichs e Karplus, 1997) $\times 100 \%$; ^c: R_{work} = $(\sum_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)||) / (\sum_{hkl} |F_{obs}(hkl)|) \times 100 \%$; ^d: Calculado usando Molprobity.

Além de apresentar forma homodimérica, a estrutura cristalográfica da Bg10 possui um enovelamento clássico do tipo TIM barril (β/α)₈, comum à outras β-glicosidases da família GH1 (Figura 4).

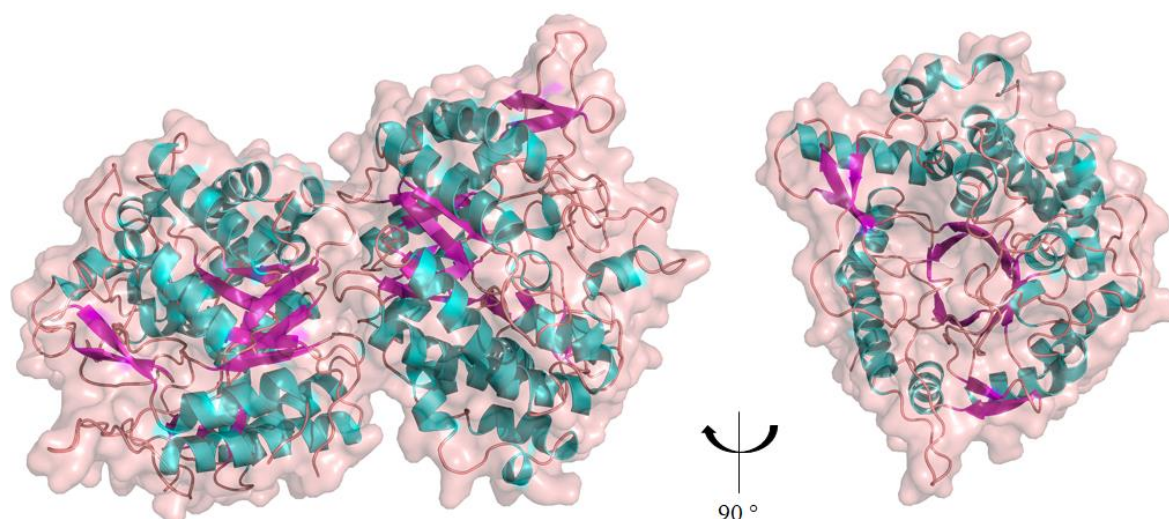


Figura 4. Estrutura geral homodimérica da Bg10. A: A estrutura é representada na forma de cartoon (com folhas beta coloridas na cor magenta e alpha-hélices coloridas na cor ciano) e superfície (colorida em rosa claro).

A fim de avaliar similaridades entre as enzimas, foi realizado um alinhamento de estruturas e um alinhamento múltiplo de sequências entre a Bg10 e β -glicosidases homólogas com diferentes níveis de tolerância e estimulação pela glicose (Tabela 2). A partir desta análise foi possível observar a posição dos resíduos catalíticos e as regiões conservadas destas enzimas (Figuras 5 e 6).

Tabela 2. Comparação entre as características da Bg10 com outras β -glicosidases GH1 e seus diferentes níveis de tolerância/estimulação pela glicose.

Enzima (PDB)	Tolerância à glicose (mM)	Estimulação por glicose (mM)	Cobertura (%)	Identidade (%)	Referências
3ahy	50	0	91	38	(JENG <i>et al.</i> , 2011)
1gnx	65	nd	95	69	(GUASCH <i>et al.</i> , 1999; PEREZ-PONS <i>et al.</i> , 1994)
4mdo	450	50	92	37	(DE GIUSEPPE <i>et al.</i> , 2014; SOUZA <i>et al.</i> , 2013)
Bg10	500	150	Este trabalho		(GOMES-PEPE <i>et al.</i> , 2016b)
5ogz	750	nd		44	(SHARMA <i>et al.</i> , 2019)
3vik	1000	200	96	35	(JENG <i>et al.</i> , 2012; UCHIMA <i>et al.</i> , 2011)
5xgz	nd	50	96	47	(MATSUZAWA; WATANABE; YAOI, 2017)

nd: não disponível

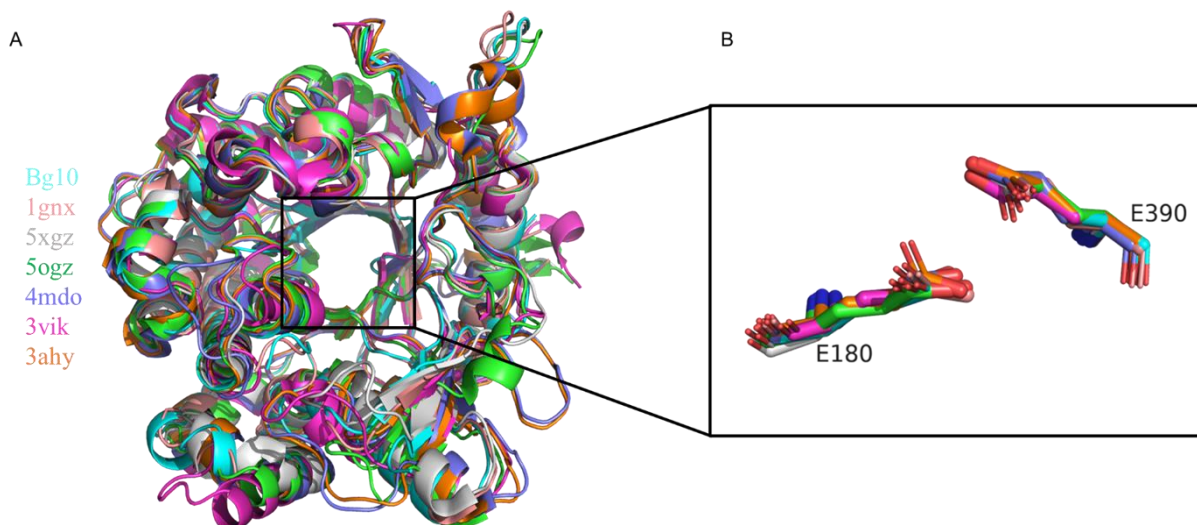


Figura 5. Propriedades estruturais e de tolerância/estímulo da Bg10 com outras β -glicosidases GH1. A: Comparação estrutural entre Bg10 e outras β -glicosidases GH1 com diferentes níveis de tolerância/estimulação pela glicose (PDB IDs 1gnx, 5xgz, 5ogz, 4mdo, 3vik e 3ahy); B: Superposição dos resíduos do sítio ativo entre a Bg10 e outras β -glicosidases GH1.

Embora seja possível observar diferenças estruturais discretas em alguns resíduos, o alinhamento entre as estruturas e a Bg10 evidenciou semelhança estrutural e posicionamento dos resíduos catalíticos mantendo o padrão do enovelamento de β -glicosidases GH1 (RMSD de 0,730 Å para 3ahy; 0,401 Å para 1gnx; 0,895 Å para 4mdo; 0,591 Å para 5ogz; 0,927 Å para 3vik e 0,711 Å para 5xgz) (Figura 5.A e 5.B). Ainda, é possível observar que as estruturas mais similares à Bg10, com RMSD de 0,401 Å (PDB 1gnx) e 0,591 Å (PDB 5ogz), apresentam diferentes níveis de tolerância à glicose (65 e 750 mM, respectivamente) (Tabela 2), assim, podemos inferir que o enovelamento proteico pode não estar diretamente relacionado aos mecanismos de tolerância e que, possivelmente, a chave para compreendê-los esteja envolvido com alguns resíduos chave (Lee et al., 2012; De Giuseppe et al., 2014; Mariano et al., 2017; Santos et al., 2019).

3.4. Sítio ativo da Bg10

O canal de ligação ao substrato é composto por diversos resíduos que formam tanto a porção glicona quanto aglicona, sendo que os resíduos que estão posicionados no subsítio -1 (glicona), responsável pela ligação da extremidade não-redutora do substrato, são extremamente conservados (Figura 6). Os resíduos que

formam a porção glicona são Q34, H135, W136, N179, E180, N237, N310, Y312, E390, W437, E444, W445 e F453, enquanto que a porção aglicona, esta por sua vez responsável pela ligação da extremidade redutora do substrato, é formada pelos resíduos W182, C183, L187, H194, H239, T313 E T315 (Figura 6).

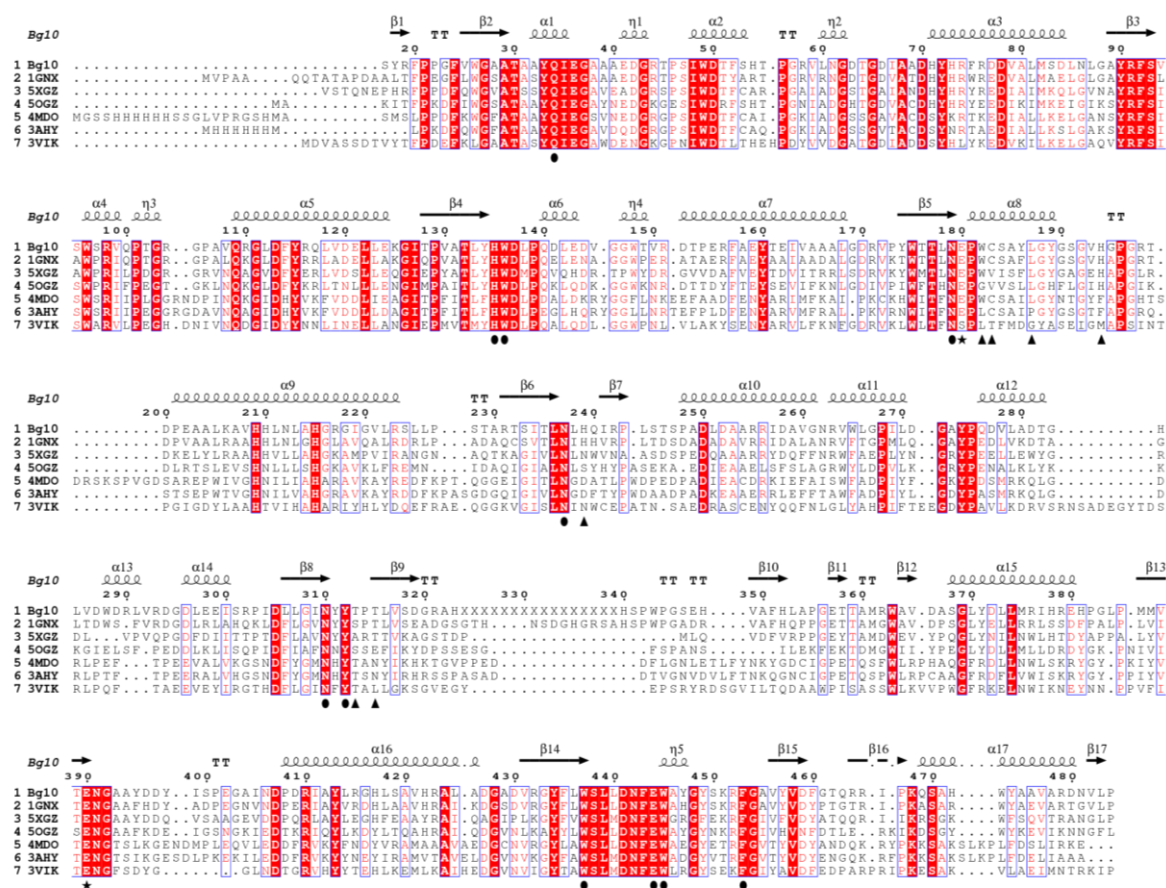


Figura 6. Alinhamento de sequências de aminoácidos da Bg10 com outras β-glicosidases da família GH1. Os números de acesso correspondem a β-glicosidase de *Streptomyces* sp. (1gnx), microrganismo não cultivável (5xgz), *Ruminiclostridium thermocellum* (5ogz), *Humicola insolens* (4mdo), *Trichoderma reesei* (3ahy) e *Neotermes koshunensis* (3vik). Caixas com fundo vermelho e letras brancas: resíduos estritamente conservados; letras vermelhas dentro de caixas azuis: regiões de sequência altamente similares. Os resíduos catalíticos de ácido glutâmico são indicados por uma estrela. •: Resíduos que formam a porção glicona; ▲: Resíduos que formam a porção aglicona. Elementos de estrutura secundária obtidos a partir da entrada Bg10 apresentados no topo (α-hélices como espirais, folhas β como setas e voltas como letras TT). O alinhamento foi realizado através do ClustalW (Thompson et al., 1994) e a figura elaborada utilizando o programa ESPript (Robert e Gouet, 2014).

Conforme esperado para o bolsão catalítico, os motivos N-E-P (resíduos 179 - 181) e V/I-T-E-N-G (388-392) (Figura 6) são regiões conservadas em β -glicosidases da família GH1, sendo que os resíduos de ácido glutâmico presentes nestes motivos conservados são responsáveis pela catálise enzimática atuando como resíduo ácido/base geral e nucleófilo catalítico, respectivamente.

Além disso, os resíduos de ácido glutâmico correspondentes na enzima Bg10 são encontrados nas posições E180 e E390 (Figura 6) e estão posicionados de maneira oposta no fundo da estrutura supersecundária formada pelo barril $(\beta/\alpha)_8$ e são separados por aproximadamente 4.4 Å (Figura 7).

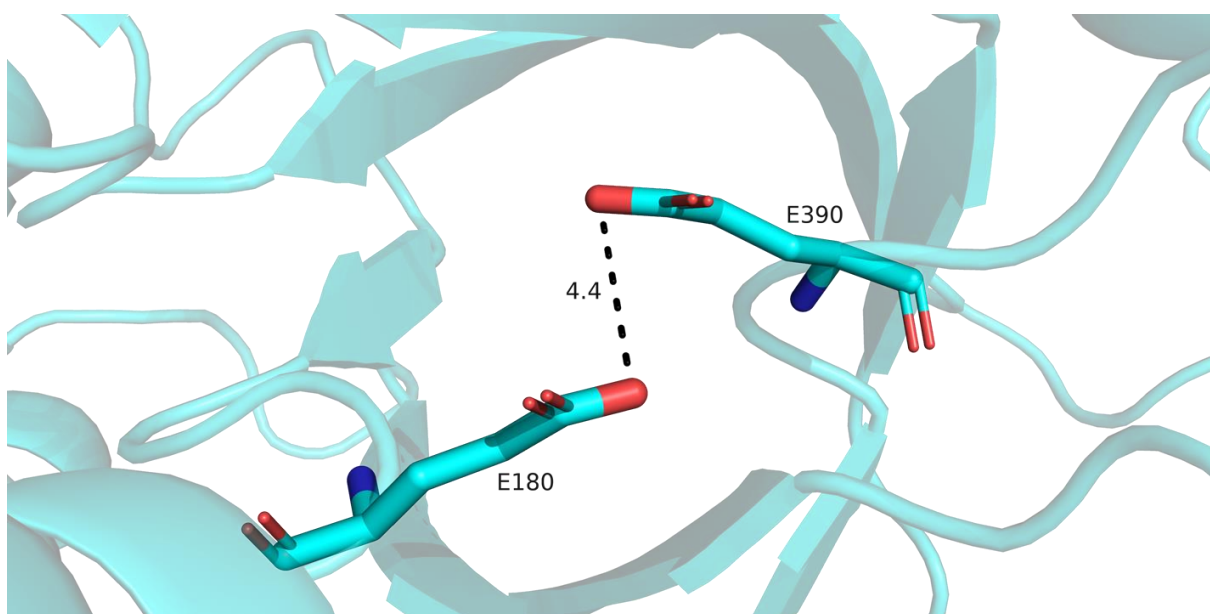


Figura 7. Distância em angstroms entre os resíduos do sítio ativo E180 e E390 da Bg10.

3.5. Bases moleculares da tolerância/estimulação pelo produto

Alguns estudos revelaram em detalhes o mecanismo de reação, cavidade e sítio catalítico presentes em algumas GH1 e GH3 permitindo compreender aspectos importantes extrapoláveis para cada família, como a presença de regiões de afinidade específica para regiões agliconas e gliconas; mecanismos de retenção do carbono anomérico, funções dos resíduos do canal catalítico, além de diferenças nos resíduos catalíticos de ação ácido/base e nucleofílico respectivo de cada família (Jeng et al., 2012; De Giuseppe et al., 2014; Yang et al., 2015; Matsuzawa et al., 2016). No

entanto, alguns aspectos peculiares presentes em algumas β -glicosidases resistentes e estimuladas pelo produto têm atraído notória atenção dos pesquisadores, uma vez que a base molecular relacionada à tais características ainda não é bem compreendida (Yang et al., 2015).

Todavia, análises estruturais revelam que o canal catalítico de β -glicosidases GH1 é mais profundo e mais estreito que o canal de β -glicosidases GH3, fazendo com que o acesso da glicose ao canal ativo seja restrito, o que suscita a hipótese que a limitação do acesso seja uma das características que contribuem para tolerância à glicose (De Giuseppe et al., 2014).

A estrutura da Bg10 exibiu um canal com 17.83 Å de profundidade (profundidade do resíduo nucleófilo) e 592 Å³ de volume (Figura 8), corroborando com os resultados encontrados por De Giuseppe et al. (2014), onde a β -glicosidase GH1 de *Humicola insolens* (HiBG, PDB 4mdo) tolerante (450 mM) e estimulada pela glicose (50 mM), características estas que são muito similares à da Bg10 (350 mM de tolerância e 100 mM de estimulação pela glicose), apresentou o canal catalítico com 17 Å de profundidade e 381 Å³ de volume, enquanto que uma GH3 de *Aspergillus aculeatus* (PDB 4iig, intolerante à glicose) apresentou um canal de 11,3 Å de profundidade e 297 Å³ de volume. Estes autores, ao avaliarem diferenças entre a profundidade do canal catalítico entre β -glicosidases GH1 com diferentes níveis de tolerância/estimulação pela glicose e uma GH3 não tolerante, evidenciaram que a largura do canal pode favorecer a entrada da glicose ao sítio ativo, afetando a atividade enzimática por inibição competitiva. Em outras palavras, pode-se dizer que, em canais mais largos e rasos, como os canais de β -glicosidases GH3, o acesso da glicose ao sítio ativo é facilitado.

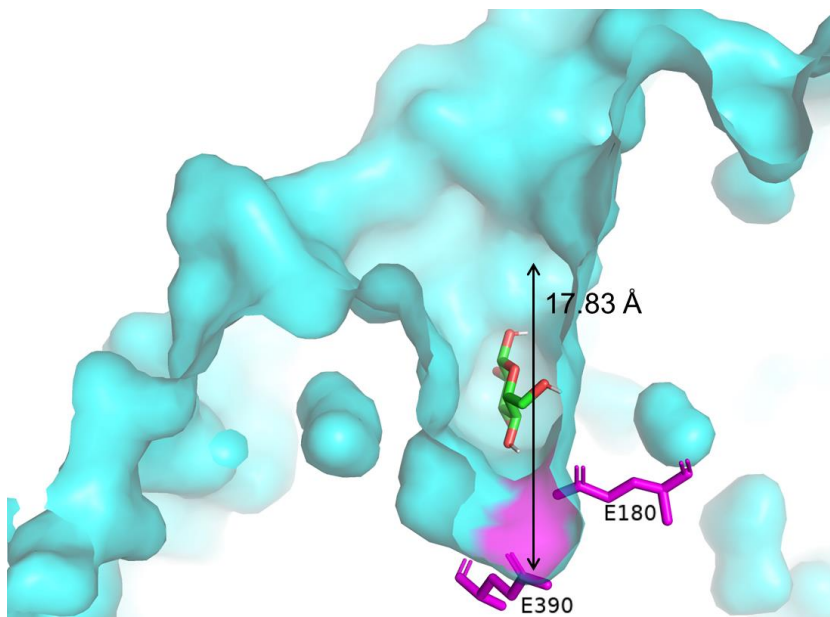


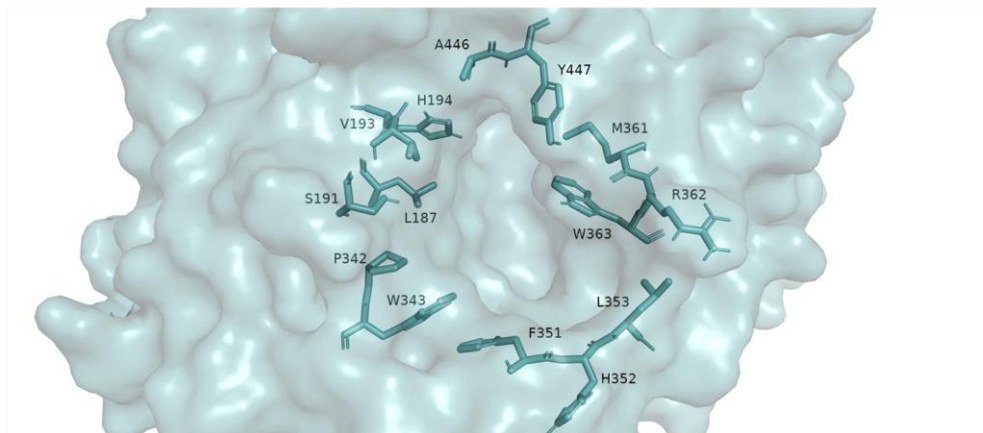
Figura 8. Vista lateral do canal ativo da Bg10 após simulação de ancoragem molecular com a glicose exibindo a profundidade do resíduo nucleófilo em angstroms. Resíduos catalíticos coloridos em “pink” e átomos de carbono da glicose coloridos em verde.

Do mesmo modo, outros fatores também devem agregar sentido a esta teoria para que a mesma seja capaz de alicerçar as bases teóricas para a tolerância e estimulação ao produto em β -glicosidases, sendo que a literatura vigente sugere fatores que vão desde aspectos relacionados à influência de resíduos específicos em proximidade ao canal catalítico (Liu et al., 2011; De Giuseppe et al., 2014; Cao et al., 2015; Guo et al., 2016), à ligação da glicose a locais secundários com diferentes afinidades e à geometria de certos locais de ligação que favorecem a clivagem do substrato através da transglicosilação (Yang et al., 2015), bem como fatores relacionados ao papel conjunto de resíduos que determinam as condições eletrostáticas específicas de regiões enzimáticas estratégicas como os “gatekeepers”, indicando que uma entrada mais neutra favorece tanto a tolerância, como o estímulo (De Giuseppe et al., 2014).

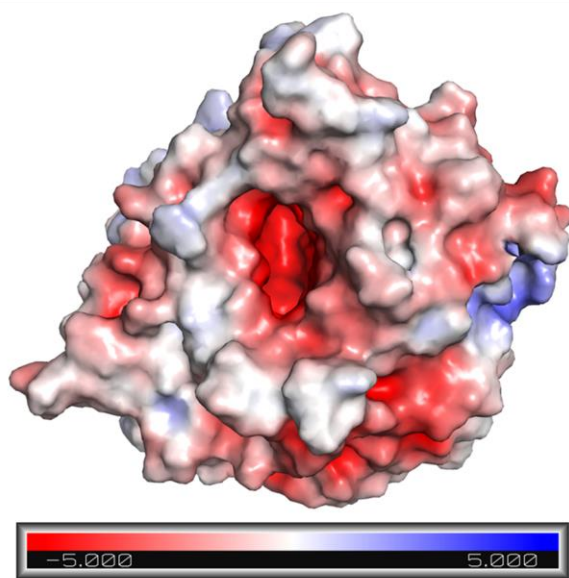
Tendo em vista estes fatores, investigamos quais destas características estavam presentes na estrutura da Bg10, visando entender o quanto estas teorias corroboravam com o fato de a enzima ser tolerante e estimulada pela glicose. Os resíduos L187, S191, V193, H194, P342, W343, F351, H352, L353, M361, R362, W363, A446 e Y447 compõem os “gatekeepers” na Bg10. Estes resíduos que regulam a entrada da glicose ao sítio ativo são principalmente polares (Figura 9.A) e fazem

parte do subsítio aglicona sendo geralmente mais neutros que o subsítio -1, região formada por resíduos negativamente carregados (Figura 9.B).

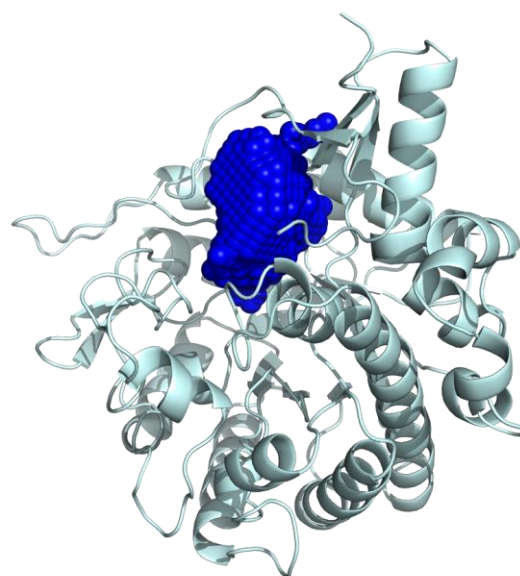
A



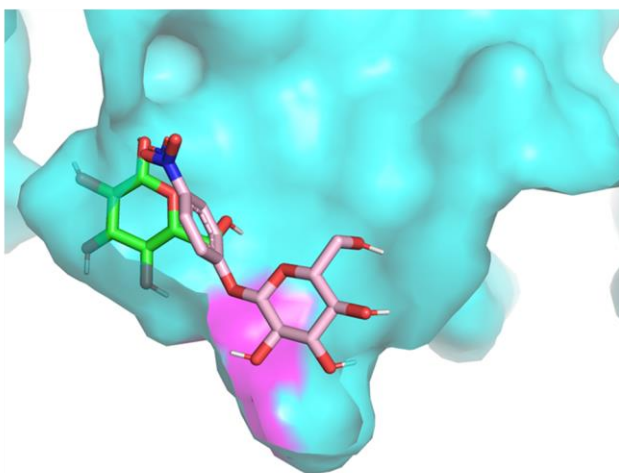
B



C



D



E

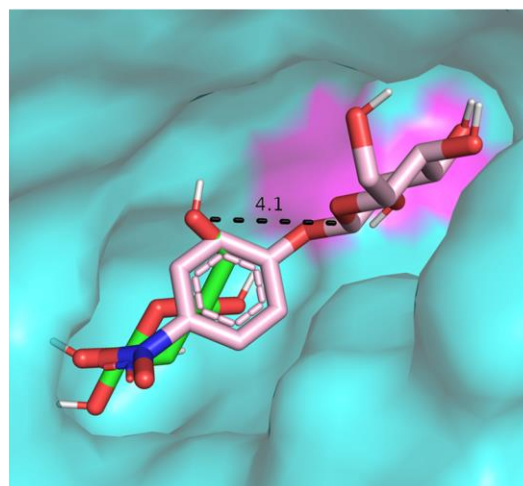


Figura 9. Canal catalítico da Bg10. A: Resíduos “gatekeepers” (resíduos que compõem a entrada do canal catalítico) preditos para Bg10. B: Superfície eletrostática da Bg10. As superfícies coloridas em vermelho representam cargas negativas, em branco cargas neutras e em azul cargas positivas. C: Canal catalítico predito com o servidor POCASA (Yu et al., 2009) (configurações padrão). D: Simulação de ancoragem molecular da glicose (com átomos de carbono coloridos em verde) e do substrato *p*NP-G (com átomos de carbono coloridos em rosa claro). E: Distância em angstroms entre o oxigênio 6-OH da glicose e o carbono da ligação glicosídica.

Além disso, evidências da literatura apontam para a influência dos resíduos W168, C170 e L173 (correspondentes à W182, C183 e L187 na Bg10), encontrados em diversas β -glicosidases tolerantes (Figura 6), que podem auxiliar os “gatekeepers” estreitando o canal que dá acesso ao sítio ativo e regulando a entrada e a orientação da glicose (Lee et al., 2012; De Giuseppe et al., 2014; Mariano et al., 2017; Santos et al., 2019). Santos e colaboradores (2019) realizaram mutagênese sítio dirigida de L167W/P172L da β -glicosidase de *Trichoderma harzianum* (ThBGI) (PDB 5bwf) (Santos et al., 2016) baseada na estrutura da β -glicosidase altamente tolerante e levemente estimulada (HiBg, PDB 4mdo (De Giuseppe et al., 2014)). Esta mutação dupla resultou em um aumento significativo na tolerância da ThBGI (até 1 M de glicose). A Bg10, por sua vez, já apresenta os resíduos de triptofano e leucina nestas posições (W182 e L187), reforçando o pressuposto sobre os resíduos “gatekeepers” neutros auxiliados por estes resíduos no mecanismo de tolerância à glicose.

Por outro lado, Yang et al. (2015) sugerem ainda que a presença do resíduo T228 (posição ocupada pelo resíduo H239 na Bg10) é um dos responsáveis pelo aumento da tolerância à glicose, no entanto, este resíduo está presente também na glicosidase 1gnx que é levemente tolerante (até 50 mM), ao contrário da Bg10 altamente tolerante e levemente estimulada. Além disso, esse resíduo não se apresentou conservado entre as β -glicosidases com diferentes níveis de tolerância e estímulo aqui analisadas (Tabela 2), sendo substituído por asparagina em 5xgz e 3vik, serina em 5ogz e aspartato em 4mdo e 3ahy (Figura 6). De fato, His239 realizou contatos do tipo hidrofóbicos e ligações de hidrogênio em todos os ligantes analisados (Figura 10, Tabela 3), o que indica sua importância durante a hidrólise. Este cenário reforça que somente a presença de determinados resíduos não é suficiente para suportar o mecanismo de tolerância destas enzimas. No entanto, ensaios de atividade

com mutantes para o resíduo H239 na Bg10 podem ajudar a compreender melhor sua função.

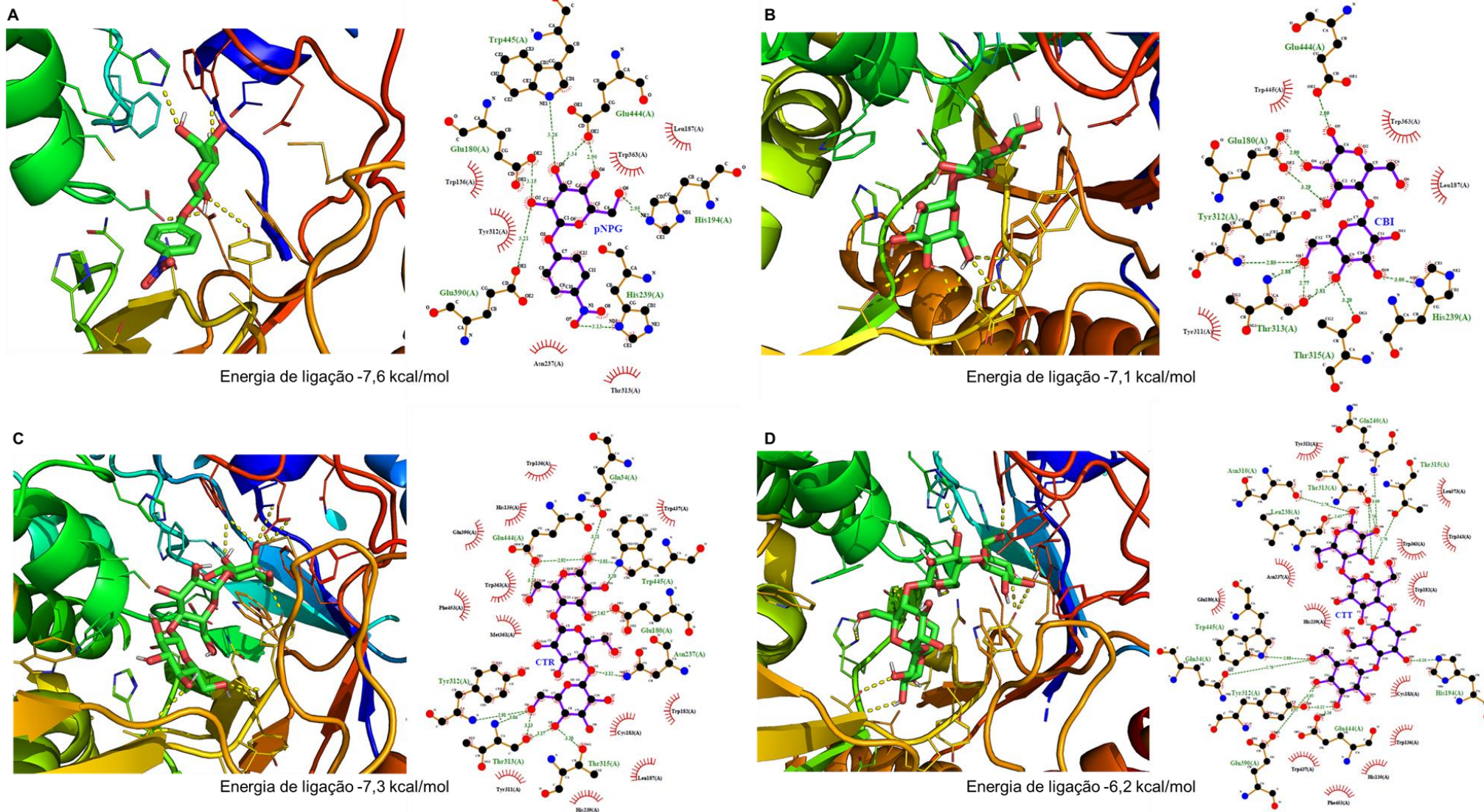


Figura 10. Diagrama de potenciais interações entre a enzima Bg10 e seus ligantes. A: pNPG; B: celobiose; C: celotriose; D: celotetraose. Linhas tracejadas verdes indicam ligações de hidrogênio e seu comprimento em angstroms; ícones vermelhos indicam resíduos envolvidos em contatos hidrofóbicos; os conectores dos ligantes estão coloridos em lilás e dos resíduos de aminoácidos em dourado. Figuras preparadas pelo Pymol e LigPlot+ v2.2 (Laskowski e Swindells, 2011).

3.6. Simulações de ancoragem molecular

Sabe-se que, não apenas a acessibilidade do canal influencia para a ligação da glicose, mas também sua preferência/afinidade são cruciais e podem variar de proteína para proteína dependendo das características de cada uma delas (Yang et al., 2015). Sendo assim, realizamos um estudo de ancoragem molecular com os ligantes *p*NP-G, glicose (BGC), celobiose (CBI), celotriose (CTR) e celotetraose (CTT) para a estrutura Bg10 sob o intuito de prever as possíveis interações entre proteína-ligante. O melhor modo de ligação escolhido para *p*NP-G, glicose, celobiose, celotriose e celotetraose foi aquele que apresentou a menor energia livre de ligação, sendo -7,6; -6; -7,1; -7,3 e -6,2 kcal/mol, respectivamente (Tabela 3).

Tabela 3. Resíduos da Bg10 envolvidos em contatos com os ligantes analisados.

Ligante	Energia de ligação (kcal/mol)	Aminoácidos envolvidos em ligações de hidrogênio	Aminoácidos envolvidos em interações hidrofóbicas
<i>p</i> NP-G	-7,6	Glu444, His194, His239, Glu390, Glu180, Trp445	Leu187, Trp363, Thr313, Asn237, Tyr312, Trp136
BGC	-6	Gln240, Thr315, His239, Asn237, Asn310, Leu238, Thr313	Tyr311, Tyr312
CBI	-7,1	Glu444, His239, Thr315, Thr313, Tyr312, Glu180	Trp363, Leu187, Tyr311, Trp445
CTR	-7,3	Gln34, Trp445, Glu180, Asn237, Thr315, Thr313, Tyr312, Glu444	Trp437, Trp182, Cys183, Leu187, His239, Tyr311, Met361, Phe453, Trp363, Glu390, His135, Trp136
CTT	-6,2	Gln240, Thr315, His194, Glu444, Glu390, Tyr312, Gln34, Trp445, Leu238, Asn310, Thr313	Leu373, Trp343, Trp363, Trp182, Cys183, Trp136, His135, Phe453, Trp437, Glu180, His239, Asn237, Tyr311

Com relação à glicose, foi possível observar que ela apresentou maior afinidade ao lado oposto ao sítio ativo, em uma região localizada no meio do canal catalítico, posicionada de forma adjacente a ligação glicosídica do *p*NP-G, onde a geometria entre estes ligantes apresentou uma distância de 4.1 Å (Figuras 9.D e 9.E). Em um estudo realizado por Yang et al. (2015), a geometria entre os mesmos ligantes foi muito semelhante (4.3 Å) para uma β -glicosidase de origem metagenômica tolerante e estimulada pela glicose (Bg11A), indicando que esta geometria não afeta a atividade

enzimática via inibição competitiva pelo local de ligação, mas pode favorecer a ocorrência de reações de transglicosilação ao invés de hidrólise (Yang et al., 2015).

A energia de ligação com celotriose foi de -7,3 e com celobiose -7,1; mas menor com celotetraose (-6,2 kcal/mol) indicando que o sítio ativo da Bg10 acomoda melhor substratos com até três unidade de glicose (Figura 10). A energia de ligação do *p*NP-G foi mais forte (-7,6 kcal/mol) que em celobiose, corroborando com estudos cinéticos realizados anteriormente, onde a atividade da Bg10 contra o substrato *p*-nitrofenil β -D-celobiosídeo (*p*NP-C) foi de somente 20% quando comparada com o substrato *p*NP-G (100%) (Gomes-Pepe et al., 2016). Os aminoácidos e os tipos de interações com os ligantes estão descritos na Tabela 3.

4. CONCLUSÕES

No presente estudo descrevemos a estrutura cristalográfica de uma β -glicosidase de origem metagenômica tolerante e estimulada pela glicose. A Bg10 apresentou um enovelamento clássico do tipo TIM barril $(\beta/\alpha)_8$, comum a outras β -glicosidases da família GH1. Os resíduos de ácido glutâmico E180 e E380 que atuam como ácido/base e nucleófilo, respectivamente, estão posicionados de maneira oposta um ao outro a uma distância de 4.4 Å no fundo da estrutura supersecundária formada pelo barril $(\beta/\alpha)_8$. Os estudos estruturais e as simulações de ancoragem molecular permitiram identificar o local de preferência de ligação da glicose localizado no lado oposto ao sítio ativo e no meio do canal catalítico. Além disso, foi possível identificar resíduos conservados e os resíduos envolvidos em ligações de hidrogênio e interações hidrofóbicas com os ligantes. Ademais, a estrutura da Bg10 revelou a presença de resíduos “gatekeepers” neutros, uma entrada do canal ativo mais neutra que o restante da estrutura; e um canal catalítico profundo e estreito. Todas estas características estão diretamente relacionadas à tolerância e ao estímulo pela glicose nesta enzima, corroborando com elementos anteriormente pressupostos pela literatura, bem como indicado pelos dados experimentais previamente obtidos para Bg10. Estes resultados permitiram compreender alguns de seus principais mecanismos de tolerância e estímulo que, aliados a novos estudos empregando técnicas de mutagênese sítio dirigida, poderão contribuir consideravelmente para o

aumento da atividade catalítica e da tolerância à glicose da Bg10 e de outras β -glicosidases GH1, úteis no processo de sacarificação da biomassa lignocelulósica, dentre outras aplicações biotecnológicas.

5. REFERÊNCIAS

- Adams, P.D., Afonine, P. V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., et al. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. **Acta Crystallographica Section D: Biological Crystallography** 66: 213–221.
- Cao, L.C., Wang, Z.J., Ren, G.H., Kong, W., Li, L., Xie, W. e Liu, Y.H. (2015). Engineering a novel glucose-tolerant β -glucosidase as supplementation to enhance the hydrolysis of sugarcane bagasse at high glucose concentration. **Biotechnology for Biofuels** 8: 1–12.
- Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. e Richardson, D.C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D: Biological Crystallography** 66: 12–21.
- Davis, B.J. (1964). Davis : Disc Electrophoresis. **Clinical Applications** 121: 404–427.
- De Giuseppe, P.O., Souza, T.D.A.C.B., Souza, F.H.M., Zanphorlin, L.M., Machado, C.B., Ward, R.J., Jorge, J.A., Furriel, R.D.P.M. e Murakami, M.T. (2014). Structural basis for glucose tolerance in GH1 β -glucosidases. **Acta Crystallographica Section D: Biological Crystallography** 70: 1631–1639.
- Diederichs, K., Karplus, P. (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. **Nature Structural Biology** 4: 269–275.
- Emsley, P., Lohkamp, B., Scott, W.G. e Cowtan, K. (2010). Features and development of Coot. **Acta Crystallographica Section D: Biological Crystallography** 66: 486–501.
- Gomes-Pepe, E.S., Sierra, E.G.M., Pereira, M.R., Castellane, T.C.L. e De Lemos, E.G.M. (2016). Bg10: A novel metagenomics alcohol-tolerant and glucose-stimulated gh1 β -glucosidase suitable for lactose-free milk preparation. **PLoS ONE** 11: 1–25.
- Gomes, E.S., Schuch, V. e Lemos, E.G. de M. (2013). Biotechnology of polyketides: New breath of life for the novel antibiotic genetic pathways discovery through metagenomics. **Brazilian Journal of Microbiology** 44: 1007–1034.

Guasch, A., Vallmitjana, M., Pérez, R., Querol, E., Pérez-Pons, J.A. e Coll, M. (1999). Cloning, overexpression, crystallization and preliminary X-ray analysis of a family 1 β -glucosidase from *Streptomyces*. **Acta Crystallographica Section D: Biological Crystallography** 55: 679–682.

Guo, B., Amano, Y. e Nozaki, K. (2016). Improvements in glucose sensitivity and stability of *trichoderma reesei* β -glucosidase using site-directed mutagenesis. **PLoS ONE** 11: 1–12.

Jeng, W.Y., Wang, N.C., Lin, C.T., Chang, W.J., Liu, C.I. e Wang, A.H.J. (2012). High-resolution structures of *Neotermes koshunensis* β -glucosidase mutants provide insights into the catalytic mechanism and the synthesis of glucoconjugates. **Acta Crystallographica Section D: Biological Crystallography** 68: 829–838.

Kabsch, W., T., B.A., K., D., A., K.P., K., D., S., M., G., R.R.B., P., E., et al. (2010). XDS. **Acta Crystallographica Section D Biological Crystallography** 66: 125–132.

Laskowski, R.A. e Swindells, M.B. (2011). LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. **Journal of Chemical Information and Modeling** 51: 2778–2786.

Lee, H.L., Chang, C.K., Jeng, W.Y., Wang, A.H.J. e Liang, P.H. (2012). Mutations in the substrate entrance region of β -glucosidase from *Trichoderma reesei* improve enzyme activity and thermostability. **Protein Engineering, Design and Selection** 25: 733–740.

Liu, J., Zhang, X., Fang, Z., Fang, W., Peng, H. e Xiao, Y. (2011). The 184th residue of β -glucosidase Bgl1B plays an important role in glucose tolerance. **Journal of Bioscience and Bioengineering** 112: 447–450.

Mariano, D.C.B., Leite, C., Santos, L.H.S., Marins, L.F., Machado, K.S., Werhli, A. V., Lima, L.H.F. e De Melo-Minardi, R.C. (2017). Characterization of glucose-tolerant β -glucosidases used in biofuel production under the bioinformatics perspective: A systematic review. **Genetics and Molecular Research** 16: 1–19.

Matsuzawa, T., Jo, T., Uchiyama, T., Manninen, J.A., Arakawa, T., Miyazaki, K., Fushinobu, S. e Yaoi, K. (2016). Crystal structure and identification of a key amino acid for glucose tolerance, substrate specificity, and transglycosylation activity of metagenomic β -glucosidase Td2F2. **FEBS Journal** 283: 2340–2353.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. e Read, R.J. (2007). Phaser crystallographic software. **Journal of Applied Crystallography** 40: 658–674.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S. e Olson, A.J. (2010). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. **Journal of Computational Chemistry** 30: 2785–2791.

Robert, X. e Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. **Nucleic Acids Research** 42: 320–324.

Santos, C.A., Morais, M.A.B., Terrett, O.M., Lyczakowski, J.J., Zanphorlin, L.M., Ferreira-Filho, J.A., Tonoli, C.C.C., Murakami, M.T., et al. (2019). An engineered GH1 β -glucosidase displays enhanced glucose tolerance and increased sugar release from lignocellulosic materials. **Scientific Reports** 9: 1–10.

Santos, C.A., Zanphorlin, L.M., Crucello, A., Tonoli, C.C.C., Ruller, R., Horta, M.A.C., Murakami, M.T. e De Souza, A.P. (2016). Crystal structure and biochemical characterization of the recombinant ThBgl, a GH1 B-glucosidase overexpressed in *Trichoderma harzianum* under biomass degradation conditions. **Biotechnology for Biofuels** 9: 1–11.

Singh, G., Verma, A.K. e Kumar, V. (2016). Catalytic properties, functional attributes and industrial applications of β -glucosidases. **3 Biotech** 6: 1–14.

Srivastava, N., Rathour, R., Jha, S., Pandey, K., Srivastava, M., Thakur, V.K., Sengar, R.S., Gupta, V.K., et al. (2019). Microbial beta glucosidase enzymes: Recent advances in biomass conversion for biofuels application. **Biomolecules** 9: 1–23.

Thompson, J.D., Higgins, D.G. e Gibson, T.J. (1994). Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. **Nucleic Acids Research** 22: 4673–4680.

Trott, O. e Olson, A. (2010). Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. **Journal of Computational Chemistry** 31: 455–461.

Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T. e Bricogne, G. (2011). Data processing and analysis with the autoPROC toolbox. **Acta Crystallographica Section D: Biological Crystallography** 67: 293–302.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., et al. (2011). Overview of the CCP4 suite and current developments. **Acta Crystallographica Section D: Biological Crystallography** 67: 235–242.

Yang, Y., Zhang, X., Yin, Q., Fang, W., Fang, Z., Wang, X., Zhang, X. e Xiao, Y. (2015). A mechanism of glucose tolerance and stimulation of GH1 β -glucosidases. **Scientific Reports** Nature Publishing Group 5: 1–12.

Yu, J., Zhou, Y., Tanaka, I. e Yao, M. (2009). Roll: A new algorithm for the detection of protein pockets and cavities with a rolling probe sphere. **Bioinformatics** 1: 46–52.

Zhang, Y., Min, Z., Qin, Y., Ye, D.Q., Song, Y.Y. e Liu, Y.L. (2019). Efficient Display of *Aspergillus niger* β -Glucosidase on *Saccharomyces cerevisiae* Cell Wall for Aroma Enhancement in Wine. **Journal of Agricultural and Food Chemistry** 67: 5169–5176.

Zhao, L., Zhou, T., Li, X., Fan, S. e You, L. (2013). Expression and characterization of GH3 β -Glucosidase from *Aspergillus niger* NL-1 with high specific activity, glucose inhibition and solvent tolerance. **Microbiology (Russian Federation)** 82: 356–363.

CAPÍTULO 3 – Caracterização estrutural *in silico* e estudos de ancoragem molecular da β -xilosidase bacteriana BXil

RESUMO – As β -xilosidases são enzimas hidrolíticas que atuam sobre a cadeia principal da xilana em sinergia com endoxilanases, aumentando a eficiência da hidrólise da xilana e, conseqüentemente, elevando a produção de açúcares fermentescíveis. Em vista disso, as β -xilosidases são consideradas essenciais para o complexo xilanolítico e têm-se tornado alvo de pesquisas acadêmicas e industriais, principalmente para o desenvolvimento de bioprocessos e produtos. O principal objetivo deste trabalho foi caracterizar através de abordagens *in silico* uma β -xilosidase (BXil) anotada no genoma da *Chitinophaga* sp. CB10. Os estudos incluíram a caracterização das estruturas primária e secundária, modelagem e caracterização da terciária e estudos de simulação de ancoragem molecular. A partir dos resultados obtidos foi possível ter acesso a estrutura tridimensional da BXil que apresentou uma repetição de domínio hélice beta atribuídos como GH43, arquitetura esta que não possui caracterização funcional ou estrutural até o momento. Apesar de ambos domínios terem sido identificados como GH43, apenas o C-terminal apresenta a estrutura secundária hélice beta com disposição de 5 lâminas características de GH43 e, considerando a qualidade superior da modelagem para o domínio C-terminal somente este domínio foi utilizado para os estudos de ancoragem molecular. Com base nestas análises, foi possível prever a disposição de seu sítio ativo e os principais resíduos envolvidos em interações com diferentes substratos, indicando ainda que a enzima BXil é capaz de acomodar substratos de tamanho intermediário em seu canal catalítico. A função do domínio N-terminal ainda não pôde ser esclarecida, uma vez que este domínio não é similar a nenhum outro previamente determinado por experimentos cristalográficos, permanecendo de função desconhecida e carecendo de estudos adicionais de cristalografia e cinética enzimática dos domínios individuais e em conjunto para a completa caracterização funcional de BXil. Não obstante, a ausência de enzimas similares a BXil na literatura pode ser evidência de uma enzima de atuação distinta das demais enzimas conhecidas, o que pode ser valioso do ponto de vista biotecnológico.

Palavras-chave: Glicosil Hidrolase 43, hélice beta, interações proteína-ligante, repetição de domínios

1. INTRODUÇÃO

A estrutura heterogênea da xilana, o principal constituinte da hemicelulose, é formada por diferentes tipos de ligações e ramificações proporciona uma barreira física que limita a penetração das forças mecânicas e microbianas. Logo, a hidrólise completa da xilana requer enzimas que atuam sobre sua estrutura principal, como endo-1,4- β -D-xilanases (comumente conhecidas como xilanases, E.C. 3.2.1.8) e β -xilosidases (E.C. 3.2.1.37), além de enzimas auxiliares que atuam sinergicamente nos pontos de ramificação (Collins et al., 2005; Wang et al., 2014; Moreira e Filho, 2016).

As β -xilosidases, ou somente xilosidases, são distribuídas em diversas famílias de GHs e realizam a hidrólise de ligações β -1,4 de xilo-oligossacarídeos e xilobiase através do mecanismo de inversão do carbono anomérico liberando unidades de xilose (Knob et al., 2010; Juturu e Wu, 2012).

Algumas xilosidases desta família apresentam atividade bifuncional de arabinofuranosidase e β -xilosidase, principalmente em substratos de aril-glicosídeo (Falck et al., 2016), devido às semelhanças espaciais das ligações glicosídicas e grupos hidroxilas entre os resíduos de D-xilopirranose e L-arabinofuranose (Shallom e Shoham, 2003; Huy et al., 2013).

A ação sinérgica das β -xilosidases e endoxilanases aumenta a eficiência da hidrólise da xilana e, conseqüentemente, eleva a produção de açúcares fermentescíveis. Isto é possível graças à ação catalítica das β -xilosidases que remove os xilo-oligossacarídeos que inibem as endoxilanases e limitam a hidrólise da xilana (Polizeli et al., 2005; Knob et al., 2010; Canilha et al., 2012; Juturu e Wu, 2012). Em vista disso, as xilosidases são consideradas essenciais para o complexo xilanolítico e têm-se tornado alvo de pesquisas acadêmicas e industriais, principalmente para o desenvolvimento de bioprocessos e produtos.

A bactéria gram-negativa *Chitinophaga* sp. CB10 (Kishi et al., 2017), foi isolada pelo grupo de pesquisa LBMP de um consórcio microbiano proveniente de uma pilha de bagaço de cana-de-açúcar para produção de etanol e, conforme o esperado considerando o ambiente de isolamento, foi capaz de degradar carboximetilcelulose (CMC), motivando o grupo de pesquisa a sequenciar o material genético e a anotar os genes relacionados à degradação de materiais lignocelulósicos. Um gene referente

a ORF CB_612456 que codifica uma β -xilosidase pertencente à família GH43 foi anotada no genoma deste organismo e, após extensiva busca nos bancos de dados através da ferramenta Blast, observou-se que a arquitetura de domínios desta enzima não possui nenhum tipo de caracterização funcional ou estrutural até o momento.

Estudos realizados utilizando ferramentas de bioinformática, ou estudos *in silico*, são mais rápidos e de menor custo que métodos experimentais e permitem a interpretação prévia das inúmeras sequências enzimáticas disponíveis em bancos de dados fornecendo informações importantes como, por exemplo, a previsão de estruturas tridimensionais, comparação de sequências e estruturas, encaixe de ligantes e interações proteicas. Além de fornecer uma visão ampla sobre estas características, as análises *in silico* permitem manipular a estrutura tridimensional de acordo com propósitos específicos facilitando os ensaios *in vitro*.

Diante do exposto, o objetivo do presente estudo foi caracterizar através de análises *in silico* o gene codificador da β -xilosidase anotado no genoma da *Chitinophaga* sp. CB10, incluindo a modelagem por homologia de sua estrutura terciária e simulações de ancoragem molecular com diferentes ligantes, a fim de conhecer os parâmetros físicos gerais e características estruturais desta enzima, como a disposição do domínio catalítico e interações proteína-ligante.

2. MATERIAL E MÉTODOS

2.1. Origem da ORF CB_612456

A bactéria gram-negativa *Chitinophaga* sp. CB10 foi isolada de um consórcio microbiano proveniente de uma pilha de bagaço de cana-de-açúcar residual da produção de etanol, tendo seu genoma parcial depositado no NCBI GenBank sob o número de acesso MLAV00000000 (Kishi et al., 2017). Considerando o ambiente de isolamento, o referido consórcio microbiano foi capaz de degradar carboximetilcelulose (CMC) motivando a anotação de genes relacionados à degradação de materiais lignocelulósicos (Funnicelli, 2018). Neste sentido, foi realizada a prospecção *in silico* de genes relacionados a degradação de biomassa a partir do genoma da *Chitinophaga* sp. CB10, levando em consideração os motivos e

domínios conservados de Glicosil Hidrolases (GHs), em especial β -xilosidases.

2.2. Análise da estrutura primária

A identificação dos domínios conservados da β -xilosidase foi observado no Pfam (<https://pfam.xfam.org/>) (Finn et al., 2016) e os dados físico-químicos teóricos gerados com a ferramenta “ProtParam ExPASy” (<https://web.expasy.org/protparam/>) (Gasteiger et al., 2005). Ademais, com intuito de identificar possíveis peptídeos sinais, foi realizada uma análise com o servidor SignalP 5.0 (<http://www.cbs.dtu.dk/services/SignalP/>) (Almagro Armenteros et al., 2019).

Com intuito de avaliar a similaridade da β -xilosidase (BXil) com outras enzimas, foi realizado alinhamento múltiplo de sequências pelo ClustalW (Thompson et al., 1994) utilizando a ferramenta Mega7 (Kumar et al., 2016). As sequências utilizadas foram obtidas a partir da base de dados “nr” (não redundante) do NCBI (Apêndice A). Após o alinhamento, uma árvore fenética foi construída através do programa Mega7 (Kumar et al., 2016), usando o método de “Neighbor Joining” (Saitou e Nei, 1987) com “bootstrap” de 5000 réplicas.

2.3. Análise da estrutura secundária

A análise da estrutura secundária da BXil foi realizada através dos servidores PSIPRED 4.0 (<http://bioinf.cs.ucl.ac.uk/psipred/>) (Buchan e Jones, 2019) e SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) (Geourjon e Deléage, 1995), a fim de observar sua organização incluindo regiões desordenadas, número de hélices alfa, folhas beta e regiões de “coil”.

2.4. Modelagem por homologia e caracterização dos domínios

A modelagem por homologia teve como finalidade criar um modelo tridimensional da BXil, elaborado a partir de um modelo de referência com alta similaridade de sequência com a proteína alvo, sendo que quanto maior a identidade entre as proteínas, melhor é a qualidade do modelo gerado. A predição da estrutura

proteica para cada domínio foi obtida através do servidor Swiss Model (<https://swissmodel.expasy.org/>). O modelo com melhor QMEAN (Qualitative Model Energy ANalysis) (Benkert et al., 2008) para cada domínio foi refinado através do servidor “RefineD” (Bhattacharya e Valencia, 2019) com modo de refinamento conservador e, em seguida, a qualidade dessas estruturas foi avaliada através do gráfico de Ramachandran obtido no “web server” Rampage (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>).

As estruturas foram visualizadas e analisadas por meio do programa Pymol (The PyMOL Molecular Graphics System, Version 2.3.0 Schrödinger, LLC.).

Após a modelagem dos domínios, o potencial eletrostático da BXil foi calculado com o APBS (“Adaptive Poisson-Boltzmann Solver”) integrado ao programa PyMOL. Além disso, o Pymol também foi utilizado para identificar a hidrofobicidade da molécula, a qual foi colorida com base na escala de hidrofobicidade de Eisenberg.

Os modelos estruturais da BXil também foram avaliados no “web server” I-TASSER (“Iterative Threading ASSEmbly Refinement”) (Yang et al., 2015; Roy e Kucukural, 2010; Zhang, 2008) a fim de prever sua estrutura e suas funções, e no “web server” CASTp (“Computed Atlas of Surface Topography of proteins”) (Tian et al., 2018) a fim de localizar e medir as propriedades geométricas e topológicas de sua estrutura.

2.5. Estudos de ancoragem molecular

O modo de ligação dos substratos xilobiose (X2), xilotriose (X3), xilotetraose (X4) e xilopentose (X5) com o sítio ativo da BXil foi avaliado por meio de simulação de ancoragem molecular. Este método testa inúmeras combinações possíveis entre o ligante e o sítio ativo, e a conformação mais provável, ou seja, com menor energia livre de ligação, é utilizada para análise do complexo proteína-ligante.

A estrutura da BXil, denominada receptor, e dos substratos, denominados ligantes, foram preparadas usando a ferramenta AutoDock Tools 1.5.6 (Morris et al., 2010). O preparo do receptor consistiu na remoção das moléculas de água e adição de átomos de hidrogênio polares, seguido pela determinação das coordenadas da grade de busca. O tamanho da grade de busca foi estipulado para x=20, y=20 e z=20

com espaçamento de 1 Å, o centro da grade nas dimensões x= 32,9, y= -20,538, z=40,744 e o parâmetro de exaustividade igual a 10. O preparo dos ligantes iniciou-se pela determinação automática do número de torsões para cada uma das estruturas, fusão dos hidrogênios polares, cálculo das cargas Gasteiger, finalizando com a determinação do ponto de torsão.

A ancoragem molecular foi realizada com o AutoDock Vina 1.1.2 (Trott e Olson, 2010) utilizando o receptor com cada um dos ligantes juntamente com os parâmetros da grade definidos anteriormente, considerando o receptor como rígido e os ligantes como flexíveis. O melhor modo de ligação foi considerado como sendo o modo com menor energia livre de ligação, o qual foi alinhado com a estrutura do receptor e escolhido para análises seguintes. A visualização da estrutura proteína-ligante e o preparo de figuras foram realizados com o software PyMol (The PyMOL Molecular Graphics System, Version 2.3.0 Schrödinger, LLC.).

As interações entre proteína-ligante foram avaliadas por meio da ferramenta LigPlot+ (Laskowski e Swindells, 2011) utilizando o complexo proteína-ligante (X2, X3 e X4) a fim de conhecer os resíduos envolvidos e o tipo de interação estabelecida com os ligantes.

3. RESULTADOS E DISCUSSÃO

3.1. Origem e descrição da ORF CB_612456

Dentre as enzimas preditas por homologia, a ORF CB_612456 apresentou 96% de cobertura e 64,84% de identidade com uma GH43 de *Chitinophagaceae* bacterium (número de acesso NCBI NCU05616.1) e tamanho compatível com sequências anotadas de enzimas similares (porém não caracterizadas) quando submetida a base de dados não curada “nr” (não redundante) do NCBI.

Entretanto, considerando os resultados obtidos no BLAST Protein (“Basic Local Alignment Search Tool”) contra a base de dados PDB (“Protein Data Bank”) e UniProt/SwissProt (“Universal Protein Resource”), os resultados de maior similaridade, respectivamente, foram 40% de cobertura com uma Endo-1,4-beta-xylanase D de *Bacteroides thetaiotaomicron* (número de acesso PDB 3qz4_A) e 39%

de cobertura com β -xilosidase/arabinosidase de *Bacteroides ovatus* (número de acesso Uniprot P49943.1), indicando que trata-se de uma β -xilosidase ainda não descrita. Dentre as enzimas depositadas nestes bancos de dados foram encontradas sequências que compartilhavam similaridade com um domínio de cada vez, seja o N-terminal ou o C-terminal.

3.2. Análise da estrutura primária

Considerando a análise da ORF CB_612456 contra a base de dados de famílias de domínios enzimáticos Pfam (<http://pfam.xfam.org/search/sequence>) a arquitetura encontrada para a sequência da BXil apresentou dois domínios GH43 (PF04616), ambos pertencentes a superfamília frutossidase (CL0143) (Naumoff, 2001) (Figura 1), sendo que os membros deste clã adotam uma dobra hélice beta de cinco lâminas (Nurizzo et al., 2002; Alberto et al., 2004) e é composto por nove famílias, dentre as quais estão GH32, GH43, GH62, GH68 e GH130. Até o momento do presente estudo, 387 sequências apresentaram a mesma arquitetura predita, composta por dois domínios sequenciais de GH43 (Figura 1), no entanto, nenhuma delas foi previamente caracterizada estrutural e funcionalmente.

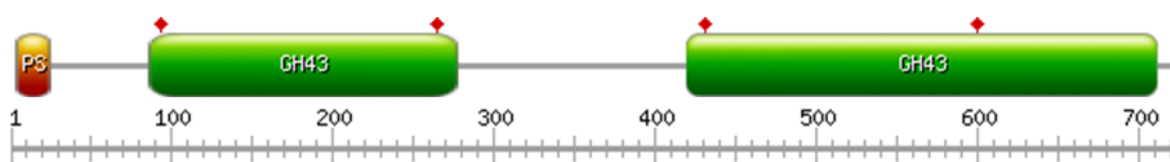


Figura 1. Representação esquemática da estrutura de domínios da BXil. A região em amarelo representa o peptídeo sinal predito por SignalP 5.0 e as regiões em verde representam cada um dos domínios, com resíduos do sítio ativo indicados por losangos vermelhos.

Os dados físico-químicos teóricos da BXil foram preditos empregando a ferramenta ProtParam ExPASy (<https://web.expasy.org/protparam/>) e resultaram em massa molecular de 81,25 kDa e pI teórico de 8. O índice de instabilidade foi de 35,7%, o que classifica esta proteína como estável. O índice alifático indica o volume relativo ocupado pelos aminoácidos alanina, valina, isoleucina e leucina nas cadeias laterais alifáticas e pode ser assumido como um indicador de termoestabilidade. A

Flavobacteria, ordem Flavobacteriales e família Flavobacteriaceae. O clado IV foi formado pela classe Sphingobacteria, ordem Sphingobacteriales e família Sphingobacteriaceae.

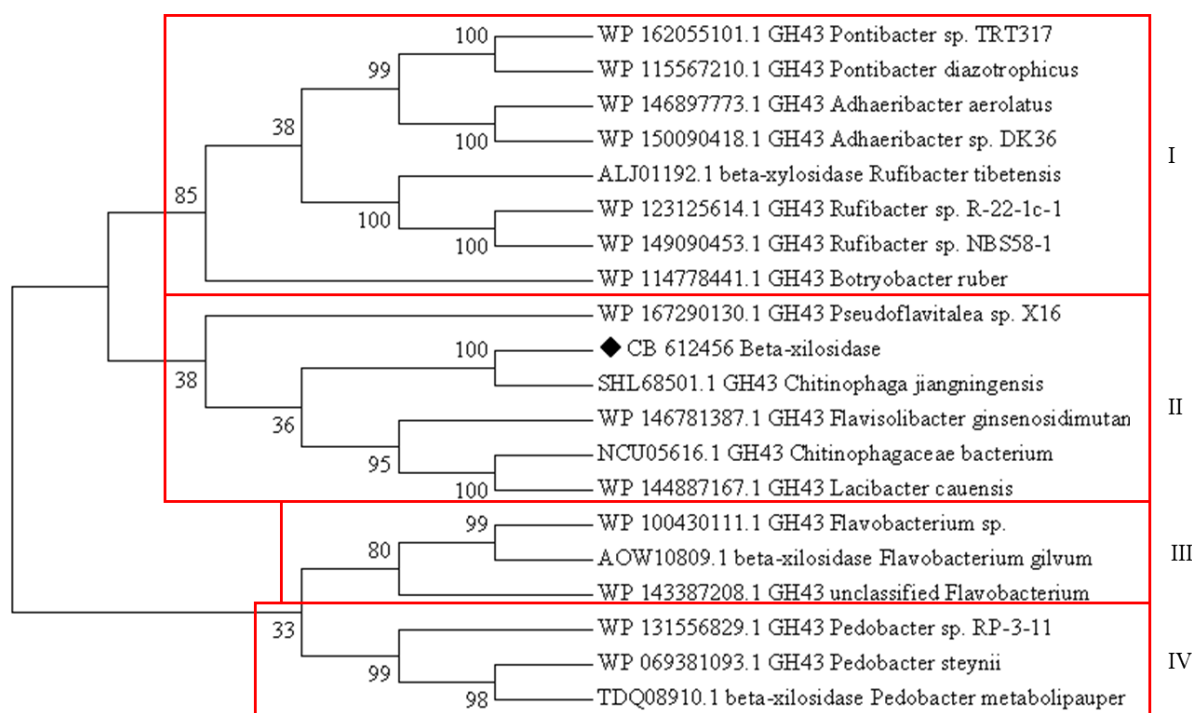


Figura 3. Fenograma de sequências de aminoácidos da BXil e de outras β -xilosidases depositadas no NCBI construído usando o método “Neighbor Joining” (Saitou e Nei, 1987). A árvore de consenso de “bootstrap” foi inferida a partir de 5000 réplicas. A porcentagem de árvores replicadas nas quais os táxons associados se agruparam no teste são mostrados próximos aos ramos. O losango preto indica a β -xilosidase de *Chitinophaga* sp. CB10 (BXil). As sequências de aminoácidos foram recuperadas do banco de dados “nr” do NCBI através da ferramenta BlastP, utilizando a ORF CB_612456 como sequência de consulta.

3.3. Análise da estrutura secundária

A estrutura secundária da β -xilosidase é composta por folhas beta (33,2%), hélices alfa (10,51%), e “random coil” (56,29%) (Figura 4.A). Além disso, uma longa região desordenada de ligação a proteínas foi observada nos primeiros 29 aminoácidos, os quais também representam o peptídeo sinal (Figura 4.B).

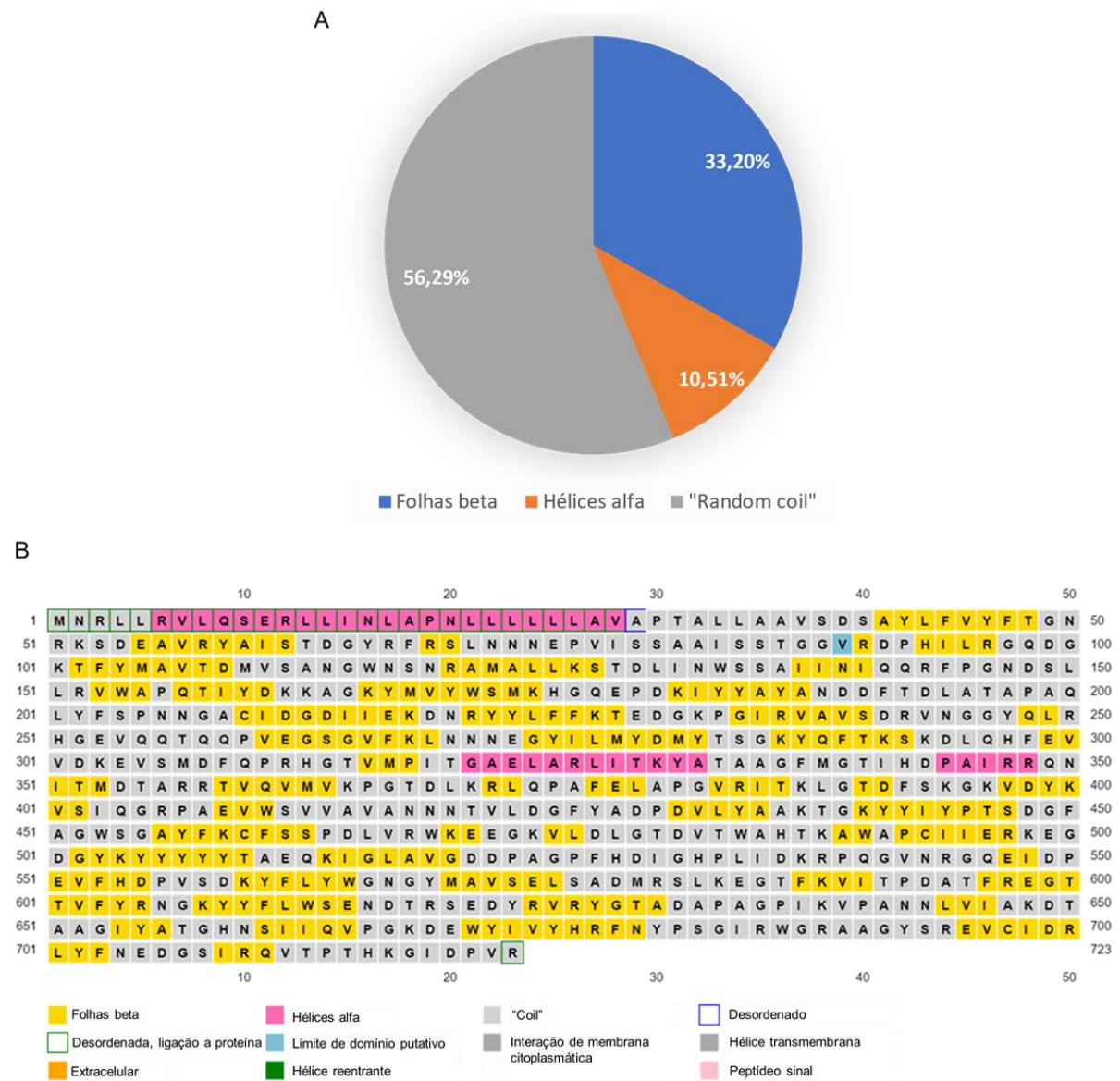


Figura 4. Predição da estrutura secundária da BXil. A: Composição da estrutura secundária predita através da ferramenta SOPMA (Geourjon e Deléage, 1995); B: Predição da estrutura secundária através da ferramenta PSIPRED (Buchan e Jones, 2019), onde regiões de hélices alfa são coloridas em rosa; folhas beta são coloridas em amarelo e regiões "coil" são coloridos em cinza claro. A região do peptídeo sinal formado por uma hélice alfa contornado em verde também representa uma região desordenada e de ligação a proteínas.

3.4. Modelagem por homologia e caracterização dos domínios

Devido à ausência de enzimas caracterizadas com a arquitetura de dois domínios GH43, a submissão da sequência de aminoácidos da β -xilosidase resultou em modelos estruturais individuais para cada domínio (N e C-terminal), ambos os

domínios formando um sítio ativo em forma de funil (Figura 5). A modelagem por homologia de proteínas multidomínios e de membrana ainda é um desafio a ser superado, assim como alcançar um $\text{RMSD} \leq 3 \text{ \AA}$ entre proteína alvo e molde, particularmente em proteínas grandes e proteínas com conteúdo significativo de estrutura β (Kessel e Ben-Tal, 2018).

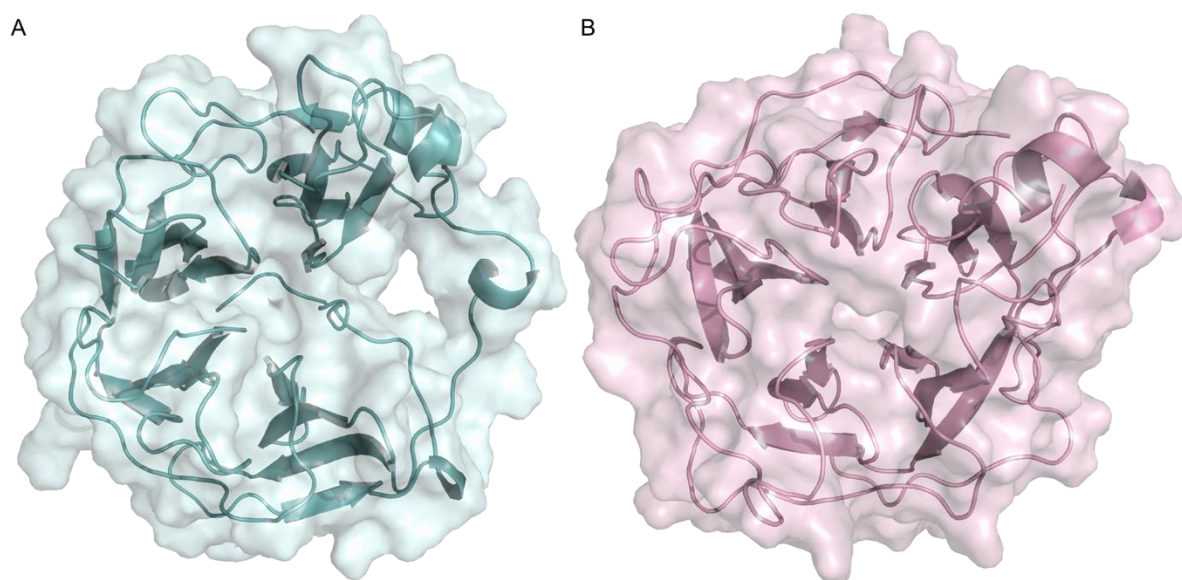


Figura 5. Estrutura geral da BXil representada em cartoon e superfície. A: N-terminal colorido em cerceta claro. B: C-terminal colorido em rosa claro.

Os modelos com melhor função de pontuação QMEAN (Benkert et al., 2008), que fornece a avaliação da qualidade do modelo, foram gerados com base nas estruturas cristalográficas de uma arabinofuranosidase GH62 de *Talaromyces pinophilus* (PDB 6f1j) para o domínio N-terminal e endo-1,4-beta-xilanase D GH43 (PDB 3qz4) de *Bacteroides thetaiotaomicron* para o C-terminal (Figura 5), compartilhando 17,24% e 56,36% de identidade com o molde, respectivamente. A baixa homologia apresentada entre o molde e o modelo gerado para o domínio N-terminal (17,24%) não representa uma boa confiabilidade preditiva.

A qualidade dos modelos estruturais para ambos os domínios foi avaliada pelo “web server” Rampage (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) antes e após a etapa de refinamento. O gráfico de Ramachandran para o modelo inicial do domínio N-terminal apresentou 89,6% dos resíduos em regiões favoráveis; 8,1% em regiões permitidas e 2,3% em outliers; enquanto que para a estrutura refinada 96,9%

dos resíduos permaneceram em regiões favoráveis; 2,3% em regiões permitidas e 0,8% em outliers. Para o modelo inicial do domínio C-terminal os resíduos do gráfico Ramachandran permaneceram 92,9% em regiões favoráveis; 5,4% em regiões permitidas e 1,7% em outliers enquanto que para a estrutura refinada esses valores chegaram em 95,6% em regiões favoráveis; 4,4% em regiões permitidas e nenhum outlier; indicando uma melhora significativa nas estruturas tridimensionais.

Os domínios N e C-terminal, quando alinhados entre si, revelaram baixa similaridade de sequência (41% de cobertura e 37,5% de identidade) e estrutural (RMSD de 5,524 Å - 5 ciclos de rejeição de outliers, 1632 de 2061 átomos), sugerindo que talvez cada um deles possua diferentes funções, uma vez que a maioria de β -xilosidases GH43 é bifuncional, exibindo atividade de β -xylosidase e -L-arabinofuranosidase simultaneamente (Moraïs et al., 2012).

No que diz respeito às demais análises da estrutura da BXil, a representação em superfície da BXil está colorida de acordo com seu potencial eletrostático (Figura 6.A e 6.C) e hidrofobicidade (Figura 6.B e 6.D) revelando que o canal do sítio ativo de ambos os domínios é mais eletronegativo que o restante da estrutura ao passo que parecem ser menos hidrofóbicos. Tal característica está diretamente relacionada a capacidade da enzima em se associar com o substrato através de interações não covalentes.

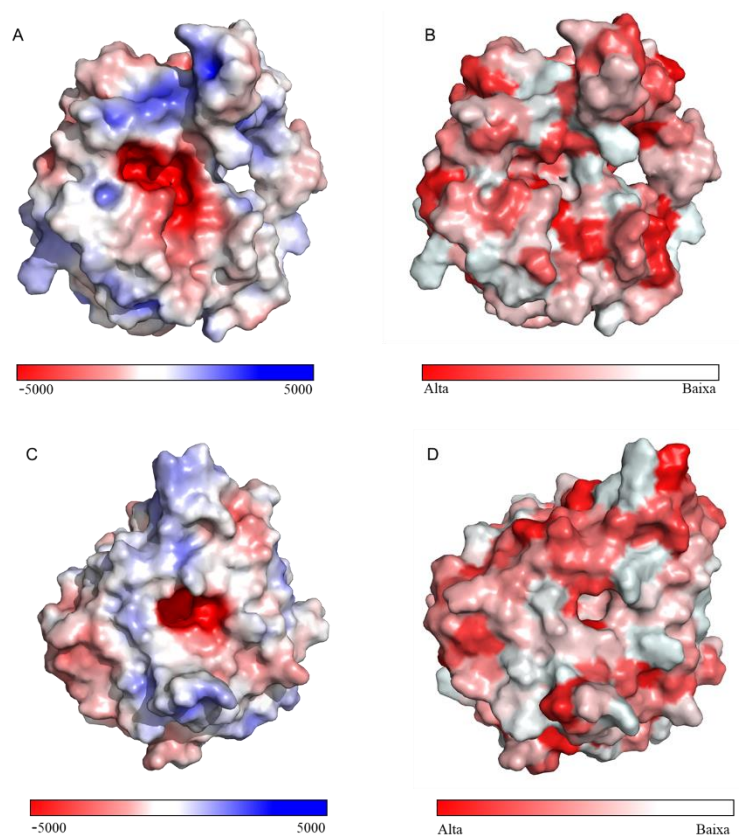


Figura 6. Superfície eletrostática e de hidrofobicidade do N-terminal (A e B) e do C-terminal (C e D) da BXil, respectivamente. Na representação de superfície eletrostática as superfícies coloridas em vermelho representam cargas negativas, em branco cargas neutras e em azul cargas positivas. Na representação de superfície de hidrofobicidade as superfícies vermelhas representam regiões hidrofóbicas e as brancas regiões não-hidrofóbicas, respectivamente.

Até o presente momento, poucas β -xilosidases GH43 possuem estrutura depositada no banco de dados PDB, sendo que a maioria delas apresenta estrutura composta por um domínio N-terminal formado por hélice beta de 5 lâminas dispostas em torno de um canal central e um C-terminal com conformação β -sanduíche de função desconhecida, e outras que apresentam domínio único.

A enzima BXil é composta por dois domínios hélice beta (Figura 5), sendo que os modelos estruturais gerados para cada um apresentaram um número diferentes de lâminas: o modelo N-terminal apresentou um domínio formado por hélice beta de 4 lâminas, contra 5 lâminas do modelo C-terminal (Figura 7). A diferença do número de lâminas entre os domínios pode ser evidência de funções diferentes, ou ainda, ser devido à baixa capacidade preditiva do modelo N-terminal (17,24% de identidade com o molde). Tipicamente, cada uma destas lâminas é formada por quatro fitas

β antiparalelas (Chen et al., 2011), no entanto cada domínio da BXil apresentou um número diferente de fitas: três e quatro fitas para o N-terminal e quatro fitas para o C-terminal (Figura 7).

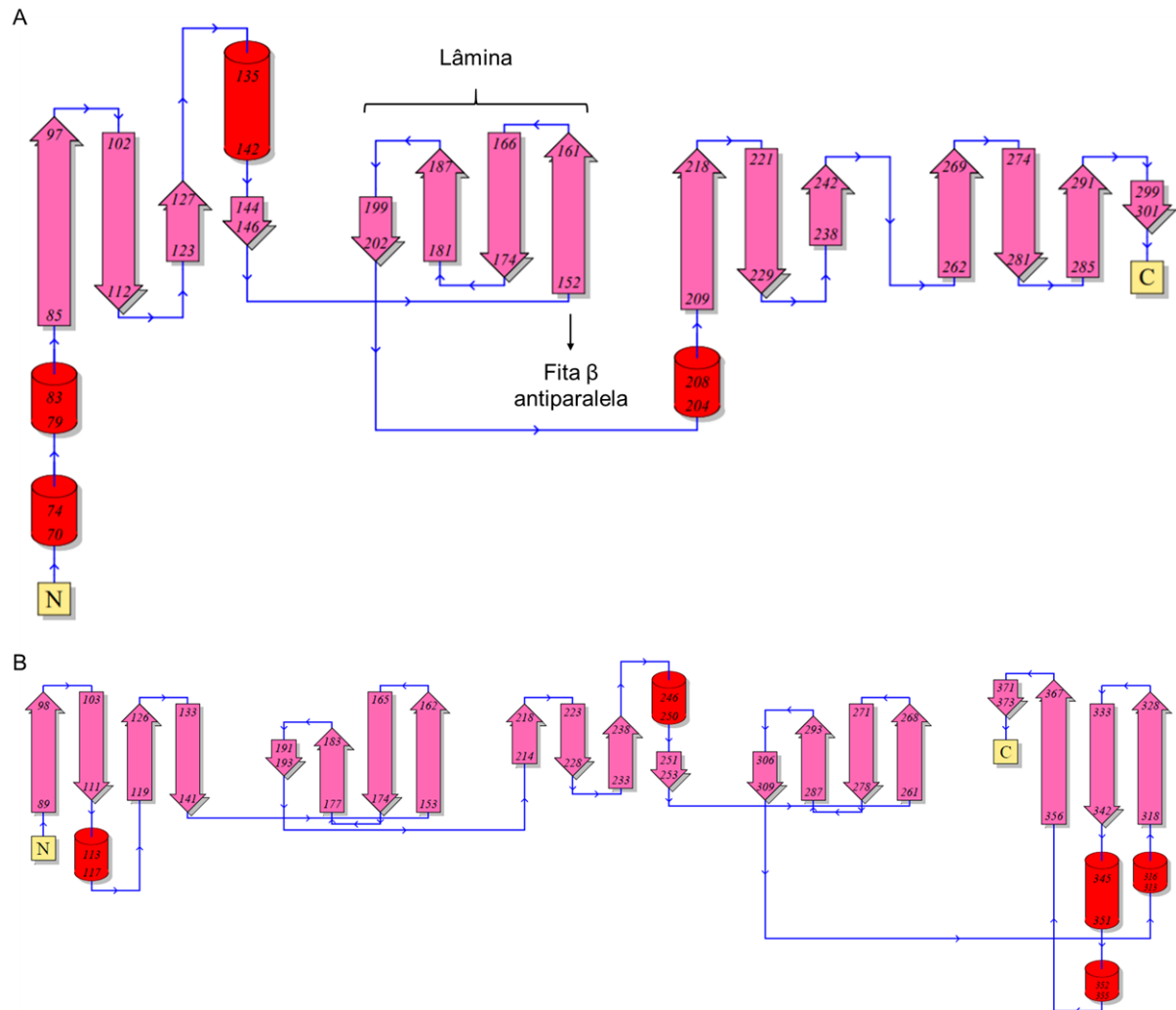


Figura 7. Diagrama de topologia dos domínios da BXil gerado pelo “web server” PDBsum Generate (<https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html>). A: Diagrama do domínio N-terminal exibindo quatro lâminas, cada uma composta por um número diferente de fitas β antiparalelas (três e quatro fitas). B: Diagrama de topologia do domínio C-terminal exibindo cinco lâminas, cada uma composta por quatro fitas β antiparalelas. Folhas beta representadas por setas coloridas em rosa; hélices beta representadas por cilindros coloridos em vermelho; “coils” representados por flechas coloridas em azul; regiões amino (N) e carboxi-terminal (C) de cada domínio são representados por caixas coloridas em amarelo.

Ainda, a estrutura hélice beta pode combinar-se com outros domínios dentro da mesma proteína de diferentes maneiras e, assim, executar funções distintas. A disposição destas estruturas é dividida em quatro grupos: sendo a) toda hélice, onde a proteína é formada por um ou dois domínios hélice com diferentes números de lâminas, como é o caso da BXil; b) hélice N-terminal, c) hélice C-terminal e d) hélice intermediária, onde a estrutura hélice beta está localizada no domínio N-terminal, C-terminal e entre eles, respectivamente (Chen et al., 2011).

Proteínas com estrutura hélice beta podem possuir de quatro a 10 lâminas, sendo que a variedade no número de lâminas pode estar diretamente associada à diversidade de funções, no entanto, como isso ocorre ainda não está completamente esclarecido (Chen et al., 2011; Afanasieva et al., 2019). Ademais, os “loops” encontrados na face superior e inferior destas proteínas possuem ampla plasticidade e também podem favorecer a diversidade funcional de proteínas (Chen et al., 2011).

Estruturas com cinco a oito lâminas são amplamente distribuídas na natureza e aquelas que possuem quatro ou cinco lâminas possuem funções limitadas às hidrolases, transferases e proteínas de ligação ao açúcar, enquanto que aquelas que possuem seis ou sete lâminas apresentam uma diversidade de funções muito mais ampla (Chen et al., 2011).

Embora não tenha sido possível prever a topologia geral da BXil e a geometria dos seus domínios, uma vez que a previsão da estrutura terciária ocorreu de maneira individual para cada domínio, quando a BXil foi avaliada pelo servidor I-Tasser/TM Align (Zhang e Skolnick, 2005) foi possível identificar 10 estruturas análogas à BXil. A endo-xiloglucanase GH74 de *Cellvibrio japonicus* (PDB 5fkq) apresentou maior similaridade estrutural (RMSD 2,01 Å) a partir do modelo BXil predito pelo próprio programa contra a base de dados PDB. Assim como a BXil, a endo-xiloglucanase (PDB 5fkq) é composta por dois domínios hélice beta, diferindo entre si no número de lâminas apresentadas, sendo um domínio típico de GH74, com sete lâminas, e outro domínio típico de GH43, com cinco lâminas. Aparentemente, o programa retornou maior índice de similaridade entre a estrutura PDB 5fkq e a BXil devido ao módulo similar GH43 (C-terminal na BXil). De acordo com Yao e colaboradores (2004), apesar de algumas proteínas compartilharem a mesma estrutura hélice beta, elas podem diferir amplamente na sequência e na função, acreditando-se ainda que essa

conformação estrutural seja responsável por fornecer uma arquitetura base para inúmeras funções.

Para as GHs 43, até o momento, não há nenhuma proteína com a mesma arquitetura de domínios encontradas para BXil caracterizada funcional ou estruturalmente.

A ausência de estudos de ancoragem molecular com ligante e de caracterização e estrutural para esta arquitetura de domínios dificulta o entendimento sobre a função de cada um dos domínios e a relação entre eles. Ambos participam em conjunto do reconhecimento e organização do substrato ou cada um realiza a hidrólise individualmente? Moraes e colaboradores (2012) realizaram um estudo com uma β -xilosidase de *Thermobifida fusca* formada por um domínio N-terminal hélice beta (domínio catalítico) e um domínio C-terminal β -sanduíche de função desconhecida, onde avaliaram o comportamento dos domínios separadamente e em conjunto. Os domínios separados levaram à completa perda da atividade catalítica do N-terminal, porém não afetou sua capacidade de se ligar à xilana, sugerindo que a presença do domínio C-terminal é necessária para correta degradação do substrato. Além disso, a reassociação dos domínios restaurou a atividade enzimática indicando que ambos funcionam em conjunto e podem apresentar funções distintas.

Considerando a conformação de hélice beta com quatro lâminas encontrada para o domínio N-terminal da BXil, essa estrutura supersecundária poderia estar relacionada a hidrolases, transferases ou mesmo proteínas de ligação ao açúcar, os CBMs ("Carbohydrate-Binding Module"), o que não permite, apenas fundamentado nessa análise a identificação de seu papel funcional. Por outro lado, a baixa similaridade entre a estrutura de referência e a proteína alvo (<20%) podem não ter resultado em um modelo realístico para o domínio N-terminal, o que dificulta ainda mais a capacidade de predição *in silico* de sua função, carecendo que futuros estudos voltados para cristalografia de raios-X e análise cinética dos domínios combinados ou separados que poderão elucidar o papel funcional deste domínio.

3.5. Ancoragem molecular e caracterização do sítio ativo da BXil

Considerando a baixa capacidade preditiva do modelo N-terminal, optamos por realizar os estudos de ancoragem molecular somente com o domínio C-terminal, o qual representa uma boa conservação estrutural entre as estruturas (56,36 % de identidade de sequência).

O sítio ativo do domínio C-terminal apresentou uma topologia de bolso e os resíduos da tríade catalítica, preditos automaticamente através do CDD/Sparcle (Lu et al., 2020) no NCBI, se encontram nas posições D429, D549, E598 no centro da cavidade formada pela hélice beta de cinco lâminas (Figura 8). Estes resíduos estão presentes em enzimas pertencentes às famílias 43 e 62 (Clan F) e 32 e 68 (clan J) de Glicosil Hidrolases (Pons et al., 2004).

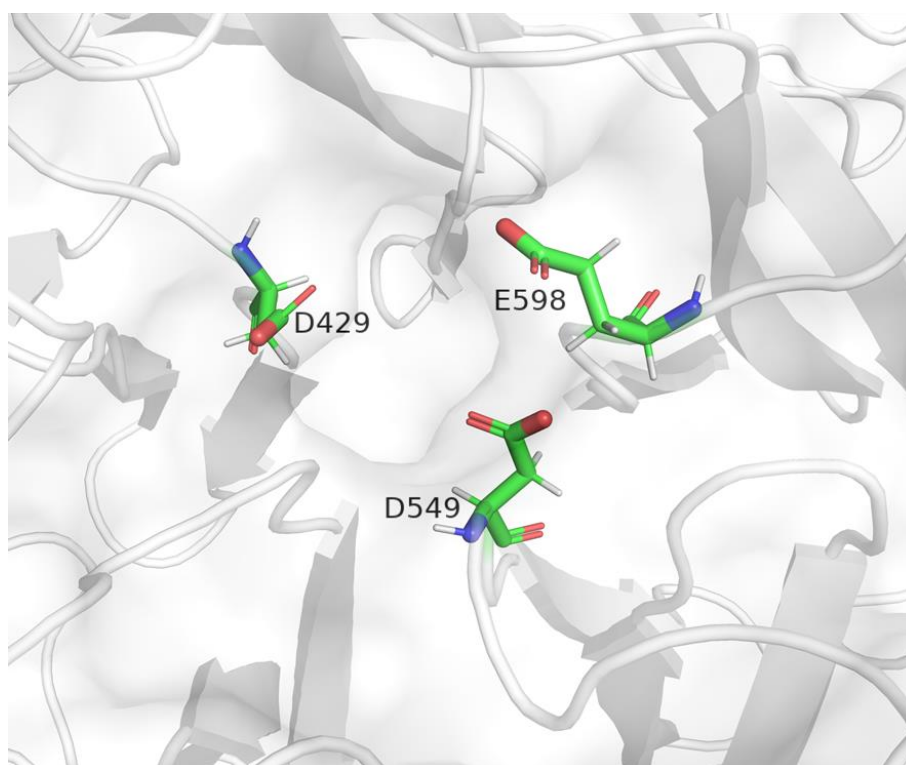


Figura 8. Resíduos da tríade catalítica do C-terminal da BXil. Os resíduos de carbono estão coloridos em verde, nitrogênios em azul, oxigênios em vermelho e hidrogênios em cinza claro.

Os estudos de ancoragem molecular visam prever a melhor conformação de uma molécula (ligante) em outra (receptor), como o sítio ativo de uma proteína por exemplo, a fim de formar um complexo estável (Lengauer e Rarey, 1996). Este método

é amplamente utilizado no desenho racional de fármacos, porém sua aplicação para elucidar processos bioquímicos de moléculas ainda é incipiente.

Para cada ligante avaliados diferentes resultados foram obtidos e somente o modo com menor energia livre de ligação (melhor afinidade) foi selecionado para as análises seguintes. A enzima BXil apresentou forte energia de afinidade com os substratos X2, X3, X4 e X5, que foram de -7,3; -8,2; -7 e 1,4 kcal/mol; respectivamente. A afinidade entre BXil e o substrato X3 foi maior do que com X2, mas significativamente menor com X5, isso demonstra que o sítio ativo da BXil pode facilmente acomodar um substrato intermediário até X4.

A ferramenta LigPlot+ v2.2 (Laskowski e Swindells, 2011) foi utilizada a fim de prever potenciais interações entre os resíduos e o ligante. Os resíduos envolvidos nas interações proteína-ligante com os diferentes substratos podem ser visualizados na Tabela 1 e na Figura 9.

Tabela 1. Resíduos da BXil envolvidos em contatos com os ligantes.

Ligante	Energia de ligação (kcal/mol)	Aminoácidos envolvidos em ligações de hidrogênio	Aminoácidos envolvidos em interações hidrofóbicas
X2	-7,3	Asp429*, Asp549*, Asn567, Glu598*, Asp617, Thr618, His659, Arg694	Trp453, Ala491, Ile548, His676
X3	-8,2	Asp429*, Asp549*, Asn567, Glu598*, Asp616, Asp617, Arg619, His659, His676	Trp453, Trp490, Ala491, Gln546, Arg597, Arg694
X4	-7	Asp429*, Asp549*, Asn567, Glu598*, Asp616, Asp617, Thr618, Arg619, His659, His676	Trp453, Trp490, Ala491, Gln546, Arg597, Arg694

*: resíduos da tríade catalítica

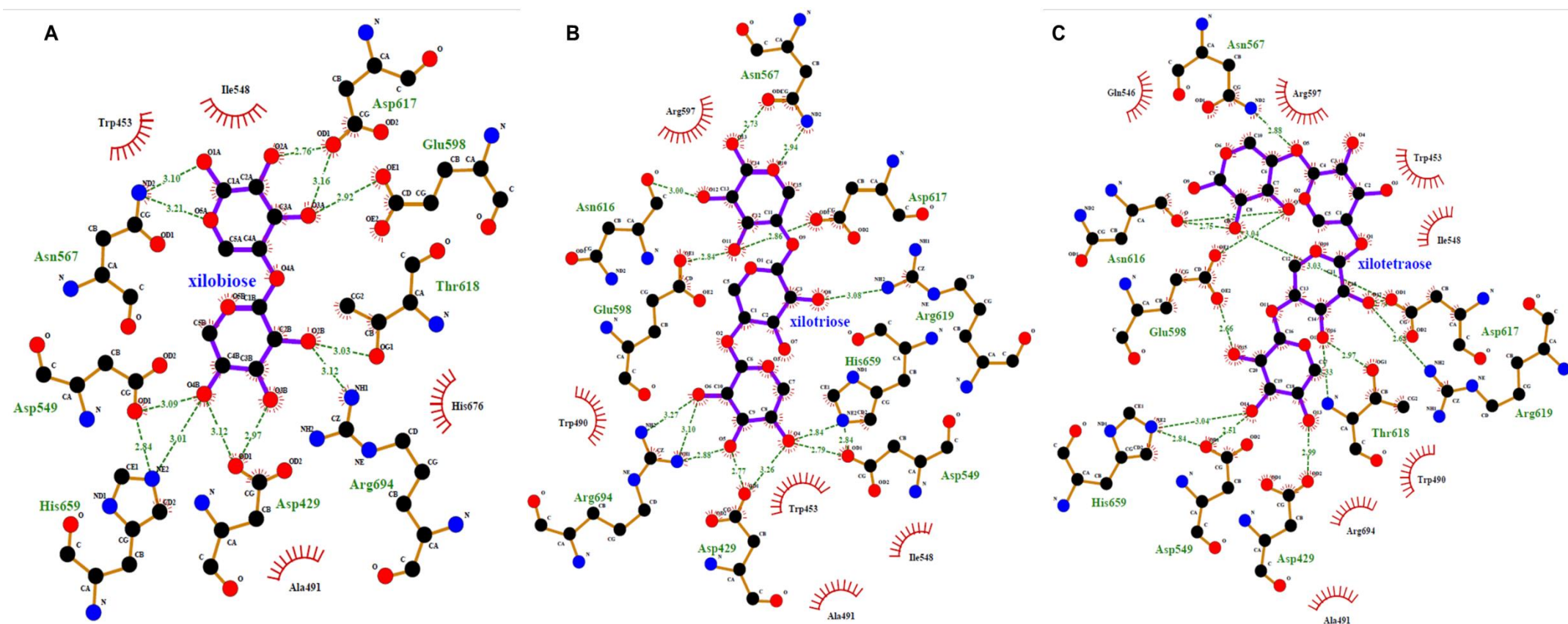


Figura 9. Diagrama de potenciais interações entre a enzima BXil e seus ligantes. A: xilobiose; B: xilotriose; C: xilotetraose. Linhas tracejadas verdes indicam ligações de H e seu comprimento em angstroms; ícones vermelhos indicam resíduos envolvidos em contatos hidrofóbicos; os conectores dos ligantes estão coloridos em lilás e dos resíduos de aminoácidos em dourado. Figuras preparadas pelo Pymol e LigPlot+ v2.2 (Laskowski e Swindells, 2011).

O servidor I-Tasser/COACH (Yang et al., 2013a; 2013b), combinado com outros programas baseados em estrutura, realiza a predição de locais de ligação proteína-ligante a partir do banco de dados de funções de proteínas BioLiP, através de comparações de subestrutura específica de ligação e de perfis de sequência (Yang et al., 2013a, 2013b). Com base nos resultados obtidos, os resíduos do local de ligação preditos para a BXil foram similares aos anotados pelo servidor para a estrutura PDB 2exj complexada com α -D-xilopirranose, sendo possível a predição do local de ligação da enzima BXil (Figura 10).

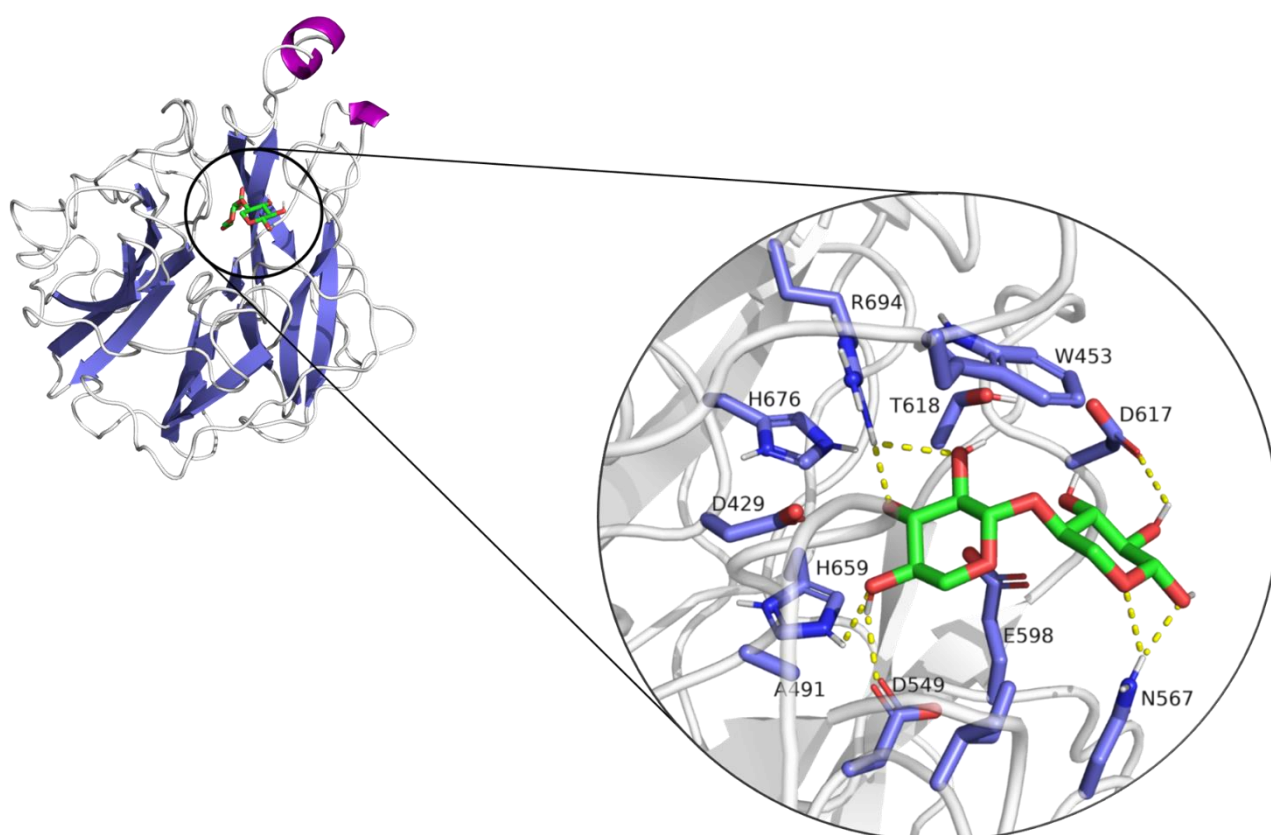


Figura 10. Resíduos da enzima BXil envolvidos na interação com xilobiose preditos a partir de análises com o servidor I-Tasser/COACH (Yang et al., 2013a; 2013b). Resíduos de aminoácidos coloridos em ardósia e xilobiose colorido em verde. Linhas tracejadas amarelas indicam ligações de hidrogênio.

O resíduo da tríade catalítica D429 atua como nucleófilo e está localizado a 5,8 Å do carbono anomérico do substrato no subsítio -1, enquanto que o E598, que atua como próton doador, está a 5,1 Å do oxigênio glicosídico (Figura 11.A). O resíduo D549, que está localizado a 5,5 Å do resíduo E598, é essencial para a modulação pKa

do E598 bem como sua correta orientação em relação ao substrato (Nurizzo et al., 2002; Br  x et al., 2006) (Figura 11.A). O pr  ton doador e a base geral est  o localizados em faces opostas da cavidade e a dist  ncia entre eles    de 7,5   , aproximadamente 1,5    abaixo da dist  ncia comum observada em outras enzimas invertases (McCarter e Stephen Withers, 1994; Br  x et al., 2006).

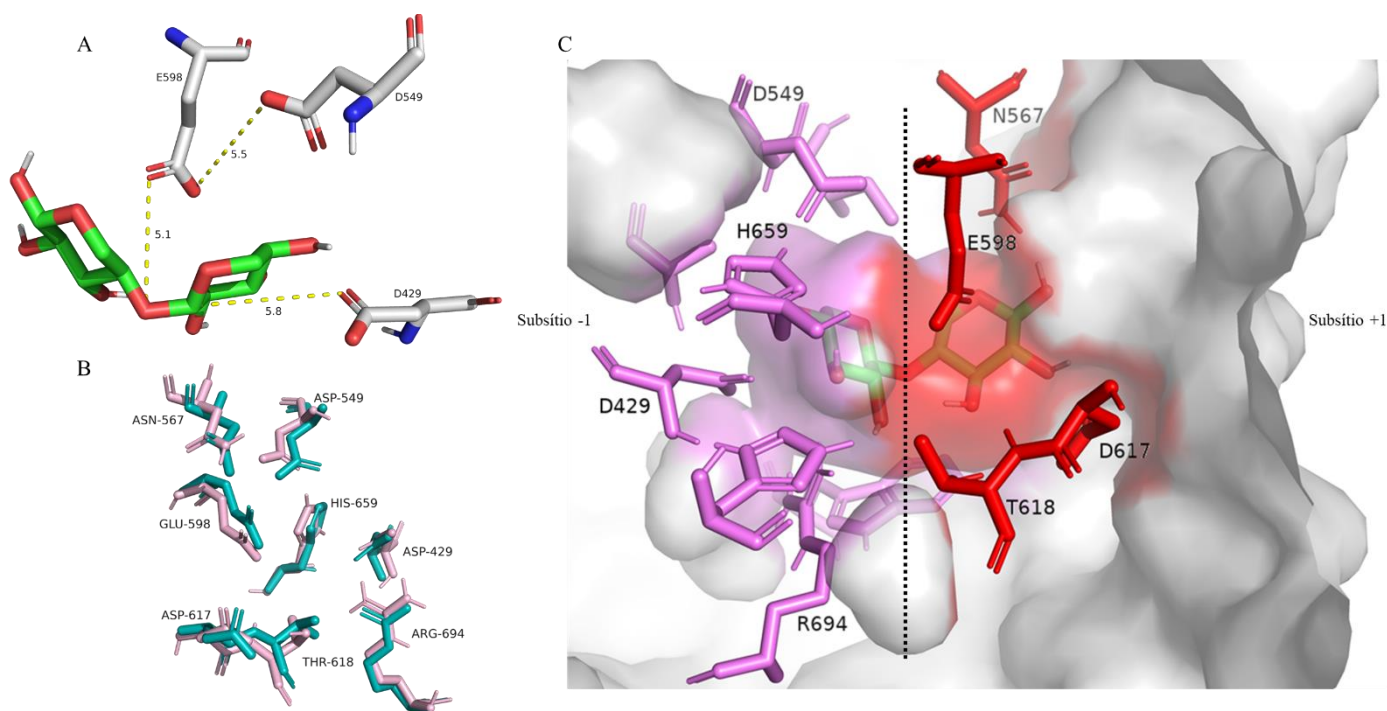


Figura 11. Representa  o esquem  tica do s  tio ativo da BXil. A: Configura  o do s  tio ativo com o ligante xilobiose indicando a dist  ncia em angstroms entre os res  duos e o ligante. O ligante xilobiose (X2) est   colorido em verde e os res  duos catal  ticos coloridos em branco; oxig  nios em vermelho, nitrog  nios em azul e hidrog  nios em cinza. B: Sobreposi  o dos res  duos da estrutura PDB 3qz4 (coloridos em cerceta) com os principais res  duos da BXil (coloridos em rosa claro) envolvidos com o ligante xilobiose. C: Subs  tios de liga  o ao substrato. Res  duos coloridos em rosa e vermelho indicam o subs  tio -1 e +1, respectivamente.

O ligante X2 foi escolhido para representar o complexo prote  na-ligante nas an  lises seguintes a fim de facilitar a an  lise das intera  es. Os principais res  duos do s  tio ativo do C-terminal foram definidos a partir da ancoragem molecular com o ligante X2 e est  o distribu  dos em dois subs  tios (Figura 11.C).

Os membros da fam  lia GH43 realizam a hidr  lise de extremidades n  o redutoras de substratos aril-glicos  dicos atrav  s do mecanismo de invers  o do carbono anom  rico e a clivagem da liga  o glicos  dica ocorre entre os subs  tios -1

e +1 (representado pela linha tracejada na Figura 11.C). O subsítio -1, ou subsítio glicona, apresenta um local de ligação para a extremidade não redutora do substrato e realiza várias ligações de hidrogênio com os resíduos D429, D549, H659 e R694 e contatos hidrofóbicos com os resíduos W453, A491, I548 e H676. No subsítio +1, ou aglicona, se localiza a extremidade redutora do substrato e os demais resíduos envolvidos na degradação de xilobiose (N567, E598, D617 e T618 realizando ligações de hidrogênio) (Figura 9, Tabela 1).

A superposição da estrutura do C-terminal da BXil com a estrutura cristalográfica PDB 3qz4 utilizada como molde para construção da estrutura alvo apresentou RMSD de 0,732 Å (5 ciclos de rejeição de outliers, 1811 de 2056 átomos) e os principais resíduos envolvidos na hidrólise de substrato foram bem sobrepostos indicando que o sítio ativo possui orientação semelhante entre as moléculas (Figura 11.B).

O servidor CASTp (Tian et al., 2018) (com um raio de sonda padrão de 1,4 Å) foi utilizado para calcular e comparar a área de superfície e o volume do canal do sítio ativo do C-terminal da BXil e da estrutura da endoxilanase PDB 3qz4. A área e o volume do canal para a estrutura BXil foram de 197,298 Å² e 122,826 Å³, respectivamente, enquanto que para a endoxilanase os valores foram de 152.725 Å² e 106.322 Å³, indicando similaridade entre o canal das duas estruturas.

Além disso, o alinhamento de sequências realizadas pelo ClustalW (Thompson et al., 1994) permitiu identificar as regiões conservadas do domínio C-terminal e, de acordo com o esperado, os resíduos envolvidos na interação com o substrato encontram-se em regiões altamente conservadas (Figura 12).

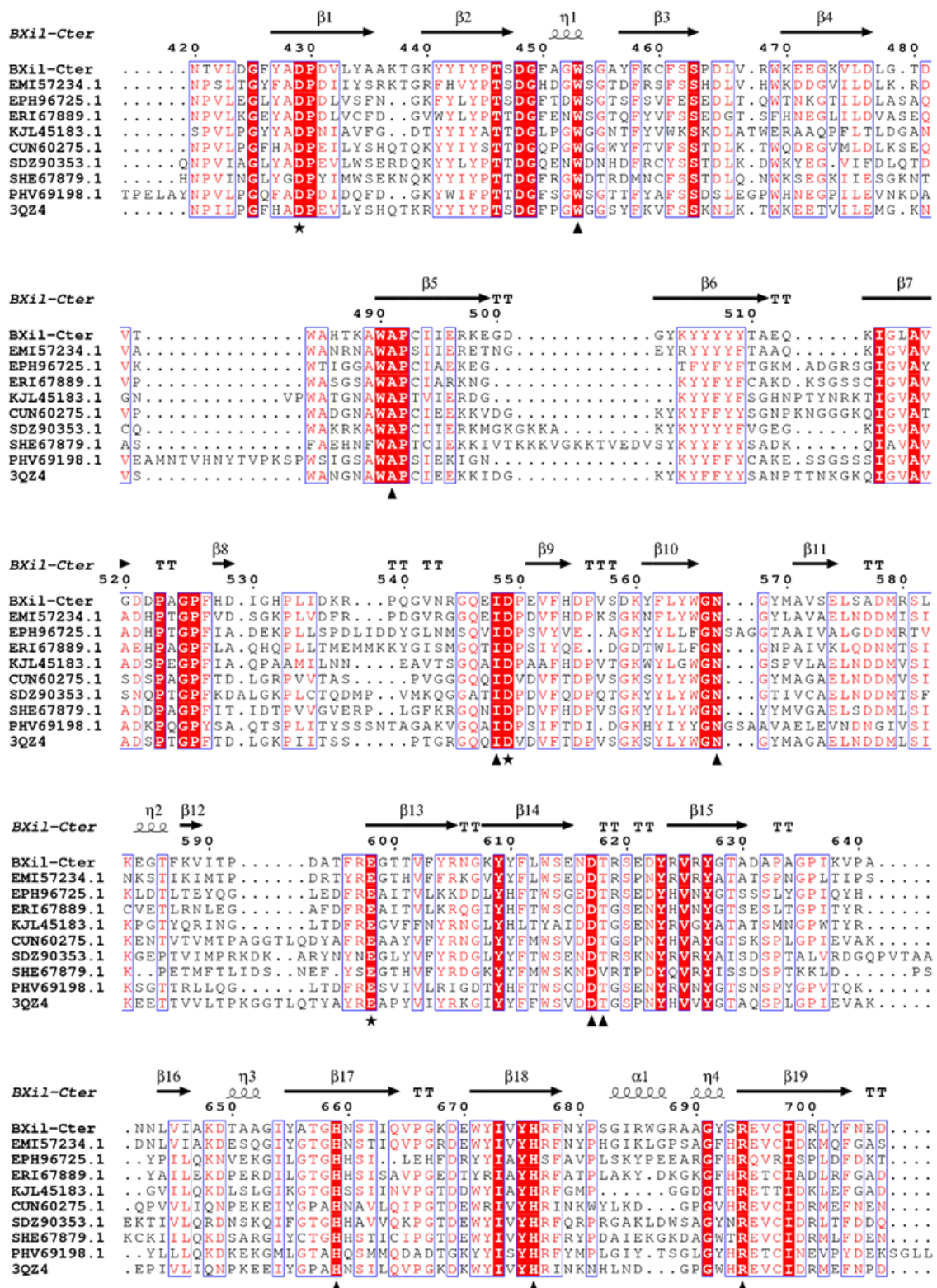


Figura 12. Alinhamento de sequências de aminoácidos da BXil com sequências similares (EMI57234.1, EPH96725.1; ERI67889.1; KJL45183.1; CUN60275.1; SDZ90353.1; SHE67879.1; PHV69198.1; 3QZ4). Caixas preenchidas em vermelho com letras brancas: resíduos estritamente conservados; letras vermelhas dentro de

caixas azuis: regiões altamente similares. Os resíduos da tríade catalítica (D429, E549 e E598) estão indicados por uma estrela, os demais resíduos envolvidos na interação com o substrato estão indicados por um triângulo abaixo da sequência. Elementos de estrutura secundária obtidos a partir do C-terminal apresentados no topo (α -hélices como espirais, folhas β como setas e voltas como letras TT). O alinhamento foi realizado através do ClustalW (Thompson et al., 1994) e a figura elaborada utilizando o programa ESPript (Robert e Gouet, 2014).

4. CONCLUSÕES

A partir deste estudo foi possível prever a estrutura tridimensional da BXil que possui uma repetição de domínio hélice beta atribuídos como GH43, distribuídos em uma arquitetura de domínios que até o momento não possui caracterização funcional e estrutural. Embora ambos domínios tenham sido identificados como GH43, apenas o C-terminal apresenta a estrutura secundária hélice beta com disposição de cinco lâminas características de GH43, enquanto que o domínio N-terminal apresenta apenas quatro lâminas, o que pode ser um artefato da predição estrutural, uma vez que este domínio não é similar a nenhum outro previamente determinado por experimentos cristalográficos, permanecendo de função desconhecida. Considerando a qualidade superior encontrada para o domínio C-terminal, foi possível prever a disposição de seu sítio ativo e os principais resíduos envolvidos em interações com diferentes substratos, fornecendo algumas percepções valiosas sobre suas propriedades em relação à sua capacidade em acomodar substratos de tamanho intermediário em seu canal catalítico, corroborando a predição funcional desta enzima com β -xilosidases GH43. Entretanto, a função do domínio N-terminal ainda não pôde ser esclarecida, carecendo de estudos adicionais de cristalografia e cinética enzimática dos domínios individuais e em conjunto para a completa caracterização funcional de BXil. Apesar disto, a ausência de enzimas similares à BXil na literatura pode ser evidência de uma enzima de atuação distinta das demais enzimas conhecidas, o que pode ser valioso do ponto de vista biotecnológico.

5. REFERÊNCIAS

Afanasieva, E., Chaudhuri, I., Martin, J., Hertle, E., Ursinus, A., Alva, V., Hartmann, M.D. e Lupas, A.N. (2019). Structural diversity of oligomeric β -propellers with different numbers of identical blades. **eLife** 8: 1–15.

Alberto, F., Bignon, C., Sulzenbacher, G., Henrissat, B. e Czjzek, M. (2004). The Three-dimensional Structure of Invertase (β -Fructosidase) from *Thermotoga maritima* Reveals a Bimodular Arrangement and an Evolutionary Relationship between Retaining and Inverting Glycosidases. **Journal of Biological Chemistry** 279: 18903–18910.

Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G. e Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. **Nature Biotechnology** 37: 420–423.

Benkert, P., Tosatto, S. C., & Schomburg, D. (2008). QMEAN: A comprehensive scoring function for model quality assessment. *Proteins: Structure, Function, and Bioinformatics* 71: 261-277.

Bhattacharya, D. e Valencia, A. (2019). RefinedD: Improved protein structure refinement using machine learning based restrained relaxation. **Bioinformatics** 35: 3320–3328.

Brüx, C., Ben-David, A., Shallom-Shezifi, D., Leon, M., Niefind, K., Shoham, G., Shoham, Y. e Schomburg, D. (2006). The Structure of an Inverting GH43 β -Xylosidase from *Geobacillus stearothermophilus* with its Substrate Reveals the Role of the Three Catalytic Residues. **Journal of Molecular Biology** 359: 97–109.

Buchan, D.W.A. e Jones, D.T. (2019). The PSIPRED Protein Analysis Workbench: 20 years on. **Nucleic Acids Research** 47: W402–W407.

Canilha, L., Kumar Chandel, A., dos Santos Milessi, T.S., Fernandes Antunes, F.A., da Costa Freitas, W.L., das Graças Almeida Felipe, M. e da Silva, S.S. (2012). Bioconversion of sugarcane biomass into ethanol: an overview about composition, pretreatment methods, detoxification of hydrolysates, enzymatic saccharification, and ethanol fermentation. **Journal of biomedicine & biotechnology** 2012: 1–15.

Chen, C.K.M., Chan, N.L. e Wang, A.H.J. (2011). The many blades of the β -propeller proteins: Conserved but versatile. **Trends in Biochemical Sciences** 36: 553–561.

Collins, T., Gerday, C. e Feller, G. (2005). Xylanases, xylanase families and extremophilic xylanases. **FEMS Microbiology Reviews** 29: 3–23.

Falck, P., Linares-Pastén, J.A., Adlercreutz, P. e Karlsson, E.N. (2016). Characterization of a family 43 β -xylosidase from the xylooligosaccharide utilizing putative probiotic *Weissella* sp. strain 92. **Glycobiology** 26: 193–202.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., et al. (2016). The Pfam protein families database: Towards a more sustainable future. **Nucleic Acids Research** 44: D279–D285.

- Funicelli, M.I.G. (2018). **Explorando uma comunidade bacteriana isolada de uma pilha de bagaço de cana-de-açúcar e seu potencial funcional na degradação de biomassa lignocelulósica**. 65 f. Dissertação (Mestrado em Microbiologia Agropecuária) – Unesp, Jaboticabal.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. e Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In: J.M., W. (Eds.), **The Proteomics Protocols Handbook**. Humana Press: p. 571–607.
- Geourjon, C. e Deléage, G. (1995). Sopma: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. **Bioinformatics** 11: 681–684.
- Huy, N.D., Thayumanavan, P., Kwon, T.H. e Park, S.M. (2013). Characterization of a recombinant bifunctional xylosidase/arabinofuranosidase from *Phanerochaete chrysosporium*. **Journal of Bioscience and Bioengineering** 116: 152–159.
- Juturu, V. e Wu, J.C. (2012). Microbial xylanases: Engineering, production and industrial applications. **Biotechnology Advances** 30: 1219–1227.
- Kessel, A. e Ben-Tal, N. (2018). **Methods of Structure Determination and Prediction**. In: Kessel A., Ben-Tal N. (Eds.) *Introduction to Proteins: Structure, Function, and Motion*, Chapman and Hall/CRC: p. 259-354.
- Kishi, L.T., Lopes, E.M., Fernandes, C.C., Fernandes, G.C., Sacco, L.P., Alves, L.M.C. e Lemos, E.G.M. (2017). Draft Genome Sequence of a Chitinophaga Strain Isolated from a Lignocellulose Biomass-Degrading Consortium. **Genome Announcements** 5: 1–2.
- Knob, A., Terrasan, C.R.F. e Carmona, E.C. (2010). β -Xylosidases from filamentous fungi: An overview. **World Journal of Microbiology and Biotechnology** 26: 389–407.
- Laskowski, R.A. e Swindells, M.B. (2011). LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. **Journal of Chemical Information and Modeling** 51: 2778–2786.
- Lengauer, T. e Rarey, M. (1996). Computational methods for biomolecular docking. **Current Opinion in Structural Biology** 6: 402–406.
- Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., et al. (2020). CDD/SPARCLE: the conserved domain database in 2020. **Nucleic acids research** 48: D265–D268.

Kumar S., Stecher G., Tamura K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Molecular Biology and Evolution Society** 33: 1870-1874.

McCarter, J.D. e Stephen Withers, G. (1994). Mechanisms of enzymatic glycoside hydrolysis. **Current Opinion in Structural Biology** 4: 885–892.

Moraís, S., Salama-Alber, O., Barak, Y., Hadar, Y., Wilson, D.B., Lamed, R., Shoham, Y. e Bayer, E.A. (2012). Functional association of catalytic and ancillary modules dictates enzymatic activity in glycoside hydrolase family 43 β -xylosidase. **Journal of Biological Chemistry** 287: 9213–9221.

Moreira, L.R.S. e Filho, E.X.F. (2016). Insights into the mechanism of enzymatic hydrolysis of xylan. **Applied Microbiology and Biotechnology** 100: 5205–5214.

Naumoff, D.G. (2001). beta-fructosidase superfamily: homology with some alpha-L-arabinases and beta-D-xylosidases. **Proteins** 42: 66–76.

Nurizzo, D., Turkenburg, J.P., Charnock, S.J., Roberts, S.M., Dodson, E.J., McKie, V.A., Taylor, E.J., Gilbert, H.J. e Davies, G.J. (2002). Cellvibrio japonicus α -l-arabinanase 43a has a novel five-blade β -propeller fold. **Nature Structural Biology** 9: 665–668.

Polizeli, M.L.T.M., Rizzatti, A.C.S., Monti, R., Terenzi, H.F., Jorge, J.A. e Amorim, D.S. (2005). Xylanases from fungi: Properties and industrial applications. **Applied Microbiology and Biotechnology** 67: 577–591.

Pons, T., Naumoff, D.G., Martínez-Fleites, C. e Hernández, L. (2004). Three Acidic Residues are at the Active Site of a β -Propeller Architecture in Glycoside Hydrolase Families 32, 43, 62, and 68. **Proteins: Structure, Function and Genetics** 54: 424–432.

Robert, X. e Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. **Nucleic Acids Research** 42: 320–324.

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. **Nature protocols** 5: 725-738.

Saitou N., Nei M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution Society** 4: 406-425.

Schrödinger, L. (2015). The PyMOL Molecular Graphics System, Version 2.0.

Shallom, D. e Shoham, Y. (2003). Microbial hemicellulases. **Current Opinion in Microbiology** 6: 219–228.

Thompson, J.D., Higgins, D.G. e Gibson, T.J. (1994). Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. **Nucleic Acids Research** 22: 4673–4680.

Tian, W., Chen, C., Lei, X., Zhao, J. e Liang, J. (2018). CASTp 3.0: Computed atlas of surface topography of proteins. **Nucleic Acids Research** 46: W363–W367.

Wang, G., Huang, X., Ng, T.B., Lin, J. e Ye, X.Y. (2014). High phylogenetic diversity of glycosyl hydrolase family 10 and 11 xylanases in the sediment of Lake Dabusu in China. **PLoS ONE** 9: e112798.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. **Nature methods** 12: 7-8.

Yang, J., Roy, A. e Zhang, Y. (2013a). Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. **Bioinformatics** 29: 2588–2595.

Yang, J., Roy, A. e Zhang, Y. (2013b). BioLiP: A semi-manually curated database for biologically relevant ligand-protein interactions. **Nucleic Acids Research** 41: 1096–1103.

Yaoi, K., Kondo, H., Noro, N., Suzuki, M., Tsuda, S. e Mitsuishi, Y. (2004). Tandem repeat of a seven-bladed β -propeller domain in oligoxyloglucan reducing-end-specific cellobiohydrolase. **Structure** 12: 1209–1217.

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. **BMC bioinformatics** 9: 40.

Zhang, Y., & Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. **Nucleic acids research** 33: 2302-2309.

CAPÍTULO 4 – Expressão, extração e ensaios preliminares de cristalização da β -xilosidase recombinante (BXil) expressa em corpos de inclusão

RESUMO – A produção de proteínas recombinantes geralmente é realizada a partir de hospedeiros heterólogos, como a *Escherichia coli*, a fim de realizar ensaios de caracterização funcional e estrutural destas macromoléculas. No entanto, durante a etapa de expressão proteica frequentemente ocorre a formação de agregados proteicos, denominados corpos de inclusão. Os corpos de inclusão afetam diretamente o rendimento de proteína solúvel, sendo que proteínas insolúveis representam um fator limitante para a biologia estrutural. A recuperação da proteína a partir de corpos de inclusão é possível através de etapas de solubilização e renaturação *in vitro*. Contudo, apesar da existência de inúmeros protocolos disponíveis para solubilização de corpos de inclusão e renaturação proteica, o sucesso desta técnica depende diretamente das características intrínsecas de cada proteína podendo variar consideravelmente de uma para outra, sendo necessário estabelecê-lo através de ensaios de tentativa e erro. Diante disso, o principal objetivo do presente estudo foi testar diferentes protocolos de expressão, extração e renaturação *in vitro* de uma β -xilosidase recombinante (BXil) a fim de estabelecer um protocolo eficiente para recuperação desta enzima a partir de corpos de inclusão e realizar ensaios de cristalografia de raios-X e determinação da sua estrutura. A BXil foi expressa em diferentes cepas bacterianas e submetida a diferentes métodos de extração e renaturação proteica sob diferentes condições físico-químicas. Após inúmeros ensaios realizados foi possível recuperar a BXil expressa em corpos de inclusão a partir de um protocolo otimizado de extração desnaturante utilizando ureia (2 M e 6 M) e guanidina (3 M) e posterior renaturação *in vitro*. A amostra renaturada e purificada foi então submetida a ensaios de cristalização, no entanto, não foi observada a formação de cristais em nenhuma das condições iniciais ensaiadas. A ausência de cristais para BXil nas condições avaliadas pode estar relacionada às características intrínsecas da molécula, ou mesmo à homogeneidade, pureza e concentração da amostra. Apesar dos métodos ensaiados não terem sido suficientes para a produção de cristais para a proteína de interesse, nossas abordagens foram efetivas para recuperação da enzima na forma solúvel e ativa, resultado obtido após uma varredura sistemática de análises utilizando diferentes combinações de métodos, o que pode ser útil para auxiliar a obtenção de outras enzimas igualmente expressas em corpos de inclusão.

Palavras-chave: expressão heteróloga, extração desnaturante, proteínas insolúveis

1. INTRODUÇÃO

A biologia estrutural é a área de investigação de estruturas e sistemas biológicos a nível molecular e atômico, aplicada em vários setores do conhecimento. Obter as informações estruturais de enzimas permite o entendimento das relações estrutura-função e de seus mecanismos de reação e, a partir destes conhecimentos, potencializar suas características de acordo com propósitos específicos, favorecendo o avanço biotecnológico nos mais diversos setores.

As informações estruturais obtidas são depositadas em bancos de dados específicos. O banco de dados RCSB PDB (<https://www.rcsb.org/>, Berman et al., 2000), por exemplo, armazena milhares de dados estruturais de proteínas e outras macromoléculas biológicas relacionadas a agricultura e a biomedicina. Esse banco é de domínio público e, seus dados, são obtidos a partir da difração de raios-X, ressonância magnética nuclear e crio-microscopia eletrônica, possuindo, atualmente, 171588 estruturas depositadas.

Atualmente, a cristalografia de raios-X é o método mais utilizado para determinação de estruturas a partir de um cristal formado pela macromolécula alvo. No entanto, o processo de cristalização não é uma tarefa simples e fácil. Inúmeros fatores podem influenciar a formação de cristais como pH, temperatura, força iônica, composição da solução de cristalização, íons metálicos, inibidores e cofatores (Alexander McPherson, 2017). Além disso, o sucesso da coleta de dados e da imagem estrutural final depende da qualidade dos cristais (Alexander McPherson, 2017).

O primeiro passo é a expressão e a purificação da proteína alvo e, geralmente, este processo é feito utilizando hospedeiros heterólogos, como a *Escherichia coli*. Este organismo é amplamente utilizado para expressão de proteínas recombinantes e apresenta protocolos bem estabelecidos, é de crescimento rápido e alta densidade celular quando comparado a outros sistemas de expressão (Rosano e Ceccarelli, 2014; Kaur et al., 2018). Contudo, durante o processo de expressão frequentemente ocorre a formação de agregados proteicos chamados corpos de inclusão, que afetam diretamente o rendimento de proteína solúvel (Goldberg et al., 1996; Cabrita e Bottomley, 2004). A expressão de proteínas na fração insolúvel ainda é um fator limitante para a biologia estrutural.

Várias alternativas para melhorar a expressão da proteína na fração solúvel em *E. coli* estão disponíveis, tais como alterar o tipo de promotor e/ou vetor de expressão para um com baixo número de cópias; reduzir a quantidade de indutor e, consequentemente, a quantidade da proteína expressa; reduzir a temperatura de expressão; utilizar meios de cultivo que não necessitam da adição de indutores químicos; adicionar etiquetas de solubilização fusionadas às sequências de proteínas; e utilizar diferentes cepas de *E. coli* desenhadas para aumentar a solubilidade de proteínas difíceis de expressar e proteínas de membrana, tais como as cepas C41 (DE3) e Lemo21 (DE3).

Outra possibilidade é realizar o truncamento incremental aleatório da sequência do gene alvo, o qual gera uma biblioteca de milhares de construções possíveis que são rastreadas a fim de detectar uma construção solúvel (Hart e Tarendeau, 2006). Esta plataforma, denominada ESPRIT (“Expression of Soluble Proteins by Random Incremental Truncation”), utiliza a Exonuclease III para truncar sequencialmente as extremidades do gene alvo de maneira unidirecional, a partir da extremidade 5’ ou da extremidade 3’, ou bidirecional, o que permite gerar milhares de construções variadas (Yumerefendi et al., 2010; Mas e Hart, 2017).

Mesmo todas estas possibilidades ainda podem ser ineficientes em alguns casos, sendo assim, recuperar a proteína a partir de corpos de inclusão pode ser a escolha mais adequada em alguns casos. A recuperação da proteína expressa em corpos de inclusão é realizada através da solubilização destes agregados utilizando altas concentrações de agentes caotrópicos, como ureia e guanidina. Durante este processo a proteína sofre desnaturação perdendo seu arranjo tridimensional, se fazendo necessário uma etapa de renaturação.

Apesar da existência de inúmeros protocolos disponíveis para solubilização de corpos de inclusão e renaturação proteica, o sucesso desta técnica depende diretamente das características intrínsecas de cada proteína e dos recursos disponíveis, uma vez que requer grandes quantidades de reagentes em todas as etapas. Além disso, esses procedimentos são em sua grande maioria morosos e complexos, de forma que a escolha dos métodos e a execução sistemática de cada abordagem visando encontrar a melhor opção para cada tipo de amostra requer inúmeras etapas de testes e validações.

Diante disso, este trabalho teve como objetivo avaliar de forma sistemática diferentes protocolos de expressão e extração da β -xilosidase recombinante, denominada BXil, anteriormente expressa em corpos de inclusão, sob o intuito inicial de obtenção de amostras em concentração suficiente ($\geq 10 \text{ mg/mL}^{-1}$) para realização de ensaios de cristalografia de raios-X para determinação da estrutura tridimensional.

2. MATERIAL E MÉTODOS

2.1. Origem da ORF CB_612456 codificadora de BXil

A bactéria gram-negativa *Chitinophaga* sp. CB10 foi isolada de um consórcio microbiano proveniente de uma pilha de bagaço de cana-de-açúcar residual da produção de etanol, tendo seu genoma parcial depositado no NCBI GenBank sob o número de acesso MLAV000000000 (Kishi et al., 2017). Considerando o ambiente de isolamento, o referido consórcio microbiano foi capaz de degradar carboximetilcelulose (CMC) motivando a anotação de genes relacionados à degradação de materiais lignocelulósicos (Funnicelli, 2018). Neste sentido, foi realizada a prospecção *in silico* de genes relacionados a degradação de biomassa a partir do genoma da *Chitinophaga* sp. CB10, levando em consideração os motivos e domínios conservados de Glicosil Hidrolases (GHs), em especial β -xilosidases.

2.2. Clonagem molecular da ORF CB_612456

O desenho dos “primers” específicos para o gene da ORF CB_612456 codificador da BXil foi realizado manualmente e levou em consideração os sítios de restrição das enzimas escolhidas (EcoRI e HindIII) e o sítio de clonagem do vetor de expressão pET-28a(+) (Tabela 1). Os “primers” foram avaliados com a ferramenta online PCR Primer Stats (“Sequence Manipulation Suits”, www.bioinformatics.org), a fim de verificar a porcentagem de conteúdo GC, temperatura de “melting” e a possível formação de estruturas secundárias.

Estes “primers” foram sintetizados na região imediatamente externa ao gene, incluindo sítios de reconhecimento para enzimas de restrição EcoRI

(oligonucleotídeos iniciadores “sense”) e HindIII (oligonucleotídeos iniciadores “antisense”), a fim de possibilitar a clonagem direcionada em vetor de expressão. Também foram desenhados primers deletando-se os nucleotídeos que representam o peptídeo sinal identificado e o códon de parada, a fim de se obter uma enzima com 2 caudas de histidina (6xHis), uma no N e outra no C-terminal. A simulação da clonagem da proteína recombinante foi realizada no software SnapGene (snapgene.com).

Tabela 1. Oligonucleotídeos utilizados na clonagem da β -xilosidase

Endonuclease de restrição	Sequência 5' – 3'	Descrição	CG %	Tm ° C	Tamanho do fragmento (pb)
F_EcoRI	TATAGAATTCATGAACCGGTTGCTGCGTGTA	Completo	41.18	62	2172
R_HindIII	TATATAAGCTTCAGCGCACGGGATCAATTC		43.33	60	
R_HindIII_stop	ATATATAAGCTTGCGCACGGGATCAATTC	Sem Stop códon	41.38	59	2169
F_EcoRI_SP	GAATTC ACTGCGCTGCTGGCGGCCGTTTCAGACAG	Sem peptídeo sinal	60	70	2061

Sublinhado: Sítio de restrição; Negrito: gene; F: forward; R: reverse.

A amplificação do gene codificador da BXil foi realizada através de reações de PCR com a PCRBIO Taq Mix (PCR Biosystems, Londres, Reino Unido). O produto desta amplificação foi digerido com as mesmas enzimas de restrição do vetor, EcoRI e HindIII, e quantificado em aparelho NanoDrop 1000 (Thermo Fisher Scientific) de acordo com as instruções do fabricante.

O DNA plasmidial foi extraído utilizando o kit “Wizard® Plus SV Minipreps DNA Purification System (Promega, Madison, Wisconsin, EUA)”, segundo especificações do fabricante, quantificado em aparelho NanoDrop e visualizado por eletroforese em gel de agarose 1% (p/v) contendo brometo de etídeo (0,5 mg/μL). A imagem do gel foi fotodocumentada em Gel Doc 100 (através do software Image Lab version 6.0.1 Bio-Rad, Hercules, Califórnia, Estados Unidos da América). Em seguida, o DNA plasmidial foi submetido a uma reação de restrição com as enzimas específicas (EcoRI e HindIII) e desfosforilação com a FastAP “Thermosensitive Alkaline Phosphatase” (Thermo Fisher Scientific). Tanto o produto de PCR quanto o DNA plasmidial foram purificados em colunas utilizando o kit “DNA Clean & Concentrator” (Zymo Research, Irvine,

Califórnia, EUA) seguindo as orientações do fabricante e, posteriormente, submetidos à reação de ligação com a enzima DNA ligase T4 (Thermo Fisher Scientific) a 16 °C durante 16 horas.

O vetor recombinante pET28a-BXil foi transformado por meio de choque térmico em células competentes de *E. coli* BL21 (DE3) (Novagen) desenvolvidas para clonagem e expressão de proteínas recombinantes. A confirmação da clonagem foi realizada a partir de PCR de colônias, bem como por sequenciamento do vetor recombinante pelo Sanger ABI 3130 XL (Applied Biosystems®) de acordo com as instruções do fabricante, utilizando os “primers” de sequenciamento do vetor (T7 “forward” e T7 “reverse”). Os vetores recombinantes contendo o gene de interesse (clones) foram coletados, adicionados em meio LB contendo o antibiótico canamicina e cultivados a 37 °C durante 16 horas. Após esse período, uma alíquota de cada clone foi coletada (800uL) e adicionada em tubos contendo glicerol (20% v/v), congeladas em nitrogênio líquido e armazenadas a -80 °C.

2.3. Expressão da proteína recombinante

A partir do estoque em glicerol as células de *E. coli* BL21 (DE3) (Novagen) contendo os vetores recombinantes foram estriadas em meio LB sólido suplementado com canamicina a 50 µg/mL⁻¹ e incubadas a 37 °C durante 16 horas. Uma colônia isolada foi inoculada em 50 mL de meio LB contendo antibiótico, e deixada sob agitação orbital constante (200 rpm) a 37 °C durante 16 horas. Uma alíquota de 5 mL foi adicionada em 500 mL de meio LB contendo o antibiótico apropriado e deixado sob agitação orbital (200 rpm) a 37 °C até atingir a fase logarítmica de crescimento, com D.O₆₀₀ entre 0,4 e 0,6 (exceto quando indicado). A seguir, a cultura foi induzida com diferentes concentrações (0,1 a 1 mM) de isopropil-β-D-tiogalactopiranosídeo (IPTG) e deixada sob agitação por 22 horas a 37 °C, com coleta de alíquotas a cada duas horas durante as 8 horas iniciais. Após as 22 horas de incubação, a última alíquota foi coletada e a cultura centrifugada a 9000 x g por 30 minutos a 4 °C. A confirmação da expressão da BXil nos diferentes tempos coletados foi realizada através da eletroforese desnaturante em gel de poliacrilamida 7% - SDS-Page (Laemmli, 1970). A eletroforese foi realizada com tampão Tris-Glicina [50 mM de Tris-HCl; 150 mM de

Glicina e 0,1% de SDS (p/v)]. O gel foi corado com “Coomassie Brilliant Blue” [40% de metanol (v/v); 10% de ácido acético glacial (v/v) e Coomassie R-250 a 0,1% (p/v)] e descorado em solução contendo 10 % de ácido acético (v/v) e 10% de etanol (v/v).

A expressão da proteína recombinante em BL21 (DE3) (Novagen) também ocorreu nas temperaturas de 20 °C, 25 °C e 30 °C, com diferentes concentrações de IPTG (0.1 mM a 1 mM) em meio LB, meio mineral M9 (Geerlof, 2010) e meio de autoindução (ZYP-5052) (Studier, 2005) (Apêndice B), a fim de verificar a influência dos diferentes meios de cultivo sob a expressão da proteína.

Os vetores recombinantes também foram inseridos em outras cepas de *E. coli* através de choque térmico na tentativa de expressar a proteína na fração solúvel. As cepas utilizadas foram Arctic Express (Agilent Technologies, Califórnia, EUA), Lemo21 (DE3) (expressa com 0, 250, 500 e 1000 µM de ramnose) (NEB, Ipswich, Massachusetts, EUA); Rosetta (DE3) (Novagen); BL21-Gold (DE3) (Agilent Technologies), Tuner (DE3) (Novagen), C41 (DE3), C43 (DE3) (Lucigen, Wisconsin, Estados Unidos) em diferentes temperaturas (10° C para Arctic Express, 20°C e 37°C para Tuner e C43, 30 °C e 37 °C para as demais) nas concentrações de IPTG de 0,1 mM e 1 mM (Apêndice B).

2.4. Extração nativa

A tentativa de extração da proteína na forma nativa foi realizada a partir de três clones, o C1, C4 e C10, sendo cada um representante de uma construção distinta: o clone C1, contendo o gene completo contendo uma extensão de 6 histidinas (6xHis) no N-terminal; o clone C10, contendo duas extensões 6xHis: uma no N- e outra no C-terminal a fim de aumentar a superfície de afinidade da proteína com a coluna de purificação; e o clone C4 com uma cauda 6xHis no N-terminal e peptídeo sinal removido. Para tanto, iniciou-se pela ressuspensão do “pellet” em tampão de extração específico (Apêndice B) contendo lisozima. As amostras foram incubadas em banho de gelo durante 30-60 minutos, submetidas à lise celular e centrifugadas por 20 minutos a 15000 x g a 4 °C, para separação de extratos solúveis e insolúveis. O “pellet” foi novamente ressuspensionado em tampão de extração com adição de detergente Triton X-100 nas concentrações de 0.25% a 1% (v/v), lisado e centrifugado

nas mesmas condições anteriores (denominado segundo extrato). Tanto o “pellet” quanto o sobrenadante resultante de cada etapa foram analisados por SDS-Page. A lise celular também foi realizada com 3 a 5 passagens do extrato proteico em um aparelho French Press (Thermo Fisher Scientific) (1000 psi). Outras condições também foram avaliadas e estão todas detalhadamente descritas no Apêndice B.

2.5. Extração desnaturante

O método desnaturante utilizado para solubilização da BXil contém diferentes concentrações de ureia (2 M e 6 M) e cloreto de guanidina 3M. Estes compostos atuam como agentes caotrópicos a fim de solubilizar os corpos de inclusão (CIs) e obter a proteína solúvel na forma desnaturada, sendo necessária posterior renaturação. A seguir, estão descritos o método de extração desnaturante e de renaturação proteica utilizados (adaptado de Kim e colaboradores, 2015).

2.5.1. Lise celular

O clone C1 foi selecionado para obtenção da BXil através de extração desnaturante. O estoque em glicerol do clone C1 foi estriado em meio LB sólido contendo canamicina ($50 \mu\text{g/mL}^{-1}$) e incubado a 37°C por 16 horas. Para o pré-inóculo, uma colônia isolada foi coletada e cultivada a 37°C em meio LB líquido durante 16 horas; em seguida, uma alíquota (10% v/v) foi adicionada em um novo meio LB líquido e cultivada a 37°C até atingir a DO_{600} de 0,4-0,6. Em seguida, a indução ocorreu com 0,1 mM de IPTG durante 6 horas.

As células foram centrifugadas a $15000 \times g$, 4°C durante 20 minutos e o “pellet” foi ressuspendido em tampão de extração contendo Tris HCl 50 mM pH 8, NaCl 500 mM, MgCl 5 mM, Imidazol 20 mM, 1 pastilha de inibidor de proteases (Pierce Protease Inhibitor Cocktails – Thermo Fisher Scientific) e DNase I ($20 \mu\text{g/mL}^{-1}$), incubado durante 30-60 minutos no gelo, homogeneizado e lisado em French Press (Thermo Fisher Scientific) a 1000 psi, com 5 passagens do extrato. O material lisado foi centrifugado a $24000 \times g$, 10 minutos, 4°C . O sobrenadante foi armazenado e o “pellet” foi submetido a duas lavagens, conforme descrito a seguir.

2.5.2. Lavagem celular e solubilização dos corpos de inclusão

Na primeira etapa de lavagem o “pellet” foi ressuspendido no tampão A (Tris HCl 20 mM, NaCl 500 mM, Ureia 2M e Triton X-100 2% v/v) e homogeneizado de modo que não restassem partes agregadas residuais dos corpos de inclusão. O material foi centrifugado a 24000 x g, 30 minutos, 4 °C, e o sobrenadante armazenado. Esta etapa foi realizada duas vezes. No final de cada etapa uma amostra de cada sobrenadante e o “pellet” final foram coletados para análise em SDS-Page.

A terceira lavagem ocorreu através da ressuspensão do “pellet” resultante das lavagens anteriores em tampão B (Tris HCl 20 mM, NaCl 500 mM, Ureia 6M, Imidazol 20 mM e DTT 2 mM), mantido sob agitação durante 30-60 minutos em temperatura ambiente. Em seguida, a solução foi centrifugada a 24000 xg por 10 minutos, a fim de remover o material insolúvel. O “pellet” foi armazenado e o sobrenadante contendo a proteína solubilizada e desnaturada foi submetido à renaturação. Quando necessário, a proteína solubilizada foi filtrada em filtro de 45 µm (Merck MilliPore) antes de seguir para a etapa de renaturação.

2.6. Renaturação *in vitro*

A renaturação proteica foi realizada por diluição de gota (1:20) em tampão Tris HCl 20 mM pH 8, NaCl 500 mM, imidazol 20 mM, glutathione reduzida 4 mM, glutathione oxidada 1 mM e L-arginina 500 mM. Durante esta etapa, a proteína desnaturada foi lentamente adicionada ao tampão de renaturação, gota-a-gota, durante o período de uma hora, a 4 °C com a solução de renaturação sob agitação. A bomba peristáltica P1 (GE Healthcare) (Figura 1) utilizada para adicionar a proteína gota-a-gota foi configurada com fluxo baixo e contínuo, de aproximadamente 10 mL/h.



Figura 1. Imagem ilustrativa da bomba peristáltica P1 utilizada para adicionar a proteína à solução de renaturação (Fonte: Sinapse Biotecnologia).

2.7. Purificação proteica

A proteína renaturada foi concentrada em tubo Amicon (Merck MilliPore) com “cutoff” de massa molecular de 50 kDa e purificada em uma coluna Superdex 200 Increase 10/300 GL (GE Healthcare) conectada a um aparelho de automatização cromatográfica ÄKTA FPLC System (GE Healthcare) equilibrada em tampão Tris HCl pH 9, 200 mM NaCl e glicerol 10% (v/v). As frações eluídas foram coletadas, analisadas em SDS-Page e armazenadas a 4 °C para análises posteriores.

2.8. Determinação da concentração da proteína

A concentração da proteína renaturada e purificada foi calculada a partir da medida de absorção a 280 nm utilizando um espectrofotômetro NanoDrop 1000 (Thermo Fisher Scientific) com base nos valores do coeficiente de extinção molar e na massa molecular do monômero preditos a partir da sequência de aminoácidos,

com auxílio do software ProtParam (“ProtParam tool - ExPASy”, <https://web.expasy.org/protparam/>).

2.9. Teste de detecção de atividade enzimática

Ensaio quantitativo de atividade enzimática em aparelho de absorvância não foram realizados uma vez que a quantidade de proteína renaturada obtida não foi suficiente para a realização destes testes. No entanto, foi realizado um teste qualitativo de atividade enzimática com a proteína renaturada e purificada a fim de verificar o sucesso da renaturação proteica. Para tanto, uma alíquota da enzima foi incubada em 2 mM de *p*-nitrofenil- β -D-xilopiranosídeo (*p*NP-X) a 37 °C durante 30 min e a atividade da β -xilosidase foi identificada pela mudança de coloração da amostra de incolor para amarelo, através da liberação de *p*-nitrofenol, que ocorre devido à reação da enzima com o substrato cromogênico, evidenciando que a renaturação proteica foi eficiente.

2.10. Análise de Espalhamento dinâmico de luz (DLS - Dynamic Light Scattering)

Para avaliar se a renaturação ocorreu adequadamente e monitorar a homogeneidade da amostra e a formação de agregados, uma placa Terasaki (96 poços) foi inicialmente preenchida com óleo de parafina a fim de evitar o contato das amostras com o ar e, em seguida, 1 μ L da proteína renaturada e purificada diluída (0,1 mg/mL⁻¹) foi adicionado à placa. A placa foi analisada no Spectro Light 600 (Xtal Concepts, Hamburg, Alemanha) em um comprimento de onda de 658 nm, a 20 °C com número de medidas igual a 20 e viscosidade de 0,345. Os resultados foram avaliados por meio de gráficos e imagens.

2.11. “Western blotting”

Inicialmente realizou-se uma eletroforese desnaturante em gel de poliacrilamida 10% com as amostras de interesse. Em seguida o gel foi lavado 3 vezes

em H₂O destilada e a membrana PVDF Immun-Blot (BioRad) incubada com metanol durante 5 minutos. O gel, o papel de filtro e a membrana PVDF foram incubados em tampão de transferência (Tris Base 48 mM, Glicina 39 mM, metanol 20% e SDS 10%) durante 10 minutos. A transferência ocorreu em sistema Trans-Blot® SD Semi-dry (BioRad) a 15 V durante 45 minutos. Em seguida, a membrana foi lavada durante 10 minutos por 3 vezes com a solução TBS-T (5% TBS 20x e 0,05% Tween 20). Após a lavagem, a membrana foi incubada durante 1 hora em temperatura ambiente sob agitação constante em uma solução bloqueadora (TBS-T acrescido de 5% (p/v) de leite em pó). Em seguida, a membrana foi incubada com o anticorpo primário anti-His (1:3000 v/v) na solução bloqueadora, sob agitação constante a 4 °C durante 16 horas. A membrana foi então lavada durante 10 minutos por 3 vezes com a solução TBS-T e incubada com o anticorpo secundário “anti-mouse” AP (1:30000 v/v) na solução bloqueadora, sob agitação constante em temperatura ambiente durante 1 hora. A membrana foi novamente lavada com TBS-T e corada com o substrato BCIP/NBT (5-bromo-4-chloro-3-indolyl-phosphate/nitro blue tetrazolium) até o surgimento de bandas de coloração roxa. A reação foi interrompida adicionando H₂O destilada à membrana.

2.12. Ensaio de cristalização

Os experimentos de cristalização foram realizados utilizando o método de difusão de vapor em gota assente em placas de 96 poços preparadas pelo robô de pipetagem Mosquito LCP (SPT Labtech). Os ensaios iniciais foram realizados utilizando os kits comerciais JCSG+ Screen, Structure I & II e Morpheus HT-96 (Molecular Dimensions, Sheffield, Inglaterra). Cada gota (condição) foi formada por 250 nL da proteína (50% v/v) e 250 nL da solução de cristalização (50% v/v) e avaliada em triplicata. As placas foram seladas e mantidas em ambientes com temperatura controlada a 4 °C, 20 °C ou 30 °C. As gotas foram examinadas sob uma lupa estereoscópica diariamente durante 60 dias.

3. RESULTADOS E DISCUSSÃO

3.1. Clonagem molecular da ORF CB_612456

O gene codificador da β -xilosidase de 2172 pb e massa molecular $\sim 81,25$ kDa foi prospectada e anotada no genoma de *Chitinophaga* sp., uma bactéria gram-negativa isolada de um consórcio microbiano oriundo de uma pilha de bagaço de cana-de-açúcar de uma indústria de etanol (Kishi et al., 2017). A enzima foi amplificada através de reações de PCR gerando as construções: pET28-BXil_FL (“primers” F_EcoRI e R_HindIII), pET28-BXil_SC (“primers” F_EcoRI e R_HindIII_stop), e pET28-BXil_SP (“primers” F_EcoRI_SP e R_HindIII) (Figura 2; Tabela 1).

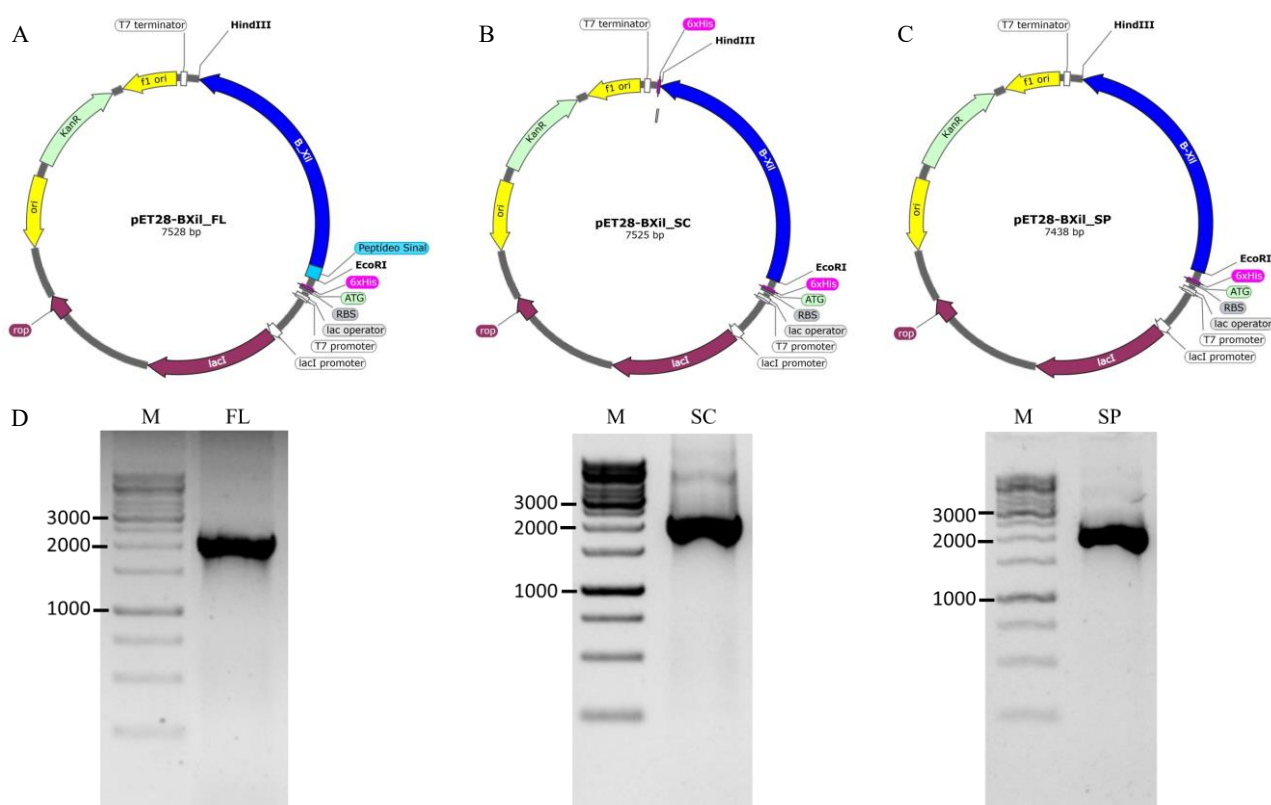


Figura 2. Representação esquemática dos mapas dos vetores recombinantes e reações de amplificação por PCR do gene codificador da β -xilosidase obtidos a partir de diferentes combinações de primers. A: Vetor recombinante contendo o gene β -xilosidase com a sequência completa (pET28-BXil_FL – Clone C1). B: Vetor recombinante contendo o gene β -xilosidase sem stop códon (6xHis no N- e no C-terminal do gene; pET28-BXil_SC – Clone C10). C: Vetor recombinante contendo o gene β -xilosidase sem peptídeo sinal (pET28-BXil_SP). Figuras elaboradas com o software SnapGene (snapgene.com). D: Reações de amplificação do gene β -xilosidase com 2.172 pb. M: Marcador molecular 1kb plus (Fermentas); FL: clone C1

sequência completa (“full length”); SC: clone C10 sem stop códon; SP: clone C4 sem peptídeo sinal.

3.2. Expressão e extração nativa

A análise de expressão da proteína recombinante em *E. coli* ocorreu em diferentes condições de temperatura (18 °C a 37 °C), diferentes concentrações de IPTG (0,03 mM a 1 mM), em meio LB, meio mineral M9 (Geerlof, 2010) e meio de auto indução (Studier, 2005), a fim de verificar a influência de cada componente dos meios de cultivo sob a expressão da proteína recombinante.

Inicialmente, os vetores recombinantes (clones) obtidos a partir das diferentes combinações de “primers” (Figura 2; Tabela 1) foram inseridos em células competentes *E. coli* BL21 (DE3) e submetidos a ensaios de expressão. Quando inseridos em *E. coli* BL21 (DE3), os clones C1, C4 e C10 foram submetidos à diferentes combinações de técnicas de expressão e extração da proteína alvo (Apêndice B).

Inicialmente, o clone C1 (pET28-BXil_FL) foi expresso a 20 °C em meio de auto indução (ZYP-5052) (Studier, 2005) durante 21 horas (dados não mostrados), e a 30 °C e 37 °C em meio LB com 0,1 e 0,2 mM de IPTG durante 6 horas (Figura 3).

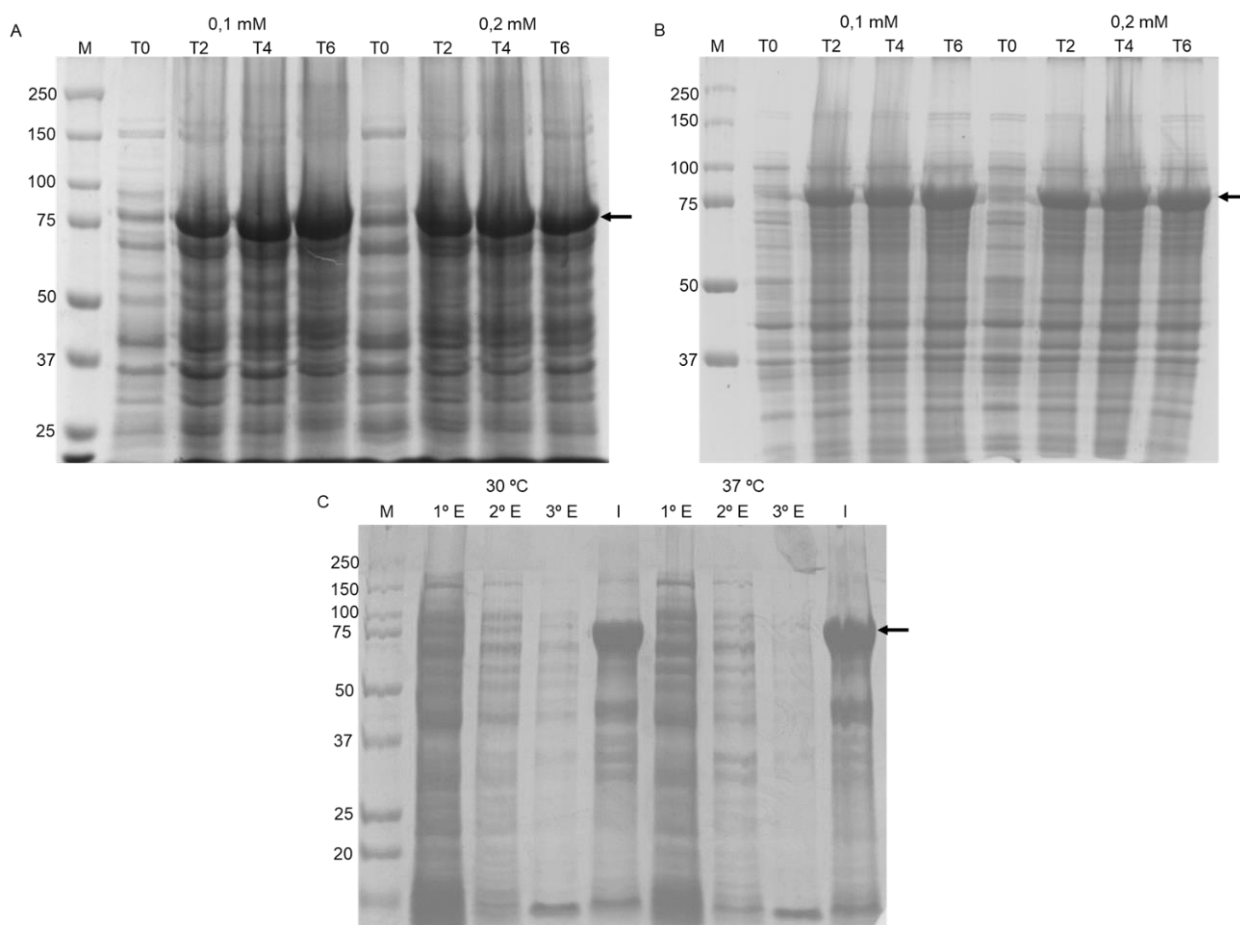


Figura 3. Eletroforese em gel de poliacrilamida (SDS-Page) das análises realizadas em meio LB com 0,1 mM de IPTG a partir do clone C1 (pET28-BXil_FL). A: Ensaio de expressão a 30 °C com 0,1 mM de IPTG e com 0,2 mM de IPTG nos tempos T0, T2, T4 e T6 horas. B: Ensaio de expressão a 37 °C com 0,1 e 0,2 mM de IPTG nos tempos T0, T2, T4 e T6 horas. C: Extração nativa da proteína expressa durante 6 horas com 0,1 mM IPTG. Extração da expressão a 30 °C e a 37 °C; 1º extrato, 2º extrato, 3º extrato e fração insolúvel (I). M: marcador de massa molecular (“Precision Plus Protein Dual Color Standards” – BioRad). A seta indica a BXil de aproximadamente 81,25 kDa.

O clone C4 (pET28-BXil_SP), quando expresso em meio LB a 37 °C e 25 °C, foi induzido com diferentes concentrações de IPTG, variando de 0,03 mM a 1 mM. A menor concentração de IPTG (0,03 mM) não foi capaz de induzir a expressão da proteína nestas condições (dados não mostrados). Além disso, a expressão a 37 °C durante 22 horas resultou na degradação da proteína alvo (Figura 4.B). A tentativa de extração nativa em ambas as temperaturas avaliadas também foi realizada, no entanto, a proteína permaneceu nos corpos de inclusão (Figura 4.C e 4.D).

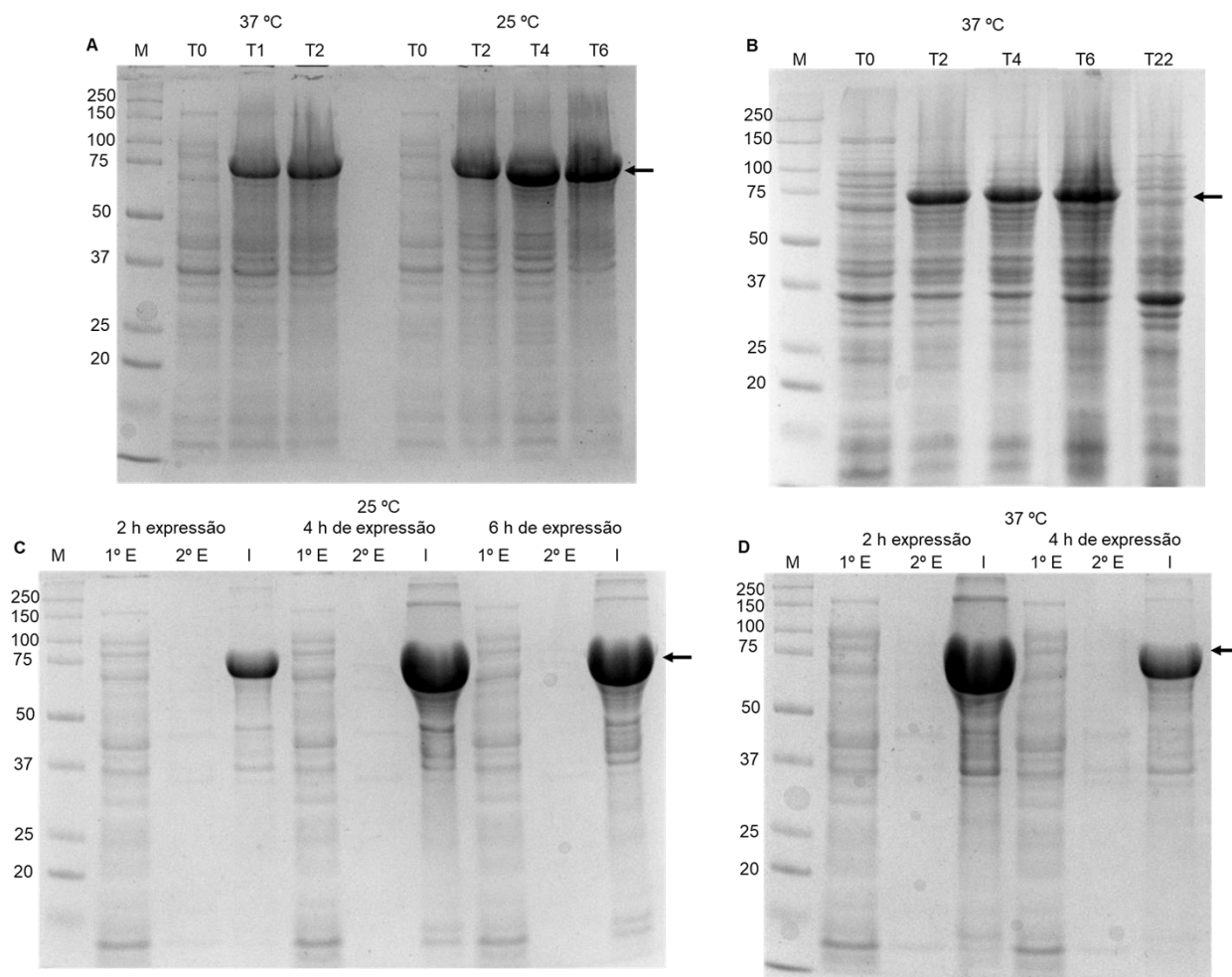


Figura 4. Eletroforese em gel de poliacrilamida (SDS-Page) das análises realizadas em meio LB com 0,1 mM de IPTG a partir do clone C4 (pET28-BXil_SP). A: Ensaio de expressão a 37 °C e a 25 °C em diferentes tempos (T0 a T6 horas) B: Ensaio de expressão em meio LB com 0,1 mM de IPTG a 37 °C em diferentes tempos (T0 a T22 horas); C: Extração nativa da expressão a 25 °C com 2 horas, 4 horas e 6 horas de expressão. D: Extração nativa da expressão a 37 °C com 2 e 4 horas de expressão. (1° extrato, 2° extrato e fração insolúvel (I)). M: marcador de massa molecular ("Precision Plus Protein Dual Color Standards" – BioRad). A seta indica a BXil de aproximadamente 81,25 kDa.

A expressão do clone C4 também ocorreu em meio mineral M9 (Geerlof, 2010) a 37 °C induzido com 0,3 e 0,5 mM de IPTG, sendo que o comportamento da expressão pareceu ser o mesmo em ambas as concentrações (Figura 5.A). O processo de extração nativa foi realizado para ambas as concentrações de indutor, no entanto, a maior parte das proteínas permaneceu na fração insolúvel (Figura 5.B). Nesse momento, houve uma tentativa de purificação da proteína solúvel presente no primeiro extrato através de cromatografia por troca iônica e por cromatografia de

afinidade ao níquel, no entanto, nenhum dos métodos obteve sucesso, pois a maior quantidade de proteínas foi expressa na fração insolúvel.

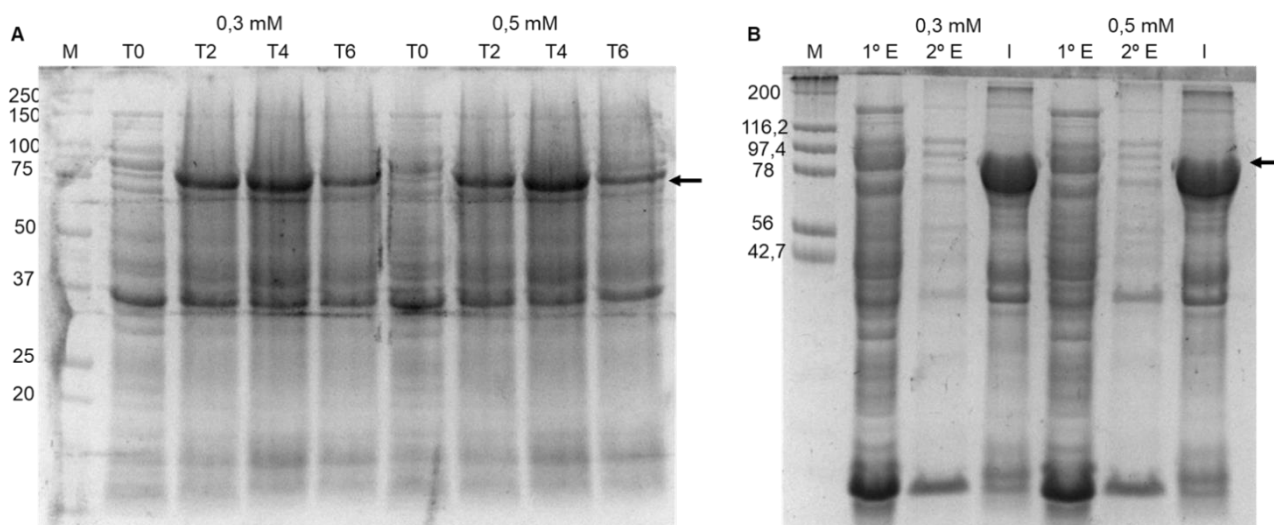


Figura 5. Eletroforese em gel de poliacrilamida (SDS-Page) das análises realizadas a 37 °C em meio M9 com 0,3 e 0,5 mM de IPTG a partir do clone C4. A: Ensaio de expressão com 0,3 mM e com 0,5 mM de IPTG em diferentes tempos (T0 a T6 horas); B: extração nativa da expressão a 37 °C com 0,3 mM e com 0,5 mM de IPTG. 1º extrato, 2º extrato e fração insolúvel (I). M: marcador de massa molecular (“Precision Plus Protein Dual Color Standards” – BioRad). A seta indica a BXil de aproximadamente 81,25 kDa.

O clone C10 (pET28-BXil_SC), por sua vez, foi expresso em meio LB a 30 °C induzido com 0,1 mM de IPTG durante 5 horas e a 37 °C induzido com 0,1 mM a 1 mM de IPTG durante 6 horas; e em meio de auto indução (ZYP- 5052) (Studier, 2005) a 18 °C durante 16 horas (Figura 6.A; Apêndice B). As condições avaliadas sob as temperaturas de 30 e 37 °C com 0,1 e 1 mM de IPTG em meio LB e em meio de auto indução foram submetidas ao processo de extração nativa, novamente sem sucesso (Figura 6.B e 6.C).

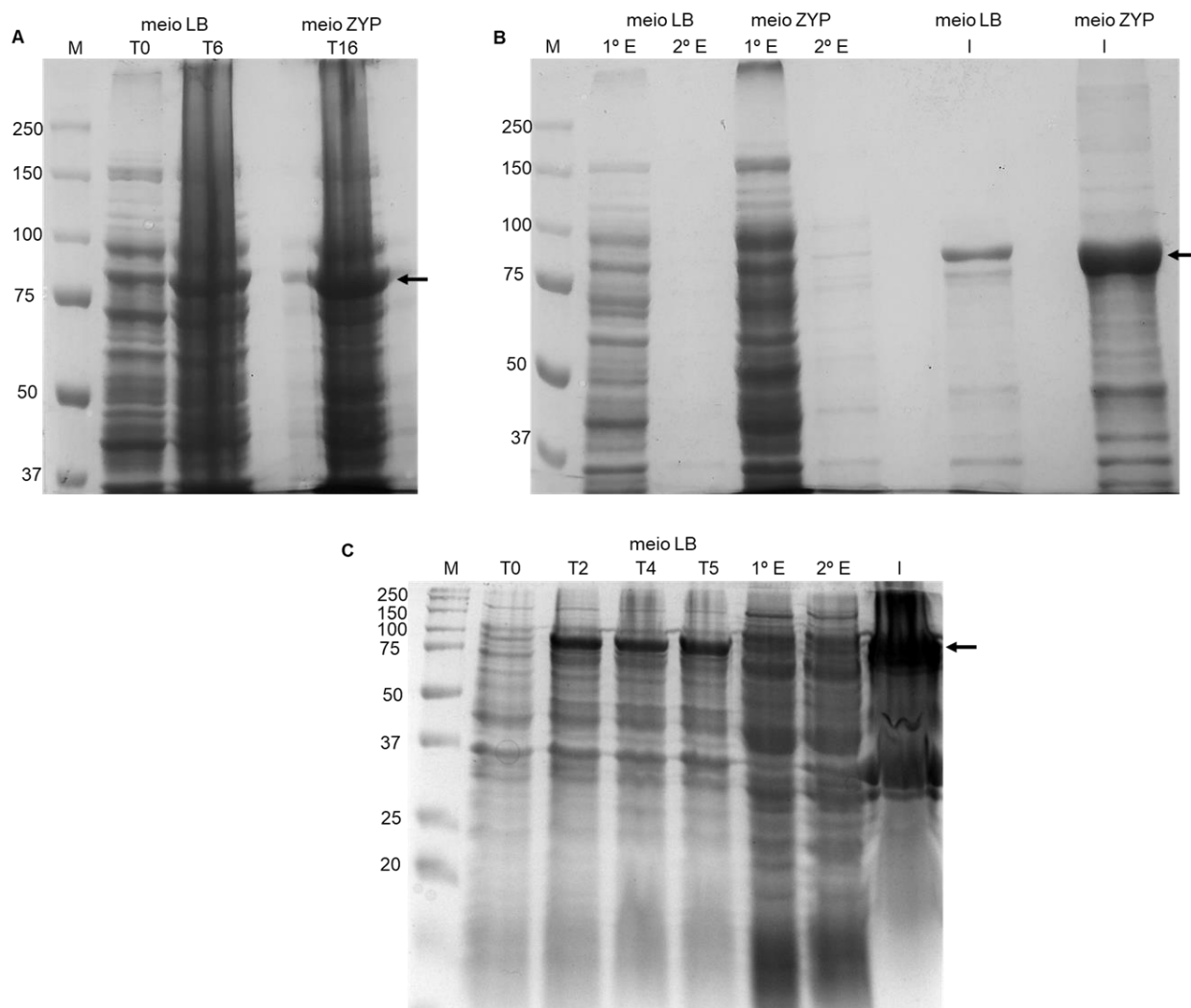


Figura 6. Eletroforese em gel de poliacrilamida (SDS-Page) das análises a partir do clone C10. A: Ensaio de expressão a 37 °C em meio LB e em meio de auto indução (ZYP) com 1 mM de IPTG a 18 °C durante 16 horas. B: Extração nativa meio LB e meio ZYP (1° e 2° extrato, fração insolúvel (I)). C: Expressão em meio LB a 30 °C com 0,1 mM de IPTG (T0 a T5 horas) e extração nativa (1° extrato, 2° extrato e fração insolúvel (I)). M: marcador de massa molecular ("Precision Plus Protein Dual Color Standards" – BioRad). A seta indica a BXil de aproximadamente 81,25 kDa.

A superexpressão de proteínas recombinantes em *E. coli* é o método mais utilizado e de baixo custo benefício, no entanto, tende a formar agregados de proteínas biologicamente inativas chamados corpos de inclusão (Cabrita e Bottomley, 2004; Vallejo e Rinas, 2004; Chura-Chambi et al., 2013).

Estes agregados ocorrem tipicamente a partir da auto associação de algumas proteínas que servem como ponto de nucleação de pequenos agregados solúveis que tendem a aumentar formando agregados maiores e insolúveis, seja por proteínas em

seu estado conformacional nativo ou em estados estruturalmente alterados (Philo e Arakawa, 2009; Lebendiker e Danieli, 2014). Este processo pode ser estimulado de acordo com as propriedades inerentes de cada proteína bem como por outros fatores como temperatura, concentração de proteína, concentração do indutor, nível de expressão do plasmídeo, ausência de chaperonas, dentre outros (Cabrita e Bottomley, 2004; Philo e Arakawa, 2009; Wang et al., 2010; Lebendiker e Danieli, 2014). Essas causas podem ocorrer de forma isolada ou em combinação durante a expressão de diferentes proteínas e, embora inúmeros avanços tenham sido alcançados quanto à melhora na expressão e na solubilidade proteica, não há nenhum protocolo estabelecido para evitar a formação destes agregados em todas as proteínas (Rinas et al., 2017).

Assim sendo, a construção C10 também foi inserida em outras cepas de *E. coli* na tentativa de expressar a proteína na fração solúvel. A expressão em Arctic Express (Agilent Technologies) ocorreu em meio LB a 12 °C sob indução de 1 mM e 3 mM de IPTG durante 24 horas (Figura 7). A Arctic Express foi projetada para aumentar o rendimento de proteínas solúveis em temperaturas de a 4-12 °C. Esta cepa co-expressa as chaperoninas Cpn10 e Cpn60 da bactéria *Oleispira antarctica*, que possuem alta identidade com as chaperoninas GroEL e GroES da *E. coli*, as quais conferem melhor processamento da proteína em baixas temperaturas melhorando o enovelamento e a solubilidade da proteína e, conseqüentemente, prevenindo a agregação.

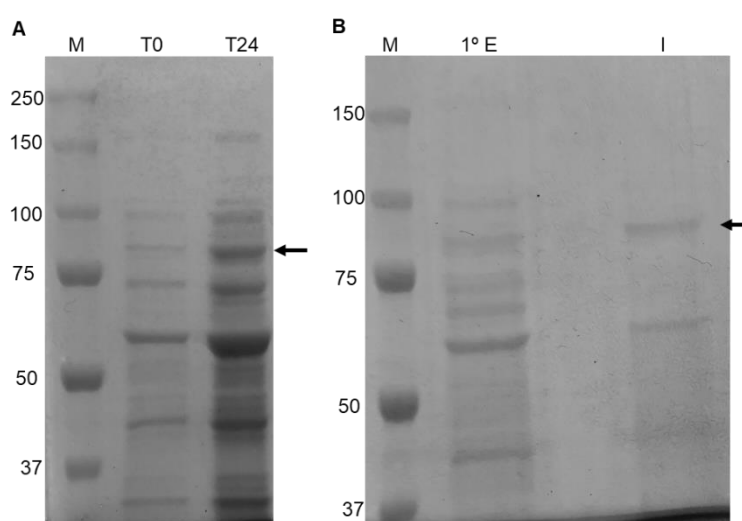


Figura 7. Eletroforese em gel de poliacrilamida (SDS-Page) da construção C10 em Arctic Express. A: Ensaio de expressão em meio LB a 12 °C com 1 mM de IPTG

durante 24 horas (T0 e T24); B: Extração nativa (1º extrato e fração insolúvel (I)). A seta indica a BXil de aproximadamente 81,25 kDa.

As cepas C41 (DE3) e C43 (DE3) (Lucigen) são frequentemente utilizadas para expressão de proteínas tóxicas e de membrana. A expressão da construção C10 em *E. coli* C41 (DE3) ocorreu sob indução de 1 mM de IPTG em meio LB a 37 °C durante 6 horas, e o gel SDS-Page da extração nativa mostrou uma banda na mesma altura da β -xilosidase (~81 kDa) (Figura 8.A e 8.B). A partir deste resultado optou-se por realizar a purificação do extrato por cromatografia de afinidade ao níquel e por exclusão de tamanho (Figura 8.C e 8.D). Sendo assim, a fim de confirmar a presença da enzima foi realizado o sequenciamento N-terminal da amostra através da Degradação de Edman e um ensaio de “Western Blotting”, os quais revelaram que a amostra purificada se tratava de uma proteína de *E. coli*.

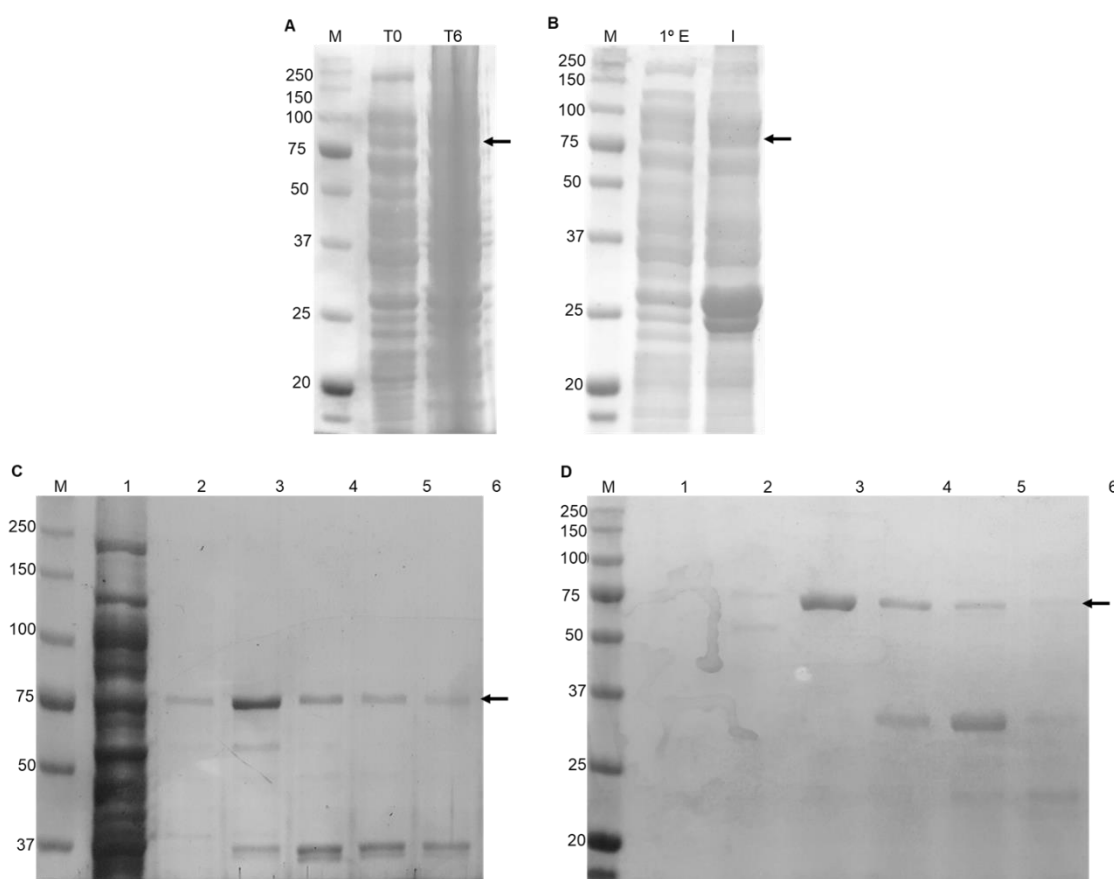


Figura 8. Eletroforese em gel de poliacrilamida (SDS-Page) da construção C10 na *E. coli* C41. A: Ensaio de expressão em meio LB a 37 °C com 1 mM de IPTG (T0 e T6 horas); B: Extração nativa (1º extrato e fração insolúvel (I)); C: Cromatografia foi afinidade ao níquel (IMAC) (Linha 1: flow through; Linhas 2 a 6: amostras purificadas referentes ao pico eluído a 225 mM de Imidazol); D: Cromatografia por exclusão de

tamanho (SEC) (Linhas 1 a 6: amostras eluídas). M: marcador de massa molecular (Precision Plus Protein Dual Color Standards – BioRad). A seta indica a BXil de aproximadamente 81,25 kDa.

A construção C10 também foi expresso em *E. coli* C43 (DE3) (Lucigen) com 1 mM de IPTG a 20 e 37 °C durante 16 horas e, novamente, a maior parte da proteína expressa permaneceu na fração insolúvel (dados não mostrados).

Em Lemo21 (DE3) (NEB) o clone C10 foi expresso a 37 °C com 1 mM de IPTG e com diferentes concentrações de ramnose (0, 250, 500 e 1000 μ M), durante 4 e 6 horas (Figura 9). Lemo21 oferece os recursos da BL21 (DE3) além de permitir a expressão ajustável de proteínas alvo difíceis (proteínas de membrana, proteínas tóxicas e proteínas insolúveis) melhorando o enovelamento e a produção de proteínas solúveis. Isso ocorre através da adição de L-ramnose ao meio de cultivo, que regula o nível de lisozima T7 (T7Lys), o qual é o inibidor natural da atividade da polimerase de RNA T7 (Wagner et al., 2008). Ainda assim, a expressão da proteína alvo nesta cepa em ambas as temperaturas avaliadas, ocorreu somente na ausência de ramnose, que é quando Lemo21 realiza o mesmo que uma cepa contendo o plasmídeo pLysS. O plasmídeo pLysS reduz os níveis de expressão basal do gene alvo, sendo adequado para expressão de proteínas tóxicas.

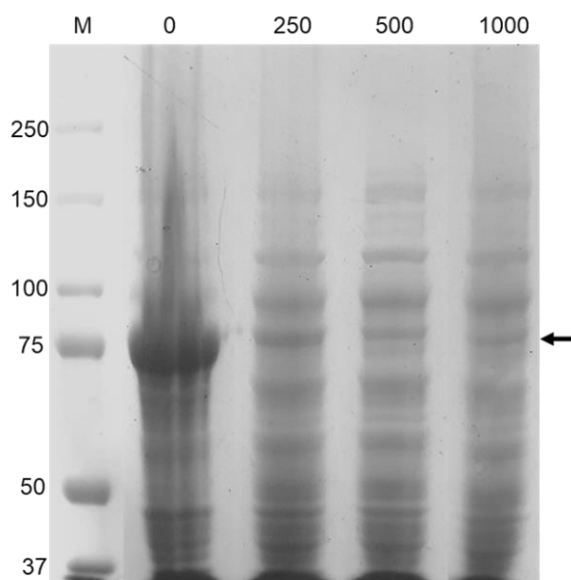


Figura 9. Eletroforese em gel de poliacrilamida (SDS-Page) da expressão da construção C10 em Lemo21 a 37 °C durante 6 horas com diferentes concentrações de ramnose (0 μ M, 250 μ M, 500 μ M e 1000 μ M). M: marcador de massa molecular ("Precision Plus Protein Dual Color Standards" – BioRad). A seta indica a BXil de aproximadamente 81,25 kDa.

O hospedeiro Rosetta (DE3) (Novagen) foi desenvolvido para expressar proteínas eucarióticas que contêm códons raramente usados em *E. coli*, através da presença de um plasmídeo que codifica os códons AGG/AGA (arginina), CGG (arginina), AUA (isoleucina), CUA (leucina), CCC (prolina) e GGA (glicina) nesta cepa. A expressão da construção C10 neste hospedeiro ocorreu em meio LB a 37 °C com 1 mM de IPTG durante 6 horas, sendo que a extração nativa também evidenciou que a proteína foi novamente expressa em corpos de inclusão (Figura 10).

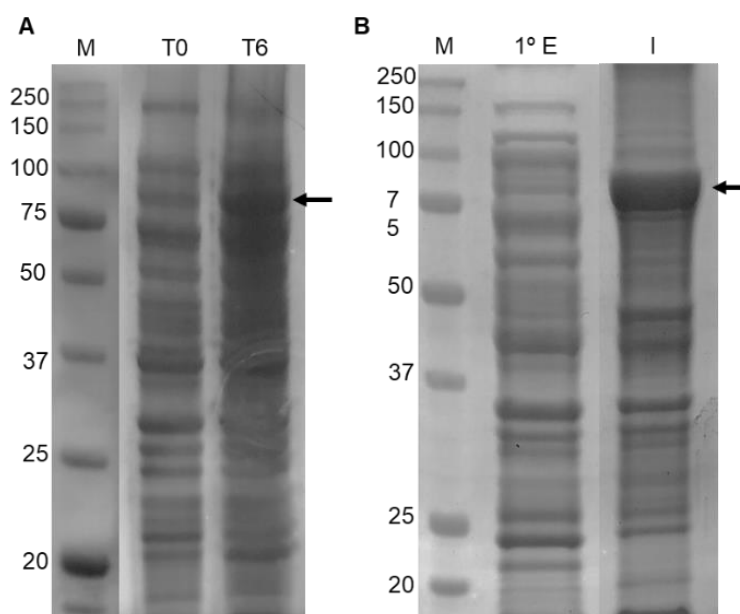


Figura 10. Eletroforese em gel de poliacrilamida (SDS-Page) da expressão da construção C10 em Rosetta. A: Expressão a 37 °C durante 6 horas (T0 e T6); B: Extração nativa (1° extrato e fração insolúvel (I)). M: marcador de massa molecular ("Precision Plus Protein Dual Color Standards" – BioRad). A seta indica a BXil de aproximadamente 81,25 kDa.

As cepas Tuner (DE3) (Novagen) possuem uma mutação de deleção do gene lac permease (lacY) e resultam em um nível de indução homogêneo, permitido através da entrada uniforme do IPTG em todas as células da população. Dependendo da concentração de IPTG, a regulação da expressão ocorre de níveis muito baixos até níveis de total indução, sendo que a expressão em níveis mais baixos é capaz de aumentar a solubilidade e a atividade de proteínas alvo difíceis. O clone C10 nesta cepa foi expresso a 20 e 37 °C com 1 mM de IPTG (Figura 11).

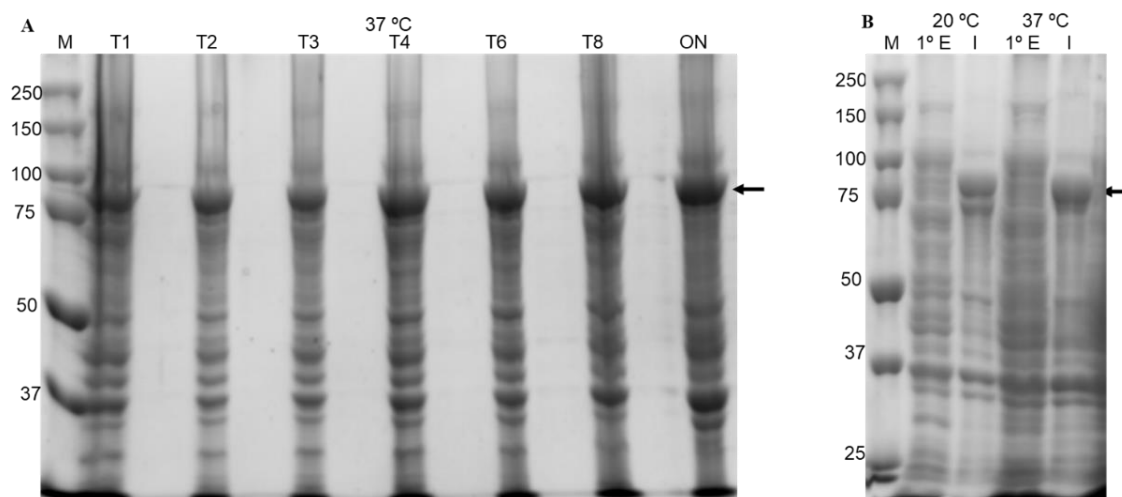


Figura 11. Eletroforese em gel de poliacrilamida (SDS-Page) da expressão da construção C10 em Tuner. A: Expressão a 37 °C com 1 mM de IPTG (T0 a T16 horas). B: Extração nativa das amostras expressas durante 16 horas e das amostras expressas a 20 °C (1° extrato e fração insolúvel (I)). M: marcador de massa molecular (“Precision Plus Protein Dual Color Standards” – BioRad). A seta indica a BXil de aproximadamente 81,25 kDa.

Embora as células competentes BL21-Gold (DE3) (Agilent Technologies) não sejam específicas para melhorar a solubilidade proteica, mas sejam versões melhoradas de *E. coli* BL21, elas também foram utilizadas para expressão da construção C10. Neste caso, a expressão ocorreu a 37 °C com 1 mM de IPTG durante 4 e 6 horas e, da mesma maneira que nas outras cepas utilizadas, a proteína alvo continuou na fração insolúvel (Figura 12).

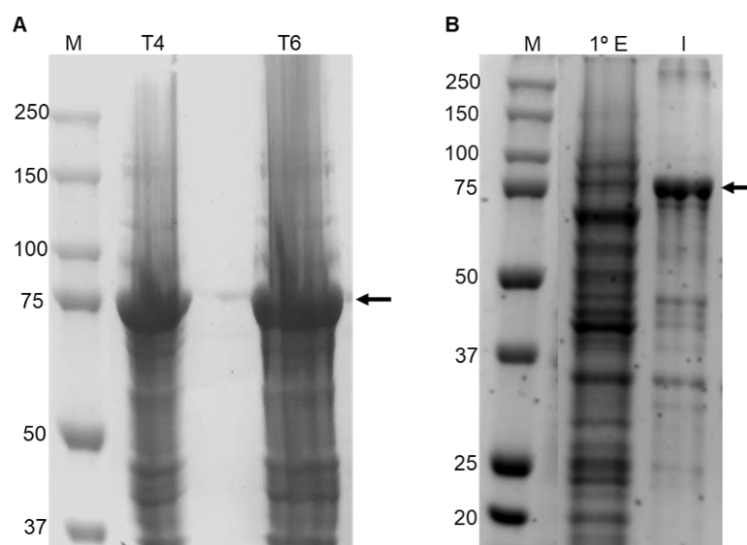


Figura 12. Eletroforese em gel de poliacrilamida (SDS-Page) da expressão da construção C10 em BL21-Gold. A: Expressão a 37 °C com 1 mM IPTG (T4 e T6 horas);

B: Extração nativa (1º extrato e fração insolúvel (I)). M: Marcador de massa molecular ("Precision Plus Protein Dual Color Standards" – BioRad). A seta indica a BxII de aproximadamente 81,25 kDa.

Apesar das inúmeras tentativas de expressão proteína alvo na fração solúvel empregando baixas temperaturas, baixas concentrações de indutor, diferentes linhagens de células competentes com características úteis para aumentar a solubilidade, além de diferentes combinações e métodos de extração (Apêndice B), a maior parte da proteína continuou sendo expressa em corpos de inclusão. Diante disso, optou-se pela extração desnaturante e posterior renaturação da proteína alvo, a partir da construção C1 em *E. coli* BL21 (DE3).

3.3. Extração desnaturante – corpos de inclusão

Quando a proteína alvo é expressa na forma de corpos de inclusão, geralmente, o método de extração da proteína na forma nativa não é capaz de desfazer estes agregados e retirar a proteína da fração insolúvel. Diante disso, recomenda-se realizar a extração desnaturante, que utiliza agentes caotrópicos para solubilizar os corpos de inclusão.

Em alguns casos a formação de corpos de inclusão aparece como vantagem, pois estes agregados são formados por grandes concentrações da proteína praticamente sem contaminantes, além de protegerem-na da degradação proteolítica (Cabrita e Bottomley, 2004; Vallejo e Rinas, 2004). No entanto, o rendimento da proteína após renaturação não é alto, portanto, do ponto de vista industrial e para cristalografia, esta condição pode tornar o processo de obtenção da enzima mais complexo e oneroso e, ainda assim, não ter êxito.

A recuperação das proteínas agregadas é realizada através do isolamento e solubilização dos corpos de inclusão por meio de agentes desnaturantes e posterior renaturação *in vitro*. Os agentes desnaturantes mais comumente utilizados são ureia e cloreto de guanidina, que desnaturam completamente a estrutura terciária da proteína, ou por métodos mais suaves que podem preservar sua estrutura nativa. Os métodos mais brandos envolvem o uso de baixas concentrações de desnaturantes, detergentes ou arginina (Vallejo e Rinas, 2004).

Enquanto o êxito para a solubilização de corpos de inclusão é mais simples e requer somente um agente caotrópico e um agente redutor, o sucesso para a correta renaturação proteica depende de várias condições físicas e químicas que envolvem as propriedades intrínsecas de cada proteína e, portanto, não há regras definidas ou um protocolo pré-determinado, mas é necessário estabelecê-lo através de tentativa e erro (Goldberg et al., 1996; Cabrita e Bottomley, 2004; Rinas et al., 2017).

Outros métodos de extração desnaturante também foram avaliados para solubilização dos corpos de inclusão e renaturação *in vitro*, no entanto, a proteína não apresentou qualquer atividade contra o substrato cromogênico (dados não mostrados). Deste modo, neste estudo o protocolo escolhido foi adaptado de Kim e colaboradores (2015), levando em consideração inúmeros testes de tentativa e erro, bem como a disponibilidade de equipamentos e material (Figura 13).

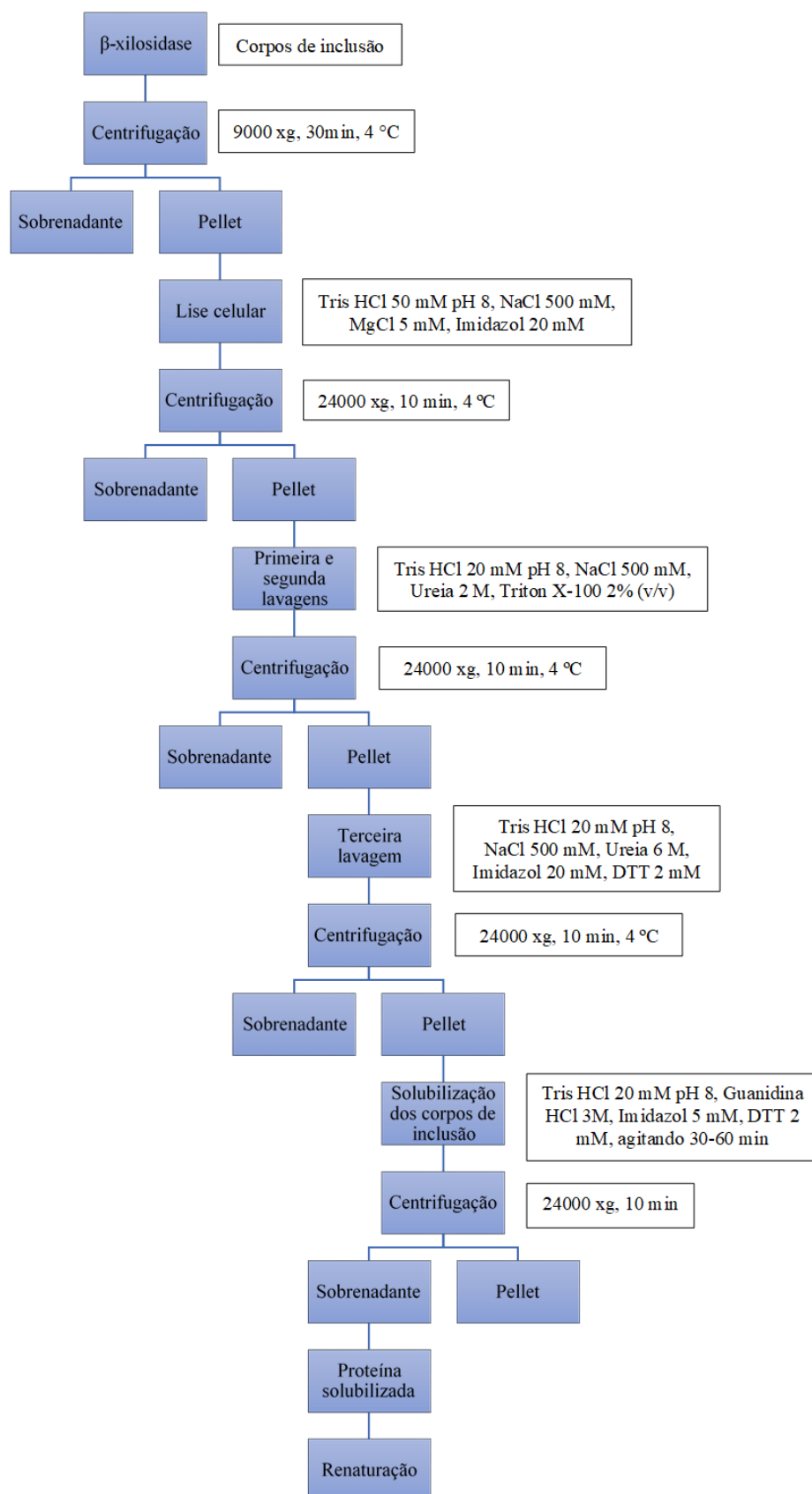


Figura 13. Esquema para obtenção de proteína solúvel a partir de corpos de inclusão utilizando agentes desnaturantes.

Após a lise celular e centrifugação, a fração insolúvel contendo os corpos de inclusão foi submetida à sucessivas lavagens a fim de remover contaminantes celulares, como proteínas da membrana externa. Para isso, o tampão utilizado durante as lavagens continha detergente Triton X-100 e ureia em baixas concentrações. Depois da lavagem, os corpos de inclusão foram solubilizados utilizando agentes caotrópicos, geralmente 4-6 M de cloreto de guanidina ou 8 M de ureia, sendo que a guanidina é mais utilizada devido à sua ação caotrópica mais forte (Cabrita e Bottomley, 2004; Vallejo e Rinas, 2004). Além disso, a presença de DTT (Ditiotreitol) durante a solubilização é necessária para manter os resíduos de cisteína no estado reduzido e impedir eventuais ligações dissulfeto, que interferem na solubilidade (Clark, 2001).

Durante a etapa de solubilização a solução permaneceu sob agitação lenta durante 30-60 minutos a fim de permitir o desdobramento completo da estrutura da proteína. Em alguns casos, quando a solubilização é mais difícil, é necessário aumentar a temperatura nesta fase, geralmente para 60 °C.

Após o término da solubilização a solução foi centrifugada e as frações coletadas, sendo que o sobrenadante final continha a proteína solubilizada. Esta, por sua vez, foi filtrada quando necessário e submetida à renaturação *in vitro* (Figura 14).

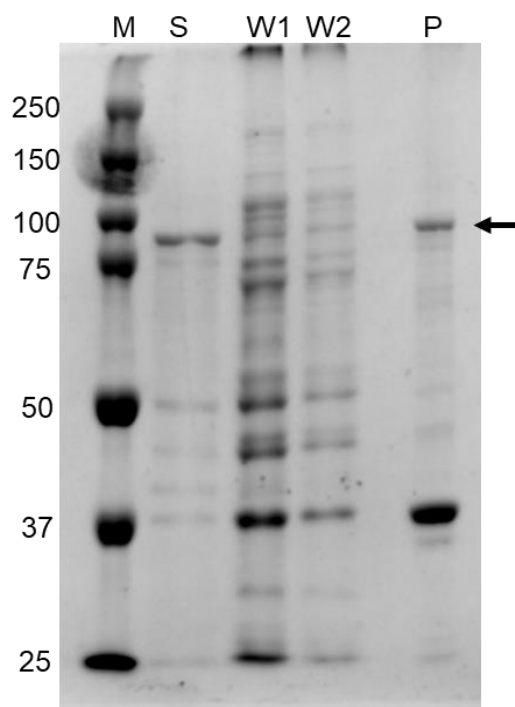


Figura 14. Eletroforese em gel de poliacrilamida (SDS-Page) da extração desnaturante da proteína alvo (clone C1). M: marcador de massa molecular; S:

proteína solubilizada; W1: sobrenadante coletado após a segunda lavagem; W2: sobrenadante coletado após terceira lavagem; P: “pellet” final (fração insolúvel). A seta indica a BxII de aproximadamente 81,25 kDa.

3.4. Renaturação *in vitro*

As proteínas provenientes da solubilização estão desprovidas da sua conformação nativa e devem ser submetidas a renaturação para que adquiram novamente a sua forma correta. Para isso, as proteínas devem ser submetidas à renaturação *in vitro*, que pode ser alcançada através de diálise, diluição gota a gota, cromatografia de afinidade ou filtração em gel. Neste estudo utilizou-se o método de diluição gota a gota, o qual é o mais utilizado em estudos de pequena escala por ser simples e não necessitar de equipamentos de alto custo.

O processo de renaturação tem o objetivo de remover os agentes desnaturantes e envolve uma solução contendo diferentes aditivos que permitem a correta renaturação e minimizam a agregação, os quais geram perda da proteína. Cada proteína possui características inerentes que devem ser consideradas durante a escolha da composição da solução renaturante. No entanto, embora os princípios de renaturação e agregação sejam amplamente discutidos na literatura, encontrar uma condição adequada para a renaturação de uma proteína específica pode levar tempo e nem sempre tem o rendimento mínimo esperado. Isto pode ser explicado pela competição que ocorre entre os estados de renaturação correta com a dobragem incorreta e a agregação da proteína alvo, sendo que, geralmente, a via da renaturação correta encontra-se em desvantagem (Cabrita e Bottomley, 2004; Vallejo e Rinas, 2004). Além disso, inúmeros fatores influenciam o sucesso da renaturação, como a concentração final de proteína a ser renaturada, baixas temperaturas, componentes do tampão e método de renaturação utilizados (Goldberg et al., 1996; Cabrita e Bottomley, 2004; Lebendiker e Danieli, 2014). Por fim, quando a correta renaturação é alcançada geralmente obtemos proteínas biologicamente ativas.

A composição do tampão de renaturação varia amplamente e depende diretamente das características de cada proteína. Geralmente se utiliza tampão Tris em pH neutro ou levemente alcalino, além de diversos aditivos com diferentes finalidades, incluindo detergentes, glicerol, NDSB, arginina, PEG 3350, ureia e

guanidina em baixas concentrações, NaCl, MgCl₂, EDTA, glutathione (GSH/GSSG), cisteína/cistina, DTT, dentre outros (Goldberg et al., 1996; Clark et al., 1999; Cabrita e Bottomley, 2004; Lebendiker e Danieli, 2014), os quais foram testados em diferentes combinações e concentrações até encontrar uma solução de renaturação adequada à proteína de interesse.

Após inúmeros testes, o tampão de renaturação final era composto por Tris HCl pH8, NaCl 500 mM, Imidazol 20 mM, glutathione reduzida 4 mM (GSH), glutathione oxidada 1 mM (GSSG) e L-arginina 500 mM. A composição do tampão de renaturação foi definida com base no processo iterativo de tentativa e erro considerando a literatura, bem como as características da proteína e os recursos disponíveis.

A renaturação foi feita utilizando o método de diluição gota a gota (1:20) onde a proteína solubilizada foi lentamente transferida para o tampão de renaturação a 4 °C com um fluxo de aproximadamente 10 mL/h. Neste caso, o tampão de renaturação estava sob agitação constante a fim de auxiliar na homogeneização da solução e, conseqüentemente, favorecer a renaturação.

Durante a renaturação *in vitro* ocorreu a formação de agregados proteicos que foram removidos por centrifugação ou filtração a 45 µm. Em seguida, a proteína renaturada foi quantificada a 280 nm em aparelho NanoDrop 1000 (Thermo Fisher Scientific) com base nos valores do coeficiente de extinção molar e massa molecular do monômero preditos anteriormente. O rendimento proteico obtido após renaturação e remoção dos agregados foi de aproximadamente 0,5 mg/mL⁻¹ por litro de cultivo. A solução contendo a proteína renaturada (aproximadamente 500 mL) foi então concentrada em célula agitada Amicon (modelo 8200) (Merck Millipore) com uma membrana de “cutoff” de massa molecular de 50 kDa até atingir a concentração de aproximadamente 11,7 mg/mL⁻¹ e, posteriormente, submetida à purificação por exclusão de tamanho a fim de aumentar a pureza da amostra, estimar a massa molecular da proteína e prepará-la para armazenamento e uso.

3.5. Purificação proteica

Para a etapa de purificação, uma alíquota de 500 µL da proteína renaturada e concentrada (aproximadamente 11,7 mg/mL⁻¹) foi submetida à cromatografia por

exclusão de tamanho (SEC) em uma coluna Superdex 200 Increase 10/300 GL (GE Healthcare) conectada a um aparelho de automatização cromatográfica ÄKTA FPLC System (GE Healthcare) e equilibrada em tampão Tris HCl pH 9, 200 mM NaCl e glicerol 10% (v/v). Além de aumentar a pureza da amostra, a SEC é a técnica mais aceitável e confiável para detectar a presença de proteínas oligomerizadas corretamente e monitorar a presença de agregados (Lebendiker e Danieli, 2014).

As frações eluídas da coluna foram coletadas e analisadas por SDS-Page (Figura 15). Em seguida, as alíquotas referentes a principal fração eluente foram unidas, concentradas através de filtros Amicon® Ultra-15 (Merck MilliPore) (“cutoff” de 30 kDa) e, posteriormente quantificadas em aparelho NanoDrop 1000 (Thermo Fisher Scientific) com base nos parâmetros físico-químicos preditos anteriormente. Após a SEC foi possível recuperar, no total, 1 mL da proteína a uma concentração de aproximadamente $2,1 \text{ mg/mL}^{-1}$, resultante da purificação da alíquota de 500 μL anteriormente injetada na coluna.

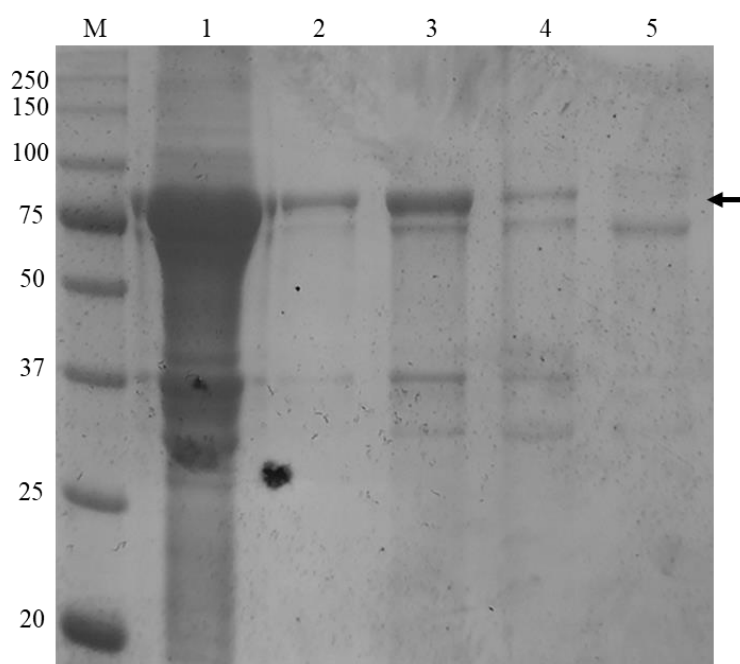


Figura 15. Eletroforese por SDS-Page da BXil após purificação por SEC. M: Marcador de massa molecular (“Precision Plus Protein Dual Color Standards” – BioRad); Linha 1: Amostra injetada na coluna de purificação a uma concentração de aproximadamente $11,7 \text{ mg/mL}^{-1}$; linhas 2 a 5: frações eluídas após a purificação. A seta indica a BXil de aproximadamente 81,25 kDa.

Ambas as amostras da proteína renaturada antes e após a purificação por SEC (11,7 e 2,1 mg/mL⁻¹, respectivamente) foram submetidas à ensaios de cristalização.

3.6. Análise de Espalhamento dinâmico de luz e detecção da atividade enzimática

Para avaliar se a renaturação ocorreu adequadamente e monitorar a homogeneidade da amostra e a formação de agregados proteicos utilizam-se diferentes métodos, tais como a presença de atividade frente à degradação de substrato, cromatografia por exclusão de tamanho, dicroísmo circular, cristalogênese e DLS (Vincentelli et al., 2009; Lebendiker e Danieli, 2014).

Diante disso, a proteína renaturada e purificada foi submetida a um teste qualitativo para detectar a presença de atividade enzimática usando o substrato cromogênico *p*-nitrofenil- β -D-xilopiranosídeo (*p*NP-X) e à análise de Espalhamento dinâmico de luz (DLS). Os testes com o substrato cromogênico permitiram detectar a presença de atividade enzimática resultando na mudança da coloração para amarelo (Figura 16.A), enquanto que os resultados obtidos a partir do DLS indicaram que a amostra era monodispersa (Figura 16.B), ou seja, todas as partículas possuíam o mesmo tamanho e não variaram de acordo com o tempo.

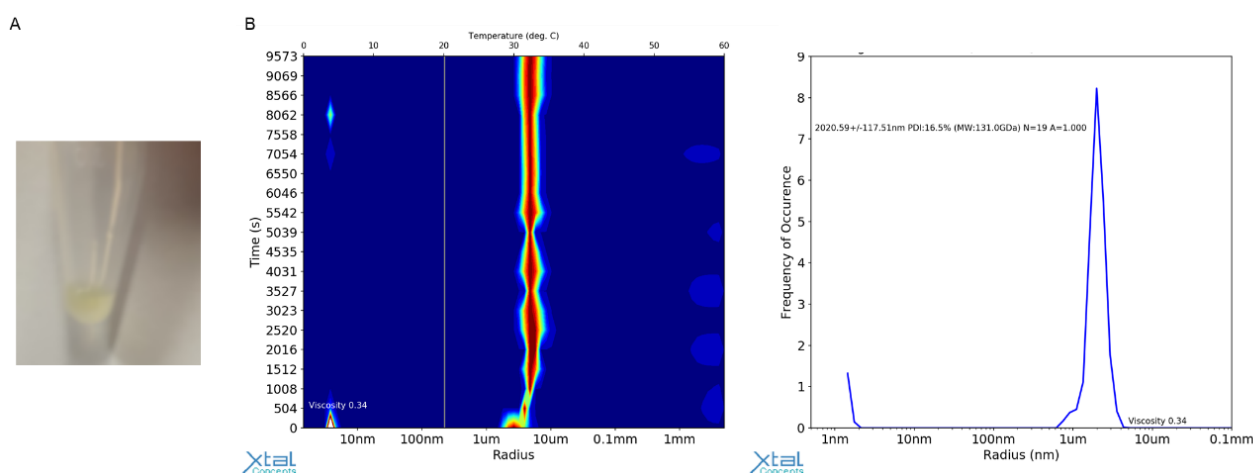


Figura 16. Análise da eficiência da etapa de renaturação *in vitro* da BXil. A: teste qualitativo de detecção de atividade enzimática com a proteína renaturada e purificada. B: Gráfico de distribuição e histograma obtidos a partir da análise de DLS realizadas no Spectro Light 600 (Xtal Concepts).

3.7. Ensaio de cristalização da BXil através do método de difusão de vapor em gota assente

As placas preparadas com os kits comerciais para os ensaios de cristalização através do método de difusão de vapor em gota assente foram seladas e mantidas em ambientes com temperatura controlada e as gotas foram examinadas sob uma lupa estereoscópica diariamente durante 60 dias.

Os kits utilizados têm como finalidade de realizar o “screening” inicial dos parâmetros de cristalização e possuem condições amplamente variadas (tampões, sais, precipitantes e pH). Além destas características, o Kit Morpheus HT-96 (Molecular Dimensions) é ideal para cristalizar proteínas de membrana. Apesar disso, em nenhuma condição avaliada observou-se a formação de cristais ou micro cristais durante o período de 60 dias; observou-se somente gotas claras e gotas com a formação de precipitados. Isso pode indicar que a concentração de proteína e do reagente não atingiu o ponto de saturação que deveria nuclear e, conseqüentemente, produzir cristais. Além disso, outros fatores também influenciam o processo de cristalização, tais como temperatura, concentração da proteína, composição e concentração do precipitante e a presença de contaminantes.

Apesar dos inúmeros esforços e diante dos resultados obtidos, conclui-se que a baixa solubilidade da proteína pode estar relacionada às características intrínsecas, como por exemplo, a presença de um “linker” demasiado grande entre os domínios, impedindo-os de assumirem uma orientação relativa uniforme, o que pode impedir a cristalização. Em alguns casos, a expressão do gene pode ter mais sucesso quando em hospedeiros eucarióticos, no entanto, estes sistemas possuem suas próprias limitações, como alto custo, longa duração e modificações pós-traducionais variadas (Yumerefendi et al., 2010).

4. CONCLUSÕES E PERSPECTIVAS

O gene anotado no genoma de uma *Chitinophaga* sp. codificou uma β -xilosidase de aproximadamente 81,25 kDa com uma arquitetura de domínios ainda não caracterizada, a qual não foi expressa com sucesso na fração solúvel apesar das

inúmeras tentativas. A extração desnaturante com concentração média de agentes caotrópicos e a posterior renaturação *in vitro* mostraram ser viáveis para obtenção da enzima ativa a partir de corpos de inclusão mesmo que com baixo rendimento. No entanto, a ausência de cristalização pode ter sido resultado de alguma falha durante o processo de extração desnaturante ou renaturação, ou ainda, devido às características intrínsecas da enzima. Apesar disso, os métodos sistematizados neste trabalho podem servir como ponto de partida para a purificação de outras enzimas para “screening” inicial de atividade de interesse, uma vez que foi capaz de resultar em proteína solúvel e ativa, conforme evidenciado pelos dados de DLS e degradação de substrato cromogênico.

Espera-se que a expressão de domínios separados possa resolver o problema de solubilidade e permitir novos ensaios de cristalografia e determinação da estrutura tridimensional. Deste modo, uma tentativa de expressão da proteína na forma solúvel será realizada no Instruct-ERIC, Grenoble – França, onde será utilizada a técnica de truncamento incremental aleatório da sequência do gene alvo, denominada ESPRIT (Yumerefendi et al., 2010; Mas e Hart, 2017).

5. REFERÊNCIAS

- Alexander McPherson. (2017). Protein Crystallization. In: Wlodawer, A., Jaskolski, M., & Dauter, Z. (Eds.), **Protein Crystallography: Methods and Protocols**. New York, NY: Springer Nature, p. 17–50.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. e Bourne, P.E. (2000). The protein data bank. **Nucleic acids research** 28: 235–242.
- Cabrita, L.D. e Bottomley, S.P. (2004). Protein expression and refolding - A practical guide to getting the most out of inclusion bodies. **Biotechnology Annual Review** 10: 31–50.
- Chura-Chambi, R.M., Cordeiro, Y., Malavasi, N. V., Lemke, L.S., Rodrigues, D. e Morganti, L. (2013). An analysis of the factors that affect the dissociation of inclusion bodies and the refolding of endostatin under high pressure. **Process Biochemistry** 48: 250–259.
- Clark, E.D.B. (2001). Protein refolding for industrial processes. **Current Opinion in Biotechnology** 12: 202–207.

Clark, E.D.B., Schwarz, E. e Rudolph, R. (1999). Inhibition of aggregation side reactions during in vitro protein folding. **Methods in Enzymology** 309: 217–236.

Funniceili, M.I.G. (2018). **Explorando uma comunidade bacteriana isolada de uma pilha de bagaço de cana-de-açúcar e seu potencial funcional na degradação de biomassa lignocelulósica**. 65 f. Dissertação (Mestrado em Microbiologia Agropecuária) – Unesp, Jaboticabal.

Geerlof, A. **M9 mineral medium**. Helmholtz Center Munich, Protocols. 15 de maio de 2010. Disponível em: < https://www.helmholtz-muenchen.de/fileadmin/PEPF/Protocols/M9-medium_150510.pdf>; Acesso em: 20 de novembro de 2019.

Goldberg, M.E., Expert-Bezançon, N., Vuillard, L. e Rabilloud, T. (1996). Non-detergent sulphobetaines: A new class of molecules that facilitate in vitro protein renaturation. **Folding and Design** 1: 21–27.

Hart, D.J. e Tarendeau, F. (2006). Combinatorial library approaches for improving soluble protein expression in Escherichia coli. **Acta Crystallographica Section D: Biological Crystallography** 62: 19–26.

Kaur, J., Kumar, A. e Kaur, J. (2018). Strategies for optimization of heterologous protein expression in E. coli: Roadblocks and reinforcements. **International Journal of Biological Macromolecules** 106: 803–822.

Kim, H.W., Lee, S.Y., Park, H. e Jeon, S.J. (2015). Expression, refolding, and characterization of a small laccase from *Thermus thermophilus* HJ6. **Protein Expression and Purification** 114: 37–43.

Kishi, L.T., Lopes, E.M., Fernandes, C.C., Fernandes, G.C., Sacco, L.P., Alves, L.M.C. e Lemos, E.G.M. (2017). Draft Genome Sequence of a Chitinophaga Strain Isolated from a Lignocellulose Biomass-Degrading Consortium. **Genome Announcements** 5: 1–2.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature** 227: 680–5.

Lebendiker, M. e Danieli, T. (2014). Production of prone-to-aggregate proteins. **FEBS Letters** Federation of European Biochemical Societies 588: 236–246.

Mas, P.J. e Hart, D.J. (2017). ESPRIT: A method for defining soluble expression constructs in poorly understood gene sequences. In: Burgess-Brown N. (Eds.), **Heterologous Gene Expression in E.coli**. New York, NY: Humana Press, p. 45–63.

Philo, J. e Arakawa, T. (2009). Mechanisms of Protein Aggregation. **Current Pharmaceutical Biotechnology** 10: 348–351.

Rinas, U., Garcia-Fruitós, E., Corchero, J.L., Vázquez, E., Seras-Franzoso, J. e Villaverde, A. (2017). Bacterial Inclusion Bodies: Discovering Their Better Half. **Trends in Biochemical Sciences** Elsevier Ltd 42: 726–737.

Rosano, G.L. e Ceccarelli, E.A. (2014). Recombinant protein expression in *Escherichia coli*: Advances and challenges. **Frontiers in Microbiology** 5: 1–17.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. **Protein expression and purification** 41: 207–34.

Vallejo, L.F. e Rinas, U. (2004). Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. 12: 1–12.

Vincentelli, R., Canaan, S., Campanacci, V., Valencia, C., Maurin, D., Frassinetti, F., Scappucini-Calvo, L., Bourne, Y., et al. (2009). High-throughput automated refolding screening of inclusion bodies. **Protein Science** 13: 2782–2792.

Wagner, S., Klepsch, M.M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Högbom, M., Van Wijk, K.J., et al. (2008). Tuning *Escherichia coli* for membrane protein overexpression. **Proceedings of the National Academy of Sciences of the United States of America** 105: 14371–14376.

Wang, W., Nema, S. e Teagarden, D. (2010). Protein aggregation-Pathways and influencing factors. **International Journal of Pharmaceutics** 390: 89–99.

Yumerefendi, H., Tarendeau, F., Mas, P.J. e Hart, D.J. (2010). ESPRIT: An automated, library-based method for mapping and soluble expression of protein domains from challenging targets. **Journal of Structural Biology** 172: 66–74.

CAPÍTULO 5 – Ensaio preliminares de cristalização da lacase metagenômica (LacMeta) prospectada em solo sob serapilheira de eucalipto

RESUMO – As lacases atuam numa ampla variedade de compostos polifenólicos, fenólicos, não fenólicos e aminas aromáticas. Estas enzimas são úteis para a biorrefinaria, indústrias alimentícias, indústrias de processamento de papel e celulose, indústrias têxteis, no tratamento de efluentes, na biorremediação de solos e águas contaminados com hidrocarbonetos aromáticos policíclicos, dentre outros. Geralmente, a maioria das lacases possui três centros de ligação ao cobre no seu centro catalítico, considerados essenciais para o correto desempenho catalítico dessas macromoléculas. Algumas lacases no entanto, apresentam somente dois domínios de cobre, sendo conhecidas como “small lacases”, ou SLACs. A caracterização cinética aliada a caracterização estrutural de SLACs pode contribuir para elucidar suas características individuais. Este estudo teve como objetivo resolver a estrutura de uma “small” lacase metagenômica (LacMeta) anteriormente caracterizada por estudos cinéticos, a fim de compreender a base molecular envolvida na evolução das SLACs e comparar suas características estruturais com as das principais lacases já caracterizadas. Em estudos anteriores a LacMeta apresentou elevado potencial biotecnológico para degradação de corantes da indústria têxtil. No presente estudo, a LacMeta foi expressa e purificada em alta concentração e submetida a ensaios de cristalização e difração de raios-X, onde os dados de difração coletados foram medidos a uma resolução de 1,8 Å. Algumas tentativas de determinação da estrutura foram realizadas, entretanto, o método de substituição molecular não foi eficiente devido à falta de um modelo de fase adequado no Banco de Dados de Proteínas. Assim, novos estudos cristalográficos utilizando cristais derivados com selenometionina serão necessários para determinar a estrutura da LacMeta.

Palavras-chave: difração de raios-X, enzima metagenômica, “multicopper oxidase”, substituição molecular

1. INTRODUÇÃO

Lacases (1.10.3.2) são oxidoreduktases que atuam numa ampla variedade de compostos como polifenólicos, fenólicos, não fenólicos e aminas aromáticas, incluindo a lignina e substâncias húmicas (Balcázar-López et al., 2016). Elas fazem parte de uma superfamília chamada “Multicopper oxidases” (MCOs), as quais possuem três centros de ligação ao cobre no seu centro catalítico e são amplamente distribuídas em fungos, plantas e bactérias (Baldrian, 2006; Balcázar-López et al., 2016).

Esses três centros de ligação ao cobre, ou centro de cupredoxina, acomodam quatro tipos de átomos de cobre e são designados por T1, T2 e T3. O centro T1 apresenta um cobre de tipo 1; o centro T2 apresenta um cobre do tipo 2 e o T3 apresenta dois cobres do tipo 3, sendo que os cobres tipo 2 e tipo 3 estão posicionados na forma de um triângulo. O centro T1 confere a coloração azul destas enzimas e é responsável pela oxidação do substrato e transferência dos elétrons para os outros centros de ligação ao cobre. Os centros T2 e T3 formam o centro trinuclear de cobre (TNC), responsável pela ligação, ativação e redução do oxigênio em água (Solomon et al., 1996; Komori et al., 2009; Giardina et al., 2010; Sirim et al., 2011; Gunne et al., 2014) (Figura 1).

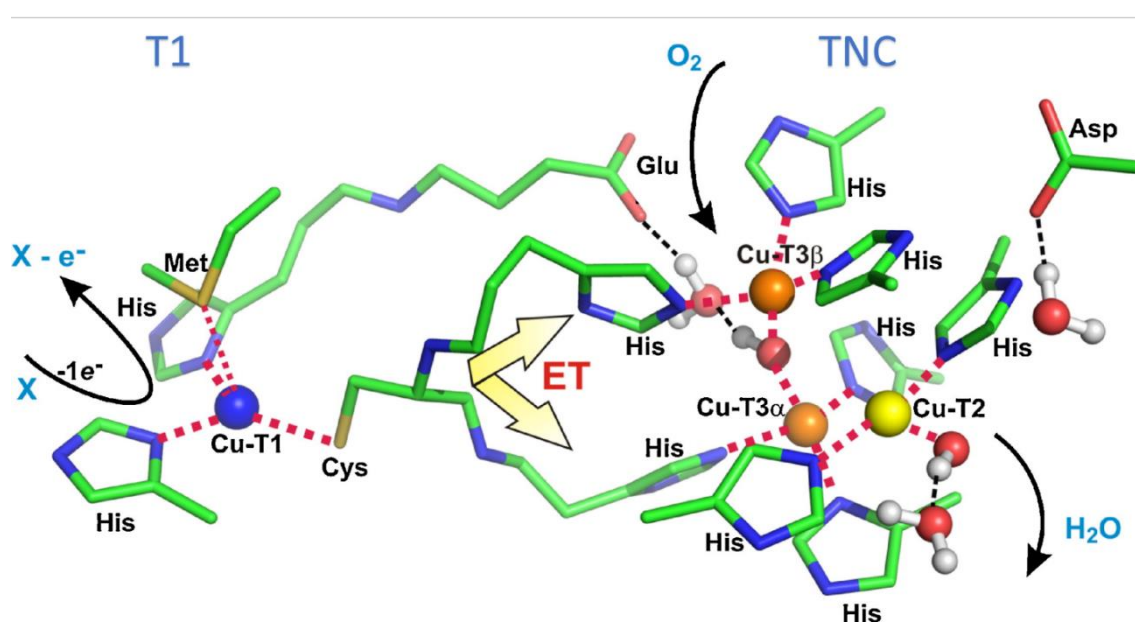


Figura 1. Arquitetura dos centros de ligação ao cobre encontrado nas MCOs (PDB 1kv7). Adaptado de (Rulíšek; Ryde, 2013).

Os íons de cobre presentes nesses centros são coordenados por um resíduo de metionina, um resíduo de cisteína e 10 resíduos de histidina altamente conservados que são considerados essenciais para o correto desempenho catalítico dessas macromoléculas. Essas histidinas estão localizadas dentro de assinaturas, ou motivos conservados, que são divididas em H-X-H-G, H-X-H, H-X(2)-H-X-H E H-C-H-X(3)-H-X(4)-[MLF] (onde o X representa qualquer resíduo de aminoácido) (Reiss et al., 2013; Gunne et al., 2014).

Normalmente, as lacases possuem estes três domínios de ligação ao cobre e são denominadas lacases canônicas típicas, porém, algumas apresentam somente dois domínios sendo chamadas de “Small lacases”, ou SLACs. Curiosamente, as SLACs necessitam possuir uma organização homotrimérica para serem biologicamente ativas. Isso porque, pelo fato de ter um dos domínios de ligação ao cobre ausente, necessitam utilizar os domínios adjacentes para reconstruir o TNC funcional (Endo et al., 2003; Machczynski et al., 2004; Murphy et al., 2008; Skálová et al., 2009).

Envolvidas em diversos processos biológicos, como homeostase de metais pesados e degradação de lignina (Giardina et al., 2010), estas enzimas têm recebido muita atenção nos últimos anos devido à sua ampla variedade de aplicações biotecnológicas. Dentre suas principais utilidades podemos destacar sua participação em indústrias alimentícias, de cosméticos, de papel e celulose, têxteis, na nanotecnologia, na biorremediação e síntese orgânica (Riva, 2006; Rodríguez Couto e Toca Herrera, 2006; Sirim et al., 2011).

A etapa de pré-tratamento realizada na biomassa a fim de alterar ou remover a lignina libera compostos fenólicos que inibem a hidrólise da celulose e hemicelulose durante o processo de conversão de biomassa em biocombustível (Wyman et al., 2005; Parawira e Tekere, 2011; Alvira et al., 2013; Bagewadi et al., 2017). A remoção destes compostos inibitórios é realizada através de métodos físicos e químicos, no entanto, a desintoxicação biológica utilizando enzimas surge como uma alternativa ecológica. Apesar disso, o uso de enzimas em muitos processos industriais ainda é considerado oneroso e complexo, sobretudo devido à baixa estabilidade e efetividade destas moléculas perante as condições adversas do próprio processo que participam (Zeng et al., 2011; Johnson, 2016).

Diante desta demanda por enzimas mais efetivas e de baixo custo, os recentes avanços da engenharia de proteínas surgiram como estratégia para superar este problema.

A caracterização cinética aliada a estudos de resolução da estrutura tridimensional de lacases pode contribuir para a compreensão das relações estrutura-função e, a partir da engenharia de enzimas, aprimorar suas principais características e, por conseguinte, ampliar suas aplicações biotecnológicas. Alguns exemplos de características desejáveis são atividade aumentada, maior especificidade e seletividade em diferentes substratos (Kunamneni et al., 2008).

Uma nova lacase de dois domínios (“small” lacase) foi prospectada no metagenoma de solo sob serapilheira de eucalipto (Gomes et al., 2013; Gomes-Pepe et al., 2016) e caracterizada cineticamente em um estudo preliminar (Lima, 2017).

A LacMeta, que apresentou similaridade de 82% com uma proteína hipotética de *Streptomyces rubidus* [WP_069465126.1], foi purificada na forma homotrimérica com 128,22 kDa. Apresentou atividade ótima a 37 °C e em pH ácido para o substrato ABTS, mas também manteve sua atividade acima de 50% em uma ampla faixa de temperatura (10-60 °C) (Lima, 2017). Além disso, a LacMeta degradou sete corantes de cinco classes distintas sem a presença de um mediador redox. Diante das características apresentadas, a LacMeta demonstra elevado potencial biotecnológico para o tratamento de águas residuais de indústria têxtil, resultando ainda no depósito de uma patente.

No presente estudo, esta lacase foi submetida à ensaios de cristalização com objetivo de compreender a base molecular envolvida na evolução das SLACs e comparar suas características estruturais com as das principais lacases já caracterizadas. No entanto, a técnica de substituição molecular utilizada para determinação de estruturas não foi bem sucedida para a enzima em questão, sendo assim, aqui descrevemos somente os ensaios preliminares de cristalização e as perspectivas futuras.

Este estudo teve como objetivo realizar ensaios de cristalografia por difração de raios-X e determinação da estrutura cristalográfica de uma “small” lacase metagenômica anteriormente caracterizada através de estudos cinéticos.

2. MATERIAL E MÉTODOS

2.1. Origem da proteína

A LacMeta foi previamente relatada em um trabalho de origem do grupo de pesquisa do LBMP (Lima, 2017). Essa enzima foi prospectada utilizando a região conservada para lacases L3 (HLHGH) em uma base de dados metagenômicos do LBMP utilizando a abordagem “sequence-driven” e teve sua caracterização cinética realizada.

2.2. Expressão e extração

Para os ensaios de expressão, uma colônia isolada foi cultivada em meio LB suplementado com canamicina ($50 \mu\text{g/mL}^{-1}$) a 37°C durante 16 horas. Em seguida, 10% (v/v) desse pré-inóculo foi adicionado em um novo meio LB contendo canamicina ($50 \mu\text{g/mL}^{-1}$) para expressão da proteína recombinante. A expressão ocorreu a 37°C , durante 16 horas, sob indução com 0,1 mM de IPTG. As células foram centrifugadas e lisadas por um homogeneizador de alta pressão (APV Model 2000 Homogenizer) com três passagens do extrato a 11600 psi e novamente centrifugada para separação dos extratos solúveis e insolúveis.

2.3. Purificação proteica

A fração solúvel contendo a LacMeta foi inicialmente purificada por Cromatografia de afinidade por íons metálicos imobilizados (IMAC) utilizando uma coluna “Sepharese” (HisTrap HP – GE Healthcare, 5 mL), eluída com um gradiente crescente de Imidazol. As frações eluídas foram reunidas e concentradas através de filtros Amicon® Ultra-15 (Merck MilliPore) com capacidade de retenção de moléculas com massa molecular superior a 10 kDa e, posteriormente, purificada por cromatografia líquida por exclusão molecular (SEC) utilizando uma coluna “Superdex 200 Increase” 10/300 GL (GE Healthcare) conectada em um aparelho de automatização cromatográfica ÄKTA FPLC System (GE Healthcare) e equilibrada em

tampão Tris HCl pH 9, 200 mM NaCl e 10% (v/v) de glicerol. As frações eluídas na purificação SEC (1 mL cada fração) referente a principal fração eluente visualizada no cromatograma foram analisadas por SDS-Page.

A concentração da proteína pura foi calculada a partir da medida de absorção a 280 nm utilizando um espectrofotômetro NanoDrop 1000 (Thermo Fisher Scientific, Waltham, Massachusetts, EUA) com base nos valores do coeficiente de extinção molar e na massa molecular da proteína calculados utilizando o programa ProtParam (<https://web.expasy.org/protparam/>).

As frações eluídas da SEC que exibiram um elevado grau de pureza e concentração desejável ($\geq 10 \text{ mg/mL}^{-1}$) foram diretamente submetidas a ensaios de cristalização.

2.4. Experimentos de cristalização

Os primeiros ensaios de cristalização e o “screening” inicial foram realizados no Laboratório Nacional de Biociências (LNBio), Campinas – SP, através de proposta de pesquisa enviada pela pesquisadora Dra. Elisângela Soares Gomes Pepe. Os cristais obtidos nas condições avaliadas não apresentaram nenhuma difração. Diante disso, a amostra da enzima LacMeta foi gentilmente cedida pela pesquisadora para a presente pesquisa.

Os ensaios descritos a seguir foram realizados durante estágio doutoral no Laboratório de Cristalografia Aplicada à Indústria e a Medicina, do Instituto de Tecnologia Química e Biológica (ITQB) da Universidade Nova de Lisboa (UNL) em Oeiras, Portugal, sob orientação do pesquisador Dr. Pedro Matias.

A fim de otimizar os ensaios realizados no LNBio, as condições iniciais foram reproduzidas manualmente em placas de 24 poços com volume de gota de 1 μL (variando a proporção da proteína e do precipitante) e volume de reservatório de 500 μL . As placas de cristalização foram conservadas em três ambientes com temperatura constante (4°C, 20°C e 30°C). As gotas foram examinadas sob uma lupa estereoscópica diariamente.

Novos kits comerciais foram testados: JCSG+ “Screen”, “Structure” I & II, “Pact Premier”, “The BCS Screen”, “Morpheus HT-96” (Molecular Dimensions) e os kits “The

Pact Suite” e “The AmSO4 Suite” (Qiagen), a fim de estabelecer uma nova condição ideal. Os novos testes foram realizados pelo método de difusão de vapor em placas de 96 poços preparadas pelo robô de pipetagem Mosquito LCP (SPT Labtech) (Figura 1). As placas foram seladas e mantidas em ambientes com temperatura constante conforme descrito anteriormente.



Figura 1. Imagem ilustrativa do robô de pipetagem Mosquito LCP (SPT Labtech) com câmara de controle de umidade que impede a evaporação da amostra.

A seguir, as condições com formação de cristais foram reproduzidas manualmente em triplicata com volume de gota de 1 μL (0,5 μL de proteína e 0,5 μL do precipitante) adicionando diferentes aditivos do kit de cristalização Additive Screen HT (Hampton Research). Os cristais considerados adequados para os testes

seguintes foram coletados, imersos em solução crioprotetora, congelados em nitrogênio líquido e enviados para difração.

2.5. Coleta de dados de difração

Os cristais LacMeta foram submetidos à difração de raios-X na Linha de Luz I03 no sincrotron “Diamond Light Source” em Oxfordshire, Reino Unido, usando um detector Eiger2 XE 16M e um comprimento de onda 0.97628 Å.

3. RESULTADOS E DISCUSSÃO

3.1. Expressão e purificação da proteína recombinante

A enzima LacMeta (~42,74 kDa) foi expressa com alto rendimento em *E. coli* e, em seguida, foi purificada por meio de cromatografia de afinidade ao níquel utilizando a resina Ni-NTA e eluída com 500 mM de imidazol. Posteriormente, as frações eluídas na purificação SEC referente a principal fração eluente visualizada no cromatograma foram analisadas por eletroforese em gel SDS-Page (Davis, 1964) (Figura 2). O rendimento total da proteína após a IMAC e após a SEC foi de aproximadamente 63 e 46 mg/mL⁻¹ a partir de 1 litro de cultivo, respectivamente.

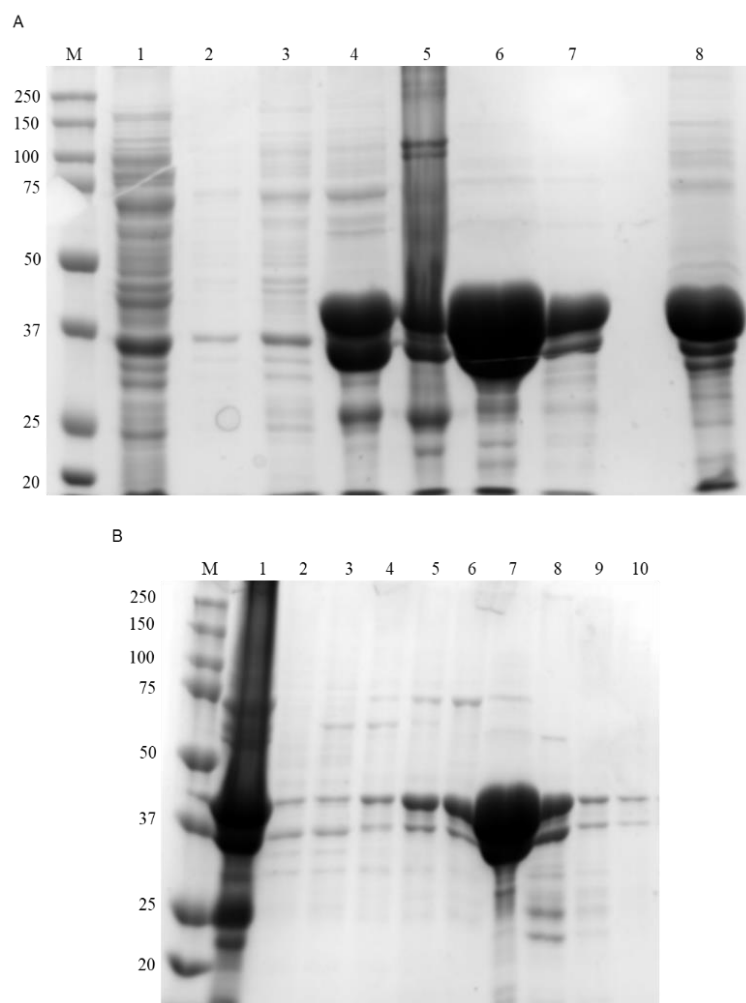


Figura 2. Eletroforese por SDS-Page da LacMeta. A: Purificação por cromatografia de afinidade IMAC; 1-8: frações eluídas em gradiente de imidazol (100 a 500 mM). B: Purificação por cromatografia por exclusão de tamanho SEC; 1: amostra injetada para purificação SEC; 2 a 10: frações eluídas. M: Marcador de massa molecular (“Precision Plus Protein Dual Color Standards” – BioRad).

3.2. Cristalização e otimização dos cristais

Os cristais formados a partir das condições iniciais avaliadas não apresentavam as características desejáveis para ensaios de difração de raios-X, pois estavam aglomerados, eram muito pequenos, de forma irregular e frágeis impedindo sua coleta. Quando a coleta era realizada com sucesso, os cristais coletados apresentaram pouca ou nenhuma difração. Diante disso, outros kits comerciais, incluindo kit de aditivos, foram utilizados para o preparo das placas pelo robô de pipetagem Mosquito LCP (SPT Labtech).

Nesse momento, as placas de cristalização foram mantidas a 20°C e observadas diariamente até que foi possível observar a formação de cristais com tamanho adequado (aproximadamente 5 dias).

Os cristais foram obtidos na condição ideal contendo 25% (v/v) PEG 3550, 200 mM sulfato de sódio, 100 mM BIS-TRIS propano pH 7,5 e 10% (v/v) de etanol 30% (v/v) na temperatura de 20 °C após 5 dias. Os cristais únicos, com faces e bordas lisas foram coletados para ensaios de difração de raios-X (Figura 3), imersos em uma solução crioprotetora (15% de glicerol (v/v), 25% (v/v) PEG 3550, 200 mM sulfato de sódio, 100 mM BIS-TRIS propano pH 7,5 e 10% (v/v) de etanol 30% (v/v)) e usados para coleta de dados.

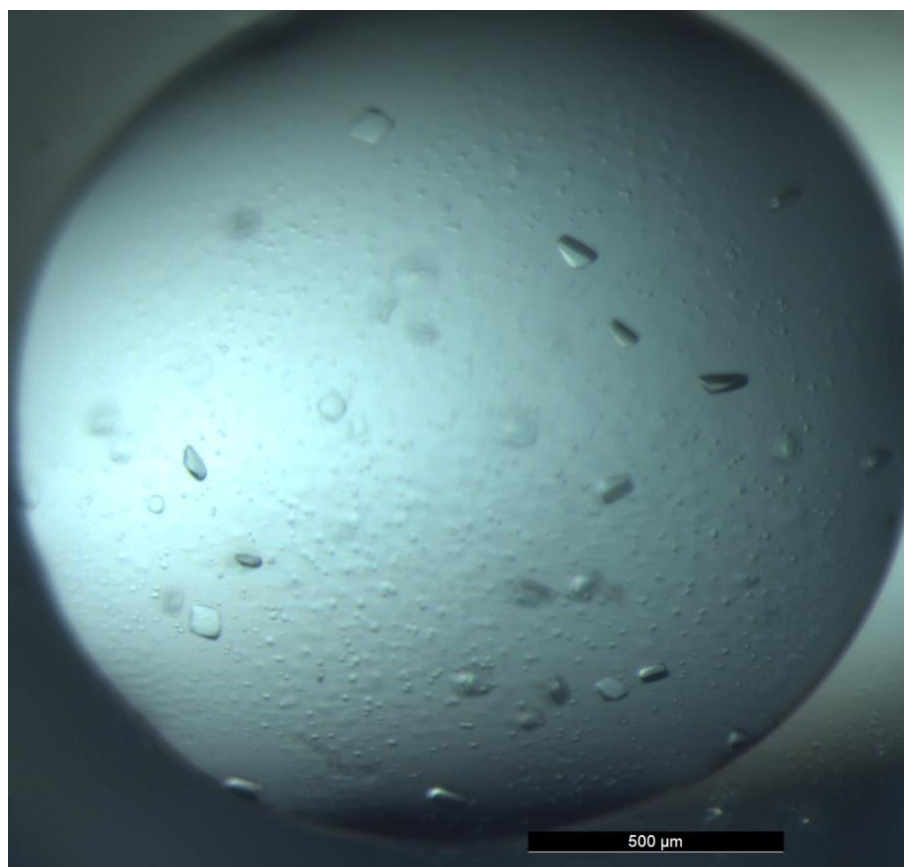


Figura 3. Cristais LacMeta obtidos sob a condição ideal contendo 25% (v/v) PEG 3550, 200 mM Sulfato de sódio, 100 mM BIS-TRIS propano pH 7,5 e 10% (v/v) de etanol 30% (v/v), crescidos em temperatura de 20 °C durante 5 dias.

3.3. Coleta e processamento de dados

Os melhores dados de difração coletados para os cristais LacMeta foram a uma resolução de 1,8 Å. A busca por sequências similares no banco de dados PDBe (“The Protein Data Bank in Europe”) resultou em uma sequência com similaridade de 29,5% com a LacMeta e algumas tentativas de determinação estrutural foram realizadas sem sucesso. Normalmente, um molde com similaridade de sequência acima de 25% pode ser utilizado para determinar a estrutura pelo método de substituição molecular, no entanto, podem haver exceções (Evans e McCoy, 2007; Taylor, 2010).

A baixa similaridade de sequência é uma característica comum entre “Multicopper” oxidases que costumam compartilhar conservação elevada restrita a domínios catalíticos. Além disso, estes resultados corroboram com os estudos anteriores que indicaram que a LacMeta se trata de uma nova enzima.

Assim, sugerimos novos ensaios utilizando um sistema de expressão estabelecido para a substituição de resíduos de metionina por selenometionina (SeMet) na proteína recombinante, o que permitirá a determinação estrutural através do método de Difração anômala de comprimento de onda único (SAD - “Single wavelength anomalous diffraction”).

4. CONCLUSÕES E PERSPECTIVAS

A LacMeta foi expressa e purificada com alto rendimento em *E. coli* e submetida à ensaios de cristalização e difração de raios-X. Os melhores dados de difração coletados foram medidos a uma resolução de 1,8 Å, no entanto, não foi possível determinar sua estrutura tridimensional através do método de substituição molecular, principalmente devido à falta de um modelo de fase adequado no Banco de Dados de Proteínas, resultado da baixa similaridade entre as sequências da proteína alvo e da proteína modelo. A baixa similaridade é uma característica comum entre “Multicopper oxidases” que costumam compartilhar conservação elevada restrita a domínios catalíticos. Além disso, estes resultados corroboram com os estudos anteriores que indicaram que a LacMeta se trata de um nova enzima.

Diante disso, novos ensaios cristalográficos utilizando cristais de um derivado de selenometionina serão necessários para determinar a estrutura da LacMeta, pois trata-se de uma lacase bacteriana pouco conhecida e de alto interesse no setor

biotecnológico. Os estudos cristalográficos permitirão a determinação da estrutura da LacMeta e o conhecimento obtido será importante para elucidar seu mecanismo de ação e, eventualmente, construir mutantes com propriedades catalíticas melhoradas.

5. REFERÊNCIAS

Alvira, P., Moreno, A.D., Ibarra, D., Sáez, F. e Ballesteros, M. (2013). Improving the fermentation performance of *saccharomyces cerevisiae* by laccase during ethanol production from steam-exploded wheat straw at high-substrate loadings. **Biotechnology Progress** 29: 74–82.

Bagewadi, Z.K., Mulla, S.I. e Ninnekar, H.Z. (2017). Optimization of laccase production and its application in delignification of biomass. **International Journal of Recycling of Organic Waste in Agriculture** 6: 351–365.

Balcázar-López, E., Méndez-Lorenzo, L.H., Batista-García, R.A., Esquivel-Naranjo, U., Ayala, M., Kumar, V.V., Savary, O., Cabana, H., et al. (2016). Xenobiotic compounds degradation by heterologous expression of a *trametes sanguineus* laccase in *trichoderma atroviride*. **PLoS ONE** 11: e0147997.

Baldrian, P. (2006). Fungal laccases – occurrence and properties. **FEMS Microbiology Reviews** 30: 215–242.

Davis, B.J. (1964). Davis : Disc Electrophoresis. **Clinical Applications** 121: 404–427.

Endo, K., Hayashi, Y., Hibi, T., Hosono, K., Beppu, T. e Ueda, K. (2003). Enzymological characterization of EpoA, a laccase-like phenol oxidase produced by *Streptomyces griseus*. **Journal of Biochemistry** 133: 671–677.

Evans, P. e McCoy, A. (2007). An introduction to molecular replacement. **Acta Crystallographica Section D: Biological Crystallography** International Union of Crystallography 64: 1–10.

Giardina, P., Faraco, V., Pezzella, C., Piscitelli, A., Vanhulle, S. e Sannia, G. (2010). Laccases: A never-ending story. **Cellular and Molecular Life Sciences** 67: 369–385.

Gomes-Pepe, E.S., Sierra, E.G.M., Pereira, M.R., Castellane, T.C.L. e De Lemos, E.G.M. (2016). Bg10: A novel metagenomics alcohol-tolerant and glucose-stimulated gh1 β -glucosidase suitable for lactose-free milk preparation. **PLoS ONE** 11: 1–25.

Gomes, E.S., Schuch, V. e Lemos, E.G. de M. (2013). Biotechnology of polyketides: New breath of life for the novel antibiotic genetic pathways discovery through metagenomics. **Brazilian Journal of Microbiology** 44: 1007–1034.

Gunne, M., Höppner, A., Hagedoorn, P.L. e Urlacher, V.B. (2014). Structural and redox properties of the small laccase Ssl1 from *Streptomyces svaceus*. **FEBS Journal** 281: 4307–4318.

Johnson, E. (2016). Integrated enzyme production lowers the cost of cellulosic ethanol. **Biofuels, Bioproducts and Biorefining** 10: 164–174.

Komori, H., Miyazaki, K. e Higuchi, Y. (2009). X-ray structure of a two-domain type laccase: A missing link in the evolution of multi-copper proteins. **FEBS Letters** 583: 1189–1195.

Kunamneni, A., Camarero, S., García-Burgos, C., Plou, F.J., Ballesteros, A. e Alcalde, M. (2008). Engineering and Applications of fungal laccases for organic synthesis. **Microbial Cell Factories** 7: 1–17.

Lima, N.S.M. (2017). **Busca de biomoléculas com potencial biotecnológico em database metagenômico por sequence-driven**. 80 f. Dissertação (Mestrado em Microbiologia Agropecuária) – Unesp, Jaboticabal.

Machczynski, M.C., Vijgenboom, E., Samyn, B. e Canters, G.W. (2004). Characterization of SLAC: A small laccase from *Streptomyces coelicolor* with unprecedented activity. **Protein Science** 13: 2388–2397.

Murphy, M.E.P., Lindley, P.F. e Adman, E.T. (2008). Structural comparison of cupredoxin domains: Domain recycling to construct proteins with novel functions. **Protein Science** 6: 761–770.

Parawira, W. e Tekere, M. (2011). Biotechnological strategies to overcome inhibitors in lignocellulose hydrolysates for ethanol production: Review. **Critical Reviews in Biotechnology** 31: 20–31.

Reiss, R., Ihssen, J., Richter, M., Eichhorn, E., Schilling, B. e Thöny-Meyer, L. (2013). Laccase versus Laccase-Like Multi-Copper Oxidase: A Comparative Study of Similar Enzymes with Diverse Substrate Spectra. **PLoS ONE** 8: e65633.

Riva, S. (2006). Laccases: blue enzymes for green chemistry. **Trends in Biotechnology** 24: 219–226.

Rodríguez Couto, S. e Toca Herrera, J.L. (2006). Industrial and biotechnological applications of laccases: A review. **Biotechnology Advances** 24: 500–513.

Rulíšek, L. e Ryde, U. (2013). Theoretical studies of the active-site structure, spectroscopic and thermodynamic properties, and reaction mechanism of “Multicopper oxidases”. **Coordination Chemistry Reviews** 257: 445–458.

Sirim, D., Wagner, F., Wang, L., Schmid, R.D. e Pleiss, J. (2011). The Laccase Engineering Database: A classification and analysis system for laccases and related “Multicopper oxidases”. **Database** 2011: 1–7.

Skálová, T., Dohnálek, J., Østergaard, L.H., Østergaard, P.R., Kolenko, P., Dušková, J., Štěpánková, A. e Hašek, J. (2009). The Structure of the Small Laccase from *Streptomyces coelicolor* Reveals a Link between Laccases and Nitrite Reductases. **Journal of Molecular Biology** 385: 1165–1178.

Solomon, E.I., Sundaram, U.M. e Machonkin, T.E. (1996). “Multicopper oxidases” and oxygenases. **Chemical Reviews** 96: 2563–2606.

Taylor, G.L. (2010). Introduction to phasing. **Acta Crystallographica Section D: Biological Crystallography** 66: 325–338.

Wyman, C.E., Dale, B.E., Elander, R.T., Holtzapple, M., Ladisch, M.R. e Lee, Y.Y. (2005). Coordinated development of leading biomass pretreatment technologies. **Bioresource Technology** 96: 1959–1966.

Zeng, J., Singh, D. e Chen, S. (2011). Biological pretreatment of wheat straw by *Phanerochaete chrysosporium* supplemented with inorganic salts. **Bioresource Technology** 102: 3206–3214.

APÊNDICES

Apêndice A. Sequências de aminoácidos provenientes do banco de dados NCBI.

Tabela 1A. Análise de similaridade da BXil com sequências de aminoácidos recuperadas do banco de dados “nr” do NCBI através da ferramenta BlastP utilizadas na análise fenética.

Número de acesso	Descrição	Organismo	Família	Número de aminoácidos	Cobertura (%)	Identidade (%)
SHL68501.1	GH43	<i>Chitinophaga jiangningensis</i>	<i>Chitinophagaceae</i>	487	67	86.7
NCU05616.1	GH43	<i>Chitinophagaceae</i> bacterium	<i>Chitinophagaceae</i>	723	96	64.84
WP_167290130.1	GH43	<i>Pseudoflavitalea</i> sp. X16	<i>Chitinophagaceae</i>	709	97	65.29
WP_144887167.1	GH43	<i>Lacibacter cauensis</i>	<i>Chitinophagaceae</i>	719	98	63.52
WP_146781387.1	GH43	<i>Flavisolibacter ginsenosidimutans</i>	<i>Chitinophagaceae</i>	719	96	63.22
WP_143387208.1	GH43	unclassified <i>Flavobacterium</i>	<i>Flavobacteriaceae</i>	728	94	64.09
AOW10809.1	beta-xilosidase	<i>Flavobacterium gilvum</i>	<i>Flavobacteriaceae</i>	735	95	63.45
WP_100430111.1	GH43	<i>Flavobacterium</i> sp.	<i>Flavobacteriaceae</i>	738	95	63.62
WP_114778441.1	GH43	<i>Botryobacter ruber</i>	<i>Hymenobacteraceae</i>	736	99	63.34
WP_162055101.1	GH43	<i>Pontibacter</i> sp. TRT317	<i>Hymenobacteraceae</i>	751	94	66.81
WP_146897773.1	GH43	<i>Adhaeribacter aerolatus</i>	<i>Hymenobacteraceae</i>	719	97	63.36
WP_123125614.1	GH43	<i>Rufibacter</i> sp. R-22-1c-1	<i>Hymenobacteraceae</i>	692	94	65.35
WP_150090418.1	GH43	<i>Adhaeribacter</i> sp. DK36	<i>Hymenobacteraceae</i>	724	94	64.82
ALJ01192.1	beta-xilosidase	<i>Rufibacter tibetensis</i>	<i>Hymenobacteraceae</i>	704	94	63.60
WP_115567210.1	GH43	<i>Pontibacter diazotrophicus</i>	<i>Hymenobacteraceae</i>	751	94	66.52
WP_149090453.1	GH43	<i>Rufibacter</i> sp. NBS58-1	<i>Hymenobacteraceae</i>	695	94	65.74
WP_131556829.1	GH43	<i>Pedobacter</i> sp. RP-3-11	<i>Sphingobacteriaceae</i>	718	99	62.22
WP_069381093.1	GH43	<i>Pedobacter steynii</i>	<i>Sphingobacteriaceae</i>	718	100	61.88
TDQ08910.1	beta-xilosidase	<i>Pedobacter metabolipauper</i>	<i>Sphingobacteriaceae</i>	722	98	61.26

	Tuner (DE3)	0,4 – 0,6	1	LB	20, 37	16	Tris HCl pH 8 (20 mM)	5	200	20	-	French Press	Não
	C43 (DE3)	0,8 – 1	1	LB	20, 37	16	Tris HCl pH 8 (20 mM)	5	200	20	-	French Press	Não
C4 (Sem peptídeo sinal)	BL21 (DE3)	0,4 - 0,6	0,1	LB	37	1, 2, 6	Fosfato pH 7 (50 mM)	10	200	20	1	Sonicador ultrassônico	Não
						22 (sem expressão)	-	-	-	-	-	-	-
		0,4 - 0,6	0,03	LB, M9	37	4 (sem expressão)	-	-	-	-	-	-	-
						6	Tris HCl pH 9 (50 mM)	10	200	20	1 (+CTAB 1% v/v) ¹	Sonicador ultrassônico	Sim (IMAC e Troca iônica)

* O detergente não iônico Triton X-100 foi utilizado somente na segunda etapa de extração, quando realizada.

‡ Todas as extrações realizadas continham lisozima a $\leq 1 \text{ mg/mL}^{-1}$.

† French press: as células foram rompidas com 3 a 5 passagens do extrato a 6894,76 KPa. Sonicador ultrassônico: As células foram rompidas por ultrassom com Sonicador ultrassônico Branson Sonifier 250 (Branson, Connecticut, EUA), com razão cíclica de 20% com 15 ciclos de 10 pulsos e intervalos de 10s entre cada ciclo.

‡: dados não mostrados.

¹: CTAB (brometo de cetrônio): detergente catiônico utilizado na solubilização de CIs.