UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

LORENZO E. R. P. P. L. D. B. BRIGANTI

Estudos estruturais e biofísicos de enzimas degradadoras de
hemicelulose visando suas aplicações na biotecnologia industrial

São Carlos
2021

LORENZO E. R. P. P. L. D. B. BRIGANTI

Estudos estruturais e biofísicos de enzimas degradadoras de hemicelulose visando suas aplicações na biotecnologia industrial

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

Área de concentração: Física Aplicada
Opção: Física Biomolecular
Orientador Prof. Dr. Igor Polikarpov

Versão Corrigida
(versão original disponível na Unidade que aloja o Programa)

São Carlos
2021

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

Briganti, Lorenzo E.R.P.P.L.D.B

Estudos estruturais e biofísicos de enzimas degradadoras de hemicelulose visando suas aplicações na biotecnologia industrial / Lorenzo E.R.P.P.L.D.B Briganti; orientador Igor Polikarpov - versão corrigida -- São Carlos, 2021.
196 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Aplicada Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2021.

1. Cristalografia de raios-X. 2. Dinâmica molecular. 3. Bioquímica de proteínas. 4. Biologia estrutural. 5. Biofísica. I. Polikarpov, Igor, orient. II. Título.

“Somos a soma de todos que vieram antes de nós.” Dedico este trabalho a todos que vieram antes de mim, em especial à minha mãe, que sempre acreditou em mim e me incentivou, e ao meu pai, um grande homem e grande exemplo pra mim.

AGRADECIMENTOS

Gostaria de agradecer ao Professor Igor Polikarpov, pela oportunidade de realizar este trabalho, por acreditar no meu potencial, pela confiança depositada e por me ensinar tanto nesse tempo.

Um agradecimento especial à minha mãe, que sempre acreditou em mim, mesmo nos momentos mais difíceis e mais incertos, sempre me incentivou e fez com que tudo isso fosse possível. Se o título fosse compartilhado, com certeza seria com ela.

Agradeço pela melhor companhia que a vida poderia me dar, em todos os momentos, em todos os sentidos, grande incentivadora, que torna a vida mais leve e as coisas mais fáceis, Jacqueline Mattos. Não foram poucos os momentos difíceis, e você sempre esteve do meu lado. Estendo meu agradecimento a toda a sua família, de coração.

Agradeço ao Professor Alessandro Nascimento, pelo incentivo, pelos ensinamentos, discussões sobre o trabalho e colaborações, e ao Professor Richard Garratt, pelos valiosos ensinamentos tanto na sala de aula quanto fora dela, pelo exemplo, e por inspirar, tornando coisas complexas (como por exemplo simetria em cristalografia) tão agradáveis.

Agradeço aos meus colegas de laboratório, pelos bons momentos vividos, risadas e colaborações também: Atílio, por todos os valiosos conselhos, pela amizade e pelas risadas; Marco, pela amizade, por ensinar tanto, sobre tudo, e me incentivar bastante; Evandro, pelos ensinamentos, conselhos, muitas risadas e amizade incondicional, além das colaborações; Danilo, pelos ótimos momentos no laboratório e fora dele, ser um grande amigo e grande companheiro de trabalho; Aline, por todo o apoio, pela amizade, pelos bandejões e por escutar, nos momentos difíceis; Melissa, pela amizade, por ser essa pessoa fantástica e sempre alegrar o dia de todos; Ana Elisa, pela amizade, confiança e confidência nos vários momentos difíceis no laboratório. São os melhores amigos que eu poderia ter em São Carlos, amizades que eu vou levar para o resto da vida. Agradeço à Vanessa, pela confiança, amizade, incentivo, apoio e pelas colaborações, e estendo os agradecimentos à Ana Gabriela e ao Caio também. Agradeço à Aissata, que mesmo sem falar muito português ilumina o dia da gente com suas risadas e seus comentários. Agradeço também ao Vinícius pelas conversas, pelas discussões acadêmicas e aleatórias, e por alegrar o ambiente do laboratório. Agradeço ao Darlan pela amizade, pelas dicas e também pelas risadas, que não foram poucas. Agradeço também aos alunos e ex-alunos do Professor Alessandro, grandes amizades que carregarei para sempre: Victor Rabesquini, Erika Chang, Stheffany Colmanetti e Raíssa Gutierrez. Foi

muito bom conviver e aprender com todos vocês. Agradeço também ao Adílson Wanderley, grande amigo do Instituto, do futebol e da vida.

Aos técnicos e especialistas do laboratório: Livia, Josimar e Possatto. Um agradecimento especial a Livia Manzine por todos os ensinamentos sobre Bioquímica e Biofísica, por sempre colaborar com o meu trabalho desde o meu primeiro dia no laboratório, sendo a melhor amiga que alguém poderia ter nesse período.

Ao IFSC, por toda a infraestrutura disponível para os alunos, e a todos os funcionários do Instituto por todo o auxílio prestado nesse tempo. Um agradecimento especial ao Ricardo e também ao Silvio da Seção de Pós Graduação, pela ajuda e orientação quanto ao que fazer, que foram fundamentais. Ao Gilmar, ao Sebá e ao Luciano do STI, que foram essenciais para este trabalho. À equipe da Biblioteca, sempre tão prestativa e eficiente. Um agradecimento especial à Neusa, por toda o apoio e atenção, pois foi fundamental para a confecção da tese. Sem vocês tudo isso seria muito mais difícil.

À toda a equipe do CNPEM - da SAU, LNNano, LNBio e LNLS-Sirius, em especial à Dra. Ana Zeri, um exemplo para todos nós, como pessoa e como profissional também, sem o seu apoio a parte cristalográfica desse trabalho não seria a mesma, meu conhecimento não seria o mesmo, e eu também não seria o mesmo. Agradeço ainda ao Dr. Andrey e ao Dr. Araújo, cujo auxílio foi fundamental para a realização desse trabalho. A Manacá tem um futuro brilhante, com uma equipe tão competente.

A quem me trouxe alegria e amor nesse período tão difícil: Zuca. Feliz é quem pode chegar em casa e receber o amor incondicional e efusivo de sua companhia canina. Uma prova de que o amor dura pra sempre.

A todos os servidores da USP, pela prestatividade, apoio e auxílio, sempre que necessário. Vocês fazem a Universidade ser o que ela é: um centro de excelência. Agradeço ainda aos membros da banca, pelos valiosos conselhos e correções desta tese.

A todos que foram responsáveis, direta ou indiretamente, para que eu chegasse até aqui.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) Código de financiamento 001.

" Out of the cradle
onto dry land
here it is
standing:
atoms with consciousness;
matter with curiosity.

Stands at the sea,
wonders at wondering: I
a universe of atoms
an atom in the Universe."

Richard Feynman

RESUMO

BRIGANTI, L.E.R.P.P.L.D.B. **Estudos estruturais e biofísicos de enzimas degradadoras de hemicelulose visando suas aplicações em biotecnologia industrial.** 2021. 195 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2021.

As paredes celulares das plantas são estruturas que possuem diversas funções, entre elas conferir sustentação e proteção às mesmas. Essas paredes são formadas basicamente por celulose, hemicelulose e lignina, além de outros compostos minoritários. A hemicelulose, o segundo componente mais abundante, que responde por cerca de 30% da parede celular, é composta por xilano, um polímero de xilose com decorações como arabinose e ácido metil-glucurônico, o que a torna um material recalcitrante. Para a degradação da hemicelulose e aproveitamento dos açúcares contidos nela, se faz necessária a hidrólise eficiente desse componente. O presente estudo teve como alvos enzimas degradadoras de xilano com a finalidade de produção de compostos de interesse industrial, como xilo-oligossacarídeos (XOS), xilose, que seria o monômero da cadeia de xilano, e xilulose, um isômero da xilose. Os XOS são conhecidos por serem prebióticos, com grande importância na indústria alimentícia e promovendo saúde e bem-estar. Já a xilulose é uma pentose aproveitada na produção de etanol por fermentação industrial, ao contrário da xilose, que não pode ser aproveitada nas mesmas condições. As enzimas estudadas foram as glicosil-hidrolases, ou GHs, sendo mais especificamente as GHs 10, 11 e 43, e a xilose isomerase. Foram utilizadas técnicas físicas para caracterização destas enzimas e entendimento sobre o mecanismo das mesmas. A GH10, que se mostrou uma boa produtora de XOS, teve desvendada o seu padrão de clivagem enzimático e conseguiu-se compreender a relação entre a estrutura cristalográfica da enzima e o perfil de produtos gerados pela mesma. A GH11, que possui um perfil enzimático incomum, indicando cooperatividade, teve a sua estrutura cristalográfica resolvida e, através de simulações de dinâmica molecular, compreendeu-se o papel do sítio secundário de ligação na estabilização da enzima e na diminuição da energia livre de interação com o ligante. A GH43, que possui afinidade maior por XOS maiores e cujo mutante F507A não apresenta atividade catalítica, teve sua estrutura cristalográfica resolvida, e através de análises estruturais e simulações de dinâmica molecular, compreendeu-se que as dobras que o ligante fazia em direção à enzima poderiam resultar em interações extras em XOS maiores, além do fato da mutação F507A estreitar o bolso catalítico da enzima, impedindo a acomodação de um ligante e consequentemente a atividade catalítica da mesma. Para a xilose isomerase, foram

feitos extensivos estudos de caracterização, com a determinação de parâmetros como temperatura ótima (60 °C), pH ótimo (cerca de 8.0), termoestabilidade (50% de atividade residual após 72h a 50 °C), além de raio de giro (36 Å), determinação da estrutura cristalográfica, além de análises de conservação dos resíduos no grupo. Ensaios preliminares de fermentação foram realizados, mas outros testes, com diferentes técnicas de fermentação, devem ser realizados para encontrar a melhor forma de utilizar industrialmente essa enzima.

Palavras-chave: Cristalografia de raios-X. Dinâmica molecular. Bioquímica de proteínas. Biologia estrutural. Biofísica.

ABSTRACT

BRIGANTI, L.E.R.P.P.L.D.B. **Structural and biophysical studies in hemicellulose-degrading enzymes for industrial biotechnology applications**. 2021. 195 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2021.

Plant cell walls are structures that have several functions, including providing support and protection to them. These walls are basically formed by cellulose, hemicellulose and lignin, in addition to other minority compounds. Hemicellulose, the second most abundant component, which accounts for about 30% of the cell wall, is composed of xylan, a xylose polymer with decorations such as arabinose and methyl-glucuronic acid, which makes it a recalcitrant material. For the degradation of hemicellulose and the use of the sugars contained in it, the efficient hydrolysis of this component is necessary. The present study aimed at xylan degrading enzymes with the purpose of producing compounds of industrial interest, such as xylo-oligosaccharides (XOS), xylose, the monomer of the xylan chain, and xylulose, an isomer of xylose. XOS are known to be prebiotics, with great importance in the food industry and promoting health and well-being. Xylulose, on the other hand, is a pentose used in the production of ethanol by industrial fermentation, unlike xylose, which cannot be used under the same conditions. The enzymes studied were glycosyl hydrolases, or GHs, more specifically GHs 10, 11 and 43, and xylose isomerase. Physical techniques were used to characterize these enzymes and understand their mechanism. GH10, which proved to be a good producer of XOS, had its enzymatic cleavage pattern unraveled and it was possible to understand the relationship between the crystallographic structure of the enzyme and the profile of products generated by it. GH11, which has an unusual enzymatic profile indicating cooperativity, had its crystallographic structure resolved and, through molecular dynamics simulations, the role of the secondary binding site in stabilizing the enzyme and lowering the free energy of interaction was comprehended and described with the binder. GH43, which has greater affinity for larger XOS and whose F507A mutant does not present catalytic activity, had its crystallographic structure resolved, and through structural analysis and molecular dynamics simulations, it was understood that the folding that the ligand made towards the enzyme could possibly result in extra interactions in larger XOS, in addition to the fact that the F507A mutation narrows the catalytic pocket, preventing the accommodation of a ligand and consequently abolishing the enzymatic catalytic activity. For xylose isomerase, extensive characterization studies were carried out, with the determination of parameters such as

optimal temperature (60 °C), optimal pH (about 8.0), thermostability (50% residual activity after 72h at 50 °C), in addition to radius of gyration (36 Å), determination of the crystallographic structure, as well as analyzes of conservation of residues in the group. Preliminary fermentation tests have been carried out, but other tests, with different fermentation techniques, must be carried out to find the best way for industrially usage of this enzyme.

Keywords: X-ray crystallography. Molecular dynamics. Protein biochemistry. Structural biology. Biophysics.

LISTA DE FIGURAS

- Figura 1 - Representação da parede celular vegetal e seus componentes. (A) Estruturação da parede e posicionamento das estruturas. (B) Distribuição e arranjo dos constituintes da parede celular. 31
- Figura 2 - Representação estrutural dos diferentes tipos de heteroxilano. (A) O-Acetil-4-O-metilglucuronoxilano de vegetais de “madeira dura”. (B) Arabino-4-O-metilglucuronoxilano de vegetais de “madeira mole”. (C) Arabinoxilano de cereais. 32
- Figura 3 - Modo de ação das Hidrolases de glicosídeos (GHs) em relação ao local de clivagem na cadeia de xilano. 34
- Figura 4 - Ilustração esquemática de um diagrama de fase de cristalização de proteínas. Parâmetros ajustáveis incluem concentração de precipitante ou aditivo, pH e temperatura. Os quatro principais métodos de cristalização, com uma rota diferente para alcançar as zonas de nucleação e metaestável, assumindo que o parâmetro ajustável é concentração de precipitante, são representados: (i) Microbatch, (ii) Difusão de vapor, (iii) Diálise e (iv) Difusão de interface livre (FID). Os círculos pretos preenchidos representam as condições iniciais. 38
- Figura 5 - Representação gráfica da cela unitária (A) e da rede cristalina (B), com seus vetores de translação. 39
- Figura 6 - As 14 redes de Bravais. 41
- Figura 7 - Representações de planos cristalinos. (A) eixos intersectados por plano hkl, (B) plano (200), (C) plano (110), (D) plano (111) e direção [111]. 44
- Figura 8 - Representação geométrica da Lei de Bragg. Duas ondas em fase, onde d é a distância entre planos atômicos. 47
- Figura 9 - Diagrama de Argand, representando o fator de estrutura F e a fase associada, evidenciando o caráter centrossimétrico da difração. 49
- Figura 10 - Representação dos tipos de energia potencial (V), com suas respectivas equações. 59
- Figura 11 - Expressão e solubilidade de TcXyn10A analisada por SDS-PAGE. (A) Etapa de expressão. M: marcador molecular; 1: Frações retiradas com 0h de indução; 2: 16h após indução; Após a lise celular 3: fração insolúvel; 4: fração solúvel. (B) Etapa final de purificação. (M) marcador molecular. 70
- Figura 12 - Figura 12: Padrão de clivagem TcXyn10A. Os perfis do produto liberado por TcXyn10A sob condições ótimas de enzima de diferentes XOS medidos por HPAEC: (A) xylotriose (X3), (B) xylotetraose (X4); (C) xilopentaose (X5) (D) xilohexaose (X6). Xn em cada perfil indica o padrão de xilooligossacarídeos variando de xilose a hexaose. X3 não é clivado pela

enzima, enquanto X4 é clivado em X2. Tanto X5 quanto X6 são degradados pela enzima predominantemente em X2 e X3.72

Figura 13 - Avaliação da produção de xilo-oligossacarídeos por *TcXyn10A*. (A) Produção de XOS a partir de glucuronoxilano de madeira de faia medida por (A) HPAEC; e (C) MALDI-TOF, e arabinoxilano de centeio medido por (B) HPAEC; e (D) MALDI-TOF. A enzima libera uma ampla gama de XOS de glucuronoxilano de madeira de faia e de arabinoxilano de centeio. Xn em cada perfil indica o padrão de XOS variando de xilose a xilohexaose.73

Figura 14 - Estrutura cristalográfica. (A) Estrutura 3D de *TcXyn10A* revelando a estrutura conservada de barril α/β (barril TIM), com xilose (XYL) no sítio ativo sobreposto com o mapa Omit correspondente no nível de 1,2 sigma. Os resíduos do sítio ativo na vizinhança da xilose são mostrados como bastões e as quatro alças são coloridas em vermelho escuro, azul, amarelo e roxo. O programa UCSF Chimera v1.13 (www.cgl.ucsf.edu/chimera)¹⁰⁷ foi usado para criar esta figura; (B) Alinhamento estrutural baseado em sequência de *TcXyn10A*, *CmXyn10B* e *SoXyn10A*. Os resíduos catalíticos são destacados em triângulos vermelhos; Os resíduos do site ativo são destacados em azul; Os resíduos de interação de arabinose são representados em triângulos rosa, enquanto os resíduos de interação de MeGlcA são representados em triângulos amarelos. Os quatro loops são mostrados em vermelho escuro, azul, amarelo e roxo, respectivamente. (C) Resíduos de aminoácidos na vizinhança de XYL observados em *TcXyn10A* (verde) são mostrados. Resíduos correspondentes de estruturas homólogas de *CmXyn10B* (PDB ids 1UR1 e 1UQZ)⁸⁸ e estrutura de *SoXyn10A* (PDB id 1V6V)¹⁰² são mostrados em amarelo, azul e marrom, respectivamente. As distâncias de interação com a xilose são fornecidas em Angstroms; (D) A sobreposição estrutural de *TcXyn10A* (mostrado em azul) com dois outros GH10s, *CmXyn10B* (PDB id 1UR1 e 1UQZ, colorido em amarelo e azul, respectivamente) e *SoXyn10A* (PDB id 1V6V, representado em marrom) é dado na representação de fita.77

Figura 15 - Bases estruturais do reconhecimento de *TcXyn10A* de xilano não decorado e glucuronoxilano. (A) Vista frontal da enzima, com a fenda catalítica em seu centro e X6 no sítio ativo, com os dois resíduos catalíticos (Glu132 e Glu239) representados em vermelho. Os subsítios da fenda do sítio ativo são numerados de -3 a +3, (B) MeGlcAX3, co-cristalizado com *SoXyn10A*, doca bem nos subsítios negativos *TcXyn10A*, (C) As cadeias laterais dos resíduos interagindo com o MeGlcAX3 (em verde) e o as distâncias para os resíduos interagentes, em Å, são mostradas. Resíduos correspondentes da estrutura de *SoXyn10A* são mostrados para comparação (em marrom), (D) MeGlcAX3, co-cristalizado com *CmXyn10B* posicionado nos subsítios positivos do sítio ativo, (E) As distâncias de interação de MeGlcAX3, para os resíduos vizinhos de *TcXyn10A* (em verde) nos subsítios positivos (+1 a +3). A posição e orientação dos resíduos correspondentes da estrutura da GH10 de *C. mixtus* são mostradas em azul para comparação. Os resíduos em (C) e (E) são numerados e coloridos de acordo com as sequências correspondentes.80

- Figura 16 - Representação esquemática de arabinoxilano e glucuronoxilano. As unidades de D-xilose de uma estrutura de xilano são mostradas como estrelas laranja, as decorações de L-arabinose como estrelas verdes e as decorações de ácido D-glucurônico são mostradas como diamantes divididos em azul e branco, seguindo a Nomenclatura de Símbolo para Representação Gráfica de Glicanos.¹¹²⁻¹¹³ 82
- Figura 17 - Interações putativas de *TcXyn10A* com arabinoxilano. (A) AraX2 observado em *CmXyn10B* ajustado dentro dos subsítios negativos do sítio ativo de *TcXyn10A* (de -2 a -1), com decoração de arabinose na posição -2; (B) As distâncias entre AraX2 refinado dentro do sítio ativo *TcXyn10A* (em verde) para os resíduos vizinhos são mostradas. Os resíduos correspondentes da estrutura homóloga *CmXyn10B* são dados a laranja; (C) AraX3 da estrutura homóloga *CmXyn10B* ancorada nos subsites negativos (-3 a -1) do sítio ativo *TcXyn10A*. A decoração arabinose está na posição -2, (D) A mesma sobreposição que em (C) mostrando as distâncias para os resíduos de interação de *TcXyn10A* (mostrado em verde) e *CmXyn10B* (representado em roxo). Os resíduos em (B) e (D) são coloridos e numerados de acordo com a respectiva enzima. (E) Arabinoxilano nos subsítios positivos (aglicona) (+1 a +3). As decorações de arabinose são modeladas nas posições +1, +2 e +3, (F) A mesma sobreposição que em (E), exibindo as distâncias para os resíduos de interação do sítio ativo de *TcXyn10A*; (G) Vista lateral de (E), mostrando o espaço disponível para decoração Ara no local +2; (H) Distâncias dos resíduos do sítio ativo ao ligante como mostrado em (G). 83
- Figura 18 - Padrão de clivagem de acordo com estudos estruturais de subsítios de *TcXyn10A*. (A) Padrão de clivagem de xilohexaose (X6) e xilotetraose (X4) e seus possíveis produtos; (B) Glucuronoxilano, com decoração MeGlcA nos subsítios -3 e +1, e os produtos; (C) Padrão de clivagem de arabinoxilano e seus possíveis produtos, de acordo com a posição da decoração de arabinose. As unidades de D-xilose de uma estrutura de xilano são mostradas como estrelas laranja, as decorações de L-arabinose como estrelas verdes e as decorações de ácido D-glucurônico são mostradas como diamantes divididos em azul e branco, seguindo a Nomenclatura de Símbolo para Representação Gráfica de Glicanos.¹¹²⁻¹¹³ 85
- Figura 19 - Purificação de *PxXyn11B*. O gel SDS a 12% mostra os resultados das etapas de purificação da *PxXyn11B*. M: marcador de baixo peso molecular; Lisado: a fração solúvel do lisado; F.I.: fração insolúvel do lisado (pellet); Flow-through: fluxo observado na purificação de IMAC; Lv: fração de lavagem; 250 mM imidazol: eluição de proteína com imidazol 250 mM. 92
- Figura 20 - Estrutura cristalográfica de *PxXyn11B*. (A) Estrutura 3D de *PxXyn11B* colorida de azul (região N-terminal) a vermelho (região C-terminal). Hélices α e fitas β , conforme definido por PDBSum são indicadas, e também os resíduos catalíticos. (B) Vista lateral de (A), com as partes N-terminal, C-terminal, palma, dedos, polegar e cordão da proteína indicadas. 95
- Figura 21 - Alinhamento de sequência baseado em estrutura. (A) As sequências de AnXyn11A, CTX, XynII, rXynA, EnXyn11A, TfGH11, NFX e XlnB2

foram alinhadas com a sequência *PxXyn11B* e o NTR (região N-terminal), resíduos catalíticos e resíduos do SBS são destacados. (B) Superposição estrutural das enzimas acima mencionadas. Azul marinho: *AnXyn11A*; Roxo: *CTX*; Verde escuro: *XynII*; Ciano escuro: *PxXyn11B*; Vermelho escuro: *rXynA*; Dourado: *EnXyn11A*; Laranja: *TfGH11*; Vermelho: *NFX*; Ciano: *XlnB2*. (C) Vista lateral de (B). (D) Vista posterior de (B).96

Figura 22 - Análise estrutural do sítio ativo de *PxXyn11B*. (A) Estrutura cristalográfica de *PxXyn11B* com X6 de *Trichoderma reesei* (PDB id 4HK8) sobreposta. (B) Sobreposição de *PxXyn11B* (ciano escuro) e GH11 de *Trichoderma reesei* (verde escuro). (C) Subsítios em *PxXyn11B*, com os respectivos resíduos. Os triângulos vermelhos indicam o local de clivagem. Os resíduos xilanopirosídicos são representados como estrelas amarelas, de acordo com a nomenclatura de símbolos dos glicanos.^{112, 113} (D) Zoom nos sítios ativos, com os resíduos exibidos com as respectivas cores.99

Figura 23 - Sobreposição das estruturas *PxXyn11B* e *XBS1*. (A) Vista frontal, mostrando o sítio catalítico e o SBS na parte traseira. (B) Vista lateral de (A), com resíduos no sítio ativo indicados. (C) Ampliação da região do SBS, com os resíduos envolvidos destacados e as distâncias entre os resíduos de *PxXyn11B* e os resíduos de xilopiranosídeo indicadas.101

Figura 24 - (A e B) Duas vistas de *PxXyn11B* ligado a xilotetraose (X4). As figuras mostram a superposição de 500 quadros tirados em intervalos de 1 ns durante a simulação de equilíbrio de DM. Os subsítios do sítio ativo (-2 a +2) são indicados abaixo do painel (A). (C e D) Duas visualizações de *PxXyn11B* ligado a X4 no sítio ativo e a uma xilotriose (X3) no SBS. Ambos X3 e X4 são mostrados na forma de palitos. As linhas indicam a mudança na mobilidade perto da fita $\beta 5$. (E) Energia de interação (potencial) calculada entre a enzima e o substrato (X4) para o complexo binário (linha preta) e para o complexo ternário (linha vermelha), ou seja, a enzima ligada a X3 no SBS. Para efeito de comparação, as energias de interação entre a enzima e o X3 no SBS também são mostradas em linhas azuis.102

Figura 25 - Comparação da mobilidade para o complexo *PxXyn11B*-X4 (topo) e complexo ternário *PxXyn11B*-X4 / X3 (parte inferior). Quatro visualizações são mostradas para cada sistema simulado. O X4 e o X3 são mostrados em palitos e as estruturas das proteínas são coloridas de acordo com sua mobilidade em uma escala de azul a vermelho.103

Figura 26 - Comparação entre o complexo binário *PxXyn11B*-X4 (azul) e o complexo ternário *PxXyn11B*-X4 / X3 (laranja). Estruturas representativas do cluster mais populoso na trajetória de DM são comparadas. Os átomos de hidrogênio foram omitidos para maior clareza. X4 e X4 são mostrados em palitos e os subsítios são numerados na parte inferior da figura. O reposicionamento de N62, Q173 e Y172 no complexo binário cria um espaço adicional que permite que os açúcares no sítio positivo flutuem, afrouxando a interação com a enzima.106

- Figura 27 – Simulação de DM de *PxXyn11B* com X5 no sítio ativo e X3 no SBS. (A) Quadros representativos: primeiro painel no tempo zero, painel 2 após 130 ns, painel 3 após 220 ns, painel 4 após 250 ns, onde X3 atinge um local adicional, próximo ao subsítio -2, estabilizando a região do polegar da enzima, conformação observada até o final da simulação, após 500 ns (quinto painel). (B) Análise da energia potencial de interação entre *PxXyn11B* e uma molécula X3, inicialmente ligada ao SBS (linha vermelha) e também entre *PxXyn11B* e uma molécula X5 ligada ao sítio ativo nos subsítios -2 a +3 ao longo de 500 ns. 107
- Figura 28 - Estrutura cristalográfica. (A) Estrutura 3D tetramérica de *EcXyl43* F507A, com as cadeias A, B, C e D representadas em vermelho, verde, amarelo e azul, respectivamente, e as esferas verdes indicando o íon Cálcio. A figura inferior indica uma rotação de 180 graus ao longo do eixo y. (B) Monômero de *EcXyl43* F507A, mostrando o domínio catalítico da β -propeller em ciano e o domínio auxiliar sanduíche- β em azul escuro. Uma molécula de glicerol foi observada no sítio ativo, densidade eletrônica representada pelo mapa 2mFo - DFC omit com contorno de 1,5 σ . (C) Mesmo monômero observado em (b), em representação de superfície, com o túnel catalítico indicado pelo círculo branco tracejado..... 117
- Figura 29 - Alinhamento de sequência baseado em estrutura. (A) As sequências de *GsXyl*, *SrXyl* e *BpXyl* foram alinhadas com a sequência *EcXyl43*, e a alça, domínio auxiliar, resíduos catalíticos e resíduos de sítio ativo são destacados. Os resíduos do sítio de ligação do cálcio são destacados em triângulos verdes. (B) Alinhamento estrutural das enzimas acima mencionadas. *EcXyl43* é representado em ciano, *GsXyl* em verde, *SrXyl* em marrom e *BpXyl* em amarelo. O loop contendo a mutação F507A é destacado em um círculo preto tracejado. (C) Ampliação da região do sítio ativo, mostrando as distâncias entre F506 e E186, e A507 e E186, e as cadeias laterais de resíduos de aminoácidos nas posições correspondentes das enzimas sobrepostas. 120
- Figura 30 - Sítio de ligação do cálcio em *EcXyl43* F507A. O íon cálcio é representado como uma esfera verde. Os resíduos envolvidos na interação cálcio-enzima são destacados e as distâncias são mostradas em Å. 122
- Figura 31 - Superposição de *EcXyl43* F507A e *GsXyl*. (A) Sobreposição de *EcXyl43* (azul) e *GsXyl* (verde), com Xilobiose de *GsXyl* no sítio ativo (quadrado tracejado). As esferas verdes representam o Cálcio em ambas as enzimas. (B) Zoom do sítio ativo (quadrado tracejado em (A)), com os resíduos (verde para *GsXyl* e azul para *EcXyl43*) e suas distâncias aos átomos do ligante. 123
- Figura 32 - Docking molecular de X6 no sítio ativo de *EcXyl43*. (A) Vista lateral de *EcXyl43* com X6 em seu site ativo (quadrado tracejado). (B) Zoom do sítio ativo, mostrando distâncias de interação entre enzima e substrato. 124
- Figura 33 - Dinâmica molecular de *EcXyl43* com X6 ancorado no sítio ativo. (A) Posição inicial de *EcXyl43* e X6 no sítio ativo, perpendicular à enzima (vista frontal). (B) Zoom de (A), mostrando a distância do último resíduo

xilosídico (X6) a H189 (23,8 Å). (C) O mesmo que (A), com rotação de 90 graus ao longo do eixo y (vista lateral). (D) Zoom de (C), mostrando a distância de X6 a H189. (E) Posição intermediária de *EcXyl43* com X6 no sítio ativo, mostrando o ligante curvado em direção à alça. (F) Zoom de (E), mostrando a distância do último resíduo xilosídico (X6) a H189 (3,4 Å). (G) Vista lateral de (E), com rotação de 90 graus ao longo do eixo y. (H) Zoom de (G). (I) Gráfico da distância entre o último resíduo xilosídico (X6) e H189 ao longo da simulação (500 ns). Os arquivos PDB foram produzidos com VMD e as figuras foram produzidas com Chimera.....126

Figura 34 - Comparação estrutural de fendas catalíticas. (A) *GsXyl* (PDB id 2EXJ) em verde claro (domínio catalítico) e verde escuro (domínio auxiliar), com Xilobiose no sítio ativo. (B) Zoom do sítio ativo, mostrando o túnel com o ligante dentro. (C) *EcXyl43* em azul claro (domínio catalítico) e azul escuro (domínio auxiliar), com glicerol no sítio ativo. (D) Zoom do sítio ativo, mostrando o túnel com o ligante dentro. (E) *SrXyl* (PDB id 3C2U) em rubi (domínio catalítico) e vermelho (domínio auxiliar), com BXP no sítio ativo. (F) Zoom do sítio ativo, mostrando o túnel com o ligante dentro. (G) *BpXyl* (PDB id 5ZQX) em amarelo (domínio catalítico) e dourado (domínio auxiliar), com X2 no sítio ativo. (H) Zoom do sítio ativo, mostrando o túnel com o ligante dentro. (I) Sobreposição de *GsXyl* (verde), *EcXyl43* (azul claro), *SrXyl* (vermelho) e *BpXyl* (amarelo) com xilobiose no sítio ativo.128

Figura 35 - Trajetórias de *EcXyl43* ao longo da Dinâmica Molecular. Representação de cluster de (A) *EcXyl43* com X6 ancorado no sítio ativo, (B) *EcXyl43* na forma Apo e (C) *EcXyl43* F507A na forma Apo ao longo de 500 ns de simulação de DM com intervalo de 1 ns, com estruturas coloridas por RMSF de azul (mínimo) a vermelho (> 3 Å). As setas indicam o loop 502-510, contendo Phe507. (D) RMSD dos sistemas ao longo da Dinâmica. (E) RMSF dos sistemas ao longo da Dinâmica, por resíduo. (F) RMSF dos sistemas, comparando a região de loop dos resíduos 502 - 510 (onde a mutação F507A está localizada).130

Figura 36 - Distâncias dos resíduos F507 e A507 a E186 ao longo da Dinâmica Molecular. (A) Superposição de *EcXyl43* nativo (azul claro), F507A (ciano escuro) e *GsXyl* (verde). (B) Zoom do sítio ativo, mostrando as posições dos resíduos nas enzimas. (C) Gráfico da distância entre os resíduos F507 e E186 em *EcXyl43* e A507 e E186 em *EcXyl43* F507A ao longo da DM.131

Figura 37 Expressão e purificação de *SdXI*. (A) SDS-PAGE indicando o peso molecular em kDa, o marcador de peso molecular (M), o extrato celular (1) e a fração após a purificação (2). (B) perfil cromatográfico de gel-filtração, mostrando o pico único referente a *SdXI*.141

Figura 38 - Perfil de temperatura ótima (a) e pH ótimo (b) de *SdXI*.142

Figura 39 - Atividade relativa ao longo do tempo (horas), em diferentes temperaturas.....143

Figura 40 - Inibição de *SdXI* por Xilitol, em diferentes concentrações.144

Figura 41 - Gráfico de Ultracentrifugação analítica.145

Figura 42 - Modelo por homologia de <i>SdXI</i> gerado pelo programa Swiss-Model. Observa-se o homotetrâmero, composto por 4 cadeias iguais, cada uma com uma cor. As esferas roxas representam o íon Mg^{2+}	147
Figura 43 - Perfil de DLS da enzima <i>SdXI</i>	148
Figura 44 - Espectro de dicroísmo circular em função da elipticidade (mdeg) da <i>SdXI</i> em tampão Na_2PO_4 25 mM NaCl 25 mM pH 8,0 a 25 °C.....	148
Figura 45 - Curva de espalhamento de raios-x a baixo ângulo para <i>SdXI</i> (a) e função $P(r)$ (b).	150
Figura 46 - Modelo 3D de <i>SdXI</i> construído pelo DAMAVER (A), com estrutura do modelo por homologia de <i>SdXI</i> sobreposta (B). (C) e (D) representam diferentes rotações ao longo do eixo X e Y, respectivamente.....	151
Figura 47 - Curvas de espalhamento de <i>SdXI</i> , os círculos abertos representando os dados experimentais, a linha vermelha representando o modelo Dam, a linha azul clara representando o modelo Gasbor, e a linha azul escura o modelo por homologia.....	152
Figura 48 - Resultado do alinhamento ESPRIPT, indicando regiões conservadas (caixas amarelas e vermelhas), região de abertura do anel (setas verdes) e região de ligação de metal (seta vermelha para o sítio 1 e seta azul para o sítio 2). O ácido glutâmico (E274 em <i>SdXI</i>) no final da folha β 10 com duas setas coordena simultaneamente dois íons metálicos.....	154
Figura 49 - Gráficos da estrutura secundária de <i>PirXI</i> (a), <i>SdXI</i> (b) e <i>A. missouriensis</i> XI (c).....	156
Figura 50 - Monômero <i>SdXI</i> , com os domínios principais indicados. Painel inferior: Representação das regiões de estrutura secundária de <i>SdXI</i> , com a região N e C terminal destacadas.	160
Figura 51 - Oligomerização em <i>SdXI</i> . Em detalhes, dímeros (EA) e (DC), tetrâmero (ACDE), interface dímero-dímero (ED) e esquema de oligomerização, mostrando área de interações, ligações (Hidrogênio e Sal - vermelho e azul, respectivamente) e contatos não ligados (laranja). O esquema no canto inferior direito indica conjuntos diméricos AE e CD (interface monômero - monômero); Interface tetramérica (AC e ED) e interações cruzadas (AD e EC).....	162
Figura 52 - Monômero <i>SdXI</i> , com 2 sítios de metal. Os painéis superiores indicam a distância atômica e os resíduos envolvidos na coordenação do metal. Os painéis inferiores mostram o mapa de densidade eletrônica (Omit) em azul, corroborando a posição correta do metal.....	163
Figura 53 - Superposição de diferentes estruturas XI. Em marrom: <i>SdXI</i> , em laranja: <i>PirXI</i> , ambas as classes II, enquanto <i>S. rubiginosus</i> e <i>A. missouriensis</i> , ambas de classe I, são representadas em azul-claro e roxo, respectivamente.	164

Figura 54 – Análise de conservação em <i>SdXI</i> . (A), (B), (C), (D) e (E) ilustram diferentes posições da molécula. Legenda indicando do resíduo mais variável (azul) ao mais conservado (vermelho).	165
Figura 55 - Monômero de <i>SdXI</i> obtido da análise CONSURF na interface dimérica AE (A) e tetramérica AC (B).....	166
Figura 56 - Análise filogenética molecular pelo método de máxima verossimilhança. O organismo fonte de <i>SdXI</i> , <i>Saccharophagus degradans</i> , é destacado em vermelho.....	167

LISTA DE TABELAS

Tabela 1 -	Sistemas cristalinos e suas especificações quanto a comprimento e ângulos.....	40
Tabela 2 -	Sistemas cristalinos, grupos pontuais e seus respectivos grupos espaciais.....	43
Tabela 3 -	Hidrólise enzimática de diferentes fontes de xilano em 24 horas.	74
Tabela 4 -	Parâmetros de coleta e refinamento cristalográfico de TcXyn10A.....	75
Tabela 5 -	Comparações entre TcXyn10A, CmXyn10B e SoXyn10A	78
Tabela 6 -	Coleta de dados cristalográficos de PxXyn11B e estatísticas do refinamento do modelo.....	93
Tabela 7 -	Enzimas utilizadas na comparação estrutural com PxXyn11B	98
Tabela 8 -	Coleta de dados cristalográficos de <i>EcXyl43</i> F507A e estatísticas do refinamento do modelo.....	115
Tabela 9 -	Resultados do PISA da interface de <i>EcXyl43</i> F507A e análises de conjuntos. As áreas de superfície (A.S.), enterrada (A.E.) e de interface (A.I.) são representadas em Å ² ; NRes: Número de resíduos; NHB: Número de potenciais ligações de hidrogênio; NSB: Número de pontes salinas.	118
Tabela 10 -	Parâmetros de alinhamento para as enzimas selecionadas de BLASTp e DALI. Valores obtidos em comparação com <i>EcXyl43</i>	119
Tabela 11 -	Distâncias interatômicas entre diferentes resíduos em <i>EcXyl43</i> , <i>GsXyl</i> , <i>SrXyl</i> e <i>BpXyl</i>	121
Tabela 12 -	Análise do túnel de GH43s com atividade de β-xilosidase. R.L.: Raio livre do gargalo do túnel (Å) - sem cadeia lateral; C.: Comprimento do túnel (Å); Hpa: Índice de hidropatia (-4,5 hidrofílico a +4,5 hidrofóbico); Hp h: Escala de hidrofobicidade (-1,140 hidrofílico a +1,810 hidrofóbico); Pos: resíduos positivos; Neg: resíduos negativos; Aro: resíduos aromáticos; Lig: ligante.	129
Tabela 13 -	Comparação dos elementos de estrutura secundária em <i>SdXI</i> computados com base na sequência protéica, no modelo por homologia e nos dados experimentais obtidos a partir dos experimentos de CD.	149
Tabela 14 -	Comparação entre diferentes técnicas e parâmetros obtidos para <i>SdXI</i> . * Rh refere-se ao raio hidrodinâmico obtido pelo experimento de DLS, e Rg refere-se ao raio de giração, obtido pelo experimento de SAXS. Estado se refere ao estado oligomérico da enzima em solução.	153
Tabela 15 -	Diferenças encontradas no alinhamento ESPRIPT entre as classes XI.....	155

Tabela 16 -	Comparação do conteúdo da estrutura secundária, em números absolutos, entre <i>SdXI</i> , <i>PirXI</i> e <i>A. missouriensis XI</i>	157
Tabela 17 –	Coleta de dados cristalográficos <i>SdXI</i> e estatísticas de refinamento do modelo.....	158
Tabela 18 -	Resultados do PISA para <i>SdXI</i> . M1: Monômero 1; M2: Monômero 2; Área: Área de interface ($\text{\AA}^2$); ΔiG : Ganho de energia livre de solvatação na formação da interface, em kcal / M; Nres: Número de resíduos na interface; NHb: Número de ligações de hidrogênio; NSb: Número de pontes de sal; CSS: Complexation Significance Score, que indica o quão significativa é a interface para a formação da estrutura quaternária. ...	161

LISTA DE ABREVIATURAS E SIGLAS

XOS	Xilo-oligossacarídeos
CAZy	Repositório de sequências de enzimas ativas em carboidratos (no inglês <i>Carbohydrate-Active enZYme database</i>)
EC	Número EC (no inglês <i>Enzyme Commission number</i>)
GH	Hidrolase de Glicosídeo
SAXS	Espalhamento de raios X a baixo ângulo
DM	Dinâmica molecular
MS	Espectrometria de Massas
FAB	Bombardeamento atômico rápido
ESI	Ionização por electrospray
MALDI	Ionização e dessorção a laser assistida por matriz
m/z	relação massa/carga
TOF	Tempo de voo
MS/MS	Espectrometria de massas in tandem
I	Intensidade medida
F	Fator de estrutura
F	Módulo do fator de estrutura
F ²	Fator de estrutura elevado ao quadrado
MR	Substituição Molecular (no inglês <i>Molecular Replacement</i>)
ML	Máxima verossimilhança (no inglês <i>Maximum Likelihood</i>)
SAD	difração anômala de comprimento de onda único
MAD	difração anômala de múltiplos comprimentos de ondas
D _{max}	Dimensão máxima da partícula
T _c	de <i>Thermobacillus composti</i>
LB	Caldo lisogênico (no inglês <i>Lysogenic Broth</i>)
DO ₆₀₀	Densidade ótica a 600 nm
IPTG	isopropil-β-D-tiogalactopiranosídeo
PMSF	fenilmetanossulfonil fluoreto
DTT	ditiotreitól
Ni-NTA	Resina de agarose contendo Ni ²⁺ e ácido nitrilotriacético
X5	Xilopentaose

X6	Xilohexaose
X2	Xilobiose
X3	Xilotriose
X4	Xilotetraose
PDB	Repositório de estruturas de biomoléculas (no inglês <i>Protein Data Bank</i>)
Cm	de <i>Cellvibrio mixtus</i>
So	de <i>Streptomyces olivaceoviridis</i>
MeGlcA	ácido 4-O-metil-D-glucurônico
SBS	Sítio secundário de ligação
Px	de <i>Paenibacillus xylanivorans</i>
Ec	de <i>Enterobacter cloacae</i>
PPP	Via das Pentoses-Fosfato
Gs	de <i>Geobacillus stearothermophilus</i>
Bp	de <i>Bacillus pumilus</i>
XI	Xilose Isomerase
Sd	de <i>Saccharophagus degradans</i>

LISTA DE SÍMBOLOS

nm	nanômetro
λ	comprimento de onda
c	velocidade da luz no vácuo
f	frequência
h	constante de Planck
F	força
m	massa
a	aceleração
q	carga
E	campo elétrico
Å	ângstrom
µg	micrograma
mL	mililitros
°C	graus Celsius
rpm	revoluções por minuto
mM	milimolar
x g	Unidade de Força Centrífuga Relativa (RCF) g = força gravidade
kDa	kilo Daltons
M	molar
cm	centímetros
pH	potencial hidrogeniônico
µL	microlitro
v/v	razão volume/volume
U	unidade internacional (enzimologia)
mol	molar
s	segundo
K	Kelvin
ps	picossegundo
ns	nanosegundo
fs	femtosegundo
µs	microssegundo
kcal	quilocaloria
Å ²	Ângstrons quadrados

min	minuto
cm ³	centímetros cúbicos
S	Svedberg

SUMÁRIO

1	INTRODUÇÃO	31
1.1	Contextualização	31
1.2	Parede Celular e sua degradação	31
1.3	Enzimas e suas aplicações	33
1.4	Técnicas Físicas aplicadas à Biologia Estrutural	35
1.4.1	Espectrometria de Massas	35
1.4.2	Cristalografia de raios X.....	37
1.4.2.1	Formação de cristais	37
1.4.2.2	Rede e cela unitária.....	38
1.4.2.3	Simetria em Cristalografia	42
1.4.2.4	Planos cristalinos e índices de Miller	43
1.4.2.5	Interação luz-matéria	44
1.4.2.6	Difração em cristais.....	46
1.4.2.7	Refinamento da estrutura obtida	50
1.4.2.8	Alguns parâmetros cristalográficos e seus significados	52
1.4.3	Espalhamento de raios X a baixo ângulo (SAXS)	55
1.4.4	Dinâmica Molecular	57
2	OBJETIVOS.....	63
2.1	Objetivos gerais.....	63
2.2	Objetivos específicos.....	63
3	CAPÍTULO 1 – TcXyn10A.....	65
3.1	Introdução	65
3.2	Material e métodos	65
3.2.1	Clonagem, expressão e purificação.....	65
3.2.2	Padrão de clivagem enzimática.....	67
3.2.2.1	Análises de cromatografia de troca aniônica de alto desempenho (HPAEC)	67
3.2.2.1.1	Produção enzimática de xilooligossacarídeos.....	68
3.2.2.2	Caracterização dos produtos por espectrometria de massas (MALDI-TOF).....	68
3.2.3	Cristalização, coleta de dados e processamento	68
3.2.4	Determinação estrutural por difração de raios-X	69
3.2.5	Análise estrutural	69
3.3	Resultados e discussão	69
3.3.1	Clonagem, expressão e purificação.....	69

3.3.2 Padrão de clivagem enzimática	71
3.3.4 Determinação e análise da estrutura cristalográfica de TcXyn10A.....	74
3.3.4.1 Coleta de dados, refinamento, construção de modelo e comparação com estruturas homólogas.....	74
3.3.4.2 Reconhecimento estrutural para decorações de xilano e perfil de produtos hidrolíticos.....	79
3.4 Conclusões.....	86
4 CAPÍTULO 2 – <i>PxXyn11B</i>	87
4.1 Introdução	87
4.2 Material e métodos	88
4.2.1 Clonagem, expressão e purificação	88
4.2.2 Cristalização, coleta de dados e processamento.....	89
4.2.3 Solução da estrutura cristalográfica	90
4.2.4 Análise estrutural	90
4.2.4.1 Comparações Estruturais.....	90
4.2.5 Simulações de Dinâmica Molecular.....	91
4.3 Resultados e discussão	91
4.3.1 Clonagem, expressão e purificação	91
4.3.2 Determinação e análise da estrutura cristalográfica de <i>PxXyn11B</i>	92
4.3.2.1 Coleta de dados, refinamento, construção de modelo e comparação com estruturas homólogas.....	92
4.3.2.2 Análise do sítio catalítico	98
4.3.2.3 Análise do SBS	100
4.3.4 Simulações de Dinâmica Molecular.....	102
4.4 Conclusões.....	108
5 CAPÍTULO 3 – <i>EcXyl43</i>	109
5.1 Introdução	109
5.2 Material e métodos	110
5.2.1 Clonagem, expressão e purificação	110
5.2.2 Cristalização, coleta de dados e processamento.....	111
5.2.3 Solução da estrutura cristalográfica por difração de raios-X	112
5.2.4 Análise estrutural	112
5.2.4.1 Alinhamento estrutural	112
5.2.4.2 Docking Molecular	113
5.2.5 Simulações de Dinâmica Molecular.....	113
5.3 Resultados e discussão	114

5.3.1 Clonagem, expressão e purificação.....	114
5.3.2 Determinação e análise da estrutura cristalográfica EcXyl43 F507A.....	114
5.3.2.1 Coleta de dados cristalográficos, refinamento, construção de modelo e comparação com estruturas homólogas	114
5.3.4 Análise do sítio de ligação de cálcio	122
5.3.5 Comparações Estruturais	122
5.3.6 Docking molecular de X6 e o sítio ativo de EcXyl43	124
5.3.7 Simulação de Dinâmica Molecular de EcXyl43.....	125
5.3.8 Por que o mutante F507A anula a atividade enzimática EcXyl43?	127
5.3.9 Simulações de dinâmica molecular para mutante F507A.....	129
5.4 Conclusões.....	132
6 CAPÍTULO 4 – SdXI	132
6.1 Introdução	133
6.2. Materiais e métodos	134
6.2.1 Clonagem de XI de <i>S. degradans</i>	134
6.2.2 Expressão heteróloga e purificação de proteína	134
6.2.3 Estudos de espalhamento dinâmico de luz (DLS)	135
6.2.4 Medições de dicroísmo circular (CD)	135
6.2.5 Ensaio Enzimáticos.....	136
6.2.6 Análise de estabilidade térmica	137
6.2.7 Medições de ultracentrifugação analítica (AUC)	137
6.2.8 Modelagem por homologia.....	138
6.2.9 Coleta e processamento de dados SAXS	138
6.2.10 Ensaio de cristalização da SdXI	139
6.2.11 Coleta e processamento de dados	139
6.2.12 Construção e validação de modelo cristalográfico	139
6.2.13 Análise estrutural	140
6.2.14 Estudos de conservação e análise filogenética	140
6.3 Resultados e discussão	141
6.3.1 pH e temperatura ideais	141
6.3.2 Estabilidade térmica.....	142
6.3.3 Ensaio cinéticos	143
6.3.4 Ultracentrifugação analítica (AUC).....	145
6.3.5 Modelagem por homologia de SdXI	146

6.3.6 Medições de DLS e CD	147
6.3.7 Análise de Espalhamento de Raios-X a Baixo Ângulo (SAXS)	149
6.3.8 Alinhamento de sequências de aminoácidos e análise de estrutura secundária	153
6.3.8.1 Análise de estrutura secundária	155
6.3.9 Coleta de dados de cristalografia, refinamento e estrutura terciária	157
6.3.10 Análise de Conservação	164
6.3.11 Análise filogenética	166
6.4 Conclusões	168
7 Conclusões gerais	169
REFERÊNCIAS	171
ANEXOS	189

1 INTRODUÇÃO

1.1 Contextualização

O reino Plantae,¹ constituído por cerca de 350 mil espécies,² compreende uma parte muito importante dos ecossistemas terrestres, apresentando uma diversidade incontável de formas, habitats e fisiologias adaptadas.³ As células vegetais que formam os tecidos das plantas apresentam, além de membrana plasmática, uma estrutura conhecida como parede celular. As paredes celulares ricas em carboidratos desempenham papéis importantes no ciclo de vida da planta,⁴ como determinar o tamanho e a forma das células, fornecer suporte físico para o crescimento e desenvolvimento, atuando como barreiras defensivas e como fonte de moléculas reguladoras biológicas.⁵ As paredes celulares compreendem a maior parte do peso seco da planta e representam a maior parte da biomassa lignocelulósica terrestre. As paredes das células vegetais de diferentes fontes (gramíneas monocotiledôneas, eudicotiledôneas lenhosas, eudicotiledôneas herbáceas e resinosas) têm composições diferentes, mas todas contêm três frações principais: celulose, hemicelulose e lignina.⁶⁻⁷ A planta seca em geral contém 40 a 50% de celulose, 25 a 35% de hemiceluloses, 20 a 25% de lignina e 5 a 10% de outros componentes.⁶

A biomassa vegetal é um recurso importante para a agricultura e produção de energia, manufatura agroindustrial, indústria à base de madeira e biocombustíveis.⁸ Apesar da grande abundância de biomassa vegetal, apenas 2% desse recurso é usado atualmente pelo homem,⁶ possuindo um grande potencial de exploração.

1.2 Parede Celular e sua degradação

A parede celular é composta, basicamente, de lamela média, parede primária e parede secundária (Figura 1). O polissacarídeo mais abundante nas paredes celulares é a celulose, e a segunda maior fração, é a hemicelulose, um polímero constituído por xilano.⁹

Figura 1 - Representação da parede celular vegetal e seus componentes. (A) Estruturação da parede e posicionamento das estruturas. (B) Distribuição e arranjo dos constituintes da parede celular.

Fonte: Adaptada de EVANGELISTA¹⁰

Estruturalmente, o xilano é um polissacarídeo com uma estrutura composta por resíduos de D-xilose covalentemente ligados por ligações β -1,4,¹¹ que podem ser ramificadas com ácido D-glucurônico ligado a (4-O-metil) α -1,2 ou decorado com D-arabinofuranose com ligação α -1,2 ou α -1,3.¹² Além disso, grupos O-acetil também podem ser ligados à cadeia principal do polímero de xilano¹² conferindo recalcitrância a este biopolímero.¹³

As propriedades moleculares do xilano, como peso molecular, identidade, distribuição e tipo de substituintes variam dependendo da espécie da planta, seu estágio de desenvolvimento e o tipo de células. Os xilanos podem ser divididos em quatro grupos de acordo com o tipo de ramificação da estrutura β -D-xilopiranosil: homoxilano, arabinoxilano, glucuronoxilano e arabinoglucuronoxilano.¹⁴ A representação dos diferentes tipos de heteroxilano pode ser observada na Figura 2.

Figura 2 - **Representação estrutural dos diferentes tipos de heteroxilano.** (A) O-Acetil-4-O-metilglucuronoxilano de vegetais de “madeira dura”. (B) Arabino-4-O-metilglucuronoxilano de vegetais de “madeira mole”. (C) Arabinoxilano de cereais.

Fonte: POLLET *et al.*¹⁵

O glucuronoxilano pode ser decorado com glucuronosil e / ou 4-O-metil glucuronosil (MeGlcA) na posição 2 dos resíduos de xilose e é o maior componente da hemicelulose em madeiras duras.^{11,14} Arabinoxilanos, que representam a maior fração de hemicelulose em grãos de cereais e gramíneas, podem ter substituições de L-arabinose na posição 3 e/ou posições 2 e 3 dos resíduos de xilopirranose e são a maior fração de hemicelulose em grãos de cereais e gramíneas.^{11,14}

1.3 Enzimas e suas aplicações

A indústria global de enzimas tem crescido significativamente devido ao seu papel fundamental em aplicações ecologicamente sustentáveis. A expectativa é de crescimento de 8% anual, chegando a US\$ 12,2 bilhões em 2027.¹⁶ Em meio aos diferentes campos de utilização, a produção de biocombustíveis pode ser destacada como uma importante aplicação, com um mercado de US\$ 136,2 bilhões em 2019 e crescimento previsto de 2,2%, chegando a US\$ 153,8 bilhões em 2024.¹⁷ Outro processo conhecido no mercado de transformações enzimáticas, a produção de xilooligossacarídeos (XOS) a partir de xilano utilizando xilanases e seu potencial uso como prebióticos são conhecidos desde os anos 2000.¹⁸ Atualmente o mercado está crescendo através de uma maior preocupação com o bem-estar físico, com adoção de dietas saudáveis e uso de XOS como prebióticos em alimentos com baixo teor de gordura, que podem ser usados seletivamente por um microrganismo probiótico do trato digestivo para conferir benefícios à saúde do hospedeiro.¹⁹ Estima-se que o mercado global de ingredientes probióticos apresente uma taxa de crescimento anual de 11,1%, atingindo um valor de US\$ 6 bilhões em 2022.²⁰

As enzimas podem ser divididas de diversas formas, dependendo do contexto. De acordo com o Repositório de sequências de enzimas ativas em carboidratos (CAZy),²¹ as enzimas podem ser classificadas:

- Quanto a forma de clivagem (figura 3): ação *exo*, clivando a extremidade (geralmente não-redutora) da cadeia; ou ação *endo*, clivando o meio da cadeia.

Figura 3 - Modo de ação das Hidrolases de glicosídeos (GHs) em relação ao local de clivagem na cadeia de xilano.

Fonte: CAZYPEDIA.²²

- Quanto à reação catalisada, o número Enzyme Commission (EC), onde diferentes enzimas podem possuir o mesmo número EC e uma enzima pode possuir mais de um número EC;
- Quanto ao mecanismo da reação. No caso das hidrolases de glicosídeos (GHs), existem dois tipos principais de mecanismo de hidrólise do polímero: O mecanismo de inversão, que ocorre em apenas um passo, e o mecanismo de retenção, que ocorre em dois passos e com a produção de um composto intermediário;
- Quanto à sequência, utilizando algoritmos para designar famílias. As GHs foram classificadas em mais de 100 famílias.²³

Uma vez que a hemicelulose é um grupo muito heterogêneo de biopolímeros, o complexo enzimático que a degrada também é muito complexo.²⁴ Inclui hidrolases de glicosídeos (GHs, EC 3.2.1.-) com diferentes atividades: endo-1,4- β -D-xilanases (EC 3.2.1.8), que hidrolisam ligações internas na cadeia de xilano; 1,4- β -D-xilosidases (EC 3.2.1.37), que atacam os xilo-oligossacarídeos da extremidade não redutora e liberam xilose; endo-1,4- β -D-mananases (EC 3.2.1.78), que clivam ligações internas em manano e 1,4- β -D-manosidases (EC 3.2.1.25), que clivam manooligossacarídeos em manose enquanto as decorações são removidas por α -D-galactosidases (EC 3.2.1.22), α -L-arabinofuranosidases (EC 3.2.1.55), α -glucuronidases (EC 3.2.1.139), acetil xilano esterases (EC 3.1.1.72) e feruloílo e p-cúmarico esterases ácidas (EC 3.1.1.73), entre outras enzimas.²⁴⁻²⁶

Muito importante na hidrólise do xilano, a atividade endo- β -1,4-xilanase (EC.3.2.1.8) é encontrada nas famílias GH 3, 5, 8, 9, 10, 11, 12, 16, 26, 30, 43, 44, 51, 62, 98 e 141,^{21,27} mas a grande maioria das endo-xilanases caracterizadas são das famílias GH 10 e 11.²¹

1.4 Técnicas Físicas aplicadas à Biologia Estrutural

A Biologia Estrutural é uma área que envolve a biologia molecular, bioquímica e biofísica, e tem por objetivo estudar a estrutura molecular das macromoléculas biológicas (em especial as proteínas, os ácidos nucleicos e as membranas lipídicas), como eles adquirem as estruturas que possuem e como as alterações em suas estruturas afetam suas funções.²⁸

Diversas ferramentas podem ser utilizadas no estudo dessas macromoléculas biológicas, tais como Espectrometria de Massas, Cristalografia de Raios X, Ressonância Magnética Nuclear, Espalhamento de Raios X a Baixo Ângulo (SAXS), Difração de Nêutrons, Crio-microscopia Eletrônica e Dicroísmo Circular, entre outras. Além disso, estudos computacionais como simulações de Dinâmica Molecular (DM) podem ser utilizados para entender melhor o comportamento das enzimas ao longo de um intervalo de tempo e sob determinadas condições.

Neste capítulo serão abordadas as técnicas utilizadas nesse trabalho, a saber: Espectrometria de Massas, Cristalografia de Raios X, Espalhamento de Raios X a Baixo Ângulo e Dinâmica Molecular.

1.4.1 Espectrometria de Massas

A Espectrometria de Massas (MS) é uma técnica instrumental bastante utilizada²⁹, com o primeiro instrumento desse tipo, conhecido como espectrógrafo de parábola, construído por J.J. Thomson em 1913, quando descobriu dois isótopos de néon, e posteriormente descobriu outros e confirmando o conceito de isótopos.³⁰ Francis Aston, seu aluno, deu continuidade ao trabalho construindo um instrumento apropriadamente denominado de espectrômetro de massa, contribuindo significativamente para a área.³⁰

A estrutura de um espectrômetro de massas consiste essencialmente em³¹:

- **Sistema de introdução de amostras**, podendo ser por inserção ou infusão direta;
- **Fonte de ionização**, onde ocorre a transferência de energia interna no processo de ionização. Exemplos de fontes: Ionização por elétrons; Ionização química; Bombardeamento atômico rápido (FAB); Ionização química a pressão atmosférica; Ionização por electrospray (ESI); Ionização e dessorção a laser assistida por matriz (MALDI);
- **Analizador de massas**, onde ocorre a separação dos íons de acordo com a relação massa/carga (m/z). Os analisadores podem ser de baixa resolução, como o Quadrupolo, Ion Trap 3D ou linear, ou de alta resolução, como o Tempo de Voo (TOF), Transformada de Fourier de Ressonância Cíclotron de Íons e Orbitrap. Para executar as experiências de espectrometria de massas em tandem, esses analisadores podem ser acoplados em conjunto, onde um primeiro analisador separa os íons, uma célula de colisão é usada para os íons do fragmento, e um segundo analisador os separa do fragmento.
- **Sistema de detecção e processamento**, onde o detector transforma a corrente dos íons em sinal elétrico, amplificando o sinal obtido, que pode ser processado por computador e criando o espectro de massas correspondente.

A técnica é baseada na ionização das moléculas da amostra na fase gasosa, seguida da separação e detecção dos íons resultantes de acordo com a razão massa-carga (m/z). Os resultados são exibidos na forma de um espectro de massa, que é uma representação gráfica da abundância de íons versus m/z . As informações obtidas a partir do espectro de massa podem incluir a massa molecular da molécula da amostra e as massas de seus fragmentos, que podem ser usadas para determinar a estrutura da molécula. Uma vez que as moléculas se fragmentam de maneira única sob determinadas condições, é possível identificar definitivamente a amostra com base no espectro de massa resultante.²⁹

A espectrometria de massa *in tandem* (MS/MS) envolve vários estágios de espectrometria de massa com fragmentação de íons precursores (MS1) para íons de fragmento altamente específico (MS2) e levou a melhorias significativas na taxa e reprodutibilidade da identificação e quantificação de proteínas.³²

1.4.2 Cristalografia de raios X

A cristalografia de raios X é uma técnica utilizada para determinar a estrutura tridimensional detalhada de moléculas, incluindo macromoléculas biológicas complexas, como proteínas e ácidos nucleicos.³³ Trata-se de uma técnica que se baseia na interação da radiação eletromagnética na faixa de 0,01–10 nm (embora normalmente 0,05–0,3 nm) com a matéria na forma cristalina, de modo que as estruturas das moléculas cristalizadas podem ser determinadas com resolução até seus átomos individuais.³³ No caso de proteínas, utiliza-se, portanto, monocristais como amostra.

1.4.2.1 Formação de cristais

A formação de cristais requer duas etapas: nucleação e crescimento. A nucleação é a etapa de iniciação para a cristalização, onde as moléculas de proteína em solução se unem como agregados para formar um núcleo sólido estável.³⁴ A etapa de nucleação é crítica para a formação do cristal, pois é a transição de fase de primeira ordem das amostras que passam de um alto grau de liberdade para obter um estado ordenado (aquoso para sólido).³⁴ Para que a etapa de nucleação seja bem-sucedida, a manipulação dos parâmetros de cristalização, como pH, temperatura e adição de compostos químicos, é essencial.³⁴ Na fase de crescimento, após a formação do núcleo, o cristal cresce através de moléculas que se ligam a esse núcleo estável. O diagrama de formação de um cristal de proteína pode ser visto na Figura 4.

Figura 4 - Ilustração esquemática de um diagrama de fase de cristalização de proteínas. Parâmetros ajustáveis incluem concentração de precipitante ou aditivo, pH e temperatura. Os quatro principais métodos de cristalização, com uma rota diferente para alcançar as zonas de nucleação e metaestável, assumindo que o parâmetro ajustável é concentração de precipitante, são representados: (i) Microbatch, (ii) Difusão de vapor, (iii) Diálise e (iv) Difusão de interface livre (FID). Os círculos pretos preenchidos representam as condições iniciais.

Fonte: Adaptada de CHAYEN; SARIDAKIS³⁵

A solubilidade é definida como a concentração de proteína no soluto que está em equilíbrio com os cristais. A curva de supersolubilidade é definida como a linha que separa as condições sob as quais nucleação (ou separação de fase ou precipitação) ocorre a partir daqueles sob os quais a cristalização solução permanece clara se não for perturbada.³⁵

1.4.2.2 Rede e cela unitária

Em estruturas cristalinas, arranjos de átomos são organizados em um padrão regular e repetido, e o menor elemento repetitivo no cristal é chamado de cela unitária. A cela unitária é, portanto, o menor volume a partir do qual o cristal pode ser inteiramente reconstruído apenas utilizando a translação. Já a unidade assimétrica é o menor volume a partir do qual a cela unitária pode ser construída pela aplicação da simetria cristalográfica.³⁶

A forma e o tamanho da cela unitária podem ser descritos pelo comprimento de três eixos (a , b , c) e os ângulos (α , β , γ) entre os eixos³⁷ (figura 5a). Uma rede é um arranjo periódico de coleção de pontos em um cristal, e essa rede pode ser representada por um conjunto de vetores de translação em duas dimensões, conforme demonstrado na Figura 5b.

Figura 5 - Representação gráfica da cela unitária (A) e da rede cristalina (B), com seus vetores de translação.

Fonte: Adaptada de AMEH³⁷

A rede espacial, um conceito desenvolvido por Auguste Bravais em 1848 para descrever o arranjo da estrutura cristalina, define um número infinito de pontos no espaço de forma que o arranjo de pontos em torno de um ponto seja idêntico em torno de quaisquer outros pontos³⁸. A rede é, portanto, definida por 3 vetores ($\vec{a}$, $\vec{b}$, $\vec{c}$) tal que:

$$r' = r + u\vec{a} + v\vec{b} + w\vec{c} \quad (1.1)$$

Onde u , v e w são inteiros, e o ponto r' é idêntico ao ponto r .

O sistema de eixos x , y , z (Fig. 5a) é usado para designar a disposição dos pontos na rede espacial. As redes de Bravais podem ser divididas em 7 sistemas cristalinos (Tabela 1), de acordo com o comprimento a , b , c e os ângulos α , β e γ em relação aos vetores de translação primitivos (representados na figura 5b).

Tabela 1 - Sistemas cristalinos e suas especificações quanto a comprimento e ângulos

Sistema Cristalino	Comprimentos	Ângulos
triclínico	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma$
monoclínico	$a \neq b \neq c$	$\alpha = \beta = 90^\circ \neq \gamma$
ortorrômbico	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90^\circ$
tetragonal	$a = b \neq c$	$\alpha = \beta = \gamma = 90^\circ$
hexagonal	$a = b \neq c$	$\alpha = \beta = 90^\circ, \gamma = 120^\circ$
trigonal	$a = b = c$	$\alpha = \beta = \gamma < 120^\circ, \neq 90^\circ$
cúbico	$a = b = c$	$\alpha = \beta = \gamma = 90^\circ$

Fonte: Adaptada de WEGENER.³⁹

Todos os arranjos possíveis de pontos no espaço da rede em três dimensões podem ser representados por 14 redes (Fig. 6).

Figura 6 - As 14 redes de Bravais.

Fonte: Adaptada de WEGENER.³⁹

1.4.2.3 Simetria em Cristalografia

De acordo com o dicionário Merriam-Webster, simetria é definida como a “correspondência em tamanho, forma e posição relativa das partes em lados opostos de uma linha divisória ou plano médio ou em torno de um centro ou eixo”.⁴⁰ Uma estrutura de cristal é dita simétrica se certa direção da geometria e propriedade física do cristal são reproduzidas por certa operação de simetria.³⁷ Como operações e elementos de simetria importantes em proteínas, podemos citar⁴¹:

- Identidade: Utiliza-se o símbolo “E”. Esta operação deixa a molécula inalterada;
- Eixo de rotação: Utiliza-se como símbolo “En”, onde o “n” equivale à ordem do eixo de rotação. A operação se dá por $2\pi / n$. Neste caso, um eixo de ordem 2, por exemplo, fornece uma rotação de 180°;
- Eixo parafuso: Utiliza-se como símbolo “En_j”, onde ocorre uma rotação de ordem n no eixo, seguida de uma translação de ordem j (1, ..., n-1) ao longo desse eixo.

Os grupos pontuais são um conjunto de operações de simetria, correspondendo a um dos grupos de pontos em três dimensões, de modo que cada operação deixaria a estrutura de um cristal inalterada. Um grupo monoclinico seria 2 (um eixo de rotação de ordem 2), e um ortorrômbico seria 222 (três eixos de rotação de ordem 2 perpendiculares entre si), na notação Hermann-Mauguin, ou notação internacional.⁴² A adição de 10 elementos básicos de simetria (m, 1, $\bar{1}$, 2, 3, 4, 6, $\bar{3}$, $\bar{4}$, $\bar{6}$) com as 22 combinações possíveis entre eles (2/m, 4/m, 6/m, 222, 32, 422, 622, 432, 23, mmm, $\bar{3}m$, 4/mmm, 6/mmm, $m\bar{3}m$, $m\bar{3}$, mm2, 3m, 4mm, $\bar{4}m2$, 6mm, $\bar{6}m2$ e $\bar{4}3m$) resulta em 32 grupos pontuais de simetria cristalina. No caso de proteínas, não existem os elementos de inversão ($\bar{n}$) nem de espelho (m), deixando apenas 11 grupos pontuais. A adição de eixos parafuso, em combinação desses grupos com as redes de Bravais levam aos 65 grupos espaciais possíveis em proteínas (Tabela 2).

Tabela 2 - Sistemas cristalinos, grupos pontuais e seus respectivos grupos espaciais.

Sistema Cristalino (Bravais)	Grupo Pontual	Grupo Espacial
Triclínico	1	P1
Monoclínico	2	P2, P2 ₁ , C2
Ortorrômico	222	P222, P222 ₁ , P2 ₁ 2 ₁ 2,
		P2 ₁ 2 ₁ 2 ₁ , C222 ₁ , C222, F222, I222, I2 ₁ 2 ₁ 2 ₁
Tetragonal	4	P4, P4 ₁ , P4 ₂ , P4 ₃ , I4, I4 ₁
	422	P422, P42 ₁ 2, P4 ₁ 22, P4 ₁ 2 ₁ 2, P4 ₂ 22, P4 ₂ 2 ₁ 2, P4 ₃ 22, P4 ₃ 2 ₁ 2 I422, I4 ₁ 22
Trigonal	3	P3, P3 ₁ , P3 ₂ , R3
	32	P312, P321, P3 ₁ 12, P3 ₁ 21, P3 ₂ 12, P3 ₂ 21, R32
Hexagonal	6	P6, P6 ₁ , P6 ₅ , P6 ₂ , P6 ₄ , P6 ₃
	622	P622, P6 ₁ 22, P6 ₅ 22, P6 ₂ 22, P6 ₄ 22, P6 ₃ 22
Cúbico	23	P23, F23, I23, P2 ₁ 3, I2 ₁ 3
	432	P432, P4 ₂ 32, F432, F4 ₁ 32 I432, P4 ₃ 32, P4 ₁ 32, I4 ₁ 32

Fonte: Elaborada pelo autor.

1.4.2.4 Planos cristalinos e índices de Miller

Os índices de Miller, $h\ k\ l$, são os inversos das intersecções dos respectivos planos com as direções a , b , c , depois de suprimidos os fatores comuns,⁴³ como pode ser observado na Figura 7. Representa-se para os planos parênteses $(h\ k\ l)$ e para as direções normais a esses planos utilizam-se colchetes $[h\ k\ l]$.

Figura 7 - Representações de planos cristalinos. (A) eixos intersectados por plano hkl, (B) plano (200), (C) plano (110), (D) plano ($1\bar{1}1$) e direção $[1\bar{1}1]$.

Fonte: Adaptada de ALCÁCER.⁴³

Os índices negativos são representados com o sinal menos sobre o índice. Os planos cristalinos são definidos pelos inversos das intersecções com os eixos cristalinos (a/h , b/k , c/l) e designados por ($h\ k\ l$). Na figura estão representados os planos (200) e (110). Nota-se por exemplo que um plano (200) corta os eixos a , b , c para os valores $a/2$, ∞ e ∞ , respectivamente, sendo, portanto, os seus inversos 2, 0, 0.

A equação de um plano pode ser escrita como:

$$x/x_0 + y/y_0 + z/z_0 = 1 \quad (1.2)$$

Sendo x_0 , y_0 , z_0 as intersecções com os respectivos eixos. Portanto a equação do plano ($h\ k\ l$) pode ser escrita como:

$$hx + ky + lz = 1 \quad (1.3)$$

1.4.2.5 Interação luz-matéria

Radiação eletromagnética refere-se às ondas do campo eletromagnético, propagando (irradiando) através do espaço, carregando energia radiante eletromagnética.⁴⁴ Nesse grupo, inclui-se ondas de rádio, microondas, infravermelho, luz (visível), ultravioleta, raios X e raios gama. Em ordem decrescente de energia, as seções do espectro eletromagnético são: raios gama, raios X, radiação ultravioleta, luz visível, radiação infravermelha e ondas de rádio, sendo microondas uma subseção desta última.⁴⁵

O vetor campo elétrico $\vec{E}$ é perpendicular ao vetor campo magnético $\vec{B}$ e ambos oscilam perpendicularmente ao vetor de propagação, portanto, com comprimento de onda λ e fase ϕ . A partícula elementar mediadora da força eletromagnética, e quantum da radiação

eletromagnética é chamada fóton, que não possui massa e se move sempre na velocidade da luz no vácuo ($c = 299.792.458 \text{ m/s}$). O fóton possui uma energia E proporcional a sua frequência, f :

$$E = hf = hc/\lambda \quad (1.4)$$

onde h é a constante de Planck, λ é o comprimento de onda e c é a velocidade da luz. Essa equação é também conhecida como equação de Planck-Einstein e demonstra que a energia dos fótons é diretamente proporcional à frequência da onda eletromagnética, fazendo com que os fótons dos raios gama, por exemplo, possuam a maior energia no espectro eletromagnético.

A força exercida no átomo na presença do campo elétrico do raio X é dada pela 2ª lei de Newton ($\vec{F} = m \vec{a}$), no caso do átomo, $\vec{F} = q \vec{E}$.

Igualando as equações, temos que $q \vec{E} = m \vec{a}$, portanto

$$\vec{a} = q \vec{E} / m \quad (1.5)$$

Levando em consideração que a massa do próton e do nêutron são cerca de 2000 vezes maiores do que a do elétron, e além disso a carga do nêutron é 0, estas contribuem de forma negligenciável, tornando significativa apenas a interação entre o raio X e a nuvem eletrônica dos átomos, dependente do número atômico (Z) destes átomos. Esses elétrons ressonam na mesma frequência dos raios X incidentes e emitem ondas secundárias com intensidade dada pela equação de Thomson:⁴⁶

$$I_s(\theta) = \beta I_p \frac{1}{r^2} (1 + \cos^2 \theta/2) \quad (1.6)$$

Onde I_s é a intensidade das ondas secundárias, 2θ é o ângulo de espalhamento, β é o raio clássico do elétron, I_p é a intensidade das ondas primárias, r é a distância do objeto ao detector e $(1 + \cos^2 \theta/2)$ é o fator de polarização. O feixe incidente pode ser denotado pelo vetor unitário $\vec{S}_0$ e o feixe espalhado por $\vec{S}_1$. A diferença de caminho entre as ondas incidente e espalhada será $\vec{S}_1 - \vec{S}_0 = \vec{S}$.⁴⁶

Na interação de um raio X com um átomo, seu campo elétrico produzirá uma força que fará com que cada átomo tenha um fator de espalhamento f , de acordo com seu número atômico. Esse fator possui a seguinte equação:

$$f = \sum_{i=1}^4 a_i \exp -b_i (\sin \theta/\lambda)^2 + c \quad (1.7)$$

Essa dependência é dada pelos coeficientes de Cromer-Mann (a_i , b_i e c), que são conhecidos e tabelados para cada átomo que ocorre naturalmente na tabela periódica⁴⁷. O fator de espalhamento também depende do comprimento de onda do raio X e cai exponencialmente com o aumento do ângulo de detecção. No ângulo de incidência ($\theta = 0^\circ$), f é igual ao número de elétrons no átomo. A unidade de f é dada em elétrons.

1.4.2.6 Difração em cristais

Uma vez que cristais são formados por unidades (celas unitárias) repetidas em três direções, e conseqüentemente um arranjo periódico de moléculas, as ondas espalhadas, ou difratadas, sofrerão fenômeno de interferência construtiva (aumento na amplitude das ondas, sem mudança de fases), ou interferência destrutiva (diminuição na amplitude ou cancelamento das ondas). Na cristalografia de raios X, a interferência construtiva é bastante importante, tirando proveito do efeito multiplicador pelo cristal das ondas secundárias dos entes espalhadores (nuvens eletrônicas dos átomos). Matematicamente, a difração em cristais pode ser descrita pela Lei de Bragg:⁴⁸

$$n \lambda = 2 d_{hkl} \sin \theta \quad (1.8)$$

onde n é um número natural, λ é o comprimento de onda, d é a distância entre os planos atômicos e θ é o ângulo de espalhamento. Geometricamente, a Lei de Bragg pode ser representada como na figura 8.

Figura 8 - Representação geométrica da Lei de Bragg. Duas ondas em fase, onde d é a distância entre planos atômicos.

Fonte: Adaptada de BRITANNICA⁴⁹

Após a difração por meio da interação do raio X com o cristal, o detector mede uma intensidade I_{hkl} , que por sua vez é proporcional ao quadrado da sua amplitude ($I \propto A^2$). Como pode-se perceber a partir da Figura 8 e da equação (1.8), a difração em cristais vai depender do comprimento de onda utilizado (com uma energia associada), do ângulo (incidência e reflexão) e da distância interplanar entre os planos cristalinos $h k l$.

Em termos práticos, convencionou-se chamar o espaço da rede de Bravais do cristal como “espaço real”, ou direto, e o espaço onde ocorre a medição da difração como “espaço recíproco”, representando a transformada de Fourier do espaço real.

O índice de Miller pode também ser definido como um ponto na rede recíproca. A rede cristalina determina a posição dos picos no padrão de difração, descritos em termos de índice de Miller (h, k, l). Cada ponto na rede recíproca possui uma intensidade I_{hkl} , proporcional ao módulo quadrático do Fator de Estrutura ($I_{hkl} \propto |F_{hkl}|^2$), que possui informações sobre a amplitude da onda associada. O fator de estrutura pode simplificarmente ser definido por:

$$\vec{F}_{hkl} = \sum_{j=1}^N f_j \exp [2\pi i (hx_j + ky_j + lz_j)] \quad (1.9)$$

onde a soma é sobre todos os átomos na célula unitária, x_j, y_j, z_j são as coordenadas posicionais do j -ésimo átomo, f_j é o fator de espalhamento (equação 1.7) do j -ésimo átomo.⁵⁰

Da mesma forma que o fator de estrutura é uma transformada de Fourier (FT) do espaço real para o recíproco, através da transformada inversa de Fourier (FT⁻¹) é possível reconstituir a função temporal original. Isso é importante na determinação de estruturas em cristalografia, pois permite reconstruir o mapa de densidade eletrônica $\rho(\vec{r})$. A equação de densidade eletrônica pode ser expressa por:

$$\rho_{xyz} = 1/V \sum_{h=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} |\vec{F}_{hkl}| \exp i\phi_{hkl} \exp -2\pi i (hx_j + ky_j + lz_j) \quad (1.10)$$

Como pode ser observado na equação (1.11) a densidade eletrônica é uma função dependente, entre outros fatores, do módulo do fator de estrutura, que pode ser calculado, e da fase associada ϕ . Entretanto, em um experimento clássico de difração por raio X, é medida a quantidade de energia transferida do campo elétrico para a área unitária do detector (em W.m⁻²), em média, ao longo do tempo. Em um segundo, 1.018 oscilações de ondas atingiram o detector. Neste caso, a fase é perdida. Isso é conhecido como o “problema das fases”.

A fase associada ao fator de estrutura pode ser definida como o ângulo (Φ) formado entre o vetor do fator de estrutura $\vec{F}_{hkl}$ e o eixo real, como podemos observar na esquematização de um diagrama de Argand (figura 9) onde é possível representar geometricamente números complexos que estão na forma algébrica representados $z = a + bi$, em que a é a parte real e b a parte imaginária.

Figura 9 - Diagrama de Argand, representando o fator de estrutura F e a fase associada, evidenciando o caráter centrossimétrico da difração.

Fonte: Adaptada de CRISTALOGRAFIA⁵¹

Para determinação das fases, e consequente resolução da estrutura tridimensional da molécula a partir de ensaios de cristalografia de raios x, alguns métodos podem ser utilizados, como Substituição Molecular e Faseamento Experimental.

Substituição Molecular (do inglês Molecular Replacement – MR): Neste procedimento, utiliza-se o padrão de difração da molécula (proteína) de estrutura desconhecida e uma proteína homóloga de estrutura já resolvida (modelo), e com identidade sequencial de pelo menos 25%. Neste caso, as coordenadas atômicas da estrutura resolvida são posicionadas no interior da cela unitária do cristal de maneira que o padrão de difração teórico do modelo se aproxima do padrão experimental, posicionando o modelo corretamente orientado na cela unitária, através de função de rotação e translação, e utilizando algoritmos

otimizados, como máxima verossimilhança (maximum likelihood, ou ML) e estatística multivariada, em vez dos métodos tradicionais como mínimos quadrados e Patterson.⁵²

Faseamento Experimental: Nesse grupo se encaixam casos onde não é possível ter um bom modelo para substituição molecular. São utilizadas técnicas como:

- Substituição Isomorfa, onde ocorre a adição de um átomo eletronicamente mais denso (Hg, Pt, Pb, Ag, Gd). A interação ordenada e consistente desses metais com regiões específicas das proteínas dentro do cristal altera de maneira significativa as intensidades das reflexões no padrão de difração. A partir da comparação das intensidades derivatizadas com as de cristais não-derivatizados, é possível inferir com precisão a posição dos átomos "pesados" e estimar um conjunto inicial de fases para todas as reflexões.

- Difração Anômala, que se baseia na variação do comprimento de onda do raio X incidente. Se a energia equivalente corresponde a uma transição entre diferentes camadas de elétrons de um tipo de átomo específico, haverá uma modificação das fases da onda difratada devido à absorção de fótons incidentes. Os átomos relativamente leves encontrados em proteínas, como carbono, nitrogênio e oxigênio não contribuem para a dispersão anômala em comprimentos de onda de raios X normais utilizados para a cristalografia de raios X (entre 0,7 e 2 Å). Assim, para se observar esse efeito, o átomo pesado deve ser nativo da proteína (enxofre, por exemplo) ou derivatizado, como no caso da substituição isomorfa. Além disso, o comprimento de onda do raio X emitido pela fonte deve ser ajustado às bordas de absorção do átomo pesado, geralmente correspondentes a energias das camadas eletrônicas mais internas (K ou L)⁵³. Pode ser do tipo SAD (difração anômala de comprimento de onda único, do inglês "Single wavelength Anomalous Diffraction"), quando utiliza apenas um comprimento de onda, ou MAD (difração anômala de múltiplos comprimentos de ondas, do inglês "Multiple wavelength Anomalous Diffraction").

1.4.2.7 Refinamento da estrutura obtida

Após a solução do problema das fases e reconstrução da densidade eletrônica da proteína, é necessário realizar etapas de refinamento da estrutura. Na determinação da estrutura, refinamento é o processo de melhorar os parâmetros de uma estrutura aproximada (trial) até que o melhor ajuste seja alcançado entre um padrão de difração observado e aquele

calculado pela transformação de Fourier a partir da estrutura experimental parametrizada numericamente.⁵⁴ O refinamento pode ser baseado em I , $|F|$ ou F^2 .

O refinamento baseado em I (as intensidades medidas) tem o mérito de usar as medidas brutas diretamente, embora requeira a incorporação no refinamento dos fatores de correção (fator de escala, polarização de Lorentz e absorção) que são aplicados durante a redução dos dados padrão. Existem, no entanto, problemas de alta correlação estatística ao refinar os parâmetros de absorção juntamente com os parâmetros de deslocamento anisotrópico.⁵⁴

O refinamento baseado em $|F|$ envolve problemas matemáticos com reflexos muito fracos ou reflexos com intensidades medidas negativas. Também existem dificuldades em estimar as incertezas padrão $\sigma(F)$ a partir dos valores $\sigma(F^2)$ para intensidades medidas fracas ou nulas.⁵⁴

O refinamento com base em F^2 evita essas dificuldades e também reduz a probabilidade das iterações de refinamento se estabelecerem em um mínimo local. Ele também simplifica o tratamento de estruturas geminadas e não centrosimétricas. Por essas razões, provavelmente é mais atualmente a técnica mais usada, embora dependa muito da atribuição de pesos razoáveis para reflexões individuais.⁵⁴

Entre os tipos de refinamentos, podemos citar:

Refinamento de mínimos quadrados: A abordagem mais comum na determinação de estruturas inorgânicas ou de moléculas pequenas é minimizar uma função $\sum \omega (Y_o - Y_c)^2$ onde Y_o representa a força de um ponto de difração observado ou reflexão de um plano de rede do cristal, às vezes conhecido como coeficiente do fator de estrutura, e pode ser I (a intensidade da reflexão medida), $|F|$ (a magnitude do fator de estrutura) ou F^2 (o quadrado do fator de estrutura). Y_c é o valor calculado a partir do modelo estrutural para a mesma reflexão, e ω é um peso atribuído refletindo a importância que essa reflexão tem para a soma. Os pesos geralmente representam uma estimativa da precisão da quantidade medida. A soma é feita sobre todas as reflexões medidas.⁵⁴

Máxima verossimilhança (Maximum Likelihood): O princípio da máxima verossimilhança formaliza a ideia de que a qualidade de um modelo é julgada por sua consistência com as observações, de maneira eficiente.⁵⁴⁻⁵⁵ Se um modelo for consistente com uma observação, então - se o modelo estiver correto - haveria uma grande probabilidade de

fazer uma observação com aquele valor. Para um conjunto de observações relevantes, a probabilidade de gerar tal conjunto é uma excelente medida da qualidade do modelo. Para observações independentes, a probabilidade conjunta de fazer o conjunto de observações é o produto das probabilidades de fazer cada observação independente.⁵⁴

Seja $P(|F_o|;|F_c|)$ a probabilidade de se obter um fator de estrutura observado F_o dado um valor calculado F_c . A probabilidade conjunta é a função de verossimilhança (likelihood) L é dada pela equação $L = \prod_{hkl} P(|F_o|;|F_c|)$. Uma vez que é mais conveniente trabalhar com somas do que produtos, normalmente trabalha-se com o logaritmo negativo da função de verossimilhança:⁵⁴

$$L = -\sum_{hkl} \log P(|F_o|;|F_c|) \quad (1.11)$$

O procedimento matemático para determinar a probabilidade máxima torna-se então o de minimizar L .⁵⁴

Alguns parâmetros, conhecidos como parâmetros refináveis, podem ser variados para melhorar o ajuste. Dentre eles podemos citar as coordenadas atômicas, os parâmetros de deslocamento atômico, um fator de escala para trazer as amplitudes ou intensidades observadas e calculadas para a mesma escala, parâmetros de extinção, fatores de ocupância e até mesmo o grupo de espaço atribuído.⁵⁴

Restrição é uma relação matemática exata que reduz o número de parâmetros livres em um modelo. Por exemplo, a posição de um átomo em uma posição geral é especificada por três coordenadas, todas as quais podem ser variadas independentemente.⁵⁴ No entanto, um átomo posicionado em uma posição de simetria especial tem uma ou mais coordenadas posicionais determinadas pela simetria (por exemplo, um átomo em um centro de inversão na célula unitária tem todas as três coordenadas fixas).⁵⁴ Restrições são regras matemáticas rígidas que devem ser respeitadas durante o refinamento; eles reduzem o número de parâmetros refináveis.⁵⁴ Um refinamento restrito é aquele que inclui outras restrições além daquelas que surgem da simetria do grupo espacial (uma vez que estas estão necessariamente sempre presentes).⁵⁴

1.4.2.8 Alguns parâmetros cristalográficos e seus significados

Na análise de estruturas cristalográficas, alguns parâmetros se mostram bastante importantes. A seguir, uma relação de alguns deles:

- Resolução: A resolução é medida em Å e pode ser definida como o espaçamento mínimo (d) de planos (hkl) de rede cristalina que ainda podem produzir difração detectável e mensurável de raios-X⁵⁶. Este termo define o nível de detalhe, ou a distância mínima entre as características estruturais que podem ser distinguidas nos mapas de densidade eletrônica. O valor de 1,5 Å corresponde a ligações covalentes C – C típicas em macromoléculas, por exemplo. Quanto maior a resolução, ou seja, quanto menor o espaçamento d , melhor, pois há mais reflexões independentes disponíveis para definir a estrutura.⁵⁶
- R-factor: Geralmente expresso como porcentagem, mas frequentemente como fração decimal, mede a discrepância relativa global entre as amplitudes do fator de estrutura obtidas experimentalmente, F_{obs} , e as amplitudes do fator de estrutura calculadas, F_{calc} , obtidas a partir do modelo. O R-factor, definido como $\sum |F_{obs} - F_{calc}| / \sum F_{obs}$, combina o erro inerente aos dados experimentais e o desvio do modelo da realidade.⁵⁶
- R_{free} : É uma função residual calculada durante o refinamento da estrutura da mesma maneira que o R-factor convencional, mas aplicada a um pequeno subconjunto de reflexões selecionadas aleatoriamente que não são usadas no refinamento do modelo estrutural.⁵⁷ O objetivo é monitorar o progresso do refinamento e verificar se o R-factor não está sendo reduzido artificialmente pela introdução de muitos parâmetros. Desse modo, se o modelo matemático da estrutura se tornar excessivamente complexo, ou seja, incluir parâmetros para os quais não há justificativa nos dados experimentais, o R_{free} não melhorará (mesmo que o R-factor possa diminuir), indicando superinterpretação dos dados. Isso ocorre porque os parâmetros supérfluos tendem a modelar os erros aleatórios do conjunto de dados de trabalho, que não estão correlacionados com os erros do conjunto R_{free} ⁵⁶. Dessa forma, R_{free} testa quão bem o modelo prediz observações experimentais que não são usadas para ajustar o modelo. Uma porcentagem fixa do número total de reflexões geralmente é atribuída ao grupo livre. É calculado da mesma forma que o R-factor convencional $R = \sum |F_{obs} - F_{calc}| / \sum |F_{obs}|$ mas usa um pequeno subconjunto de reflexos selecionados aleatoriamente que são colocados à parte desde o início e não usados no refinamento do modelo estrutural.⁵⁷

Um R_{free} ponderado também pode ser calculado sobre o conjunto de reflexões não utilizadas no refinamento:⁵⁶

$$\omega R = (\sum |\omega| Y_o - Y_c|^2 / \sum |\omega Y_o|^2)^{1/2} \quad (1.12)$$

onde Y representa F, F^2 ou I.

Após cada ciclo de refinamento, o R_{free} e o R-factor para o conjunto de trabalho de reflexões são calculados. À medida que o refinamento converge, os fatores R funcional e livre se aproximam de valores estáveis.

- R_{merge} : A precisão das intensidades médias pode ser avaliada a partir da propagação das medições individuais de reflexões equivalentes pelo residual R_{merge} . Sua fórmula simplificada é $R_{\text{merge}} = \sum_h \sum_i (|<I_h> - I_{h,i}| / \sum_h \sum_i I_{h,i})$ onde h enumera as reflexões únicas e i seus contribuintes de simetria equivalente. Na camada de resolução mais alta, o R_{merge} pode atingir 30–40% para cristais de baixa simetria e até 60% para cristais de alta simetria, visto que, no último caso, a redundância geralmente é maior.⁵⁶
- R_{meas} : É uma medida como o R_{merge} , mas leva em consideração a redundância no cálculo⁵⁸. Enquanto R_{merge} possui a fórmula $R_{\text{merge}} = \sum_{hkl} \sum_j |I_{hkl,j} - <I_{hkl}>| / \sum_{hkl} \sum_j I_{hkl,j}$, o R_{meas} possui a fórmula $R_{\text{meas}} = \sum_{hkl} (n/n-1)^{1/2} \sum_{j=1}^n |I_{hkl,j} - <I_{hkl}>| / \sum_{hkl} \sum_j I_{hkl,j}$ onde $<I_{hkl}>$ é a média de observações relacionadas à simetria (ou pares de Friedel) de uma reflexão única.^{56, 58}
- B-factor: O termo B-factor, às vezes chamado de fator de Debye-Waller, fator de temperatura ou parâmetro de deslocamento atômico, é usado na cristalografia de proteínas para descrever a atenuação de raios-X ou espalhamento de nêutrons causado pelo movimento térmico.⁵⁹ Os B-factors são geralmente expressos em $\text{\AA}^2$ e variam de ~ 2 a ~ 100 . O modelo do B-factor usado é geralmente isotrópico, ou seja, descreve apenas a amplitude do deslocamento, mas modelos mais elaborados descrevem o deslocamento anisotrópico individual de cada átomo.⁵⁶
- Mosaicidade: Definida como uma medida da propagação das orientações do plano do cristal. No modelo de um cristal real, este é descrito como um conglomerado de blocos cristalinos minúsculos, inclinados entre si por frações de um minuto de arco.⁶⁰ A mosaicidade também é definida como a largura total na metade do máximo (FWHM) dos picos de difração e é uma

propriedade intrínseca de cada cristal. Um alto valor de mosaicidade, implica em pontos de difração mais largos, complicando a indexação e determinação de células unitárias devido à sobreposição de pontos.

1.4.3 Espalhamento de raios X a baixo ângulo (SAXS)

O espalhamento de raios-X a baixo ângulo é uma técnica que permite a análise de macromoléculas em solução e que, portanto, não depende da formação de cristais para análise, como no caso da cristalografia de raios-X. Trata-se de uma técnica bastante dinâmica, uma vez que permite a utilização de diferentes valores de pH⁶¹ e temperatura,⁶² parâmetros importantes em se tratando de estrutura e função de enzimas.⁶³

A técnica de SAXS se baseia na interação da radiação eletromagnética com a matéria a ser estudada, e consequente espalhamento de raios-X através dessa amostra. A função de espalhamento $F(\vec{q})$, dependente do número de espalhadores N e do vetor de espalhamento $\vec{q}$, é representada por:⁶⁴

$$F(\vec{q}) = \sum_{i=1}^N f_j(\vec{q}) \exp i\vec{q} \cdot \vec{r} \quad (1.13)$$

onde $\vec{r}$ é a distância do centro de espalhamento ao espalhador, $\vec{q}$ é o vetor de espalhamento no espaço recíproco, onde $|q| = 4\pi \sin\theta / \lambda$ com 2θ sendo o ângulo de espalhamento e λ o comprimento de onda e $f_j(\vec{q})$ o fator de forma atômico, semelhante ao definido na equação (1.7).

De maneira semelhante ao que ocorre em cristalografia de raios-X, no espalhamento em SAXS não é possível medir a $F(\vec{q})$, e sim detecta-se a intensidade I , que não carrega as informações das fases das ondas. Além disso, é conveniente assumir que o espalhamento seja isotrópico, o meio seja homogêneo, o contraste $\Delta\rho = \rho_{amostra} - \rho_{solvente}$ não seja variável e que as partículas tenham movimento independente entre si, a média sobre todas as orientações é dada por $\langle \exp -i\vec{q} \cdot \vec{r}_{ij} \rangle = \sin qr_{ij}/qr_{ij}$.

Através de I pode-se obter a $P(r)$.⁶⁵⁻⁶⁷ A $P(r)$ é um histograma das distâncias na partícula representando uma função de distribuição de pares de distância, ou seja, representa a distribuição de distâncias entre pares de elétrons limitados pelo volume da partícula.⁶⁸ A $P(r)$ está, portanto, limitada entre zero e a dimensão máxima $D_{\max}$ da partícula ($0 \leq r \leq D_{\max}$), e contém informações sobre o tamanho e forma da molécula no espaço real.⁶⁴

A $P(r)$ pode fornecer parâmetros importantes, tais como o raio de giração e a intensidade a ângulo zero $I(q \rightarrow 0)$. O raio de giração (R_g) é definido como a média da raiz quadrada média da distância de todos os elementos de dispersão a partir do centro de massa da molécula⁶⁹ e descreve a distribuição da massa da partícula em torno do centro de massa.⁶⁴

A $I(q \rightarrow 0)$ pode trazer informações muito importantes sobre as partículas em solução. Para sistemas monodispersos e com q tendendo a zero, a Intensidade $I(\vec{q})$ pode ser aproximada por uma expansão de Taylor da seguinte forma:

$$I(\vec{q}) = 4 \pi \int_0^{D_{max}} P(\vec{r}) \frac{1}{qr} [qr - \frac{1}{3!}(qr)^3 - \frac{1}{5!}(qr)^5 + \dots] dr \quad (1.14)$$

Nesta expansão apenas os termos $\leq q^3$ vão contribuir para $I(\vec{q})$.⁶⁴ Dessa forma, através de aproximações obtém-se a equação conhecida como Lei de Guinier:

$$I(\vec{q}) = I(q \rightarrow 0) \exp (-q^2 R_g^2 / 3) \quad (1.15)$$

Com essa equação, é possível obter o raio de giração, utilizando $\ln I(\vec{q}) = \ln I(q \rightarrow 0) - q^2 R_g^2 / 3$. Os valores de R_g e $I(q \rightarrow 0)$ podem ser obtidos com um ajuste linear dessa equação, plotando $\ln I(q)$ versus q^2 em um gráfico.

Dentre outros parâmetros calculados a partir da curva de intensidades do SAXS, podemos citar:

Invariante de Porod e volume: Na situação em que q tende a zero, Porod propôs que um termo na equação de $P(r)$ é invariante sob quaisquer modificações do sistema ou da sua função de espalhamento.^{64, 68} Com o invariante $Q = \int_0^\infty I(q) q^2 dq$ chega-se a equação:

$$V = 2 \pi^2 I(q \rightarrow 0) / Q \quad (1.16)$$

onde V é o volume da partícula.

Massa molecular: Através de programas como o SAXSMoW2, é possível calcular a massa molecular de uma proteína através de sua curva de espalhamento de SAXS.⁷⁰

Informações sobre flexibilidade e alterações conformacionais: Através do gráfico de Kratky $I(q)q^2$ versus q ,⁶⁸ informações sobre a flexibilidade podem ser obtidas de acordo com o perfil da curva do gráfico. Proteínas globulares tendem a formar um pico proeminente a baixo ângulo, indicando pouca ou nenhuma flexibilidade. Já proteínas não-estruturadas apresentam um aumento de forma contínua de $I(q)q^2$ a medida em que q aumenta, enquanto proteínas com múltiplos domínios podem apresentar uma combinação desses perfis.^{64, 71-72}

Além disso, através do pacote de programa ATSAS⁶⁷ é possível construir um modelo de envelope molecular 3D da proteína estudada, utilizando programas como o GASBOR⁷³ e sobrepor ao modelo cristalográfico através do SUPCOMB.⁷⁴

1.4.4 Dinâmica Molecular

A Dinâmica Molecular (DM) pode ser definida como uma técnica de simulação computacional que permite a previsão da evolução temporal de um determinado sistema de interação, envolvendo a geração de trajetórias atômicas de um sistema usando integração numérica da equação de movimento de Newton para um potencial interatômico específico definido por uma condição inicial e condição de contorno (limite).⁷⁵ Essa técnica simula o movimento natural do sistema molecular⁷⁶ e está se tornando uma ferramenta inestimável para estudar a relação entre estrutura e dinâmica biomolecular, crucial para a compreensão de processos celulares importantes.⁷⁷

A dinâmica molecular (DM) é outra abordagem para a investigação da localização do átomo no espaço. Nessa abordagem, um modelo de ponto único é substituído por um modelo dinâmico no qual o sistema nuclear é forçado a se mover. A simulação do movimento é realizada pela solução numérica das equações dinâmicas newtonianas clássicas. O conjunto de possíveis localizações de átomos dá, por exemplo, o perfil de conjunto conformacional para uma dada molécula. A DM também pode fornecer informações sobre as propriedades termodinâmicas e dinâmicas das moléculas e também pode ser usada para refinamento de estruturas de raios-X.⁷⁸ Por se basear, essencialmente, na solução da 2ª Lei de Newton (equação 1.17), a DM faz parte dos denominados “Métodos Clássicos”, em oposição aos métodos baseados na Física Quântica (métodos de mecânica quântica).⁷⁹

$$Fx_i = m_i d^2x_i / dt^2 = m_i \Delta v_i / \Delta t = m_i a_i \quad (1.17)$$

onde Fx_i é a força aplicada ao átomo i na posição x , t é o tempo, v a velocidade, m_i a massa do átomo i e a_i a aceleração do átomo i .

A temperatura fornece energia para que os átomos sofram uma aceleração, mudando suas posições espaciais. Entretanto, como os átomos não estão isolados, estão sujeitos a forças interatômicas e inter-moleculares, e o cálculo destas forças é realizado por uma outra função matemática, denominada campo de força.⁷⁹ No contexto de simulações de dinâmica molecular de proteínas, o termo “campo de força” se refere à combinação de uma fórmula matemática e

parâmetros associados que são usados para descrever a energia da proteína em função de suas coordenadas atômicas.⁸⁰ Cada campo de força estabelece um conjunto de equações matemáticas dedicadas a reproduzir aspectos do comportamento molecular, como o estiramento de ligações químicas, a deformação de um ângulo de ligação ou a torção de um diedro, como podemos observar em um espectro de infravermelho.⁷⁹ As equações podem ser divididas em termos intramoleculares, ou internos (equação 1.18) e intermoleculares, ou externos (equação 1.19).⁸¹

$$E_{ligado} = \sum_{\text{ligações}} K_b (b - b_0)^2 + \sum_{\text{ângulos}} K_\theta (\theta - \theta_0)^2 + \sum_{\text{diedros impróprios}} \quad (1.18)$$

$$K_\phi (\phi - \phi_0)^2 + \sum_{\text{diedros}} \sum_{n=1}^6 K_{\phi,n} (1 + \cos(n \phi - \delta_n))$$

$$E_{\text{não-ligado}} = \sum_{\text{pares } i,j \text{ não-ligados}} q_i q_j / 4 \pi \epsilon_0 r_{ij} + \sum_{\text{pares } i,j \text{ não-ligados}} \epsilon_{ij} [(R_{\min,ij} / r_{ij})^{12} - 2(R_{\min,ij} / r_{ij})^6] \quad (1.19)$$

onde K_b = constante de força de ligação; b_0 = comprimento de ligação de referência; K_θ = constante de força angular; θ_0 = ângulo de valência de referência; K_ϕ = constante de força diédrica imprópria; ϕ_0 = valor de referência do ângulo diédrico impróprio (geralmente 0); $K_{\phi,n}$ = amplitude diédrica; n = multiplicidade diédrica; δ_n = fase diédrica; q_i, q_j = cargas parciais; ϵ_{ij} = profundidade do poço de potencial Lennard-Jones; $R_{\min,ij}$ = raio de Lennard-Jones.⁸¹

Consequentemente, a energia potencial do sistema pode ser representada por:

$$U = U_{\text{ligação}} + U_{\text{ângulo}} + U_{\text{diedros}} + U_{\text{não-ligado}} \quad (1.20)$$

Na equação 1.20, a $U_{\text{ligação}}$ representa as oscilações em torno do comprimento de ligação no equilíbrio, a $U_{\text{ângulo}}$ representa as oscilações de 3 átomos em torno do ângulo de ligação no equilíbrio, a U_{diedros} representa a rotação torsional de 4 átomos em torno de uma ligação central e a $U_{\text{não-ligado}}$ representa os termos de energia não ligada (eletrostática e Lennard-Jones).

As energias potenciais e as respectivas ilustrações dos termos podem ser observadas na figura 10.

Figura 10 - Representação dos tipos de energia potencial (V), com suas respectivas equações

Fonte: VERLI.⁷⁹

Estas equações (Figura 10), por sua vez, são calibradas (ou seja, parametrizadas) para reproduzir o comportamento dos compostos de interesse.

Em uma simulação de dinâmica molecular, existem diferentes tipos de descrição dos átomos, dependendo do nível de simplificação: Todos os átomos (all-atom), onde o campo de força descreve todos os átomos do sistema, com um custo computacional muito maior⁷⁹; átomo unido (united-atom), um primeiro nível de simplificação, onde os átomos de hidrogênio apolares (ligados a carbono) são unidos aos átomos de carbono, dando origem a pseudoátomos, que representam as propriedades de CH, CH₂ e CH₃, com exceção do CH de anéis aromáticos, com descrição própria nos campos de força,⁷⁹ coarse-grained (CG) onde vários átomos são agrupados em uma única partícula. No caso de peptídeos, o esqueleto peptídico seria uma partícula única, e a cadeia lateral de zero (glicina) a quatro (triptofano) partículas, por exemplo.⁷⁹

O sistema pode ter diferentes montagens (ensembles) dependendo do objeto e condições de estudo. Entre essas montagens, podemos citar: a Microcanônica (NVE), onde o sistema é isolado de mudanças em moles de substância (N), volume (V) e energia (E), correspondendo a um processo adiabático sem troca de calor, com a energia total sendo

conservada;⁸² a Canônica (NVT), onde a quantidade de substância (N), volume (V) e temperatura (T) são conservados e a energia dos processos endotérmicos e exotérmicos é trocada por um termostato; a Isotérmica-Isobárica (NPT), onde a quantidade de substância (N), pressão (P) e temperatura (T) são conservados, e além de um termostato, um barostato é necessário. Uma vez que a maioria das observações experimentais são realizadas em temperatura e pressão constantes, a montagem NPT tem sido amplamente utilizada em simulações moleculares.⁸³

A simulação de dinâmica molecular ocorre geralmente em etapas, que podem ser divididas da seguinte forma:

Preparação: Na preparação, a tipagem atômica, ou seja, o nome que é dado para cada átomo, pode ser diferente. Diferentes tipos de átomos são atribuídos a diferentes elementos, mas também a diferentes estados de hibridização do mesmo elemento e, geralmente, até mesmo a átomos no mesmo estado de hibridização, mas em um ambiente químico diferente. Por exemplo, no caso de campos de força de proteína, $C\alpha$ normalmente obteria um tipo de átomo diferente dos outros carbonos sp^3 , um carbono de carbonila obteria um tipo de átomo diferente dependendo se ele faz parte de um grupo amida ou carboxilato, e assim por diante.⁸¹ Apesar disso influenciar a precisão da simulação do sistema, as particularidades nessa etapa fazem com que não haja muita interconversão entre os diferentes programas de DM.

Além disso, durante a preparação do sistema, é feito o assinalamento de cargas. Nessa etapa, cargas são assinaladas a diferentes resíduos, de acordo com o pH de interesse do estudo. Programas como o CHARMM-GUI⁸⁴ fazem a protonação do sistema de acordo com o pH desejado.

No fim da preparação, serão gerados, além do arquivo .pdb inicial com as coordenadas atômicas da proteína alvo do estudo, um arquivo .psf que contém seis seções principais de interesse: átomos, ligações, ângulos, diédricos, impróprios (termos de força diédrica usados para manter a planaridade) e termos cruzados.⁸⁵ Além disso, são necessários o arquivo de topologia (necessário para gerar o .psf), que define os tipos de átomos usados no campo de força; os nomes dos átomos, tipos, ligações e cargas parciais de cada tipo de resíduo,⁸⁵ e o arquivo de parâmetros, que fornece um mapeamento entre as interações ligadas e não ligadas envolvendo as várias combinações de tipos de átomos encontrados no arquivo de topologia e parâmetros para todos os termos de ligação, ângulo, diédrico, impróprio e van der Waals na função potencial do campo de força.⁸⁵ Na preparação a proteína ainda é

solvatada (utilizando solvente, como por exemplo água, com os modelos TIP3P ou TIP4P⁸⁶), em uma caixa (cúbica, octaédrica ou dodecaédrica) de tamanho definido, com contra-íons (como NaCl ou KCl) para estabilizar a carga total do sistema, ou deixar o sistema em uma molaridade de sal pré-definida.

Uma vez preparado o sistema, a simulação é feita em, basicamente, quatro etapas:

Minimização: Na etapa de solvatação, algumas moléculas de água podem ser colocadas em posições desfavoráveis energeticamente, como por exemplo o hidrogênio da água muito próximo do hidrogênio da cadeia lateral de uma arginina, promovendo repulsão eletrostática que se não for corrigida antes da simulação de DM, pode gerar uma explosão na simulação, ou mudanças na conformação da proteína, ou desnaturações.⁷⁹ Nesta etapa o solvente se adapta ao redor do soluto. É importante ressaltar, ainda, que são estabelecidas as chamadas “condições periódicas de contorno”, Neste caso, as caixas contendo solvente, íons e proteína são replicadas em todas as dimensões, de forma periódica.⁷⁹ Dessa forma, caso uma molécula saia da caixa central, ela não entrará em contato com o vácuo, e sim com solvente. Nesse caso, uma de suas imagens entrará pela face oposta, mantendo o número de moléculas constante.⁷⁹

Aquecimento: Nessa etapa, ocorre o aumento de temperatura do sistema até a temperatura desejada (geralmente 300 K) de maneira linear, para fornecer Energia Potencial (U) ao sistema.

Equilibração: Durante o aquecimento, além de energia potencial, energia cinética também é inserida no sistema. O estágio de equilibração é usado para equilibrar as energias cinética e potencial, ou seja, para distribuir a energia cinética "bombeada" para o sistema durante o aquecimento entre todos os graus de liberdade. Em outros termos, a energia cinética deve ser transferida para energia potencial.⁸²

Produção: Nesta etapa, o sistema é analisado ao longo de um determinado tempo. No presente trabalho, essa etapa teve uma montagem do tipo NPT, com a quantidade de partículas constante, pressão e temperatura constantes, com uso de termostato e barostato, se aproximando, assim, das condições desejadas.⁸³

2 OBJETIVOS

2.1 Objetivos gerais

O presente trabalho teve por objetivo caracterizar enzimas degradadoras de xilano, em particular as hidrolases de glicosídeo (GHs) da família 10, 11, 43 e uma xilose isomerase (XI), através de técnicas como cristalografia de raios X, SAXS, espectrometria de massas e simulações de dinâmica molecular.

2.2 Objetivos específicos

- Para as xilanases da família GH10 os objetivos foram identificar os produtos de hidrólise através de espectrometria de massas e utilizar a cristalografia de raios X para determinação da estrutura e consequente compreensão da relação estrutura-padrão de clivagem;
- Para as xilanases da família GH11 os objetivos foram utilizar a cristalografia de raios X e simulações de dinâmica molecular para compreender o perfil enzimático incomum apresentado por essa enzima;
- Para as xilosidases da família GH43, os objetivos foram utilizar a cristalografia de raios X e simulações de dinâmica molecular para entender o papel da mutação F507A na atividade enzimática, além da maior atividade dessa enzima em XOS maiores;
- Para a xilose isomerase foram realizados ensaios para definir os parâmetros enzimáticos, tais como pH e temperatura ótimos, constante de Michaelis-Mentem (K_M), constante de inibição (K_I), além de ensaios de cristalografia de raios X para determinação da estrutura, SAXS para determinar o envelope molecular, além de análises filogenéticas.

3 CAPÍTULO 1 – *TcXyn10A*

3.1 Introdução

A descoberta e caracterização de enzimas capazes de transformar agro-resíduos, inclusive xilano, em produtos sustentáveis que promovam o crescimento da bioeconomia é extremamente importante, como já discutido anteriormente. Nesta nova realidade, é crucial encontrar novos microorganismos que possam transformar o xilano, por exemplo, em compostos de interesse comercial. *Thermobacillus composti* é uma bactéria gram-negativa, aeróbia, formadora de esporos e termotolerante isolada de um reator de compostagem.⁸⁷ Apesar de algumas caracterizações prévias do mecanismo de ligação de GH10 a substratos decorados,^{15,88} houve uma demanda, no presente trabalho, para investigar a base estrutural de atividade xilanolítica, a fim de explicar os produtos bioquimicamente gerados. Aqui, descrevemos a caracterização dos produtos de hidrólise e a estrutura cristalográfica de uma xilanase recombinante de *Thermobacillus composti* GH10 (*TcXyn10A*) (GenBank ID: AGA58864.1) com resolução de 2,1 Å com o objetivo de desvendar detalhes moleculares do reconhecimento de glucurono e arabinoxilano, explicando assim os produtos de interesse comercial, os XOS, gerados na hidrólise do xilano. Esse trabalho foi realizado em colaboração com a Dra. Vanessa Pellegrini e a BSc. Ana Gabriela Sepulchro, com a caracterização bioquímica e enzimática realizadas por elas, com resultados que deram origem, junto com os resultados obtidos pelo autor desta tese e apresentados neste capítulo, a um artigo de co-autoria dividida pelos três (Anexo A2).

3.2 Material e métodos

3.2.1 Clonagem, expressão e purificação

A etapa de clonagem foi realizada pelo Dr. César Moises Camilo, enquanto as etapas de expressão e purificação foram realizadas pela Dra. Vanessa Pellegrini e pela Ana Gabriela Sepulchro.

A sequência genômica para a xilanase GH10 de *Thermobacillus composti* KWC4 (*TcXyn10A*) (GenBank ID: AGA58864.1) foi amplificada e clonada no vetor de expressão

pETTRXA-1a / LIC.⁸⁹ A cepa *Escherichia coli* DH5 α foi usada como receptor para transformações durante a construção, propagação e armazenamento de plasmídeo, enquanto *E. coli* Rosetta (DE3) (Novagen, Darmstadt, Alemanha) foi usada para expressão de proteína.

A expressão heteróloga foi realizada em meio líquido LB na presença de canamicina e cloranfenicol nas concentrações finais de 50 $\mu\text{g/mL}$ e 34 $\mu\text{g/mL}$ respectivamente. A expressão foi realizada a 37 °C e 200 rpm em um agitador por 4 horas até que a densidade celular atingisse uma DO₆₀₀ de 0,6. Em seguida, as células cultivadas foram induzidas com 0,5 mM de isopropil- β -D-tiogalactopiranosídeo (IPTG) sob agitação contínua a 20 °C em um agitador a 200 rpm durante 16 horas. A cultura foi centrifugada a 7.878 x g a 4 °C durante 15 minutos e o pellet celular foi ressuspensão em um tampão de lise (500 mM NaCl, 50 mM Tris-HCl a pH 7,5, 50 mg/mL lisozima, 1 mM fenilmetanossulfonil fluoreto (PMSF) e 1 mM ditioneitol (dTT)). As células ressuspensas foram lisadas por oito ciclos de sonicação em banho de gelo usando um 550 Sonic Dismembrator Sonifier (Fisher Scientific, Hampton, EUA) em 40% de amplitude. O lisado foi clarificado por centrifugação a 20.200 x g durante 20 minutos a 4 °C para remover resíduos celulares e o sobrenadante foi utilizado para a purificação de TcXyn10A.

O sobrenadante foi aplicado a uma coluna Ni-NTA Superflow de 5 mL (Qiagen, Hilden, Alemanha), previamente equilibrada com um tampão contendo Tris-HCl 50 mM a pH 7,5 e NaCl 150 mM. Foi utilizado um gradiente crescente de imidazol (5 mM, 20 mM, 200 mM e 500 mM) para a eluição do material que não era de interesse, e a fração de 200 mM de imidazol contendo a TcXyn10A foi concentrada em um concentrador de corte molecular de 50 kDa para reduzir a concentração de imidazol para cerca de 10 mM. A amostra resultante foi incubada com protease recombinante Tobacco Etch Virus (TEV) a 4 °C por 16 h para a remoção da cauda 6xHis-tiorredoxina da enzima alvo.

Em seguida, uma segunda etapa de cromatografia de afinidade de Ni²⁺ foi utilizada para separar a TcXyn10A sem tag de contaminantes, incluindo a protease TEV. Uma etapa final de cromatografia de exclusão por tamanho foi realizada no sistema Äkta Purifier 10 (GE Healthcare, Chicago, EUA) usando uma coluna HiLoad 16/600 Superdex 75 pg (GE Healthcare, Chicago, EUA) e o tampão mencionado acima. A purificação foi avaliada por um volume eluído vs gráfico de absorvância, e as frações de proteína purificadas foram submetidas a análise de eletroforese em gel de dodecil sulfato de sódio a 15% (SDS-PAGE) sob condições desnaturantes⁹⁰ e subsequentemente coradas com Azul Brilhante de Coomassie

R-250. A enzima purificada foi posteriormente concentrada e sua concentração determinada medindo a absorção a 280 nm usando o espectrofotômetro Nanodrop 1000 (Thermo Scientific, Waltham, EUA) e aplicando o coeficiente de extinção calculado ($\epsilon = 81360 \text{ M}^{-1} \text{ cm}^{-1}$).

3.2.2 Padrão de clivagem enzimática

Produtos solúveis clivados enzimaticamente por *TcXyn10A* de glucuronoxilano de madeira de faia (Sigma-Aldrich, St. Louis, EUA) e arabinoxilano de centeio (Megazymes, Wicklow, Irlanda) foram analisados usando o sistema HPAEC ICS-5000 (Dionex, Sunnyvale, EUA) e espectrômetro de massa Microflex LT MALDI-TOF (Bruker Daltonics, Massachusetts, EUA). As reações foram realizadas em concentrações e condições ideais de pH (7.0) e temperatura (65 °C) estabelecidas previamente por Vanessa Pellegrini e Ana Gabriela Sepulchro.⁹¹

3.2.2.1 Análises de cromatografia de troca aniônica de alto desempenho (HPAEC)

Para analisar a formação do produto em função do tempo de reação, foram retiradas alíquotas após 5, 15 e 30 minutos e 1, 2, 4, 8 e 24 horas. As amostras foram fervidas a 95 °C por 5 minutos, centrifugadas a $9.600 \times g$ por 10 minutos a 4 °C e analisadas pelo sistema HPAEC (ICS-5000, Dionex, Sunnyvale, EUA), equipado com coluna analítica CarboPAC1 250 mm $\times$ 2 mm (Dionex, Sunnyvale, EUA).

Para avaliar a formação de produtos de clivagem por *TcXyn10A* a partir de xilopentaose (X5) e xilohexaose (X6), os XOS foram preparados nas concentrações iniciais de 0,1 mg/mL e os produtos foram analisados após 0, 1, 5 e 10 minutos de tempo de reação. Duas soluções foram utilizadas para a eluição das amostras: uma sendo NaOH 100 mM (solução A) e outra sendo acetato de sódio 500 mM mais NaOH 100 mM (solução B). A taxa de fluxo foi fixada em 0,5 mL / min a 30 °C. O volume de injeção foi de 1 μL e a eluição foi realizada usando os procedimentos das etapas usando as seguintes etapas: separação isocrática (20 min, 100% A), separação por gradiente (7 min, 12-100% B), lavagem da coluna (3 min, 100% B). Essa etapa foi realizada pela Ana Gabriela Sepulchro e Dra. Vanessa Arnoldi Pellegrini.

3.2.2.1.1 Produção enzimática de xilooligossacarídeos

Para quantificar o açúcar redutor liberado e poder determinar o rendimento de conversão, xilose, xilobiose (X2), xylotriose (X3), xilopentose e xilohexose puras (todas da marca Sigma-Aldrich, St. Louis, EUA) foram usadas como os padrões para identificação e quantificação dos produtos gerados de hidrólise de glucuronoxilano de madeira de faia (Sigma-Aldrich, St. Louis, EUA) e arabinoxilano de centeio (Megazymes, Wicklow, Irlanda).

3.2.2.2 Caracterização dos produtos por espectrometria de massas (MALDI-TOF)

Posteriormente, os produtos solúveis da reação de 24 horas foram analisados por dessorção / ionização a laser assistida por matriz com espectrometria de massa com detecção de tempo de voo (MALDI-TOF) usando Microflex LT MALDI-TOF (Bruker Daltonics, Massachusetts, EUA). Uma alíquota do produto da reação enzimática (0,4 µL) foi misturada com 0,8 µL de matriz de ácido 2,5-di-hidroxibenzoico (DHB) a 10 mg/mL e deixada secar sobre a placa de amostra. As medições foram feitas no modo de íon positivo linear dentro de um intervalo de 400–3300 m/z. As massas médias foram atribuídas e processadas usando o software flexAnalysis™ (Bruker Daltonics, Massachusetts, EUA) e MMass.⁹²

3.2.3 Cristalização, coleta de dados e processamento

Para os ensaios de cristalização, a *TcXyn10A* purificada foi concentrada a 10 mg/mL e xilose foi adicionada à solução (100 mM; Sigma-Aldrich, St. Louis, EUA). Os ensaios de cristalização foram realizados usando um robô de cristalização Honeybee (DIGILAB, Massachusetts, EUA) e os cristais foram formados e crescidos por técnica de difusão de vapor em gotas com diferentes proporções de proteína e solução. Os cristais foram detectados em condições de cristalização consistindo em 0,1 M HEPES, 70% (v/v) 2-metil-2,4-pentanodiol (MPD), pH 7,5 a 20 °C, com 2 µL de proteína misturada com 1 µL de tampão na gota de cristalização.

Os cristais de *TcXyn10A* resultantes foram congelados rapidamente em nitrogênio líquido após imersão em óleo mineral. Os dados de difração de raios X foram coletados na linha de luz W01B-MX2 do Laboratório Nacional de Luz Síncrotron (Campinas, Brasil),⁹³

utilizando um detector Pilatus2M (Dectris, Filadélfia, EUA), com um tempo de exposição de 3 s por imagem e um ângulo de oscilação de 0,1 °. Os dados da imagem de difração foram indexados, integrados e escalonados usando o pacote XDS⁹⁴ e AIMLESS.⁹⁵

3.2.4 Determinação estrutural por difração de raios-X

A xilanase GH10 de *Geobacillus stearothermophilus* (código PDB 1N82), com 61,6% de identidade sequencial de aminoácidos e uma cobertura de 96%, foi identificada como o modelo de estrutura de homologia mais próxima de acordo com o BLASTP.⁹⁶

A substituição molecular foi realizada por Phaser⁹⁷ usando as coordenadas atômicas da xilanase de *G. stearothermophilus* como modelo de pesquisa. A estrutura foi então submetida pelo Autobuild⁹⁸ e os ciclos de refinamento foram realizados usando os programas Coot⁹⁹ e PHENIX-refine.¹⁰⁰ A qualidade do modelo foi avaliada usando o programa Molprobit.¹⁰¹

As coordenadas atômicas foram depositadas no Protein Data Bank (PDB) sob o código de acesso 6WQW.

3.2.5 Análise estrutural

A fim de compreender o padrão de clivagem, a estrutura do cristal de TcXyn10A foi sobreposta com xilanases GH10 de *Cellvibrio mixtus* (CmXyn10B)⁸⁸ e *Streptomyces olivaceoviridis* (SoXyn10A)¹⁰² (PDB ids 1UR1, 1UQZ e 1V6V), complexadas com o XOS decorados em seus sítios ativos. O programa PyMol (The PyMOL Molecular Graphics System, Versão 2.0 Schrödinger, LLC.) foi usado para comparações estruturais, sobreposições e medidas. As sequências SoXyn10A e CmXyn10B foram alinhadas contra TcXyn10A em Clustal Omega¹⁰³ e, em seguida, submetidas ao programa ESPript¹⁰⁴, para comparações de sequências baseadas em estrutura.

3.3 Resultados e discussão

3.3.1 Clonagem, expressão e purificação

O gene da xilanase de *Thermobacillus composti* da família GH10 (GenBank: AGA58864.1) foi clonado com sucesso no vetor de expressão pETTRXA-1a / LIC e transformado em células de *E. coli* Rosetta. O TcXyn10A foi expresso como uma proteína de 49 kDa fusionada com His-Trx e a solubilidade da proteína expressa foi avaliada por SDS PAGE (Figura 11a).

Figura 11 - Expressão e solubilidade de TcXyn10A analisada por SDS-PAGE. (A) Etapa de expressão. M: marcador molecular; 1: Frações retiradas com 0h de indução; 2: 16h após indução; Após a lise celular 3: fração insolúvel; 4: fração solúvel. (B) Etapa final de purificação. (M) marcador molecular.

Fonte: Adaptada de SEPULCHRO *et al.*⁹¹

A purificação da proteína foi realizada por cromatografia de afinidade ao Ni^{+2} . Após a remoção do His-Trx-tag com a protease TEV, a proteína de interesse foi coletada após uma segunda etapa de cromatografia de afinidade de Ni^{+2} , onde ela foi eluída imediatamente após a inserção na coluna, enquanto o His-Trx-tag e a TEV permaneceram ligadas à resina da coluna. Em seguida, uma cromatografia de filtração em gel foi realizada para garantir a pureza da amostra. A massa teórica de TcXyn10A prevista com base em sua sequência de aminoácidos é 38,54 kDa e a proteína purificada migrou como uma única banda em SDS PAGE, garantindo assim sua pureza (Figura 11b).

3.3.2 Padrão de clivagem enzimática

O padrão de clivagem enzimática foi avaliado na presença de diferentes substratos: xilotriose (X3), xilotetraose (X4), xilopentaose (X5) e xilohexaose (X6) (Figura 12), e também glucuronoxilano de faia e arabinoxilano de centeio (Figura 13).

Para reunir informações detalhadas sobre o padrão de hidrólise de XOS pela *TcXyn10A*, Ana Gabriela Sepulchro e Vanessa Pellegrini analisaram a clivagem de X3, X4, X5 e X6. O padrão obtido foi: X3 não foi hidrolisado pela enzima (Figura 12A), X4 foi hidrolisado pela enzima em X2 (Figura 12B), X5 foi hidrolisado pela enzima em X2 e X3, enquanto apenas pequenas quantidades de X4 foram observadas (Figura 12C). X6 levou a uma formação de pequenas e decrescentes quantidades de X4, enquanto que nenhum X5 foi observado (Figura 12D).

Um padrão hidrolítico final semelhante para estes dois últimos oligossacarídeos foi observado, com a enzima liberando X2 e X3 como os principais produtos finais de sua hidrólise (Figura 12C e D). Estes resultados indicam que *TcXyn10A* preferencialmente reconhece e cliva oligossacarídeos mais longos (de X4 em diante). Xilo-oligossacarídeos mais curtos (como X2 e X3) não são clivados por *TcXyn10A*.⁹¹

Figura 12 - Figura 12: Padrão de clivagem TcXyn10A. Os perfis do produto liberado por TcXyn10A sob condições ótimas de enzima de diferentes XOS medidos por HPAEC: (A) xilotriose (X3), (B) xilotetraose (X4); (C) xilopentaose (X5) (D) xilohexaose (X6). Xn em cada perfil indica o padrão de xilooligossacarídeos variando de xilose a hexaose. X3 não é clivado pela enzima, enquanto X4 é clivado em X2. Tanto X5 quanto X6 são degradados pela enzima predominantemente em X2 e X3.

Fonte: Adaptada de SEPULCHRO *et al.*⁹¹

Os padrões de produtos de hidrólise enzimática de glucurono- e arabinoxilanos foram avaliados em função do tempo por HPAEC, em experimentos realizados por Ana Gabriela Sepulchro e Vanessa Pellegrini, revelando a formação de xilobiose (X2), xilotriose (X3), xilotetraose (X4), xilopentaose (X5) e xilohexaose (X6) (Figura 13C e D). Quantidades de xilose (X1) estavam presentes entre os produtos em tempos de reação mais longos, especialmente em arabinoxilano (Figura 13C e D).

A espectrometria de massa MALDI-TOF foi realizada pelo autor da tese e também foi utilizada para avaliar os padrões de clivagem de amostras após 24 horas de clivagem (Figura 13A e B). Os picos correspondentes a XOS X2 e X3 não decorado e o produto decorado MeGlcA²Xyl3^{Na+} foram detectados entre os produtos de clivagem de xilano de madeira de faia por *TcXyn10A* (Figura 13A). O espectro de massa da hidrólise de arabinoxilano revelou X2 e X3 como os principais produtos não decorados (Figura 13B). A gama de produtos decorados é consideravelmente mais complexa e inclui AraX2, AraX3, Ara2X3 e Ara2X4.

Figura 13 - Avaliação da produção de xilo-oligosacarídeos por *TcXyn10A*. (A) Produção de XOS a partir de glucuronoxilano de madeira de faia medida por (A) HPAEC; e (C) MALDI-TOF, e arabinoxilano de centeio medido por (B) HPAEC; e (D) MALDI-TOF. A enzima libera uma ampla gama de XOS de glucuronoxilano de madeira de faia e de arabinoxilano de centeio. Xn em cada perfil indica o padrão de XOS variando de xilose a xilohexaose.

Fonte: Adaptada de SEPULCHRO *et al.*⁹¹

Além disso, o rendimento de conversão em XOS foi medido por Ana Gabriela Sepulchro e Vanessa Pellegrini. Após 24 horas de reação, os rendimentos totais de conversão de xilano em XOS foram de 29,09% e 16,98% para glucuronoxilano e arabinoxilano, respectivamente.⁹¹ Além disso, a hidrólise enzimática do glucuronoxilano por TcXyn10A produziu quantidades significativamente menores de xilose (0,49%) em comparação com arabinoxilano (2,34%), e o principal produto hidrolítico para ambas os xilanos foi X2 seguido por X3 (Tabela 2).

Tabela 3 - Hidrólise enzimática de diferentes fontes de xilano em 24 horas.

Fonte do Xilano	Taxa de conversão total de XOS (%)	Distribuição de XOS (%)					
		X1	X2	X3	X4	X5	X6
Glucuronoxilano de faia	29,58	0,49	19,56	6,25	2,43	0,85	-
Arabinoxilano de centeio	19,32	2,34	13,62	3,36	-	-	-

Fonte: SEPULCHRO *et al.*⁹¹

A produção enzimática do XOS por TcXyn10A possui, portanto, um potencial biotecnológico promissor que vai além do etanol de segunda geração, uma vez que os XOS possuem ampla gama de aplicações na indústria farmacêutica, agricultura, veterinária e como ingredientes para rações e alimentos.⁹¹ Assim, a enzima TcXyn10A apresenta um potencial promissor para transformar glucurono- e arabinoxilanos em XOS que podem contribuir para a valorização de resíduos subutilizados, como a biomassa vegetal resultante da produção de combustíveis de primeira geração, originando produtos sustentáveis de maior valor agregado.⁹¹

3.3.4 Determinação e análise da estrutura cristalográfica de TcXyn10A

3.3.4.1 Coleta de dados, refinamento, construção de modelo e comparação com estruturas homólogas

Os dados de difração de raios-X foram coletados no Laboratório Nacional de Síncrotron (LNLS),⁹³ utilizando uma rotação de 0,1 graus, resultando em 3600 imagens para coleta 360°. O processamento dos dados com XDS revelou que os cristais de TcXyn10A obtidos pertencem ao grupo espacial P4₃2₁2.

A substituição molecular foi realizada com Phaser usando a estrutura de xilanase GH10 de *G. stearothermophilus* (PDB id 1N82) como modelo, resultando em uma única solução. A estrutura foi refinada utilizando o programa phenix.refine, e submetida ao Autobuild, a fim de colocar as cadeias laterais em falta. Isso levou a um declínio no R-free e no R-factor, e mais ciclos de construção e refinamento de modelo foram realizados usando Coot e PHENIX. A estrutura final foi avaliada com o MolProbity. Os parâmetros de coleta e refinamento são fornecidos na Tabela 4.

Tabela 4 - Parâmetros de coleta e refinamento cristalográfico de TcXyn10A

Conjunto de dados	TcXyn10A
Coleta de dados	LNLS MX2
Comprimento de onda (Å)	1,458
Grupo espacial	P 4 ₃ 2 ₁ 2
Parâmetros de cela	
$a=b, c$ (Å)	81,264; 92,363
Faixa de resolução (Å)	33,82 – 2,10 (2,18 – 2,1)
R_{meas}	0,1133 (1,181)
CC _{1/2}	1 (0,868)
Reflexões totais	422.460 (35.127)
Reflexões únicas	18.595 (1.809)
Multiplicidade	22,7 (19,4)
Completeza (%)	98,40 (99,78)
Mosaicidade	0,3
$\langle I/\sigma_I \rangle$	18,26 (2,23)
Refinamento	
No. mol. unidade assimétrica	1
Resíduos por molécula	305
	(continua)

(continuação)

R_{work}	0,2345 (0,3095)
R_{free}	0,2674 (0,3723)
CC geral	1 (0,961)
Ligantes	XYP
<i>B-factors</i> ($\text{\AA}^2$)	
Médio	48,22
Macromoléculas	48,58
Ligantes	54,99
<i>R.m.s. deviations</i>	
Comprimentos de ligação ($\text{\AA}$)	0,004
Ângulos de ligação ($^\circ$)	0,89
<i>Ramachandran</i>	
Ramachandran favorecidos (%)	96,99
Ramachandran permitidos (%)	3,01
Ramachandran outliers (%)	0,00
Rotâmeros outliers (%)	0
<i>Escore geral do MolProbity</i>	1,52

Os valores em parênteses são relativos à camada de mais alta resolução.

Fonte: Adaptada de SEPULCHRO *et al.*⁹¹

O monômero TcXyn10A (Figura 14a) apresenta um barril TIM α/β óctuplo clássico¹⁰⁵, típico para enzimas GH10 (PDB- <http://www.rcsb.org/>).¹⁰⁶ No decorrer do refinamento, a densidade de elétrons forte e contínua no sítio ativo da enzima foi identificada e interpretada como xilose (ver mapa Omit na Figura 14a).

Figura 14 - Estrutura cristalográfica. (A) Estrutura 3D de *TcXyn10A* revelando a estrutura conservada de barril α/β (barril TIM), com xilose (XYL) no sítio ativo sobreposto com o mapa Omit correspondente no nível de 1,2 sigma. Os resíduos do sítio ativo na vizinhança da xilose são mostrados como bastões e as quatro alças são coloridas em vermelho escuro, azul, amarelo e roxo. O programa UCSF Chimera v1.13 (www.cgl.ucsf.edu/chimera)¹⁰⁷ foi usado para criar esta figura; (B) Alinhamento estrutural baseado em sequência de *TcXyn10A*, *CmXyn10B* e *SoXyn10A*. Os resíduos catalíticos são destacados em triângulos vermelhos; Os resíduos do site ativo são destacados em azul; Os resíduos de interação de arabinose são representados em triângulos rosa, enquanto os resíduos de interação de MeGlcA são representados em triângulos amarelos. Os quatro loops são mostrados em vermelho escuro, azul, amarelo e roxo, respectivamente. (C) Resíduos de aminoácidos na vizinhança de XYL observados em *TcXyn10A* (verde) são mostrados. Resíduos correspondentes de estruturas homólogas de *CmXyn10B* (PDB ids 1UR1 e 1UQZ)⁸⁸ e estrutura de *SoXyn10A* (PDB id 1V6V)¹⁰² são mostrados em amarelo, azul e marrom, respectivamente. As distâncias de interação com a xilose são fornecidas em Angstroms; (D) A sobreposição estrutural de *TcXyn10A* (mostrado em azul) com dois outros GH10s, *CmXyn10B* (PDB id 1UR1 e 1UQZ, colorido em amarelo e azul, respectivamente) e *SoXyn10A* (PDB id 1V6V, representado em marrom) é dado na representação de fita.

A comparação de *TcXyn10A* com estruturas de xilanase GH10 mostrou que as enzimas compartilham uma arquitetura estrutural muito conservada. A fenda catalítica *TcXyn10A* tem 18,2 Å de comprimento e compreende 6 subsítios. A altura média da fenda do sítio ativo é 7,6 Å e sua profundidade é 5,5 Å. Um alinhamento de sequência baseado em estrutura de *TcXyn10A*, *CmXyn10B* e *SoXyn10A* foi feito com ESPript, e o resultado pode ser visto na Figura 6b, e a Tabela 4 mostra a comparação entre essas enzimas, incluindo a porcentagem de identidade de sequência e cobertura.

Tabela 5 - Comparações entre *TcXyn10A*, *CmXyn10B* e *SoXyn10A*

Enzima	PDB id	Identidade (%)	Cobertura (%)
<i>TcXyn10A</i>	6WQW	100	100
<i>CmXyn10B</i>	1UQZ, 1UQY, 1UR1	39.77	96
<i>SoXyn10A</i>	1V6V	34.98	89

Fonte: Adaptada de SEPULCHRO *et al.*⁹¹

Apesar da conservação estrutural considerável dentro da família GH10 de enzimas, uma série de diferenças, especialmente no comprimento e conformação de algumas alças na estrutura *TcXyn10A* são óbvias. Quando comparado com seu homólogo mais próximo, *Xyn10* de *G. stearothermophilus* (PDB id 1N82), *TcXyn10A* tem um RMSD de 0,585 Å, e quando sobreposto com *CmXyn10B* e *SoXyn10A*, *TcXyn10A* tem um RMSD de 0,808 Å e 1,046 Å, para as posições Cα, respectivamente. O sítio catalítico contém dois resíduos Glu (Glu132 e Glu239 na enzima *TcXyn10A* e Glu134 e Glu241 na xilanase de *G. stearothermophilus*). Glu 132 é descrito como resíduo de ácido/base e é precedido por uma asparagina (Asn 131 na sequência de aminoácidos *TcXyn10A*) em uma xilanase GH10 típica.¹⁰⁸ A asparagina está normalmente envolvida em importantes interações de ligações de hidrogênio com o substrato.¹⁰⁸

Glu239 é o nucleófilo catalítico, presente na sequência conservada ITELD¹⁰⁹ como observado na sequência GH10 de *G. stearothermophilus*. No entanto, para *TcXyn10A* a isoleucina inicial é substituída por valina, um resíduo semelhante, mas um pouco mais curto. Esses resíduos catalíticos realizam a reação de retenção seguindo o mecanismo de deslocamento duplo de Koshland.¹¹⁰⁻¹¹¹

A partir do alinhamento estrutural, torna-se claro que a sequência do resíduo ácido/base é estritamente conservada na parte Asn - Glu, com pequenas diferenças no Ala subsequente (Val em alguns casos) e alta variabilidade no Val 134. Uma vez que esses resíduos estão na área de sítio ativo, isso pode interferir nas interações com ligantes, principalmente com os decorados. Na região ao redor do nucleófilo (Glu 239), a sequência Thr - Glu - Leu - Asp é estritamente conservada.

3.3.4.2 Reconhecimento estrutural para decorações de xilano e perfil de produtos hidrolíticos

Para entender os determinantes estruturais do reconhecimento de *TcXyn10A* de xilano decorado (glucurono e arabinoxilano) e não decorado, a estrutura cristalográfica da enzima foi sobreposta com *CmXyn10B* (código PDB 1UQY) (Figura 15A) usando Pymol, copiando o ligante para *TcXyn10A*. Esta estrutura particular foi escolhida porque foi determinada com uma xiloheptaose (X7) ligada ao seu sítio ativo e pode ser sobreposta com *TcXyn10A* com RMSD de 0,825 Å.

Diferentes XOS foram sobrepostos a partir de estruturas existentes, como *CmXyn10B* e *SoXyn10A*, com a estrutura da enzima usando o X7 de *CmXyn10B* como guia, usando o programa Coot para avaliar a posição correta do xilosídeo no subsítio -1. O X6 não decorado se encaixa bem no sítio ativo *TcXyn10A* e se sobrepõe com a xilose observada experimentalmente na posição -1 (Figura 15 A e B).

Figura 15 - Bases estruturais do reconhecimento de *TcXyn10A* de xilano não decorado e glucuronoxilano. (A) Vista frontal da enzima, com a fenda catalítica em seu centro e X6 no sítio ativo, com os dois resíduos catalíticos (Glu132 e Glu239) representados em vermelho. Os subsítios da fenda do sítio ativo são numerados de -3 a +3, (B) MeGlcAX3, co-cristalizado com *SoXyn10A*, doca bem nos subsítios negativos *TcXyn10A*, (C) As cadeias laterais dos resíduos interagindo com o MeGlcAX3 (em verde) e o as distâncias para os resíduos interagentes, em Å, são mostradas. Resíduos correspondentes da estrutura de *SoXyn10A* são mostrados para comparação (em marrom), (D) MeGlcAX3, co-cristalizado com *CmXyn10B* posicionado nos subsítios positivos do sítio ativo, (E) As distâncias de interação de MeGlcAX3, para os resíduos vizinhos de *TcXyn10A* (em verde) nos subsítios positivos (+1 a +3). A posição e orientação dos resíduos correspondentes da estrutura da GH10 de *C. mixtus* são mostradas em azul para comparação. Os resíduos em (C) e (E) são numerados e coloridos de acordo com as sequências correspondentes.

Fonte: SEPULCHRO *et al.*⁹¹

A análise dos subsítios negativos do sítio ativo *TcXyn10A* deixa claro que os possíveis produtos hidrolíticos são X2 e X3, que ocupam os subsítios -2, -1, +1, +2 e +3, respectivamente (Figura 15A). A enzima também aceita bem MeGlcA3X3 com uma decoração no subsítio -3.¹⁵ O resultado é consistente com os resultados de espectrometria de massas e análises HPAEC dos produtos da degradação de glucuronoxilano de madeira de faia por *TcXyn10A* (Figura 13). A análise detalhada da fenda catalítica de *TcXyn10A* mostra impedimento estérico em uma série de posições do ligante (posições -1 e -2), resultando em poucas decorações permitidas. Além disso, *TcXyn10A* possui Arg304, cuja cadeia lateral interage com a decoração MeGlcA no subsítio -3, indicando possível estabilização do glucuronoxilano decorado dentro da fenda do sítio ativo de *TcXyn10A*, quando a decoração é encontrada no subsítio -3. Nos subsítios positivos, a decoração MeGlcA na posição +1 permitida e estabelece novas interações com a enzima, particularmente com as cadeias laterais dos resíduos Asp136 e Tyr177 (Figura 15E). Isso corrobora observações anteriores de que as xilanases com um resíduo Tyr no subsítio +1 têm maior atividade sobre o glucuronoxilano.¹⁵

As diferenças estruturais entre *TcXyn10A* e *CmXyn10B*, em particular uma interação extra feita pelo resíduo Arg304 da *TcXyn10A* e a decoração MeGlcA, podem sugerir porquê *TcXyn10A* apresenta uma afinidade maior para glucuronoxilano do que *CmXyn10B*, com um K_M menor (K_M *TcXyn10A* = 1.8 mg.mL⁻¹; K_M *CmXyn10B* = 3.79 mg.mL⁻¹).^{88, 91} Em termos de especificidade de substrato, das duas enzimas utilizadas para o alinhamento estrutural na figura 15, apenas os dados de *CmXyn10B* estão disponíveis na literatura. Esta enzima tem especificidade para glucuronoxilano de 17.782 U/mol, o que é aproximadamente três vezes maior do que para arabinoxilano de centeio, 5.882 U/mol. Convertendo as unidades para poder realizar uma análise comparativa, os valores para *TcXyn10A* tornam-se 4×10^6 U/mol para glucuronoxilano e $4,5 \times 10^6$ U/mol para arabinoxilano de centeio. Observa-se então que em *CmXyn10B* a especificidade para glucuronoxilano e arabinoxilano de centeio é consideravelmente menor (200 e 700 vezes, respectivamente) do que em *TcXyn10A*, indicando que a *TcXyn10A* seja muito mais eficiente e tenha um desempenho mais homogêneo nos dois tipos de xilano.⁹¹

A fim de investigar o reconhecimento de arabinoxilano, XOS decorados foram construídos com base na estrutura de *CmXyn10B*. Para locais negativos, a arabinoxilobiose e a arabinoxilotriose foram posicionadas sobrepondo a estrutura cristalográfica de *TcXyn10A* a *CmXyn10B* com ligantes (PDB ids 1UR2 e 2CNC). Para subsítios positivos, uma vez que atualmente não há estrutura cristalográfica com este tipo de ligante, o programa Coot foi usado para construir decorações de arabinose em um backbone de xilotriose de *CmXyn10B*, e o programa Pymol foi usado para verificar o ligante. Diferentemente do glucuronoxilano, onde as decorações de ácido D-glucurônico ou ácido 4-O-metil-D-glucurônico são ligadas ao O-2 do resíduo de xilose dentro do backbone do xilano, o arabinoxilano pode ter decorações de arabinose ligadas ao O-2 e/ou O-3 de resíduos de xilose dentro da estrutura de xilano, como pode ser observado na figura 16.

Figura 16 - Representação esquemática de arabinoxilano e glucuronoxilano. As unidades de D-xilose de uma estrutura de xilano são mostradas como estrelas laranja, as decorações de L-arabinose como estrelas verdes e as decorações de ácido D-glucurônico são mostradas como diamantes divididos em azul e branco, seguindo a Nomenclatura de Símbolo para Representação Gráfica de Glicanos.¹¹²⁻¹¹³

Fonte: SEPULCHRO *et al.*⁹¹

Nossa análise estrutural mostra que as decorações de arabinose são permitidas em quase qualquer subsítio dentro do sítio ativo de *TcXyn10A*, exceto no subsítio -1. Na posição -1, a decoração Ara entra em conflito com a cadeia lateral do resíduo Trp297 e, portanto, não é permitida. Sabe-se que os subsítios -2 e -3 das xilanases GH10 têm capacidade para acomodar arabinoxilanos decorados e esta característica é conservada nas xilanases GH10.¹⁵ A região aglicona (subsítios positivos em GH10) é menos conservada e a permissibilidade para decorações de arabinoxilano varia de xilanase a xilanase.^{88,102,114-115}

É conhecido que as xilanases GH10 sem um subsítio +3 podem acomodar arabinose no subsítio +2, enquanto as xilanases com um subsítio +3 não podem.⁸⁸ Ao mesmo tempo, para *CmXyn10B*, as decorações de L-arabinofuranosídeo ligadas em 3' são observadas no subsítio +2 e poderiam, muito provavelmente, ser toleradas quando ligadas covalentemente aos resíduos de xilopirranose nas posições -3 e +4. Nossa análise do sítio de ligação de *TcXyn10A* indica que virtualmente qualquer subsítio positivo (aglicona) pode acomodar decorações de arabinose (Figura 17), e nos subsítios negativos a única posição proibida é -1. Nossas análises estruturais também mostram que, apesar do *TcXyn10A* ter um subsítio +3, a decoração de arabinose no subsítio +2 é permitida.

Figura 17 - Interações putativas de *TcXyn10A* com arabinosilano. (A) AraX2 observado em *CmXyn10B* ajustado dentro dos subsítios negativos do sítio ativo de *TcXyn10A* (de -2 a -1), com decoração de arabinose na posição -2; (B) As distâncias entre AraX2 refinado dentro do sítio ativo *TcXyn10A* (em verde) para os resíduos vizinhos são mostradas. Os resíduos correspondentes da estrutura homóloga *CmXyn10B* são dados a laranja; (C) AraX3 da estrutura homóloga *CmXyn10B* ancorada nos subsítios negativos (-3 a -1) do sítio ativo *TcXyn10A*. A decoração arabinose está na posição -2, (D) A mesma sobreposição que em (C) mostrando as distâncias para os resíduos de interação de *TcXyn10A* (mostrado em verde) e *CmXyn10B* (representado em roxo). Os resíduos em (B) e (D) são coloridos e numerados de acordo com a respectiva enzima. (E) Arabinosilano nos subsítios positivos (aglicona) (+1 a +3). As decorações de arabinose são modeladas nas posições +1, +2 e +3, (F) A mesma sobreposição que em (E), exibindo as distâncias para os resíduos de interação do sítio ativo de *TcXyn10A*; (G) Vista lateral de (E), mostrando o espaço disponível para decoração Ara no local +2; (H) Distâncias dos resíduos do sítio ativo ao ligante como mostrado em (G).

Fonte: SEPULCHRO *et al.*⁹¹

De acordo com estudos anteriores,¹⁵ subsítios positivos podem ou não acomodar decorações. A decoração de arabinose é geralmente bem acomodada dentro do subsítio +1 dos sítios ativos de xilanases GH10 e há uma permissibilidade de tais decorações no subsítio +2, previamente demonstrada por Fujimoto e colegas,¹⁰² mas geralmente não é conservada entre enzimas da família GH10. No mesmo estudo, nenhuma decoração é mencionada para o subsítio +3, mas nossa análise estrutural também indica que *TcXyn10A* pode aceitar decorações de arabinose do substrato dentro do subsítio +3. Embora o grupo O-2 do ligante não esteja completamente exposto ao solvente, parece que o subsítio pode, no entanto, acomodar a decoração de arabinose, que pode formar interações do tipo ligações de hidrogênio com a cadeia lateral de Glu137 (a uma distância de 3,4 Å) e a cadeia lateral de Asp241 (a uma distância de 2,7 Å) (Figura 17).

O resíduo Tyr295 da estrutura de *TcXyn10A* interage com a decoração Ara no subsítio -2, enquanto a xilanase *CmXyn10B* homóloga tem uma substituição Ala na posição correspondente. O reconhecimento de decorações Ara por *TcXyn10A* pode indicar interações mais fortes da enzima com o arabinoxilano decorado. A comparação das análises estruturais com nossos estudos bioquímicos experimentais demonstra claramente que a clivagem de *TcXyn10A* de glucuronoxilano de faia e arabinoxilano de centeio em pequenos XOS depende das decorações do substrato. A enzima cliva os substratos estudados produzindo preferencialmente X2 e X3 não decorados, o que foi confirmado por técnicas analíticas e explicado pelas análises estruturais dos substratos no sítio ativo da enzima.

Além disso, a enzima produz MeGlcA3X3 como o principal XOS decorado liberado por *TcXyn10A* a partir de glucuronoxilano de faia. O fato da presença de decorações serem mais abundante em arabinoxilano e a maior permissibilidade de *TcXyn10A* para decorações Ara em seu sítio ativo resulta na produção de um número maior de XOS decorados ao agir sobre arabinoxilano (Figura 18).

Figura 18 - Padrão de clivagem de acordo com estudos estruturais de subsítios de *TcXyn10A*. (A) Padrão de clivagem de xilohexaose (X6) e xilotetraose (X4) e seus possíveis produtos; (B) Glucuronoxilano, com decoração MeGlcA nos subsítios -3 e +1, e os produtos; (C) Padrão de clivagem de arabinoxilano e seus possíveis produtos, de acordo com a posição da decoração de arabinose. As unidades de D-xilose de uma estrutura de xilano são mostradas como estrelas laranja, as decorações de L-arabinose como estrelas verdes e as decorações de ácido D-glucurônico são mostradas como diamantes divididos em azul e branco, seguindo a Nomenclatura de Símbolo para Representação Gráfica de Glicanos.¹¹²⁻¹¹³

Fonte: SEPULCHRO *et al.*⁹¹

De acordo com Charnock *et al.*,¹¹⁴ o subsítio -2 exibe uma forte afinidade por resíduos de xilose, o que o torna essencial para a clivagem do xilano. Isso também é válido para *TcXyn10A*. Na verdade, a hidrólise de X6 e X5 por *TcXyn10A* não resultou em quantidades significativas de X1, indicando a necessidade de o subsítio -2 ser ocupado como um pré-requisito para uma clivagem enzimática eficiente. Curiosamente, a produção de X1 foi observada quando a enzima agia sobre arabino e glucuronoxilano decorado. Isso é consistente com a noção de que as interações diretas entre as decorações nos subsítios positivos (+1 a +3) e os substratos podem ter um efeito positivo na clivagem do substrato pela enzima. Além disso, uma quantidade de xilose liberada parece ser diretamente proporcional a um grau de decoração do substrato. Claramente *TcXyn10A* libera quantidade significativamente maior de X1 de arabinoxilano de centeio (substituição de Ara de cerca de 38%, com liberação de cerca de 2,34% de X1) do que de xilano de faia (substituição de MeGlcA de aproximadamente 13%, com liberação de cerca de 0,49% de X1) (Figura 13 e Tabela 3). Isso pode indicar que a

produção enzimática de XOS usando *TcXyn10A* poderia ser mais eficiente a partir de glucuronoxilano do que em arabinoxilano.

3.4 Conclusões

O padrão de clivagem enzimática de *TcXyn10A* foi avaliado experimentalmente usando HPAEC e MALDI-TOF e revelou a liberação predominante de X2 e X3 a partir de xilanos decorados e não decorados pela enzima. Quanto aos oligossacarídeos decorados, *TcXyn10A* libera predominantemente MeGlcA3X3 a partir de glucuronoxilano de faia e uma variedade de XOS decorados a partir de arabinoxilano de centeio. De uma maneira geral, a utilização da *TcXyn10A* se mostrou bastante interessante do ponto de vista biotecnológico, uma vez que produziu uma variedade considerável de XOS a partir de glucuronoxilano e arabinoxilano. Além disso, a estrutura cristalográfica da enzima foi resolvida com resolução de 2,1Å, desvendando os determinantes estruturais que definem uma gama de produtos hidrolíticos e explicam a permissibilidade para as decorações no sítio ativo da enzima. A análise estrutural conseguiu identificar diferenças entre os resíduos de interação do ligante com *TcXyn10A* e *CmXyn10B*, fornecendo informações sobre a atividade enzimática e também enriquecendo a literatura anterior, propondo o subsítio +3 como capaz de acomodar decorações Arabinose.⁹¹

4 CAPÍTULO 2 – PxXyn11B

4.1 Introdução

Embora o mecanismo catalítico canônico para as enzimas GH11 seja bem descrito em termos estruturais e bioquímicos,^{108,116} a descoberta de um sítio secundário de ligação (SBS) levantou novas e interessantes questões sobre o mecanismo molecular pelo qual este sítio de ligação modula a catálise promovida por membros GH11.¹¹⁷ Estudos anteriores atribuíram algumas funções putativas ao SBS, que incluem (i) direcionar a enzima para o substrato, (ii) guiar o substrato para o sítio ativo, (iii) alterar a especificidade da enzima e (iv) agir como um sítio alostérico.¹¹⁸ Até o momento, poucas estruturas cristalográficas de GH11 identificaram glicanos ligados ao SBS, incluindo as endo-xilanases de *Bacillus subtilis* (PDB id 2QZ3), *Aspergillus niger* (PDB id 2QZ2)¹¹⁸⁻¹¹⁹ e *Bacillus circulans* (BcX) (PDB id 1XNB),¹²⁰ mas o mecanismo molecular pelo qual a ligação ao SBS influencia o sítio ativo da enzima permanece desconhecido.

Anteriormente descrita como GH11XynB, PxXyn11B, uma xilanase uni-modular do gênero recentemente descrito *Paenibacillus xylanivorans*,¹²¹ é ativo em xilano de madeira de faia (85 IU/mg) e arabinoxilano de trigo (131,5 IU/mg), exibindo características cinéticas muito peculiares.¹²² Em vez de uma curva hiperbólica típica esperada para uma cinética enzimática do tipo Michaelis-Menten, uma curva sigmoidal não-Michaeliana foi observada com um coeficiente de Hill (η_H) de 4, sugerindo um alto grau de cooperatividade,¹²² bastante incomum para membros da família GH11. Achados semelhantes também foram descritos para uma quimera de proteína de ligação xilanase-xilose, com um coeficiente de Hill ($\eta_H > 2,0$)¹²³ e uma GH11 de *Pseudozyma brasiliensis*.¹²⁴ Conforme discutido por Ghio *et al.*,¹²² este resultado poderia sugerir a possível presença de sítios secundários de ligação (SBSs) em PxXyn11B. Apesar das evidências bioquímicas consistentes, os mecanismos moleculares detalhados pelos quais as interações dos xilo-oligossacarídeos (XOS) com o SBS podem resultar em um aumento na atividade enzimática de forma cooperativa ainda não são compreendidos.

Ludwiczek e colegas¹²⁰ mostraram que para a GH11 de *Bacillus circulans* (BcX) a enzima se liga a XOS no sítio ativo e ao SBS com constantes de dissociação semelhantes (K_D), mas diferenças marcantes nas taxas de koff.¹²⁰ Os autores também descobriram que mutações no SBS aumentaram a afinidade por XOS no sítio ativo, sugerindo algum grau de

interferência entre os sítios e indicando alguma cooperatividade.¹²⁰ Esses achados indicam que o aumento da atividade enzimática devido às interações entre XOS e SBS pode ser mediado por uma diminuição na constante K_M .

Uma ativação das enzimas por seus produtos tem uma relevância biotecnológica considerável. Uma vez que muitas enzimas do tipo GH são fortemente inibidas por seus produtos hidrolíticos,¹²⁵ seu uso em aplicações industriais é restrito a um número relativamente pequeno de enzimas conhecidas ou projetadas para serem mais tolerantes à inibição do produto e/ou às condições em que os produtos são (parcialmente) removidos da reação.¹²³ Por outro lado, para as enzimas GH11 com o SBS, o XOS produzido durante a hidrólise enzimática do xilano causa a ativação da enzima via interação com o SBS. Isso poderia resultar em um aumento na eficiência quando usado em bioprocessos sob condições industrialmente relevantes. Dada a importância das xilanases GH11 para uma série de aplicações biotecnológicas, este efeito pode ser de relevância geral para a aplicação industrial das enzimas da família GH11.

No presente estudo, foi determinada a primeira estrutura cristalina de uma xilanase GH11 do gênero *Paenibacillus* e foram conduzidas simulações de Dinâmica Molecular (DM) para fornecer informações estruturais sobre o papel do SBS na atividade enzimática de PxXyn11B. O resultado deste trabalho foi publicado pelo autor desta tese.¹²⁶

4.2 Material e métodos

4.2.1 Clonagem, expressão e purificação

A etapa de clonagem foi realizada pelos colaboradores da Argentina, Eleonora Campos e Silvina Ghio. As etapas de expressão e purificação foram realizadas no laboratório do Grupo de Biotecnologia Molecular, através do aluno Caio Capetti, com a supervisão da Dra. Vanessa Pellegrini.

A clonagem de genes e a expressão de proteínas foram previamente descritas em Ghio *et al.*¹²² Resumidamente, *Escherichia coli* cepa DH5 α foi utilizada como um receptor para transformação durante a construção, propagação e armazenamento do plasmídeo, enquanto *E. coli* BL21 foi utilizada para expressão da proteína. A cepa foi gentilmente cedida pela Professora Eleonora Campos (IABIMO, Argentina).

Um pré-inóculo de *E. coli* BL21 contendo o plasmídeo pJExpress (NdeI/XbaI) *PxXyn11B* foi inoculado em meio LB-ampicilina (100 µg/mL) e incubado a 37 °C e 200 rpm em um agitador, durante a noite. Em seguida, 5 mL da pré-cultura foram usados para inocular 500 mL do meio de expressão. Depois de atingir uma OD₆₀₀ de 0,6 a 37 °C e 200 rpm, a expressão da proteína foi induzida com 0,5 mM de isopropil-β-D-tiogalactopiranosídeo (IPTG) a 21 °C sob agitação contínua a 200 rpm por 16 horas. A cultura foi centrifugada a 4.000 x g a 4 °C por 20 minutos e o pellet celular foi ressuspenso em um tampão de lise, pH 8.0 (20 mM Tris-HCl, 20 mM KCl, 1 mM EDTA, 0,1% (v/v) Triton-X e imidazol 10 mM). Em seguida, a lisoizima foi adicionada, a uma concentração final de 1 mg / mL. Após isso, seis ciclos de 10 segundos de sonicação em banho de gelo usando 550 Sonic Dismembrator Sonifier (Fisher Scientific, Hampton, EUA) a 40% da amplitude máxima foram realizados. O lisado foi clarificado por centrifugação a 10.000 x g por 30 minutos a 4 °C para remover detritos celulares e o sobrenadante foi usado para a etapa posterior de purificação de *PxXyn11B*.

O sobrenadante foi aplicado em uma coluna Ni-NTA Superflow de 5 mL (Qiagen, Hilden, Alemanha), previamente equilibrada com o tampão de lise. Três volumes de tampão de lavagem (20 mM Tris-HCl, 20 mM KCl, 1 mM EDTA, 0,1% (v/v) Triton-X, 20 mM Imidazol, pH 7.0) foram usados para lavar a coluna. A proteína foi eluída com dois tampões de eluição, ambos com Tris-HCl 20 mM, pH 8.0, KCl 20 mM e EDTA 1 mM, mas com diferentes concentrações de imidazol: o primeiro tampão tinha 250 mM enquanto o segundo tinha 500 mM de imidazol. As frações de proteína resultantes foram reunidas e concentradas em um concentrador de corte molecular de 50 kDa para reduzir a concentração de imidazol para 20 mM. A purificação foi avaliada por análise de eletroforese em gel de dodecilsulfato de sódio a 15% (SDS-PAGE) sob condições desnaturantes⁹⁰ e subsequentemente corada com Coomassie Brilliant Blue R-250. A enzima purificada foi posteriormente concentrada usando um concentrador Vivaspín de 10 kDa (GE-Healthcare, Chicago, IL, EUA) a 1.500 x g e sua concentração foi determinada medindo a absorção a 280 nm usando espectrofotômetro Nanodrop 1000 (Thermo Scientific, Waltham, EUA) e aplicando o coeficiente de extinção calculado ($\epsilon = 75860 \text{ M}^{-1} \text{ cm}^{-1}$).

4.2.2 Cristalização, coleta de dados e processamento

Para os ensaios de cristalização, a *PxXyn11B* purificada foi concentrada a 10 mg/mL. Os ensaios de cristalização foram realizados utilizando um robô de cristalização Honeybee (DIGILAB, Massachusetts, EUA) e os cristais foram cultivados por técnica de difusão de vapor em gotas com diferentes proporções de proteína para solução, usando o kit Crystal Screen (Hampton Research, Califórnia, EUA). Os cristais cresceram em condições de cristalização consistindo em citrato tribásico de sódio 1,6 M pH 6,5 a 20 °C, na proporção de 2:1, com 2 µL de proteína e 1 µL de tampão na gota de cristalização.

Os cristais de *PxXyn11B* resultantes foram congelados rapidamente em nitrogênio líquido após imersão em etilenoglicol (30% v/v). Os dados de difração de raios X foram coletados na linha de luz W01B-MX2 do Laboratório Nacional de Luz Síncrotron (Campinas, Brasil),⁹³ utilizando um detector Pilatus2M (Dectris, Filadélfia, EUA), com um tempo de exposição de 2 s por imagem e um ângulo de rotação do cristal de 0,5 °. Os dados da imagem de difração foram indexados e integrados utilizando o pacote XDS⁹⁴ e escalonados com AIMLESS.⁹⁵

4.2.3 Solução da estrutura cristalográfica

A substituição molecular foi realizada utilizando o programa BALBES¹²⁷, utilizando a xilanase GH11 de *Bacillus subtilis* (PDB id 2Z79¹¹⁹), que possui 85,2% de similaridade de sequência de aminoácidos. A estrutura resultante foi então submetida ao Phenix AutoBuild⁹⁸ e os ciclos de refinamento foram realizados utilizando os programas Coot⁹⁹ e phenix.refine.¹⁰⁰ A qualidade do modelo foi avaliada usando Molprobity,¹⁰¹ e as coordenadas atômicas foram depositadas no Protein Data Bank (PDB) sob o código de acesso 7KV0.

4.2.4 Análise estrutural

4.2.4.1 Comparações Estruturais

Para fins de comparação estrutural, a estrutura cristalográfica de *PxXyn11B* foi sobreposta com uma xilanase GH11 de *Trichoderma reesei* complexada com xilohexaose (X6) no sítio ativo (PDB id 4HK8¹²⁸), e uma estrutura cristalográfica de uma GH11 de *Bacillus subtilis* mutante E172A em complexo com xilotetraose com SBS (XBS1, PDB id 2QZ3,¹¹⁹). O software de visualização UCSF Chimera¹⁰⁷ foi usado para sobreposições e comparações.

4.2.5 Simulações de Dinâmica Molecular

As simulações de dinâmica molecular de *PxXyn11B* e posteriores análises foram realizadas sob supervisão do Professor Alessandro Nascimento, utilizando o cluster Nascimento e o programa AMBER18.¹²⁸⁻¹³⁰ A estrutura cristalográfica de *PxXyn11B* sem moléculas de água foi usada como modelo inicial para simulações de DM. Complexos de *PxXyn11B* ligados a X4 no sítio ativo com e sem X3 no SBS foram criados utilizando a mesma estrutura inicial. Os estados de protonação em pH 6,5 e na presença NaCl 150 mM foram estimados usando servidor H++.¹³² Os sistemas apo, xylotetraose ligado no sítio ativo (*PxXyn11B*-X4) e xylotetraose ligado no sítio ativo com xylotriose ligado no SBS (*PxXyn11B*-X4-X3SBS) foram solvatados em caixas octaédricas com preenchimento de solvente de 12 Å. Os íons cloreto foram adicionados para atingir a neutralidade na caixa de simulação. As simulações foram feitas utilizando o campo de força AMBERFF 14SB,¹³³ o modelo de água TIP3P⁸⁶ e o campo de força de carboidrato GLYCAM06¹³⁴ para os xilosídeos. Cada sistema foi minimizado em 20.000 etapas e, em seguida, aquecido a 300 K em 50 ps na montagem NVT. Posteriormente, a densidade do sistema foi equilibrada em 50 ps na montagem NPT. Restrições harmônicas foram mantidas para minimização, aquecimento e equilíbrio de densidade com pesos harmônicos de 10,0, 10,0 e 5,0 kcal/(mol.Å²), respectivamente. Finalmente, o sistema foi equilibrado por 2 ns sem restrições. Simulações produtivas foram executadas por 500 ns na montagem NPT. Um intervalo de tempo de 2 fs foi usado e as ligações covalentes envolvendo átomos de hidrogênio foram restringidas usando o algoritmo SHAKE. As trajetórias de DM foram analisadas usando VMD,¹³⁵ CPPTRAJ¹³⁶ e scripts/pequenos códigos desenvolvidos internamente.

4.3 Resultados e discussão

4.3.1 Clonagem, expressão e purificação

PxXyn11B foi expressa e purificada com cromatografia de afinidade de Ni⁺². Em seguida, uma cromatografia de filtração em gel foi realizada para garantir a pureza da amostra. A massa teórica de *PxXyn11B* prevista com base em sua sequência de aminoácidos é de 20,8 kDa e a proteína purificada migrou como uma única banda no SDS PAGE, com o peso molecular esperado e uma pureza necessária para estudos estruturais e bioquímicos (Figura 19).

Figura 19 - Purificação de *PxXyn11B*. O gel SDS a 12% mostra os resultados das etapas de purificação da *PxXyn11B*. M: marcador de baixo peso molecular; Lisado: a fração solúvel do lisado; F.I.: fração insolúvel do lisado (pellet); Flow-through: fluxo observado na purificação de IMAC; Lv: fração de lavagem; 250 mM imidazol: eluição de proteína com imidazol 250 mM.

Fonte: BRIGANTI *et al.*¹²⁶

4.3.2 Determinação e análise da estrutura cristalográfica de *PxXyn11B*

4.3.2.1 Coleta de dados, refinamento, construção de modelo e comparação com estruturas homólogas

Os dados de difração cristalográfica foram coletados no Laboratório Nacional de Síncrotron (LNLS) na linha MX2,⁹³ utilizando uma rotação angular de 0,2 graus por imagem, resultando em 1.800 imagens para coleta 360°. O processamento dos dados com XDS revelou que o cristal de *PxXyn11B* pertence ao grupo espacial C222₁.

A substituição molecular foi realizada com o programa BALBES utilizando estrutura de xilanase GH11 de *B. subtilis* (PDB id 2Z79) como um modelo, resultando em uma única

solução com pontuação MR de 0,66 e um fator Q de 0,845. A estrutura foi refinada até 2,5 Å e submetida ao AutoBuild,⁹⁸ a fim de adicionar as cadeias laterais ausentes, ou modifica-las. Isso levou a uma diminuição no R_{free} e R_{work} , e mais ciclos de construção e refinamento do modelo foram realizados usando Coot⁹⁹ e phenix.refine.¹⁰⁰ A estrutura final foi avaliada com o programa MolProbity.¹⁰¹ Ambos os parâmetros de coleta e refinamento são fornecidos na Tabela 1.

Tabela 6 - Coleta de dados cristalográficos de PxXyn11B e estatísticas do refinamento do modelo.

Conjunto de dados	PxXyn11B
Coleta de dados	LNLS MX2
Comprimento de onda (Å)	1,458
Grupo especial	C 2 2 2 ₁
Parâmetros de cela	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	62,426 62,995 183,200
Faixa de resolução (Å)	45,80 – 2,50 (2,59 – 2,50)
R_{pim}	0,0593 (0,5444)
CC _{1/2}	0,996 (0,564)
Reflexões totais	164.962 (15.900)
Reflexões únicas	12.880 (1.266)
Multiplicidade	12,8 (12,6)
Completeza (%)	99,86 (99,53)
Mosaicidade	0,3
$\langle I/\sigma_I \rangle$	10,30 (1,39)
Refinamento	
No. mol. unidade assimétrica	2
Resíduos por molécula	186

(continua)

(continuação)

R_{work}	0,2149 (0,3094)
R_{free}	0,2480 (0,3254)
CC geral	0,999 (0,849)
Ligantes	EDO
<i>B-factors</i> ($\text{\AA}^2$)	
Médio	45,65
Macromoléculas	45,67
Ligantes	47,95
<i>R.m.s. deviations</i>	
Comprimentos de ligação ($\text{\AA}$)	0,002
Ângulos de ligação ($^\circ$)	0,56
<i>Ramachandran</i>	
Ramachandran favorecidos (%)	96,47
Ramachandran permitidos (%)	3,53
Ramachandran outliers (%)	0,00
Rotâmeros outliers (%)	1,01
<i>Escore geral MolProbity</i>	1,62

Os valores em parênteses são relativos à camada de mais alta resolução.

Fonte: BRIGANTI *et al.*¹²⁶

PxXyn11B apresenta duas moléculas na unidade assimétrica, mas uma vez que sua comparação não indicou diferenças significativas e a sobreposição das moléculas A e B mostrou um RMSD de 0,212 $\text{\AA}$, as análises e representações adicionais foram feitas sobre a molécula A. Cada monômero (Figura 1a) tem um formato beta jelly roll, típico do clã GH-C,¹³⁷ formado por 11 fitas β e uma alfa hélice, e uma fenda de aproximadamente 2.097 $\text{\AA}^3$, onde o sítio catalítico está localizado, de acordo com PDBsum.¹³⁸

Figura 20 - Estrutura cristalográfica de *PxXyn11B*. (A) Estrutura 3D de *PxXyn11B* colorida de azul (região N-terminal) a vermelho (região C-terminal). Hélices α e fitas β , conforme definido por PDBSum são indicadas, e também os resíduos catalíticos. (B) Vista lateral de (A), com as partes N-terminal, C-terminal, palma, dedos, polegar e cordão da proteína indicadas.

Fonte: BRIGANTI *et al.*¹²⁶

PxXyn11B tem a forma sanduíche β canônica das xilanases da família GH11 de "mão direita", na qual as fitas β se estendem para frente e para trás entre as duas folhas β empacotadas para formar os domínios do dedo e da palma¹³⁹ (Figura 1). Um laço estendido entre as fitas $\beta 8$ e $\beta 9$ faz parte do chamado domínio do polegar, e um laço entre as fitas $\beta 6$ e $\beta 7$ forma o cordão entre os dois domínios da folha β .

Para comparar a estrutura de *PxXyn11B* com outras enzimas da família GH11, a sequência da proteína foi submetida ao BLASTp⁹⁶ e oito sequências homólogas foram selecionadas (Tabela 2). As sequências das enzimas selecionadas foram alinhadas com *PxXyn11B* usando Clustal Omega¹⁰³ e um alinhamento de sequência baseado na estrutura foi realizado com ESPrpt¹⁰⁴ (Figura 21).

Figura 21 - Alinhamento de sequência baseado em estrutura. (A) As sequências de AnXyn11A, CTX, XynII, rXynA, EnXyn11A, TfGH11, NFX e XlnB2 foram alinhadas com a sequência PxXyn11B e o NTR (região N-terminal), resíduos catalíticos e resíduos do SBS são destacados. (B) Superposição estrutural das enzimas acima mencionadas. Azul marinho: AnXyn11A; Roxo: CTX; Verde escuro: XynII; Ciano escuro: PxXyn11B; Vermelho escuro: rXynA; Dourado: EnXyn11A; Laranja: TfGH11; Vermelho: NFX; Ciano: XlnB2. (C) Vista lateral de (B). (D) Vista posterior de (B).

Fonte: BRIGANTI *et al.*¹²⁶

Como pode ser observado na Figura 21, existem algumas diferenças entre as enzimas, principalmente no que diz respeito à sua região N-terminal (NTR). Esta região é mais longa em algumas enzimas, como XlnB2, EnXyn11A e AnXyn11A. Outras xilanases GH11 também possuem NTR mais longo, embora não tão longo como nas anteriores. Tem sido amplamente aceito que a região N-terminal (NTR) do GH11 é importante para a termoestabilidade da xilanase.¹⁴⁰ Por exemplo, as xilanases GH11 de *Nonomuraea flexuosa* e

Neocallimastix patriciarum têm alta termoestabilidade e NTRs longos, com o último tendo o NTR mais longo relatado para GH11.¹³⁹⁻¹⁴¹

Comparando o SBS dentro das sequências de xilanases da família GH11, é possível observar que vários resíduos desempenham papéis importantes neste local de ligação (Figura 2a). A análise de alinhamento de sequências de aminoácidos revela que N54 está presente apenas em *PxXyn11B* e *XBS1*, enquanto a mesma posição é identificada como S82 em *AnXyn11A*. Outros resíduos importantes, N139 e W183, novamente estão apenas presentes em *PxXyn11B* e *XBS1* também. Isso indica que os resíduos N54, N139 e W183 podem ser responsáveis pela formação de SBS. A comparação das estruturas da enzima, em relação à porcentagem de identidade, cobertura e desvios da raiz média quadrada (RMSD), é mostrada na Tabela 7.

Tabela 7 - Enzimas utilizadas na comparação estrutural com PxXyn11B

Microorganismo	Enzima	Identidade (%)	Cobertura (%)	RMSD (Å)	PDB id	Referência
<i>Panibacillus</i> <i>xylanivorans</i> A59	PxXyn11B	-	-	-	7KV0	Este estudo
<i>B. subtilis</i>	XBS1	85.33	96	0.343	2QZ3	Vandermarliere ¹¹⁹
<i>Thermopolyspora</i> <i>flexuosa</i>	NFX	66.47	89	0.643	1M4W	Hakulinen ¹⁴¹
<i>Thermobifida</i> <i>fusca</i>	TfGH11	65.22	96	0.667	3ZSE	Van Bueren ¹⁴³
<i>Streptomyces</i> <i>lividans</i>	XlnB2	63.69	94	0.468	5EJ3	Gagné ¹⁴⁴
<i>Escherichia coli</i>	EnXyn11A	60.44	94	0.594	2VGD	Vardakou ¹⁴⁵
<i>Chaetomium</i> <i>thermophilum</i>	CTX	59.12	94	0.554	1H1A	Hakulinen ¹⁴¹
<i>Trichoderma</i> <i>reesei</i>	XynII	53.04	94	0.542	4HK8	Wan ¹²⁸
<i>Aspergillus niger</i>	AnXyn11A	43.45	86	0.833	6QE8	Schroder ¹⁴⁶

Fonte: BRIGANTI *et al.*¹²⁶

4.3.2.2 Análise do sítio catalítico

A fim de analisar o sítio ativo de PxXyn11B, a enzima foi sobreposta a GH11 de *Trichoderma reesei* (PDB id 4HK8)¹²⁸ complexada com xilohexaose (X6), ligada ao seu sítio ativo (Figura 22).

Figura 22 - Análise estrutural do sítio ativo de PxXyn11B. (A) Estrutura cristalográfica de PxXyn11B com X6 de *Trichoderma reesei* (PDB id 4HK8) sobreposta. (B) Sobreposição de PxXyn11B (ciano escuro) e GH11 de *Trichoderma reesei* (verde escuro). (C) Subsítios em PxXyn11B, com os respectivos resíduos. Os triângulos vermelhos indicam o local de clivagem. Os resíduos xilanopirosídicos são representados como estrelas amarelas, de acordo com a nomenclatura de símbolos dos glicanos.^{112, 113} (D) Zoom nos sítios ativos, com os resíduos exibidos com as respectivas cores.

Fonte: BRIGANTI *et al.*¹²⁶

A hidrólise enzimática da ligação glicosídica requer um doador de prótons e um nucleófilo/base (Glu 170 e Glu 77 em PxXyn11B, respectivamente). A hidrólise por hidrolases de glicosídeo pode ocorrer por dois mecanismos distintos: retenção ou inversão da configuração anomérica do substrato.¹¹⁰ Em ambos os mecanismos de retenção e inversão, a posição do doador de prótons é idêntica, dentro da distância de ligação de hidrogênio do oxigênio glicosídico, mas em enzimas de retenção, a base catalítica nucleofílica está na vizinhança próxima do carbono anomérico do açúcar.^{116,146} No caso da GH11, como a PxXyn11B, o mecanismo é de retenção.¹⁴⁸

O sítio ativo de *PxXyn11B* compreende 6 subsítios, semelhantes ao observado na GH11 de *Trichoderma reesei* (PDB id 4HK8). De acordo com Madan e Lee,¹⁴⁹ três resíduos de aminoácidos não são conservados dentro do sítio ativo: Y96, Y171 e S16. Em *PxXyn11B*, eles foram identificados como Y87, Y164 e Q6, ou seja, apenas o último resíduo não é conservado.

A análise do sítio ativo também indica a presença de dois importantes resíduos previamente descritos na literatura: Y87 e Y172 nos subsítios + 2 e + 3. A estrutura cristalográfica indica que esses resíduos de Tyr contribuem para a reação enzimática ao estabilizar os subsítios positivos. De acordo com esta descoberta, as substituições por alanina nos resíduos Y180 no subsítio + 2 e Y96 no subsítio + 3 mostraram uma redução nas atividades enzimáticas em até 95% em GH11 de *Thermomyces lanuginosus*, TlXynA.¹⁵⁰

4.3.2.3 Análise do SBS

O sítio secundário de ligação foi relatado anteriormente para estruturas de GH11 bacterianas para apenas uma enzima, a xilanase de *Bacillus subtilis* (XBS1, PDB id: 2QZ3).¹¹⁷⁻¹¹⁸ Para fins de comparação, *PxXyn11B* foi sobreposta à estrutura da xilanase de *Bacillus subtilis* e os resíduos dentro do SBS foram analisados.

Figura 23 - Sobreposição das estruturas *PxXyn11B* e *XBS1*. (A) Vista frontal, mostrando o sítio catalítico e o SBS na parte traseira. (B) Vista lateral de (A), com resíduos no sítio ativo indicados. (C) Ampliação da região do SBS, com os resíduos envolvidos destacados e as distâncias entre os resíduos de *PxXyn11B* e os resíduos de xilopiranosídeo indicadas.

Fonte: BRIGANTI *et al.*¹²⁶

Conforme observado no alinhamento mostrado na Figura 21 e também na Figura 23, *PxXyn11B* tem um SBS estritamente conservado, também presente na estrutura da xilanase de *Bacillus subtilis*. Além da similaridade de sequência e estrutura entre as enzimas *PxXyn11B* e *Bacillus subtilis*, a caracterização cinética prévia revelou uma curva sigmoideal, típica para enzimas cooperativas, semelhante ao que já foi observado para *PbXynA*.¹²⁴ Além disso, coeficientes de Hill de 4 e 6 foram determinados para *PxXyn11B*¹²² e *PbXynA*,¹²⁴ respectivamente. Portanto, pretendeu-se avaliar os papéis putativos da interação de XOS no SBS e seu papel na atividade enzimática usando simulações de dinâmica molecular.

4.3.4 Simulações de Dinâmica Molecular

A estrutura cristalográfica de *PxXyn11B* determinada experimentalmente foi usada para simulações de DM. Para uma investigação mais aprofundada, três sistemas foram preparados e simulados: uma apo enzima, determinada pela estrutura cristalográfica; a enzima ligada à xilotetraose (X4) em seu sítio ativo, e a enzima ligada a X4 no sítio ativo e a uma molécula de xilotriose (X3) no SBS. Em todos os casos, uma simulação produtiva de 0,5 μ s foi analisada usando CPPTRAJ,¹³⁶ AmberEnergy ++¹⁵¹ e também AMBERTOOLS.

As análises das simulações revelaram que *PxXyn11B* é uma enzima notavelmente estável, sendo observadas apenas pequenas flutuações na estrutura durante as simulações. Usando o pacote MDLovoFit,¹⁵² mapeamos o núcleo mais estável da enzima (a fração 30% menos móvel da estrutura da enzima), que inclui a maioria das β -fitas. Esta porção rígida da proteína é mostrada em cartoon azul na Figura 24 abaixo, enquanto a fração móvel da enzima é mostrada em vermelho na figura.

Figura 24 - (A e B) Duas vistas de *PxXyn11B* ligado a xilotetraose (X4). As figuras mostram a superposição de 500 quadros tirados em intervalos de 1 ns durante a simulação de equilíbrio de DM. Os subsítios do sítio ativo (-2 a +2) são indicados abaixo do painel (A). (C e D) Duas visualizações de *PxXyn11B* ligado a X4 no sítio ativo e a uma xilotriose (X3) no SBS. Ambos X3 e X4 são mostrados na forma de palitos. As linhas indicam a mudança na mobilidade perto da fita $\beta 5$. (E) Energia de interação (potencial) calculada entre a enzima e o substrato (X4) para o complexo binário (linha preta) e para o complexo ternário (linha vermelha), ou seja, a enzima ligada a X3 no SBS. Para efeito de comparação, as energias de interação entre a enzima e o X3 no SBS também são mostradas em linhas azuis.

O mapeamento da flutuação atômica através da estrutura da enzima, também mostrado na Figura 24, indica uma mobilidade ligeiramente reduzida para a enzima ligada a X3 no SBS em comparação com a enzima ligada a X4 no sítio ativo, mas com SBS vazio (Figura 25). A análise mostra que a porção C-terminal da enzima é dramaticamente estabilizada pelas interações com X3 ligado ao SBS.

Figura 25 - Comparação da mobilidade para o complexo PxxYn11B-X4 (topo) e complexo ternário PxxYn11B-X4 / X3 (parte inferior). Quatro visualizações são mostradas para cada sistema simulado. O X4 e o X3 são mostrados em palitos e as estruturas das proteínas são coloridas de acordo com sua mobilidade em uma escala de azul a vermelho.

Fonte: BRIGANTI *et al.*¹²⁶

Esta estabilização se propaga para as fitas β da região C-terminal, em particular $\beta 11$, a última fita β na estrutura, que abrange o sítio ativo e contém o resíduo catalítico Glu170 e Tyr172, conhecido fazer parte dos subsítios positivos +1 e +2 (Figura 22). Além disso, a palma da mão e a região do polegar mostram uma maior estabilização quando X3 está ligado ao SBS. É importante destacar que as coordenadas iniciais eram exatamente as mesmas para as simulações dos sistemas enzimáticos apo e ligado a X4 e a X3. Essa descoberta sugere que a estabilização da porção terminal da enzima por meio de interações com XOS pode reduzir diretamente a mobilidade do substrato por meio da estabilização dos resíduos no sítio ativo. Este resultado também é suportado pela ligeira diminuição na energia de interação entre a enzima e o substrato (X4) quando a simulação do complexo PxxYn11B-X4 é comparada com

o complexo ternário P_xXyn11B-X4-X3^{SBS} (Figura 24). Os resultados mostraram que a energia potencial média de interação entre a enzima e a xilotriose ligada no SBS foi de aproximadamente -57 kcal/mol, enquanto o substrato (xilotetraose) se liga à enzima com uma energia de interação de -95 e -125 kcal/mol para o binário e complexo ternário, respectivamente. Esses dados sugerem um melhor ajuste do substrato no sítio ativo quando o sítio é estabilizado pela interação de XOS no SBS na superfície da proteína.

Em consonância com esses resultados, a estimativa de energia de ligação com base no MM-GBSA (do inglês *Molecular mechanics with generalised Born and surface area solvation*) implementado no MM-PBSA.py¹⁵³ mostrou uma energia de interação de $-49,7 \pm 0,4$ kcal / mol para a interação de X4 com GH11-X3, enquanto para o complexo binário (X4-GH11), uma energia de interação de $-35,4 \pm 0,2$ kcal/mol foi calculada a partir de dados de DM, reforçando a melhor interação da enzima com o substrato na presença de um xilooligossacarídeo no SBS.

Outro ponto importante é que o domínio polegar-alça, conhecido por sua importância no mecanismo de catálise da GH11 por seu movimento ter papel fundamental na glicosilação/desglicosilação e liberação do produto,¹⁵³⁻¹⁵⁴ também é estabilizado pelo X3 em SBS (Figura 25). É importante notar também que os resultados relatados, baseados em simulações clássicas de DM, são consistentes com as medições de NMR de xilanase de *Bacillus circulans* feitas por Ludwiczek e colaboradores,¹²⁰ que relataram um aumento na afinidade pelo substrato após a ligação de XOS para o SBS. Além disso, Paes e colaboradores¹⁵⁴ estabeleceram valores de energia de ligação mais altos para os subsítios positivos, com -17,7 e -13,2 kcal/mol para os subsítios +1 e +2, respectivamente, do que para os subsítios negativos, com -14,0 e -23,8 kcal/mol para subsítios -1 e -2, respectivamente. Isso é consistente com a simulação de DM realizada neste estudo, onde os resíduos de xilopiranosídeo em subsítios positivos demonstraram maior mobilidade do que aqueles localizados em subsítios negativos. Além disso, a energia de interação total do complexo binário, com GH11 de *Thermobacillus xylanilyticus* e o ligante, foi calculada como -86,3 kcal/mol, o que é consistente com a energia calculada para o complexo binário P_xXyn11B, mas superior à energia encontrada para o complexo ternário P_xXyn11B.

Em termos mecanísticos, como a ligação de uma molécula X3 ao SBS se propaga como um efeito estabilizador do sítio ativo da enzima? Uma possível explicação para esse efeito estabilizador seria uma mudança na conformação do açúcar no subsítio -1, que é

conhecido por ser relevante para a atividade da enzima. Foram medidos os ângulos de dobramento de Cremer-Pople para o subsítio -1 em ambas as simulações, ou seja, com o X3 ligado no SBS e na ausência do X3 no SBS e em ambos os cenários foi constatado que a xilose -1 está na conformação de barco 4C_1 de baixa energia, semelhante às conformações observadas em complexos de Michaelis para outras β -xilanases.¹⁵⁶ Assim, a distorção da conformação de açúcar no subsítio -1 não parece ser a força motriz determinante nesse contexto.

Um exame mais aprofundado das trajetórias de DM revelou um grande reposicionamento nas cadeias laterais de Tyr172 e Asn62. Esses resíduos, que estão localizados nas fitas $\beta 11$ e $\beta 5$, que são as mesmas fitas- β que são diretamente estabilizadas pela interação do XOS no SBS (ver triângulos azuis na Figura 21). Além disso, esses resíduos desempenham um papel importante na coordenação da interação entre o substrato e os subsítios positivos da enzima +1 a +3 (Figura 22). Na ausência de um ligante ligado no SBS, essas interações tornam-se mais soltas, resultando em uma diminuição na energia de interação (enzima-substrato) e, muito provavelmente, na catálise menos eficiente (Figura 26). Curiosamente, esses achados estão de acordo com o estudo de Ludwiczek e colaboradores,¹²⁰ que relataram que o aumento na atividade da enzima devido à interação no SBS foi devido a uma diminuição no K_M para substratos solúveis e insolúveis. A análise dos dados de simulação sugere então que o mecanismo envolvido no aumento da atividade envolve a estabilização direta das fitas- β 11 e 5, com a consequente estabilização dos resíduos nos subsítios positivos do sítio ativo. A estabilização desses resíduos resulta em uma interação mais estreita com o substrato, aumentando a afinidade (ou a energia de interação), diminuindo o K_M e levando a um aumento da eficiência.

Figura 26 - Comparação entre o complexo binário PxXyn11B-X4 (azul) e o complexo ternário PxXyn11B-X4 / X3 (laranja). Estruturas representativas do cluster mais populoso na trajetória de DM são comparadas. Os átomos de hidrogênio foram omitidos para maior clareza. X4 e X4 são mostrados em palitos e os subsítios são numerados na parte inferior da figura. O reposicionamento de N62, Q173 e Y172 no complexo binário cria um espaço adicional que permite que os açúcares no sítio positivo flutuem, afrouxando a interação com a enzima.

Fonte: BRIGANTI *et al.*¹²⁶

Também é interessante notar que X2 e X3 são alguns dos principais produtos gerados pelas xilanases da família GH11. Portanto, tomados em conjunto, os estudos estruturais e simulações de dinâmica molecular revelam um ciclo de feedback bioquímico pelo qual os principais produtos da reação hidrolítica de xilano, X2 e X3, catalisada pela xilanase GH11, se liga a seu SBS e causa mudanças conformacionais que se propagam e estabilizam o sítio ativo da enzima. Esta regulação alostérica induzida por produto pode ser comum para xilanases GH11 com o SBS evoluído para regular a afinidade do substrato em enzimas de domínio único, de alguma forma paralela ao que é observado em GHs contendo CBM.¹²⁰

Foram também simulados diferentes substratos na presença de uma molécula X3 ligada ao SBS. Vale a pena notar, na simulação de PxXyn11B ligado a X5 no sítio ativo e X3

no SBS, foi observado que X3 escapou do SBS e se difundiu para o solvente. Após cerca de 100 ns, a molécula X3 atingiu um local adicional próximo à região do polegar e permaneceu ligada a este novo local até o final da simulação (Figura 27a).

Figura 27 – Simulação de DM de *PxXyn11B* com X5 no sítio ativo e X3 no SBS. (A) Quadros representativos: primeiro painel no tempo zero, painel 2 após 130 ns, painel 3 após 220 ns, painel 4 após 250 ns, onde X3 atinge um local adicional, próximo ao subsítio -2, estabilizando a região do polegar da enzima, conformação observada até o final da simulação, após 500 ns (quinto painel). (B) Análise da energia potencial de interação entre *PxXyn11B* e uma molécula X3, inicialmente ligada ao SBS (linha vermelha) e também entre *PxXyn11B* e uma molécula X5 ligada ao sítio ativo nos subsítios -2 a +3 ao longo de 500 ns.

Fonte: BRIGANTI *et al.*¹²⁶

Nota-se que esta interação também contribuiu para a estabilização do substrato no sítio ativo, conforme observado a partir das energias de interação (Figura 27b). Esses resultados sugerem que mais de um local pode existir na superfície da enzima contribuindo

para a estabilização da mesma e também para a estabilização do substrato. Em linha com este achado, ambas as estruturas cristalinas de GH11 de *B. subtilis* ligado à xilotriose, 2QZ2 e 2QZ3, têm uma molécula X3 ligada na superfície da enzima, mas em diferentes locais de ligação, sugerindo que mais de um local pode existir e contribuir para a enzima estabilização.

4.4 Conclusões

A estrutura cristalográfica de raios X de *PxXyn11B* e seu alinhamento de sequência de aminoácidos revelam a conservação da região do SBS. As comparações estruturais e simulações de DM fornecem uma explicação para a comunicação molecular entre SBS e sítio ativo, desvendando a base molecular do comportamento cooperativo de *PxXyn11B*. Um ciclo de feedback positivo mediado por produto promove interações alostéricas entre o sítio ativo e o SBS da enzima, envolvendo a estabilização estrutural da região C-terminal da enzima. Este mecanismo molecular poderia ser empregado para a engenharia molecular de enzimas da família GH11 para uma melhor performance industrial, através da ativação das mesmas por seus produtos hidrolíticos, aumentando a eficiência do processo.

5 CAPÍTULO 3 – *EcXyl43*

5.1 Introdução

GH43, uma das 164 famílias de GH no banco de dados CAZy,^{21,156} inclui enzimas com várias atividades hemicelulolíticas, sendo as mais comuns β -xilosidase,¹⁵⁸ α -L-arabinofuranosidases¹⁵⁹ e endo- α -L-arabinanases.¹⁵⁹⁻¹⁶⁰ Recentemente, a família GH43 foi dividida em 37 subfamílias.¹⁶² As β -xilosidases da família GH43 empregam um mecanismo catalítico que leva à inversão da configuração de carbono anomérico do produto durante a conversão de xilo-oligossacarídeos curtos (XOS) em xilose, impedindo a síntese de oligossacarídeos longos por transglicosilação, que promove a síntese de oligossacarídeos. Isso é uma característica vantajosa para a produção de bioetanol de segunda geração, uma vez que este processo requer a produção de monômeros como a xilose a partir da biomassa lignocelulósica.¹⁶³

O gênero *Enterobacter* pertence à família *Enterobacteriaceae* e é formado por bactérias Gram-negativas, anaeróbias facultativas, em forma de bastonete e não formadoras de esporos.¹⁶⁴ De acordo com Mezzatesta *et al.*,¹⁶⁵ *E. cloacae* é amplamente encontrada na natureza, mas tem importância clínica como bactéria oportunista com efeitos significativos em pacientes de terapia intensiva, especialmente por causa de sua resistência a antibióticos. *EcXyl43* é a único GH43 de *Enterobacter* que foi expressa e caracterizada até o momento.¹⁶⁵⁻¹⁶⁶

Ao contrário de algumas outras GH43, como WXyn43m por exemplo, que tem maior atividade para xilooligossacarídeos menores (XOS) e atividade decrescente para XOS mais longos (de xilobiose para xilotetraose, nesta ordem),¹⁶⁸ a caracterização bioquímica de *EcXyl43* demonstrou maior atividade para XOS maiores, e um baixo nível de atividade em xilano comercial.¹⁶⁷ Além disso, foi demonstrado experimentalmente utilizando mutagênese sítio-dirigida, que os mutantes D14A, W73G e F507A são completamente desprovidos da atividade enzimática.¹⁶⁷ O resíduo D14 é catalítico na *EcXyl43*, logo sua mutação abolir a atividade enzimática é compreensível, assim como o resíduo W73, que faz parte do sítio ativo, fazendo interações do tipo Van der Waals com os resíduos xilopiranosídicos, e sua mutação também influencia na atividade. Entretanto até o momento não se sabe o motivo da mutação do resíduo F507, que faz parte do domínio auxiliar na *EcXyl43*, causar a perda da atividade enzimática.

A xilose é o segundo carboidrato mais abundante na natureza¹⁶⁹ e, quando é transportada para o citoplasma bacteriano e isomerizada em D-xilulose, entra na via da pentoses-fosfato (PPP), que é a principal via bioquímica para o metabolismo da xilose.¹⁷⁰ A PPP é encontrada em todos os organismos celulares, onde fornece D-ribose para a biossíntese de ácidos nucleicos, D-eritrose 4-fosfato para a síntese de aminoácidos aromáticos e NADPH para reações anabólicas.¹⁷¹ Por este motivo, é fundamental compreender as enzimas envolvidas no metabolismo da xilose para melhor compreender a capacidade de adaptação a diferentes ambientes, utilizando diferentes fontes de carbono.

O objetivo deste trabalho foi caracterizar estruturalmente o mutante F507A da β -xilosidase GH43 de *Enterobacter cloacae* (EcXyl43 F507A), e compreender os determinantes estruturais para aumento de atividade da enzima para XOS mais longo (com o grau de oligomerização até 6). Além disso, o papel funcional de um resíduo não catalítico, Phe507, do domínio auxiliar, foi investigado em detalhes.

5.2 Material e métodos

5.2.1 Clonagem, expressão e purificação

A etapa de clonagem foi realizada pelos colaboradores da Argentina, Dra. Eleonora Campos e Dra. Ornella Ontañón. As etapas de expressão e purificação foram realizadas no laboratório do Grupo de Biotecnologia Molecular, pelo aluno Caio Capetti e supervisão da Dra. Vanessa Pellegrini.

A clonagem de genes e a expressão de proteínas foram previamente descritas em Ontañón *et al.*¹⁶⁷ *Escherichia coli* cepa DH5 α foi usada como hospedeira para transformações durante a construção, propagação e armazenamento de plasmídeo, enquanto *E. coli* Rosetta (DE3) (Novagen, Darmstadt, Alemanha) foi usada para expressão de proteína.

Uma pré-cultura de *E. coli* BL21-pET28a EcXyl43 foi inoculada em meio LB-canamicina (50 μ g/mL) e incubada a 37 °C e 200 rpm em um agitador, durante a noite. Foi utilizada uma proporção de 5 mL de pré-inóculo para 500mL de expressão. A expressão foi realizada a 37 °C e 200 rpm em um agitador por aproximadamente 4 horas até que a densidade celular atingiu uma DO₆₀₀ de 0,6. Em seguida, as células cultivadas foram induzidas com 1 mM de isopropil β -D-tiogalactopiranosídeo (IPTG) a 18 °C sob agitação

contínua em um agitador a 200 rpm durante 16 horas. A cultura foi centrifugada a 4000 x g a 4 °C durante 20 minutos e o pellet celular foi ressuspensão num tampão de lise, pH 8.0 (NaCl 300 mM, NaH₂PO₄ 50 mM e imidazol 10 mM). Em seguida, a lisozima foi adicionada ao extrato, em banho de gelo, de modo que a concentração final fosse de 1mg / mL. Após isso, foram realizados oito ciclos de 10 segundos de sonicação em banho de gelo usando um 550 Sonic Dismembrator Sonifier (Fisher Scientific, Hampton, EUA) com amplitude de 40%. O lisado foi clarificado por centrifugação a 10.000 x g durante 30 minutos a 4 °C para remover resíduos celulares e o sobrenadante foi usado para purificação da *EcXyl43*.

O sobrenadante foi aplicado em uma coluna Ni-NTA Superflow de 5 mL (Qiagen, Hilden, Alemanha), previamente equilibrada com o tampão de lise. Três volumes de tampão de lavagem pH 8.0 (300mM NaCl, 50mM NaH₂PO₄ e 20 mM imidazol) foram usados para lavar a coluna. A proteína foi eluída com dois tampões de eluição pH 8.0, ambos com NaCl 300 mM e NaH₂PO₄ 50 mM, mas com diferentes concentrações de imidazol: o primeiro tampão tinha 250 mM enquanto o segundo tinha 500 mM. Em seguida, as frações de proteína foram reunidas e concentradas em um concentrador de corte molecular de 50 kDa para reduzir a concentração de imidazol para 20 mM. A purificação foi avaliada por análise de eletroforese em gel de dodecil sulfato de sódio 15% (SDS-PAGE) sob condições desnaturantes⁹⁰ e subsequentemente corada com Coomassie Brilliant Blue R-250. A enzima purificada foi posteriormente concentrada usando um concentrador Vivaspín de 10 kDa (GE-Healthcare) a 1.500 x g e sua concentração foi determinada medindo a absorção a 280 nm usando o espectrofotômetro Nanodrop 1000 (Thermo Scientific, Waltham, EUA) e aplicando o coeficiente de extinção calculado ($\epsilon = 150.690 \text{ M}^{-1}\text{cm}^{-1}$). Esta metodologia também foi usada para expressão e purificação de mutantes D14A e F507A.

5.2.2 Cristalização, coleta de dados e processamento

Para os ensaios de cristalização, a *EcXyl43* purificada e sua mutante *EcXyl43* F507A cataliticamente inativa foram concentradas a 10 mg/mL. Os ensaios de cristalização foram realizados usando um robô de cristalização Honeybee (DIGILAB, Massachusetts, EUA) e os cristais foram formados e crescidos por técnica de difusão de vapor em gotas com diferentes proporções de proteína e solução, usando o kit Crystal Screen (Hampton Research, Califórnia, EUA). Os cristais de *EcXyl43* F507A foram detectados sob condições de cristalização consistindo em formato de sódio 4 M, pH 8,0 a 20 °C, com 2 µL de proteína misturada com 1

μL de tampão na gota de cristalização. Apesar dos esforços contínuos, não fomos capazes de cristalizar a *EcXyl43* nativa recombinante.

Os cristais de *EcXyl43* F507A resultantes foram congelados rapidamente em nitrogênio líquido após imersão em glicerol (30% v/v). Os dados de difração de raios-X foram coletados na linha de luz W01B-MX2 do Laboratório Nacional de Luz Síncrotron (Campinas, Brasil)⁹³, usando um detector Pilatus2M (Dectris, Filadélfia, EUA), com um tempo de exposição de 2 s por imagem e um ângulo de rotação do cristal de 0,2 °. Os dados da imagem de difração foram indexados e integrados utilizando o pacote XDS⁹⁴ e escalonados com AIMLESS.⁹⁵

5.2.3 Solução da estrutura cristalográfica por difração de raios-X

A β -xilosidase GH43 de *Geobacillus stearothermophilus* (*GsXyl43*,¹⁷² código PDB 2EXH), com 56,4% de identidade de sequência de aminoácidos e uma cobertura de 99%, foi identificada como o modelo de estrutura de homologia mais próxima de acordo com BLASTP.⁹⁶

A substituição molecular foi realizada por Phaser⁹⁷ usando as coordenadas atômicas de *GsXyl43* como um modelo de busca. A estrutura resultante foi então submetida ao Autobuild⁹⁸ e os ciclos de refinamento foram realizados usando os programas Coot⁹⁹ e phenix.refine.¹⁰⁰ A qualidade do modelo foi avaliada usando Molprobity,¹⁰¹ e os locais de ligação do íon cálcio foram confirmados usando CheckMyMetal.¹⁷²⁻¹⁷³ As coordenadas atômicas foram depositadas no Protein Data Bank (PDB) sob o código de acesso 7K1R.

5.2.4 Análise estrutural

5.2.4.1 Alinhamento estrutural

A sequência *EcXyl43* foi submetida ao BLASTP, usando o Protein Data Bank (PDB)¹⁷⁵ como banco de dados. O principal resultado foi a GH43 de *Geobacillus stearothermophilus* (*GsXyl*), com valor de identidade de 56,45% e cobertura de 99%. Para estudos comparativos estruturais, a estrutura cristalográfica *EcXyl43* F507A foi submetida ao

DALI,¹⁷⁶ para comparar com outras estruturas 3D na base de dados do PDB. As 3 principais enzimas com atividade de β -xilosidase coincidiram com os principais resultados na pesquisa BLASTP e foram selecionadas para análise posterior. Três sequências de aminoácidos de GH43 de *Geobacillus stearothermophilus*, *Bacillus pumilus* e *Selenomonas ruminantium* (códigos de acesso PDB 2EXH_A, 5ZQX e 3C2U, respectivamente)^{171,176-177} foram alinhados contra *EcXyl43* usando o programa Clustal Omega¹⁰³ e, em seguida, submetidos ao ESPRIPT.¹⁰⁴

Para fins de comparação estrutural, a estrutura cristalográfica de *EcXyl43* F507A foi sobreposta com β -xilosidases GH43 de *G. stearothermophilus* (*GsXyl*, código PDB 2EXH) e *GsXyl* complexado com xilobiose (código PDB 2EXJ). O programa Chimera¹⁰⁷ foi usado para comparações estruturais. Para estudos do modelo *EcXyl43* nativo, o programa Coot foi usado para reverter manualmente a mutação F507A.

5.2.4.2 Docking Molecular

A fim de entender as interações moleculares entre *EcXyl43* F507A e xilo-oligossacarídeos, uma molécula de xilohexaose (X6) foi encaixada em *EcXyl43* usando o programa LIBELA.¹⁷⁹ As coordenadas para os substratos foram geradas usando o servidor web GLYCAM¹⁸⁰ e o campo de força GLYCAM¹³⁴ foi usado para gerar arquivos SYBYL MOL2 usando o programa TLEAP disponível no AMBERTOOLS e a estrutura cristalográfica da enzima foi utilizada para docking. Para este propósito, átomos de hidrogênio foram adicionados na ferramenta UCSF Chimera DockPrep¹⁰⁷ e cargas atômicas foram atribuídas usando campo de força AMBER FF14SB.¹³³ Por fim, o encaixe do substrato foi realizado utilizando os parâmetros padrão do LIBELA. Essa etapa foi realizada pelo Professor Alessandro Nascimento.

5.2.5 Simulações de Dinâmica Molecular

As simulações de dinâmica molecular do *EcXyl43* foram realizadas com a supervisão do Professor Alessandro Nascimento, utilizando o cluster Nascimento e o programa AMBER16.¹²⁹ A estrutura cristalográfica de *EcXyl43* F507A e *EcXyl43* nativa modelada não-mutada sem moléculas de água foram usadas como modelos iniciais para simulações de DM. Os complexos de *EcXyl43* e *EcXyl43* F507A ligados a X6 foram criados

da mesma maneira. Os estados de protonação em pH 6,5 e na presença de NaCl 150 mM foram estimados usando servidor H++.¹³² As estruturas apo e ligadas a xilohexaose foram solvatadas em caixas octaédricas de modo que a camada de solvente tivesse pelo menos 12 Å de espessura. Os íons sódio e cloreto foram adicionados para atingir a neutralidade e garantir uma concentração de NaCl 150 mM dentro da caixa de simulação. As simulações foram configuradas usando o campo de força AMBERFF 19SB,¹⁸¹ modelo de água TIP3P⁸⁶ e campo de força de carboidrato GLYCAM06¹³⁴ para descrever os xilosídeos. Cada sistema foi minimizado em 20.000 etapas e, em seguida, aquecido a 300 K em 50 ps na montagem NVT. Posteriormente, a densidade do sistema foi equilibrada em 50 ps na montagem NPT. Restrições harmônicas foram mantidas para minimização, aquecimento e equilíbrio de densidade com pesos harmônicos de 10,0, 10,0 e 5,0 kcal/(mol*Å²), respectivamente. Finalmente, o sistema foi equilibrado por 2 ns sem restrições. Simulações produtivas foram executadas por 500 ns no conjunto NPT. Um intervalo de tempo de 2fs foi usado e as ligações covalentes envolvendo átomos de hidrogênio foram restringidas usando o algoritmo SHAKE. As trajetórias de MD foram analisadas usando VMD,¹³⁵ MDLovofit,¹⁵² MDanalysis¹⁸¹⁻¹⁸² e códigos computacionais desenvolvidos internamente.

5.3 Resultados e discussão

5.3.1 Clonagem, expressão e purificação

EcXyl43 foi expressa e purificada usando cromatografia de afinidade ao Ni²⁺. A massa teórica da enzima prevista com base em sua sequência de aminoácidos é de 64,4 kDa e a proteína purificada migrou como uma única banda em SDS PAGE, garantindo assim sua pureza. A enzima purificada foi utilizada para ensaios de cristalização. Como mencionado anteriormente, apenas a *EcXyl43* F507A formou cristais, e estes seguiram para as etapas posteriores.

5.3.2 Determinação e análise da estrutura cristalográfica *EcXyl43* F507A

5.3.2.1 Coleta de dados cristalográficos, refinamento, construção de modelo e comparação com estruturas homólogas

Os cristais de *EcXyl43* F507A apareceram em 2 dias e cresceram até as dimensões máximas em 10 dias. Os dados de difração cristalográfica foram coletados no Laboratório Nacional de Síncrotron (LNLS), usando uma rotação angular de 0,2 graus por imagem, resultando em 1.800 imagens para uma coleta de dados de 360°. O processamento dos dados com XDS revelou que o cristal de *EcXyl43* pertencia ao grupo espacial $P2_12_12_1$.

A substituição molecular foi realizada com o programa Phaser usando a estrutura da β -xilosidase GH43 de *G. stearothermophilus* (PDB id 2EXH) como modelo, resultando em uma única solução com um LLG de 4032,656 e um TFZ de 50,7. A estrutura foi refinada utilizando phenix.refine, e submetida ao Autobuild, a fim de colocar as cadeias laterais faltantes. Isso levou a uma queda tanto no R_{free} quanto no R_{work} , e mais ciclos de construção e refinamento de modelos foram realizados usando Coot e PHENIX. A estrutura final foi avaliada com MolProbity, e ambos os parâmetros de coleta e refinamento são fornecidos na Tabela 8.

Tabela 8 – Coleta de dados cristalográficos de *EcXyl43* F507A e estatísticas do refinamento do modelo

Conjunto de dados	<i>EcXyl43</i> F507A
Coleta de dados	LNLS MX2
Comprimento de onda (Å)	1,458
Grupo espacial	$P 2_1 2_1 2_1$
Parâmetros de cela	
a, b, c (Å)	87,12 153,10 182,40
Faixa de resolução (Å)	22,71 – 2,4 (2,486 – 2,4)
R_{pim}	0,07546 (0,6614)
$CC_{1/2}$	0,994 (0,437)
Reflexões totais	1.253.638 (118.376)
Reflexões parciais	95.858 (9.404)
Multiplicidade	13,1 (12,5)
Completeza (%)	99,74 (99,11)
$\langle I/\sigma_I \rangle$	8,73 (1,13)
	(continua)

(continuação)

Refinamento

No. mol. unidade assimétrica	4
Resíduos por molécula	536
R_{work}	0,1924 (0,2977)
R_{free}	0,2246 (0,3136)
CC geral	0,998 (0,78)
Ligantes	GOL, CA
<i>B-factors</i> ($\text{\AA}^2$)	
Médio	42,10
Macromoléculas	42,23
Ligantes	43,09
<i>R.m.s. deviations</i>	
Comprimentos de ligação ($\text{\AA}$)	0,003
Ângulos de ligação ($^\circ$)	0,65
<i>Ramachandran</i>	
Ramachandran favoráveis (%)	95,47
Ramachandran permitidos (%)	4,11
Ramachandran outliers (%)	0,42
Rotâmeros outliers (%)	0
<i>MolProbity Overall Score</i>	1,48

Os valores em parênteses são referentes à camada de mais alta resolução.

Fonte: Elaborada pelo autor

EcXyl43 F507A é um tetrâmero (dímero de dímeros) na unidade assimétrica (Figura 1a). Cada protômero, ou cadeia, consiste em dois domínios: o domínio catalítico, formado por uma beta-propeller de 5 pás típica para os membros da família GH43, e um domínio auxiliar

adicional de sanduíche beta de 6 dobras (Figura 1b). Uma alça contendo os resíduos 502 a 510, que pertence ao domínio auxiliar, fecha o túnel catalítico (Figura 1c).

Figura 28 - Estrutura cristalográfica. (A) Estrutura 3D tetramérica de *EcXyl43* F507A, com as cadeias A, B, C e D representadas em vermelho, verde, amarelo e azul, respectivamente, e as esferas verdes indicando o íon Cálcio. A figura inferior indica uma rotação de 180 graus ao longo do eixo y. (B) Monômero de *EcXyl43* F507A, mostrando o domínio catalítico da β -propeller em ciano e o domínio auxiliar sanduíche- β em azul escuro. Uma molécula de glicerol foi observada no sítio ativo, densidade eletrônica representada pelo mapa $2mFo - DFc$ Omit com contorno de $1,5 \sigma$. (C) Mesmo monômero observado em (b), em representação de superfície, com o túnel catalítico indicado pelo círculo branco tracejado.

Fonte: Elaborada pelo autor.

A unidade tetramérica de *EcXyl43* F507A é composta por dois dímeros que são posicionados a 90° um do outro, com simetria pontual D2 (Figura 1A). Um tetrâmero antiparalelo com estrutura quaternária semelhante foi previamente observado para a β -xilosidase *GsXyl* de *Geobacillus stearothermophilus*.¹⁷² Phe507, um resíduo essencial para a atividade catalítica, está localizado em uma alça do domínio auxiliar (azul escuro nas Figuras 1b e 1c). A mutação F507A extingue a atividade da enzima.¹⁶⁷

Uma comparação dos monômeros na unidade assimétrica revelou que duas subunidades possuem moléculas de glicerol ligadas ao sítio ativo, enquanto as outras duas cadeias de proteína não. Sobrepondo cada subunidade com suas contrapartes com PyMol e analisando os valores de RMSD, descobrimos que eles são apenas ligeiramente diferentes um do outro, com valores de RMSD variando de 0,151 Å a 0,201 Å.

Para uma investigação mais aprofundada da estrutura quaternária, a estrutura cristalográfica de *EcXyl43* F507A foi submetida ao servidor PISA [59]. Os resultados são apresentados na Tabela 9.

Tabela 9 - Resultados do PISA da interface de *EcXyl43* F507A e análises de conjuntos. As áreas de superfície (A.S.), enterrada (A.E.) e de interface (A.I.) são representadas em Å²; NRes: Número de resíduos; NHB: Número de potenciais ligações de hidrogênio; NSB: Número de pontes salinas.

Conjunto	A.S. (Å ²)	A.E. (Å ²)	A.I. (Å ²)	NRes	NHB	NSB
ABCD	70160	11200	-	-	-	-
AD	36710	4170	1853.4	53	32	17
BC	36310	4170	1853.6	56	33	19
AB	38700	1890	709.6	20	2	0
CD	38880	1900	718.7	20	2	0

Fonte: Elaborada pelo autor.

As áreas de interface calculadas com o servidor PISA sugerem que *EcXyl43* F507A forma um dímero de dímeros. As interfaces AD e BC são muito maiores do que as interfaces AB e CD. Além disso, o número de resíduos envolvidos na interface (53 e 56, para AD e BC, respectivamente) é maior, e essas interfaces têm muito mais ligações de hidrogênio e pontes salinas.

Para comparar a estrutura de *EcXyl43* com outras enzimas do grupo GH43, a sequência proteica foi submetida a BLASTp e a estrutura cristalográfica de *EcXyl43* F507A ao DALI, resultando em três estruturas de enzimas homólogas (*GsXyl*, *SrXyl* e *BpXyl*; Tabela 10).

Tabela 10 - Parâmetros de alinhamento para as enzimas selecionadas de BLASTp e DALI. Valores obtidos em comparação com *EcXyl43*.

Enzima	PDB id	Identidade (%)	Cobertura (%)	R.M.S.D. (Å)
<i>GsXyl</i>	2EXH	56.06	99	1.2
<i>SrXyl</i>	3C2U	53.99	100	1.2
<i>BpXyl</i>	5ZQX	51.95	99	1.1

Fonte: Elaborada pelo autor.

As sequências das enzimas selecionadas foram alinhadas a *EcXyl43* utilizando o Clustal Omega e um alinhamento de sequência baseado na estrutura foi realizado com ESPript (Figura 29A). A superposição estrutural mostra uma notável conservação da estrutura geral com maiores diferenças observadas na alça envolvendo os resíduos 145-155 (numeração da *EcXyl43*) e também na região envolvendo os resíduos 502-510. Curiosamente, essas duas regiões abrangem o bolso da região do sítio ativo e podem estar relacionadas com rearranjos estruturais no sítio ativo da enzima na presença/ausência de um substrato.

Figura 29 - Alinhamento de sequência baseado em estrutura. (A) As sequências de *GsXyl*, *SrXyl* e *BpXyl* foram alinhadas com a sequência *EcXyl143*, e a alça, domínio auxiliar, resíduos catalíticos e resíduos de sítio ativo são destacados. Os resíduos do sítio de ligação do cálcio são destacados em triângulos verdes. (B) Alinhamento estrutural das enzimas acima mencionadas. *EcXyl143* é representado em ciano, *GsXyl* em verde, *SrXyl* em marrom e *BpXyl* em amarelo. O loop contendo a mutação F507A é destacado em um círculo preto tracejado. (C) Ampliação da região do sítio ativo, mostrando as distâncias entre F506 e E186, e A507 e E186, e as cadeias laterais de resíduos de aminoácidos nas posições correspondentes das enzimas sobrepostas.

Fonte: Elaborada pelo autor

A análise da Figura 2 indicou que, apesar da conservação entre o loop 502-510, uma mutação como F507A pode promover rearranjos espaciais, aproximando o loop do túnel catalítico e, assim, tornando-o mais estreito. A comparação das distâncias é mostrada na Tabela 11.

Tabela 11 - Distâncias interatômicas entre diferentes resíduos em *EcXyl43*, *GsXyl*, *SrXyl* e *BpXyl*.

Enzima	Resíduo 1	Átomo 1	Resíduo 2	Átomo 2	Distância (Å)
<i>EcXyl43</i>	F506	CZ	E186	CD	6.5
	A507	CB		CD	12.3
<i>GsXyl</i>	A505	CB	E187	CD	13.8
	F506	CZ		CD	7.8
<i>SrXyl</i>	F507	CZ	E186	CD	9.8
	F508	CZ		CD	8.2
<i>BpXyl</i>	F503	CZ	Q186	CD	9.6
	F504	CZ		CD	8.6

Fonte: Elaborada pelo autor

Com base na Tabela 11, pode-se inferir que uma distância típica entre F507 e o E186 catalítico seria em torno de 8 Å (7,8 Å para F506 em *GsXyl*, 8,2 Å para F508 em *SrXyl* e 8,6 Å para F504 em *BpXyl*). Em contraste, *EcXyl43* F507A mostra uma distância de 12,3 Å entre A507 e a cadeia lateral do resíduo catalítico E186. Além disso, o resíduo anterior (F506) aproxima-se do E186, em direção ao túnel catalítico, estreitando-o, a uma distância de apenas 6,5 Å. Em comparação com outros membros GH43, os resíduos anteriores (A505 em *GsXyl*, F507 em *SrXyl* e F503 em *BpXyl*) permanecem a uma distância maior do que o resíduo Phe seguinte correspondente.

5.3.4 Análise do sítio de ligação de cálcio

A estrutura cristalográfica de *EcXyl43* F507A indicou um sítio de ligação de cálcio, localizado no domínio auxiliar. Para posteriores investigações, a estrutura foi submetida à ferramenta CheckMyMetal (CMM). Íons divalentes, como o cálcio, são conhecidos por terem função estrutural¹⁸³⁻¹⁸⁵ e seu sítio de ligação exibe uma geometria octaédrica. Em alguns casos, como em *GsXyl*, um íon de cálcio é coordenado por três resíduos e duas moléculas de água, exibindo a mesma geometria octaédrica. O sítio de ligação de cálcio de *EcXyl43* é mostrado na Figura 3.

Figura 30 - Sítio de ligação do cálcio em *EcXyl43* F507A. O íon cálcio é representado como uma esfera verde. Os resíduos envolvidos na interação cálcio-enzima são destacados e as distâncias são mostradas em Å.

Fonte: Elaborada pelo autor.

5.3.5 Comparações Estruturais

EcXyl43 apresenta o mecanismo enzimático catalítico do tipo inversão¹⁸⁷ com Glu186 sendo um ácido geral e Asp14 uma base geral. No mecanismo de inversão, a base geral coordena uma molécula de água que ataca o átomo glicônio C1', enquanto o ácido geral perde seu hidrogênio ionizável, protonando o grupo de saída no átomo de oxigênio glicosídico e, ao final do processo, um monômero de açúcar é lançado como um produto.^{146,187} Portanto, a análise das interações entre os resíduos no túnel do sítio ativo e o substrato torna-se muito relevante e para tanto, foi realizada a sobreposição estrutural de *GsXyl* (código PDB 2EXJ) e

EcXyl43 F507A utilizando o programa Chimera, revelando as distâncias entre os resíduos e o substrato (Figura 31) e permitindo inferir possíveis interações entre enzima e substrato, bem como interações entre os resíduos de *EcXyl43* F507A.

Figura 31 - Superposição de *EcXyl43* F507A e *GsXyl*. (A) Sobreposição de *EcXyl43* (azul) e *GsXyl* (verde), com Xilobiose de *GsXyl* no sítio ativo (quadrado tracejado). As esferas verdes representam o Cálcio em ambas as enzimas. (B) Zoom do sítio ativo (quadrado tracejado em (A)), com os resíduos (verde para *GsXyl* e azul para *EcXyl43*) e suas distâncias aos átomos do ligante.

Fonte: Elaborada pelo autor.

O ácido geral, Glu186 em *EcXyl43* F507A, forma duas interações do tipo ligação de hidrogênio com o substrato. Primeiramente, como um resíduo catalítico, está 3,0 Å de distância do átomo de oxigênio na ligação glicosídica entre os subsítios -1 e 1. Nessa distância, a interação é possível e é crucial para a atividade da enzima. Além disso, Glu186 está a uma distância de 2,9 Å do O2 do subsítio +1, estabilizando o substrato neste subsítio.

Dada a importância do resíduo Asp14, diretamente envolvido no mecanismo de catálise como base geral, seria de se esperar que a mutação desse resíduo causasse perda de atividade enzimática. De fato, o mutante D14A não possui atividade enzimática contra os substratos da enzima nativa,¹⁶⁷ uma vez que como base geral ativa uma molécula de água para um ataque de deslocamento único no carbono anomérico, resultando na inversão da configuração anomérica, essencial para a clivagem.¹⁷² A média das quatro distâncias possíveis entre quatro átomos de oxigênio de Glu186 e de Asp14 foi encontrada em 8,0 Å, o que é consistente com a distância normalmente encontrada para enzimas de inversão (na faixa de 8,0 Å -10,0 Å).¹⁸⁷ A distância observada entre Asp14 (OD1) e o carbono anomérico foi de 5,2 Å, típica para seu papel de base geral. Além disso, Asp14 está 2,7 Å distante do O4 do resíduo xilosídico do subsítio -1. Além das interações feitas por Arg289 (2,5 Å de O4) e

Asp127 (3,4 Å de O2) com o resíduo xilosídico no subsítio -1, Asp14 também está envolvido em importantes interações com o substrato.

Outro resíduo chave para a catálise é Asp127, descrito como um modulador pKa em membros da família GH43. Este resíduo também ajuda a manter a posição relativa e a orientação de Glu186 em relação ao substrato.¹⁸⁹ Em pH ótimo (6,5) o ácido geral (Glu186) tenderia a ser desprotonado, devido ao seu pKa que é 4,25 para a cadeia lateral. O Asp127, posicionado a uma distância de 4,4 Å do ácido geral, aumenta o pKa de Glu186, de forma que este é protonado mesmo em pH mais alto e pode, então, fazer parte da reação. Estudos cinéticos e bioquímicos mostram que em *GsXyl*, o pKa do ácido geral aumenta para 7,1 na presença do modulador de pKa em discussão.¹⁵⁸

5.3.6 Docking molecular de X6 e o sítio ativo de *EcXyl43*

A fim de entender melhor a afinidade mais alta de *EcXyl43* para XOS mais longo (com um grau de polimerização de até 6), foi realizado o docking molecular de X6 (Figura 32).

Figura 32 - Docking molecular de X6 no sítio ativo de *EcXyl43*. (A) Vista lateral de *EcXyl43* com X6 em seu site ativo (quadrado tracejado). (B) Zoom do sítio ativo, mostrando distâncias de interação entre enzima e substrato.

Fonte: Elaborada pelo autor.

Como poderia ser esperado, pode-se ver claramente que as interações entre o substrato e a enzima nos subsítios -1 e +1 são exatamente as mesmas observadas para X2. Além disso, interações adicionais são formadas no subsítio +2, incluindo ligações de hidrogênio com Lys99 e His150. A análise de nossos resultados de docking sugere interações

mais fortes de *EcXyl43* com o substrato xilopiranosídeos nas posições -1, +1 e +2. Para resíduos de açúcar nas posições +3, +4 e +5, nenhuma interação foi observada no modelo de docking inicial.

Dois outros resíduos altamente conservados entre os membros da família GH43 são Trp73 e Phe507. Trp73 é orientado paralelamente ao primeiro resíduo de xilose (subsítio -1) e estabiliza esta unidade de xilose por meio de interações do tipo van der Waals. É parte de uma sequência conservada 73WAP75 que forma um bolso para guiar os substratos para se ligarem na proximidade dos resíduos catalíticos.¹⁹⁰ Em arabinanase GH43 de *Bacillus subtilis*, a mutação de Trp74 em alanina afetou negativamente a atividade catalítica,¹⁹¹ e para *EcXyl43*, a mutação W73G aboliu a atividade catalítica.¹⁶⁷ Ao mesmo tempo, o resíduo Phe507 não é um resíduo catalítico e nem pertence ao domínio catalítico, ele está localizado na alça 502-510 do domínio auxiliar que se estende para dentro do domínio catalítico beta-propeller e envolve o túnel do local ativo. Phe507 desempenha um papel importante no sítio catalítico, uma vez que sua mutação F503Y em BpXynB demonstrou que a ligação mais forte do segundo (ou aglicona) açúcar é importante na promoção da reação reversa da hidrolase.¹⁷⁷ Além disso, a mutação de Phe507 para alanina em *EcXyl43* anulou completamente sua atividade enzimática.¹⁶⁷

5.3.7 Simulação de Dinâmica Molecular de *EcXyl43*

Por que *EcXyl43* tem maior atividade enzimática em XOS mais longos (como X6, por exemplo) em comparação com os mais curtos? De acordo com os resultados de docking molecular com X6, não há interações entre os resíduos xilosídicos 4, 5 e 6 e a enzima, havendo, teoricamente, apenas interação desta com os resíduos 1, 2 e 3. Isto trouxe a necessidade de buscar ferramentas que expliquem essa característica enzimática de *EcXyl43*. Haveria a possibilidade de serem criadas interações adicionais ao longo do tempo entre a enzima e o substrato? Já haviam sido observadas em diversas outras simulações de DM dobras e torções em XOS ao longo do tempo, demonstrando flexibilidade do ligante. Para entender melhor as interações de *EcXyl43* com ligantes, foram realizadas simulações de Dinâmica Molecular (MD) com X6 como um ligante acoplado à enzima (Figura 33).

Figura 33 - Dinâmica molecular de *EcXyl43* com X6 ancorado no sítio ativo. (A) Posição inicial de *EcXyl43* e X6 no sítio ativo, perpendicular à enzima (vista frontal). (B) Zoom de (A), mostrando a distância do último resíduo xilosídico (X6) a H189 (23,8 Å). (C) O mesmo que (A), com rotação de 90 graus ao longo do eixo y (vista lateral). (D) Zoom de (C), mostrando a distância de X6 a H189. (E) Posição intermediária de *EcXyl43* com X6 no sítio ativo, mostrando o ligante curvado em direção à alça. (F) Zoom de (E), mostrando a distância do último resíduo xilosídico (X6) a H189 (3,4 Å). (G) Vista lateral de (E), com rotação de 90 graus ao longo do eixo y. (H) Zoom de (G). (I) Gráfico da distância entre o último resíduo xilosídico (X6) e H189 ao longo da simulação (500 ns). Os arquivos PDB foram produzidos com VMD e as figuras foram produzidas com Chimera.

Fonte: Elaborada pelo autor.

Simulações de dinâmica molecular de *EcXyl43* com o X6 ligado à enzima (Figura 33) mostram que o ligante de fato se mantém bastante estabilizado no bolso catalítico, em particular os resíduos 1, 2 e em menor grau, 3 do xilooligossacarídeo, enquanto os demais resíduos (4, 5 e 6) apresentam grande mobilidade, até que em determinado momento o ligante se dobra sobre a superfície da enzima por um período de tempo considerável (cerca de 60% do tempo de simulação), fazendo interações adicionais com a enzima. Isso forneceria uma possível explicação de por que *EcXyl43* mostrou maior atividade em XOS mais longo, até um grau de oligomerização de 6 (X6).

5.3.8 Por que o mutante F507A anula a atividade enzimática *EcXyl43*?

A análise estrutural mostrou que a entrada do túnel catalítico é significativamente mais estreita no mutante *EcXyl43* F507A em comparação com a β -xilosidase GH43 de *G. stearothermophilus* (PDB id: 2EXH), potencialmente interferindo no acesso do substrato ao sítio ativo da enzima (Figura 34).

Figura 34 - Comparação estrutural de fendas catalíticas. (A) *GsXyl* (PDB id 2EXJ) em verde claro (domínio catalítico) e verde escuro (domínio auxiliar), com Xilobiose no sítio ativo. (B) Zoom do sítio ativo, mostrando o túnel com o ligante dentro. (C) *EcXyl43* em azul claro (domínio catalítico) e azul escuro (domínio auxiliar), com glicerol no sítio ativo. (D) Zoom do sítio ativo, mostrando o túnel com o ligante dentro. (E) *SrXyl* (PDB id 3C2U) em rubi (domínio catalítico) e vermelho (domínio auxiliar), com BXP no sítio ativo. (F) Zoom do sítio ativo, mostrando o túnel com o ligante dentro. (G) *BpXyl* (PDB id 5ZQX) em amarelo (domínio catalítico) e dourado (domínio auxiliar), com X2 no sítio ativo. (H) Zoom do sítio ativo, mostrando o túnel com o ligante dentro. (I) Sobreposição de *GsXyl* (verde), *EcXyl43* (azul claro), *SrXyl* (vermelho) e *BpXyl* (amarelo) com xilobiose no sítio ativo.

te: Elaborada pelo autor.

Na verdade, o túnel catalítico de *EcXyl43* é mais estreito não apenas em comparação com *GsXyl*, mas também com *SrXyl* e *BpXyl*, como pode ser observado nas figuras 34 B, D, E e H. O loop onde F507 está localizado, tem uma sequência característica ⁵⁰⁵GFF⁵⁰⁷ (Gly505 - Phe506 - Phe 507) em *EcXyl43* e se Phe507 é mutado, todo o loop se desloca em direção ao sítio catalítico. O F506 anterior move-se em direção aos resíduos do sítio ativo, estreitando o túnel catalítico e interferindo na posição de ligação do ligante (Figura 34I). Isso explica porque a mutação F507A anula a atividade enzimática em *EcXyl43* F507A.

Para comparação detalhada entre os diferentes túneis de membros da família GH43 com atividade β -xilosidase, as estruturas de *EcXyl43* e *GsXyl43* foram submetidas ao programa PDBsum.¹³⁸ Os resultados foram analisados e podem ser vistos na Tabela 12.

Tabela 12 - Análise do túnel de GH43s com atividade de β -xilosidase. R.L.: Raio livre do gargalo do túnel (Å) - sem cadeia lateral; C.: Comprimento do túnel (Å); Hpa: Índice de hidropatia (-4,5 hidrofílico a +4,5 hidrofóbico); Hph: Escala de hidrofobicidade (-1,140 hidrofílico a +1,810 hidrofóbico); Pos: resíduos positivos; Neg: resíduos negativos; Aro: resíduos aromáticos; Lig: ligante.

Enzima	R.L.	C.	HPa	HPh	Pos	Neg	Aro	Lig
<i>EcGH43 F507A</i>	1.88	18.7	-1.34	0.41	5	2	5	GOL
<i>GsXyl</i>	3.43	16.8	-2.46	-0.34	5	2	2	
<i>BpXyl</i>	2.54	15.9	0.15	0.45	2	2	5	XYP

Fonte: Elaborada pelo autor

Os resultados do PDBSum indicam que *EcXyl43 F507A* de fato tem um túnel consideravelmente mais estreito do que *GsXyl*, comparando o raio livre do gargalo de túnel (R.L.), que têm valores de 1,88 Å e 3,43 Å, respectivamente. *BpXyl* também tem um gargalo de túnel mais largo do que *EcXyl43* (2,54 Å), indicando que mesmo com a presença do ligante (xilobiose para *BpXyl*), o túnel formado por resíduos de fenilalanina em posições equivalentes a Phe 507 na estrutura de *EcXyl43* (Phe 506 em *GsXyl* e Phe 504 em *BpXyl*) é mais largo do que o túnel de *EcXyl43* com mutação F507A. Isso dá indícios sobre os determinantes estruturais da falta de atividade enzimática para o mutante *EcXyl43 F507A*.

5.3.9 Simulações de dinâmica molecular para mutante F507A

A fim de explorar ainda mais o papel dos resíduos Phe507 em *EcXyl43* e Ala507 no mutante F507A e seus papéis na estrutura do túnel do sítio ativo, três sistemas moleculares foram gerados e analisados: *EcXyl43* na forma apo, o mutante *EcXyl43 F507A* na forma apo e *EcXyl43* com X6 docado no sítio ativo. O RMSD e o RMSF no processo de simulações de DM conduzidas são apresentados na Figura 35.

Figura 35 - Trajetórias de *EcXyl43* ao longo da Dinâmica Molecular. Representação de cluster de (A) *EcXyl43* com X6 ancorado no sítio ativo, (B) *EcXyl43* na forma Apo e (C) *EcXyl43* F507A na forma Apo ao longo de 500 ns de simulação de DM com intervalo de 1 ns, com estruturas coloridas por RMSF de azul (mínimo) a vermelho ($> 3\text{\AA}$). As setas indicam o loop 502-510, contendo Phe507. (D) RMSD dos sistemas ao longo da Dinâmica. (E) RMSF dos sistemas ao longo da Dinâmica, por resíduo. (F) RMSF dos sistemas, comparando a região de loop dos resíduos 502 - 510 (onde a mutação F507A está localizada).

Fonte: Elaborada pelo autor.

Curiosamente, a presença de um substrato no sítio ativo não resulta em uma redução do RMSD, em comparação com a estrutura inicial (experimental) (Figura 35d). Isso pode ser devido ao fato de o sítio ativo da β -xilosidase *EcXyl43* ser um bolso, com volume suficiente para acomodar apenas 3 unidades xilopiranosídicas. Ao contrário, em β -xilanases das famílias GH10 e GH11, por exemplo, a fenda catalítica é mais ampla e mais longa, com muito mais subsítios e, conseqüentemente, mais interações entre enzima e substrato.

Na direção oposta, uma análise de RMSF (flutuação da raiz quadrada média) (Figura 35e) indica que alguns resíduos têm maior flutuação para a estrutura *EcXyl43* ligada ao X6, em comparação com a enzima livre. Estes resultados, mostrados nas Figuras 35a a 35c, sugerem que as alças de *EcXyl43*, em particular aquelas em contato com o solvente, possuem alta mobilidade. Uma vez que o loop 502 - 510 contendo F507 é de interesse particular para

nosso estudo, o RMSF da região foi analisado em detalhes (Figura 35f). A análise demonstra que tanto a enzima nativa na forma apo como na forma ligada a X6 apresentam RMSF mais alto nesta região, sugerindo que o loop se move muito mais para acomodar o substrato. Pode-se notar que a diferença de RMSF entre as formas apo mutante F507A e a apo nativa de *EcXyl43* é de cerca de 3Å em alguns resíduos do loop, o que pode sugerir que a mutação F507A pode diminuir essa mobilidade de alça. Este, inclusive, pode ser um dos motivos pelos quais não foi possível obter cristais de *EcXyl43* sem a mutação F507A, uma vez que essa mobilidade da enzima nativa pode prejudicar o processo de cristalização.

Após a análise RMSF, pares de distâncias foram analisados ao longo das trajetórias de MD da *EcXyl43*. Para monitorar o tamanho do túnel, duas distâncias foram medidas em dois sistemas: F507 - E186 no sistema *EcXyl43* nativo e A507 - E186 no sistema *EcXyl43* mutante F507A (Figura 36).

Figura 36 - Distâncias dos resíduos F507 e A507 a E186 ao longo da Dinâmica Molecular. (A) Superposição de *EcXyl43* nativo (azul claro), F507A (ciano escuro) e *GsXyl* (verde). (B) Zoom do sítio ativo, mostrando as posições dos resíduos nas enzimas. (C) Gráfico da distância entre os resíduos F507 e E186 em *EcXyl43* e A507 e E186 em *EcXyl43* F507A ao longo da DM.

Fonte: Elaborada pelo autor.

O monitoramento de distância ao longo das trajetórias de DM revela que o resíduo A507 na estrutura mutante está muito mais próximo do E186 com uma distância média de 16,1 Å ao longo de toda a simulação (variação entre 14,1 Å a 18,1 Å). Por outro lado, o monitoramento da distância ao longo das trajetórias de MD revela que o resíduo F507 na estrutura nativa possui uma distância maior para E186, variando de 16,1 Å a 23,4 Å e com uma distância média de 19,5 Å. Estes dados corroboram com o que foi observado e inferido através das análises da estrutura cristalográfica, ou seja, um túnel mais estreito para o mutante F507A através de uma aproximação desse loop em direção ao sítio ativo.

5.4 Conclusões

O presente trabalho descreveu uma análise da estrutura cristalográfica de *EcXyn43* e suas simulações de Dinâmica Molecular, visando desvendar os determinantes estruturais da maior atividade enzimática de *EcXyn43* em substratos mais longos, como X6 por exemplo. Nesse caso, foi observado que ocorre uma espécie de dobra no ligante, aumentando a interação enzima-substrato, através da diminuição de distâncias entre a *EcXyl43* e o ligante. Além disso, este trabalho também fornece explicações para a falta de atividades enzimáticas do mutante F507A de *EcXyn43*. Neste caso, essa mutação gera uma aproximação do loop, estreitando a entrada do sítio catalítico. Esse tipo de informação pode fornecer bases para engenharia de proteínas visando mutações não-catalíticas.

6.1 Introdução

As Xilose Isomerases (XI) (EC 5.3.1.5), também conhecidas como Glicose isomerases, são enzimas que catalisam a interconversão de uma aldose (xilose) em uma cetose (xilulose). As XIs foram descritas pela primeira vez em 1953¹⁹² e têm sido extensivamente estudadas desde então. Essas enzimas são as únicas isomerases comercialmente produzidas devido ao seu papel fundamental na indústria de alimentos para a produção de xarope de milho rico em frutose (do inglês *High Fructose Corn Syrup*, ou HFCS) a partir do amido, por meio da conversão de D-glicose em D-frutose.¹⁹³ No entanto, existe uma ampla gama de substratos usados pelos XIs. Segundo diversos autores,¹⁹³⁻¹⁹⁵ as XIs catalisam a epimerização do Carbono 2 de açúcares de 4 a 6 carbonos, ou seja, tetroses, pentoses e hexoses. Porém, a taxa de reação varia consideravelmente entre essas enzimas.¹⁹⁷

Os parâmetros cinéticos de XIs são dependentes do substrato, com valores de K_M tão baixos quanto 5 mM e tão altos quanto 93 mM para D-xilose e uma faixa muito mais alta de valores de K_M para D-glicose (86 a 920 mM).¹⁹⁸

Nos últimos anos, uma variedade de trabalhos prospectando microrganismos capazes de degradar polissacarídeos foram conduzidos e uma série de resultados promissores e interessantes foram encontrados para estudos utilizando bactérias marinhas.¹⁹⁸⁻²⁰¹ O ciclo do carbono em ambiente marinho ainda não é totalmente compreendido, e a degradação microbiana das paredes das células vegetais não é tão bem caracterizada como em sistemas terrestres, como solo e bactérias ruminantes,²⁰³ o que indica que essa área de estudo possui um potencial muito grande a ser explorado.

Saccharophagus degradans (previamente conhecido como *Microbulbifer degradans*) é uma bactéria gram-negativa, móvel e aeróbica, inicialmente isolada de uma cultura obtida de *Spartina alterniflora* parcialmente decomposta em um pântano salino da Baía de Chesapeake em Matthews County, Va.²⁰⁴ Essa bactéria é capaz de utilizar carboidratos complexos, como quitina, celulose, pectina, amido e xilano, entre outros, como únicas fontes de carbono.²⁰⁵ Estudos anteriores relataram que este microrganismo pode degradar pelo menos 10 tipos diferentes de polissacarídeos complexos.²⁰⁶ Sua temperatura de crescimento é em torno de 30 °C e o seu pH ótimo é de 7,5.²⁰³⁻²⁰⁴

No presente estudo foi proposto caracterizar bioquímica e biofisicamente e determinar o envelope molecular de baixa resolução da XI de um microrganismo marinho, *Saccharophagus degradans* (SdXI). Além disso, foi determinada a estrutura cristalográfica de raios-X de SdXI com resolução de 2,5 Å, que serviu para estudos posteriores, como análise estrutural, conservação dentro do grupo e análises filogenéticas da família.

6.2. Materiais e métodos

6.2.1 Clonagem de XI de *S. degradans*

A etapa de clonagem de XI de *S. degradans* foi realizada previamente pelo Dr. Cesar Camilo, no laboratório do Grupo de Biotecnologia Molecular.

A metodologia de clonagem, incluindo desenho e construção de primers, condições da PCR, purificação do vetor pET Trx-1a/LIC, anelamento e transformação foi baseada no protocolo de Camilo & Polikarpov.⁸⁹ O DNA genômico de *S. degradans* cepa 2–40 foi usado como modelo para amplificação por PCR da inserção alvo e os oligonucleotídeos específicos do gene foram adquiridos da EXXTEND (Campinas, Brasil) (forward: 5'-CAGGGCGCCATGAGCGTAGTACTTGG - 3' e reverse: 5'-GACCCGACGCGGTTAACGAATATATTGG - 3'). O gene foi clonado e inserido em células de *E. coli* Rosetta (DE3) pLys para expressão da proteína.

6.2.2 Expressão heteróloga e purificação de proteína

As células de *E. coli* transformadas com o plasmídeo recombinante foram cultivadas em placas de meio de ágar LB suplementadas com canamicina (30 µg/mL) e cloranfenicol (34 µg/mL). Em seguida, um pré-inóculo foi realizado incubando uma colônia colhida em 5 mL de meio LB contendo os mesmos antibióticos sob agitação *overnight* (130 rpm), a 37 °C. Em seguida, o pré-inóculo foi transferido para um frasco de 2 L com 1 L de meio LB, também contendo os mesmos antibióticos e incubado a 37 °C, 150 rpm até atingir uma DO₆₀₀ de aproximadamente 0,6. Nesse ponto, a temperatura foi baixada para 18 °C e foi adicionado IPTG 0,5 mM (concentração final) para induzir a expressão da proteína, sob agitação de 150 rpm por 16 horas. As células foram colhidas por centrifugação (7000 rpm, 30 minutos, 4 °C) e ressuspensas em 20 mL de tampão de lise (N-lauroilsarcosina 0,4%, Tris-HCl 25 mM pH 8,0,

NaCl 200 mM, imidazol 5 mM, β -mercaptoetanol 10 mM, lisozima 0,25 mg/mL e PMSF 1 mM) mantendo em gelo durante 1 hora. A suspensão de células foi sonicada por dez ciclos alternando 30 segundos ligada e 45 segundos desligada com amplitude de 40%, o extrato foi clarificado por centrifugação (16.000 rpm, 30 minutos, 4 °C) e então inserido em uma coluna com 2,5 mL de resina Ni-NTA (Qiagen) equilibrada com tampão (imidazol 5 mM, NaCl 100 mM, Tris-HCl 50 mM, pH 8,0). O material não absorvido foi lavado da coluna utilizando 10 vezes o volume da coluna (cv) e a proteína purificada foi então eluída com 5 volumes da coluna de tampão de eluição (imidazol 300 mM, NaCl 100 mM, Tris-HCl 50 mM a pH 8,0). A fração eluída foi dialisada e concentrada. A tag His-Trx foi removida por digestão durante a noite com a protease do Tobacco Etch Virus (TEV) a 4 °C, e a proteína sem a tag foi purificada por uma segunda cromatografia de afinidade através de resina Ni-NTA, coletando o fluxo contendo a enzima alvo. As amostras de eluição foram analisadas por SDS-PAGE 15% ⁹⁰ coradas com azul brilhante de Coomassie R-250.²⁰⁷ A proteína foi ainda purificada por cromatografia de exclusão por tamanho, que foi realizada usando uma coluna Superdex-200 (16/60) com tampão NaCl 100 mM e Tris-HCl 50 mM a pH 8,0. A concentração de proteína foi determinada através do coeficiente de extinção ($\epsilon = 55.935 \text{ M}^{-1}\text{cm}^{-1}$) a 280 nm usando espectrofotômetro NanoDrop 2000 (Thermo Scientific).

6.2.3 Estudos de espalhamento dinâmico de luz (DLS)

SdXI purificada foi ressuspensa em tampão fosfato-NaCl 25 mM, pH 8,0. A dispersão dinâmica de luz (DLS) foi realizada para analisar o grau de dispersão, bem como a estabilidade de *SdXI* em solução antes das medições de SAXS. As medidas de DLS foram realizadas em sistema DLS (Zetasizer Nano-Malvern) em temperatura ambiente e com concentração de proteína de 1 mg/mL. Os dados foram avaliados usando o Zetasizer Nano Range.

6.2.4 Medições de dicroísmo circular (CD)

Os experimentos de Dicroísmo Circular foram realizados no Laboratório de Espectroscopia e Calorimetria (LEC), sob supervisão da Dra. Fernanda Batista e Dra. Ana Carolina Figueira.

Os espectros de Far-UV CD foram medidos usando uma célula de quartzo com comprimento de caminho óptico de 0,1 cm a 25 °C e concentração de proteína de 1,5 μ M, em um espectropolarímetro Jasco J-810 (Jasco Co.). Os espectros de UV distante da proteína foram registrados ao longo de uma faixa de comprimento de onda de 190-260 nm utilizando uma média de sinal de 8 espectros para o tampão e 20 espectros para a amostra. O instrumento foi ajustado para velocidade de varredura de 100 nm/min e resolução de etapa de 1 nm. O sinal da proteína foi corrigido para a contribuição do tampão e a previsão dos elementos da estrutura secundária foi obtida utilizando o servidor web CAPITO (<http://capito.nmr.leibniz-fli.de/>).²⁰⁸

6.2.5 Ensaio Enzimáticos

Os ensaios de atividade enzimática foram realizados utilizando o método colorimétrico de ácido sulfúrico-cisteína-carbazol,²⁰⁹ com xilulose como padrão para as curvas de calibração. Os ensaios de pH e temperatura ótimos foram realizados em triplicata com xilose como substrato. Para o pH ideal, a mistura de reação contendo 10 μ L de enzima a 0,2 mg/mL, 10 μ L de xilose 1 M, 5 μ L de CoCl_2 0,02 M e 5 μ L de MgSO_4 0,2 M foram diluídos em tampão citrato/fosfato/glicina 40 mM (tampão de McIlvaine) e incubado em diferentes pHs variando de 2,0 a 10,0. Após a determinação do pH ideal (pH = 8,0), o ensaio de temperatura foi realizado incubando a mesma mistura de reação de 20 a 80 °C por 30 min e parando as reações em banho de gelo. Depois disso, 25 μ L de cada reação foram levados para outra placa e 150 μ L de H_2SO_4 71%, 5 μ L de cisteína 1,5% e 5 μ L de carbazol 0,12% (em etanol) foram adicionados. As placas foram tampadas e deixadas em temperatura ambiente por 1 hora. A absorbância foi medida em 540 nm com espectrofotômetro. Os parâmetros cinéticos foram determinados utilizando concentrações crescentes de xilose (de 5 mM a 300 mM). As reações foram realizadas em Tris-HCl 25 mM, tampão NaCl 200 mM (pH 8,0) a 60 °C durante 10 minutos, e também medidas pelo método do ácido sulfúrico-cisteína-carbazol. As constantes cinéticas foram determinadas por regressão não linear usando OriginPro 8.0. Os ensaios de inibição foram conduzidos com o mesmo tampão, utilizando concentração crescente de xilose (de 10 mM a 100 mM) e diferentes concentrações de inibidor (xilitol 70, 140 e 210 mM). Os dados foram tratados pelos programas Excel e Origin e os parâmetros estabelecidos assumindo que uma unidade de enzima (U) produz 1,0 μ mol de xilulose por minuto.

6.2.6 Análise de estabilidade térmica

SdXI foi incubada com tampão Tris-HCl 25 mM, NaCl 200 mM (pH 8,0) por diferentes períodos de tempo, de 0 a 72 horas, a 30, 50 e 60 °C. Após o respectivo período de incubação, a reação contendo a enzima e o tampão Tris-HCl 25 mM, NaCl 200 mM (pH 8,0) foi incubada a 60 °C por 10 minutos e medida pelo método do carbazol-cisteína-ácido sulfúrico.

6.2.7 Medições de ultracentrifugação analítica (AUC)

Os experimentos de ultracentrifugação analítica foram realizados no Laboratório de Espectroscopia e Calorimetria (LEC), sob supervisão da Dra. Fernanda Batista e Dra. Ana Carolina Figueira. O tratamento dos dados foi realizado com o auxílio do Dr. Evandro Ares de Araújo. A intenção foi realizar experimentos de velocidade de sedimentação (VS) e equilíbrio de sedimentação (ES), mas por motivos operacionais foi possível realizar apenas experimentos de VS.

A estrutura quaternária de *SdXI* em solução foi analisada conduzindo experimentos de velocidade de sedimentação (VS) usando uma ultracentrífuga analítica ProteomeLab XL-I equipada com um rotor AN-60Ti e sistema óptico de absorvância (Beckman Coulter, Brea, CA). Os experimentos foram realizados no Laboratório de Espectroscopia e Calorimetria (CNPEM, Campinas, Brasil). Quatrocentos µL de solução de proteína a 0,86 mg/mL pré-equilibrada em uma solução tampão (fosfato de sódio 25 mM pH 8,0, NaCl 25 mM) foram carregados em uma peça central de epóxi de setor duplo no conjunto de células com um comprimento de caminho óptico de 1,2 cm.

Antes das medições de VS, as amostras foram centrifugadas a 17.000 x g por 5 minutos a 4 °C e o rotor foi resfriado a 20 °C por 2 horas. Os experimentos foram realizados a 20 °C usando uma velocidade de rotor de 20.000 rpm. Foi realizado modo de varredura contínua (de 5,9 - 7,2 cm) com aquisição de quadros de varredura em intervalos de 3 minutos. A análise e processamento dos dados foram conduzidos usando o programa SEDFIT (<https://sedfitsedphat.nibib.nih.gov/software>)²¹⁰ e os parâmetros biofísicos usados no modelo

de ajuste foram calculados pelo programa SEDNTERP (<http://www.jphilo.mailway.com/download.htm>).

6.2.8 Modelagem por homologia

A sequência de aminoácidos de *SdXI* foi obtida a partir da base de dados UniProt (<http://www.uniprot.org/>).²¹¹ Usando a ferramenta Swiss Model Modeling (<https://swissmodel.expasy.org/interactive/>),²¹² 50 modelos foram selecionados a partir da sequência original e um modelo foi gerado com base na estrutura cristalográfica da xilose isomerase de *Piromyces* E2 (PDB id: 5NH4)²¹³ que apresenta 61.75% de identidade e 97% de cobertura. Alternativamente, também foi utilizado o programa Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>)²¹⁴ para a construção do modelo por homologia. Neste caso, foi utilizada a estrutura 3D da xilose isomerase de *Thermoanaerobacterium thermosulfurigenes* como modelo (PDB id: 1A0C)²¹⁵ com 50.69% de identidade e 97% de cobertura.

6.2.9 Coleta e processamento de dados SAXS

Os dados de SAXS foram coletados na linha de luz D02A-SAXS2 do Laboratório Nacional de Luz Síncrotron (Campinas, Brasil) usando um detector Pilatus 300k com faixa de transferência de momento de $0,01 \text{ \AA}^{-1} < q < 0,45 \text{ \AA}^{-1}$. Três concentrações diferentes de proteína (1, 2 e 6 mg/mL) foram medidas e 10 imagens foram coletadas em intervalo de exposição de 30 segundos cada. A imagem média foi gerada utilizando o programa FIT2D.²¹⁶ Foi utilizada a definição de transferência de momento q como $q = 4\pi \sin\theta/\lambda$, onde 2θ é o ângulo de espalhamento observado entre o feixe incidente e espalhado e λ é o comprimento de onda do feixe de raios-X incidente.

Para cada conjunto de dados o raio de giração (R_g) foi estimado utilizando a equação de Guinier $I(q) = I(0) \cdot \exp[(-q^2 \cdot R_g^2)/3]$, para a qual a região de Guinier foi definida como o limite $(q_{min}, q_{max}) = (\pi/D_{max}, 1.3/R_g)$.²¹⁶⁻²¹⁸ A função $P(r)$ de distribuição de pares de distância da partícula e dimensão máxima D_{max} foram estimadas utilizando o programa GNOM.⁶⁶

Modelos de baixa resolução foram gerados utilizando uma abordagem ab initio. Esta modelagem ab initio de modelos de “átomos falsos” (DAMs) foi realizada usando o programa

DAMMIF²²⁰ e um modelo médio foi calculado pelo programa DAMAVER,²²¹ e o modelo final foi aprimorado com o software DAMMIN.²²² Para selecionar o modelo mais típico, a discrepância mínima entre as curvas de SAXS experimentais e calculadas foi quantificada usando o parâmetro χ minimizado. Com um modelo DAM para cada amostra de proteína, pelo menos 15 modelos foram calculados em média usando o programa DAMAVER e as figuras das representações do modelo foram geradas com o programa Pymol Molecular Graphics (Schrödinger, LLC.). Além disso, foram gerados modelos utilizando o programa GASBOR⁷³, para comparação e escolha do melhor modelo de envelope molecular. O peso molecular da proteína em solução foi calculado a partir dos dados SAXS utilizando o programa SAXS_MoW.^{70, 223}

6.2.10 Ensaios de cristalização da *SdXI*

A *SdXI* purificada foi concentrada a 10 mg/mL, com base em seu coeficiente de extinção ($\epsilon = 55.935 \text{ M}^{-1}\text{cm}^{-1}$) a 280 nm, conforme determinado utilizando o espectrofotômetro NanoDrop 2000 (Thermo Scientific). Os ensaios de cristalização foram realizados por um robô de cristalização (Honeybee, DIGILAB), usando três kits de cristalização diferentes: PEG II Suite (QIAGEN), Crystal Screen HT (Hampton Research) e Index HT (Hampton Research). Os cristais foram formados e crescidos a 20 °C, por técnica de difusão de vapor em gota sentada e 3 diferentes condições de relação proteína-poço foram testadas: 0,5: 1; 1:1 e 2:1.

6.2.11 Coleta e processamento de dados

Os cristais de proteína foram congelados rapidamente em nitrogênio líquido após imersão em óleo mineral. Os dados de difração de raios X foram coletados na linha de luz W01B-MX2 do Laboratório Nacional de Luz Síncrotron (Campinas, Brasil),⁹³ usando um detector Pilatus2M (Dectris), com tempo de exposição de 5 s por imagem e ângulo de rotação de 0,33° resultando em 1080 imagens. Os dados das imagens de difração foram indexados, integrados e escalonados utilizando o programa XDS⁹⁴ e alternativamente o programa DIALS,²²⁴ além do programa AIMLESS.⁹⁵

6.2.12 Construção e validação de modelo cristalográfico

A enzima *BtXI* de *Bacteroides thetaiotaomicron* (código PDB 4XKM)²²⁵ foi identificada como o modelo de homologia mais próximo com estrutura cristalográfica conhecida de acordo com BLASTp,⁹⁶ com 98% de cobertura e 62% de identidade de sequência de aminoácidos.

A substituição molecular foi realizada através do programa Phaser⁵² utilizando as coordenadas atômicas de *BtXI* como modelo de busca e os resultados provenientes do programa XDS. A estrutura foi então submetida ao Autobuild⁹⁸ e os ciclos de refinamento foram realizados usando os programas Coot⁹⁹ e phenix.refine.¹⁰⁰ A qualidade do modelo foi avaliada usando o programa Molprobity.¹⁰¹

Como uma abordagem alternativa, o resultado obtido através do DIALS foi submetido ao programa BALBES,¹²⁷ a fim de encontrar a melhor solução. O resultado foi então analisado com o Coot, submetido ao Autobuild, refinado através do phenix.refine e avaliado utilizando Molprobity. As coordenadas atômicas serão ainda depositadas no Protein Data Bank (PDB).

6.2.13 Análise estrutural

Para alinhar *SdXI* com outras sequências de XIs e comparar suas estruturas secundárias, foram usados os programas Toffee²²⁶ e ESPRIT.¹⁰⁴ Para comparações estruturais, como área de interação, topologia, mapa de domínio e resíduos de interação, foram utilizados os programas PDBSum generate¹³⁸ e PISA.²²⁷

6.2.14 Estudos de conservação e análise filogenética

Para estudos de conservação a sequência *SdXI* foi submetida ao BLAST contra 1000 outras sequências e um arquivo de alinhamento foi criado com ClustalW.¹⁰³ O programa ConSurf²²⁸ foi utilizado para identificar resíduos conservados, através da submissão do arquivo PDB de *SdXI* e do arquivo de alinhamento (Método ConSurf).²²⁹ Para a análise filogenética as sequências foram alinhadas através do programa MUSCLE²³⁰ e o programa MEGA7²³¹ foi utilizado para construção e análise da árvore filogenética.

6.3 Resultados e discussão

O gene que codifica a XI de *Saccharophagus degradans* foi clonado e expresso com sucesso em *E. coli* (DE3)Rosetta pLys. A sequência do gene contém um quadro de leitura aberto (ORF) de 1332 bp que codifica um polipeptídeo de 443 aminoácidos. Segundo o ProtParam,²³² a massa molecular teórica do monômero da proteína é de 49,5 kDa, e após as etapas de purificação, o rendimento da *SdXI* recombinante foi em torno de 30 mg por litro. A etapa de purificação de filtração em gel e o SDS PAGE do *SdXI* resultante estão na figura 37.

Figura 37 - Expressão e purificação de *SdXI*. (A) SDS-PAGE indicando o peso molecular em kDa, o marcador de peso molecular (M), o extrato celular (1) e a fração após a purificação (2). (B) perfil cromatográfico de gel-filtração, mostrando o pico único referente a *SdXI*.

Fonte: Elaborada pelo autor.

6.3.1 pH e temperatura ideais

A *SdXI* recombinante apresentou atividade significativa na faixa de pH de 7,5 a 9,0 (Figura 38a), e uma temperatura ideal para a atividade enzimática de 60 °C (Figura 38b). Um declínio acentuado na atividade enzimática de *SdXI* é observada a temperaturas a partir de 70 °C (Figura 38b).

Figura 38 - Perfil de temperatura ótima (a) e pH ótimo (b) de *SdXI*.

Fonte: Elaborada pelo autor.

Considerando que a sequência codificante de DNA (no inglês, *coding DNA sequence*, ou CDS) de *SdXI* foi clonada a partir de um microrganismo aquático da Baía de Chesapeake era de se esperar que essa enzima apresentasse um pH levemente alcalino, uma vez que o pH da Baía de Chesapeake está entre 7,0 a 9,0.²³³ Porém, em relação à temperatura ótima, o valor relativamente alto obtido nos experimentos claramente não corresponde à temperatura média da água da Baía (15-25 °C).²³³ Impulsionada pela adaptação do microrganismo às condições de seu nicho, a temperatura ótima das atividades enzimáticas tende a seguir as condições normais de temperatura ambiente. A temperatura ótima da enzima consideravelmente mais alta pode indicar uma característica conservada dentro deste grupo, uma vez que o hospedeiro raramente, ou nunca, experimentaria tais condições de alta temperatura, e muitas XIs costumam ter temperaturas ótimas mais altas.²³⁴

6.3.2 Estabilidade térmica

Posteriormente, os dados de estabilidade térmica do *SdXI* foram medidos e analisados (Figura 39). A enzima é consideravelmente estável a 30 °C e mantém 50% de sua atividade inicial após 72 horas a 50 °C. No entanto, sendo incubada a 60 °C, sua atividade enzimática diminui para 72% da atividade original após 4 horas e cai para menos de 40% após 32 horas de incubação (Figura 39).

Figura 39 - Atividade relativa ao longo do tempo (horas), em diferentes temperaturas.

Fonte: Elaborada pelo autor

6.3.3 Ensaios cinéticos

Após a determinação das condições ideais para a atividade enzimática, ensaios cinéticos foram realizados. O valor de K_M foi determinado usando 10 minutos de reação a 60 °C. Uma pesquisa no banco de dados de enzimas BRENDA (<https://www.brenda-enzymes.org/>)²³⁴ mostra que a faixa de K_M exibida por XIs caracterizadas bioquimicamente está entre 5 mM e 95 mM, com alguns valores superiores a 100 mM. O valor de K_M determinado experimentalmente para *SdXI* foi 58 mM, que está dentro da faixa esperada de valores determinados para outras XIs. Staudigl *et al.*²³⁵ determinou um K_M de 177,4 mM para uma XI de *Lactobacillus reuteri*. Os valores de K_M para XIs de *Piromyces* sp,²³⁶ *Streptomyces rubiginosus*²³⁷ e *Thermus thermophilus*²³⁸ foram relatados como 86,97 mM, 83 mM e 605 mM, respectivamente. Esses valores de K_M são consideravelmente maiores do que o valor encontrado para *SdXI*. Por outro lado, o valor de K_M para XIs de *Streptomyces olivochromogenes*²³⁹ *Thermus aquaticus*²⁴⁰ e *Thermoanaerobacter ethanolicus*²⁴¹ foram calculados como 33, 15 e 11 mM, respectivamente. Isso evidencia a necessidade de mais estudos estruturais, que possam relacionar eventuais modificações estruturais à afinidade da enzima pelo substrato.

A SdXI tem uma $V_{\max}$ tão alta quanto 15 UI/mg, comparável ao valor reportado para a XI de *Thermoanaerobacter ethanolicus* (17 UI/mg)²⁴¹. O valor de k_{cat} para SdXI foi calculado como $9,95 \text{ s}^{-1}$, consideravelmente maior do que o descrito para *Streptomyces rubiginosus* ($5,52 \text{ s}^{-1}$).²³⁷ O valor k_{cat}/K_M de SdXI -e de $0,17 \text{ mM}^{-1}\text{s}^{-1}$, semelhante à eficiência catalítica do mutante D57N da XI de *Streptomyces rubiginosus*, com um valor de k_{cat}/K_M de $0,18 \text{ mM}^{-1}\text{s}^{-1}$.²³⁷

Após a definição dos parâmetros cinéticos (K_M , $V_{\max}$, k_{cat} e k_{cat}/K_M), ensaios de inibição foram realizados para determinar o valor de K_i para um inibidor bem conhecido de Xilose Isomerase, o xilitol. Três concentrações diferentes de xilitol foram testadas contra um gradiente de xilose. As curvas (figura 40) mostraram que o xilitol é um inibidor competitivo, levando ao aumento do K_M aparente, com o mesmo valor de $V_{\max}$. Com o K_M app obtido para as diferentes concentrações de inibidores, o valor de K_i determinado para esta enzima foi de 99 mM.

Figura 40 - Inibição de SdXI por Xilitol, em diferentes concentrações.

Fonte: Elaborada pelo autor.

O K_i obtido para o xilitol é significativamente maior do que o K_i determinado para a XI de *Bacteroides stercoris* (5,1 mM).²⁴² Valores de K_i de 4,6 mM, 68,7 mM e 33,2 mM, respectivamente, para uma enzima de tipo selvagem, mutantes E372G e E372G/V379A de XI de *Thermus thermophilus*.²⁴³ Para a inibição competitiva, o valor do K_i é igual à concentração do inibidor que duplica o valor K_M aparente. Diante disso, valores mais elevados de K_i são preferíveis, pois representam uma inibição mais fraca.

6.3.4 Ultracentrifugação analítica (AUC)

Como mencionado anteriormente no item 6.2.7, o intuito inicial foi de realizar experimentos de ultracentrifugação analítica de velocidade de sedimentação e equilíbrio de sedimentação, uma vez que este último pode fornecer diversos parâmetros da enzima, tais como constantes cinéticas. Entretanto, por motivos operacionais, apenas os experimentos de velocidade de sedimentação foram realizados para avaliar o comportamento molecular de SdXI em solução. Os parâmetros biofísicos (densidade do tampão $\rho = 1,0028 \text{ g/cm}^3$, viscosidade do tampão $\eta = 0,0180$, volume específico parcial da proteína $V\text{-bar} = 0,73000 \text{ cm}^3/\text{g}$) usados no modelo de ajuste foram calculados pelo programa SEDNTERP. Parâmetros como coeficiente de sedimentação $c(S)$, distribuição e razão de atrito (f/f_0), bem como a massa molecular da proteína (equação 1.21), foram obtidos a partir da análise utilizando o SEDFIT. A distribuição $c(S)$ exibiu um único pico que cobriu a faixa de 10,77S - 11,03S e um coeficiente de sedimentação máximo ($S_{20,w}$) de 10,9S, demonstrando a presença de um único componente proteico (Figura 41).

Figura 41 - Gráfico de Ultracentrifugação analítica.

Fonte: Cedida por Dr. Evandro Araújo.

Para o cálculo da massa molecular, temos a seguinte equação relacionando o coeficiente de sedimentação e a massa molecular:²⁴⁴

$$s = u / \omega^2 r = M (1 - \bar{v}\rho) / Nf \quad (1.21)$$

Onde s é o coeficiente de sedimentação, u é a velocidade, ω é a velocidade angular (em radianos/segundo), r é a distância radial ao centro de rotação, M é a massa molar do soluto (g/mol), $\bar{v}$ é o volume específico parcial, ρ é a densidade da solução, N é o número de Avogadro e f é o coeficiente friccional.

Uma vez que *SdXI* tem uma massa molecular teórica de 49,6 kDa, calculada com base em sua sequência de aminoácidos, a massa molecular obtida de 205 kDa indica claramente que o *SdXI* forma tetrâmeros em solução. Uma discrepância entre a massa teórica para o tetrâmero *SdXI* e sua massa molecular experimentalmente determinado pela AUC é de apenas 3%. Esta análise também mostra uma razão de atrito (f/f_0) de 1,13, sugerindo que *SdXI* adota a conformação de um tetrâmero altamente compacto e estável em solução (figura 41).

6.3.5 Modelagem por homologia de *SdXI*

Utilizando as ferramentas Swiss Model e Phyre2, foram construídos diferentes modelos de *SdXI* a partir da sua sequência. A sobreposição de ambas as estruturas resultou em um desvio da raiz quadrática média (RMSD) de 1,26 Å. A ferramenta da web SAVES (<http://services.mbi.ucla.edu/SAVES/>) foi usada para a estimativa e validação da qualidade do modelo. SAVES é um meta-servidor que executa um conjunto de 6 programas (CRYST1, PROVE, ERRAT, VERIFY3D, PROCHECK, WHATCHECK) para analisar os modelos. Os resultados indicaram que o modelo de homologia feito através do SWISS-MODEL (Figura 42) é o mais adequado, pois os fatores de qualidade gerais fornecidos pelo ERRAT foram 94,601 e 86,571 para a estrutura do SWISS-Model e para o modelo Phyre2, respectivamente.

Figura 42 - Modelo por homologia de SdXI gerado pelo programa Swiss-Model. Observa-se o homotetrâmero, composto por 4 cadeias iguais, cada uma com uma cor. As esferas roxas representam o íon Mg²⁺.

Fonte: Elaborada pelo autor.

6.3.6 Medições de DLS e CD

Os experimentos de DLS foram realizados anteriormente aos ensaios de SAXS, para confirmar a monodispersividade da amostra, uma vez que a polidispersividade da mesma poderia dificultar as análises dos resultados dos experimentos. Além disso, foi possível também estimar o tamanho e o raio hidrodinâmico da enzima. Como até o momento não havia sido possível a cristalização da enzima, foram realizados experimentos de CD a fim de estimar as porcentagens de conteúdo de estrutura secundária da *SdXI*.

Os estudos de DLS revelaram uma proteína monodispersa com um raio hidrodinâmico de $50 \pm 14 \text{ \AA}$ e um peso molecular estimado (MW) de $164,4 \pm 45 \text{ kDa}$. A distribuição de massa experimental pode ser observada na Figura 43.

Figura 43 - Perfil de DLS da enzima *SdXI*.

Fonte: Elaborada pelo autor.

Para analisar a estrutura secundária do *SdXI*, dados experimentais de CD foram plotados usando o programa Origin (Figura 44) e processados usando o software CAPITO.²⁰⁸

Figura 44 - Espectro de dicroísmo circular em função da elipcticidade (mdeg) da *SdXI* em tampão Na_2PO_4 25 mM NaCl 25 mM pH 8,0 a 25 °C.

Fonte: Elaborada pelo autor.

O modelo estrutural de homologia construído usando a análise de estrutura secundária SWISS-MODEL foi processado usando 2Struc (<http://2struc.cryst.bbk.ac.uk/>).²⁴⁵ A comparação revelou que os resultados experimentais eram comparáveis aos dados previstos de Chou-Fassmann, mas muito mais semelhantes ao modelo gerado por homologia *SdXI*, conforme observado na Tabela 1.

Tabela 13 - Comparação dos elementos de estrutura secundária em *SdXI* computados com base na sequência protéica, no modelo por homologia e nos dados experimentais obtidos a partir dos experimentos de CD.

Dados	Hélice	Fita- β	Irregular
Sequência	0,54	0,25	0,21
Modelo por homologia	0,42	0,19	0,38
Espectro de CD	0,41	0,30	0,43

Fonte: Elaborada pelo autor

6.3.7 Análise de Espalhamento de Raios-X a Baixo Ângulo (SAXS)

Com o objetivo de obter mais informações moleculares sobre a enzima *SdXI* em solução, incluindo seu tamanho, forma e estado oligomérico, experimentos de SAXS foram conduzidos utilizando radiação síncrotron na Fonte Nacional de Luz Síncrotron (LNLS, Campinas, Brasil). A curva de espalhamento de raios X e sua análise inicial são mostradas na Figura 45. A região de Guinier apresentou um sistema monodisperso, sem evidências de agregação ou dano por radiação. Usando a transformada de Fourier indireta da curva de espalhamento, uma função de distribuição de distâncias de pares $P(r)$ foi obtida, que mostrou um perfil simétrico típico para macromoléculas compactas / globulares (Fig. 45). O tamanho máximo da proteína, D_{max} , obtido por esta função foi 108 Å e o valor do raio de giração calculado a partir da função de distribuição de distância ($R_g = 36,71$ Å) está de acordo com o valor obtido pela análise de Guinier ($36,65 \pm 0,28$ Å).

Figura 45 - Curva de espalhamento de raios-x a baixo ângulo para *SdXI* (a) e função $P(r)$ (b).

Fonte: Elaborada pelo autor.

Uma forma 3D de baixa resolução da proteína foi construída utilizando modelagem ab initio realizada com pacotes DAMMIN e GASBOR. Dez modelos de átomos falsos (DAMs) foram calculados usando DAMAVER para os modelos DAMMIN e GASBOR.

Os modelos médios gerados com DAMMIN e GASBOR foram analisados, e o melhor modelo médio, construído com o DAMAVER, foi sobreposto (Figura 46) com o modelo de homologia *SdXI* construído usando SWISS-MODEL. A sobreposição foi realizada por meio do programa SUPCOMB.

Figura 46 - Modelo 3D de *SdXI* construído pelo DAMAVER (A), com estrutura do modelo por homologia de *SdXI* sobreposta (B). (C) e (D) representam diferentes rotações ao longo do eixo X e Y, respectivamente.

Fonte: Elaborada pelo autor.

A intensidade de espalhamento do DAM final pode ser comparada diretamente com a curva de espalhamento experimental. A Figura 47 mostra a intensidade de espalhamento teórica calculada a partir de DAMs simulados ab initio como uma linha vermelha, ajustada com a curva SAXS experimental (círculo preto aberto) e a curva de espalhamento derivada do GASBOR (linha azul claro). Uma curva de espalhamento do modelo por homologia (*SdXI* modelado em SWISS-MODEL) foi gerada usando FoXS (<https://modbase.compbio.ucsf.edu/foxs/about.html>),²⁴⁶ e é apresentada em azul escuro na Figura 47.

Figura 47 - Curvas de espalhamento de *SdXI*, os círculos abertos representando os dados experimentais, a linha vermelha representando o modelo Dam, a linha azul clara representando o modelo Gasbor, e a linha azul escura o modelo por homologia.

Fonte: Elaborada pelo autor.

A curva de espalhamento reconstruída com base no modelo DAMs médios e a curva de espalhamento simulada calculada a partir do modelo de homologia ($\chi = 0,91$) estão em muito bom acordo com os dados experimentais SAXS. O DAM final obtido e o modelo construído por homologia confirmam claramente que a xilose isomerase de *Saccharophagus degradans* tem um arranjo tetramérico e uma forma molecular aproximadamente globular em solução. Além disso, o R_g (~ 36 Å) e o D_{max} (~ 108 Å) são próximos aos de XIs típicos. De acordo com o Small Angle Scattering Biological Data Bank (SASBDB, <https://www.sasbdb.org/data/>), uma XI de *Streptomyces rubiginosus* (entrada SASBDB: SASDAK6) tem R_g e D_{max} de 34 Å e 97 Å, respectivamente. Além disso, a curva de espalhamento de XI de *Streptomyces rubiginosus* e a $P(r)$ estão muito próximas das obtidas para *SdXI* (Figura 45). A estimativa de peso molecular *SdXI* com base em dados SAXS calculados usando SAXS_MoW é de 198 kDa, que é quase exatamente o peso molecular

teórico para o tetrâmero *SdXI* (uma discrepância de 0,11% de o MW esperado). Estes resultados, juntamente com estudos de DLS e AUC (Tabela 2), confirmam a estrutura quaternária tetramérica de *SdXI* e mostram sua forma aproximadamente globular em solução.

Tabela 14 - Comparação entre diferentes técnicas e parâmetros obtidos para *SdXI*. * R_h refere-se ao raio hidrodinâmico obtido pelo experimento de DLS, e R_g refere-se ao raio de giração, obtido pelo experimento de SAXS. Estado se refere ao estado oligomérico da enzima em solução.

	R_g ou R_h^*	MW Estimada	Dispersividade	Estado
DLS	$50 \pm 14 \text{ \AA}$	$164.4 \pm 45 \text{ kDa}$	Monodispersa	Tetramérica
AUC	-	205 kDa	Monodispersa	Tetramérica
SAXS	$36 \text{ \AA}$	198 kDa	Monodispersa	Tetramérica

Fonte: Elaborada pelo autor.

De acordo com o PROTPARAM, a *SdXI* apresenta um peso molecular teórico na sua forma tetramérica de 198,3 kDa. Isto demonstra que as técnicas utilizadas convergiram para uma mesma informação, e nesse caso, a técnica de SAXS foi a mais sensível em relação a aferição da massa molecular da enzima.

6.3.8 Alinhamento de sequências de aminoácidos e análise de estrutura secundária

Cinco sequências de aminoácidos XIs (de *Piromyces* sp., *Thermoanaerobacterium thermosulfurigenes*, *Streptomyces rubiginosus*, *Actinoplanes missouriensis* e *Thermus thermophilus*; código de acesso GENBANK CAB76571.1, P19148, AAA26838.1, WP_014441152.1 e WP_014677.152.1 e WP_0146.152.1 e WP_0146.1651, respectivamente) foram alinhadas à *SdXI* utilizando o programa T-Coffee²²⁶ no modo Expresso e depois submetido ao ESPRIPT.¹⁰⁴ Os resultados (Figura 48) indicaram uma diferença clara entre as Xilose Isomerases de Classe I e Classe II.

Figura 48 - Resultado do alinhamento ESPRIT, indicando regiões conservadas (caixas amarelas e vermelhas), região de abertura do anel (setas verdes) e região de ligação de metal (seta vermelha para o sítio 1 e seta azul para o sítio 2). O ácido glutâmico (E274 em *SdXI*) no final da folha $\beta 10$ com duas setas coordena simultaneamente dois íons metálicos.

Fonte: Elaborada pelo autor.

Como pode ser observado na Figura 48, apesar do grande grau de conservação no sítio ativo de XI, existem diferenças estruturais entre as duas classes de Xilose Isomerases, como a região N-terminal e comprimento de algumas hélices, como pode ser observado na Tabela 15. A *SdXI* claramente se insere na Classe II de Xilose Isomerases com uma região N-terminal notadamente maior do que as XIs de Classe I.

Tabela 15 - Diferenças encontradas no alinhamento ESPRIPT entre as classes XI.

Classe	Microorganismos	Loop N-terminal	hélice $\eta 1$	hélice $\alpha 3$	hélice $\alpha 5$
<i>S. rubiginosus</i>					
I	<i>A. missouriensis</i>	Curto	ausente	20 resíduos	21 resíduos
	<i>T. termophilus</i>				
<i>S. degradans</i>					
II	<i>Piromyces sp.</i>	Longo	presente	14 resíduos	17 resíduos
	<i>T. thermosulfurigenes</i>				

Fonte: Elaborada pelo autor.

As diferenças encontradas entre as duas classes são consistentes com os estudos anteriores.²²⁵ Considerando o alinhamento, as diferenças entre as duas classes de XI são essencialmente em relação ao comprimento, com a Classe II sendo de 40 a 50 resíduos mais longa, em média. No entanto, apesar das diferenças de comprimento (mostradas na Tabela 1), o sítio ativo, bem como os sítios de metal, são muito conservados, conforme mostrado na Figura 1.

6.3.8.1 Análise de estrutura secundária

No intuito de compreender melhor as diferenças estruturais entre as diferentes classes de XI, a estrutura secundária de *SdXI* também foi analisada utilizando o PDBSum generate (<http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html>)¹³⁸ e comparada com outras estruturas, particularmente da Classe I (*A. missouriensis*)²⁴⁷ e II (*Piromyces sp.*)²¹³. Os gráficos da estrutura secundária dessas três enzimas são apresentados na Figura 49.

Figura 49 - Gráficos da estrutura secundária de *PirXI* (a), *SdXI* (b) e *A. missouriensis XI* (c).

Fonte: Elaborada pelo autor.

Comparando os padrões apresentados na Figura 49, é possível identificar diferenças que não foram evidentes no alinhamento. A figura indica que, além da diferença no comprimento, existem também algumas diferenças nos elementos de estrutura secundária entre as classes de XI. Uma comparação entre enzimas das diferentes classes pode ser observada na Tabela 16, fornecida pelo ProMotif no PDBSum.

Tabela 16 - Comparação do conteúdo da estrutura secundária, em números absolutos, entre *SdXI*, *PirXI* e *A. missouriensis XI*.

	<i>SdXI</i>	<i>PirXI</i>	<i>A. missouriensis XI</i>
Folhas	4	3	3
Unidades beta-alfa-beta	5	5	3
Grampo beta	1	1	-
Protuberâncias beta	3	3	-
Fitas	14	12	10
Hélices	20	20	22
Interações hélice-hélice	43	49	41
Voltas beta	27	25	19
Voltas gamma	10	8	5

Fonte: Elaborada pelo autor.

Esta análise mostrou que *SdXI* e *PirXI* apresentam conteúdos muito semelhantes de elementos da estrutura secundária. Esse resultado também revela que, apesar de algumas semelhanças, como conteúdo de Hélices e Folhas, estruturas do tipo protuberâncias beta (beta bulges) aparecem apenas nas enzimas Classe II, estando ausentes nas enzimas Classe I de XI. Este padrão não explica as diferenças nos parâmetros cinéticos por si só, levando essa investigação para uma análise da estrutura terciária.

6.3.9 Coleta de dados de cristalografia, refinamento e estrutura terciária

O cristal de *SdXI* surgiu nas seguintes condições: 0.1 M TRIS pH 8.5, 2 M $(\text{NH}_4)_2\text{SO}_4$, em aproximadamente uma semana. Os dados cristalográficos foram coletados no Laboratório Nacional de Síncrotron (LNLS).⁹³ O cristal pertence ao grupo espacial P1 com oito cadeias de proteínas em unidade assimétrica, formando dois tetrâmeros com subunidades idênticas. A solução obtida através do programa BALBES foi a melhor, utilizando como modelo a estrutura de XI de *Thermotoga neapolitana* (PDB 1A0E) e gerando uma solução com fator Q de 0,796 e probabilidade de 99% de ser uma solução. O resultado proveniente da

substituição molecular foi então refinado utilizando o programa phenix.refine, e submetido ao Autobuild, a fim de inserir corretamente os átomos ausentes, em particular nas cadeias laterais. Estas etapas levaram a uma diminuição do R-free e o refinamento adicional foi realizado usando também o Coot (inserção de metais, análise de rotâmeros e plotagem de Ramachandran) e uma estrutura final foi avaliada através do MolProbity. Ambos os parâmetros de coleta e refinamento são fornecidos na Tabela 17.

Tabela 17 – Coleta de dados cristalográficos *SdXI* e estatísticas de refinamento do modelo.

Conjunto de dados	<i>SdXI</i>
Coleta de dados	LNLS MX2
Comprimento de onda (Å)	1,4586
Grupo espacial	P 1
Parâmetros de cela	
a, b, c (Å)	86,41 91,68 136,20
α, β, γ (°)	72,27 82,19 65,62
Faixa de resolução (Å)	39,35 – 2,70 (2,75 – 2,70)
R_{meas}	0,238 (0,627)
$CC_{1/2}$	0,931 (0,682)
Reflexões totais	199.760 (9.949)
Reflexões únicas	95.539 (4.595)
Multiplicidade	2,1 (2,2)
Completeza (%)	96,0 (94,3)
Mosaicidade	0,3
Delta B Anisotrópico (Å ²)	15,03
$\langle I/\sigma_I \rangle$	5,2 (2,1)

(continua)

 (continuação)
Refinamento

No. Mol. Unidade assimétrica	8
Resíduos por molécula	443
R_{work}	0,195 (0,247)
R_{free}	0,235 (0,277)
CC geral	0,854
Ligantes	Mg
<i>B-factors</i> ($\text{\AA}^2$)	
Médio	14,80
Macromoléculas	14,78
Ligantes	16,58
<i>R.m.s. deviations</i>	
Comprimentos de ligação ($\text{\AA}$)	0,003
Ângulos de ligação ($^\circ$)	0,500
<i>Ramachandran</i>	
Ramachandran favoráveis (%)	97,02
Ramachandran permitidos (%)	2,67
Ramachandran outliers (%)	0,31
Rotâmeros outliers (%)	0
<i>MolProbity Overall Score</i>	1,60

*Os valores em parênteses são referentes à camada de mais alta resolução.

Fonte: Elaborada pelo autor.

O monômero de *SdXI* (Figura 50) consiste em dois domínios principais: um barril TIM (consistindo em um barril β paralelo de oito fitas) localizado na região N-terminal com um grande loop inicial, e uma cauda C-terminal, importante no processo de tetramerização.

Figura 50 - Monômero *SdXI*, com os domínios principais indicados. Painel inferior: Representação das regiões de estrutura secundária de *SdXI*, com a região N e C terminal destacadas.

Fonte: Elaborada pelo autor.

Sabe-se que os XIs são encontradas principalmente no estado tetramérico em solução e que as estruturas terciárias e quaternárias são muito importantes na atividade da XI.²⁴⁸ A

interface dimérica também é considerada muito importante, pois o estado dimérico pode apresentar atividade, ao contrário do monômero, sem atividade alguma.²⁴⁸ Para melhor compreensão da formação da interface em relação a parâmetros como o número de resíduos dentro de cada interface, área de contato, ΔiG , número e tipo de ligações e Pontuação de Significância de Complexação, foi realizada a análise utilizando o PISA (Tabela 18).

Tabela 18 - Resultados do PISA para *SdXI*. M1: Monômero 1; M2: Monômero 2; Área: Área de interface ($\text{\AA}^2$); ΔiG : Ganho de energia livre de solvatação na formação da interface, em kcal / M; Nres: Número de resíduos na interface; NHb: Número de ligações de hidrogênio; NSb: Número de pontes de sal; CSS: Complexation Significance Score, que indica o quão significativa é a interface para a formação da estrutura quaternária.

M1	M2	Área ($\text{\AA}^2$)	ΔiG	Nres	NHb	NSb	CSS
A	E	4084.7	- 42.9	106	47	19	0.639
C	D	4115.8	- 42.9	105	48	21	0.639
A	C	2080.6	3.7	54	45	11	0.127
D	E	2106.6	4.4	55	42	13	0.127
A	D	1370.6	-8.7	38	15	10	0.143
C	E	1359.6	-13.9	38	16	9	0.143

Fonte: Elaborada pelo autor.

De acordo com os dados obtidos, a interface dimérica apresentou a maior área, seguida da interface dímero-dímero. No entanto, as interações (AD) e (CE), embora com menos resíduos dentro da região de interface, também contribuem para a estabilização da estrutura quaternária, para o valor delta G do tetrâmero (ΔG_{int} de -177,5 kcal/Mol e ΔG_{diss} de 149,7 kcal/Mol). Os valores Delta G para os conjuntos (dímero, tetrâmero e interação cruzada) indicaram que os últimos desempenham um papel muito importante na manutenção da estrutura quaternária. Isso pode ser corroborado pelos valores de CSS, uma vez que a pontuação é definida “como uma fração máxima da energia livre total de ligação que pertence à interface em montagens estáveis” [40], variando de 0 a 1. Um valor positivo para cross-interação, ligeiramente superior à interação tetrâmero, destaca que este tipo de interação pode ser mais importante para a manutenção do tetrâmero do que a própria interface tetramérica. Isso pode fornecer uma visão inicial sobre a coevolução da estrutura, uma vez que um

monômero inativo tem três regiões de interface, formando um tetrâmero muito compacto e ativo.

A análise com o programa CCP4MG²⁴⁹ indicou a montagem do tetrâmero (Figura 51).

Figura 51 - Oligomerização em *SdXI*. Em detalhes, dímeros (EA) e (DC), tetrâmero (ACDE), interface dímero-dímero (ED) e esquema de oligomerização, mostrando área de interações, ligações (Hidrogênio e Sal - vermelho e azul, respectivamente) e contatos não ligados (laranja). O esquema no canto inferior direito indica conjuntos diméricos AE e CD (interface monômero - monômero); Interface tetramérica (AC e ED) e interações cruzadas (AD e EC).

Fonte: Elaborada pelo autor.

A análise através do CCP4MG evidencia o formato do dímero, também conhecido como 69, ou yin-yang, ilustrado nas interfaces AE e CD, além da interface tetramérica ED. É possível perceber o aspecto anti-paralelo das moléculas na interface tetramérica, com os domínios C-terminal apontando para direções opostas.

Após o refinamento final, foi realizada a determinação dos sítios de ligação aos metais (Figura 52).

Figura 52 - Monômero *SdXI*, com 2 sítios de metal. Os painéis superiores indicam a distância atômica e os resíduos envolvidos na coordenação do metal. Os painéis inferiores mostram o mapa de densidade eletrônica (Omit) em azul, corroborando a posição correta do metal.

Fonte: Elaborada pelo autor.

Os locais de ligação de metal são altamente conservados e em *SdXI* compreendem resíduos E238, E274, D302 e D345 (sítio metálico 1) e E274, H277, D313 e D315 (sítio metálico 2). Uma sobreposição de *SdXI* e outras estruturas XI conhecidas (XIs de *Streptomyces rubiginosus*, *Piromyces* sp E2 e *Actinoplanes missouriensis*, códigos 3KBN, 5NH4 e 5XIM)^{212,248-249} foi realizada para observar o perfil de alinhamento terciário e examinar a posição e orientação dos resíduos catalíticos (Figura 53).

Figura 53 - Superposição de diferentes estruturas XI. Em marrom: *SdXI*, em laranja: PirXI, ambas as classes II, enquanto *S. rubiginosus* e *A. missouriensis*, ambas de classe I, são representadas em azul-claro e roxo, respectivamente.

Fonte: Elaborada pelo autor.

Apesar das diferenças, observa-se no painel da Figura 53 a região do sítio catalítico, que não apresenta diferenças significativas na conformação dos resíduos. Esta análise evidencia a conservação do sítio catalítico mesmo entre as diferentes classes de XI.

6.3.10 Análise de Conservação

Com o intuito de compreender a questão da conservação dos resíduos entre as diferentes XIs sequenciadas até o momento, o programa CONSURF²²⁸ foi utilizado para determinar o nível de conservação desses resíduos ao longo da enzima, do mais variável ao mais conservado, visualizando essas diferenças na própria estrutura. Para as entradas utilizadas no programa, foi gerado um arquivo de alinhamento através do PSI-BLAST⁹⁶ e do Clustal Omega,¹⁰³ e a estrutura cristalográfica de *SdXI* foi utilizada como modelo de entrada de arquivo pdb. Foram obtidos diferentes graus de conservação, conforme observado na figura 54.

Figura 54 – Análise de conservação em *SdXI*. (A), (B), (C), (D) e (E) ilustram diferentes posições da molécula. Legenda indicando do resíduo mais variável (azul) ao mais conservado (vermelho).

Fonte: Elaborada pelo autor.

Como pode ser observado na Figura 54, a análise de conservação foi feita utilizando o monômero de *SdXI*, mais especificamente a cadeia A, para facilitar a visualização dos resíduos. Os resultados indicaram que houve um maior grau de conservação dentro da região de interfaces e sítio ativo. A superposição das interfaces (dímero e tetrâmero) com os resultados de conservação são mostradas na Figura 55.

Figura 55 - Monômero de *SdXI* obtido da análise CONSURF na interface dimérica AE (A) e tetramérica AC (B).

Fonte: Elaborada pelo autor.

6.3.11 Análise filogenética

Para estabelecer as relações filogenéticas entre as XIs e verificar se a distinção de classes se repete no padrão filogenético, diferentes sequências de XIs de 21 táxons foram alinhadas usando MUSCLE²³⁰ e a filogenia foi construída usando Phylo.fr²⁵² e MEGA7.²³¹ Foi criado um workflow, onde a curadoria foi realizada utilizando Gblocks. A árvore filogenética foi reconstruída usando o método de máxima verossimilhança implementado no programa PhyML (v3.1/3.0 aLRT). A confiabilidade para a ramificação interna foi avaliada pelo teste aLRT (SH-Like). A representação gráfica (Figura 56) e a edição da árvore filogenética foram realizadas com TreeDyn (v198.3).

Figura 56 - Análise filogenética molecular pelo método de máxima verossimilhança. O organismo fonte de *SdXI*, *Saccharophagus degradans*, é destacado em vermelho.

Fonte: Elaborada pelo autor.

Na figura 56 pode-se observar a árvore com a maior probabilidade de log (-8588,68). A árvore inicial para a busca heurística foram obtidas automaticamente pela aplicação dos algoritmos Neighbour-Join e BioNJ a uma matriz de distâncias pareadas estimadas usando um modelo JTT²⁵³ e então selecionando a topologia com valor de log verossimilhança superior. A árvore foi desenhada em escala, com comprimentos de ramificações medidos em número de substituições por local. A análise envolveu 21 sequências de aminoácidos, e todas as posições contendo lacunas e dados ausentes foram eliminadas. Havia um total de 331 posições no conjunto de dados final, e para o teste de filogenia, 1000 replicações de bootstraps foram realizadas.

As análises filogenéticas mostram que existe uma distinção clara entre as duas classes XI (100% de confiança). Embora não seja possível determinar onde as duas classes divergem, é claro que há evidências filogenéticas para essa separação. Além disso, observa-se também uma grande proximidade filogenética entre PirXI e *SdXI*, como pode ser observado também nas comparações estruturais. Entretanto, com essas análises, não é possível descrever a história natural de XI, evidenciar qual a classe mais primitiva, e como ocorreu sua derivação. Isso fornece insights para futuros estudos filogenéticos, visando identificar um possível ancestral, confirmar sua classe e lançar mais luz sobre o processo evolutivo dentro da família XIs.

6.4 Conclusões

O presente estudo de *SdXI* caracterizou bioquimicamente e estruturalmente uma enzima de potencial interesse biotecnológico, de papel crucial na conversão de Xilose a Xilulose, um açúcar fermentescível que pode ser aproveitado em fermentações industriais. O pH ótimo é 8.0, e sua temperatura ótima é de 60 °C. A *SdXI* apresentou uma boa termoestabilidade, mantendo cerca de 50% de sua atividade máxima após 72 horas a 50°C. A enzima se encontra em forma tetramérica em solução, com um raio de giro de 36 Å e massa molecular calculada através do SAXSMoW de 198 KDa. A estrutura cristalográfica mostrou que a *SdXI* apresenta a forma clássica de XI, com um barril TIM na região do sítio ativo, e uma alça na região C-terminal. Os dímeros são anti-paralelos, no formato “yin-yang”, e na formação dos tetrâmeros, além das interações tetraméricas, participam também as ligações cruzadas. A região do sítio ativo e das interfaces de oligomerização são bastante conservadas ao longo das XIs, e estudos filogenéticos mostraram que a *SdXI* pertence à Classe II das XIs. Apesar da extensiva caracterização, mais estudos serão necessários, principalmente na parte operacional de fermentações industriais, para melhor aproveitamento desta enzima em bioprocessos.

7 Conclusões gerais

O conjunto de trabalhos apresentado visou a caracterização de enzimas importantes na degradação de Xilano, como endoxilanases (GH10 e GH11), xilosidases (GH43) e xilose isomerase (XI). Através da cristalografia de raios X foi possível determinar a estrutura das diferentes enzimas, e no caso da GH10 estudada (*TcXyn10A*), foi possível compreender a relação entre estrutura cristalográfica e padrão de clivagem, contribuindo para o enriquecimento da literatura existente sobre o padrão de clivagem do xilano e produção de compostos de alto valor agregado, como os XOS. Além disso, foi possível também compreender a importância de sítios além do ativo na atividade enzimática de uma GH11 (*PxXyn11B*), utilizando a dinâmica molecular como ferramenta, oferecendo informações importantes para o desenvolvimento de enzimas e coquetéis enzimáticos mais eficientes para a indústria. Foi também evidenciada a importância de resíduos não-catalíticos na atividade de uma GH43 (*EcXyl43*), contribuindo para o aprofundamento do conhecimento da mesma e possibilitando uma base para a engenharia racional dessa família. A caracterização bioquímica e estrutural de uma de uma Xilose Isomerase (*SdXI*) indicou uma enzima com temperatura e pH ótimos acima do utilizado no processo de fermentação etanólica, porém também evidenciou uma boa termoestabilidade da enzima, fato que pode ser explicado pela alta compactação do tetrâmero em solução. Este é um estudo que serve de base para posteriores estudos de aplicação em processos tecnológicos.

Dada a complexidade estrutural do Xilano, é importante que as enzimas que participam da sua degradação sejam estudadas e bem compreendidas, e o presente estudo contribuiu nesse sentido, fornecendo informações sobre a formação de produtos de hidrólise e contribuindo para a literatura sobre o tema, podendo resultar em novas abordagens pela indústria biotecnológica.

REFERÊNCIAS

- 1 CAVALIER-SMITH, T. Eukaryote kingdoms: Seven or nine? **BioSystems**, v. 14, n. 3–4, p. 461–481, 1 jan. 1981.
- 2 **The Plant List**. Disponível em: <http://www.theplantlist.org/>. Acesso em: 8 set. 2020.
- 3 SØRENSEN, I.; DOMOZYCH, D.; WILLATS, W. G. T. How have plant cell walls evolved? **Plant Physiology**, v. 153, n. 2, p. 366–372, 1 jun. 2010.
- 4 HOUSTON, K. *et al.* The plant cell wall: a complex and dynamic structure as revealed by the responses of genes under stress conditions. **Frontiers in Plant Science**, 2016. DOI: 10.3389/fpls.2016.00984.
- 5 BACIC, A.; HARRIS, P. J.; STONE, B. A. Structure and function of plant cell walls. In: STUMPF, P. K.; CONN, E. E. **The biochemistry of plants**. San Diego: Academic Press, 1988. p. 297
- 6 PAULY, M.; KEEGSTRA, K. Cell-wall carbohydrates and their modification as a resource for biofuels. **Plant Journal**, v. 54, n. 4, p. 599, 1 maio 2008. DOI: 10.1111/j.1365-313X.2008.03463.x
- 7 VOGEL, J. Unique aspects of the grass cell wall. **Current Opinion in Plant Biology**, v. 11, n. 3, p. 301–307, 1 jun. 2008.
- 8 SCARLAT, N. *et al.* The role of biomass and bioenergy in a future bioeconomy: Policies and facts. **Environmental Development**, v. 15, p. 3–34, 1 jul. 2015.
- 9 SMITH, P. J. *et al.* Designer biomass for next-generation biorefineries: Leveraging recent insights into xylan structure and biosynthesis. **Biotechnology for Biofuels**, v. 10, p. 286, 30 nov. 2017. DOI: 10.1186/s13068-017-0973-z.
- 10 EVANGELISTA, D. E. **Estudos funcionais e estruturais de pectinases e xilanases com potencial para aplicações biotecnológicas**. 2017. 155 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2017.
- 11 EBRINGEROVÁ, A.; HEINZE, T. Xylan and xylan derivatives - Biopolymers with valuable properties, 1: Naturally occurring xylans structures, isolation procedures and properties. **Macromolecular Rapid Communications**, v. 21, n. 9, p. 542–556, 20 jun. 2000.
- 12 DE VRIES, R. P.; VISSER, J. Aspergillus Enzymes Involved in Degradation of Plant Cell Wall Polysaccharides. **Microbiology and Molecular Biology Reviews**, v. 65, n. 4, p. 497–522, 1 dez. 2001.
- 13 SCHELLER, H. V.; ULVSKOV, P. Hemicelluloses. **Annual Review of Plant Biology**, v. 61, n. 1, p. 263–289, 2 jun. 2010.
- 14 NAIDU, D. S.; HLANGOTHI, S. P.; JOHN, M. J. Bio-based products from xylan: A review. **Carbohydrate Polymers**, v. 179, p. 28–41, 1 jan. 2018. DOI: 10.1016/j.carbpol.2017.09.064.
- 15 POLLET, A.; DELCOUR, J. A.; COURTIN, C. M. Structural determinants of the substrate specificities of xylanases from different glycoside hydrolase families. **Critical**

Reviews in Biotechnology, v. 30, n. 3, p. 176, set. 2010. DOI: 10.3109/07388551003645599

16 **GLOBAL enzymes market to grow at 8% annually to reach USD 12.2 billion valuation by 2027**: transparency market research - bloomberg. Disponível em: <<https://www.bloomberg.com/press-releases/2020-02-06/global-enzymes-market-to-grow-at-8-annually-to-reach-usd-12-2-billion-valuation-by-2027-transparency-market-research>>. Acesso em: 1 fev. 2021.

17 **BIOFUELS market size, share, outlook & industry growth report**. Disponível em: <https://www.bccresearch.com/market-research/energy-and-resources/liquid-biofuels-outside-north-america-report.html>. Acesso em: 1 fev. 2021.

18 VÁZQUEZ, M. J. *et al.* Xylooligosaccharides: manufacture and applications. **Trends in Food Science and Technology**, v. 11, n. 11, p. 387–393, 1 nov. 2000.

19 GIBSON, G. R. *et al.* Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. **Nature Reviews Gastroenterology and Hepatology**, v. 14, p. 491, 1 ago. 2017. DOI: 10.1038/nrgastro.2017.75.

20 **PREBIOTICS**: ingredients, product and market research report to 2022. Disponível em: <https://www.bccresearch.com/market-research/food-and-beverage/prebiotics-ingredients-applications-and-global-markets.html>. Acesso em: 1 fev. 2021.

21 LOMBARD, V. *et al.* The carbohydrate-active enzymes database (CAZy) in 2013. **Nucleic Acids Research**, v. 42, p. D490, 2014. DOI: 10.1093/nar/gkt1178.

22 CAZYPEDIA CONTRIBUTORS. **Glycoside hydrolases**. Disponível em: http://www.cazypedia.org/index.php?title=Glycoside_hydrolases&oldid=11575. Acesso em: 2 fev. 2021.

23 HENRISSAT, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochemical Journal**, v. 280, n. 2, p. 309–316, 1991.

24 JØRGENSEN, H.; KRISTENSEN, J. B.; FELBY, C. Enzymatic conversion of lignocellulose into fermentable sugars: Challenges and opportunities. **Biofuels, Bioproducts and Biorefining**, v. 1, n. 2, p. 119, 1 out. 2007. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/bbb.4>. Acesso em: 15 jul. 2020.

25 BEG, Q. K. *et al.* Microbial xylanases and their industrial applications: a review. **Applied Microbiology and Biotechnology**, v. 56, p. 326, 2001. Disponível em: <https://link.springer.com/article/10.1007/s002530100704>. Acesso em: 15 jul. 2020

26 SHALLOM, D.; SHOHAM, Y. Microbial hemicellulases. **Current Opinion in Microbiology**, v. 6, n. 3, p. 219, 1 jun. 2003.

27 NGUYEN, S. T. C. *et al.* Function, distribution, and annotation of characterized cellulases, xylanases, and chitinases from CAZy. **Applied Microbiology and Biotechnology**, v. 102, n. 4, p. 1629, 1 fev. 2018.

28 BANASZAK, L. **Foundations of structural biology**. San Diego: Academic Press, 2000.

29 SMITH, R. W. Mass spectrometry. *In*: **Encyclopedia of forensic sciences**. 2nd ed. . Berlin: Elsevier Inc., 2013. p. 603–608.

- 30 SHARMA, K. S. Mass spectrometry-the early years. **International Journal of Mass Spectrometry**, v. 349–350, n. 1, p. 3–8, 1 set. 2013.
- 31 DE HOFFMANN, E.; STROOBANT, V. **Mass spectrometry: principles and applications**. 3rd. ed. [New York: Wiley, 2007].
- 32 COTTER, D. R.; SABHERWAL, S.; SCHUBERT, K. O. Proteomics for diagnostic and therapeutic blood biomarker discovery in schizophrenia and other psychotic disorders. *In: Personalized psychiatry*. Berlin: Elsevier, 2019. p. 307–317.
- 33 HUXFORD, T. X-ray crystallography. *In: Brenner's encyclopedia of genetics*. 2nd ed. Berlin: Elsevier Inc., 2013. p. 366–368.
- 34 MCPHERSON, A.; GAVIRA, J. A. Introduction to protein crystallization. **Acta Crystallographica Section F: structural biology communications**, v 70, Pt 1, p. 2–20, 2014.
- 35 CHAYEN, N. E.; SARIDAKIS, E. Protein crystallization: From purified protein to diffraction-quality crystal. **Nature Methods**, v. 5, n. 2, p. 147–153, fev. 2008.
- 36 COWTAN, K. **Crystals and symmetry**. Disponível em: <http://www.ysbl.york.ac.uk/~cowtan/workshop/CrystYork2009-Lect1-sym-for-website.pdf>. Acesso em: 5 fev. 2021.
- 37 AMEH, E. S. A review of basic crystallography and x-ray diffraction applications. **International Journal of Advanced Manufacturing Technology**, v. 105, n. 7–8, p. 3289–3302, 1 dez. 2019.
- 38 DEKKER, A. J. The crystalline state. *In: DEKKER, A. J. Solid state physics*. London: Macmillan Education UK, 1981. p. 1–31.
- 39 WEGENER, T. **Symmetry, crystal systems and Bravais lattices: physics in a nutshell**. Disponível em: <https://www.physics-in-a-nutshell.com/article/6/symmetry-crystal-systems-and-bravais-lattices>. Acesso em: 5 fev. 2021.
- 40 MERRIAM-WEBSTER. **Symmetry**. Disponível em: <https://www.merriam-webster.com/dictionary/symmetry>. Acesso em: 20 fev. 2021.
- 41 DE WOLFF, P. M. *et al.* Definition of symmetry elements in space groups and point groups report of the International Union of Crystallography Ad-Hoc Committee on the nomenclature of symmetry. **Acta Crystallographica**, v. A45, Part 7, p. 494–499, 1989.
- 42 GRENIER, B. *et al.* Crystallography: Symmetry groups and group representations. **EPJ Web of Conferences**, v. 22, n. 6, 2012. DOI: 10.1051/epjconf/20122200006.
- 43 ALCÁCER, L. **Física do estado sólido**. Disponível em: http://www.lx.it.pt/~alcacer/FisicaEstadoSolido_v4.pdf. Acesso em: 6 fev. 2021.
- 44 PURCELL, E. M.; MORIN, D. J. **Electricity and magnetism**. 3rd. ed. Cambridge : Cambridge University Press, 2013.
- 45 UCAR CENTER FOR SCIENCE EDUCATION. **Electromagnetic (EM) spectrum** | Disponível em: <https://scied.ucar.edu/em-spectrum>. Acesso em: 7 fev. 2021.
- 46 BLUNDELL, T. L.; JOHNSON, L. **Protein crystallography**. Cambridge: Academic Press, 1976.

- 47 CROMER, D. T.; MANN, J. B. X-ray scattering factors computed from numerical Hartree–Fock wave functions. **Acta Crystallographica Section A**, v. 24, n. 2, p. 321–324, 1 mar. 1968.
- 48 MCPHERSON, A. **Introduction to macromolecular crystallography**. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2009.
- 49 EDITORS OF ENCYCLOPEDIA BRITANNICA. **Bragg law**. Disponível em: <https://www.britannica.com/science/Bragg-law>. Acesso em: 8 fev. 2021.
- 50 DICTIONARY OF CRYSTALLOGRAPHY. **Structure factor**. Disponível em: https://dictionary.iucr.org/Structure_factor. Acesso em: 8 fev. 2021.
- 51 CRISTALOGRAFÍA. **Resolución estructural**. dispersión anómala. Disponível em: https://www.xtal.iqfr.csic.es/Cristalografia/parte_07_4.html. Acesso em: 16 fev. 2021.
- 52 MCCOY, A. J. *et al.* Phaser crystallographic software. **Journal of Applied Crystallography**, v. 40, n. 4, p. 658–674, 13 jul. 2007.
- 53 AMBROSIO, A. L. B.; FRANCHINI, K. G. Cristalografia macromolecular: a biologia sob a ótica dos raios X. **Ciência e Cultura**, v. 69, n. 3, p. 29–36, jul. 2017.
- 54 DICTIONARY OF CRYSTALLOGRAPHY. **Refinement**. Disponível em: <https://dictionary.iucr.org/Refinement>. Acesso em: 9 fev. 2021.
- 55 ENDERS, C. K. Maximum likelihood estimation. *In*: **Encyclopedia of statistics in behavioral science**. Chichester, UK: John Wiley & Sons, Ltd, 2005. v. 3, p. 1164–1170
- 56 WLODAWER, A. *et al.* Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. **FEBS Journal**, v. 275, n. 1, p. 1–21, 2008.
- 57 DICTIONARY OF CRYSTALLOGRAPHY. **Free R factor**. Disponível em: https://dictionary.iucr.org/Free_R_factor. Acesso em: 15 fev. 2021
- 58 DIEDERICH, K.; KARPLUS, P. A. Improved R-factors for diffraction data analysis in macromolecular crystallography. **Nature Structural Biology**, v. 4, n. 4, p. 269–275, abr. 1997.
- 59 SUN, Z. *et al.* Utility of B-factors in protein science: interpreting rigidity, flexibility, and internal motion and engineering thermostability. **Chemical Reviews**, v. 119, n. 3, p. 1626–1665, 2019.
- 60 DICTIONARY OF CRYSTALLOGRAPHY. **Mosaic crystal**. Disponível em: https://dictionary.iucr.org/Mosaic_crystal. Acesso em: 16 fev. 2021.
- 61 BARBOSA, L. R. S. *et al.* The importance of protein-protein interactions on the pH-induced conformational changes of bovine serum albumin: a small-angle x-ray scattering study. **Biophysical Journal**, v. 98, n. 1, p. 147–157, 6 jan. 2010.
- 62 CHALAL, M. *et al.* SAXS investigation of the effect of temperature on the multiscale structure of a macroporous poly(N-isopropylacrylamide) gel. **Macromolecules**, v. 43, n. 4, p. 2009–2017, 2010.
- 63 ROBINSON, P. K. Enzymes: principles and biotechnological applications. **Essays in**

Biochemistry, v. 59, p. 1–41, 15 nov. 2015. DOI: 10.1042/bse0590001.

64 DE ARAÚJO, E. A. **Explorando as relações entre estrutura e função das hidrolases de glicosídeo das famílias 9, 48 e 74**. 2018. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2018.

65 GLATTER, O. A New Method for the Evaluation of Small-Angle Scattering Data. **Journal of Applied Crystallography**, v. 10, pt. 5, p. 415–421, 1977.

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

67 FRANKE, D. *et al.* ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. **Journal of Applied Crystallography**, v. 50, n. pt 4, p. 1212–1225, 2017.

68 GLATTER, O.; KRATKY, O. **Small Angle X-ray Scattering**. London: Academic Press, 1982.

69 RHODES, G. Other Diffraction Methods. *In*: RHODES, G. **Crystallography made crystal clear**. Amsterdam: Elsevier, 2006. p. 211–235.

70 PIIADOV, V. *et al.* SAXSMoW 2.0: Online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale. **Protein Science**, v. 28, n. 2, p. 454–463, 1 fev. 2019.

71 MERTENS, H. D. T.; SVERGUN, D. I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. **Journal of Structural Biology**, v. 172, n. 1, p. 128–141, 1 out. 2010.

72 RAMBO, R. P.; TAINER, J. A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. **Biopolymers**, v. 95, n. 8, p. 559–571, 1 ago. 2011.

73 SVERGUN, D. I.; PETOUKHOV, M. V.; KOCH, M. H. J. Determination of domain structure of proteins from x-ray solution scattering. **Biophysical Journal**, v. 80, n. 6, p. 2946–2953, 2001.

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

75 ROY, K.; KAR, S.; DAS, R. N. Computational chemistry. *In*: ROY, K.; KAR, S.; DAS, R. N. **Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment**. Amsterdam: Elsevier, 2015. p. 151–189.

76 PITMAN, M. R.; MENZ, R. I. Methods for protein homology modelling. **Applied Mycology and Biotechnology**, v. 6, p. 37–59, 2006. DOI: 10.1016/S1874-5334(06)80005-5.

77 CHOW, E. *et al.* New technologies for molecular dynamics simulations. **Comprehensive Biophysics**, v. 9, p. 86–104, 2012. DOI: 10.1016/B978-0-12-374920-8.00908-5.

78 POLANSKI, J. Chemoinformatics. *In*: WALCZAK, B. (ed.). **Comprehensive Chemometrics**. Berlin: Elsevier, 2009. p. 459–506.

- 79 VERLI, H. Dinâmica Molecular. *In*: VERLI, H. (org.). **Bioinformática: da biologia à flexibilidade moleculares**. São Paulo: Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq, 2014. p. 173–187.
- 80 GUVENCH, O.; MACKERELL, A. D. Comparison of protein force fields for molecular dynamics simulations. **Methods in Molecular Biology**, v. 443, p. 63–88, 2008. DOI: 10.1007/978-1-59745-177-2_4.
- 81 VANOMMESLAEGHE, K.; GUVENCH, O.; MACKERELL, A. D. Molecular mechanics. **Current Pharmaceutical Design**, v. 20, n. 20, p. 3281–3292, 4 jun. 2014.
- 82 BARSEGOV, V. **Part I. Strategy for running molecular dynamics simulations**. Disponível em: <http://faculty.uml.edu/vbarsegov/teaching/bioinformatics/lectures/MDSimulationsModified.pdf>. Acesso em: 19 fev. 2021.
- 83 ULINE, M.; CORTI, D. Molecular dynamics at constant pressure: allowing the system to control volume fluctuations via a “Shell” particle. **Entropy**, v. 15, n. 12, p. 3941–3969, 23 set. 2013.
- 84 JO, S. *et al.* CHARMM-GUI: A web-based graphical user interface for CHARMM. **Journal of Computational Chemistry**, v. 29, n. 11, p. 1859–1865, ago. 2008.
- 85 **PSF FILES**. Disponível em: <https://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node23.html>. Acesso em: 19 fev. 2021.
- 86 JORGENSEN, W. L. *et al.* Comparison of simple potential functions for simulating liquid water. **Journal of Chemical Physics**, v. 79, n. 2, p. 926–935, 15 jul. 1983.
- 87 WATANABE, K. *et al.* *Thermobacillus composti* sp. nov., a moderately thermophilic bacterium isolated from a composting reactor. **International Journal of Systematic and Evolutionary Microbiology**, v. 57, n. 7, p. 1473–1477, jul. 2007.
- 88 PELL, G. *et al.* The mechanisms by which family 10 glycoside hydrolases bind decorated substrates. **Journal of Biological Chemistry**, v. 279, n. 10, p. 9597–9605, 5 mar. 2004.
- 89 CAMILO, C. M.; POLIKARPOV, I. High-throughput cloning, expression and purification of glycoside hydrolases using Ligation-Independent Cloning (LIC). **Protein Expression and Purification**, v. 99, p. 35–42, 2014.
- 90 LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, v. 227, n. 5259, p. 680–685, 1970.
- 91 SEPULCHRO, A. G. V. *et al.* Transformation of xylan into value-added biocommodities using *Thermobacillus composti* GH10 xylanase. **Carbohydrate Polymers**, v. 247, p. 116714, 2020. DOI: 10.1016/j.carbpol.2020.116714.
- 92 NIEDERMEYER, T. H. J.; STROHALM, M. mMass as a Software Tool for the Annotation of Cyclic Peptide Tandem Mass Spectra. **PLoS ONE**, v. 7, n. 9, p. e44913, 13 set. 2012.
- 93 GUIMARÃES, B. G. *et al.* The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS. **Journal of Synchrotron Radiation**, v. 16, n. 1, p. 69–75, 2008.

- 94 KABSCH, W. XDS. **Acta Crystallographica Section D: biological crystallography**, v. 66, n. 2, p. 125–132, 1 feb. 2010.
- 95 EVANS, P. R.; MURSHUDOV, G. N. How good are my data and what is the resolution? **Acta Crystallographica Section D: biological crystallography**, v. 69, n. 7, p. 1204–1214, 13 jul. 2013.
- 96 ALTSCHUL, S. F. *et al.* Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. **Nucleic Acids Research** v. 25, n. 17, p. 3389–3402, 1 set. 1997.
- 97 MCCOY, A. J. *et al.* Phaser crystallographic software. **Journal of Applied Crystallography**, v. 40, n. 4, p. 658–674, 13 jul. 2007.
- 98 TERWILLIGER, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. **Acta Crystallographica Section D: biological crystallography**, v. 64, pt 1, p. 61–69, 4 dez. 2007.
- 99 EMSLEY, P. *et al.* Features and development of Coot. **Acta Crystallographica Section D: biological crystallography**, v. 66, n. 4, p. 486–501, 24 mar. 2010.
- 100 AFONINE, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. **Acta Crystallographica Section D: biological crystallography**, v. 68, n. 4, p. 352–367, 16 abr. 2012.
- 101 CHEN, V. B. *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D: Biological Crystallography**, v. 66, n. 1, p. 12–21, 2010.
- 102 FUJIMOTO, Z. *et al.* Crystal Structures of Decorated Xylooligosaccharides Bound to a Family 10 Xylanase from *Streptomyces olivaceoviridis* E-86. **Journal of Biological Chemistry**, v. 279, n. 10, p. 9606–9614, 5 mar. 2004.
- 103 SIEVERS, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**, v. 7, n. 1, p. 539, jan. 2011.
- 104 ROBERT, X.; GOUET, P. Deciphering key features in protein structures with the new ENDscript server. **Nucleic Acids Research**, v. 42, n. W1, p. W320–W324, jul. 2014.
- 105 BANNER, D. W. *et al.* Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution: Using amino acid sequence data. **Nature**, v. 255, n. 5510, p. 609–614, 1975.
- 106 BERMAN, H. M. *et al.* **The protein data bank** **nucleic acids research**. Oxford: Oxford University Press, 2000.
- 107 PETTERSEN, E. F. *et al.* UCSF chimera - a visualization system for exploratory research and analysis. **Journal of Computational Chemistry**, v. 25, p. 1605–1612, 2004.
- 108 MACLEOD, A. M. *et al.* The acid/base catalyst in the exoglucanase/xylanase from *Cellulomonas fimi* is glutamic acid 127: evidence from detailed kinetic studies of mutants. **Biochemistry**, v. 33, n. 20, p. 6371–6376, 1 maio 1994.
- 109 TULL, D. *et al.* Glutamic acid 274 is the nucleophile in the active site of a “retaining” exoglucanase from *Cellulomonas fimi*. **Journal of Biological Chemistry**, v. 266, n. 24, p.

15621–15625, 25 ago. 1991.

110 KOSHLAND, D. E. Stereochemistry and the mechanism of enzymatic reactions. **Biological Reviews**, v. 28, n. 4, p. 416–436, 1 nov. 1953.

111 TERWISSCHA VAN SCHELTINGA, A. C. *et al.* Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and X-ray structure of a complex with allosamidin: evidence for substrate assisted catalysis. **Biochemistry**, v. 34, n. 48, p. 15619–15623, 1995.

112 VARKI, A. *et al.* Symbol nomenclature for graphical representations of glycans. **Glycobiology**, v. 25, n. 12, p. 1323–1324, 1 dez. 2015.

113 NEELAMEGHAM, S. *et al.* Updates to the symbol nomenclature for glycans guidelines. **Glycobiology**, v. 29, n. 9, p. 620–624, 1 set. 2019.

114 CHARNOCK, S. J. *et al.* The topology of the substrate binding clefts of glycosyl hydrolase family 10 xylanases are not conserved. **Journal of Biological Chemistry**, v. 273, n. 48, p. 32187–32199, 27 nov. 1999.

115 VARDAKOU, M. *et al.* A family 10 thermoascus aurantiacus xylanase utilizes arabinose decorations of xylan as significant substrate specificity determinants. **Journal of Molecular Biology**, v. 352, n. 5, p. 1060–1067, 7 out. 2005.

116 DAVIES, G.; HENRISSAT, B. Structures and mechanisms of glycosyl hydrolases. **Structure**, v. 3, n. 9, p. 853–859, 1 set. 1995.

117 COCKBURN, D. *et al.* Analysis of surface binding sites (SBSs) in carbohydrate active enzymes with focus on glycoside hydrolase families 13 and 77 - a mini-review. **Biologia**, v. 69, p. 705–712. 2014. DOI: 10.2478/s11756-014-0373-9.

118 CUYVERS, S. *et al.* Secondary substrate binding strongly affects activity and binding affinity of *Bacillus subtilis* and *Aspergillus niger* GH11 xylanases. **FEBS Journal**, v. 278, n. 7, p. 1098–1111, 1 abr. 2011.

119 VANDERMARLIERE, E. *et al.* Crystallographic analysis shows substrate binding at the -3 to +1 active-site subsites and at the surface of glycoside hydrolase family 11 endo-1,4- β -xylanases. **Biochemical Journal**, v. 410, n. 1, p. 71–79, 15 fev. 2008.

120 LUDWICZEK, M. L. *et al.* A Secondary Xylan-binding Site Enhances the Catalytic Activity of a Single-domain Family 11 Glycoside Hydrolase. **Journal of Molecular Biology**, v. 373, n. 2, p. 337–354, 2007.

121 GHIO, S. *et al.* *Paenibacillus xylanivorans* sp. nov., a xylan-degrading bacterium isolated from decaying forest soil. **International Journal of Systematic and Evolutionary Microbiology**, v. 69, n. 12, p. 3818–3823, 1 dez. 2019.

122 -. GHIO, S. *et al.* *Paenibacillus* sp. A59 GH10 and GH11 extracellular endoxylanases: Application in biomass bioconversion. **Bioenergy Research**, v. 11, n. 1, p. 174–190, 2018

123 RIBEIRO, L. F. *et al.* A xylose-stimulated xylanase-xylose binding protein chimera created by random nonhomologous recombination. **Biotechnology for Biofuels**, v. 9, n. 1, p. 119, 6 dez. 2016.

124 BORGES, T. A. *et al.* Biochemical characterization of an endoxylanase from

Pseudozyma brasiliensis sp. nov. strain GHG001 isolated from the intestinal tract of Chrysomelidae larvae associated to sugarcane roots. **Process Biochemistry**, v. 49, n. 1, p. 77–83, 2014.

125 POLIZELI, T. M. *et al.* xylanases from fungi: properties and industrial applications. **Applied Microbiology and Biotechnology**, v. 67, n. 5, p. 577–591, 2005.

126 BRIGANTI, L. *et al.* Structural and molecular dynamics investigations of ligand stabilization via secondary binding site interactions in *Paenibacillus xylanivorans* GH11 xylanase. **Computational and Structural Biotechnology Journal**, v. 19, p. 1557–1566, 1 jan. 2021.

127 LONG, F. *et al.* BALBES: a molecular-replacement pipeline. **Acta Crystallographica Section D: biological crystallography**, v. 64, pt 1, p. 125–132, 2007. DOI: 10.1107/S0907444907050172.

128 WAN, Q. *et al.* X-ray crystallographic studies of family 11 xylanase Michaelis and product complexes: Implications for the catalytic mechanism. **Acta Crystallographica Section D: biological crystallography**, v. 70, n. 1, p. 11–23, 1 jan. 2014.

129 CASE, D. A. *et al.* The amber biomolecular simulation programs. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1668–1688, 2005.

130 SALOMON-FERRER, R. *et al.* Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. **Journal of Chemical Theory and Computation**, v. 9, n. 9, p. 3878–3888, set. 2013.

131 LE GRAND, S.; GÖTZ, A. W.; WALKER, R. C. SPFP: Speed without compromise - a mixed precision model for GPU accelerated molecular dynamics simulations. **Computer Physics Communications**, v. 184, n. 2, p. 374–380, 1 fev. 2013.

132 GORDON, J. C. *et al.* H++: A server for estimating pKas and adding missing hydrogens to macromolecules. **Nucleic Acids Research**, v. 33, n. SUPPL. 2, p. W368-71, jul. 2005.

133 MAIER, J. A. *et al.* ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. **Journal of Chemical Theory and Computation**, v. 11, n. 8, p. 3696–3713, ago. 2015.

134 KIRSCHNER, K. N. *et al.* GLYCAM06: a generalizable biomolecular force field. carbohydrates. **Journal of Computational Chemistry**, v. 29, n. 4, p. 622–655, 2008.

135 HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: Visual molecular dynamics. **Journal of Molecular Graphics**, v. 14, n. 1, p. 33–38, 1996.

136 ROE, D. R.; CHEATHAM, T. E. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. **Journal of Chemical Theory and Computation**, v. 9, n. 7, p. 3084–3095, 2013.

137 CAZYPEDIA CONTRIBUTORS. **Glycoside hydrolase family 11**. Disponível em: https://www.cazypedia.org/index.php/Glycoside_Hydrolase_Family_11. Acesso em: 5 out. 2020.

138 LASKOWSKI, R. A. *et al.* PDBsum: structural summaries of PDB entries. **Protein Science**, v. 27, n. 1, p. 129–134, 1 jan. 2018.

- 139 MURAKAMI, M. T. *et al.* Correlation of temperature induced conformation change with optimum catalytic activity in the recombinant G/11 xylanase A from *Bacillus subtilis* strain 168 (1A1). **FEBS Letters**, v. 579, n. 28, p. 6505–6510, 21 nov. 2005.
- 140 HAN, N. *et al.* Improving the thermostability of a fungal GH11 xylanase via site-directed mutagenesis guided by sequence and structural analysis. **Biotechnology for Biofuels**, v. 10, n. 1, p. 133, 23 dez. 2017.
- 141 HAKULINEN, N. *et al.* Three-dimensional structures of thermophilic β -1,4-xylanases from *Chaetomium thermophilum* and *Nonomuraea flexuosa*: comparison of twelve xylanases in relation to their thermal stability. **European Journal of Biochemistry**, v. 270, n. 7, p. 1399–1412, 1 abr. 2003.
- 142 CHENG, Y. S. *et al.* Structural analysis of a glycoside hydrolase family 11 xylanase from *Neocallimastix patriciarum*: Insights into the molecular basis of a thermophilic enzyme. **Journal of Biological Chemistry**, v. 289, n. 16, p. 11020–11028, 2014.
- 143 LAMMERTS VAN BUEREN, A. *et al.* Three-dimensional structure of a thermophilic family GH11 xylanase from *Thermobifida fusca*. **Acta Crystallographica Section F: structural biology and crystallization communications**, v. 68, n. 2, p. 141–144, 25 fev. 2012.
- 144 GAGNÉ, D. *et al.* Ligand binding enhances millisecond conformational exchange in xylanase B2 from *Streptomyces lividans*. **Biochemistry**, v. 55, n. 30, p. 4184–4196, 2016.
- 145 VARDAKOU, M. *et al.* Understanding the structural basis for substrate and inhibitor recognition in eukaryotic GH11 xylanases. **Journal of Molecular Biology**, v. 375, n. 5, p. 1293–1305, 1 fev. 2008.
- 146 SCHRÖDER, S. P. *et al.* Dynamic and functional profiling of xylan-degrading enzymes in aspergillus secretomes using activity-based probes. **ACS Central Science**, v. 5, n. 6, p. 1067–1078, 26 jun. 2019.
- 147 MCCARTER, J. D.; STEPHEN WITHERS, G. Mechanisms of enzymatic glycoside hydrolysis. **Current Opinion in Structural Biology**, v. 4, n. 6, p. 885–892, 1 jan. 1994.
- 148 GEBLER, J. *et al.* Stereoselective hydrolysis catalyzed by related β -1,4-glucanases and β -1,4-xylanases. **Journal of Biological Chemistry**, v. 267, n. 18, p. 12559–12561, 25 jun. 1992.
- 149 MADAN, B.; LEE, S. G. Sequence and structural features of subsite residues in GH10 and GH11 xylanases. **Biotechnology and Bioprocess Engineering**, v. 23, n. 3, p. 311–318, 1 jun. 2018.
- 150 WU, X. *et al.* The contribution of specific subsites to catalytic activities in active site architecture of a GH11 xylanase. **Applied Microbiology and Biotechnology**, v. 104, n. 20, p. 8735–8745, 1 out. 2020.
- 151 NASCIMENTO, A. S. **AmberEnergy++**. Disponível em: <https://github.com/alessandronascimento/amberenergy>. Acesso em: 15 fev. 2021
- 152 MARTÍNEZ, L. Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. **PLOS ONE**, v. 10, n. 3, p. e0119264, 27 mar. 2015.

- 153 MILLER, B. R. *et al.* MMPBSA.py: an efficient program for end-state free energy calculations. **Journal of Chemical Theory and Computation**, v. 8, n. 9, p. 3314–3321, 11 set. 2012.
- 154 PAËS, G. *et al.* Thumb-loops up for catalysis: a structure/function investigation of a functional loop movement in a GH11 xylanase. **Computational and Structural Biotechnology Journal**, v. 1, n. 2, p. e201207001, 1 jul. 2012.
- 155 PAËS G. *et al.* New insights into the role of the thumb-like loop in GH-11 xylanases. **Protein Engineering, Design and Selection**, v. 20, n. 1, p. 15–23, 1 jan. 2007.
- 156 IGLESIAS-FERNÁNDEZ, J. *et al.* The complete conformational free energy landscape of β -xylose reveals a two-fold catalytic itinerary for β -xylanases. **Chemical Science**, v. 6, n. 2, p. 1167–1177, 2015.
- 157 HENRISSAT, B.; DAVIES, G. Structural and sequence-based classification of glycoside hydrolases. **Current Opinion in Structural Biology**, v. 7, n. 5, p. 637–644, 1 out. 1997.
- 158 SHALLOM, D. *et al.* Biochemical characterization and identification of the catalytic residues of a family 43 β -D-xylosidase from *Geobacillus stearothermophilus* T-6. **Biochemistry**, v. 44, n. 1, p. 387–397, 11 jan. 2005.
- 159 FLIPPHI, M. J. A. *et al.* Cloning of the *Aspergillus niger* gene encoding α -l-arabinofuranosidase A. **Applied Microbiology and Biotechnology**, v. 39, n. 3, p. 335–340, jun. 1993.
- 160 FLIPPHI, M. J. A. *et al.* Molecular cloning, expression and structure of the endo-1,5- α -l-arabinase gene of *Aspergillus niger*. **Applied Microbiology and Biotechnology**, v. 40, n. 2–3, p. 318–326, nov. 1993.
- 161 MCKIE, V. A. *et al.* Arabinanase A from *Pseudomonas fluorescens* subsp. cellulosa exhibits both an endo- and an exo- mode of action. **Biochemical Journal**, v. 323, n. 2, p. 547–555, 15 abr. 1997.
- 162 MEWIS, K. *et al.* Dividing the large glycoside hydrolase family 43 into subfamilies: a motivation for detailed enzyme characterization. **Applied and Environmental Microbiology**, v. 82, n. 6, p. 1686–1692, 2016.
- 163 SMAALI, I.; RÉMOND, C.; O'DONOHUE, M. J. Expression in *Escherichia coli* and characterization of β -xylosidases GH39 and GH-43 from *Bacillus halodurans* C-125. **Applied Microbiology and Biotechnology**, v. 73, n. 3, p. 582–590. 2006. DOI: 10.1007/s00253-006-0512-5.
- 164 DAVIN-REGLI, A.; PAGÈS, J. M. *Enterobacter aerogenes* and *Enterobacter cloacae*; Versatile bacterial pathogens confronting antibiotic treatment. **Frontiers in Microbiology**, v. 6, p. 392. 2015. DOI:10.3389/fmicb.2015.00392.
- 165 MEZZATESTA, M. L.; GONA, F.; STEFANI, S. *Enterobacter cloacae* complex: Clinical impact and emerging antibiotic resistance. **Future Microbiology**, v. 7, n. 7, p. 887–902. 2012. DOI: 10.2217/fmb.12.61.
- 166 CAMPOS, E. *et al.* Purification and characterization of a GH43 β -xylosidase from

Enterobacter sp. identified and cloned from forest soil bacteria. **Microbiological Research**, v. 169, n. 2–3, p. 213–220, 2014.

167 ONTAÑÓN, O. M. *et al.* EcXyl43 β -xylosidase: molecular modeling, activity on natural and artificial substrates, and synergism with endoxylanases for lignocellulose deconstruction. **Applied Microbiology and Biotechnology**, v. 102, n. 16, p. 6959–6971, 1 ago. 2018.

168 FALCK, P. *et al.* Characterization of a family 43 β -xylosidase from the xylooligosaccharide utilizing putative probiotic Weissella sp. strain 92. **Glycobiology**, v. 26, n. 2, p. 193–202, 16 mar. 2015.

169 BENGTSSON, O.; HAHN-HÄGERDAL, B.; GORWA-GRAUSLUND, M. F. Xylose reductase from *Pichia stipitis* with altered coenzyme preference improves ethanolic xylose fermentation by recombinant *Saccharomyces cerevisiae*. **Biotechnology for Biofuels**, v. 2, n. 1, p. 1–10, 5 maio 2009.

170 KIM, D.; WOO, H. M. Deciphering bacterial xylose metabolism and metabolic engineering of industrial microorganisms for use as efficient microbial cell factories. **Applied Microbiology and Biotechnology**, v. 102, n. 22, p. 9471–9480. 2018. DOI: 10.1007/s00253-018-9353-2.

171 JEFFRIES, T. W. Engineering yeasts for xylose metabolism. **Current Opinion in Biotechnology**, v. 17, n. 3, p. 320–326. 2006. DOI: 10.1016/j.copbio.2006.05.008.

172 BRÜX, C. *et al.* 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**, v. 359, n. 1, p. 97–109, 26 maio 2006.

173 ZHENG, H. *et al.* Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. **Nature Protocols**, v. 9, n. 1, p. 156–170, jan. 2014.

174 . ZHENG, H. *et al.* CheckMyMetal: a macromolecular metal-binding validation tool. **Acta Crystallographica Section D: structural biology**, v. 73, n. 3, p. 223–233, 22 fev. 2017.

175 BERMAN, H. M. *et al.* **The protein data bank nucleic acids research**. 2000. Disponível em: <http://www.rcsb.org/pdb/status.html>. Acesso em: 24 fev. 2021

176 HOLM, L. DALI and the persistence of protein shape. **Protein Science**, v. 29, n. 1, p. 128–140, 5 jan. 2020.

177 HONG, S. *et al.* Structure-based protein engineering of bacterial β -xylosidase to increase the production yield of xylobiose from xylose. **Biochemical and Biophysical Research Communications**, v. 501, n. 3, p. 703–710, 27 jun. 2018.

178 BRUNZELLE, J. S. *et al.* Structure of the two-subsite β -d-xylosidase from *Selenomonas ruminantium* in complex with 1,3-bis[tris(hydroxymethyl)methylamino]propane. **Archives of Biochemistry and Biophysics**, v. 474, n. 1, p. 157–166, 1 jun. 2008.

179 DOS SANTOS MUNIZ, H.; NASCIMENTO, A. S. Ligand- and receptor-based docking with LiBELa. **Journal of Computer-Aided Molecular Design**, v. 29, n. 8, p. 713–723, ago. 2015.

180 WOODS GROUP. **Glycam web**. Disponível em: <http://glycam.org>. Acesso em: 26 fev. 2021.

- 181 TIAN, C. *et al.* Ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. **Journal of Chemical Theory and Computation**, v. 16, n. 1, p. 528–552, 14 jan. 2020.
- 182 MARTÍNEZ, L. **MDAnalysis**. version 19-340. Disponível em: <http://leandro.iqm.unicamp.br/mdanalysis/>. Acesso em: 26 fev. 2021.
- 183 MARTÍNEZ, L.; SHIMIZU, S. Molecular interpretation of preferential interactions in protein solvation : a solvent-shell perspective by means of minimum-distance distribution functions. **Journal of Chemical Theory and Computation**, v. 12, p. 6358-6372, 8 nov. 2017. DOI: 10.1021/acs.jctc.7b00599.
- 184 EPTING, K. L. *et al.* Influence of divalent cations on the structural thermostability and thermal inactivation kinetics of class II xylose isomerases. **FEBS Journal**, v. 272, n. 6, p. 1454–1464, 3 mar. 2005.
- 185 ADALBERTO *et al.* Effect of divalent metal ions on the activity and stability of β -galactosidase isolated from *Kluyveromyces lactis*. **Journal of Basic and Applied Pharmaceutical Sciences**, v. 31, n. 3, p. 143–150, 2010.
- 186 ZENG, J. *et al.* Effects of metal ions on stability and activity of hyperthermophilic pyrolysins and further stabilization of this enzyme by modification of a Ca^{2+} -binding site. **Applied and Environmental Microbiology**, v. 80, n. 9, p. 2763–2772, 2014.
- 187 ROHMAN, A.; DIJKSTRA, B. W.; PUSPANINGSIH, N. N. T. β -Xylosidases: structural diversity, catalytic mechanism, and inhibition by monosaccharides. **International Journal of Molecular Sciences**, v. 20, n. 22, p. 5524, 6 nov. 2019.
- 188 BARKER, I. J.; PETERSEN, L.; REILLY, P. J. Mechanism of xylobiose hydrolysis by GH43 β -xylosidase. **Journal of Physical Chemistry B**, v. 114, n. 46, p. 15389–15393, 25 nov. 2010.
- 189 NURIZZO, D. *et al.* *Cellvibrio japonicus* α -l-arabinanase 43a has a novel five-blade β -propeller fold. **Nature Structural Biology**, v. 9, n. 9, p. 665–668, 2002.
- 190 JORDAN, D. B.; LI, X.-L. Variation in relative substrate specificity of bifunctional β -d-xylosidase/ α -l-arabinofuranosidase by single-site mutations: Roles of substrate distortion and recognition. **Biochimica et Biophysica Acta (BBA) - proteins and proteomics**, v. 1774, n. 9, p. 1192–1198, 6 set. 2007.
- 191 PONS, T. *et al.* 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 bioinformatics**, v. 54, n. 3, p. 424–432, 15 fev. 2004.
- 192 MITSUHASHI, S.; LAMPEN, J. O. Conversion of D-xylose to D-xylulose in extracts of *Lactobacillus pentosus*. **Journal of Biological Chemistry**, v. 204, n. 2, p. 1011–1018, 1 out. 1953.
- 193 JOKELA, J.; PASTINEN, O.; LEISOLA, M. Isomerization of pentose and hexose sugars by an enzyme reactor packed with cross-linked xylose isomerase crystals. **Enzyme and Microbial Technology**, v. 31, n. 1–2, p. 67–76, 1 jul. 2002.
- 194 YAMANAKA, K. D-Xylose Isomerase from *Lactobacillus brevis*. **Methods in Enzymology**, v. 41, n. C, p. 466–471, 1 jan. 1975.

- 195 PASTINEN, O.; SCHOEMAKER, H. E.; LEISOLA, M. Xylose isomerase catalysed novel hexose epimerization. **Biocatalysis and Biotransformation**, v. 17, n. 5, p. 393–400, 1999.
- 196 KARIMAKI, J. *et al.* Engineering the substrate specificity of xylose isomerase. **Protein Engineering Design and Selection**, v. 17, n. 12, p. 861–869, 16 fev. 2005.
- 197 VUOLANTO, A. *et al.* C-2 epimer formation of tetrose, pentose and hexose sugars by xylose isomerase. **Biocatalysis and Biotransformation**, v. 20, n. 4, p. 235–240, 2002.
- 198 BHOSALE, S. H.; RAO, M. B.; DESHPANDE, V. V. Molecular and industrial aspects of glucose isomerase. **Microbiological Reviews**, v. 60, n. 2, p. 280–300. 1996.
- 199 MÜHLING, M.; JOINT, I.; WILLETTS, A. J. The biodiscovery potential of marine bacteria: An investigation of phylogeny and function. **Microbial Biotechnology**, v. 6, n. 4, p. 361–370, 2013.
- 200 NUMATA, K.; MORISAKI, K. Screening of marine bacteria to synthesize polyhydroxyalkanoate from lignin: Contribution of lignin derivatives to biosynthesis by *oceanimonas doudoroffii*. **ACS Sustainable Chemistry and Engineering**, v. 3, n. 4, p. 569–573, 6 abr. 2015.
- 201 TRINCONE, A. Enzymatic processes in marine biotechnology. **Marine Drugs**, v. 15, n. 4, p. 93. DOI: 10.3390/md15040093.
- 202 RAO, T. E.; IMCHEN, M.; KUMAVATH, R. Marine enzymes: production and applications for human health. **Advances in Food and Nutrition Research**, v. 80, p. 149–163, 2017. DOI: 10.1016/bs.afnr.2016.11.006.
- 203 TAYLOR, L. E. *et al.* Complete cellulase system in the marine bacterium *Saccharophagus degradans* strain 2-40T. **Journal of Bacteriology**, v. 188, n. 11, p. 3849–3861, jun. 2006.
- 204 ANDRYKOVITCH, G.; MARX, I. Isolation of a new polysaccharide-digesting bacterium from a salt marsh. **Applied and Environmental Microbiology**, v. 54, n. 4, p. 1061–1062, 1988.
- 205 EKBORG, N. A. *et al.* *Saccharophagus degradans* gen. nov., sp. nov., a versatile marine degrader of complex polysaccharides. **International Journal of Systematic and Evolutionary Microbiology**, v. 55, n. 4, p. 1545–1549, 1 jul. 2005.
- 206 HOWARD, M. B. *et al.* Genomic analysis and initial characterization of the chitinolytic system of *Microbulbifer degradans* strain 2-40. **Journal of Bacteriology**, v. 185, n. 11, p. 3352–3360, 1 jun. 2003.
- 207 NEUHOFF, V.; STAMM, R.; EIBL, H. Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: a systematic analysis. **Electrophoresis**, v. 6, n. 9, p. 427–448, 1 jan. 1985.
- 208 WIEDEMANN, C.; BELLSTEDT, P.; GÖRLACH, M. CAPITO - a web server-based analysis and plotting tool for circular dichroism data. **Bioinformatics**, v. 29, n. 14, p. 1750–1757, 15 jul. 2013.
- 209 DISCHE, Z.; BORENFREUND, E. A new spectrophotometric method for the detection

- and determination of keto sugars and trioses. **Journal of Biological Chemistry**, v. 192, n. 2, p. 583-587. 1951.
- 210 SCHUCK, P. *et al.* Size-distribution analysis of proteins by analytical ultracentrifugation: Strategies and application to model systems. **Biophysical Journal**, v. 82, n. 2, p. 1096-1111, 2002.
- 211 BATEMAN, A. *et al.* UniProt: the universal protein knowledgebase in 2021. **Nucleic Acids Research**, v. 49, n. D1, p. D480-D489, 8 jan. 2021.
- 212 BIASINI, M. *et al.* SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. **Nucleic Acids Research**, v. 42, n. W1, 1 jul. 2014. DOI: 10.1093/nar/gku340.
- 213 LEE, M. *et al.* Metal dependence of the xylose isomerase from *piromyces* sp. E2 explored by activity profiling and protein crystallography. **Biochemistry**, v. 56, n. 45, p. 5991-6005, 14 nov. 2017.
- 214 KELLEY, L. A. *et al.* The Phyre2 web portal for protein modeling, prediction and analysis. **Nature Protocols**, v. 10, n. 6, p. 845-858, 30 jun. 2015.
- 215 LLOYD, L. F. *et al.* Crystallization and preliminary x-ray diffraction studies of xylose isomerase from thermoanaerobacterium thermosulfurigenes strain 4B. **Journal of Molecular Biology**, v. 240, n. 5, p. 504-506, 28 jul. 1994.
- 216 HAMMERSLEY, A. P. **FIT2D**: an introduction and overview. Disponível em: http://www.esrf.eu/computing/scientific/FIT2D/FIT2D_INTRO/fit2d.html. Acesso em: 1 mar. 2021.
- 217 RICE, S. A. Small angle scattering of X-rays. A. Guinier and G. Fournet. translated by C. B. Wilson and with a bibliographical appendix by K. L. Yudowitch. Wiley, New York, 1955. 268 pp. \$7.50. **Journal of Polymer Science**, v. 19, n. 93, p. 594-594, mar. 1956.
- 218 KONAREV, P. V.; SVERGUN, D. I. A posteriori determination of the useful data range for small-angle scattering experiments on dilute monodisperse systems. **IUCrJ**, v. 2, n. pt 3, p. 352-360, 30 abr. 2015.
- 219 PERRY, J. J. P.; TAINER, J. A. Developing advanced X-ray scattering methods combined with crystallography and computation. **Methods**, v. 59, n. 3, p. 363-371. 2013. DOI: 10.1016/j.ymeth.2013.01.005.
- 220 FRANKE, D.; SVERGUN, D. I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. **Journal of Applied Crystallography**, v. 42, n. 2, p. 342-346, 2009.
- 221 VOLKOV, V. V.; SVERGUN, D. I. Uniqueness of *ab initio* shape determination in small-angle scattering. **Journal of Applied Crystallography**, v. 36, n. 3, p. 860-864, 1 jun. 2003.
- 222 SVERGUN, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. **Biophysical Journal**, v. 76, n. 6, p. 2879-2886, 1999.
- 223 FISCHER, H. *et al.* Determination of the molecular weight of proteins in solution from a

single small-angle X-ray scattering measurement on a relative scale. **Journal of Applied Crystallography**, v. 43, n. 1, p. 101–109, 1 dez. 2010.

224 WINTER, G. *et al.* DIALS: Implementation and evaluation of a new integration package. **Acta Crystallographica Section D: structural biology**, v. 74, n. Pt 2, p. 85–97, 2018.

225 HAN, B. G. *et al.* Crystal structure of a class 2 D-xylose isomerase from the human intestinal tract microbe *Bacteroides thetaiotaomicron*. **Korean Society for Structural Biology**, p. 41–47, 2015.

226 NOTREDAME, C.; HIGGINS, D. G.; HERINGA, J. T-coffee: a novel method for fast and accurate multiple sequence alignment. **Journal of Molecular Biology**, v. 302, n. 1, p. 205–217, 8 set. 2000.

227 KRISSINEL, E.; HENRICK, K. Inference of macromolecular assemblies from crystalline state. **Journal of Molecular Biology**, v. 372, n. 3, p. 774–797, 21 set. 2007.

228 ASHKENAZY, H. *et al.* ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. **Nucleic Acids Research**, v. 38, n. SUPPL. 2, 15 maio 2010. DOI: 10.1093/nar/gkq399.

229 LANDAU, M. *et al.* ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. **Nucleic Acids Research**, v. 33, n. SUPPL. 2, jul. 2005. DOI: 10.1093/nar/gki370.

230 EDGAR, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v. 32, n. 5, p. 1792–1797, 1 mar. 2004.

231 KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. **Molecular Biology and Evolution**, v. 33, n. 7, p. 1870–1874, 1 jul. 2016.

232 GASTEIGER, E. *et al.* Protein identification and analysis tools on the ExPASy server. In: **The Proteomics Protocols Handbook**. Totowa: Humana Press, 2005. p. 571–607.

233 STATION HOME PAGE. **Chesapeake bay bridge tunnel, Va.** Disponível em: <https://tidesandcurrents.noaa.gov/stationhome.html?id=8638863>. Acesso em: 1 mar. 2021.

234 JESKE, L. *et al.* BRENDA in 2019: a European ELIXIR core data resource. **Nucleic Acids Research**, v. 47, n. D1, p. D542–D549. 2019. DOI: 10.1093/nar/gky1048.

235 STAUDIGL, P.; HALTRICH, D.; PETERBAUER, C. K. L-arabinose isomerase and d-xylose isomerase from *Lactobacillus reuteri*: characterization, coexpression in the food grade host *Lactobacillus plantarum*, and application in the conversion of d-galactose and d-glucose. **Journal of Agricultural and Food Chemistry**, v. 62, n. 7, p. 1617–1624, 19 fev. 2014.

236 LEE, S. M.; JELLISON, T.; ALPER, H. S. Directed evolution of xylose isomerase for improved xylose catabolism and fermentation in the yeast *Saccharomyces cerevisiae*. **Applied and Environmental Microbiology**, v. 78, n. 16, p. 5708–5716, ago. 2012.

237 WALTMAN, M. J. *et al.* Engineering acidic *Streptomyces rubiginosus* D-xylose isomerase by rational enzyme design. **Protein Engineering, Design and Selection**, v. 27, n. 2, p. 59–64, fev. 2014.

- 238 PATEL, D. H. *et al.* Engineering of the catalytic site of xylose isomerase to enhance bioconversion of a non-preferential substrate. **Protein Engineering, Design and Selection**, v. 25, n. 7, p. 331–336, jun. 2012.
- 239 SUEKANE, M.; TAMURA, M.; TOMIMURA, C. Physico-chemical and enzymatic properties of purified glucoseisomerases from streptomyces olivochromogenes and bacillus stearothermophilus. **Agricultural and Biological Chemistry**, v. 42, n. 5, p. 909–917, 1978.
- 240 LEHMACHER, A.; BISSWANGER, H. Comparative kinetics of D-xylose and D-glucose-isomerase activities of the D-xylose Isomerase from *Thermus aquaticus* HB8. **Biological Chemistry Hoppe-Seyler**, v. 371, n. 1, p. 527–536, 1990.
- 241 FAN, L. *et al.* Cloning and analysis of the xylAB operon and characterization of xylose isomerase from *Thermoanaerobacter ethanolicus*. **Biotechnology Letters**, v. 33, n. 3, p. 593–598, 1 mar. 2011.
- 242 HA, S. J. *et al.* Xylitol does not inhibit xylose fermentation by engineered *Saccharomyces cerevisiae* expressing xylA as severely as it inhibits xylose isomerase reaction in vitro. **Applied Microbiology and Biotechnology**, v. 92, n. 1, p. 77–84, out. 2011.
- 243 LÖNN, A. *et al.* Cold adaptation of xylose isomerase from *Thermus thermophilus* through random PCR mutagenesis: Gene cloning and protein characterization. **European Journal of Biochemistry**, v. 269, n. 1, p. 157–163, 2002.
- 244 RALSTON, G. **Introduction to analytical ultracentrifugation**. Brea, CA, USA: Beckman Coulter, 2016.
- 245 KLOSE, D. P.; WALLACE, B. A.; JANES, R. W. 2Struc: the secondary structure server. **Bioinformatics**, v. 26, n. 20, p. 2624–2625, 24 ago. 2010.
- 246 SCHNEIDMAN-DUHOVNY, D. *et al.* Accurate SAXS profile computation and its assessment by contrast variation experiments. **Biophysical Journal**, v. 105, n. 4, p. 962–974, 20 ago. 2013.
- 247 VAN DEN BROECK, A. *et al.* Arginine residues as stabilizing elements in proteins. **Biochemistry**, v. 31, n. 8, p. 2239–2253, 1 fev. 1992.
- 248 GHATGE, M. S. *et al.* Unfolding and refolding of glucose/xylose isomerase from *Streptomyces* sp. NCIM 2730. **Enzyme and Microbial Technology**, v. 16, n. 4, p. 323–327, 1 abr. 1994.
- 249 MCNICHOLAS, S. *et al.* Presenting your structures: the CCP4mg molecular-graphics software. **Acta Crystallographica Section D: biological crystallography**, v. 67, n. 4, p. 386–394, abr. 2011.
- 250 KOVALEVSKY, A. Y. *et al.* Metal ion roles and the movement of hydrogen during reaction catalyzed by d-xylose isomerase: a joint x-ray and neutron diffraction study. **Structure**, v. 18, n. 6, p. 688–699, jun. 2010.
- 251 JENKINS, J. *et al.* Protein engineering of xylose (Glucose) isomerase from *actinoplanes missouriensis*. 1. crystallography and site-directed mutagenesis of metal binding sites. **Biochemistry**, v. 31, n. 24, p. 5449–5458, 1 fev. 1992.
- 252 DEREPPER, A. *et al.* Phylogeny.fr: robust phylogenetic analysis for the non-specialist.

Nucleic Acids Research, v. 36, n. Web Server issue, 2008.

253 JONES, D. T.; TAYLOR, W. R.; THORNTON, J. M. The rapid generation of mutation data matrices from protein sequences. **Bioinformatics**, v. 8, n. 3, p. 275–282, jun. 1992.

ANEXOS

Anexo A – Artigos publicados

- A1 Structural and Molecular Dynamics Investigations of Ligand Stabilization via Secondary Binding Site Interactions in *Paenibacillus xylanivorans* GH11 Xylanase**

Structural and molecular dynamics investigations of ligand stabilization via secondary binding site interactions in *Paenibacillus xylanivorans* GH11 xylanase

Lorenzo Briganti^a, Caio Capetti^a, Vanessa O.A. Pellegrini^a, Silvina Ghio^b, Eleonora Campos^b, Alessandro S. Nascimento^{a,*}, Igor Polikarpov^{a,*}

^a Instituto de Física de São Carlos, Universidade de São Paulo, Avenida Trabalhador São-carlosense 400, 13566-590 São Carlos, SP, Brazil

^b Instituto de Agrobiotecnología y Biología Molecular (IABIMO), CONICET-INTA, Las Reservas y Nicolas Repetto (qto), Hurlingham (1686), Prov. Buenos Aires, Argentina

ARTICLE INFO

Article history:

Received 2 January 2021

Received in revised form 27 February 2021

Accepted 1 March 2021

Available online 7 March 2021

Keywords:

Glycosyl hydrolase

Secondary binding site

Molecular dynamics

Crystallographic structure

GH11 xylanase

Paenibacillus sp

ABSTRACT

Glycoside hydrolases (GHs) are essential for plant biomass deconstruction. GH11 family consist of endo- β -1,4-xylanases which hydrolyze xylan, the second most abundant cell wall biopolymer after cellulose, into small bioavailable oligomers. Structural requirements for enzymatic mechanism of xylan hydrolysis is well described for GH11 members. However, over the last years, it has been discovered that some enzymes from GH11 family have a secondary binding sites (SBS), which modulate the enzymes activities, but mechanistic details of the molecular communication between the active site and SBS of the enzymes remain a conundrum. In the present work we structurally characterized GH11 xylanase from *Paenibacillus xylanivorans* A57 (PxXyn11B), a microorganism of agricultural importance, using protein crystallography and molecular dynamics simulations. The PxXyn11B structure was solved to 2.5 Å resolution and different substrates (xylo-oligosaccharides from X3 to X6), were modelled in its active and SBS sites. Molecular Dynamics (MD) simulations revealed an important role of SBS in the activity and conformational mobility of PxXyn11B, demonstrating that binding of the reaction products to the SBS of the enzyme stabilizes the N-terminal region and, consequently, the active site. Furthermore, MD simulations showed that the longer the ligand, the better is the stabilization within active site, and the positive subsites contribute less to the stabilization of the substrates than the negative ones. These findings provide rationale for the observed enzyme kinetics, shedding light on the conformational modulation of the GH11 enzymes via their SBS mediated by the positive molecular feedback loop which involve the products of the enzymatic reaction.

© 2021 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Plant biomass is an important resource for production of energy, agro-industrial manufacture, wood-based industry, and biofuels [1]. Plant cell walls from different sources (monocot grasses, woody dicots, herbaceous dicots, and softwoods) have different compositions, but all contain three main fractions: cellulose, hemicellulose and lignin [2,3]. Despite of a huge abundance of plant biomass, only 2% of this resource is currently used by humans [2]. The most abundant polysaccharide in cell walls is cellulose, making up 40% to 50% of the wall material, and the second largest

fraction, representing 25% to 35% of the walls, comprises the hemicelluloses [2]. A large portion, of hemicellulose is constituted by xylan [4], which can be decorated with glucuronic acid and/or arabinose [5].

According to Carbohydrate-Active Enzymes database (CAZy), endo- β -1,4-xylanases (EC.3.2.1.8) are found in GH families 3, 5, 8, 9, 10, 11, 12, 16, 26, 30, 43, 44, 51, 62, 98, and 141 [6,7], but the vast majority of characterized xylanases are from GH families 10 and 11 [7]. GH11 xylanases are endo-acting retaining enzymes which employ a classical Koshland double-displacement mechanism [8].

Although the canonical catalytic mechanism for GH11 enzymes is well described in structural and biochemical terms [9,10], discovery of a secondary binding site (SBS) on the surface of the GH11 xylanases raised new and interesting questions regarding

* Corresponding authors.

E-mail addresses: asnascimento@fisc.usp.br (A.S. Nascimento), ipolikarpov@fisc.usp.br (I. Polikarpov).

<https://doi.org/10.1016/j.csbj.2021.03.002>

2001-0370/© 2021 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

- A2** Transformation of xylan into value-added biocommodities using *Thermobacillus composti* GH10 xylanase.

Contents lists available at ScienceDirect

Carbohydrate Polymers

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

Transformation of xylan into value-added biocommodities using *Thermobacillus composti* GH10 xylanase

Ana Gabriela Veiga Sepulchro¹, Vanessa O.A. Pellegrini¹, Lorenzo Briganti¹,
Evandro A. de Araujo, Simara S. de Araujo, Igor Polikarpov^{*}

Instituto de Física de São Carlos, Universidade de São Paulo, Avenida Trabalhador São-carlense 400, 13566-590, São Carlos, SP, Brazil

ARTICLE INFO

Keywords:

Xylan
Crystallographic structure
Prebiotics
Cleavage patterns
Bioconversion

ABSTRACT

Enzymatic transformation of xylans into renewable fuels and value-added products is mediated by xylanases. Here we describe the biochemical and X-ray structural characterization of *Thermobacillus composti* GH10 xylanase (TcXyn10A) at 2.1 Å resolution aiming to unravel details of its recognition of glucurono- and arabinoxylan at a molecular level. TcXyn10A improves the efficiency of pretreated lignocellulosic biomass hydrolysis by a commercial enzyme cocktail causing a 15.35 % increase in xylose release and 4.38 % glucose release after 24 h of reaction. The enzyme releases predominantly xylobiose and xylofuranose, as well as MeGlcA3 × 3 (from beechwood glucuronoxylan) and a range of decorated xylooligosaccharides (XOS) from rye arabinoxylan, with Am2 × 2 being the major product. The enzyme liberates XOS with the yields of 29.09 % for beechwood glucuronoxylan and 16.98 % for rye arabinoxylan. Finally, TcXyn10A has a high thermal stability, halotolerance, and resistance to ethanol, biochemical properties that can be desirable for a number of industrial applications.

1. Introduction

Plant cell walls, representing the majority of terrestrial lignocellulosic biomass, are complex structures consisting of cellulose, hemicellulose, pectins, lignin, and a minor fraction of proteins. Production of renewable fuels, green chemicals, and sustainable materials from lignocellulosic plant biomass relies on enzymatic depolymerization of the major cell wall biopolymers to be efficient and cost-effective (Himmel et al., 2007; Ragauskas et al., 2006; Steen et al., 2010).

Xylan, which constitutes the largest portion of hemicellulose, is a diverse group of polysaccharides that represent 20–35% of lignocellulosic biomass (Smith, Wang, York, Peña, & Urbanowicz, 2017). This decorated polymer with a backbone of β -(1→4)-linked xylose residues, contributes to a high recalcitrance of the biomass and therefore presents itself as a challenge as well as an opportunity for the bioconversion of lignocellulosic feedstock into biocommodities (Scheller & Ulvskov, 2010).

The molecular properties of xylan such as molecular weight, identity, distribution and type of substituents vary depending on the species of plant, its stage of development, and the type of cells. Xylans can be divided into four groups according to the type of branching from the β -

D-xylopyranosyl backbone: homoxylan, arabinoxylan, glucuronoxylan, and arabinoglucuronoxylans (Naidu, Hlangathi, & John, 2018). Here we investigate enzymatic conversion of two major groups of xylans: glucuronoxylan and arabinoxylan (Fig. 1). Glucuronoxylan can be decorated with glucuronosyl and/or 4-O-methyl glucuronosyl (MeGlcA) at a position 2 of the xylose residues and is the largest component of hemicellulose in hardwoods (Ehringerová & Heinze, 2000; Naidu et al., 2018). Arabinoxylans, that represent the largest fraction of hemicellulose in cereal grains and grasses, may bear D-arabinose substitutions at the position 3 and/or positions 2 and 3 of the xylopyranose residues and are the largest fraction of hemicellulose in cereal grains and grasses (Fig. 1) (Ehringerová & Heinze, 2000; Naidu et al., 2018).

Due to structural and chemical variability of xylan, complete and efficient enzymatic hydrolysis of the polymer requires a simultaneous action of a number of hydrolytic enzymes (Gilbert, Ståhlbrand, & Brumer, 2008). The main enzymes involved in xylan hydrolysis are xylanases, widely distributed between prokaryotes and eukaryotes (Dekker & Richards, 1976), that cleave the β -1,4 bonds between xylose residues of the hemicellulose backbone (Kallim, Böhringer, Ali, & Schäberle, 2015).

Enzymatic degradation of xylan generates products of biotechnological interest, of which bioethanol and xylooligosaccharides stand

^{*} Corresponding author.

E-mail address: ipolikarpov@ifsc.usp.br (I. Polikarpov).

¹ These authors contributed equally to this work.

<https://doi.org/10.1016/j.carbpol.2020.116714>

Received 1 May 2020; Received in revised form 13 June 2020; Accepted 1 July 2020

Available online 03 July 2020

0144-8617/ © 2020 Elsevier Ltd. All rights reserved.

- A3** **Low-resolution molecular shape, biochemical characterization and emulsification properties of a halotolerant esterase from *Bacillus licheniformis*.**

ORIGINAL ARTICLE

Low-resolution molecular shape, biochemical characterization and emulsification properties of a halotolerant esterase from *Bacillus licheniformis*

Ana Elisa T. Leite¹ · Lorenzo Briganti¹ · Evandro Ares de Araújo¹ · Vanessa de Oliveira Arnoldi Pellegrini¹ · Cesar Moyses Camilo¹ · Igor Polikarpov¹

Received: 5 December 2019 / Revised: 28 June 2020 / Accepted: 8 July 2020 / Published online: 18 July 2020
 © European Biophysical Societies' Association 2020

Abstract

Bacterial esterases are highly versatile enzymes, currently widely used in detergents, biosurfactants, bioemulsifiers and as biocatalysts in paper and food industries. Present work describes heterologous expression, purification, and biophysical and biochemical characterization of a halotolerant esterase from *Bacillus licheniformis* (BIEstA). BIEstA preferentially cleaves *p*NP-octanoate and both activity and stability of the enzyme increased in the presence of 2 M NaCl, and also with several organic solvents (ethanol, methanol and DMSO). Furthermore, BIEstA has considerable emulsifying properties, particularly with olive oil as substrate. Our studies also show that the enzyme is monomeric in solution and its small-angle X-ray scattering low-resolution molecular envelope fits well its high-resolution homology model.

Keywords Esterases · *Bacillus licheniformis* · Biochemical characterization · Small-angle X-ray scattering

Introduction

Modern biotechnology is unthinkable without applications of enzymes. Currently, over 200 different enzymes, including carboxylic ester hydrolases, are used in industrial applications (Sharma 2011) for manufacturing a wide range of industrial products, in diverse areas such as (bio)chemical synthesis, pharmaceutical industry, food, feed and detergent additives (Adrio 2013). The commercial use of lipases and esterases has a wide variety of different applications including synthesis of biopolymer and biodiesel, production of pharmaceuticals, agrochemicals, cosmetics and flavors, having billion dollars value (Joseph, Ramtke and Thomas 2008).

Carboxylic ester hydrolases (triacylglycerol acyl hydrolases EC 3.1.1.3) is a group of enzymes that frequently act at

the lipid-water interface, catalyzing the disruption of triglyceride chains via hydrolysis of ester bonds. These reactions cause release of fatty acids and organic alcohols (Nthangeni 2001; KhyamiHorani 1996; Barros 2010). The enzymes can be further divided into lipases and esterases, which are present in diverse organisms, such as plants, yeasts, fungi and bacteria (Kumar 2005; Sharma 2001). Carboxylic ester hydrolases are widely applied in flavor and fragrances production, drug synthesis, transesterification and hydrolysis of fats and oils and also in biosensor manufacturing (Pandey 1999; Wakelin 1997). Within the recent years, lipases have also been used as biosurfactants and bioemulsifiers for oil recovery processes by petroleum industries (Banat 2000; Joshi 2016).

Bacillus licheniformis is a gram-positive bacterium used for production of several enzymes, such as lipases, esterases, proteases and α -amylases, among others (Eveleigh 1981). Here, we describe the recombinant protein production, biophysical and biochemical characterization and low-resolution SAXS-derived molecular shape in solution of an esterase from *Bacillus licheniformis* (BIEstA).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00249-020-01448-7>) contains supplementary material, which is available to authorized users.

✉ Igor Polikarpov
 ipolikarpov@ifsc.usp.br

¹ São Carlos Institute of Physics, University of São Paulo, São Carlos, SP, Brazil