

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

ANDRE SCHÜTZER DE GODOY

**Estudos estruturais e funcionais das enzimas beta-
galactosidases de bactérias**

São Carlos

2016

ANDRE SCHÜTZER DE GODOY

Estudos estruturais e funcionais das enzimas beta-galactosidases de bactérias

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

Área de concentração: Física Aplicada

Opção: Física Biomolecular

Orientador: Prof. Dr. Igor Polikarpov

Co-orientador: Dr. César Moises Camillo

São Carlos
2016

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

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

Godoy, Andre Schutzer de
Estudos estruturais e funcionais das enzimas beta-
galactosidases de bactérias / Andre Schutzer de
Godoy; orientador Igor Polikarpov; co-orientador
César Moises Camilo -- São Carlos, 2016.
123 p.

Tese (Doutorado - Programa de Pós-Graduação em
Física Biomolecular) -- Instituto de Física de São
Carlos, Universidade de São Paulo, 2016.

1. Beta-galactosidases. 2. Biotecnologia. 3.
Enzimas. 4. Raio-X. I. Polikarpov, Igor, orient. II.
Camilo, César Moises, co-orient. III. Título.

FOLHA DE APROVAÇÃO

Andre Schutzer de Godoy

Tese apresentada ao 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.

Aprovado(a) em: 29/08/2016

Comissão Julgadora

Dr(a). Alessandro Silva Nascimento

Instituição: (IFSC/USP)

Dr(a). Raghuvir Krishnaswamy Arni

Instituição: (UNESP/São José do Rio Preto)

Dr(a). Flávio Henrique da Silva

Instituição: (UFSCar/São Carlos)

Dr(a). Yvonne Primerano Mascarenhas

Instituição: (IFSC/USP)

Dr(a). Mario Tyago Murakami

Instituição: (CNPEM/Campinas)

Aos meus pais

AGRADECIMENTOS

Á família, que sempre me incentivou e fortaleceu a trilhar meu caminho

Á Rafaela, pelos anos de suporte, carinho e respeito que se passaram e aos que estão por vir

Ao professor Igor, pela orientação e confiança nesse e outros projeto

Aos professores César, Alessandro, JR e Murakami pelos inestimáveis ensinamentos que possibilitaram esse trabalho

Aos colaboradores e amigos Kadowaki, Amanda, Melissa e Heloísa, não só pelas contribuições de cada um mas também pela amizade

A todos os amigos da sala 8 (Atílio, Aline, Bruno, Caio, Danilo, Evandro, Grazi, Má, Rê, Vasi e Van), pelos anos de amizade e risadas

A todos os amigos que passaram pelo laboratório

Aos técnicos e funcionários do IFSC, em especial a Livia, Maria, Possato e Josimar

Aos amigos do campus 2, que estiveram juntos em todos os momentos

A meus grandes amigos e amigas da melhor turma da bio de todos os tempos

A FAPESP pela concessão da bolsa de doutorado direto e as grandes oportunidades que essa me proporcionou

Ao grande Milton

Aos meus pais, por tudo

*“Gostaria que isso não tivesse acontecido na minha época” - disse Frodo.
“Eu também - disse Gandalf. - Como todos os que vivem nestes tempos. Mas
a decisão não é nossa. Tudo o que temos de decidir é o que fazer com o
tempo que nos é dado”*

J.R.R. Tolkien, A Sociedade do Anel

RESUMO

GODOY, A. S. **Estudos estruturais e funcionais das enzimas beta-galactosidases de bactérias**. 2016. 123 p. Tese (Doutorado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2016.

As β -galactosidases são dissacaridasas capazes de realizar a reação de hidrólise das ligações $\beta(1\rightarrow4)$ de um galactosídeo, tendo a lactose como principal substrato natural. Essas enzimas são amplamente utilizadas na ciência e na indústria, apresentando um alto potencial biotecnológico. Além das propriedades hidrolíticas, as β -galactosidases possuem a característica de sintetizar açúcares complexos chamados galactooligossacarídeos, que são conhecidos como prebióticos. Nesse projeto, nos propusemos a estudar enzimas do tipo β -galactosidase com alto potencial biotecnológico. Foram escolhidos genes dos organismos *Xanthomonas campestris pv. campestris* e *Bifidumbacterium bifidum*, os quais possuem alta atividade β -galactosídica, conforme a literatura. Foram clonados nove genes, dos quais o produto de quatro foram purificados, cristalizados e tiveram sua estrutura cristalográfica determinada. A enzima BbgII foi resolvida pela técnica de *single anomalous diffraction*. Sua estrutura revelou um trímero em forma de barril, no qual foi possível observar interações entre os resíduos do sítio ativo e a galactose. Adicionalmente, realizamos a caracterização bioquímica e cinética da enzima nativa e de mutantes pontuais. Também foram resolvidas as estruturas cristalográficas das enzimas XCC_1754, XCC_2404 e XCC_2895. A enzima XCC_1754 exibiu uma significativa alteração no *loop* 11 nas cadeias entre os dois monômeros da unidade assimétrica. Esse *loop* exibe as conformações aberta e fechada sobre o sítio de interação com os substratos e, através de ensaios de mutação, propomos que essas diferenças são mediadas pelas glicinas 294 e 302, que atuam como uma dobradiça. Apesar de apresentar menor afinidade pelo substrato, o mutante G294P exibiu uma atividade 50% maior do que a enzima nativa. Enquanto isso, o mutante G302P, apesar de exibir um ganho em sua afinidade, perdeu a capacidade de processar eficientemente o substrato. A enzima XCC_2895 também possui três domínios, porém características bioquímicas similares a XCC_1754. Apesar de haver ainda a necessidade de mais estudos para podermos comparar ambas, acreditamos que o fato da enzima XCC_2404 possuir uma estabilidade térmica mais elevada que a enzima XCC_1754, pode estar relacionado com a formação de grandes oligômeros.

Palavras-chave: Beta-galactosidases. Biotecnologia. Enzimas. Raios X.

ABSTRACT

GODOY, A. S. Structural and functional studies of beta-galactosidases enzymes from bacteria. 2016. 123 p. Tese (Doutorado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2016.

The β -galactosidases are glycosyl hydrolases that act at the $\beta(1\rightarrow4)$ bonds from galactosides, with lactose as the main natural substrate. The use of such enzymes in both science and industry is very common, due its high biotechnological applicability. Beside its hydrolytic capacity, the β -galactosidases are also commonly used for the synthesis of galactoligosacharydes, well-known prebiotics. The focus of this project was to study β -galactosidases with high biotechnological potential. For that, genes from the organisms *Xanthomonas campestris pv. campestris*, e *Bifidumbacterium bifidum* were selected due the high β -galactosidic activity of those organisms, according previous works. Such genes were cloned, and four of them were expressed, crystalized and had its x-ray structure determined. The enzyme BbgII was solved applying the single anomalous diffraction method. Its structure shows a trimer forming a barrel, in which it was also possible to observe interactions between residues of the active site and galactose. The native and site direct mutants were biochemically characterized, revealing important features and the different roles of amino acids of active site. We also solved the structure of the enzymes XCC_1754, XCC_2404 and XCC_2895. The enzyme XCC_1754 showed a significant difference between the loop 11 of the two monomers at the asymmetric unit. This loop presented both open and closed conformations, which we believe it was caused by glycine's 294 and 302, acting as a hinge. Despite the mutant G294P exhibited a decrease in enzyme affinity for the substrate, its general activity increased up to 50%. Meanwhile, the mutant G302P has gain affinity for the substrates, but it fails into process the substrate efficiently. The enzyme XCC_2895 shows three domains, but its biochemical properties are similar to the enzyme XCC_1754. Although there is still a need for further studies to be able to compare the enzymes, we believe that the higher thermal stability of XCC_2404 compared to XCC_1754 could be related to the formation of large oligomers.

.
Keywords: Beta-galactosidases. Biotechnology. Enzymes. X-ray.

LISTA DE FIGURAS

- Figura 1 - Aplicações da β -galactosidase na ciência. Em A, células humanas indicam níveis diferentes de expressão com base na ativação do *operon lac* (cor azul). Em B, cristais de lacZ corados com X-gal. Em C, a lacZ resolvida por Jacobson em 1994. Em D, o mapa de crioeletromicroscopia da enzima LacZ à 2,2 Å de resolução publicado em 2015. 30
- Figura 2 - Estrutura da enzima lacZ. No modelo cada domínio está destacado em uma cor e marcado de acordo com a ordem a partir da região N terminal. PDBid 4TTG. 32
- Figura 3 - β -galactosidases multidomínios da família GH35, pertencentes aos eucariotos *Penicillium sp.* e *Trichoderma reesei*. Em destaque temos os domínios barril TIM, onde se encontram os resíduos catalíticos. 33
- Figura 4 - β -galactosidases contendo dois domínios da família GH42, pertencentes aos organismos *Thermus sp.* e *Bacillus circulans*. Em destaque temos os domínios barril TIM, onde encontramos os resíduos catalíticos. 34
- Figura 5 - Esquema do mecanismo de retenção utilizado pelas β -galactosidases. No passo 1, o galactosídeo sofre o ataque nucleofílico por uma água, formando um intermediário covalente com a enzima, e liberando o radical hidrolisado. A partir daí a reação pode seguir dois caminhos: a hidrólise completa do galactosídeo consumindo uma água (Passo 2); ou a transgalactosilação, onde o doador de prótons da hidrólise é proveniente do terminal redutor de um carboidrato. Nesse caso, ocorre a formação de galactooligosacarídeos como produto da enzima. 36
- Figura 6 - Esquema ilustrativo do método LIC utilizado para clonagem. Em A, temos o gene de interesse (verde) amplificado por primers (salmão) contendo uma extensão de bases complementares ao vetor. Em B, o produto do PCR é tratado com a enzima Phusion polimerase na presença de dATP, deixando para trás um fragmento de fita simples. Em C, vemos o pareamento do produto tratado com o vetor de interesse (azul), tratado de maneira recíproca. 44
- Figura 7 - Esquema do método de difusão de vapor. Uma mistura de proteína e solução-mãe é selada no sistema contendo a solução mãe pura. A diferença nas concentrações de precipitante entre as soluções favorece a troca de vapor no sentido mistura-solução mãe. 48
- Figura 8 - Esquema do método de matriz esparsa e diagrama de fases de uma macromolécula. Em A, vemos a representação de um universo de possibilidades (cubo) onde algumas regiões foram exploradas (esferas) através do método; Em B, o diagrama de fases de uma proteína hidrofílica. 48

- Figura 9 - Amplificação do fragmento contendo o gene da enzima BbgII. Em 1, temos o marcador molecular; Em 2, vemos o produto da reação de PCR contendo o fragmento amplificado referente ao gene de interesse. Experimentos marcados com * não estão relacionados aos discutidos. 57
- Figura 10 - Amplificação dos fragmentos contendo os genes das enzimas XCC_1754, XCC_2404 e XCC_2895. Em 1, temos o marcador molecular; em 2, o fragmento contendo o gene de XCC_2895 (1773 pb); em 3, o fragmento contendo o gene de XCC_1754 (1668 pb); em 4, o fragmento contendo o gene de XCC_2404 (1533 pb). Experimentos marcados com * não estão relacionados aos discutidos..... 58
- Figura 11 - Testes de expressão da enzima BbgII. Em M, temos o marcador molecular; Em 1, 2, 3 e 4 vemos respectivamente a alíquota não induzida, a alíquota após a indução, a fração solúvel e a fração insolúvel do teste a 30 °C. Em 5, 6, 7 e 8 vemos respectivamente a alíquota não induzida, a alíquota após a indução, a fração solúvel e a fração insolúvel do teste a 18 °C. A seta vermelha aponta para a proteína BbgII (77,2 kDa) encontrada na fração solúvel..... 59
- Figura 12 - Testes de expressão à 18 °C das enzimas XCC_1754, XCC_2404 e XCC_2895. Em 1 vemos o marcador molecular; Em 2 a fração solúvel da enzima XCC_2895 (65,6 kDa); Em 3 a fração solúvel da enzima XCC_1754 (60,1 kDa); Em 4 a fração solúvel da enzima XCC_2404 (55,2 kDa). As proteínas estão demarcadas pelo círculo vermelho. Experimentos marcados com * não estão relacionados aos discutidos..... 61
- Figura 13 - Cromatograma da purificação por afinidade à níquel pela enzima BbgII. A linha preta representa a absorção da amostra a 280 nm. Em azul, a porcentagem do tampão B na amostra. 62
- Figura 14 - Cromatogramas das purificações por exclusão molecular e géis SDS-Page das enzimas BbgII, XCC_1754, XCC_2404 e XCC_2895. Em A e B temos respectivamente o cromatograma e o gel da purificação da enzima BbgII. Em C e D o cromatograma e o gel da purificação da enzima XCC_1754. Em E e F temos o cromatograma e o gel da purificação da enzima XCC_2404. Em G e H temos o cromatograma e o gel da purificação da enzima XCC_2895. O quadro vermelho destaca a amostra no cromatograma. 62
- Figura 15 - Regressão linear do log da massa molecular (MW) pelo volume de eluição (mL) das amostras padrões da gel filtração analítica. O quadro mostra os valores de coeficiente angular (a) e o linear (b) calculados..... 63
- Figura 16 - Cristais das β -galactosidases. Em A, temos os cristais de BbgII na condição di-tartarato de sódio 0,2 M, PEG 3350 20% e 1-propanol 4%. Em B vemos os cristais de BbgII três dias após o *soaking* com mercúrio. Em C temos os cristais da enzima XCC_2895 na condição cloreto de lítio 0,2 M

	e PEG 3350 25%. Em D, os cristais de XCC_2404 na condição acetato de cálcio 0,2 M e PEG 3350 20%. Em E, vemos os cristais de XCC_1754 na condição citrato de tri-lítio 0,2 M e PEG 3350 20%.....	65
Figura 17 -	Coeficientes de dispersão (f' e f'') em diferentes energias para o elemento mercúrio. Triângulos e linha tracejada demarcam os valores para a energia produzida no ânodo de cobre.	66
Figura 18 -	Marcos de difração dos cristais de BbgII. Em A), o marco inicial dos cristais nativos à 1.7 Å de resolução. Em B), o marco inicial dos cristais derivados com mercúrio à 2.5 Å de resolução.	68
Figura 19 -	Marcos de difração da enzima XCC_1754. Em A), o marco de difração inicial exhibe anéis de gelo e baixa qualidade. Em B), o marco após o <i>annealing</i> mostra uma significativa melhora na qualidade dos dados.	69
Figura 20 -	Parâmetro de resolução da estrutura da enzima BbgII. Em A), vemos os valores de I_{ano}/Sig_{ano} por faixa de resolução, que sugerem um alto sinal anômalo abaixo de 3,0 Å. Em B), os coeficientes de correlação mostram dois grupos distintos de soluções. Em C), os valores de ocupância para os átomos pesados encontrados mostram uma queda brusca a partir do décimo segundo átomo. Em D), a figura de mérito (FOM) do mapa durante a modificação de densidade mostra uma distinção acentuada entre os soluções original e invertida.....	73
Figura 21 -	Passos de resolução da estrutura da enzima BbgII. Em A vemos os 12 sítios de átomos pesados encontrados no conjunto derivado. Em B, vemos a densidade eletrônica da “mão-invertida”. Em C, o modelo inicial construído automaticamente.....	74
Figura 22 -	Estrutura cristalográfica da Enzima BbgII. Em A vemos a forma monomérica da enzima onde temos o domínio 1 (em verde) formado por um barril-TIM, o domínio 2 (em azul claro) e o domínio 3 (em laranja). Em B, temos o trímero em formato de barril, que é tanto a unidade assimétrica quanto a unidade biológica da enzima.	76
Figura 23 -	Localização e visão geral do sítio ativo da enzima BbgII. Em A observamos a interação entre duas moléculas de BbgII e a localização do sítio entre essas moléculas. Em B, a visão da superfície do sítio e a posição das moléculas de galactose com suas respectivas densidades eletrônicas. Em C, a coordenação da galactose no sítio ativo da BbgII.	78
Figura 24.	Testes de estabilidade da BbgII. Em A, vemos a atividade relativa da BbgII em diferentes temperaturas. Em B, a atividade relativa da enzima em diferentes pH's	79
Figura 25 -	Curvas da cinética da enzima BbgII a 25 °C. Em A), a curva utilizando lactose como substrato. Em B), a curva com o substrato <i>o</i> NPG.....	80

Figura 26 - Plot das curvas de Lineweaver-Burk da enzima BbgII em presença de diferentes concentrações de galactose.	81
Figura 27 - Temperatura de desenovelamento da enzima BbgII e seus mutantes.....	83
Figura 28 - Comparação entre BbgII e seus homólogos. Em A), a região do sítio de zinco de A4- β -Gal (verde) e Bca- β -gal (rosa) mostra substituição de todos os resíduos em BbgII (azul). Em B), a entrada do sitio as moléculas de galactose, com destaque (amarelo) para região rica em resíduos com enxofre. Em C) o sitio de zinco de B1Gal42A (laranja) comparado com a região em BbgII mostra substituição dos resíduos. Em D), o alinhamento entre os sitio ativos de BbgII (azul), A4- β -Gal (verde) e Bca- β -gal (rosa).	85
Figura 29 - Modelo cristalográfico da enzima XCC_1754. A enzima possui um domínio catalítico no formato de barril TIM, e um domínio C-terminal formado por fitas.....	93
Figura 30 - Diferenças não cristalográficas observadas na estrutura de XCC_1754. Em A, vemos o loop da cadeia A na posição “aberta”. Em B, vemos o mesmo loop na posição “fechada” na cadeia B. O quadro a direita mostra detalhes de cada uma das respectivas regiões.....	94
Figura 31 - Alinhamento da enzima XCC_1754 e homólogos da família 35. Em azul, a enzima XCC_1754 na posição fechada, enquanto que o <i>loop</i> em vermelho mostra a posição aberta da mesma. Em verde, a enzima Bgl35A. Em laranja, a enzima de <i>Caulobacter crescentus</i>	95
Figura 32 - Ramachandran e visão geral da região do <i>loop</i> 11 da enzima XCC_1754. Em A), o gráfico de Ramachandran para os resíduos 290-314 mostra que as glicinas 294 e 302 possuem ângulos bastante distintos. Em B), a região do <i>loop</i> com destaque para os resíduos catalíticos e o ligante sobreposto da Bgl35A, mostrando possíveis interações com o resíduo Gln300 do <i>loop</i> 11.	96
Figura 33 - Atividade relativa da enzima XCC_1754 em diferentes condições. Em A), a atividade relativa da enzima em diferentes temperaturas. Em B), a atividade relativa da enzima em diferentes pH's.	97
Figura 34 - Termo flourimetria da enzima XCC_1754 e seus mutantes. Os testes foram feitos em pH's 5, 6 e 7 e os T_m obtidos estão mostrados em barras.	97
Figura 35 - Estrutura cristalográfica da enzima XCC_2895. Em A), a unidade assimétrica da estrutura cristalográfica. Em B), a visão do monômero com os três domínios em destaque.	103
Figura 36 - Alinhamento e sítio ativo da enzima XCC_2895. Em A) o alinhamento da enzima XCC_2895 (verde) com os PBDs 4MAD (azul), 4E8C (laranja), 3D3A (amarelo) e 3THC (rosa). Em B), o sítio da XCC_2895 com os resíduos catalíticos em destaque. Também temos em B) o ligante da enzima de <i>Streptococcus pneumoniae</i> alinhado ao sitio da XCC_2895.....	104

Figura 37 - Atividade relativa da enzima XCC_2895 em diferentes pH's e temperaturas. Em A), a atividade da enzima em diferentes pH's. Em B), a atividade da enzima em diferentes temperaturas.....	105
Figura 38 - Modelo cristalográfico inicial da enzima XCC_2404.Em A), a unidade assimétrica do modelo cristalográfico da enzima XCC_2404 onde cada cadeia está representada por uma cor diferente. Em B), o monômero da enzima formado por dois domínios.....	109
Figura 39 - Gel Nativo da enzima XCC_2404. Em 1, o marcador. Em 2 a enzima nativa. Em 3, a enzima em 10 mM de DTT. Em 4, a enzima em 10 mM de EDTA. Em 5, a enzima em 2,0 M de NaCl.....	110
Figura 40 - Caracterização bioquímica da enzima XCC_2404. Em A), a atividade relativa da enzima em diferentes temperaturas. Em B), a atividade relativa da enzima em diferentes pH's.	111
Figura 41 - Curvas de termo fluorimetria da enzima XCC_2404. Cada curva mostra a fluorescência da amostra em um pH's diferente, de acordo com a legenda.	112

LISTA DE TABELAS

Tabela 1 - Informações sobre os genes de β -galactosidases selecionados	43
Tabela 2 - Detalhes dos vetores utilizados.	45
Tabela 3 - Informações gerais sobre os alvos de estudo.....	56
Tabela 4 - Resultados da gel filtração analítica.	64
Tabela 5 - Estatísticas das coletas de dados das β -galactosidases. Em parênteses os valores para a alta resolução.	67
Tabela 6 - Estatísticas de refinamento da enzima BbgII e de seu complexo com galactose.....	75
Tabela 7 - Parâmetros cinéticos dos mutantes de BbgII a 25 °C. Onde temos * a atividade foi muito baixa para ser quantificada.	83
Tabela 8 - Estatísticas do refinamento das enzimas XCC_1754 e XCC_2895. Os valores entre parênteses são para as últimas faixas de resolução.	90
Tabela 9 - Parâmetros cinéticos da enzima XCC_1754 e seus mutantes.	98

LISTA DE ABREVIACÕES E SIGLAS

oNPG	orto-nitrofenil- β -galactosídeo
X-gal	5-bromo-4-cloro-3-indoxil- β -D-galactopiranosídeo
CAZy	<i>Carbohydrate-Active enZyme databan</i>
PDB	<i>Protein Data Bank</i>
GOS	galactooligossacarídeos
NCBI	National Center for Biotechnology Information
PCR	<i>Polymerase chain reaction</i>
DNA	ácido desoxirribonucleico
LIC	<i>Ligation independent cloning</i>
LB	Luria-Bertani
IPTG	Isopropil β -D-1-tiogalactopiranosídeo
SAD	<i>Single anomalous diffraction</i>
RMSD	Root-mean-square deviation
EDTA	ácido etilenodiamino tetra-acético
DTT	ditiotreitól
DLS	Espalhamento Dinâmico de Luz e Potencial Zeta

SUMÁRIO

1	Introdução e Objetivos	29
1.1	Um breve histórico das β -galactosidases na ciência	29
1.2	Classificação e arquitetura das β -galactosidases	31
1.3	Mecanismo e atividade das β -galactosidases	34
1.4	As β -galactosidases na medicina e na indústria	37
1.5	As β -galactosidases na produção de biocombustíveis	38
1.6	Os organismos de estudo e objetivos do projeto	38
2	Materiais e Métodos	43
2.1	Seleção dos genes das β -galactosidases	43
2.2	Clonagem através do método <i>LIC</i>	44
2.3	Transformação bacteriana.	46
2.4	Mutação sítio-dirigida.....	46
2.5	Expressão em larga escala.	46
2.6	Purificação das enzimas.....	47
2.7	Cristalização, co-cristalização, <i>soaking</i> e coleta de dados das enzimas.	47
2.8	Resolução das Estruturas	49
2.9	Atividade e cinética das enzimas	50
2.10	Medidas de termo fluorimetria.....	50
3	Resultados e Discussão	55
3.1	Análise das sequências e considerações gerais	55
3.2	Clonagem dos genes	56
3.3	Testes de expressão	58
3.4	Expressão em larga escala e purificação das enzimas.....	60
3.5	Análise do estado oligomérico das enzimas.....	63
3.6	Cristalização e das enzimas	64
3.7	Coleta e processamento de dados.....	66
4	Estrutura e função da enzima BbgII.....	73
4.1	Resolução e qualidade da estrutura cristalográfica da enzima BbgII.....	73
4.2	Estrutura tridimensional da enzima BbgII.	75
4.3	Sítio ativo da enzima BbgII.....	77
4.4	Parâmetros bioquímicos da BbgII.....	79
4.5	Cinética e inibição da enzima BbgII.	80
4.6	Atividade e Estabilidade dos mutantes de BbgII.	82
4.7	Discussão sobre a enzima BbgII	84
4.8	Conclusões e perspectivas em relação a enzima BbgII.....	86
5	Resolução e qualidade do refinamento das estrutura das enzimas XCC_1754, XCC_2404 e XCC_2895.	89

6	Estrutura e função da enzima XCC_1754.....	93
6.1	Estrutura tridimensional da enzima XCC_1754.....	93
6.2	Caracterização bioquímica da enzima XCC_1754.....	96
6.3	Cinética da enzima XCC_1754 e seus mutantes.....	98
6.4	Conclusões em relação a enzima XCC_1754.....	98
7	Estrutura e função da enzima XCC_2895.....	103
7.1	Estrutura tridimensional da enzima XCC_2895.....	103
7.2	Caracterização bioquímica, cinética e considerações da enzima XCC_2895.....	104
8	Estrutura e Função da enzima XCC_2404.....	109
8.1	Estrutura cristalográfica da enzima XCC_2404.....	109
8.2	Estado oligomérico da enzima XCC_2404.....	110
8.3	Parâmetros bioquímicos da enzima XCC_2404.....	111
8.4	Discussão sobre a enzima XCC_2404.....	112
9	Conclusões e Perspectivas.....	114
	Referências	115
	ANEXO	123

CAPÍTULO 1

Introdução & Objetivos

1 Introdução e Objetivos

1.1 Um breve histórico das β -galactosidases na ciência

As β -galactosidases (lactases, EC 3.2.1.23) foram descritas pela primeira vez em 1889 pelo professor Martinus Beijerinck, do Departamento de Microbiologia da Universidade de Delft.¹ Nesse período, Beijerinck percebeu que algumas leveduras do gênero *Saccharomyces* eram capazes de fermentar a lactose em monossacarídeos. Mesmo em condições rudimentares, ele foi capaz de precipitar e isolar as β -galactosidases das culturas de leveduras, e passou cerca de 30 anos desenvolvendo ensaios que acabaram sendo fundamentais para a microbiologia e a enzimologia moderna.¹ Supreendentemente, mesmo após 90 anos algumas de suas amostras de enzima foram testadas e exibiram atividade significativa em substratos sintéticos, como o orto-nitrofenil- β -galactosídeo (oNPG).¹

Durante os anos seguintes, vários avanços foram descritos nos ramos da biologia molecular, incluindo a descrição do mecanismo de atividade, no qual o *operon lac* regula a expressão gênica. Tal descoberta foi feita por Jacob & Monod, que observaram que a bactéria *Escherichia coli* (*E. coli*) regulava a expressão de β -galactosidase, de acordo com as quantidades de lactose disponíveis no meio.² Nos anos 70, essas descobertas abriram as portas para um novo ramo da biologia, a produção de proteínas recombinantes.

Talvez a propriedade mais conhecida das β -galactosidases no ramo da ciência seja a habilidade dessa enzima em degradar o X-gal (5-bromo-4-cloro-3-indoxil- β -D-galactopiranosídeo), um composto incolor com alta afinidade por essa enzima. Quando clivado pela enzima, a porção 5-bromo-4-cloro-3-indoxil do composto é liberada e se dimeriza, formando o 5,5'-dibromo-4,4'-dicloro-índigo, um corante azul utilizado em diversas práticas na biologia molecular,³ especialmente na área de microbiologia, embriologia e técnicas gerais de biologia molecular (Figura 1).⁴⁻⁶ Na cristalografia de proteínas, uma das questões mais fundamentais é, se de fato, a macromolécula cristalizada está em uma conformação similar a da encontrada em solução. Para verificarmos isso, frequentemente, os professores demonstram que os cristais de β -galactosidase são capazes de processar o X-gal e adquirir uma coloração azul (Figura 1), não só demonstrando que estes são formados por enzimas ativas, mas ao mesmo tempo mostrando que a catálise não requer grandes mudanças conformacionais por parte da enzima (de outra maneira, os cristais seriam destruídos).⁷

As β -galactosidases também tiveram um papel importante na história da biologia estrutural. Em 1994, a β -galactosidase de *E. coli* (LacZ) foi a primeira proteína a ter sua estrutura resolvida com mais de 1000 aminoácidos, quebrando um paradigma da época e mostrando ser possível estudar cristais não-virais com eixos de célula maiores que 500 Å.⁸ Outro recorde mais recente foi na crioeletromicroscopia, onde a LacZ bateu os recordes de resolução para a técnica, atingindo os fantásticos 2,2 Å, e desenhando paradigmas para esse novo ramo da biologia estrutural (Figura 1).⁹

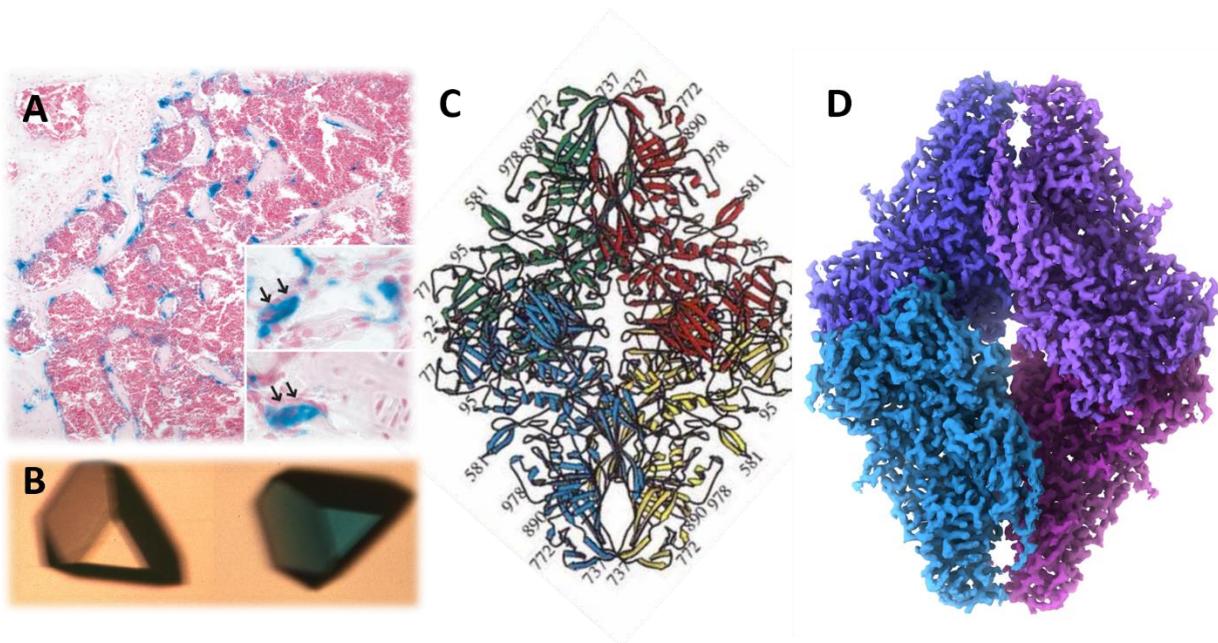


Figura 1 - Aplicações da β -galactosidase na ciência. Em A, células humanas indicam níveis diferentes de expressão com base na ativação do *operon lac* (cor azul). Em B, cristais de lacZ corados com X-gal. Em C, a lacZ resolvida por Jacobson em 1994. Em D, o mapa de crioeletromicroscopia da enzima LacZ à 2,2 Å de resolução publicado em 2015.

Fonte: Adaptada de Rouwenhorst¹, Dimri⁴, Jacobson⁸ e Bartesaghi⁹

Por fim, podemos concluir que a lacZ teve um papel ativo no desenvolvimento da ciência nos últimos 150 anos, principalmente, nas áreas de cinética enzimática, expressão gênica e biologia molecular, celular e estrutural. Sua alta massa molecular e habilidade de formar oligômeros, alta capacidade de regulação gênica, simplicidade reacional e alta estabilidade fizeram essa enzima um padrão para o desenvolvimento de técnicas, criação de modelos e estudos de caso, possibilitando um grande avanço em diversas áreas do conhecimento.⁷

1.2 Classificação e arquitetura das β -galactosidases

Os carboidratos são a base da biomassa de qualquer organismo e estão envolvidos em, praticamente, todas as funções metabólicas conhecidas. Apesar de terem uma composição química semelhante, os carboidratos apresentam tanto uma grande variedade de formas estereo-químicas, como também a capacidade de se arranjar com praticamente qualquer outro carboidrato. Dessa maneira, um simples hexassacarídeo pode ter algo em torno de 10^{12} arranjos distintos.¹⁰ Durante nossa história natural, a evolução explorou essa capacidade dos carboidratos, a fim de criar uma variedade infinita de estruturas derivadas de carboidratos.

Tanto os carboidratos complexos, como a hemicelulose, ou extremamente simples, como monossacarídeos, são amplamente distribuídos na natureza e participam de todas as interações inter e intra organismos. Tais interações permitiram que milhões de enzimas surgissem para realizar as mais variadas funções, como a formação de carboidratos (transferases de glicosídeos) e a quebra (hidrolases de glicosídeos, lipases, esterases, etc). Devido a enorme variedade de enzimas encontradas na natureza, foi-se necessário que estas fossem classificadas de maneira rápida e eficiente, para que mais pesquisadores pudessem acompanhar as descobertas da área. Para isso, criou-se em 1999 o *Carbohydrate-Active enZyme databank* (CAZy), a fim de proporcionar online os avanços e novidades na área.¹¹⁻¹² Com base na semelhança entre aminoácidos, o CAZy agrupa as diferentes enzimas em famílias com pelo menos um membro caracterizado bioquimicamente entre esses.^{11,13} Atualmente (julho de 2016), o grupo das hidrolases de glicosídeos (GH) contém cerca de 300.000 membros, divididos entre 135 famílias, e constitui cerca de 50% das enzimas já classificadas pelo CAZy.

As β -galactosidases são enzimas com mais de um domínio, sendo que os ácidos glutâmicos catalíticos geralmente estão localizados nas fitas beta 4 e 7 de domínio do tipo barril-TIM.¹⁴ Essas enzimas foram nomeadas de tal maneira pela sua capacidade em hidrolisar pequenos galactosídeos (como a lactose) em riboses. Apesar disso, as β -galactosidases podem também atuar em diversos carboidratos contendo galactosídeos, como polímeros e oligossacarídeos.¹⁵ Dessa maneira, estão envolvidas em passos fundamentais do metabolismo de carboidratos, sendo encontradas em praticamente todos os organismos, havendo representantes em todos os reinos da vida. Com base na identidade sequencial, o CAZy distribuiu os enovelamentos distintos de β -galactosidases nas famílias das GH 2, 35 e 42.¹²

As GH2 incluem enzimas com atividades do tipo β -galactosidases, (EC 3.2.1.23), β -mannosidase (EC 3.2.1.25), β -glucuronidase (EC 3.2.1.31), α -L-arabinofuranosidase (EC 3.2.1.55), endo- β -mannosidase (EC 3.2.1.152) e exo- β -glucosaminidase (EC 3.2.1.165). O reino protista concentra a maior parte dos representantes da família, porém podemos encontrar representantes dessas enzimas em todas os reinos. No caso das β -galactosidases, vemos que as enzimas possuem cinco domínios, sendo que o terceiro domínio contém um barril do tipo TIM (Figura 2). Tais β -galactosidases geralmente possuem em torno de 1000 aminoácidos e formam tetrâmeros em solução. Não se sabe ao certo o motivo pelo qual essas enzimas formam grandes complexos, porém foi sugerido que tais enzimas são derivadas de β -galactosidases mais simples, com um sítio alongado capaz de se ligar a uma grande variedade de oligossacarídeos. Nesse contexto, outros domínios acessórios foram somados a arquitetura da molécula a fim de restringir o tamanho do sítio ativo, favorecendo a atividade da enzima sob dissacarídeo, como a alolactose.⁷ Outra possibilidade é de que, por estarem diretamente envolvidas com a ativação gênica, grandes complexos facilitariam a célula regular os níveis de enzima citosólicos.¹⁶ Apesar de não haver uma função clara para cada um dos domínios acessórios, já foi demonstrado que a remoção de pequenos fragmentos das regiões N e C terminal da proteína levam a formação de dímeros inativos.¹⁷ A família GH2 possui atualmente 20 estruturas cristalográficas depositadas no *Protein Data Bank* (PDB), dentre as quais, cinco são β -galactosidases, incluindo a enzima lacZ de *E. coli*.^{8,18}

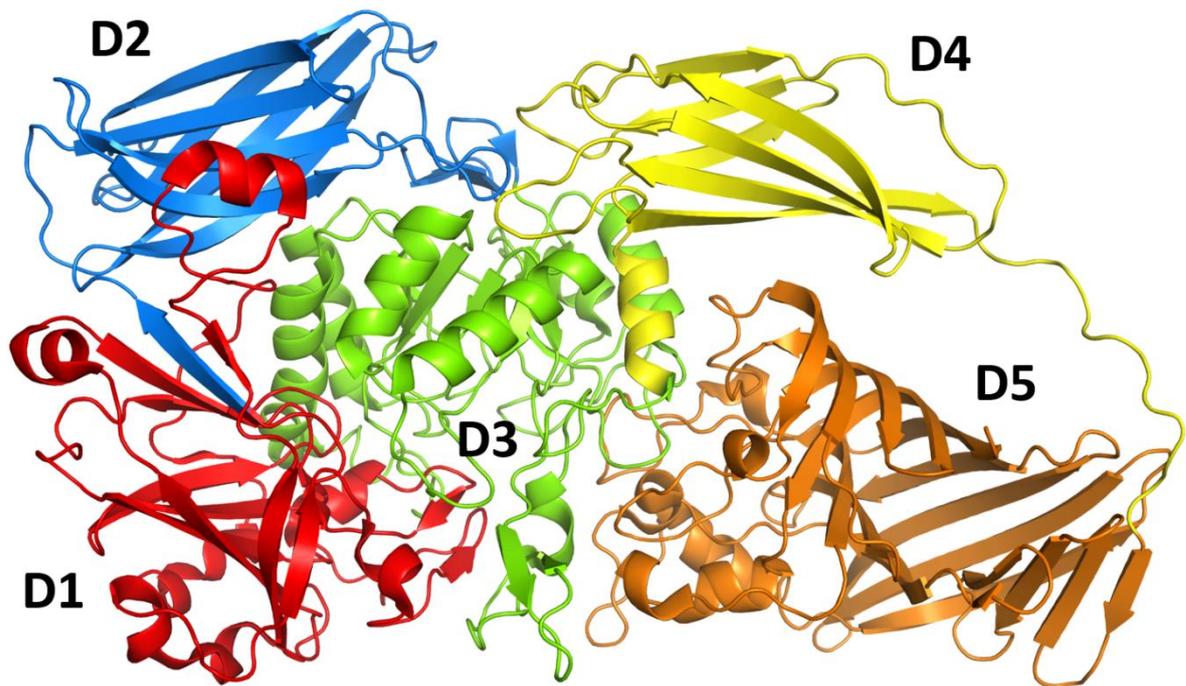


Figura 2 - Estrutura da enzima lacZ. No modelo cada domínio está destacado em uma cor e marcado de acordo com a ordem a partir da região N terminal. PDBid 4TTG.

Fonte: Elaborada pelo autor.

Além de β -galactosidase, são também encontradas na família 35 das hidrolases de glicosídeos outras $\text{exo-}\beta$ -glucosaminidase (EC 3.2.1.165), $\text{exo-}\beta$ -1,4-galactanase (EC 3.2.1.-) e β -1,3-galactosidase (EC 3.2.1.-). Nessa família encontramos a maioria das β -galactosidas de eucariotos, incluindo as de plantas, fungos e animais. Possuem uma quantidade variável de domínios, sendo que, podemos encontrar exemplares com dois, três ou cinco domínios (Figura 3).¹⁹⁻²⁰ Mesmo assim, os resíduos catalíticos invariavelmente aparecem em um domínio do tipo barril TIM. Atualmente a família possui oito estruturas de β -galactosidase depositadas no PDB, incluindo a enzima de humanos (PDBid 3THC, não publicada).

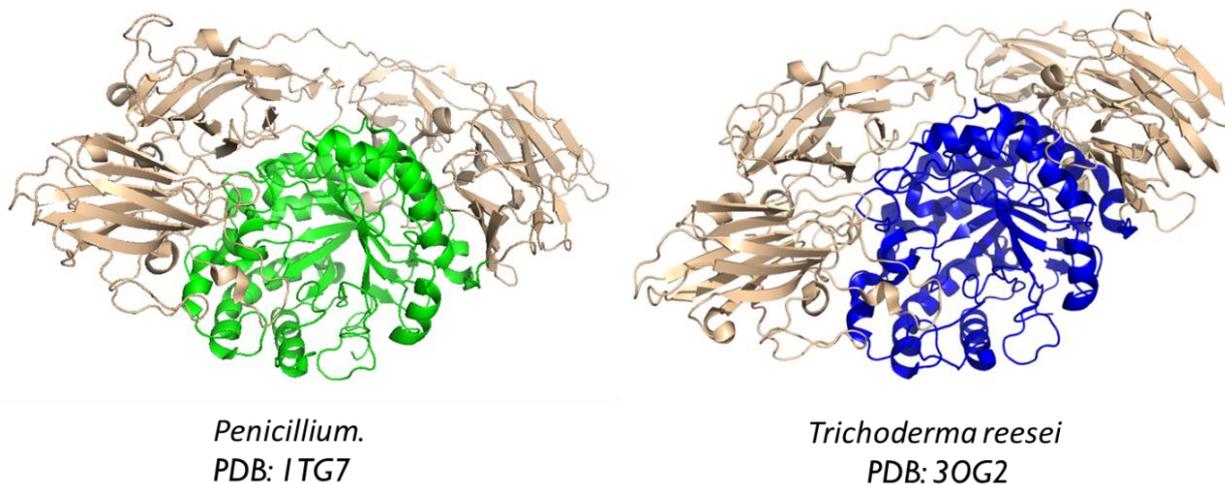


Figura 3 - β -galactosidases multidomínios da família GH35, pertencentes aos eucariotos *Penicillium sp.* e *Trichoderma reesei*. Em destaque temos os domínios barril TIM, onde se encontram os resíduos catalíticos.

Fonte: Elaborada pelo autor

As β -galactosidases da família 42 são encontradas principalmente em organismos procariotos, sendo que, poucos exemplos aparecem em fungos e arqueobactérias. Formadas por três domínios (Figura 4), essas enzimas possuem como característica marcante o fato de se organizarem em uma estrutura trimérica em forma de barril.²¹⁻²² Atualmente, existem quatro exemplares dessas enzimas com estruturas depositadas no PDB.

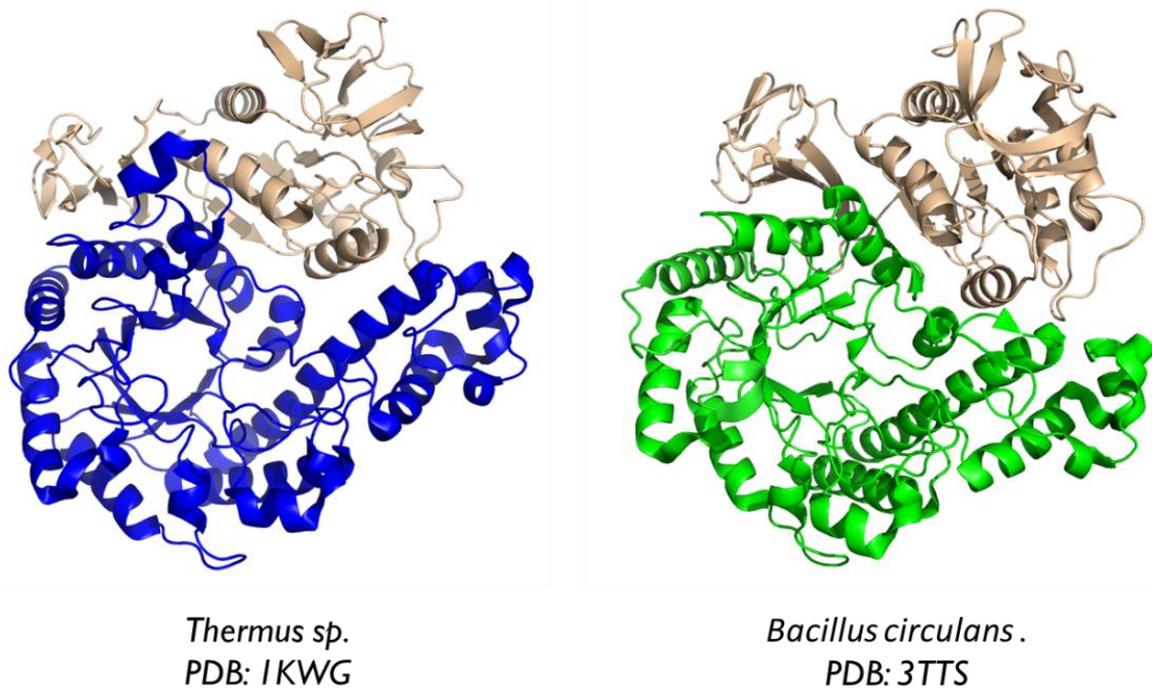


Figura 4 - β -galactosidases contendo dois domínios da família GH42, pertencentes aos organismos *Thermus sp.* e *Bacillus circulans*. Em destaque temos os domínios barril TIM, onde encontramos os resíduos catalíticos.

Fonte: Elaborada pelo autor.

1.3 Mecanismo e atividade das β -galactosidases

Durante seus estudos sobre a estereo química de produtos de reação enzimática, Daniel Koshland,²³ propôs um mecanismo que explicasse como as enzimas eram capazes de processar seus substratos. Segundo ele, a hidrólise enzimática de uma ligação glicosídica, feita por qualquer enzima, tem como base dois aminoácidos fundamentais: um aminoácido que atue como o doador de prótons (ácido), e outro como o nucleófilo (base) da reação. Esse processo pode ocorrer através de dois mecanismos distintos: o mecanismo de retenção e o mecanismo de inversão do carbono anomérico.²³

Em ambos os mecanismos, a posição do doador de prótons é a mesma, a cerca de uma ligação de hidrogênio de distância (2,2 ~ 2,6 Å) do oxigênio glicosídico. No caso das enzimas com mecanismo do tipo retenção, o nucleófilo está também a uma distância similar a uma ligação de hidrogênio, possibilitando que este atue diretamente na interação substrato-doador de prótons. Nesse caso, o carbono anomérico sofre uma dupla inversão (uma para cada interação com o doador de prótons), o que faz com que o carbono anomérico do produto tenha

uma estereo-química similar ao do substrato.²⁴ Já no caso das enzimas com o mecanismo do tipo inversão, a interação nucleófilo-carbono anomérico é mediada por uma água, uma vez que o nucleófilo está a uma maior distância do carbono anomérico. Isso faz com que haja uma inversão na conformação estereo-química do carbono anomérico do produto em relação ao substrato. Nas enzimas do tipo retenção, os aminoácidos catalíticos estão a cerca de 5,5 Å de distância, enquanto que nas do tipo inversão essas distâncias são da ordem de ~10 Å.²⁴ Apesar de haverem exceções, na maioria dos casos esses aminoácidos são tipicamente ácidos glutâmicos ou aspárticos.²⁴

As β -galactosidases são, tipicamente, enzimas com o mecanismo do tipo retenção. No primeiro passo dessa reação, o resíduo ácido ataca o carbono anomérico do substrato, formando o intermediário covalente e liberando uma molécula de água. Em seguida, a base catalítica age como um catalisador, facilitando a hidrólise do complexo enzima-substrato por uma molécula de água (Figura 5). O primeiro passo dessa reação é chamado de etapa de glicosilação, enquanto que o segundo passo é chamado de deglicosilação.²³

Apesar de comumente agirem como hidrolases, as β -galactosidases são também conhecidas por sua capacidade de agir como transferases. Isso ocorre pelo mesmo mecanismo da hidrólise, com a diferença de que neste caso a base catalítica utiliza o próton proveniente do terminal redutor de um outro açúcar no meio. Neste caso, o complexo enzima-galactosil recebe o radical, que pode ter tamanhos de cadeia variados, e forma diferentes tipos de galactooligossacarídeos (GOS).²⁵ Essa via do mecanismo é chamada de transglicosilação (Figura 5). Quando a lactose é utilizada como produto, os GOS são geralmente formados por galactose e glicose, podendo variar em sua composição e tamanho de cadeia (geralmente 3 à 8 monossacarídeos).

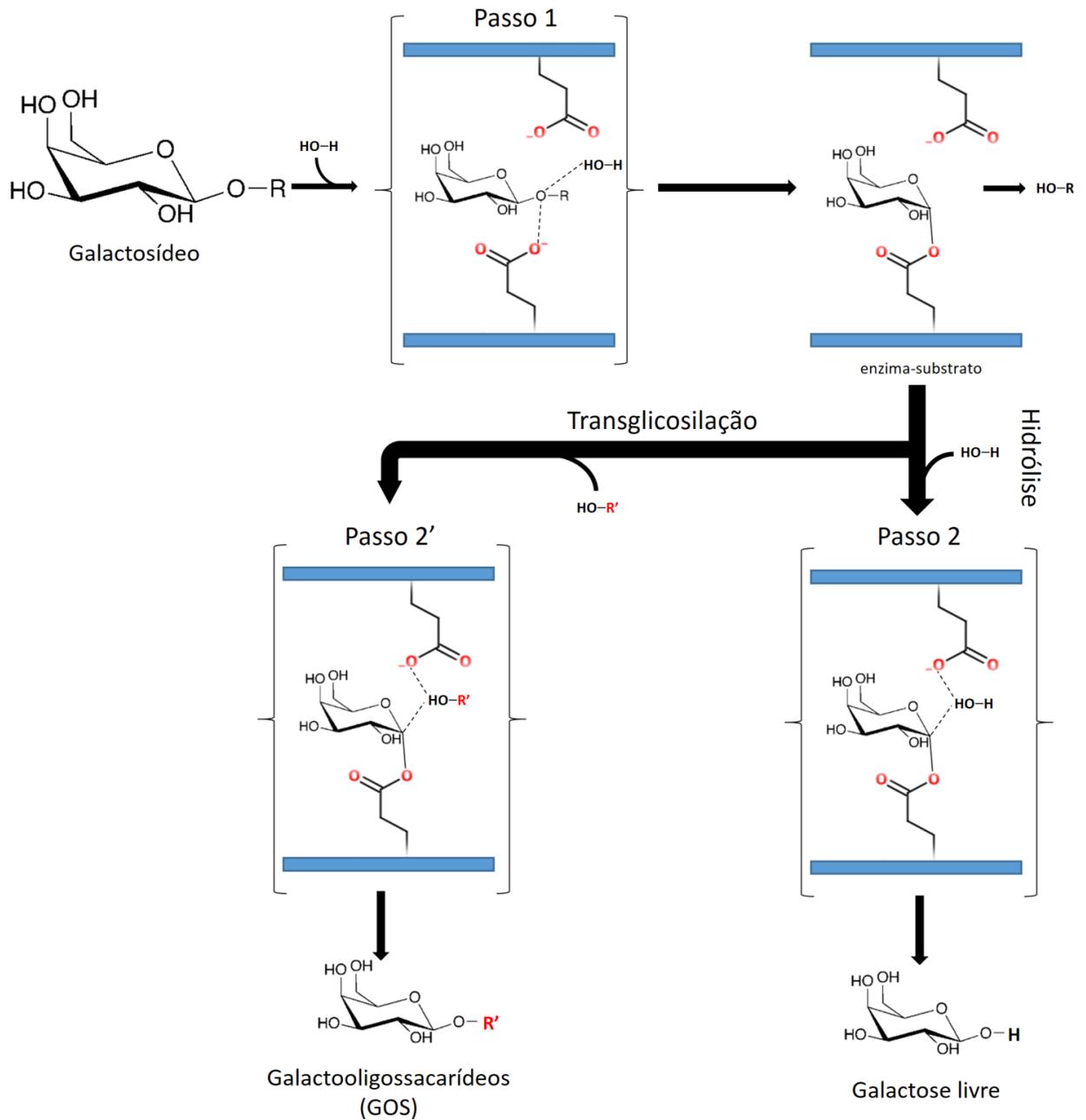


Figura 5 - Esquema do mecanismo de retenção utilizado pelas β -galactosidases. No passo 1, o galactosídeo sofre o ataque nucleofílico por uma água, formando um intermediário covalente com a enzima, e liberando o radical hidrolisado. A partir daí a reação pode seguir dois caminhos: a hidrólise completa do galactosídeo consumindo uma água (Passo 2); ou a transgalactosilação, onde o doador de prótons da hidrólise é proveniente do terminal redutor de um carboidrato. Nesse caso, ocorre a formação de galactooligosacarídeos como produto da enzima.

Fonte: Elaborada pelo autor.

1.4 As β -galactosidases na medicina e na indústria

A intolerância a lactose, ou hipolactasia, é uma doença comum em que o organismo tem dificuldades de processar a lactose. A principal causa dessa doença são as baixas concentrações de β -galactosidases na parede da mucosa do intestino delgado.²⁶ A lactose é o carboidrato mais comum encontrado no leite, e é basicamente o dissacarídeo formado pela galactose e a glucose. Quando o intestino delgado não digere devidamente a lactose, ela passa para o intestino grosso, onde é fermentada pela microbiota do cólon. Esse processo leva a produção de ácidos graxos de cadeia curta, além de gases como H₂, CO₂ e metano.²⁷ Esses produtos podem causar sintomas que incluem dor abdominal, cólica, diarreia e náusea.²⁸ A prevalência dessa doença varia na população, podendo ser inferior a 5% em alguns locais (como em países do norte da Europa), mas também podendo atingir 90% da população em outros (como países da África e Ásia). Numa perspectiva mundial, estima-se que cerca de 75% da população adulta exiba algum grau de intolerância a lactose.¹⁵

Numa perspectiva industrial, a hidrólise enzimática é a principal ferramenta para a diminuição das quantidades de lactose em produtos derivados do leite.¹⁵ Esse processo geralmente utiliza β -galactosidases imobilizadas em um reator de banho.²⁹ Apesar de ser um processo simples, e considerado relativamente caro, principalmente devido aos altos custos e quantidades de enzimas necessárias no protocolo.³⁰ Portanto, a caracterização de enzimas mais eficientes, estáveis e com perfis variados de atividade é extremamente importante para a indústria de laticínios.

Além do seu importante papel na hidrólise da lactose, as β -galactosidases também são amplamente utilizadas na indústria de alimentos funcionais. Isso porque os GOS são considerados moléculas prebióticas, sendo seletivamente fermentadas e causando alterações específicas na composição e/ou atividade da microflora intestinal, promovendo benefícios ao hospedeiro.³¹ Por não serem digeridos por humanos, os GOS representam uma rica fonte de substrato para bifidobactérias e outros organismos probióticos, como do gênero *Lactobacillus*. Esses organismos inibem o crescimento de microrganismos indesejáveis e patogênicos,^{32; 33} além de regularem o pH do trato digestivo e facilitarem a digestão.³⁴⁻³⁵

Moléculas de GOS são tipicamente sintetizadas a partir da atividade de β -galactosidases em sobre a lactose pelo processo de transglicosilação. Em 2005, somente o Japão comercializou cerca de seis mil toneladas de GOS, sintetizados a partir da atividade das β -galactosidases.³⁶ Além dos efeitos prebióticos dos GOS, a transglicosilação também diminui a

concentração de açúcar nos alimentos, a aumenta o conteúdo de fibras.¹⁵ Por essas razões, as β -galactosidases com alta capacidade de produção de GOS são um importante alvo comercial.

1.5 As β -galactosidases na produção de biocombustíveis

Nas últimas décadas, a desconstrução da parede celular das plantas tem sido um dos principais desafios enfrentados por pesquisadores ao redor do mundo para a produção de biocombustíveis. Isso não é uma surpresa, uma vez que a parede celular é o principal tecido de proteção das plantas, e sua rica e complexa rede de carboidratos evoluiu durante milhões de anos a fim de defender o organismo contra as interações com o ambiente.³⁷⁻³⁸ Apesar da complexidade de sua composição, a parede celular tem uma arquitetura padrão, formada por microfibrilas de celulose envolta por uma complexa rede de glicanos, chamada hemicelulose.³⁸ Essa rede é formada por uma complexa matriz de polissacarídeos (como glucano, xilano, xiloglucano, etc), além de glicoproteínas e pectinas, e muitas vezes lignina.^{39;}⁴⁰ Superar a complexidade química imposta pela parede celular tem sido a grande barreira para a tecnologia da produção de biocombustíveis.⁴¹⁻⁴²

Para superar esse problema são necessárias não uma, mas complexos coquetéis enzimáticos agindo sinergicamente.⁴³ Atualmente a hidrólise da parede celular requer um enorme número de enzimas diferentes, como endoglucanases, exoglucanases, β -glucosidases, galactanases, xilanases, β -manosidase, β -galactosidases, β -manase, α -L-arabinase, além de enzimas auxiliares e expansinas.⁴⁴⁻⁴⁷ No caso específico das β -galactosidases, já foi demonstrado que sua ação sinérgica com as endo-xiloglucanases são fundamentais para a desconstrução dos oligossacarídeos provenientes do xiloglucano em açúcares fermentescíveis.⁴⁸ Apesar dos esforços, a desconstrução da parede celular ainda é um passo limitante para a produção de biocombustíveis a um preço competitivo no mercado. Por esse motivo, a caracterização de novas enzimas permanece uma área de grande interesse para a indústria biotecnológica.

1.6 Os organismos de estudo e objetivos do projeto

O organismo *Xanthomonas campestris pv. campestris* é uma bactéria gran-negativa, conhecida por sua patogenicidade a culturas de citrus. Essa bactéria é causadora do cancro cítrico, também chamado de “*black root disease*”, uma praga que afeta culturas de citros no mundo todo e causa elevados prejuízos para a agricultura.⁴⁹ Além dos interesses comerciais, já foi demonstrado que esse organismo possui uma elevada atividade β -galactosídica em meios ricos em lactose.⁵⁰ Curiosamente, tal organismo possui três diferentes β -galactosidasas da família 35, o que não é comum em procariotos. Sendo assim, decidimos investigar as diferenças funcionais dessas enzimas e tentar entender o papel das mesmas no metabolismo de carboidratos do organismo.

Outro organismo que foi estudado foi a bactéria *Bifidumbacterium bifidum s17*. As Bifidobactérias são organismos sacarolíticos, com as habilidades de fermentar glicose, galactose e frutose, que normalmente, habitam o trato gastrointestinal,⁵¹ constituindo um dos maiores grupos de bactérias da microbiota fecal (em torno de 4 % do total) em adultos na população européia.⁵² Dados fisiológicos confirmam que essas bactérias são capazes de utilizar uma ampla variedade de carboidratos, como GOS,³³ xilo-oligossacarídeos,⁵³ pectina,⁵⁴ e fruto-oligossacarídeos.⁵⁵ Por tais razões, acreditamos que as β -galactosidasas de *Bifidumbacterium bifidum* possuem um grande potencial biotecnológico.

Tendo isso em vista, o objetivo do nosso estudo é compreender as bases moleculares que garantem as funções e propriedades bioquímicas das β -galactosidasas de *Xanthomonas campestris pv. campestris* e *Bifidumbacterium bifidum* com potencial biotecnológico. Mais especificamente, nosso objetivo é realizar a produção em larga escala, seguida da caracterização bioquímica destas enzimas. Em seguida utilizaremos a técnica de difração de raios X por monocristais para a elucidação do modelo estrutural para estas enzimas. Por fim, nossas análises terão como objetivo relacionar a estrutura das enzimas as singularidades funcionais que possam vir a aparecer em cada alvo caracterizado.

CAPÍTULO 2

Materials & Métodos

2 Materiais e Métodos

2.1 Seleção dos genes das β -galactosidases.

A seleção dos genes de possíveis enzimas com atividade β -galactosídica foi feita com base na anotação por homologia realizada pelo NCBI. Na Tabela 1, podemos ver o resumo de informações sobre tais genes, bem como a nomenclatura que utilizaremos para esse trabalho (tanto para as enzimas quanto para os genes).

Tabela 1 - Informações sobre os genes de β -galactosidases selecionados

Código de Acesso (GenBank)	Família (CAZY)	Nomenclatura adotada	Organismo
AAM42167.1	GH35	XCC_2895	<i>Xanthomonas campestris pv. campestris</i>
AAM40554.1	GH2	XCC_1256	<i>Xanthomonas campestris pv. campestris</i>
AAM43337.1	GH2	XCC_4116	<i>Xanthomonas campestris pv. campestris</i>
NP_637758.1	GH35	XCC_2404	<i>Xanthomonas campestris pv. campestris</i>
NP_637121.1	GH35	XCC_1754	<i>Xanthomonas campestris pv. campestris</i>
ADO52393.1	GH2	Bb_lacZ1	<i>Bifidumbacterium bifidum</i>
ADO53518.1	GH42	BbgII	<i>Bifidumbacterium bifidum</i>
ADO53611.1	GH2	Bb_lacZ3	<i>Bifidumbacterium bifidum</i>
YP_003938286.1	GH2	Bb_lacZ4	<i>Bifidumbacterium bifidum</i>

Fonte: Elaborada pelo autor.

Tais genes foram clonados a partir dos DNAs genômicos da cepa ATCC33913 de *Xanthomonas campestris pv. campestris* (disponível na biblioteca de organismos de nosso laboratório) e da cepa S17 do organismo *Bifidobacterium bifidum* (gentilmente cedido pelo professor Christian Riedel - Ulm University, Alemanha). Os DNAs genômicos dos organismos foram extraídos com o *kit* comercial “Wizard® Genomic DNA Purification Kit” (Promega®), seguindo as recomendações do fabricante.

2.2 Clonagem através do método LIC.

O sistema de clonagem adotado foi método independente de ligação (LIC – *Ligation-Independent Cloning*).⁵⁶ Esse sistema se baseia na atividade exonucleásica 3'→5' da enzima T4 DNA Polimerase para criar regiões de fita simples de 10 a 15 nucleotídeos nas pontas 5' do vetor. Paralelamente, são gerados produtos de PCR dos genes desejados, com caudas de fitas-simples a 5', complementares às do vetor. Então são adicionadas extensões apropriadas nos *primers* de tal modo que o fragmento amplificado não tenha timina (no caso do vetor) ou adenina (no caso do gene) nos 10 últimos nucleotídeos. Após o tratamento com T4 DNA Polimerase na presença somente de dATP (para o gene) ou somente dTTP (para o vetor), a atividade exonuclease 3'→5' se encarrega de remover bases no sentido 3'→5', até se deparar com numa adenina (gene) ou timina (vetor). O pareamento do vetor com o inserto é então feito na ausência de ligase, simplesmente misturando-se os dois fragmentos de DNA (Figura 6). Tal mistura é então utilizada para transformar bacterias *E. coli* DH5α competentes com os respectivos plasmídeos.

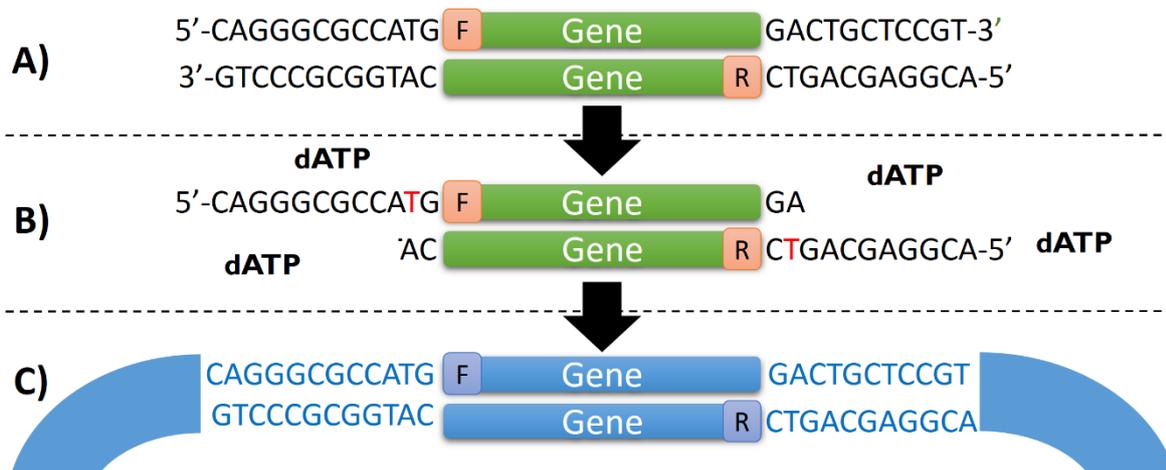


Figura 6 - Esquema ilustrativo do método LIC utilizado para clonagem. Em A, temos o gene de interesse (verde) amplificado por primers (salmão) contendo uma extensão de bases complementares ao vetor. Em B, o produto do PCR é tratado com a enzima Phusion polimerase na presença de dATP, deixando para trás um fragmento de fita simples. Em C, vemos o pareamento do produto tratado com o vetor de interesse (azul), tratado de maneira recíproca.

Fonte: Elaborada pelo autor.

Ao planejar os *primers* para a reação, incluímos a sequência 5'-CAGGGCGCCATG-3' na extremidade 5' de cada *primer senso (fw)*, e a sequência 5'-GACCCGACGCGGTTA-3' na

extremidade 5' de cada *primer anti-senso (rv)*. Os genes de interesse foram amplificados por PCR com a enzima Phusion® High-Fidelity DNA Polymerase (NEB, EUA) segundo as especificações do fabricante. Após a amplificação, os fragmentos foram tratados com a enzima T4 DNA Polimerase (NEB, USA) na presença de T4 polimerase buffer 1x, DTT 4,0 mM e dTTP 2,5 mM, num volume final de 20 µL. A reação foi incubada por 30 min a 22 °C e inativada por 20 min a 75 °C. O mesmo tratamento é feito com o vetor, porém utilizando dTTP 2,5 mM, de forma a gerar um fragmento complementar as extensões dos *primers*.

Os genes BbgII e Bb_lacZ3 foram clonados pelo método LIC no vetor pPROEX-Hta (Invitrogen®), adaptado para as extensões de *primers* descritas acima. Já os outros genes foram clonados no vetor de expressão pET-TrxA/LIC (gentilmente cedido por Arie Geerlof - Helmholtz Zentrum, Alemanha).⁵⁷ Detalhes sobre os vetores utilizados podem ser vistos na Tabela 2. As sequencias dos clones foram previamente checadas por sequenciamento.

Tabela 2 - Detalhes dos vetores utilizados.

	Promotor	Antibiótico	Tag(s)	Sítio de clivagem	Origem de replicação
pPROEX-HTa	trc	Amp	N-His	TEV	pBR322
pET-TrxA/LIC	T7/lac	Kan	N-TrxA + N-His	TEV	pBR322

Fonte: Elaborada pelo autor.

2.3 Transformação bacteriana.

Os plasmídeos contendo os fragmentos de interesse foram utilizados para transformar células de *Escherichia coli* pelo método de choque-térmico em diferentes linhagens de expressão. Utilizamos a linhagem DH10B para a propagação dos vetores logo após a transformação. Os plasmídeos de interesse foram então purificados pelo *kit* comercial PureYield™ Plasmid Miniprep System (PROMEGA®), segundo as especificações do fabricante. Em seguida, transferimos esses plasmídeos para as linhagens de expressão Rosetta(DE3) e BL21(DE3) (New England Biolabs®).

2.4 Mutação sítio-dirigida.

Os *primers* utilizados no passo de mutação sítio-dirigida foram baseados no método de PCR inverso. Tal método se baseia na amplificação do plasmídeo inteiro utilizando-se oligonucleotídeos que pareiam entre si no plasmídeo, dirigindo a amplificação em sentidos opostos. A amplificação é realizada com a enzima Phusion High-Fidelity DNA Polymerase (Thermo-Scientific) gerando o plasmídeo na forma linear com pontas não fosforiladas. Em seguida, o produto deve ser fosforilado com a enzima T4 Polynucleotide Kinase (NEB), circularizado por ligação com a T4 DNA ligase (NEB) e utilizado para a transformação de bactérias *E. coli* DH10B. Nesse método, o plasmídeo selvagem é propagado aritmeticamente enquanto que o plasmídeo contendo a mutação é propagado exponencialmente, resultando em praticamente 100% de cópias com a mutação após 30 ciclos de PCR.⁵⁸ A mutação então é confirmada através do sequenciamento dos plasmídeos após sua propagação.

2.5 Expressão em larga escala.

Para a expressão em larga escala do clone BbgII utilizamos os clones transformados na célula competente BL21(DE3). Para isso, preparamos inóculos de 5,0 mL de meio LB (triptona 1%, extrato de levedura 0,5% e NaCl 1%) contendo 100 $\mu\text{g}\cdot\text{mL}^{-1}$ de ampicilina, que foram mantidos à 37 °C durante 12 h, sob a agitação de 250 rpm. Após esse período, as células desse foram transferidas para o mesmo meio, com um volume inicial 200 vezes maior que o anterior. Esse meio foi mantido sob agitação à 37 °C, até a D.O. 600 μm atingir um valor entre 0,6 e 1,0. Então adicionamos 1,0 mM de isopropil β -D-1-thiogalactopiranosídeo (IPTG) ao meio, e continuamos a agitação à 18 °C por cerca de 12 h.

Após a expressão, as bactérias foram sedimentadas por centrifugação (4000 rpm em rotor SLA-3000, por 15 min à 4 °C). O sedimento foi então ressuspensionado no tampão A (fosfato 50 mM, NaCl 50 mM, pH 7,4). A lise celular foi feita com adição de lisozima (10 $\text{mg}\cdot\text{L}^{-1}$ cultura), seguido de sonicação. A fração solúvel foi então separada por centrifugação (17.000 rpm, 50 min, à 4 °C) e filtrada, onde o sobrenadante foi utilizado para os passos de purificação.

Para a expressão em larga escala dos clones XCC_1754, XCC_2404 e XCC_2895 utilizamos os clones transformados em células Rosetta (DE3). O crescimento e a expressão foram realizados no meio autoindutor ZYM5052 ((NH₄)₂SO₄ 0,33%; KH₂PO₄ 0,68%; Na₂PO₄ 0,7%; Glicerol 0,5%; Glicose 0,5%; Lactose 0,2%; Triptona 1%; Extrato de Levedura 0,5%),⁵⁹, durante 24 h à 20 °C, sob agitação constante. A sedimentação e lise celular foi realizada da mesma maneira que para a enzima BbgII.

2.6 Purificação das enzimas.

Os extratos celulares contendo as enzimas BbgII, XCC_1754, XCC_2404 e XCC_2895 foram purificados em dois passos. Ambos os vetores utilizados acrescentam uma cauda de histidina na região N-terminal da proteína, sendo que o vetor pET-TrxA/LIC acrescenta ainda uma proteína acessória (tioredoxina) entre a cauda e o produto de expressão. Portanto o primeiro passo de purificação realizado foi a cromatografia de afinidade por níquel. Para isso, utilizamos a resina Níckel Talon Superflow (Clontech®), equilibrada tampão A. Após a lavagem com o mesmo tampão, a eluição foi realizado com um gradiente crescente de tampão B (Na₂HPO₄ 6,4%, NaH₂PO₄ 1,6%, NaCl 150 mM, imidazol 300 mM, pH 7,4). A remoção da tioredoxina foi feita pela enzima TEV protease à 18 °C, durante um período de 24 horas. A cauda de histidina da enzima BbgII não foi removida.

Para o segundo passo, utilizamos a coluna HiLoad Superdex 200 16/60 (GE Healthcare®), equilibrada com o tampão C (Hepes 30 mM, NaCl 150 mM, pH 7,5). A pureza das amostras foi checada através do gel SDS 15%

2.7 Cristalização, co-cristalização, *soaking* e coleta de dados das enzimas.

Atualmente o método mais utilizado para a cristalização de macromoléculas se baseia no método de difusão de vapor associado a estratégia de matriz esparsa.⁶⁰ No método de difusão de vapor, a proteína em questão é misturada a uma solução contendo no mínimo um agente precipitante. Essa mistura é então selada em um sistema contendo a solução precipitante, de forma que haja troca de vapor entre a mistura e a solução mãe. Por causa da diluição, a

concentração de precipitante da mistura é menor do que a solução mãe, o que leva a uma difusão maior de vapor no sentido mistura-solução mãe (Figura 7).⁶⁰

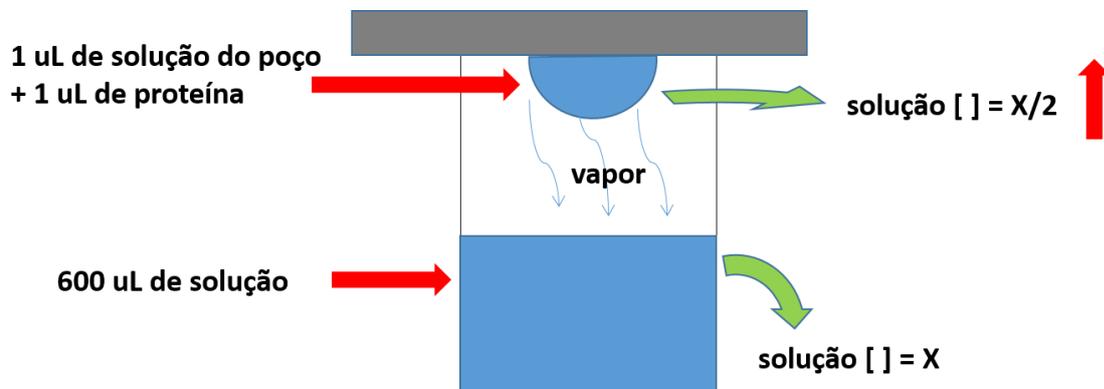


Figura 7 - Esquema do método de difusão de vapor. Uma mistura de proteína e solução-mãe é selada no sistema contendo a solução mãe pura. A diferença nas concentrações de precipitante entre as soluções favorece a troca de vapor no sentido mistura-solução mãe.

Fonte: Elaborada pelo autor.

Nessa estratégia, utilizamos *kits* comerciais com soluções diversas para investigar dentro de um universo de condições possíveis o comportamento da macromolécula em questão. No caso, empresas se baseiam na taxa de sucesso e reincidência de precipitantes e tampões diversos para desenvolver *kits*, possibilitando uma enorme quantidade de informação sobre o comportamento físico-químico da macromolécula e, frequentemente, até mesmo cristais (Figura 8).⁶⁰ No caso de macromoléculas hidrofílicas em solução aquosa, o comportamento em uma condição específica geralmente pode ser previsto pelo diagrama de fases de sólidos em solução (Figura 8).

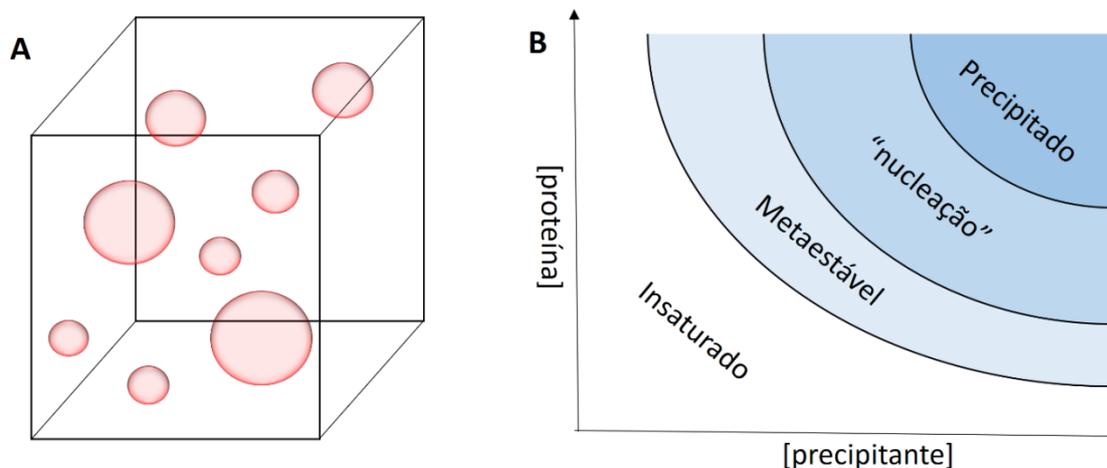


Figura 8 - Esquema do método de matriz esparsa e diagrama de fases de uma macromolécula. Em A, vemos a representação de um universo de possibilidades (cubo) onde algumas regiões foram exploradas (esferas) através do método; Em B, o diagrama de fases de uma proteína hidrofílica.

Fonte: Elaborada pelo autor.

Um experimento ideal de cristalização começa com uma concentração de proteína suficiente para que o sistema se encontre próximo a fase metaestável do diagrama (Figura 8). Nessas condições, todas as moléculas do sistema (tanto macromoléculas, quanto precipitantes) se encontram solvatadas. Uma vez que o sistema perde água por difusão de vapor, as camadas de solvatação de ambos, proteína e precipitante, começam a se reduzir, até o ponto onde estes começam a competir pelas moléculas de água do sistema. Quando não há mais moléculas de água disponível para satisfazer o sistema, este se torna instável e passa a reduzir a energia formando interações entre seus componentes. Tais interações frequentemente levam a precipitação dos componentes do sistema. Porém, em condições particulares, essas interações podem vir a levar a formação de sólidos organizados de forma periódica, tais como cristais.

As enzimas BbgII, XCC_2404, XCC_1754 e XCC_2895 foram cristalizadas pelo método de matriz esparsa, utilizando *kits* comerciais e o robô de cristalização *HoneyBee* (Digilab®). Os cristais foram otimizados com base nas condições iniciais. Os cristais da enzima BbgII foram também co-cristalizados com o ligante natural β -D-galactose, que foi adicionada a uma concentração de 10 mM nas gotas iniciais. Para a derivatização dos cristais de BbgII, utilizamos a técnica de *soaking*, onde acrescentamos cerca de 2,0 mM de metilmercúrio a cristais já formados.

Antes da coleta os cristais foram crio-protetidos com 15% de etileno glicol (no caso das enzimas ou XCC_2404, XCC_1754 e XCC_2895), ou com 8% polietilenoglicol (PEG) 200 (no caso dos cristais de BbgII). Os cristais de BbgII foram coletados em um micromax-007 R-AXIS IV (Rigaku®), com fluxo constante de nitrogênio (-173 °C). Os outros cristais foram enviados para o *National Synchrotron Light Source* (NSLS - Nova Iorque, Estados Unidos) e coletados na linha de luz X4C, equipado com um detector MAR 165 CCD.

2.8 Resolução das Estruturas

Os conjuntos de dados foram indexados, integrados e escalonados com o programa XDS,⁶¹ onde os limites de resolução foram estabelecidos pelo método de Karplus.⁶² O programa Aimless,⁶³ foi utilizado na conversão das intensidades em amplitudes e cálculo das estatísticas. A qualidade dos dados foi checada com o programa phenix.Xtriage.⁶⁴ Para as substituições moleculares, utilizamos o programa Phaser.⁶⁵ Já para a resolução da estrutura da

enzima BbgII, utilizamos a técnica de *single anomalous diffraction* (SAD), onde utilizamos os programas SHELXC/D,⁶⁶ para encontrar os sítios de átomos pesados, e o programa autoSHARP,⁶⁷ para o cálculo das fases e para passos de modificação de densidade. A construção dos modelos, refinamento e validação das estruturas foram conduzidas pelos programas Autobuild,⁶⁴ phenix.refine,⁶⁸ Coot,⁶⁹ e Molprobitry,⁷⁰ As figuras e análises foram feitas com uso do programa PyMOL.

2.9 Atividade e cinética das enzimas

A atividade das enzimas foi medida por dois métodos. O primeiro, utilizando o substrato sintético orto-nitrofenil- β -D-galactopíranosídeo (*o*NPG), um substrato sintético análogo à lactose. Esse substrato possui uma ligação glicosídica entre uma molécula de galactose e uma molécula de nitrofenil. Quando hidrolisada, a molécula de nitrofenil absorve fortemente na faixa de 420 nm, o que facilita nossas medições. Nessa reação, 90 μ L do substrato foram adicionados em 10 μ L de enzima diluída a 1 unidade por mL (definida como a quantidade de enzima que degrada 1 μ m de substrato em 1 minuto). Em seguida, a reação foi parada com adição de 150 μ L de 1,0 M de carbonato de sódio. Essas reações foram então lidas na faixa dos 420 nm. Devido à simplicidade, esse método foi utilizado para os testes de estabilidade térmica em tampão Z (Na_2HPO_4 6,0%, NaH_2PO_4 0,16%, pH 6,0). Também utilizamos esse método para medirmos a estabilidade das enzimas em diferentes pH.

Para medirmos a capacidade hidrolítica das enzimas para com seu substrato natural, a lactose, utilizamos o *kit* comercial glicose-oxidase-peroxidase (Sigma®), segundo as especificações do fabricante. Nos estudos cinéticos, as reações foram medidas durante 10 minutos, em diferentes concentrações de substratos. Para os cálculos consideramos a velocidade de reação (V) como sendo igual a velocidade inicial (V_0) da enzima. O protocolo foi adaptado no robô Freedom EVO (TECAN®). Todos os testes cinéticos foram feitos em tampão fosfato 50 mM pH 6,0 e as constantes de Michaelis-Menten (K_M), inibição (K_i) e velocidade máxima (V_{max}) foram calculadas pelo com ajuda do programa OriginPro (versão 8.0).

2.10 Medidas de fluorimetria de varredura diferencial (*ThermoFluor*)

A estabilidade térmica da enzimas e seus mutantes foi medida em experimentos de termo fluorimetria.⁷¹ Para isso, 10 μM das enzimas foram misturados com diferentes tampões, e então foi acrescentado 1,0 μL da sonda Sypro Orange (Invitrogen®) na diluição recomendada pelo fabricante (1:2000). As misturas foram analisadas em um CFX96 Real Time System (Bio-Rad®). As variações de fluorescência foram monitoradas a 490 μm (excitação) e a 530 μm (emissão). Todas as medidas foram feitas em triplicata. O cálculo da temperatura de *melting* (T_m) foi realizado com o software GraphPad Prism (versão 5.0).

2.11 Determinação dos produtos de reação das enzimas

Para verificarmos a produção de GOS pelas enzimas, utilizamos o método de *High-Performance Anion-Exchange Chromatography* (HPAEC), utilizando o aparelho Dionex BioLC® acoplado com uma bomba GP50, e o eletrodo ED50 (ThermoScientific, USA). Para isso, as amostras contendo 200 g.L^{-1} de lactose em tampão Z foram incubadas com as 0.1 mg.mL^{-1} de enzima (ou seus respectivos mutantes). As reações foram mantidas à 37 °C durante 1 h, e fervidas em seguida durante 5 min. Tais amostras foram injetadas na coluna CarboPac PA1 equilibrada com NaOH 125 mM, e em seguida eluidas com NaOH 125 mM e $\text{C}_2\text{H}_3\text{NaO}_2$ 500 mM.

CAPÍTULO 3

Resultados & Discussão I

Clonagem, purificação, cristalização & coleta de dados

3 Resultados e Discussão

3.1 Análise das sequências e considerações gerais

Os genes identificados na Tabela 1 foram analisados com ferramentas de bioinformática, a fim de estimarmos a melhor estratégia para a clonagem e produção dos mesmos. De agora em diante iremos focar nos resultados obtidos com as enzimas expressas pelos genes BbgII, XCC_1754, XCC_2404 e XCC_2895. Para buscarmos homólogos sequencias e/ou estruturais, foram utilizados os algoritmos de busca local, como *Blast* e o *HHPred*.⁷²⁻⁷³ Alinhamentos e comparações entre as sequencias foram realizadas no programa *CLUSTAL W*.⁷⁴

Utilizamos a ferramenta *SignalP*,⁷⁵ para predizer a presença de peptídeo sinal na porção N-terminal das sequências de estudo. Caso presente ($p > 0,15$), o mesmo foi desconsiderado no planejamento dos *primers*. Para o planejamento dos *primers* fw e *primers* rv, tanto para amplificação dos genes nativas quanto dos mutantes elaborados posteriormente, utilizamos o programa *HTP-OligoDesigner*⁷⁶ desenvolvido em nosso grupo.

As sequências dos genes de interesse foram utilizadas então para calcularmos o número de pares de bases (pb) presentes em cada gene, e o número de aminoácidos de cada proteína madura. Outros parâmetros físico-químicos, como massa molecular (MM) e o coeficiente de extinção molar (ϵ) das sequencias proteicas foram estimados teoricamente com o programa *ProtParam*.⁷⁷ Essas informações, assim outras estão resumidas na Tabela 3.

Tabela 3 - Informações gerais sobre os alvos de estudo.

Alvos	Peptídeo sinal	pb	Aminoácidos	MM (kDa)	k^e ($M^{-1}.m^{-1}$)
BbgII					
N160A	Não	2067	689	77,2	178,1
E161A	Não	2067	689	77,2	178,1
Y289F	Não	2067	689	77,2	178,1
Y371F	Não	2067	689	77,2	178,1
E320A	Não	2067	689	77,2	178,1
E368A	Não	2067	689	77,2	178,1
XCC_1754					
G294P	Não	1668	557	60,1	104,2
G302P	Não	1668	557	60,1	104,2
P295_P301dele	Não	1668	557	60,1	104,2
XCC_2404					
	Sim	1533	510	55,2	92,8
XCC_2895					
	Sim	1773	590	65,6	115,2

Fonte: Elaborado pelo autor

3.2 Clonagem dos genes

Os genes foram clonados pelo método LIC, sendo que os genes BbgII e Bb_lacZ3 foram clonados no vetor pPROEX-Hta, enquanto os outros genes foram clonados no vetor pET-TrxA/LIC. Tais vetores contendo os genes de interesse foram utilizado para transformação de bactérias *E. coli* próprias para expressão. A expressão da enzima BbgII foi realizada na cepa BL21(DE3), e o restante das enzimas foi expressa na cepa Rosetta(DE3).

Como foi discutido no item 2.2, uma das vantagens do método LIC é o fato de ser possível realizar a clonagem em poucos passos e sem o uso de enzimas de restrição. Outra vantagem é que o mesmo *overhang* pode ser adicionado em genes diferentes, permitindo a clonagem de muitos alvos em paralelo, sem a necessidade de utilizarmos protocolos alvo específicos.⁷⁸ O primeiro gene que testamos esse método em nosso laboratório foi o gene da BbgII. Com os

primers contendo os respectivos *overhangs* (Tabela 3), amplificamos o fragmento a partir do DNA genômico de *B. bifidum*, utilizando a enzima *Phusion Polymerase*. O resultado da amplificação pode ser visto na Figura 9.

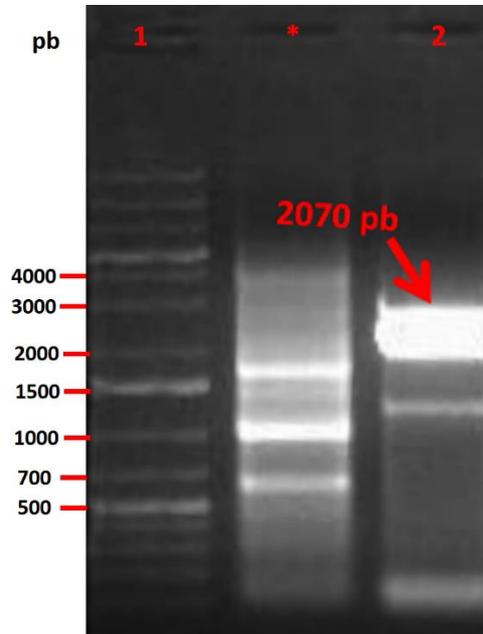


Figura 9 - Amplificação do fragmento contendo o gene da enzima BbgII. Em 1, temos o marcador molecular; Em 2, vemos o produto da reação de PCR contendo o fragmento amplificado referente ao gene de interesse. Experimentos marcados com * não estão relacionados aos discutidos.

Fonte: Elaborada pelo autor

A banda contendo o gene de interesse foi então seccionada do gel de agarose, e purificada com o *kit* comercial *Wizard® SV Gel and PCR Clean-Up System*. Tanto o fragmento, quanto o vetor de interesse, foram tratados de acordo com o item 2.2, e misturados nas proporções de 1:3, respectivamente. Em seguida bactérias da cepa DH10B foram transformadas com os insertos, e a presença do plasmídeo foi checada por *PCR* de colônia.

Após confirmarmos a eficiência do método de LIC no gene de BbgII, nosso laboratório adotou uma nova plataforma de clonagem baseada em tal método, porém, realizada em larga escala e para os múltiplos genes. Nesse contexto estão inseridos os genes de XCC_1754, XCC_2404 e XCC_2895. Como foi descrito para a BbgII, os fragmentos contendo os genes de interesse foram amplificados a partir do DNA genômico de *X. campestris* (Figura 10).

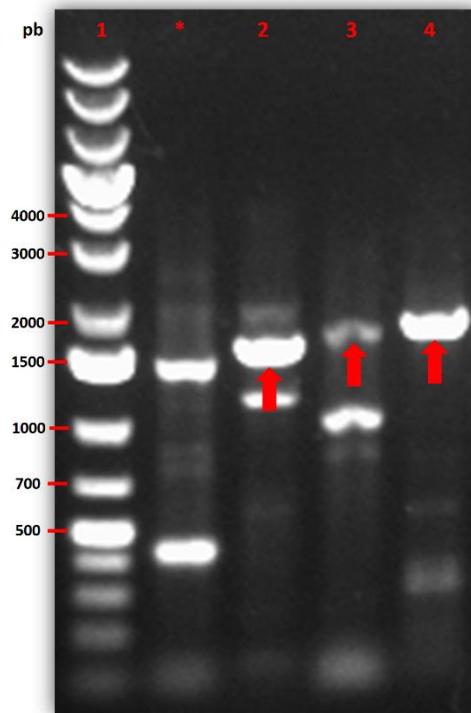


Figura 10 - Amplificação dos fragmentos contendo os genes das enzimas XCC_1754, XCC_2404 e XCC_2895. Em 1, temos o marcador molecular; em 2, o fragmento contendo o gene de XCC_2895 (1773 pb); em 3, o fragmento contendo o gene de XCC_1754 (1668 pb); em 4, o fragmento contendo o gene de XCC_2404 (1533 pb). Experimentos marcados com * não estão relacionados aos discutidos.

Fonte: Elaborada pelo autor

3.3 Testes de expressão

Após a confirmação da presença dos plasmídeos contendo os genes de interesse, células da cepa DH10B foram propagadas e tiveram os respectivos plasmídeos extraídos, utilizando o *kit* comercial *QIAprep Spin Miniprep (QIAGEN®)*. Estes foram então utilizados para a transformação das cepas de expressão *Rosetta(DE3)*. A cepa contendo o gene de BbgII foi crescida à 37 °C até a D_{600nm} atingir valores próximos a 0,6. Em seguida, adicionamos 1,0 mM de IPTG na cultura e mantivemos uma alíquota sob agitação à 30 °C, e outra à 18 °C. Após 12 h foi feita a lise dessas células, conforme descrito no item 2.5., e tais frações foram submetidas a análise através do gel SDS-Page 15% (Figura 11).

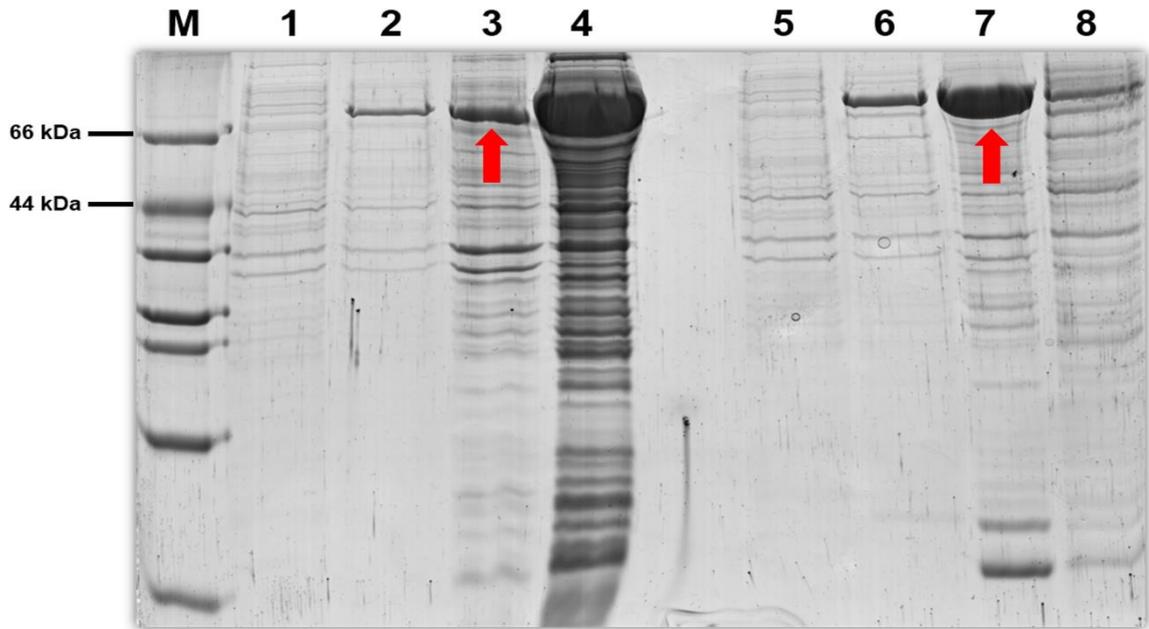


Figura 11 - Testes de expressão da enzima BbgII. Em M, temos o marcador molecular; Em 1, 2, 3 e 4 vemos respectivamente a alíquota não induzida, a alíquota após a indução, a fração solúvel e a fração insolúvel do teste a 30 °C. Em 5, 6, 7 e 8 vemos respectivamente a alíquota não induzida, a alíquota após a indução, a fração solúvel e a fração insolúvel do teste a 18 °C. A seta vermelha aponta para a proteína BbgII (77,2 kDa) encontrada na fração solúvel.

Fonte: Elaborada pelo autor

Como ficou claro na Figura 11, os testes à 18 °C exibiram uma banda muito mais significativa de proteína solúvel do que os testes realizados à 30 °C. Sendo assim, definimos aquela com sendo a temperatura padrão para experimentos futuros.

No caso das enzimas XCC_1754, XCC_2404 e XCC_2895, os testes de expressão foram realizados na plataforma de testes em larga escala do laboratório. Essa plataforma realiza o mesmo protocolo de lise discutido no item 2.5., porém de forma automatizada utilizando o robô Freedom EVO (TECAN®). Após a lise celular, apenas as frações solúveis foram submetidas a análise no sistema de eletroforese capilar LabChip® GX (Figura 12).

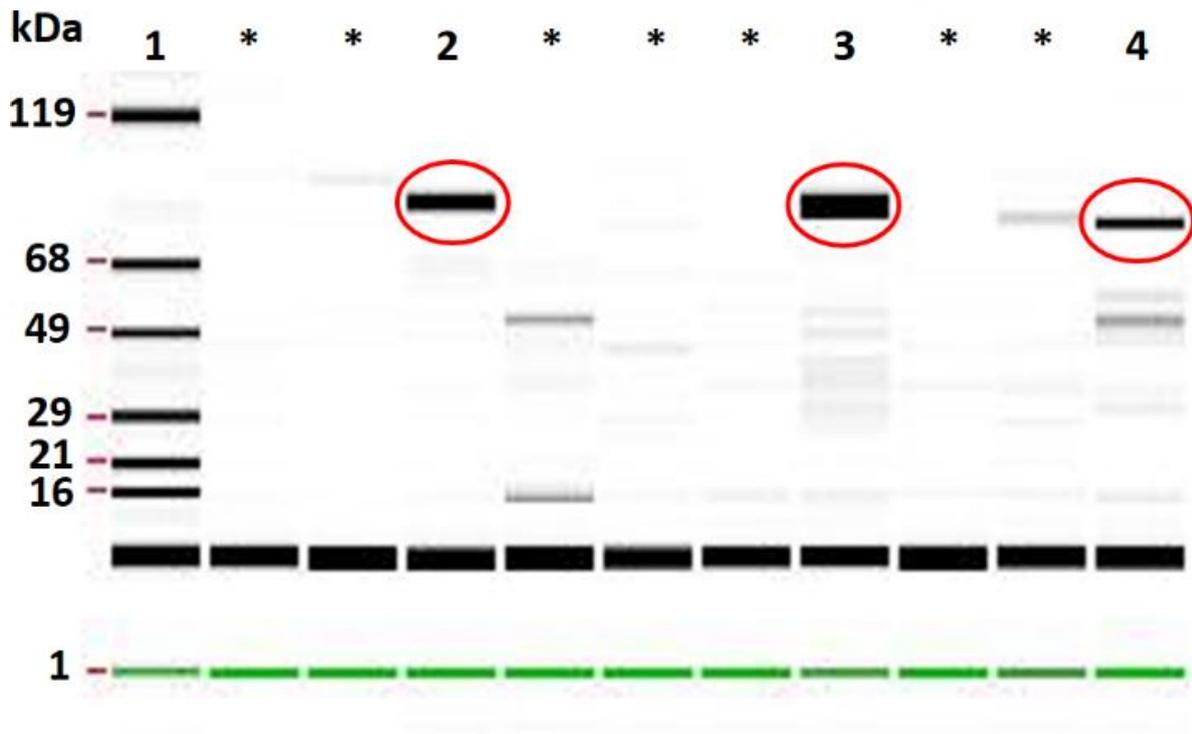


Figura 12 - Testes de expressão à 18 °C das enzimas XCC_1754, XCC_2404 e XCC_2895. Em 1 vemos o marcador molecular; Em 2 a fração solúvel da enzima XCC_2895 (65,6 kDa); Em 3 a fração solúvel da enzima XCC_1754 (60,1 kDa); Em 4 a fração solúvel da enzima XCC_2404 (55,2 kDa). As proteínas estão demarcadas pelo círculo vermelho. Experimentos marcados com * não estão relacionados aos discutidos.

Fonte: Elaborada pelo autor

Para as enzimas XCC_1754, XCC_2404 e XCC_2895, a temperatura de 18 °C, e a concentração de 1,0 mM de IPTG também produziu uma quantidade significativa de proteínas solúveis (Figura 12). Essas condições foram consideradas ótimas para a expressão e foram replicadas em larga escala.

3.4 Expressão em larga escala e purificação das enzimas

Com base nos resultados obtidos no item anterior, as proteínas BbgII, XCC_1754, XCC_2404 e XCC_2895 foram expressas em larga escala (cerca de 2,0 L de cultura) e submetidas a um protocolo de purificação baseado em duas etapas. Na primeira etapa, utilizamos a afinidade por níquel conferida as enzimas pela adição da cauda de histidina na porção N-terminal de cada molécula. A lise celular foi realizada como descrito no item 2.5, e em seguida cada amostra foi inserida em uma coluna *HisTrap FF* (GE Healthcare Life

Sciences®). Após a lavagem, as amostras foram eluídas através de um gradiente com um tampão contendo imidazol 300 mM, conforme observado na Figura 13.

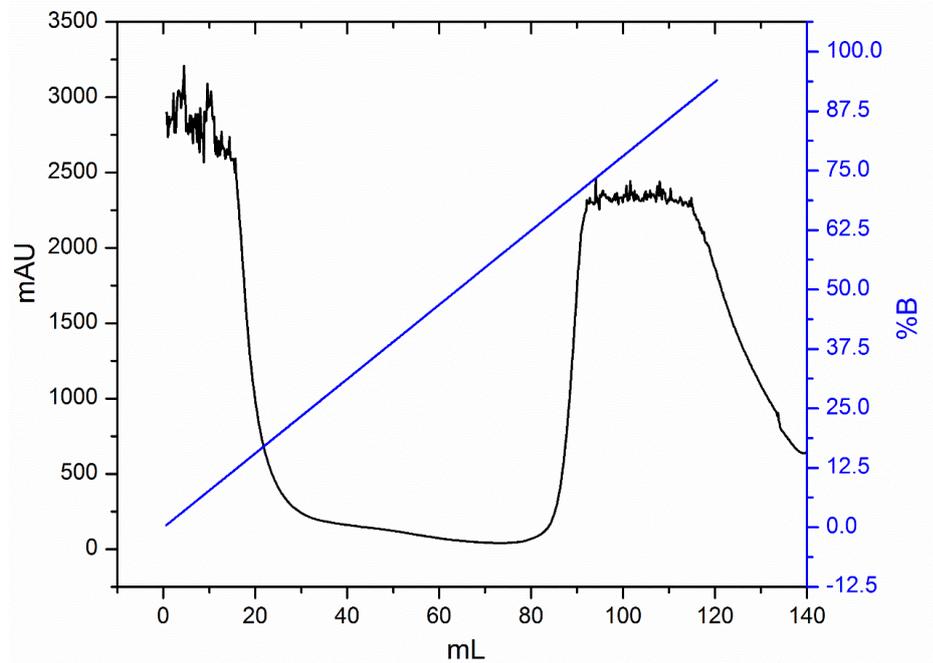


Figura 13 - Cromatograma da purificação por afinidade à níquel pela enzima BbgII. A linha preta representa a absorção da amostra a 280 nm. Em azul, a porcentagem do tampão B na amostra.

Fonte: Elaborada pelo autor.

A etapa de purificação por afinidade removeu a maior parte dos contaminantes das amostras. Porém, como o intuito era a cristalização e caracterização funcional das enzimas, foi necessário acrescentarmos um segundo passo de purificação por cromatografia de exclusão por tamanho ou gel-filtração. Abaixo podemos ver a Figura 14 que contém os cromatogramas oriundos da gel-filtração e os géis SDS-Page 15% das amostras puras.

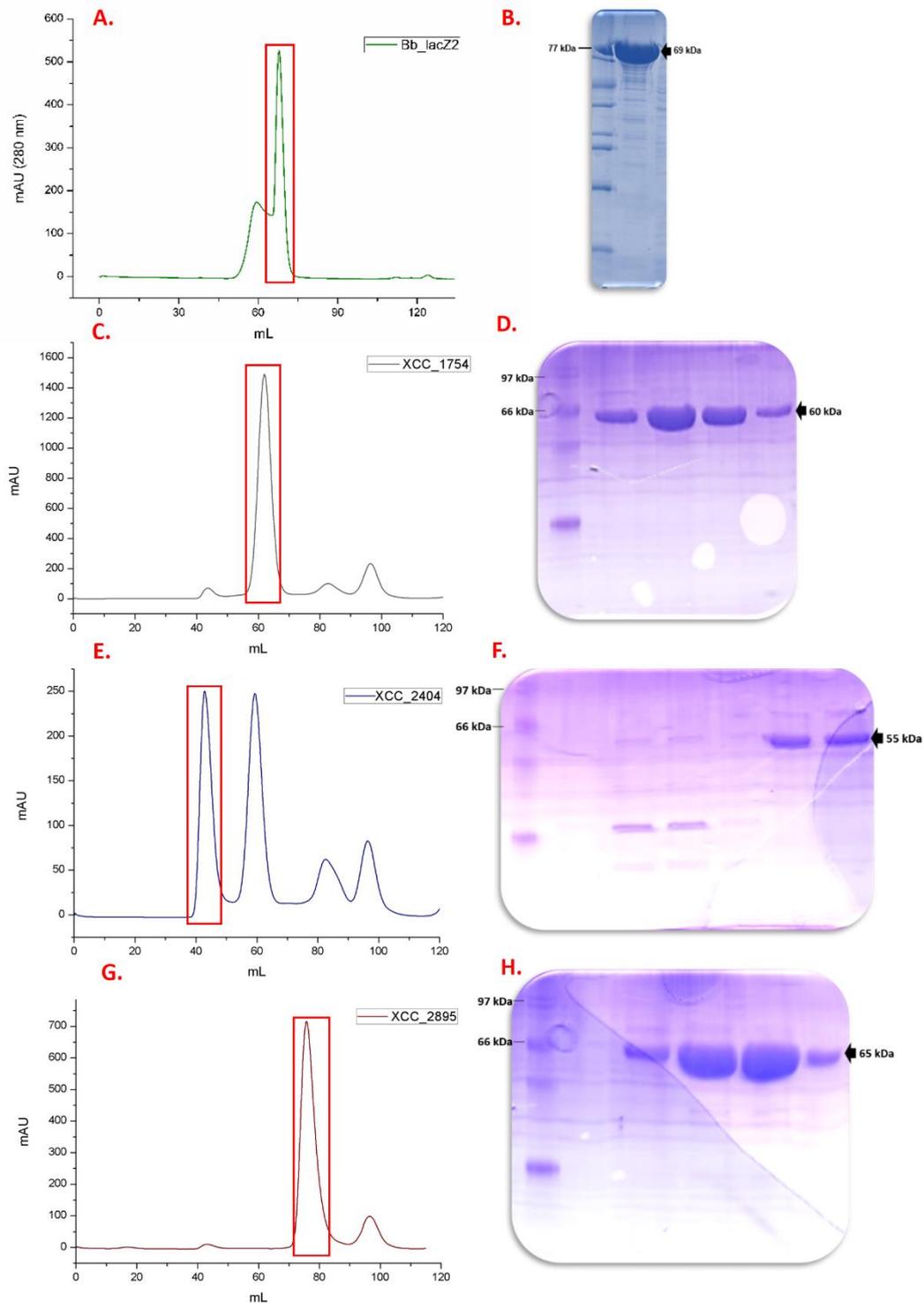


Figura 14 - Cromatogramas das purificações por exclusão molecular e géis SDS-Page das enzimas BbgII, XCC_1754, XCC_2404 e XCC_2895. Em A e B temos respectivamente o cromatograma e o gel da purificação da enzima BbgII. Em C e D o cromatograma e o gel da purificação da enzima XCC_1754. Em E e F temos o cromatograma e o gel da purificação da enzima XCC_2404. Em G e H temos o cromatograma e o gel da purificação da enzima XCC_2895. O quadro vermelho destaca a amostra no cromatograma.

Fonte: Elaborada pelo autor

As purificações das enzimas BbgII, XCC_1754, XCC_2404 e XCC_2895 foram realizadas com sucesso. Os rendimentos finais destas enzimas (por litro de cultura) foram de 40,00 mg, 12,00 mg, 2,00 mg e 12,15 mg, respectivamente.

3.5 Análise do estado oligomérico das enzimas

A partir dos resultados obtidos na purificação por exclusão molecular é possível estimarmos o valor de massa molecular em solução das amostras com base no volume de eluição de cada uma. Para isso, utilizamos a técnica de gel filtração analítica, onde o volume de eluição de cada amostra é comparado com o de padrões de proteínas com massa molecular em solução conhecidas. Para isso, utilizamos o *kit Gel Filtration HMW Calibration (GE Healthcare Life Sciences®)* de acordo com as especificações do fabricante. Com base no volume de eluição de cada proteína utilizada, construímos um gráfico do log da massa molecular pelo volume de eluição de cada amostra. Em seguida, utilizamos a regressão linear para calcular o coeficiente angular (a) e o linear (b) da relação linear entre os pontos

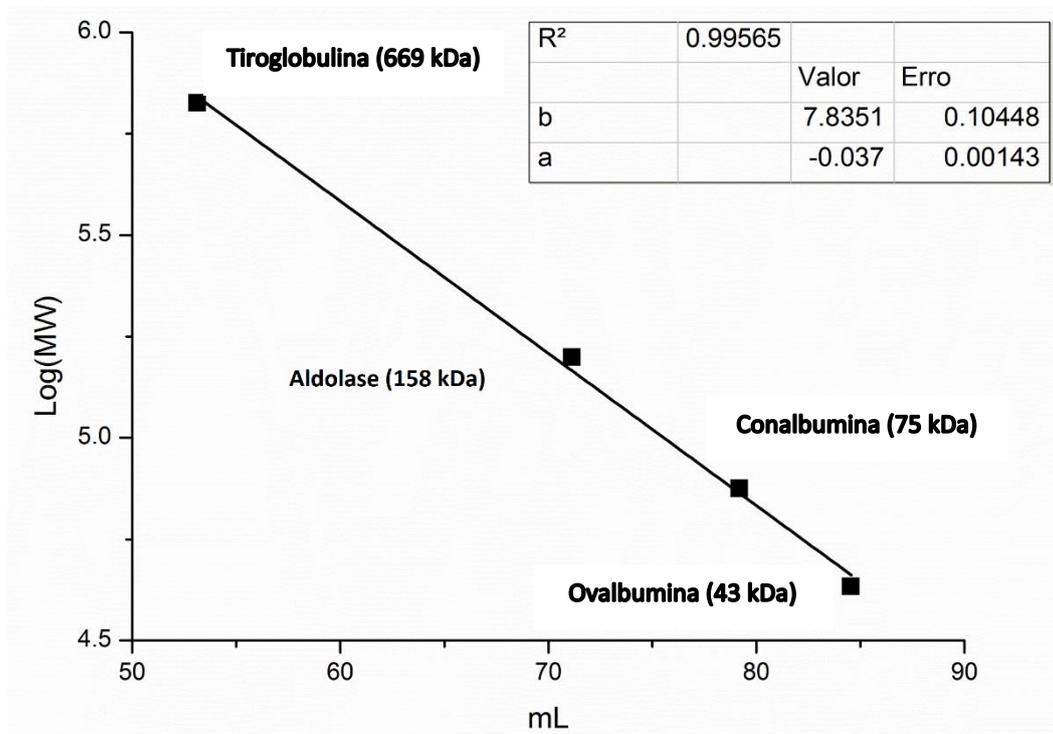


Figura 15 - Regressão linear do log da massa molecular (MW) pelo volume de eluição (mL) das amostras padrões da gel filtração analítica. O quadro mostra os valores de coeficiente angular (a) e o linear (b) calculados.

A partir dessa curva foi possível calcular os valores de massa molecular para os picos de cada enzima purificada na Figura 14. Com base nesses resultados, foi também possível estimar o estado oligomérico de cada enzima. Os resultados calculados estão resumidos na Tabela 4.

Tabela 4 - Resultados da gel filtração analítica.

	BbgII	XCC_1754	XCC_2404	XCC_2895
Pico (mL)	67,6	62,0	59,0	75,6
Massa teórica (kDa)	69,0	60,0	55,0	65,0
Massa experimental (kDa)	198,6	226,0	298,6	64,7
Estado oligomérico	Trímero	Trímero	Hexâmero	Monômero

Fonte: Elaborada pelo autor

A enzima BbgII apresentou uma massa molecular em solução de cerca de 198 kDa, cerca de o triplo da massa teórica da proteína monomérica. Esse resultado condiz com o estado trimérico já descrito em outras enzimas homólogas da família GH42.²¹⁻²² Já a enzima XCC_1754, apresentou a massa de 226 kDa, sugerindo que haja pelo menos três, ou possivelmente quatro moléculas formando uma unidade biológica. Assim como a XCC_1754, a enzima XCC_2404 também apresentou um estado multimérico. A massa calculada sugere que essa enzima forme um hexâmero em solução. Ambas essas enzimas pertencem a família das GH35. Por fim, a enzima XCC_2895 foi a única a apresentar um valor de massa que sugere o estado monomérico. Ainda assim, seu principal homólogo, a BgaC de *Streptococcus pneumoniae*, foi descrita como sendo um homodímero em solução.⁷⁹ Apesar de servirem como indícios, mais testes são necessários para afirmarmos qual o real estado oligomérico de cada uma das enzimas.

3.6 Cristalização e das enzimas

Para a cristalização da BbgII, dialisamos a enzima no tampão D (Bis-Tris 100 mM, NaCl 50 mM, MgSO₄ 1,0 mM, pH 6,5), que foi determinado a partir de experimentos fracassados de cristalização.⁸⁰ Após a diálise, a enzima foi concentrada até cerca de 10 mg/mL e os testes de cristalização foram repetidos. Os cristais foram obtidos na condição di-tartarato de sódio 0,2 M e PEG 3350 25%, e otimizados com solventes orgânicos. Os melhores cristais cresceram na condição di-tartarato de sódio 0,2 M, PEG 3350 20% e 1-propanol 4% (Figura 16). Para a derivatização desses cristais, utilizamos a técnica de *soaking* com 2,0 mM de metilmercúrio. A mesma técnica foi utilizada para a incorporação da galactose no cristal.

No caso das enzimas XCC_1754, XCC_2404 e XCC_2895, as concentrações necessárias para a cristalização foram respectivamente 28,0; 1,5 e 10,0 mg/mL. Os cristais de XCC_1754 cresceram na condição de citrato de tri-lítio 0,2 M e PEG 3350 20% (Figura 16). Os cristais de XCC_2404 cresceram na condição acetato de cálcio 0,2 M e PEG 3350 20% (Figura 16). Já os cristais de XCC_2895 cresceram na condição cloreto de lítio 0,2 M e PEG 3350 25% (Figura 16).

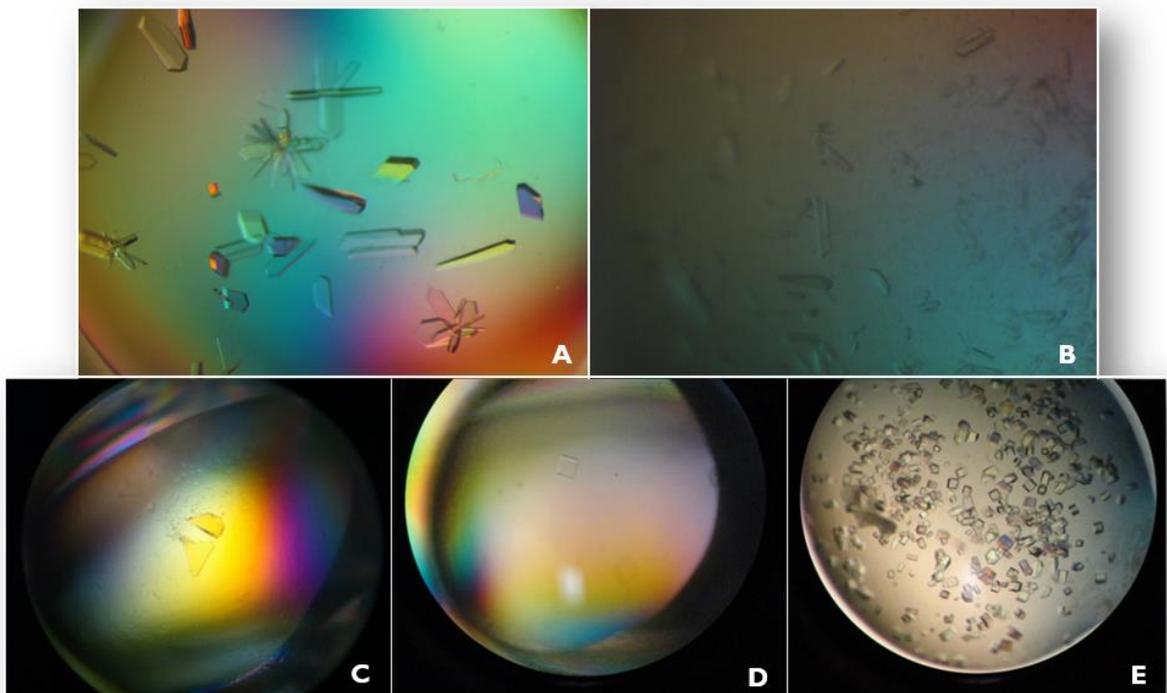


Figura 16 - Cristais das β -galactosidasas. Em A, temos os cristais de BbgII na condição di-tartarato de sódio 0,2 M, PEG 3350 20% e 1-propanol 4%. Em B vemos os cristais de BbgII três dias após o *soaking* com mercúrio. Em C temos os cristais da enzima XCC_2895 na condição cloreto de lítio 0,2 M e PEG 3350 25%. Em D, os cristais de XCC_2404 na condição acetato de cálcio 0,2 M e PEG 3350 20%. Em E, vemos os cristais de XCC_1754 na condição citrato de tri-lítio 0,2 M e PEG 3350 20%.

Fonte: Elaborada pelo autor.

3.7 Coleta e processamento de dados

Os cristais obtidos no item anterior foram crio-protetidos de acordo com o descrito no item 2.7. Para a enzima BbgII, utilizamos o ânodo rotatório do Instituto de Física da USP para a coleta de dados de um conjunto nativo e de um complexo proteína-galactose, incubada pelo método de *soaking*. Como tal enzima não possuía homólogos com alta identidade sequencial, optamos pela técnica de *SAD* para o faseamento do conjunto. Essa técnica utiliza a medida da dispersão da intensidade espalhada (estimada pela diferença entre f' e f'') para um determinado elemento em um dado comprimento de onda (λ). Para isso, os cristais de BbgII foram imersos na solução de cristalização contendo 10 mM de metil mercúrio (Figura 16). O mercúrio é comum em protocolos de faseamento experimental *in-house* devido ao seu alto espalhamento anômalo produzido na faixa de energia gerada pelo ânodo de cobre (cerca de 12 e para 8014 eV, Figura 17), e também pelo fato deste metal se ligar covalentemente a cisteínas livres da proteína.

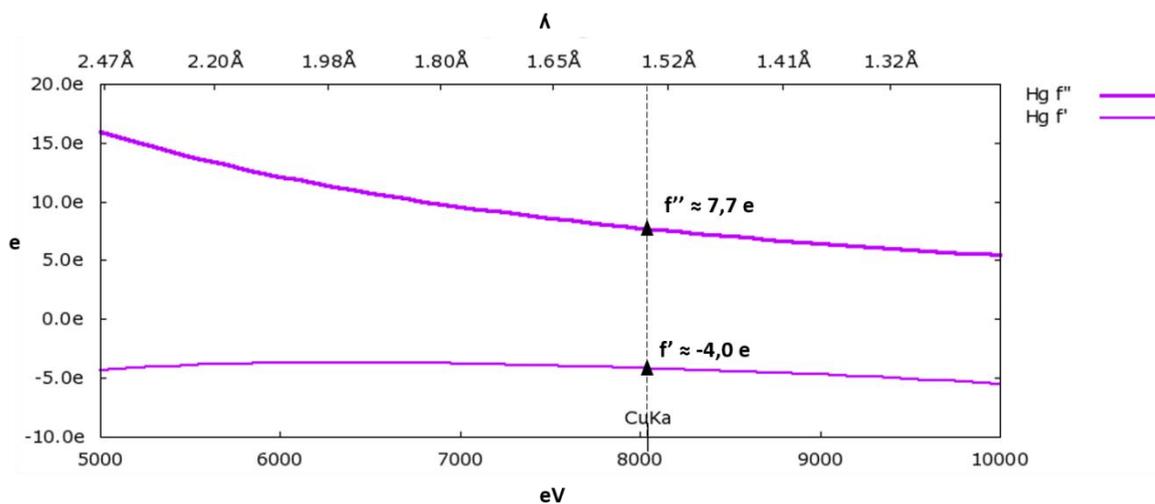


Figura 17 - Coeficientes de dispersão (f' e f'') em diferentes energias para o elemento mercúrio. Triângulos e linha tracejada demarcam os valores para a energia produzida no ânodo de cobre.

Fonte: <http://skuld.bmsc.washington.edu/scatter/> (acesso, 19/05/2016)

As estratégias de coleta foram focadas na maximização da completeza e resolução dos conjuntos de dados com base no grupo espacial de cada cristal.⁸¹ A integração e redução dos

dados foram realizadas pelo programa XDS,⁶¹ e as estatísticas de escalonamento foram geradas pelo programa Aimless,⁶³ e podem ser observadas na Tabela 5.

Tabela 5 - Estatísticas das coletas de dados das β -galactosidases. Em parênteses os valores para a alta resolução.

	BbglI	BbglI (Derivado)	Complexo BbglI-galactose	XCC_1754	XCC_2404	XCC_2895
Fonte de Raio-X	Ânodo	Ânodo	Ânodo	X4C (NSLS)	X4C (NSLS)	X4C (NSLS)
Detector	ImagePlate	ImagePlate	ImagePlate	MAR 165 CCD	MAR 165 CCD	MAR 165 CCD
Comprimento de Onda (Å)	1,54	1,54	1,54	1,10	1,10	1,1
Resolução (Å)	35,8 (1,7)	46,2 (2,6)	42,8 (1,9)	46,0 (1,7)	49,1 (2,5)	42,8 (1,7)
Grupo Espacial	P 1 2, 1	P 1 2, 1	P 1 2, 1	P 4 2, 2	P 1 2, 1	P 1 2, 1
Célula Unitária (a b c α β γ)	95,7 99,2 110,7 90,0 106,5 90,0	100,1 100,7 107,7 90,0 106,4 90,0	96,7 101,6 114,6 90,0 105,4 90,0	158,1 158,1 94,8 90,0 90,0 90,0	128,3 163 235,5 90,0 92,8 90,0	80,8 84,7 99,8 90,0 96,6 90,0
Reflexões únicas	200160 (18207)	64119 (3985)	153059 (13379)	131650 (12988)	332223 (15162)	283540 (27563)
Multiplicidade	3,1 (2,7)	6,1 (4,9)	3,7 (3,4)	16,3 (16,0)	4,2 (4,0)	4,3 (2,6)
Completeness (%)	98,5 (89,9)	99,1 (87,8)	97,0 (85,0)	99,9 (99,6)	99,1 (91,8)	97,6 (96,1)
I/ σ (I)	11,9 (1,9)	14,7 (1,8)	13,4 (2,0)	14,5 (2,5)	7,4 (1,5)	13,70 (1,4)
R _{int} (%)	8,7 (65,0)	9,6 (105,4)	9,7 (66,3)	4,4 (30,6)	18,3 (99,0)	4,2 (67,0)
CC _{1/2}	--	--	--	0,99 (0,73)	0,98 (0,53)	0,99 (0,61)

Fonte: Elaborada pelo autor.

Como podemos ver na Figura 16 e na Figura 18, o *soaking* com o mercúrio teve um efeito deletério na qualidade dos cristais. Mesmo assim, foi possível obter dados com resolução suficiente (cerca de 2,6 Å) e boa qualidade ($R_{\text{meas}} \sim 9.6\%$) para a resolução da estrutura pelo método de *SAD*.

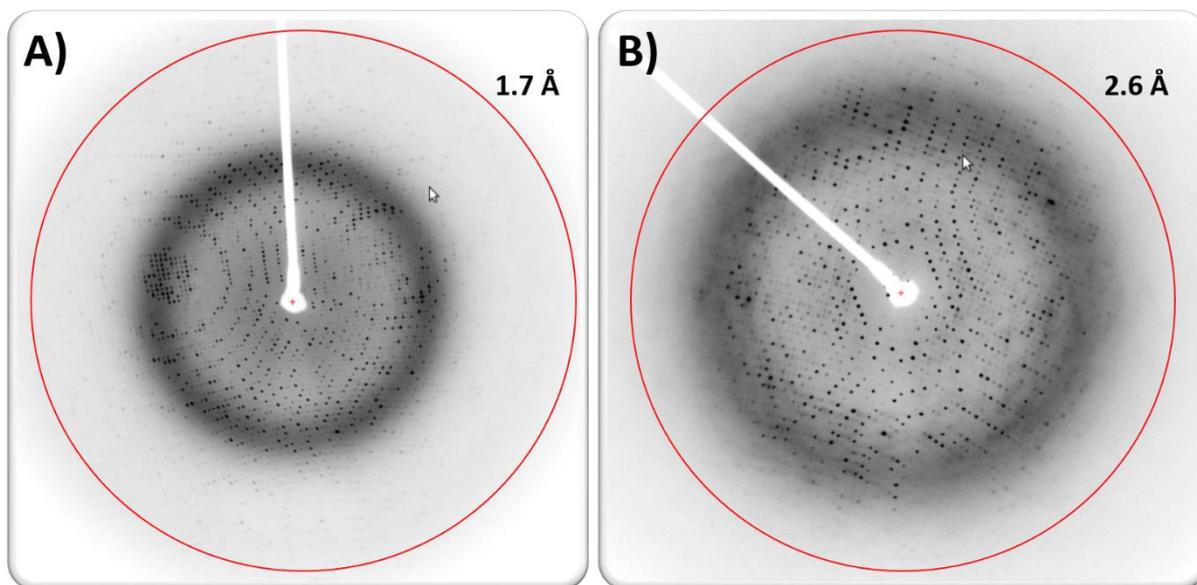


Figura 18 - Marcos de difração dos cristais de BbgII. Em A), o marco inicial dos cristais nativos à 1.7 Å de resolução. Em B), o marco inicial dos cristais derivados com mercúrio à 2.5 Å de resolução.

Fonte: Elaborada pelo autor.

Os cristais de XCC_1754, XCC_2404 e XCC_2895 foram coletados durante o curso Rapidata (2012), no *National Synchrotron Light Source* (NSLS), Nova Iorque, Estados Unidos, na linha de luz X4C. O limite de resolução desses conjuntos foi determinado pelo CC1/2 dos conjuntos.^{62; 63} Os cristais apresentaram respectivamente 1,7; 2,5 e 1,7 Å de resolução, e ótimas estatísticas após o processamento (Tabela 5). Podemos ver abaixo, o exemplo do marco de difração da enzima XCC_1754, que mostrou formação de gelo durante o envio, e foi submetida a um processo de *annealing* durante 10 segundos, o que permitiu a remoção do gelo e aumento da qualidade da difração do cristal (Figura 19).

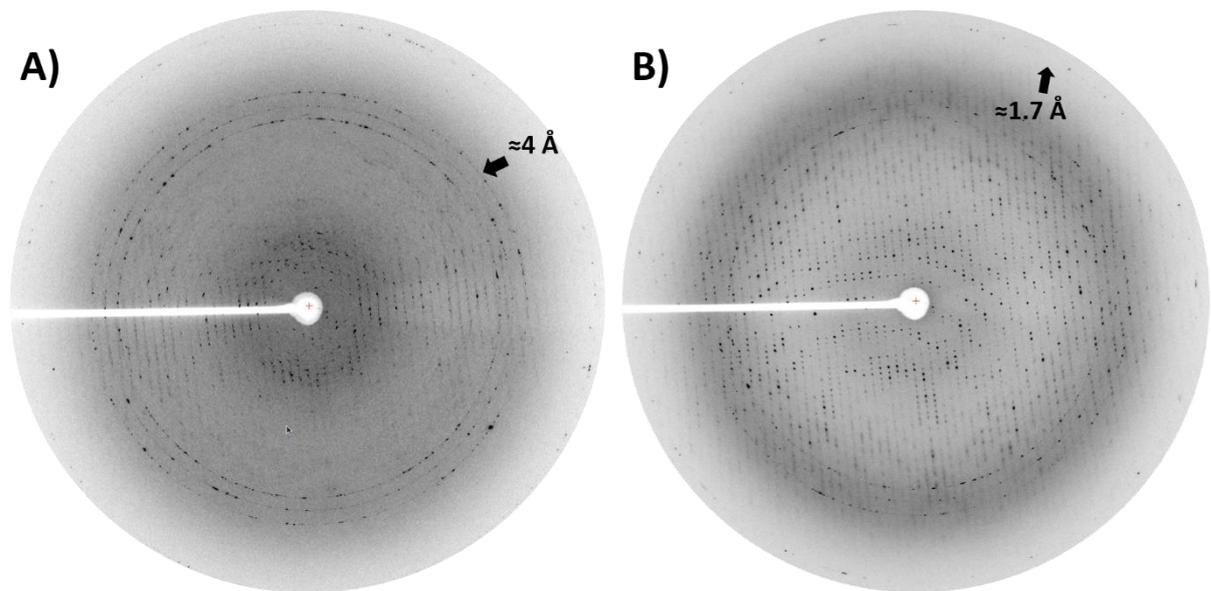


Figura 19 - Marcos de difração da enzima XCC_1754. Em A), o marco de difração inicial exibe anéis de gelo e baixa qualidade. Em B), o marco após o *annealing* mostra uma significativa melhora na qualidade dos dados.

Fonte: Elaborada pelo autor.

Em geral, os dados obtidos para todas as enzimas apresentaram boa qualidade e foram utilizados para a determinação das estruturas cristalográficas das enzimas, discutidos nos tópicos posteriores.

CAPÍTULO 4

Resultados & Discussão II

Enzima BbgII

4 Estrutura e função da enzima BbgII

4.1 Resolução e qualidade da estrutura cristalográfica da enzima BbgII.

Os dados de difração, provenientes da enzima BbgII foram integrados no grupo espacial $P1\ 2_1\ 1$, na célula unitária com eixos a, b, c medindo 95,7 Å, 99,2 Å e 110,7 Å, e um ângulo β de 106,5 °. Com base nesses parâmetros, utilizamos o coeficiente de Matthews,⁸² para estimar o conteúdo da célula unitária. No caso, ele indicou que a célula unitária é formada por três moléculas de BbgII e 47,7% de solvente. A estrutura foi resolvida pelo método de SAD, com uma combinação dos programas SHELXC/D,⁶⁶ e AutoSharp.⁶⁷

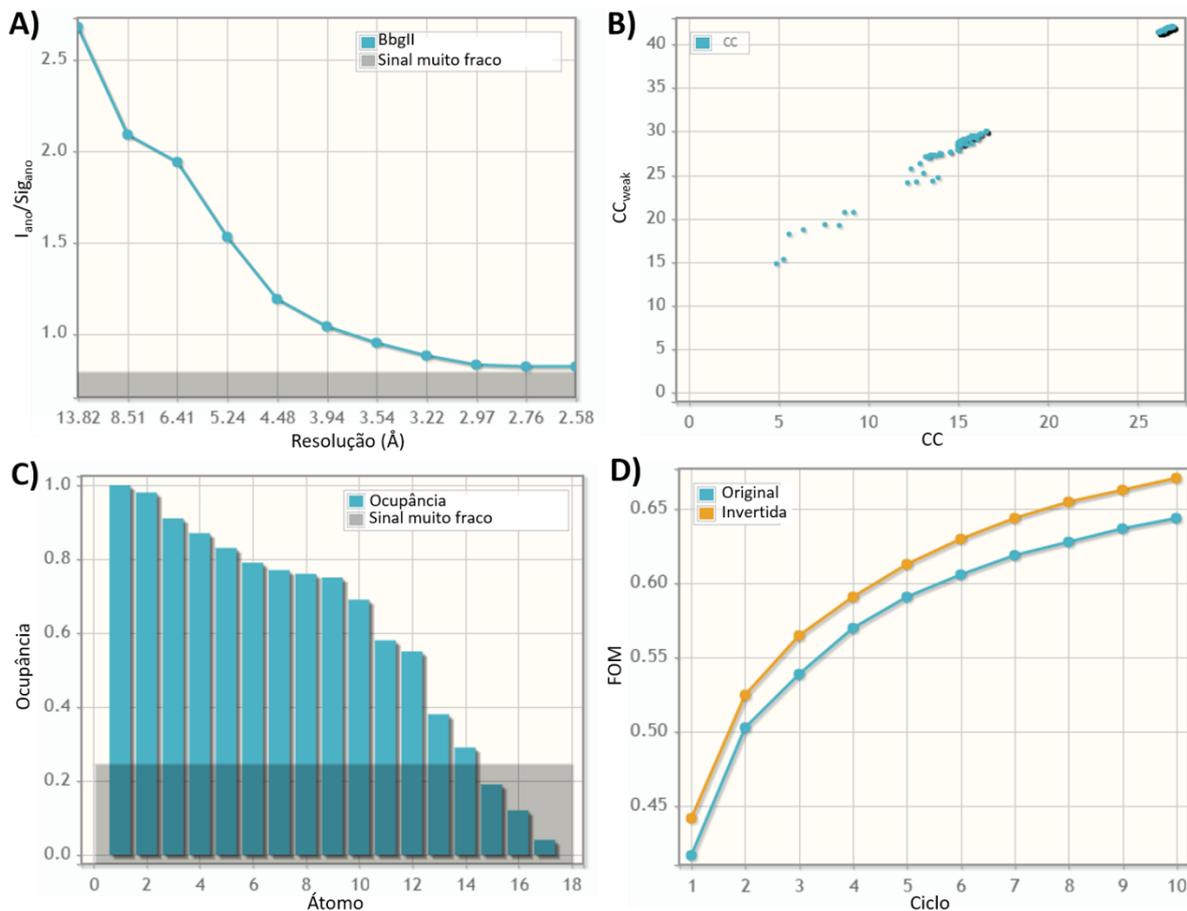


Figura 20 - Parâmetro de resolução da estrutura da enzima BbgII. Em A), vemos os valores de I_{ano}/Sig_{ano} por faixa de resolução, que sugerem um alto sinal anômalo abaixo de 3,0 Å. Em B), os coeficientes de correlação mostram dois grupos distintos de soluções. Em C), os valores de ocupância para os átomos pesados encontrados mostram uma queda brusca a partir do décimo segundo átomo. Em D), a figura de mérito (FOM) do mapa durante a modificação de densidade mostra uma distinção acentuada entre os soluções original e invertida.

Fonte: Elaborada pelo autor

O programa SHELXC foi utilizado para preparar os dados (cálculo do f'' , e estimação do ângulo α) e extrair as diferenças de intensidade dos pares de Friedel (I_{ano}/Sig_{ano}), enquanto que o programa SHELXD foi utilizado para o mapeamento dos átomos pesados (Figura 20). Foram encontrados 12 átomos de mercúrio no conjunto derivado, com um coeficiente de correlação (CC) de 0,44 à 3,10 Å de resolução. Em seguida realizamos o refinamento dos sítios e cálculo das fases (combinado com a modificação de densidade) com o programa AutoSharp. O achatamento de solvente descartou a solução de “mão-direita” como correta, enquanto que as fases de “mão-invertida” apresentaram características de densidade eletrônica de cadeias. Essa densidade foi utilizada para a construção do modelo com o programa AutoBuild,⁶⁴ (Figura 21). Os átomos de mercúrio foram encontrados ligados nas cisteínas 140, 356, 389 e 474, em todas as três cadeias.

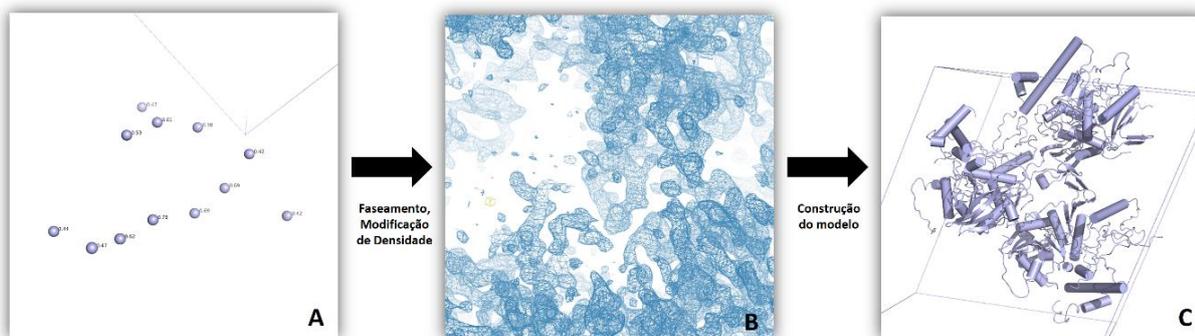


Figura 21 - Passos de resolução da estrutura da enzima BbgII. Em A vemos os 12 sítios de átomos pesados encontrados no conjunto derivado. Em B, vemos a densidade eletrônica da “mão-invertida”. Em C, o modelo inicial construído automaticamente.

Fonte: Elaborada pelo autor

Após o cálculo inicial das fases utilizamos o programa Coot,⁶⁹ para a modelagem manual e vários ciclos do programa AutoBuild para obter um modelo completo. Esse modelo foi então refinado e validado com os dados da enzima nativa, a 1,75 Å, até um modelo final com valores de R_{work}/R_{free} de 15,3/19,2 (%). O modelo final apresenta três cadeias da enzima BbgII, que se iniciam na Lys3 e terminam na Arg689, onde apenas pequenos fragmentos não apresentaram densidade eletrônica (resíduos 624-626 e 639-642). A validação do modelo mostrou que 97,15% dos resíduos se encontram em regiões favoráveis no diagrama de Ramachandran, e somente 0,34% resíduos estão em regiões não permitidas. Nas três cadeias, os resíduos Ser324 e Pro473 estão em regiões não permitidas. O resíduo Ser324 possivelmente está envolvido na trimerização da BbgII, onde vemos a formação de um contato polar (2,6 Å) entre o OG do resíduo e o OE1 do Glu530 da molécula de outra cadeia.

O modelo de BbgII em complexo com galactose apresentou qualidade similar ao da enzima nativa, onde o RMSD entre esses modelos é de apenas 0.26 Å (para 689 C α). As estatísticas de refinamento e validação para ambos os modelos podem ser vistas na Tabela 6.

Tabela 6 - Estatísticas de refinamento da enzima BbgII e de seu complexo com galactose.

	BbgII	Complexo BbgII-galactose
R_{work}	0,15 (0,26)	0,15 (0,24)
R_{free}	0,19 (0,31)	0,20 (0,29)
Número de átomos		
Ligantes	101	107
Águas	2696	2023
Resíduos	2051	2052
RMSD (ligações) (Å)	0,009	0,019
RMSD (ângulos) (°)	1,17	1,83
Ramachandran favoráveis (%)	97	96
Ramachandran não permitidos (%)	0,34	0,39
Clashscore	2,08	2,82
B-fatores médios (Å²)		
Macromoléculas	16,80	24,10
Ligantes	29,30	42,00
Solvente	27,80	32,50

Fonte: Elaborada pelo autor

4.2 Estrutura tridimensional da enzima BbgII.

A estrutura cristalográfica da enzima BbgII é composta por três monômeros, onde cada monômero é formado por três domínios distintos (Figura 22). O primeiro domínio, formado por 12 α -hélices e 8 fitas- β , vai do início da cadeia polipeptídica até o resíduo 398 e possui o enovelamento do tipo barril-TIM, bastante similar ao de outras estruturas de GH42, como as β -galactosidases de *Thermus thermophilus* A4 (A4- β -gal) e *Bacillus circulans* sp. *alkalophilus* (Bca- β -gal).^{21; 22} Essas enzimas possuem respectivamente 34% e 36% de identidade sequencial com BbgII, e seus RMSD são respectivamente 1,8 Å e 1,5 Å. Porém, quando comparamos somente o primeiro domínio desses homólogos com o primeiro domínio da BbgII, esses números baixam para 1,2 Å e 1,0 Å, sugerindo que esse domínio é mais conservado do que a média global da proteína. Isso é provavelmente devido ao fato de que esse domínio contém os principais resíduos catalíticos das β -galactosidases das GH42. No

caso da BbgII encontramos os resíduos Glu160 e Glu320 que são, respectivamente, o doador de prótons e o nucleófilo do mecanismo de inversão.

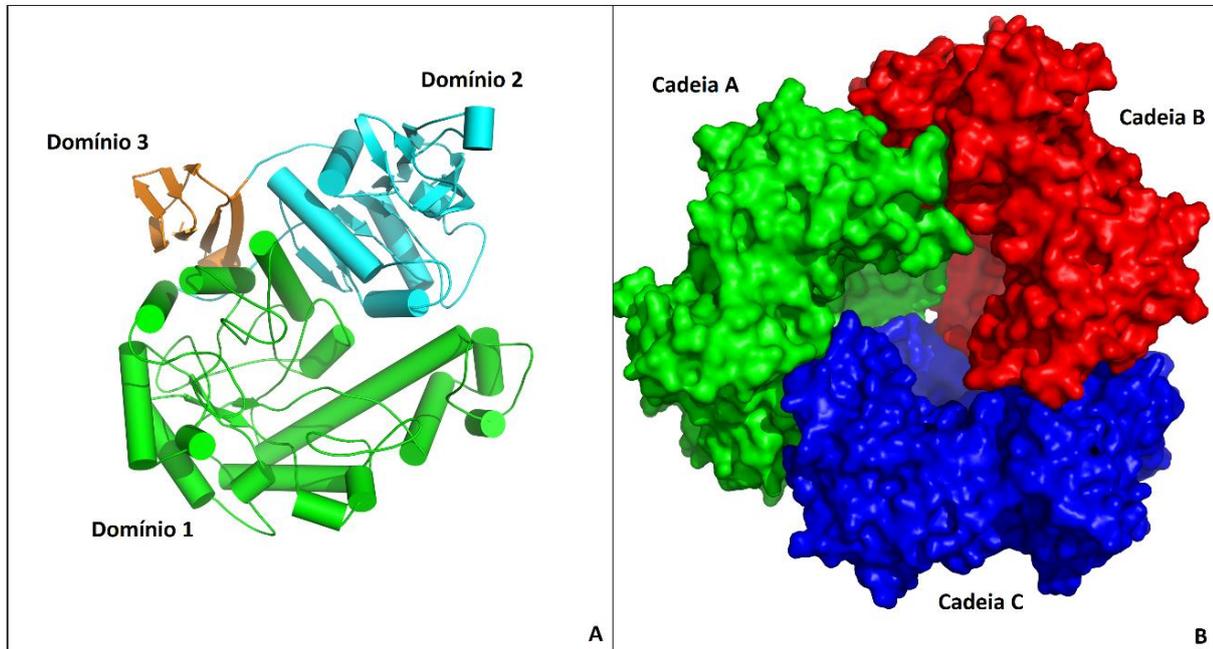


Figura 22 - Estrutura cristalográfica da Enzima BbgII. Em A vemos a forma monomérica da enzima onde temos o domínio 1 (em verde) formado por um barril-TIM, o domínio 2 (em azul claro) e o domínio 3 (em laranja). Em B, temos o trímero em formato de barril, que é tanto a unidade assimétrica quanto a unidade biológica da enzima.

Fonte: Elaborada pelo autor

Apesar de bastante parecidos, existem diferenças significativas entre o primeiro domínio da BbgII e os de A4- β -gal e Bca- β -gal. Por exemplo, em A4- β -gal e Bca- β -gal observamos quatro metioninas formando um bolsão de enxofre no núcleo hidrofóbico do barril-TIM.²² No caso da BbgII, três dessas metioninas foram substituídas, onde somente a Met319 permaneceu conservada. As outras diferenças estruturais acontecem principalmente nos *loops*, como por exemplo o *loop* entre os resíduos Pro214 e Met223. Porém a principal diferença entre os primeiros domínios das enzimas está no fato da BbgII não possuir um sítio de zinco. Nas estruturas de A4- β -gal e Bca- β -gal vemos um sítio de zinco, formado por quatro cisteínas, que foi descrito como importante na estabilização da enzima.²² Já na BbgII, essas cisteínas foram substituídas pelos resíduos Arg126, Asp129, Ser171 e Ala174, e o sítio se encontra ausente.

O segundo domínio da BbgII se inicia no resíduo 399 e vai até o 627. Esse domínio é formado por seis α -hélices, seis fitas- β paralelas e cinco fitas- β antiparalelas. É o mais diferente dos domínios, sendo que seu RMSD com a enzima A4- β -gal é de 3,2 Å. Esse domínio faz a interface com o primeiro domínio, estando provavelmente envolvido no processo de oligomerização. O terceiro domínio vai do resíduo 628 até o C-terminal e é formado por sete fitas- β antiparalelas, com um enovelamento do tipo sanduíche- β . Esse domínio não parece estar envolvido no processo de oligomerização ou na atividade da enzima.

A enzima BbgII foi resolvida como um trímero em sua unidade assimétrica (Figura 22). No caso dessa enzima, podemos observar um eixo de simetria de ordem três no centro da unidade assimétrica, onde os três monômeros estão empacotados, formando uma estrutura em forma de barril. Análises no programa PISA,⁸³ demonstraram que existem 35 pontes de hidrogênio, 7 pontes salinas e pelo menos 70 resíduos realizando algum tipo de interação entre a interface de dois monômeros de BbgII, sugerindo fortemente que o empacotamento cristalino é também o empacotamento biológico da enzima. Além disso, a estrutura trimérica concorda com a massa molecular experimental encontrada (item 2.6).

O fato das β -galactosidases da família GH42 formarem trímeros em solução havia sido reportado.^{21; 84} A estrutura da β -galactosidase de *Thermus thermophilus* A4 (A4- β -gal) foi resolvida como um monômero, porém análises do empacotamento e dados experimentais revelaram que a estrutura quaternária da enzima é composta por três monômeros em forma de barril, assim como a BbgII.²¹ Outro exemplo é a enzima de *Bacillus circulans sp. alkalophilus* (Bca- β -gal) que foi resolvida como um hexâmero, porém análises demonstraram que a superfície de contato entre dois trímeros (também em forma de barril) não é estável o suficiente para manter a conformação hexamérica em solução.²²

4.3 Sítio ativo da enzima BbgII

O sítio ativo da enzima BbgII está localizado no interior do barril, próximo à região de interface entre duas moléculas (Figura 23A). A estrutura da enzima BbgII em complexo com a galactose revelou detalhes sobre esse sítio. Primeiramente, observamos que a galactose aparece na isoforma α , apesar do *soaking* ter sido feito com β -D-galactose. Isso provavelmente acontece devido às redes de ligações de hidrogênio que estabilizam a molécula de galactose na forma α -D-galactose, especialmente a interação entre o O1 da galactose e o

OH da Tyr289 (2,9 Å) (Figura 23C). Podemos observar também as interações da α -D-galactose com os resíduos Arg122, Asn160, Glu161, Asp287, Glu320, Trp328 e Glu368. Os resíduos Glu161 e Glu320 são, respectivamente, o doador de prótons e o nucleófilo da reação de inversão, e se encontram à 1,9 Å e 2,8 Å de distância do O1 da galactose. Enquanto o b-fator dos átomos da galactose varia em torno de 16 Å², o b-fator do O1 chega a 29 Å², sugerindo que esse átomo possui alta mobilidade e pode interagir com ambos os glutâmicos durante a reação.

Uma segunda densidade eletrônica próxima ao sítio revelou uma segunda molécula de galactose próxima ao Trp328 (Figura 23B). Essa molécula possui um b-fator médio bastante elevado (cerca de 60 Å²), sugerindo sua alta mobilidade na estrutura cristalina. Mesmo assim, podemos assumir que essas interações são responsáveis por coordenar a molécula de galactose até o sítio ativo da enzima, sendo então uma posição transiente durante a reação.

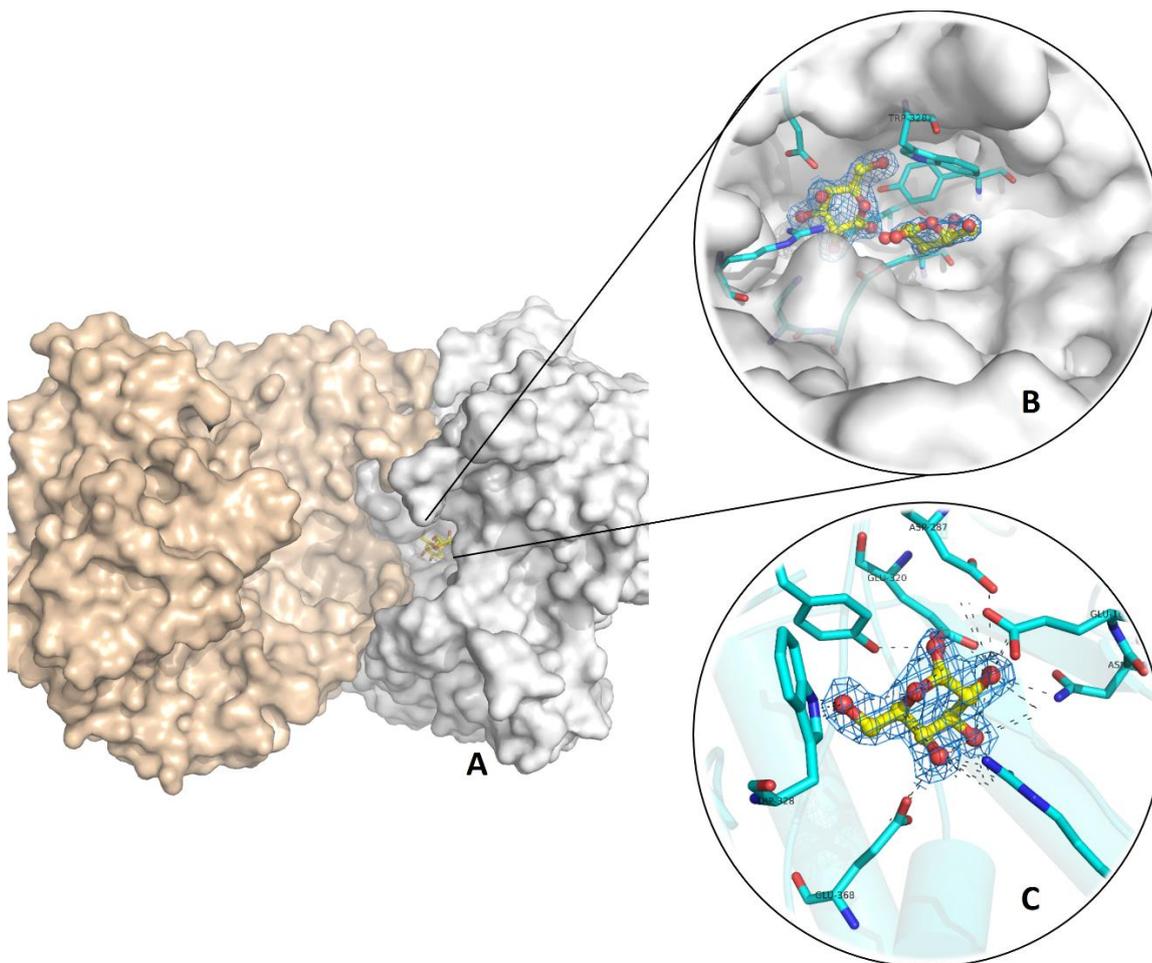


Figura 23 - Localização e visão geral do sítio ativo da enzima BbgII. Em A observamos a interação entre duas moléculas de BbgII e a localização do sítio entre essas moléculas. Em B, a visão da superfície do sítio e a posição das moléculas de galactose com suas respectivas densidades eletrônicas. Em C, a coordenação da galactose no sítio ativo da BbgII.

Outra coordenação particular da enzima BbgII é a interação entre o Trp201 da cadeia B com o O3 da galactose localizada no sítio da cadeia A. Essa interação é uma ponte de hidrogênio intermediada por uma água, a cerca de 3 Å de distância de ambos, e sugere que a trimerização não somente é importante para a estabilização, mas também para a atividade da enzima. Comparando os sítios de A4- β -gal e Bca- β -gal, observamos que os todos os resíduos se mantêm conservados nas três estruturas. As posições dos resíduos em relação a galactose são também bastante similares, sendo que a única mudança ocorre no Trp378 de BbgII e o triptofano relativo de A4- β -gal, onde o rotâmero aparece girado 180°. Mesmo assim, observamos que o grau de conservação do sítio é extremamente elevado dentro da família das GH42.

4.4 Parâmetros bioquímicos da BbgII.

Começamos os estudos funcionais da enzima BbgII determinando qual a temperatura e o pH ótimo da enzima em solução. Para isso, foi utilizado o substrato sintético *o*NPG. Para determinarmos o pH ótimo, utilizamos o tampão acetato-borato-fosfato (ABF) 40 mM, entre os pH 2,0 e 10.

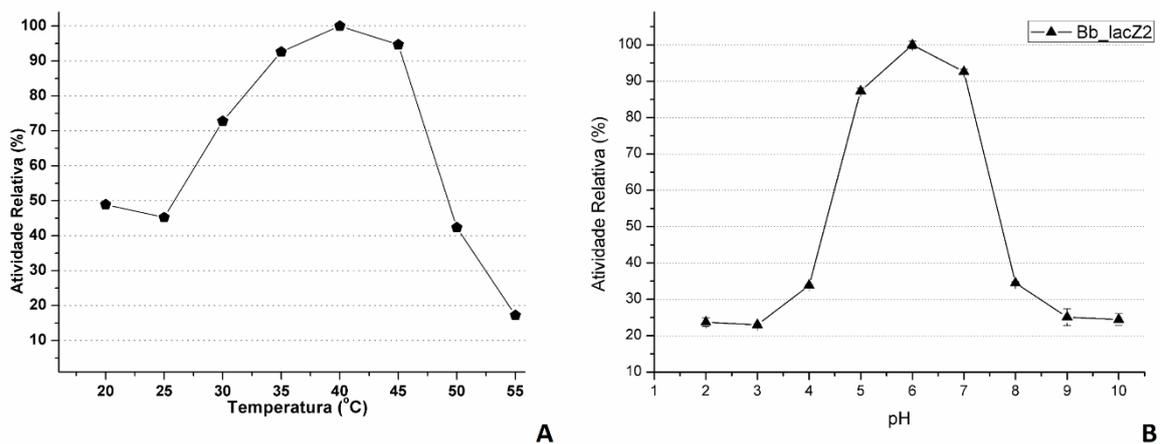


Figura 24 - Testes de estabilidade da BbgII. Em A, vemos a atividade relativa da BbgII em diferentes temperaturas. Em B, a atividade relativa da enzima em diferentes pH's

Fonte: Elaborada pelo autor.

O pH ótimo da enzima foi determinado como sendo 6,0, sendo que a enzima se mantém com quase 90% da atividade entre os pH's 5,0 e 7,0. A atividade da enzima cai drasticamente fora desse intervalo. No caso da temperatura ótima, a enzima mostrou um pico

de atividade a 40 °C, trabalhando com cerca de 90% da atividade entre as temperaturas de 35 e 45 °C. Considerando-se que as bactérias do gênero *Bifidumbacterium* ocorrem no trato digestivo humano, era de se esperar que suas enzimas trabalhassem em condições semelhantes às encontradas no nosso organismo.⁵²

4.5 Cinética e inibição da enzima BbgII.

A cinética da enzima BbgII foi medida com o substrato sintético *o*NPG e com lactose, seu substrato natural. Os testes foram realizados a 25 °C para evitar o decaimento da atividade enzimática devido à desnaturação térmica. Para o *o*NPG obtivemos um K_m de $5,95 \pm 1,8$ mM, com um V_{max} de $9,65 \pm 0,8$ $\mu\text{M}\cdot\text{s}^{-1}$. Com base nesses valores, calculamos o K_{cat} de 1860 s^{-1} , e o K_{cat}/K_m de 312 $\text{mM}^{-1}\cdot\text{s}^{-1}$. Quando a lactose foi utilizada como substrato, os valores de K_m foram de $23,84 \pm 5,7$ mM, e o V_{max} foi cerca de $45,0 \pm 3,0$ $\mu\text{M}\cdot\text{s}^{-1}$. Para a lactose, o K_{cat} observado foi de $0,7$ s^{-1} , e o K_{cat}/K_m foi cerca de $0,003$ $\text{mM}^{-1}\cdot\text{s}^{-1}$ (Figura 25).

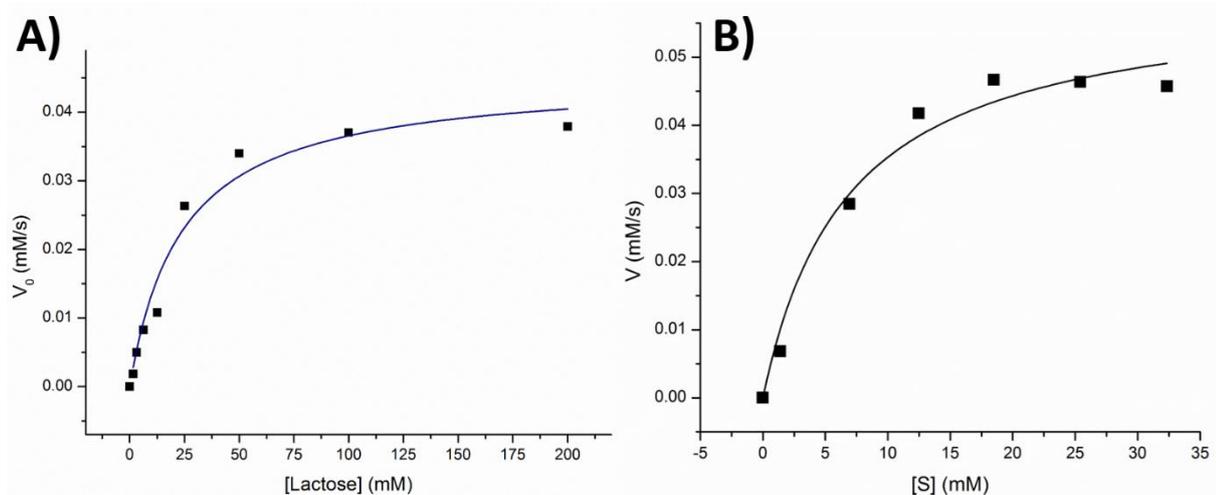


Figura 25 - Curvas da cinética da enzima BbgII a 25 °C. Em A), a curva utilizando lactose como substrato. Em B), a curva com o substrato *o*NPG.

Fonte: Elaborada pelo autor.

Quando comparamos a eficiência catalítica (K_{cat}/K_m), vemos que a enzima possui uma eficiência 10^5 vezes maior pelo substrato *o*NPG do que pela lactose. Várias enzimas da família GH42 mostraram resultados semelhantes a BbgII,^{21,84} sugerindo uma preferência pelo *o*NPG em relação ao substrato natural. Quando medimos as constantes a 40 °C, observamos que o K_m subiu para 6,84 mM, mostrando que a eficiência enzimática diminuiu

(provavelmente devido à aceleração da reação inversa). Ainda assim, o K_{cat} dessa situação aumentou para 6880 s^{-1} , mostrando que a processividade da enzima é maior nessa temperatura.

A estrutura cristalográfica da BbgII em complexo com a α -galactose sugere que essa molécula seja um ligante forte da enzima. Assim sendo, nos testamos o efeito da galactose como possível inibidor da enzima. Para isso utilizamos o substrato *o*NPG para o cálculo das constantes de Michaelis-Menten contendo diferentes concentrações de galactose (0; 12,5; 25,5 e 32,5 mM, respectivamente). Os dados foram então utilizados para observarmos o efeito do aumento da concentração de galactose através da regressão linear pela fórmula de Lineweaver-Burk (Figura 26).

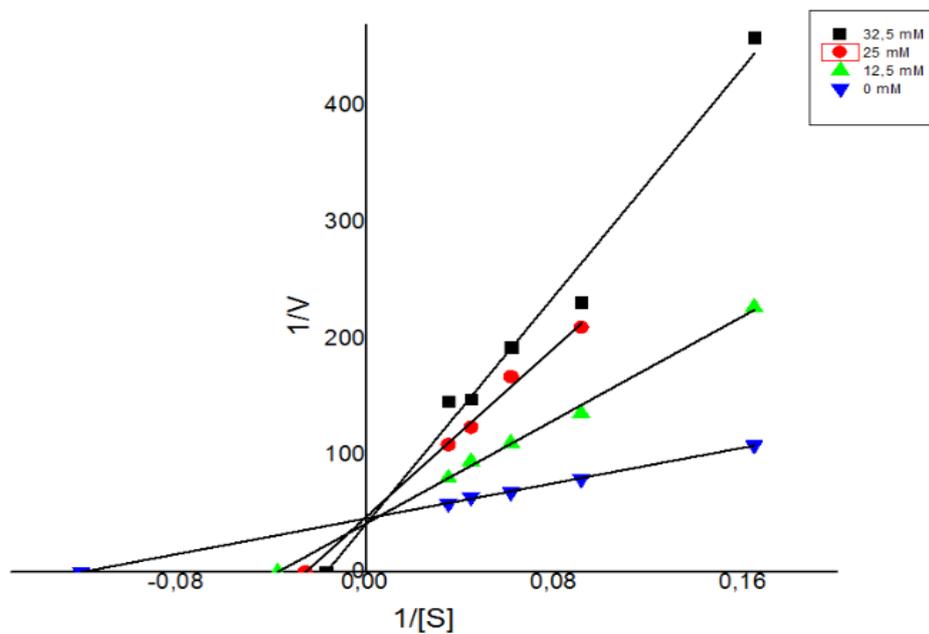


Figura 26 - Plot das curvas de Lineweaver-Burk da enzima BbgII em presença de diferentes concentrações de galactose.

Fonte: Elaborada pelo autor.

Como esperado, a presença da galactose reduziu a taxa de conversão de substrato para a enzima. Porém, podemos observar pela intercessão das retas no eixo y da curva que a presença de galactose não afeta a velocidade máxima da reação e, portanto, somente afeta a afinidade da enzima pelo substrato. Portanto, podemos concluir que a galactose é um inibidor competitivo da reação. Com base nos valores de K_m , calculados por essa curva, foi possível estimarmos a constante de inibição da enzima (K_i). Para isso, definimos o coeficiente α como

sendo a razão entre o K_m da enzima pelo K_m da enzima com o inibidor ($[I]$). Em seguida, o K_i foi calculado de acordo com a Equação 1.⁸⁵

$$K_i = \frac{[I]}{\alpha - 1} \quad (1)$$

Para a BbgII, o valor calculado para K_i da galactose foi de $6,0 \pm 1,5$ mM, sugerindo que ela reduz a eficiência enzimática da enzima pela metade em concentrações bem baixas.

4.6 Atividade e Estabilidade dos mutantes de BbgII.

Com o objetivo de melhor compreendermos o funcionamento da enzima BbgII, realizamos mutações sítio-dirigidas nos resíduos Asn160, Glu161, Glu320 e Glu368, que foram substituídos por alaninas. Os mutantes foram nomeados N160A, E161A, E320A e E368A. Os resíduos Tyr289 e His371 foram substituídos por fenilalaninas e nomeados chamados Y289F e H371F. Ambos os resíduos se localizam na entrada do sítio ativo. Utilizamos o experimento de termo-fluorimetria para medir o quanto essas mutações afetaram a estabilidade térmica da enzima em solução. O resultado pode ser visto na Figura 27.

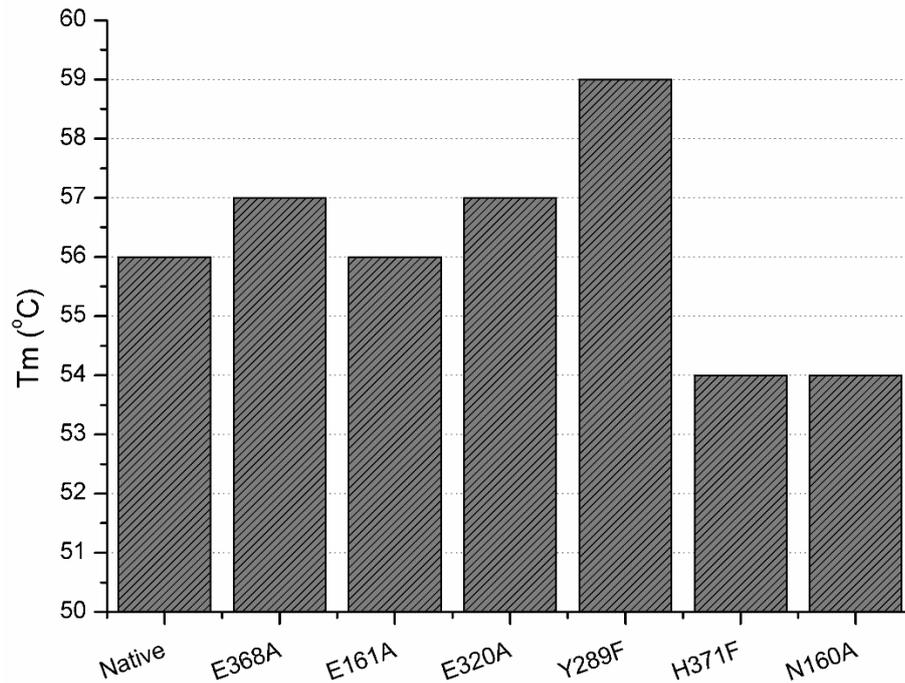


Figura 27 - Temperatura de desenovelamento da enzima BbgII e seus mutantes.

Fonte: Elaborada pelo autor.

As mutações E368A, E161A e E320A pareceram não afetar significativamente a estabilidade térmica da enzima. Já os mutantes H371F e N160A demonstraram uma estabilidade térmica reduzida, o que pode ser um indício de que esses resíduos afetam a estabilidade do sítio. O mutante Y289F apresentou um ganho significativo de estabilidade (+3 °C) quando comparado com a enzima nativa. Os motivos para esse fenômeno ainda estão sendo analisados. Após verificarmos a estabilidade, utilizamos esses mutantes para testes cinéticos com o substrato *o*NPG, para verificar o efeito de cada mutação na capacidade hidrolítica da enzima. Os resultados podem ser vistos na Tabela 7.

Tabela 7 - Parâmetros cinéticos dos mutantes de BbgII a 25 °C. Onde temos * a atividade foi muito baixa para ser quantificada.

	V_{max} ($\mu\text{M}\cdot\text{s}^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	V_{max}/K_m (s^{-1})	k_{cat}/K_m ($\text{mM}^{-1}\cdot\text{s}^{-1}$)
Nativa	$9,65 \pm 0,8$	$5,95 \pm 1,8$	1862	1,6	312,4
N160A	$1,8 \pm 0,1$	$4,27 \pm 1,6$	20,7	0,04	4,8
E161A	0	--	--	--	--
Y289F	$16,2 \pm 1,83$	$14,4 \pm 3,8$	188,4	1,12	13,0
H371F	$15,6 \pm 2,33$	$40,2 \pm 9,4$	180,8	0,39	4,4
E320A	0	--	--	--	--
E368A	*	--	--	--	--

Fonte: Elaborada pelo autor

Como era de se esperar, os mutantes E161A e E320A não apresentaram nenhuma atividade hidrolítica. O mutante E368A apresentou atividade bastante reduzida. O glutâmico 368 aparece interagindo com o O4 da molécula de galactose em nossa estrutura (Figura 23), provavelmente não estando diretamente envolvido na reação de inversão. Seu efeito deve estar envolvido com a estabilização e/ou alocação do substrato no sítio ativo, sendo que a sua falta diminui drasticamente a eficiência enzimática.

O mutante N160A apresentou um aumento na afinidade pelo substrato (K_m de 4,2 mM), porém uma queda de cerca de 10^2 vezes na constante catalítica da enzima. Esse resíduo se localiza no interior no fundo do sítio ativo (Figura 23C), sendo provável que esse resíduo esteja envolvido na estabilização do substrato no sítio, ou auxilie a reação de inversão. Já os mutantes H371F e Y289F demonstraram uma diminuição na afinidade pelo substrato, o que sugere que esses resíduos estejam envolvidos na alocação do substrato para o sítio. Esses mutantes apresentaram uma K_{cat} cerca de 10 vezes menor que a da enzima nativa.

4.7 Discussão sobre a enzima BbgII

A microbiota humana é composta por uma grande variedade de organismos pertencentes a diversos filos, incluindo os filos *Firmicutes* (79.4%), *Bacteroides* (16.9%), *Actinobacteria* (1%), *Proteobacteria* (0.1%) or *Verrucomicrobia* (0.1%).⁸⁶ Os *Bacteroides* são conhecidos por seus efeitos probióticos, em especial aqueles pertencentes ao gênero das bifidobactérias.⁸⁶ A capacidade desse grupo em utilizar β -galactosídeos de cadeia curta como fonte de energia (como os GOS) é um fator chave na regulação e controle da população do gênero no organismo.⁸⁷ De fato, vemos que cerca de 8% do genoma das bifidobactérias está relacionado ao metabolismo de carboidratos,⁸⁸⁻⁸⁹ porém pouco se entende a nível molecular sobre tal metabolismo.

BbgII foi anteriormente descrita como um dímero em solução,⁹⁰ o que é inconsistente com a literatura já publicada para a família 42,^{21-22,84,91} e com as evidências encontradas nesse trabalho. Nossos dados sugerem que a enzima BbgII tem o clássico enovelamento trimérico observado em outros membros da família, além de uma estrutura formada por três domínios, sendo que os resíduos catalíticos estão contidos no domínio I do tipo barril TIM.

Uma diferença significativa entre BbgII e seus homólogos é a ausência de um sítio de zinco no domínio I, visto em A4- β -Gal, Bca- β -gal and GanB (Figura 28).^{21-22,84,91} Tal sítio foi

descrito como importante para a estabilidade da enzima, mas não necessariamente afeta a atividade das mesmas.²¹⁻²² O homólogo mais próximo da BbgII, a BIGal42A, também possui um sítio de zinco na região do eixo de simetria do trímero, que também está ausente em BbgII (Figura 28).⁹²

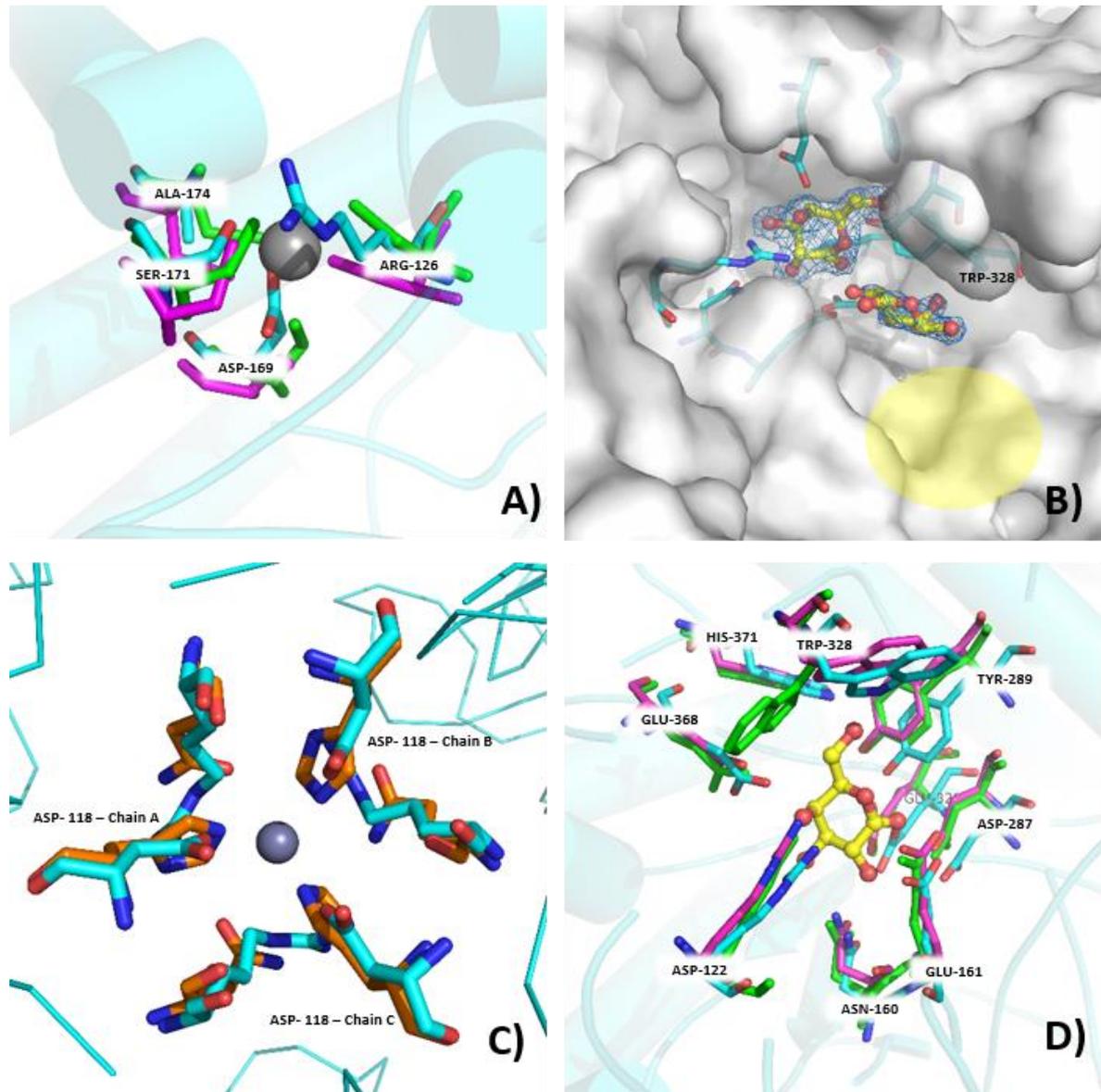


Figura 28 - Comparação entre BbgII e seus homólogos. Em A), a região do sítio de zinco de A4-β-Gal (verde) e Bca-β-gal (rosa) mostra substituição de todos os resíduos em BbgII (azul). Em B), a entrada do sítio as moléculas de galactose, com destaque (amarelo) para região rica em resíduos com enxofre. Em C) o sítio de zinco de BIGal42A (laranja) comparado com a região em BbgII mostra substituição dos resíduos. Em D), o alinhamento entre os sítios ativos de BbgII (azul), A4-β-Gal (verde) e Bca-β-gal (rosa).

Fonte: Elaborada pelo autor.

O complexo BbgII-galactose revelou duas moléculas de galactose ligadas ao sítio ativo: uma ligada fortemente aos principais resíduos catalíticos, e outra na entrada do sítio. O alto B-

fator ($\sim 40 \text{ \AA}^2$) sugere que essa segunda molécula está em uma posição transiente do mecanismo de reação. Essa molécula faz interface com o Trp328 e a His371, e também com uma região rica em aminoácidos contendo enxofre, provavelmente envolvidos na ancoragem e estabilização do substrato.

4.8 Conclusões e perspectivas em relação a enzima BbgII.

A enzima BbgII foi purificada, cristalizada, caracterizada bioquimicamente e teve sua estrutura cristalográfica resolvida com sucesso. Além de importantes informações sobre a estrutura nativa, também obtivemos a estrutura da enzima em complexo com seu ligante natural, a galactose. Essa estrutura permitiu-nos observar quais resíduos estão diretamente envolvidos na reação, e por meio de mutações sítio dirigidas e estudos cinéticos, conseguimos propor funções para a maioria dos resíduos do sítio na hidrolise.

Como já era esperado os resíduos Glu161 e Glu320 são fundamentais para a reação. As análises sequenciais e estruturais sugerem que estes são os pares nucleófilo e doador de prótons, respectivamente. Já o resíduo Glu368 também se mostrou essencial para a reação. O papel de tal aminoácido na reação pode estar relacionado ao equilíbrio de dos pKa's dos resíduos catalíticos, sendo que o mutante não seria capaz de atingir um estado de cargas em que a reação se torne favorável. O mutante H371F mostrou uma diminuição da afinidade pelo substrato ($K_m \sim 40 \text{ mM}$) em relação a enzima nativa ($K_m \sim 6,0 \text{ mM}$), o que corrobora as observações estruturais. O mesmo efeito foi observado para o mutante Y289F ($K_m \sim 14 \text{ mM}$). Curiosamente, o mutante do Trp328 não gerou nenhuma proteína solúvel, indicando que ele pode estar também envolvido na estabilidade global da estrutura da molécula. Já o mutante N160A manteve a afinidade pelo substrato similar a da enzima nativa ($K_m \sim 4,3 \text{ mM}$), porém vimos uma queda de quase 10^2 na eficiência catalítica. Isso pode indicar que tal resíduo tem um papel importante no processo catalítico e/ou na estabilização do substrato.

CAPÍTULO 5

Resultados & Discussão III

Resolução e Refinamento das estruturas de *Xanthomonas*

5 Resolução e qualidade do refinamento das estrutura das enzimas XCC_1754, XCC_2404 e XCC_2895.

Os dados cristalográficos provenientes da cristalização da enzima XCC_1754 foram integrados no grupo espacial P4 2₁ 2, revelando a célula unitária com valores a, b, c de 158,1 Å, 158,1 Å e 94,8 Å, e seu coeficiente de Matthews ⁸² sugeriu a presença de duas moléculas na unidade assimétrica. Já os dados provenientes dos cristais da enzima XCC_2404 foram integrados no grupo espacial P1 2₁ 1, com parâmetros da célula unitária a, b, c de 80,8 Å, 84,7 Å, 99,8 Å, e um ângulo β de 96,6°. O coeficiente de Matthews também sugeriu um conteúdo celular de duas moléculas para essa enzima. Os dados de difração de raios X da enzima XCC_2895 também foram integrados no grupo espacial P1 2₁ 1, com parâmetros da célula 128,3 Å, 163 Å e 235,5 Å para a, b e c, e um ângulo β de 92,8°. Para essa enzima, porém, o coeficiente de Matthews indicou que podem haver entre 16 e 25 moléculas na unidade assimétrica.

As três estruturas foram resolvidas por substituição molecular com o programa Phaser.⁶⁵ As enzimas XCC_1754 e XCC_2404 foram resolvidas com base na β -galactosidase de *Caulobacter crescentus* (PDB id: 3U7V, 58% e 44% de identidade sequencial com XCC_1754 e XCC_2404, respectivamente). A enzima XCC_2895 foi resolvida com base na β -galactosidase de *Streptococcus pneumoniae* (PDB id: 4E8C, com 41% de identidade sequencial).⁷⁹ Os modelos foram construídos pelos programas Coot,⁶⁹ e Autobuild,⁹³ e, concomitantemente, refinados e validados pelos programas phenix.refine (Afonine *et al.*, 2012) e MolProbity.⁷⁰ Os parâmetros finais dos refinamentos das enzimas XCC_1754, XCC_2404 e XCC_2895 podem ser vistos na Tabela 8.

Tabela 8 - Estatísticas do refinamento das enzimas XCC_1754 e XCC_2895. Os valores entre parênteses são para as últimas faixas de resolução.

	XCC_1754	XCC_2404	XCC_2895
R_{work}	0,16 (0,22)	0,20 (0,30)	0,16 (0,28)
R_{free}	0,19 (0,26)	0,24 (0,35)	0,20 (0,32)
Número de átomos			
Ligantes	41	0	60
Águas	1626	2580	1336
Resíduos	1023	8152	1170
RMSD (ligações) (Å)	0,011	0,006	0,012
RMSD (ângulos) (°)	1,33	1,05	1,23
Ramachandran favoráveis (%)	96,6	98,0	96,8
Ramachandran não permitidos (%)	0	0,03	0
Clashscore	1,77	4,95	2,04
B-fatores médios (Å²)			
Macromoléculas	17,80	32,80	35,50
Ligantes	15,50	--	39,10
Solvente	29,20	29,30	40,10

Fonte: Elaborada pelo autor.

A estrutura cristalográfica da enzima XCC_1754 foi resolvida em 1,7 Å de resolução. Todos os resíduos foram modelados em regiões permitidas no diagrama de Ramachandran. Mesmo após vários ciclos, quatro resíduos continuaram com ângulos não favoráveis. Na cadeia A, a Met47 (CB-CG-SD: 5,1 σ), Lys56 (CA-CB-CG: 4,2 σ) e Trp379 (CA-CB-CG: 4,6 σ). Na cadeia B, somente o Trp379 (CA-CB-CG: 4.014 σ) apresentou problemas. Uma porcentagem dos rotâmeros (0,7%) também apresentaram problemas, porém dentro do limite tolerável para essa resolução. A enzima XCC_2895 foi resolvida em 1,7 Å de resolução, com um R_{work}/R_{free} final de 0,16/0,20. Essa enzima também apresentou uma unidade assimétrica composta por duas moléculas, porém o programa PISA indica que as interfaces não são estáveis em solução. A enzima XCC_2404 apresentou uma unidade assimétrica incomum. Isso porque o enorme volume (cerca de 4.900.000 Å³), e a baixa simetria (P1 2₁ 1) da cela unitária sugerem que a unidade assimétrica seja composta por mínimo 16 moléculas. Utilizando-se da simetria não cristalográfica, foi possível refinar essa estrutura até valores de R_{work}/R_{free} 20/24 (Tabela 8).

CAPÍTULO 6

Resultados & Discussão IV

Enzima XCC_1754

6 Estrutura e função da enzima XCC_1754

6.1 Estrutura tridimensional da enzima XCC_1754

A estrutura final da enzima XCC_1754 apresentou dois domínios, um deles sendo um barril-TIM (domínio catalítico). O segundo domínio (domínio C-terminal) é composto por apenas fitas β . A unidade assimétrica da enzima é composta por um dímero, como demonstrado na Figura 29.

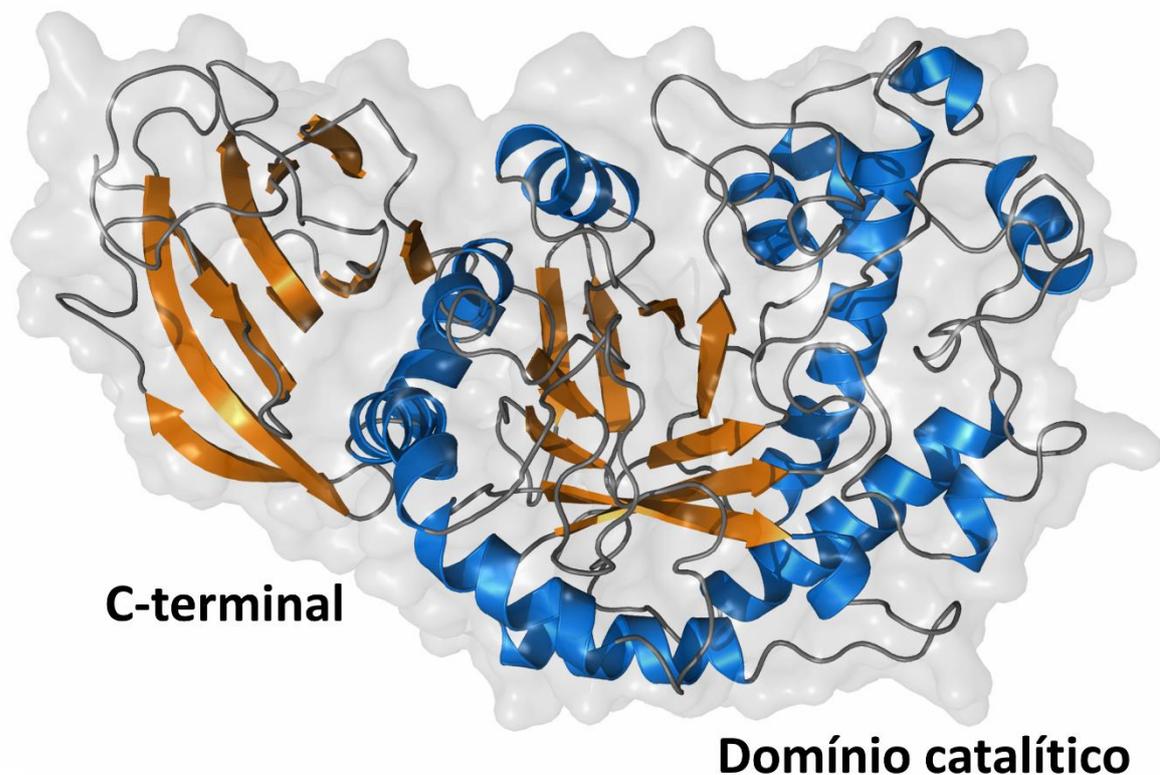


Figura 29 - Modelo cristalográfico da enzima XCC_1754. A enzima possui um domínio catalítico no formato de barril TIM, e um domínio C-terminal formado por fitas.

Fonte: Elaborada pelo autor.

Apesar da unidade assimétrica formada por um dímero, as análises realizadas pelo programa PISA,⁸³ nenhum dos contatos cristalinos observados é estável suficiente para manter a estrutura quaternária da enzima em solução. Na Figura 14 podemos ver que a gel filtração analítica sugere a presença de um trímero em solução. Porém, este fato pode ser resultado de aglomeração inespecífica, comum em proteínas em altas concentrações.

Um fato curioso sobre a estrutura cristalográfica da enzima XCC_1754 foram as diferenças entre as cadeias A e B da estrutura. O *loop* 11, que vai da Ala290 até a Leu314, aparece em conformações diferentes nas cadeias A e B. Na cadeia B, ele forma uma pequena hélice e se aloja na parte superior do barril-TIM (posição “fechada”). Já na cadeia A, ele não possui uma estrutura secundária definida (posição “aberta”), como pode ser visto na Figura 30.

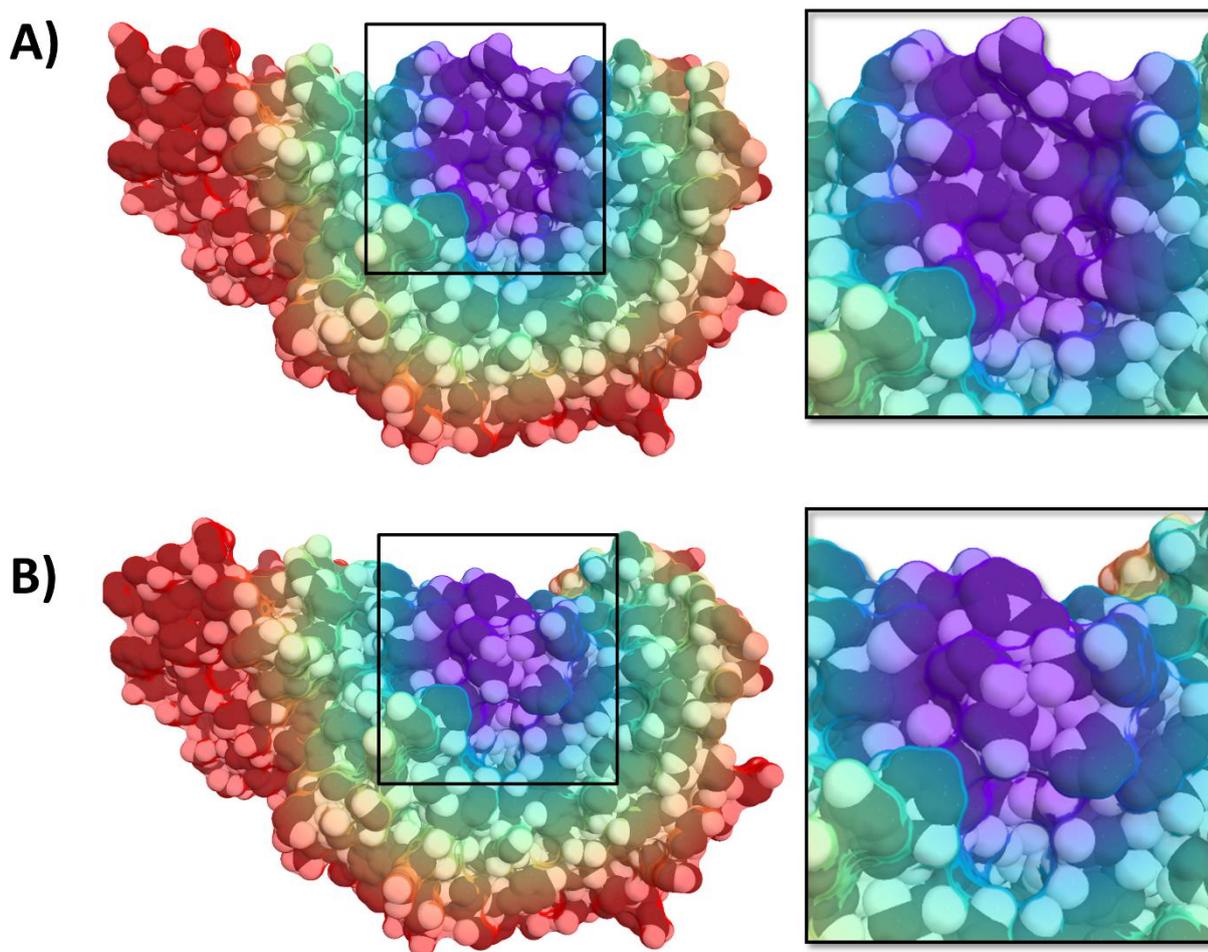


Figura 30 - Diferenças não cristalográficas observadas na estrutura de XCC_1754. Em A, vemos o loop da cadeia A na posição “aberta”. Em B, vemos o mesmo loop na posição “fechada” na cadeia B. O quadro a direita mostra detalhes de cada uma das respectivas regiões.

Fonte: Elaborada pelo autor.

A posição desse *loop* parece afetar a estabilidade de várias partes da enzima. A posição “aberta” parece estabilizar as hélices 6 e 7, enquanto que a posição “fechada” estabiliza a hélice 14 (Figura 30). A mobilidade de loops já foi observado em outras β -galactosidases da família 2,⁹⁴ porém nunca na família 35. Os homólogos da XCC_1754 são as enzimas de *Cellvibrio japonicus* (PDB 4D1I,⁴⁸) e a enzima de *Caulobacter crescentus* (PDB

3U7V, não publicada). Essas enzimas possuem respectivamente 55 e 58% de identidade sequencial, além de ambas terem cerca de 0,72 Å de RMSD (para 2884 e 2814 átomos, respectivamente) com cadeia A da enzima XCC_1754 (Figura 31).

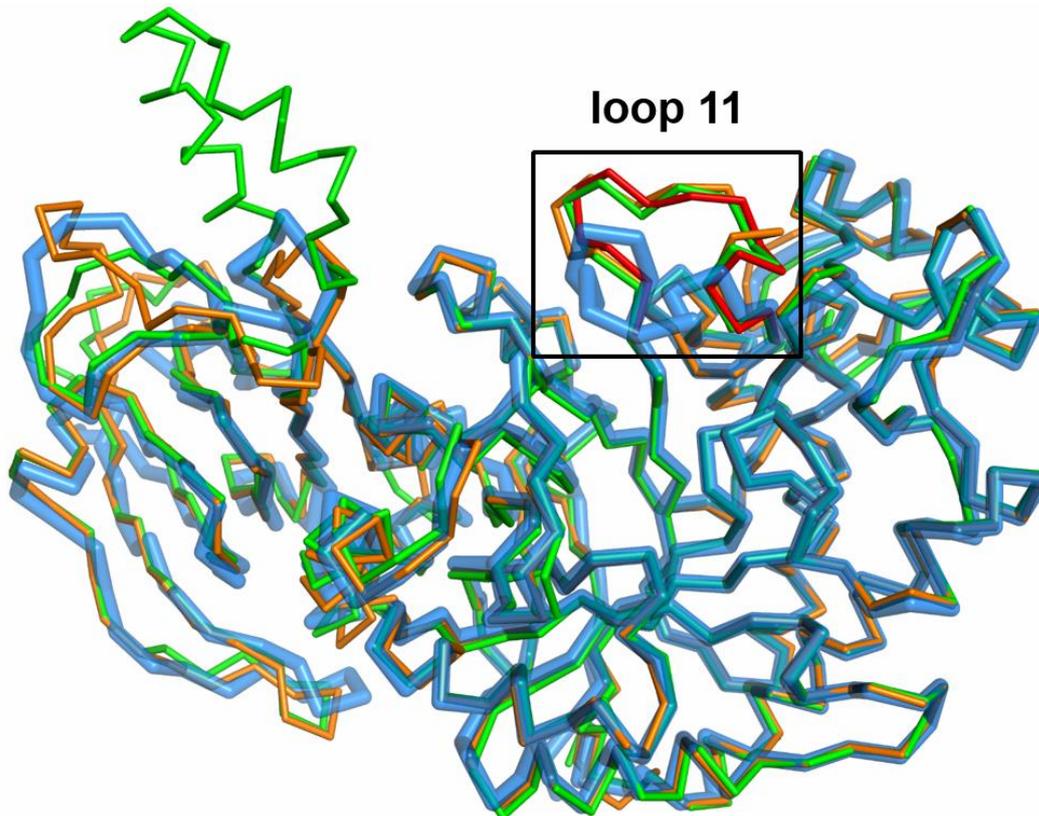


Figura 31 - Alinhamento da enzima XCC_1754 e homólogos da família 35. Em azul, a enzima XCC_1754 na posição fechada, enquanto que o *loop* em vermelho mostra a posição aberta da mesma. Em verde, a enzima Bgl35A. Em laranja, a enzima de *Caulobacter crescentus*.

Fonte: Elaborada pelo autor.

Podemos ver também pela estrutura que ambos os homólogos apresentaram o *loop* 11 na posição aberta, como a cadeia A de XCC_1754 (Figura 31). O *loop* 11 é formado por aminoácidos que mantem os ângulos phi (ϕ) e psi (ψ) semelhantes em ambas as cadeias, com exceção para os resíduos Gly294 e Gly302 (Figura 32). Nesses resíduos nós vemos uma torção completa dos ângulos no elemento de simetria não cristalográfico respectivo, sugerindo que as glicinas possam atuar como uma espécie de “dobradiça” para o *loop* como um todo. Curiosamente, quando alinhamos a estrutura da XCC_1754 com a Bgl35A contendo o ligante 1-deoxigalactonojirimicina (PDBid 4D1J), observamos que a posição fechada do *loop* 11 teria interações diretas com o ligante através da Gln300 (Figura 32).⁴⁸

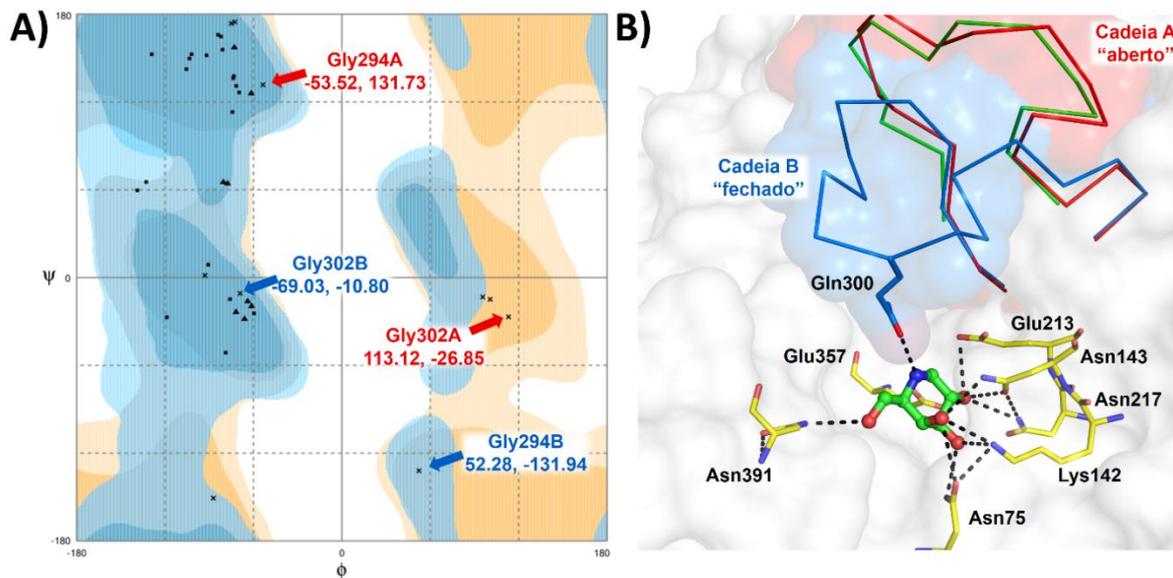


Figura 32 - Ramachandran e visão geral da região do *loop* 11 da enzima XCC_1754. Em A), o gráfico de Ramachandran para os resíduos 290-314 mostra que as glicinas 294 e 302 possuem ângulos bastante distintos. Em B), a região do *loop* com destaque para os resíduos catalíticos e o ligante sobreposto da Bgl35A, mostrando possíveis interações com o resíduo Gln300 do *loop* 11.

Fonte: Elaborada pelo autor.

Como podemos observar, o *loop* 11 está próximo a cavidade onde se encontram os resíduos catalíticos, e pode ter um papel no processo enzimático que ainda não foi explorado na família 35. Para explorarmos a função deste, três mutantes da enzima XCC_1754 foram construídos: o mutante G294P, G302P e o mutante P295_P301dele. Nos dois primeiros, as respectivas glicinas foram mudadas para prolinas, a fim de remover a mobilidade dos ângulos phi e psi. Já no último o *loop* 11 foi totalmente removido da proteína, para verificarmos o efeito da ausência de tal *loop* nas propriedades físico-químicas da proteína.

6.2 Caracterização bioquímica da enzima XCC_1754

A atividade relativa da enzima XCC_1754 foi testada em diferentes condições de pH e temperatura a fim de determinar os valores ótimos para ambos. Para isso utilizamos o substrato *o*NPG. Os resultados dessa caracterização podem ser vistos na Figura 33.

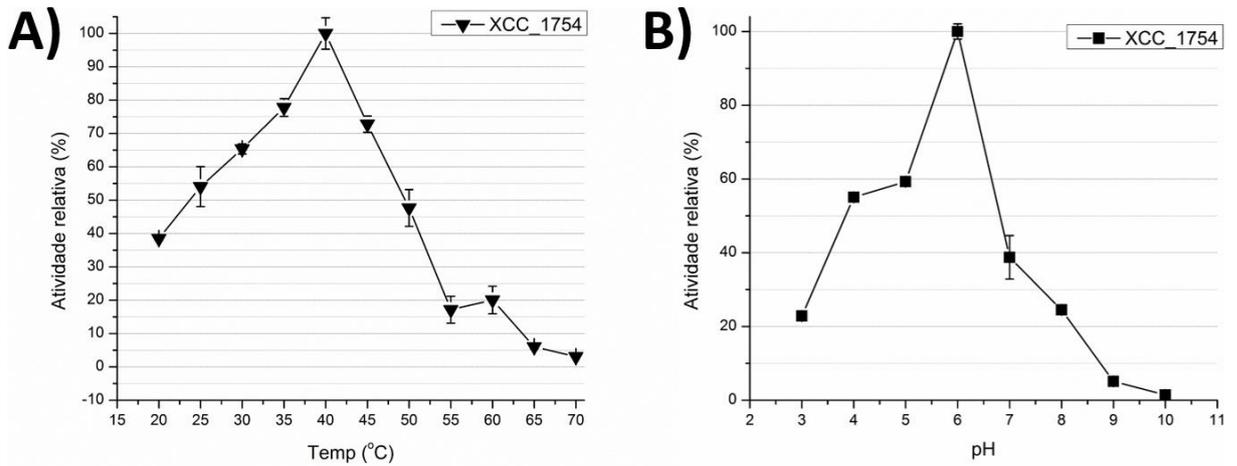


Figura 33 - Atividade relativa da enzima XCC_1754 em diferentes condições. Em A), a atividade relativa da enzima em diferentes temperaturas. Em B), a atividade relativa da enzima em diferentes pH's.

Fonte: Elaborada pelo autor.

Podemos observar que a enzima tem uma atividade ótima em 40 °C, sendo que a atividade cai pela metade quando a temperatura passa de 50 °C. O pH ótimo da enzima é por volta de 6,0, sendo que a atividade cai para cerca de 40% quando a enzima é colocada em pH 7,0. Com base nestes resultados nos testamos a estabilidade térmica da enzima e seus mutantes através do experimento de termo fluorimetria em tampões próximos ao pH ótimo. Os resultados podem ser vistos na Figura 34.

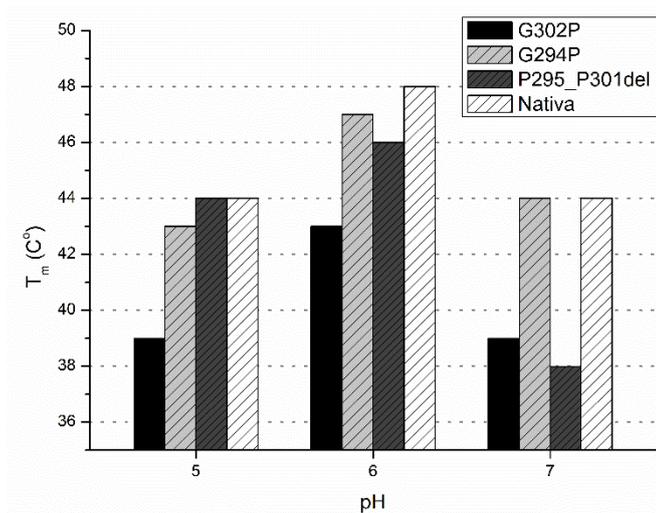


Figura 34 - Termo flourimetria da enzima XCC_1754 e seus mutantes. Os testes foram feitos em pH's 5, 6 e 7 e os T_m obtidos estão mostrados em barras.

Fonte: Elaborada pelo autor.

Como podemos observar na Figura 34, a enzima XCC_1754 possui um T_m de cerca de 48 °C no pH 6,0, e cerca de 44 °C nos pH's 5,0 e 7,0. O mutante G294P sofreu uma pequena queda no T_m nestas temperaturas, com no máximo um grau de diferença para cada pH. Já o mutante G302P sofreu uma acentuada diminuição na estabilidade térmica, sendo que no pH 6,0 houve a diminuição de 5 °C na T_m em relação a nativa. Por fim, o mutante P295_P301dele também apresentou uma significativa diminuição na T_m , especialmente no pH 7,0 onde a diferença é de 6 °C.

6.3 Cinética da enzima XCC_1754 e seus mutantes

Para compreendermos melhor o papel do *loop 11*, nós fizemos a caracterização cinética da enzima nativa e dos mutantes G294P, G302P e P295_P301dele. Todos os testes foram feitos à 30 °C com o substrato *o*NPG, e os resultados estão listados na Tabela 9.

Tabela 9 - Parâmetros cinéticos da enzima XCC_1754 e seus mutantes.

	V_{max} ($\mu\text{M/s}$)	K_m (mM)	k_{cat} (s^{-1})	V_{max}/K_m (s^{-1})	k_{cat}/K_m ($\text{mM}^{-1}.\text{s}^{-1}$)	Atividade (%)
Nativa	$6,44 \pm 0,35$	$2,9 \pm 0,7$	265	0,002	89,8	100
G294P	$18,25 \pm 0,5$	$5,5 \pm 0,5$	752	0,003	135,7	151
G302P	$0,37 \pm 0,005$	$0,65 \pm 0,06$	1,52	0,0005	2,3	2,6
P295_P301dele	$1,22 \pm 0,04$	$0,95 \pm 0,17$	10	0,0013	10,6	11,8

Fonte: Elaborada pelo autor

6.4 Conclusões em relação a enzima XCC_1754

A enzima XCC_1754 apresentou uma estrutura similar a outras estruturas de β -galactosidases pertencentes a família 35. Sua estrutura é formada por dois domínios, sendo que o primeiro forma um barril do tipo TIM e contém todos os resíduos catalíticos conservados nos homólogos. Com uma unidade assimétrica formada por dois monômeros, os contatos cristalinos revelaram duas isoformas distintas do *loop 11* da enzima, na qual a posição “fechada” estaria em contato próximo ao substrato ligado ao sitio ativo. Quando

comparamos as diferentes posições dos *loops*, vemos que as glicinas 294 e 302 tem ângulos phi e psi bastante divergentes entre as duas cadeias, sugerindo que tais resíduos possam atuar como uma dobradiça para tal *loop*. Com isso em mente tais resíduos foram mutados para prolinas, a fim de eliminar qualquer mobilidade nesses ângulos, e os efeitos da mutação foram avaliados e comparados com a nativa.

A enzima nativa apresentou um K_m de cerca de 3,0 mM, e um k_{cat} de cerca de 265 s^{-1} , sendo que nesse caso a eficiência catalítica (k_{cat}/K_m) é cerca de 90 $mM^{-1}.s^{-1}$. O mutante G294P por sua vez apresentou uma diminuição na afinidade pelo substrato, com um K_m de cerca de 5,5 mM. Apesar disso, o k_{cat} desse mutante é cerca de três vezes maior que a enzima nativa, e a eficiência catalítica é quase 50% maior. Podemos ver o efeito oposto a esse nos outros dois mutantes (G302P e P295_P301dele), onde para o ambos a afinidade da enzima pelo substrato é superior a nativa, mas a eficiência catalítica é drasticamente reduzida. Com isso, acreditamos que as mutações tiveram efeitos opostos na posição do *loop* 11, sendo que é provável que o mutante G294P manteve o *loop* em uma posição próxima a fechada, dificultando o acesso do substrato ao sítio, mas ao mesmo tempo facilitando a clivagem. Já os mutantes G302P e P295_P301dele possivelmente travaram o *loop* na posição próxima a aberta, permitindo o fácil reconhecimento do substrato, porém, dificultando a clivagem e, assim, diminuindo a eficiência da enzima.

CAPÍTULO 7

Resultados & Discussão V

Enzima XCC_2895

7 Estrutura e função da enzima XCC_2895.

7.1 Estrutura tridimensional da enzima XCC_2895

A enzima XCC_2895 foi resolvida em 1,7 Å de resolução, com um $R_{\text{work}}/R_{\text{free}}$ final de 0,16/0,20. Essa enzima também apresentou uma unidade assimétrica composta por duas moléculas, porém o programa PISA indica que as interfaces não são estáveis em solução. O enovelamento da enzima revelou a presença de três domínios. O primeiro forma um barril-TIM, e contém os resíduos catalíticos da enzima. Os outros dois são sanduiches de fitas β e possuem função desconhecida (Figura 35).⁹⁵

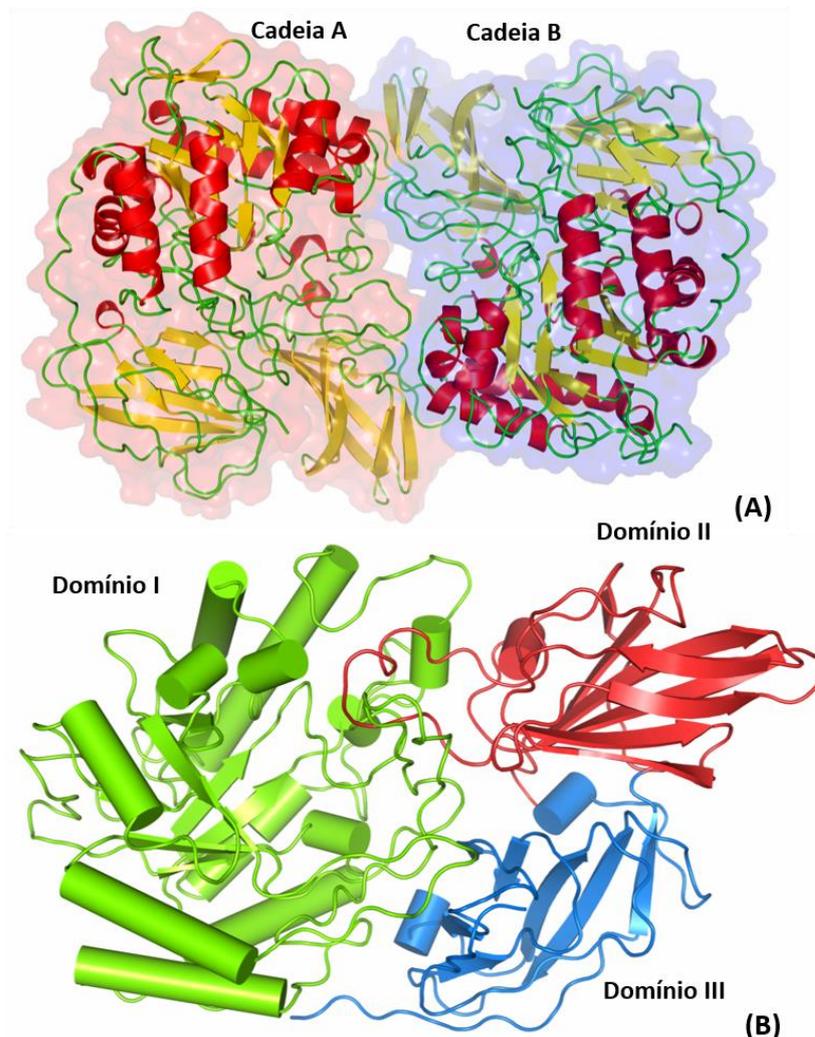


Figura 35 - Estrutura cristalográfica da enzima XCC_2895. Em A), a unidade assimétrica da estrutura cristalográfica. Em B), a visão do monômero com os três domínios em destaque.

Fonte: Elaborada pelo autor

A enzima XCC_2895 possui quatro homólogas com estruturas depositadas no PDB, incluindo a enzima de *Bacillus circulans* (PDB 4MAD, 43% de identidade),⁹⁵ a enzima de *Streptococcus pneumoniae* (PDB 4E8C, 40% de identidade),⁷⁹ a enzima de *Bacteroides thetaiotaomicron* (PDB 3D3A, 36% identidade, não publicada) e a β -galactosidase de humanos (PDB 3THC, 35% identidade).⁹⁶ Essas enzimas possuem respectivamente 0,85, 1,00, 1,00 e 0,94 Å de RMSD (para 585 carbonos alfa) com a XCC_2895 (Figura 36).

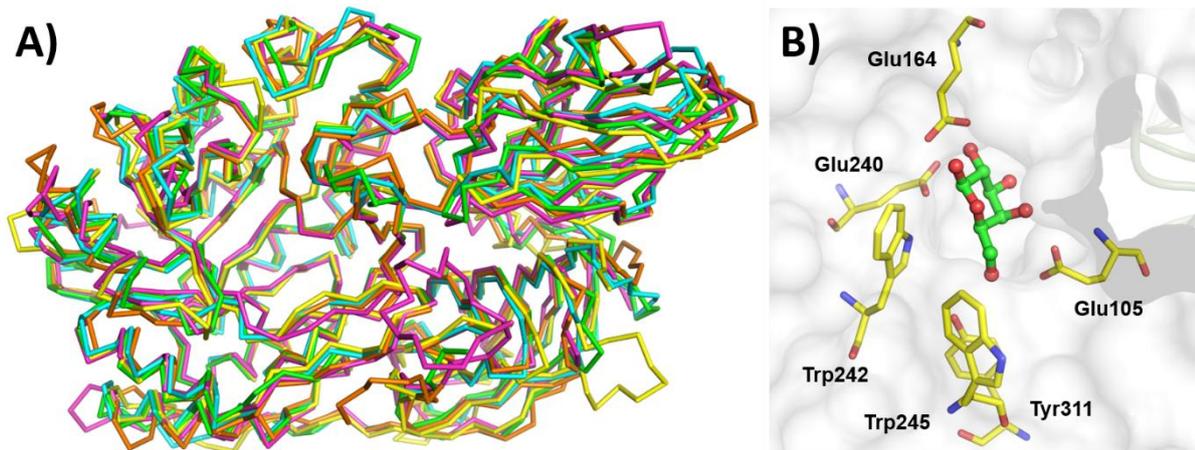


Figura 36 - Alinhamento e sítio ativo da enzima XCC_2895. Em A) o alinhamento da enzima XCC_2895 (verde) com os PDBs 4MAD (azul), 4E8C (laranja), 3D3A (amarelo) e 3THC (rosa). Em B), o sítio da XCC_2895 com os resíduos catalíticos em destaque. Também temos em B) o ligante da enzima de *Streptococcus pneumoniae* alinhado ao sítio da XCC_2895.

Fonte: Elaborada pelo autor.

7.2 Caracterização bioquímica, cinética e considerações da enzima XCC_2895.

A atividade relativa da enzima XCC_2895 foi testada em diferentes condições de pH e temperatura a fim de determinar os valores ótimos para ambos. Para isso utilizamos o substrato *o*NPG. Os resultados dessa caracterização podem ser vistos na (Figura 37).

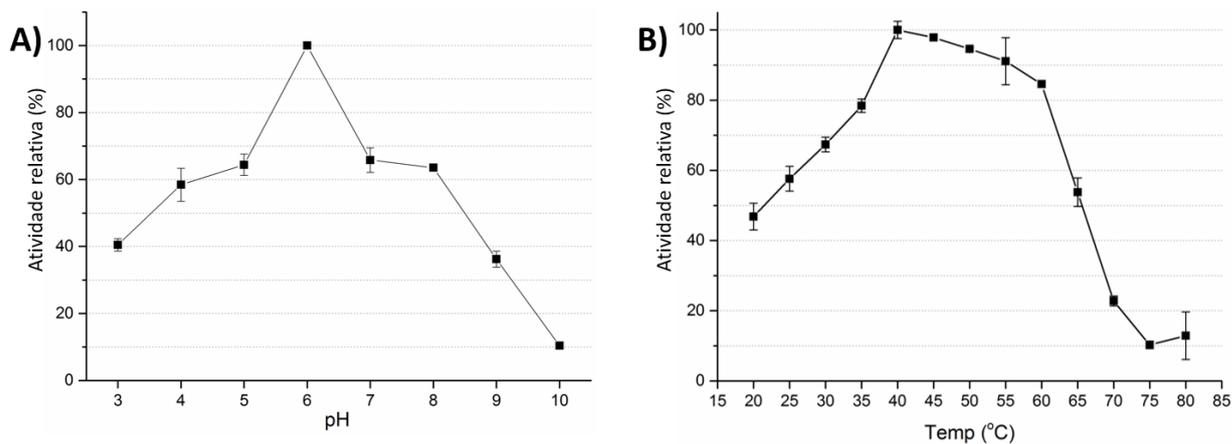


Figura 37 - Atividade relativa da enzima XCC_2895 em diferentes pH's e temperaturas. Em A), a atividade da enzima em diferentes pH's. Em B), a atividade da enzima em diferentes temperaturas.

Fonte: Elaborada pelo autor.

A enzima XCC_2895 apresenta um pH ótimo por volta de 6,0, enquanto que a temperatura é cerca de 40 °C. Porém diferente da enzima XCC_1754, essa enzima mostra maior estabilidade térmica em temperaturas mais elevadas, sendo que mais de 80% da atividade é mantida ainda à 60 °C. O K_m da enzima para o substrato *o*NPG foi de cerca de $1,5 \pm 0,6$, enquanto que a V_{max} foi cerca de $2,2 \mu\text{M}\cdot\text{s}^{-1}$. O k_{cat} da enzima para tal substrato é cerca de 269 s^{-1} , e a especificidade cerca de $178,0 \text{ mM}^{-1}\cdot\text{s}^{-1}$.

A enzima XCC_2895 possui um perfil bioquímico semelhante a enzima XCC_1754, porém com uma arquitetura bastante diferente. Enquanto a XCC_1754 possuiu dois domínios, a primeira é formada por três domínios diferentes. Os experimentos realizados até o momento não mostraram nenhuma diferença que possa ser explicada por esta diferença na arquitetura. Portanto, no futuro, pretendemos comparar a atividade dessas enzimas em diferentes substratos, como xiloglucano e pectina, a fim de verificar se existe diferenças entre as enzimas.

CAPÍTULO 8

Resultados & Discussão VI

Enzima XCC_2404 e Considerações Finais

8 Estrutura e Função da enzima XCC_2404

8.1 Estrutura cristalográfica da enzima XCC_2404

A enzima XCC_2404 apresentou uma unidade assimétrica incomum. Isso porque o enorme volume (cerca de 4.900.000 Å³), e a baixa simetria (P1 2₁ 1) da cela unitária sugerem que a unidade assimétrica seja composta por 16 cadeias. Com ajuda da simetria não cristalográfica foi possível refinar essa estrutura até valores de R_{work}/R_{free} de 20/24 (%). A visão da unidade assimétrica, bem como a do monômero da enzima pode ser vista na Figura 32.

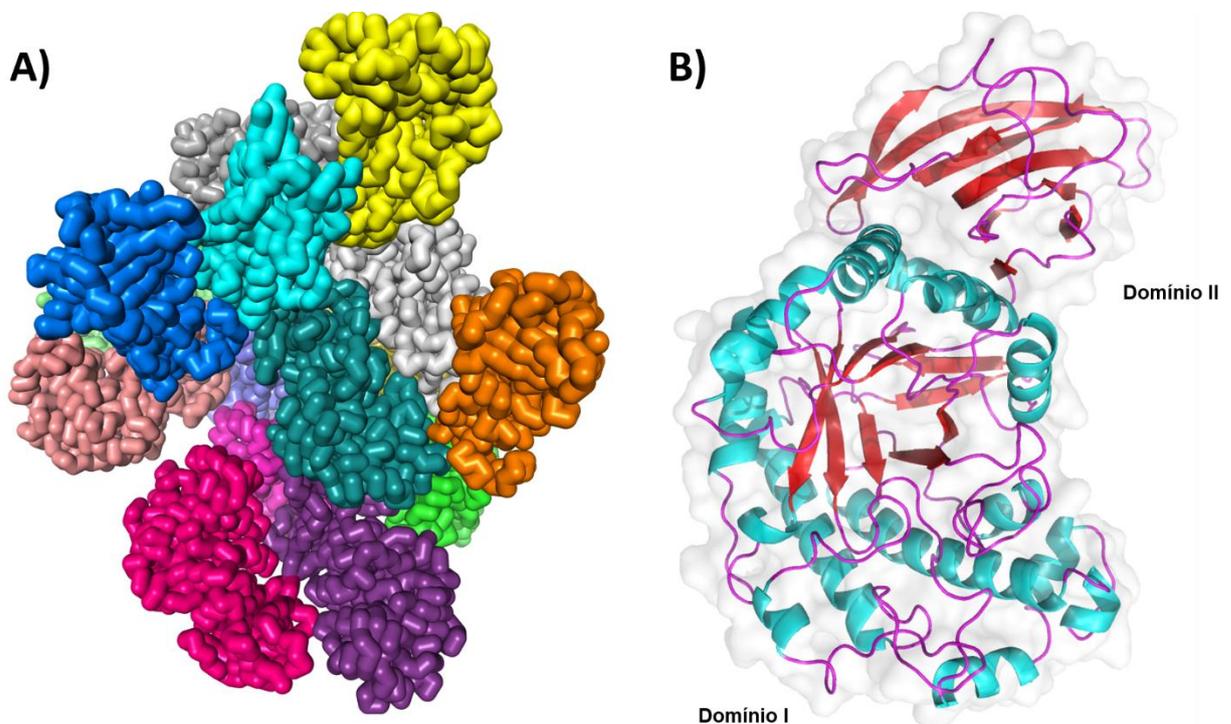


Figura 38 - Modelo cristalográfico inicial da enzima XCC_2404. Em A), a unidade assimétrica do modelo cristalográfico da enzima XCC_2404 onde cada cadeia está representada por uma cor diferente. Em B), o monômero da enzima formado por dois domínios.

Fonte: Elaborada pelo autor.

Assim como a XCC_1754, a enzima XCC_2404 possui dois domínios, sendo que o primeiro tem o enovelamento do tipo barril TIM, e o segundo domínio formado por fitas. Tais enzimas possuem 41% de identidade entre si, sendo que a XCC_2404 também tem 44% e 41% de identidade com as enzimas de *Caulobacter crescentus* e *Cellvibrio japonicus*,⁴⁸

respectivamente. Apesar do grande número de moléculas na unidade assimétrica, análises com o programa PISA,⁸³ não revelaram nenhum contato formando uma organização estável nessa determinada conformação.

8.2 Estado oligomérico da enzima XCC_2404

Tanto os resultados observados na gel filtração analítica (item 3.5), quanto a unidade assimétrica formada por múltiplas cadeias observada na estrutura cristalográfica nos fizeram querer estudar o estado oligomérico dessa enzima. Para verificarmos se o tamanho observado em solução não se trata de uma agregação inespecífica, a enzima foi tratada com diferentes compostos a fim de eliminar qualquer forma de interação comum entre proteínas. Para isso, a proteína foi tratada com 10 mM de ácido etilenodiamino tetra-acético (EDTA), 10 mM de ditioneitol (DTT) ou 2,0 M de NaCl, a fim de eliminar qualquer interação inespecífica. Após 30 min, a proteína foi analisada em gel nativo para verificarmos o estado oligomérico. O resultado pode ser visto na Figura 33.

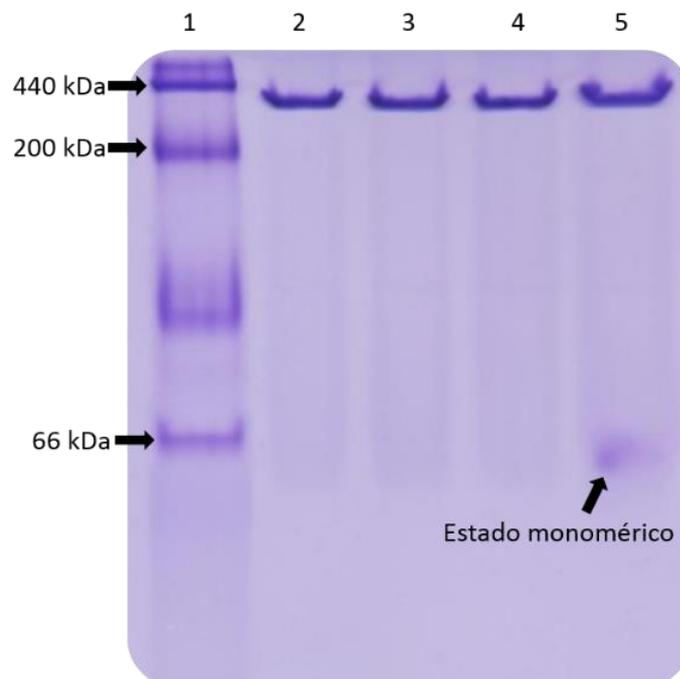


Figura 39 - Gel Nativo da enzima XCC_2404. Em 1, o marcador. Em 2 a enzima nativa. Em 3, a enzima em 10 mM de DTT. Em 4, a enzima em 10 mM de EDTA. Em 5, a enzima em 2,0 M de NaCl.

Fonte: Elaborada pelo autor

Considerando que o monômero possui uma massa teórica de 55 kDa (Tabela 3), podemos ver que a enzima XCC_2404 forma aglomerados de alta massa molecular, com mais de 300 kDa, sugerindo a formação de pelo menos hexâmeros em solução (Figura 33). Podemos também ver na figura 39 que esse estado de oligomerização é estável mesmo após tratamentos drásticos, sendo que, somente o tratamento com NaCl 2,0 M foi capaz de liberar enzima no estado monomérico.

Para estimar melhor o tamanho esperado do oligômero utilizamos a técnica de Espalhamento Dinâmico de Luz e Potencial Zeta (DLS), a fim de calcular o raio hidrodinâmico da proteína em solução. Os dados sugerem que a proteína possui um raio hidrodinâmico de $7,85 \pm 0,37 \mu\text{m}$, e o oligômero teria uma massa molecular de 374,1 kDa, equivalente a 6,8 moléculas de XCC_2404. A massa equivalente a um heptâmero seria condizente com a massa observada no gel nativo (Figura 33), porém mais estudos são necessários para concluirmos a respeito destes valores.

8.3 Parâmetros bioquímicos da enzima XCC_2404

A atividade relativa da enzima XCC_2404 foi testada em diferentes condições de pH e temperatura a fim de determinar os valores ótimos para ambos. Para isso utilizamos o substrato *o*NPG. Os resultados dessa caracterização podem ser vistos na Figura 34.

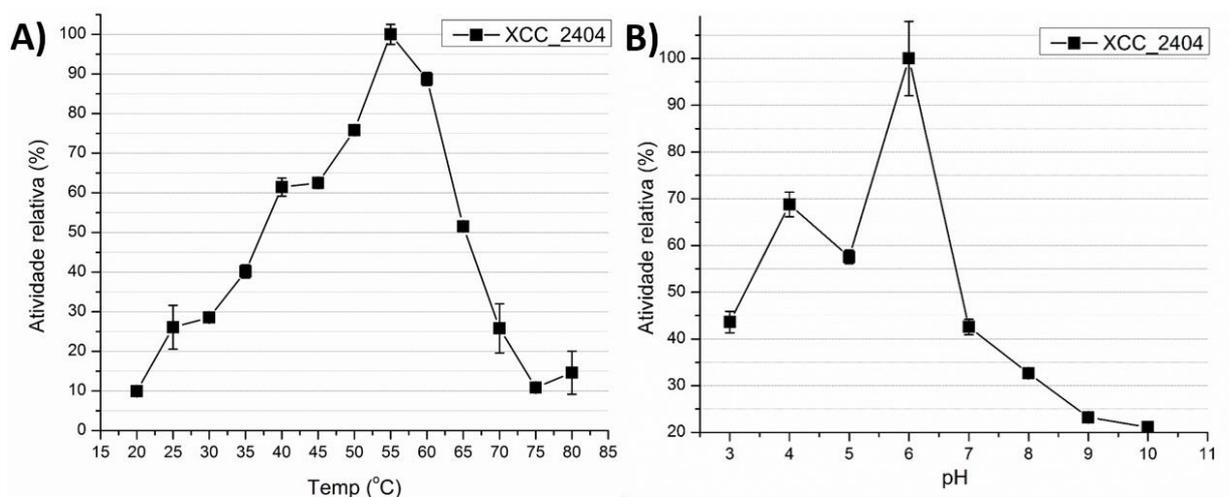


Figura 40 - Caracterização bioquímica da enzima XCC_2404. Em A), a atividade relativa da enzima em diferentes temperaturas. Em B), a atividade relativa da enzima em diferentes pH's.

Podemos ver através da figura 40 que a enzima XCC_2404 trabalha em pHs similares a seu homólogo XCC_1754, com atividade máxima em pH 6,0 e que decai rapidamente em outros valores. Em contrapartida, podemos ver que essa enzima possui uma temperatura ótima bastante elevada em relação ao seu homólogo em *Xanthomonas*. Enquanto a enzima XCC_1754 tem uma temperatura ótima à 40 °C (item 3.10.2), a enzima XCC_2404 possui uma temperatura ótima de 55 °C, além de reter cerca de 90% da atividade em 60 °C (Figura 34). Para testar a estabilidade da enzima, utilizamos também a técnica de termo fluorimetria em quatro pHs diferentes

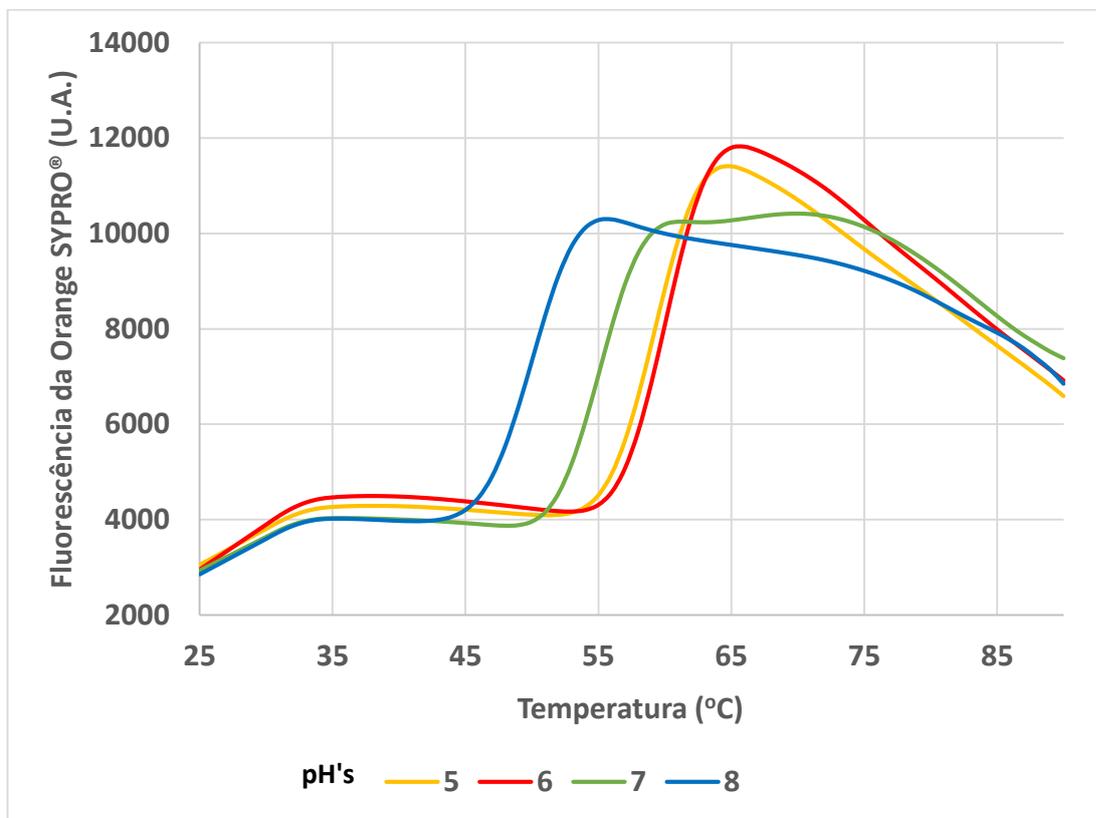


Figura 41 - Curvas de termo fluorimetria da enzima XCC_2404. Cada curva mostra a fluorescência da amostra em um pH's diferente, de acordo com a legenda.

Fonte: Elaborada pelo autor.

Podemos ver pela curva que a enzima possui maior estabilidade nos pHs 5 e 6 ($T_m \sim 60$ °C) que nos pHs 7 e 8 (T_m 55 e 50 °C, respectivamente). Isso mostra que a enzima XCC_2404 é mais estável que seu homólogo XCC_1754.⁷¹ Além disso, o decaimento lento da fluorescência sugere que ocorre a formação de oligômeros, que se dissociam somente em temperaturas próximas a T_m da enzima.

8.3 Discussão sobre a enzima XCC_2404

A estrutura da enzima XCC_2404 possui um enovelamento típico visto em homólogos da família 35, incluindo a enzima XCC_1754, também pertencente a *Xanthomonas*. A molécula é formada por dois domínios, sendo que um deles forma um barril e contém todos os resíduos catalíticos. Entretanto, a enzima aparenta formar grandes oligômeros em solução, bastante diferente de suas homólogas.⁷¹ A gel filtração analítica sugere que enzima forma hexâmeros, enquanto o DLS e o gel nativo sugerem que a proteína forma heptâmeros em solução. Como esses resultados divergem, ainda são necessários mais experimentos a fim de afirmar o estado oligomérico dessa enzima. Mesmo assim, podemos ver que apesar da similaridade com a enzima XCC_1754, a enzima XCC_2404 é significativamente mais termo estável em solução, o que pode estar relacionado ao fato dessa enzima formar grandes oligômeros.⁹⁷ Por fim, vemos também que a organização quaternária resistiu a tratamentos diversos, indicando que esse estado é relativamente estável em solução.

9 Conclusões e Perspectivas

Nesse projeto foi possível realizar a clonagem de todos os alvos selecionados inicialmente em diversos vetores de expressão. Dos nove alvos iniciais, realizamos a purificação e cristalização de quatro β -galactosidases (sendo uma de *B. bifidum* e três de *X. campestris*). A enzima BbgII, pertencente à família 42, teve sua estrutura cristalográfica resolvida e seus estudos funcionais concluídos. Também foi possível realizar mutações sítio-dirigidas em resíduos catalíticos importantes, e a análise dessas mutações nas propriedades cinéticas da enzima.

Quanto às enzimas de *X. campestris*, pertencentes a família 35, elas foram purificadas e cristalizadas. As enzima XCC_1754 exibiu uma diferença cristalográfica que nos permitiu propor um modelo para o funcionamento do *loop* 11, o que não havia sido descrito para a família 35. A enzima XCC_2895 possui três domínios, mas não observamos nenhuma diferença na atividade dessa enzima quando comparada com o homólogo XCC_1754, que possui apenas dois. Nesse caso, estudos mais específicos para mapear a preferência por substratos e tipos de ligação deverão ser realizados no futuro.

Por fim, a enzima XCC_2404 apresentou um modelo cristalográfico complexo, com 16 moléculas na unidade assimétrica. Ensaio bioquímicos, como a gel filtração analítica e gel nativo indicam que essa enzima forma grandes oligômeros, de pelo menos seis subunidades. Essa estrutura quaternária ainda não foi elucidada, porém as análises bioquímicas indicam que tal conformação traz um significativo aumento na estabilidade térmica da enzima quando comparada ao homólogo XCC_1754.

Referências

- 1 ROUWENHORST, R. J.; PRONK, J. T.; VAN DIJKEN, J. P. The discovery of beta-galactosidase. **Trends in Biochemical Sciences**, v. 14, n. 10, p. 416-418, 1989.
- 2 JACOB, F. et al. Operon: a group of genes with the expression coordinated by an operator. **Comptes Rendus Hebdomadaires des Séances de L'Académie des Sciences**, v. 250, p. 1727-9, 1960.
- 3 SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. **Molecular cloning: a laboratory manual**. Plainview: Cold Spring Harbour, 1989.
- 4 DIMRI, G. P. et al. A biomarker that identifies senescent human cells in culture and in aging skin in vivo. **Proceedings of the National Academy of Sciences**, v. 92, n. 20, p. 9363-9367, 1995.
- 5 GARY, R. K.; KINDELL, S. M. Quantitative assay of senescence-associated beta-galactosidase activity in mammalian cell extracts. **Analytical Biochemistry**, v. 343, n. 2, p. 329-334, 2005.
- 6 ITAHANA, K.; CAMPISI, J.; DIMRI, G. P. Methods to detect biomarkers of cellular senescence: the senescence-associated beta-galactosidase assay. **Methods in Molecular Biology**, v. 371, p. 21-31, 2007. doi: 10.1007/978-1-59745-361-5_3.
- 7 JUERS, D. H.; MATTHEWS, B. W.; HUBER, R. E. LacZ beta-galactosidase: structure and function of an enzyme of historical and molecular biological importance. **Protein Science**, v. 21, n. 12, p. 1792-1807, 2012.
- 8 JACOBSON, R. H. et al. Three-dimensional structure of beta-galactosidase from *E. coli*. **Nature**, v. 369, n. 6483, p. 761-766, 1994.
- 9 BARTESAGHI, A.; MERK A.; BANERJEE S.; MATTHIES D.; WU X.; MILNE J. L.; SUBRAMANIAM S. 2.2 a resolution cryo-EM structure of beta-galactosidase in complex with a cell-permeant inhibitor. **Science**, v. 348, n. 6239, p. 1147-1151, 2015.
- 10 LAINE, R. A. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. **Glycobiology**, v. 4, n. 6, p. 759-767, 1994.
- 11 CANTAREL, B. L. et al. The carbohydrate-active enzymes database (CAZy): an expert resource for glycomics. **Nucleic Acids Research**, v. 37, n. 1, p. D233-D238, 2009.
- 12 LOMBARD, V. et al. The carbohydrate-active enzymes database (CAZy) in 2013. **Nucleic Acids Research**, v. 42, n. 1, p. D490-D495, 2014.
- 13 LEVASSEUR, A. et al. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. **Biotechnology for Biofuels**, v. 6, n. 1, p. 41, 2013.

14 CUPPLES, C. G.; MILLER, J. H.; HUBER, R. E. Determination of the roles of Glu-461 in beta-galactosidase (*Escherichia coli*) using site-specific mutagenesis. **Journal of Biological Chemistry**, v. 265, n. 10, p. 5512-5518, 1990.

15 HUSAIN, Q. Beta galactosidases and their potential applications: a review. **Critical Reviews in Biotechnology**, v. 30, n. 1, p. 41-62, 2010.

16 JUERS, D. H. et al. A structural view of the action of *Escherichia coli* (lacZ) beta-galactosidase. **Biochemistry**, v. 40, n. 1, p. 14781-14794, 2001.

17 ULLMANN, A. Complementation in beta-galactosidase: from protein structure to genetic engineering. **Bioessays**, v. 14, n. 3, p. 201-205, 1992.

18 JUERS, D. H. et al. High resolution refinement of beta-galactosidase in a new crystal form reveals multiple metal-binding sites and provides a structural basis for alpha-complementation. **Protein Science**, v. 9, n. 9, p. 1685-1699, 2000.

19 ROJAS, A. L. et al. Crystal structures of beta-galactosidase from *Penicillium sp.* and its complex with galactose. **Journal of Molecular Biology**, v.343, n. 5, p.1281-1292, 2004.

20 MAKSIMAINEN, M. et al. Crystal structures of *Trichoderma reesei* beta-galactosidase reveal conformational changes in the active site. **Journal of Structural Biology**, v.174, n. 1, p. 156-163, 2011.

21 HIDAKA, M. et al. Trimeric crystal structure of the glycoside hydrolase family 42 beta-galactosidase from *Thermus thermophilus* A4 and the structure of its complex with galactose. **Journal of Molecular Biology**, v.322, n. 1, p.79-91, 2002.

22 MAKSIMAINEN, M. et al. Structural analysis, enzymatic characterization, and catalytic mechanisms of beta-galactosidase from *Bacillus circulans sp. alkalophilus*. **FEBS Journal**, v. 279, n. 10, p. 1788-1798, 2012.

23 KOSHLAND, D. E. Stereochemistry and the mechanism of enzymatic reactions. **Biological Reviews**, v. 28, n. 4, p. 416-436, 1953.

24 DAVIES, G.; HENRISSAT, B. Structures and mechanisms of glycosyl hydrolases. **Structure**, v. 3, n. 9, p. 853-859, 1995.

25 HUBER, R. E.; KURZ, G.; WALLENFELS, K. A quantitation of the factors which affect the hydrolase and transgalactosylase activities of β -galactosidase (*E. coli*) on lactose. **Biochemistry**, v. 15, n. 9, p. 1994-2001, 1976.

26 RODRÍGUEZ-NOGALES, J. M.; LÓPEZ, A. D. A novel approach to develop β -galactosidase entrapped in liposomes in order to prevent an immediate hydrolysis of lactose in milk. **International Dairy Journal**, v. 16, n. 4, p. 354-360, 2006.

27 WILSON, J. Milk intolerance: lactose intolerance and cow's milk protein allergy. **Newborn and Infant Nursing Reviews**, v. 5, n. 4, p. 203-207, 2005.

- 28 CAPPELLO, G.; MARZIO, L. Rifaximin in patients with lactose intolerance. **Digestive and Liver Disease**, v. 37, n. 5, p. 316-319, 2005.
- 29 HAIDER, T.; HUSAIN, Q. Immobilization of β galactosidase from *Aspergillus oryzae* via immunoaffinity support. **Biochemical Engineering Journal**, v. 43, n. 3, p. 307-314, 2009.
- 30 RIOS, G. M. et al. Progress in enzymatic membrane reactors – a review. **Journal of Membrane Science**, v. 242, n. 1–2, p. 189-196, 2004.
- 31 ROBERFROID, M. Prebiotics: the concept revisited. **Journal of Nutrition**, v. 137, n. 3, p. 830-837, 2007.
- 32 HOPKINS, M. J.; MACFARLANE, G. T. Nondigestible oligosaccharides enhance bacterial colonization resistance against *Clostridium difficile* in vitro. **Applied Environmental Microbiology**, v. 69, n. 4, p. 1920-1927, 2003.
- 33 PARK, A. R.; OH, D. K. Galacto-oligosaccharide production using microbial β -galactosidase: current state and perspectives. **Applied Microbiology and Biotechnology**, v. 85, n. 5, p. 1279-1286, 2010.
- 34 COLLINS, M. D.; GIBSON, G. R. Probiotics, prebiotics, and synbiotics: approaches for modulating the microbial ecology of the gut. **American Journal of Clinical Nutrition**. v. 65, n. 5, p. 1052-1057, 1999.
- 35 MACFARLANE, S.; MACFARLANE, G. T.; CUMMINGS, J. H. Review article: prebiotics in the gastrointestinal tract. **Alimentary Pharmacology & Therapeutics**, v. 24, n. 5, p.701-14, 2006
- 36 TANIGUCHI, H. Carbohydrate active enzymes for the production of oligosaccharides. In: HOUM, C. T. (Ed.) **Handbook of industrial biocatalysis**. New York: CRC Press, 2005. p. 20-1-20-25.
- 37 CARPITA, N. C.; GIBEAUT, D. M. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. **Plant Journal**, v. 3, n. 1, p. 1-30, 1993.
- 38 SØRENSEN, I.; DOMOZYCH, D.; WILLATS, W. G. T. How have plant cell walls evolved? **Plant Physiology**, v. 153, n. 2, p. 366-372, 2010.
- 39 GRABBER, J. H. How do lignin composition, structure, and cross-linking affect degradability? a review of cell wall model studies. **Crop Science**, n. 3, p. 820-831, 2005. doi:10.2135/cropsci2004.0191.
- 40 COSGROVE, D. J. Growth of the plant cell wall. **Nature Reviews Molecular and Cellular Biology**, v. 6, n. 11, p. 850-861, 2005.
- 41 LYND, L. R. et al. Microbial cellulose utilization: fundamentals and biotechnology. **Microbiology and Molecular Biology Reviews**. v. 66, n. 3, p. 506-577, 2002.

- 42 LYND, L. R. How biotech can transform biofuels. **Nature Biotechnology**, v. 26, n. 2, p. 169-172, 2008.
- 43 MOHANRAM, S. et al. Novel perspectives for evolving enzyme cocktails for lignocellulose hydrolysis in biorefineries. **Sustainable Chemical Processes**, v. 1, n. 1, p. 1-12, 2013.
- 44 SCHUSTER, B.; CHINN, M. Consolidated bioprocessing of lignocellulosic feedstocks for ethanol fuel production. **Bioenergy Research**, v. 6, n. 2, p. 416-435, 2013.
- 45 MARTIN, I. et al. In vivo expression of a *Cicer arietinum* beta-galactosidase in potato tubers leads to a reduction of the galactan side-chains in cell wall pectin. **Plant Cell Physiology**, v. 46, n. 10, p. 1613-22, 2005.
- 46 SHALLOM, D.; SHOHAM, Y. Microbial hemicellulases. **Current Opinion in Microbiology**, v. 6, n. 3, p. 219-228, 2003.
- 47 HEMSWORTH, G. R. et al. Discovery and characterization of a new family of lytic polysaccharide monoxygenases. **Nature Chemical Biology**, v. 10, n. 2, p. 122-126, 2014.
- 48 LARSBRINK, J. et al. A complex gene locus enables xyloglucan utilization in the model saprophyte *Cellvibrio japonicus*. **Molecular Microbiology**, v. 94, n. 2, p. 418-433, 2014.
- 49 QIAN, W. et al. Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. *campestris*. **Genome Research**, v. 15, n. 6, p. 757-767, 2005.
- 50 YANG, T. C.; WU, G. H.; TSENG, Y. H. Isolation of a *Xanthomonas campestris* strain with elevated beta-galactosidase activity for direct use of lactose in xanthan gum production. **Letters in Applied Microbiology**, v.35, n. 5, p. 375-379, 2002.
- 51 BIAVATI, B. et al. *Bifidobacteria*: history, ecology, physiology and applications. **Annals of Microbiology**, v. 50, n. 2, p. 117-131, 2000.
- 52 LAY, C. et al. Colonic microbiota signatures across five northern european countries. **Applied Environmental Microbiology**, v.71, n. 7, p.4153-4155, 2005.
- 53 PALFRAMAN, R. J.; GIBSON, G. R.; RASTALL, R. A. Carbohydrate preferences of *Bifidobacterium* species isolated from the human gut. **Current Issues in Intestinal Microbiology**, v. 4, n. 2, p. 71-75, 2003.
- 54 SLOVAKOVA, L.; DUSKOVA, D.; MAROUNEK, M. Fermentation of pectin and glucose, and activity of pectin-degrading enzymes in the rabbit caecal bacterium *Bifidobacterium pseudolongum*. **Letters in Applied Microbiology**, v. 35, n. 2, p. 126-130, 2002.
- 55 KAPLAN, H.; HUTKINS, R. W. Fermentation of fructooligosaccharides by lactic acid bacteria and bifidobacteria. **Applied Environmental Microbiology**, v. 66, n. 6, p. 2682-2684, 2000.

56 ASLANIDIS, C.; DE JONG, P. J. Ligation-independent cloning of PCR products (LIC-PCR). **Nucleic Acids Research**, v. 18, n. 20, p. 6069-6074, 1990.

57 HAUN, R. S.; SERVENTI, I. M.; MOSS, J. Rapid, reliable ligation-independent cloning of PCR products using modified plasmid vectors. **Biotechniques**, v. 13, n. 4, p. 515-518, 1992.

58 RABHI, I. et al. A novel simple and rapid PCR-based site-directed mutagenesis method. **Molecular Biotechnology**, v. 26, n. 1, p. 27-34, 2004.

59 WILLIAM, F. S. Protein production by auto-induction in high-density shaking cultures. **Protein Expression and Purification**, v. 41, n. 1, p. 207-234, 2005.

60 MCPHERSON, A. **Crystallization of biological macromolecules**. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1944.

61 KABSCH, W. XDS. **Acta Crystallographica D: biological crystallography**, v. 66, n. 2, p. 125-132, 2010.

62 KARPLUS, P. A.; DIEDERICHS, K. Linking crystallographic model and data quality. **Science**, v. 336, n. 6084, p. 1030-1033, 2012.

63 EVANS, P. R.; MURSHUDOV, G. N. How good are my data and what is the resolution? **Acta Crystallographica D: biological crystallography**, v. 69, n. 7, p. 1204-1214, 2013.

64 ADAMS, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. **Acta Crystallographica D: biological crystallography**, v. 66, n. 2, p. 213-221, 2010.

65 MCCOY, A. J. et al. Phaser crystallographic software. **Journal of Applied Crystallography**, v. 40, n. 4, p. 658-674, 2007.

66 SHELDRIK, G. A short history of SHELX. **Acta Crystallographica A**, v. 64, n. 1, p. 112-122, 2008.

67 VONRHEIN, C. et al. Automated structure solution with autoSHARP. **Methods in Molecular Biology**, v. 364, p. 215-230, 2007. doi: 10.1385/1-59745-266-1:215

68 AFONINE, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. **Acta Crystallographica D: biological crystallography**, v. 68, n. 4, p. 352-367, 2012.

69 EMSLEY, P. et al. Features and development of Coot. **Acta Crystallographica D: biological crystallography**, v. 66, n. 4, p. 486-501, 2010.

70 CHEN, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallographica D: biological crystallography**, v. 66, n. 1, p. 12-21, 2010.

71 ERICSSON, U. B. et al. Thermofluor-based high-throughput stability optimization of proteins for structural studies. **Analytical Biochemistry**, v. 357, n. 2, p.289-98, 2006.

72 ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403-410, 1990.

73 SÖDING, J.; BIEGERT, A.; LUPAS, A. N. The HHpred interactive server for protein homology detection and structure prediction. **Nucleic Acids Research**, v. 33, n. 2, p. W244-W248, 2005.

74 THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v. 22, n. 22, p. 4673-4680, 1994.

75 PETERSEN, T. N. et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, v. 8, n. 10, p. 785-786, 2011.

76 CAMILO, C. M. et al. HTP-oligodesigner: an online primer design tool for high-throughput gene cloning and site-directed mutagenesis. **Journal of Computational Biology**, v. 23, n. 1, p. 27-29, 2016.

77 WILKINS, M. R. et al. Protein identification and analysis tools in the ExPASy server. **Methods in Molecular Biology**, v. 112, p. 531-52, 1999. doi: 10.1385/1-59259-584-7:531.

78 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. doi:10.1016/j.pep.2014.03.008.

79 CHENG, W. et al. Structural Insights into the substrate specificity of *streptococcus pneumoniae* beta(1,3)-Galactosidase BgaC. **Journal of Biological Chemistry**, v.287, n. 27, p. 22910-22918, 2012.

80 COLLINS, B.; STEVENS, R. C.; PAGE, R. Crystallization optimum solubility screening: using crystallization results to identify the optimal buffer for protein crystal formation. **Acta Crystallographica F: structural biology communications**, v. 61, n. 12, p.1035-1038, 2005.

81 DAUTER, Z. Data-collection strategies. **Acta Crystallographica D: biological crystallography**, v. 55, n. 10, p. 1703-1717, 1999.

82 MATTHEWS, B. W. Solvent content of protein crystals. **Journal of Molecular Biology**, v. 33, n. 2, p. 491-497, 1968.

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

84 VIBORG, A. H. et al. A beta1-6/beta1-3 galactosidase from *Bifidobacterium animalis* subsp. *lactis* B1-04 gives insight into sub-specificities of beta-galactoside catabolism within *Bifidobacterium*. **Molecular Microbiology**, v. 94, n. 5, p. 1024-1040, 2014.

85 COPELAND, R. A. **Enzymes**: a practical introduction to structure, mechanism, and data analysis. New York: John Wiley & Sons, 2004.

86 POKUSAEVA, K.; FITZGERALD, G.; VAN SINDEREN, D. Carbohydrate metabolism in *Bifidobacteria*. **Genes & Nutrition**, v. 6, n. 3, p. 285-306, 2011.

87 PEACOCK, K. S. et al. Isomer-specific consumption of galactooligosaccharides by bifidobacterial species. **Journal of Agricultural and Food Chemistry**, v. 61, n. 51, p. 12612-12619, 2013.

88 KLIJN, A.; MERCENIER, A.; ARIGONI, F. Lessons from the genomes of bifidobacteria. **FEMS Microbiology Reviews**, v. 29, n. 3, p. 491-509, 2005.

89 KLEEREBEZEM, M.; VAUGHAN, E. E. Probiotic and gut lactobacilli and bifidobacteria: molecular approaches to study diversity and activity. **Annual Reviews in Microbiology**, v. 63, p. 269-290, 2009. doi: 10.1146/annurev.micro.091208.073341.

90 GOULAS, T. et al. Comparative analysis of four beta-galactosidases from *Bifidobacterium bifidum* NCIMB41171: purification and biochemical characterisation. **Applied Microbiology Biotechnology**, v. 82, n. 6, p. 1079-1088, 2009.

91 DONG, Y. N. et al. A differentially conserved residue (Ile42) of GH42 beta-galactosidase from *Geobacillus stearothermophilus* BgaB is involved in both catalysis and thermostability. **Journal of Dairy Science**, v. 98, n. 4, p. 2268-2276, 2015.

92 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, 2014.

93 TERWILLIGER, T. C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. **Acta Crystallographica D: biological crystallography**, v. 64, n. 1, p. 61-69, 2008.

94 JANCEWICZ, L. J. et al. Ser-796 of β -galactosidase (*Escherichia coli*) plays a key role in maintaining a balance between the opened and closed conformations of the catalytically important active site loop. **Archives of Biochemistry and Biophysics**, v. 517, n. 2, p. 111-122, 2012.

95 HENZE, M. et al. Rational design of a glycosynthase by the crystal structure of β -galactosidase from *Bacillus circulans* (BgaC) and its use for the synthesis of N-acetyllactosamine type 1 glycan structures. **Journal of Biotechnology**, v. 191, p. 78-85, 2014. doi: 10.1016/j.jbiotec.2014.07.003.

96 OHTO, U. et al. Crystal structure of human beta-galactosidase: structural basis of Gm1 gangliosidosis and morquio B diseases. **Journal of Biological Chemistry**, v. 287, n. 3, p. 1801-1812, 2012.

97 TANAKA, Y. et al. How oligomerization contributes to the thermostability of an archaeon protein: protein l-isoaspartyl-o-methyltransferase from *Sulfolobus tokodaii*. **Journal of Biological Chemistry**, v. 279, n. 31, p. 32957-32967, 2004.

ANEXOS

ANEXO A - Artigos completos publicados em periódicos

1 **GODOY, A. S.**; DE LIMA, M. Z.; CAMILO, C. M.; POLIKARPOV, I. Crystal structure of a putative exo-beta-1,3-galactanase from *Bifidobacterium bifidum* S17. **Acta Crystallographica F: structure biology communications**, v. 72, p. 288-293, 2016.

2 PELLEGRINI, V. A.; SERPA, V.; **GODOY, A. S.**; CAMILO, C.; BERNARDES, A.; REZENDE, C.; PEREIRA JUNIOR, N.; CAIRO, J. F.; SQUINA, F.; POLIKARPOV, I. Recombinant *Trichoderma harzianum* endoglucanase I (Cel7B) is a highly acidic and promiscuous carbohydrate-active enzyme. **Applied Microbiology and Biotechnology**, v. 99, n. 22, p. 9591–9604, 2015

3 EVANGELISTA, D. E.; **GODOY, A. S.**; PAULA, F. F. P.; HENRIQUE-SILVA, F.; POLIKARPOV, I. Expression, purification, crystallization and preliminary X-ray diffraction analysis of the pectin methylesterase from the sugar cane weevil *Sphenophorus levis*. **Acta Crystallographica F: structural biology communications**, v. 70, n. 3, p. 331-334, 2014.

4 BERNARDES, A.; SOUZA, P. C.; MUNIZ, J. R.; RICCI, C. G.; AYERS, S. D.; PAREKH, N. M.; **GODOY, A. S.**; TRIVELLA, D. B.; REINACH, P.; WEBB, P., SKAF, M. S.; POLIKARPOV, I. Molecular mechanism of peroxisome proliferator-activated receptor alpha activation by WY14643: a new mode of ligand recognition and receptor stabilization. **Journal of Molecular Biology**. v. 425, n. 16, p. 2878-2893, 2013.