



UNIVERSIDADE FEDERAL DO CEARÁ

CENTRO DE CIÊNCIAS

DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR

BACHARELADO EM BIOTECNOLOGIA

**GABRIELA MESQUITA LOPES DE
LIMA**

**ENZIMAS PARA DESCONSTRUÇÃO DA QUITINA CODIFICADAS NO GENOMA
DE *Chromobacterium violaceum*: CARACTERIZAÇÃO *in silico* E VALIDAÇÃO
EXPERIMENTAL DO SEU POTENCIAL QUITINOCLÁSTICO**

FORTALEZA

2021

GABRIELA MESQUITA LOPES DE LIMA

ENZIMAS PARA DESCONSTRUÇÃO DA QUITINA CODIFICADAS NO GENOMA
DE *Chromobacterium violaceum*: VALIDAÇÃO EXPERIMENTAL, AVALIAÇÃO DO
POTENCIAL QUITINOCLÁSTICO *in silico*

Monografia apresentada ao Curso de Biotecnologia do Departamento de Bioquímica e Biologia Molecular da Universidade Federal do Ceará, como requisito parcial para obtenção do Título de Bacharel em Biotecnologia.

Orientador: Prof. Dr. Thalles Barbosa Grangeiro.

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

L698e Lima, Gabriela Mesquita Lopes de.

Enzimas para desconstrução da quitina codificadas no genoma de *Chromobacterium violaceum* : validação experimental, avaliação do potencial quitinoclástico *in silico* / Gabriela Mesquita Lopes de Lima. – 2021.
102 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Ciências, Curso de Biotecnologia, Fortaleza, 2021.

Orientação: Prof. Dr. Thalles Barbosa Grangeiro.

1. *Chromobacterium violaceum*. 2. *Escherichia coli*. 3. Quitina. 4. Quitinases. 5. Proteínas recombinantes. I. Título.

CDD 661

GABRIELA MESQUITA LOPES DE LIMA

ENZIMAS PARA DESCONSTRUÇÃO DA QUITINA CODIFICADAS NO GENOMA
DE *Chromobacterium violaceum*: VALIDAÇÃO EXPERIMENTAL, AVALIAÇÃO DO
POTENCIAL QUITINOCLÁSTICO *in silico*

Monografia apresentada ao Curso de
Biotecnologia da Universidade Federal do
Ceará, como requisito parcial para obtenção do
título de Bacharel em Biotecnologia

Aprovada em: ____/____/_____.

BANCA EXAMINADORA

Prof. Dr. Thalles Barbosa Grangeiro (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dra. Suelen Carneiro de Medeiros
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

Prof. Dr. Eduardo Henrique Silva de Sousa
Universidade Federal do Ceará (UFC)

A Deus.

Aos meus familiares e amigos pelo apoio e
pelo suporte.

AGRADECIMENTOS

Ao CNPq, pelo apoio financeiro.

Ao Prof. Dr. Thalles Barbosa Grangeiro, pela excelente orientação e pela paciência e disponibilidade para tirar dúvidas.

Aos participantes da Banca Examinadora, pelo tempo concedido para avaliar este trabalho e pelas colaborações.

Aos colegas do Laboratório de Genética Molecular, principalmente, Simone, Edvar e Mayara, pelo suporte e pelas sugestões recebidas.

Aos meus colegas de graduação, especialmente, Laís, Bruno e Natália, por me darem suporte em todos os momentos.

Aos professores do Curso de Bacharelado em Biotecnologia, por todo conhecimento utilizado para tornar meu aprendizado enriquecedor.

À minha família, especialmente, minha mãe (Klecia) e minha avó (Aldeni), por sempre torcerem pelo meu sucesso.

RESUMO

Enzimas capazes de degradar quitina, um polissacarídeo abundante na natureza, estão presentes em muitos microrganismos e têm potencial para serem utilizadas no desenvolvimento de produtos inovadores para a agricultura, medicina e indústria. O objetivo deste trabalho, portanto, foi transformar células de *Escherichia coli* BL21(DE3), com plasmídeos codificantes para certas proteínas de *Chromobacterium violaceum* ATCC 12472, e verificar se essas são produzidas tanto em escala reduzida, quanto em escala maior e se possuem atividade enzimática na degradação da quitina. Foram realizadas análises *in silico*, a fim de avaliar a probabilidade de diferentes enzimas quitinoclásticas serem produzidas de forma solúvel, quando expressas em *E. coli*. Após essas análises, algumas proteínas foram escolhidas, representando membros das famílias AA10 (3 proteínas), GH3 (2 proteínas) e GH23 (3 proteínas). Plasmídeos contendo as sequências de DNA que codificam essas proteínas, foram, então, introduzidos em *E. coli* BL21(DE3) e a expressão das moléculas recombinantes foi realizada, de forma piloto, em diferentes temperaturas (20 °C, 25 °C, 30 °C e 35 °C). A análise por eletroforese (SDS-PAGE) das frações intracelulares de células induzidas, revelou, em todas as amostras, a presença de bandas com massas moleculares aparentes que correspondiam aos valores esperados para as massas das proteínas. foram realizados, para todas as proteínas estudadas, testes de atividades quitinoclástica e quitosanásica, seguindo protocolos bem estabelecidos anteriormente. Apenas duas proteínas, ambas pertencentes à família GH3 (CV0259 e CV2065), obtiveram resultados para atividade quitinoclástica e nenhuma obteve resultado para atividade quitosanásica. A proteína com melhor resultado na indução piloto (CV2065) foi produzida em maior quantidade e purificada por cromatografia de afinidade por íons metálicos, utilizando uma resina de cobalto. Por fim, ambas as proteínas passaram por análises de bioinformática, a fim de entender seus domínios proteicos e gerar uma estrutura 3D das mesmas utilizando modelagem por homologia. Assim, a produção, purificação e análise de proteínas quitinoclásticas foram bem-sucedidas, o que abre precedentes para a produção e purificação, em larga escala, de proteínas recombinantes com diversos potenciais econômicos e industriais.

Palavras-chave: *Chromobacterium violaceum*, *Escherichia coli*, quitina, quitinases, proteínas recombinantes.

ABSTRACT

Enzymes capable of degrading chitin, a recalcitrant polysaccharide abundant in nature, are present in many microorganisms and have the potential to be used in different applications in agriculture, medicine and industry. Therefore, the main objective of this study was to transform *Escherichia coli* BL21 (DE3) cells with plasmids coding for proteins from *Chromobacterium violaceum* ATCC 12472, and verify if these proteins are produced both on a small and large scale and if they have enzymatic activity in chitin degradation. *In silico* analyzes were carried out in order to assess if different chitinolytic enzymes would be produced in a soluble way when expressed in *E. coli*. After that, some proteins from the families AA10 (3 proteins), GH3 (2 proteins) and GH23 (3 proteins) were chosen. Plasmids carrying the DNA sequences encoding these proteins were then introduced into *E. coli* BL21 (DE3) cells and the expression of the recombinant molecules was carried out, in a small scale, at different temperatures (20 ° C, 25 ° C, 30 ° C and 35 ° C). An electrophoresis analysis (SDS-PAGE) of the lysed cells revealed the presence of bands with an apparent molecular mass that correspond to the expected values of the protein mass. After verifying the proteins production, tests for enzymatic activity were performed for all the studied proteins. Only two of them, both members of the GH3 family (CV0259 and CV2065), showed chitinolytic activity while none of the studied proteins tested positive for chitosan degradation. The protein with the best yield on previous tests (CV2065) was produced in a slightly large scale and purified by metal affinity chromatography using a cobalt resin. Finally, both proteins were analyzed using bioinformatic tools in order to understand their protein domains and generate a 3D structure of them by homology-modeling. Overall, the production, purification and analysis of chitinolytic proteins have been successful, which sets precedents for a large-scale production and purification of recombinant proteins with promising industrial applications.

Keyword: *Chromobacterium violaceum*, *Escherichia coli*, chitin, chitinases, recombinant proteins.

LISTA DE FIGURAS

Figura 1 –	Estrutura química da quitina	20
Figura 2 –	Estrutura química da quitosana	21
Figura 3 –	Estrutura química da N-acetilglicosamina	22
Figura 4 –	Aplicações industriais de quitinases	28
Figura 5 –	Eletroforese em gel de agarose (0,8%).	55
Figura 6 –	Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 20 °C (fração solúvel).	58
Figura 7 –	Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 20 °C (fração insolúvel)	59
Figura 8 –	Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 25 °C (fração solúvel)	60
Figura 9 –	Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 25 °C (fração insolúvel)	61
Figura 10 –	Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 30 °C (fração solúvel)	62
Figura 11 –	Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 30 °C (fração insolúvel)	63
Figura 12 –	Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 35 °C (fração solúvel)	64
Figura 13 –	Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 35 °C (fração insolúvel)	65
Figura 14 –	Resultado do teste de atividade quitinoclástica de CV2065 e CV0259 contra quitina coloidal	69
Figura 15 –	Resultado do teste de atividade quitinoclástica de CV2065 e CV0259 contra quitina insolúvel	70
Figura 16 –	Perfil da cromatografia de afinidade por íons metálicos utilizando uma resina de cobalto, HisPur™	72
Figura 17 –	Domínios proteicos, obtidos a partir de análise no CDD para a proteína CV0259, da família GH3	74
Figura 18 –	Domínios proteicos, obtidos a partir de análise no CDD para a proteína CV2065, da família GH3	75
Figura 19 –	Representação gráfica da arquitetura de domínios a proteína CV0259 utilizando o programa IBS	76
Figura 20 –	Representação gráfica da arquitetura de domínios a proteína CV2065 utilizando o programa IBS	77
Figura 21 –	Resultados obtidos para a proteína CV0259 após a submissão da mesma ao servidor SWISS-MODEL	81
Figura 22 –	Resultados obtidos para a proteína CV2065 após a submissão da mesma ao servidor SWISS-MODEL	82
Figura 23 –	Plot de Ramachandran para a proteína CV0259, obtido a partir do servidor MolProbity	84

Figura 24 –	Plot de Ramachandran para a proteína CV2065, obtido a partir do servidor MolProbity.	85
Figura 25 –	Sobreposição da estrutura 3D de CV0259 com seu template (4ZM6)	89
Figura 26 –	Sobreposição da estrutura 3D de CV2065 com seu template (6JTI)	90

LISTA DE TABELAS

Tabela 1 –	Áreas de aplicação e usos potenciais de quitina e seus derivados.	27
Tabela 2 –	Identificação das proteínas recombinantes de fusão, expressas a partir dos insertos clonados em pET32a(+), codificando hidrolases de glicosídeos de <i>C. violaceum</i> ATCC 12472	51
Tabela 2 –	Famílias e massas moleculares teóricas (monoisotópicas) das proteínas recombinantes de fusão, expressas a partir dos insertos clonados em pET32a(+), codificando hidrolases de glicosídeos de <i>C. violaceum</i> ATCC 12472	52
Tabela 4 –	Solubilidades teóricas das proteínas recombinantes de fusão, expressas a partir dos insertos clonados em pET32a(+), codificando hidrolases de glicosídeos de <i>C. violaceum</i> ATCC 12472	53
Tabela 5 –	Temperaturas de indução adequadas para a produção heteróloga, em células de <i>E. coli</i> BL21(DE3), de proteínas recombinantes de <i>C. violaceum</i> ATCC 12472	67
Tabela 6 –	Validação dos modelos 3D para a proteína CV0259 por diferentes servidores	86
Tabela 7 –	Validação dos modelos 3D para a proteína CV2065 por diferentes servidores	87

LISTA DE ABREVIATURAS E SIGLAS

3D	Tridimensional
AA	<i>Auxiliary Activity</i>
ATCC	<i>American Type Culture Collection</i>
BLAST	<i>Basic Local Alignment Search Tool</i>
BLASTP	<i>Protein Blast</i>
CAZy	<i>Carbohydrate-Active Enzymes</i>
CDD	<i>Conserved Domains Database</i>
GH	<i>Glycoside Hydrolase</i>
GlcNAc	N-Acetyl-Glicosamina
IBS	<i>Illustrator for Biological Sequences</i>
IPTG	Isopropil β -D-1-tiogalactopiranósideo
LB	Caldo Luria Berthani
LPMO	Lytic polysaccharide monooxygenases
NCBI	<i>National Center for Biotechnology Information</i>
ORF	<i>Open Reading Frame</i>
pI	Ponto Isoelétrico
ProSA-web	<i>Protein Structure Analysis</i>
RMSD	<i>Root-mean-square-deviation</i>
SDS	<i>Sodium Dodecyl Sulfate</i> (Dodecil Sulfato de Sódio)
SDS-PAGE	Eletroforese de proteínas em gel de poliacrilamida sob condições desnaturantes (acrúscimo de SDS)
SOC	<i>Super Optimal broth with Catabolite Repression</i>
UFC	Universidade Federal do Ceará

SUMÁRIO

1	INTRODUÇÃO	16
2	REVISÃO DE LITERATURA.....	19
2.1	Quitina.....	19
2.1.1	<i>Quitosana.....</i>	20
2.1.2	<i>N-Acetyl-D-Glicosamina.....</i>	21
2.2	Enzimas quitinoclásticas.....	23
2.2.1	<i>Aplicações de enzimas quitinoclásticas e potenciais usos de quitina e seus derivados.....</i>	25
2.3	<i>C. violaceum</i> como fonte de enzimas para degradação da quitina.....	29
2.4	Produção de proteínas por sistemas heterólogos.....	30
2.5	Bioinformática para o estudo de proteínas.....	31
2.5.1	<i>Estudo de domínios proteicos.....</i>	32
2.5.2	<i>Modelagem de proteínas.....</i>	33
3	PERGUNTAS DE PARTIDA.....	36
4	JUSTIFICATIVA.....	37
5	OBJETIVOS.....	38
5.1	Objetivo Geral.....	38
5.2	Objetivos Específicos.....	38
6	MATERIAIS E MÉTODOS.....	39
6.1	Análise <i>in silico</i> de proteínas e construção de plasmídeos sintéticos	39
6.2	Transformação de células para a produção de proteínas recombinantes de <i>C. violaceum</i>.....	39

	14	
6.2.1	<i>Transformação de E. coli DH5α por choque térmico</i>	39
6.2.2	<i>Extração de DNA Plasmidial</i>	40
6.2.3	<i>Digestão de DNA Plasmidial</i>	41
6.3.4	<i>Eletroforese de DNA em gel de agarose</i>	41
6.2.5	<i>Transformação de E. coli BL21(DE3) por choque térmico</i>	42
6.3	Produção de proteínas recombinantes	43
6.3.1	<i>Triagem de temperatura de indução da expressão de proteínas recombinantes em E. coli BL21(DE3)</i>	43
6.3.2	<i>Lise celular</i>	43
6.3.3	<i>Eletroforese de proteínas em gel de poliacrilamida sob condições desnaturantes</i>	44
6.4	Teste para atividade quitinoclástica das proteínas recombinantes	45
6.4.1	Teste para atividade quitinoclástica utilizando quitina coloidal	45
6.4.2	Teste para atividade quitinoclástica utilizando quitina insolúvel	45
6.5	Teste para atividade quitosanásica das proteínas recombinantes	46
6.6	Produção de uma proteína recombinante em E. coli em escala maior	47
6.6.1	<i>Indução de proteína recombinante</i>	47
6.6.2	<i>Purificação de proteína recombinante</i>	47
6.7	Ferramentas de bioinformática para estudo de proteínas	48
6.7.1	<i>Análise dos domínios das proteínas estudadas</i>	48
6.7.2	<i>Predição da estrutura 3D de proteínas</i>	48
6.7.2.1	<i>Escolha de modelos e modelagem por homologia</i>	48
6.7.2.1	<i>Avaliação dos modelos</i>	49
7	RESULTADOS E DISCUSSÃO	50
7.1	Seleção de proteínas a serem expressas em E. coli	50
7.2	Transformação de células de E. coli com insertos sintéticos	54

7.3	Indução piloto de proteínas recombinantes.....	57
7.4	Teste para atividades quitinoclástica e quitosanásica.....	68
7.5	Produção e purificação de uma proteína recombinante.....	71
7.6	Avaliação dos domínios das proteínas com atividade enzimática.....	73
7.7	Construção e avaliação de modelos 3D para proteínas com atividade enzimática.....	80
8	CONCLUSÃO.....	92
	REFERÊNCIAS.....	93
	ANEXO A - SEQUÊNCIAS DAS PROTEÍNAS ESTUDADAS.....	102
	ANEXO B - MAPA PLASMIDIAL.....	104

1 INTRODUÇÃO

A quitina é um polímero insolúvel em água, linear e não ramificado de N-acetil-beta-D-glucosamina, com os resíduos de GlcNAc unidos por ligações beta-(1,4) O-glicosídicas. Depois da celulose, é o polissacarídeo mais abundante na natureza, estando presente em diferentes seres vivos (exoesqueleto de insetos, carapaça de crustáceos, ovos de nematoídes e conchas de moluscos) (MUZZARELLI, 1999). Quitina e seus derivados, como quitosana e GlcNAc, apresentam aplicações na indústria alimentícia (aumento de tempo de prateleira de produtos e conservação de alimentos), na indústria farmacêutica (desenvolvimento de sistemas de entrega de fármacos) e na agricultura, a fim de impedir crescimento de microrganismos (BARIKANI *et al.*, 2014; DASH *et al.*, 2011; HAMED *et. al.*, 2016; SETHULEKSHMI, 2014; XIA *et al.*, 2011; ZENG; MEI; WU, 2010). A quitina é difícil de ser purificada e modificada quimicamente e, desse modo, há a necessidade de identificar enzimas que facilitem a utilização desse polissacarídeo (HOWARD, 2003; MUZZARELLI, 1999).

A degradação e modificação de quitina e seus derivados, especialmente, quitosana, podem ser realizadas por organismos vivos, como bactérias e fungos, por meio de enzimas específicas (HOWARD, 2003). A biodegradação de quitina é realizada por hidrolases (quitinases e quitobiases) e por enzimas oxidativas as quais são chamadas, de modo geral, de enzimas quitinoclásticas. Quitinases (E.C. 3.2.1.14) são enzimas que hidrolisam ligações covalentes β -1,4 entre os resíduos de GlcNAc das cadeias que formam a quitina e a quitodextrina (LOBO, 2013). Monooxigenases líticas de polissacarídeos (LPMOs; EC 1.14.99.53) constituem outra classe de enzimas, denominada de Atividades Auxiliares ou AAs (VAAJE-KOLSTAD *et al.*, 2010). Essas enzimas possuem aplicação em diversas áreas como na medicina, na agricultura e na indústria (AAM, 2010; CHEN, 2010; LOBO, 2013; VAN LOON, 2006).

A clivagem de quitina por quitinases gera multímetros solúveis de N-acetilglicosamina, ou GlcNAc, como quitotetraose, quitotriose e quitobiose (BHATTACHARYA, 2008). Comercialmente, o GlcNAc possui algumas aplicações que justificam a pesquisa de enzimas que produzam esse composto. Na área médica, N-acetilglicosamina pode ser utilizada para tratar pacientes que possuem danos nas cartilagens dos ossos e problemas como artrose e artrite reumatoide, por exemplo (KAMEL, 1992). Quitosana, obtida pela desacetilação alcalina da quitina, também possui diversas aplicações em diferentes indústrias. Pode ser utilizada na inibição de proteases, na preparação de

hidrogéis, na engenharia de tecidos, na produção de alimentos (conservativo e emulsificante), entre outros (SANDEEP *et al.*, 2013). Além de aplicações utilizando os produtos da quitina, também existem aquelas que utilizam diretamente quitinases. Podem, por exemplo, ser utilizadas na degradação de resíduos de quitina, gerados pela indústria de crustáceos, e no impedimento do crescimento de insetos e larvas, os quais possuem quitina na sua constituição e causam danos extensos em lavouras (NAMPOOTHIRI 2004; STOYKOV, 2014; VU, 2009). Enzimas quitinoclásticas podem ser obtidas em diferentes organismos, tais como plantas, fungos e, principalmente, bactérias.

Bactérias são fontes de enzimas relevantes que podem ser produzidas de forma recombinante, ou seja, que podem ser produzidas pela clonagem de genes em vetores de expressão (CLARK, 2016). Entre esses microrganismos, existe a *Chromobacterium violaceum*, organismo saprófito e de vida livre, a qual possui sua estirpe ATCC 12472 totalmente analisada e sequenciada (CALDAS, 1978; CALDAS, 1990; VASCONCELOS *et al.*, 2003). Esse genoma possui vários genes codificando prováveis quitinases e outras enzimas quitinoclásticas, pertencentes a diferentes famílias de hidrolases de glicosídeos (GHs) e atividades auxiliares (AAs), em especial GH23, GH3 e AA10. No entanto, apenas poucas proteínas desse organismo foram produzidas e purificadas. Por exemplo, a proteína CV2935, da família GH18, foi produzida de maneira recombinante em células de *Escherichia coli* BL21(DE3) e avaliada para a presença de atividade quitinolítica (LOBO, 2013). Isso justifica a realização de estudos que envolvam a expressão heteróloga em *E. coli* de proteínas quitinoclásticas de *C. violaceum*, visto que esse microrganismo pode ser fonte de enzimas bastante relevantes economicamente.

Bioinformática faz referência a uma ciência multidisciplinar que utiliza o emprego de ferramentas computacionais, a fim de estudar problemas e questões biológicas, com aplicações em diferentes áreas (VERLI, 2014). O estudo de proteínas utilizando essas ferramentas permite a descoberta de informações sobre, por exemplo, a estrutura tridimensional e funções principais dessas. Estruturas de proteínas podem ser deduzidas por meio de modelagem por homologia. Nesse caso, a proteína (alvo) terá sua estrutura 3D utilizando a estrutura tridimensional de outra proteína similar, através de alinhamentos da sequência alvo com sequências presentes em bancos de dados (VERLI, 2014). Domínios são unidades funcionais de uma proteína, sendo responsáveis pelas funções e interações da proteína com outras estruturas como, por exemplo, substratos de enzimas (BAGOWSKI *et al.*, 2010). Desse modo, domínios e a estrutura 3D de uma proteína são importantes para entender a dinâmica de proteínas, o que justifica estudo destes por, principalmente, ferramentas de bioinformática.

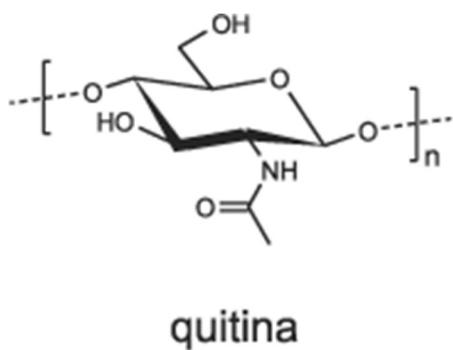
Como dito anteriormente, quitinases e os produtos derivados de quitina e quitosana possuem diversas aplicações em áreas economicamente estratégicas e, desse modo, é importante que existam estudos envolvendo a purificação e a caracterização dessas enzimas, bem como sua produção em larga escala. A indústria de produção de enzimas atrai, em particular, biotecnologistas, visto que há a necessidade de desenvolver processos enzimáticos mais eficazes e da descoberta de novas enzimas, as quais podem ser aplicadas em processos diversos (DAHIYA, 2006). Isso justifica a realização de estudos que envolvam a expressão heteróloga em *E. coli* de proteínas de *C. violaceum*, visto que esse microrganismo pode ser fonte de enzimas bastante relevantes economicamente.

2 REVISÃO DE LITERATURA

2.1 Quitina

A quitina (Figura 1) é um polímero insolúvel em água, linear e não ramificado de N-acetil-beta-D-glicosamina, com os resíduos de GlcNAc unidos por ligações beta-(1,4) O-glicosídicas (MUZZARELLI, 1999). Depois da celulose, a quitina é o polissacarídeo mais abundante na natureza e, portanto, esses polímeros estão presentes em diferentes seres vivos. Além de possuir essas características, a quitina existe em três formas polimórficas, chamadas de α -quitina, β -quitina e γ -quitina. A primeira é a forma mais presente na natureza, sendo composta por folhas com cadeias paralelas e antiparalelas alternadas. Assim α -quitina faz parte, principalmente, da constituição do exoesqueleto de insetos, da carapaça de crustáceos, de ovos de nematoides; além de estar presente em conchas de moluscos e na parede celular de fungos como ascomicetos, zigomicetos, basidiomicetos e deuteromicetos (AAM *et al.*, 2010; GARDNER *et al.*, MUZZARELLI, 1999). A β -quitina, formada por folhas paralelas, e a γ -quitina, na qual cada terceira folha tem a direção oposta às duas folhas anteriores, são encontradas mais raramente na natureza em, principalmente, estruturas mais flexíveis como a sépia de lulas (KAYA *et al.*, 2017; SINGH; SHITIZ; SINGH, 2017). Esse polissacarídeo também está presente em algumas algas diatomáceas (Bacillariophyta) e nas espécies *Thalassiosira fluviatilis* e *Cylotella cryptica* a quitina se apresenta constituída por unidades de 2-acetamido-2-desoxi-D-glicopiranose (CAMPANA-FILHO *et al.*, 2007).

Figura 1 – Estrutura química da quitina.



Fonte: Silva, 2006

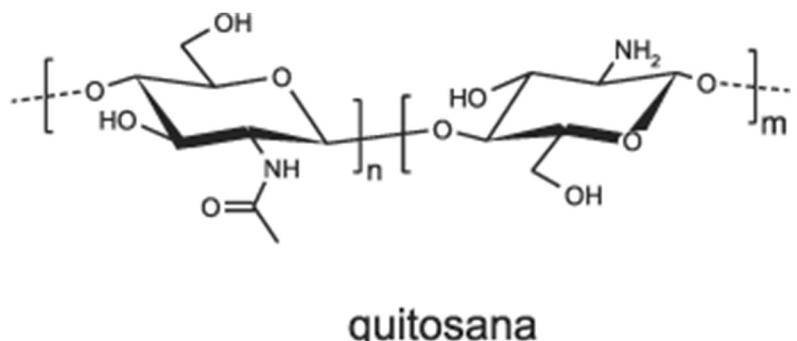
Quitina foi isolada pela primeira vez por Braconnot, em 1811, a partir de uma fração extraída de fungos superiores. No entanto, apesar de ter sido isolada antes da celulose, não existiam estudos mais aprofundados sobre as propriedades e aplicações industriais do polissacarídeo, visto que não havia conhecimento básico sobre suas propriedades e reatividade química. Foi a partir da década de 1970, com os livros escritos por Muzzarelli, que as propriedades desse polímero puderam ser estudadas extensivamente (ROBERTS, 1992).

Apesar de relevante em diversas áreas industriais, a quitina, por ser um biopolímero, é difícil de ser purificada e modificada quimicamente (MUZZARELLI, 1999). Dito isso, é necessária a identificação de enzimas que modifiquem ou degradem esse polissacarídeo, a fim de facilitar atividades que utilizam produtos específicos derivados desse biopolímero e entender as propriedades químicas da quitina (HOWARD, 2003).

2.1.1 *Quitosana*

Quitosana (Figura 2) é um polímero formado, principalmente, por unidades de D-glicosamina o qual está presente em alguns fungos como *Mucor* e *Zygomycetes* (SILVA, 2006). O polímero pode ser obtido a partir da desacetilação da quitina, através de tratamento por álcalis, especialmente solução aquosa de hidróxido de sódio (NaOH). Quitosana possui unidades de 2-amino-2-deoxiglicose, o que leva à formação de sais em solução e de polieletrolitos. Além disso, pode haver também a formação de filmes a partir de unidades de quitosana, característica que a distingui de alguns outros polissacarídeos, como a celulose (KUMAR, 2004).

Figura 2 – Estrutura química da quitosana.



Fonte: Silva, 2006

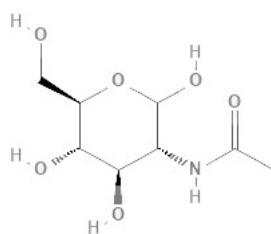
Assim como a quitina, a maior parte da quitosana disponível comercialmente é extraída de cascas de caranguejo e de camarões, o que pode tornar o processamento de resíduos da indústria pesqueira e alimentícia rentável (KUMAR, 2000). A quitosana é conhecida por ser segura e não-tóxica para seres humanos e, desse modo, seu uso é cada vez mais expressivo em indústrias alimentícias e farmacêuticas, por exemplo. No entanto, o polímero é insolúvel em soluções aquosas neutras, meio no qual a maior parte das enzimas fisiológicas possuem atividade. Desse modo, derivados hidrossolúveis de quitosana são necessários, a fim de expandir o uso da mesma (SILVA, 2006). Em diversos estudos, derivados hidrossolúveis de quitosana foram obtidos a partir de versões do polímero com baixo grau de acetilação ou a partir da incorporação de ácido maléico à hidroxipropilquitosana e carboximetilquitosana (XIE, 2001; DUNG, 1993).

Devido ao teor de nitrogênio em sua composição (6,89%), a quitosana pode ser utilizada como um agente quelante (KUMAR, 2000). Atividade antimicrobiana, causada por esses grupamentos amínicos, é uma das suas características. Isso permite que o polissacarídeo iniba a atividade de microrganismos como *E. coli* e *Pseudomonas aeruginosa*, por exemplo. Além disso, possui propriedades analgésicas e anticoagulantes e permite a aceleração da cicatrização (SILVA, 2006). Todas essas características podem fazer com que o uso industrial da quitosana e, por conseguinte, de enzimas que permitem sua transformação seja viável economicamente.

2.1.2 *N*-acetil-D-Glicosamina

N-acetil-D-Glicosamina (Figura 3), também conhecida como *N*-acetilglicosamina (GlcNac), 2-acetamino-2-deoxi- β -D-glicose ou 2-(acetilamino)-2-deoxi-D-glicose, é um monossacarídeo derivado da glicose. O peso molecular desse monossacarídeo é 221.21 g/mol, fórmula molecular igual a C8H15NO6 e solubilidade em água de 25%. A polimerização do GlcNac, por meio de ligações glicosídicas β -1,4, dá origem à quitina. (CHEN *et al.*, 2010).

Figura 3 – Estrutura química da N-acetilglicosamina.



Fonte: <https://pubchem.ncbi.nlm.nih.gov/compound/N-Acetyl-D-Glucosamine>.

Além de estar presente na quitina, GlcNac também está presente na constituição de outros polissacarídeos e, portanto, é amplamente distribuída na natureza. A mureína (peptidoglicano), constituída por resíduos de ácido murâmico e N-acetilglicosamina, é a principal constituinte da parede celular de bactérias (ASHRY *et al.*, 2007). Como outro exemplo de polissacarídeo pode-se destacar o ácido hialurônico (hialuronano), biopolímero formado por ligações glicosídicas entre ácido glucurônico e GlcNac. O ácido hialurônico está presente na constituição da matriz extracelular, sendo responsável pela viscoelasticidade de cartilagens e do líquido sinovial, além de funcionar, para bactérias patogênicas, como mecanismo de defesa contra respostas imunológicas de hospedeiros (ASHRY *et al.*, 2007).

Como foi dito anteriormente, a N-acetilglicosamina é um monossacarídeo amplamente distribuído na natureza. NOPE1 é um transportador de GlcNac presente em milho e arroz. Esse transportador é necessário para iniciar a simbiose, termo utilizado para definir uma relação entre duas espécies que resulta em vantagens mútuas, entre micorriza arbuscular (AM ou *Arbuscular mycorrhiza*) e as duas plantas citadas previamente (NADAL *et al.*, 2017). N-acetilglicosamina também está envolvida na produção de biofilmes (acúmulo de bactérias em superfícies sólidas) de *Staphylococcus aureus*. A partir do uso de cepas de *S. aureus* que não continham um operon contendo genes para a produção de poli-n-acetilglicosamina (PNAG), pode-se constatar que o polissacarídeo, formado por resíduos de GlcNac, contribui, tanto para a aderência das células bacterianas, quanto para o estabelecimento de infecções (LIN *et al.*, 2015). N-acetilglicosamina também está presente em uma série de hormônios existentes em mamíferos. Os hormônios folículo estimulante (FSH), luteinizante (LH) e o estimulador da tireoide (TSH) são alguns exemplos (CHEN *et al.* 2010).

Apesar de estar presente na constituição de diversos polissacarídeos, a maior fonte de N-acetilglicosamina continua sendo a quitina, derivada da biomassa, tanto dos exoesqueletos de camarão e de caranguejo, quanto de fungos como *Actinomucor taiwanensisfungus* (CHEN

et al., 2001; CHEN *et al.*, 2010). A produção desse monossacarídeo acontece, principalmente, por hidrolise ácida da quitina, um método que além de ser de alto custo, possui um baixo rendimento e uma grande quantidade de resíduos de ácido clorídrico (HCl) (SASHIWA *et al.*, 2003). Desse modo, a utilização de hidrólise enzimática da biomassa de quitina é um método mais adequado à produção de N-acetilglicosamina. Foi descrita por Sashiwa *et al.* (2002), a produção de GlcNac a partir de α -quitina por um extrato de enzimas obtido de *Aeromonas hydrophila* H2330. O estudo obteve bons resultados na purificação do monossacarídeo, porém com uma baixa produção do mesmo. Em Das *et al.* (2018), por outro lado, foi verificada a produção de GlcNac com até 92% de rendimento, resultado obtido a partir da hidrólise enzimática de quitina utilizando quitinases isoladas de *Aspergillus terreus*. Percebe-se, a partir desses resultados, que enzimas quitinoclásticas são necessárias para a digestão completa da quitina, o que leva a rendimentos maiores de GlcNac.

2.2 Enzimas quitinoclásticas

A degradação e modificação de quitina e seus derivados podem ser realizadas por organismos vivos, como bactérias e fungos, por meio de enzimas específicas. Estas enzimas estão envolvidas no transporte, despolimerização e metabolismo de quitina e de quitoligossacarídeos (HOWARD, 2003) e, portanto, são necessárias para diferentes atividades nos organismos nos quais estão presentes. Por exemplo, podem ser utilizadas para remodelar o exoesqueleto de insetos durante o crescimento e desenvolvimento, para digerir, sob a forma de metabólitos que podem ser absorvidos e digeridos, outros organismos que possuem quitina em sua composição e para degradar microrganismos patogênicos com quitina constituindo parede celular (RATHORE, 2015). Bactérias são consideradas os organismos mais importantes na biodegradação de quitina, especialmente no solo, onde a taxa de hidrólise de quitina é relacionada à presença de bactérias, e em ambientes marinhos, já que, mesmo sendo ubíqua, quase nenhuma quitina pode ser encontrada em sedimentos marinhos (BEIER, 2013; HOWARD, 2003; KIELAK *et al.*, 2013). Além disso, algumas diatomáceas e algumas colônias de fungos demonstraram também possuir capacidade de degradar quitina presente em carapaças de outros organismos (WURZBACHER *et al.*, 2010).

A biodegradação de quitina é realizada por hidrolases (quitinases e quitobiases) e por enzimas oxidativas (mono-oxigenases líticas de polissacarídeos), as quais são chamadas, de modo geral, de enzimas quitinoclásticas. Desse modo, são necessárias múltiplas enzimas para

a degradação desse polímero em quitoligossacarídeos, N-acetil glucosamina (GlcNAc) e quitobiose (dímero de unidades de glucosamina ligadas por ligações β -1,4) (HOWARD, 2003).

Hidrolases de glicosídeos (GH, glicosidases ou carboidrases, EC 3.2.1) compõem um vasto grupo de enzimas capazes de degradar ligações *O*-glicosídicas entre dois ou mais açúcares ou entre um açúcar e outro composto de natureza distinta. Genes para codificar GHs são encontrados na maioria dos organismos, especialmente, bactérias, e a quantidade desses genes depende do nicho ecológico em que esses microrganismos estão inseridos (NAUMOFF, 2011). Atualmente, existem 153 famílias de GHs (GH1-GH153), constituindo aproximadamente 47% das enzimas classificadas no Carbohydrate-Active Enzyme (CAZy), nas quais o domínio catalítico está associado a um ou mais domínios de ligação a carboidrato ou módulos de ligação a carboidrato (carbohydrate-binding modules, CBMs), que reconhecem e se ligam ao substrato, mas não possuem atividade catalítica sobre o mesmo (BORASTON *et al.*, 2004; CANTAREL, 2008).

Quitinases (EC 3.2.1.14) são enzimas que hidrolisam ligações covalentes β -1,4 entre os resíduos de GlcNAc das cadeias que formam a quitina e a quitodextrina (Lobo, 2013). Quitinases podem ser classificadas em endoquitinases e exoquitinases. As primeiras clivam quitina em locais aleatórios gerando oligômeros de N-acetilglicosamina enquanto exoquitinases liberam N-acetilglicosamina a partir da extremidade não redutora da quitina. Exoquitinases podem ser divididas em quitobiases (permite a conversão de quitobiose em monômeros de N-acetilglicosamina) e β -1,4- N-acetilglicosaminidases (permite a conversão multímeros, como di-acetilquitobiose, quitotriose e quitotetraose em monômeros de N-acetilglicosamina) (ZHOU *et al.*, 2019). As quitinases de estruturas conhecidas estão presentes nas famílias 18 e 19 das glicosil hidrolases, apesar de quitinases pertencentes às famílias GH23 e GH48 já terem sido descritas (ARIMORI *et al.*, 2013; ADRANGI, 2013; FUJITA *et al.*, 2006). Quitinases GH18 estão presentes em vírus, bactérias, fungos, plantas e diversos grupos de insetos e mamíferos e as GH19 estão presentes em plantas superiores, nematóides, vírus e algumas bactérias (Funkhouser, 2007; Henrissat, 1991).

Além dessas enzimas, mono-oxigenases líticas de polissacarídeos (LPMOs; EC 1.14.99.53) constituem outra classe de enzimas capazes de degradar quitina e outros polissacarídeos, tais como amido e celulose. Após alguns estudos, essas enzimas foram agrupadas no CAZy em uma nova categoria denominada de Atividades Auxiliares (Auxiliary Activities, AAs), a fim de agrupar as famílias de LPMOs e outras famílias de enzimas oxidativas envolvidas na degradação de polissacarídeos (famílias AA1-AA15). Apesar de a maioria das proteínas presentes nas famílias AA mostrarem maior atividade em conjunto com

hidrolases de glicosídeo, essas proteínas permitem a oxidação de ligações glicosídicas 1,4 em polissacáideos, de modo a permitirem a degradação destes (YANG; FUKAMIZO, 2019). Desse modo, os membros da família AA9 agem diretamente sobre celulose (VAAJE-KOLSTAD *et al.*, 2010), enquanto os membros da família AA11 atuam sobre quitina. Membros da família AA10 podem atuar sobre quitina ou sobre celulose, enquanto que aqueles da família AA15 são bifuncionais, atuando sobre celulose e quitina (VAAJE-KOLSTAD *et al.*, 2005). Das famílias que atuam sobre quitina (AA10, AA11 e AA15), apenas AA10 ocorre em bactérias.

Hidrolases de glicosídeos, especialmente quitinases, têm atraído o interesse de biotecnologistas pelas aplicações dessas enzimas em medicina, agricultura e indústria (Lobo, 2013). Por exemplo, algumas quitinases possuem propriedades anti-fúngicas e alguns N-acetil quitoligossacáideos (propriedades antibacterianas) e GlcNAc, produzidos pela degradação da quitina, podem ser utilizados na indústria farmacêutica e médica (AAM, 2010; CHEN, 2010; VAN LOON, 2006).

2.2.1 Aplicações de enzimas quitinolíticas e potenciais usos de quitina e seus derivados

A indústria de produção de enzimas é um campo bastante extenso e, desse modo, existem diversas áreas nas quais enzimas de diferentes organismos podem ser aplicadas. Essa área atrai, em particular, biotecnologistas, visto que há a necessidade de desenvolver processos enzimáticos mais eficazes e da descoberta de novas enzimas, as quais podem ser aplicadas em processos diversos (DAHIYA, 2006). Nesse contexto, enzimas que participam da degradação da quitina, em especial hidrolases de glicosídeos (quitinases), se apresentam como interessantes do ponto de vista industrial e médico. Desse modo, organismos que são capazes de produzir essas enzimas, tanto de forma natural, como de forma heteróloga, também se tornam alvos de estudos e pesquisas.

A clivagem enzimática de quitina por quitinases ocorre, na natureza, de forma aleatória, podendo promover a clivagem em diferentes áreas do polímero. Esse processo gera multímetros solúveis de N-acetilglicosamina, ou GlcNAc, como quitotetraose, quitotriose e quitobiose. Bactérias, portanto, produzem essas enzimas, a fim de degradarem GlcNAc em acetato, e frutose-6-fosfato, os quais são as principais fontes de nutrição desse organismo (BHATTACHARYA, 2008).

Quitina, bem como seus derivados (obtidos pela catálise enzimática utilizando enzimas quitinoclásticas) possuem uma série de características que podem justificar aplicações em indústrias economicamente estratégicas. Entre as diversas particularidades, destacam-se o fato de serem de origem natural e, portanto, biodegradáveis (de modo a não poluir o ambiente), serem praticamente atóxicas e serem compatíveis com tecidos animal e vegetal (THARANATHAN e KITTUR, 2003). Desse modo, existem inúmeros estudos que levam em consideração os potenciais usos de quitina e quitosana em diversas áreas.

Como exemplificado na Tabela 1, a quitina pode ser aplicada, por exemplo, na indústria alimentícia, a fim de aumentar qualidade e tempo de prateleira de diversos produtos (HAMED *et al.*, 2016) ou de ser utilizada como conservante de alimentos (SETHULEKSHMI, 2014, BARIKANI *et al.*, 2014). Em relação à quitosana, esta pode ser utilizada como prebiótico, por ser uma fonte de fibras não digeridas por enzimas estomacais (CHEN, 2010) e pode ser utilizada como suporte para imobilização de enzimas empregadas na produção de alimentos e no tratamento de resíduos com poluentes (KRAJEWSKA, 2004). Na área farmacêutica, biopolímeros de quitina e quitosana são utilizados para entrega de fármacos, devido a sua biocompatibilidade, biodegradabilidade e baixa toxicidade (DASH *et al.*, 2011), além de serem utilizados na indústria de cosméticos, aumentando a absorção de compostos ativos pela pele (MORGANTI *et al.*, 2008), e na agricultura a partir do tratamento de sementes, a fim de diminuir ataques de microrganismos (ZENG; MEI; WU, 2010). Assim como quitina e quitosana, N-acetil-D-glicosamina também possuem diversas aplicações. Na medicina, o composto pode ser utilizado no tratamento de artrose e artrite reumatoide, de modo a reduzir danos nos ossos e diminuir a liberação de superóxido por leucócitos (KAMEL, 1992). Existem também estudos que analisam o uso de GlcNAc no tratamento de doença inflamatória intestinal (DII), nos quais ficou comprovado que esse composto ajuda na restauração da camada protetora do trato gastrointestinal (BURTAN, 1998). Na indústria, N-acetilglicosamina, produzida a partir da degradação da quitina, pode ser utilizada em cosméticos, com a finalidade de aumentar a hidratação da pele e pode ser utilizada também como aditivos em leites, cervejas e vinhos (CHEN; SHEN; LIU, 2010).

Tabela 1 - Áreas de aplicação e usos potenciais de quitina e seus derivados.

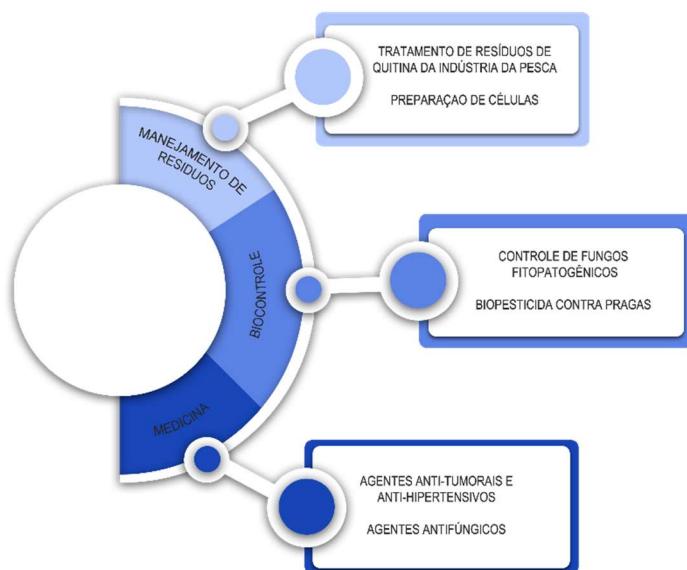
Área	Exemplos de aplicações
Farmacêutica e Biomédica	Entrega de fármacos Liberação gradual de medicamentos Regeneração de tecidos Barreira contra infecção microbiana
Cosmética e dermatológica	Transporte de ativos Desenvolvimento de produtos para cabelo, pele e higiene oral Agentes hidratantes e anti-envelhecimento
Alimentos	Barreira contra deterioração de alimentos Aditivos para conservação de alimentos e desenvolvimento de biofilmes comestíveis Estabilizantes Ingredientes para pré-bióticos
Têxtil	Fibras antimicrobianas e não-alergênicas Agentes coagulantes e floculantes para tingimentos e remoção de compostos químicos de resíduos
Agricultura	Aumento da defesa de plantas e proteção contra microrganismos e pragas Substituto para pesticidas químicos Acelerador da germinação de sementes
Indústria de papel	Aumenta a resistência do papel à umidade Embalagens biodegradáveis para alimentos
Imobilização de enzimas	Imobilização de enzimas utilizando quitina, quitosana e derivados Desenvolvimento de biossensores para medir níveis de contaminantes no meio ambiente

Fonte: Adaptado de Hamed et al., 2016.

Além de haver aplicabilidade para os produtos gerados a partir da atividade enzimática das quitinases, também existem aplicações que utilizam diretamente essas enzimas. Quitinases podem, por exemplo, ser utilizadas na degradação de resíduos de quitina gerados pela indústria de crustáceos (STOYKOV, 2014). Além disso, podem ser utilizadas também para degradar a quitina presente no exoesqueleto de alguns insetos que causam danos extensos em lavouras. Existem estudos que avaliaram a sinergia entre quitinases e endotoxinas de *Bacillus*

thuringiensis para provocar efeitos tóxicos em larvas do gênero *Choristoneura* (NAMPOOTHIRI 2004; VU, 2009). Aplicações de quitinases para o controle do crescimento de fungos fitopatogênicos, em plantações, também são estudadas extensamente. Por exemplo, a expressão de uma quitinase em plantas de batata e tabaco permitiu que essas linhagens se tornassem totalmente resistentes a alguns fungos, tais como *Alternaria alternata*, *Alternaria solani*, e *Botrytis cinerea* (HOWELL, 2003). Alguns usos de quitinases estão expostos na Figura 2, adaptada de Rathore, 2015.

[Figura 4 - Aplicações industriais de quitinases.



Fonte: Adaptado de Rathore, 2015

Quitinases possuem diversas aplicações em áreas diferenciadas e economicamente estratégicas e, desse modo, é importante que existam estudos envolvendo a purificação e a caracterização de enzimas quitinolásticas, bem como sua produção em larga escala (preferencialmente, de forma heteróloga).

2.3 *C. violaceum* como fonte de enzimas para degradação da quitina

C. violaceum é uma bactéria saprófita e anaeróbia facultativa pertencente à classe β -Proteobactéria, presente, principalmente, em ecossistemas tropicais e subtropicais (KOBURGER, 1982). Esse microrganismo tem como principal característica a produção de um pigmento violeta, o qual é bastante utilizado de forma dermatológica, além de estar presente em abundância no Rio Negro da floresta Amazônica brasileira, motivo pelo qual é estudado extensivamente por pesquisadores brasileiros (CALDAS, 1978; CALDAS, 1990; VASCONCELOS, 2003). Alguns microrganismos, como *C. violaceum*, possuem uma imensa quantidade de proteínas que com potencial de uso em diferentes indústrias e, desse modo, podem ser produzidas em larga escala.

O genoma da estirpe ATCC 12472 de *C. violaceum* foi totalmente sequenciado e analisado em 2003 pelo Brazilian National Genome Sequencing Consortium e, a partir de análises feitas desse genoma, vários genes codificando prováveis quitinases e outras enzimas quitinoclásticas, pertencentes a diferentes famílias de GHs (GH3, GH18, GH19, GH20 e GH23) e AAs (AA10), foram identificados e anotados (VASCONCELOS, 2003). A proteína CV2935, da família GH18, por exemplo, foi produzida de maneira recombinante em células de *E. coli* BL21(DE3) e avaliada para a presença de atividade quitinoclásticas (LOBO, 2013). Apesar desse estudo, ainda existem muitas proteínas de *C. violaceum* que ainda não tiveram suas capacidades enzimáticas avaliadas.

Proteínas recombinantes são, geralmente, proteínas de interesse comercial, médico, entre outros, produzidas por microrganismos que normalmente não possuem os genes que codificam a produção da proteína de interesse em questão. Desse modo, para produzir essas proteínas, é necessário isolar o gene que as codificam e clonar esse gene em um vetor de expressão. Esse vetor será utilizado para transformar as células de um microrganismo, especialmente *E. coli*, a fim de induzir a produção heteróloga dessas proteínas (CLARK, 2016). Ao levar em consideração a ausência de informação sobre a funcionalidade da maioria dos membros das famílias de GHs de *C. violaceum* envolvidas na desestruturação da quitina e a vasta quantidade de enzimas quitinoclásticas desse microrganismo que ainda não foram caracterizadas e avaliadas, percebe-se que devem haver estudos que envolvam a expressão heteróloga em *E. coli* de proteínas quitinoclásticas de *C. violaceum*, visto que esse microrganismo pode ser fonte de enzimas bastante relevantes economicamente.

2.4. Produção de proteínas por sistemas heterólogos

Organismos vivos, como plantas e microrganismos, podem ser fontes de proteínas relevantes, tanto do ponto de vista comercial e industrial, quanto terapêutico. No entanto, apenas uma pequena quantidade de proteínas pode ser isolada, de forma abundante, de seus hospedeiros nativos (HARTLEY, 2006). Proteínas heterólogas são proteínas produzidas em organismos diferentes do seu organismo de origem. A produção dessas proteínas envolve clonagem de um gene que a codifique em um sistema que maximize a produção e a purificação da mesma. Desse modo, a expressão de proteínas em sistemas heterólogos é necessária para a obtenção, em larga escala, de proteínas economicamente interessantes para diversas áreas industriais.

A expressão de proteínas a partir da transformação de microrganismos é uma técnica fundamental para a biologia molecular e para a pesquisa científica envolvendo proteínas em geral (QORONFLEH, 2007). Organismos, eucarióticos ou procarióticos, são constantemente utilizados para essa atividade por possuírem características que permitem a super expressão de proteínas de interesse. Entre as características mais comuns se destacam: rápido crescimento celular; capacidade de crescimento em meio de cultivo simples e de baixo custo; atoxicidade; manipulação e propagação fáceis e disponibilidade de diferentes estirpes bem definidas geneticamente (LODGE, 2007).

O organismo utilizado com mais frequência para a expressão de proteínas é *E. coli* por possuir todas as características citadas anteriormente, além de possuir uma série de mecanismos utilizados para otimizar a clonagem e a expressão proteica (LODGE, 2007). A criação de cepas modificadas, a fim de produzir melhorias no processo, é um exemplo de mecanismo. A cepa DH5 α possui modificações que conferem maior estabilidade ao genoma bacteriano e facilitam a transformação celular com vetores contendo genes de interesse. Uma mutação no gene endA1 permite a desativação de uma endonuclease e permite que o vetor inserido não seja degradado no citoplasma. Além disso, uma mutação em recA1 permite diminuir a recombinação homóloga do vetor com o genoma bacteriano. Isso é necessário para permitir que o plasmídeo se replique no citoplasma bacteriano (BRYANT, 1986). Outra cepa disponível comercialmente e utilizada para expressão heteróloga de proteínas é a BL21(DE3). Essa cepa possui um gene para a codificação da RNA polimerase T7 de bacteriófago DE3 sob um promotor lacUV5, induzido por IPTG (DAVANLOO, 1984). Além disso, possui elevado nível de expressão devido à ausência de duas proteases, Lon e OmpT. Essa ausência permite um maior acúmulo

de proteínas no citoplasma bacteriano, ou seja, contribui para a superexpressão proteica (PHUE, 2008).

Apesar do uso de *E. coli* permitir a alta expressão de proteínas recombinantes, existe uma série de entraves que dificultam a atividade. Por exemplo, pode haver a formação de corpos de inclusão, estruturas insolúveis formadas por agregados de proteínas fora da conformação nativa. Apesar das razões para a formação dessas estruturas não serem completamente compreendidas, a falta de mecanismos de formação de pontes dissulfeto aliada à maquinaria celular do microrganismo poder não suportar altos níveis de expressão, podem contribuir para a formação de corpos de inclusão (QORONFLEH, 2007). Um método adequado para aumentar a solubilidade de corpos de inclusão é a utilização de proteínas de fusão, as quais aumentam a produção de proteínas na fração solúvel (YOUNG, 2012). A tiorredoxina (TRX), especialmente, é utilizada em proteínas expressas em *E. coli*. Além de aumentar a solubilidade e evitar, na maioria das vezes, a formação de corpos de inclusão, a presença dessa tag de fusão pode aumentar a atividade da proteína expressa (LAVALLIE, 1993). Portanto, a otimização da produção de proteínas recombinantes deve levar em consideração a solubilidade proteica em meio aquoso.

2.5 Bioinformática para o estudo de proteínas

A bioinformática, também conhecida como biocomputação, surgiu na década de 1980 e se estabilizou como mais uma tecnologia computacional com impactos sociais, tanto individuais quanto coletivos (ARAÚJO, 2008). De modo geral, a bioinformática pode ser considerada como uma área de conhecimento que utiliza técnicas computacionais e matemáticas para apresentar novas soluções para a análise de dados biológicos (LESK, 2008). Desse modo, a bioinformática é uma poderosa ferramenta para a resolução de problemas complexos, contribuindo para o avanço das Ciências Biológicas.

Por se tratar de uma área multidisciplinar, a bioinformática reúne diferentes áreas de conhecimento, tais como ciência da computação, matemática, estatística e biologia molecular. A reunião de abordagens diferentes permite, principalmente, a análise da grande quantidade de dados obtidos a partir do sequenciamento de DNA e de proteínas (PROSDOCIMI, 2007). Além disso, a bioinformática permite que haja um compartilhamento maior de informações através da internet. Isso permite que pesquisadores, do mundo todo, possam ter acesso a dados e a ferramentas que ajudam nos estudos relacionados à, principalmente, biologia molecular.

(ARAÚJO, 2008). Diante disso, a bioinformática possui uma gama extensa de aplicações em diversas áreas da ciência.

Em relação a sua atuação na biologia molecular, a bioinformática é utilizada, principalmente, para execução de tarefas que fornecem subsídios para a pesquisa realizada nessa área. Desse modo a bioinformática pode ser utilizada, por exemplo, para a análise de sequências nucleotídicas ou proteicas, para buscar por similaridades entre sequências, para anotar genomas inteiros de organismos e para a realização de filogenia e modelagem molecular (ALVES, 2013). Ferramentas de bioinformática também podem ser utilizadas para análises em diferentes “ômicas”, tais como genômica e proteômica, a fim de responder questões complexas envolvendo cada uma dessas áreas (FACHIN, 2017). Uma aplicação direta na genômica, por exemplo, pode envolver a descoberta de diferenças entre sequências nucleotídicas entre um organismo saudável e um doente, o que influencia o diagnóstico e o tratamento de doenças (ARAÚJO, 2008). Em relação à proteômica, que estuda o conjunto de proteínas expressas em uma célula ou tecido, a bioinformática pode ser utilizada, por exemplo, para identificar sequências de aminoácidos e para predizer estruturas terciárias de proteínas (BRUM, 2007).

Ferramentas de bioinformática permitem a análise de sequências de aminoácidos e de estruturas proteicas, além de fornecerem dados sobre a interação de proteínas com seus possíveis substratos. Portanto, a bioinformática é extremamente importante em estudos que envolvem proteínas por permitir um estudo mais aprofundado das mesmas.

2.5.1 Estudo de domínios proteicos

Proteínas são formadas por resíduos de aminoácidos, com cada um destes unidos por ligações covalentes chamadas de ligações peptídicas. Essas biomoléculas, formadas tanto por extensas cadeias peptídicas quanto por um pequeno número de resíduos de aminoácidos, exercem diferentes funções biológicas e controlam, praticamente, todos as reações e processos que acontecem nas células. Em seu estado nativo, uma proteína possui uma estrutura tridimensional que vai ao encontro de sua função principal. Dentro dessas estruturas tridimensionais, existe o conceito de domínio. Domínios são definidos como uma parte da cadeia polipeptídica que é estável isoladamente, de modo a poder manter sua função, caso seja separado do restante da proteína (NELSON e COX, 2014).

Domínios são importantes para manter a função das proteínas. E isso pode ser comprovado, principalmente, pelo papel dos mesmos em enzimas. Por exemplo, a proximidade

de dois domínios pode isolar o substrato de uma enzima do ambiente, permitindo, desse modo, que a reação enzimática ocorra de forma eficiente (HAYWARD, 1999). Domínios também são importantes para manter estável a conformação nativa de proteínas. Em um estudo realizado por Xu e Nussinov (1999), foi observado que domínios são estáveis com 61 a 143 resíduos de aminoácidos. Cadeias muito curtas de aminoácidos podem não ser favoráveis o suficiente para a formação de interações entre os resíduos, de modo a prejudicar a estrutura do domínio. Por outro lado, cadeias muito grandes também não são favoráveis à formação de estruturas dos domínios, de modo que o número médio de resíduos é igual a 100.

Além de contribuir para a estrutura tridimensional correta de proteínas, domínios também são utilizados em estudos de processos evolutivos. A comparação de sequências de ácidos nucleicos ou de aminoácidos permite a observação de partes das sequências que apresentam concordância. Essas partes representam, geralmente, domínios proteicos. Desse modo, por possuírem a mesma sequência de aminoácidos em proteínas diferentes, percebe-se que domínios proteicos são estruturas altamente conservadas e que exercem funções chave nas proteínas em que estão presentes (BAGOWSKI, 2010; NELSON e COX, 2014).

O estudo de domínios proteicos utilizando, principalmente, ferramentas de bioinformática, é vital em pesquisas envolvendo proteínas desconhecidas para elucidar informações importantes sobre as mesmas.

2.5.2 Modelagem de proteínas

O arranjo dos átomos de uma proteína, de maneira tridimensional, é chamado de estrutura terciária. Quando proteínas possuem segmentos das cadeias polipeptídicas interagindo entre si e se organizando em complexos tridimensionais, fala-se em estrutura quaternária (NELSON e COX, 2014). Proteínas realizam determinadas funções, de acordo com essas estruturas tridimensionais e, desse modo, a função de uma determinada proteína está ligada a sua forma nativa. Assim, a descoberta da conformação espacial dessas biomoléculas é uma das principais funções da bioinformática (PRODOSCIMI, 2007).

A modelagem molecular consiste, principalmente, no estudo da geometria e de propriedades de proteínas através da utilização de ferramentas computacionais. O método é utilizado, principalmente, em estudos biológicos realizados com o intuito de elucidar estruturas e mecanismos de ação de proteínas desconhecidas ou difíceis de serem obtidas e purificadas em larga escala (FIGUEIREDO, 2005). No entanto, a modelagem molecular utiliza padrões de ligações químicas e moleculares e dados obtidos de experimentos físicos, tais como cristalografia com difração de raios-X ou estruturas deduzidas de ressonância magnética

nuclear (NMR). Apesar de existirem avanços nesses métodos, eles geralmente possuem problemas que persistem até hoje. Em relação ao uso da cristalografia, a obtenção de amostras em quantidade suficiente para formar os cristais é muito difícil na maioria dos casos. Além disso, cerca de apenas uma em cada 20 proteínas produzem cristais adequados, o que torna a determinação estrutural um grande desafio (FILHO e ALENCASTRO, 2003; MAGGIO e RAMNARAYAN, 2001). Tendo em vista as dificuldades em descobrir a estrutura de proteínas desconhecidas por métodos clássicos de modelagem molecular, a modelagem de proteínas por homologia se destaca como um método mais simples de ser utilizado.

Modelagem por homologia ou modelagem comparativa utiliza dados cristalográficos já existentes, presentes em bancos de dados de proteínas ou de sequências de aminoácidos, para determinar estruturas e funções de proteínas desconhecidas (FIGUEIREDO, 2005). A abordagem baseia-se em padrões gerais conservados evolutivamente. Desse modo, segundo Branden e Tooze (1991):

(a) homologia entre sequências de aminoácidos significa semelhança estrutural e funcional; (b) proteínas homólogas apresentam regiões internas conservadas, especialmente estruturas relacionadas à estrutura secundária (α -hélices e folhas- β , por exemplo); (c) as principais diferenças estruturais entre proteínas homólogas ocorrem nas regiões constituídas, principalmente, por alças (“loops”), que ligam os elementos de estruturas secundárias.

A abordagem de modelagem por homologia, portanto, é mais simples e mais utilizada por usar uma sequência de aminoácidos da proteína alvo e comparar esta com dados cristalográficos de uma proteína já estudada anteriormente, chamada de “template”, a qual possui semelhanças e homologia com a proteína estudada (SOARES, 2016). Para utilizar a modelagem por homologia de modo adequado, uma série de passos deve ser seguida. Segundo Prodoscimi (2007), esses passos envolvem:

- (1) Identificação das proteínas já modeladas que apresentam sequências similares à proteína estudada. Essa identificação ocorre, geralmente, pelo alinhamento de sequências da proteína alvo com sequências presentes em bancos de dados de proteínas;
- (2) Seleção de proteínas que serão utilizadas como molde, baseada em algoritmos de semelhança entre essas e a proteína alvo;
- (3) Alinhamento da sequência utilizada com as sequências obtidas no banco de dados;
- (4) Construção de modelos da proteína alvo através da similaridade desta com as proteínas escolhidas como molde;
- (5) Avaliação dos modelos utilizando algoritmos de teste.

A modelagem por homologia possui algumas limitações. Por exemplo, para criar um modelo tridimensional confiável, é necessário encontrar um *template* que possua mais de 30% de semelhança com a proteína alvo. No entanto, moldes para proteínas grandes, ou seja, que possuem mais de 200 resíduos de aminoácidos, geralmente, não estão presentes nos bancos de dados de proteínas visto que a maior parte dos dados cristalográficos presentes nestes são de fragmentos de proteínas maiores. *Templates* com resíduos faltando ou com rotâmeros em posições desfavoráveis podem contribuir para que o modelo tridimensional da proteína desconhecida apresente erros (MARTZ; HASS, 2020). No entanto, apesar dessas limitações, a modelagem por homologia ainda é a ferramenta mais utilizada em estudos de bioinformática que utilizam proteínas com estruturas desconhecidas por sua simplicidade e eficácia na maioria dos casos.

3 PERGUNTAS DE PARTIDA

Vasconcelos *et al.* (2003) sequenciaram completamente o genoma da estirpe ATCC 12472 de *C. violaceum*. Essa análise permitiu a descoberta de vários genes codificando prováveis quitinases e outras enzimas quitinoclásticas, pertencentes a diferentes famílias de GHs (GH3, GH18, GH19, GH20 e GH23) e AAs (AA10), foram identificados e anotados. Apesar disso, existem poucas informações sobre produção dessas proteínas de forma recombinante e sobre a verdadeira capacidade de degradação de quitina e quitosana das mesmas.

Levando em consideração que a estirpe ATCC 12472 é uma fonte promissora de enzimas com alto potencial econômico e que as informações sobre essas proteínas e seus substratos ainda são bastante obscuras, as seguintes questões podem ser elaboradas:

- Proteínas codificadas nas famílias GH3, GH19, GH20, GH23 e AA10 podem ser produzidas, de forma eficiente, de forma heteróloga em células de *E. coli*?
- As proteínas estudadas, presentes no genoma de *C. violaceum* ATCC 12472, são capazes de degradar quitina e quitosana?
- Quais são os mecanismos de catálise utilizados pelas proteínas que possuem atividade enzimática?

4 JUSTIFICATIVA

Enzimas quitinoclásticas ou quitosanásica possuem um amplo espectro de aplicações econômicas e biotecnológicas tais como: tratamento de doenças como asma e sinusite crônica, utilização em drogas antitumorais, preservação de alimentos e controle de insetos e fungos em lavouras (HAMID *et al.*, 2013). Essas proteínas, presentes em bactérias como, por exemplo, *C. violaceum*, podem ser produzidas em larga escala de forma recombinante, trazendo inúmeros benefícios a diferentes indústrias e áreas da ciência.

A degradação de quitina é realizada, industrialmente, por métodos químicos envolvendo ácidos fortes. Estas reações, no entanto, além de fornecerem um baixo rendimento, produzem resíduos tóxicos em grande quantidade (ZHOU *et. al.*, 2019). A obtenção de N-acetilglicosamina, realizada por métodos químicos, apresenta os mesmos problemas de produção de resíduos poluentes e de baixo rendimento e toxicidade (ZHANG *et al.*, 2016). Assim, a degradação enzimática de resíduos de quitina para a produção de N-acetilglicosamina e outros derivados é bastante atrativa.

A descoberta da estrutura e do mecanismo de catálise de proteínas com atividades de degradação de quitina ou quitosana é tão importante quanto à produção em larga escala das mesmas. A função de uma proteína é determinada por sua estrutura tridimensional, visto que o arranjo de átomos na mesma determina, por exemplo, o sítio catalítico de uma enzima. A descoberta de como uma proteína está arranjada no espaço, portanto, permite um maior entendimento de como uma proteína funciona, permitindo o controle e modificação da mesma, de acordo com diferentes objetivos (DARNELL, 2020).

Assim, ao levar em consideração essas informações, fica claro que a realização de estudos que envolvam a expressão heteróloga de proteínas, que degradam quitina ou quitosana, de *C. violaceum*, bem como a descoberta da estrutura tridimensional das mesmas, é importante, tanto cientificamente, quanto economicamente.

5 OBJETIVOS

5.1 Objetivo Geral

Producir, de maneira heteróloga, proteínas codificadas no genoma de *C. violaceum* ATCC 12472 e avaliar as proteínas recombinantes obtidas para a presença de atividade enzimática. Além disso, caracterizar, *in silico*, proteínas que possuem essa atividade, a fim de entender mais sobre o modo como as mesmas interagem com seus substratos.

5.2 Objetivos Específicos

- Transformar células de *E. coli* DH5 α e BL21(DE3) com plasmídeos sintéticos codificantes para as proteínas alvo (proteínas de *C. violaceum* ATCC 12472 das famílias GH3, GH19, GH20, GH23 e AA10).
- Avaliar a expressão das proteínas recombinantes em diferentes temperaturas de cultivo.
- Avaliar se as proteínas recombinantes apresentam atividade para degradação de quitina (quitinoclástica) e de quitosana (quitosanásica).
- Escolher de acordo com critérios a proteína recombinante mais promissora, para produção e purificação em larga escala.
- Estudar os domínios conservados das proteínas com atividade enzimática, a fim de entender o modo como essas proteínas degradam substratos e investigar semelhanças entre proteínas da mesma família de hidrolases de glicosídeos (GHs).
- Gerar um modelo tridimensional para proteínas com atividade quitinoclástica ou quitosanásica.

6 MATERIAIS E MÉTODOS

6.1 Análise *in silico* de proteínas e construção de plasmídeos sintéticos

Foram realizadas análises *in silico* de diferentes proteínas presentes em 10 famílias de proteínas (AA10, GH3, GH16, GH18, GH19, GH20, GH23, GH24, GH46 e GH73), a fim de escolher proteínas promissoras para a expressão heteróloga. As análises foram feitas de modo a obter informações sobre algumas proteínas de cada família, com o propósito de escolher as que, teoricamente, terão uma melhor expressão em células de *E. coli*. Desse modo, número de nucleotídeos, massa molecular, ponto isoelétrico, presença de peptídeo sinal e solubilidade foram obtidos para cada proteína. Massa molecular, número de nucleotídeos e pI de cada proteína foram medidos, utilizando a ferramenta Compute pI/Mw tool do ExPASy (GASTEIGER *et al.*, 2005). A presença de peptídeo sinal foi determinada ao submeter as sequências de aminoácidos no servidor SignalP (NIELSEN *et al.*, 1997). Para determinar a solubilidade de cada proteína, foram utilizados quatro servidores diferentes: PROSO (SMIALOWSKI *et al.*, 2007), PROSO II (SMIALOWSKI *et al.*, 2012), Protein-Sol (HEBDITCH *et al.*, 2017) e SCRATCH Protein Predictor (CHENG *et al.*, 2005). Após a escolha, houve a síntese de plasmídeos codificando cada proteína escolhida. Esses plasmídeos foram construídos com otimização de códons para expressão em *E. coli* e os fragmentos de DNA foram clonados, cada um, nos vetores pET-32a, com os sítios de restrição para BamHI e XhoI. As proteínas foram fusionadas com tags de tiorredoxina (Trx), a fim de aumentar a solubilidade das mesmas, e as construções foram produzidas pela empresa GenOne.

6.2 Transformação de células para a produção de proteínas recombinantes de *C. violaceum*

6.2.1 Transformação de *E. coli* DH5 α por choque térmico

Com o intuito de realizar a transformação por choque térmico, foi necessário preparar células competentes de *E. coli* por tratamento com cloreto de cálcio, a partir do protocolo desenvolvido por Sambrook e Russel (2001). As células de DH5 α foram plaqueadas em meio

LB ágar, sem antibiótico, e cultivadas, em estufa a 37 °C, por 16 horas. Após o crescimento, houve a escolha de uma colônia isolada e a inoculação desta colônia em 5 mL de meio LB caldo, também sem antibiótico. A cultura foi incubada a 37 °C, 180 rpm e, após 16 h, 1 mL da cultura foi transferido para um frasco contendo 100 mL de meio LB caldo. O frasco foi incubado a 37 °C, sob agitação de 180 rpm até atingir uma densidade óptica 0,2 a 0,3, observada por espectrofotômetro à 600 nm (D.O.600). Após atingirem a D.O.600 desejada, as células foram coletadas por centrifugação a 3.000 x g por 10 min a 4 °C. Posterior à centrifugação, as células foram ressuspensas em 20 mL de solução gelada de CaCl₂ 50 mM e incubadas no gelo por 30 minutos. As células foram novamente centrifugadas a 3.000 x g por 10 min a 4 °C e ressuspensas em 2 mL de CaCl₂ 50 mM. Em seguida, foram aliquotadas em porções de 200 µL, em microtubos estéreis com capacidade para 1,5 mL. Para iniciar o processo de transformação foram adicionados, a cada microtubo, 4 µL (5 ng/Ml, totalizando 20 ng) das soluções de plasmídeos (GenOne). Além dos tubos contendo as construções, um tubo contendo 4 µL de água deionizada estéril e 200 µL da suspensão de células de *E. coli* foi preparado para ser utilizado como controle. Os microtubos foram incubados em gelo por 30 min. Posteriormente, foram incubados em banho-maria a 42 °C por 3 min e, em seguida, incubados em gelo por 2 min. Foram adicionados 800 µL de meio SOC a cada microtubo e estes foram incubados, novamente em banho-maria a 37° C por 1h e 30 min, com agitação ocasional. Em seguida, foram plaqueados 150 µL de cada transformação em placas contendo LB ágar contendo carbenicilina 100 µg/mL. As placas foram incubadas em posição invertida a 37 °C por, no mínimo, 12 horas.

6.2.2 Extração de DNA Plasmidial

A extração de DNA plasmidial foi feita utilizando o kit NucleoSpin® Plasmid (Machery-Nagel), de acordo com o manual do fabricante. Duas colônias isoladas, de cada transformação, foram inoculadas em frascos diferentes, contendo 5 mL de meio LB caldo com carbenicilina em uma concentração de 100 mg/mL. Essas culturas foram incubadas a 37 °C, sob agitação de 180 rpm, por 16 horas. Em seguida, as células de cada transformação foram transferidas para microtubos estéreis de 2 mL e coletadas por centrifugação, em microcentrífuga, a 13.000 rpm por 30 segundos. O sobrenadante de cada microtubo foi descartado e o segmento celular resultante foi ressuspensão em 250 µL de Tampão A1 (tampão de ressuspensão contendo RNase), fornecido pelo kit. Após verificar se não existe nenhum

agregado celular, o Tampão A2 (tampão de lise, contendo solução de hidróxido de sódio a 0,5-2,0 %) foi adicionado. O conteúdo dos microtubos foi misturado, suavemente, por inversão e, em seguida, foram incubados em temperatura ambiente por 4 min. Foram adicionados, a cada tubo, 300 μ L de tampão A3 (tampão de neutralização com hidrocloreto de guanidina 36-50%) e esses foram misturados por inversão até o desaparecimento da cor azul. Os microtubos foram centrifugados durante 5 min a 13.000 rpm. Após a centrifugação, 750 μ L do sobrenadante foram adicionados às colunas NucleoSpin® Plasmid fornecidas pelo kit. As colunas, juntamente com os tubos de coleta de 2 mL foram centrifugadas por 1 min a 13.000 rpm e, em seguida, houve o descarte do fluxo do tubo coletor e adição de 600 μ L do Tampão A4 (tampão de lavagem contendo etanol). As colunas foram, novamente, centrifugadas por 1 min a 13.000 rpm. O fluxo presente no tubo de coleta após a centrifugação foi descartado e as colunas foram centrifugadas por 2 min a 13.000 rpm, a fim de retirar resíduos de etanol. Em seguida, as colunas foram encaixadas em microtubos de 1,5 mL estéreis e, posteriormente, houve a adição de 25 μ L de Tampão TE pH 8,0 estéril. Após a adição, os microtubos foram incubados por 3 minutos em temperatura ambiente e centrifugados durante 1 min a 13.000 rpm (etapa que foi realizada duas vezes). O DNA plasmidial obtido foi quantificado por um espectrofotômetro Nanodrop, com absorbâncias de 260 e 280nm, com os resultados expressos em ng/ μ L.

6.2.3 Digestão de DNA Plasmidial

A reação de digestão foi realizada utilizando 2 μ L de Tampão R 10X (Fermentas), 2 μ L da enzima BamHI 15 U/ μ L (Invitrogen), 2 μ L da enzima XhoI 10 U/ μ L (Fermentas) e 16 μ L de DNA plasmidial, totalizando uma solução de 20 μ L. As soluções foram incubadas a 37 °C por 3 h e, em seguida, a 80 °C por 20 min em banho-maria. Após a digestão, houve a confirmação da mesma por eletroforese em gel de agarose 0,8%.

6.2.4 Eletroforese de DNA em gel de agarose

O gel de agarose, com concentração de 0,8% (m/v), foi preparado de acordo com o protocolo descrito por Sambrook e Russel (2001). Portanto, o gel, com a concentração desejada, foi preparado a partir da dissolução de 0,8 g de agarose (Vetec) em 100 mL de solução de água destilada com 2 mL de tampão TAE (Tris-acetato 40 mM, EDTA 1 mM pH

8,0). Após o preparo do gel, foi acrescentado ao mesmo 5 μ L de brometo de etídio (0,5 μ g/mL). As amostras resultantes da digestão de DNA plasmidial foram misturadas a 4 μ L de tampão Loading Dye Green 5x (Promega), totalizando 20 μ L. Cada uma das amostras foi aplicada a um poço no gel de agarose e a eletroforese foi realizada a uma voltagem constante de 100 v. Utilizou-se uma cuba horizontal e o gel foi submerso em tampão de corrida TAE 1x. O marcador utilizado foi o 1 Kb DNA Ladder (Thermo Fisher Scientific). Após o término da eletroforese, as bandas de DNA foram visualizadas em um transiluminador (*Vilber Lourmat*) a partir da exposição do gel à luz UV. As imagens foram obtidas por um fotodocumentador (*Vilber Lourmat*) acoplado ao equipamento.

6.2.5 Transformação de *E. coli* BL21(DE3) por choque térmico

A transformação por choque térmico de células de *E. coli* BL21(DE3) foi feita utilizando o mesmo protocolo descrito no item 6.2.1. Células competentes de BL21 foram preparadas por protocolo com cloreto de cálcio. De cada suspensão celular foram retirados 200 μ L e adicionados a tubos estéreis com capacidade de 1,5 mL. Houve a adição de 2 μ L de DNA plasmidial (aproximadamente 5 ng/ μ L, totalizando 10 ng) obtido a partir da extração feita a cada microtubo para cada transformação (foram obtidos dois clones de cada transformação, mas apenas um foi utilizado). Novamente, houve o preparo de um microtubo contendo 2 μ L de água deionizada estéril e 200 μ L da suspensão de células para ser utilizado como controle. Os microtubos foram incubados em gelo por 30 min. Posteriormente, foram incubados em banho-maria a 42 °C por 3 min e, em seguida, incubados em gelo por 2 min. Foram adicionados 800 μ L de meio SOC a cada microtubo e estes foram incubados, novamente em banho-maria a 37° C por 1h e 30 min, com agitação ocasional. Em seguida, foram plaqueados 150 μ L de cada transformação em placas contendo LB ágar contendo carbenicilina 100 mg/mL. As placas foram incubadas em posição invertida a 37 °C por, no mínimo, 12 horas.

6.3 Produção de proteínas recombinantes

6.3.1 *Triagem de temperatura de indução da expressão de proteínas recombinantes em E. coli BL21(DE3)*

As células de *E. coli* BL21(DE3), contendo as nove estirpes diferentes para a produção de hidrolases de glicosídeos, foram plaqueadas em meio LB ágar e incubadas invertidas a 37 °C por, no mínimo, 12 horas. Uma colônia isolada de cada placa foi inoculada em 3 mL de meio LB caldo contendo carbenicilina em uma concentração de 200 µg/mL. Foram utilizados tubos tipo Falcon com capacidade igual a 15 mL. Os tubos contendo as culturas foram incubados por 37 °C com agitação orbital constante de 180 rpm por um período de três horas. Após esse período, 1,5 mL de cada cultura foram transferidos para outros tubos Falcon com capacidade para 15 mL. Houve a adição de IPTG (Isopropil β -D-1-tiogalactopiranosida) para uma concentração final de 0,4 mM e, posteriormente, tanto os tubos com IPTG quanto os tubos sem IPTG adicionado foram incubados a 25 °C sob agitação de 180 rpm por 3 h. Essa indução foi repetida para temperaturas de 20 °C, 30°C e 35 °C.

6.3.2 *Lise celular*

Posteriormente às induções em diferentes temperaturas, as células foram coletadas por centrifugação a 6.000 x g por 20 min a 5 °C e ressuspensas em 150 µL de tampão de lise TES [Tris-HCl 50 mM pH 8,0, NaCl 0,15 M, EDTA 2 mM e glicerol 10% (v/v)]. Após a adição do tampão de lise, as células ressuspensas foram transferidas para microtubos com capacidade de 1,5 mL, nos quais foram adicionados 15 µL de Triton X-100 1% (v/v) e lisozima para uma concentração final de 100 µg/mL (a partir de um estoque de 10 mg/mL). Em seguida, os tubos foram incubados em banho-maria a 35 °C por 30 min para desenvolverem viscosidade. Após esse tempo, houve a adição de 2 µL de DNase I, RNase- free, HC (50 U/L) (Thermo Fisher Scientific) e de cloreto de magnésio (MgCl₂) e cloreto de cálcio (CaCl₂) para uma concentração final de 8 mM e 10 mM, respectivamente. Posterior à adição, os microtubos foram incubados, novamente, a 37 °C por 30 min até o desaparecimento da viscosidade. Ao serem retirados do banho-maria, os microtubos foram centrifugados, em microcentrífuga, a 10.000 rpm por 20 min. O sobrenadante obtido, tanto para células induzidas com IPTG, quanto para células não

induzidas, foi transferido para microtubos com capacidade de 1,5 mL. Esse sobrenadante representa a fração solúvel das proteínas, enquanto o pellet de células obtido representa a fração insolúvel.

6.3.3 Eletroforese de proteínas em gel de poliacrilamida sob condições desnaturantes

A eletroforese em gel de poliacrilamida, contendo dodecil sulfato de sódio (SDS-PAGE), foi realizada de acordo com o protocolo descrito por Laemmli (1970), adaptado para géis montados entre placas de vidro. O gel de empacotamento foi preparado com 3,5% de acrilamida, 1% de SDS e tampão Tris-HCl 1,0 M pH 6,8, enquanto para o gel de resolução foram utilizados 15% de acrilamida, 1% de SDS e tampão Tris-HCl 1,5 M pH 8,8. Os extratos totais de proteínas (30 μ L de cada) foram diluídos em 10 μ L tampão de amostra [Tris-HCl 0,0625 M, pH 6,8; SDS 2% (m/v); 2- β -mercaptoetanol 5% (v/v) e azul de bromofenol 0,001% (m/v)] quatro vezes concentrado. Os corpos de inclusão, contendo a fração insolúvel de proteínas, também foram submetidos à eletroforese após serem diluídos com 20 μ L do mesmo tampão de amostra. Após serem misturadas com tampão, as amostras foram aquecidas a 100 °C por 10 min. A eletroforese para as amostras produzidas pelas induções em temperaturas diferentes (20 °C, 25 °C, 30 °C e 35° C) foram realizadas em sistema vertical Dual Gel Caster (Hoefer, São Francisco, USA) sob voltagem constante de 120 V. O tampão utilizado para as corridas foi Tris-HCl 0,025 M, pH 8,3; glicina 0,192 M; SDS 0,1 % (m/v). Foram utilizados, nas corridas, dois marcadores de peso molecular diferentes. O primeiro foi o Low Molecular Weight (GE Healthcare) e o segundo foi o Page Ruler™ Prestained Protein Ladder (Thermo Fisher Scientific). As bandas de proteínas, tanto para fração solúvel quanto para fração insolúvel, foram coradas com Comassie Coloidal com CBB250, preparado de acordo com o protocolo estabelecido por Dyballa e Metzger (2009), e descoradas com água destilada. Os géis obtidos foram fotodocumentados e as bandas de proteínas foram identificadas.

6.4 Teste para atividade quitinoclástica das proteínas recombinantes

6.4.1 *Teste para atividade quitinoclástica utilizando quitina coloidal*

Os extratos proteicos de todas as proteínas recombinantes foram analisados para a presença de atividade quitinoclástica em quitina coloidal. Como controles, foram utilizados água destilada, tampão acetato de sódio 50 Mm, pH 5,2 e o extrato proteico obtido pela indução de células de BL21(DE3) não transformadas. Para tanto, houve a adição, em tubos eppendorf de tampa rosqueada, à temperatura ambiente de 250 µL das amostras e de 250 µL de quitina coloidal, preparada seguindo o protocolo determinado por Boller (1992) com modificações de Martins-Miranda (2002). Os tubos tiveram seus conteúdos homogeneizados e foram incubados a 37 °C sob leve agitação e, em seguida, foram incubados a 100 °C por 5 minutos. Após esses procedimentos, os tubos foram resfriados, imediatamente, em banho de gelo e centrifugados a 10.000 x g por 10 minutos. Com o objetivo de realizar a próxima etapa do experimento, foram adicionados, em novos tubos (tampa rosqueada, 2 mL) 300 µL do sobrenadante obtido após a centrifugação e 10 µL de β -Glucoronidase. Os tubos foram incubados a 37 °C por 60 min, sob leve agitação. Após esse tempo, foram incubados a 100 °C por 5 minutos e resfriados, imediatamente, em banho de gelo. Em seguida, houve a adição de 190 µL de tampão acetato de sódio 50 mM, pH 5,2 e de 100 µL de tetraborato de potássio 0,6 M. Os tubos foram, novamente, incubados a 100 °C por 5 minutos e resfriados em banho de gelo por 5 minutos. Por último, houve a adição de DMAB diluído em ácido acético (1:2). Os tubos tiveram sua solução homogeneizada e foram incubados a 37 °C por 20 minutos. As absorbâncias dos referidos tubos foram medidas em espectrofotômetro (Genesys 10 UV Scanning) foram medidas a 585 nm, imediatamente após a incubação final. A atividade foi feita em duplicata.

6.4.2 *Teste para atividade quitinoclástica utilizando quitina insolúvel*

Os extratos proteicos de todas as proteínas recombinantes também foram analisados para a presença de atividade quitinoclástica com quitina insolúvel. O protocolo seguido foi determinado por Boller (1992) com modificações de Martins-Miranda (2002) e de Fernandes

(2017). Os controles foram os mesmos utilizados na atividade com quitina coloidal (água destilada, tampão acetato de sódio 50 mM, pH 5,2 e o extrato proteico obtido pela indução de células de BL21(DE3) não transformadas). Em tubos eppendorf de tampa rosqueada, foram adicionados 2,5 mg de quitina de casca de caranguejo (Merck KGaA, Darmstadt, Germany) e 100 μ L de tampão acetato de sódio 50 mM, pH 5,2. Em seguida, houve a adição de 200 μ L de cada amostra (extratos proteicos das proteínas recombinantes) e de cada controle, totalizando 300 μ L para cada reação. Os tubos foram mantidos a 37°C, sob agitação orbital de 180 rpm, por 24 h. Após esse período, a reação foi parada por aquecimento a 100°C por 5 min e imediato resfriamento em banho de gelo e centrifugados a 10.000 x g por 10 minutos. Foram adicionados, em novos tubos (tampa rosqueada, 2 mL), o sobrenadante obtido após a centrifugação e 10 μ L de β -Glucoronidase. Os tubos foram incubados a 37 °C por 60 minutos, sob leve agitação. Após esse tempo, foram incubados a 100 °C por 5 minutos e resfriados, imediatamente, em banho de gelo. Houve a adição de 190 μ L de tampão acetato de sódio 50 mM, pH 5,2 e de 100 μ L de tetraborato de potássio 0,6 M. Os tubos foram, novamente, incubados a 100 °C por 5 minutos e resfriados em banho de gelo por 5 minutos. Por último, houve a adição de DMAB diluído em ácido acético. Os tubos tiveram sua solução homogeneizada e foram incubados a 37 °C por 20 minutos. As absorbâncias dos tubos foram medidas em espectrofotômetro (Genesys 10 UV Scanning) foram medidas a 585 nm, imediatamente após a incubação final. Esse ensaio também foi realizado em duplicata.

6.5 Teste para atividade quitosanásica das proteínas recombinantes

Assim como para a atividade quitinoclástica, houve a necessidade de analisar os extratos das proteínas seguindo um protocolo, adaptado de Miller (1959), que utiliza ácido 3,5 dinitrossalicílico (DNS) para medir a atividade quitosanásica. Os controles utilizados foram água destilada e o extrato proteico, obtido pela indução de células de BL21(DE3) não transformadas. Em tubos eppendorf de tampa rosqueada, foram adicionados 250 μ L de quitosana coloidal e 250 μ L de cada amostra. A mistura foi incubada por 37°C, em banho maria, por 1h. Após a incubação, foram adicionados 500 μ L do reagente DNS a cada tubo, os quais foram incubados 100 °C por 5 minutos e resfriados em banho de gelo por 5 minutos. Após esse período, foi adicionado 1 mL de solução de tartarato duplo de sódio e potássio a cada tubo. Por fim, houve a centrifugação dos tubos a 10.000 x g por 5 minutos. Os sobrenadantes obtidos para cada tubo tiveram suas absorbâncias medidas em espectrofotômetro (Genesys 10 UV Scanning)

a 540 nm após zerar o aparelho com o branco (água destilada). O teste para atividade quitosanásica foi realizado em duplicata para cada amostra de proteína recombinante e para os controles.

6.6 Produção de uma proteína recombinante em *E. coli* em escala maior

6.6.1 Indução de proteína recombinante

As células de *E. coli* BL21(DE3), contendo o plasmídeo para uma das hidrolases de glicosídeo estudadas, foram plaqueadas em meio LB ágar e incubadas invertidas a 37 °C por, no mínimo, 12 horas. Uma colônia isolada de cada placa foi inoculada em 10 mL de meio LB caldo, contendo carbenicilina em uma concentração de 200 µg/mL. Foi utilizado um Erlenmeyer com capacidade igual a 100 mL, a fim de produzir um pré-inóculo. O Erlenmeyer, contendo a cultura, foi incubado por 37 °C com agitação orbital constante de 180 rpm por um período de três horas. Após esse período, 1,0 ml do pré-inóculo foi transferido para 250 ml de meio LB caldo presente em um Erlenmeyer com capacidade para 3 L. Houve a adição de IPTG (Isopropil β-D-1-tiogalactopiranosida) para uma concentração final de 0,4 mM e, posteriormente, o Erlenmeyer, contendo o inóculo, foi incubado por um período de três horas a uma temperatura de 30 °C sob agitação orbital constante de 180 rpm. Posterior à indução, as células foram coletadas por centrifugação a 6.000 x g por 20 min a 5 °C e submetidas ao processo de lise celular, já descrito anteriormente, utilizando, porém, um tampão diferente do utilizado para a indução piloto (fosfato de sódio 50mM, cloreto de sódio 300mM, imidazol 10mM, pH 7,4).

6.6.2 Purificação de proteína recombinante

Com a finalidade de purificar a proteína expressa de forma heteróloga em células de BL21(DE3), foi realizado um processo de cromatografia de afinidade por íons metálicos utilizando uma resina de cobalto, HisPur™ Cobalt Resin (Thermo Fisher Scientific). A matriz contém o volume de 3 ml e foi previamente equilibrada com o tampão de equilíbrio contendo fosfato de sódio 50mM, cloreto de sódio 300m, imidazol 10mM, pH 7.4. Após esse tampão ser

totalmente drenado da matriz, a fração solúvel da proteína, obtida a partir da lise celular, foi aplicada na matriz. O preparado foi passado duas vezes na matriz, a fim de obter o máximo de ligação. Em seguida, a matriz foi lavada com o mesmo tampão de equilíbrio e, posteriormente, com um tampão de eluição, contendo fosfato de sódio 50mM, cloreto de sódio 300m, imidazol 150mM, pH 7,4. Os picos obtidos foram dialisados contra o tampão fosfato de sódio 50 mM, NaCl 500 mM e, em seguida, analisados por eletroforese em gel de poliacrilamida (SDS-PAGE). O fluxo de eluição da coluna foi de 1 mL/min e as frações coletadas foram de 2 mL por tubo e ensaio. As leituras das frações foram feitas por um espectrofotômetro Genesys 10 UV Scanning e suas absorbâncias foram medidas a 280 nm.

6.7 Ferramentas de bioinformática para estudo de proteínas

6.7.1 *Análise dos domínios das proteínas estudadas*

Com o propósito de obter informações sobre o modo como as proteínas que obtiveram resultados na degradação de quitina ou quitosana interagem com seus substratos, foram realizados alguns estudos de bioinformática sobre os domínios que compõem essas. Para tanto, as sequências de aminoácidos para cada proteína, obtidas no banco de dados do NCBI foram convertidas no formato FASTA e submetidas ao Banco de Domínios Conservados ou CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>), também do NCBI (MARCHLER-BAUER, 2002). Após obter os resultados para cada proteína, foram preparadas representações gráficas da arquitetura de domínios de cada proteína escolhida para ser produzida em *E. coli* utilizando um programa chamado IBS (Liu et. al., 2015).

6.7.2 *Predição da estrutura 3D de proteínas*

6.7.2.1 *Escolha de modelos e modelagem por homologia*

A sequência de aminoácidos, para as proteínas que apresentaram atividade quitinoclástica ou quitosanásica, foi submetida à análise no SWISS-MODEL (BIASINI, 2014)

e no BLASTP (MCGINNIS, 2004), a fim de obter modelos que possuam alta semelhança com as proteínas alvo. Além de fornecer templates, a estrutura prevista para as proteínas em questão também pode ser fornecida pelo Swiss-Model, o qual apresenta parâmetros que avaliam as melhores estruturas preditas.

6.7.2.2 Avaliação dos modelos

As estruturas preditas foram avaliadas utilizando o servidor MolProbity (CHEN, 2010). O servidor, em questão, utiliza o *plot* de Ramachandran, o qual relaciona a posição de resíduos de aminoácidos com regiões energeticamente favoráveis, baseado em ângulos de rotação (SOARES, 2016). Além disso, as estruturas preditas foram submetidas ao ERRAT2 (COLOVOS, 1993) e ao VERIFY-3D (BOWIE, 1991), com o propósito de avaliar a qualidade das mesmas. O servidor ProSA-web (Protein Structure Analysis (WIEDERSTEIN, 2007) também foi utilizado para avaliar a qualidade geral dos modelos utilizando Z-score. As estruturas preditas ainda foram submetidas ao software PyMOL (The PyMolMolecular Graphics System, Versão 2.0 Schrödinger, LLC) a fim de compará-las com os templates utilizados.

7 RESULTADOS E DISCUSSÃO

7.1 Seleção de proteínas a serem expressas em *E. coli*

Após análises *in silico*, foram escolhidas nove proteínas para serem clonadas em células de *E. coli* DH5 α e expressas de forma solúvel por células de *E. coli* BL21(DE3). Essa escolha foi baseada em dados de, principalmente, solubilidade, visto que proteínas, produzidas em corpos de inclusão, possuem dificuldade maior em serem expressas e purificadas. Os dados sobre o nome e a identificação de cada proteína, massa molecular teórica e a solubilidade de cada proteína estão representados nas Tabelas 2, 3 e 4, respectivamente. Ressalta-se que as proteínas serão identificadas pela ORF (*open reading frame*) de cada inserto devido ao nome extenso de cada uma delas.

Tabela 2 - Identificação das proteínas recombinantes de fusão, expressas a partir dos insertos clonados em pET32a(+), codificando hidrolases de glicosídeos de *C. violaceum* ATCC 12472.

Inserto (ORF)	Identificação
CV2592	AAQ60262.1 carbohydrate-binding protein [<i>Chromobacterium violaceum</i> ATCC 12472]
CV0259	AAQ57938.1 probable glycosyl hyrolase, family 3 [<i>Chromobacterium violaceum</i> ATCC 12472]
CV0554	AAQ58230.1 probable carbohydrate- binding protein [<i>Chromobacterium violaceum</i> ATCC 12472]
CV0096	AAQ57775.1 hypothetical protein CV_0096 [<i>Chromobacterium violaceum</i> ATCC 12472]
CV3323	AAQ60987.1 carbohydrate-binding protein [<i>Chromobacterium violaceum</i> ATCC 12472]
CV2065	AAQ59745.2 probable beta-N- acetylhexosaminidase [<i>Chromobacterium violaceum</i> ATCC 12472]
CV0348	AAQ58026.2 conserved hypothetical protein [<i>Chromobacterium violaceum</i> ATCC 12472]
CV2641	AAQ60311.1 cell invasion protein [<i>Chromobacterium violaceum</i> ATCC 12472]
CV2601	AAQ60281.2 peptidoglycan N- acetylmuramoylhydrolase [<i>Chromobacterium violaceum</i> ATCC 12472]

Tabela 3 – Famílias e massas moleculares teóricas (monoisotópicas) das proteínas recombinantes de fusão, expressas a partir dos insertos clonados em pET32a(+), codificando hidrolases de glicosídeos de *C. violaceum* ATCC 12472.

Inserto (ORF)	Massa molecular teórica (em Da)	Família de GH ou AA a qual pertence o domínio catalítico
CV2592	59640.28	AA10
CV0259	72982.23	GH3
CV0554	58441.73	AA10
CV0096	43888.71	GH16
CV3323	59726.24	AA10
CV2065	57001.57	GH3
CV0348	77092.94	GH23
CV2641	33818.72	GH23
CV2601	40690.61	GH23

Fonte: elaborada pela autora

Tabela 4 – Solubilidades teóricas das proteínas recombinantes de fusão, expressas a partir dos insertos clonados em pET32a(+), codificando hidrolases de glicosídeos de *C. violaceum* ATCC 12472.

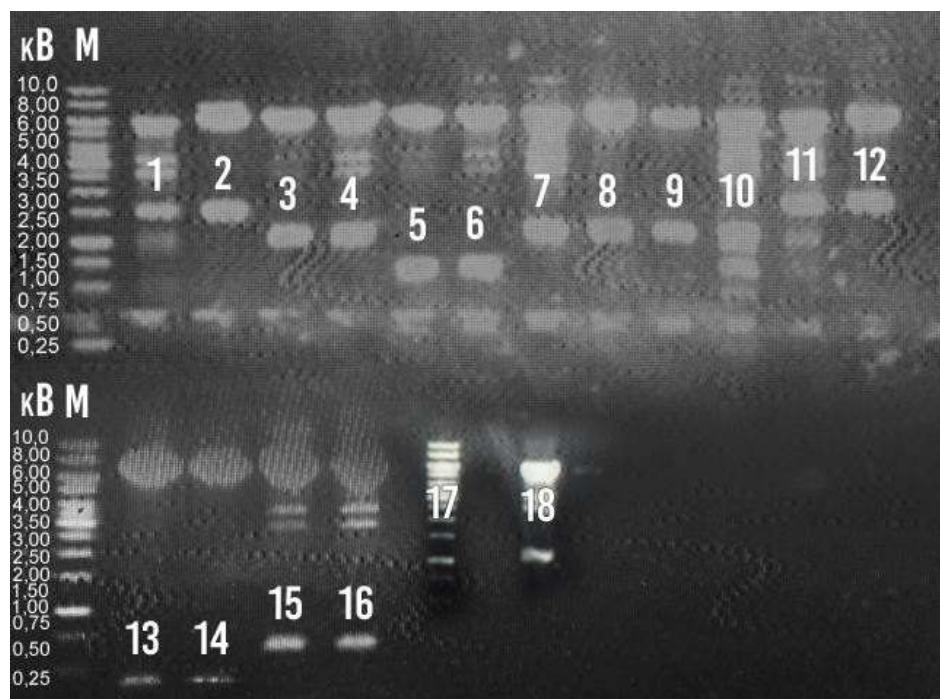
Inserto (ORF)	Solubilidade (PROSO)	Solubilidade (PROSO II)	Solubilidade (Protein-Sol)	Solubilidade (SCRATCH Protein Predictor)
CV2592	0.721	0.443	0.385	0.977889
CV0259	0.815	0.438	0.342	0.582346
CV0554	0.635	0.490	0.366	0.649400
CV0096	0.655	0.480	0.358	0.882515
CV3323	0.583	0.554	0.347	0.945102
CV2065	0.556	0.653	0.494	0.716236
CV0348	0.703	0.655	0.763	0.830785
CV2641	0.695	0.287	0.476	0.672451
CV2601	0.572	0.686	0.673	0.887462

PROSO e PROSO II: utilizam dados extraídos do Protein Expression and Purification Core Facility (pepcDB) e do Protein Data Bank (PDB) para desenvolver um algoritmo com mais de 70% de acurácia. Resultados maiores que 0,6 são mais favoráveis. Protein-Sol: utiliza dados prévios de solubilidade de diferentes proteínas expressas em células de *E. coli*. Resultados maiores que 0,45 são mais favoráveis. SCRATCH Protein Predictor: utiliza dados de proteínas superexpressas em células de *E. coli*. Resultados maiores que 0,5 são mais favoráveis. Fonte: elaborada pela autora

7.2 Transformação de células de *E. coli* com insertos sintéticos

Foram sintetizados nove plasmídeos a partir das sequências de nucleotídeos das proteínas escolhidas, a fim de produzi-las de modo recombinante em células de *E. coli*. Para permitir essa produção, houve a clonagem de células de *E. coli* da estirpe DH5 α . A transformação foi verificada, de forma inicial, ao plaquear as células em meio LB sólido com carbenicilina 100 μ g/mL. Todas as placas apresentaram crescimento ao serem submetidas à incubação por 37 °C por um período de 12 horas. A placa que continha células não transformadas, ou seja, o controle, não apresentou crescimento. Em seguida, com o objetivo de confirmar a presença dos insertos no genoma bacteriano de *E. coli* DH5 α transformadas com os plasmídeos sintéticos, houve a extração de DNA plasmidial das mesmas. Posterior à quantificação do DNA por Nanodrop e à digestão por enzimas de restrição (BamHI 15 U/ μ L e XhoI 10 U/ μ L), houve a aplicação das amostras em poços presentes em gel de agarose 0,8%. Foram escolhidas duas colônias de cada placa contendo células de DH5 α transformadas, totalizando 18 amostras de DNA plasmidial (Figura 5).

Figura 5 - Eletroforese em gel de agarose (0,8%).



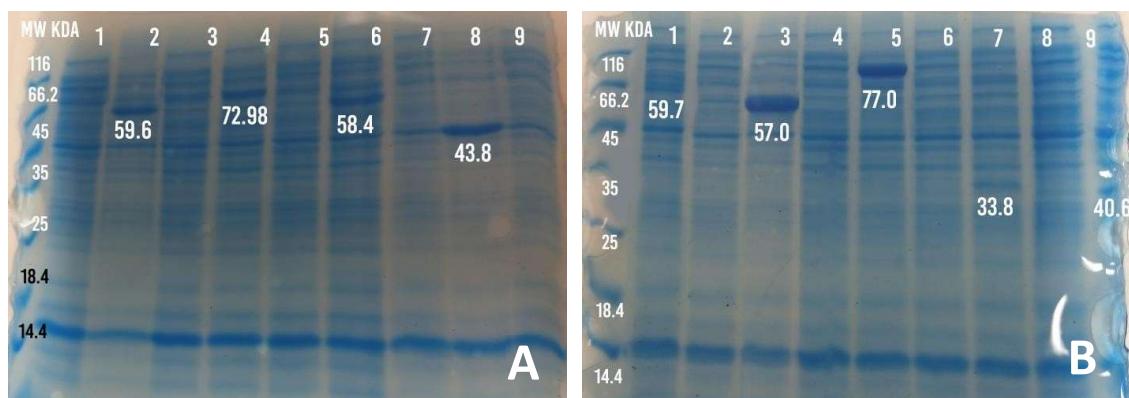
Todos os poços contêm DNA plasmidial digerido por enzimas de restrição (BamHI 15 U/ μ L e XhoI 10 U/ μ L).
 Poço 1: DNA plasmidial para CV0259 (clone 1). Poço 2: DNA plasmidial para CV0259 (clone 2). Poço 3: DNA plasmidial para CV0554 (clone 1). Poço 4: DNA plasmidial para CV0554 (clone 2). Poço 5: DNA plasmidial para CV0096 (clone 1). Poço 6: DNA plasmidial para CV0096 (clone 2). Poço 7: DNA plasmidial para CV3323 (clone 1). Poço 8: DNA plasmidial para CV3323 (clone 2). Poço 9: DNA plasmidial para CV2065 (clone 1). Poço 10: DNA plasmidial para CV2065 (clone 2). Poço 11: DNA plasmidial para 0348 (clone 1). Poço 12: DNA plasmidial para CV0348 (clone 2). Poço 13: DNA plasmidial para CV2641 (clone 1). Poço 14: DNA plasmidial para CV2641 (clone 2). Poço 15: DNA plasmidial para CV2601 (clone 1). Poço 16: DNA plasmidial para CV2601 (clone 2). Poço 17: DNA plasmidial para CV2592 (clone 1). Poço 18: DNA plasmidial para CV2592 (clone 2). M: Marcador 1 Kb Plus DNA Ladder (Thermo Fisher Scientific). Fonte: elaborada pela autora.

Após a transformação em células de DH5 α , houve a extração do DNA plasmidial dessas células, com o intuito de transformar células de BL21(DE3) e permitir a indução das proteínas recombinantes escolhidas. A transformação dessas células também foi comprovada pelo plaqueamento das células transformadas para cada plasmídeo em meio LB sólido contendo carbenicilina 100 μ g/mL. Todas as placas contendo as células transforadas para cada um dos plasmídeos apresentaram crescimento após condições de incubação de 37 °C por 12 horas.

7.3 Indução piloto de proteínas recombinantes

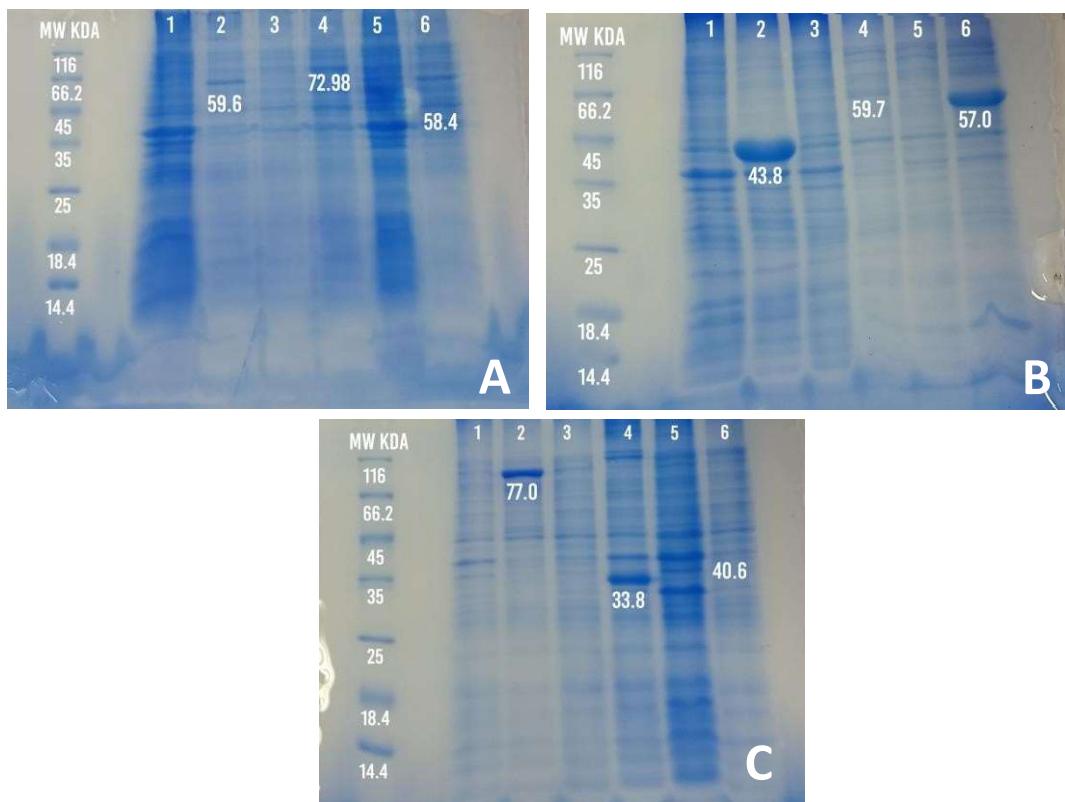
As células de BL21(DE3) foram cultivadas em meio LB líquido contendo carbenicilina 200 µg/mL a 37 °C, por 3 horas e sob agitação orbital de 180 rpm. Metade da suspensão de células (1,5 mL) foi induzida com IPTG (0,4 mM) por 3 horas e sob agitação orbital de 180 rpm. Foram utilizadas 4 temperaturas de incubação diferentes, 20 °C, 25 °C, 30 °C e 35°C, a fim de avaliar se há mudança na produção de proteínas recombinantes na fração solúvel. O restante da suspensão não foi induzido com o objetivo de avaliar se as proteínas são, realmente, recombinantes ou se são nativas de *E. coli*. As células foram lisadas e as proteínas resultantes da lise foram submetidas à eletroforese em gel de poliacrilamida 15% em condições desnaturantes (SDS-PAGE). Foram produzidos géis mostrando, para cada temperatura de indução, tanto as proteínas recombinantes em fração solúvel (Figura 6, Figura 8, Figura 10 e Figura 12) quanto em fração insolúvel (Figura 7, Figura 9, Figura 11 e Figura 13).

Figura 6 - Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 20 °C (fração solúvel).



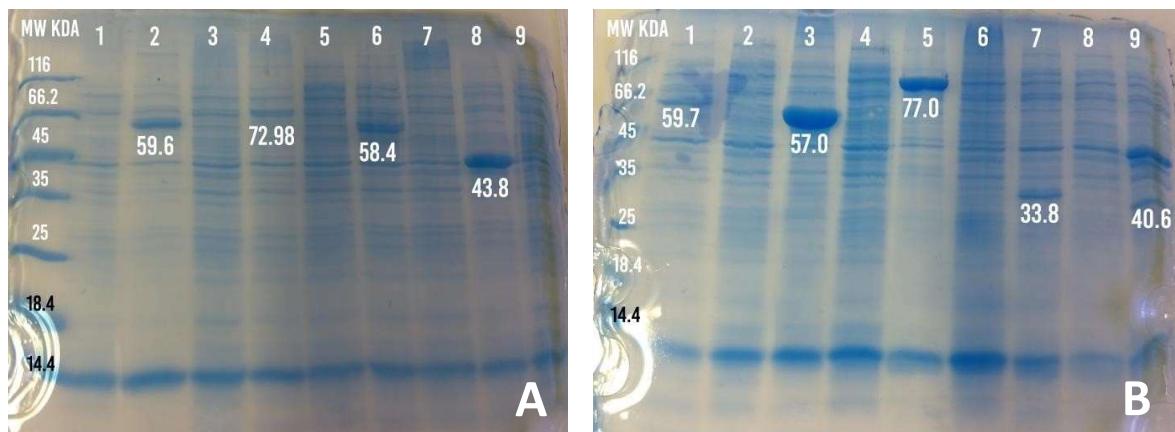
(A) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 20 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: *CV2592* (não induzidas). Poço 2: *CV2592* (induzidas). Poço 3: *CV0259* (não induzidas). Poço 4: *CV0259* (induzidas). Poço 5: *CV0554* (não induzidas). Poço 6: *CV0554* (induzidas). Poço 7: *CV0096* (não induzidas). Poço 8: *CV0096* (induzidas). Poço 9: *CV3323* (não induzidas). (B) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 20 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: *CV3323* (induzidas). Poço 2: *CV2065* (não induzidas). Poço 3: *CV2065* (induzidas). Poço 4: *CV0348* (não induzidas). Poço 5: *CV0348* (induzidas). Poço 6: *CV2641* (não induzidas). Poço 7: *CV2641* (induzidas). Poço 8: *CV2601* (não induzidas). Poço 9: *CV2601* (induzidas). MW: marcador LMW (Low Molecular Weight-GE Healthcare). Fonte: elaborada pela autora.

Figura 7 - Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 20 °C (fração insolúvel).



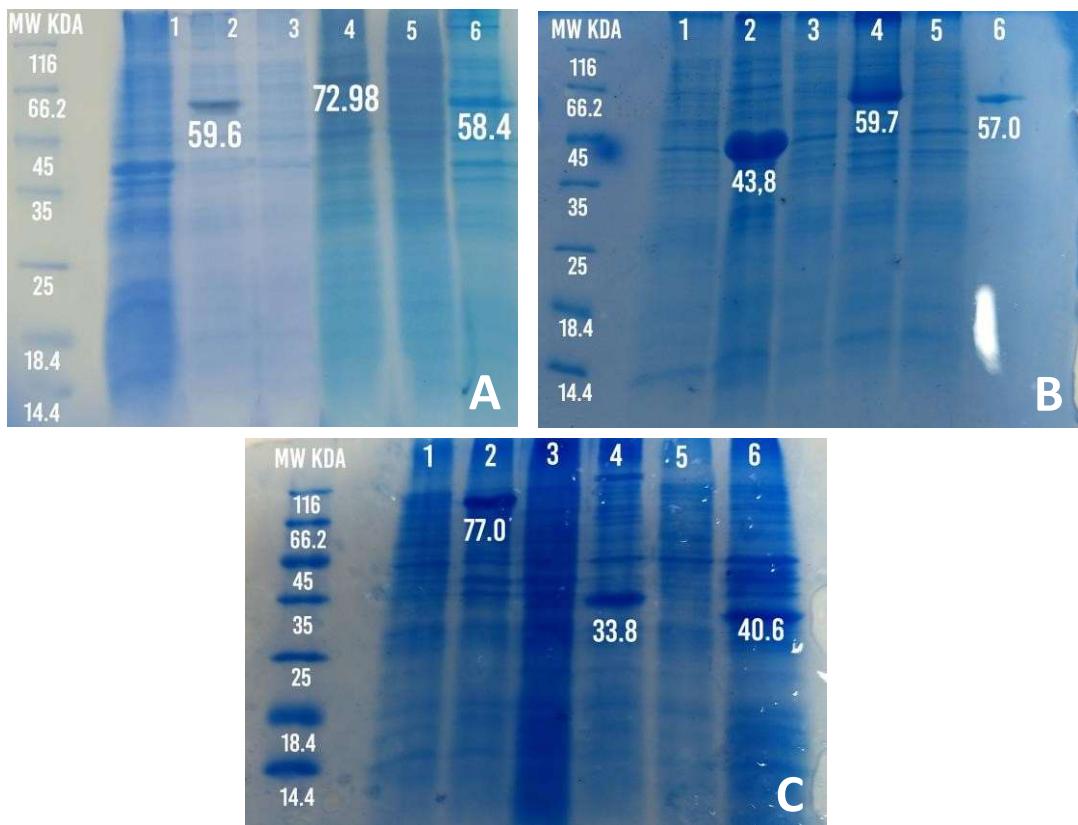
Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 20 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV2592 (não induzidas). Poço 2: CV2592 (induzidas). Poço 3: CV0259 (não induzidas). Poço 4: CV0259 (induzidas). Poço 5: CV0554 (não induzidas). Poço 6: CV0554 (induzidas). (B) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 20 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV0096 (não induzidas). Poço 2: CV0096 (induzidas). Poço 3: CV3323 (não induzidas). Poço 4: CV3323 (induzidas). Poço 5: CV2065 (não induzidas). Poço 6: CV2065 (induzidas). (C) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 20 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV0348 (não induzidas). Poço 2: CV0348 (induzidas). Poço 3: CV2641 (não induzidas). Poço 4: CV2641 (induzidas). Poço 5: CV2601 (não induzidas). Poço 6: CV2601 (induzidas). MW: marcador LMW (Low Molecular Weight-GE Healthcare). Fonte: elaborada pela autora.

Figura 8 - Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 25 °C (fração solúvel).



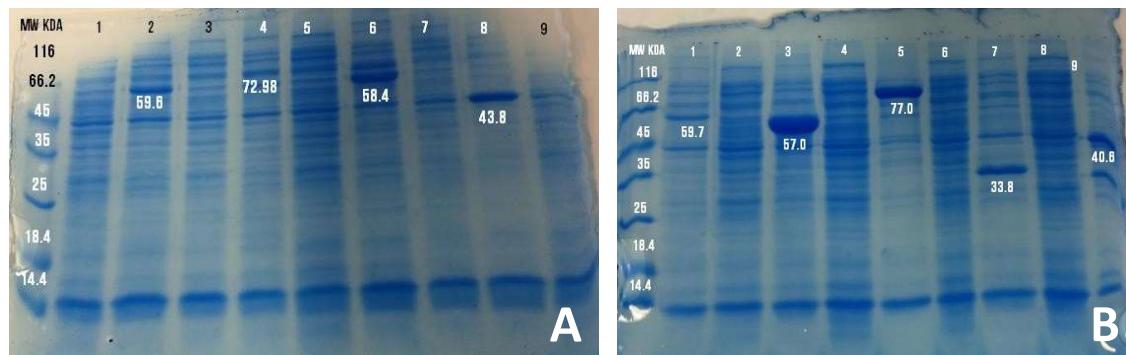
(A) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 25 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV2592 (não induzidas). Poço 2: CV2592 (induzidas). Poço 3: CV0259 (não induzidas). Poço 4: CV0259 (induzidas). Poço 5: CV0554 (não induzidas). Poço 6: CV0554 (induzidas). Poço 7: CV0096 (não induzidas). Poço 8: CV0096 (induzidas). Poço 9: CV3323 (não induzidas). (B) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 25 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: (induzidas). Poço 2: fração solúvel de CV2065 (não induzidas). Poço 3: CV2065 (induzidas). Poço 4: CV0348 (não induzidas). Poço 5: CV0348 (induzidas). Poço 6: CV2641 (não induzidas). Poço 7: CV2641 (induzidas). Poço 8: CV2601 (não induzidas). Poço 9: CV2601 (induzidas). MW: marcador LMW (Low Molecular Weight-GE Healthcare). Fonte: elaborada pela autora.

Figura 9 - Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 25 °C (fração insolúvel).



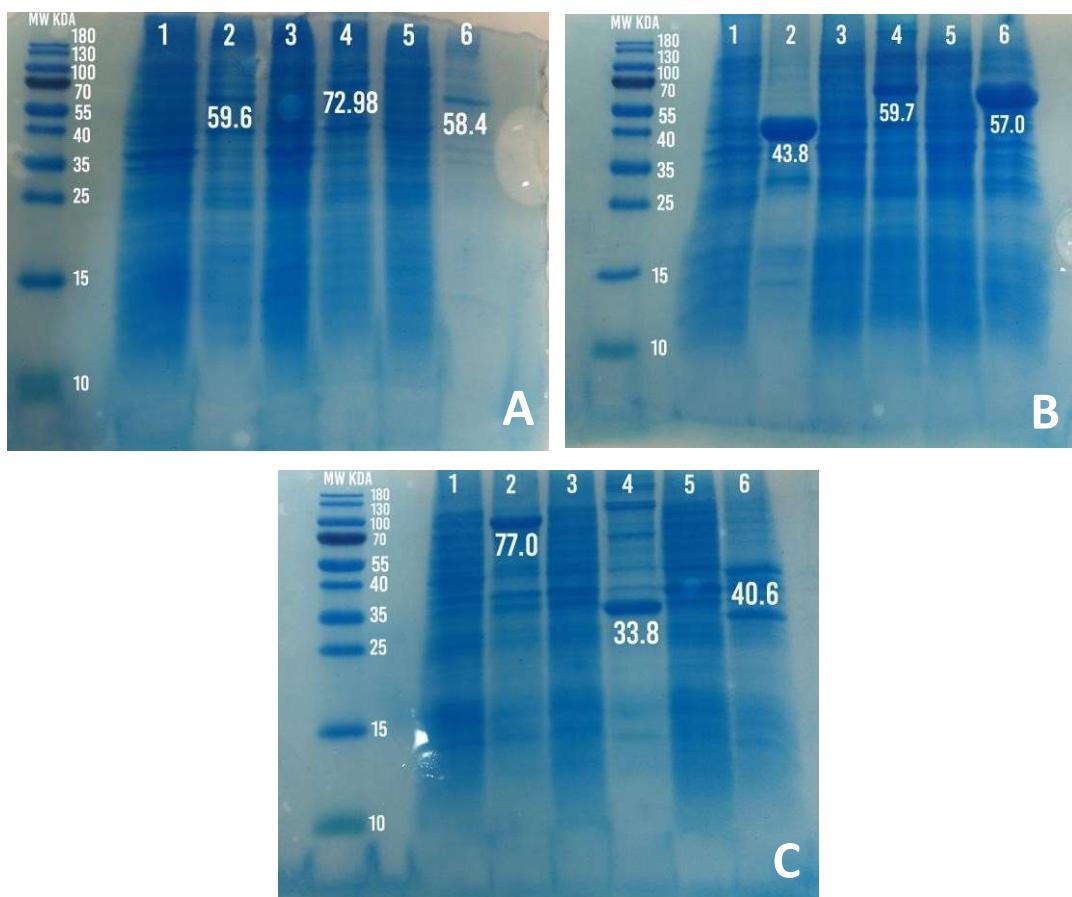
(A) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 25 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: de *CV2592* (não induzidas). Poço 2: *CV2592* (induzidas). Poço 3: células para *CV0259* (não induzidas). Poço 4: insolúvel *CV0259* (induzidas). Poço 5: insolúvel *CV0554* (não induzidas). Poço 6: insolúvel *CV0554* (induzidas). (B) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 25 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: insolúvel de *CV0096* (não induzidas). Poço 2: insolúvel *CV0096* (induzidas). Poço 3: células para *CV3323* (não induzidas). Poço 4: insolúvel *CV3323* (induzidas). Poço 5: insolúvel *CV2065* (não induzidas). Poço 6: *CV2065* (induzidas). (C) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 25 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: fração insolúvel de *CV0348* (não induzidas). Poço 2: *CV0348* (induzidas). Poço 3: *CV2641* (não induzidas). Poço 4: *CV2641* (induzidas). Poço 5: *CV2601* (não induzidas). Poço 6: *CV2601* (induzidas). MW: marcador LMW (Low Molecular Weight-GE Healthcare). Fonte: elaborada pela autora.

Figura 10 - Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 30 °C (fração solúvel).



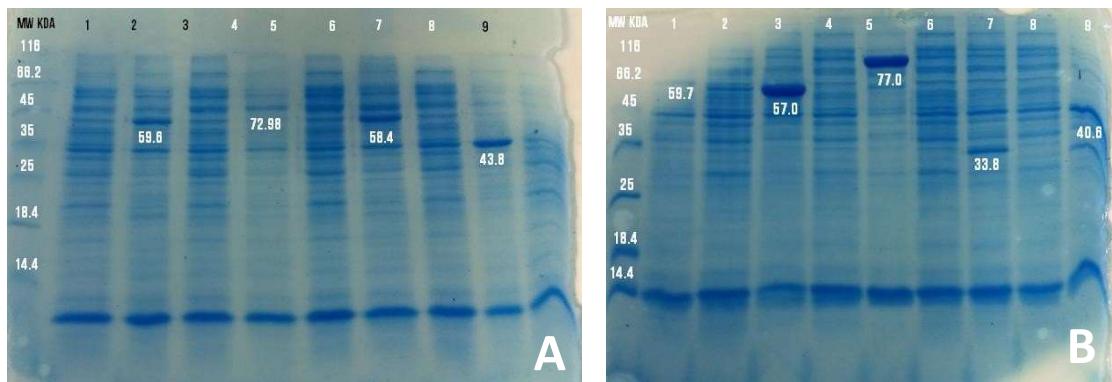
A) Células de BL21(DE3), transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 30 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV2592 (não induzidas). Poço 2: CV2592 (induzidas). Poço 3: CV0259 (não induzidas). Poço 4: CV0259 (induzidas). Poço 5: CV0554 (não induzidas). Poço 6: de CV0554 induzidas). Poço 7: CV0096 (não induzidas). Poço 8: CV0096 (induzidas). Poço 9: fração CV3323 (não induzidas). (B) Células de BL21(DE3), transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 30 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV3323 (induzidas). Poço 2: CV2065 (não induzidas). Poço 3: CV2065 (induzidas). Poço 4: CV0348 (não induzidas). Poço 5: CV0348 (induzidas). Poço 6: CV2641 (não induzidas). Poço 7: CV2641 (induzidas). Poço 8: CV2601 (não induzidas). Poço 9: CV2601 (induzidas). MW: marcador LMW (Low Molecular Weight-GE Healthcare). Fonte: elaborada pela autora

Figura 11 - Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 30 °C (fração insolúvel).



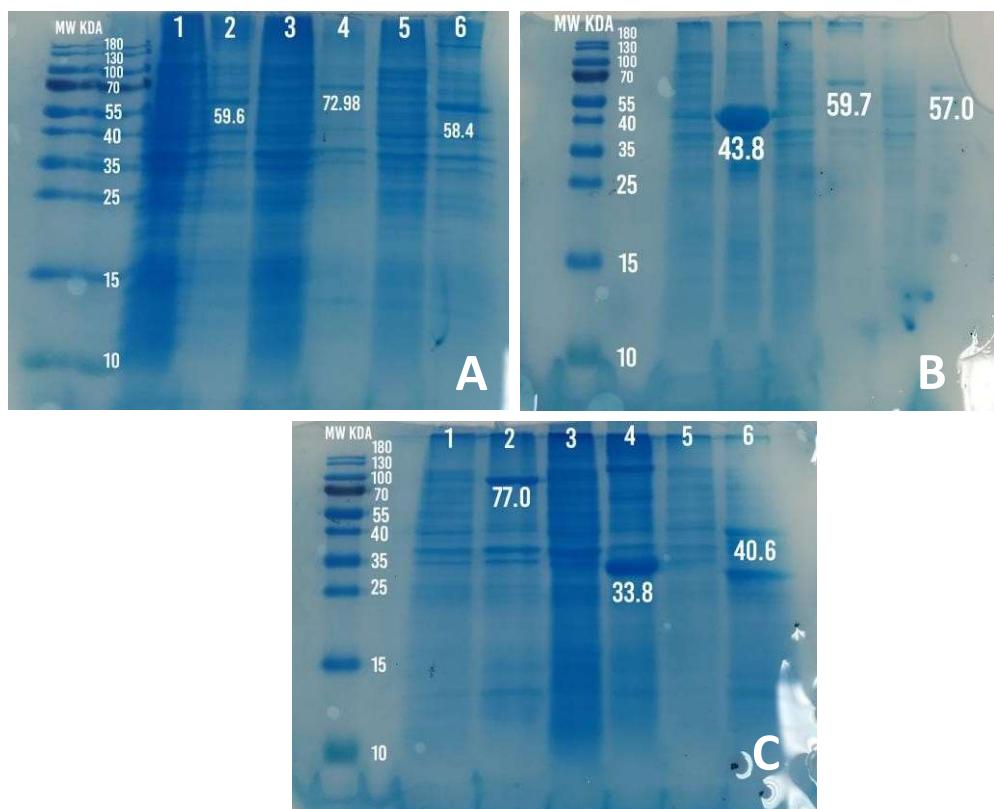
(A) Células de BL21(DE3), transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 30 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV2592 (não induzidas). Poço 2: CV2592 (induzidas). Poço 3: CV0259 (não induzidas). Poço 4: CV0259 (induzidas). Poço 5: CV0554 (não induzidas). Poço 6: CV0554 (induzidas). (B) Células de BL21(DE3), transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 30 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV0096 (não induzidas). Poço 2: CV0096 (induzidas). Poço 3: CV3323 (não induzidas). Poço 4: CV3323 (induzidas). Poço 5: CV2065 (não induzidas). Poço 6: CV2065 (induzidas). (C) Células de BL21(DE3), transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 30 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV0348 (não induzidas). Poço 2: CV0348 (induzidas). Poço 3: CV2641 (não induzidas). Poço 4: CV2641 (induzidas). Poço 5: CV2601 (não induzidas). Poço 6: CV2601 (induzidas). MW: marcador Page RulerTM Prestained Protein Ladder (Thermo Fisher Scientific). Fonte: elaborada pela autora.

Figura 12 - Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 35 °C (fração solúvel).



(A) Células de BL21(DE3), transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 35° C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV2592 (não induzidas). Poço 2: CV2592 (induzidas). Poço 3: CV0259 (não induzidas). Poço 4: CV0259 (induzidas). Poço 5: CV0554 (não induzidas). Poço 6: CV0554 (induzidas). Poço 7: CV0096 (não induzidas). Poço 8: CV0096 (induzidas). Poço 9: CV3323 (não induzidas). (B) Células de BL21(DE3), transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 35 ° C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: CV3323 (induzidas). Poço 2: CV2065 (não induzidas). Poço 3: CV2065 (induzidas). Poço 4: CV0348 (não induzidas). Poço 5: CV0348 (induzidas). Poço 6: CV2641 (não induzidas). Poço 7: CV2641 (induzidas). Poço 8: CV2601 (não induzidas). Poço 9: CV2601 (induzidas). MW: marcador LMW (Low Molecular Weight-GE Healthcare). Fonte: elaborada pela autora.

Figura 13 - Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para indução a 35 °C (fração insolúvel).



(A) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 35 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: *CV2592* (não induzidas). Poço 2: *CV2592* (induzidas). Poço 3: *CV0259* (não induzidas). Poço 4: *CV0259* (induzidas). Poço 5: *CV0554* (não induzidas). Poço 6: *CV0554* (induzidas). (B) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 35 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: *CV0096* (não induzidas). Poço 2: *CV0096* (induzidas). Poço 3: *CV3323* (não induzidas). Poço 4: *CV3323* (induzidas). Poço 5: *CV2065* (não induzidas). Poço 6: *CV2065* (induzidas). (C) Células de *BL21(DE3)*, transformadas com plasmídeos sintéticos, não induzidas e induzidas para a produção de proteínas com IPTG a 35 °C. O peso molecular identificando as proteínas é mostrado em cada banda. Poço 1: *CV0348* (não induzidas). Poço 2: *CV0348* (induzidas). Poço 3: *c CV2641* (não induzidas). Poço 4: *CV2641* (induzidas). Poço 5: *CV2601* (não induzidas). Poço 6: *CV2601* (induzidas). MW: marcador Page Ruler™ Prestained Protein Ladder (Thermo Fisher Scientific). Fonte: elaborada pela autora.

A expressão das proteínas por células de *E. coli* BL21(DE3) foi evidenciada ao realizar induções em pequena escala (piloto). A realização desses experimentos, em diferentes temperaturas, (20 °C, 25 °C, 30 °C e 35 °C) também forneceu resultados satisfatórios, os quais serão utilizados, futuramente, para a otimização da produção e purificação dessas proteínas em escala maior. A partir da análise dos géis de poliacrilamida, pode-se inferir, tanto que há a produção de proteínas, quanto qual temperatura de indução entre as quatro analisadas permitiu a produção mais acentuada, na forma solúvel, de cada uma delas. Os valores para cada temperatura estão expostos na Tabela 5.

Levando em consideração esses dados, percebe-se que temperatura de 30 °C provocou uma maior produção para a maioria das proteínas estudadas (CV2592, CV0554, CV3323, CV2065, CV0348, CV2641 e CV2601). As temperaturas de 20 °C e 25 °C permitiram uma produção mais acentuada das proteínas CV0259 e CV0096, respectivamente. Supõe-se que a maior parte das proteínas não teve bons resultados com a temperatura de 20 °C por essa temperatura ser muito abaixo da utilizada para o crescimento de células de *E. coli* BL21(DE3), a qual é cerca de 27 °C. Os resultados para a produção de proteínas, sob indução a 35 °C, foram similares aos obtidos para a temperatura de 30 °C, apesar dos géis para essa temperatura apresentarem bandas levemente menos intensas. Portanto, percebe-se que houve a produção, em menor escala, de proteínas recombinantes em fração solúvel, apesar de a produção de forma insolúvel ainda estar presente. Esses dados são coerentes com as análises *in silico* realizadas inicialmente, visto que todas as proteínas são solúveis.

Tabela 5 - Temperaturas de indução adequadas para a produção heteróloga, em células de *E. coli* BL21(DE3), de proteínas recombinantes de *C. violaceum* ATCC 12472.

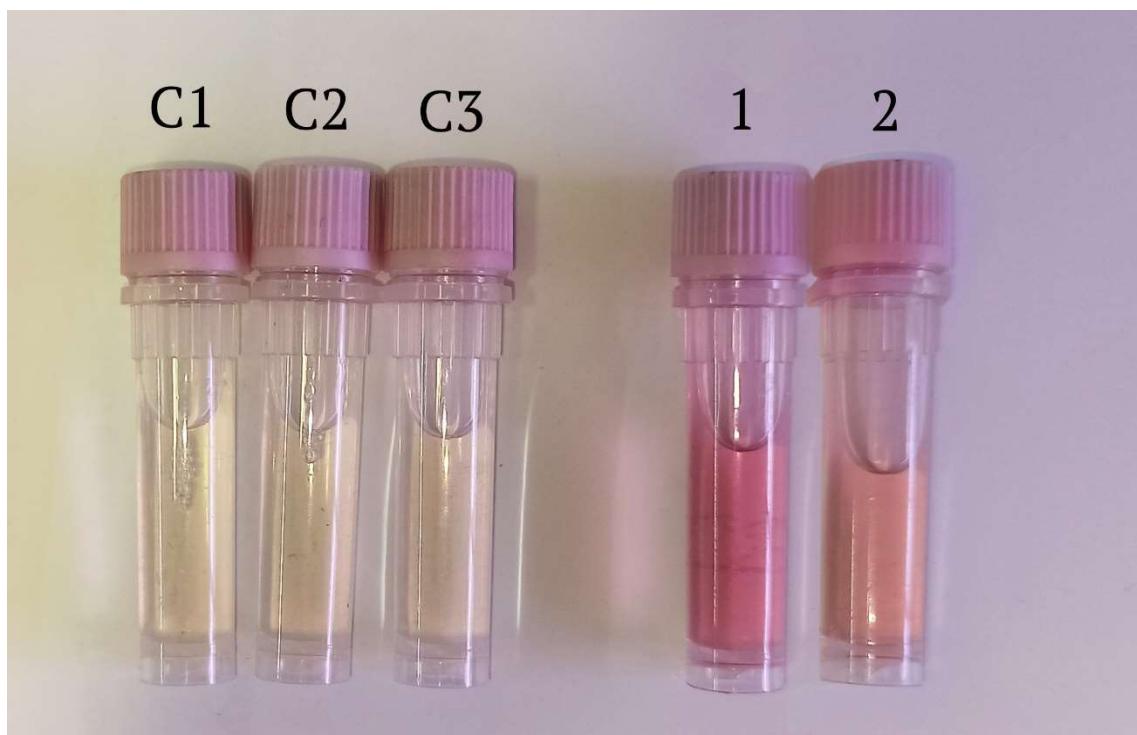
Inserto (ORF)	Temperatura de Indução
CV2592	30 °C
CV0259	20 °C
CV0554	30 °C
CV0096	25 °C
CV3323	30 °C
CV2065	30 °C
CV0348	30 °C
CV2641	30 °C
CV2601	30 °C

Fonte: elaborada pela autora

7.4 Teste para atividades quitinoclástica e quitosanásica

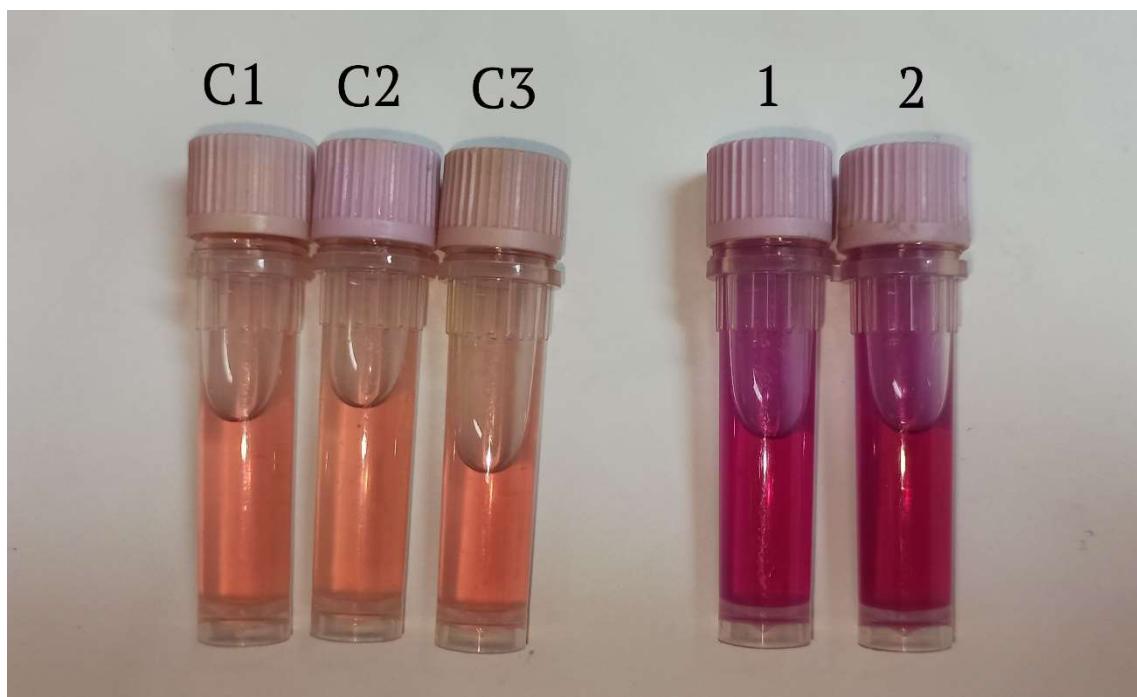
Os extratos proteicos das proteínas estudadas foram submetidos a testes, a fim de avaliar a presença de atividade para degradação de quitina ou quitosana. Nenhum dos extratos possuiu atividade quitosanásica. Entretanto, duas proteínas apresentaram coloração rosada ao serem testadas para atividade quitinoclástica, tanto utilizando quitina coloidal, quanto insolúvel. Uma delas foi a proteína CV2065 e a outra foi a proteína CV0259, ambas da família de hidrolases de glicosídeo 3 (GH3). A partir da observação visual, percebe-se que as duas proteínas em questão podem degradar quitina. As figuras 14 e 15 mostram o resultado para os testes envolvendo quitina coloidal e quitina insolúvel, respectivamente. A coloração rosa mais forte, observada na atividade contra quitina insolúvel, deve-se ao fato de os tubos ficarem incubados por mais tempo em relação à atividade realizada contra quitina coloidal (24 h e 1 h, respectivamente). Desse modo, pode-se concluir que a atividade quitinoclástica é proporcional ao tempo de contato da enzima contra seu substrato.

Figura 14 – Resultado do teste de atividade quitinoclástica de CV2065 e CV0259 contra quitina coloidal.



C1: controle contendo água destilada. C2: controle contendo tampão acetato de sódio 50 mM pH 5,2. C3: controle contendo extrato de proteínas produzido por células de BL21(DE3) não transformadas. 1: resultado da atividade quitinoclástica contra quitina coloidal para CV0259. 2: resultado da atividade quitinoclástica contra quitina coloidal para CV2065. Fonte: elaborada pela autora.

Figura 15 – Resultado do teste de atividade quitinoclástica de CV2065 e CV0259 contra quitina insolúvel.

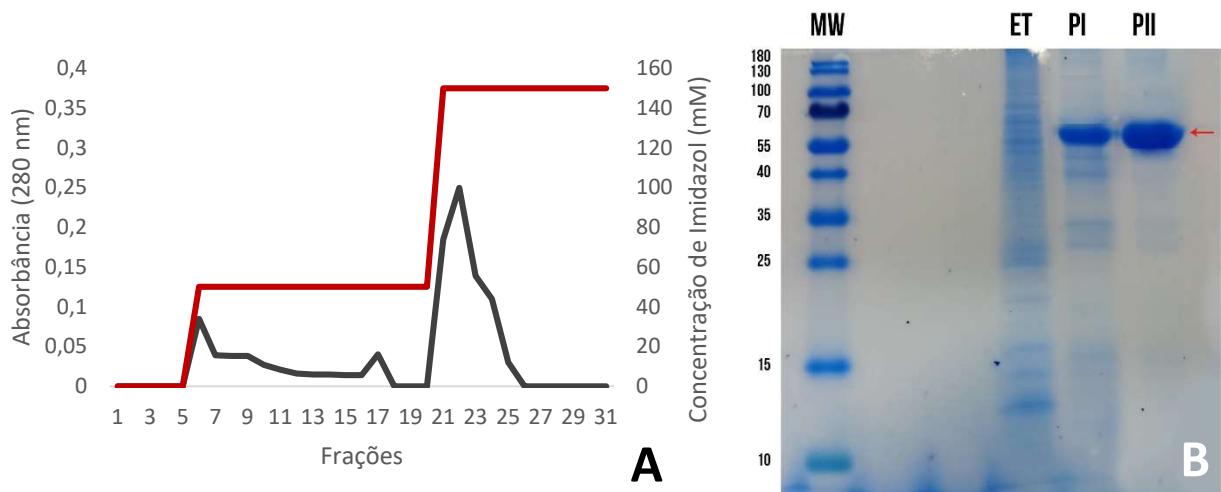


C1: controle contendo água destilada. C2: controle contendo tampão acetato de sódio 50 mM pH 5,2. C3: controle contendo extrato de proteínas produzido por células de BL21(DE3) não transformadas. 1: resultado da atividade quitinoclástica contra quitina insolúvel para CV0259. 2: resultado da atividade quitinoclástica contra quitina insolúvel para CV2065. Fonte: elaborada pela autora.

8.5 Produção e purificação de uma proteína recombinante

A proteína CV2065, uma hidrolase de glicosídeo pertencente à família GH3, foi escolhida para ser produzida em uma escala maior por possuir a maior produção aparente na fração solúvel. As células de BL21(DE3) contendo o plasmídeo para a produção dessa hidrolase, foram cultivadas em meio LB líquido contendo carbenicilina 200 µg/mL a 37 °C, por 3 horas e sob agitação orbital de 180 rpm. A suspensão de células foi induzida com IPTG (0,4 mM) também por 3 horas e sob agitação orbital de 180 rpm a uma temperatura de 30° C (temperatura escolhida por produzir melhores resultados para a produção da proteína em questão). As células foram lisadas e as proteínas, resultantes da lise, foram submetidas à cromatografia de afinidade por íons metálicos, utilizando uma resina de cobalto, HisPur™ Cobalt Resin (Thermo Fisher Scientific). Após o processo de cromatografia e de diálise das amostras em tampão Tris-HCl 50 mM pH 8,0, NaCl 500 mM e, os picos obtidos foram analisados por eletroforese em gel de poliacrilamida (SDS-PAGE). Através da análise do gel, é possível perceber que a proteína CV2065 foi eluída com o tampão contendo imidazol na concentração de 150 mM (Figura 16). Além disso, pode-se constatar que a proteína foi purificada, mas que pode ser necessária mais uma cromatografia, com a finalidade de purificá-la de forma mais sofisticada.

Figura 16 - Perfil da cromatografia de afinidade por íons metálicos utilizando uma resina de cobalto, HisPur™.

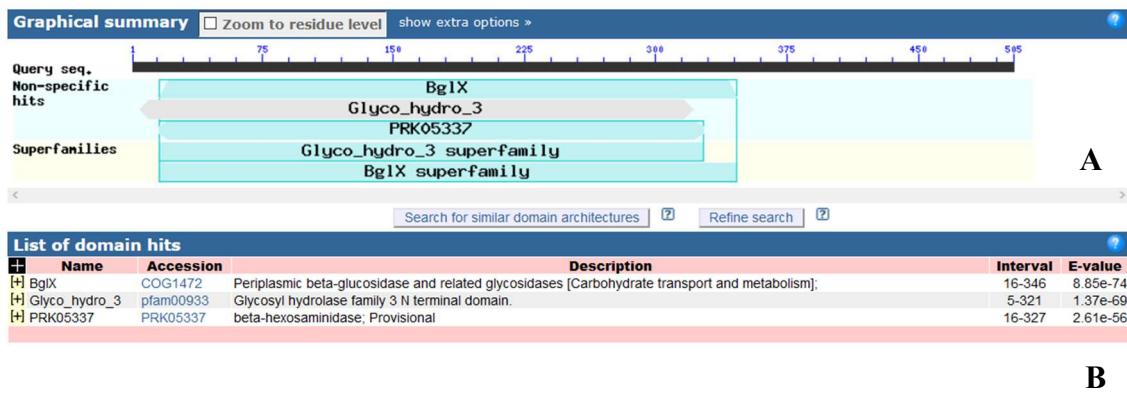


(A) A cromatografia permitiu que fossem eluidos dois picos, um com imidazol 10 mM e outro com imidazol 150 mM. A proteína foi eluida no pico retido 2. (B) Eletroforese em gel de poliacrilamida (15%) em condições desnaturantes (SDS-PAGE) para purificação por cromatografia de CV2065. São observados dois picos, um eluido com o tampão de imidazol 50 mM (PI) e outro eluido com tampão de imidazol 150 mM (PII). A proteína foi eluida no pico II (marcação vermelha). MW: marcador Page Ruler™ Prestained Protein Ladder (Thermo Fisher Scientific); ET: Extrato Total; PNR: Pico Não Retido; PI: Pico retido 1; PII: Pico Retido 2. Fonte: Elaborada pela autora.

8.6 Avaliação dos domínios das proteínas com atividade enzimática

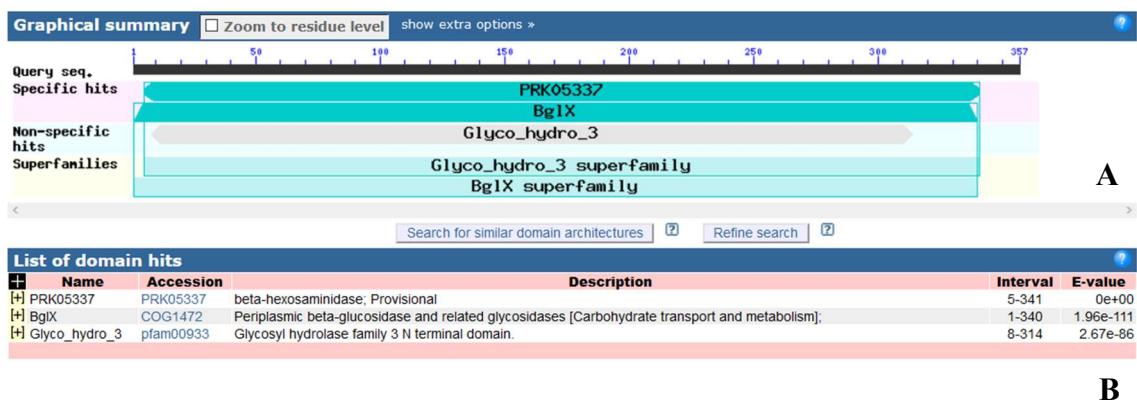
Para obter informações sobre os domínios proteicos das proteínas CV0259 e CV2065, as sequências de aminoácidos para cada uma delas foram submetidas ao CDD. Os resultados obtidos, para cada proteína, estão demonstrados nas figuras 17 e 18. Além disso, um esquema dos domínios foi feito, utilizando o programa IBS, a fim de facilitar a visualização dos domínios de cada proteína. O esquema para a proteína CV0259 está exposto na figura 19 e o esquema para a proteína CV2065 na figura 20.

Figura 17 – Domínios proteicos, obtidos a partir de análise no CDD para a proteína CV0259, da família GH3.



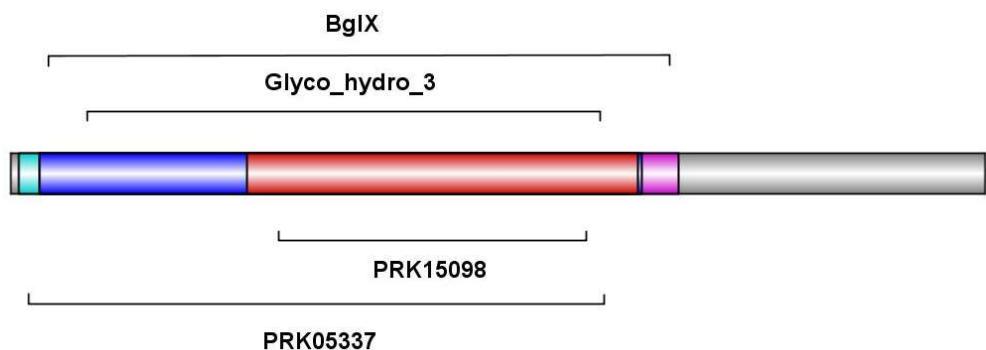
(A) Representação da posição dos domínios proteicos na sequência de aminoácidos da proteína CV0259. (B) Lista com os nomes detalhados dos domínios presentes na proteína CV0259. A lista mostra também o número de acesso do domínio, o intervalo e aminoácidos no qual cada domínio está presente e o E-value (score de confiabilidade do alinhamento). Fonte: Retirado do Banco de Domínios Conservados (CDD) (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>).

Figura 18 – Domínios proteicos, obtidos a partir de análise no CDD para a proteína CV2065, da família GH3.



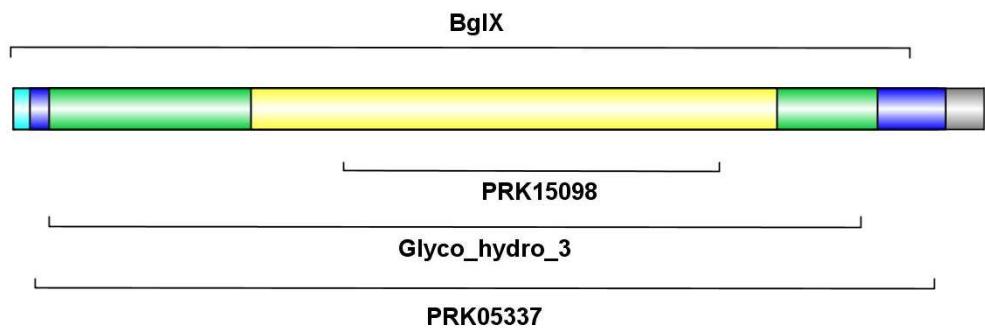
(A) Representação da posição dos domínios proteicos na sequência de aminoácidos da proteína CV2065. (B) Lista com os nomes detalhados dos domínios presentes na proteína CV2065. A lista mostra também o número de acesso do domínio, o intervalo e aminoácidos no qual cada domínio está presente e o E-value (score de confiabilidade do alinhamento). Fonte: Retirado do Banco de Domínios Conservados (CDD) (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>).

Figura 19 – Representação gráfica da arquitetura de domínios a proteína CV0259 utilizando o programa IBS.



Azul claro: PRK05337. Azul escuro: Glyco_Hydro_3. Rosa: BgIX. Vermelho: PRK15098. Fonte: Elaborada pela autora utilizando o programa IBS.

Figura 20 – Representação gráfica da arquitetura de domínios a proteína CV2065 utilizando o programa IBS.



Amarelo: PRK15098. Azul claro: BgIX. Azul escuro: PRK05337. Verde: Glyco_Hydro_3. Fonte: Elaborada pela autora utilizando o programa IBS.

Ao analisar os resultados obtidos, percebe-se que as proteínas, apesar de possuírem uma quantidade de aminoácidos diferente, possuem os mesmos domínios proteicos, o que vai ao encontro da informação de que ambas fazem parte da mesma família proteica. A família de hidrolases de glicosídeo 3, ou GH3, agrupa β -D-glicosidases, α -L-arabinofuranosidases, β -D-xilopiranosidases, N-acetil-beta-D-glicosaminidases, e N-acetil- β -D-glicosamina fosforilases. A maioria dessas proteínas está amplamente distribuída em bactérias, fungos e plantas, estando envolvidas, majoritariamente, na degradação de biomassa, na remodelagem de parede celulares e na proteção dos indivíduos contra patógenos (HARVEY, 2000; MACDONALD, 2015). As N-acetil- β -D-glicosaminidases, em especial, possuem atividades na degradação de quitina e na reciclagem de peptidoglicanos (ARA, 2020). Bactérias que possuem enzimas quitinoclásticas são as maiores responsáveis pela degradação de quitina na natureza, seja no solo, onde as taxas de hidrólise de quitina podem ser relacionadas com a presença de bactérias, ou em ambientes aquáticos. Em ambos os casos, a degradação promovida por enzimas de bactérias contribui para tornar a quitina biologicamente utilizável, contribuindo para a ciclagem de nutrientes (BEIER, 2013; SOUZA, 2011). A partir dessas informações, pode-se perceber o motivo pelas proteínas CV0259 e CV2065, ambas da família GH3, possuírem atividade de degradação de quitina.

A proteína CV0259 é uma possível hidrolase de glicosídeo e a CV2065 é uma β -hexosaminidase (ou N-acetil-B-D-glicosaminidase). Como foi dito anteriormente, ambas as proteínas possuem os mesmos domínios proteicos. O domínio BgIX, presente no intervalo 16-346 em CV0259 e no intervalo 5-341 em CV2065, corresponde a uma β -glicosidases periplasmática. Em *E. coli*, BgIX pode ajudar a bactéria a sobreviver através da hidrólise de açúcares, a fim de obter glicose. Além disso, pode ser utilizada como celulase ou para degradação de peptidoglicano e outros polissacarídeos complexos (YANG, 1996). Outro domínio, presente em CV0259 (no intervalo 5-321 aminoácidos) e em CV2065 (no intervalo 8-314) é glyco_hydro_3, correspondente ao domínio terminal da família GH3. Esse domínio, como o próprio nome diz, é componente de todos os membros de hidrolases de glicosídeo 3 e está relacionado à hidrólise de ligações glicosídicas. O último domínio, presente nos intervalos 16-327 e 5-341 nas proteínas CV0259 e CV2065, respectivamente, é o PRK05337, correspondente a uma hexosaminidase. Como já foi dito, hexosaminidases são enzimas envolvidas na catálise de resíduos N-acetil-D-hexosamina terminais (frutosamina, galactosamina, glicosamina e manosamina) (CABEZAS, 1989).

Através dos estudos domínios das proteínas que obtiveram atividade enzimática para a degradação de quitina, ou seja, CV0259 e CV2065, observa-se que essas possuem domínios que justificam sua atividade quitinoclástica. Porém, esses domínios não são específicos para

quitina, de modo que as proteínas em questão podem participar da degradação de outros polissacarídeos complexos, como celulose e petidoglicano.

8.7 Construção e avaliação modelos 3D para proteínas com atividade enzimática

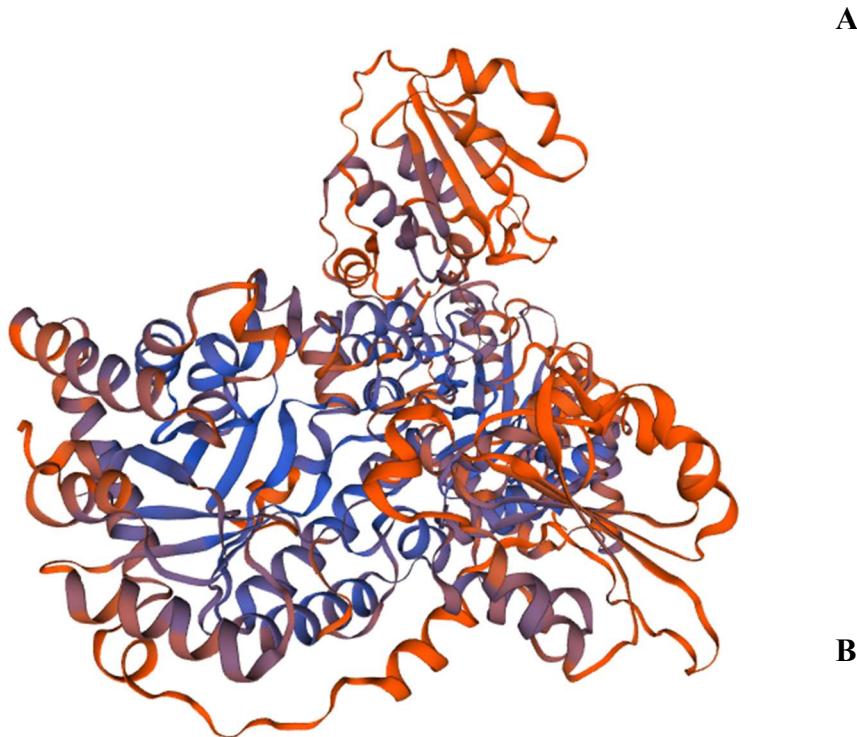
Como foi dito anteriormente, a modelagem de proteínas por homologia é mais fácil de ser utilizada por permitir que uma proteína alvo seja comparada com sequências presentes em diferentes bancos de dados. Para prever a estrutura 3D das proteínas que obtiveram algum tipo de atividade enzimática, ou seja, CV0259 e CV2065, houve a necessidade de obter, primariamente, modelos ou templates com alta semelhança com as proteínas estudadas (representada, neste caso, por valores baixos de *E-value* e identidade entre sequências).

Para obter os modelos que serviram como base para a predição das estruturas proteicas das proteínas estudadas, as sequências dessas foram submetidas a dois servidores diferentes, BLASTP e Swiss-Model. Um bom modelo ou *template* deve possuir o menor *E-value* possível, além de possuir mais de 30% de semelhança com a proteína alvo, este último valor sendo fornecido pelo BLASTP. Desse modo, para a proteína CV0259, o modelo 4ZM6 foi o que obteve maior semelhança e para a proteína CV2065, o modelo com maior semelhança foi o 6JT1. O resultado foi obtido em ambos os servidores.

O servidor Swiss-Model permite que a estrutura da proteína alvo seja construída a partir dos templates. O alinhamento das sequências das proteínas alvo com os templates, bem como as estruturas preditas para cada proteína estão representados nas figuras 21 e 22. A proteína CV0259 obteve uma semelhança de cerca de 35,02% com o modelo 4ZM6 e a proteína CV2065 obteve uma semelhança de 61,13%.

Figura 21 – Resultados obtidos para a proteína CV0259 após a submissão da mesma ao servidor SWISS-MODEL.

Model_01:A	MTLSDERLDQLAGRC ^{LMV} DVAGPVLTEEEAARLQRMVR ^{RAVC} I>FRRNVPDAASTRRLVADLKCAL---GH	67
Model_01:B	MTLSDERLDQLAGRC ^{LMV} DVAGPVLTEEEAARLQRMVR ^{RAVC} I>FRRNVPDAASTRRLVADLKCAL---GH	67
4zm6.1.A	---NDDNLDK ^{EIGQILM} GFDGLEPTPGIIDIENHNLGSIIDFSRNIA ^{EKQVQKLTHSLQQIARNAGH}	71
Model_01:A	D--V ^{LIGID} QEGGAVMRT-RFLPQAPAAMALAAVGDEALARRVGGAVARGLKS ^{LGVNWDFAPVLDLN} NNP	134
Model_01:B	D--V ^{LIGID} QEGGAVMRT-RFLPQAPAAMALAAVGDEALARRVGGAVARGLKS ^{LGVNWDFAPVLDLN} NNP	134
4zm6.1.A	KRPLFIA ^{DQENG} VVRLGDSGYI ^{PG} MA ^L GA ^D GS ^S TAARNVAMAI ^S KE ^L LT ^I GMN ^W LA ^P LDVNNNP	141
Model_01:A	ANPVIGERSFGAD ^P ARA ^A ALARAWMEGH ^L AEGV ^I CCV ^{KH} D ^P GHGDT ^T DSHLDLP ^V DKP ^Q AELRDY ^E LA	204
Model_01:B	ANPVIGERSFGAD ^P ARA ^A ALARAWMEGH ^L AEGV ^I CCV ^{KH} D ^P GHGDT ^T DSHLDLP ^V DKP ^Q AELRDY ^E LA	204
4zm6.1.A	LNPVIGVR ^S YQ ^G D ^E ELVAR ^M GIA ^A Q ^V E ^G YQ ^B SKV ^I S ^T ^K H ^D PGHDT ^A TD ^S HLDV ^P INKT ^L EELD ^T E ^L	211
Model_01:A	PF-KALSAH ^A HPA---LMSAHI ^R PALDADW ^P ATL ^S P ^A IL ^T G ^L LR ^R E ^L G ^F FRG ^V AIT ^I D ^A L ^N M ^R A ^I R ^E WG ^O	269
Model_01:B	PF-KALSAH ^A HPA---LMSAHI ^R PALDADW ^P ATL ^S P ^A IL ^T G ^L LR ^R E ^L G ^F FRG ^V AIT ^I D ^A L ^N M ^R A ^I R ^E WG ^O	269
4zm6.1.A	PFKKAL ^E AGGIAC ^P T ^S M ^G H ^M ^D PH ^E IX ^D V ^S S ^I A ^P E ^I VR ^D LL ^R RF ^G Y ^K G ^V I ^I DC ^L E ^M DA ^V E ^T V ^G I	281
Model_01:A	PAGAVQ ^T TLK ^A GAD ^L ALV ^L Q ^F ADE ^M E ^A FS ^D AL ^R AA ^L R ^N G ^E LA ^Q AR ^L EE ^A AR ^V D ^A L ^I R ^R Y ^P S ^R N ^D C ^G Y ^P A ^E	339
Model_01:B	PAGAVQ ^T TLK ^A GAD ^L ALV ^L Q ^F ADE ^M E ^A FS ^D AL ^R AA ^L R ^N G ^E LA ^Q AR ^L EE ^A AR ^V D ^A L ^I R ^R Y ^P S ^R N ^D C ^G Y ^P A ^E	339
4zm6.1.A	PKGALM ^Q A ^G ND ^M A ^M S ^H TL ^A FQKDA ^F KV ^L LYS ^A Q ^E G ^Q L ^D K ^E I ^R Q ^S L ^Q R ^V A ^Q L ^K D ^Q F ^L N ^W D ^D V ^L Q ^A D	351
Model_01:A	RRA---ADE ^A LFAD ^F ADAM ^T R ^A ---R ^G E ^I P ^T LAAGQ ^R L ^R L ^V V ^Q D ^Q AP ^S D ⁻ GV ^S E ^E G ^L A ^D A ^L I ^A A ^L S ^R	400
Model_01:B	RRA---ADE ^A LFAD ^F ADAM ^T R ^A ---R ^G E ^I P ^T LAAGQ ^R L ^R L ^V V ^Q D ^Q AP ^S D ⁻ GV ^S E ^E G ^L A ^D A ^L I ^A A ^L S ^R	400
4zm6.1.A	L ^K I ^M G ^S EAH ^A T ^L SK ^E L ^I DR ^V E ^T V ^V I ^R KNT ^I L ^P I ^R F ^A Q ^T D ^K I ^L FL ^A H ^V P ^Q T ^L A ^D S ^E K ^E P ^F I ^N S ^F H ^A S ^L K	421
Model_01:A	RHA---LEVVR ^R Q ^E R ^A M ^L D ^W ST ^L LP ^Q D ^G L ^F TV ^L A ^S T ^T R ^E RY ^G ---Q ^R E ^{QQ} S ^W K ^P D ^L H ^L ^L L ^W N ^P Y ^A A ^D	461
Model_01:B	RHA---LEVVR ^R Q ^E R ^A M ^L D ^W ST ^L LP ^Q D ^G L ^F TV ^L A ^S T ^T R ^E RY ^G ---Q ^R E ^{QQ} S ^W K ^P D ^L H ^L ^L L ^W N ^P Y ^A A ^D	461
4zm6.1.A	R ^H T ^N L ^E Y ^I ^T NE ^E T ^P -D ^I S ^Q K ^I Q ^A D ^N V ^I C ^E A ^M A ^N I ^Y E ^F Q ^V R ^M V ^Q Q ^A Q ^K I ^A K ^R L ^V V ^A W ^M N ^P Y ^D Q ^M C ^F P	490
Model_01:A	I ^A A ^P A ^L I ^S Y ^G F ^A D ^A AL ^F A ^S R ^C L ^A G ^E I ^A A ^G R ^L P ^V A ^L G ^Q P ^Y Q ^E Q	505
Model_01:B	I ^A A ^P A ^L I ^S Y ^G F ^A D ^A AL ^F A ^S R ^C L ^A G ^E I ^A A ^G R ^L P ^V A ^L G ^Q P ^Y Q ^E Q	505
4zm6.1.A	Q ^V D ^T Y ^L T ^Y E ^Y T ^P P ^A H ^E A ^A V ^R L ^I F ^G E ^I E ^T R ^S R ^L P ^I S ^I -----	527

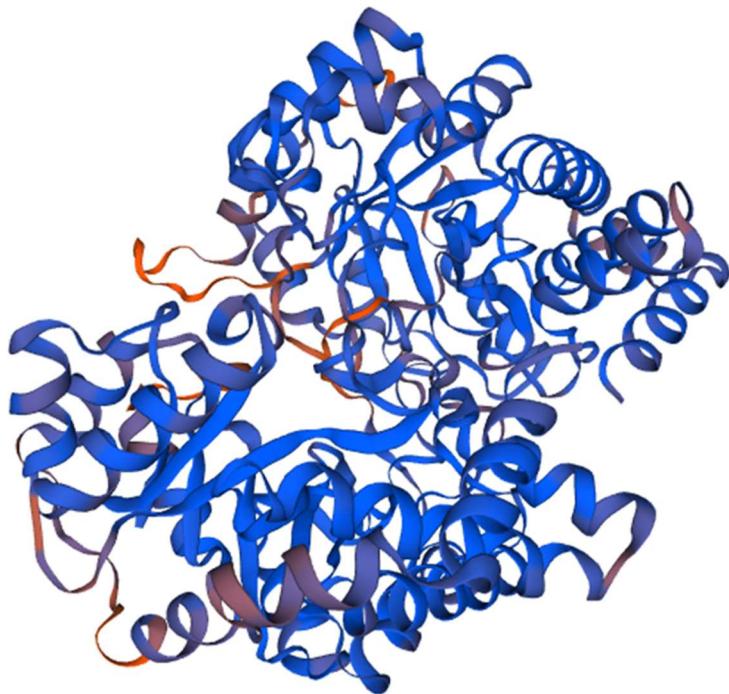


(A) Alinhamento entre as sequências de CV0259 e o template escolhido, 4ZM6. (B) Estrutura predita por homologia para a proteína CV0259. Fonte: Retirado de Swiss-Model (<https://swissmodel.expasy.org>).

Figura 22 – Resultados obtidos para a proteína CV2065 após a submissão da mesma ao servidor SWISS-MODEL.

Model_01:A	MNLPRGPVMV	DIAGFALTEAERARL	SHPLVGGV	I	FRRNFHN	IEQLRALTAEIRALRTPHL	LIAV	DHEGG	70		
Model_01:B	MNLPRGPVMV	DIAGFALTEAERARL	SHPLVGGV	I	FRRNFHN	IEQLRALTAEIRALRTPHL	LIAV	DHEGG	70		
6jti.1.A	-----	PVMADIAAFRLT	EEEKQRL	LDPAIGG	I	FRRNFQN	IEQLKILTAEIKALR	PEL	LIAV	DHEGG	109
Model_01:A	RV	DRFLDGFTRLPPM	RVLGEAWDADR	-DQALKLAETVGYVLAEL	SACGID	L	SFTP	LDLDWERCAVIGN	139		
Model_01:B	RV	DRFLDGFTRLPPM	RVLGEAWDADR	-DQALKLAETVGYVLAEL	SACGID	L	SFTP	LDLDWERCAVIGN	139		
6jti.1.A	RV	DRFIEGFTRLPAM	NVLGQIWWDKD	gASA	AAETAAGQVGRVLATE	SACGID	L	SFTP	LDLDWGNCAVIGN	179	
Model_01:A	RAFHRD	PEAVSALAEALQQGLGRGGMM	CGKH	PGHGYVEGDSHLMP	DDRTLAQIE	QD	DLPFARLAD	209			
Model_01:B	RAFHRD	PEAVSALAEALQQGLGRGGMM	CGKH	PGHGYVEGDSHLMP	DDRTLAQIE	QD	DLPFARLAD	209			
6jti.1.A	RSFHRN	PEAVARLALALQKGLA	GGMKS	CGKH	PGHGYVEGDSHLVL	I	DGR	LEDELE	RDLEPFRIMSE	249	
Model_01:A	AGMGAVMPAHV	L	OPAVDSQPAGE	SKVWLTDILRG	RIGFDGVIFSD	A	DMAGAAGACTYVQRADAAL	AAGC	279		
Model_01:B	AGMGAVMPAHV	L	OPAVDSQPAGE	SKVWLTDILRG	RIGFDGVIFSD	ALDMAGAAGACTYVQRADAAL	AAGC	279			
6jti.1.A	EGMAAVMPAHV	W	OPQVDTK	PEIWLQILRR	IGFKGVIFSD	DLTMEGACGAGG	EKERARISFE	AGC	319		
Model_01:A	DMV	VCNQPEEADAMLAALAPP	-PQPO	LAERLERMAGKSRA	AEDWQR	LIA	TPDFAAAQAAVRQ	DAMPKD	348		
Model_01:B	DMV	VCNQPEEADAMLAALAPP	-PQPO	LAERLERMAGKSRA	AEDWQR	LIA	TPDFAAAQAAVRQ	DAMPKD	348		
6jti.1.A	D	V	VCNR	FDLVDEL	RDGFTIPDNQ	EL	-AGR	WQ	HEAVQAVMOTMGFQAAQAFVAGLA	SPQDTA	388
Model_01:A	AGPQVGEAH								357		
Model_01:B	AGPQVGEAH								357		
6jti.1.A	GGVKVGEAF								397		

A



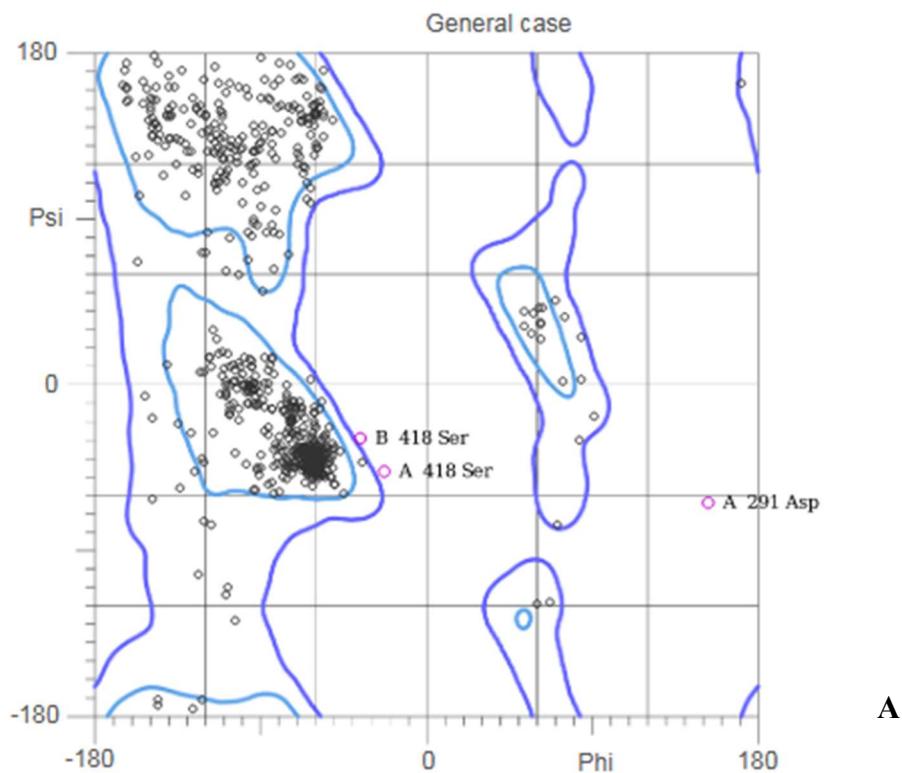
B

(A) Alinhamento entre as sequências de CV2065 e o template escolhido, 6JTI. (B) Estrutura predita por homologia para a proteína CV2065. Fonte: Retirado de Swiss-Model (<https://swissmodel.expasy.org>).

Em 1963, Gopalasamudram Narayan Ramachandran descreveu que a forma de dois aminoácidos vizinhos são fruto de dois ângulos de torção, denominados φ e ψ . Desse modo, algumas combinações de ângulos, correspondentes a estruturas secundárias de proteínas, são favorecidas em detrimento de outras (VERLI, 2014). O gráfico que mostra essas combinações e a distribuição dos aminoácidos em diferentes estruturas secundárias se chama gráfico de Ramachandran. Os gráficos para as proteínas CV0259 e CV2065 estão representados nas Figuras 23 e 24, respectivamente, obtidos a partir da submissão das estruturas no servidor MolProbity.

Observando os gráficos para as estruturas tridimensionais das proteínas alvo, percebe-se que ambas possuem desvios em ângulos de torção (CV0259 com 17 e CV2065 com 2). Ao submeter as estruturas aos servidores ERRAT, VERIFY-3D e ProSA-web, entretanto, percebeu-se que os modelos foram bem avaliados apesar dos desvios indicados pelo *plot* de Ramachandran (Tabelas 6 e 7). Desse modo, conclui-se que os modelos 3D obtidos foram bem preditos, sendo, provavelmente, o mais próximo da realidade possível.

Figura 23 – Plot de Ramachandran para a proteína CV0259, obtido a partir do servidor MolProbability.



92.7% (908/979) of all residues were in favored (98%) regions.
 98.3% (962/979) of all residues were in allowed ($\geq 99.8\%$) regions.

B 446 Lys (-80.1, -88.5)

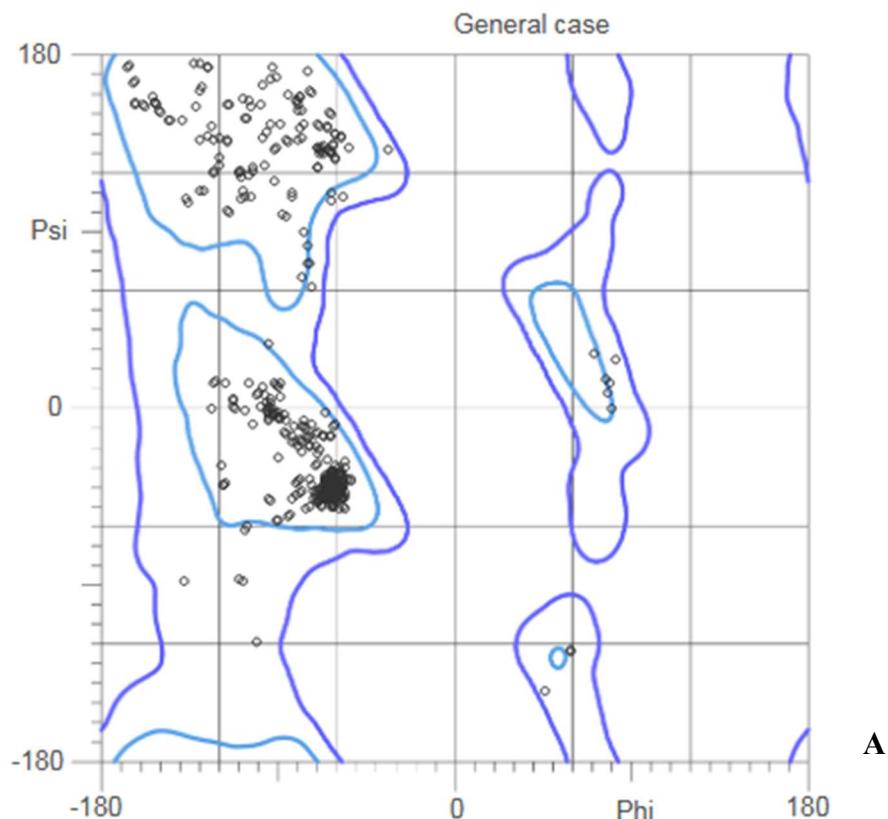
There were 17 outliers (phi, psi):

- | | |
|---------------------------|---------------------------|
| A 23 Pro (-110.3, 22.9) | A 418 Ser (-24.0, -47.8) |
| A 49 Pro (-93.5, -32.1) | A 447 Pro (-106.3, 127.3) |
| A 87 Pro (-108.3, 162.5) | B 49 Pro (-123.2, -31.6) |
| A 137 Pro (-96.3, -46.1) | B 137 Pro (-87.4, -38.2) |
| A 214 Pro (-111.0, 10.1) | B 139 Ile (76.1, 50.3) |
| A 291 Asp (153.4, -64.3) | B 214 Pro (-105.3, -12.3) |
| A 328 Tyr (-131.7, -49.6) | B 328 Tyr (-138.2, -64.1) |
| A 337 Pro (27.5, 146.3) | B 418 Ser (-36.2, -29.8) |

B

(A) Plot de Ramachandran para o caso geral da proteína CV0259, obtido da análise pelo MolProbability. (B) Desvios nos ângulos dos aminoácidos presentes na proteína CV0259. A imagem também mostra a procentagem de resíduos de aminoácidos presentes em regiões favoráveis ou não favoráveis. Fonte: Retirado de MolProbability (<http://molprobity.biochem.duke.edu/>).

Figura 24 – Plot de Ramachandran para a proteína CV2065, obtido a partir do servidor MolProbity.



95.4% (640/671) of all residues were in favored (98%) regions.
 99.7% (669/671) of all residues were in allowed (>99.8%) regions.

There were 2 outliers (phi, psi):

A 303 Gln (-102.1, 7.0)
 B 303 Gln (-94.6, 5.7)

B

(A) Plot de Ramachandran para o caso geral da proteína CV2065, obtido da análise pelo MolProbity. (B) Desvios nos ângulos dos aminoácidos presentes na proteína CV2065. A imagem também mostra a procentagem de resíduos de aminoácidos presentes em regiões favoráveis ou não favoráveis. Fonte: Retirado de MolProbity (<http://molprobity.biochem.duke.edu/>).

Tabela 6 – Validação dos modelos 3D para a proteína CV0259 por diferentes servidores.

Servidor	Tipo de análise	Resultado
ERRAT2	Fator de qualidade global	90.715
VERIFY-3D	<i>Averaged 3D-1D score</i>	92.47 (%)
ProSA-web	Qualidade geral ou Z-score	-8.69

Errat 2: analisa a frequência de interações não covalentes entre átomos, ou seja, avalia a qualidade geral do modelo. Estruturas com valores acima de 91 são confiáveis. VERIFY-3D: determina a compatibilidade do modelo com sua sequência de aminoácidos atribuindo um score que deve ser maior que 0,2. Estruturas com mais de 80% dos resíduos com valores acima de 0,2 são confiáveis. ProSA-web: indica a qualidade geral do modelo a partir da comparação com dados de proteínas presentes no PDB, as quais possuem valores de Z-score entre -10 e 5, geralmente. Estruturas com valores de Z-score entre -10 e 5 são confiáveis. Fonte: elaborada pela autora

Tabela 7 – Validação dos modelos 3D para a proteína CV2065 por diferentes servidores.

Servidor	Tipo de análise	Resultado
ERRAT2	Fator de qualidade global	91.8712
VERIFY-3D	<i>Averaged 3D-1D score</i>	90.96 (%)
ProSA-web	Qualidade geral ou Z-score	-9.38

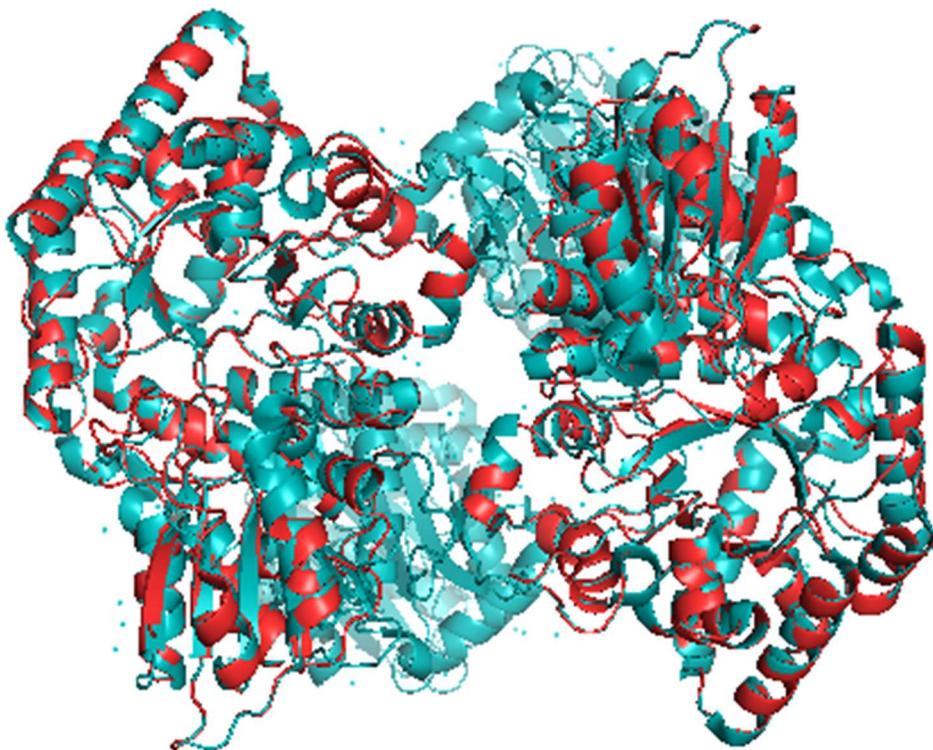
Errat 2: analisa a frequência de interações não covalentes entre átomos, ou seja, avalia a qualidade geral do modelo. Estruturas com valores acima de 91 são confiáveis. VERIFY-3D: determina a compatibilidade do modelo com sua sequência de aminoácidos atribuindo um score que deve ser maior que 0,2. Estruturas com mais de 80% dos resíduos com valores acima de 0,2 são confiáveis. ProSA-web: indica a qualidade geral do modelo a partir da comparação com dados de proteínas presentes no PDB, as quais possuem valores de Z-score entre -10 e 5, geralmente. Estruturas com valores de Z-score entre -10 e 5 são confiáveis. Fonte: elaborada pela autora

PyMol é um software utilizado para visualizar estruturas tridimensionais de proteínas, produzindo imagens de alta qualidade em diferentes representações (YUAN; CHAN; HU, 2017). Além de permitir a simples visualização, PyMol pode alinhar duas estruturas e mostrar o grau de identidade entre as mesmas, expresso por *root mean square deviation* (RMSD). O valor de RMSD é sempre maior que zero e, quanto mais próximo desse valor, melhor é o alinhamento entre as duas estruturas. Desse modo, os valores de RMSD obtidos são de 0,263 e 0,139 para as proteínas CV0259 e CV2065, respectivamente (Figuras 25 e 26). Esses valores vão ao encontro da informação de que a proteína CV2065 possui maior similaridade com seu *template* (61,14).

Visto que as estruturas previstas possuem alta similaridades com seus *templates*, pode-se fazer inferências sobre o método de catálise das proteínas alvo. A proteína 4ZM6, *template* utilizado para obter a estrutura tridimensional de CV0259, é uma β -N-acetilglicosaminidase de *Rhizomucor miehei* (RmNag). A proteína em questão faz parte da família GH3, possuindo resíduos Asp80, Arg149, Lys179, His192, Met229, Asp268 e Met297 no sítio catalítico. Em vários estudos utilizando proteínas da família GH3, o resíduo aspartato é identificado como um resíduo chave no sítio ativo da enzima, junto de um glutamato utilizando em um mecanismo de catálise ácido/base. RmNag, no entanto, utiliza uma diade His/Asp como mecanismo de catálise ácido/base, composta pelos resíduos His192 e Asp268, uma combinação incomum para N-acetilglicosaminidases da família GH3 (QIN *et al.*, 2015). Desse modo, pode-se supor que a proteína CV0259 possui um mecanismo de catálise parecido com o do seu *template*.

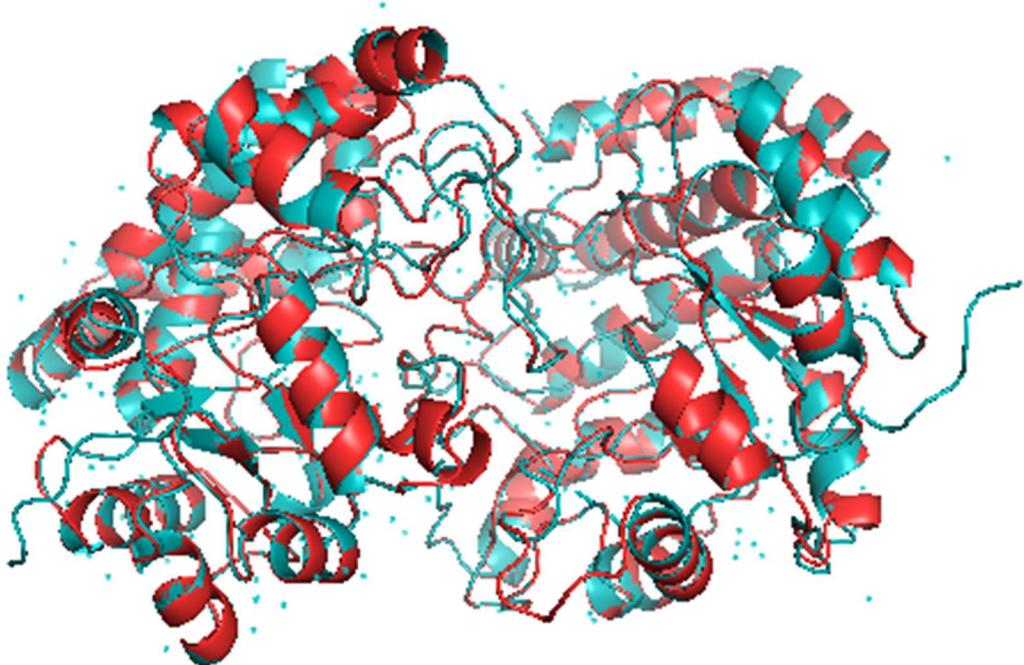
A proteína 6JTI, utilizada como *template* para a CV2065, é uma β -hexosaminidase de *Neisseria gonorrhoeae* codificada por um gene chamado de NagZ (que dá nome a proteína). Essas enzimas estão envolvidas na reciclagem de paredes celulares de bactérias. Em *Bacillus subtilis*, por exemplo, proteínas codificadas pelo gene NagZ possuem uma diade de His/Asp no seu sítio catalítico, com o resíduo de histidina estando na posição adequada para o mecanismo de catálise ácido/base. A presença desses dois aminoácidos é comum no sítio ativo de proteínas codificadas por esse gene porém não é a regra geral para β -hexosaminidases, que geralmente possuem um resíduo de glutamato no sítio ativo (SU; SHENG; LIU, 2016). Assim, pode-se supor que esses mecanismos também estão presentes na proteína CV2065, visto que a mesma possui alta similaridade com 6JTI.

Figura 25– Sobreposição da estrutura 3D de CV0259 com seu *template* (4ZM6).



O modelo para a proteína CV0259 está representado em vermelho e o modelo para 4ZM6 está representado em azul. Fonte: elaborada pela autora utilizando o programa PyMol.

Figura 26 – Sobreposição da estrutura 3D de CV2065 com seu *template* (6JTI).



O modelo para a proteína CV2065 está representado em vermelho e o modelo para 6JTI está representado em azul. Fonte: elaborada pela autora utilizando o programa PyMol.

Analizando os dados obtidos utilizando ferramentas de bioinformática, percebe-se que podem ser realizados estudos posteriores com o objetivo de investigar a fundo os sítios catalíticos das proteínas presentes neste trabalho, em especial aquelas que foram confirmadas para a presença de atividade enzimática. Além disso, podem ser feitos testes de degradação de quitina presente em diferentes formas, tais como paredes celular de bactérias e de fungos. Testes para atividade enzimática em outros polissacarídeos, tais como peptidoglicanos e celulose, também podem ser realizados, futuramente, utilizando tanto as enzimas com atividade quitinoclástica quanto as proteínas que não apresentaram nenhuma atividade para a degradação de quitina.

8 CONCLUSÃO

Utilizando técnicas de biologia molecular e análises de solubilidade em diversos softwares diferentes, foi possível produzir, de forma heteróloga, proteínas de *C. violaceum* ATCC 12472 em células de *E. coli* DH5 α e BL21(DE3). Apenas duas proteínas obtiveram resultados positivos em testes para atividade quitinoclástica, utilizando quitinas coloidal e insolúvel, chamadas de CV0259 e CV2065. Através do uso de servidores e softwares de bioinformática, pode haver a predição da estrutura tridimensional de ambas. Informação útil para inferir informações sobre o mecanismo de catálise de ambas.

REFERÊNCIAS

- AAM, B. B.; HEGGSET, E. B.; NORBERG, A. L.; SØRLIE, M.; VÅRUM, K. M.; EIJSINK, V. G. Production of chitooligosaccharides and their potential applications in medicine. **Mar Drugs**, v. 8, n. 5, p. 1482-1517, 2010.
- ADRANGI, S.; FARAMARZI, M. A. From bacteria to human: a journey into the world of chitinases. **Biotechnol. Adv.**, v. 31, p. 1786-1795, 2013.
- ALVES, S. M. A bioinformática e sua importância para a biologia molecular. **Revista Brasileira de Educação e Saúde**, v. 3, n. 4, p. 18-25, 2013.
- SANDEEP, A.; SANGAMESHWAR, K.; MUKESH, G.; CHANDRAKANT, R.; AVINASH, D. A brief overview on chitosan applications. **Indo American Journal of Pharmaceutical Research**, v. 3, n. 12, 2013.
- ARAÚJO N. D.; FARIA R. P.; PEREIRA P. B.; FIGUEIRÊDO, F. M.; MORAIS, A. M. B.; SALDANHA, L. C.; GABRIEL, J. E. A Era da Bioinformática: seu potencial e suas implicações para as Ciências Da Saúde. **Estudos de Biologia**, v. 30, n. 70, p. 143-148, 2008.
- ARA, K. Z. G.; MÅNBERGER, A.; GABRIŠKO, M.; LINARES-PASTÉN, J. A.; JASILIONIS, A.; FRÍÐJÓNSSON, O. H.; HREGGVIÐSSON, G. Ó.; JANEČEK, S./ KAR, E. N. Characterization and diversity of the complete set of GH family 3 enzymes from *Rhodothermus marinus* DSM 4253. **Scientific Reports**, v. 10, n. 1329, 2020.
- ARIMORI, T.; KAWAMOTO, N.; SHINYA, S.; OKAZAKI, N.; NAKAZAWA, M.; MIYATAKE, K.; FUKAMIZO, T.; UEDA, M.; TAMADA, T. Crystal structures of the catalytic domain of a novel glycohydrolase family 23 chitinase from *Ralstonia* sp. A-471 reveals a unique arrangement of the catalytic residues for inverting chitin hydrolysis. **J. Biol. Chem.**, v. 288, p. 18696-18706, 2013.
- BARIKANI, M.; OLIAEI, E.; SEDDIQI, H.; HENGAMEH. H. Preparation and application of chitin and its derivatives: a review. **Iranian Polymer Journal**, v. 23, n. 307, p. 307-326, 2014.
- BAGOWSKI, C. P.; BRUINS, W.; VELTHUIS, A. J. W. The Nature of Protein Domain Evolution: Shaping the Interaction Network. **Current Genomics**, v. 11, p. 368-376, 2010.
- BEIER, S.; BERTILSSON, S. Bacterial chitin degradation—mechanisms and ecophysiological strategies. **Frontiers in Microbiology**, v. 4, 2013.
- BHATTACHARYA, D.; NAGPURE, A.; GUPTA, R. K. Bacterial Chitinases: Properties and Potential. **Critical Reviews in Biotechnology**, v. 27, n.1, p. 21-28, 2008.
- BOLLER, T. Biochemical analysis of chitinases and β -1, 3-glucanases. **Molecular plant pathology: a practical approach**, v. 2, p. 23–30, 1992.
- BORASTON, A. B.; BOLAM, D. N.; GILBERT, H. J.; DAVIES, G. J. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. **Biochem J**, v. 382, n. 3, p. 769-781, 2004.
- BOWIE, J.; LUTHY, R.; EISENBERG, D. A method to identify protein sequences that fold into a known three-dimensional structure. **Science**, v. 253, n. 5016, p. 164-170, 1991.
- BRANDEN, C.; TOOZE, J. **Introduction to protein structure**, Gerland: New York, 1991.

BRYANT, F. R. Construction of a Recombinase-deficient Mutant recA Protein That Retains Single-stranded DNA-dependent ATPase Activity. **The Journal Of Biological Chemistry**, v. 263, n. 18, p. 8716-8723, 1986.

BRUM, I. J. B. Ferramentas de Bioinformática para Proteômica. Dissertação (Mestrado em Biologia Funcional e Molecular) – Instituto de Biologia, Universidade Estadual de Campinas. São Paulo, p.83. 2007.

BURTAN, A. F. N-AcetylGlucosamine as a Cytoprotective Agent. KR PatentNO. 0145715 B, 1998.

CABEZAS, J. A. Some comments on the type references of the official nomenclature (IUB) for beta-N-acetylglucosaminidase, beta-N-acetylhexosaminidase and beta-N-acetylgalactosaminidase. **Biochemical Journal**, v. 261, n. 3, p. 261-263, 1989.

CALDAS, L. R. Um pigmento nas águas negras. **Ciência Hoje** 11, 56-57, 1990.

CALDAS, L. R.; LEITÃO, A. A. C.; SANTOS, S. M.; TYRRELL, R. M. Preliminary experiments on the photobiological properties of violacein. In *Proceedings of the International Symposium on Current Topics in Radiology and Photobiology*, ed. Tyrrell, R. M. (Academia Brasileira de Ciências, Rio de Janeiro), pp. 121–126, 1978.

CAMPANA-FILHO, S. P; DE BRITTO D.; CURTI, E.; ABREU, F. R.; CARDOSO, M. B.; BATTISTI, M. V.; SIM, P. C.; GOY, R. C.; SIGNINI, R.; LAVALL, R. L. Extração, estruturas e propriedades de A- e B-quitina. **Química Nova**, v. 30, n. 20, p. 644-650, 2007.

CANTAREL, B. L.; COUTINHO, P. M.; RANCUREL, C.; BERNARD, T.; LOMBARD, V.; HENRISSAT, B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. **Nucleic Acids Res.**, v. 37, 2009.

CHEN, V. B.; ARENDALL 3RD, W. B.; HEADD, J. J.; KEEDY, D. A.; IMMORMINO, R. M.; KAPRAL, G. J.; MURRAY, L. W.; RICHARDSON, J. S.; RICHARDSON, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallographica**, v. 8, p. 2493-2516, 2010.

CHEN, J. K.; SHEN, C. R.; LIU, C. L. N-Acetylglucosamine: Production and Applications **Mar. Drugs**, v. 8, n. 9, p. 2493-2516, 2010.

CHENG, J.; RANDALL, A. Z.; SWEREDOSKI, M. J.; BALDI, P. SCRATCH: a protein structure and structural feature prediction server. **Nucleic Acid Research**, v. 33, 2005.

COLOVOS, C.; YEATES, T. O. Verification of protein structures: Patterns of nonbonded atomic interactions. **Protein Science**, v. 2, p. 1511-1519, 1993.

DAHIYA, N.; TEWARI, R.; HOONDAL, G. S. Biotechnological aspects of chitinolytic enzymes: a review. **Appl Microbiol Biotechnol**, v. 71, n. 6, p. 773-782, 2006.

DARNELL, S. Why Structure Prediction Matters. 2020. Disponível em: <https://www.dnastar.com/blog/structural-biology/why-structure-prediction-matters>. Acesso em: 03 de mar. de 2021.

DASH, M.; CHIELLINI, F.; OTTENBRITE, R. M.; CHIELLINI, E. Chitosan-a versatile semi-synthetic polymer in biomedical applications. **Progress in Polymer Science**, v. 36, n. 8, p. 981-1014, 2011.

DAVANLOO, P.; ROSENBERG, A. H.; DUNN, J. J.; STUDIER, F. W. Cloning and expression of the gene for bacteriophage T7 RNA polymerase. **Proc Natl Acad Sci U S A**, v. 81, n. 7, p. 2035-2039, 1984.

DUNG, P. L. E.; MILAS, M.; RINAUDO, M.; & DESBRIÈRES, J. Water soluble derivatives obtained by controlled chemical modifications of chitosan. **Carbohydrate Polymers**, v. 24, n. 3, p. 209–214, 1994.

DYBALLA N.; METZGER S. Fast and Sensitive Colloidal Coomassie G-250 Staining for Proteins in Polyacrylamide Gels. **Journal of Visualized Experiments**, v. 30, n. 1431, 2009.

EL SAYED, H. EL.; ASHRY, E. S. H. E.; MOHAMED R. E. ALY, M. R. E. Synthesis and biological relevance of N-acetylglucosamine-containing oligosaccharides. **Pure and Applied Chemistry**, v. 79, n. 12, pp. 2229-22242, 2007.

FACHIN, L. P. A integração das “ômicas” através de ferramentas de Bioinformática. In: Reunião Anual da Sbpc, 69., 2017, Belo Horizonte. Anais Eletrônicos...Belo Horizonte: SBPC, 2017. Disponível em: http://www.sbpcnet.org.br/livro/69ra/PDFs/arq_5350_2821.pdf. Acesso em: 22 fev. 2021.

FERNANDES, M. R. Degradação de quitina em *Chromobacterium*: proposição da via catabólica e expressão heteróloga de uma monooxigenase lítica de polissacarídeo de *C. violaceum* ATCC 12472. Monografia (Bacharelado em Ciências Biológicas) – Departamento de Biologia, Universidade Federal do Ceará. Fortaleza, p. 113, 2017

FIGUEIREDO, J. E. F.; ROCHA, W. D.; BENEDITO, V. A.; COELHO, V. T. S. Estudo de modelagem molecular por homologia para análise genômica funcional. **Embrapa Milho e Sorgo, Comunicado Técnico 129**, Sete Lagoas, 2005. Disponível em: <https://www.infoteca.cnptia.embrapa.br/bitstream/doc/489727/1/Com129.pdf>. Acesso em: 25 fev. 2021.

FILHO, O. A. S.; ALENCASTRO, R. B. Modelagem de proteínas por homologia. **Química Nova**, v. 26, n. 2, p. 253-259, 2003.

FUJITA, K.; SHIMOMURA, K.; YAMAMOTO, K.; YAMASHITA, T.; SUZUKI, K. A chitinase structurally related to the glycoside hydrolase family 48 is indispensable for the hormonally induced diapause termination in a beetle. **Biochem Biophys Res Commun.**, v. 345, n. 1, p. 502-507, 2006.

FUNKHOUSER, J. D.; ARONSON, N. N. JR. Chitinase family GH18: evolutionary insights from the genomic history of a diverse protein family. **BMC Evol Biol.**, v. 7, n. 96, 2007.

GARDNER, K. H.; BLACKWELL, J. Refinement of structure of beta-chitin. **Biopolymers**, v. 14, p.1581-1595, 1975.

GASTEIGER, E.; HOOGLAND, C.; GATTIKER, A.; DUVAUD S.; WILKINS, M. R.; APPEL, R. D.; BAIROCH, A. Protein Identification and Analysis Tools on the ExPASy Server. **The Proteomics Protocols Handbook**. 1. ed. Humana Press, 2005.

HAMED I., ÖZOGUL F.; REGENSTEIN J. M. Industrial applications of crustacean by-products (chitin, chitosan, and chitooligosaccharides): A review. **Trends in Food Science & Technology**, v. 48, p. 40-50, 2016.

- HAMID, R.; A. KHAN, M. A.; AHMAD, M.; AHMAD M. M.; ABDIN, M. Z.; MUSARRAT, J.; JAVED, S. Chitinases: An update. **Journal of Pharmacy & BioAllied Sciences**, v. 5, n. 1, p. 21-29, 2013.
- HARTLEY, J. L. Cloning technologies for protein expression and purification. **Current Opinion in Biotechnology**, v. 17, n. 4, p. 359–366, 2006.
- HARVEY, A. J.; HRMOVA, M.; GORI, R.; N. VARGHESE, J. N.; FINCHER, G. B. Comparative modeling of the three-dimensional structures of family 3 glycoside hydrolases. **Proteins: Structure, Function and Bioinformatics**, v. 41, n. 2, p. 257-269, 2000.
- HAYWARD, S. Structural Principles Governing Domain Motions in Proteins. **PROTEINS: Structure, Function, and Genetics**, v. 36, p. 425-435, 1999.
- HEBDITCH, M.; CARBALLO-AMADOR, M. A.; CHARONIS, S.; CURTIS, R.; WARWICKER, J. Protein-Sol: a web tool for predicting protein solubility from sequence. **Bioinformatics**, v. 33, n. 19, p. 3098-3100, 2017.
- HENRISSAT, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochem J.**, v. 289, p. 309-316, 1991.
- HOWARD, M. B.; EKBORG, N. A.; WEINER, R. M.; HUTCHESON, S. W. Detection and characterization of chitinases and other chitin-modifying enzymes. **Journal of Industrial Microbiology and Biotechnology**, v. 30, n. 11, p. 627-635, 2003.
- HOWELL, C. R. Mechanisms Employed by *Trichoderma* Species in the Biological Control of Plant Diseases: The History and Evolution of Current Concepts. **Plant Disease**, v. 87, n. 1, p. 4-10, 2003.
- KAMEL, M.; ALNAHDI, M. Inhibition of superoxide anion release from human polymorphonuclear leukocytes by N-acetyl-galactosamine and N-acetyl-glucosamine. **Clinical Rheumatology**, v. 11, n. 2, p. 254-260, 1992.
- KAYA, M.; MUJTABA, M.; EHRLICH, H.; SALABERRIA, A. M.; BARAN, T.; AMEMIYA, C. T.; GALLI, R.; AKYUZ, L.; SARGIN, I.; LABIDI, J. On chemistry of γ -chitin. **Carbohydrate Polymers**, v. 176, p. 177-186. 2017.
- KIELAK, A. M.; CRETOIU, M. S.; SEMENOV, A. V.; SØRENSEN, S. J.; VAN ELSAS, J. D. Bacterial chitinolytic communities respond to chitin and pH alteration in soil. **Appl Environ Microbiol**, v. 79, n. 1, p. 263-272, 2013.
- KOBURGER, J. A; MAY, S. O. Isolation of *Chromobacterium* spp. from foods, soil, and water. **Appl Environ Microbiol**, v. 44, n. 6, p. 1463-1465. 1982.
- KRAJEWSKA, B. Application of chitin- and chitosan-based materials for enzyme immobilizations: a review. **Enzyme and Microbial Technology**, v. 35, n. 2, p. 126-139, 2004.
- KUMAR, M. N. R. A review of chitin and chitosan applications. **Reactive and Functional Polymers**, v. 46, n. 1, p. 1–27, 2000.
- KUMAR, M. N. V. R.; MUZZARELLI, R. A. A.; MUZZARELLI, C.; SASHIWA, H.; DOMB, A. J. Chitosan Chemistry and Pharmaceutical Perspectives. **Chemical Reviews**, v. 104, n. 12, p. 6017-6084, 2004.

- LAEMMLI, U. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, v. 227, n. 5259, p. 680, 1970.
- LAVALLIE, E. R.; DIBLASIO, E. A.; KOVACIC, S.; GRANT, K. L.; SCHENDEL, P. F.; MCCOY, J. M. A Thioredoxin Gene Fusion Expression System That Circumvents Inclusion Body Formation in the *E. coli* Cytoplasm. **Nature Biotechnology**, v. 11, n. 2, p. 187–193, 1993.
- LESK, A. M. *Introdução à bioinformática*. 2. ed. Artmed, 2008.
- LIN, M. H.; SHU, J. C.; LIN, L. P.; CHONG, K, Y.; CHENG, Y. W. *et al.* Elucidating the Crucial Role of Poly N-Acetylglucosamine from *Staphylococcus aureus* in Cellular Adhesion and Pathogenesis. **PLOS ONE**, v. 10, n. 4, 2015.
- LIU, W.; XIE, Y.; MA, J.; LUO, X.; NIE, P.; ZUO, X.; LAHRMANN, U.; ZHAO, Q.; ZHENG, Y.; ZHAO, Y.; XUE, Y.; REN, J. IBS: an illustrator for the presentation and visualization of biological sequences. **Bioinformatics**, v. 31, n. 20, p. 3359-3361, 2015.
- LOBO, M. D. P.; SILVA, F. D. A.; LANDIM, P. G. D. C. *et al.* Expression and efficient secretion of a functional chitinase from *Chromobacterium violaceum* in *Escherichia coli*. **BMC Biotechnology**, v. 13, n. 46, 2013.
- LODGE, J.; LUND, P.; MINCHIN, S. *Gene Cloning: Principles and Applications*. 1. ed. Reino Unido: Taylor & Francis Group, 2007.
- MACDONALD, S. S.; BLAUKOPF, M.; WITHERS, S. G. N-Acetylglucosaminidases from CAZy Family GH3 Are Really Glycoside Phosphorylases, Thereby Explaining Their Use of Histidine as an Acid/Base Catalyst in Place of Glutamic Acid. **Glycobiology and Extracellular Matrices**, v. 290, n. 8, p. 4487-4895, 2015.
- MAGGIO, E. T.; RAMNARAYAN, K. Recent developments in computational proteomics. **Trends in Biotechnology**, v. 19, n. 7, p. 266-272, 2001.
- MARTINS-MIRANDA, A. Atividade de enzimas relacionadas com estresses bióticos e abióticos em plântulas de feijão-de-corda [*Vigna unguiculata* (L.) walp.] exposta à salinidade e deficiência hídrica. **Dissertação** (Mestrado em Bioquímica) – Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará. Fortaleza, p. 85, 2002.
- MARTZ, E.; HAAS, J. Practical Guide to Homology Modeling. **Proteopedia**, 2020. Disponível em: <https://proteopedia.org/wiki/index.php/Practical_Guide_to_Homology_Modeling>. Acesso em: 16 de abr. de 2021.
- MILLER, G. L. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. **Analytical Chemistry**, v. 31, n. 3, p. 426-428, 1959.
- MINKE, R.; BLACKWELL, J. The structure of alpha-chitin. **J Mol Biol**, v. 120, p. 167-181, 1978.
- MUZZARELLI, R.A.A. Native, industrial, and fossil chitins. In *Chitin and Chitinases*, P. Jollès, and R.A.A. Muzzarelli, eds. (Basel: Birkhäuser Basel), pp. 1-6, 1999.
- MORGANTI, P.; FABRIZI, G.; PALOMBO, P.; PALOMBO, M.; RUOCCHI, E.; CARDILLO, A.; GIANLUCA, MORGANTI. Chitin-nanofibrils: A new active cosmetic carrier. **Journal of Applied Cosmetology**, v. 26, 2008).

- MARCHLER-BAUER A.; PANCHENKO, A. R.; SHOEMAKER, B. A.; THIESSEN, P. A.; GEER, L. Y.; BRYANT, S. H. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. **Nucleic Acid Research**, v. 30, n. 281, p. 281-283, 2002.
- NADAL, M.; SAWERS, R.; NASEEM, S. *et al.* An N-acetylglucosamine transporter required for arbuscular mycorrhizal symbioses in rice and maize. **Nature Plants**, v. 3, 2017.
- NAMPOOTHIRI, K. M.; BAIJU, T. V.; SANDHYA, C.; SABU, A.; SZAKACSB, G.; PANDEY, A. Process optimization for antifungal chitinase production by *Trichoderma harzianum*. **Process Biochemistry**, v. 39, n. 11, p. 1583-1590, 2004.
- NAUMOFF, D. Hierarchical classification of glycoside hydrolases. **Biochemistry**. v. 76, p. 622-35, 2011.
- NELSON, D. L.; COX, M. M. Princípios de Bioquímica de Lehninger. Tradução de Ana Beatriz Gorini da Veiga et al. Porto Alegre: Artmed, 2014 (Título original: Lehninger principles of biochemistry).
- NIELSEN, H.; ENGELBRECHT, J.; BRUNAK, S.; VON HEIJNE, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. **Protein Engineering Design and Selection**, v.10, n. 1, p. 1-6, 1997.
- PICHYANGKURA, R.; KUDAN, S.; KUTTIYAWONG, K.; SUKWATTANASINITT, M.; AIBA, S. Quantitative production of 2-acetamido-2-deoxy-D-glucose from crystalline chitin by bacterial chitinase. **Carbohydrate Research**, v. 337, n. 6, p. 557-559, 2002.
- PHUE, J. N.; LEE, S. J.; TRINH, L.; SHILOACH, J. Modified *Escherichia coli* B (BL21), a superior producer of plasmid DNA compared with *Escherichia coli* K (DH5alpha). **Biotechnol Bioeng**, v. 101, n. 4, p. 831-836, 2008.
- PROSDOCIMI, F. Introdução à Bioinformática, 2007. Disponível em: http://www.iq.usp.br/setubal/bmc/2013/FProsdocimi07_CursoBioinfo.pdf. Acesso em: 22 fev. 2021.
- QIN, Z.; XIAO, Y.; YANG, X.; MESTERS, J. R.; YANG, S.; JIANG, Z. A unique GCN5-related glucosamine N-acetyltransferase region exist in the fungal multi-domain glycoside hydrolase family 3 β -N-acetylglucosaminidase. **Scientific Reports**, v. 5, n. 18292, 2015.
- QORONFLEH, M. W.; HESTERBERG, L. K.; SEEFELDT, M. B. Confronting high-throughput protein refolding using high pressure and solution screens. **Protein Expression and Purification**, v. 55, n. 2, p. 209–224, 2007.
- RATHORE, A. S.; GUPTA, R. D. Chitinases from Bacteria to Human: Properties, Applications, and Future Perspectives. **Enzyme Research**, v. 2015, 2015.
- ROBERTS G. A. F. Chitin Chemistry, Mc Millan Press Ltd: London, 1992, cap 1.
- ROVATI, L.; CASULA, P.; MASCHERPA, S. N-Acetylglucosamine for Treating Degenerative Afflictions of the Joints. 3697652. US Patent. 1972.
- SAMBROOK, J., RUSSELL, D. W. **Molecular cloning: A Laboratory Manual**, vol. 1-3. Cold Spring Habour Laboratory Press, New York, 2001.
- SASHIWA, H.; FUJISHIMA, S.; YAMANO, N.; KAWASAKI, N.; NAKAYAMA, A.; MURAKI, E.; HIRAGA, K.; ODA, K.; AIBA, S. Production of N-acetyl-d-glucosamine from

α -chitin by crude enzymes from *Aeromonas hydrophila* H-2330. **Carbohydrate Research**, v. 337, n. 8, p. 761-763, 2002.

SASHIWA, H.; FUJISHIMA, S.; YAMANO, N.; KAWASAKI, N.; NAKAYAMA, A.; MURAKI, E.; SUKWATTANASINITT, M.; PICHYANGKURA, R.; AIBA, S. Enzymatic production of N-acetyl-D-glucosamine from chitin. Degradation study of N-acetylchitooligosaccharide and the effect of mixing of crude enzymes. **Carbohydrate Polymers**, v. 51, n. 4, p. 391-395, 2003.

SETHULEKSHMI, C. Chitin--and its benefits. **International Journal of Advanced Research in Biological Sciences**, v. 1, p. 171-175, 2014.

SILVA, H. S. R. C.; DOS SANTOS, K. S. C. R; FERREIRA, E. I. Quitosana: derivados hidrossolúveis, aplicações farmacêuticas e avanços. **Química Nova**, v. 29, n. 4, 2006.7

SINGH, R.; SHITIZ, K.; SINGH, A. Chitin and chitosan: biopolymers for wound management. **International Wound Journal**, v. 14, n. 6, p. 1276-1289. 2017

SMIALOWSKI, P.; DOOSE, G.; TORKLER, P.; KAUFMANN, S.; FRISHMAN, D. PROSO II - a new method for protein solubility prediction. **FEBS Journal**, v. 279, n. 12, p. 2192-2200, 2012.

SMIALOWSKI, P.; MARTIN-GALIANO, A. J.; MIKOLAJKA, A.; GIRSCHICK, T.; HOLAK, T. A.; FRISHMAN, D. Protein solubility: sequence based prediction and experimental verification. **Bioinformatics**, v. 23, p. 2536–2542, 2007.

SOARES, J. M. Caracterização *in silico* e docking molecular de uma N-acetil- β -D-glicosaminidase de *Chromobacterium violaceum*: estudos preliminares para o design de um bioprocesso. **Monografia** (Bacharelado em Biotecnologia) - Centro de Ciências, Universidade Federal do Ceará. Fortaleza, p. 92, 2016.

SOUZA, C. P.; ALMEIDA, B. C.; COLWELL, R. R.; RIVERA, I. N. G. The Importance of Chitin in the Marine Environment. **Marine Biotechnology**, v. 13, n. 5, p. 823–830, 2011.

STOYKOV, Y. M.; PAVLOV, A. I.; KRASTANOV, A. I. Chitinase biotechnology: Production, purification, and application. **Engineering in life science**, v. 15, n. 1, p. 30-38, 2014.

SU, H.; SHENG, X.; LIU, Y. Insights into the catalytic mechanism of N-acetylglucosaminidase glycoside hydrolase from *Bacillus subtilis*: a QM/MM study. **Organic & Biomolecular Chemistry**, v. 14, n. 13, p. 3432–3442. 2016

VAAJE-KOLSTAD, G.; WESTERENG, B.; HORN, S. J.; LIU, Z.; ZHAI, H.; SØRLIE, M.; EIJSINK, V. G. H. An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. **Science**, v. 30, n. 6001, 2010.

VAAJE-KOLSTAD, G.; HORN, S. J.; VAN AALTEN, D. M.; SYNSTAD, B.; EIJSINK, V. G. The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. **J Biol Chem**, v. 280, n. 31, p 28492-28497, 2995.

VAN LOON, L. C.; REP, M.; PIETERSE, C. M. J. Significance of inducible defense-related proteins in infected plants. **Annu Rev Phytopathol**, v. 44, p. 135-162, 2006.

VASCONCELOS, A. T. R.; ALMEIDA DE, D. F.; HUNGRIA, M.; GUIMARÃES, C. T.; ANTÔNIO, R. V.; ALMEIDA, F. C.; ALMEIDA, L. G. P.; ALMEIDA, R.; AVLES-GOMES,

- J. A.; ANDRADE, E. M.; ARARIPE, J.; ARAÚJO, M. F. F.; ASTOLFI-FILHO, S.; AZEVEDO, V.; BAPTISTA, A. J.; BATAUS, L. A. M.; BATISTA DA, J. S.; BELÓ, A.; VAN DEN BERG, C.; BOGO, M.; BONATTO, S.; BORDIGNON, J.; BRIGIDO, M. M.; BRITO, C. A.; BROCCHI, M.; BURITY, H. A.; CAMARGO, A. A.; CARDOSO, D. DAS D. DE P.; CARNEIRO, N. P.; CARRARO, D. M., *et al.* The complete genome sequence of *Chromobacterium violaceum* reveals remarkable and exploitable bacterial adaptability. *Proc Natl Acad Sci U S A*, v. 100, n. 20, p. 11660-11665, 2003.
- VERLI, H. Bioinformática: da Biologia à flexibilidade molecular. Porto Alegre: FAPERGS, 2014.
- VU, K. D.; YAN, S.; TYAGI, R. D.; VALÉRO, J. R.; SURAMPALLI, R. Y. Induced production of chitinase to enhance entomotoxicity of *Bacillus thuringiensis* employing starch industry wastewater as a substrate. *Bioresource Technology*, v. 100, n. 21, p. 5260-5269, 2009.
- WIEDERSTEIN, M.; SIPPL, M. J. ProSA-web: interactive web service for the recognition or errors in three-dimensional structures of proteins. *Nucleic Acids Research Web Server*, v. 35, p. 407-410, 2007.
- WURZBACHER, C.; GROSSART, H. Improved detection and identification of aquatic fungi and chitin in aquatic environments. *Mycologia*, v. 104, 2012.
- XIA, W.; LIU, P.; JZHANG, J.; CHEN, J. Biological activities of chitosan and chitooligosaccharides. *Food Hydrocolloids*, v. 25, n. 2, p. 170-179, 2011.
- XIE, W.; XU, P.; LIU, Q. Antioxidant activity of water-soluble chitosan derivatives. *Bioorganic & Medicinal Chemistry Letters*, v. 11, n. 13, p. 1699–1701, 2001.
- XU, D.; NISSINOV, R. Favorable domain size in proteins. *Folding & Design*, v. 3, n. 1, p. 11-17, 1997.
- YANG, Q.; FUKAMIZO, T. Targeting Chitin-containing organisms. 1. ed. Singapore: Springer, 2019.
- YANG, M.; LUOH, S. M.; GODDARD, A.; REILLY, D.; HENZEL, W.; BASS, S. The *bglX* gene located at 47.8 min on the *Escherichia coli* chromosome encodes a periplasmic beta-glucosidase. *Microbiology (Reading)*, v. 142, n. 7, p. 1659-1665, 1996.
- YOUNG, C. L.; BRITTON, Z. T.; ROBINSON, A. S. Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. *Biotechnology Journal*, v. 7, p. 620-634, 2012.
- YUAN, S.; CHAN, H. C. S.; HU, Z. Using PyMOL as a platform for computational drug design. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, v. 7, n. 2, 2017.
- ZENG, D.; MEI, X.; WU, J. Effects of an environmentally friendly seed coating agent on combating head smut of corn *Sphacelotheca reiliana* caused by and corn growth. *Journal of Agricultural Biotechnology and Sustainable Development*, v. 2, p. 108-112, 2010.
- ZHANG, A.; GAO, C.; WANG, J.; CHEN, K.; OUYANG, P. An efficient enzymatic production of N-acetyl-d-glucosamine from crude chitin powders. *Green Chemistry*, v. 18, n. 7, p. 2147-2154, 2016.

ZHOU, J.; CHEN, J.; XU, N.; ZHANG, A.; CHEN, K.; XIN, F.; ZHANG, W.; MA, J.; FANG, Y.; JIANG, M.; DONG, W. The broad-specificity chitinases: their origin, characterization, and potential application. **Applied Microbiology and Biotechnology**, v. 103, p. 3289–3295, 2019.

ANEXO A - SEQUÊNCIAS DAS PROTEÍNAS ESTUDADAS

**a) Sequência da proteína CV2592 - AAQ60262.1 carbohydrate-binding protein
[*Chromobacterium violaceum* ATCC 12472]**

MKLAFLMPAGIGIALAGLSGASWAHGTMEVPVNRYSCYQEGAESPCTPACQEAKEVRGGTQAM
YDWNGINQNPPGDNHQAVVPDGMLCAAGQSKFKGFLARTDWPATNIVPNASGNFEFIYKAP
APHATKYFRFYVTKNGWNPSQPLKWSDLELFGTYNGNPLDASQRYHMTMKLPTGKTGRHII
YNVWKRSDEEAFYSCSDVNFGNGGPQPEPPPISNPWKEGVSVTAHENLPNKSSVTLRIFDG
SGRDVESHKVDSLATTGQAANWPYELGVVNASSQIGRIGVISSKQRTVTINPVRSATANRV
WLNERYSGYRYQIDIKKGDDGVTPPPAGDAWREGVAYAVGQVVSYQNRRYRCLQAHTAWVGA
GWTPSTTPALWTPV

**b) Sequência da proteína CV0259 - AAQ57938.1 probable glycosyl hyrolase, family 3
[*Chromobacterium violaceum* ATCC 12472]**

MTLSDERLDQLAGRCLMVDAGPVLTEEEAARLQRMVRRAVCLFRRNVPDAASTRRLVADLK
CALGHDVLIGIDQEGGAVMRTRFLPQAPAAMALAAGDEALARRVGGAVARGLKSLGVNW
APVLDLNNNPANPVIGERSFGADPARAAALARAWMEGHLAEGVACCVKHFPGHDHTDSHL
DLPVVDKPQAEELRDYELAPFKALSAHAPALMSAHRFPALDADWPATLSPAILTGLLRREL
FRGVAITDALNMRAIRERWGQPAGAVQTLKAGADIALVLQFADEMEASFDALRAALRN
QARLEEAARVDALIRRYPSRNDCGYPAERRA
DEALFADAWTRALTARGEIPTLAAGQRLR
LVVQDQAPS
DGVSEEGLAADALIAALSRRHALEV
VRFQERAMLDWSTLPQDGLFTV
LASTTR
ERYGQREQQSWKPD
LHLALWN
PYAA
ADIA
AAPAL
ISYGF
ADA
ALFA
ASR
CLAGE
IAA
AGRL
PV
ALGQPYQE

**c) Sequência da proteína CV0554 - AAQ58230.1 probable carbohydrate-binding protein
[*Chromobacterium violaceum* ATCC 12472]**

MKLHWHLPAALALSCLSAPLWAHGTMEIPINRTYSCFLEGPEAPKTPACQAAKQAGGTQAMY
DWNGINQNPSGDNHQAVVPDGTLGGGKA
EFKGFNLAREDWRTTNIVPDANGNYEF
IYKATA
PHATKYFKFYVTKNGWNPTQALKW
SDLELF
GT
VNGNP
QQDAG
KRYH
MTL
KLPQ
GKTG
KHIIF
NVWKR
RSDEEAFYSC
SDVN
FGGD
AAP
PT
AN
PW
KE
IG
SVA
AREN
LP
AG
SSAT
LRV
FDG
QGR
DA
EKHT
VAL
NV
SSQ
TAN
WP
YELA
QKV
NAAS
RAIR
IGV
LKS
QQRA
VS
IV
PV
K
DATA
N
RV
Y
LND
Y
PGY
GFA
IDLN
KG
DGG
VAT
PPA
ANS
W
KEGG
K
YAG
QV
VS
YQ
GR
SYR
CL
QA
HT
AW
AG
AG
W
TPS
TQ
STL
WQ
A
I

**d) Sequência da proteína CV0096 - AAQ57775.1 hypothetical protein CV_0096
[*Chromobacterium violaceum* ATCC 12472]**

MKTA
FALAC
ALSA
AL
A
Q
P
A
S
A
K
T
L
R
F
S
G
Y
D
W
E
V
R
N
Q
G
S
G
P
G
P
N
W
D
E
N
N
V
W
L
D
A
Q
G
R
L
H
L
K
I
A
K
R
G
G
Q
W
T
A
A
E
L
Y
L
P
Q
R
L
G
F
G
R
Y
Q
F
K
L
I
G
R
P
D
R
F
D
P
N
V
V
L
G
F
N
Y
P
T
P
D
V
G
P
D
G
S
N
E
I
D
I
E
F
A
R
W
G
S
P
A
N
P
M
G
N
Y
T
V
Y
P
A
Q
P
G
P
A
P
T
S
Y
F
D
F
A
L
G
G
D
Y
S
T
H
R
F
D
W
N
A
A
V
E
Y
Q
S
L
H
G
F
Q
D
G
D
R
Y
P
I
A
H
W
R
F
A
P
V
D
A
Q
R
V
P
Q
Q
P
L
P
V
H
L
N
L
W
L
F
Q
G
K
P
A
D
D
K
E
V
E
V
V
I
Q
E
F
K
F
V
P
A
A
R

**e) Sequência da proteína CV3323 - AAQ60987.1 carbohydrate-binding protein
[*Chromobacterium violaceum* ATCC 12472]**

MKPHFSIVMMLAGLSTS
SWA
HGT
MEVP
INRTY
SCSK
EGAESP
CTPAC
QEA
KEVR
GGTQAM
YDW
NGINQN
PPGD
NHQ
SVV
PDG
TLC
GGG
QAK
FK
GFL
ART
DW
PTT
NIV
P
DAK
GN
FE
FIY
T
ATMPH
ATKY
FKFY
VT
RNG
W
NP
QPL
K
WS
D
LE
P
FGT
YNG
N
P
PL
DDK
Q
RYH
MT
MKL
PT
GKT
GRH
II
Y
NV
W
K
RS
D
SE
EAFY
SC
SD
VN
FT
NDG
K
PE
PPP
I
SN
P
W
KE
AG
SV
TA
HEN
LP
DK
SS
VT
LR
IF
D
SH
GRD

VESHKVDLSASSGQAANWPYELGVKVNAASQIGRIGVISSKQRAVTINPVRSATANRVWLNE
RYSGYRYQIDIKKGDGGVNPPVPGDEREGVAYTVQIVSYQGRRYRCLQGHTAWTGAGWTP
STQPALWTPA

f) Sequência da proteína CV0348 - AAQ58026.2 conserved hypothetical protein [Chromobacterium violaceum ATCC 12472]

MQYKTGDSAGKHKYPIHSHVETDEETYKKYISDIVSGNISLSRPNNIGQHQKNEKKEDQAKP
KLPESSTPLGNKHESKNTPPKGNEHKKNEKERIERIRKNKEIIKKQAIESNVPVKDFLT
FADIEGGTYNEKAYNKSGAKGLFQFVPTAKAYKINGKEFDPTENTKAALQLYKDNLKAI
EKKKQKNNLEYLSGEQNPNGLDLYLAHQQGAGGYSIIQKVNAGLQSLGEKTKSNLIGNLPSK
DKKKYAAALNDRDLTKEYLKYWRENYKKRQGAIEKELSSVSQPNTEGKNEHAADIKHEVDQKK
PQADVDAVRQEAKDQAKVKQHADAEEAKQAAEQAKAKQRADAEEAKQAAEQAKAKQRADA
EEAKQKAAEQAKAKQRADAEEAKQAAEQAKAKQRADVEAAKQKAAEQAKAKQGVDAEAVKR
KEAEQAQTKIEAVGKEPVQPKQKAAAKPAQKQPAQAKPAQPAGPGNLDAQLAQLAALLGQL
NAALAPLAKPIQITVDVQNGNIVAAVNAANSQQQRRN

g) Sequência da proteína CV2065 - AAQ59745.2 probable beta-N-acetylhexosaminidase [Chromobacterium violaceum ATCC 12472]

MNLPRGPVMVDIAGFALTEAERARLSPHPLVGGVILFRRNFHNIEQLRALTAEIRALRTPHLL
IAVDHEGGRVQRFLDGFTRLPPMRLGEAWDADRDQALKLAETVGYVLAELSACGIDLSFT
PVLDLDWERCAVIGNRAFHRSPEAVSALAEALQQGLGRGGMMSCGKHYPGHGYVEGDSHHLM
PQDDRTLAQIEQDDLVPFARLADAGMGAVMPAHVLYPAVDSQPAGFSKVWLTDILRGRIGFD
GVIFSDALDMAGAAGAGTYVQRADAALAAGCDMVLVCNQPEEADAMLAALAPPQPQPLAERL
ERMAGKSRAEDWQRLIATPDFAAAQAAVRQLAMPKDLAGPQVGEAH

h) Sequência da proteína CV2641 - AAQ60311.1 cell invasion protein [Chromobacterium violaceum ATCC 12472]

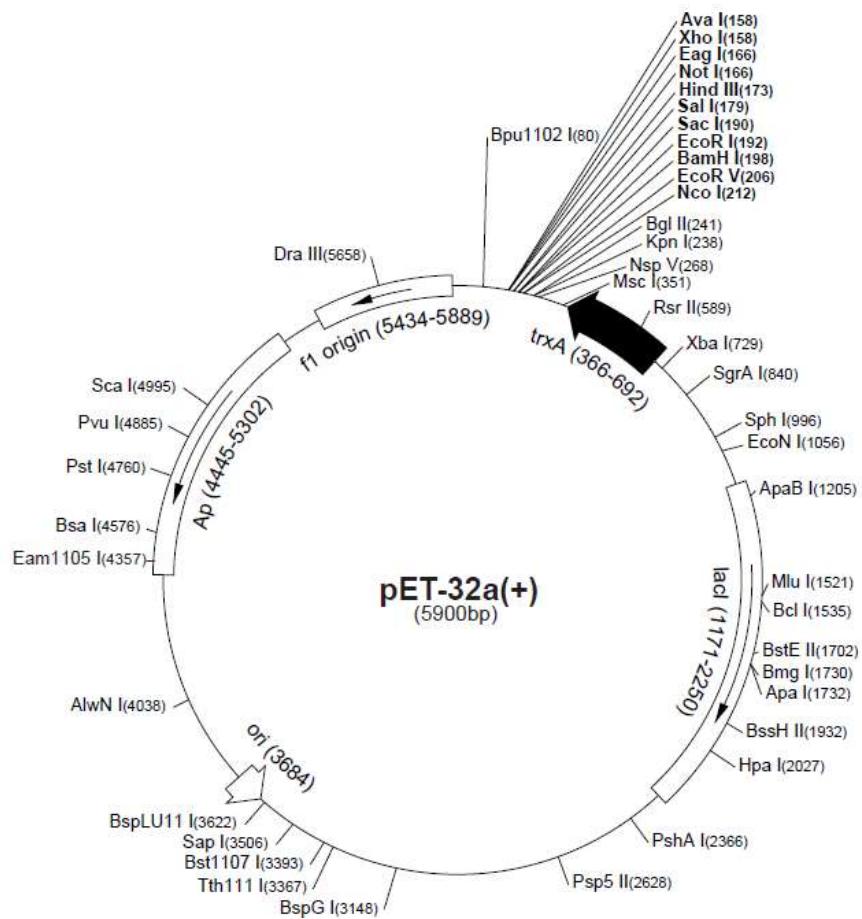
MRAALLAALLLASPLARADCWDYAGTMFNISPDLIYAIQQESGMKPDAIGRNRDGSRDLGL
MQINSALPRLRQLGVNERQLLESPCLSVVVGASILSDMMKRYGYSWEAVGAYNAGLSPNRH
ALRMRYARQVWQRYQKLQAPESARQQP

i) Sequência da proteína CV2601 - AAQ60281.2 peptidoglycan N-acetylmuramoylhydrolase [Chromobacterium violaceum ATCC 12472]

MVWTRGPEPSARAGSEPAASSQPVSSGPETKTAPQDGIAAHLPLAHKISHELDLDPKLLHAI
IKVESGYRPRAISSKGAGQGLMQVMPETGKRGFTDLMNPENNLRAGASYLKWLHNHFENDLE
LAIAGYNAGEGAVKKHGRKIPPPETQNYVKKVMASYSEPESHPLPPSPAPRKTQTPPPA
GKSIPAKLFGLLLSSPLAPS

ANEXO B - MAPA PLASMIDIAL

Os fragmentos sintéticos das sequências de DNA foram克lonados no vetor pET32a(+), utilizando sítios para as endonucleases de restrição BamHI e XhoI, de tal forma a colocar cada sequência em frame com a sequência codificando a proteína Trx.



Fonte: GenOne