

UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL

Prospecção in silico, clonagem e expressão de uma Endoglucanase de *Bacillus bombysepticus*

GEOVANA MONDINI VALVERDE
Ciências Biológicas

Jaboticabal-SP

2026

UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL

Prospecção in silico, clonagem e expressão de uma Endoglucanase de *Bacillus bombysepticus*

GEOVANA MONDINI VALVERDE

Orientador(a): Profa. Dra. Eliana Gertrudes de Macedo Lemos

Coorientador(a): Dra. Camila Cesário Fernandes Sartini

Trabalho apresentado à Faculdade de Ciências Agrárias e veterinárias-UNESP, câmpus de Jaboticabal, para obtenção do título de Bacharel em Ciências Biológicas.

Jaboticabal

1º Semestre de 2026

V215p

Valverde, Geovana Mondini

Prospecção in silico, clonagem e expressão de uma Endoglucanase de *Bacillus bombysepticus* / Geovana Mondini Valverde. -- Jaboticabal, 2026

43 p. : il., tabs., fotos

Trabalho de conclusão de curso (Bacharelado - Ciências Biológicas) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientadora: Eliana Gertrudes de Macedo Lemos

Coorientadora: Camila Cesário Fernandes Sartini

1. Análise enzimática. 2. Bioinformática. 3. Clonagem molecular. 4. Expressão gênica. 5. Produtos biológicos. I. Título.

Geovana Mondini Valverde

Prospecção in silico, clonagem e expressão de uma endoglucanase de *Bacillus bombysepticus*

Trabalho de Conclusão de Curso apresentado à Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, para obtenção do título de Bacharel em **Ciências Biológicas**.

Orientador: Prof. Dr. Eliana G. de Macedo Lemos

Coorientador (se houver): Dr. Camila Cesário Fernandes Sartini

Área de Concentração: Bioquímica de microrganismos

Trabalho aprovado em 10/06/2026

Banca Examinadora:

Documento assinado digitalmente
 **CAMILA CESARIO FERNANDES SARTINI**
Data: 10/06/2026 14:24:35-0300
Verifique em <https://validar.iti.gov.br>

Dr. Camila Cesário Fernandes Sartini

Unesp - Faculdade de Ciências Agrárias e Veterinárias - campus de Jaboticabal

Documento assinado digitalmente
 **MICHELLI INACIO GONCALVES FUNNICELLI**
Data: 10/06/2026 14:29:34-0300
Verifique em <https://validar.iti.gov.br>

Dr. Michelli Inácio Gonçalves Funniceili

Unesp - Faculdade de Ciências Agrárias e Veterinárias - campus de Jaboticabal

Documento assinado digitalmente
 **BARBARA BONFA BUZZO**
Data: 10/06/2026 14:54:55-0300
Verifique em <https://validar.iti.gov.br>

Dr. Bárbara Bonfá Buzzo

Unesp - Faculdade de Ciências Agrárias e Veterinárias - Campus de Jaboticabal

Documento assinado digitalmente
 **HIRASILVA BORBA**
Data: 11/06/2026 09:05:46-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Hirasilva Borba
Chefe do Departamento

AGRADECIMENTOS

À Universidade Estadual Paulista “Júlio de Mesquita Filho” Faculdade de Ciências Agrárias e Veterinárias (UNESP–FCAV), expresso minha sincera gratidão pela formação acadêmica de excelência que me foi proporcionada ao longo desta trajetória. Agradeço a todos(as) os(as) docentes com quem tive a oportunidade de aprender, pelos valiosos ensinamentos, pela dedicação ao ensino e pela contribuição essencial para a minha formação profissional.

Registro também meus agradecimentos à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo apoio financeiro concedido por meio do Processo FAPESP nº 2023/07808-5, e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), cujos auxílios financeiros foram fundamentais para a minha permanência na universidade e para o desenvolvimento das minhas iniciações científicas.

Agradeço aos laboratórios nos quais atuei durante a graduação, em especial ao Laboratório de Bioquímica de Microrganismos e Plantas (LBMP), onde realizei minhas atividades de pesquisa e tive a oportunidade de ser contemplada com bolsas de iniciação científica. Estendo meu reconhecimento a todos(as) os(as) alunos(as) e pesquisadores(as) que compartilharam comigo esse ambiente e que, direta ou indiretamente, contribuíram para a minha formação acadêmica e científica.

Dirijo um agradecimento especial à minha orientadora, Profa. Dra. Eliana Gertrudes de Macedo Lemos, pela orientação dedicada, pelos ensinamentos e pela confiança depositada ao longo deste percurso. Agradeço também à minha coorientadora, Profa. Dra. Camila Cesário Fernandes Sartini, pelo apoio técnico e acadêmico que foi essencial para a realização deste trabalho.

Sou igualmente grata por todas as experiências vivenciadas durante essa trajetória, tanto as positivas quanto as desafiadoras, pois todas contribuíram para meu aprendizado, amadurecimento e crescimento pessoal e profissional.

Aos(às) colegas com quem compartilhei esta etapa da graduação, agradeço a convivência, pelas trocas de conhecimento e pelos momentos compartilhados ao longo desses anos.

Por fim, expresso minha profunda gratidão à minha família, base fundamental para que eu pudesse ingressar e permanecer em uma universidade pública.

SUMÁRIO

RESUMO	XI
ABSTRACT	XII
1. INTRODUÇÃO.....	14
2. OBJETIVOS.....	17
2.1 Objetivo geral	17
2.2 Objetivos específicos.....	17
3. MATERIAL E MÉTODOS	18
3.1 Caracterização in silico da endo-1,4- β -glucanase prospectada.....	18
3.2 Avaliação da capacidade celulolítica da bactéria <i>Bacillus bombysepticus</i>	19
3.3 Amplificação do gene da endoglucanase	19
3.4 Vetor de expressão pET-SUMO.....	20
3.5 Construção do plasmídeo recombinante	20
3.6 Transformação em <i>Escherichia coli</i> ER2566	20
3.7 Cultivo e armazenamento dos clones	21
3.8 Extração plasmidial e sequenciamento	21
3.9 Indução e expressão da proteína recombinante.....	21
3.10 Extração das proteínas recombinantes	21
3.11 Análise da expressão proteica por SDS PAGE	22
3.12 Confirmação da expressão por <i>Western Blotting</i>	22
3.13 Purificação da proteína recombinante por cromatografia de afinidade e exclusão por tamanho	23
4. RESULTADOS.....	23
4.1 Caracterização in silico da endo-1,4- β -glucanase prospectada.....	23
4.2 Avaliação da capacidade celulolítica da bactéria <i>Bacillus bombysepticus</i>	26
4.3 Extração de DNA da bactéria <i>Bacillus bombysepticus</i>	27
4.4 Amplificação do gene de endoglucanase por PCR	27
4.5 Clonagem plasmidial e sequenciamento	28
4.6 Indução e expressão da proteína recombinante.....	30
5. DISCUSSÃO	34
6. CONCLUSÃO	37
REFERÊNCIAS.....	38

LISTA DE FIGURAS

- Figura 1-** Alinhamento múltiplo da sequência da endo-1,4- β -glucanase prospectada (M9428-19985) com endoglucanases pertencentes à família das glicosil hidrolases GH5 obtidas do banco de dados CAZy. As regiões coloridas representam resíduos conservados entre as sequências, indicando a preservação de regiões estrutural e funcionalmente importantes. **24**
- Figura 2-** Predição dos domínios conservados da proteína M9428-19985 utilizando a plataforma InterPro (PAYSAN-LAFOSSE et al., 2023). A proteína foi classificada como pertencente à família das glicosil hidrolases GH5, apresentando os domínios Glyco_hydro_5 e Cellulase, além da classificação na superfamília Glycoside Hydrolase e da identificação de sítios conservados característicos dessa família enzimática. **26**
- Figura 3-** Avaliação da capacidade celulolítica da bactéria *Bacillus bombysepticus*. Cultivo pelo período de 24 horas em Meio de cultura LB sólido suplementado com substrato Carboximetilcelulose (CMC) (1%). Os halos de hidrólise foram revelados pela aplicação de Vermelho Congo (1%) por 15 minutos, seguido de lavagem das placas com NaCl (1M) por 10 minutos. **26**
- Figura 4-** Gel de eletroforese em agarose 1% do DNA genômico extraído de *Bacillus bombysepticus*. M: marcador de peso molecular de DNA 1 kb (K9-1000L, Kasvi, São José dos Pinhais, PR, Brasil); pistas 1 e 2 correspondem às amostras de DNA genômico obtidas após o procedimento de extração. A concentração de DNA das amostras foi estimada por comparação visual da intensidade das bandas com o marcador de peso molecular. **27**
- Figura 5-** PCR em gradiente utilizada para otimização da temperatura de anelamento dos primers para amplificação do gene de endoglucanase e posterior clonagem no vetor pETSUMO. Os produtos da PCR foram analisados por eletroforese em gel de agarose 1%. M: marcador de peso molecular de DNA 1 kb (K9-1000L, Kasvi, São José dos Pinhais, PR, Brasil). As colunas 1–12 correspondem às reações realizadas nas temperaturas de anelamento entre 54 e 64 °C, observando-se a amplificação de um fragmento com tamanho esperado de aproximadamente 1,5 kb. **28**
- Figura 6-** Análise dos clones bacterianos por PCR de colônia após clonagem do gene de endoglucanase no vetor pETSUMO. Os produtos da amplificação foram analisados por eletroforese em gel de agarose 1%. M: marcador de peso molecular de DNA 1 kb (K9-1000L, Kasvi, São José dos Pinhais, PR, Brasil). As colunas 1–10 correspondem aos clones bacterianos analisados. Observa-se a amplificação do fragmento esperado de aproximadamente 1,5 kb em todos os clones, indicando a presença do inserto. O clone 7 apresentou uma banda adicional de maior tamanho, sugerindo amplificação inespecífica ou presença de DNA plasmidial residual. ... **29**
- Figura 7 -** Resultado da análise de similaridade da sequência consenso (contig) obtida após o sequenciamento do clone recombinante por meio da ferramenta BLASTn (NCBI). Os principais alinhamentos apresentaram cobertura de consulta (Query Cover) de 100%, identidade de aproximadamente 98,3% e valor E igual a 0,0, evidenciando elevada similaridade com sequências do gênero *Bacillus*, incluindo *Bacillus bombysepticus*. **30**
- Figura 8-** Análise da expressão da β -endoglucanase recombinante em *Escherichia coli*. (A) Perfil de expressão avaliado por SDS-PAGE após indução com 0,1 mM de IPTG nos tempos de 0, 2, 4, 6 e 24 horas. A seta indica a banda correspondente à proteína recombinante (~50 kDa). (B) Western blot das mesmas amostras confirmando a identidade da proteína recombinante por imunodeteção. M: marcador de massa molecular Precision Plus Protein™ Dual Color Standards (Bio-Rad Laboratories, Hercules, CA, EUA). **31**
- Figura 9-** Avaliação da distribuição da β -endoglucanase recombinante entre as frações solúvel e insolúvel após expressão em *Escherichia coli*. (A) SDS-PAGE das amostras analisadas. (B) Western blot das mesmas amostras confirmando a presença da proteína recombinante. M: marcador de massa molecular Precision Plus Protein™ Dual Color Standards (Bio-Rad Laboratories, Hercules, CA, EUA). Poços: (1) sobrenadante após 6 h de indução; (2) amostra antes da indução (0 h); (3) pellet após 6 h de indução; (4) poço vazio; (5) amostra correspondente ao tempo 0 h; (6) pellet após 24 h de indução. **32**
- Figura 10-** Purificação da β -endoglucanase recombinante por cromatografia de afinidade em resina Ni-NTA. (A) SDS-PAGE das diferentes etapas da purificação: FT (flow-through), W (lavagem) e frações eluídas com 20, 50, 100, 200, 500 mM e 1 M de imidazol. A seta indica a proteína

recombinante (~50 kDa). (B) Western blot das mesmas frações confirmando a identidade da proteína recombinante ao longo do processo de purificação. M: marcador de massa molecular Precision Plus Protein™ Dual Color Standards (Bio-Rad Laboratories, Hercules, CA, EUA).....**33**

Figura 11- Avaliação qualitativa das frações eluídas durante a purificação por cromatografia de afinidade utilizando o reagente de Bradford. As frações correspondem às eluições com 20, 50, 100, 200, 500 mM e 1 M de imidazol. A maior intensidade de coloração foi observada na fração eluída com 200 mM de imidazol, indicando maior concentração relativa de proteínas nessa condição.**34**

LISTA DE QUADROS

Quadro 1- Resultados das buscas de similaridade da sequência M9428-19985 em diferentes bancos de dados, apresentando a proteína mais similar identificada, o comprimento das sequências comparadas, a identidade e a cobertura dos alinhamentos obtidos.....	24
---	-----------

RESUMO

A prospecção de novas enzimas celulolíticas representa uma estratégia promissora para ampliar o repertório de biocatalisadores empregados na conversão da biomassa lignocelulósica em biocombustíveis e outros bioprodutos. Nesse contexto, este trabalho teve como objetivo prospectar *in silico*, clonar, expressar e purificar uma endo-1,4- β -glucanase codificada pelo genoma de *Bacillus bombysepticus*, visando à obtenção da proteína recombinante para futuras caracterizações bioquímicas. Inicialmente, foi realizada a caracterização *in silico* da sequência gênica por meio de análises de similaridade, alinhamento múltiplo e identificação de domínios conservados. A atividade celulolítica da bactéria foi avaliada em meio contendo carboximetilcelulose, seguida pela extração do DNA genômico, amplificação do gene por reação em cadeia da polimerase, clonagem no vetor de expressão pET-SUMO, transformação em *Escherichia coli* ER2566, confirmação por sequenciamento, expressão heteróloga e purificação por cromatografia de afinidade. As análises *in silico* classificaram a proteína como uma endoglucanase pertencente à família das glicosil hidrolases GH5, enquanto o ensaio em meio sólido confirmou a capacidade celulolítica da bactéria. A clonagem do gene foi confirmada por PCR de colônia e sequenciamento. A expressão da proteína recombinante foi detectada por SDS-PAGE e Western blotting, evidenciando uma banda compatível com a massa molecular esperada, embora também tenham sido observados fragmentos de menor massa molecular, sugerindo possível degradação proteolítica ou processamento incompleto da proteína. A purificação por cromatografia de afinidade permitiu recuperar a proteína recombinante, com maior rendimento na eluição com 200 mM de imidazol. Conclui-se que o estudo permitiu identificar e clonar uma nova endoglucanase de *B. bombysepticus* e obter resultados iniciais de sua expressão heteróloga e purificação, estabelecendo uma base para otimizações futuras.

Palavras-chave: Análise enzimática; Bioinformática, Clonagem molecular; Expressão gênica; Produtos biológicos.

ABSTRACT

The search for new cellulolytic enzymes represents a promising strategy to expand the repertoire of biocatalysts used in converting lignocellulosic biomass into biofuels and other bioproducts. In this context, this study aimed to prospect *in silico*, clone, express, and purify an endo-1,4- β -glucanase encoded by the *Bacillus bombysepticus* genome, with the goal of obtaining the recombinant protein for future biochemical characterization. Initially, the gene sequence was characterized *in silico* through similarity analyses, multiple sequence alignment, and the identification of conserved domains. The bacterium's cellulolytic activity was evaluated in a medium containing carboxymethylcellulose, followed by genomic DNA extraction, gene amplification via polymerase chain reaction (PCR), cloning into the pET-SUMO expression vector, transformation into *Escherichia coli* ER2566, sequence confirmation, heterologous expression, and purification via affinity chromatography. *In silico* analyses classified the protein as an endoglucanase belonging to the GH5 glycosyl hydrolase family, while the solid-medium assay confirmed the bacterium's cellulolytic capability. Gene cloning was confirmed by colony PCR and sequencing. Recombinant protein expression was detected via SDS-PAGE and Western blot, revealing a band consistent with the expected molecular mass; however, lower molecular mass fragments were also observed, suggesting possible proteolytic degradation or incomplete protein processing. Affinity chromatography purification enabled the recovery of the recombinant protein, with the highest yield obtained during elution with 200 mM imidazole. In conclusion, this study successfully identified and cloned a new endoglucanase from *B. bombysepticus* and yielded initial results regarding its heterologous expression and purification, establishing a foundation for future optimization.

Keywords: Molecular cloning; gene expression; enzymatic analysis; biological products; bioinformatics.

1. INTRODUÇÃO

A biomassa lignocelulósica constitui uma matéria-prima renovável, abundante e sustentável, com grande potencial de aplicação em processos fermentativos voltados à produção de bioprodutos. Estruturalmente, é composta principalmente por celulose, hemicelulose e lignina (BASERA; CHAKRABORTY; SHARMA, 2024), cujas proporções relativas variam de acordo com o tipo de biomassa, os métodos de cultivo e as condições climáticas (KUMAR et al., 2026). Essa organização estrutural confere elevada recalcitrância ao material, limitando a acessibilidade aos polissacarídeos e reduzindo a eficiência dos processos de conversão em bioprodutos e biocombustíveis (CHAKRABORTY et al., 2024).

A lignina é reconhecida como um dos principais fatores responsáveis pela recalcitrância da biomassa lignocelulósica. Por essa razão, estratégias recentes de pré-tratamento têm sido direcionadas à modificação e à remoção seletiva desse polímero, visando não apenas reduzir a resistência do material à conversão, mas também possibilitar sua recuperação e valorização (YOO et al., 2020). Atualmente, existem diversas técnicas de pré-tratamento baseadas em métodos físicos, químicos, biológicos e suas combinações que vêm sendo desenvolvidas com o objetivo de remover a lignina, aumentar a área superficial acessível da biomassa e reduzir a cristalinidade da celulose, favorecendo assim o processo de digestão e contribuindo para uma maior eficiência na produção de bioprodutos (ZHANG, 2026). Apesar dos avanços tecnológicos, o alto custo dos processos de pré-tratamento e de produção de coquetéis enzimáticos ainda representa um dos principais entraves à viabilidade econômica da conversão da biomassa lignocelulósica (SINGH NEE' NIGAM et al., 2009).

Diante desse cenário, o pré-tratamento biológico tem se destacado como uma alternativa sustentável aos métodos físicos, químicos e físico-químicos, por empregar microrganismos ou enzimas capazes de degradar seletivamente os componentes da biomassa sob condições brandas de temperatura, pressão e pH. Em comparação aos processos convencionais, essa estratégia reduz o consumo de energia, dispensa o uso de reagentes químicos agressivos e minimiza a formação de compostos inibidores que podem prejudicar as etapas subsequentes de hidrólise enzimática e fermentação (SINDHU; BINOD; PANDEY, 2016; ZHANG et al., 2025).

Entre as abordagens biológicas, a hidrólise enzimática é um dos métodos mais eficientes para converter celulose em açúcares simples. Esse processo envolve o uso de enzimas celulolíticas, como as celulasas, que catalisam a quebra das ligações β -1,4-glicosídicas na celulose, liberando glicose (WOŻNIAK et al., 2025). Bactérias e fungos produzem celulasas, essas enzimas são capazes de degradar o polímero recalcitrante (HOUFANI et al., 2023).

As celulasas pertencem, em sua maioria, ao grupo das glicosil hidrolases (Glycoside Hydrolases – GHs), uma das principais classes de enzimas carboidrato-ativas (CAZymes). Essas enzimas catalisam a hidrólise de ligações glicosídicas presentes em polissacarídeos, oligossacarídeos e glicoconjugados, sendo classificadas em famílias de acordo com a similaridade de suas sequências de aminoácidos, sua organização estrutural e seu mecanismo catalítico. Essa classificação, organizada no banco de dados CAZy (Carbohydrate-Active enZymes Database), constitui uma importante ferramenta para estudos evolutivos, funcionais e para a prospecção de novos biocatalisadores (LOMBARD et al., 2014; CAZY, 2025). As principais celulasas envolvidas na degradação da celulose são as endoglucanases (EC 3.2.1.4), as exoglucanases ou celobiohidrolases (EC 3.2.1.91 e EC 3.2.1.74) e as β -glicosidases (EC 3.2.1.21), que juntas desempenham papel essencial na valorização de resíduos lignocelulósicos e atuam em ampla faixa de temperatura e pH, além de apresentarem notável estabilidade (FERRANDO et al., 2025; ZHU et al., 2025). Devido a essa versatilidade, as enzimas celulolíticas têm despertado grande interesse em diversas aplicações industriais, como na produção de biocombustíveis, no setor agroindustrial, na indústria de alimentos e nas indústrias de papel e celulose, sendo desejáveis enzimas que apresentem especificidade, estabilidade e alta eficiência catalítica (RONO et al., 2025).

Entre essas enzimas, as endoglucanases merecem destaque devido ao seu papel central no início da hidrólise enzimática, uma vez que atuam internamente nas cadeias de celulose, promovendo a clivagem das ligações β -1,4-glicosídicas e a degradação de substratos como β -glucanos e carboximetilcelulose. São amplamente distribuídas na natureza, sendo encontradas em bactérias, fungos e insetos, nos quais desempenham papéis importantes na degradação da parede celular vegetal, além de possuírem múltiplas aplicações biotecnológicas, sobretudo na valorização de subprodutos agroindustriais e na produção de glicose destinada à fermentação em bioetanol (LV et al., 2025; ESCUDER-RODRÍGUEZ et al., 2022).

Diante do crescente interesse industrial por endoglucanases com elevada atividade catalítica e estabilidade em diferentes condições de processo, torna-se necessária a identificação de novas enzimas com essas características. Nesse contexto, a prospecção *in silico* tem se consolidado como ferramenta essencial para a descoberta de novos biocatalisadores, pois permite selecionar, por meio de análises computacionais, candidatos com maior potencial de interesse para posterior validação experimental, reduzindo tempo, custos e o número de candidatos a serem testados (ARIAEENEJAD et al., 2024). Essas abordagens incluem ferramentas de bioinformática, análises de similaridade de sequências, estudos filogenéticos e modelagem estrutural, que auxiliam na identificação e priorização de enzimas (NOBLE; CROWE; O'DONOHUE, 2021). Nesse sentido, o avanço das tecnologias de sequenciamento e o crescente número de genomas microbianos depositados em bancos de dados públicos têm ampliado a capacidade de explorar computacionalmente a diversidade genética de microrganismos, consolidando estratégias de prospecção genômica (*genome mining*) como ferramentas eficientes para identificar genes codificadores de enzimas ainda não caracterizadas e acelerar a descoberta de novas glicosil hidrolases (NGUYEN, 2024; ARIAENEJAD et al., 2024).

Uma vez identificados os genes candidatos por meio de análises computacionais, a clonagem molecular e a expressão heteróloga constituem etapas essenciais para validar experimentalmente as predições obtidas *in silico*. A produção recombinante possibilita a obtenção de quantidades suficientes da proteína de interesse para sua caracterização bioquímica, determinação de parâmetros cinéticos e avaliação da estabilidade, permitindo converter predições genômicas em biocatalisadores efetivamente caracterizados (MAXEINER et al., 2023; AVALON et al., 2024).

Nesse contexto, a prospecção de microrganismos capazes de produzir celulasas tem recebido crescente atenção devido à sua relevância para a conversão da biomassa lignocelulósica em açúcares fermentáveis. As bactérias do gênero *Bacillus* se destacam nesse cenário por produzirem diversos compostos biologicamente ativos, entre eles enzimas hidrolíticas (LEGEIN et al., 2020; WEISBURG et al., 1991), além de apresentarem características propícias ao cultivo industrial, como taxas de crescimento elevadas, eficiente secreção de proteínas e ampla adaptabilidade (ROBSON; CHAMBLISS, 1989; SIMONE; PALVA, 1993; REN et al., 2022). Diversos estudos já identificaram, em espécies desse gênero, genes que participam ativamente

da degradação da celulose (QADAR et al., 2009; JOO et al., 2012; PAN et al., 2021; AFFANDI et al., 2021; AVCI et al., 2021; PAN et al., 2023).

Dentre as espécies do gênero, *Bacillus bombysepticus*, originalmente isolada por Hartman (1931) em cadáveres de larvas de *Bombyx mori* (bicho-da-seda), desperta particular interesse. Análises filogenéticas baseadas no gene 16S rRNA revelaram sua proximidade taxonômica com outras espécies do grupo, como *Bacillus cereus* e *Bacillus thuringiensis*, e análises de sintonia de aminoácidos e nucleotídeos entre *B. bombysepticus* e *B. thuringiensis* evidenciaram identidade genômica de até 96,49% (CHENG et al., 2014). Esse elevado grau de similaridade sugere que *B. bombysepticus* pode abrigar genes associados à degradação de polissacarídeos vegetais, ainda pouco explorados, o que a torna uma candidata promissora para a prospecção de novas celulasas.

Diante desse cenário, a identificação e a caracterização de novas endoglucanases provenientes de espécies bacterianas representam uma estratégia promissora para ampliar o repertório de enzimas disponíveis para aplicações industriais. Assim, este trabalho teve como objetivo investigar uma enzima codificada pelo genoma de *Bacillus bombysepticus*, por meio de sua prospecção *in silico*, clonagem, expressão heteróloga e purificação, visando à obtenção da proteína recombinante para futuras caracterizações.

2. OBJETIVOS

2.1 Objetivo geral

Prospectar, clonar, expressar e purificar uma endoglucanase codificada pelo genoma de *Bacillus bombysepticus*, visando à obtenção da proteína recombinante por expressão heteróloga em *Escherichia coli*.

2.2 Objetivos específicos

- ✓ Identificar o gene da enzima endoglucanase no genoma da bactéria *Bacillus bombysepticus*, isolada da rizosfera de *Mimosa pudica* L.;
- ✓ Amplificar o gene da enzima a partir da Reação em cadeia pela Polimerase;
- ✓ Clonar o gene da endoglucanase em vetor de expressão apropriado para produção heteróloga;
- ✓ Expressar a enzima recombinante em *Escherichia coli*;
- ✓ Purificar a proteína recombinante.

3. MATERIAL E MÉTODOS

3.1 Caracterização in silico da endo-1,4- β -glucanase prospectada

O genoma da bactéria *Bacillus bombysepticus* JAB01 foi sequenciado usando as tecnologias Illumina NovaSeq e Oxford Nanopore (MinION) derivado de estudos prévios pelo grupo de pesquisa do Laboratório de Bioquímica de Microrganismos e Plantas (LBMP), a bactéria foi isolada da rizosfera da planta *Mimosa pudica* L. no Campus da UNESP- FCAV de Jaboticabal. A sequência codificadora da enzima endo-1,4- β -glucanase foi obtida a partir do genoma da bactéria *Bacillus bombysepticus* JAB01. A sequência gênica obtida foi empregada para a identificação de uma Open Reading Frame (ORF) putativa correspondente ao gene de interesse. A seleção dessa ORF foi baseada na presença de características estruturais e funcionais compatíveis com enzimas da família das glicosil hidrolases (GHs), especialmente da família GH5, conforme descrito no banco de dados CAZy (Carbohydrate-Active enZymes Database) (Lombard et al., 2014).

A sequência de aminoácidos correspondente à ORF selecionada foi submetida à análise de similaridade utilizando o algoritmo BLASTp (Basic Local Alignment Search Tool for proteins), disponível no National Center for Biotechnology Information (NCBI) (Altschul et al., 1997), com o objetivo de identificar proteínas homólogas previamente caracterizadas. As buscas foram realizadas utilizando os bancos de dados UniProtKB/Swiss-Prot, Protein Data Bank (PDB) e Patented Protein Sequences (PAT), sendo avaliados parâmetros como porcentagem de identidade, cobertura da sequência, valor E (*E-value*) e similaridade com proteínas funcionalmente caracterizadas. Essas análises permitiram verificar o grau de conservação da sequência e sua relação com enzimas pertencentes à família das endoglucanases.

Adicionalmente, foi realizado o alinhamento múltiplo da sequência proteica com sequências homólogas pertencentes à família GH5 obtidas do banco de dados CAZy, utilizando o software BioEdit (Hall, 1999), com a finalidade de identificar regiões conservadas e possíveis resíduos envolvidos na atividade catalítica da enzima. A identificação de domínios conservados e a classificação funcional da proteína foram conduzidas utilizando as bases de dados InterPro (Blum et al., 2021) e Pfam (Mistry et al., 2021), permitindo confirmar a presença de domínios característicos de glicosil hidrolases da família GH5 e sua associação com a atividade celulolítica.

3.2 Avaliação da capacidade celulolítica da bactéria *Bacillus bombysepticus*

A capacidade celulolítica da bactéria *Bacillus bombysepticus* foi avaliada por meio do cultivo em meio sólido contendo carboximetilcelulose (CMC), conforme metodologia descrita por Kasana et al. (2008), com adaptações. A bactéria foi previamente cultivada em meio Luria-Bertani (LB) líquido a 30°C por 16 horas a 180 rpm. Posteriormente, foi inoculada em placas contendo meio LB sólido suplementado com carboximetilcelulose (CMC) a 1% (m/v), empregada como substrato celulósico para a detecção da atividade celulolítica. As placas foram incubadas em estufa bacteriológica a 37°C por 24 horas, em condições aeróbias.

Após o período de incubação, as placas foram coradas com solução de Vermelho Congo a 1% (m/v), permanecendo em contato com o corante por 15 minutos à temperatura ambiente, com o objetivo de permitir a ligação do corante ao polissacarídeo não degradado. Em seguida, o excesso de corante foi removido, e as placas foram lavadas com solução de Cloreto de Sódio (NaCl) 1M por 10 minutos, a fim de promover o contraste entre as regiões contendo celulose intacta e as áreas onde ocorreu a hidrólise do substrato.

A atividade celulolítica foi determinada por meio da observação da formação de halos claros ao redor das colônias bacterianas, indicando a degradação da carboximetilcelulose pela ação de enzimas celulolíticas extracelulares. A presença desses halos foi utilizada como indicativo da capacidade da bactéria em produzir enzimas capazes de degradar substratos celulósicos.

3.3 Amplificação do gene da endoglucanase

A bactéria *Bacillus bombysepticus* foi cultivada em meio de cultura Luria-Bertani (LB) líquido sob agitação de 180 rpm a 30°C por aproximadamente 16 horas. Após o cultivo, a cultura foi submetida à centrifugação a $12.000 \times g$ por 10 minutos a 4°C. O sobrenadante foi descartado e o pellet bacteriano obtido foi utilizado para a extração do DNA genômico. A extração foi realizada utilizando o Wizard® Genomic DNA Purification Kit, conforme as instruções do fabricante (Promega, 2019). Posteriormente, as amostras foram quantificadas utilizando o Qubit Fluorometer 2.0 em conjunto com o Qubit dsDNA BR Assay Kit, de acordo com as recomendações do fabricante (Thermo Fisher Scientific, 2015).

O gene codificador da endoglucanase identificado no genoma de *Bacillus bombysepticus* foi amplificado por reação em cadeia da polimerase (PCR). Para a

amplificação do gene foram utilizados primers específicos. Os parâmetros físico-químicos dos oligonucleotídeos, incluindo temperatura de melting (T_m), formação de estruturas secundárias, auto complementaridade e formação de dímeros, foram avaliados utilizando a ferramenta online PCR Primer Stats, disponível no Sequence Manipulation Suite (Stothard, 2000), assegurando condições adequadas para amplificação eficiente do gene de interesse. Para otimização das condições de anelamento dos primers, foi conduzido um gradiente de temperatura variando de 54°C a 64°C. As reações de PCR foram feitas em termociclador e os produtos amplificados foram analisados por eletroforese em gel de agarose a 1% (p/v) para confirmação do tamanho esperado do fragmento amplificado.

3.4 Vetor de expressão pET-SUMO

O vetor de expressão pET-SUMO (Invitrogen™, Thermo Fisher Scientific), contendo o promotor T7, foi utilizado para a clonagem do gene de interesse, permitindo sua expressão heteróloga em sistemas bacterianos. A ligação do fragmento amplificado ao vetor foi realizada conforme o protocolo descrito no manual do Champion™ pET SUMO Protein Expression System (Invitrogen, 2010).

3.5 Construção do plasmídeo recombinante

A construção do plasmídeo recombinante foi realizada por meio da ligação do fragmento amplificado por PCR ao vetor de expressão pET-SUMO utilizando a enzima T4 DNA ligase. A reação de ligação foi preparada em volume final de 10 µL contendo 1,0 µL do inserto (121 ng/µL), 2,0 µL do vetor pET-SUMO (25 ng/µL), 1,0 µL de tampão de ligação 10×, 1,0 µL da enzima T4 DNA ligase e 5,0 µL de água ultrapura. A proporção molar entre inserto e vetor foi ajustada para 3:1. Em seguida, a mistura foi incubada a 15 °C por aproximadamente 16 horas (overnight), conforme as recomendações do fabricante.

3.6 Transformação em *Escherichia coli* ER2566

O plasmídeo recombinante foi utilizado para transformação em células competentes de *Escherichia coli* ER2566 por meio do método de choque térmico. Inicialmente as células foram incubadas em gelo por 2 minutos, seguidas de incubação a 42°C por 90 segundos e retorno ao gelo por mais 2 minutos.

Posteriormente foram adicionados 970 µL de meio SOC ((triptona 2% p/v, extrato de levedura 0,5% p/v, NaCl 10 mM, KCl 2,5 mM, MgCl₂ 10 mM, MgSO₄ 10 mM e glicose 20 mM) estéril, e as células foram incubadas sob agitação orbital (150 rpm) a 37°C por 1

h e 30 minutos para recuperação celular. Após a etapa de recuperação, alíquotas de 50 µL da cultura foram plaqueadas em placas de Petri contendo meio LB sólido suplementado com canamicina (50 µg/mL) e incubadas a 37°C por aproximadamente 16 horas.

3.7 Cultivo e armazenamento dos clones

As colônias obtidas foram inoculadas em meio LB líquido contendo canamicina (50 µg/mL) e incubadas sob agitação a 37°C por aproximadamente 16 horas. Alíquotas das culturas foram misturadas com glicerol estéril até concentração final de 20% (v/v) e armazenadas a -80°C para preservação dos clones.

3.8 Extração plasmidial e sequenciamento

Os clones obtidos foram submetidos à extração de DNA plasmidial utilizando o Kit Wizard® Plus Minipreps DNA Purification System. Os plasmídeos purificados foram posteriormente encaminhados para sequenciamento pelo método de Sanger, com o objetivo de confirmar a presença e a correta inserção do gene da endoglucanase no vetor de expressão pET-SUMO.

3.9 Indução e expressão da proteína recombinante

A partir do estoque em glicerol, células de *Escherichia coli* ER2566 contendo o plasmídeo recombinante foram inoculadas em 50 mL de meio LB suplementado com canamicina (50 µg/mL) e incubadas sob agitação orbital (200 rpm) a 37°C por 16 horas. Em seguida, uma alíquota correspondente a 1% da cultura foi transferida para 200 mL de meio LB contendo o mesmo antibiótico e incubada nas mesmas condições até atingir densidade óptica (DO_{600 nm}) entre 0,4 e 0,6.

A expressão da proteína recombinante foi induzida pela adição de IPTG na concentração final de 0,1 mM. A cultura foi incubada a 28°C e as amostras foram coletadas nos tempos de 0, 2, 4, 6, 8, 10 e 24 horas após a indução. Ao final do experimento, a cultura foi centrifugada a 9000 × g por 30 minutos a 4°C, e o pellet celular foi armazenado a -20°C para posterior extração proteica.

3.10 Extração das proteínas recombinantes

O pellet celular foi ressuspenso em tampão de lise composto por 10 mM Tris-HCl (pH 8,0), 100 mM NaCl e 10% (v/v) de glicerol. A lise celular foi realizada por ultrassonicação utilizando o equipamento LB-750 (Labometric, Leiria, Portugal), com 15 ciclos de 10 pulsos intercalados por intervalos de 10 segundos.

Após a lise, o material foi centrifugado a $17000 \times g$ por 20 minutos a 4°C. O pellet foi submetido a dois ciclos adicionais de ultrassonicação, seguidos de centrifugação nas mesmas condições. As frações solúveis (sobrenadantes) e insolúveis (pellets) obtidas em cada etapa foram utilizadas nas análises posteriores.

3.11 Análise da expressão proteica por SDS PAGE

As frações proteicas foram analisadas por eletroforese em gel de poliacrilamida contendo SDS (SDS-PAGE), conforme descrito por Laemmli (1970). Foi utilizado gel de separação a 12% e gel de empilhamento a 5%.

As amostras foram misturadas ao tampão de amostra contendo SDS e β -mercaptoetanol, desnaturadas a 95°C por 5 minutos e aplicadas juntamente com marcador de massa molecular. A corrida eletroforética foi realizada em sistema vertical utilizando tampão contendo 25 mM Tris, 192 mM glicina e 0,1% SDS, sob voltagem constante de 120 V.

Após a separação, os géis foram corados com Coomassie Brilliant Blue R-250 por aproximadamente 1 hora e posteriormente descorados em solução contendo 10% de ácido acético e 10% de etanol até a visualização das bandas proteicas. A análise permitiu verificar a distribuição da proteína recombinante entre as frações solúvel e insolúvel.

3.12 Confirmação da expressão por *Western Blotting*

As proteínas separadas por SDS-PAGE foram transferidas para membrana de PVDF previamente ativada em metanol, utilizando sistema de transferência úmida. Após a transferência, a membrana foi bloqueada com solução contendo 5% de leite desnatado em tampão TBS-T por aproximadamente 1 hora em temperatura ambiente. Em seguida, a membrana foi incubada com o anticorpo primário anti-histidina diluído em solução de bloqueio, por 2 horas em temperatura ambiente. Após sucessivas lavagens com TBS-T, realizou-se incubação com o anticorpo secundário conjugado à enzima peroxidase por 1 hora em temperatura ambiente. A detecção foi realizada utilizando reagente quimioluminescente, e as imagens foram obtidas em sistema apropriado para análise da intensidade das bandas, permitindo confirmar a expressão da proteína recombinante.

3.13 Purificação da proteína recombinante por cromatografia de afinidade e exclusão por tamanho

A fração solúvel contendo a endoglucanase recombinante foi inicialmente submetida à purificação por cromatografia de afinidade por íons metálicos imobilizados (Immobilized Metal Affinity Chromatography – IMAC), utilizando resina Ni-NTA (Qiagen, Hilden, Alemanha). A eluição da proteína foi realizada por meio de um gradiente crescente de imidazol, com concentrações variando de 10 a 1000 mM.

As frações eluídas foram submetidas à etapa de dessalinização (desalting), com o objetivo de remover o imidazol e promover a troca do tampão, minimizando possíveis interferências nas análises posteriores.

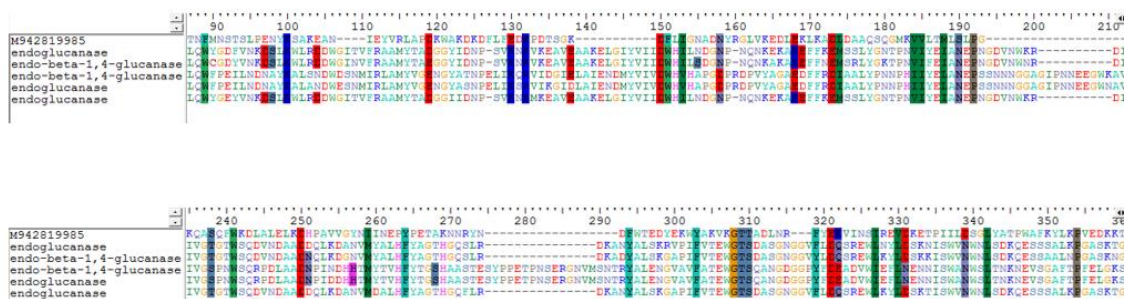
Como etapa adicional de purificação, as frações de interesse foram submetidas à cromatografia por exclusão de tamanho (Size Exclusion Chromatography – SEC), utilizando uma coluna HiLoad 16/60 Superdex 200 (GE Healthcare Bio-Sciences, Uppsala, Suécia) acoplada ao sistema ÄKTApure (GE Healthcare, EUA). A coluna foi previamente equilibrada com tampão Tris-HCl pH 9,0, contendo 200 mM de NaCl e 10% (v/v) de glicerol. Após a purificação, a concentração proteica das frações foi determinada pelo método de Bradford. Em seguida, as amostras obtidas nas etapas de IMAC e SEC foram analisadas por eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio (SDS-PAGE), com o objetivo de avaliar a eficiência da purificação, o grau de pureza e o perfil eletroforético da proteína recombinante.

4. RESULTADOS

4.1 Caracterização *in silico* da endo-1,4- β -glucanase prospectada

Com o objetivo de caracterizar a sequência codificante da proteína M9428-19985 e confirmar sua classificação funcional, foram realizadas análises *in silico* empregando diferentes ferramentas de bioinformática e bancos de dados especializados. Inicialmente, a sequência foi alinhada com endoglucanases previamente caracterizadas pertencentes à família das glicosil hidrolases GH5, obtidas do banco de dados CAZy, utilizando o software BioEdit. O alinhamento múltiplo revelou a presença de diversas regiões conservadas distribuídas ao longo da sequência (Figura 1), evidenciando a manutenção de resíduos estruturais importantes entre a proteína prospectada e endoglucanases já descritas. Embora existam regiões variáveis, a conservação observada sugere que a proteína compartilha características estruturais típicas das enzimas pertencentes à família GH5.

Figura 1- Alinhamento múltiplo da sequência da endo-1,4- β -glucanase prospectada (M9428-19985) com endoglucanases pertencentes à família das glicosil hidrolases GH5 obtidas do banco de dados CAZy. As regiões coloridas representam resíduos conservados entre as sequências, indicando a preservação de regiões estrutural e funcionalmente importantes.



Fonte: Elaborado pelo próprio autor.

A fim de complementar sua caracterização, a sequência foi comparada com proteínas depositadas em diferentes bancos de dados por meio da ferramenta BLAST (Quadro 1).

Quadro 1 Resultados das buscas de similaridade da sequência M9428-19985 em diferentes bancos de dados, apresentando a proteína mais similar identificada, o comprimento das sequências comparadas, a identidade e a cobertura dos alinhamentos obtidos.

Banco de Dados	Entrada mais similar (função/ID)	Comprimento (aa): ORF/Similar	Identidade	Cobertura
Swiss-Prot	Endoglucanase/ A3DJ77.1	422/343	20,86%	71%
Pataa-NCBI	Endoglucanase/ AAR51771.1	422/407	41,13%	91%
Protein Data Bank (PDB)	Endoglucanase/ 8B3Y	422/335	27,37%	67%

Fonte: Elaborado pelo próprio autor.

Na base UniProtKB/Swiss-Prot, a melhor correspondência foi obtida com uma endoglucanase de *Acetivibrio thermocellus* (A3DJ77.1), apresentando identidade de 20,86% e cobertura de 71%. Embora a identidade entre as sequências seja relativamente baixa, a elevada cobertura indica que grande parte da proteína apresenta alinhamento com uma enzima funcionalmente caracterizada, sugerindo conservação da arquitetura estrutural típica das endoglucanases.

A busca realizada na base de sequências patenteadas do NCBI (PAT) identificou maior similaridade com uma sequência patenteada derivada de *Klebsiella*

pneumoniae (AAR51771.1), apresentando identidade de 41,13% e cobertura de 91%. A elevada cobertura observada indica que praticamente toda a extensão da proteína apresenta correspondência com a sequência patenteada, reforçando sua similaridade com proteínas previamente descritas.

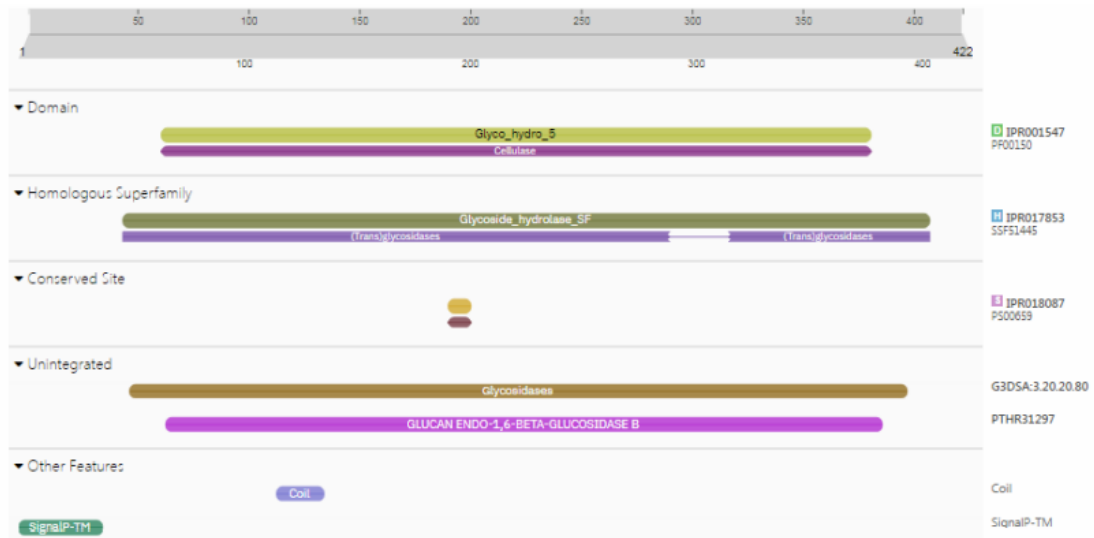
A comparação com estruturas tridimensionais depositadas no Protein Data Bank (PDB) revelou como melhor correspondência uma endoglucanase de *Thermogutta terrifontis* (PDB ID: 8B3Y), apresentando identidade de 27,37% e cobertura de 67%. Apesar da identidade moderada, a correspondência estrutural observada fornece evidências adicionais de que a proteína prospectada compartilha o dobramento característico das endoglucanases pertencentes à família GH5.

A caracterização funcional foi reforçada pelas análises realizadas na plataforma InterPro (Figura 2), que classificou a proteína como pertencente à família das glicosil hidrolases 5 (GH5). Foram identificados os domínios Glyco_hydro_5 e Cellulase, além da sua inserção na superfamília Glycoside Hydrolase, classificação corroborada simultaneamente pelas bases InterPro (IPR001547), Pfam (PF00150) e SUPERFAMILY (SSF51445). Observa-se que o domínio catalítico ocupa praticamente toda a extensão da proteína, característica típica de endoglucanases da família GH5.

Além da identificação do domínio catalítico, foram detectados sítios conservados previstos para proteínas dessa família, reforçando a classificação funcional obtida nas análises de similaridade. Também foi identificada uma pequena região estrutural predita como *coiled-coil*, não sendo observados domínios acessórios adicionais. Esses resultados indicam que a proteína apresenta uma organização estrutural predominantemente composta pelo domínio catalítico responsável pela hidrólise das ligações β -1,4-glicosídicas da celulose.

Em conjunto, os resultados obtidos pelo alinhamento múltiplo, pelas buscas de similaridade em bancos de proteínas, estruturas tridimensionais e domínios conservados demonstram que a sequência prospectada apresenta as principais características estruturais e funcionais de uma endo-1,4- β -glucanase pertencente à família GH5. Essas evidências sustentam sua seleção para as etapas subsequentes de clonagem, expressão heteróloga e caracterização bioquímica.

Figura 2- Predição dos domínios conservados da proteína M9428-19985 utilizando a plataforma InterPro (PAYSAN-LAFOSSE et al., 2023). A proteína foi classificada como pertencente à família das glicosil hidrolases GH5, apresentando os domínios Glyco_hydro_5 e Cellulase, além da classificação na superfamília Glycoside Hydrolase e da identificação de sítios conservados característicos dessa família enzimática.



Fonte: Elaborado pelo próprio autor.

4.2 Avaliação da capacidade celulolítica da bactéria *Bacillus bombysepticus*

O ensaio para determinar a atividade celulolítica da bactéria *Bacillus bombysepticus* mostrou que a bactéria possui atividade enzimática atuando na hidrólise de CMC, conforme ilustrado na Figura 3.

Figura 3- Avaliação da capacidade celulolítica da bactéria *Bacillus bombysepticus*. Cultivo pelo período de 24 horas em Meio de cultura LB sólido suplementado com substrato Carboximetilcelulose (CMC) (1%). Os halos de hidrólise foram revelados pela aplicação de Vermelho Congo (1%) por 15 minutos, seguido de lavagem das placas com NaCl (1M) por 10 minutos.



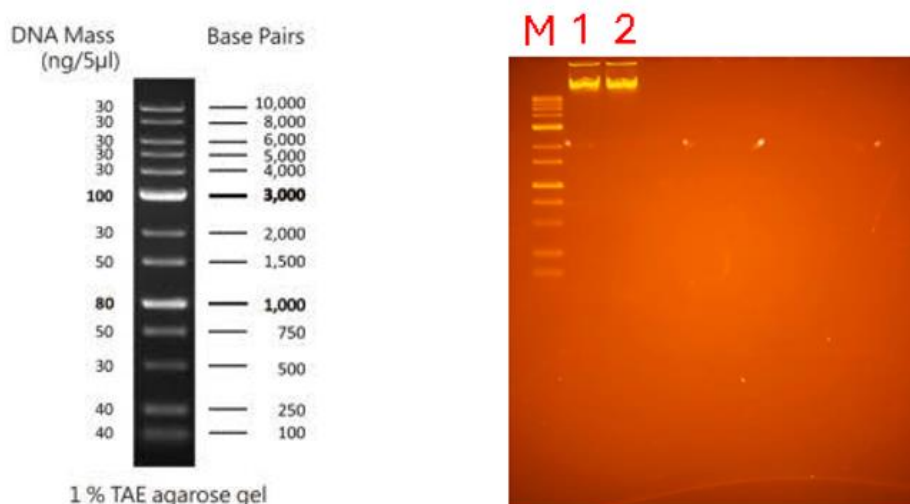
Fonte: Elaborado pelo próprio autor.

Por meio da observação da placa corada com Vermelho Congo, é possível constatar o processo de degradação, manifestado pela formação de áreas mais claras em contraste com a região circundante.

4.3 Extração de DNA da bactéria *Bacillus bombysepticus*

A extração de DNA genômico de *Bacillus bombysepticus* foi bem-sucedida, sendo a integridade do material confirmada por eletroforese em gel de agarose. Observou-se a presença de bandas intensas de alto peso molecular, próximas ao poço de aplicação, sem evidências aparentes de degradação (Figura 4).

Figura 4- Gel de eletroforese em agarose 1% do DNA genômico extraído de *Bacillus bombysepticus*. M: marcador de peso molecular de DNA 1 kb (K9-1000L, Kasvi, São José dos Pinhais, PR, Brasil); pistas 1 e 2 correspondem às amostras de DNA genômico obtidas após o procedimento de extração. A concentração de DNA das amostras foi estimada por comparação visual da intensidade das bandas com o marcador de peso molecular.



Fonte: Elaborado pelo próprio autor.

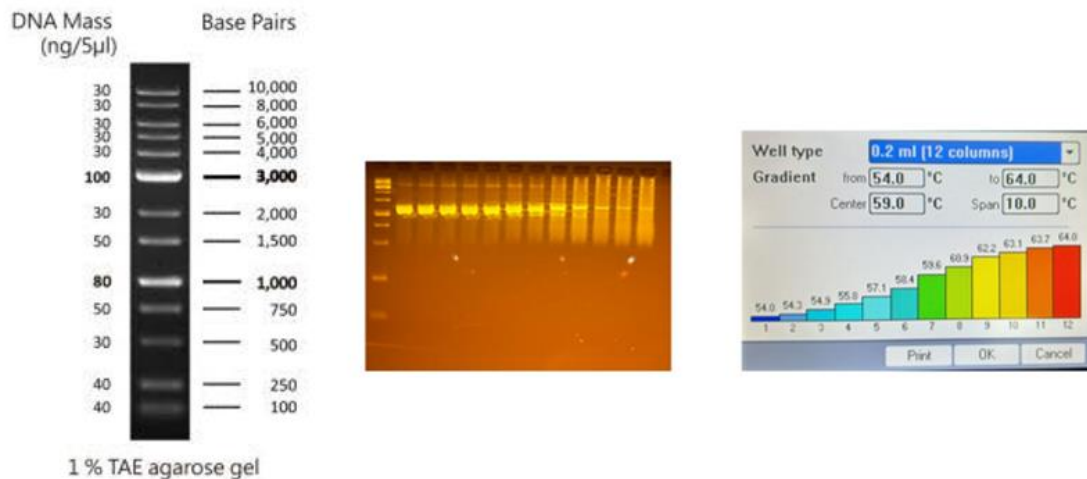
Após a extração, o DNA genômico foi quantificado utilizando o fluorímetro Qubit. As amostras apresentaram concentrações de 101 ng/µL (amostra 1) e 132 ng/µL (amostra 2), indicando rendimento adequado para as etapas subsequentes de amplificação por PCR.

4.4 Amplificação do gene de endoglucanase por PCR

A PCR em gradiente de temperatura de anelamento (54,0–64,0 °C) resultou na amplificação do fragmento-alvo em todas as condições avaliadas, evidenciada pela presença de bandas únicas e bem definidas. Não foram observados produtos inespecíficos ou formação evidente de dímeros de primers, indicando boa

especificidade da reação. A temperatura de anelamento adotada para os experimentos subsequentes foi de 54,3 °C (Figura 5).

Figura 5- PCR em gradiente utilizada para otimização da temperatura de anelamento dos primers para amplificação do gene de endoglucanase e posterior clonagem no vetor pETSUMO. Os produtos da PCR foram analisados por eletroforese em gel de agarose 1%. M: marcador de peso molecular de DNA 1 kb (K9-1000L, Kasvi, São José dos Pinhais, PR, Brasil). As colunas 1–12 correspondem às reações realizadas nas temperaturas de anelamento entre 54 e 64 °C, observando-se a amplificação de um fragmento com tamanho esperado de aproximadamente 1,5 kb.

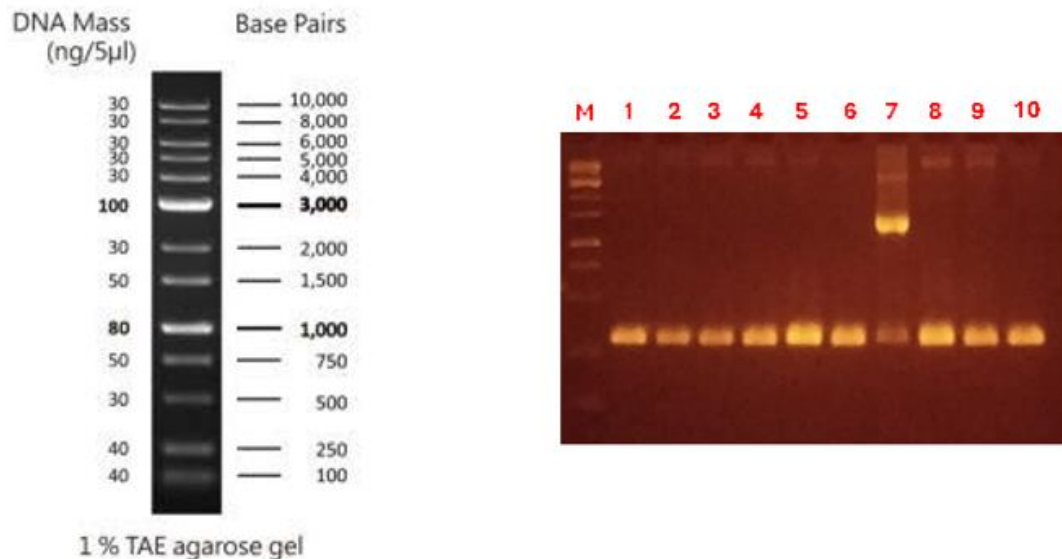


Fonte: Elaborado pelo próprio autor.

4.5 Clonagem plasmidial e sequenciamento

Após a etapa de clonagem no vetor de expressão pET-SUMO, colônias bacterianas foram selecionadas para confirmação da presença do inserto por meio de PCR de colônia. Foram analisados dez clones independentes, numerados de 1 a 10. Os produtos da amplificação foram submetidos à eletroforese em gel de agarose, sendo observadas bandas com o tamanho esperado em diversos clones, indicando a presença do inserto no plasmídeo recombinante (Figura 6).

Figura 6- Análise dos clones bacterianos por PCR de colônia após clonagem do gene de endoglucanase no vetor pETSUMO. Os produtos da amplificação foram analisados por eletroforese em gel de agarose 1%. M: marcador de peso molecular de DNA 1 kb (K9-1000L, Kasvi, São José dos Pinhais, PR, Brasil). As colunas 1–10 correspondem aos clones bacterianos analisados. Observa-se a amplificação do fragmento esperado de aproximadamente 1,5 kb em todos os clones, indicando a presença do inserto. O clone 7 apresentou uma banda adicional de maior tamanho, sugerindo amplificação inespecífica ou presença de DNA plasmidial residual.



Fonte: Elaborado pelo próprio autor.

O clone selecionado foi submetido ao sequenciamento, e as leituras obtidas foram utilizadas para a montagem de uma sequência consenso (contig). Posteriormente, essa sequência foi analisada utilizando a ferramenta BLASTn do NCBI para verificar sua identidade.

A análise por BLASTn revelou alinhamentos significativos com sequências pertencentes ao gênero *Bacillus*, apresentando cobertura de consulta (Query Cover) de 100%, identidade de aproximadamente 98,3% e valor E igual a 0,0 (Figura 7).

Figura 7 -Resultado da análise de similaridade da sequência consenso (contig) obtida após o sequenciamento do clone recombinante por meio da ferramenta BLASTn (NCBI). Os principais alinhamentos apresentaram cobertura de consulta (Query Cover) de 100%, identidade de aproximadamente 98,3% e valor E igual a 0,0, evidenciando elevada similaridade com sequências do gênero *Bacillus*, incluindo *Bacillus bombysepticus*.

Sequences producing significant alignments		Download	Select columns	Show	100				
☒ select all 100 sequences selected		GenBank	Graphics	Distance tree of results	MSA Viewer				
	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☒	Bacillus sp. A260 chromosome, complete genome	Bacillus sp. A260	2034	2034	100%	0.0	98.26%	5286434	CP046398.1
☒	Bacillus cereus strain DFO40 chromosome, complete genome	Bacillus cereus	2034	2034	100%	0.0	98.26%	5327326	CP089228.1
☒	Bacillus cereus strain FORC087 chromosome, complete genome	Bacillus cereus	2034	2034	100%	0.0	98.26%	5271204	CP029454.1
☒	Bacillus cereus strain K8 chromosome, complete genome	Bacillus cereus	2034	2034	100%	0.0	98.26%	5329684	CP016595.1
☒	Bacillus cereus strain lycx chromosome, complete genome	Bacillus cereus	2034	2034	100%	0.0	98.26%	5484345	CP129005.1
☒	Bacillus thuringiensis serovar indiana strain HD521, complete genome	Bacillus thuringiensis serovar indiana	2034	2034	100%	0.0	98.26%	5429688	CP010106.1
☒	Bacillus cereus strain 2-6A chromosome, complete genome	Bacillus cereus	2034	2034	100%	0.0	98.26%	5191018	CP117421.1
☒	Bacillus bombysepticus strain F12 chromosome, complete genome	Bacillus bombysepticus	2034	2034	100%	0.0	98.26%	5244749	CP085406.1
☒	Bacillus cereus strain A22 chromosome, complete genome	Bacillus cereus	2034	2034	100%	0.0	98.26%	5310791	CP085498.1
☒	Bacillus bombysepticus str. Wang, complete genome	Bacillus bombysepticus str. Wang	2034	2034	100%	0.0	98.26%	5295783	CP007512.1
☒	Bacillus cereus strain RC33 chromosome, complete genome	Bacillus cereus	2034	2034	100%	0.0	98.26%	5325341	CP072774.1
☒	Bacillus cereus strain DLOU-Tangshan chromosome, complete genome	Bacillus cereus	2028	2028	100%	0.0	98.20%	5542157	CP040340.1
☒	Bacillus cereus strain FWZ-1 chromosome, complete genome	Bacillus cereus	2028	2028	100%	0.0	98.20%	5346470	CP158363.1
☒	Bacillus cereus strain Ta-LAA-64 chromosome	Bacillus cereus	2028	2028	100%	0.0	98.20%	5147492	CP174498.1
☒	Bacillus sp. JAS24-2 chromosome, complete genome	Bacillus sp. JAS24-2	2028	2028	100%	0.0	98.20%	5225480	CP031068.1
☒	Bacillus cereus strain G1-1 chromosome, complete genome	Bacillus cereus	2028	2028	100%	0.0	98.20%	5544716	CP042929.1

Fonte: Elaborado pelo próprio autor.

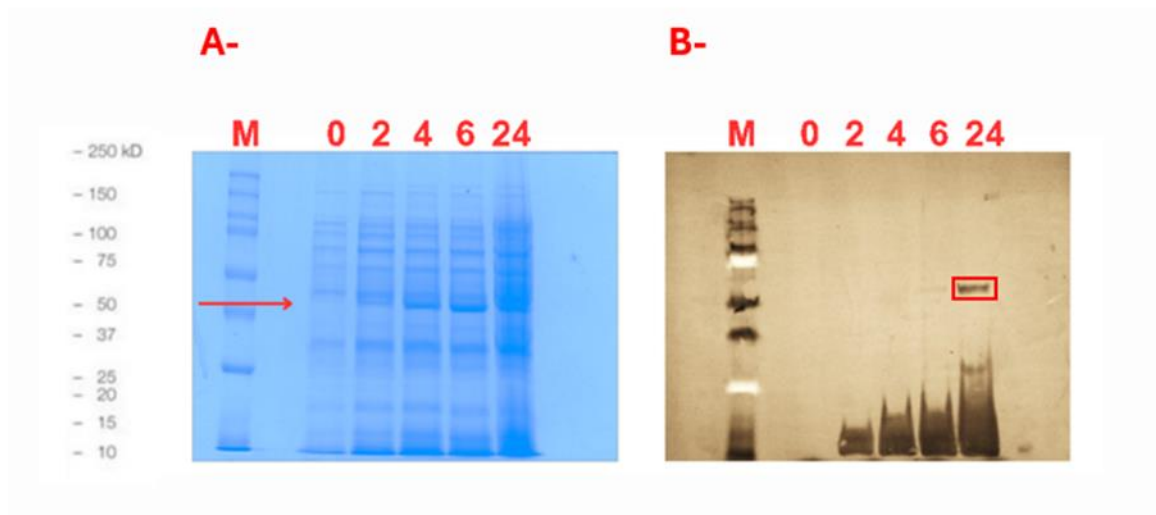
Entre os alinhamentos obtidos encontram-se sequências de *Bacillus bombysepticus*, espécie de origem do gene utilizado neste estudo, além de outras espécies filogeneticamente próximas do gênero *Bacillus*. A ocorrência de múltiplos alinhamentos com valores de identidade semelhantes é esperada, uma vez que genes pertencentes à mesma família podem apresentar elevado grau de conservação entre espécies relacionadas. Esses resultados corroboram a identidade do inserto clonado e confirmam a obtenção da sequência esperada para as etapas subsequentes de expressão heteróloga.

4.6 Indução e expressão da proteína recombinante

Após a confirmação do clone recombinante, iniciou-se a expressão heteróloga da β -endoglucanase em *Escherichia coli*. A produção da proteína foi acompanhada durante 24 horas por meio de SDS-PAGE (Figura 8A). Observou-se o surgimento e aumento gradual da intensidade de uma banda com massa molecular aproximada de 50 kDa ao longo do período de indução, sendo sua maior intensidade observada após 24 horas de cultivo. A identidade da banda correspondente à proteína recombinante foi confirmada por *Western blotting* (Figura 8B), no qual foi detectado sinal específico na região correspondente ao peso molecular esperado para a β -endoglucanase fusionada à proteína SUMO. A intensidade do sinal também foi maior na amostra

coletada após 24 horas de indução, confirmando que esse período proporcionou a maior produção da proteína recombinante.

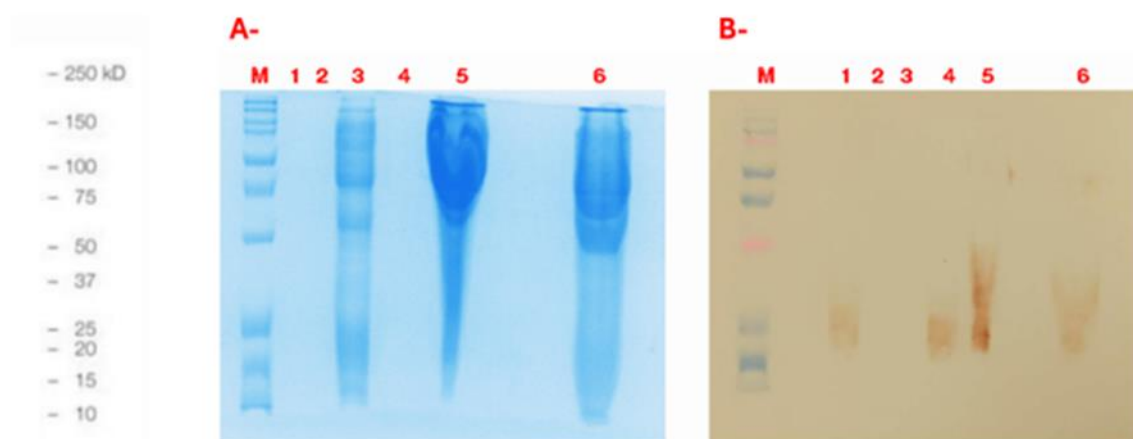
Figura 8- Análise da expressão da β -endoglucanase recombinante em *Escherichia coli*. (A) Perfil de expressão avaliado por SDS-PAGE após indução com 0,1 mM de IPTG nos tempos de 0, 2, 4, 6 e 24 horas. A seta indica a banda correspondente à proteína recombinante (~50 kDa). (B) Western blot das mesmas amostras confirmando a identidade da proteína recombinante por imunodeteção. M: marcador de massa molecular Precision Plus Protein™ Dual Color Standards (Bio-Rad Laboratories, Hercules, CA, EUA).



Fonte: Elaborado pelo próprio autor.

Após a definição do tempo de expressão, as frações solúvel e insolúvel foram analisadas por SDS-PAGE e Western blotting (Figura 9).

Figura 9-Avaliação da distribuição da β -endoglucanase recombinante entre as frações solúvel e insolúvel após expressão em *Escherichia coli*. (A) SDS-PAGE das amostras analisadas. (B) Western blot das mesmas amostras confirmando a presença da proteína recombinante. M: marcador de massa molecular Precision Plus Protein™ Dual Color Standards (Bio-Rad Laboratories, Hercules, CA, EUA). Poços: (1) sobrenadante após 6 h de indução; (2) amostra antes da indução (0 h); (3) pellet após 6 h de indução; (4) poço vazio; (5) amostra correspondente ao tempo 0 h; (6) pellet após 24 h de indução.

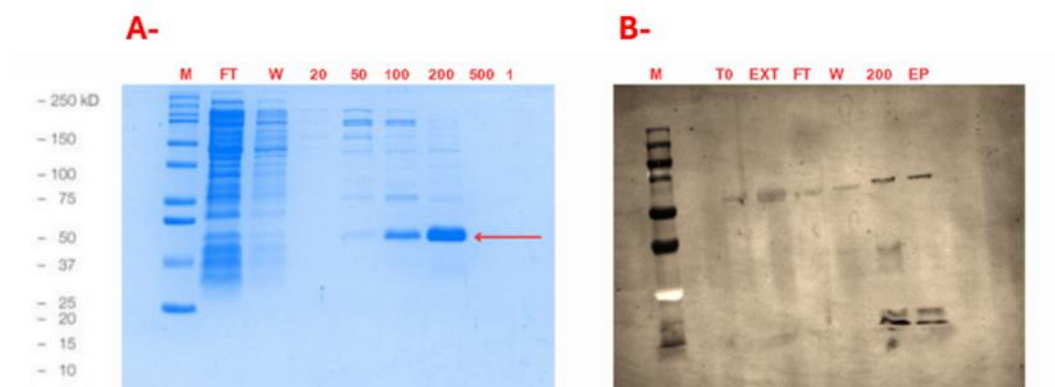


Fonte: Elaborado pelo próprio autor.

A proteína recombinante foi detectada tanto no sobrenadante quanto no pellet celular, indicando que parte da proteína permaneceu na fração solúvel enquanto outra parte foi acumulada na forma insolúvel. Como a proteína de interesse foi detectada na fração solúvel, essa fração foi utilizada nas etapas subsequentes de purificação por cromatografia de afinidade.

A fração solúvel contendo a β -endoglucanase recombinante foi submetida à cromatografia de afinidade em resina Ni-NTA utilizando gradiente crescente de imidazol para eluição da proteína. As diferentes etapas da purificação foram monitoradas por SDS-PAGE e confirmadas por Western blotting (Figura 10).

Figura 10-Purificação da β -endoglucanase recombinante por cromatografia de afinidade em resina Ni-NTA. (A) SDS-PAGE das diferentes etapas da purificação: FT (flow-through), W (lavagem) e frações eluídas com 20, 50, 100, 200, 500 mM e 1 M de imidazol. A seta indica a proteína recombinante (~50 kDa). (B) Western blot das mesmas frações confirmando a identidade da proteína recombinante ao longo do processo de purificação. M: marcador de massa molecular Precision Plus Protein™ Dual Color Standards (Bio-Rad Laboratories, Hercules, CA, EUA).



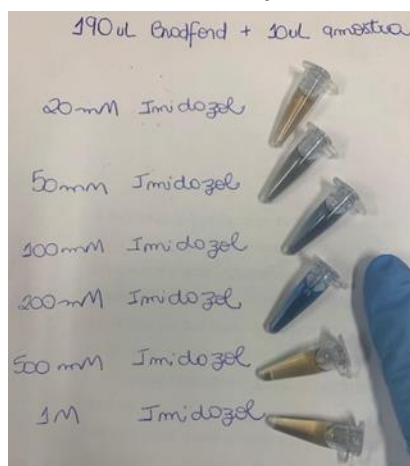
Fonte: Elaborado pelo próprio autor.

Na análise por SDS-PAGE (Figura 10A), observou-se a presença da proteína recombinante principalmente nas frações eluídas com imidazol, sendo a banda correspondente ao peso molecular esperado mais intensa na eluição com 200 mM de imidazol. Nas concentrações superiores (500 mM e 1 M), verificou-se redução da intensidade da banda, indicando que a maior parte da proteína havia sido recuperada anteriormente.

Os resultados obtidos por Western blotting (Figura 10B) confirmaram que a banda observada correspondia à β -endoglucanase recombinante, demonstrando maior concentração da proteína na fração eluída com 200 mM de imidazol. Também foi possível detectar pequenas quantidades da proteína em outras etapas da purificação, indicando perdas discretas durante o processo de lavagem e eluição.

As frações eluídas foram posteriormente avaliadas pelo ensaio de Bradford (Figura 11), utilizado como análise qualitativa da presença de proteínas nas diferentes concentrações de imidazol.

Figura 11- Avaliação qualitativa das frações eluídas durante a purificação por cromatografia de afinidade utilizando o reagente de Bradford. As frações correspondem às eluições com 20, 50, 100, 200, 500 mM e 1 M de imidazol. A maior intensidade de coloração foi observada na fração eluída com 200 mM de imidazol, indicando maior concentração relativa de proteínas nessa condição.



Fonte: Elaborado pelo próprio autor.

Observou-se maior intensidade da coloração azul na fração eluída com 200 mM de imidazol, corroborando os resultados obtidos por SDS-PAGE e Western blotting, que indicaram essa condição como a mais eficiente para recuperação da proteína recombinante.

5. DISCUSSÃO

A expressão heteróloga da β -endoglucanase proveniente de *Bacillus bombysepticus* em *Escherichia coli* ER2566, utilizando o vetor pET-SUMO, resultou na produção de uma proteína recombinante com massa molecular próxima à esperada (~49 kDa), conforme observado por SDS-PAGE e confirmado por Western blot. Entretanto, também foram observados fragmentos menores (~10 kDa e ~25 kDa), o que motivou a investigação de possíveis explicações para esse padrão.

Uma primeira hipótese é a degradação proteolítica por proteases endógenas de *E. coli*, que podem reconhecer a proteína heteróloga como anômala durante sua superexpressão e direcioná-la para degradação. Alternativamente, esses fragmentos poderiam resultar do processamento pós-traducional da enzima, que pode sofrer clivagem específica por auto-proteólise ou pela ação de proteases celulares, originando produtos de menor peso molecular (SASKA; DAVID, 2008). Essa possibilidade é reforçada por MAURIZI (1992), que afirma que, em *E. coli*, a degradação de proteínas desempenha papéis importantes na regulação dos níveis de

proteínas específicas e na eliminação de proteínas danificadas ou anormais; a bactéria possui um número muito grande de enzimas proteolíticas distribuídas no citoplasma, na membrana interna e no periplasma, sendo que a atividade da maioria delas *in vitro* não depende de energia.

Além da degradação proteolítica, não se pode descartar uma terceira explicação: a presença de múltiplas bandas pode estar relacionada à expressão de isoformas truncadas, geradas por diferentes mecanismos celulares e que constituem proteoformas distintas da proteína canônica. Em Western blottings, essas bandas frequentemente são interpretadas como produtos de degradação quando, na verdade, podem resultar de início alternativo da tradução, fenômeno associado à instabilidade do mRNA ou à presença de regiões secundárias na sequência da proteína (MCWHITE et al., 2024; KOBAYASHI et al., 2009).

Independentemente da origem dos fragmentos, observou-se que a proteína de interesse foi detectada tanto na fração solúvel quanto na insolúvel, o que já indica algum grau de instabilidade ou dificuldade de dobramento correto. Nesse contexto, a recuperação da proteína por cromatografia de afinidade em resina Ni-NTA apresentou maior rendimento na eluição com 200 mM de imidazol.

Diante desse cenário de instabilidade, uma estratégia comumente empregada para reduzir a degradação e favorecer a forma solúvel é o ajuste das condições de indução. A redução da temperatura de indução pode diminuir a atividade proteolítica, favorecendo a expressão de proteínas mais estáveis e solúveis, uma vez que a indução em temperaturas reduzidas melhora significativamente a solubilidade, a estabilidade e o dobramento adequado das proteínas recombinantes (GUPTA; SHUKLA, 2016). Essa relação entre temperatura e estabilidade também foi observada por Ma et al. (2020), cuja endoglucanase GH5 Cel-5A apresentou maior atividade e estabilidade em temperaturas moderadas (30–60 °C), perdendo atividade em temperaturas mais elevadas — o que indica adaptação da enzima a condições mais brandas.

Esse tipo de sensibilidade às condições de expressão e purificação não é exclusivo do presente estudo, sendo relatado por diversos trabalhos com outras endoglucanases da família GH5. Dung et al. (2022), por exemplo, expressaram e purificaram com sucesso uma endoglucanase recombinante GH5-CBM72-CBM72 proveniente do rúmen de cabras em *E. coli*, obtendo atividade contra carboximetilcelulose (CMC) e elevada pureza. Nesse estudo, observou-se pela

primeira vez que o imidazol provocou a precipitação da enzima quando expressa em *E. coli* BL21, enquanto a expressão em *E. coli* Rosetta 1 não apresentou esse efeito, evidenciando que a escolha da cepa hospedeira e das condições de purificação influencia diretamente a estabilidade da proteína recombinante.

De forma semelhante, Guan et al. (2018) demonstraram, em seu trabalho com uma enzima GH5, que a otimização das condições de expressão e a remoção da cauda 6×His aumentaram significativamente a atividade de uma nova endoglucanase de *Bacillus subtilis* SB13, associada também a cinco substituições de aminoácidos relacionadas ao desempenho da enzima.

Já Wu et al. (2022) relataram um mecanismo ainda mais peculiar: a endoglucanase GsCelA de *Geobacillus sp.* sofre um processo espontâneo de auto-truncamento, que aumenta sua atividade catalítica e termoestabilidade sem depender da ação de proteases externas, sendo influenciado pela estrutura da enzima, pH, temperatura e íons metálicos — configurando um novo mecanismo de ativação para endoglucanases da família GH5. Esses exemplos reforçam que a formação de fragmentos de menor massa molecular pode, em certos casos, representar não um artefato indesejado, mas um processo funcional de maturação da enzima.

Vale ainda considerar que os padrões observados no SDS-PAGE e no Western blotting também podem ser influenciados por fatores puramente técnicos, como a concentração de acrilamida do gel e as condições de preparação da amostra, incluindo desnaturação, redução e tempo de incubação. A otimização desses parâmetros melhora a resolução das bandas e a reprodutibilidade da separação proteica, contribuindo para uma análise mais precisa dos fragmentos de diferentes massas moleculares (IDOWU et al., 2025).

Diante do conjunto de resultados e hipóteses discutidos, experimentos futuros devem ser conduzidos para identificar a origem exata dos fragmentos menores observados. A adição de inibidores de proteases e a alteração das condições de expressão podem ajudar a testar a hipótese de degradação proteolítica, enquanto análises estruturais e de sequência podem indicar eventos de processamento pós-traducional ou expressão de isoformas truncadas. Por fim, ajustes na eletroforese podem garantir melhor separação das bandas, auxiliando na interpretação dos resultados. A compreensão desses aspectos é essencial para otimizar a expressão e a estabilidade da endoglucanase recombinante, permitindo sua aplicação em estudos futuros.

6. CONCLUSÃO

O presente trabalho teve como objetivo prospectar, clonar e iniciar a expressão de uma endo-1,4- β -glucanase identificada no genoma de *Bacillus bombysepticus*. As etapas de prospecção *in silico* permitiram confirmar a presença de uma ORF pertencente à família GH5, indicando potencial funcional para degradação de celulose e justificando sua escolha como alvo para clonagem heteróloga.

A clonagem do gene no vetor pET-SUMO foi confirmada por PCR de colônia e sequenciamento, possibilitando o avanço para os testes de expressão em *Escherichia coli*. Entretanto, embora tenham sido observadas bandas na faixa esperada em SDS-PAGE e Western blot, os padrões obtidos não permitiram afirmar de forma conclusiva que a proteína recombinante foi expressa de maneira íntegra ou funcional. A presença de fragmentos menores e sinais inespecíficos sugere que podem ter ocorrido degradação proteolítica, baixo rendimento, processamento incompleto ou limitações do sistema de expressão utilizado.

A purificação por cromatografia de afinidade resultou em frações contendo proteínas reativas ao anticorpo anti-His, porém a integridade e a funcionalidade da endoglucanase ainda não puderam ser validadas. Dessa forma, a etapa de caracterização enzimática, incluindo ensaios cinéticos, determinação de pH e temperatura ótimos e avaliação da atividade sobre CMC, permanece dependente de otimizações adicionais no processo de expressão e purificação.

São necessários de ajustes experimentais para confirmar a produção estável e funcional da proteína recombinante. Os próximos passos são testar diferentes temperaturas de indução, avaliar concentrações alternativas de IPTG, utilizar cepas de *E. coli* com menor atividade proteolítica, adicionar inibidores de proteases, otimizar o tampão de lise e, se necessário, avaliar vetores ou sistemas de expressão alternativos.

REFERÊNCIAS

AFFANDI, M. et al. Enzymatic activity of bacteria isolated from the gut of *Cylindroiulus* sp.: future prospect for decomposing agent. *IOP Conference Series: Earth and Environmental Science*, Bristol, v. 762, n. 1, art. 012042, 2021. DOI: <https://doi.org/10.1088/1755-1315/762/1/012042>

ALTSCHUL, Stephen F. et al. Basic local alignment search tool. *Journal of Molecular Biology*, London, v. 215, n. 3, p. 403–410, 1990. DOI: [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)

ARIAEENEJAD, Shohreh; GHARECHAH, Javad; FOROOZANDEH SHAHRAKI, Mehdi et al. Precision enzyme discovery through targeted mining of metagenomic data. *Natural Products and Bioprospecting*, v. 14, n. 1, art. 7, 2024. DOI: <https://doi.org/10.1007/s13659-023-00426-8>

ARIAEENEJAD, Shohreh; GHARECHAH, Javad; FOROOZANDEH SHAHRAKI, Mehdi et al. Precision enzyme discovery through targeted mining of metagenomic data. *Natural Products and Bioprospecting*, v. 14, n. 1, art. 7, 2024. DOI: <https://doi.org/10.1007/s13659-023-00426-8>

AVALON, Joshua et al. Bacterial natural product discovery by heterologous expression. *Journal of Industrial Microbiology and Biotechnology*, v. 51, n. 1, 2024.

AVCI, Ayse; DEMIR, Selin; AKÇAY, Fikriye Alev. Production, properties and some applications of alkaliphilic *Bacillus* sp. protease. *Preparative Biochemistry & Biotechnology*, v. 51, n. 8, p. 803–810, 2021.

BASERA, Priyanka; CHAKRABORTY, Shuchishloka; SHARMA, Neha. Lignocellulosic biomass: insights into enzymatic hydrolysis, influential factors, and economic viability. *Discover Sustainability*, v. 5, n. 1, art. 311, 2024.

BIO-RAD LABORATORIES. *Precision Plus Protein™ Dual Color Standards: instruction manual*. Hercules, CA: Bio-Rad Laboratories, 2021.

BLUM, Matthias et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Research*, Oxford, v. 49, n. D1, p. D344–D354, 2021. DOI: <https://doi.org/10.1093/nar/gkaa977>

CAZY. *Glycoside Hydrolase family classification*. Disponível em: <https://www.cazy.org/Glycoside-Hydrolases>. Acesso em: 27 jun. 2026.

CHAKRABORTY, Prasenjit et al. Technological advancements in the pretreatment of lignocellulosic biomass for effective valorization: a review of challenges and prospects. *Journal of Industrial and Engineering Chemistry*, v. 137, p. 29–60, 2024.

CHENG, Tingcai et al. Complete genome sequence of *Bacillus bombysepticus*, a pathogen leading to black chest septicemia in *Bombyx mori*. *Genome Announcements*, v. 2, n. 3, e00312-14, 2014. DOI: <https://doi.org/10.1128/genomeA.00312-14>

DUNG, Cao Thi; KHOA, Dao Trong; HUYEN, Do Thi; QUY, Nguyen Thi; HAI, Truong Nam. Purification of recombinant endoglucanase GH5-CBM72-CBM72 expressed in *Escherichia coli*. *Vietnam Journal of Biotechnology*, v. 20, n. 4, p. 623–631, 2022. DOI: <https://doi.org/10.15625/1811-4989/16953>

ESCUDE-RODRÍGUEZ, Juan-José et al. Characterization of a novel thermophilic metagenomic GH5 endoglucanase heterologously expressed in *Escherichia coli* and *Saccharomyces cerevisiae*. *Biotechnology for Biofuels and Bioproducts*, v. 15, n. 1, art. 76, 2022. DOI: <https://doi.org/10.1186/s13068-022-02120-X>

FERRANDO, Jordi et al. From mold to mill: *StachCel5*, a novel thermoalkaliphilic endoglucanase from *Stachybotrys chartarum* for pulp fiber biorefining. *International Journal of Biological Macromolecules*, v. 280, art. 145969, 2025. DOI: <https://doi.org/10.1016/j.ijbiomac.2025.145969>

GUAN, Xuefang; CHEN, Penglian; XU, Qingxian; QIAN, Lei; HUANG, Juqing; LI, Bin. Expression, purification and molecular characterization of a novel endoglucanase protein from *Bacillus subtilis* SB13. *International Journal of Biological Macromolecules*, v. 108, p. 1032–1040, 2018. DOI: <https://doi.org/10.1016/j.ijbiomac.2017.11.057>

GUPTA, Sanjeev K.; SHUKLA, Pratyosh. Advanced technologies for improved expression of recombinant proteins in bacteria: perspectives and applications. *Critical Reviews in Biotechnology*, v. 36, n. 6, p. 1089–1098, 2016. DOI: <https://doi.org/10.3109/07388551.2015.1084264>

HARTMAN, E. A flacherie disease of the silkworm caused by *Bacillus bombysepticus*. *Lingnan Science Journal*, v. 10, p. 279–289, 1931.

HOUFANI, Aicha Asma et al. Lignocellulosic biomass degrading enzymes and characterization of cellulase and xylanase from *Bosea* sp. FBZP-16. *Biomass Conversion and Biorefinery*, v. 13, n. 13, p. 12085–12103, 2023. DOI: <https://doi.org/10.1007/s13399-021-02022-1>

IDOWU, Omogbolahan Samson et al. Optimizing SDS-PAGE for accurate protein characterization in nutritional research and food quality assessment. *International Journal of Innovative Science and Research Technology*, v. 10, n. 1, 2025.

JOO, Han Seung; JANG, Won Choi. Purification and characterization of a novel alkaline protease from *Bacillus horikoshii*. *Journal of Microbiology and Biotechnology*, v. 22, n. 1, p. 58–68, 2012. DOI: <https://doi.org/10.4014/jmb.1108.08036>

KASANA, Ramesh Chand et al. A rapid and easy method for the detection of microbial cellulases on agar plates using Gram's iodine. *Current Microbiology*, v. 57, n. 5, p. 503–507, 2008. DOI: <https://doi.org/10.1007/s00284-008-9276-8>

KASVI. *Catálogo Geral 2023*. São José dos Pinhais: Kasvi, 2023. Disponível em: <https://kasvi.com.br/wp-content/uploads/2023/06/KASVI-Catalogo-2023-Geral-WEB-compactado.pdf>. Acesso em: 30 jun. 2026.

KOBAYASHI, R.; PATENIA, R.; ASHIZAWA, S.; VYKOUKAL, J. Targeted mass spectrometric analysis of N-terminally truncated isoforms generated via alternative translation initiation. *FEBS Letters*, v. 583, n. 14, p. 2441–2445, 2009. DOI: <https://doi.org/10.1016/j.febslet.2009.05.042>

KUMAR, Kishan et al. Cellulases and hemicellulases in lignocellulose degradation: a comparative analysis of wild-type and recombinant systems. *Biochemical and Biophysical Research Communications*, v. 758, art. 153675, 2026. DOI: <https://doi.org/10.1016/j.bbrc.2025.153675>

LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, London, v. 227, n. 5259, p. 680–685, 1970. DOI: <https://doi.org/10.1038/227680a0>

LEGEIN, Marie et al. Modes of action of microbial biocontrol in the phyllosphere. *Frontiers in Microbiology*, v. 11, art. 1619, 2020. DOI: <https://doi.org/10.3389/fmicb.2020.01619>

LOMBARD, Vincent et al. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, Oxford, v. 42, n. D1, p. D490–D495, 2014. DOI: <https://doi.org/10.1093/nar/gkt1178>

LV, Huaiyu et al. An endoglucanase from *Erwinia amylovora* induces broad-spectrum immune response in plants. *International Journal of Biological Macromolecules*, v. 308, art. 142550, 2025. DOI: <https://doi.org/10.1016/j.ijbiomac.2025.142550>.

MA, Lingling; RAKHMANOVA, Aizhan; WANG, Xin; YI, Yanglei; SHAN, Yuanyuan; LIU, Bianfang; ZHOU, Yuan; LÜ, Xin. Cloning and characterization of low-temperature adapted GH5-CBM3 endo-cellulase from *Bacillus subtilis* 1AJ3 and their application in the saccharification of switchgrass and coffee grounds. *AMB Express*, v. 10, art. 42, 2020. DOI: <https://doi.org/10.1186/s13568-020-00975-y>

MAXEINER, Stephan; KRASSTEVA-CHRIST, Gabriela; ALTHAUS, Mike. Pitfalls of using sequence databases for heterologous expression studies: a technical review. *The Journal of Physiology*, v. 601, n. 9, p. 1611–1623, 2023. DOI: <https://doi.org/10.1113/JP283546>.

MCWHITE, C. D. et al. Alternative proteoforms and proteoform-dependent assemblies in humans and plants. *Molecular Systems Biology*, v. 20, n. 8, p. 933–951, 2024. DOI: <https://doi.org/10.1038/s44320-024-00048-3>.

NCBI RESOURCE COORDINATORS. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, Oxford, v. 46, n. D1, p. D8–D13, 2018. DOI: <https://doi.org/10.1093/nar/gkx1095>.

NGUYEN, Dinh T. Genome mining for new enzyme chemistry. *ACS Catalysis*, v. 14, n. 7, p. 4536–4553, 2024. DOI: <https://doi.org/10.1021/acscatal.4c00469>

NOBLE, C.; CROWE, J. D.; O'DONOHUE, M. J. A roadmap for metagenomic enzyme discovery. *Natural Product Reports*, v. 38, n. 11, p. 1999–2023, 2021. DOI: <https://doi.org/10.1039/D1NP00006C>

PAN, Ieshita et al. Exploration of compost soil for the production of thermostable *Bacillus* protease to synthesize bioactive compounds through soy protein hydrolysis. *Agronomy*, v. 13, n. 4, art. 1019, 2023. DOI: <https://doi.org/10.3390/agronomy13041019>

PAN, Ieshita et al. Exploration for thermostable β -amylase of a *Bacillus* sp. isolated from compost soil to degrade bacterial biofilm. *Microbiology Spectrum*, v. 9, n. 2, art. e00647-21, 2021. DOI: <https://doi.org/10.1128/Spectrum.00647-21>

PAYSAN-LAFOSSE, T. et al. InterPro in 2022. *Nucleic Acids Research*, Oxford, v. 51, n. D1, p. D418–D427, 2023. DOI: <https://doi.org/10.1093/nar/gkac993>.

PROMEGA CORPORATION. Wizard® Genomic DNA Purification Kit: technical manual. Madison, WI: Promega Corporation, 2019.

REN, Wei et al. Cross-habitat distribution pattern of *Bacillus* communities and their capacities of producing industrial hydrolytic enzymes in Paracel Islands: habitat-dependent differential contributions of the environment. *Journal of Environmental Management*, v. 323, art. 116252, 2022. DOI: <https://doi.org/10.1016/j.jenvman.2022.116252>

ROBSON, Lori M.; CHAMBLISS, Gleen H. Cellulases of bacterial origin. *Enzyme and Microbial Technology*, v. 11, n. 10, p. 626–644, 1989. DOI: [https://doi.org/10.1016/0141-0229\(89\)90042-3](https://doi.org/10.1016/0141-0229(89)90042-3).

RONO, Justice Kipkorir et al. Biochemical characterization of a **bifunctional** endoglucanase/glucomannanase derived from mountain soil. *Biotechnology Letters*, v. 47, n. 2, art. 33, 2025. DOI: <https://doi.org/10.1007/s10529-024-03561-3>.

SASKA, Ivana; CRAIK, David J. Protease-catalysed protein splicing: a new post-translational modification? *Trends in Biochemical Sciences*, v. 33, n. 8, p. 363–368, 2008. DOI: <https://doi.org/10.1016/j.tibs.2008.05.007>.

SIMONEN, Marjo; PALVA, Ilkka. Protein secretion in *Bacillus* species. *Microbiological Reviews*, v. 57, n. 1, p. 109–137, 1993.

SINDHU, R.; BINOD, P.; PANDEY, A. Biological pretreatment of lignocellulosic biomass: an overview. *Bioresource Technology*, v. 199, p. 76–82, 2016. DOI: <https://doi.org/10.1016/j.biortech.2015.08.030>.

SINGH NEE' NIGAM, Poonam; GUPTA, Nutan; ANTHWAL, Ashish. Pre-treatment of agro-industrial residues. In: **SINGH NEE' NIGAM, Poonam; PANDEY, Ashok** (ed.). *Biotechnology for agro-industrial residues utilisation*. Dordrecht: Springer, 2009. p. 13–33. DOI: https://doi.org/10.1007/978-1-4020-9942-7_2

STOTHARD, Paul. The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques*, New York, v. 28, n. 6, p. 1102–1104, 2000. DOI: <https://doi.org/10.2144/00286ir01>.

THE UNIPROT CONSORTIUM. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research*, Oxford, v. 49, n. D1, p. D480–D489, 2021. DOI: <https://doi.org/10.1093/nar/gkaa1100>

THERMO FISHER SCIENTIFIC. *Qubit™ dsDNA HS Assay Kits user guide*. Waltham, MA: Thermo Fisher Scientific, 2015.

WEISBURG, William G. et al. 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology*, v. 173, n. 2, p. 697–703, 1991. DOI: <https://doi.org/10.1128/jb.173.2.697-703.1991>

WOŹNIAK, Adrian et al. Review of lignocellulosic biomass pretreatment using physical, thermal and chemical methods for higher yields in bioethanol production. *Sustainability*, v. 17, n. 1, art. 287, 2025. DOI: <https://doi.org/10.3390/su17010287>

WU, Mei-Huey; KAO, Mu-Rong; LI, Chen-Wei; YU, Su-May; HO, Tuan-Hua David. A unique self-truncation of bacterial GH5 endoglucanases leads to enhanced activity and thermostability. *Biotechnology for Biofuels and Bioproducts*, v. 15, art. 65, 2022. DOI: <https://doi.org/10.1186/s13068-022-02156-z>.

YOO, Chang Geun et al. The critical role of lignin in lignocellulosic biomass conversion and recent pretreatment strategies: a comprehensive review. *Bioresource Technology*, v. 301, art. 122784, 2020. DOI: <https://doi.org/10.1016/j.biortech.2020.122784>.

ZHANG, Guangming et al. Recent advances in anaerobic digestion of lignocellulosic biomass with rumen microbes for biogas production. *Biomass and Bioenergy*, v. 212, art. 109320, 2026. DOI: <https://doi.org/10.1016/j.biombioe.2025.109320>.

ZHANG, Yanbin et al. Different biomass pretreatments for gasification: a critical review. *Journal of the Energy Institute*, v. 118, art. 101992, 2025. DOI: <https://doi.org/10.1016/j.joei.2025.101992>.

ZHU, Tianxiang et al. Functional characterization of a highly efficient endoglucanase from *Bacillus licheniformis* BJ2022 and its application in the preparation of low-molecular-weight konjac glucomannan. *Journal of Agricultural and Food Chemistry*, v. 73, n. 14, p. 8367–8380, 2025. DOI: <https://doi.org/10.1021/acs.jafc.5c01079>.