
CIÊNCIAS BIOLÓGICAS

GABRIEL FELICIO SILVA

**AVALIAÇÃO DE DESPOLIMERASES FÁGICAS
CONTRA *XANTHOMONAS CITRI* SUBSP. *CITRI***

Rio Claro - SP
2025

GABRIEL FELICIO SILVA

**AVALIAÇÃO DE DESPOLIMERASES FÁGICAS CONTRA
XANTHOMONAS CITRI SUBSP. CITRI**

Trabalho de Conclusão de Curso apresentado ao Instituto de Biociências – Câmpus de Rio Claro, da Universidade Estadual Paulista “Júlio de Mesquita Filho”, para obtenção do grau de Bacharel em Ciências Biológicas.

Orientador: Prof. Dr. Henrique Ferreira

Coorientador: Dr. Caio Felipe Cavicchia Zamunér

Rio Claro - SP
2025

S586a Silva, Gabriel Felicio
Avaliação de despolimerases fágicas contra *Xanthomonas citri* subsp. *citri* / Gabriel Felicio Silva. -- Rio Claro, 2025
47 p. : il., tabs., fotos

Trabalho de conclusão de curso (Bacharelado e licenciatura - Ciências Biológicas) - Universidade Estadual Paulista (UNESP), Instituto de Biociências, Rio Claro
Orientador: Henrique Ferreira
Coorientador: Caio Felipe Cavicchia Zamunér

1. Microbiologia molecular. 2. Bacteriófagos. 3. Cancro citríco. 4. Bioinformática. 5. Clonagem molecular. I. Título.

GABRIEL FELICIO SILVA

AVALIAÇÃO DE DESPOLIMERASES FÁGICAS CONTRA XANTHOMONAS CITRI SUBSP. CITRI

Trabalho de Conclusão de Curso apresentado ao Instituto de Biociências – Câmpus de Rio Claro, da Universidade Estadual Paulista “Júlio de Mesquita Filho”, para obtenção do grau de Bacharel em Ciências Biológicas.

BANCA EXAMINADORA:

Dr. Caio Felipe Cavicchia Zamunér

Dr. Giovane Böerner Hypolito

Dr. Igor Vinicius Ramos Otero

Aprovado em: 07 de novembro de 2025

Documento assinado digitalmente
 **GABRIEL FELICIO SILVA**
Data: 24/11/2025 11:46:24-0300
Verifique em <https://validar.iti.gov.br>

Assinatura do discente

Documento assinado digitalmente
 **HENRIQUE FERREIRA**
Data: 25/11/2025 07:48:00-0300
Verifique em <https://validar.iti.gov.br>

Assinatura do orientador

Documento assinado digitalmente
 **CAIO FELIPE CAVICCHIA ZAMUNER**
Data: 26/11/2025 08:41:52-0300
Verifique em <https://validar.iti.gov.br>

Assinatura do coorientador

*Dedico esse trabalho em memória da minha amada avó,
Neuza, que permanece viva em minhas lembranças e em cada
conquista da minha vida.*

AGRADECIMENTOS

Agradeço, em primeiro lugar, aos meus pais, Andreia e Carlos, que amo incondicionalmente. Sem o apoio, o incentivo e a confiança de vocês, nada disso teria sido possível. Cada gesto de cuidado e cada palavra de encorajamento me permitiram chegar até aqui, e serei eternamente grato.

Ao Professor Henrique Ferreira, agradeço pela oportunidade de orientação, pela confiança e pelas valiosas contribuições, fundamentais para meu crescimento acadêmico e pessoal. À equipe do Laboratório de Genética de Microrganismos, sou grato pela colaboração, pelo apoio técnico e pelo ambiente de troca científica que enriqueceu minha formação. Em especial, ao meu coorientador Caio Zamunér, obrigado pelos ensinamentos, pela paciência e pela dedicação durante o desenvolvimento desta pesquisa e da minha Iniciação Científica.

Aos meus amigos Ana, João, Caio e Maria, expresso profunda gratidão. Vocês tornaram essa jornada mais leve, oferecendo apoio, incentivo e momentos de alegria que transformam desafios em experiências memoráveis. Com vocês, essa trajetória acadêmica não foi apenas de aprendizado, mas também um período repleto de lembranças e vínculos que levarei para a vida toda. Vocês foram o motivo pelo qual nunca desisti e jamais esquecerei cada instante vivido juntos.

O presente trabalho foi realizado com o apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brasil. Processo nº 2024/13851-3. As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade do autor e não necessariamente refletem a visão da FAPESP.

"The next morning, on opening the incubator, I experienced one of those rare moments of intense emotion ...I saw that the broth culture, which the night before had been very turbid, was perfectly clear: all the bacteria had vanished, they had dissolved away like sugar in water. As for the agar spread, it was devoid of all growth ... in a flash, I had understood: what caused my clear spots was in fact an invisible microbe, a filterable virus, but a virus parasitic on bacteria."

(Felix d'Herelle, 1948)

RESUMO

Este trabalho buscou investigar o potencial de despolimerases fágicas como alternativa no controle do cancro cítrico, doença de grande impacto econômico causada por *Xanthomonas citri* subsp. *citri*. A pesquisa concentrou-se na predição, clonagem, expressão e avaliação funcional de genes candidatos a despolimerases identificados em genomas de bacteriófagos associados a *X. citri*. Foram analisados 72 genomas fágicos por meio de um *pipeline* bioinformático que integrou ferramentas recentes de predição baseadas em aprendizado de máquina (PhageDPO, DePP e DepoScope) com análises de homologia estrutural (HHpred), modelagem tridimensional (AlphaFold) e busca de domínios conservados (Foldseek). A partir dessas etapas, foram selecionadas duas proteínas com alto grau de identidade estrutural e indícios de domínios de ligação a açúcares, escolhida a proteína XacP77_034 como alvo para clonagem e expressão. Na etapa experimental, fragmentos correspondentes ao gene completo (XP77G) e ao domínio de degradação de açúcares (XP77SD) foram clonados em vetores de expressão, seguidos de ensaios de produção em *E. coli*. Embora tenha sido possível obter proteínas recombinantes solúveis após protocolos de purificação e *refolding*, os testes funcionais contra linhagens de *X. citri* não demonstraram atividade de degradação de exopolissacarídeos, sugerindo que a proteína não apresenta função de despolimerase nas condições avaliadas. O resultado indica que, embora o *pipeline* bioinformático e a estratégia de expressão tenham se mostrado eficientes na seleção de candidatas e na superação de dificuldades de solubilidade, as ferramentas de predição atuais ainda apresentam limitações importantes, principalmente quanto aos seus bancos de dados de usados para o seu treinamento. Também é possível que a atividade enzimática dependa de condições específicas ou da presença de cofatores, fatores não investigados neste estudo. Nesse contexto, a principal contribuição deste trabalho foi a validação de um *pipeline* integrado para a prospecção de despolimerases, evidenciando, ao mesmo tempo, a necessidade de expandir os bancos de dados com validações funcionais em uma gama mais ampla de bacteriófagos, a fim de aumentar a confiabilidade das predições em estudos futuros.

Palavras-chave: Bacteriófago; Cancro cítrico; Despolimerase; Exopolissacarídeos.

ABSTRACT

This study aimed to investigate the potential of phage-derived depolymerases as an alternative strategy for the control of citrus canker, a phytopathology of major economic impact caused by *Xanthomonas citri* subsp. *citri* (*X. citri*). The research focused on the prediction, cloning, expression, and functional evaluation of candidate depolymerase genes identified in bacteriophage genomes associated with *X. citri*. A total of 72 phage genomes were analyzed through a bioinformatic pipeline that integrated recent machine learning-based prediction tools (PhageDPO, DePP, and DepoScope) with structural homology analyses (HHpred), three-dimensional modeling (AlphaFold), and conserved domain searches (Foldseek). From these steps, two proteins were selected due to their high structural identity and evidence of sugar-binding domains, with protein XacP77_034 chosen as the target for cloning and expression. In the experimental stage, fragments corresponding to the full gene (XP77G) and the sugar-degrading domain (XP77SD) were cloned into expression vectors and subjected to production assays in *Escherichia coli*. Although soluble recombinant proteins were obtained after purification and refolding protocols, functional tests against *X. citri* strains did not demonstrate exopolysaccharide-degrading activity, suggesting that the protein does not exhibit depolymerase function under the evaluated conditions. These findings indicate that, while the bioinformatic pipeline and expression strategy were effective in candidate selection and in overcoming solubility challenges, current prediction tools still present important limitations, particularly concerning the scope and representativeness of their training datasets. It is also possible that enzymatic activity depends on specific conditions or the presence of cofactors, factors not explored in this study. In this context, the main contribution of this work was the validation of an integrated pipeline for depolymerase prospecting, while simultaneously highlighting the need to expand databases with functional validations across a broader range of bacteriophages in order to improve the reliability of predictions in future studies.

Keywords: Bacteriophage; Citrus canker; Depolymerase; Exopolysaccharides.

Title in english: Evaluation of Phage-Derived Depolymerases Against *Xanthomonas citri* subsp. *citri*.

LISTA DE FIGURAS

Figura 1 – Painel representativo compreendendo as etapas das análises de bioinformática.....	24
Figura 2 – Alinhamento múltiplo das proteínas P27_002 e XacP77_034.....	26
Figura 3 – Modelagem da proteína P27_002 construída com o AlphaFold.....	28
Figura 4 – Modelagem da proteína XacP77_034 construída com o AlphaFold.....	29
Figura 5 – Eletroforese em gel de agarose da PCR de <i>E. coli</i> DH10B contendo a construção pET28-XP77SD.....	31
Figura 6 – Microscopia óptica por contraste de interferência diferencial de <i>E. coli</i> BL21 contendo a construção pET28-XP77SD após 4h de indução.....	32
Figura 7 – SDS-PAGE da expressão e <i>refolding</i> de XP77SD após 4h de indução.....	33
Figura 8 – SDS-PAGE da expressão e purificação de XP77G após 24h de indução..	35
Figura 9 – SDS-PAGE da expressão e purificação do plasmídeo pTSL vazio após 24h de indução.....	36
Figura 10 – Testes enzimáticos para a proteína XP77G contra as linhagens de <i>X. citri</i> 306 e FDC2.....	37

LISTA DE TABELAS

Tabela 1 – 72 fagos isolados utilizados no estudo.....	17
Tabela 2 – <i>Primers</i> desenhados para as estratégias de clonagem idealizadas.....	19
Tabela 3 – Representação das proteínas identificadas no genoma do fago XacP77 após análise das ferramentas DePP e PhageDPO.....	25
Tabela 4 – Representação das proteínas identificadas no genoma do fago P27 após análise das ferramentas DePP e PhageDPO.....	25
Tabela 5 – Construções gênicas empregadas.....	30

SUMÁRIO

1 INTRODUÇÃO.....	10
1.1 Objetivo geral.....	15
1.2 Objetivos específicos.....	15
2 MATERIAIS E MÉTODOS.....	16
2.1 Bacteriófagos, Bactérias e Condições de Cultivo.....	16
2.2 Identificação das Despolimerases.....	16
2.3 Isolamento e Clonagem dos Genes.....	19
2.3.1 Reação de PCR.....	19
2.4 Expressão e purificação das proteínas.....	20
2.4.1 Checagem inicial da expressão de proteínas.....	20
2.4.2 Minipreps de proteínas.....	21
2.4.3 Expansão do cultivo bacteriano.....	22
2.4.4 Purificação a partir das culturas de 500 ML.....	22
2.5 Avaliação da atividade enzimática em placa de cultivo.....	22
3 RESULTADOS.....	24
3.1 Análises de Bioinformática.....	24
3.2 Clonagem e Expressão.....	30
4 DISCUSSÃO.....	38
5 CONCLUSÃO.....	41
REFERÊNCIAS.....	42
ANEXO A – CONSTRUÇÃO PET28-XP77SD.....	46
ANEXO B – CONSTRUÇÃO PTSL-XP77G.....	47

1 INTRODUÇÃO

Comumente referidos como fagos, os bacteriófagos são vírus naturais que infectam especificamente bactérias. A sua estrutura é geralmente caracterizada por uma cabeça proteica icosaédrica que contém o material genético, uma cauda proteica contrátil ou não-contrátil, longa ou curta, responsável pelo reconhecimento da célula hospedeira e à inserção do material genético (Cuervo; Carrascosa, 2012; Nobrega *et al.*, 2018).

Os fagos foram inicialmente identificados pelo microbiologista Felix d'Herelle, que os denominou bacteriófagos, e posteriormente, aplicou com sucesso preparações de fagos para tratar crianças com disenteria bacteriana em 1919 (Strathdee *et al.*, 2023). Após a sua descoberta, os fagos passaram a ser utilizados para o tratamento de infecções bacterianas em humanos e animais na década de 1930 (Summers, 2012). Com o surgimento da penicilina na década de 1940, a fagoterapia caiu em desuso, porém, desde a última década, devido ao crescente problema da resistência microbiana aos antibióticos, o uso de fagos para o tratamento de doenças ocasionadas por bactérias tem ganhado atenção novamente (Melo *et al.*, 2020).

Podemos dividir os fagos em dois grupos, relacionados aos seus ciclos de replicação e propagação. O primeiro grupo inclui os fagos de ciclo lítico, cuja replicação do fago culmina com a lise da célula bacteriana, resultando na liberação de novos vírions (Hobbs; Abedo, 2016). Em contrapartida, existem os fagos de ciclo temperado, anteriormente chamados de lisogênicos, os quais podem integrar seu genoma ao da bactéria hospedeira como um profago, replicando-se em sincronia com o DNA bacteriano sem causar lise imediata (Strathdee *et al.*, 2023). Dessa forma, os fagos líticos são comumente considerados mais adequados para o tratamento de infecções bacterianas, pois estes matam uma proporção muito alta das células que infectam.

Um dos principais desafios à eficácia dos fagos é a presença de biofilmes bacterianos, agregados complexos de micro-organismos embebidos em uma matriz de substâncias poliméricas extracelulares (EPS). Essa matriz é composta por polissacarídeos, proteínas, lipídeos e ácidos nucleicos que envolve as células bacterianas, protegendo-as de estresses ambientais, oferecendo assim uma defesa robusta contra a ação de fagos e antibióticos (Grigson *et al.*, 2023). A

composição do EPS varia significativamente entre bactérias Gram-positivas e Gram-negativas, refletindo as diferenças estruturais e funcionais de suas paredes celulares. As bactérias Gram-positivas possuem uma parede celular espessa composta principalmente por peptidoglicano, enquanto as Gram-negativas têm uma parede celular mais complexa, com uma camada fina de peptidoglicano e uma membrana externa contendo lipopolissacarídeos (LPS) (Latham, 2021).

No caso de bactérias Gram-positivas, tais como na família *Staphylococci*, o EPS frequentemente contém polissacarídeos catiônicos (Donlan, 2002). Essas substâncias contribuem para a aderência e estabilidade do biofilme, facilitando a colonização de superfícies. Além disso, o EPS de Gram-positivas também contém quantidades consideráveis de proteínas, como enzimas e estruturas proteicas como pili e fímbrias, auxiliando na formação e manutenção do biofilme (Muhammad *et al.*, 2020). Por outro lado, em bactérias Gram-negativas, o EPS é frequentemente composto por polissacarídeos neutros ou polianiónicos, proteínas e lipídios derivados da membrana externa (Muhammad *et al.*, 2020). A propriedade aniônica é conferida, principalmente, por ácidos urônicos (como ácido D-glucurônico, ácido D-galacturônico e ácido manurônico) e piruvatos ligados por cetil, os quais permitem a associação de cátions divalentes como cálcio e magnésio, que se ligam às cadeias poliméricas, fortalecendo a adesão no biofilme maduro (Donlan, 2002).

Para superar essa barreira, alguns fagos codificam enzimas denominadas despolimerases. Essas enzimas desempenham um papel fundamental na potencialização do desempenho dos bacteriófagos contra biofilmes bacterianos, degradando polissacarídeos capsulares (CPS), lipopolissacarídeos (LPS) ou EPS presentes nos biofilmes, facilitando a adsorção dos fagos às células bacterianas e promovendo sua difusão através da matriz do biofilme (Oliveira; Drulis-Kawa; Azevedo, 2022). As despolimerases fágicas se apresentam como componentes integrais das partículas fágicas ou como proteínas solúveis liberadas durante a lise das células hospedeiras (Hobbs; Abedo, 2016). Essas enzimas frequentemente apresentam uma estrutura homotrimérica, o que possibilita múltiplos sítios de ligação para interagir com os polissacarídeos específicos da superfície bacteriana, com estruturas terciárias altamente estáveis (Knecht, Veljkovic, Fieseler, 2020).

O mecanismo de ação das despolimerases envolve etapas sequenciais distintas. Primeiramente, a enzima se liga ao polissacarídeo na superfície bacteriana. Essa ligação é mediada por interações específicas entre a enzima e o polissacarídeo, facilitadas pelo sítio ativo da enzima. Uma vez ligada, a despolimerase catalisa a clivagem das cadeias de polissacarídeos em locais específicos, destruindo as barreiras protetoras das bactérias e permitindo ao fago injetar seu DNA e infectar a célula bacteriana (Knecht, Veljkovic, Fieseler, 2020).

A atividade das despolimerases pode ser observada através da formação de zonas opacas (halos) ao redor de círculos claros em placas de lise. Esses halos resultam na degradação dos EPS da cápsula bacteriana, expondo as bactérias aos fagos. O halo se expande com o tempo devido à contínua degradação da cápsula bacteriana e pode ser influenciado pela difusão de enzimas livres ou mudanças no estado fisiológico das bactérias (Knecht, Veljkovic, Fieseler, 2020).

Essas enzimas podem ser agrupadas em dois principais tipos com base no modo de ação: hidrolases e liases. Hidrolases são enzimas que catalisam a hidrólise de ligações glicosídicas, adicionando moléculas de água, quebrando a ligação entre as moléculas de açúcar (Pires *et al.*, 2016). Este grupo inclui as sialidases, levanases, xilanases, dextranases, rhamnosidases e peptidases. O tipo específico de ligação glicosídica alvo depende da especificidade do substrato da enzima; por exemplo, sialidases clivam a ligação α de ácidos siálicos terminais, enquanto xilanases clivam ligações glicosídicas β -1,4 no xilano (Pires *et al.*, 2016). Em contraste, as liases catalisam a clivagem de ligações glicosídicas em polissacarídeos sem adição de água, quebrando ligações glicosídicas (1,4) através de um mecanismo de β -eliminação (Pires *et al.*, 2016). Este processo envolve a eliminação de uma molécula, geralmente água, resultando na formação de uma ligação dupla. Especificamente, as liases catalisam a clivagem das ligações glicosídicas ao romper a ligação entre o átomo de carbono anomérico e o átomo de carbono adjacente na molécula de açúcar. Exemplos de liases incluem liases de alginato e pectina/pectato, as quais exibem atividades específicas em relação a diferentes tipos de substâncias poliméricas (Pires *et al.*, 2016).

É importante ressaltar que despolimerases não matam diretamente as bactérias, mas removem as camadas de polissacarídeos protetores das células bacterianas, expondo-as e sensibilizando-as aos componentes do sistema

imunológico ou a agentes antibacterianos, podendo assim potencializar a penetração destes no biofilme, demonstrando um efeito sinérgico com alguns antibióticos (Guo; Liu; Zhang, 2023). Diante deste cenário, existe um grande potencial para preparações de despolimerases fágicas para reduzir significativamente o uso de antibióticos na medicina, agricultura, aquicultura, pecuária e medicina veterinária, auxiliando no combate a doenças de origem bacteriana que afetam humanos, animais ou plantas.

Dentre estas, podemos citar o cancro cítrico, causado pela bactéria Gram-negativa *Xanthomonas citri* subsp. *citri* (*X. citri*), uma doença sem cura conhecida que afeta todos os tipos comerciais de citros, resultando na queda prematura de folhas e frutos, diminuindo a produtividade (Naqvi *et al.*, 2022). Em *X. citri*, um dos principais EPS produzidos é a goma xantana, a qual possui a capacidade de se ligar a certos polissacarídeos da parede celular das plantas (Rudolph, 1993). Essa interação leva ao encharcamento das bactérias, as quais ficam embebidas em um limo semelhante a um gel no espaço intercelular. Essa camada de gel proporciona um ambiente propício para a colonização bacteriana, oferece uma barreira física contra compostos bacteriostáticos e facilita o seu crescimento. Além disso, a alta capacidade de retenção de água da goma xantana ajuda as bactérias a sobreviverem durante condições desfavoráveis, garantindo sua persistência e sucesso na colonização de hospedeiros vegetais (Rudolph, 1993).

A produção da goma xantana envolve múltiplos genes, incluindo o agrupamento gênico *gum* (*gumBCDEFGHIJKLM*) e o gene *xanA* (Picchi *et al.*, 2021). Testes anteriores demonstraram que a deleção dos genes *gumD* e *xanA* afeta a sobrevivência epifítica de *X. citri* e leva a diminuição dos sintomas causados pela bactéria (Dunger *et al.*, 2007; Picchi *et al.*, 2021). Isto pois, a redução na produção de goma xantana pelos mutantes afeta a sua capacidade de se aderir aos tecidos do hospedeiro, dificultando a sua sobrevivência sob condições de estresse o que afeta a sua capacidade de colonizar a planta (Picchi *et al.*, 2021). Dessa forma, a aplicação de despolimerases em plantas infectadas com cancro cítrico poderia levar a degradação da goma xantana das bactérias. Isto possivelmente pode causar a diminuição da sua sobrevivência epifítica, bem como no subsequente desenvolvimento dos sintomas associados ao cancro. Além disso, a degradação da goma xantana poderia levar a maior

susceptibilidade das bactérias à ação de substâncias bactericidas como o cobre. A pulverização de soluções contendo esse metal são, atualmente, o principal meio para o controle dessa doença (FERENCE *et al.*, 2018). Porém, a sua eficácia depende da aplicação com alta frequência, o que pode ser prejudicial ao meio ambiente, dado o caráter biocumulativo do cobre, e levar a seleção de linhagens resistentes (Behlau *et al.*, 2012). Assim, a utilização de despolimerases fágicas poderia oferecer uma ação sinérgica ao cobre, reduzindo a quantidade e frequência da aplicação deste metal.

A identificação de genes candidatos para sequências de despolimerases é realizada principalmente por meio de ferramentas de bioinformática. No entanto, a literatura envolvendo as despolimerases codificadas pelos bacteriófagos que infectam *X. citri* ainda é muito defasada o que torna um desafio a identificação e principalmente a validação das despolimerases em *X. citri*. O processo tradicional para a identificação de despolimerases foi inicialmente descrito por Latka *et al.*, (2019). Esse incluiu a filtragem de proteínas de ligação ao hospedeiro (RBPs) de fagos, seguida pela busca por homologia, utilizando as ferramentas BlastP (Altschul *et al.*, 1990), Phyre2 (Kelley *et al.*, 2015), SWISS-MODEL (Bordoli e Schwede, 2011), HMMER (Finn, Clements e Eddy, 2011) e HHPred (Soding, Biegert e Lupas, 2005). Os critérios para a seleção de proteínas com potencial atividade de despolimerase incluíam: tamanho (>200 resíduos), anotação como *tail fibers*, *tail spikes* ou proteínas hipotéticas no banco de dados NCBI, e homologia com domínios enzimáticos conhecidos (liase ou hidrolase). A homologia com um desses domínios deveria abranger pelo menos 100 resíduos, e uma estrutura β -helicoidal típica deveria ser sugerida pelo Phyre2 (Vieira *et al.*, 2023).

Contudo, este método é considerado trabalhoso e, dessa forma, ferramentas mais baseadas em *machine learning*, alimentadas com bancos de dados de enzimas validadas experimentalmente foram desenvolvidas como método alternativo para facilitar a identificação das despolimerases com uma maior acurácia. Dentre elas temos a PhageDPO, DepoScope e PhageDePP (Vieira *et al.*, 2023; Concha-Eloko *et al.*, 2024; Magill; Skvortsov, 2023), que propõe facilitar a busca e identificação de sequências de despolimerases.

No entanto, tais ferramentas de bioinformática são muito recentes e, portanto, poucas pesquisas para a sua validação foram realizadas, principalmente voltada

para bactérias de interesse agrícola. Assim, o objetivo deste projeto é identificar candidatos a despolimerases através das ferramentas de bioinformática disponíveis atualmente e, assim, e avaliar a atividade destas enzimas contra *X. citri*.

1.1 Objetivo geral

Avaliar a atividade de despolimerases de novos isolados de bacteriófagos que possuem a bactéria *X. citri* como hospedeiro.

1.2 Objetivos específicos

- a) Identificar genes, a partir de ferramentas de bioinformática, que codificam possíveis despolimerases em novos isolados de bacteriófagos que infectam *X. citri*.
- b) Realizar anotações das sequências identificadas com o a ferramenta Phold e confirmação das predições utilizando o ColabFold seguido do Foldseek.
- c) Expressar e purificar despolimerases codificadas por fagos de *X. citri*

2 MATERIAIS E MÉTODOS

2.1 Bacteriófagos, Bactérias e Condições de Cultivo

A fim de propagar os bacteriófagos, foi utilizada a linhagem de *Xanthomonas citri* subsp. *citri* FDC2 (Ragupathy *et al.*, 2023), sensível aos bacteriófagos. As bactérias foram cultivadas utilizando meio NYG (peptona 5 g/L, extrato de levedura 3 g/L e glicerol 20 g/L; ágar 15g/L quando necessário) a 29°C, como também em meio líquido a 200 rpm. Ademais, a linhagem de *X. citri* 306 (Da Silva *et al.*, 2002), resistente aos bacteriófagos, também foi utilizada em testes enzimáticos, sendo cultivadas nas mesmas condições.

Para a propagação dos bacteriófagos, 100µl de uma suspensão já purificada, juntamente com 100 µl de uma cultura de *X. citri* FDC2, cultivada *overnight*, foram adicionadas à uma placa já contendo um pequeno filme de NYG sólido com 1,5% de ágar e, sobre esta, foi vertido 3 mL de NYG soft-ágar (5 g/L de peptona; 3 g/L de extrato de levedura; 20 g/L de glicerol e 0,65% de ágar (w/v)). O crescimento foi feito a 29 °C por 24 horas. Após este período, foi adicionado 5mL de tampão SM (100 mM NaCl; 10 mM MgSO₄; 50 mM Tris-HCl pH 7.4; 0.01% ágar (v/v)). As placas foram incubadas a 12 °C por 24 horas e, em seguida, essa solução foi filtrada em filtro 0,22 µm e estocadas a 12 °C.

Para a clonagem e expressão dos genes foram utilizadas, respectivamente, as linhagens de *E. coli* DH10B e BL21. Estas foram cultivadas em meio LB (extrato de levedura 5 g/L, triptona 10 g/L, NaCl 10 g/L; ágar 15 g/L quando necessário) a 37°C, como também em meio líquido a 200 rpm.

2.2 Identificação das Despolimerases

Foram obtidas as sequências codificantes das despolimerases a partir de dados genômicos isolados de 72 fagos (Tabela 1), os quais incluíam o re-sequenciamento das linhagens dos fagos CP1 e CP2 (Ahmad *et al.*, 2014). A anotação funcional das sequências do genoma foi realizada com a utilização do *software* Pharokka v. 1.7.2 (Bouras, 2024).

Tabela 1 – 72 fagos isolados utilizados no estudo

ID do Fago	Linhagem Hospedeira	Sequenciado	No. de Genes	Tamanho (nt)	Familia Prevista	Origem	Fonte	GenBank/ RefSeq
CP1	FL71 /FL72	Re-sequencia do	4	43870	Autographiviridae	Japão	Ahmad <i>et al.</i> , 2014	AB720063.2
CP2	FL72/FL75	Re-sequenciado	5	42963	Autographiviridae	Japão	Ahmad <i>et al.</i> , 2014	NC_020205.1
XacP77 (OL ccΦ7)	FL75	Este estudo	0	76109	Schitoviridae	Estados Unidos	Balogh <i>et al.</i> , 2008	
XacP157 (Φxaac A2003-1)	FL75	Este estudo	9	44166	Autographiviridae	Argentina	J.B. Jones ^a	
XacP165 (Φxaac A2003-9)	FL75	Este estudo	9	43808	Autographiviridae	Argentina	J.B. Jones ^a	
XacP180 (Φxaac A2003-24)	FL75	Este estudo	8	43604	Autographiviridae	Argentina	J.B. Jones ^a	
XacP186 (Φxaac A2003-30)	FL75	Este estudo	6	44113	Autographiviridae	Argentina	J.B. Jones ^a	
P1	FDC2	Este estudo	6	43697	Autographiviridae	Brasil	Este estudo	
P2	FDC2	Este estudo	5	43385	Autographiviridae	Brasil	Este estudo	
P3	FDC2	Este estudo	5	41312	Autographiviridae	Brasil	Este estudo	
P4	FDC2	Este estudo	6	43398	Autographiviridae	Brasil	Este estudo	
P5	FDC2	Este estudo	6	42357	Autographiviridae	Brasil	Este estudo	
P6	FDC2	Este estudo	6	43510	Autographiviridae	Brasil	Este estudo	
P7	FDC2	Este estudo	5	42357	Autographiviridae	Brasil	Este estudo	
P8	FDC2	Este estudo	5	43398	Autographiviridae	Brasil	Este estudo	
P9	FDC2	Este estudo	6	43725	Autographiviridae	Brasil	Este estudo	
P10	FDC2	Este estudo	6	42239	Autographiviridae	Brasil	Este estudo	
P11	FDC2	Este estudo	6	39034	Autographiviridae	Brasil	Este estudo	
P12	FDC2	Este estudo	6	43304	Autographiviridae	Brasil	Este estudo	
P13	FDC2	Este estudo	8	43398	Autographiviridae	Brasil	Este estudo	
P14	FDC2	Este estudo	6	43326	Autographiviridae	Brasil	Este estudo	
P15	FDC2	Este estudo	5	43582	Autographiviridae	Brasil	Este estudo	
P16	FDC2	Este estudo	5	43582	Autographiviridae	Brasil	Este estudo	
P17	FDC2	Este estudo	6	43398	Autographiviridae	Brasil	Este estudo	
P18	FDC2	Este estudo	8	43725	Autographiviridae	Brasil	Este estudo	
P19	FDC2	Este estudo	6	43697	Autographiviridae	Brasil	Este estudo	
P20	FDC2	Este estudo	6	43522	Autographiviridae	Brasil	Este estudo	
P21	FDC2	Este estudo	5	43398	Autographiviridae	Brasil	Este estudo	
P23	FDC2	Este estudo	5	43582	Autographiviridae	Brasil	Este estudo	
P24	FDC2	Este estudo	5	43398	Autographiviridae	Brasil	Este estudo	
P25	FDC2	Este estudo	6	43522	Autographiviridae	Brasil	Este estudo	
P26	FDC2	Este estudo	6	43398	Autographiviridae	Brasil	Este estudo	

P27	FDC2	Este estudo	5	43398	Autograph iviridae	Brasil	Este estudo
P29	FDC102	Este estudo	6	43385	Autograph iviridae	Brasil	Este estudo
P30	FDC102	Este estudo	4	43089	Autograph iviridae	Brasil	Este estudo
P31	FDC102	Este estudo	6	43398	Autograph iviridae	Brasil	Este estudo
P32	FDC102	Este estudo	6	43398	Autograph iviridae	Brasil	Este estudo
P33	FDC102	Este estudo	6	43579	Autograph iviridae	Brasil	Este estudo
P34	FDC102	Este estudo	6	43398	Autograph iviridae	Brasil	Este estudo
P35	FDC102	Este estudo	6	43398	Autograph iviridae	Brasil	Este estudo
P37	FDC102	Este estudo	6	43698	Autograph iviridae	Brasil	Este estudo
P38	FDC102	Este estudo	6	43313	Autograph iviridae	Brasil	Este estudo
P40	FDC102	Este estudo	4	43161	Autograph iviridae	Brasil	Este estudo
P41	FDC102	Este estudo	6	43326	Autograph iviridae	Brasil	Este estudo
P42	FDC102	Este estudo	7	43579	Autograph iviridae	Brasil	Este estudo
P43	FDC102	Este estudo	6	43398	Autograph iviridae	Brasil	Este estudo
P44	FDC102	Este estudo	4	43398	Autograph iviridae	Brasil	Este estudo
P45	FDC102	Este estudo	5	43579	Autograph iviridae	Brasil	Este estudo
P46	FDC102	Este estudo	6	43698	Autograph iviridae	Brasil	Este estudo
P47	FDC102	Este estudo	6	44072	Autograph iviridae	Brasil	Este estudo
P48	FDC102	Este estudo	4	43304	Autograph iviridae	Brasil	Este estudo
P50	FDC102	Este estudo	6	43579	Autograph iviridae	Brasil	Este estudo
P51	FDC102	Este estudo	5	43398	Autograph iviridae	Brasil	Este estudo
P52	FDC102	Este estudo	6	43398	Autograph iviridae	Brasil	Este estudo
P53	FDC102	Este estudo	6	43579	Autograph iviridae	Brasil	Este estudo
P54	FDC102	Este estudo	6	43398	Autograph iviridae	Brasil	Este estudo
P55	FDC2	Este estudo	4	43398	Autograph iviridae	Brasil	Este estudo
P56	FDC2	Este estudo	6	43398	Autograph iviridae	Brasil	Este estudo
P57	FDC2	Este estudo	6	43398	Autograph iviridae	Brasil	Este estudo
P58	FDC2	Este estudo	4	43326	Autograph iviridae	Brasil	Este estudo
P59	FDC102	Este estudo	5	43398	Autograph iviridae	Brasil	Este estudo
P60	FDC102	Este estudo	5	43326	Autograph iviridae	Brasil	Este estudo
P61	FDC102	Este estudo	5	43398	Autograph iviridae	Brasil	Este estudo
P70	FDC2	Este estudo	6	43398	Autograph iviridae	Brasil	Este estudo
P71	FDC2	Este estudo	7	42934	Autograph iviridae	Brasil	Este estudo
P72	FDC2	Este estudo	8	42934	Autograph iviridae	Brasil	Este estudo
P73	FDC2	Este estudo	5	43398	Autograph iviridae	Brasil	Este estudo
P74	FDC102	Este estudo	4	43522	Autograph iviridae	Brasil	Este estudo
P75	FDC2	Este estudo	6	43522	Autograph iviridae	Brasil	Este estudo
P76	FDC102	Este estudo	6	43698	Autograph iviridae	Brasil	Este estudo

P77	FDC2	Este estudo	0	43626	Autograph iviridae	Brasil	Este estudo	
LeafA	FDC2	Este estudo	6	43348	Autograph iviridae	Brasil	Este estudo	Virulenta

^aUniversity of Florida, Plant Pathology Department, 1453 Fifield Hall, Gainesville, FL 32611

Fonte: Dados da pesquisa

A busca por sequências de possíveis despolimerases foi realizada com diferentes ferramentas de bioinformática, desenvolvidas com base em *machine-learning* disponíveis (PhageDPO, DePP e DepoScope) (Vieira *et al.*, 2023; Magill; Skvortsov, 2023; Concha-Eloko *et al.*, 2024). Ainda, as sequências selecionadas foram submetidas a busca por homólogos estruturais utilizando a ferramenta HHpred. Foram desenhados *primers* (Tabela 2) para as sequências de despolimerases selecionadas, e estes foram avaliados quanto à sua estabilidade, formação de estrutura secundária, Tm, etc., usando o *software* *Generunner*.

Tabela 2 - *Primers* desenhados para as estratégias de clonagem idealizadas.

Alvos	Primer	Enzima de Restrição	Sequência (5' – 3')
XP77G	Foward	EcoRI	AAAGAATTCATGACCAGCCCGCACATCCTC
	Reverse	HindIII	TTTAAGCTTTCATGCGAGCGTTCCAGAAAG
XP77SD	Foward	EcoRI	AAAGAATTCATGACCAGCCCGCACATCCTC
	Reverse	XhoI	TTTCTCGAGTCACACGGGCGGGATGCG

Fonte: Dados da pesquisa

2.3 Isolamento e Clonagem dos Genes

Todas as enzimas de restrição e modificação, bem como as polimerases usadas no PCR, foram obtidas da Thermo-Scientific. A amplificação dos genes de despolimerases foi feita com a enzima Taq ou DNA polimerase de alta fidelidade, de acordo com as especificações do fabricante.

2.3.1 Reação de PCR

1 U de Taq DNA polimerase, 1 µM de cada *primer*, 100 µM de dNTPs, DMSO 3%, tampão enzimático conforme recomendado e 2,5 mM de MgCl₂ (quando necessário). O protocolo de PCR touchdown foi conduzido nas seguintes condições.

(i) Para o fragmento XP77G, iniciou-se com desnaturação a 94 °C por 2:30 min,

seguida por nove ciclos compostos por 94 °C por 0:45 s, 64 °C por 1:00 min e 72 °C por 2:30 min. A cada ciclo, a temperatura de anelamento foi reduzida em 0,5 °C, até atingir 60 °C. Em seguida, foram realizados 30 ciclos convencionais de PCR, consistindo em 95 °C por 0:45 s, 63 °C por 1:00 min e 72 °C por 2:30 h. A reação foi finalizada com extensão final a 72 °C por 5:00 min. (ii) Para o fragmento XP77SD, a reação seguiu o mesmo esquema, exceto pelo ciclo de anelamento do *touchdown*, no qual a temperatura variou de 68 °C a 64 °C (com decréscimo de 0,5 °C por ciclo), e pelo anelamento dos ciclos convencionais, conduzido a 64 °C. Os produtos de PCR foram analisados através de eletroforese em géis de agarose (0,7% em 1X TAE), e bandas dos tamanhos esperados foram extraídas dos géis para clonagem subsequente usando o kit GeneJet Gel Extraction (Thermo-Scientific). Em seguida, produtos de PCR isolados foram cortados usando as enzimas de restrição *EcoRI/HindIII/XhoI* e, com as mesmas enzimas, o excerto foi ligado ao vetor de expressão pET-28a(+) (Novagen) e/ou pTSL, construído com a proteína com atividade de chaperona *SlyD* (Taylor *et al.*, 2016).

Tais produtos de ligação foram transformados em linhagens quimicamente competentes de *E. coli* DH10B (Invitrogen) para seleção (a preparação de linhagens de *E. coli* quimiocompetentes e a transformação seguiram Sambrook; Fritsch; Maniatis, 1989). Por fim, a integridade dos vetores de expressão foi avaliada por perfil de restrição e sequenciamento de DNA.

2.4 Expressão e purificação das proteínas

A fim de realizar a expressão das despolimerases, vetores construídos foram transformados em linhagens de *E. coli* BL21 (DE3) quimiocompetentes.

2.4.1 Checagem inicial da expressão de proteínas

A partir de um pré-inóculo cultivado por 12h, as células foram aplicadas em 50 mL de meio LB suplementado com 100 µg/mL do antibiótico Ampicilina, sendo mantidas sob agitação à 37 °C e 200 rpm, com uma D.O. 600nm inicial de ~0.1. Quando a D.O. 600 nm atingiu ~0,5, foi adicionado 1 mM de isopropil-β-D-tiogalactopiranosídeo (IPTG), e culturas foram cultivadas por 24 horas a 30 °C. Após este período, foram coletadas alíquotas de 1 mL no tempo zero (antes da indução), 4 horas e 24 horas de indução, a fim de verificar a expressão da

proteína. As alíquotas foram centrifugadas a 12.000 xg/2 min, e os *pellets* celulares foram processados para a realização do SDS-PAGE. Este processo, bem como o próprio SDS-PAGE e a coloração com Coomassie foram realizados conforme Sambrook; Fritsch; Maniatis (1989), com a utilização de um sistema de mini-gel vertical BioRad. Após certificar a expressão das nossas proteínas em SDS-PAGE, foram feitos *minipreps* de proteínas para verificar a expressão/capacidade funcional correta das His6-TAGs.

2.4.2 *Minipreps de proteínas*

Ao final da expressão, os *pellets* das culturas de 50 mL (após indução por 24h à 30°C) foram coletados por centrifugação a 7.000 xg/15 min. Em caso de estes não serem processados diretamente, eles foram congelados e mantidos a -80 °C. *Pellets* de células frescas ou descongeladas foram mantidos no gelo e a estes foi adicionado 1 ml de tampão A (30 mM Tris-HCl pH= 8; 60 mM NaCl; 10% glicerol). Tais suspensões celulares acomodadas em tubos de microcentrífuga de 1,5 mL foram colocadas em gelo e submetidas a 20 ciclos de lise por ultrassom (cada ciclo composto por: 10s ligado/60s desligado) utilizando-se o sonicador VibraCell (VCX-130), equipado com uma sonda (ponta de 3 mm). Os lisados celulares foram clarificados a 15.000 xg/10 min e 4 °C usando uma microcentrífuga de bancada. Os lisados (sobrenadante) foram transferidos para novos tubos, e os *pellets* foram dissolvidos em tampão de amostra para proteínas contendo SDS (Sambrook; Fritsch; Maniatis, 1989) para verificar a possível presença de proteínas como precipitados em SDS-PAGE (corpos de inclusão). 200 µL de Sepharose rápida quelante (GE), carregada com Ni²⁺ (leito de 50%) foi adicionada sobre o sobrenadante (fase solúvel). A ligação ocorreu em *batch*, sendo mantida sob leve agitação no gelo por 1h. A mistura foi então centrifugada por 30 seg a 2000 xg para decantar a resina, e o sobrenadante (*flowthrough*) foi transferido para um tubo de microcentrífuga de 1,5 mL e reservado para posterior análise por SDS-PAGE. Em seguida, a resina foi submetida a uma etapa de lavagem com 900 µL do Tampão B (mesma formulação do Tampão A, acrescido de imidazol em concentrações variáveis), contendo 20 mM de imidazol, empregando-se novamente o sistema de *batch*/centrifugação. Para avaliação da *miniprep*, alíquotas de 40 µL de cada fração (*flowthrough* e lavagens) foram adicionadas a 10 µL de tampão de amostra para

SDS-PAGE (Sambrook; Fritsch; Maniatis, 1989) e submetidas à análise de por eletroforese.

2.4.3 *Expansão do cultivo bacteriano*

A melhor linhagem/sistema de expressão foi selecionada e cultivada em 20 mL de meio LB suplementado com 100 µg/mL de Ampicilina, a 37 °C, sendo mantida sob agitação a 200 rpm por 12h. Após este período, as culturas foram diluídas para 1:10 em 500 mL de LB/Ampicilina e incubadas a 37 °C e 200 rpm. Ao atingirem a D.O. 600nm de ~0,5, a indução ocorreu conforme definido nos experimentos iniciais.

2.4.4 *Purificação a partir das culturas de 500 ML*

Após a indução, as células foram coletadas por centrifugação a 10.000 rpm/30 min, e mantidas a -80 °C até o processamento. Estas células foram descongeladas, e foi adicionado 5 mL de tampão A. Após, estas foram submetidas à 20 ciclos de sonicação (cada ciclo composto por: 10s ligado/60s desligado), com a sonda (ponta de 6 mm). A clarificação foi feita por centrifugação a 20.000 rpm por 60 minutos a 9 °C. Os precipitados, quando houve necessidade, foram solubilizados em 8M de uréia (razão: 3 mL de 8M de uréia/1g de *pellet*) e centrifugados novamente para clarificação. Neste caso, a ureia foi removida por diálise sequencial contra o Tampão A com concentração decrescente de uréia em um esquema de duas dobras (6M, 4M, 2M, 1M e 0M) usando membranas de diálise (12K; SIGMA). Por fim, para a purificação da fase solúvel ou amostras solubilizadas, os lisados clarificados foram carregados em colunas de purificação carregadas com Ni²⁺. Após o empacotamento da fase estacionária, a coluna foi equilibrada com Tampão A e as amostras foram eluídas em um gradiente de Tampão B (20, 40, 80 e 200 mM de imidazol). Todos os volumes de eluição foram coletados e verificados em SDS-PAGE.

2.5 **Avaliação da atividade enzimática em placa de cultivo**

A fim de testar a atividade de lise das proteínas fágicas extraídas em relação às linhagem de *X. citri* FDC02 e 306, respectivamente sensível e resistente aos bacteriófagos, foi utilizado o método *spot-on-lawn*, de acordo com Oliveira *et al.*, (2019). Uma alíquota de 100 µL de inóculos de bactérias na fase *log* mediana foram adicionados em placas de Petri contendo meio NYG *soft*-ágar, a fim de formar

tapetes bacterianos. Sobre estas foram adicionados 5 μ L de proteína purificada, e as placas foram incubadas *overnight* à 29 °C. A determinação da atividade das despolimerases foi observada pela presença ou não de pontos claros ou halos (zonas opacas) no tapete bacteriano.

3 RESULTADOS

3.1 Análises de Bioinformática

Inicialmente, foram submetidas à anotação funcional 72 sequências genômicas de diferentes fagos (Tabela 1), com o objetivo de identificar proteínas potencialmente envolvidas na degradação de exopolissacarídeos, como as despolimerases. As anotações primárias foram obtidas por meio da ferramenta *Phold* (Bouras, 2024), a qual forneceu predições funcionais para os ORF codificados por cada genoma analisado. As análises seguiram o seguinte pipeline (Figura 1).

Figura 1 - Painel representativo compreendendo as etapas das análises de bioinformática.



Fonte: Dados da pesquisa

A partir dessas anotações, foram selecionadas sequências proteicas candidatas para análises em ferramentas especializadas em predição de despolimerases, como *PhageDPO*, *DePP*. A ferramenta *DepoScope* foi testada, porém não retornou resultados relevantes. Considerou-se como *hits* significativos aqueles anotados como *tail proteins* ou *tail fiber proteins* e com valores de similaridade superiores a 50% (Tabelas 3 – 4).

Tabela 3 - Representação das proteínas identificadas no genoma do fago XacP77 após análise das ferramentas DePP e PhageDPO.

ORF	Anotação	DePP (%)	PhageDPO (%)
28	<i>tail protein</i>	75	1
27	<i>tail protein</i>	75	4
34	<i>tail fiber protein</i>	64	3
35	<i>tail fiber protein</i>	54	2

Fonte: Dados da pesquisa

Tabela 4 - Representação das proteínas identificadas no genoma do fago P27 após análise das ferramentas DePP e PhageDPO.

ORF	Anotação	DePP (%)	PhageDPO (%)
10	<i>tail protein</i>	74	12
9	<i>tail protein</i>	73	62
2	<i>tail fiber protein</i>	65	39
3	<i>tail fiber protein</i>	56	50

Fonte: Dados da pesquisa

As proteínas selecionadas foram então submetidas a validação por homologia estrutural com a ferramenta HHPred (Zimmermann *et al.*, 2018), para verificação de correspondência com proteínas previamente caracterizadas. Em seguida, foi realizada a modelagem tridimensional das estruturas selecionadas utilizando o AlphaFold 2 (Jumper *et al.*, 2021; Mirdita *et al.*, 2022), com o objetivo de avaliar a conformação estrutural e a confiabilidade preditiva.

Por fim, as estruturas modeladas foram analisadas com a ferramenta Foldseek (Van Kempen *et al.*, 2024), a fim de identificar a possível presença de regiões relacionadas à ligação à açúcares, característica funcional presente em

muitas despolimerases (Pires *et al.*, 2016). A partir dessas análises, foi possível selecionar duas proteínas candidatas com alto potencial para atividade de despolimerase: P27_002 (fago P27) e XacP77_034

Em seguida, as duas sequências foram comparadas por alinhamento múltiplo utilizando Clustal Omega (Madeira *et al.*, 2022) (Figura 2), apresentando uma identidade de sequência de 92,48%. Com o auxílio da ferramenta HHPred, encontramos os seus possíveis domínios de ligação a açúcares (P27_002: aminoácidos 87-239; XacP77_034: aminoácidos 34-239).

Figura 2 - Alinhamento múltiplo das proteínas P27_002 e XacP77_034 realizado com a ferramenta Clustal Omega, com identidade de sequência de 92,48%. Domínios putativos de ligação e degradação de açúcares, preditos por HHPred, estão destacados em cinza. Diferenças entre os aminoácidos das sequências estão marcadas em vermelho.

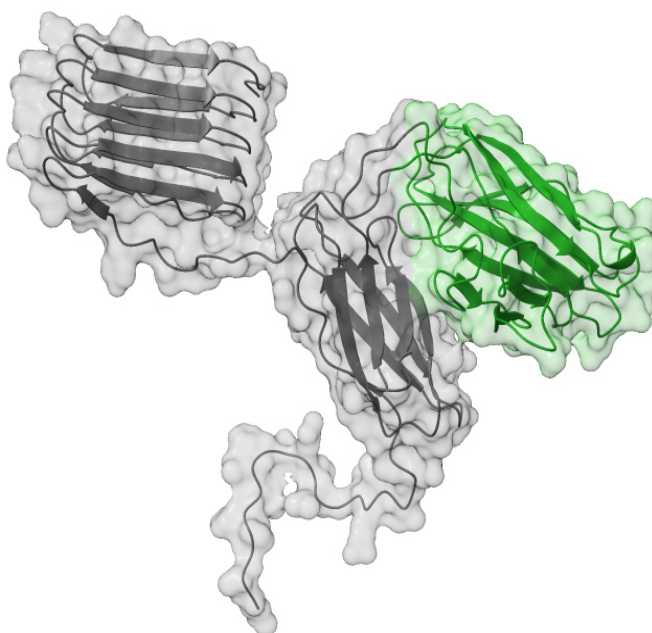
P27_002	MTSPHILFTRPLGLRFVGPPLQVVGQLDGYTAGAEAYEGRLQIINNIGDCTVQWVSGTMP	60
XacP77_034	MTSPHILFTRPLGLRFVGPPLQVVGQLDGYTAGAEAYEGRLQIINNIGDCTVQWVSGTMP	60
P27_002	PGTSVRVDNSTHEVVVAWPEYEPDPNTTVVPNGSFEAGDVNWSKGAGWLIGQSGGFDAY	120
XacP77_034	PGTSVRVDNSTHEVVVAWPEYEPDPNTTVVPNGSFEAGDVNWSKGAGWLIGQSGGFDAY	120
P27_002	DGSWSAAYTGRGTSSLEGTRVPVVVGTSITGSIQVQQGGSSKGNVAGVMLLWYDAAGNL	180
XacP77_034	DGSWSAAYTGRGTSSLEGTRVPVVVGTSITGSIQVQQGGSSKGNVAGVMLLWYDAAGNL	180
P27_002	IDYDVGNIVRSGSNGAWHISTVTASAPAGSFYVSLGVTANRKRQNRRLWVDQGRWNHSYA	240
XacP77_034	IDYDVGNIVRSGSNGAWHISTVTAAAPAGSFYVALGVTANRKRQNRRLWVDQGRWNHSYA	240
P27_002	SGANSDDDYCLVLLVRDAAGREALWEGCVGERAVFLTQTYPIDAIEELDSSFIVDGGFL	300
XacP77_034	SGANSDDDYCLVLLVRDAAGREALWEGCVGERAVFLTSQTYPIDAVEEVDSSFIVDGGFL	300
P27_002	RVPPVLSLTEGADASAVIISGELRTPLLSYVVAPEGVDADAVALSGELRVVQRFYTMVPE	360
XacP77_034	RIPPVFALDEGTDASAVIISGELRTPLLGYSIPPEGIDADMVILSGELRVVQRFYTMVPE	360
P27_002	AVDAALAIQSGTLAQVLITTTMTPEGVDPSITLLSGTLA	399
XacP77_034	AVDAGMLLPSTLAQVLITTTMTPEGVDSSLTLLSGTLA	399

Fonte: Dados da pesquisa

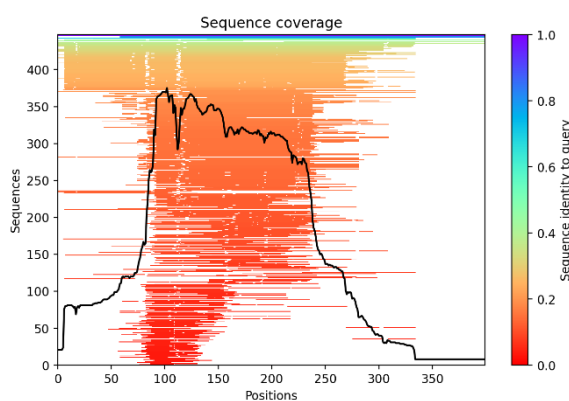
Os modelos estruturais das proteínas P27_002 (Figura 3) e XacP77_034 (Figura 4), gerados por meio do AlphaFold, foram posteriormente analisados utilizando o ChimeraX (Meng *et al.*, 2023), permitindo a observação de uma alta semelhança na organização terciária entre ambas. Em cada uma delas, foi possível identificar a presença de três estruturas β -helicoidais (β -helix), sendo que o domínio de quebra de açúcar está localizado principalmente na β -helix terminal em ambos os casos. Na proteína XacP77_034, esse domínio apresenta uma extensão adicional, abrangendo também parte da β -helix intermediária.

Figura 3 - Modelagem da proteína P27_002 construída com o AlphaFold. A) Estrutura tridimensional da proteína P27_002, visualizada no ChimeraX. A proteína apresenta três estruturas β -helix, com o domínio de quebra de açúcar destacado em verde (aminoácidos: 87-239). B) *Predicted Align Error* (PAE). C) Cobertura da sequência. D) Valor previsto de IDDT em cada posição

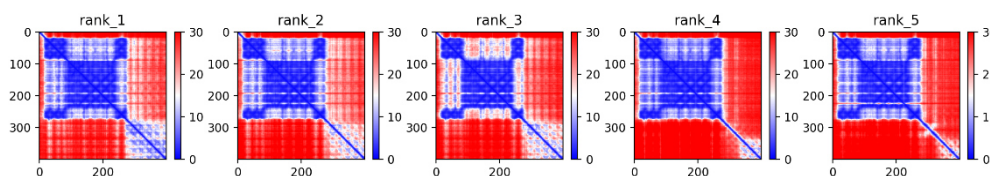
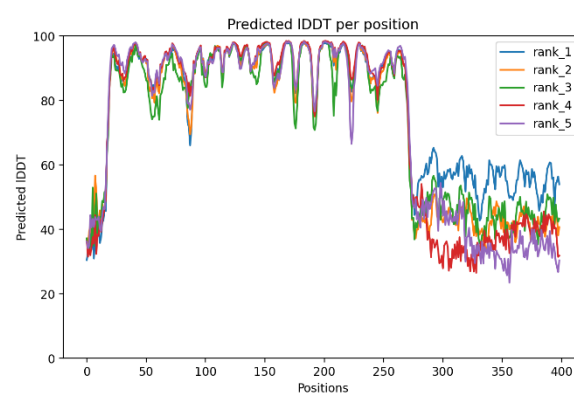
A)



B)



C)

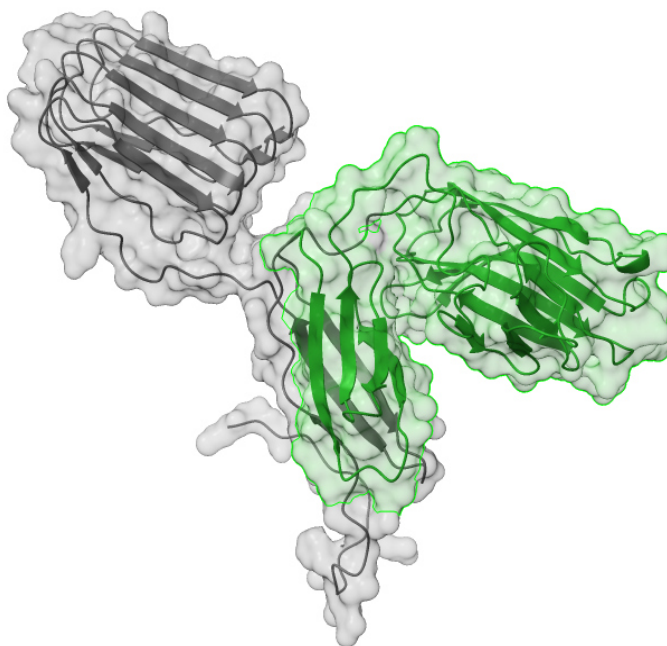


D)

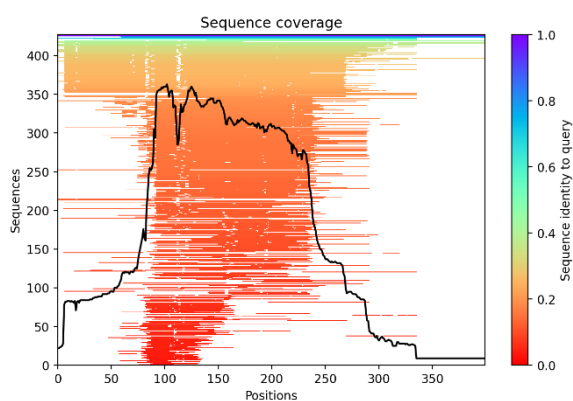
Fonte: Dados da pesquisa

Figura 4 - Modelagem da proteína XacP77_034 construída com o AlphaFold. A) Estrutura tridimensional da proteína XacP77_034, visualizada no ChimeraX. A proteína apresenta três estruturas β -helix, com o domínio de quebra de açúcar destacado em verde (aminoácidos: 34-239). B) *Predicted Align Error* (PAE). C) Cobertura da sequência. D) Valor previsto de IDDT em cada posição.

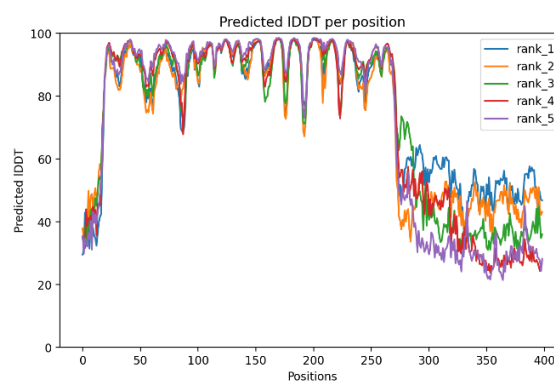
A)



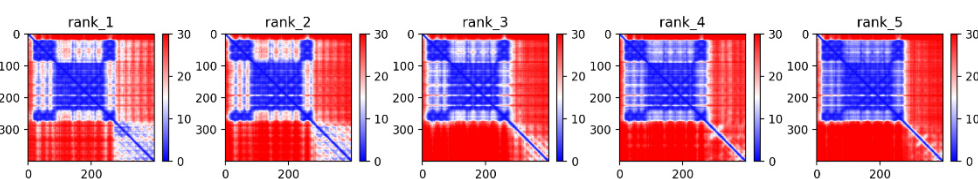
B)



C)



D)



Fonte: Dados da pesquisa

3.2 Clonagem e Expressão

Objetivando a análise de eficácia das ferramentas na identificação de proteínas com atividades de despolimerase, a proteína XacP77_034 foi escolhida como alvo deste trabalho. Assim, os fragmentos correspondentes ao gene completo (XP77G) e ao seu domínio de degradação de açúcares (XP77SD) foram amplificados por PCR a partir do DNA molde obtido do fago XacP77 purificado. Os fragmentos amplificados foram posteriormente utilizados em diferentes estratégias de clonagem e expressão. A Tabela (5) apresenta um resumo das construções gênicas utilizadas no estudo (Anexos A; B):

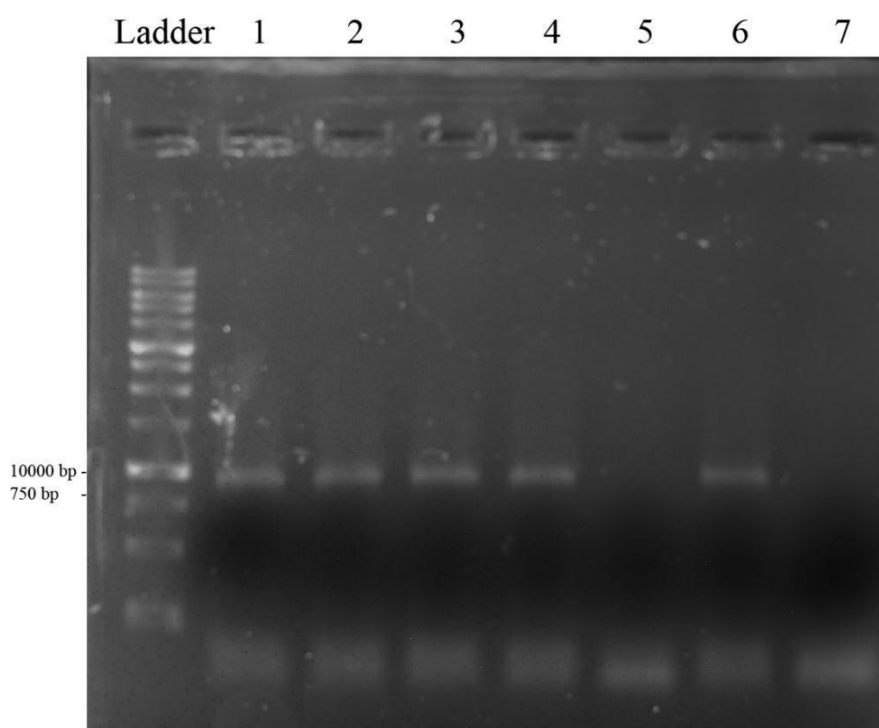
Tabela 5 - Construções gênicas empregadas.

Nome da Construção	Inserto	Vetor de Expressão	Linhagem de Expressão	Principal Resultado da Expressão
pET28-XP77SD	Domínio putativo de degradação de açúcares (XP77SD)	pET-28a(+)	<i>E. coli</i> BL21 (DE3)	Proteína recombinante expressa na fração insolúvel
pTSL-XP77G	Gene completo da proteína XacP77_034 (XP77G)	pTSL (com fusão da chaperona <i>SlyD</i>)	<i>E. coli</i> BL21 (DE3)	Proteína recombinante expressa na fração solúvel
pTSL (Controle)	Nenhum (vetor vazio)	pTSL	<i>E. coli</i> BL21 (DE3)	Utilizado como controle para confirmar a expressão da chaperona <i>SlyD</i> isoladamente

Fonte: Dados da pesquisa

Inicialmente, realizamos a clonagem dos fragmentos no vetor pET-28a(+), obtendo-se cinco clones referentes ao domínio de degradação de açúcares, os quais foram transformados em *E. coli* DH10B (Figura 5). Destes clones, três foram selecionados para sequenciamento, que confirmou a inserção correta do fragmento no vetor.

Figura 5 - Análise por eletroforese em gel de agarose 0,7% da reação de PCR realizada com colônias de *E. coli* DH10B transformadas contendo a construção pET28-XP77SD (936 bp). A corrida foi conduzida por 1 hora a 75 V. Nos poços 1 a 6 foram carregadas amostras de seis colônias distintas obtidas após transformação. O poço 7 corresponde ao controle branco (reação sem inserto). A marcação de peso molecular utilizada foi 1kb ladder (GoldBio). Observa-se a presença de bandas específicas com aproximadamente 936 bp nos poços 1, 2, 3, 4 e 6, indicando clones positivos para o inserto de interesse. O controle branco (poço 7) não apresenta amplificação.

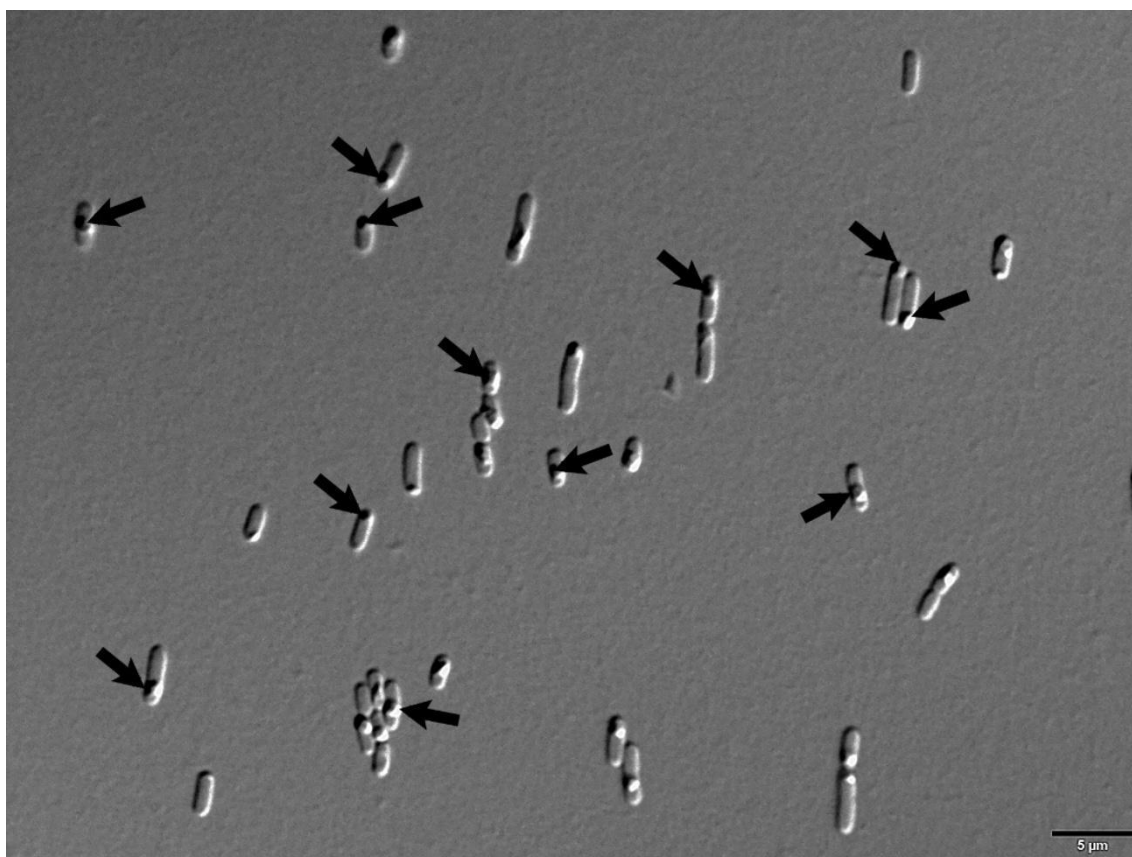


Fonte: Dados da pesquisa

O plasmídeo de expressão pET28-XP77SD foi introduzido na linhagem de *E. coli* BL21, para a produção da proteína. Após indução com 1mM de IPTG por 4 horas a 30 °C, a análise por SDS-PAGE confirmou a presença de uma banda com peso molecular esperado (36,31 kDA). Entretanto, a proteína acumulou-se exclusivamente na fração insolúvel do lisado celular (Figura 7 - poço 4), o que poderia indicar a formação de corpúsculos de inclusão.

Para corroborar essa observação, as células induzidas foram analisadas por microscopia de contraste de interferência diferencial (DIC), que evidenciou a presença de estruturas refringentes no citoplasma, compatíveis com agregados proteicos insolúveis (Figura 6)

Figura 6 - Microscopia óptica por contraste de interferência diferencial (DIC) dos clones de *E. coli* BL21 contendo a construção pET28-XP77SD após 4 h de indução. Observa-se a presença de corpúsculos de inclusão no citoplasma bacteriano (indicados pelas setas pretas), evidenciando o acúmulo de proteína recombinante em forma insolúvel. Aumento: 100x; escala: 5 μ m.

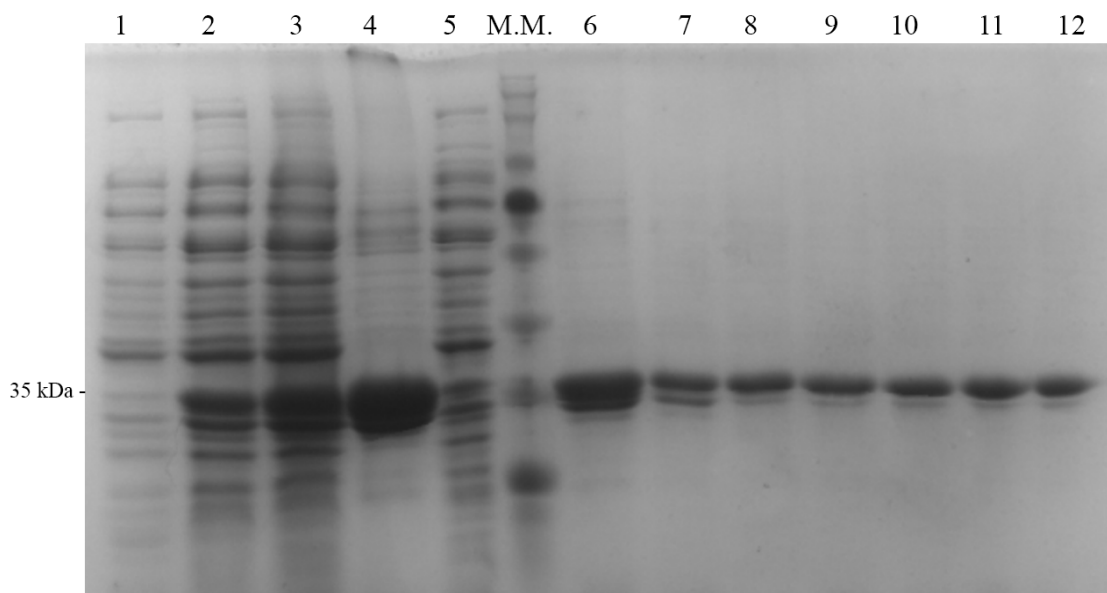


Fonte: Dados da pesquisa

Para recuperar a proteína na sua forma solúvel, foi realizado um protocolo de *solubilização*. Inicialmente a fração insolúvel foi solubilizada em tampão contendo o agente desnaturante uréia 8 M. A remoção gradual da ureia foi então realizada por diálise sequencial. Ao final do processo, a proteína tornou-se detectável na fração solúvel (Figura 7, poço 7), indicando que o protocolo foi capaz de restaurar, ao menos parcialmente, sua solubilidade.

Figura 7 - Análise por SDS-PAGE da expressão e *refolding* do domínio de degradação de açúcar da proteína XacP77_034 após 4 horas de indução com IPTG.

A proteína recombinante foi expressa em *E. coli* BL21 e submetida a extração, solubilização e diálise em diferentes concentrações de ureia. O gel foi corado com Azul de Coomassie. Os poços correspondem às seguintes condições: poço 1 – células não induzidas; poço 2 – 2 horas pós-indução; poço 3 – 4 horas pós-indução; poço 4 – lisado celular (pellet – fração insolúvel); poço 5 – lisado clarificado (sobrenadante); M.M. – marcador de peso molecular; poço 6 – pellet resolubilizado em 8 M de uréia; poço 7 – proteína solubilizada em 8 M de uréia; poços 8 a 12 – proteína submetida à *refolding* em 6 M, 4 M, 2 M, 1 M e 0 M de uréia, respectivamente.



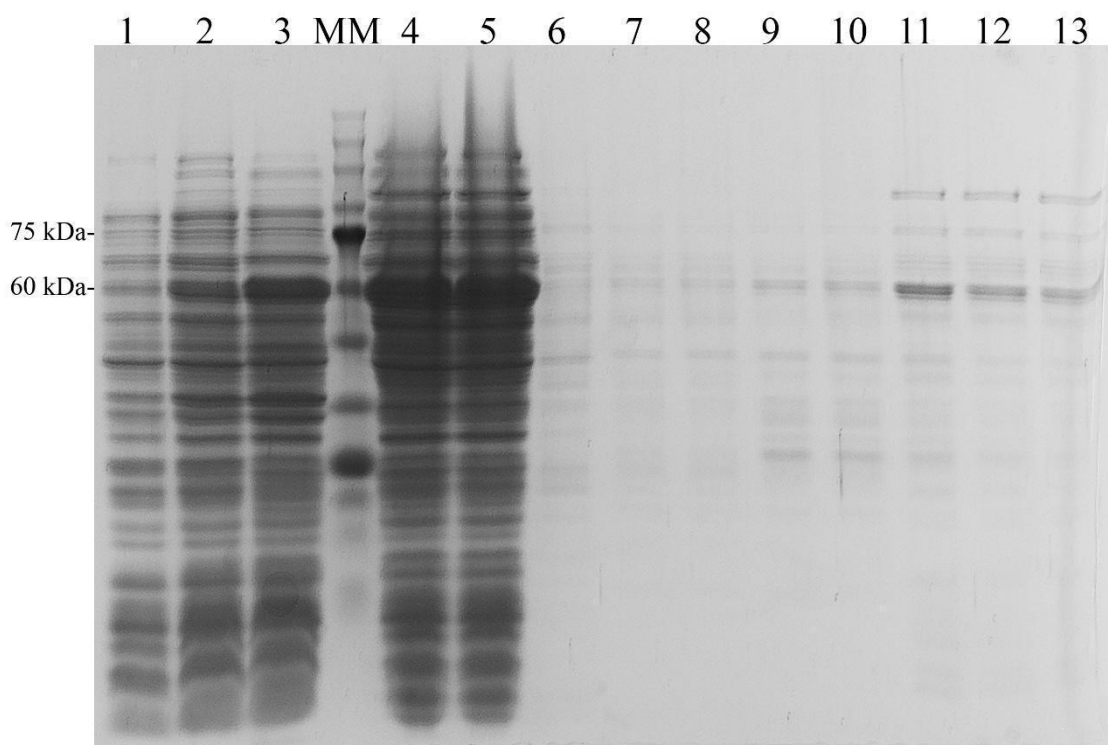
Fonte: Dados da pesquisa

Afim de verificar a atividade de despolimerase do domínio, foram realizados ensaios enzimáticos preliminares, nos quais a proteína foi testada quanto a sua capacidade de degradar os EPS de *X. citri* FDC2, mas não foi detectada atividade sob as condições testadas, caracterizando a proteína como inativa em termos de degradação de açúcares nesta etapa.

A expressão do domínio XP77SD no sistema pET28a(+) resultou exclusivamente na produção de corpúsculos de inclusão insolúveis. Diante disso, a estratégia foi modificada visando aumentar a solubilidade da proteína. Optou-se pela clonagem do gene completo (XP77G) no vetor pTSL, que promove a fusão da proteína alvo com a chaperona *SlyD*, uma abordagem projetada para auxiliar no enovelamento correto e prevenir a agregação.

A expressão da proteína de fusão *SlyD*-XP77G foi induzida com 1 mM de IPTG por 24 horas a 30°C. A análise do lisado celular por SDS-PAGE confirmou que a estratégia foi bem-sucedida, com a proteína recombinante localizada predominantemente na fração solúvel (Figura 8, poço 5). Subsequentemente, a proteína foi submetida à purificação por cromatografia de afinidade, utilizando um gradiente de imidazol. A eluição mostrou-se mais eficiente na concentração de 200 mM, resultando em uma fração de maior pureza e concentração (Figura 8, poços 11-13).

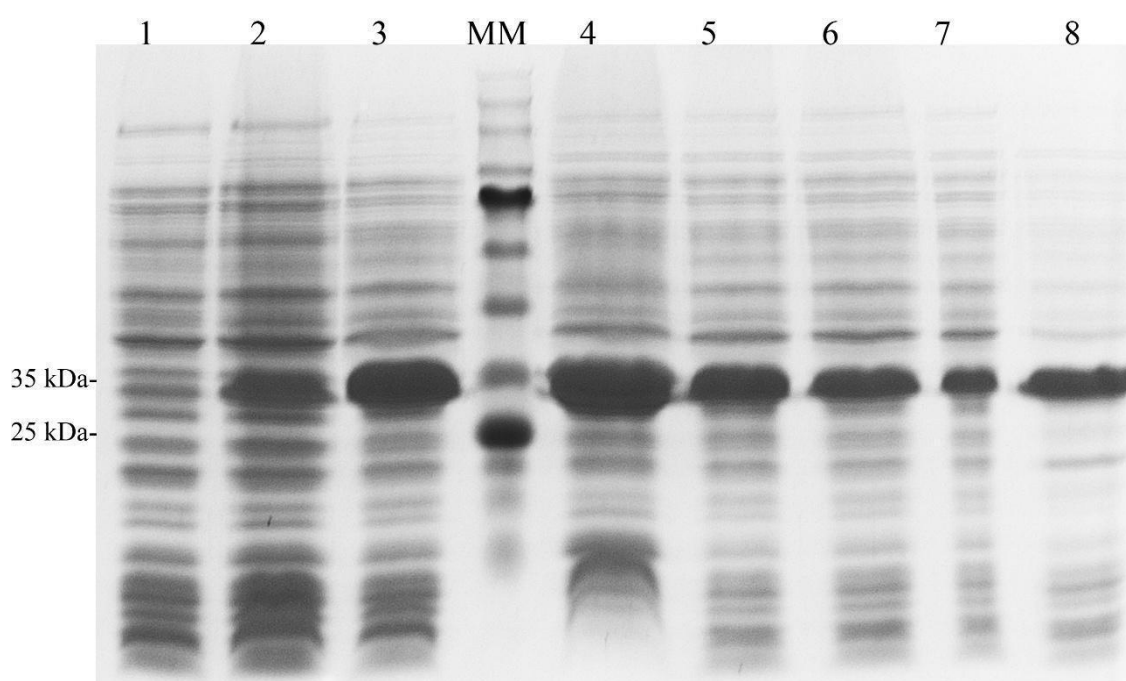
Figura 8 - Análise por SDS-PAGE da expressão e purificação da proteína XacP77_034 após 24 horas de indução com IPTG. O vetor foi expresso em *E. coli* BL21 e, após a lise do *pellet* celular, a amostra foi submetida à purificação utilizando de 20 a 200 mM de imidazol. O gel foi corado com Coomassie Blue. Os poços correspondem aos seguintes: poço 1 – células não induzidas; poço 2 – 4 horas pós-indução; poço 3 – 24 horas pós-indução; M.M. – marcador de peso molecular (kDa); poço 4 – lisado celular (pellet – fração insolúvel); poço 5 – lisado clarificado (sobrenadante); poço 6 – *flowthrough* ; poço 7 – eluição com 20 mM de imidazol; poço 8 – eluição com 40 mM de imidazol; poços 9 e 10 – eluições com 80 mM de imidazol; poços 11, 12 e 13 - eluição com 200 mM de imidazol.



Fonte: Dados da pesquisa

Para validar o sistema de expressão e servir como controle, o vetor pTSL vazio foi induzido sob as mesmas condições. A análise confirmou a expressão da proteína *SlyD* isoladamente, que foi detectada tanto na fração solúvel quanto na insolúvel (Figura 9, poços 4 e 5). Notavelmente, a purificação da *SlyD* ocorreu com uma baixa concentração de imidazol (20 mM, Figura 9, poço 7), confirmando a funcionalidade da cauda de histidina do vetor e fornecendo um perfil de eluição distinto da proteína de fusão completa.

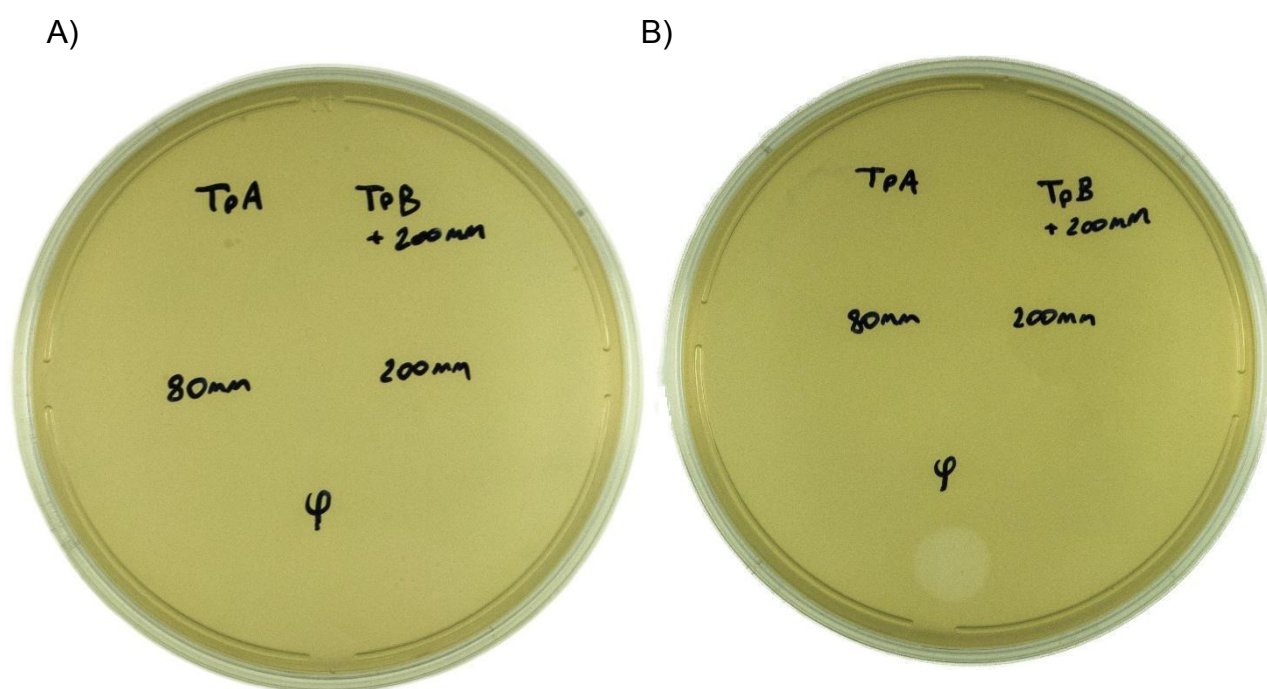
Figura 9 - Análise por SDS-PAGE da expressão e purificação do plasmídeo pTSL vazio após 24 horas de indução com IPTG. O vetor foi expresso em *E. coli* BL21 e, após a lise do *pellet* celular, a amostra foi submetida à purificação utilizando 20 mM de imidazol. O gel foi corado com Coomassie Blue. Os poços correspondem aos seguintes: poço 1 – células não induzidas; poço 2 – 4 horas pós-indução; poço 3 – 24 horas pós-indução; M.M. – marcador de peso molecular (kDa); poço 4 – lisado celular (*pellet* – fração insolúvel); poço 5 – lisado clarificado (*sobrenadante*); poço 6 – *flowthrough* ; poço 7 – eluição com 20 mM de imidazol; poço 8 – resina utilizada na purificação.



Fonte: Dados da pesquisa

Por fim, a fim de confirmar se a proteína XP77G havia atividade de despolimerase, realizamos testes enzimáticos contra as linhagens de *X. citri* 306 e FDC2, as quais são, respectivamente, resistente e sensível ao fago XacP77 (Figura 10). Notou-se que, em ambas as linhagens, a proteína não mostrou atividade na degradação de EPS, enquanto em FDC2 (B) o fago formou uma placa de lise, o que leva à desqualificação da proteína encontrada como uma despolimerase nas condições testadas.

Figura 10 - Testes enzimáticos para a proteína XP77G contra as linhagens de *X. citri* 306 (A) e FDC02 (B). Para os testes, 100 μ L de inóculo das linhagens foram plaqueados em meio NYG-ágar e recobertos com uma fina camada de ágar-soft. Em seguida, aplicaram-se 5 μ L da proteína purificada, eluída com 80 mM e 20 mM de imidazol. Como controles, foram utilizados 5 μ L dos tampões A (sem imidazol) e B (com 200 mM de imidazol), além de 2 μ L do fago XacP77 (φ). Observou-se formação de placa de lise apenas na linhagem FDC02, sensível ao fago, correspondente ao controle positivo. Não foram detectados halos de atividade enzimática (despolimerase) para XP77G em nenhuma das linhagens. Na linhagem 306, resistente ao fago, não houve formação de lise ou halos enzimáticos.



Fonte: Dados da pesquisa

4 DISCUSSÃO

A busca por despolimerases associadas a fagos de *Xanthomonas citri* se coloca como um potencial alternativa biotecnológica para o controle da fitopatologia, uma vez que o exopolissacarídeo produzido por *X. citri* (goma-xantana) possui um papel essencial na formação de biofilmes, virulência e proteção contra estresses ambientais (Rudolph, 1993; Rigano *et al.*, 2007). Com isso, a prospecção de enzimas virais, proveniente de um inimigo natural de *X. citri*, que sejam capazes de degradar esses polímeros podem não apenas auxiliar no controle da bactéria, mas também servir como ferramentas biotecnológicas promissoras (Pires *et al.*, 2016). Nesse contexto, o atual trabalho visou estabelecer um *pipeline* bioinformático que integra a predição *in silico* com a validação experimental de despolimerases candidatas.

A triagem computacional partiu da anotação funcional dos genomas fágicos identificou proteínas de cauda (*tail proteins*) e fibras de cauda (*tail fiber proteins*) como alvos preferenciais, regiões do fago onde as despolimerases tipicamente se localizam (Liu *et al.*, 2025). No entanto, a aplicação das ferramentas de predição baseadas em *machine-learning* (PhageDPO, DePP e DepoScope) revelou uma notável discrepância nos resultados. A ferramenta DePP foi mais permissiva, enquanto o PhageDPO foi mais restritivo, destacando-se apenas algumas ORFs, e o DepoScope não identificou candidatos.

Essas diferenças refletem a forma como cada ferramenta foi desenvolvida, tanto em termos de algoritmos quanto dos bancos de dados utilizados em seus treinamentos. O DepoScope combina modelos de linguagem de proteínas (ESM-2) com redes neurais convolucionais, permitindo não apenas identificar depolimerases, mas também delimitar seus domínios enzimáticos, a partir de sequências da base INPHARED, domínios catalíticos do InterPro e dobras estruturais descritas no CAZy (Concha-Eloko *et al.*, 2024). O DePolymerase Predictor (DePP), por sua vez, utiliza um *random forest* treinado com um banco reduzido composto por 50 depolimerases experimentalmente validadas e 50 sequências negativas oriundas de metagenomas de solo, baseando-se em 424 descritores físico-químicos e de composição de aminoácidos (Magill; Skvortsov, 2023). Já o PhageDPO emprega algoritmos de Support Vector Machines (SVM) e redes neurais artificiais (ANN), com treinamento realizado em 1437 sequências positivas contendo domínios associados a

depolimerases no NCBI e até 5748 sequências negativas, a partir das quais foram extraídas 578 características físico-químicas e estruturais preditas (Vieira *et al.*, 2023).

A assimetria observada entre os preditores pode ser explicada pelo fato de que proteínas de ligação a receptores (RBPs) frequentemente compartilham o mesmo arcabouço estrutural em β -hélice (Van Raaij *et al.*, 2001; Spinelli *et al.*, 2006) característico das despolimerases, o que pode confundir os preditores (Pires *et al.*, 2016). Por essa razão, a validação por homologia estrutural foi um passo de conferência adicionado na análise.

Os modelos tridimensionais gerados para as candidatas P27_002 e XacP77_034 confirmaram a presença de estruturas β -helicoidais, um traço típico de despolimerases (Squeglia *et al.*, 2020; Huang *et al.*, 2024). Essa conformação estrutural está associada à formação de um sítio de ligação específico para polissacarídeos de cadeia longa, o que possibilita a interação e a degradação eficiente de carboidratos. Tais achados reforçam tanto o papel funcional das proteínas identificadas quanto à relevância da análise estrutural na validação de predições (Knecht; Veljkovic; Fieseler, 2019). Além disso, a alta identidade de sequência observada entre as duas proteínas (92,48%) sugere forte conservação estrutural e funcional, e a organização em *tailspike* em ambas as candidatas fortalece, ainda, a escolha de XacP77_034 para a validação experimental.

A etapa experimental iniciou-se com a expressão do domínio catalítico putativo (XP77SD) em um sistema de alta expressão como o pET-28a(+). Esperávamos observar com essa escolha de utilizar somente o domínio (i) identificar se essa região da proteína possuiria um sítio catalítico para a degradação de açúcar e (ii) aumentar a possibilidade de obter uma proteína solúvel. Essa abordagem, embora tenha resultado em alta produção, ainda assim levou à formação massiva de corpúsculos de inclusão insolúveis. Tal achado não é inesperado, pois a arquitetura de β -hélice e alta taxa de tradução frequentemente levam à agregação de proteínas fágicas em *E. coli* (Rosano; Ceccarelli, 2014). Ainda, muitas vezes proteínas fágicas requerem dobramento mediado por cofatores específicos ou chaperonas ausentes em *E. coli* (Baneyx; Mujacic, 2004). Embora a recuperação por *refolding* tenha sido tentada como estratégia, a proteína permaneceu inativa, provavelmente devido à falha em reconstituir a conformação trimérica nativa necessária para atividade catalítica (Singh; Panda, 2005).

Para contornar essa limitação, adotou-se uma nova estratégia baseada em duas alterações fundamentais. Optamos por utilizar o gene completo (XP77G), partindo da hipótese de que as regiões N e C-terminais seriam cruciais para a estabilidade estrutural e possível conformação trimérica necessária para atividade. Ainda, o vetor pET-28a(+) foi substituído pelo pTSL, oferecendo uma fusão com a chaperona *SlyD* de *E. coli* (Taylor *et al.*, 2016). A função desta chaperona é assistir ativamente o enovelamento da proteína recombinante, prevenindo sua agregação, em estratégia semelhante à utilizada por Pires *et al.*, (2016). Essa abordagem combinada foi bem-sucedida, resultando na obtenção da proteína solúvel e purificada. Apesar do sucesso na obtenção de uma proteína solúvel, os testes funcionais em placa não foram suficientes para demonstrar atividade de degradação de exopolissacarídeos para XacP77_034. Outros autores já descreveram atividades em testes semelhantes a partir da expressão de despolimerases (Cornelissen *et al.*, 2011; Harper *et al.*, 2014; Pires *et al.*, 2016). Por conta disso, não seguimos para testes específicos de degradação de biofilme e atividade em EPS purificado.

A análise conjunta dos resultados revela que, por um lado, a metodologia provou-se robusta, uma vez que nosso *pipeline* bioinformático priorizou candidatas estruturalmente adequadas e a estratégia de expressão contornou os desafios de solubilidade. No entanto, o resultado funcional que obtivemos foi negativo e ressalta uma limitação crítica das ferramentas de predição atuais: a confiabilidade de todo *pipeline* de predição está condicionada à representatividade dos bancos de dados empregados. Apesar das ferramentas utilizadas se propõem a construir seu treinamento baseando-se em proteínas com atividade de despolimerase, validadas experimentalmente, ainda assim estes são amplamente povoados por proteínas provenientes de bacteriófagos que têm bactérias Gram-positivas clínicas como hospedeiras. Sendo *X. citri* uma bactéria Gram-negativa, cujos exopolissacarídeos apresentam composição e estrutura distintas (Latham, 2021), a predição funcional enfrenta uma limitação intrínseca, o que pode explicar eventuais falhas na identificação de enzimas verdadeiramente ativas. Ainda, outras estratégias de experimentação podem ser implementadas para identificar a necessidade de cofatores ou condições em que possam ser identificadas atividade catalítica – o que não foi possível explorar no período deste trabalho devido aos desafios encontrados.

5 CONCLUSÃO

A principal contribuição deste trabalho reside na validação de um *pipeline* integrado para a prospecção de despolimerases, ao mesmo tempo que deixa claro os desafios inerentes às ferramentas de predição atuais. Passos futuros cruciais para que outras investigações possam traduzir em maior sucesso devem envolver a expansão dos bancos de dados com validações funcionais em uma diversidade maior de bacteriófagos.

REFERÊNCIAS

- AHMAD, A. A.; OGAWA, M.; KAWASAKI, T.; FUJIE, M. et al., Characterization of bacteriophages Cp1 and Cp2, the strain-typing agents for *Xanthomonas axonopodis* pv. citri. **Appl Environ Microbiol**, 80, n. 1, p. 77-85, Jan 2014.
- ALTSHUL, Stephen F. Basic local alignment search tool. **J. Mol. Biol.**, v. 215, p. 403-410, 1990.
- BANEYX, F.; MUJACIC, M. Recombinant protein folding and misfolding in *Escherichia coli*. **Nature Biotechnology**, v. 22, n. 11, p. 1399-1408, 2004. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/15065692/>. Acesso em: 28 ago. 2025.
- BEHLAU, F. et al. Copper resistance genes from different xanthomonads and citrus epiphytic bacteria confer resistance to *Xanthomonas citri* subsp. *citri*. **European Journal of Plant Pathology**, v. 133, n. 4, p. 949-963, 2012.
- BOURAS, Georgios. **Phold**. 2024. Disponível em: <https://github.com/gbouras13/phold>. Acesso em: 14 maio 2025.
- CONCHA-ELOKO, Robby *et al.*, DepoScope: accurate phage depolymerase annotation and domain delineation using large language models. **bioRxiv**, p. 2024.01. 15.575807, 2024.
- CORNELISSEN, Anneleen *et al.*, The T7-related *Pseudomonas putida* phage ϕ 15 displays virion-associated biofilm degradation properties. **PloS one**, v. 6, n. 4, p. e18597, 2011.
- CUERVO, Ana; CARRASCOSA, José L. Bacteriophages: structure. **eLS**, 2012.
- DA SILVA, AC Rasera *et al.*, Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. **Nature**, v. 417, n. 6887, p. 459-463, 2002.
- D'HERELLE, F. The bacteriophage. *Atomes*, v. 3, n. 33, p. 399-403, 1948.
- DONLAN, Rodney M. Biofilms: microbial life on surfaces. **Emerging infectious diseases**, v. 8, n. 9, p. 881, 2002.
- DUNGER, Germán et al. Xanthan is not essential for pathogenicity in citrus canker but contributes to *Xanthomonas* epiphytic survival. **Archives of microbiology**, v. 188, p. 127-135, 2007.
- ERENCE, Christopher M. et al. Recent advances in the understanding of *Xanthomonas citri* ssp. *citri* pathogenesis and citrus canker disease management. **Molecular plant pathology**, v. 19, n. 6, p. 1302, 2018.
- FINN, Robert D.; CLEMENTS, Jody; EDDY, Sean R. HMMER web server: interactive sequence similarity searching. **Nucleic acids research**, v. 39, n. suppl_2, p. W29-W37, 2011.

GRIGSON, Susanna R. et al. Knowing and Naming: Phage Annotation and Nomenclature for Phage Therapy. **Clinical Infectious Diseases**, v. 77, n. Supplement_5, p. S352-S359, 2023.

GUO, Zhimin; LIU, Mengmeng; ZHANG, Dan. Potential of phage depolymerase for the treatment of bacterial biofilms. **Virulence**, v. 14, n. 1, p. 2273567, 2023.

HARPER, David R. *et al.*, Bacteriophages and biofilms. **Antibiotics**, v. 3, n. 3, p. 270-284, 2014.

HOBBS, Zack; ABEDON, Stephen T. Diversity of phage infection types and associated terminology: the problem with 'Lytic or lysogenic'. **FEMS microbiology letters**, v. 363, n. 7, p. fnw047, 2016.

HUANG, T. *et al.*, Structural and functional basis of bacteriophage K64-ORF41 depolymerase. **International Journal of Biological Macromolecules**, v. 231, p. 130917, 2024.

JUMPER, John *et al.*, Highly accurate protein structure prediction with AlphaFold. **Nature**, v. 596, n. 7873, p. 583-589, 2021.

KELLEY, Lawrence A. et al. The Phyre2 web portal for protein modeling, prediction and analysis. **Nature protocols**, v. 10, n. 6, p. 845-858, 2015.

KNECHT, Leandra E.; VELJKOVIC, Marjan; FIESELER, Lars. Diversity and function of phage encoded depolymerases. **Frontiers in microbiology**, v. 10, p. 2949, 2020.

LATHAM, Kate. Gram-Positive vs. Gram-Negative. **Biology Dictionary**, 19 jan. 2021. Disponível em: <https://biologydictionary.net/gram-positive-vs-gram-negative/>. Acesso em: 15 jul. 2024.

LATKA, Agnieszka et al. Modeling the architecture of depolymerase-containing receptor binding proteins in Klebsiella phages. **Frontiers in microbiology**, v. 10, p. 2649, 2019.

LORENZONI, A. S. G.; DANTAS, G. C.; BERGSMA, T.; FERREIRA, H. et al. *Xanthomonas citri* MinC Oscillates from Pole to Pole to Ensure Proper Cell Division and Shape. **Front Microbiol**, 8, p. 1352, 2017.

LIU, Shenshen; LEI, Tao; TAN, Yujing; HUANG, Xiaoyi; ZHAO, Wenxin; ZOU, Huanhuan; SU, Jianhui; ZENG, Ji; ZENG, Haiyan. Discovery, structural characteristics and evolutionary analyses of functional domains in *Acinetobacter baumannii* phage tail fiber/spike proteins. **BMC Microbiology**, v. 25, n. 1, p. 73, 2025

MADEIRA, Fábio *et al.*, The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. **Nucleic acids research**, v. 52, n. W1, p. W521-W525, 2024.

MAGILL, Damian J.; SKVORTSOV, Timofey A. DePolymerase Predictor (DePP): a machine learning tool for the targeted identification of phage depolymerases. **BMC bioinformatics**, v. 24, n. 1, p. 208, 2023.

MELO, Luís DR et al. Phage therapy efficacy: a review of the last 10 years of preclinical studies. **Critical reviews in microbiology**, v. 46, n. 1, p. 78-99, 2020.

MENG, Elaine C. *et al.*, UCSF ChimeraX: Tools for structure building and analysis. **Protein Science**, v. 32, n. 11, p. e4792, 2023.

MIRDITA, Milot *et al.*, ColabFold: making protein folding accessible to all. **Nature methods**, v. 19, n. 6, p. 679-682, 2022.

MUHAMMAD, Musa Hassan et al. Beyond risk: bacterial biofilms and their regulating approaches. **Frontiers in microbiology**, v. 11, p. 928, 2020.

NAQVI, Syed Atif Hasan et al. Citrus Canker—Distribution, Taxonomy, Epidemiology, Disease Cycle, Pathogen Biology, Detection, and Management: A Critical Review and Future Research Agenda. **Agronomy**, v. 12, n. 5, p. 1075, 2022.

NOBREGA, Franklin L. et al. Targeting mechanisms of tailed bacteriophages. *Nature Reviews Microbiology*, v. 16, n. 12, p. 760-773, 2018.

OLIVEIRA, Hugo *et al.*, Functional analysis and antivirulence properties of a new depolymerase from a myovirus that infects *Acinetobacter baumannii* capsule K45. **Journal of virology**, v. 93, n. 4, p. 10.1128/jvi. 01163-18, 2019.

OLIVEIRA, Hugo; DRULIS-KAWA, Zuzanna; AZEREDO, Joana. Exploiting phage-derived carbohydrate depolymerases for combating infectious diseases. **Trends in Microbiology**, v. 30, n. 8, p. 707-709, 2022.

PICCHI, Simone Cristina et al. Modified monosaccharides content of xanthan gum impairs citrus canker disease by affecting the epiphytic lifestyle of *Xanthomonas citri* subsp. *citri*. **Microorganisms**, v. 9, n. 6, p. 1176, 2021.

PIRES, Diana P. *et al.*, Bacteriophage-encoded depolymerases: their diversity and biotechnological applications. **Applied microbiology and biotechnology**, v. 100, p. 2141-2151, 2016.

RAGUPATHY, R. *et al.*, Core-genome multilocus sequence typing for epidemiological and evolutionary analyses of phytopathogenic *Xanthomonas citri*. **Applied and Environmental Microbiology**, v. 89, n. 5, p. e02101-22, 2023.

RIGANO, L. E. *et al.*, Biofilm formation, epiphytic fitness, and canker development in *Xanthomonas axonopodis* pv. *citri*. **Molecular Plant-Microbe Interactions**, v. 20, n. 6, p. 633-642, 2007

ROSANO, G. L.; CECCARELLI, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. **Microbial Cell Factories**, v. 13, p. 63, 2014.

RUDOLPH, K. Infection of the plant by *Xanthomonas*. In: **Xanthomonas**. Dordrecht: Springer Netherlands, 1993.

SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. **Molecular Cloning: A laboratory manual**. 2nd Edition ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1989.

SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, 30, n. 14, p. 2068-2069, Jul 15 2014.

SINGH, S. M.; PANDA, A. K. Solubilization and refolding of bacterial inclusion body proteins. **Journal of Bioscience and Bioengineering**, v. 99, n. 4, p. 303–310, 2005.

SODING, Johannes; BIEGERT, Andreas; LUPAS, Andrei N. The HHpred interactive server for protein homology detection and structure prediction. **Nucleic acids research**, v. 33, n. suppl_2, p. W244-W248, 2005.

SPINELLI, Silvia *et al.*, Modular structure of the receptor binding proteins of Lactococcus lactis phages: the RBP structure of the temperate phage TP901-1. **Journal of Biological Chemistry**, v. 281, n. 20, p. 14256-14262, 2006.

SQUEGLIA, F. *et al.*, Structural and functional studies of a Klebsiella phage tailspike protein revealing a parallel β -helix domain. **Structure**, v. 28, n. 8, p. 1045-1054.e3, 2020.

STRATHDEE, Steffanie A. *et al.* Phage therapy: From biological mechanisms to future directions. **Cell**, v. 186, n. 1, p. 17-31, 2023.

TAYLOR, Nicholas MI *et al.*, Structure of the T4 baseplate and its function in triggering sheath contraction. **Nature**, v. 533, n. 7603, p. 346-352, 2016.

SUMMERS, William C. The strange history of phage therapy. **Bacteriophage**, v. 2, n. 2, p. 130-133, 2012.

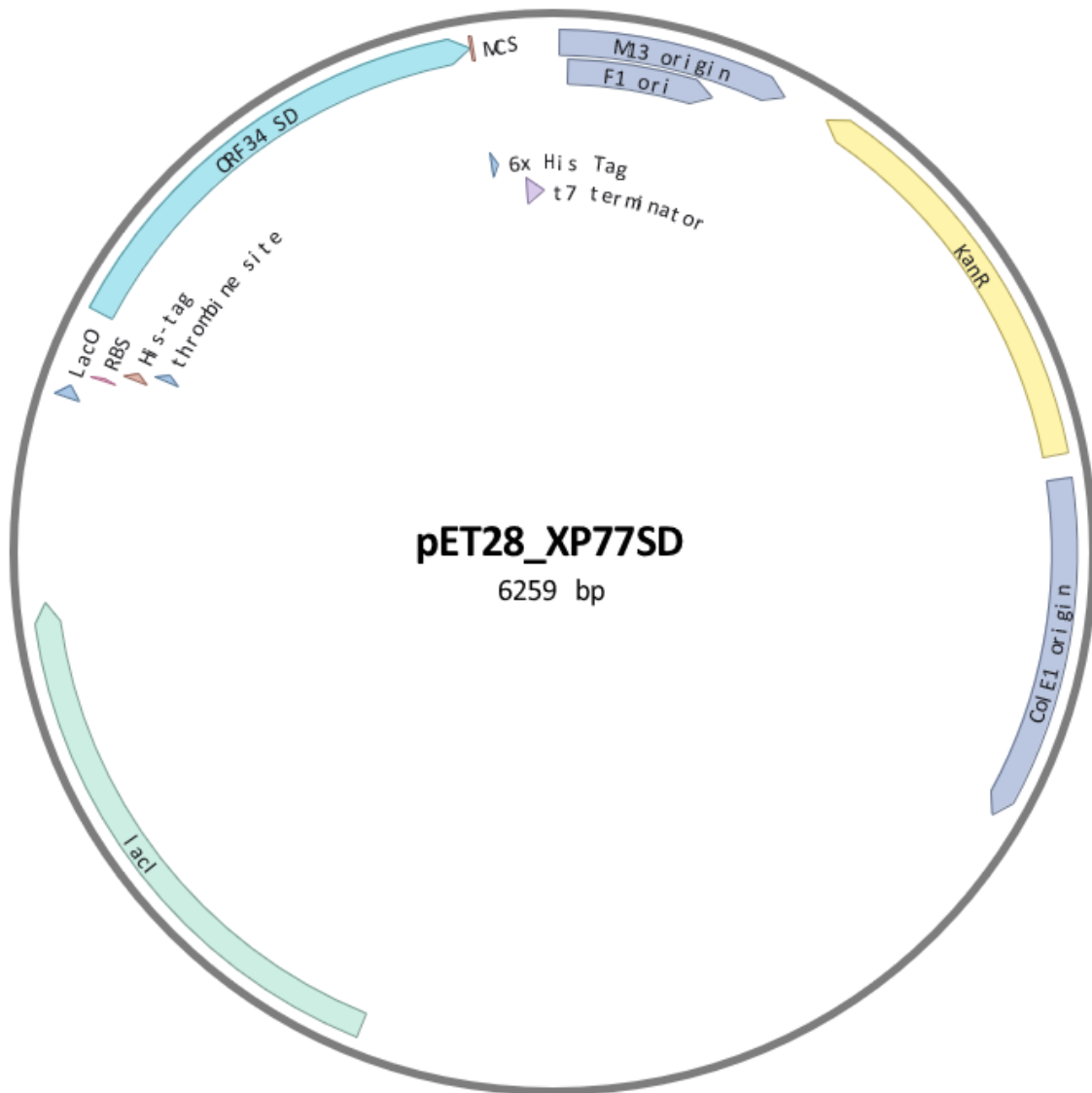
VAN KEMPEN, Michel *et al.*, Fast and accurate protein structure search with Foldseek. **Nature biotechnology**, v. 42, n. 2, p. 243-246, 2024.

VAN RAAIJ, Mark J. *et al.*, Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. **Journal of molecular biology**, v. 314, n. 5, p. 1137-1146, 2001.

VIEIRA, Maria *et al.*, PhageDPO: phage depolymerase finder. **Biorxiv**, p. 2023.02.24.529883, 2023.

YAN, J.; MAO, J.; XIE, J. Bacteriophage polysaccharide depolymerases and biomedical applications. **BioDrugs**, 28, n. 3, p. 265-274, Jun 2014.

ZIMMERMANN, Lukas *et al.*, A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. **Journal of molecular biology**, v. 430, n. 15, p. 2237-2243, 2018.

ANEXO A – CONSTRUÇÃO PET28-XP77SD

ANEXO B – CONSTRUÇÃO PTSL-XP77G