



## Pedido nacional de Invenção, Modelo de Utilidade, Certificado de Adição de Invenção e entrada na fase nacional do PCT

Número do Processo: BR 10 2021 008663 7

### Dados do Depositante (71)

---

Depositante 1 de 1

**Nome ou Razão Social:** FUNDAÇÃO UNIVERSIDADE FEDERAL DE SÃO CARLOS

**Tipo de Pessoa:** Pessoa Jurídica

**CPF/CNPJ:** 45358058000140

**Nacionalidade:** Brasileira

**Qualificação Jurídica:** Instituição de Ensino e Pesquisa

**Endereço:** Rodovia Washington Luiz, Km 235, 13565-905, Monjolinho

**Cidade:** São Carlos

**Estado:** SP

**CEP:** 13565-905

**País:** Brasil

**Telefone:** 16 3351-9040

**Fax:**

**Email:** inovacao@ufscar.br

## Dados do Pedido

---

**Natureza Patente:** 10 - Patente de Invenção (PI)

**Título da Invenção ou Modelo de Utilidade (54):** USO DE COMPOSTOS QUÍMICOS INIBIDORES DA FOSFOMANOSE ISOMERASE PARA CONTROLE DO CANCRO CÍTRICO E FITOPATOLOGIAS ASSOCIADAS AO GÊNERO XANTHOMONAS

**Resumo:** A presente invenção pertence ao campo da agroquímica e química medicinal. O presente invento refere-se ao uso de compostos químicos como insumos agrícolas promissores e inéditos para o combate ao cancro cítrico que ocorre em plantações de citros, causado pela bactéria *Xanthomonas citri* subsp. *citri* (XAC). O presente invento consiste na elaboração de composições que inibam a fosfomanose isomerase (PMI), que está associada com a patogenicidade de XAC. O invento também apresenta a comprovação funcional de que a PMI é essencial para a patogenicidade de XAC, demonstrando que compostos planejados para inibir seletivamente a PMI são eficazes no controle do cancro cítrico e fitopatologias associadas ao gênero *Xanthomonas*.

**Figura a publicar:** 18

## Dados do Inventor (72)

---

### Inventor 1 de 6

**Nome:** MARIA TERESA MARQUES NOVO MANSUR

**CPF:** 07186367836

**Nacionalidade:** Brasileira

**Qualificação Física:** Professor do ensino superior

**Endereço:** Rua Henrique Lupo, 719 – apto 14

**Cidade:** Araraquara

**Estado:** SP

**CEP:** 14802-304

**País:** BRASIL

**Telefone:** (16) 335 19040

**Fax:**

**Email:** inovacao@ufscar.br

### Inventor 2 de 6

**Nome:** ANDRÉ VESSONI ALEXANDRINO

**CPF:** 36943420879

**Nacionalidade:** Brasileira

**Qualificação Física:** Estudante de Pós Graduação

**Endereço:** Rua Luiz Saia, 220, Apto 106

**Cidade:** Itápolis

**Estado:** SP

**CEP:** 13564-010

**País:** BRASIL

**Telefone:** (16) 335 19040

**Fax:**

**Email:** inovacao@ufscar.br

### Inventor 3 de 6

**Nome:** CARLOS HENRIQUE TOMICH DE PAULA DA SILVA

**CPF:** 13789165816

**Nacionalidade:** Brasileira

**Qualificação Física:** Professor do ensino superior

**Endereço:** Rua Campos Salles, 330/94 - Centro

**Cidade:** Ribeirão Preto

**Estado:** SP

**CEP:** 14015-110

**País:** BRASIL

**Telefone:** (16) 335 19040

**Fax:**

**Email:** inovacao@ufscar.br

**Inventor 4 de 6**

**Nome:** MARIANA PEGRUCCI BARCELOS

**CPF:** 23034158882

**Nacionalidade:** Brasileira

**Qualificação Física:** Estudante de Pós Graduação

**Endereço:** Travessa Major Antão, 121

**Cidade:** Batatais

**Estado:** SP

**CEP:** 14307-008

**País:** BRASIL

**Telefone:** (16) 335 19040

**Fax:**

**Email:** inovacao@ufscar.br

**Inventor 5 de 6**

**Nome:** LEONARDO BRUNO FEDERICO

**CPF:** 22247285813

**Nacionalidade:** Brasileira

**Qualificação Física:** Estudante de Pós Graduação

**Endereço:** Rua Alexandre Herculano nº84 apto. 23A

**Cidade:** Santos

**Estado:** SP

**CEP:** 11050-030

**País:** BRASIL

**Telefone:** (16) 335 19040

**Fax:**

**Email:** inovacao@ufscar.br

**Inventor 6 de 6**

**Nome:** FRANKLIN BEHLAU

**CPF:** 27723250835

**Nacionalidade:** Brasileira

**Qualificação Física:** Pesquisador

**Endereço:** Av. Oswaldo Gonçalves de Jesus 387, casa K7 – Buona Vita

**Cidade:** Araraquara

**Estado:** SP

**CEP:** 14805-396

**País:** BRASIL

**Telefone:** (16) 335 19040

**Fax:**

**Email:** inovacao@ufscar.br

## Documentos anexados

---

<b>Tipo Anexo</b>	<b>Nome</b>
Relatório Descritivo	P2021-001_RELATORIO DESCRITIVO_DEFINITIVO_040521.pdf
Reivindicação	P2021- 001_REIVINDICACOES_DEFINITIVO_040521.p df
Desenho	P2021-001_FIGURAS_DEFINITIVO_040521.pdf
Resumo	P2021-001_RESUMO_DEFINITIVO_040521.pdf
GRU	P2021_001_GRU200_150421.pdf
Comprovante de pagamento de GRU 200	P2021_GRU_COMPROVANTE_PAGAMENTO_ 160421.pdf
Declaração de período de graça	P2021-001_PUBLICACAO_200220.pdf
Declaração de período de graça	P2021-001_PUBLICACAO_060520.pdf

## Acesso ao Patrimônio Genético

---

- Declaração Positiva de Acesso - Declaro que o objeto do presente pedido de patente de invenção foi obtido em decorrência de acesso à amostra de componente do Patrimônio Genético Brasileiro, realizado a partir de 30 de junho de 2000, e que foram cumpridas as determinações da Lei 13.123 de 20 de maio de 2015, informando ainda:

**Número da Autorização de Acesso:** AB42B48

## Declaração de Divulgação Anterior Não Prejudicial

---

- Artigo 12 da LPI - Período de Graça.

## Declaração de veracidade

---

- Declaro, sob as penas da lei, que todas as informações acima prestadas são completas e verdadeiras.

Relatório descritivo para a patente de invenção: "USO DE COMPOSTOS QUÍMICOS INIBIDORES DA FOSFOMANOSE ISOMERASE PARA CONTROLE DE CANCRO CÍTRICO E FITOPATOLOGIAS ASSOCIADAS AO GÊNERO XANTHOMONAS"

### **CAMPO DA INVENÇÃO**

[001] O presente invento pertence ao campo da agroquímica e química medicinal. O presente invento refere-se ao uso de compostos químicos como insumos agrícolas promissores e inéditos para o combate ao cancro cítrico que ocorre em plantações de citros, causado pela bactéria *Xanthomonas citri* subsp. *citri* (XAC). O presente invento consiste na elaboração de composições que inibem a fosfomanose isomerase (PMI), que está associada com a patogenicidade de XAC. Este invento também apresenta a comprovação funcional de que a PMI é essencial para a patogenicidade de XAC, demonstrando que compostos planejados para inibir seletivamente a PMI são eficazes no controle do cancro cítrico e fitopatologias associadas ao gênero *Xanthomonas*.

### **DESCRIÇÃO DO ESTADO DA TÉCNICA**

[002] O Brasil é o maior produtor mundial de laranja, com uma produção estimada em mais de 17 milhões de toneladas em 2017 (FAO STAT, 2017), sendo o estado de São Paulo responsável por 78,7% de toda a produção nacional

(IBGE, 2018). Em 2016 as exportações de suco de laranja concentrado somaram US\$ 1,43 bilhão, valor que colocou tal commodity na décima posição no ranking de exportações brasileiras (FAO STAT, 2017).

[003] A produção de citros influencia também o mercado interno brasileiro, uma vez que em 2016 o PIB gerado por todos os elos da cadeia produtiva cítrica foi de US\$ 6,5 bilhões e, além disso, a citricultura gerou 200 mil empregos diretos e indiretos (NEVES; TROMBIN, 2017). Entretanto, o potencial econômico da citricultura é limitado, em parte, pelas doenças que afetam a sua produtividade, como o cancro cítrico, causado principalmente pela bactéria *Xanthomonas citri* subsp. *citri* (XAC) (KOLLER, 1994; LANZA et al., 2018), que é uma doença causadora de prejuízos ao setor citricultor devido à queda na produtividade e qualidade dos frutos (FERENCE et al., 2018).

[004] Desde 1995, o cancro cítrico vem se alastrando de maneira alarmante, principalmente no Brasil e nos EUA, evidenciando a suscetibilidade dos citros frente a esse fitopatógeno (BEHLAU; FONSECA; BELASQUE JUNIOR, 2016). Até 2009, o manejo da doença era realizado por meio de medidas preventivas com a erradicação de plantas sintomáticas e suspeitas, concomitantemente ao uso de inseticidas para controle do minador dos citros, um inseto facilitador da infecção por XAC (BEHLAU; FONSECA; BELASQUE JUNIOR, 2016; SANCHES et al., 2014).

[005] Porém, o abrandamento da legislação a partir de 2013 e a extinção do programa de erradicação de plantas



contaminadas no estado de São Paulo e em outros estados em 2017, mudaram a estratégia de enfrentamento da doença. O foco do controle, que era voltado à prevenção de plantas doentes, passou a ser destinado à produção de frutas sem sintomas a partir da utilização de medidas de manejo como variedades menos suscetíveis, plantio de quebra-ventos ao redor dos pomares, controle do minador dos citros e, principalmente, aplicações de cobre (FERENCE et al., 2018; BEHLAU, 2021). Apesar de eficiente para o controle do cancro cítrico e prevenção de perdas, o uso regular de cobre apresenta desvantagens importantes como acúmulo no solo com efeitos nocivos ao sistema radicular das plantas cítricas, contaminação ambiental e risco de desenvolvimento de linhagens de XAC resistentes ao cobre (FERENCE et al., 2018). Assim, há a necessidade eminente de desenvolvimento de medidas alternativas e mais sustentáveis de controle da doença.

[006] O presente invento apresenta a utilização de compostos químicos para controle de cancro cítrico, com base na inibição da fosfomanose isomerase (PMI) da bactéria XAC. O presente invento apresenta em seu relatório descritivo a comprovação funcional de que a PMI é essencial para a patogenicidade de XAC e a eficácia dos compostos planejados para a inibição dos sintomas da doença em citros.

[007] Os documentos de referência de anterioridades à invenção proposta são listados abaixo.

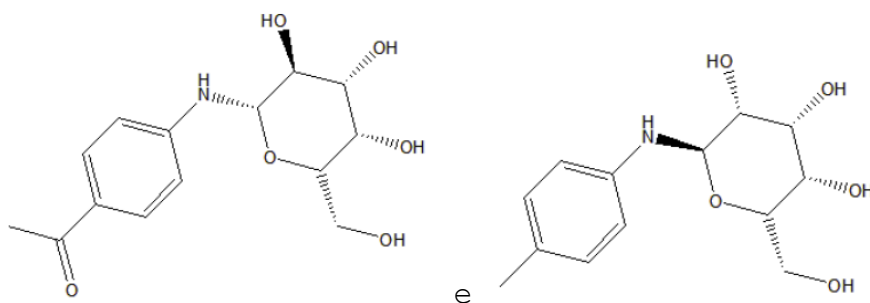
[008] A tese de doutorado "Análise funcional dos genes XANB e XYLA2 de *Xanthomonas citri* subsp. *citri*, agente causal do cancro cítrico" destaca a obtenção e

caracterização de mutante de deleção da PMI e de seu complementante para avaliação e validação da proteína PMI como essencial à patogenicidade de XAC, explicitando a potencialidade de controle do cancro cítrico por meio de compostos que sejam capazes de inibir a PMI de XAC, para os quais é solicitada a proteção.

[009] A dissertação de mestrado "Planejamento computacional e avaliação *in vitro* de novos potenciais inibidores" informa a descoberta *in silico* de compostos com o potencial de inibição da PMI de XAC para finalidade de controle do cancro cítrico, bem como as metodologias para a mesma. Consta também teste de inibição *in vitro* da PMI recombinante de XAC; entretanto, não é abordado na dissertação o teste de patogenicidade *in vivo* com os inibidores.

### **SUMÁRIO DA INVENÇÃO**

[010] Os compostos químicos para controle de cancro cítrico e fitopatologias associadas ao gênero *Xanthomonas* são compostos ou sais aceitáveis do mesmo, que consistem em



, sendo estes exemplos de inibidores da PMI, e são utilizados para fabricação de insumo agrícola para controle de cancro cítrico.

[011] Os compostos químicos para controle de cancro cítrico e fitopatologias associadas ao gênero *Xanthomonas* são utilizados para a inibição seletiva da Fosfomanose isomerase (PMI).

[012] Os compostos químicos para controle de cancro cítrico e fitopatologias associadas ao gênero *Xanthomonas* são utilizados para fabricação de insumos agrícolas para fitopatologias causadas por *Xanthomonas* sp.

[013] A Fosfomanose isomerase (PMI) é uma proteína essencial para a patogenicidade da bactéria XAC e assim pode ser alvo promissor para controle do cancro cítrico.

#### **OBJETIVOS DA INVENÇÃO**

[014] Um objetivo da presente invenção é apresentar a utilização de compostos químicos para a utilização no combate ao cancro cítrico.

#### **BREVE DESCRIÇÃO DAS FIGURAS**

[015] A Figura 1 - Desenhos dos oligonucleotídeos resultantes, sendo indicados os vetores pET28 (ou 29) e pNPTS138 (negrito) nos quais os fragmentos amplificados com os respectivos oligonucleotídeos foram clonados. Bases em vermelho indicam regiões de sobreposição com o fragmento gênico de interesse. Em letras minúsculas estão as bases adicionadas para otimização nas regiões 5', Sítios de restrição estão grifados em amarelo, com as respectivas enzimas de restrição identificadas acima e pontos de clivagem representados por "/". As bases que são

transcritas nos códons de fim de tradução (*stop codon*) estão grifadas em azul. Os resíduos de aminoácidos das traduções conceituais previstas da região N-terminal de cada proteína de interesse estão mostrados em verde. As temperaturas de *melting* (TM) dos oligonucleotídeos estão indicadas. FW e RV se referem aos oligonucleotídeos direto e reverso para o fragmento indicado. KO indica os oligonucleotídeos utilizados na confirmação da deleção e complementação gênicas. cPMI indica a região codificante do gene *xanB*, clonada em pNPTS138 para fins de complementação gênica do mutante de deleção. 1kb\_Up e 1kb\_Down consistem nos fragmentos flanqueadores de 1 kb da ORF a ser deletada.

[016] Figura 2 - Análise dos produtos de amplificação por PCR a partir de DNA genômico de XAC. Aliquotas de 3 µL de cada reação de PCR foram analisadas em eletroforese em gel de TAE agarose 0,8%. M | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). 1 | Produto de PCR referente à região codificante da PMI madura (1404 pb), utilizada para clonagem no vetor pET28a. 2 | Produto de PCR referente à região codificante da PMI madura sem stop códon nativo (1401 pb), utilizada para clonagem no vetor pET29b. 3 e 4 | Produtos de PCR referentes aos fragmentos de 1kb *upstream* e *downstream* ao gene *xanB*, respectivamente. 5 | Produto de PCR referente à região codificante da PMI madura (1404 pb), utilizada na construção do vetor de complementação pNPTS138\_CxanB. 6 | Produto de PCR referente à região codificante da PMI madura (1404 pb), utilizada na construção do vetor de expressão pET41a\_GST\_PMI.

[017] Figura 3 - Análise de restrição dos DNAs plasmidiais para confirmação da clonagem dos fragmentos de

interesse no vetor de propagação pJET 1.2. Entre 500 e 1000 ng de DNA plasmidial dos vetores de propagação foram submetidos à dupla digestão com as endonucleases específicas e analisados em eletroforese em gel de TAE agarose 0,8%. M | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). Cada vetor de propagação pJET 1.2 (2974 pb) foi digerido para excisão dos seguintes fragmentos. 1 | Região codificante do gene *xanB*. 2 | Região codificante do gene *xanB* sem o respectivo stop códon nativo. 3 | Região de 1 kb *upstream* ao gene *xanB*. 4 | Região de 1 kb *downstream* ao gene *xanB*. 5 | Região codificante do gene *xanB* para montagem do vetor de complementação pNPTS\_CxanB. 6 | Região codificante do gene *xanB* para montagem do vetor de expressão pET41a\_GST\_PMI. Uma seta indica a altura da banda do vetor de propagação pJET 1.2 (aproximadamente 3 kb), comum em todas as amostras.

[018] Figura 4 - Análise de restrição dos DNAs plasmidiais para confirmação da clonagem das regiões codificadoras da PMI nos vetores de expressão da série pET. Os vetores de expressão recombinante foram digeridos com as endonucleases específicas e analisados em eletroforese em gel de TAE agarose 0,8%. M | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). 1 | Vetor de expressão pET28a\_PMI digerido com *XhoI* e *NdeI*. 2 | Vetor de expressão pET29b\_PMI\_C digerido com *XhoI* e *NdeI*. 3 | Vetor de expressão pET41a\_GST\_PMI digerido com *EcoRI* e *XhoI*. Tamanhos de bandas do padrão mais próximos ao inserto e vetor estão indicados. Os plasmídeos pET28a e pET29b possuem cerca de 5,3 kb, já o pET41a possui 5,9 kb e os

insertos referentes às regiões codificantes da PMI possuem 1404 ou 1401 pb.

[019] Figura 5 - Análise da expressão, solubilidade e purificação da XACPMI por SDS-PAGE. A figura mostra a purificação da XACPMI (aproximadamente 53 kDa), obtida a partir do vetor de expressão pET28a\_PMI. I | Fração insolúvel do lisado. S | Fração solúvel do lisado. D | Depois da indução por IPTG. A | Antes da indução por IPTG. M | Padrão de massa molecular *Pierce Unstained Protein MW Marker* (Thermo). PC | Fração solúvel após passagem pela coluna de purificação. A proteína foi eluída pela aplicação de um gradiente crescente de concentrações de imidazol (concentrações em mM, mostradas em azul e ordem decrescente de concentração). As eluições foram realizadas com volumes de imidazol referentes a cinco volumes de coluna. O ensaio de expressão teve um rendimento de 5 mg de proteína recombinante por litro de meio de cultura. Massas do padrão estão indicadas à direita.

[020] Figura 6 - Análise da expressão, solubilidade e purificação da XACPMI\_C sob SDS-PAGE. A figura mostra a purificação da XACPMI\_C (aproximadamente 53 kDa), sintetizada a partir do vetor de expressão pET29b\_PMI\_C. M | Padrão de massa molecular *Precision Plus Protein™ Dual Color* (Bio-Rad). A | Antes da indução por IPTG. D | Depois da indução por IPTG. S | Fração solúvel do lisado. I | Fração insolúvel do lisado. PC | Fração solúvel após passagem pela coluna de purificação. A proteína foi eluída pela aplicação de um gradiente crescente de concentrações de imidazol (concentrações em mM, mostradas em azul). As eluições foram realizadas com volumes de imidazol

referentes a cinco volumes de coluna. O ensaio de expressão teve um rendimento de 5 mg de proteína recombinante por litro de meio de cultura. Massas do padrão estão indicadas à esquerda.

[021] Figura 7 - Análise da expressão, solubilidade e purificação da XACGST\_PMI por SDS-PAGE. A figura mostra a purificação da XACGST\_PMI (aproximadamente 79 kDa) em coluna de glutatona imobilizada, a partir do vetor de expressão pET41a\_GST\_PMI. M | Padrão de massa molecular *Precision Plus Protein™ Dual Color* (Bio-Rad). A | Antes da indução por IPTG. D | Depois da indução por IPTG. S | Fração solúvel do lisado. I | Fração insolúvel do lisado. A proteína foi eluída pela aplicação de 5 volumes de coluna (5 ml) de glutatona reduzida a 10 mM, sendo E1-E5 (mostrados em vermelho) os eluatos referentes a cada um dos volumes aplicados (1 ml). Massas do padrão estão indicadas à esquerda. O ensaio de expressão teve um rendimento de 6 mg de proteína recombinante por litro de meio de cultura.

[022] Figura 8 - Avaliação da atividade enzimática da XACGST\_PMI. A XACGST\_PMI purificada também foi submetida ao teste qualitativo de atividade enzimática utilizando o reagente de Seliwanoff. Branco | Composto por 0,1 M de D-manose-6-fosfato, tampão e o reagente de Seliwanoff. Reação | Reação contendo 0,1 M de D-manose-6-fosfato, XACGST\_PMI (0,3 mg/ml) em tampão e o reagente de Seliwanoff. F6P | Controle positivo contendo 0,1 M de D-frutose-6-fosfato, tampão e o reagente de Seliwanoff. A coloração vermelho-amarronzada no tubo "Reação" permitiu a conclusão de que a XACGST\_PMI também possui a atividade catalítica predita de conversão de D-manose-6-fosfato em D-frutose-6-fosfato.

Todos os tampões utilizados foram constituídos de 50 mM Tris-HCl (pH 8,0), 100 mM NaCl.

[023] Figura 9 - Análise de restrição do vetor de deleção pNPTS138\_xanB. DNA plasmidial do vetor de deleção do gene *xanB* foi submetido a diversas digestões com as endonucleases utilizadas para a sua construção e analisado em eletroforese em gel de TAE agarose 0,8%. Painel A | Representação gráfica de um segmento do vetor de deleção pNPTS138\_xanB. Painel B | Análise de restrição do vetor de deleção pNPTS138\_xanB. M | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). 1 | pNPTS138\_xanB digerido somente com *EcoRI*. 2 | pNPTS138\_xanB digerido com *HindIII* e *EcoRI*. 3 | pNPTS138\_xanB digerido com *EcoRI* e *NheI*. 4 | pNPTS138\_xanB digerido com *HindIII* e *NheI*. 5 | pNPTS138\_xanB digerido com *HindIII*, *EcoRI* e *NheI*. Tamanhos de bandas do padrão estão indicados.

[024] Figura 10 - Metodologia para seleção dos possíveis mutantes deletados. Diferentes cultivos foram realizados para facilitar o *screening* de linhagens mutantes deletadas. 1-2 | Células de XAC eletrocompetentes foram transformadas por eletroporação com o vetor de deleção e, em seguida, plaqueadas em ágar-LB adicionado de canamicina (Meio A). 3 | Colônias transformantes foram repicadas em ágar-LB contendo canamicina e sacarose (Meio B) e ágar-LB contendo somente sacarose (Meio C). 4-5 | Colônias que não apresentaram crescimento nos meios B e C foram cultivadas em caldo-LB (Meio D) e então estriadas em ágar-LB (Meio E) para seu isolamento. 6 | Colônias isoladas foram repicadas nos meios A, B e C, para seleção dos possíveis mutantes deletados (aqueles que apresentaram crescimento apenas no



meio C).

[025] Figura 11 - Confirmação da deleção do gene *xanB* por PCR. Análise dos produtos de PCR resultantes das reações utilizando oligonucleotídeos que hibridizam em regiões externas às regiões *upstream* e *downstream* de 1 kb ao gene *xanB*, com DNA molde de colônias de XAC selvagem, XccΔ*xanB* e o vetor de deleção pNPTS138\_*xanB*. M | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas); 1 | Produto de PCR utilizando XAC selvagem como molde. 2 | Produto de PCR utilizando XccΔ*xanB* como molde. 3 | Produto de PCR utilizando XccΔ*xanB* como molde, digerido com a endonuclease *EcoRI*. 4 | Produto de PCR utilizando o vetor pNPTS138\_*xanB* como molde. As bandas relevantes do padrão de tamanho molecular estão indicadas.

[026] Figura 12 - Construção do vetor de complementação pNPTS\_C*xanB* e confirmação da orientação da clonagem do gene *xanB*. A confirmação da construção do vetor de complementação pNPTS\_C*xanB* foi realizada por digestão com endonuclease, seguida pela análise por eletroforese em gel de TAE agarose 0,8%. No painel superior: M | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). 1 | pNPTS\_C*xanB* digerido com a enzima de restrição *EcoRI*. O padrão de digestão do vetor mostra que a montagem do vetor foi realizada com sucesso, uma vez que a região codificante apresenta 1404 pb, um tamanho compatível com o observado no gel. 2 | Produto de PCR utilizando o par de oligonucleotídeos Up\_PMI\_FW-pNPTS\_cPMI\_RV e o vetor pNPTS\_C*xanB* como molde. 3 | Produto de PCR utilizando o par de oligonucleotídeos Up\_PMI\_FW-pNPTS\_PMI\_FW e o vetor pNPTS\_C*xanB* como molde. A observação de uma banda na altura

de 2,4 kb na coluna 2 e a inexistência de bandas na coluna 3 comprova que a orientação da clonagem da região codificante do gene *xanB* no vetor de complementação é a esperada. No painel inferior está representada a organização genômica da região do gene *xanB*, bem como os oligonucleotídeos utilizados na PCR de confirmação da orientação da clonagem.

[027] Figura 13 - Confirmação da complementação gênica do mutante *XccΔxanB*. Análise dos produtos de PCR resultantes das reações utilizando oligonucleotídeos que hibridizam em regiões externas às regiões *upstream* e *downstream* de 1 kb ao gene *xanB*, com DNA genômico de XAC selvagem e *XccΔCxanB* como molde. M | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas); 1 | Produto de PCR utilizando XAC selvagem como molde. 2 | Produto de PCR utilizando XAC selvagem como molde, digerido com a enzima *EcoRI*. 3 | Produto de PCR utilizando *XccΔCxanB* como molde. 4 | Produto de PCR utilizando *XccΔCxanB* como molde, digerido com a enzima *EcoRI*. A restrição por *EcoRI* do fragmento amplificado ocorreu para *XccΔCxanB* e não para XAC, conforme esperado, resultando em duas bandas (um fragmento de 1,4 kb e dois fragmentos de 1 kb), o que é compatível com o tamanho de *xanB* e das duas regiões flanqueadoras. As bandas relevantes do padrão de tamanho molecular estão indicadas.

[028] Figura 14 - Ensaio de patogenicidade *in vivo* de XAC (*Xcc*), *XccΔxanB* e *XccΔCxanB*, inoculadas por aspensão em *Citrus aurantifolia*. Plantas de *Citrus aurantifolia* foram utilizadas para avaliação comparativa da patogenicidade e virulência de XAC, *XccΔxanB* e *XccΔCxanB*. Foram pulverizados

10 ml de cada cultivo bacteriano e o mesmo volume de solução salina (controle negativo) em quadruplicatas de plantas. Após serem destacadas das plantas, as folhas foram fotografadas no vigésimo oitavo dia após a aspersão.

[029] Figura 15 - Ensaio de patogenicidade *in vivo* de XAC (Xcc), Xcc $\Delta$ xanB e Xcc $\Delta$ CxanB, inoculadas por infiltração em *Citrus aurantifolia*. Plantas de *Citrus aurantifolia* foram utilizadas para avaliação comparativa da patogenicidade e virulência de XAC, Xcc $\Delta$ xanB e Xcc $\Delta$ CxanB. Para cada uma das condições foram infiltradas quatro folhas de ramos independentes, atentando-se para que todas as plantas tivessem folhas infiltradas com as três linhagens bacterianas e o controle negativo com solução salina (0,9% de NaCl). As fotografias foram obtidas no vigésimo dia após a infiltração.

[030] Figura 16 - Efeito de XanB na motilidade de XAC. As linhagens XAC (Xcc), Xcc $\Delta$ xanB e Xcc $\Delta$ CxanB foram avaliadas quanto à motilidade segundo metodologia padronizada por Malamud e colaboradores (2011). As linhagens foram cultivadas até que atingissem DO<sub>595 nm</sub> igual a 0,3. Três  $\mu$ l de cada cultivo foram pipetados no centro das placas de Petri de 5,1 cm de diâmetro com ágar-LB 0,7%. Após isso, as placas foram incubadas a 30°C por 48 horas sem agitação e registradas digitalmente (Painel superior). O diâmetro das colônias obtidas foi mensurado utilizando o software ImageJ e a média das medidas está apresentada no painel inferior. Barras de erro indicam o desvio-padrão absoluto de cada uma das triplicatas. A análise estatística dos dados foi realizada pelo teste de Tukey, com nível de confiança de 95% por meio do software Minitab (15.2).

[031] Figura 17 - Efeito de XanB na formação de biofilme. As linhagens XAC (Xcc), Xcc $\Delta$ xanB e Xcc $\Delta$ CxanB foram avaliadas quanto à formação de biofilme segundo metodologia padronizada por O'Toole e Kolter (1998). Os resultados estão apresentados como a média da razão entre a absorvância de cristal violeta em 595 nm e a densidade óptica de cada cultivo bacteriano imediatamente antes da mensuração da formação de biofilme ( $A_{595 \text{ nm}}/DO_{595 \text{ nm}}$ ). As medidas foram auferidas em 24, 48 e 72 horas. Barras de erro indicam o desvio-padrão absoluto de cada uma das sextuplicatas.

[032] Figura 18 - Efeito de XanB na resistência à radiação ultravioleta. As linhagens XAC (Xcc), Xcc $\Delta$ xanB e Xcc $\Delta$ CxanB foram avaliadas quanto à sobrevivência após exposição à radiação ultravioleta por 15 minutos a uma distância de 60 cm da fonte de luz. Os resultados estão apresentados em porcentagem de sobrevivência (% de sobrevivência), a qual foi calculada a partir dos controles de cada linhagem bacteriana não exposta à radiação ultravioleta. A análise estatística dos dados foi realizada pelo teste de Tukey, com nível de confiança de 95% por meio do software Minitab (15.2).

[033] Figura 19 - Alinhamento múltiplo final das sequências da Fosfomanose isomerase e as respectivas às estruturas indicadas como homólogas pelo BLASTp (PDBs IDs 1H5R, 2CU2, 2X5S e 2QH5) após o refinamento. Os aminoácidos com coloração azul escuro pertencem a folhas- $\beta$ , enquanto que aqueles com coloração verde pertencem a  $\alpha$ -hélices. Os aminoácidos que tiveram seus rotâmeros modificados estão realçados em azul. A simbologia de asterisco (\*) representa

a presença de um único resíduo totalmente conservado, dois pontos (:) indica a conservação entre grupos de propriedades fortemente semelhantes, e um ponto (.) indica a conservação entre grupos de propriedades fracamente semelhantes.

[034] Figura 20 - Validação do modelo gerado para a Fosfomanose isomerase de *Xanthomonas citri* subsp. *citri* mostrando a posição do ligante glicose-1-fosfato e sua posição em relação ao ligante cristalográfico, quando utilizados os programas **GOLD** e **GLIDE**. A estrutura em verde corresponde ao modelo gerado (Hipótese 4) e a em rosa corresponde à estrutura cristalográfica PDB ID 1H5R.

[035] Figura 21 - Curva ROC obtida para o estudo dos 30 compostos relatados como mais potentes e menos potentes do BindingDB com respeito ao modelo aqui obtido para a proteína Fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*.

[036] Figura 22 - Representação dos modos de interação dos compostos selecionados DL03874928, DL03999102, DL04091967 e DL35935889 na região do sítio ativo do modelo de XAC em duas dimensões obtidas como resultado do programa GOLD.

[037] Figura 23 - Inibidor DL04091967 (em representação *stick*) docado no sítio ativo do modelo da PMI de *Xanthomonas citri* subsp *citri* (em representação *Ribbons*).

[038] Figura - 24 - Alinhamento múltiplo final das sequências da Fosfomanose isomerase humana (PMI) e as respectivas às estruturas indicadas como homólogas pelo BLASTp. (PDBs IDs 5NW7 e 3H1M). Os aminoácidos com coloração azul escuro pertencem a folhas- $\beta$  enquanto aqueles

com coloração verde pertencem a  $\alpha$ -hélices. A simbologia de asterisco (\*) representa a presença de um único resíduo totalmente conservado, dois pontos (:) indica a conservação entre grupos de propriedades fortemente semelhantes e um ponto (.) indica a conservação entre grupos de propriedades fracamente semelhantes.

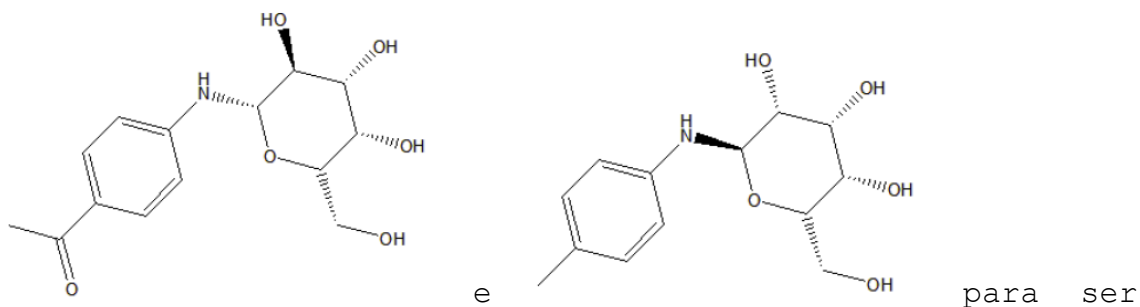
[039] Figura 25 - Curva ROC obtida para o estudo dos 30 compostos relatados como ativos e inativos no **BindingDB** para a proteína PMI de *Homo sapiens*.

[040] Figura 26 - Inibidores de PMI de XAC selecionados por triagem computacional utilizando o modelo tridimensional da PMI de XAC: 1) 2-(hidroximetil)-6-[2-(hidroximetil)phenoxy]oxano-3,4,5-triol (DL03874928); 2) 2-(hidroximetil)-6-[(piridina-2-il)amino]oxano-3,4,5-triol (DL35935889); 3) 2-(hidroximetil)-6-[(4-metilfenil)amino]oxano-3,4,5-triol (DL03999102) e 4) 1-(4-[[3,4,5-trihidroxi-6-(hidroximetil)oxa-2-il]amino}fenil)eta-1-ona (DL04091967).

[041] Figuras 27, 28 e 29 - Avaliação do efeito preventivo dos sintomas do cancro cítrico por inibidores selecionados por triagem computacional e aplicados por aspersão em folhas de plantas de limão galego, seguida da aspersão de XAC. No controle foi aspergido apenas água, seguido da aspersão de XAC.

#### **DESCRIÇÃO DETALHADA DA INVENÇÃO**

[042] O presente invento consiste na utilização de um composto ou de um sal aceitável do mesmo, em que o referido composto é selecionado a partir do grupo que consiste em:



para ser utilizado no controle de cancro cítrico, através da inibição seletiva da Fosfomanose isomerase (PMI).

[043] O conteúdo a seguir apresenta modalidades de etapas principais e testes realizados que chegam à proposição do uso dos compostos acima como inibidores de PMI no controle do cancro cítrico e não devem ser considerados como limitantes do presente invento.

[044] As etapas de proposição do uso de inibidores de PMI no controle do cancro cítrico consistiram em:

- Etapa I: Análise Funcional para comprovação da PMI como proteína essencial para a patogenicidade de XAC.
- Etapa II: Ensaios e experimentos realizados *in silico* para planejamento de inibidores seletivos da PMI de XAC.

Etapa I - Análise Funcional para comprovação da PMI como proteína essencial para a patogenicidade de XAC

#### **Linhagens bacterianas, meios e condições de cultivo**

[045] A linhagem *Xanthomonas citri* subsp. *citri* 306 (XAC) foi armazenada a  $-80^{\circ}\text{C}$  em meio Luria-Bertani (LB, Sigma) com 10% de glicerol. Os cultivos de XAC, *E. coli* DH5 $\alpha$  e *E. coli* BL21(DE3) em meios sólido e líquido foram realizados em ágar-LB (Sigma) e caldo LB (Sigma),

respectivamente. Para XAC, os cultivos foram incubados a 30°C e 200 rpm, e para as linhagens de *E. coli*, a temperatura foi de 37°C e 250 rpm.

### **Enzimas, kits e procedimentos em biologia molecular**

[046] Para a extração de DNA genômico de XAC foi utilizado o *PureLink™ Microbiome DNA Purification Kit* (Invitrogen). DNA polimerase (Phusion High Fidelity DNA Polimerase), enzimas de restrição, vetor de propagação e demais *kits* de extração de DNA foram adquiridos da Fermentas-Thermo Scientific e da Cellco Biotec. Os antibióticos ampicilina, canamicina e gentamicina foram obtidos da Sigma.

[047] Quando não detalhados, os procedimentos de biologia molecular foram conduzidos exatamente como descritos anteriormente (AUSUBEL, 2002).

### **Desenho e síntese de oligonucleotídeos**

[048] Os oligonucleotídeos utilizados na amplificação por PCR (*Polymerase Chain Reaction*, Reação em Cadeia da DNA Polimerase) da região codificante da fosfomanose isomerase de XAC foram projetados com base na sequência depositada no GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>), cujo código de acesso é XAC3580/AAM35493.1 (GDP-mannose pyrophosphorylase), de tal forma a coincidirem 15-20 bases da região 5' da fita codificante e da sua respectiva fita reversa complementar (3' do fragmento gênico).

[049] As análises *in silico* da ORF XAC3580 utilizando o software SignalP 4.0 (PETERSEN et al., 2011) (<http://www.cbs.dtu.dk/services/SignalP>) indicaram ausência



de peptídeo sinal na sequência peptídica predita.

[050] Para a construção do plasmídeo de deleção do gene-alvo (PMI, ORF XAC3580), foram projetados oligonucleotídeos designados à amplificação por PCR das regiões de 1 kb *upstream* e *downstream* à ORF XAC3580, com base na sequência do genoma completo de XAC, depositada no GenBank sob o código de acesso NC\_003919.1.

[051] Para a confirmação da deleção e complementação gênicas, foram desenhados pares de oligonucleotídeos que hibridizam em segmentos adjacentes às regiões flanqueadoras *upstream* e *downstream* de 1 kb. Como tais oligonucleotídeos não hibridizam nos vetores de deleção e complementação, tal estratégia foi adotada criteriosamente no sentido de certificar que os produtos de PCR obtidos a partir das linhagens mutante e complementante são oriundos de hibridização dos oligonucleotídeos no DNA cromossômico das linhagens e não no vetor plasmidial utilizado.

[052] Todos os oligonucleotídeos possuem sítios de restrição estrategicamente posicionados para a posterior clonagem dos produtos de PCR nos vetores de propagação, expressão, deleção e/ou complementação.

[053] Foi utilizado o software Lasergene 7.1 ([www.dnastar.com](http://www.dnastar.com)) para a análise das sequências nucleotídicas, bem como para a escolha e posicionamento dos sítios de restrição das endonucleases adotadas, de modo que tais sítios não estivessem presentes nas sequências nucleotídicas a serem clonadas.

[054] Foram inseridas bases adicionais às extremidades 5' de alguns oligonucleotídeos para possibilitarem, caso necessário, clonagem direta no vetor de expressão e,

também, otimização dos parâmetros termodinâmicos como temperatura de transição de fase (*melting temperature*) e temperatura de formação de superestruturas indesejáveis como *hairpins*, homo e hetero dímeros de hibridização oligo-oligo. Tais otimizações foram feitas com o software OligoAnalyzer 3.1

(<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer>)

. Os oligonucleotídeos foram sintetizados pela Exxtend (Paulínia-SP, Brasil).

[055] A Figura 1 apresenta a sequência de todos os oligonucleotídeos e os respectivos vetores utilizados e a Tabela 1 apresenta os pares de oligonucleotídeos e seus respectivos produtos de PCR e usos.

Tabela 1 - PARES DE OLIGONUCLEOTÍDEOS E SEUS RESPECTIVOS PRODUTOS DE PCR (pPCR), FUNÇÕES E VETORES NOS QUAIS FORAM CLONADOS

Par de oligonucleotídeos	Produto de PCR (pPCR)	Função
pET2829PMI_FW e pET2829PMI_RV	Região codificante do gene <i>xanB</i> (pPCR1)	Expressão Heteróloga
pET2829PMI_FW e pET2829PMI_RV_C	Região codificante do gene <i>xanB</i> sem o <i>stop codon</i> nativo do gene (pPCR2)	Expressão Heteróloga
pNPTS_cPMI_FW e pET2829PMI_RV	Região codificante do gene <i>xanB</i> a ser clonada no vetor pET41a (pPCR3)	Expressão Heteróloga
Up_PMI_FW e Up_PMI_RV	Região de 1 kb <i>upstream</i> ao gene <i>xanB</i> (pPCR4)	Deleção

		Gênica
<b>Down_PMI_FW e Down_PMI_RV</b>	Região de 1 kb <i>downstream</i> ao gene <i>xanB</i> <b>(pPCR5)</b>	Deleção Gênica
<b>KO_PMI_FW e KO_PMI_RV</b>	Região de 2 kb (linhagem mutante XccΔXanB) ou de 3,4 kb (linhagem selvagem XAC)	Confirmação da deleção e complementação gênicas
<b>pNPTS_cPMI_FW e pNPTS_cPMI_RV</b>	Região codificante do gene <i>xanB</i> <b>(pPCR6)</b>	Complementação gênica

[056] Foram projetados dois oligonucleotídeos reversos que hibridizam na região 3' do fragmento gênico codificante da fosfomanose isomerase, objetivando o aumento da variabilidade de produtos de expressão quanto à presença e posicionamento da(s) cauda(s) de histidinas. Além disso, um par de oligonucleotídeos permite a amplificação de um fragmento gênico que resulta na obtenção da PMI recombinante fusionada à glutathione-S-transferase (GST) quando clonado no vetor pET41a.

#### **Extração de DNA genômico**

[057] Para a extração de DNA genômico de XAC, uma colônia isolada foi inoculada em 10 ml de caldo LB, seguida de incubação a 30°C e 200 rpm até que o cultivo atingisse a DO<sub>595nm</sub> igual a 0,5. As células foram coletadas por centrifugação e lavadas três vezes com PBS (Tampão 50 mM fosfato de sódio pH 7,4; 150 mM NaCl).

[058] Em seguida, o pellet de bactérias foi utilizado para a extração de DNA genômico por meio do *kit PureLink™ Microbiome DNA Purification Kit* (Invitrogen), segundo

instruções do fabricante. O DNA foi suspenso em água mili-Q e quantificado por espectrometria a 260 nm.

#### **Amplificação e clonagem dos fragmentos gênicos de interesse**

[059] Os fragmentos gênicos de interesse foram amplificados por PCR utilizando DNA genômico de XAC como molde e os oligonucleotídeos correspondentes. As amplificações foram realizadas em um termociclador C1000 Touch (Bio-Rad) segundo o descrito a seguir:

- 50 ng de DNA genômico
- 0,25 µM de oligonucleotídeo FW
- 0,25 µM de oligonucleotídeo RV (ou RV\_C)
- 1 U de Phusion DNA polimerase
- 1,5 mM de MgCl<sub>2</sub>
- 0,2 mM de dNTP
- H<sub>2</sub>O MilliQ q.s.p 50 µl

[060] Os ciclos utilizados nas amplificações por PCR destinadas ao isolamento dos fragmentos gênicos de interesse estão descritos na Tabela 2 e os produtos de PCR analisados por eletroforese em gel de agarose estão mostrados na Figura 2.

TABELA 2 - CICLOS DE PCR UTILIZADOS NA AMPLIFICAÇÃO DOS FRAGMENTOS GÊNICOS

<b>Etapa da PCR</b>	<b>Produtos de PCR</b>	
	<b>Temperatura</b>	<b>Tempo</b>
Desnaturação inicial	98°C	10 minutos
Desnaturação	98°C	30 segundos

		(35x)
Hibridização	62°C	30 segundos (35x)
Extensão	72°C	1,5 minuto (35x)
Extensão final	72°C	10 minutos

[061] Após a purificação por eletroforese em gel de TAE agarose 0,8% contendo brometo de etídeo, os produtos de PCR foram utilizados em reação de ligação (com T4 DNA ligase) para clonagem no vetor pJET 1.2 (Fermentas), seguindo a razão molar de 3:1 (inserto:vetor).

[062] As reações de ligação foram utilizadas na transformação de células de *E. coli* DH5 $\alpha$  competentes por tratamento com cloreto de cálcio (AUSUBEL, 2002). As células transformadas foram plaqueadas em meio ágar-LB contendo ampicilina 100  $\mu\text{g}.\text{ml}^{-1}$ .

[063] Colônias bacterianas resistentes à ampicilina foram submetidas à extração de DNA plasmidial para averiguação de clones positivos para cada um dos insertos por dupla digestão do DNA plasmidial com as endonucleases específicas (Figura 1), seguida de análise por eletroforese em gel de TAE agarose 0,8% contendo brometo de etídeo (Figura 3).

[064] Os clones positivos (insertos com tamanho próximo ao esperado) tiveram o DNA plasmidial sequenciado pelo método do didesoxinucleotídeo marcado de Sanger (SANGER; NICKLEN; COULSON, 1977) no Centro de Pesquisas sobre o Genoma Humano e Células-tronco (CEGH-CEL), do Instituto de

Biociências da USP (São Paulo-SP). As reações foram feitas utilizando o *BigDye® Terminator v3.1 Cycle Sequencing Kit* (Applied Biosystems™), segundo as instruções do fabricante, em um equipamento 3730 *DNA Analyser* (Applied Biosystems™).

[065] Os insertos referentes aos produtos de PCR 1, 2 e 3 (Tabela 1) foram digeridos com as endonucleases específicas (Figura 1), purificados por eletroforese em gel de TAE agarose 0,8% e clonados nos vetores pET28a (Novagen), pET29b (Novagen), pET41a (Novagen), respectivamente, em reações de ligação com T4 DNA ligase (razão molar inserto:vetor de 3:1) com os vetores previamente digeridos com as enzimas de restrição específicas para cada uma das construções (*NdeI* e *XhoI* para pET28a e pET29b; *EcoRI* e *XhoI* para pET41a).

[066] As reações de ligação foram utilizadas na transformação de células de *E. coli* DH5 $\alpha$  quimiocompetentes. As células transformantes foram plaqueadas em meio ágar-LB contendo canamicina 30  $\mu\text{g.ml}^{-1}$  e os clones resistentes foram cultivados em caldo LB adicionado de canamicina 30  $\mu\text{g.ml}^{-1}$  para fins de extração do DNA plasmidial, seguida da confirmação da presença dos insertos por eletroforese em gel de TAE agarose 0,8% após a dupla digestão com as endonucleases específicas.

[067] Após a confirmação do sucesso das clonagens (Figura 4), os vetores foram denominados pET28a\_PMI, pET29b\_PMI\_C e pET41a\_GST\_PMI, os quais possibilitam a obtenção das seguintes proteínas recombinantes: PMI fusionada a uma cauda de histidinas N-terminal, PMI fusionada a uma cauda de histidinas C-terminal e PMI

fusionada à GST na região N-terminal.

### **Expressão heteróloga e purificação por cromatografia de afinidade**

[068] Células quimiocompetentes de *E. coli* BL21(DE3) (Novagen) foram separadamente transformadas com os vetores de expressão obtidos e plaqueadas em ágar-LB contendo canamicina 30  $\mu\text{g.ml}^{-1}$ .

[069] Colônias isoladas de cada uma das construções foram inoculadas em caldo LB adicionado de canamicina 30  $\mu\text{g.ml}^{-1}$  e submetidas aos ensaios de expressão induzidos por IPTG 0,1 mM, adicionado na fase *mid-log* de crescimento populacional bacteriano ( $\text{DO}_{595\text{nm}} = 0,4$ ). Foi adicionado cloreto de zinco (Sigma) 0,2  $\mu\text{M}$  ao meio de cultura de expressão heteróloga, uma vez que o zinco é um cofator da PMI de *Xanthomonas campestris*, proteína com elevada identidade em relação à PMI de XAC (PAPOUTSOPOULOU; KYRIAKIDIS, 1997).

[070] Os ensaios de expressão foram incubados por 16 horas a 18°C em agitador orbital a 250 rpm. Ao final da indução, as células foram coletadas por centrifugação, ressuspensas em tampão 50 mM Tris-HCl (pH 8), NaCl 100 mM e lisadas por cinco pulsos de ultrassom (*Sonic Dismembrator 500, Fisher Scientific*) com duração de 30 segundos e amplitude de 30%, com intervalos de um minuto entre os pulsos, sob banho de gelo. Para a separação da fração solúvel, o lisado celular foi submetido à centrifugação a 4°C, 12.000 x g por 30 minutos.

[071] A purificação de todas as proteínas recombinantes foi realizada por cromatografia de afinidade utilizando

ions metálicos imobilizados (IMAC, *Immobilized Metal Affinity Chromatography*) para as proteínas fusionadas somente à cauda de histidinas, ou glutathione para a proteína fusionada à GST.

[072] Para IMAC, a fração solúvel do lisado celular foi aplicada em uma coluna de níquel imobilizado Ni-NTA (Novagen) e as proteínas foram eluídas da coluna por meio da aplicação de um gradiente crescente de concentrações de imidazol (de 5 mM a 500 mM). Alíquotas das frações solúvel e insolúvel, bem como alíquotas do pós-coluna (após a passagem da fração solúvel do lisado pela coluna cromatográfica de afinidade) e dos eluatos em imidazol foram submetidas à análise em SDS-PAGE (LAEMMLI, 1970) para averiguar a funcionalidade dos sistemas de expressão e purificação (Figuras 5 e 6).

[073] Para a PMI fusionada à GST na região N-terminal (XACGST\_PMI), o procedimento de purificação foi realizado utilizando 800 µl de uma resina de sefarose imobilizada com glutathione (*Glutathione Sepharose® 4 Fast Flow, GE Healthcare*). A coluna foi lavada com cinco volumes de coluna de água destilada, seguida de uma etapa de equilíbrio com dez volumes de coluna com tampão 50 mM Tris-HCl (pH 8,0) 100 mM NaCl.

[074] Em seguida, a fração solúvel do lisado celular foi aplicada à coluna por gravidade, e a eluição foi realizada com 5 ml de glutathione reduzida (Sigma) a uma concentração de 10 mM. Nesta etapa de eluição, foram coletadas alíquotas de 1 ml para averiguação da eficiência do processo de purificação, por SDS-PAGE (Figura 7).

[075] Para os experimentos de atividade enzimática, os



ensaios de expressão foram repetidos e otimizados quanto à purificação, utilizando apenas as soluções de imidazol 25 mM ou de tampão 50 mM Tris-HCl (pH 8,0), NaCl 100 mM para lavagem da coluna de níquel ou de glutatona imobilizados, respectivamente, para eliminação de contaminantes, e 500 mM de imidazol para eluição das proteínas XACPMI e XACPMI\_C purificadas, além de 10 mM de glutatona reduzida para a eluição da proteína XACGST\_PMI.

[076] Essas últimas frações foram dialisadas contra tampão 50 mM Tris-HCl (pH 8,0), 100 mM NaCl, e quantificadas pela medida da absorção no UV em comprimento de onda de 280 nm, em um espectrofotômetro *NanoVue Plus™* (Biorchrom) utilizando como base o coeficiente de extinção molar de cada uma das proteínas recombinantes obtidas.

#### **Avaliação da atividade enzimática da PMI**

[077] A PMI catalisa a isomerização reversível de D-manose-6-fosfato em D-frutose-6-fosfato e, portanto, o teste de Seliwanoff foi usado para detectar a formação desta cetose.

[078] O teste de Seliwanoff é um teste convencionalmente utilizado para a detecção de cetoses, como frutose, xilulose, ribulose e tagatose (CHAWLA, 2014), e é baseado em um reagente composto por resorcinol (benzeno-1,3-diol) 0,05% (m/v, concentração final) em ácido clorídrico (HCl) diluído em água (1:1). A função do ácido clorídrico é desidratar a cetose e produzir um composto chamado hidroximetilfurfural, o qual reage com o resorcinol formando uma molécula de coloração avermelhada (ou esverdeada, dependendo da cetose), cuja intensidade é

proporcional à quantidade de cetose presente na solução (ROE, 1934).

[079] As reações foram constituídas por 30 µg de XACGST\_PMI (fusionada à GST e com cauda de histidinas na extremidade N-terminal), D-manose-6-fosfato (sal dissódico, Sigma) em concentração final de 0,5 mM, considerando volume final de 500 µl. Após 2 horas de reação, foram adicionados 3 ml de reagente Seliwanoff com a coloração formada sendo registrada por fotografia. Para a reação de controle negativo, o volume da enzima foi substituído pelo tampão no qual a mesma foi solubilizada (50 mM Tris-HCl (pH 8,0), 100 mM NaCl).

[080] O controle positivo teve a mesma constituição da reação anterior, substituindo D-manose-6-fosfato por D-frutose-6-fosfato, também em concentração 0,5 mM. Os testes foram realizados em triplicata (Figura 8).

### **Deleção gênica - Construção do vetor de deleção**

[081] As linhagens mutantes foram obtidas usando o vetor suicida pNPTS138 (Alley Dickon, resultados não publicados), gentilmente fornecido pelo Prof. Dr. Henrique Ferreira (UNESP - Rio Claro), seguindo metodologia padronizada por nosso grupo de pesquisa (ALEXANDRINO; GOTO; NOVO-MANSUR, 2016; CABREJOS et al., 2019; GOTO et al., 2016). O gene-alvo *xanB* foi deletado utilizando uma metodologia baseada na dupla recombinação homóloga entre o vetor pNPTS138 (5,36 kb), no qual foram clonados *in tandem* dois fragmentos de aproximadamente 1 kb correspondentes, no DNA cromossomal de XAC, às regiões *upstream* e *downstream* ao gene-alvo a ser deletado, resultando nos vetores de deleção

do gene *xanB*, denominado pNPTS138\_*xanB*.

[082] Para isso, os insertos referentes aos produtos de PCR 4 e 5 (Tabela 1) foram excisados dos vetores de propagação pJET 1.2 (Fermentas) pela digestão com as endonucleases específicas para cada um deles seguida de purificação por eletroforese em gel de TAE agarose 0,8%.

[083] Para iniciar a construção do vetor de deleção, o plasmídeo pNPTS138 foi previamente digerido com as endonucleases *HindIII* e *EcoRI* e purificado por eletroforese em gel de TAE agarose 0,8%. Em seguida, o fragmento de 1 kb *upstream* purificado foi utilizado em reação de ligação com T4 DNA ligase (razão molar inserto:vetor de 3:1) juntamente com o vetor pNPTS138 digerido.

[084] Tal reação de ligação foi utilizada na transformação de células de *E. coli* DH5 $\alpha$  quimiocompetentes (AUSUBEL, 2002) seguida de plaqueamento em ágar-LB contendo canamicina 30  $\mu\text{g}.\text{ml}^{-1}$ .

[085] Colônias transformantes foram cultivadas em LB contendo canamicina para extração do DNA plasmidial, seguida de digestão com as endonucleases específicas para confirmação da clonagem por eletroforese em gel de TAE agarose 0,8%.

[086] Para a clonagem do fragmento de 1 kb *downstream* ao gene-alvo, seguiu-se o mesmo procedimento, utilizando o vetor já clonado com a região de 1 kb *upstream*, digerindo-o com *EcoRI* e *NheI* para *xanB*.

[087] Novamente, o vetor digerido e purificado compôs reação de ligação com o fragmento de 1 kb *downstream*, seguida da transformação de células de *E. coli* DH5 $\alpha$  quimiocompetentes (AUSUBEL, 2002) e posterior plaqueamento

em ágar-LB contendo canamicina 30  $\mu\text{g}.\text{ml}^{-1}$ .

[088] Ao final da clonagem *in tandem*, colônias transformantes foram cultivadas em LB contendo canamicina 30  $\mu\text{g}.\text{ml}^{-1}$  e tiveram o DNA plasmidial extraído e digerido com as endonucleases específicas, permitindo a confirmação da construção do vetor de deleção como foi planejado (Figura 9).

#### **Deleção gênica - Transformação de XAC por eletroporação**

[089] Cem nanogramas do vetor de deleção pNPTS138\_xanB foram utilizados na eletroporação de células de XAC eletrocompetentes. A eletroporação foi realizada com cubetas de 0,2 cm em um eletroporador *Gene Pulser Xcell Electroporation Systems* (Bio-Rad), sob as seguintes condições: 2,5 kV, 50  $\Omega$  e 50  $\mu\text{F}$ .

[090] Após o pulso elétrico, foram adicionados 2 ml de caldo LB isento de antibióticos e o cultivo foi incubado por 2 horas sob agitação de 200 rpm a 30°C. Em seguida, o cultivo foi plaqueado em ágar-LB contendo canamicina (30  $\mu\text{g}.\text{ml}^{-1}$ ).

#### **Deleção gênica - Seleção de mutantes deletados**

[091] O vetor pNPTS138 possui uma cópia do gene *sacB*, que codifica para a enzima levansucrase (*SacB* de *Bacillus subtilis*), que converte a sacarose em um composto tóxico para as células (BRAMUCCI; NAGARAJAN, 1996). Desse modo, o crescimento em sacarose 10% somente ocorre para as bactérias que o eliminaram (ou seja, quando há excisão do plasmídeo com o gene do cromossomo seguido de cura plasmidial).

[092] Além disso, células resultantes de recombinação única também são eliminadas por possuírem o plasmídeo integrado no genoma. Para aumentar as chances de cura plasmidial nos mutantes que sofreram o duplo *crossing-over*, os mesmos foram cultivados repetidamente em meio de cultivo na ausência de canamicina.

[093] Colônias mutantes de XAC obtidas foram analisadas quanto à deleção por PCR utilizando oligonucleotídeos que hibridizam em regiões adjacentes aos fragmentos flanqueadores. Alíquotas dos cultivos bacterianos oriundos dos mutantes confirmados foram armazenadas a  $-80^{\circ}\text{C}$ .

[094] O processo de *screening* de possíveis mutantes deletados envolve uma série de plaqueamentos e repiques, os quais estão descritos a seguir. Após a eletroporação, colônias transformantes resistentes à canamicina foram repicadas em duas condições diferentes: ágar-LB contendo canamicina a  $30\ \mu\text{g}\cdot\text{ml}^{-1}$  e sacarose 10% (Meio B) e ágar-LB contendo sacarose 10% (Meio C).

[095] Em seguida, colônias que não tiveram crescimento nos meios B e C foram cultivadas em caldo LB isento de antibióticos (Meio D) por 12 horas, sob agitação de 200 rpm, a  $30^{\circ}\text{C}$ , para possibilitar o primeiro evento de recombinação, que resulta na integração do vetor de deleção ao DNA cromossomal de XAC.

[096] Para cada processo de deleção, tal cultivo foi estriado em ágar-LB para isolamento de colônias, as quais foram repicadas em três diferentes condições:

- Meio A: Ágar-LB contendo canamicina a  $30\ \mu\text{g}\cdot\text{ml}^{-1}$ ;
- Meio B: Ágar-LB contendo canamicina a  $30\ \mu\text{g}\cdot\text{ml}^{-1}$  e sacarose 10%;

- Meio C: Ágar-LB contendo sacarose 10%.

[097] O repique de cada uma das colônias nos três meios supracitados facilita o *screening* dos possíveis mutantes deletados, uma vez que colônias resistentes à sacarose e sensíveis à canamicina podem ser resultantes da dupla recombinação seguida da eliminação do plasmídeo de deleção.

[098] A Figura 10 ilustra o procedimento adotado na seleção das possíveis colônias mutantes deletadas. Foram isoladas colônias com capacidade de crescimento no meio contendo sacarose (meio C, etapa 6) e também incapazes de crescer em meio contendo canamicina (meio A, etapa 6), sendo que tais colônias tinham se originado inicialmente de colônia capaz de crescer nesse antibiótico (meio A, etapa 2); isto denota que a célula adquire o plasmídeo mas o "elimina" posteriormente, eventos estes essenciais para a obtenção do mutante e sua confirmação, considerando que no intermédio desses dois eventos vários repiques são realizados.

#### **Deleção gênica - Confirmação da deleção do gene-alvo**

[099] Os oligonucleotídeos KO\_PMI\_FW e KO\_PMI\_RV que hibridizam nas regiões do DNA cromossomal adjacentes às regiões flanqueadoras de 1 kb ao gene-alvo (distantes cerca de 50 pb) e não hibridizam no vetor de deleção pNPTS138\_xanB foram projetados para a confirmação das deleções por PCR.

[0100] A PCR utilizando esse par de oligonucleotídeos permite diferenciar uma colônia mutante deletada de uma selvagem de acordo com os tamanhos dos produtos de PCR obtidos.

[0101] Pelo *screening* citado no item anterior, foram isoladas colônias com capacidade de crescimento no meio contendo sacarose e incapazes de crescer nos meios contendo canamicina. Cada uma dessas colônias foi ressuspensa em 49 µl de água destilada estéril e fervida por 5 minutos.

[0102] Cinco µl desse material foram utilizados como DNA molde juntamente com 5 µl de um *mix* de PCR preparado suficiente para 10 reações (descrito abaixo) de volume final de 10 µl:

- 150 pMol de oligonucleotídeo KO\_PMI\_FW
- 150 pMol de oligonucleotídeo KO\_PMI\_RV
- 0,8 mM de dNTP (0,2 mM de cada)
- 5x *Phusion Green High-Fidelity Buffer*
- 2 U de *Phusion High-Fidelity DNA polimerase*
- H<sub>2</sub>O MiliQ q.s.p 19 µl

[0103] As amplificações foram realizadas em um termociclador *C1000 Touch* (Bio-Rad), segundo as etapas descritas na Tabela 3, sendo o produto da amplificação analisado por eletroforese em gel de TAE agarose 0,8% (Figura 11).

TABELA 3 - ETAPAS DE PCR UTILIZADAS NA CONFIRMAÇÃO DA DELEÇÃO GÊNICA

<b>Etapa da PCR</b>	<b>Temperatura</b>	<b>Tempo</b>
Desnaturação inicial	98°C	10 minutos
Desnaturação	98°C	30 segundos (45x)
Hibridização	60°C	30 segundos (45x)
Extensão	72°C	4 minutos (45x)

Extensão final	72°C	10 minutos (45x)
----------------	------	---------------------

**Complementação gênica com o vetor pNPTS138 - Construção do vetor de complementação pNPTS138\_CxanB**

[0104] Foi realizada uma tentativa inédita em XAC de reinserção do gene *xanB* no mesmo locus cromossomal do qual foi originalmente deletado, utilizando o mutante de deleção e o vetor de deleção do gene *xanB* (pNPTS138\_xanB). A metodologia aplicada, também baseada na dupla recombinação homóloga entre o inserto no vetor (regiões flanqueadoras ao gene de interesse) e o DNA cromossomal bacteriano, foi, no entanto, a inversa à utilizada para a deleção no sentido de que o gene de interesse está agora inserido no vetor pNPTS138 e ausente no cromossomo. Assim, foi realizada a construção do vetor de complementação pela clonagem do gene *xanB* no vetor recombinante pNPTS138\_xanB (utilizado na deleção de *xanB*, o qual já contém as regiões flanqueadoras ao referido gene, Figura 9A).

[0105] A região codificante do gene *xanB* foi amplificada por PCR a partir dos oligonucleotídeos pNPTS\_cPMI\_FW e pNPTS\_cPMI\_RV (ver Tabela 1), ambos com sítios de restrição para enzima *EcoRI* (Figura 1), sob as mesmas condições utilizadas para o fragmento obtido para a construção do vetor de expressão pET41a\_GST\_PMI.

[0106] O produto de PCR foi clonado no vetor pJET 1.2, como descrito anteriormente. Em seguida, este fragmento gênico foi clonado no vetor de deleção pNPTS138, gerando o vetor recombinante pNPTS138\_xanB, previamente digerido com *EcoRI*, cujo único sítio de restrição se localiza entre os



dois fragmentos flanqueadores de 1 kb (Figura 9A), produzindo o vetor de complementação pNPTS138\_CxanB. A confirmação dessa construção foi realizada por análise de restrição por eletroforese em gel de TAE agarose 0,8% (Figura 12).

**Complementação gênica com o vetor pNPTS138 - Confirmação da orientação de clonagem de *xanB* no pNPTS138\_CxanB**

[0107] Como o fragmento gênico de *xanB* utilizado na clonagem que resultou no pNPTS138\_CxanB apresenta sítio de restrição para *EcoRI* em ambas as extremidades, seriam possíveis duas orientações de clonagem no vetor pNPTS138\_xanB.

[0108] Para certificar de que a clonagem ocorreu na orientação esperada, foi realizada uma PCR utilizando oligonucleotídeos que hibridizam nas seguintes regiões: codificante do gene *xanB*, 1 kb *upstream* e 1 kb *downstream* (Figura 12).

[0109] Por meio dessa estratégia, foi possível confirmar a orientação da clonagem realizada no vetor pNPTS138\_CxanB, uma vez que a PCR utilizando os oligonucleotídeos Up\_PMI\_FW e pNPTS\_cPMI\_RV resulta em um produto de PCR com 2,4 kb no caso da orientação correta da clonagem, e nenhum produto de PCR no caso de inserção de *xanB* em orientação não apropriada.

[0110] A combinação dos oligonucleotídeos Up\_PMI\_FW e pNPTS\_cPMI\_FW resulta na ausência de produtos de PCR em caso de orientação correta da clonagem, e um produto de PCR de 2,4 kb no caso da orientação inversa.

### **Complementação gênica com o vetor pNPTS138 - Seleção de possíveis complementantes e confirmação da complementação gênica**

[0111] Após a confirmação da correta clonagem no pNPTS138\_CxanB, células de XccΔxanB eletrocompetentes foram transformadas com o referido vetor de complementação, seguindo-se o mesmo procedimento utilizado na obtenção da linhagem mutante.

[0112] Após o *screening* das possíveis colônias complementantes (resistentes à canamicina, marcador de resistência do pNPTS138), foi utilizada a mesma metodologia de confirmação da deleção gênica (PCR) para atestar a obtenção da linhagem complementante pela reinserção do gene *xanB* no mutante (Figura 13).

### **Avaliação do gene *xanB* como essencial para a patogenicidade de XAC**

[0113] As linhagens XAC, XccΔxanB e XccΔCxanB foram avaliadas quanto à patogenicidade e virulência em plantas de lima ácida 'Galego' *Citrus aurantifolia* por meio de duas metodologias distintas - aspersão e infiltração. Para ambos os experimentos, as linhagens foram cultivadas em caldo LB e ressuspendidas em solução salina (0,9% de NaCl) de modo que a concentração final ficasse em  $10^6$  UFC.ml<sup>-1</sup>.

[0114] No ensaio de aspersão, 10 ml de cada suspensão bacteriana e de solução salina, para o controle positivo, foram pulverizados com o uso de um borrifador doméstico estéril em quadruplicata de plantas. As folhas foram registradas digitalmente após 28 dias da aspersão (Figura 14).

[0115] Para o ensaio de infiltração, as linhagens ou a solução salina (controle negativo) foram infiltradas na superfície abaxial de folhas de quatro plantas, de forma que todas as plantas receberam os quatro tratamentos diferentes, de forma a minimizar o efeito do hospedeiro sobre os resultados, quando da comparação entre os fenótipos apresentados. Após vinte dias das infiltrações, as folhas foram destacadas e registradas digitalmente para comparação da infecção (Figura 15).

[0116] Os ensaios de aspersão e infiltração apresentaram um padrão fenotípico em relação aos sintomas de cancro cítrico observados, a partir do qual foi possível constatar que a deleção do gene *xanB* fez com que a bactéria perdesse a patogenicidade. Além disso, a reversão do fenótipo observada nos ensaios com a linhagem complementante XccΔ*xanB* corrobora a hipótese de que o gene *xanB* é essencial para a patogenicidade de XAC.

[0117] Comparando-se os fenótipos apresentados pelas linhagens selvagem (XAC) e complementante (XccΔ*xanB*), observou-se semelhança nos graus de encharcamento e necrose do tecido vegetal, além da mesma frequência e intensidade das pústulas acastanhadas, sintomas característicos do cancro cítrico.

[0118] Embora sejam necessários estudos posteriores para a elucidação dos efeitos da deleção na expressão de fatores de virulência, a presente constatação de que o gene *xanB* é essencial para a patogenicidade de XAC possibilita a futura prospecção de ferramentas biotecnológicas de controle para doença, por meio da utilização de inibidores da PMI e, também, do desenvolvimento de variedades de citros que

neutralizem a ação de tal proteína no processo infeccioso.

**Envolvimento da *xanB* na motilidade, formação de biofilme e resistência à radiação UV**

[0119] A capacidade de motilidade foi avaliada nas linhagens XAC, XccΔ*xanB* e XccΔC*xanB* (Figura 16), segundo metodologia padronizada por (MALAMUD et al., 2011). Resumidamente, as linhagens foram cultivadas até que atingissem a DO<sub>595nm</sub> igual a 0,3 e 3 µl destes cultivos foram pipetados no centro das placas de Petri com LB 0,7%.

[0120] O diâmetro das colônias de XAC, XccΔ*xanB* e XccΔC*xanB* foi de 2,703 ± 0,016, 1,05 ± 0,02 e 2,716 ± 0,093 cm, respectivamente. Assim, em média a linhagem XccΔ*xanB* apresentou motilidade 61,15% e 61,34% inferior às linhagens selvagem XAC e complementante XccΔC*xanB*, respectivamente, indicando que o gene *xanB* está relacionado a esta importante característica do processo infeccioso.

[0121] A redução da motilidade pode ter contribuído para a perda de patogenicidade de XccΔ*xanB*. Os fitopatógenos utilizam mecanismos de motilidade para atingirem superfícies ricas em nutrientes e os tecidos vegetais. Especificamente para XAC, a motilidade está relacionada ao seu único flagelo polar, que a permite deslizar no ambiente epifítico e, também no apoplasto (GRAHAM et al., 2004; HARSHEY, 2003).

[0122] O envolvimento da PMI na síntese de GDP-D-manose, um precursor nas vias de produção de lipopolissacarídeos (LPS), também pode contribuir no entendimento da diminuição da motilidade de XccΔ*xanB*. Em XAC, a disrupção do gene *wxacO*, participante da síntese de LPS, afetou negativamente

a motilidade do fitopatôgeno (LI; WANG, 2011a).

[0123] Assim, embora seja necessária a comprovação de que Xcc $\Delta$ xanB apresenta defectividade na produção de LPS e na formação do flagelo polar, é possível sugerir que a redução de motilidade tenha relação com esses dois componentes atuantes na patogênese do cancro cítrico.

[0124] A formação de biofilme é uma capacidade que boa parte das bactérias tem de formar agregados em matrizes que permitem adesão a superfícies, possibilitando um modo de vida dinâmico às populações bacterianas (FLEMMING; WINGENDER, 2010).

[0125] As linhagens XAC, Xcc $\Delta$ xanB e Xcc $\Delta$ CxanB foram avaliadas quanto à formação de biofilme ao longo de 24, 48 e 72 horas, segundo metodologia padronizada por O'Toole e Kolter (1998). Após 24 horas, a linhagem Xcc $\Delta$ xanB apresentou formação de biofilme consideravelmente superior às linhagens selvagem XAC e complementante Xcc $\Delta$ CxanB (Figura 18). Em *Photorhabdus luminescens*, a PMI foi relacionada somente ao estágio de maturação do biofilme, e não ao estágio inicial do seu desenvolvimento (AMOS et al., 2011).

[0126] Em 48 e 72 horas houve uma inversão significativa desse padrão, com a linhagem Xcc $\Delta$ xanB apresentando formação de biofilme 4,5 e 14,5 vezes menor, respectivamente, em relação à linhagem selvagem XAC (Figura 17).

[0127] É importante ressaltar que o comportamento da linhagem complementante Xcc $\Delta$ CxanB foi extremamente semelhante ao de XAC, indicando que os efeitos depletivos da formação de biofilme em Xcc $\Delta$ xanB após 48 e 72 horas parecem ser, de fato, decorrentes da deleção do gene *xanB*.

[0128] Li e Wang (2011b) utilizaram uma abordagem de mutagênese em larga escala por meio de *transposons* e identificaram 92 genes de XAC relacionados à formação de biofilme, dentre os quais está incluído *xanB*, o que reforça os dados aqui obtidos utilizando uma metodologia diferente para a mutação do gene-alvo.

[0129] O processo de formação e maturação dos biofilmes tem a participação de EPS e LPS, substâncias poliméricas que atuam na coesão das células e, também, na comunicação intercelular (DONLAN, 2002).

[0130] Especificamente para XAC, estudos recentes têm demonstrado que a motilidade e a adesão são características necessárias aos estágios iniciais da formação de biofilme, enquanto maturação do biofilme depende de LPS e EPS, como a goma xantana (LI; WANG, 2011a, 2011b). Portanto, a participação do gene *xanB* nas vias de biossíntese de LPS e EPS também pode explicar, em parte, o fenótipo apresentado pela linhagem mutante deletada no referido gene.

[0131] Ademais, a formação de biofilme atua como uma estratégia para a sobrevivência epifítica de XAC (RIGANO et al., 2007), e a redução na formação de biofilme apresentada por *XccΔxanB* pode explicar a perda de patogenicidade desta linhagem após a deleção do gene *xanB*.

[0132] Os biofilmes atuam efetivamente na proteção contra estresses ambientais, como a resistência à radiação ultravioleta, e também na defesa contra os mecanismos responsivos do hospedeiro (KARATAN; WATNICK, 2009).

[0133] Assim, complementarmente, as linhagens XAC, *XccΔxanB* e *XccΔC<sub>xanB</sub>* foram avaliadas quanto à resistência à radiação ultravioleta (UV). Para isso, as linhagens foram

cultivadas até que atingissem a mesma densidade celular e foram expostas à luz ultravioleta por 15 minutos a 60 cm da fonte de luz. A porcentagem de sobrevivência de cada linhagem foi calculada em triplicata (Figura 18).

[0134] A linhagem XccΔxanB apresentou uma sobrevivência três vezes menor do que a linhagem selvagem XAC, um fenótipo que foi restaurado pela complementação gênica.

[0135] A maior suscetibilidade de XccΔxanB à radiação ultravioleta pode ser explicada pela atuação da PMI na via de biossíntese de xantana, uma vez que esse EPS tem sido relacionado à resistência a diversos estresses ambientais, incluindo a exposição à radiação UV (MEYER; BOGDANOVE, 2009).

[0136] Após os testes e análises da primeira etapa, para a seleção de compostos químicos do presente invento, de análise genética funcional para a comprovação de que a PMI é essencial para a XAC causar cancro cítrico, concluiu-se que:

- ORF XAC3580 codifica para uma fosfomanose isomerase;
- O gene *xanB* é essencial para a patogenicidade in vivo de XAC no hospedeiro *Citrus aurantifolia*, e está também envolvido na formação de biofilme, motilidade e resistência à radiação ultravioleta, aspectos relevantes aos processos de sobrevivência, disseminação, infecção e colonização do hospedeiro por XAC.
- *xanB* codifica uma proteína candidata altamente promissora para o desenho de moléculas e, conseqüentemente, para o desenvolvimento de novos produtos e formulações promissores (as) para o controle do cancro cítrico, inclusive pelo fato da proteína-alvo XanB estar possivelmente acessível na

superfície bacteriana;

- Embora a infecção causada por XAC esteja relacionada a inúmeros genes e proteínas, inclusive do sistema de secreção tipo III e a mecanismos bastante complexos, a ausência de um único gene, no caso *xanB*, foi capaz de comprometer toda a capacidade infectiva de XAC no hospedeiro cítrico, o que possivelmente pode indicar um envolvimento desse gene nas fases iniciais da infecção.

[0137] ETAPA II - Ensaaios e experimentos realizados *in silico* para planejamento de inibidores seletivos da PMI de XAC

#### **Novos inibidores de PMI**

[0138] Para a pesquisa de novos inibidores da PMI, foi usada a técnica de triagem virtual (ou *in silico*) por similaridade de forma e eletrostática (potencial), além de simulações de *docking* e análises preditivas de propriedades farmacocinéticas e toxicológicas (ADME/Tox), e seus potenciais modos de interação com a enzima foram analisados por inspeção visual, selecionando-se, por fim, os compostos mais promissores. Essas moléculas selecionadas foram adquiridas para serem testadas em ensaios de atividade *in vitro* e *in vivo*.

#### **Construção do modelo da fosfomanose isomerase para *Xanthomonas citri* subsp. *citri***

[0139] O primeiro passo para esse estudo foi a construção de um modelo da enzima Fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*, utilizando a técnica de 'Modelagem molecular por homologia estrutural', já que tal estrutura ainda não existe nos bancos de dados de



estruturas de proteínas como o PDB. A modelagem por homologia estrutural inclui três passos, a saber: (1) alinhamento das sequências das proteínas-molde com a proteína-problema (que é uma etapa crucial do processo), (2) construção do modelo e (3) validação do modelo.

[0140] A criação de um modelo tridimensional para uma proteína através da homologia estrutural consiste na comparação entre duas ou mais sequências de proteínas similares. Para que esse método seja aplicado é necessário que pelo menos três situações ocorram: (a) uma estrutura proteica seja experimentalmente resolvida; (b) compartilhar um valor de identidade sequencial acima de 25% com a proteína desejada; e (c) correto alinhamento entre as sequências envolvidas (FOMBY; CHERLIN, 2011; MARTÍ-RENOM et al., 2000; SCHWEDE et al., 2003).

[0141] A criação de um modelo é possível devido à relativa conservação que ocorre no enovelamento de uma determinada família de proteínas ao longo da evolução. Assim, caso a estrutura primária de duas proteínas sejam similares, o modelo tridimensional terá grandes chances de ser similar também (MARTÍ-RENOM et al., 2000; SANCHEZ; SAIL, [s.d.]).

[0142] Nos próximos tópicos são apresentadas as 4 etapas necessárias à modelagem de proteínas por homologia estrutural e aqui aplicadas para construção do modelo de PMI de XAC.

### **Pesquisa e alinhamento de sequências homólogas**

[0143] A etapa inicial do processo de predição de estruturas tridimensionais de proteínas é a busca e

identificação de estruturas de proteínas homólogas, a partir de suas sequências primárias, as quais possuam coordenadas cristalográficas depositadas no Protein Data Bank (PDB). Para maior confiabilidade do modelo gerado, é desejável que as identidades sequenciais compartilhadas entre a proteína-problema e as proteínas-molde sejam superiores a 25% (RODRIGUEZ; VRIEND, 1997).

[0144] Para a construção do modelo da PMI, inicialmente obtivemos a sequência de aminoácidos da mesma a partir do banco de dados GenBank (AAM38423.1) ("NCBI-National Center for Biotechnology Information", 2018). Em seguida, utilizando o algoritmo do servidor online BLASTp foi possível identificar proteínas com estruturas cristalinas disponíveis no PDB e que detêm identidades sequenciais acima de 25% compartilhadas com a PMI de interesse (proteína-problema ou sequência-modelo).

[0145] O algoritmo **BLAST** (*Basic Local Alignment Search Tool*) disponibilizado pelo NCBI (*National Center for Biotechnology*) é útil para se detectar similaridades entre sequências nucleotídicas ou de proteínas. Os programas da 'família' **BLAST** foram feitos para comparar: a) sequências de consultas de DNA; b) proteínas com bancos de dados de DNA; e c) proteínas em qualquer combinação (ALTSCHUL et al., 1990).

[0146] Uma das variantes dos algoritmos da família **BLAST** é o **BLASTp**, utilizado para comparações entre sequências de proteínas, através de uma busca nos bancos de dados por uma "palavra" de três "letras", as quais são três aminoácidos sequenciais que devem ser semelhantes (hits) à sequência-modelo. Os alinhamentos desses *hits* recebem uma pontuação

chamada de escore e, depois, a análise é expandida, nas duas extremidades da "palavra", para o restante da sequência-modelo, com o intuito de se obter um maior escore.

[0147] Outra pontuação analisada pelo **BLAST** é o *e-value*, com o intuito de determinar a aleatoriedade com que uma sequência homóloga foi obtida. Essa pontuação indica a probabilidade de se encontrar um alinhamento em outra sequência, aleatoriamente. Por isso, quanto menor o valor de *e-value*, mais estatisticamente significativo é o resultado obtido.

[0148] Por fim, outro resultado importante fornecido pelo **BLAST** é a porcentagem de identidade sequencial, a qual é responsável por mostrar o grau de correspondência entre duas subsequências. Nesse caso, valores entre 18-25% indicam uma potencial semelhança estrutural, enquanto valores iguais ou maiores a 25% indicam adicionalmente uma potencial semelhança de função.

[0149] Assim, neste trabalho, o algoritmo **BLASTp** foi útil para detectarmos 3 estruturas homólogas à PMI (com PDBs IDs 2CU2, 2X5S e 2QH5), com percentuais de identidade sequencial de 36% e 38%, além de altas resoluções cristalográficas (de 1,9 a 2,35 Å), tidas como adequadas para a modelagem por homologia (ou comparativa). Foi detectada também uma quarta estrutura, sendo essa a única contendo um composto ligante no sítio ativo (PDB ID 1H5R), a qual apresenta 29% de identidade compartilhada com a PMI de interesse.

[0150] As estruturas selecionadas (PDBs IDs 2QH5, 2X5S, 2CU2 e 1H5R) referem-se às seguintes enzimas: manose-6-

fosfato isomerase, GDP-manose pirofosforilase, glicose-1-fosfato timidiltransferase e uma timidililtransferase.

[0151] A manose-6-fosfato isomerase e a GDP-manose pirofosforilase realizam ambas atividades desempenhadas pela PMI de XAC, isto é, a isomerização entre manose-6-fosfato e frutose-6-fosfato e a transferência de manose-1-fosfato (M1P) para a guanosina trifosfato (GTP). Por fim, a glicose-1-fosfato timidiltransferase e a timidililtransferase realizam a transferência de timidila produzindo, por exemplo, a DTDP-glicose.

[0152] Devido ao fato do modelo a ser gerado ser de utilidade para os estudos de descoberta de novas moléculas baseados em estruturas, a escolha da estrutura cristalográfica para sua construção levou em consideração a presença de um ligante complexado no sítio da proteína. Assim, a escolha da proteína homóloga usada para prever uma estrutura PMI XAC foi a cadeia B da proteína com código PDB 1H5R (*Escherichia coli*, cepa K12, timidililtransferase). Como todas as estruturas apresentam uma identidade compartilhada com a PMI de interesse superior a 25%, elas portanto também compartilham com a estrutura de interesse uma potencial semelhança estrutural, além da mesma função.

[0153] Em posse das sequências homólogas à sequência em estudo, o próximo passo é determinado pelo alinhamento múltiplo das sequências e avaliação do modelo assim gerado. Para aproximar o maior número possível de opções, permitindo a construção de um modelo mais confiável, o alinhamento foi realizado com as cinco sequências obtidas.

### **Alinhamento múltiplo das sequências**

[0154] Após a obtenção das sequências do PDB, foi realizado o alinhamento múltiplo dessas sequências, para o qual foi utilizado o servidor *online* **Clustal Omega** (SIEVERS et al., 2011). Esse servidor é disponibilizado pela *European Molecular Biology Laboratory - European Bioinformatics Institute* (EMBL-EBO).

[0155] O **Clustal Omega** inclui cinco etapas principais para gerar o alinhamento de múltiplas sequências. O primeiro passo consiste em produzir um alinhamento aos pares (*pairwise*) para, em seguida, serem agrupados (BLACKSHIELDS et al., 2010). A terceira etapa é a construção de uma 'árvore guia', em que os dois grupos mais próximos são combinados e repetidos até que a árvore final possa ser avaliada. Então, na etapa final, o alinhamento de múltiplas sequências é, de fato, produzido (EDDY, 2011).

[0156] Um dado importante analisado durante o alinhamento das sequências diz respeito a quais alterações de resíduos de aminoácidos essas proteínas sofreram ao longo da Evolução, estimando que a função deveria ser preservada nos trechos onde a identidade (ou mesmo a similaridade) sequencial deveria ser maior.

[0157] A partir dos resultados obtidos com o **BLASTp**, foram feitas cinco propostas de alinhamentos. Quatro dos cinco alinhamentos propostos foram feitos entre a sequência da proteína de interesse e as sequências das proteínas provenientes do **BLASTp** individualmente. O outro alinhamento refere-se a um alinhamento múltiplo aliado ao refinamento que foi feito posterior e manualmente para refutar deleções e/ou inserções em elementos de estrutura secundária.

[0158] O alinhamento das sequências foi realizado

utilizando-se o servidor online **Clustal Omega**, aliado à sobreposição das quatro estruturas provenientes do PDB (PDBs IDs: 2QH5\_A, 2X5S\_A, 2CU2\_A e 1H5R\_B), feita por intermédio do programa **Discovery Studio**, de modo a verificar a qualidade do alinhamento.

[0159] A sobreposição estrutural nos permitiu visualizar os consensos de todas as estruturas secundárias das proteínas de referência, entre hélices-alfa e fitas de folha-beta. Devido a isso, pudemos analisar se as inserções ou deleções sugeridas pelo **Clustal Omega** estavam de acordo com o esperado, estruturalmente falando. Após a análise dos alinhamentos, juntamente com as sobreposições das estruturas de referência, obteve-se o alinhamento final.

[0160] Comparando ambos os alinhamentos, aquele obtido pelo **Clustal Omega** e aquele após a etapa de refinamento, pode-se observar algumas diferenças. Foram acrescentados alguns GAPs (ou deleções) entre as estruturas de PDBs IDs 1H5R, 2CU2, 2X5S e 2QH5, devido às sobreposições, porém, o mais evidente, ocorre na estrutura PDB ID 2QH5, em que é necessário um GAP entre os aminoácidos glicina (G11) e tirosina (Y12), tal como indicado na Figura 19.

[0161] A parte final da sequência da XAC (a partir do aminoácido 345) foi descartada devido ao não alinhamento com as outras sequências selecionadas e tal passo é possível pois essa região não está associada ao sítio ativo do modelo a ser gerado. Além disso, o alinhamento apresenta valores de 58% de identidade e 33% de similaridade no sítio ativo.

### **Construção do modelo e avaliação da qualidade**

[0162] Após a conclusão da etapa de alinhamento, a próxima etapa a ser conduzida é a construção do próprio modelo. Tal passo foi possível através do programa **MODELLER**, que gerou 5 hipóteses de modelo.

[0163] Além da construção do modelo, com o auxílio do **MODELLER** se pode realizar tarefas auxiliares adicionais, incluindo a atribuição de "voltas" e "loops", alinhamento de duas sequências de proteínas ou seus perfis (MARTÍ-RENOM et al., 2000), alinhamento múltiplo de sequências de proteínas e/ou estruturas (ESWAR et al., 2007), cálculo de árvores filogenéticas e modelagem de alças ou loops em estruturas de proteínas (FISER; ŠALI, 2003).

[0164] Para a construção de um modelo tridimensional, o **MODELLER** utiliza algumas restrições espaciais definidas entre as sequências escolhidas e o molde. Algumas dessas restrições são: distâncias e ângulos diedros, comprimento e ângulos de ligação e restrições criadas manualmente como, por exemplo, enovelamento da estrutura secundária, ligações cruzadas, entre outros.

[0165] Primeiramente, são analisadas as distâncias existentes entre os resíduos das sequências alinhadas, enquanto, que os comprimentos de ligações, ângulos de ligação e de diedro são provenientes do alinhamento com o modelo. Em seguida, utiliza-se uma função objetiva para determinar os comprimentos e ângulos de ligação entre átomos não-ligados. Após esses passos, surge um modelo por satisfação dessas restrições espaciais. Subsequente ao surgimento do modelo, pode ocorrer uma avaliação e modelagem das regiões de alças/loops através de um algoritmo, seguido pela otimização estrutural do modelo

gerado (ESWAR et al., 2007; SANTOS FILHO; BICCA DE ALENCASTRO, 2003).

[0166] Em posse dos modelos gerados, o próximo passo é a etapa de avaliação para determinar o modelo mais confiável e adequado para novos estudos. Para a etapa de avaliação foram utilizados os servidores online **WHATIF** (VRIEND, 1990) para a validação dos contatos atômicos e **SAVE** v5.0, o qual contém os programas **PROCHECK**, **VERIFY3D**, **ERRAT** e **PROVE**.

[0167] Na etapa de avaliação, é aconselhável avaliar a presença de contatos atômicos adequados e de maus contatos, e também comparar os resíduos do modelo com base nos parâmetros observados nas estruturas proteicas cristalográficas de alta resolução. Esses estudos foram desenvolvidos com o servidor on-line **WHATIF** e o programa **PROCHECK**, respectivamente, e os principais resultados são aqui discutidos.

[0168] O servidor *online* **WHATIF** analisa contatos atômicos dos resíduos de aminoácidos, gerando pontuações para cada deles e, adicionalmente, proporciona um ambiente flexível para exibir, manipular e analisar pequenas moléculas, proteínas, ácidos nucleicos e suas interações, para validação e correção de estruturas proteicas.

[0169] O **PROCHECK** (LASKOWSKI et al., 1993) realiza uma validação estereoquímica através da comparação entre os resíduos do modelo com base em parâmetros observados em estruturas de proteínas de alta-resolução cristalográfica. Um dos índices gerados por esse programa é o General Factor (G-Factor), o qual, este último, consiste em uma média da estereoquímica geral do modelo.

[0170] Dados os resultados obtidos pelo **WHATIF**, pode-se



observar que a hipótese avaliada obteve uma pontuação total de - 0,758, sendo que o valor de referência para este estudo é de - 0,813, que corresponde ao escore geral da estrutura cristalográfica (PDB ID 1H5R) aqui usada como referência.

[0171] Outro ponto de atenção são as regiões identificadas no Gráfico de Ramachandran. O Diagrama Ramachandran é dividido em quatro regiões: altamente favorável, favoravelmente permitida, generosamente permitida e proibida. Quando um aminoácido está na região proibida, isso significa que qualquer distância interatômica entre átomos não ligados é menor do que a distância ótima de Van der Waals correspondente a esses átomos, implicando em repulsão/impedimento estérico. No caso da hipótese escolhida, nenhum resíduo de aminoácido é encontrado na região proibida, favorecendo sua seleção como modelo.

[0172] Um ponto importante a considerar é o Fator G, uma pontuação média e final baseada nos parâmetros estereoquímicos observados de uma determinada estrutura, e quanto menor o seu valor, mais refinada é considerada a estrutura. No caso do modelo da PMI, as hipóteses apresentaram um Fator G de - 0,1, um valor que está dentro da região para uma proteína com uma resolução de 1,9 Å (referentã da PDB ID 1H5R).

[0173] Devem ser analisadas, nesta fase do processo, as conformações dos resíduos de aminoácidos que fazem parte do local de interação em estudo. Essa inspeção visual tem como objetivo analisar e, se necessário, modificar a conformação de determinados resíduos baseado em uma de rotâmeros. A

escolha do melhor rotâmero deve levar em consideração a direção dos resíduos da estrutura utilizada como referência para o desenvolvimento do modelo e a ausência de contato com os resíduos vizinhos.

[0174] Destacamos que, para o modelo PMI gerado, foi necessário modificar alguns resíduos e, com essa mudança de rotâmeros, o valor de **WHATIF** para a hipótese escolhida foi alterado para - 0,782, aproximando-se do valor da estrutura de referência (1H5R), além de estar entre - 1 e - 0,5 indicando que se trata de um bom modelo.

[0175] A sobreposição do modelo gerado com a estrutura-molde ou 'de referência' (PDB ID 1H5R) revela ainda dados importantes que aumentam a confiabilidade do modelo obtido. Pode ser observado, na região do sítio ativo, que os resíduos de aminoácidos descritos por *Zuccotti et al. 2001*) como responsáveis pela formação do complexo ligante-PMI, Gln147, Glu162 e Val173, permanecem alocados na mesma região, quando comparado o modelo com a estrutura de referência de PDB ID 1H5R. Porém, no modelo criado baseado na sequência da XAC, o aminoácido valina é substituído por uma asparagina, mas ambos (asparagina do modelo e valina da estrutura referência) apresentam-se na mesma região e sobre a mesma conformação de cadeia principal.

[0176] O Diagrama de Ramachandran, fornecido pelo **PROCHECK**, é uma ferramenta de análise importante pois indica prováveis erros na estrutura terciária dessas macromoléculas. O diagrama é dividido em regiões energeticamente favoráveis e desfavoráveis e mostra quais resíduos se encontram em tal regiões, permitindo uma avaliação de qualidade dos modelos teóricos ou

experimentais (RAMACHANDRAN et al., 1963).

[0177] O gráfico obtido para o modelo PMI gerado não indicou ângulos phi e psi desfavoráveis para os resíduos presentes no modelo, especialmente na região do local de interação. O modelo apresentou 91,6% de seus resíduos em regiões altamente favoráveis, 7,9% em regiões adicionais permitidas e 0,4% em regiões proibidas. Embora o resíduo ASN88 apareça em uma região "generosamente permitida", seus ângulos de torção foram "herdados" das estruturas cristalográficas utilizadas como referência.

[0178] Outras ferramentas importantes para analisar um modelo de proteína predito podem ser encontradas no pacote **SAVE v5**. O uso do **PROCHECK** permite a análise da planaridade das ligações peptídicas, contatos atômicos fracos, distorção tetraédrica dos carbonos alfa, energia das ligações hidrogênio e fator G. Para o modelo PMI gerado, todos os valores estão dentro dos valores esperados para estruturas de proteína de resolução de 1,9 Å (referindo-se à estrutura PDB ID 1H5R, usada aqui como referência para a modelagem) e a distorção tetraédrica dos carbonos alfa apresentou um valor mais acurado do que o da estrutura cristalográfica.

[0179] Os parâmetros apresentados anteriormente ajudam a avaliar a qualidade estereoquímica geral do modelo selecionado. Os gráficos mostram que o modelo obteve valores dentro do intervalo de referência **PROCHECK** e, ao analisar os valores apresentados pelo **WHATIF** para os aminoácidos pertencentes ao local de interação do PMI, nenhum apresentou valores abaixo de - 5, que, se existirem, devem ser revisados e seus respectivos rotâmeros alterados.

[0180] Além das ligações atômicas que ocorrem na cadeia peptídica, há também interações entre átomos não-ligados. A estrutura terciária de uma proteína é altamente dependente dessas interações e existem seis tipos de interações prováveis entre átomos não ligados e elas não ocorrem aleatoriamente. Assim, é possível, com uso do programa **ERRAT**, identificar regiões e fazer uma análise estatística dessas interações.

[0181] O **ERRAT** (COLOVOS; YEATES, 1993) analisa as interações intermoleculares entre átomos diferentes no modelo em comparação com estruturas altamente refinadas, gerando um gráfico de uma função de erro por uma janela formada por nove resíduos. De um modo geral, é um método mais eficaz que o **Perfil 3D** proposto por Bowie e colaboradores (BOWIE; LTCY; EISENBERG, 1990), porém não é melhor que o **PROCHECK**.

[0182] Foi possível identificar três regiões distintas (verde, amarela e vermelha) através do gráfico desenvolvido pelo **ERRAT** para o modelo. A região verde mostra que todos os aminoácidos estão dentro dos valores esperados para estruturas altamente resolvidas e refinadas. A região amarela corresponde aos aminoácidos que excedem o resultado esperado e podem ser rejeitados com 95% de confiança. Espera-se que uma boa estrutura proteica tenha 5% de sua estrutura nessa região. Finalmente, as regiões vermelhas são aquelas que podem ser rejeitadas com 99% de confiança. O modelo gerado para o PMI possui poucas regiões amarelas e apenas um único aminoácido na região vermelha (Ala167) no **ERRAT**. Nenhum aminoácido presente nessas regiões faz parte do sítio ativo; portanto, não haverá problemas para testes

futuros. Além desses dados, o fator de qualidade, valor relacionado às interações atômicas não ligadas de 85,5, indica uma estrutura de alta qualidade (resultado esperado para valores superiores a 50).

[0183] Outra análise interessante é verificar se um resíduo de aminoácido corresponde ao ambiente químico em que se encontra, ou seja, se há compatibilidade do modelo com sua sequência de aminoácidos. Essa análise pode ser realizada usando-se o **VERIFY 3D**.

[0184] O **VERIFY 3D** (LUTHY; BOWIE; EISENBERG, 1992) analisa a compatibilidade entre o modelo 3D e a sua sequência de aminoácidos, baseado na localização (alfa, beta, loop) e uma sequência de ambientes químicos (polar, apolar), utilizando uma matriz proposta por Bowie e colaboradores (BOWIE; LTCY; EISENBERG, 1990). Essa matriz permite uma análise de probabilidade de se encontrar os 20 aminoácidos em diferentes ambientes químicos, com base no tipo de aminoácido e natureza do ambiente químico. Em seguida, é realizado o alinhamento entre o modelo 3D e a sequência de aminoácidos; permitindo detectar as regiões mal resolvidas.

[0185] No estudo do PMI, pode-se observar que pelo menos 80% dos aminoácidos tiveram uma pontuação igual ou superior a 0,2 no perfil 3D-1D. Além disso, há uma região (Ala66 a Leu82) que caiu abaixo do limite, indicando uma provável dobra nessa região, no entanto, não é motivo de preocupação porque não está na região do sítio ativo do modelo.

[0186] Finalmente, como última análise, é possível verificar o volume dos átomos. O programa **PROVE** correlaciona o volume obtido para cada aminoácido da

proteína com os volumes de proteínas de alta resolução.

[0187] Como resultado, o **PROVE** fornece um valor de pontuação chamado Z-score que corresponde ao desvio estatístico programado de estruturas altamente resolvidas e refinadas.

[0188] No caso do modelo PMI em estudo, tanto o Z-score quanto o Z-score RMS estavam acima da região de interesse; sendo que o Z-score possui uma média de 0,274 e o Z-score RMS com valor de 1,539. Além disso, 4,6% dos resíduos analisados foram considerados *outliers*, indicando que mais atenção deve ser dada ao modelo. No entanto, juntamente com todas as outras análises realizadas, pode-se determinar que este é um modelo confiável para o PMI de XAC, e permite dar continuidade aos estudos.

#### **Validação do modelo da PMI para uso com simulações de *docking***

[0189] Após a criação do modelo, foi realizado um experimento para a validação do mesmo, utilizando-se os programas **GOLD** (SUITE, 2015) e **GLIDE** (FRIESNER et al., 2004). Definiu-se, a partir do ligante da estrutura cristalográfica original (1H5R), a melhor região do sítio ativo na qual seriam realizados os cálculos/simulações. Em seguida, foram feitas simulações de *docking* com o ligante da estrutura PDB ID 1H5R, para definir qual seria o melhor programa para os próximos estudos. Escolhido o programa, experimentos de *docking* foram realizados com a estrutura de PMI, utilizando-se inibidores reportados no *web servidor BindingDB* (CHEN; LIU; GILSON, 2001), um repositório de estruturas e dados de atividade focado na interação entre

proteínas e moléculas com potenciais atividades terapêuticas.

[0190] Após a criação do modelo, foi realizada uma simulação para a validação do mesmo, utilizando-se os programas **GOLD** e **GLIDE**. Definiu-se, a partir do ligante da estrutura cristalográfica original (de PDB ID 1H5R), a melhor região do sítio ativo na qual seriam realizados os cálculos.

[0191] As simulações de *docking* são cálculos que analisam a potencial afinidade entre o ligante e o sítio de ligação do receptor, propondo a pose (conformação + orientação), ou predição do modo de ligação com a maior afinidade entre ligante e receptor, estimada por uma pontuação calculada com base em uma função empírica associada ao método de mecânica molecular associado. A seguir serão descritos os métodos utilizados pelos programas **GOLD** (SUITE, 2015) e **GLIDE**.

[0192] O posicionamento dos ligantes no sítio do receptor resulta de um algoritmo adotado pelo programa escolhido para as simulações de *docking*. Os algoritmos mais usados são os genéticos, baseados em Monte Carlo ou busca sistemática para a flexibilização dos ligantes e cadeias laterais dos resíduos de aminoácidos da proteína. O **GOLD** utiliza o algoritmo genético enquanto o **GLIDE** utiliza um algoritmo de busca sistemática associada ao método de Monte Carlo.

[0193] No programa **GOLD**, o processo para gerar possibilidades conformacionais para a molécula ligante (e também para seletas cadeias laterais de resíduos de aminoácidos) se inicia gerando populações contendo

conformações aleatórias de um dado ligante, em que cada indivíduo dessa população é codificado e chamado de "cromossomo". Os cromossomos possuem, então, as informações dos átomos dos ligantes como possibilidades de fazer interações. Além disso, eles também possuem informações sobre liberdades translacionais, rotacionais e conformacionais, chamadas de "genes".

[0194] Com a ideia de gerar conformações mais "evoluídas", os cromossomos utilizam mecanismos chamados operadores genéticos, no qual haverá uma análise de mutações e recombinações. Assim, diversas etapas como essa ocorrem até que seja obtida a população final (SUITE, 2015), evoluída de seus "ancestrais", em processo tendenciado à maior afinidade com o sítio receptor. Após a criação da população final, cada cromossomo/ligante é ancorado diversas vezes no sítio, com diferentes orientações, e cada pose é pontuada através de uma função chamada *fitness score*.

[0195] O primeiro passo do método utilizado pelo **GLIDE** consiste na construção de um grid no sítio de interação do receptor, determinando sua forma e propriedades. Nesse *grid* são feitas análises da presença de átomos existentes, fornecendo informações sobre os tipos de interações presentes e a pontuação para a pose do ligante. Em seguida, ocorre a geração de confôrmeros, os quais serão analisados nessa região.

[0196] A geração desses confôrmeros pelo programa **GLIDE** é baseado nas conformações de cada ligante, a qual varia conforme do número de ligações livres. A busca pelos melhores ligantes para um determinado receptor começa com a



procura por regiões em que podem ocorrer interações, tanto no ligante quanto no receptor. Se ocorrer uma correspondência entre a conformação e o sítio de ancoragem, essa conformação segue para a próxima etapa.

[0197] A segunda etapa analisa os átomos do ligante, em relação à cavidade do receptor, de forma a evitar repulsões/impedimentos 'estéricos'. Além disso, são feitas avaliações em relação ao diâmetro do ligante e suas rotações. O ligante recebe uma pontuação e caso ela seja 'satisfatória' de acordo com os critérios estabelecidos, ele segue em frente. Em seguida, é calculada uma nova pontuação, chamada de "gananciosa", que irá avaliar a posição do ligante ao receptor quando esse se move em torno de 1 Å nas coordenadas x, y e z. Esse movimento gera uma nova pontuação chamada de *ChemScore*, a qual penaliza contatos estéricos além de reconhecer diversos tipos de interações como hidrofóbica, metálicas e ligações de hidrogênio. Após isso, o ligante é novamente movimentado nas coordenadas e sua pontuação é recalculada. Nessa segunda etapa ocorre a redução do número de confôrmeros. Por fim, no terceiro passo, o algoritmo Monte Carlos explora a minimização de energia dos ligantes, pontuando as poses novamente através da função *GlideScore*.

[0198] O primeiro passo para a etapa de validação do modelo consistiu na escolha do programa para as simulações de *docking*, o **GOLD** ou o **GLIDE**. Para tanto foi feito um *redocking*, para ver qual programa conseguia posicionar melhor o ligante cristalográfico, glicose-1-fosfato (G1P), em relação à sua posição original na estrutura cristalográfica PDB ID 1H5R. Foram executadas 10 simulações

para cada programa. Como pode ser observado na Figura 20, o **GOLD** foi o programa com o qual se conseguiu um melhor posicionamento, tendo sido ele selecionado para futuras simulações de *docking*.

[0199] O sucesso de um programa na previsão de uma pose de ligante é geralmente medido pelo desvio médio quadrático, ou RMSD (Root Mean Square Deviation), com o qual é possível se determinar o grau de similaridade entre duas estruturas (COLE et al., 2005). Esse dado é calculado entre a posição do ligante observada experimentalmente e a posição do ligante prevista pelo programa, com base nas distâncias interatômicas de seus átomos, sendo o valor obtido através da seguinte expressão matemática (1):

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \|r_i^A - r_i^B\|^2} \quad (1)$$

onde N é o número total de átomos pareados entre as duas estruturas,  $r_i$  é o vetor posição do átomo i, A é o modelo gerado e B é a estrutura molde.

[0200] A estatística mais usada é a que os modelos gerados deveriam obter um RMSD < 2 Å e, embora seja comumente utilizada, é um limite empírico e arbitrário (GOHLKE; HENDLICH; KLEBE, 2000; KONTOYIANNI; MCCLELLAN; SOKOL, 2004; KRAMER; RAREY; LENGAUER, 1999). Por outro lado, alguns autores como Vieth e Bursulaya (BURSULAYA et al., 2003; VIETH et al., 1998), consideram valores de RMSD entre 2 Å e 3 Å como sucessos parciais. Para o modelo gerado da PMI de XAC foi encontrado um RMSD = 2,5 Å que, de acordo com os autores citados acima, seria considerado como um sucesso parcial.

[0201] Uma vez tendo sido selecionado e refinado o modelo final, avaliamos sua eficiência em estudos de triagem virtual. Para essa avaliação, foram escolhidos 30 inibidores conhecidos da PMI, depositados e provenientes do servidor online **BindingDB**, contendo valores descritos de potência (IC<sub>50</sub>), reportados na literatura. O objetivo dessa etapa era avaliar a capacidade que o modelo possui de recuperar os ligantes sabidamente ativos e ordená-los de forma coerente em relação ao seus IC<sub>50</sub>, com base em função de pontuação (**GOLD Fitness**) proveniente do programa **GOLD**.

[0202] Para que o objetivo fosse alcançado, primeiramente era necessária a definição do sítio de *docking*, a qual foi obtida através do posicionamento do ligante proveniente da PDB ID 1H5R em relação ao modelo. Com essa posição, foi possível obter-se uma esfera com as seguintes coordenadas espaciais: x = 23,977, y = 15,659, z = 43,886, com uma esfera de raio de 8 Å (região onde seriam realizadas as simulações de *docking*). Os resultados dessa etapa são mostrados na Tabela 4, em que os inibidores mais potentes foram os inibidores 1, 2, 3, 5, 7, 8, 9, 10, 12, 13, 16, 18, 19, 21 e 30.

TABELA 4 - RESULTADO DA RECUPERAÇÃO DE INIBIDORES (15 MAIS POTENTES E 15 MENOS POTENTES) DO BINDINGDB PELA FOSFOMANOSE ISOMERASE DE XANTHOMONAS CITRI SUBSP. CITRI.

Número do Inibidor	Nome PCid	Gold Fitness	IC50
1	2998795	62.216	1.30E+03
2	1256813	60.097	1.00E+03
3	5771379	59.096	1.00E+03
4	25199560	57.191	>1,00E+5

5	4089709	56.362	707
6	25199559	55.438	>1,00E+5
7	2311082	54.774	270
8	2248499	48.79	829
9	1489253	48.566	1.23E+03
10	3093671	45.896	1.27E+03
11	4120895	44.867	>1,00E+5
12	22416235	44.529	1.25E+03
13	4917291	44.379	709
14	387196	44.267	>1,00E+5
15	25241622	43.919	>1,00E+5
16	1510389	43.654	985
17	3311395	41.868	>1,00E+5
18	25181243	41.807	1.03E+03
19	1519965	41.795	1.00E+03
20	4879298	41.788	>1,00E+5
21	3194	41.062	190
22	25181239	36.225	>1,00E+5
23	25181212	35.122	>1,00E+5
24	22765927	35.095	>1,00E+5
25	25241626	33.333	>1,00E+5
26	6112173	32.755	>1,00E+5
27	25241624	32.246	>1,00E+5
28	6033619	27.892	>1,00E+5
29	5951502	26.357	>1,00E+5
30	25199533	22.889	1.07E+03

[0203] Observando-se os dados obtidos pelo estudo de triagem virtual, o próximo passo consistiu na obtenção de curvas ROC (*Receiver Operating Characteristic*), de predição e de enriquecimento para melhor entendimento dos resultados. Para a criação desses gráficos, foi utilizado o *web servidor Screening Explorer* (EMPEREUR-MOT; ZAGURY; MONTES, 2016) com os parâmetros selecionados como padrões.

[0204] Em estudos de triagem virtual, a análise dos dados é feita com base em uma curva ROC. Trata-se de um gráfico dividido em frações positivas verdadeiras (FVP) localizadas no eixo y e frações falso-positivas (FFP) localizadas no eixo x. Um par de FVP/FFP é representado por cada ponto dentro da curva ROC e corresponde a uma fração específica de um conjunto de moléculas. As curvas ROC resumem a capacidade geral de um método para distinguir entre compostos ativos e inativos (EMPEREUR-MOT et al., 2015; EMPEREUR-MOT; ZAGURY; MONTES, 2016).

[0205] Uma função de pontuação com perfeita discriminação é aquela em que no canto superior esquerdo do gráfico, temos uma sensibilidade perfeita (FVP = 1) e uma especificidade perfeita (FFP = 0). A sensibilidade corresponde a capacidade do sistema em predizer corretamente quais moléculas possuem certas características (proporção de positivos verdadeiras); enquanto especificidade corresponde em se predizer a ausência de características em algumas moléculas (proporção de falsos-positivos) (BERTRAND; TRIBALLEAU, 2005).

[0206] A linha diagonal presente no gráfico corresponde a experimentos em que a função de pontuação não seria discriminada. Uma curva entre a linha diagonal e a parte

superior esquerda do gráfico é considerada boa, sendo que quanto maior a distância da linha diagonal, melhor será o sistema. De forma qualitativa, quanto mais próximo à esquerda do gráfico estiver a curva, maior será a precisão geral do estudo (BERTRAND; TRIBALLEAU, 2005; EMPEREUR-MOT et al., 2015; EMPEREUR-MOT; ZAGURY; MONTES, 2016). Abaixo, na Figura 21, é mostrado o gráfico de curva ROC obtido para o estudo realizado com os 30 inibidores do **BindingDB** e o modelo de PMI XAC gerado. Nele é possível observar que grande parte do gráfico se encontra distante da linha diagonal, mostrando que o estudo apresenta uma boa precisão, conseguindo diferenciar os compostos mais ativos dos menos ativos.

[0207] Caso se torne difícil analisar uma curva ROC, é possível utilizar uma medida padrão chamada de área sob a curva (AUC). O AUC é um número que varia de 0 a 1, obtido por métodos de integração numérica e mostra o desempenho geral do estudo, isto é, a probabilidade do modelo classificar um dado positivo como negativo e vice-versa. Teoricamente, quanto maior o valor obtido para AUC, melhor é o sistema utilizado (EMPEREUR-MOT; ZAGURY; MONTES, 2016). O valor de AUC ROC obtido para esse estudo foi de 0.760, um bom valor para o estudo realizado já que se encontra próximo a 1 (valor ideal).

[0208] Um outro dado obtido com o *web servidor* é a curva de previsão, a qual quantifica a associação entre as variações de pontuação e detecção dos compostos ativos. No eixo y do gráfico é mostrado a probabilidade de atividade para cada composto presente no conjunto de dados e no eixo x, esses compostos são mostrados (EMPEREUR-MOT et al.,

2015). Consequentemente, a curva de previsão permite detectar a variação emitidas por uma função de pontuação na detecção dos compostos ativos, isto é, se as variações de probabilidade de atividade são importantes o suficiente para induzir a seleção de um limiar para triagens virtuais (EMPEREUR-MOT et al., 2015). A curva de previsão para esse estudo se encontra acima da linha limite, tornando-se um bom resultado e mostrando que o modelo foi capaz de realizar uma discriminação entre os compostos utilizados.

[0209] Um dado calculado a partir da curva de previsão é o ganho total padronizado (TG) (BURA; GASTWIRTH, 2001), que resume a discriminação dos compostos ativos em relação à variação das pontuações ao longo do conjunto de dados das moléculas (EMPEREUR-MOT et al., 2015).

[0210] A TG possui seu valor entre 0 e 1, sendo que sua análise deve estar associada ao valor obtido para a AUC. Valores de TG acima de 0,25 combinados com uma AUC ROC acima de 0,5 significa que as variações de pontuação são relevantes na discriminação dos compostos ativos; já valores de TG acima de 0,4 associados com uma AUC ROC acima de 0,5 indica que o método de triagem utilizado teve um bom desempenho e seria reproduzível em condições experimentais semelhantes (EMPEREUR-MOT; ZAGURY; MONTES, 2016).

[0211] No nosso estudo, o valor de TG obtido foi de 0,312 que, associado a um AUC ROC de 0,760, mostra que o sistema foi capaz de discriminar os compostos ativos para tal modelo proteico.

[0212] Por fim, um último dado obtido é a curva de enriquecimento, a qual permite avaliar a recuperação precoce de compostos ativos presentes em estudos de triagem

virtual. O gráfico em escala logarítmica mostra no eixo y as FVP e no eixo x o conjunto de dados (TRUCHON; BAYLY, 2007).

[0213] Assim, como as curvas ROC, as de enriquecimento possuem uma linha diagonal que inicia no canto inferior esquerdo, subindo até o canto superior direito e que representa uma distribuição aleatória. Quanto mais distante a curva gerada estiver da linha diagonal, melhor é o estudo realizado (BERTRAND; TRIBALLEAU, 2005).

[0214] Embora as curvas de enriquecimento sejam mais fáceis de traçar, elas possuem uma grande desvantagem. Essas curvas dependem diretamente da proporção de moléculas ativas presentes no conjunto; portanto, quando a proporção desses ativos aumenta, a curva de enriquecimento ideal se aproxima da linha de classificação aleatória (BERTRAND; TRIBALLEAU, 2005). Além disso, o espaço é limitado pelas curvas ideal e aleatória, prejudicando comparações com outros dados. A curva de enriquecimento do estudo em questão encontra-se acima da linha limite.

[0215] Um dado relacionado à curva de enriquecimento é o fator de enriquecimento (EF), o qual corresponde à fração de compostos ativos obtidos quando uma determinada porcentagem de compostos inativos é encontrada (TRUCHON; BAYLY, 2007). Esse dado possui uma alta variabilidade, e seu valor máximo depende diretamente da proporção de moléculas ativas no conjunto (TRUCHON; BAYLY, 2007) e do valor de compostos inativos considerados (PICCIRILLO; DO AMARAL, 2018). Nesse estudo, foi considerado a porcentagem de 10% dos compostos inativos e o valor obtido para EF foi de 2,00, mostrando que, para 10 compostos inativos



recuperados, 2 deles são compostos ativos.

[0216] Além dessas curvas, temos alguns outros dados para levar em consideração: discriminação reforçada por Boltzmann do ROC (BEDROC) (MUEGGE; ENYEDY, 2004) e o reforço inicial robusto (RIE) (CORNELL, 2006). Truchon & Bayly (2007) discutiram vários métodos para abordar o reconhecimento precoce, como o BEDROC e o RIE. Ambas as métricas são utilizadas para quantificar o reconhecimento precoce dos compostos ativos utilizados em um estudo de triagem virtual (EMPEREUR-MOT; ZAGURY; MONTES, 2016). O BEDROC é limitado por um valor de 0 a 1 e trata da probabilidade de um ativo ser classificado antes de um composto selecionado aleatoriamente. Por outro lado, o RIE, desenvolvido por Sheridan (SHERIDAN et al., 2001), utiliza a ponderação exponencial, atribuindo um maior peso para os ativos reconhecidos precocemente (ZHAO et al., 2009).

[0217] Uma das desvantagens do RIE é que o seu valor mínimo e máximo depende diretamente da quantidade de compostos ativos e a quantidade de compostos totais presente nos conjuntos. O valor obtido para essa métrica foi 1,832. Por outro lado, o valor obtido pra BEDROC foi 0,916 encontra-se muito próximo a 1, mostrando que o modelo é capaz de separar os compostos ativos e inativos.

[0218] Com isso, através desse resultado de validação e daqueles de avaliação mostrados anteriormente, podemos afirmar que o modelo para a PMI XAC está apto para ser utilizado nos posteriores estudos de triagem virtual.

### **Triagem virtual de novos inibidores da Fosfomanose Isomerase**

[0219] Acessando-se o *web servidor* **BindingDB**, um total de 959 hits são descritos como inibidores de PMI, sendo que o mais potente possui IC<sub>50</sub> de somente 4 µM. Sendo assim, torna-se desejável a descoberta e o desenvolvimento de novas classes de compostos e que possam inibir mais efetivamente a enzima aqui investigada. O objetivo é realizar experimentos de triagem virtual (por *docking* e similaridade) com a PMI em diferentes bases de dados de compostos disponíveis comercialmente. As bases de dados de compostos incluem: ZINC (subcoleções Natural Stock, Drug-database e Drug-like), BindingDB subdivisão Drugs FDA, ChemBridge (subcoleções DIVERSet™-EXPEXPRESS-PickTM Collection(DIVERSet™-EXP) e DIVERSet Core Library (DIVERSet™-CL)), Maybridge subdivisão Screening Collection, Princeton e IBS (subdivisão Natural e Synthetic), totalizando uma quantidade próxima a 19 milhões de compostos.

[0220] Algumas dessas bases contêm os compostos em formato *spatial data format* (sdf), os quais já foram reescritos em formato mais apropriado, mol2, contendo cargas atômicas, tipos de átomos, ordens de ligação e estruturas tridimensionais bem definidas e mais adequadas para as simulações de triagem virtual e propósitos gerais de modelagem molecular. Também foram geradas novas bases/bibliotecas de confôrmeros, a partir dessas, utilizando-se o programa **OMEGA** (BOSTRÖM; GREENWOOD; GOTTFRIES, 2003).

#### **Geração de confôrmeros (com OMEGA)**

[0221] Para a geração de confôrmeros, o programa **OMEGA**

da empresa **OpenEye** foi aqui utilizado. Esse programa realiza a geração de confôrmeros em três dimensões com baixo desprendimento computacional e um alto efeito no alcance da conformação bioativa entre os confôrmeros gerados.

[0222] No presente estudo, os parâmetros utilizados no **OMEGA** para a geração de confôrmeros foram: até 300/molécula, com uma energia de tensão de até 9 kcal/mol e um desvio médio quadrático de 0.6 Å, segundo protocolo otimizado por Carlos H.T.P. Silva (DA SILVA; TAFT, 2017).

[0223] O primeiro passo foi preparar as 6 bases de dados, descritas anteriormente, com o programa **OMEGA**. Esse passo tem o intuito de corrigir e criar diversas conformações das moléculas presentes nas bases de dados. Esse programa realiza a geração de confôrmeros em três dimensões, com baixo desprendimento computacional e um alto efeito no alcance da conformação bioativa entre os confôrmeros gerados. O **OMEGA** obtém os confôrmeros em três passos. O primeiro passo, para a obtenção dos confôrmeros, consiste ao acesso à biblioteca de fragmentos, gerada através do fragmento de moléculas presente em um extenso banco de dados, para a construção da estrutura em três dimensões (3D). Os fragmentos são ligados com base em valores de distância de ligações, sendo ajustados posteriormente e minimizados por um campo de força ajustado.

[0224] Em seguida, todas as torções rotativas são numeradas de acordo com valores de um banco dados presente no programa. As rotações são ajustadas com base em ângulos e energias torcionais derivadas de um conjunto de

estruturas cristalinas. Novas torções são geradas no ligante, prestando atenção a choques e duplicações, com a finalidade de gerar uma grande quantidade de confôrmeros.

[0225] Por fim, na última fase, ocorre a análise da estrutura em relação aos seus valores energéticos e geometria. Nessa etapa, a molécula pertencente ao mínimo global tem o seu RMSD comparado a outras, eliminando aquelas não pertencentes à parametrização fornecida pelo usuário (HAWKINS et al., 2010; HAWKINS; NICHOLLS, 2012).

#### **Seleção de inibidores baseada em ligantes (com ROCS e EON)**

[0226] Após a preparação das bases de dados, um segundo tipo de triagem virtual foi ainda realizado, o qual é baseado em diferentes tipos de similaridade: de forma e eletrostática. Ambos os programas utilizados nessa etapa, **ROCS** e **EON**, da **OpenEye**, tem revelado compostos promissores para o tratamento de várias doenças (NAYLOR et al., 2009).

[0227] O programa **ROCS** realiza uma triagem virtual por similaridade de forma, através da comparação entre os volumes e conformação de dois compostos, a molécula de referência (glicose-1-fosfato (G1P)) e as moléculas presentes em bancos de dados. Os cálculos dos volumes são feitos através da superfície exposta de esferas de inserção colocadas na molécula referência.

[0228] Após o **ROCS**, uma outra triagem de similaridade foi feita comparando os mapas de potencial eletrostático das moléculas. O programa utilizado nessa etapa foi o **EON**, sendo que seus arquivos de entrada foram os provenientes do **ROCS**. Com esse programa se realiza a comparação e classificação dos compostos bases em seus mapas de

superfície eletrostática. Os parâmetros adotados para ambos programas não foram alterados.

**Seleção de inibidores baseada em estrutura (com GOLD) e predição das propriedades farmacocinéticas e toxicológicas (com QIKPROP e DEREK)**

[0229] Os programas de *docking* analisam a afinidade ligante-proteína e, para essa etapa do estudo, foi utilizado o programa **GOLD**. Após os experimentos de triagem virtual realizados, somente as moléculas que passaram pelo crivo de propriedades adequadas a um potencial defensivo agrícola avançaram para a próxima etapa, avaliando-se, principalmente, a potencial toxicidade dos compostos frente ao organismo humano. As moléculas foram avaliadas com respeito às suas propriedades ADME/Tox, utilizando-se os programas **QikProp** (SCHRODINGER, 2020) (da empresa Schrodinger) e **DEREK** (MARCHANT; BRIGGS; LONG, 2008) (da empresa Lhasa), respectivamente.

[0230] Os compostos "sobreviventes" à predição de propriedades toxicológicas, foram submetidos a análises farmacocinéticas realizadas com o programa **QikProp**, e os descritores analisados foram: log CACO2, log MDCK, porcentagem de absorção oral e qualidade do modelo de absorção oral humana. Foram rejeitados todos os compostos que tiveram valores de log CACO2 e log MDCK maiores ou iguais a 500, potencial de absorção oral maior ou igual a 80% e qualidade do modelo de absorção humana igual a 'alta'. Na análise toxicológica realizada com o programa **DEREK**, os compostos rejeitados foram aqueles que apresentaram algum alerta estrutural de toxicidade segundo

*Custom Prediction e Lhasa Prediction.*

[0231] Para a obtenção dos mais promissores inibidores possíveis para a PMI de XAC, foi utilizada uma estratégia que se inicia com uma triagem virtual por similaridade, utilizando-se os programas **ROCS** e **EON**, seguida pelo *docking* e pelas predições ADME/Tox. O último passo correspondeu à análise visual da interação dos compostos com o receptor.

[0232] Nessa estratégia, a primeira triagem feita foi em relação à similaridade por forma entre a molécula referência e as bases de dados. Essa triagem foi realizada com o programa **ROCS**, e foram obtidos os 2000 melhores (ou mais promissores) resultados oriundos das bases de dados Chembridge CL, Chembridge EXP, DrugDatabase, DrugLike, FDA Binding, Maybridge e Natural Stock. Os compostos selecionados nessa etapa seguiram para os estudos de triagem baseada em similaridade de potencial eletrostático, com o programa **EON**, no qual foram selecionados os 1000 compostos mais bem classificados.

[0233] Em seguida, os arquivos obtidos na etapa anterior foram submetidos ao *docking*, utilizando-se o centroide de  $x = 23.977$ ,  $y = 15.659$  e  $z = 43.886$ , e um raio de esfera de 6 Å e mínimo de 60 (de fitness). O tamanho da esfera teve que ser modificado devido a problemas encontrados para alguns ligantes em relação a sua torção e aumento do tamanho das ligações. Os arquivos finais de cada base de dados saíram com 500 compostos e foram utilizados em estudos de predições ADME/Tox. Em seguida, esses compostos passaram por uma análise visual para verificar quais possuíam uma maior interação com o modelo da PMI de XAC, em potencial. Abaixo (Tabela 5) são mostrados quantos compostos

"sobreviveram" a todas as etapas.

TABELA 5 - RELAÇÃO ENTRE OS COMPOSTOS "SOBREVIVENTES" DAS BASES DE DADOS APÓS ESTUDOS DE TRIAGEM VIRTUAL POR SIMILARIDADE (COM **ROCS** E **EON**), DE DOCKING, PREDIÇÕES ADME/TOX (COM **DEREK** E **QikProp**) E INSPEÇÃO VISUAL, UTILIZANDO-SE A SEGUNDA ESTRATÉGIA.

BASE DE DADOS	ROCS	EON	DOCKING	QikProp	INSPEÇÃO VISUAL	
<b>Chembridge CL</b>	2000	1000	500	8	0	0
<b>Chembridge EXP</b>	2000	1000	500	1	0	0
<b>DrugDatabase</b>	2000	1000	500	13	13	0
<b>DrugLike</b>	2000	1000	500	265	83	38
<b>FDA Binding</b>	2000	1000	500	31	18	1
<b>Maybridge</b>	2000	1000	0	0	0	0
<b>Natural Stock</b>	2000	1000	500	3	3	1

[0234] Após as análises farmacocinéticas e toxicológicas *in silico*, os compostos de todas as bases foram submetidos a uma análise visual, a qual possui o intuito de verificar uma maior interação com o modelo da PMI XAC através da quantidade de interações químicas. As análises dos resultados gerados serviram como critério final para a seleção de compostos, sendo um total de 45 inibidores. Ao final dos estudos, a relação dos compostos sobreviventes em cada base de dado é mostrada nas tabelas acima.

[0235] Os resultados das simulações de *docking* nos permitiram visualizar como os compostos se comportam no sítio ativo do modelo. Sendo assim, a seleção final se baseou prioritariamente na possibilidade dos compostos interagirem com os resíduos Arg173 Asn192, Asp249, Asp270 e Gly155; sendo as interações com os aminoácidos arginina (173) e ácido aspártico (249) as mais importantes.

[0236] Os resultados obtidos com o programa GOLD indicaram que alguns compostos podem interagir de forma satisfatória no sítio ativo do modelo e, dessa forma, elegemos estes compostos como os mais promissores. Foram selecionados 1 composto oriundo da base de dados Natural Stock, 6 compostos da FDA Binding e 38 compostos da DrugLike.

[0237] De todos os compostos selecionados, alguns foram comprados e em seguida serão submetidos a ensaios *in vitro* e *in vivo*, para a comprovação da ação inibidora da enzima PMI de *Xanthomonas citri* subsp. *citri*. Doze inibidores foram selecionados para compra e quatro deles foram utilizados em ensaios biológicos: **DL03874928**, **DL03999102**, **DL04091967** e **DL35935889**. Esses inibidores que foram utilizados nos ensaios biológicos foram selecionados com base nas interações mais fortes observadas com aminoácidos interessantes que compõem o sítio ativo e quantidade dessas interações; além de serem moléculas com perfis polares em uma extremidade e hidrofóbicos em outra, e não muito rígidas.

[0238] A seguir são mostradas as estruturas dos compostos utilizados nos ensaios biológicos (Figura 22) e a estrutura de um dos inibidores (**DL04091967**) docado no sitio



ativo do modelo da PMI de *Xanthomonas citri* subsp *citri* (XAC) (Figura 23).

### **Visualização Gráfica**

[0239] Os programas utilizados para a visualização gráfica foram o **Discovery Studio** (DASSAULT SYSTEMES BIOVIA, 2015) da empresa Biovia bem como o **Maestro** (SCHRODINGER, 2018), da empresa Schroedinger.

[0240] O **Discovery Studio** foi empregado na visualização da sobreposição dos modelos obtidos pela modelagem de proteínas por homologia estrutural. Já o **Maestro** foi utilizado na análise visual dos resultados obtidos a partir dos estudos de simulações de *docking* e, conseqüentemente, para a seleção das moléculas finais.

### **Construção do modelo da Fosfomanose isomerase (PMI) de *Homo sapiens*.**

[0241] Após a construção do modelo para a PMI de XAC foi decidido se realizar um estudo em relação também à proteína PMI de *Homo sapiens*, para fins de comparação. Os passos utilizados para a construção do modelo foram iguais aos utilizados para a XAC. O primeiro passo foi realizar o alinhamento entre as sequências de aminoácidos de ambas proteínas, com o intuito de analisar as porcentagens de identidade e similaridade. Elas compartilham uma identidade de 7% e uma similaridade de 11%.

[0242] Após essa etapa, foi criado o modelo para a PMI humana para comprovar que se trata de estruturas totalmente diferentes. O modelo também foi construído por modelagem de homologia estrutural, utilizando todos os mesmos passos que

foram utilizados anteriormente, isto é, alinhamento, construção e avaliação do modelo e avaliação do modelo.

#### **ETAPA DE IDENTIFICAÇÃO DA SEQUÊNCIA DE APROVAÇÃO**

[0243] A partir dos resultados obtidos com o **BLASTp** foi feito o alinhamento das sequências através do servidor online **Clustal Omega**. A sobreposição estrutural das duas estruturas provenientes do PDB: PDBs Ids 5NW7 e 3H1M, feita por intermédio do programa **Maestro** e permitiu visualizar os consensos de todas as estruturas secundárias das proteínas de referência, entre hélices-alfa e fitas de folha-beta. Devido a isso, pudemos analisar se as inserções ou deleções sugeridas pelo **Clustal Omega** estavam de acordo com o esperado. Após a análise dos alinhamentos juntamente com as sobreposições das estruturas de referência, obteve-se o alinhamento final indicado na Figura 24.

#### **Etapa de construção e avaliação do modelo**

[0244] Após a etapa de alinhamento ter sido concluída, a próxima etapa foi conduzida através da construção do modelo em si. Tal passo foi possível através do programa **MODELLER**, que foi utilizado e, assim, gerou 5 hipóteses de modelo. De posse dos modelos gerados, surge a etapa de avaliação para determinar o modelo mais confiável e adequado para novos estudos. As análises de avaliação foram desenvolvidas com o servidor online **WHATIF** e o programa **PROCHECK**.

[0245] Dados os resultados obtidos pelo **WHATIF**, a melhor hipótese apresenta um escore total no valor de  $-0,675$ , e um Fator G de  $-0,2$ , valores compatíveis com uma proteína de resolução de  $1,85 \text{ \AA}$  (referente ao PDB ID 5NW7).

[0246] Para que possamos realizar uma melhor seleção da hipótese escolhida, devemos analisar outros parâmetros. Em seguida, foi feita a análise dos ângulos diédricos e do Diagrama de Ramachandran, para poder seguir para outras análises.

[0247] Os diagramas de Ramachandran obtidos para as hipóteses apresentaram alguns resíduos em regiões proibidas. A hipótese escolhida apresenta 89,2% de seus resíduos em regiões altamente favoráveis e apenas 0,8% em regiões proibidas sendo os resíduos Ala379, Asn91 e Ser184.

[0248] Embora a hipótese apresente resíduos em regiões proibidas no Ramachandran, nenhum deles se encontra na região do sítio ativo. Apenas utilizando os dados apresentados até agora, não foi possível escolher qual o melhor modelo gerado. Para que a escolha seja feita corretamente, foram utilizados outros parâmetros para a análise.

[0249] Outras ferramentas importantes para analisar um modelo de proteína predito podem ser encontradas no pacote **SAVE v5.0**. Para a hipótese selecionada, todos os valores estão dentro dos valores esperados para estruturas de proteína de resolução de 1,85 Å (referindo-se à estrutura PDB ID 5NW7) e a distorção tetraédrica dos carbonos alfa apresentou um melhor valor que o do cristalográfico.

[0250] Além das ligações atômicas que ocorrem na cadeia peptídica, há também interações entre átomos não-ligados. A estrutura terciária de uma proteína é altamente dependente dessas interações e existem seis tipos de interações prováveis entre átomos não ligados e elas não ocorrem aleatoriamente. Assim, é possível através do programa **ERRAT**

identificar regiões e fazer uma análise estatística dessas interações.

[0251] Observando o gráfico desenvolvido pela **ERRAT**, é possível identificar três regiões distintas: uma verde, uma amarela e uma vermelha. A hipótese gerada apresenta grande quantidade de aminoácidos listados como vermelho, porém esses aminoácidos não estão na região do sítio ativo.

[0252] O fator de qualidade, valor relacionado às interações atômicas não ligadas, apresentados pela hipótese (61,6867) estão acima de 50, mostrando que é uma estrutura de alta qualidade.

[0253] Outra análise interessante é verificar se um resíduo de aminoácido corresponde ao ambiente em que está, ou seja, se há compatibilidade do modelo com sua sequência de aminoácidos. Essa análise pode ser realizada usando o **VERIFY 3D**. No estudo da Hipótese 3, pode-se observar que 95,51% dos aminoácidos teve uma pontuação igual ou superior a 0,2 no perfil 3D-1D. Além disso, há regiões que caiu abaixo do limite, indicando uma provável dobragem nessa região; no entanto, não é motivo de preocupação porque não está na região do site ativo das hipóteses.

[0254] Finalmente, como última análise, é possível verificar o volume de átomos. O programa **PROVE** correlaciona o volume obtido para cada aminoácido da proteína com os volumes de proteína de alta resolução. Como resultado, o **PROVE** fornece um valor de pontuação chamado escore Z que corresponde ao desvio estatístico programado de estruturas altamente resolvidas e refinadas. No caso da hipótese escolhida do modelo PMI em estudo, o Z-score encontra-se dentro da faixa esperada e o Z-score RMS acima da região de

interesse. No entanto, juntamente com todas as outras análises realizadas, pode-se determinar que se trata da melhor hipótese de modelo para a PMI humana.

#### **Etapa de validação do modelo da PMI Humana**

[0255] Uma vez selecionado e refinado o modelo final, avaliamos sua eficiência em estudos de triagem virtual. Para essa avaliação, foram escolhidos 30 inibidores conhecidos da PMI, depositados e provenientes do servidor online **BindingDB**, contendo valores descritos de potência ( $IC_{50}$ ), reportados na literatura. O objetivo dessa etapa era avaliar a capacidade que o modelo possui de recuperar os ligantes sabidamente ativos (ditos "contaminantes") e ordená-los de forma coerente em relação ao seus  $IC_{50}$ , com base em função de pontuação (**GOLD Fitness**) proveniente do programa GOLD.

[0256] Para que o objetivo fosse alcançado, primeiramente era necessária a definição do sítio de *docking*, uma esfera com as seguintes coordenadas espaciais:  $x = 20.21$ ,  $y = 16.63$ ,  $z = 10.00$ , com uma esfera de raio de 7 Å. Os resultados dessa etapa são mostrados na Tabela 6, em que os inibidores mais potentes foram os inibidores 1, 2, 3, 6, 7, 8, 10, 11, 13, 14, 15, 18, 19, 20 e 28.

TABELA 6 - RESULTADO DA RECUPERAÇÃO DA SELEÇÃO DE INIBIDORES DO BindingDB.

<b>Número dos inibidores</b>	<b>Inibidores</b>	<b>GOLD PLP Fitness</b>	<b>IC50 (nM)</b>
1	2998795	76,169	1.30E+03
2	5771379	76,061	1.00E+03
3	4089709	72,416	707

4	25199560	67,963	>1,00E+6
5	25199559	67,128	>1,00E+5
6	2248499	66,702	829
7	4917291	61,567	709
8	1256813	60,232	1.00E+03
9	3311395	58,643	>1,00E+5
10	1489253	58,532	>1,00E+5
11	1519965	58,281	1.00E+03
12	4120895	58,255	>1,00E+5
13	2311082	57,179	270
14	25181243	56,910	1.03E+03
15	1510389	54,689	985
16	25241622	54,330	>1,00E+5
17	387196	53,533	>1,00E+5
18	3093671	53,365	1.27E+03
19	22416235	52,628	1.25E+03
20	3194	51,576	190
21	4879298	46,232	>1,00E+5
22	25181212	45,228	>1,00E+5
23	25181239	44,972	>1,00E+5
24	6112173	44,039	>1,00E+5
25	22765927	43,708	>1,00E+5
26	25241626	43,102	>1,00E+5
27	25241624	42,484	>1,00E+5
28	25199533	41,919	>1,00E+5
29	6033619	40,274	>1,00E+5
30	5951502	40,144	>1,00E+5

[0257] Observando-se os dados obtidos pelos estudos de

triagem virtual, o próximo passo consistiu na obtenção de curvas ROC, de predição e de enriquecimento, para melhor entendimento dos resultados. Para a criação desses gráficos, foi utilizado o *web servidor* Screening Explorer (EMPEREUR-MOT; ZAGURY; MONTES, 2016) com os parâmetros selecionados como padrões.

[0258] Em estudos de triagem virtual, a análise dos dados é feita com base em uma curva ROC (*Receiver Operating Characteristic*), resumem a capacidade geral de um método para distinguir entre compostos ativos e inativos (EMPEREUR-MOT et al., 2015; EMPEREUR-MOT; ZAGURY; MONTES, 2016). Utilizando as mesmas análises feitas no item anterior, quanto maior a distância entre a linha do gráfico e a linha diagonal, melhor será a precisão geral do estudo (BERTRAND; TRIBALLEAU, 2005; EMPEREUR-MOT et al., 2015; EMPEREUR-MOT; ZAGURY; MONTES, 2016).

[0259] Abaixo, na Figura 25, é mostrado o gráfico de curva ROC obtido para o estudo realizado com os 30 inibidores do **BindingDB** e o modelo de PMI *Homo sapiens* gerado. Nele é possível observar que grande parte do gráfico se encontra distante da linha diagonal, mostrando que o estudo apresenta uma boa precisão, conseguindo diferenciar os compostos mais ativos dos menos ativos. Inclusive uma parte da linha pertencente ao gráfico encontra-se na parte superior, mostrando que esses inibidores se encontra com a sensibilidade e especificidade perfeitas.

[0260] Outro dado que pode ser utilizado é a AUC, que mostra o desempenho geral do estudo. O valor de AUC ROC obtido para esse estudo foi de 0,792, um bom valor para o

estudo realizado já que se encontra próximo a 1 (valor ideal).

[0261] Um outro dado obtido pelo *web server* é a curva de previsão, que quantifica se as variações de probabilidade de atividade são importantes o suficiente para induzir a seleção de um limiar para triagens virtuais (EMPEREUR-MOT et al., 2015). A curva de previsão para esse estudo se encontra acima da linha limite, tornando-se um bom resultado e mostrando que o modelo foi capaz de realizar uma discriminação entre os compostos utilizados.

[0262] Um dado calculado a partir da curva de previsão é o ganho total padronizado (TG) (BURA; GASTWIRTH, 2001), que resume a discriminação dos compostos ativos em relação à variação das pontuações ao longo do conjunto de dados das moléculas (EMPEREUR-MOT et al., 2015).

[0263] TG possui seu valor entre 0 e 1, sendo que sua análise deve estar associada ao valor obtido para a AUC. Valores de TG acima de 0,4 associados com uma AUC ROC acima de 0,5 indica, que o método de triagem utilizado teve um bom desempenho e seria reprodutível em condições experimentais semelhantes (EMPEREUR-MOT; ZAGURY; MONTES, 2016).

[0264] No estudo, o valor de TG obtido foi de 0,444 que, associado a um AUC ROC de 0,792, mostra que o método utilizado teve um bom desempenho, podendo ser reprodutível em condições experimentais semelhantes.

[0265] Por fim, um último dado obtido é a curva de enriquecimento, a qual permite avaliar a recuperação precoce de compostos ativos presentes em estudos de triagem virtual. Assim, quanto mais distante a curva gerada estiver



da linha diagonal, melhor é o estudo realizado (BERTRAND; TRIBALLEAU, 2005). A curva de enriquecimento também se encontra acima da linha limite.

[0266] Um dado relacionado à curva de enriquecimento é o fator de enriquecimento (EF), que corresponde à fração de compostos ativos obtidos quando uma determinada porcentagem de compostos inativos é encontrada (TRUCHON; BAYLY, 2007). Nesse estudo foi considerado a porcentagem de 10% dos compostos inativos e o valor obtido para EF foi de 2,31.

[0267] Além dessas curvas, temos alguns outros dados para levar em consideração: discriminação reforçada por Boltzmann do ROC (BEDROC) (MUEGGE; ENYEDY, 2004) e o reforço inicial robusto (RIE) (CORNELL, 2006). Ambas as métricas são utilizadas para quantificar o reconhecimento precoce dos compostos ativos utilizados em um estudo de triagem virtual (EMPEREUR-MOT; ZAGURY; MONTES, 2016). O valor obtido para o RIE foi 2,069; enquanto que o valor obtido pra BEDROC foi 0,897 encontra-se muito próximo a 1, mostrando que o modelo é capaz de separar os compostos ativos e inativos.

[0268] Com isso, através desse resultado de validação e daqueles de avaliação mostrados anteriormente, podemos afirmar que o modelo para a PMI *Homo sapiens* provou ser capaz de diferenciar os inibidores mais potentes dos menos potentes escolhidos no **BindingDB**.

[0269] Após as análises dos resultados da segunda etapa do presente invento, o modelo de PMI de XAC construído se demonstrou confiável através das validações e da avaliação dos resíduos do sítio de interação. Dessa forma, tal modelo pôde ser utilizado em triagens virtuais em bases de dados

de compostos, na busca por novos quimiotipos ou classes químicas que poderiam interagir com o mesmo alvo, visando ao controle do cancro cítrico.

[0270] As estratégias de triagem virtual utilizadas foram validadas *in silico* e mostraram-se eficazes para a obtenção de diversas moléculas com potencial de inibição para a PMI. Um total de 45 compostos foram obtidos na etapa de triagem virtual, porém alguns foram selecionados para compra e estão passando por ensaios biológicos para validar e quantificar suas atividades biológicas assim estimadas. Os resultados serão racionalizados do ponto de vista estrutura-atividade, com vistas a um posterior refinamento ou otimização de suas propriedades ADME/Tox bem como potência inibitória da PMI.

[0271] Vale mencionar que o modelo da PMI humana, também desenvolvido por homologia estrutural e aqui descrito, foi planejado e já tem sido útil para revelar diferenças em relação à PMI da bactéria, em especial na região do sítio ativo, as quais viabilizam o planejamento de inibidores seletivos para a PMI de XAC, mas não para a homóloga humana, assim conferindo um adicional potencial atóxico aos nossos futuros bactericidas que tenham como mecanismo de ação a inibição da PMI da bactéria do cancro cítrico, como uma das vantagens em relação aos defensivos agrícolas atualmente utilizados.

#### **Testes de patogenicidade *in vivo* de XAC na presença dos inibidores selecionados**

[0272] Os inibidores selecionados por triagem computacional (denominados 1, 2, 3 e 4 - Figura 26) foram adquiridos comercialmente e ensaiados para o seu efeito

preventivo dos sintomas do cancro cítrico produzidos por XAC em plantas de limão galego. Para cada inibidor foram testadas 3 concentrações, em três ramos diferentes de cada planta, sendo utilizadas réplicas das plantas. As concentrações de inibidores (testadas em réplicas) foram: 25 µM (C1), 250 µM (C2) e 2,5 mM (C3) (para o inibidor 1) e 100 µM (C1), 1mM (C2) e 10 mM (C3) (para inibidores 2, 3 e 4), tendo sido todos os compostos dissolvidos em água. O controle foi realizado na ausência de inibidores. Os inibidores (ou água para o controle) foram aspergidos nas folhas até gotejamento e a suspensão bacteriana de XAC ( $10^6$  UFC/mL) foi aspergida após 40 minutos, tempo suficiente para "secagem" da solução de inibidores. As folhas foram documentadas após 21 dias, por escaneamento tanto da face abaxial quanto adaxial das folhas após serem destacadas das plantas (Figuras 27, 28 e 29).

**PMI como proteína-alvo para controle de doenças causadas por *Xanthomonas* sp.**

[0273] Neste invento é proposto utilizar a proteína PMI como um alvo para novos compostos de controle do cancro cítrico. Esta descoberta possui valor ainda mais significativo quando se verifica que as espécies *Citrus sinensis* (laranja), *Citrus deliciosa* (tangerina) e *Citrus limon* (limão), que são citros de interesse econômico, não codificam em seus respectivos genomas a proteína PMI; de fato não há similaridade significativa da sequência da PMI de XAC com nenhuma sequência desses hospedeiros cítricos, o que foi confirmado pela busca com a sequência XanB usando a ferramenta BLASTp no genoma dessas espécies

disponibilizadas no NCBI (<http://www.ncbi.nlm.nih.gov>). Este pode ser um indicativo importante de que compostos para inibição do alvo bacteriano não terão efeito sobre as culturas de citros.

[0274] A utilização de PMI como alvo para o delineamento de novos compostos envolve não apenas o controle do cancro cítrico mas uma enorme gama de fitopatologias causadas pelas mais diversas espécies do gênero *Xanthomonas* e outros gêneros relacionados. Esta possibilidade fica evidente por meio da busca realizada com a ferramenta **BLASTp** em bancos de proteínas anotadas a partir das sequências genômicas (NCBI), e utilizando a sequência da proteína PMI de XAC (ORF XAC3580) como "sonda": houve uma identidade sequencial de mais de 95% com a mesma proteína de muitas espécies do gênero *Xanthomonas*, dentre elas: *X. phaseoli* (99%), *X. euvesicatoria* (99%), *X. prunicola* (99%), *X. vasicola* (98%), *X. oryzae* (98%), *X. perforans* (98%), *X. bromi* (97%), *X. pisi* (97%), *X. codiae* (96%), *X. dyei* (96%), *X. cucurbitae* (96%), *X. campestris* (95-96%), *X. vesicatoria* (96-97%), *X. cucurbitae* (96%), *X. nasturtii* (97%), *X. cassavae* (96%), *X. floridensis* (96%), *X. arboricola* (96%), *X. melonis* (96%), *X. hortorum* pv. *carotae* (94%). Essas espécies são agentes causadores de patologias em uma grande variedade de culturas economicamente importantes, dentre elas: feijão, tomate, arroz, pimentão, repolho, couve, couve-flor, cenoura, abóbora, pepino, melão, nectarina, agrião, mandioca, dentre outras.

[0275] Com respeito à comprovação da enzima PMI como enzima vital à infecção da XAC em citros e, portanto, como alvo único suficiente para o planejamento de inibidores

seletivos, não há menção na literatura, especialmente em *Xanthomonas*. Como comentado, há apenas um artigo (do nosso próprio grupo de inventores) que mostra que a PMI foi detectada em análise proteômica de superfície e que, dentre outras 26 proteínas, está potencialmente relacionada com a patogenicidade bacteriana mas não há comprovação funcional na literatura. A abordagem utilizada de produção de mutante e complementante para PMI em XAC (Alexandrino, 2020) foi de fundamental importância para se confirmar a enzima como essencial à patogenicidade de XAC e, dessa forma, inferir a sua potencialidade como alvo altamente promissor para se chegar a inibidores efetivos contra o cancro cítrico. Assim, tais inibidores poderiam ter também efeito contra muitas outras espécies de *Xanthomonas* que causam fitopatologias em uma grande diversidade de culturas economicamente importantes, bem como em espécies de outros gêneros que possam também ter a enzima PMI como essencial ou relacionada com a sua patogenicidade.

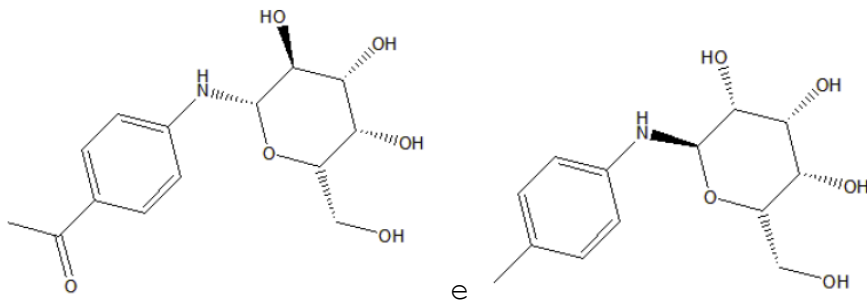
[0276] A produção recombinante da PMI utilizando técnicas de Biologia Molecular permitiu comprovar a atividade catalítica prevista para o produto do gene da PMI, gene este que foi deletado (retirado) no mutante e reintroduzido no complementante, correlacionando a ausência de patogenicidade com a ausência do gene da PMI. Como comentado, a análise proteômica disponível em si (Carnielli et al., 2017) foi apenas um indicativo inicial de inúmeras proteínas-alvo em potencial, no entanto a validação da análise proteômica pela análise genética foi essencial para se inferir que PMI é essencial à patogenicidade de XAC e em que grau de intensidade isto ocorre.

[0277] Dos dois compostos propostos como inibidores da PMI, não há qualquer menção conhecida quanto à sua possível aplicação para controle do cancro cítrico. Os compostos estão disponíveis comercialmente, com números de catálogo da empresa fornecedora dos mesmos, mas não há como chegar ao conhecimento de que são inibidores da PMI e muito menos que eles atuem como preventivos do cancro cítrico, sem realizar um extenso trabalho tal como foi desenvolvido computacionalmente, a partir da estrutura 3D da PMI de XAC, também desenvolvida no laboratório da FCFRP-USP, com modelagem estrutural por homologia estrutural. Não há estrutura 3D da PMI que esteja disponibilizada em coordenadas cartesianas no Protein Data Bank (PDB), para fim de trabalhos computacionais.

## REIVINDICAÇÕES

1. Uso de um composto ou de um sal aceitável do mesmo, **caracterizado pelo** fato de ser para a fabricação de um composto para inibição seletiva de Fosfomanose isomerase (PMI).

2. Uso de um composto ou de um sal aceitável do mesmo, de acordo com a reivindicação 1, em que o referido composto é selecionado a partir do grupo que consiste em



**caracterizado pelo** fato de ser para a fabricação de um insumo agrícola para controle de cancro cítrico.

3. Uso, de acordo com a reivindicação 1, **caracterizado** pelo fato de ser para a fabricação de insumos agrícolas para controle de fitopatologias causadas por *Xanthomonas* sp.

4. Uso, de acordo com a reivindicação 1, 2 e 3, **caracterizado pelo** fato de a Fosfomanose isomerase (PMI) ser uma proteína essencial para a patogenicidade da bactéria *Xanthomonas* sp e ser alvo promissor para controle do cancro cítrico.

**pET2829\_PMI**

pET2829PMI\_FW 5' gtaa <sup>*NdeI*</sup>CA/TATGAGCGACGTCCTACCC 3' TM 65,1 °C  
 M S D V L P

pET2829PMI\_RV 5' tatata <sup>*XhoI*</sup>C/TCGAGTTAGGCGCGGCCGTA 3' TM 68,8 °C

pET2829PMI\_RV\_C 5' tatata <sup>*XhoI*</sup>C/TCGAGGGCGCGGCCGTAG 3' TM 69,4 °C

**1kb\_Up\_PMI**

Up\_PMI\_FW 5' tatata <sup>*HindIII*</sup>A/AGCTTATACCGCAGCGCCAGGC 3' TM 68,8 °C

Up\_PMI\_RV 5' cgtata <sup>*EcoRI*</sup>G/AATTCGGGCGGGGGAAACTC 3' TM 68,7 °C

**1kb\_Down\_PMI**

Down\_PMI\_FW 5' ttaaca <sup>*EcoRI*</sup>G/AATTCTCGCGGTCTCAACGTAG 3' TM 68,7 °C

Down\_PMI\_RV 5' ataatt <sup>*NheI*</sup>G/CTAGCTTGTACGACAACCTGTACCC 3' TM 68,5 °C

**KO\_PMI**

KO\_PMI\_FW 5' GATCAGCTCGGACACCGG 3' TM 63,2 °C

KO\_PMI\_RV 5' GGACCGACAACATCATCTTCTCG 3' TM 63,8 °C

**pNPTS\_cPMI**

pNPTS\_cPMI\_FW 5' tatata <sup>*EcoRI*</sup>G/AATTCATGAGCGACGTCCTACCCAT 3' TM 67,3 °C  
 M S D V L P I

pNPTS\_cPMI\_RV 5' tatata <sup>*EcoRI*</sup>G/AATTCITAGGCGCGGCCGTAGG 3' TM 67,8 °C

Figura 1

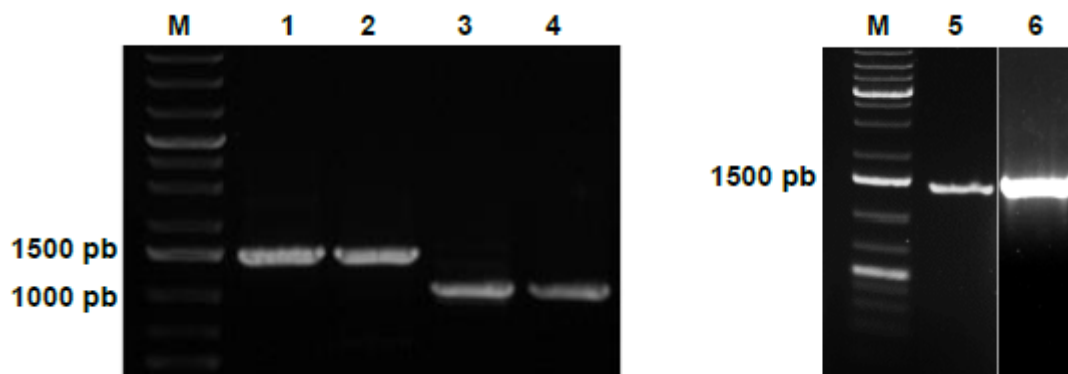


Figura 2



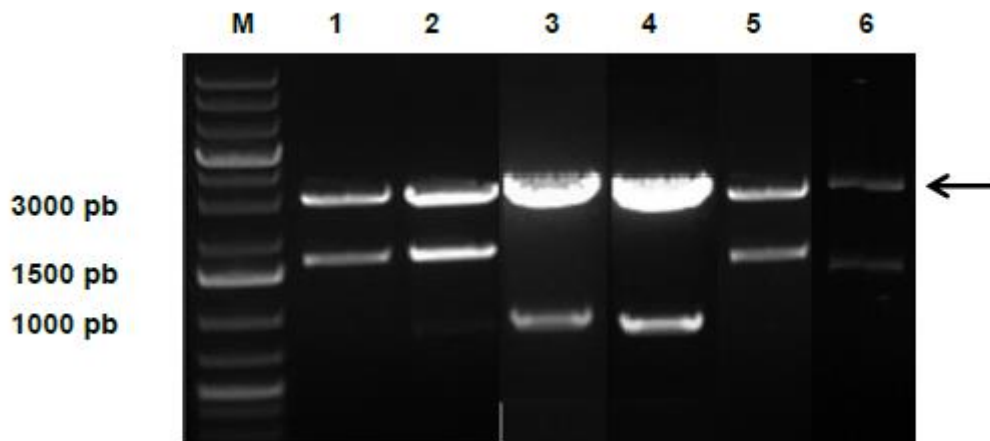


Figura 3

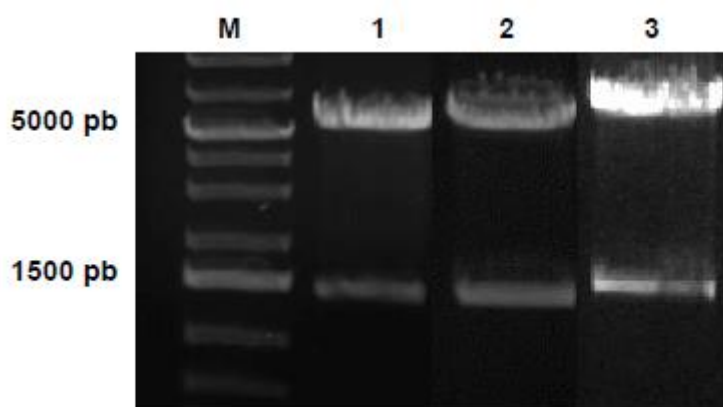


Figura 4

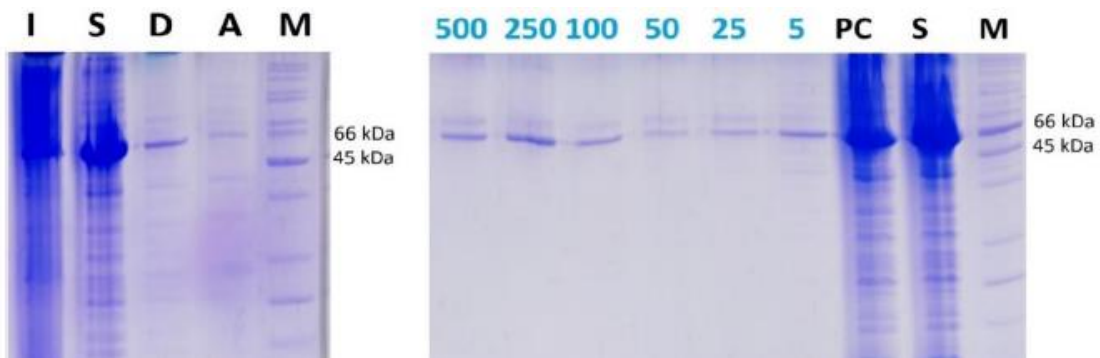


Figura 5

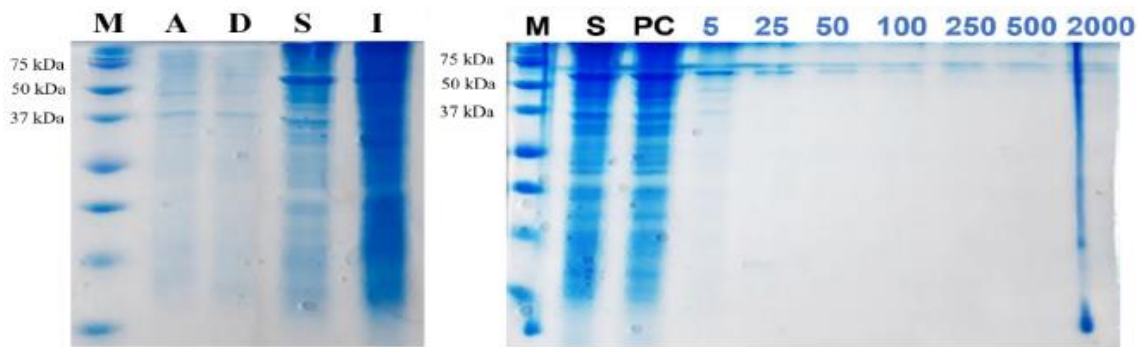


Figura 6

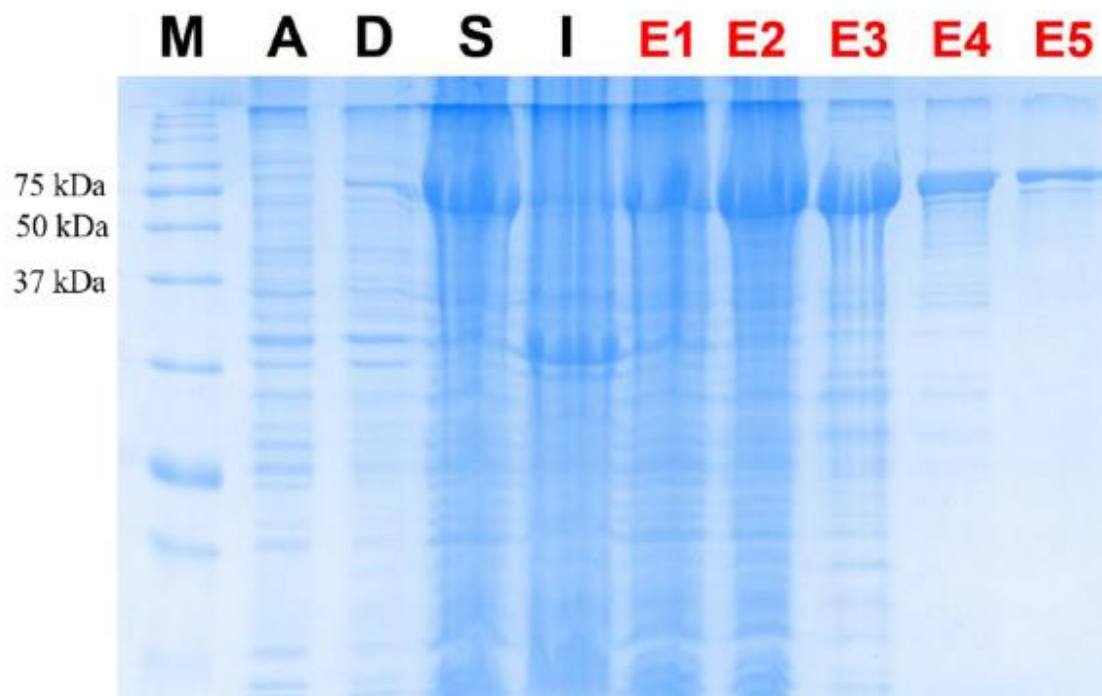


Figura 7

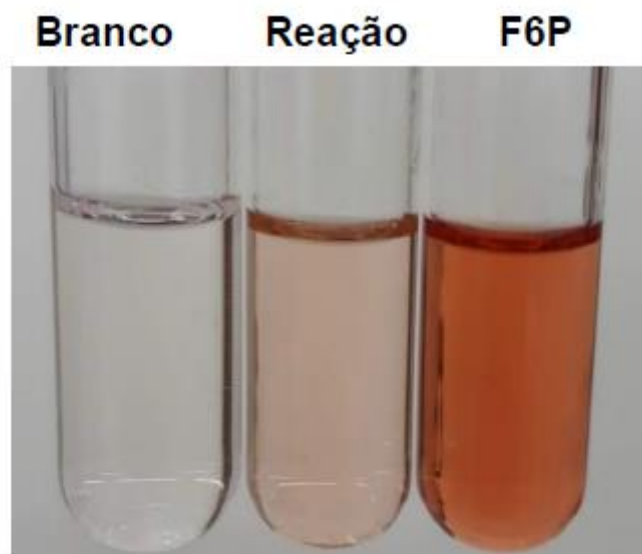


Figura 8

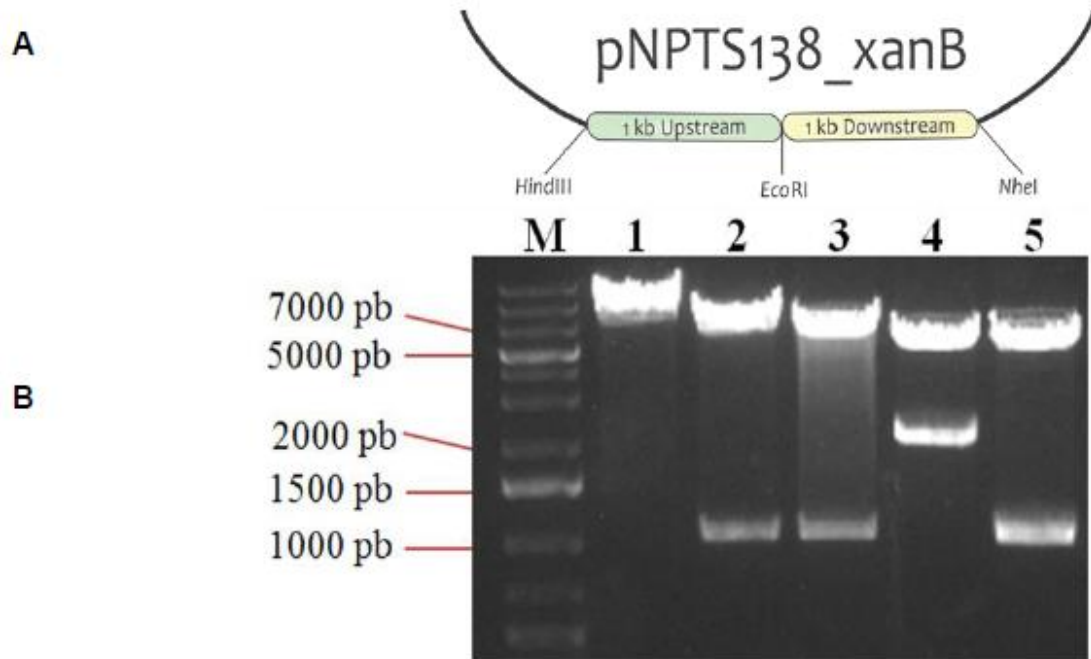


Figura 9

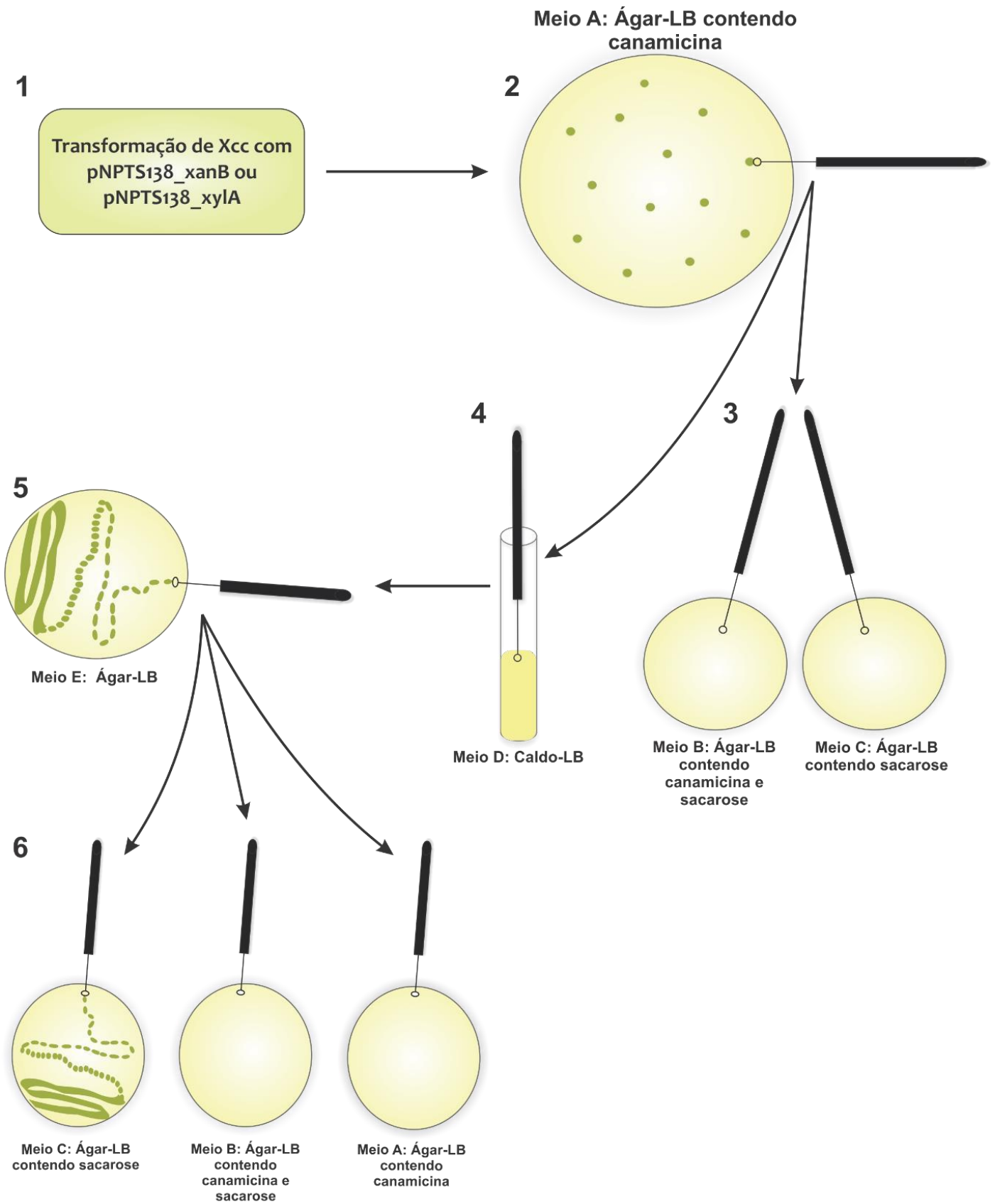


Figura 10

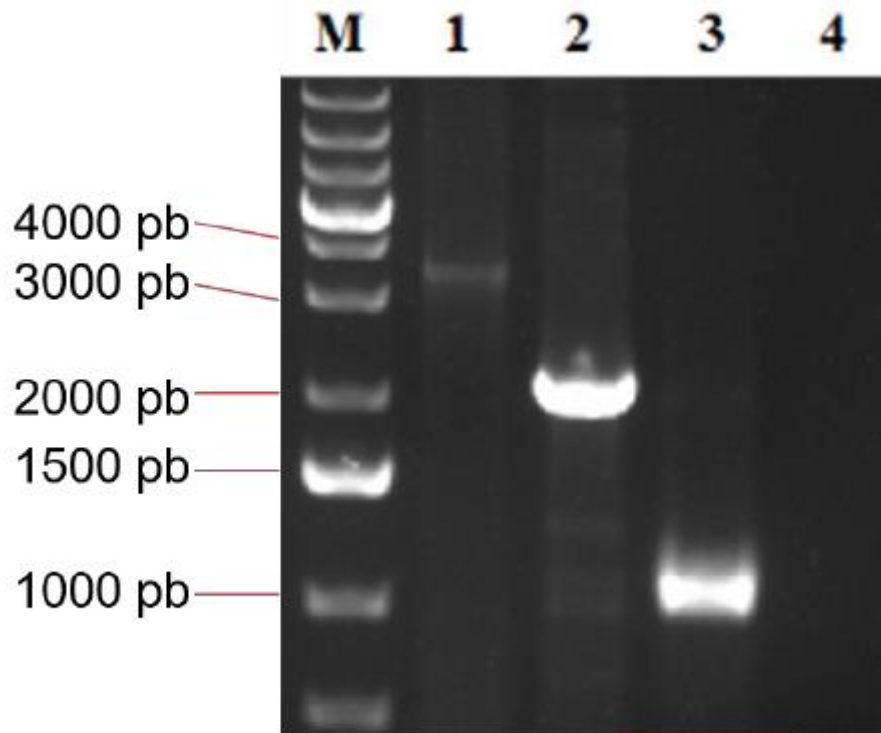


Figura 11

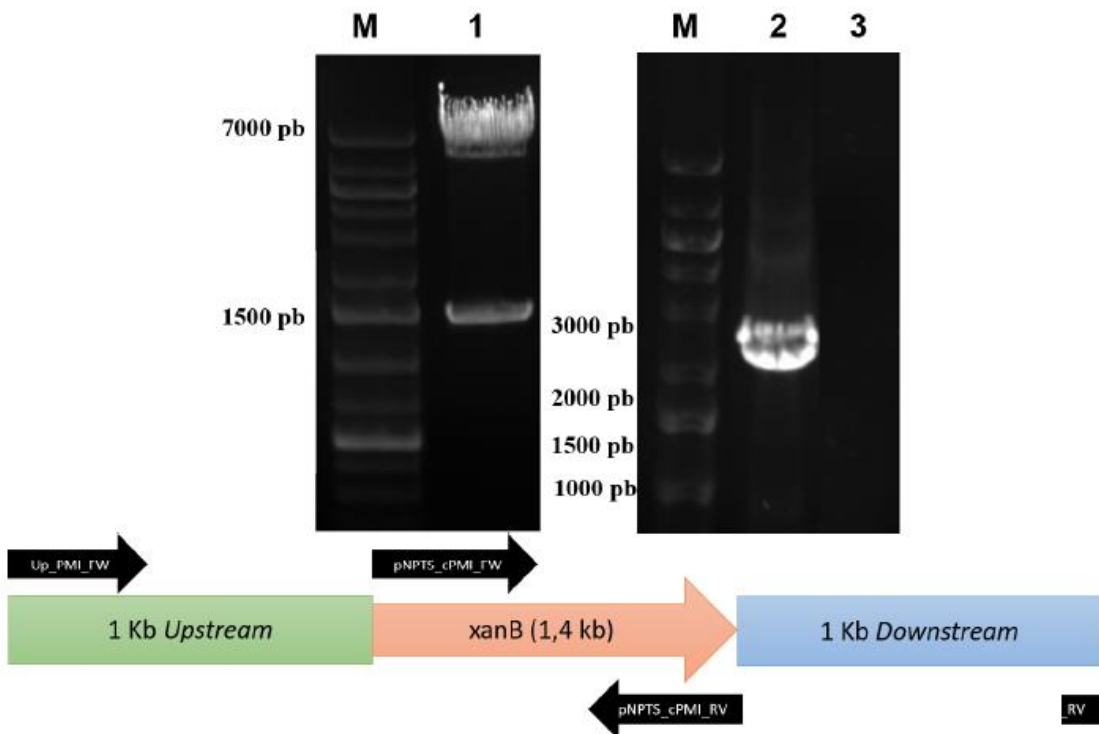


Figura 12

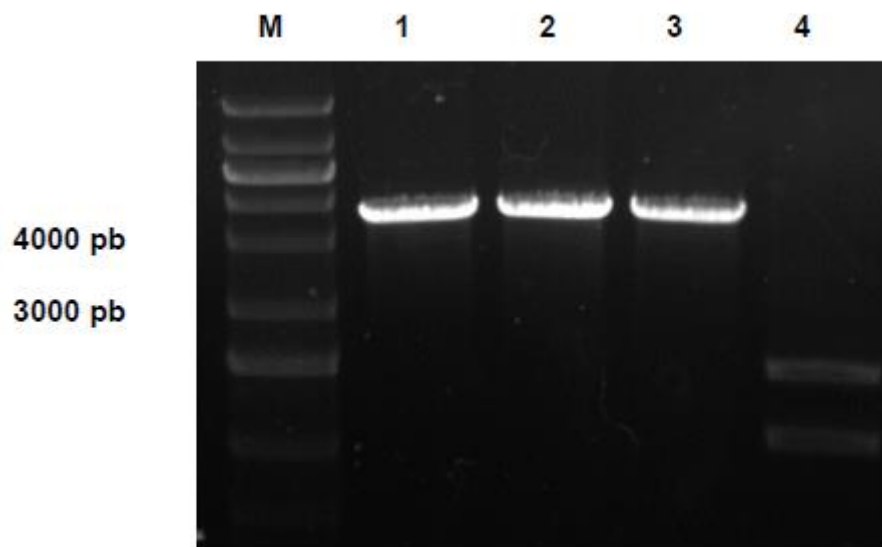


Figura 13

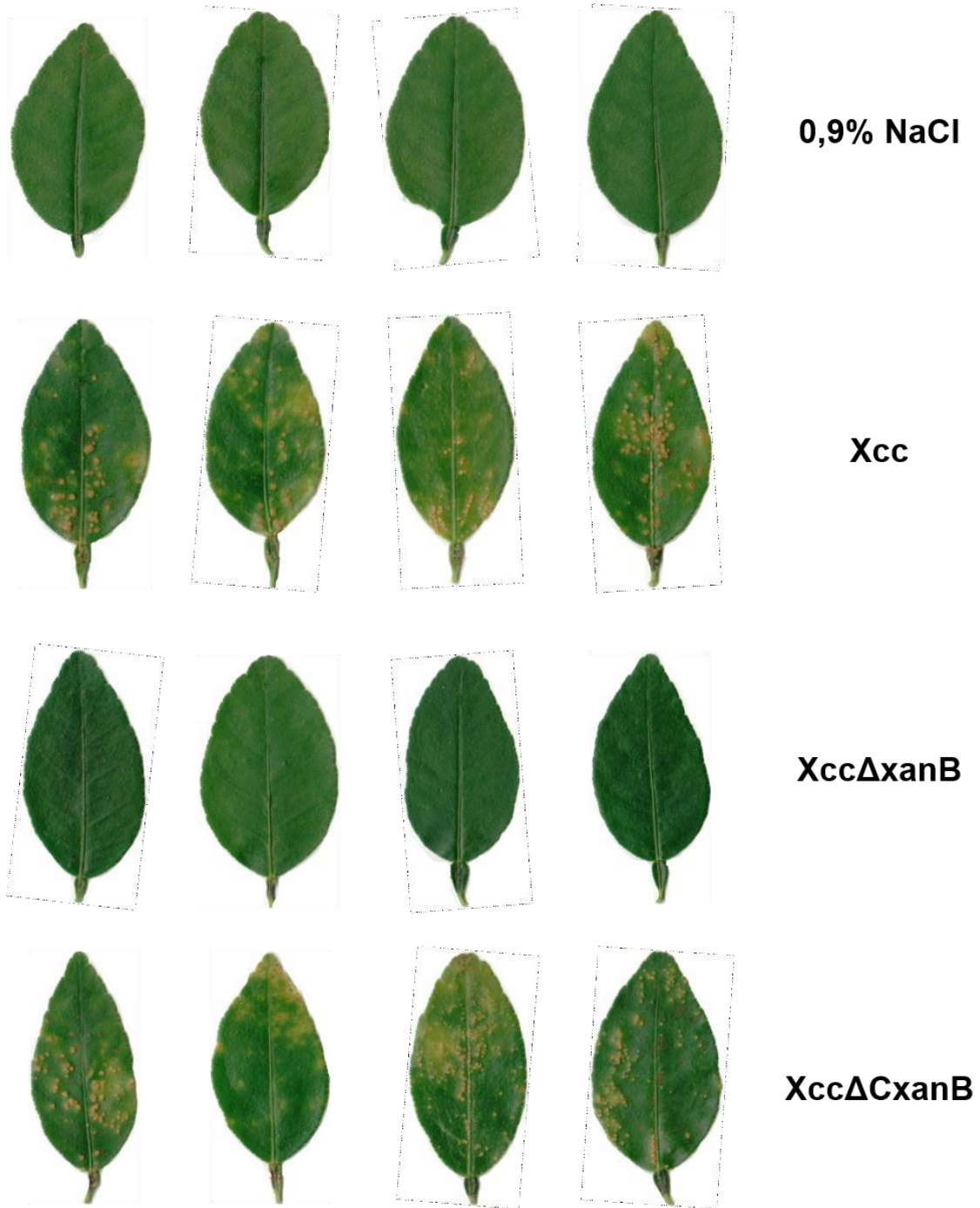


Figura 14

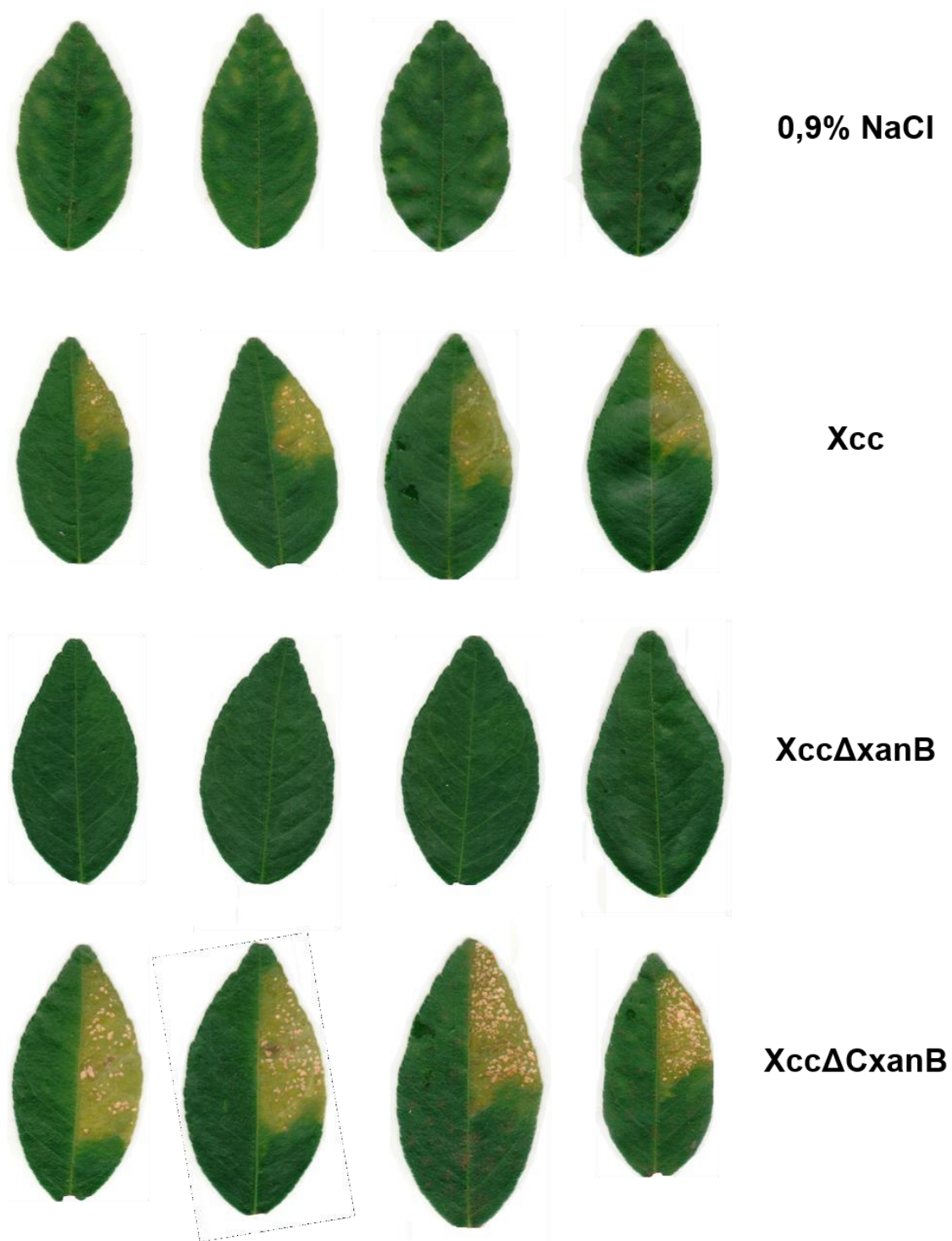


Figura 15



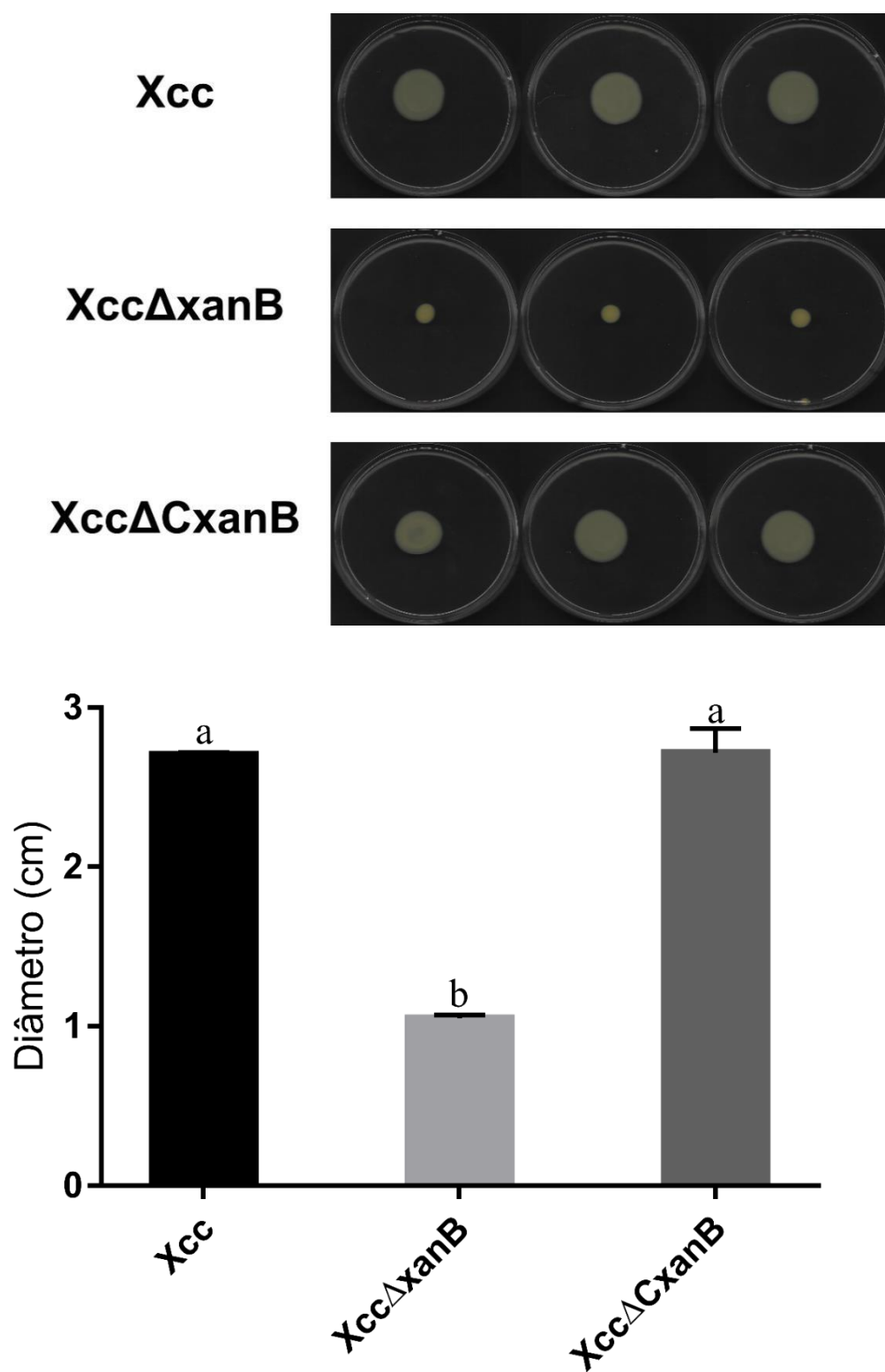


Figura 16

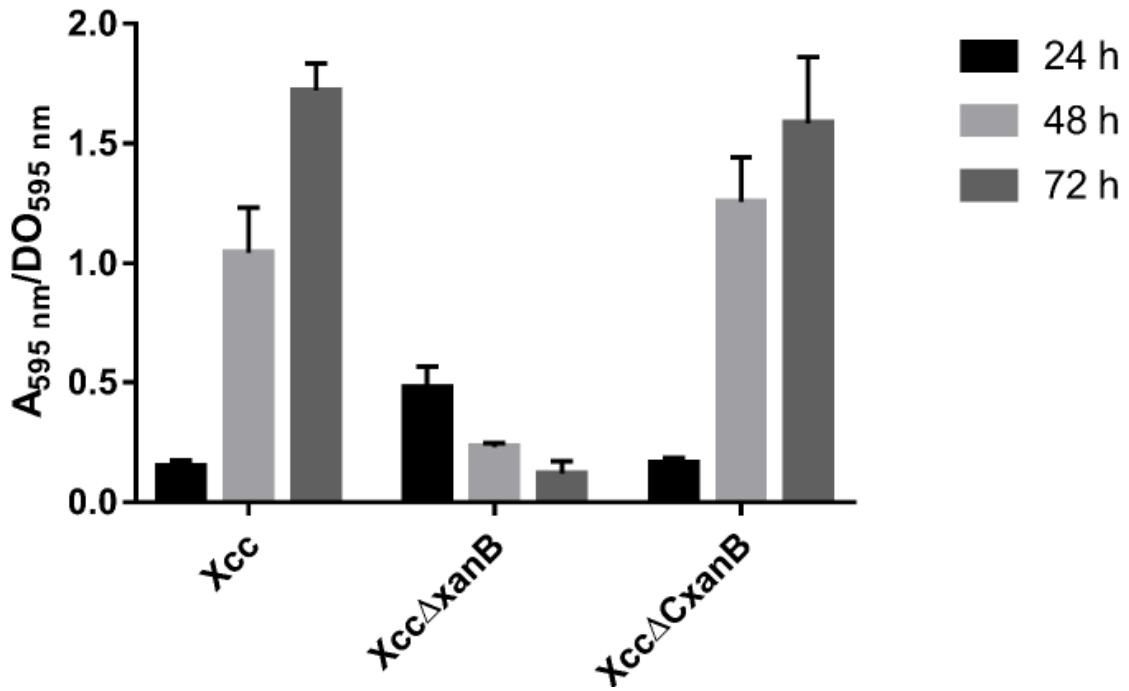


Figura 17

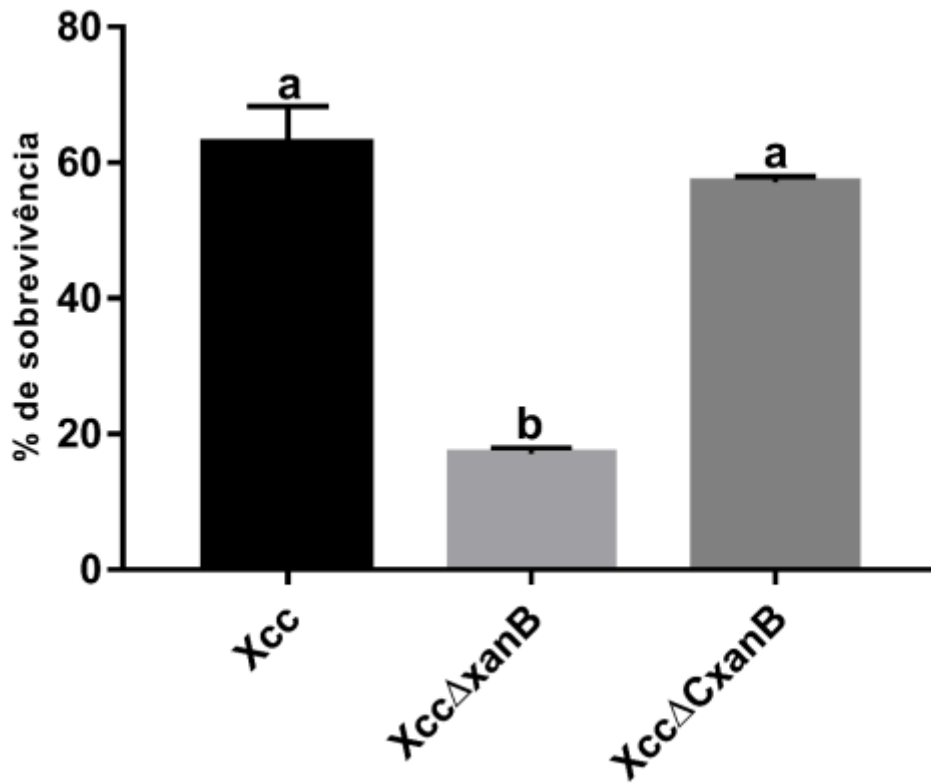


Figura 18

1H5R	-KMRKGIILAGGSGTRLYPVTMA-VSKQLLPIYD-KPMIYYPLSTLMLAGIR--DILIIST	57
2CU2	-MKTYALVMAGGRGERLWPLSREDRPKPFLPLFEGKTLLEATLERLAPL-VPPERTLLAVR	59
2X5S	--VMKALILAGGSGERFWPLSTPETPKQFLKLFGNKSLMRWTFERVLEE-MDPKDVIVVTH	58
2QH5	-LKIKNILLSG-----YPKQFLKLFDHKSLFELSFKRNASL-VD--ETLIVCN	57
<b>PMI</b>	<b>MSDVLPIILSGGSGTRLWPLSRESYPKQFLPLVGDKSMLOSTWLRAPV-AG-HAPIVWAN</b>	<b>59</b>
	:::*** * *::*:               * : * *                               ::	
1H5R	PQDTPRPFQQLLGDGSGWGLNL---QYKVQSPDGLAQAFIIGEE---FIG---GDICALVLG	110
2CU2	RDQEAVAR-PYADG-----I---RLLLEPLGRDTAGAVLLGVA---EALKEGAERLLVLPAL	108
2X5S	KDYVERTKKEPE-----LPDENIAEPMKKNTAPACFIGTKL---AD--DD-EPVLVLPAL	108
2QH5	EKHYFLALEEIKNEIK--NKSVMGLLESLSKNTANAIALSALM---SD--KE-DLLIVTPS	110
<b>PMI</b>	<b>EHRFMAAEQLDQIG--VKPS-AILLEPFRNTAPAIAVALEATRDG--AD-PLLLVLPAL</b>	<b>114</b>
	: . : : : :               . : : :	
1H5R	DNIFYG-HDLPKLMEEAVNK-ES--GATVFAYHVNDE-RYGVVEFDKN-----GTAISLEE	162
2CU2	DHYVGDDEAYREALATMLEAAEEG-FVALGLRPTPETEYGYIRLGPREG-AWYRGEQFVE	168
2X5S	DHRIPDTKFKFWKTVKKAALDALEKYDGLFTFGVPTPETGYGYIEIGEELEEGVHKVAQFRE	170
2QH5	DHLIKDLQAYENAIIKKAIDLAQKG-FLVTFGVSIDKPNTTFGYIESP----NGLD-VKRFIE	175
<b>PMI</b>	<b>DHVIQNEAFAQAVTAAATAEQG-KLVTFGIKPTAPETGYGYIKAS--AGAGASAVERFVE</b>	<b>182</b>
	*. . : : : : . : : :               *:::* : .               : : *	
1H5R	KP-----LEP--K-SNYAVTGLYFYDN-DVVMKKNLKPSAR-----	195
2CU2	KPSYAEALEYIRK-GYVWNGGVFAFAPATMAELFRRHLPSSHE-ALER-LLAGAS-----	220
2X5S	KPDLETAKKFVESGRFLWNMGFLWVKAREFTIEEVKVCESPIYENLKDVPDPRNFEE-----	225
2QH5	KPSLDKAIEFQKSGGFYFNNGMFVFOAGVFLDELKKHAPTILK-GCER-AFESLENAYFFEK	226
<b>PMI</b>	<b>KPDLATAQSYLASGEYWNVSGMFLFRASYLEELRKHHPALAD-ACOK-AWENGKRDADFT-</b>	<b>232</b>
	**     * . . . : . . *:::               . :     *.	
1H5R	-----GELEITD-INRIYLEQGRLSVAMMGRGYAWLDTGTHQSLIEASN	238
2CU2	-----LEEYVAGLPKISIDYGVMEKA-E--RVRVVLGR--FPWDDVGNWRALERVEFS	267
2X5S	-----LKKAYEKVPSISVDYAVMEKS-K--KVRVKAD--FEWSDLGNWSSVREIEG	272
2QH5	KIARLSEKSMQDLEDMSIDIALMQQS-H--KTKMVELN--AKWSDLGNFNAL-----	273
<b>PMI</b>	<b>--TRLDRDAFAASPSDSIDYAVMEKT-A--DAVVPLD--AGWVDVGSWSSL-----</b>	<b>277</b>

Figura 19

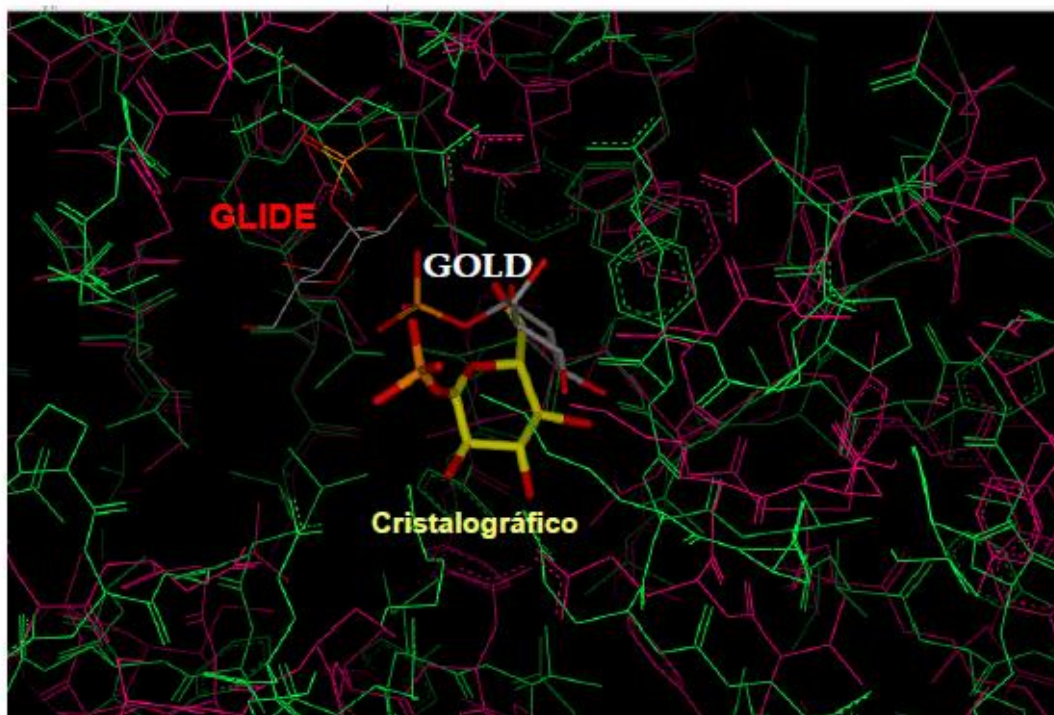


Figura 20

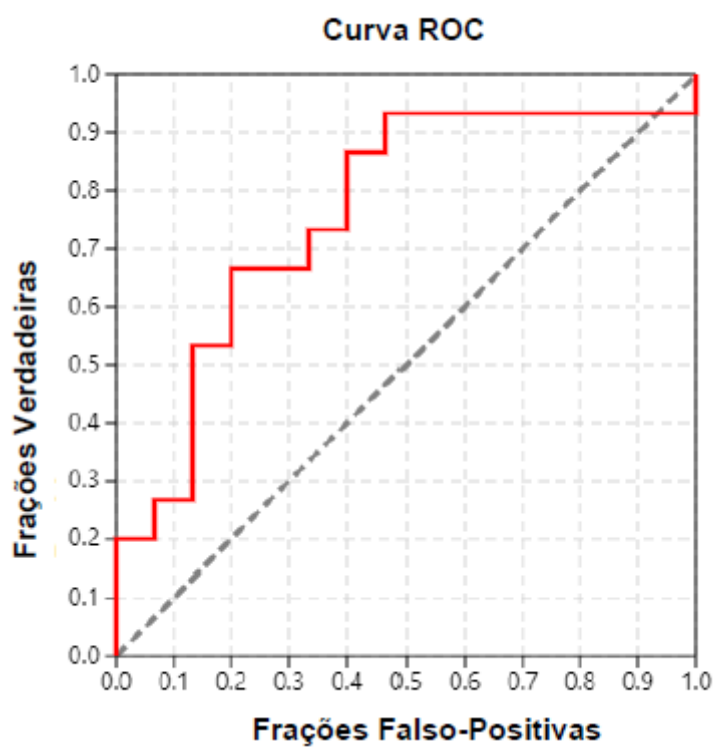


Figura 21

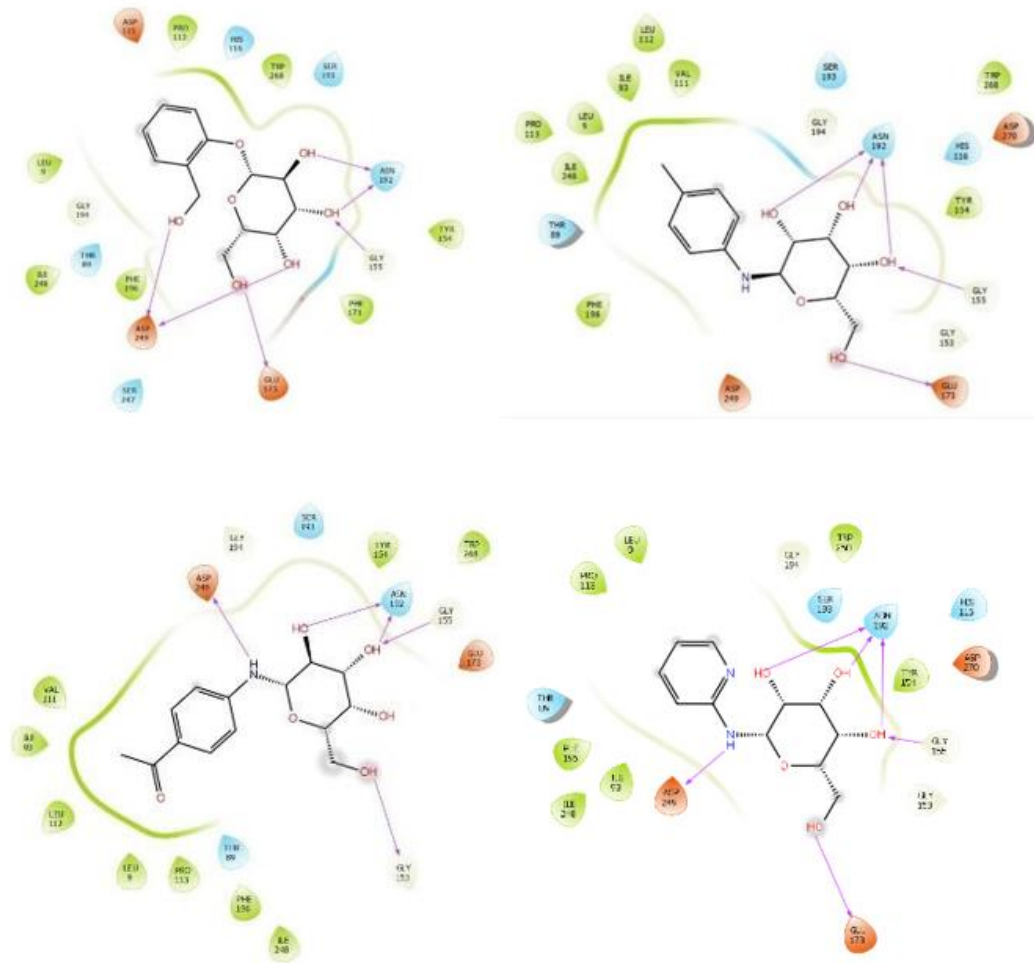


Figura 22

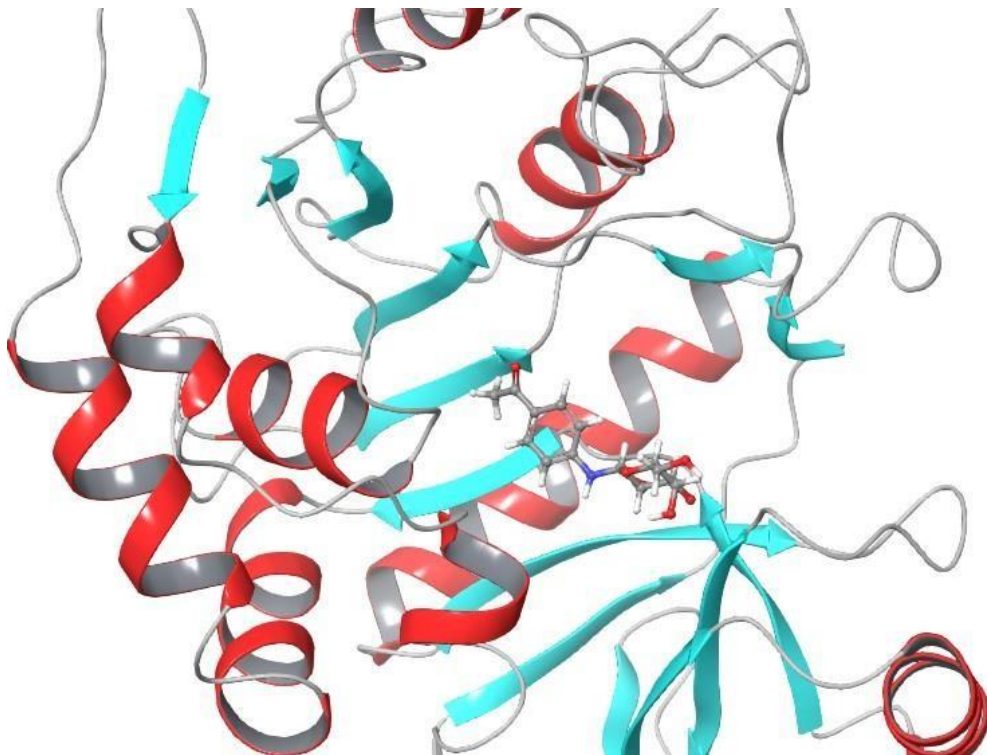


Figura 23

PMI H.	-----MAAPRVFPLSCAVQYAWGKMGSNSEVARLLASSDPLAQIAEDKPY	46
3H1M:A	-----ASMQLINSVQNYAWGSK-TALTELYG----- IANPQQQPM	35
5NW7:A	MRGSHHHHHHGMASMSSEKLFRIQCGYQNYDWGKIGSSSAVAQFVHNSDPSITIDETKPY	60
	: : . *:* ** . : : : : : *	
PMI H.	AELWMGTHPRGDAKILDNRISQKTLISQ-WIAENQDSLGSVKVDTF--NGNLPLFKVLSV	103
3H1M:A	AELWMGAHPKSSSRITTANGETVSLRDAIEKNKTAMLGEAVANRFG---ELPFLFKVLCA	92
5NW7:A	AELWMGTHPSVPSKAID--LNNQTLRDLVTAKPQEYLGESIITKFGSSKELPFLFKVLSI	118
	*****:** : : . :* : : : ** . : * :*****.	
PMI H.	ETPLSIQAHPNKELAEKHLQAPQ ----- HYPDANHKPEMAIALTPFGGLCGFRP	153
3H1M:A	AQPLSIQVHPNKRNSEIGFAKENAAGIPMDAAERNYKDPNHRKPELVFALTPFLAMNAFRE	152
5NW7:A	EKVLSIQAHDPKKLGAQLHAADPK ----- NYPDDNHRKPEMAIAVTFEGFCGFKP	168
	****.*:* . : : * * *****.::* * * . : . * :	
PMI H.	VEEIVTFKVKVPEFQFLIGDEAATHLKQTM ----- SHDSQAVASSLQSCFSLMKSEK	206
3H1M:A	PSDIVSLLQP---VAGA--HSAIAHPLQVP ----- NAERLSQLFASLLNMQG	194
5NW7:A	LDQLAKTLATVPELNEIIGQELVDEFISGIKLPAEVSQDDVNNRKLQKVFGLMNTDD	228
	..:.. * * : * . : : . * . * . *:: :	
PMI H.	KVVVEQLNLLVKRISQQAAGNMMEDIFGELLQLHQYPGDIGCFAL-YFLNLLTLKPG	265
3H1M:A	EEKSRALAVLKAALNSQQ ----- GEPWQTIIRVISEYYPDDSGLFSP-LLLNWVKLNPG	246
5NW7:A	DVIRKQQTAKLLERTDREPQVFKDIDSRLPELIQRLNRQFPNDIGLFCGCLLLNHVGLNKG	288
	. . * : : . : : : : *:* * * . :** : * : *	
PMI H.	EAMFLEANVPHAYLKGDCVECMACSDNTVRAGLTPKPIDVPTLCEMLSYPSSSKDRLEFL	324
3H1M:A	EAMFLPAETPHAYLQGVALEVMANSDNVLRAGLTPKYIDIPELVANVKFEPKPAGELLTA	305
5NW7:A	EAMFLQAKDPHAYISGDIIECM AASDNVVRAGFTPKPKDVKNLVEMLTYSYESVERQRMP	348
	***** * : *****.* :* ** ** .:***:***: * : * : : . .	
PMI H.	PTRSQ----EDPYLSIYDPPVPDFTIMKTEVPGSVTEYKVL---ALDSASILMMVQGTVI-A	379
3H1M:A	PVKSG-----AELDFPIPVDDFA--PSLHD LALQETSIGQH --- SAAILFCVEGEAVLR	355
5NW7:A	LQEFPRSKGDAVKSVLYDPPIAEFVVLQTIIDKS--KGGKQVIEGLNGPSIVVIATNGKGTIO	408
	* . : * : * : . . : . . . :* : : *	
PMI H.	STPTTQTPIPLQRGGVLFIGANE-SVSLKLTETP---KDLLIFRACLL	423
3H1M:A	-----KDEQRVLKPGESAFIGADESPVNASGT ---- GRLARVYNKL	393
5NW7:A	ITGDDSTKQKIDTGYVFFVAPG-SSIELTADSANQDQDFTTYRAFVEA	455

Figura 24

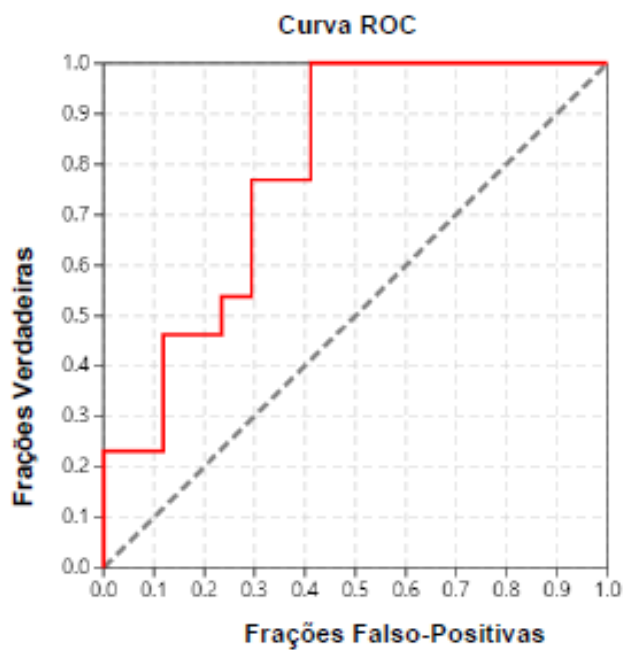


Figura 25

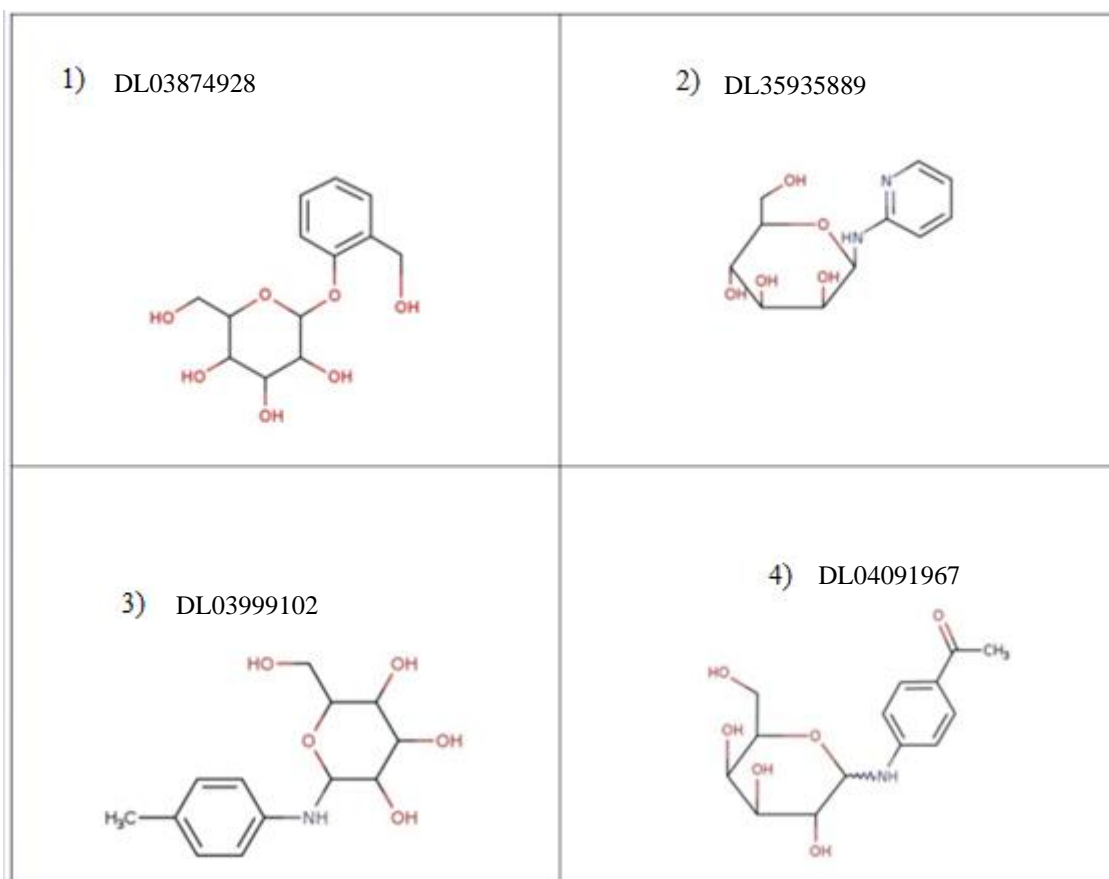


Figura 26

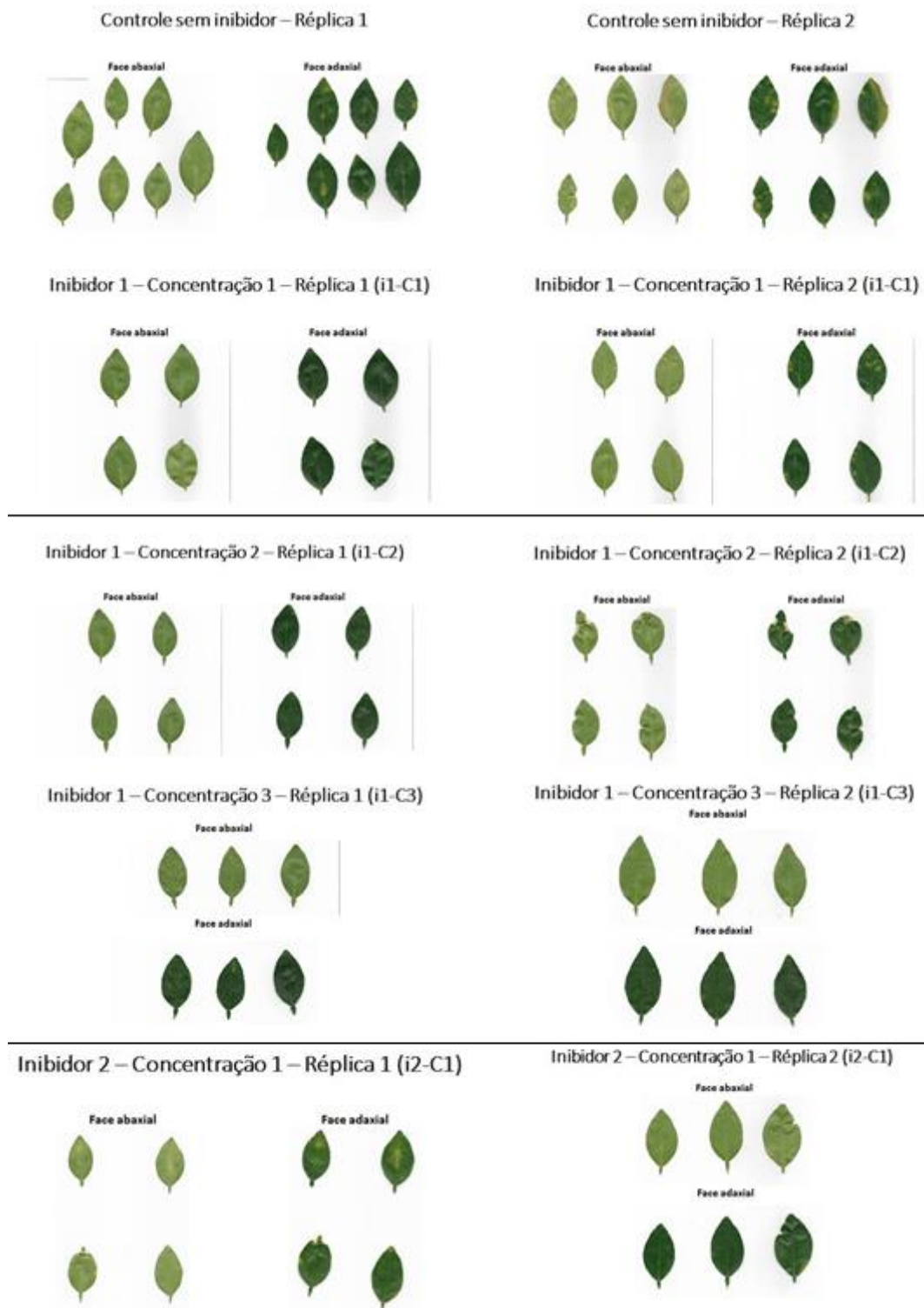


Figura 27



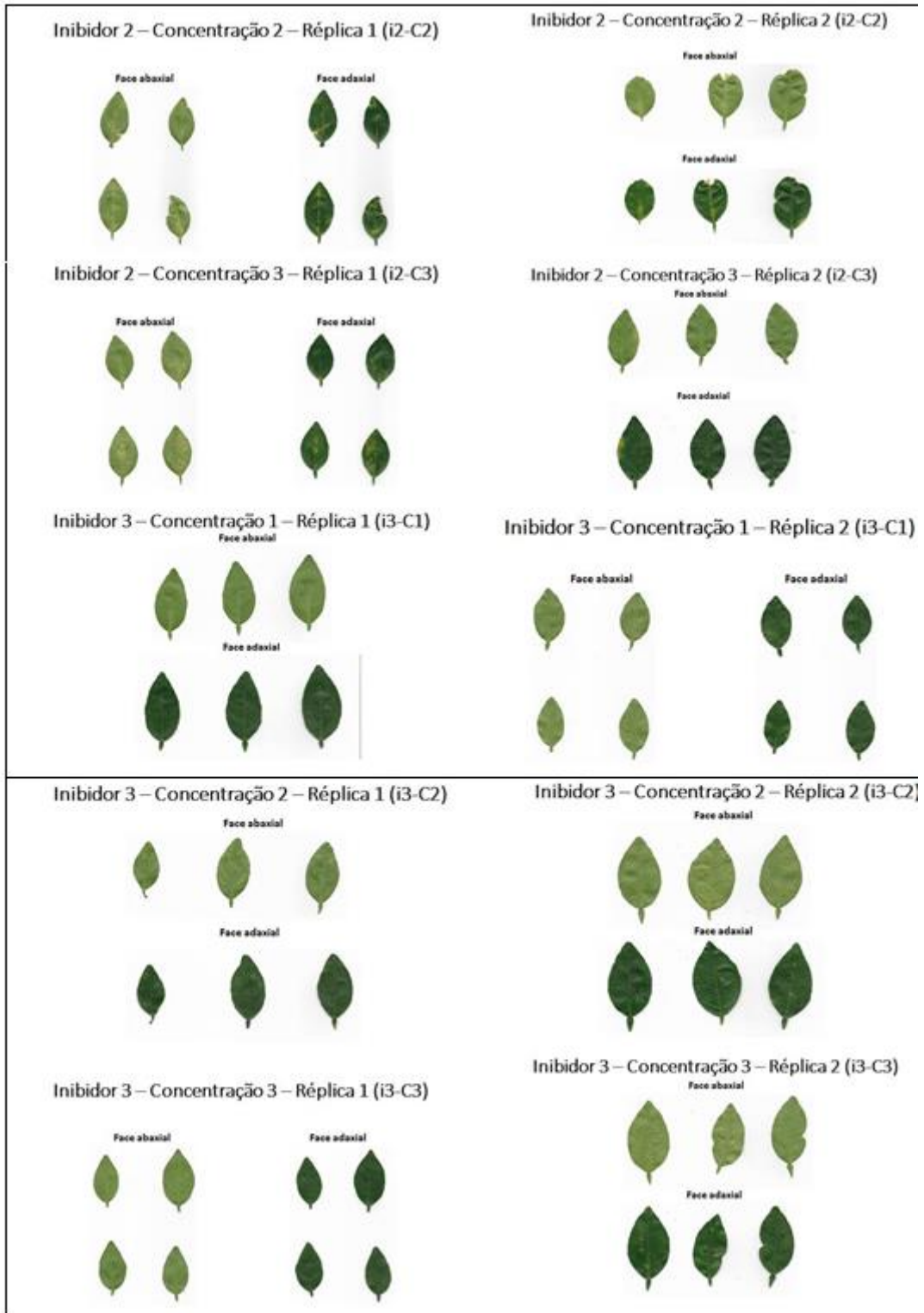


Figura 28

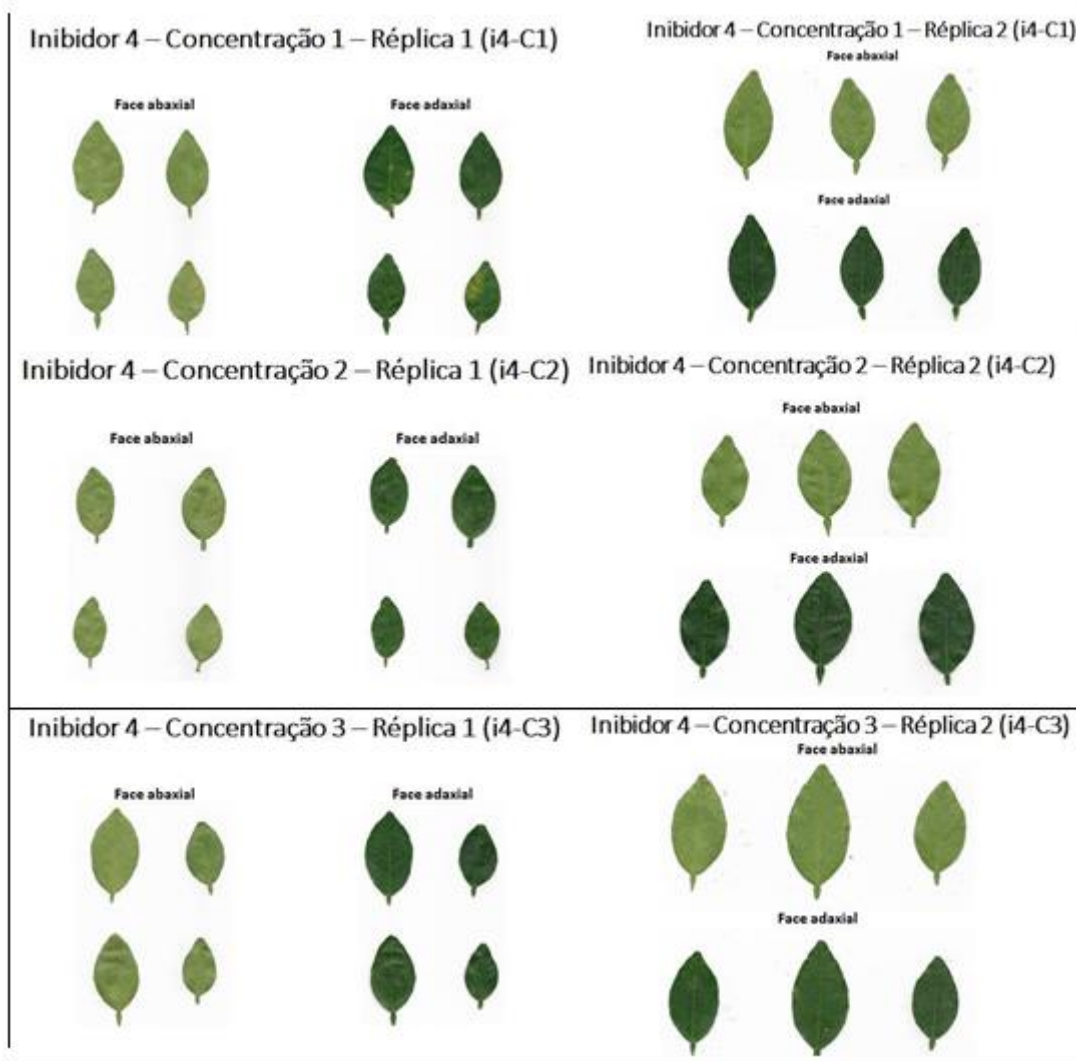


Figura 29

**RESUMO**

“USO DE COMPOSTOS QUÍMICOS INIBIDORES DA FOSFOMANOSE ISOMERASE PARA CONTROLE DO CANCRO CÍTRICO E FITOPATOLOGIAS ASSOCIADAS AO GÊNERO *XANTHOMONAS*”

A presente invenção pertence ao campo da agroquímica e química medicinal. O presente invento refere-se ao uso de compostos químicos como insumos agrícolas promissores e inéditos para o combate ao cancro cítrico que ocorre em plantações de citros, causado pela bactéria *Xanthomonas citri* subsp. *citri* (XAC). O presente invento consiste na elaboração de composições que inibam a fosfomanose isomerase (PMI), que está associada com a patogenicidade de XAC. O invento também apresenta a comprovação funcional de que a PMI é essencial para a patogenicidade de XAC, demonstrando que compostos planejados para inibir seletivamente a PMI são eficazes no controle do cancro cítrico e fitopatologias associadas ao gênero *Xanthomonas*.

**INSTRUÇÕES:**

**A data de vencimento não prevalece sobre o prazo legal. O pagamento deve ser efetuado antes do protocolo. Órgãos públicos que utilizam o sistema SIAFI devem utilizar o número da GRU no campo Número de Referência na emissão do pagamento. Serviço: 200-Pedido nacional de Invenção, Modelo de Utilidade, Certificado de Adição de Invenção e entrada na fase nacional do PCT**

**Clique aqui e pague este boleto através do Auto Atendimento Pessoa Física.**

**Clique aqui e pague este boleto através do Auto Atendimento Pessoa Jurídica.**

Recibo do Pagador

**BANCO DO BRASIL** | 001-9 | 00190.00009 02940.916196 34057.583170 2 86200000007000

Nome do Pagador/CPF/CNPJ/Endereço				
FUNDAÇÃO UNIVERSIDADE FEDERAL DE SÃO CARLOS CPF/CNPJ: 45358058000140				
RODOVIA WASHINGTON LUIZ KM 235 13565 905 MONJOLINHO, SÃO CARLOS -SP CEP:13565905				
Sacador/Avalista				
Noosso-Número	Nr. Documento	Data de Vencimento	Valor do Documento	(=) Valor Pago
29409161934057583	29409161934057583	14/05/2021	70,00	
Nome do Beneficiário/CPF/CNPJ/Endereço				
INSTITUTO NACIONAL DA PROPRIEDADE INDUST CPF/CNPJ: 42.521.088/0001-37				
RUA MAYRINK VEIGA 9 24 ANDAR ED WHITE MARTINS , RIO DE JANEIRO - RJ CEP: 20090910				
Agência/Código do Beneficiário			Autenticação Mecânica	
2234-9 / 333028-1				

**BANCO DO BRASIL** | 001-9 | 00190.00009 02940.916196 34057.583170 2 86200000007000

Local de Pagamento					Data de Vencimento	
<b>PAGÁVEL EM QUALQUER BANCO ATÉ O VENCIMENTO</b>					14/05/2021	
Nome do Beneficiário/CPF/CNPJ					Agência/Código do Beneficiário	
INSTITUTO NACIONAL DA PROPRIEDADE INDUST CPF/CNPJ: 42.521.088/0001-37					2234-9 / 333028-1	
Data do Documento	Nr. Documento	Espécie DOC	Aceite	Data do Processamento	Nosso-Número	
15/04/2021	29409161934057583	DS	N	15/04/2021	29409161934057583	
Uso do Banco	Carteira	Espécie	Quantidade	xValor	(=) Valor do Documento	
29409161934057583	17	R\$			70,00	
Informações de Responsabilidade do Beneficiário					(-) Desconto/Abatimento	
A data de vencimento não prevalece sobre o prazo legal.						
O pagamento deve ser efetuado antes do protocolo.						
Órgãos públicos que utilizam o sistema SIAFI devem utilizar o número da GRU n					(+ Juros/Multa	
o campo Número de Referência na emissão do pagamento.						
Serviço: 200-Pedido nacional de Invenção, Modelo de Utilidade, Certificado de						
Adição de Invenção e entrada na fase nacional do PCT					(-) Valor Cobrado	

Nome do Pagador/CPF/CNPJ/Endereço					Código de Baixa	
FUNDAÇÃO UNIVERSIDADE FEDERAL DE SÃO CARLOS CPF/CNPJ: 45358058000140					Autenticação Mecânica -	
RODOVIA WASHINGTON LUIZ KM 235 13565 905 MONJOLINHO,					Ficha de Compensação	
SAO CARLOS-SP CEP:13565905						
Sacador/Avalista						





## Cobrança / Títulos

G335161044716656007  
16/04/2021 10:49:22

SISBB - SISTEMA DE INFORMACOES BANCO DO BRASIL  
16/04/2021 AUTOATENDIMENTO 10.49.22  
189701897 0002

## PAGAMENTO DE CONTAS NO CARTAO

CLIENTE: FUNDAÇÃO DE AMPARO A PESQ  
AGENCIA: 1897-X CONTA: 22.000-0

-----  
NUMERO DO DOCUMENTO: 000000  
VALOR SOLICITADO: 70,00  
QUANTIDADE DE PARCELAS MENSAS: 01  
VALOR TOTAL DA OPERACAO: 70,00  
VALOR ESTIMADO DAS PARCELAS: 70,00  
TAXA MENSAL DE JUROS ESTIMADA: 0,00%  
TAXA ANUAL DE JUROS ESTIMADA: 0,00%

PAGAMENTO SERA COBRADO NA FATURA DE 01/06/2021

-----  
O VALOR SOLICITADO DESTINA-SE, EXCLUSIVAMENTE, A  
LIQUIDACAO DE PAGAMENTOS.  
-----

16/04/2021 - BANCO DO BRASIL - 10:49:22  
189701897 0002

## COMPROVANTE DE PAGAMENTO DE TITULOS

CLIENTE: FUNDAÇÃO DE AMPARO A PESQ  
CARTAO.: 4674 \*\*\*\* \* 4870

=====

BANCO DO BRASIL

-----

00190000090294091619634057583170286200000007000

BENEFICIARIO:  
INSTITUTO N P I - INPI  
NOME FANTASIA:  
INSTITUTO NACIONAL DA PROPRIEDADE I  
CNPJ: 42.521.088/0001-37  
PAGADOR:  
FUNDAÇÃO UNIVERSIDADE FEDERAL DE SA  
CNPJ: 45.358.058/0001-40

-----  
NR. DOCUMENTO 41.659  
NOSSO NUMERO 29409161934057583  
CONVENIO 02940916  
DATA DE VENCIMENTO 14/05/2021  
DATA DO PAGAMENTO 16/04/2021  
VALOR DO DOCUMENTO 70,00  
VALOR COBRADO 70,00  
=====

Pagamento com cartao de credito  
sujeito a cobranca de encargos vigentes na data  
da liquidacao.

NR.AUTENTICACAO F.905.E1A.EA2.2A9.E62

-----  
Central de Atendimento BB  
4004 0001 Capitais e regioes metropolitanas  
0800 729 0001 Demais localidades.  
Consultas, informacoes e servicos transacionais.

SAC BB  
0800 729 0722  
Informacoes, reclamacoes, cancelamento de  
produtos e servicos.

**Ouvidoria**

**0800 729 5678**

**Reclamacoes nao solucionadas nos canais habituais agencia, SAC e demais canais de atendimento.**

**Atendimento a Deficientes Auditivos ou de Fala**

**0800 729 0088**

**Informacoes, reclamacoes, cancelamento de cartao, outros produtos e servicos de Ouvidoria.**

---

Transação efetuada com sucesso por: JD428496 RAFAEL VIDAL AROCA.



UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
Programa de Pós-Graduação em Biotecnologia



**“ANÁLISE FUNCIONAL DOS GENES *XANB* E *XYLA2* DE *XANTHOMONAS CITRI* SUBSP.  
*CITRI*, AGENTE CAUSAL DO CANCRO CÍTRICO”**

**André Vessoni Alexandrino\***

Tese apresentada como parte dos requisitos para obtenção do título de DOUTOR EM BIOTECNOLOGIA do Programa de Pós-Graduação em Biotecnologia da Universidade Federal de São Carlos.

***Orientadora:***

Profa. Dra. Maria Teresa Marques Novo Mansur

**\*Bolsista CAPES – 88882.426495/2019-01**

**São Carlos – SP  
2020**

ANDRÉ VESSONI ALEXANDRINO

**“Análise funcional dos genes *xanB* e *xylA2* de *Xanthomonas citri* subsp. *citri*, agente causal do cancro cítrico”**

Tese apresentada como parte dos requisitos para obtenção do título de DOUTOR EM BIOTECNOLOGIA do Programa de Pós-Graduação em Biotecnologia da Universidade Federal de São Carlos.

***Orientadora:***

Profa. Dra. Maria Teresa Marques Novo Mansur

BANCA EXAMINADORA

*Profa. Dra. Dulce Helena Ferreira de Souza*  
Departamento de Química – UFSCar

*Prof. Dr. Flávio Henrique Silva*  
Departamento de Genética e Evolução – UFSCar

*Prof. Dr. Franklin Behlau*  
Fundo de Defesa da Citricultura – Fundecitrus

*Prof. Dr. Henrique Ferreira*  
Departamento de Bioquímica e Microbiologia (Instituto de Biociências) – UNESP Rio Claro

*Profa. Dra. Maria Teresa Marques Novo Mansur*  
Departamento de Genética e Evolução – UFSCar

São Carlos – SP  
2020



## Agradecimentos

À minha esposa, Taís, pelo companheirismo, cumplicidade, amizade, carinho, amor, compreensão e paciência, **muita paciência**.

Aos meus pais e avós, pelo carinho, amor, incentivos, conselhos e apoio em todas as fases e decisões da minha vida.

Ao meu irmão, Daniel, pela cumplicidade, amizade, amor e confiança.

Aos meus amigos Alberto Zanata, Douglas Álvaro Dias e Marcos Bravi, pela amizade, parceria e por me proporcionarem agradabilíssimas conversas.

À Profa. Dra. Maria Teresa Marques Novo Mansur, pela orientação, oportunidade, confiança e apoio, que foram essenciais para realização deste trabalho. Agradeço, também, pela relação de amizade que desenvolvemos ao longo desses mais de oito anos.

Ao Dr. Leandro Seiji Goto, pelo apoio e ensinamentos.

Ao Dr. Humberto D’Muniz Pereira (IFSC/USP), pela imensa colaboração nos estudos de estrutura de proteínas.

A todos os professores que disponibilizaram seus equipamentos e espaço, especialmente aos professores Dra. Dulce Helena Ferreira de Souza, Dr. Felipe Roberti Teixeira, Dr. Flávio Henrique da Silva, Dr. Gilberto Moraes, Dr. Iran Malavazi, Dr. Marco Antonio Del Lama e Dra. Silvia Nassif Del Lama.

Ao FUNDECITRUS, pela disponibilização de espaço para realização dos ensaios *in vivo*, especialmente ao Dr. Franklin Behlau e à Me. Tamiris Garcia.

Ao Prof. Dr. Henrique Ferreira e à Me. Lúcia Bonci Cavalca pela colaboração nos experimentos *in vivo* na UNESP – Rio Claro.

Aos colegas de laboratório, Beatriz, Bianca, Carol, Evandro, Inaê, Jéssica, Mariana, Naiane, Nicole, Solange, Vinícius e Yuri pela colaboração, amizade, parceria e momentos de descontração.

À Célia Regina Câmara, pela amizade e por todo o apoio na realização dos experimentos.

Aos meus colegas de trabalho da UNICEP e Collegium Sapiens, pela amizade e paciência nos meus momentos de hiperatividade exigidos pelo doutoramento.

À FAPESP e CAPES, pelo auxílio financeiro ao projeto.

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

Alexandrino, André Vessoni. **Análise funcional dos genes *xanB* e *xyIA2* de *Xanthomonas citri* subsp. *citri*, agente causal do cancro cítrico**. 2020. Tese de Doutorado (Programa de Pós-Graduação em Biotecnologia) – UFSCar, São Carlos, 2020.

## Resumo

A bactéria *Xanthomonas citri* subsp. *citri* (Xcc) é a principal causadora do cancro cítrico, uma doença que acomete todas as cultivares cítricas de interesse comercial, ocasionando prejuízos ao setor citricultor devido à queda na produtividade e qualidade dos frutos, e à ausência de medidas eficazes de controle e cura. Em trabalhos de análise proteômica diferencial previamente conduzidos por nosso grupo de pesquisa, fosfomanose isomerase (PMI) e xilose isomerase (XI) foram proteínas identificadas como potencialmente envolvidas na patogênese do cancro cítrico, sendo a PMI detectada pela primeira vez na superfície de Xcc. A PMI de Xcc é predita como uma enzima bifuncional com atividades catalíticas de interconversão de D-manose-6-fosfato e D-frutose-6-fosfato, bem como de D-manose-1-fosfato em GDP-D-manose. A XI de Xcc também é classificada como uma enzima bifuncional que interconverte D-xilose em D-xilulose e D-glicose em D-frutose. O presente trabalho teve como objetivo geral a caracterização funcional dos genes codificantes da PMI e da XI (*xanB* e *xyIA2*, respectivamente), especialmente no tocante à relação com a patogenicidade de Xcc, partindo de estratégias que envolveram a construção de sistemas de expressão heteróloga induzíveis por IPTG (isopropil- $\beta$ -D-tiogalactosídeo) em *E. coli*, os quais permitiram a confirmação das atividades biológicas preditas das proteínas-alvo. A dupla recombinação homóloga entre o DNA genômico bacteriano e o vetor suicida pNPTS138 contendo regiões flanqueadoras do gene-alvo a ser deletado foi utilizada para a construção de linhagens mutantes deletadas nos genes *xanB* e *xyIA2* (Xcc $\Delta$ *xanB* e Xcc $\Delta$ *xyIA2*, respectivamente). Uma linhagem complementante para o primeiro mutante (Xcc $\Delta$ C*xanB*) foi obtida com sucesso pela reinserção do gene no locus original utilizando o pNPTS138, uma metodologia inédita em Xcc. Ensaio *in vivo* em *Citrus aurantifolia* foram conduzidos utilizando as linhagens Xcc, Xcc $\Delta$ *xanB*, Xcc $\Delta$ C*xanB* e Xcc $\Delta$ *xyIA2*, os quais permitiram a avaliação da relação dos genes-alvo com a patogenicidade de Xcc. A deleção do gene *xyIA2* desencadeou um aumento na virulência de Xcc, o que pode ser explicado pelo consequente acúmulo de xilose e um possível aumento da expressão de genes *hrp*. A deleção do gene *xanB*, por sua vez, resultou na perda de patogenicidade de Xcc, com redução de suas propriedades relacionadas ao processo infeccioso, como motilidade, formação de biofilme e resistência à radiação ultravioleta, com todos esses fenótipos sendo integralmente restaurados pela complementação gênica. Sendo assim, a constatação de que o gene *xanB* é essencial para a patogenicidade de Xcc o torna um alvo inédito e bastante promissor do ponto de vista biotecnológico, propiciando uma perspectiva de controle e/ou cura do cancro cítrico por meio da utilização de inibidores da PMI e, também, do desenvolvimento de cultivares cítricas que neutralizem a ação de tal enzima durante o processo infeccioso.

**Palavras-chave:** *Xanthomonas citri* subsp. *citri*; fosfomanose isomerase; xilose isomerase; patogenicidade; virulência.

Alexandrino, André Vessoni. **Functional analysis of the *xanB* and *xylA2* genes of *Xanthomonas citri* subsp. *citri*, causative agent of citrus canker.** 2020. Tese de Doutorado (Programa de Pós-Graduação em Biotecnologia) – UFSCar, São Carlos, 2020.

## Abstract

*Xanthomonas citri* subsp. *citri* (Xcc) is the main bacterium that causes citrus canker, a disease that affects all citrus cultivars of commercial interest, causing damages to the citrus sector due to the decrease in fruit productivity and quality, and the lack of effective control and cure measures. In differential proteomic analysis previously conducted by our research group, phosphomannose isomerase (PMI) and xylose isomerase (XI) were proteins identified as potentially involved in the pathogenesis of citrus canker, the first being detected for the first time on the Xcc surface. Xcc PMI is predicted as a bifunctional enzyme with catalytic activity of interconverting D-mannose-6-phosphate into D-fructose-6-phosphate and D-mannose-1-phosphate to GDP-D-mannose. Xcc XI is also classified as a bifunctional enzyme that interconverts D-xylose into D-xylulose and D-glucose into D-fructose. The present work had as general objective the functional characterization of the coding genes of PMI and XI (*xanB* and *xylA2*, respectively), especially regarding the relationship with the Xcc pathogenicity, starting from strategies that involved the construction of IPTG-inducible heterologous expression systems in *E. coli*, which allowed the confirmation of the predicted biological activities of the target proteins. Homologous double recombination between the genomic DNA and the suicide vector pNPTS138 containing flanking regions of the target gene to be deleted was used to construct mutant strains deleted in the *xanB* and *xylA2* genes (Xcc $\Delta$ *xanB* and Xcc $\Delta$ *xylA2*, respectively). A complementing strain for the first mutant (Xcc $\Delta$ C*xanB*) was successfully obtained by reinserting the gene in the original locus using pNPTS138, an unprecedented methodology in Xcc. In vivo assays in *Citrus aurantifolia* were conducted using the Xcc, Xcc $\Delta$ *xanB*, Xcc $\Delta$ C*xanB* and Xcc $\Delta$ *xylA2* strains, which allowed the evaluation of the relationship of the target genes with the Xcc pathogenicity. Deletion of the *xylA2* triggered an increase in Xcc virulence, which may be explained by the consequent accumulation of xylose and a possible increase in *hrp* gene expression. Deletion of the *xanB* resulted in loss of Xcc pathogenicity, with reduction of its properties related to the infectious process, such as motility, biofilm formation and resistance to ultraviolet radiation, with all these phenotypes being fully restored by gene complementation. Thus, the finding that the *xanB* is essential for the Xcc pathogenicity makes it a novel and promising target from a biotechnological point of view, providing the perspective of control and/or cure of citrus canker through the use of PMI inhibitors and also the development of citrus cultivars that neutralize the action of such protein during the infectious process.

**Keywords:** *Xanthomonas citri* subsp. *citri*; phosphomannose isomerase; xylose isomerase; pathogenicity; virulence.

## Lista de Figuras

### Capítulo 1 – Introdução

<b>Figura 1.1</b> A possível relação metabólica entre PMI e XI	16
<b>Figura 1.2</b> Lesões características do cancro cítrico	20
<b>Figura 1.3</b> Estrutura química da goma xantana	24
<b>Figura 1.4</b> A relação entre a PMI e biossíntese de xantana	25
<b>Figura 1.5</b> Esquema da estrutura geral do xilano	26
<b>Figura 1.6</b> Alinhamento entre sequências peptídicas de PMI	30
<b>Figura 1.7</b> Esquema das reações catalisadas pela xilose isomerase	32
<b>Figura 1.8</b> Contextos genômicos dos genes <i>xyIA</i> em <i>Xcc</i>	33

### Capítulo 2 – Material e Métodos

<b>Figura 2.1</b> Esquema das reações envolvidas no Teste de Seliwanoff	46
<b>Figura 2.2</b> Esquema simplificado do método de detecção de D-glicose	47
<b>Figura 2.3</b> Esquema simplificado da metodologia de deleção gênica	49
<b>Figura 2.4</b> Esquemas simplificados dos plasmídeos de deleção	50
<b>Figura 2.5</b> Metodologia para seleção dos possíveis mutantes deletados	54
<b>Figura 2.6</b> Representação do vetor de complementação pLAL6_CxanB	59
<b>Figura 2.7</b> Representação dos oligonucleotídeos utilizados na confirmação da orientação da clonagem no vetor pNPTS138_CxanB	61
<b>Figura 2.8</b> Exemplos de <i>Citrus aurantifolia</i> , utilizadas como hospedeiro no ensaio de patogenicidade <i>in vivo</i> por infiltração	63
<b>Figura 2.9</b> Ensaio de patogenicidade <i>in vivo</i> das linhagens selvagem, mutantes de deleção e complementante por aspersão	64

### Capítulo 3 – Resultados e Discussão

<b>Figura 3.1</b> Análise dos produtos de amplificação por PCR a partir de DNA genômico de <i>Xcc</i>	68
<b>Figura 3.2</b> Análise de restrição dos DNAs plasmidiais para confirmação da clonagem dos fragmentos de interesse no vetor de propagação pJET 1.2	69
<b>Figura 3.3</b> Análise de restrição dos DNAs plasmidiais para confirmação da clonagem das regiões codificadoras da PMI e da XI nos vetores de expressão da série pET	70
<b>Figura 3.4</b> Análise da expressão, solubilidade e purificação da XccPMI por SDS-PAGE	71

<b>Figura 3.5</b> Análise da expressão, solubilidade e purificação da XccPMI_C sob SDS-PAGE	72
<b>Figura 3.6</b> Análise da expressão, solubilidade e purificação da XccGST_PMI por SDS-PAGE	74
<b>Figura 3.7</b> Análise da expressão, solubilidade e purificação <i>in tandem</i> da XccGST_PMI por SDS-PAGE	75
<b>Figura 3.8</b> Análise da expressão, solubilidade e purificação da XccXI por SDS-PAGE	76
<b>Figura 3.9</b> Avaliação da atividade enzimática da XccPMI_C	77
<b>Figura 3.10</b> Avaliação da atividade enzimática da XccGST_PMI	78
<b>Figura 3.11</b> Avaliação da atividade enzimática da XccXI	79
<b>Figura 3.12</b> Avaliação da atividade enzimática da XccXI	80
<b>Figura 3.13</b> Análise de restrição do vetor de deleção pNPTS138_xanB	81
<b>Figura 3.14</b> Análise de restrição do vetor de deleção pNPTS138_xylA	82
<b>Figura 3.15</b> Confirmação da deleção do gene <i>xanB</i> por PCR	84
<b>Figura 3.16</b> Confirmação da deleção do gene <i>xylA2</i> por PCR	84
<b>Figura 3.17</b> Confirmação da depleção da síntese da PMI por <i>Western blot</i>	85
<b>Figura 3.18</b> Análise de restrição do vetor de complementação pLAL6_xanB	86
<b>Figura 3.19</b> <i>Screening</i> das colônias transformantes quanto à capacidade de degradar amido	87
<b>Figura 3.20</b> Construção do vetor de complementação pNPTS_CxanB e confirmação da orientação da clonagem do gene <i>xanB</i>	88
<b>Figura 3.21</b> Confirmação da complementação gênica do mutante XccΔxanB	89
<b>Figura 3.22</b> Curva de crescimento das linhagens Xcc e XccΔxanB	90
<b>Figura 3.23</b> Ensaio de patogenicidade <i>in vivo</i> por aspersão de Xcc, XccΔxanB e XccΔCxanB em <i>Citrus aurantifolia</i>	92
<b>Figura 3.24</b> Ensaio de patogenicidade <i>in vivo</i> por infiltração de Xcc, XccΔxanB e XccΔCxanB em <i>Citrus aurantifolia</i>	94
<b>Figura 3.25</b> Ensaio de motilidade	96
<b>Figura 3.26</b> Ensaio de formação de biofilme	98
<b>Figura 3.27</b> Ensaio de resistência à radiação ultravioleta	99
<b>Figura 3.28</b> Ensaio de patogenicidade <i>in vivo</i> por infiltração de Xcc e XccΔxylA2 em <i>Citrus aurantifolia</i>	101

## Lista de Tabelas

### Capítulo 2 – Material e Métodos

<b>Tabela 2.1</b> Desenho dos oligonucleotídeos	37
<b>Tabela 2.2</b> Pares de oligonucleotídeos e seus respectivos produtos de PCR (pPCR), funções e vetores nos quais foram clonados	39
<b>Tabela 2.3</b> Variações de posicionamento da cauda de histidinas e fusão à glutathione-S-transferase (GST) na fosfomanose isomerase recombinante utilizando diferentes combinações de oligonucleotídeos	40
<b>Tabela 2.4</b> Variações de posicionamento da cauda de histidinas na xilose isomerase recombinante utilizando diferentes combinações de oligonucleotídeos	40
<b>Tabela 2.5</b> Ciclos de PCR utilizados na amplificação dos fragmentos gênicos	42
<b>Tabela 2.6</b> Etapas de PCR utilizadas na confirmação das deleções gênicas	56

## Lista de Abreviaturas e Siglas

2D-PAGE	" <i>Two-dimensional polyacrylamide gel electrophoresis</i> ", eletroforese bidimensional em gel de poliacrilamida.
BSA	" <i>Bovine Serum Albumin</i> ", albumina de soro bovino.
DO <sub>595 nm</sub>	Densidade óptica em 595 nm.
dNTP	Trifosfato de desoxinucleotídeo.
EDTA	" <i>Ethylenediamine tetraacetic acid</i> ", ácido etilenodiamino tetraacético.
FAO	" <i>Food and Agriculture Organization</i> ", Organização das Nações Unidas para a Alimentação e Agricultura.
Fundecitrus	Fundo de Defesa da Citricultura.
GenBank	Banco de dados de sequências de DNA e proteínas no NCBI.
GOD	Glicose oxidase.
GST	Glutathione-S-transferase.
IMAC	" <i>Immobilized metal affinity chromatography</i> ", cromatografia de afinidade por íons metálicos imobilizados.
IPTG	Isopropil- $\beta$ -D-tiogalactosídeo.
kb	Quilopares de bases.
kDa	Quilodalton.
LB	Meio de cultura Lúria-Bertani.
NCBI	" <i>National Center for Biotechnology Information</i> ", Centro Nacional de Informação Biotecnológica.
ORF	" <i>Open Reading Frame</i> ", quadro aberto de leitura.
pb	Pares de bases.
PBS	" <i>Phosphate Buffer Saline</i> ", tampão fosfato de sódio.
PCR	" <i>Polymerase Chain Reaction</i> ", Reação em cadeia da DNA polimerase.
PDB	" <i>Protein Data Bank</i> ", banco de dados de estruturas resolvidas de proteínas.



PMI	Fosfomanose isomerase.
POD	Peroxidase.
pPCR	Produtos de PCR.
rpm	Rotações por minuto.
SDS-PAGE	" <i>Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis</i> ", eletroforese em gel de poliacrilamida em presença de dodecilsulfato de sódio.
SAA	Secretaria de Agricultura e Abastecimento.
SMR	Sistema de Mitigação de Riscos.
TAE	Tampão Tris-Acetato-EDTA.
TM	Temperatura de <i>melting</i> ou temperatura de transição de fase.
UFC	Unidades formadoras de colônia.
XAM-M	Meio de cultura XAM-1 modificado, conhecido por ser indutor de patogenicidade.
XauB	<i>Xanthomonas fuscans</i> subsp. <i>aurantifolii</i> tipo B.
XauC	<i>Xanthomonas fuscans</i> subsp. <i>aurantifolii</i> tipo C.
Xcc	<i>Xanthomonas citri</i> subsp. <i>citri</i> linhagem 306.
XccΔxanB	Linhagem mutante de <i>Xanthomonas citri</i> subsp. <i>citri</i> com deleção do gene que codifica a fosfomanose isomerase.
XccΔCxanB	Linhagem complementante do mutante de <i>Xanthomonas citri</i> subsp. <i>citri</i> com deleção do gene que codifica a fosfomanose isomerase.
XccΔxylA	Linhagem mutante de <i>Xanthomonas citri</i> subsp. <i>citri</i> com deleção do gene que codifica a xilose isomerase.
XccGST_PMI	Fosfomanose isomerase recombinante de <i>Xanthomonas citri</i> subsp. <i>citri</i> fusionada à glutationa-S-transferase na região N-terminal.
XccPMI	Fosfomanose isomerase recombinante de <i>Xanthomonas citri</i> subsp. <i>citri</i> com cauda de histidinas N-terminal.
XccPMI_C	Fosfomanose isomerase recombinante de <i>Xanthomonas citri</i> subsp. <i>citri</i> com cauda de histidinas C-terminal.
XccXI	Xilose isomerase recombinante de <i>Xanthomonas citri</i> subsp. <i>citri</i> com cauda de histidinas N-terminal.
XI	Xilose isomerase.

## SUMÁRIO

<b>Resumo</b>	4
<b>Abstract</b>	5
<b>Lista de Figuras</b>	6
<b>Lista de Tabelas</b>	8
<b>Lista de Abreviaturas e Siglas</b>	9
<b>Capítulo 1 – Introdução</b>	15
1.1 Origem, controle e incidências do cancro cítrico	17
1.2 Características gerais e ciclo do cancro cítrico	19
1.3 O gênero <i>Xanthomonas</i>	22
1.4 Goma xantana	23
1.5 Xilano: estrutura e vias de degradação	25
1.6 Estudos genômicos e pós-genômicos	28
1.7 Genes-alvo e seus produtos	29
1.7.1 <i>xanB</i>	29
1.7.2 <i>xyIA</i>	31
1.8 Objetivos	34
1.8.1 Objetivo geral	34
1.8.2 Objetivos específicos	34
<b>Capítulo 2 – Material e Métodos</b>	35
2.1 Linhagens bacterianas, meios e condições de cultivo	35
2.2 Enzimas, kits e procedimentos em biologia molecular	35
2.3. Desenho e síntese de oligonucleotídeos	35
2.4 Extração de DNA genômico	41
2.5 Amplificação e clonagem dos fragmentos gênicos de interesse	41

<b>2.6</b>	Construção dos sistemas de expressão	43
<b>2.7</b>	Expressão heteróloga e purificação por cromatografia de afinidade	43
<b>2.8</b>	Ensaio de atividade enzimática	45
<b>2.8.1</b>	Avaliação da atividade enzimática da PMI	45
<b>2.8.2</b>	Avaliação da atividade enzimática da XI	46
<b>2.9</b>	Deleções gênicas	48
<b>2.9.1</b>	Construção dos vetores de deleção	48
<b>2.9.2</b>	Preparação de células de Xcc eletrocompetentes	51
<b>2.9.3</b>	Transformação de Xcc por eletroporação	51
<b>2.9.4</b>	Seleção de mutantes deletados	52
<b>2.9.5</b>	Confirmação das deleções dos genes-alvo	55
<b>2.10</b>	Confirmação da depleção da síntese de PMI no mutante de deleção	56
<b>2.11</b>	Complementação gênica com o vetor pLAL6	57
<b>2.11.1</b>	Construção do vetor de complementação pLAL6_CxanB	57
<b>2.11.2</b>	Seleção de complementantes	59
<b>2.12</b>	Complementação gênica com o vetor pNPTS138	60
<b>2.12.1</b>	Construção do vetor de complementação pNPTS138_CxanB	60
<b>2.12.2</b>	Confirmação da orientação de clonagem de <i>xanB</i> no pNPTS138_CxanB	61
<b>2.12.3</b>	Seleção de possíveis complementantes e confirmação da complementação gênica	62
<b>2.13</b>	Estudos funcionais das linhagens mutantes e complementante	62
<b>2.13.1</b>	Testes de patogenicidade em <i>Citrus aurantifolia</i>	62
<b>2.13.2</b>	Curvas de crescimento	64
<b>2.13.3</b>	Ensaio de formação de biofilme	65

2.13.4	Avaliação de motilidade	65
2.13.5	Sobrevivência à radiação ultravioleta	66
<b>Capítulo 3</b>	<b>– Resultados e Discussão</b>	<b>67</b>
3.1	Amplificação por PCR e clonagem	67
3.2	Expressão heteróloga e purificação das proteínas recombinantes	69
3.2.1	Construção dos vetores de expressão	69
3.2.2	Análise da expressão heteróloga e da purificação por cromatografia de afinidade	71
3.3	Avaliação das atividades enzimáticas das proteínas recombinantes	76
3.4	Deleção gênica	80
3.4.1	Construção dos vetores de deleção	80
3.4.2	Confirmação das deleções gênicas	82
3.4.3	<i>Xcc</i> Δ <i>xanB</i> é incapaz de sintetizar a PMI	85
3.5	Complementação gênica utilizando o vetor pLAL6	85
3.5.1	Construção do vetor pLAL6_ <i>CxanB</i>	85
3.5.2	<i>Screening</i> de possíveis complementantes	86
3.6	Complementação gênica utilizando o vetor pNPTS138	87
3.6.1	Construção do vetor pNPTS_ <i>CxanB</i>	87
3.6.2	Confirmação da complementação gênica	89
3.7	Caracterização fenotípica das linhagens mutantes e complementante	90
3.7.1	Deleção de <i>xanB</i> não afeta o perfil de crescimento	90
3.7.2	<i>xanB</i> é essencial para a patogenicidade de <i>Xcc</i>	91
3.7.3	<i>xanB</i> está envolvido na motilidade, formação de biofilme e resistência à radiação UV	95
3.7.4	A deleção do gene <i>xyIA2</i> aumenta a virulência de <i>Xcc</i>	100

<b>Capítulo 4 – Conclusões</b>	105
<b>Referências</b>	106

## Capítulo 1 – Introdução

O Brasil é o maior produtor mundial de laranja, com uma produção estimada em mais de 17 milhões de toneladas em 2017 (FAO STAT, 2017), sendo o estado de São Paulo responsável por 78,7% de toda a produção nacional (IBGE, 2018). Em 2016 as exportações de suco de laranja concentrado somaram US\$ 1,43 bilhão, valor que colocou tal *commoditie* na décima posição no *ranking* de exportações brasileiras (FAO STAT, 2017).

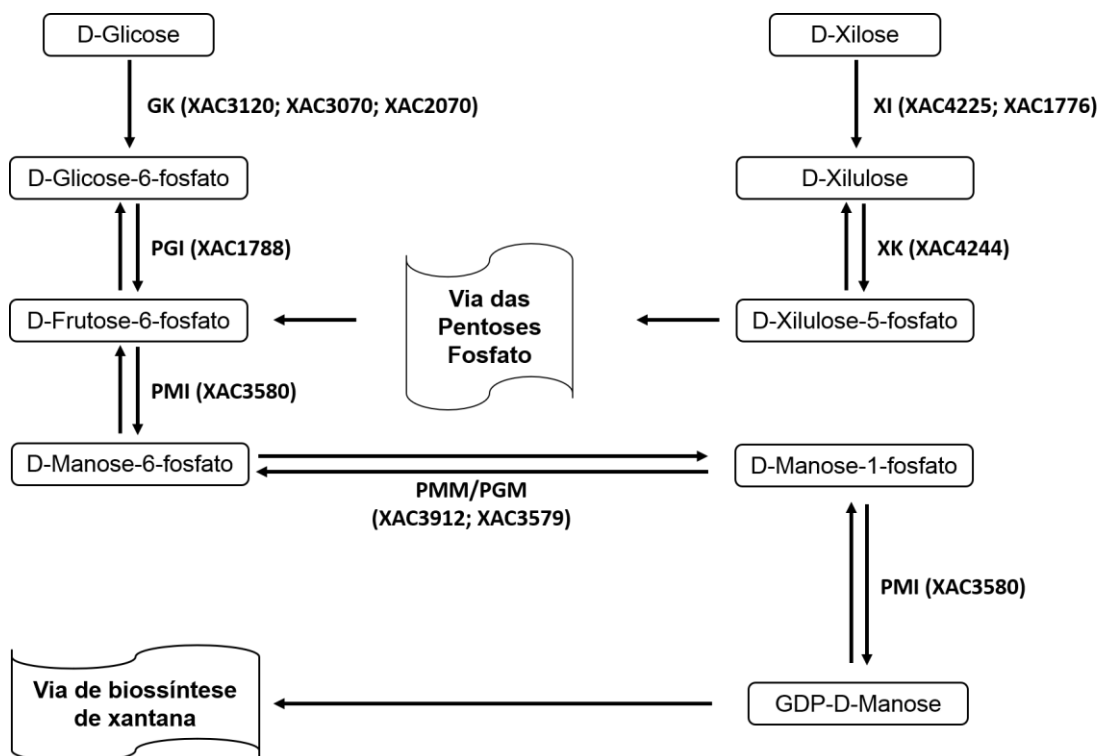
A produção de citros influencia também o mercado interno brasileiro, uma vez que em 2016 o PIB gerado por todos os elos da cadeia produtiva cítrica foi de US\$ 6,5 bilhões e, além disso, a citricultura gerou 200 mil empregos diretos e indiretos (NEVES; TROMBIN, 2017). Entretanto, o potencial econômico da citricultura é limitado, em parte, pelas doenças que afetam a sua produtividade, como o cancro cítrico, causado principalmente pela bactéria *Xanthomonas citri* subsp. *citri* (Xcc) (KOLLER, 1994; LANZA et al., 2018), que é uma doença causadora de prejuízos ao setor citricultor devido à queda na produtividade e qualidade dos frutos e à ausência de medidas eficazes de controle e cura (FERENCE et al., 2018).

Desde 1995, o cancro cítrico vem se alastrando de maneira alarmante, principalmente no Brasil e nos EUA, evidenciando a suscetibilidade dos citros frente a esse fitopatógeno (BEHLAU; FONSECA; BELASQUE JUNIOR, 2016). Até 2009, o manejo da doença era realizado por meio de medidas preventivas com a erradicação de plantas sintomáticas e suspeitas, concomitantemente ao uso de inseticidas para controle do minador dos citros, um inseto facilitador da infecção por Xcc (BEHLAU; FONSECA; BELASQUE JUNIOR, 2016; SANCHES et al., 2014).

Porém, o abrandamento da legislação e a extinção do programa de erradicação de plantas contaminadas ocorridos a partir de 2013, juntamente com a ineficiência dos setores de fiscalização, refizeram do cancro cítrico um problema para o setor citricultor, uma vez que a incidência da doença em 2018 atingiu preocupantes 16,43% dos talhões contaminados (BEHLAU; BELASQUE JUNIOR, 2014; FUNDECITRUS, 2018). Assim, independentemente da adoção de uma estratégia de manejo integrado ou de erradicação, atualmente não há estratégias eficientes no controle ou cura do cancro cítrico (FERENCE et al., 2018).

Diante do contexto apresentado, estudos que contribuam para um melhor entendimento da infecção bacteriana, da patogênese do cancro cítrico e,

adicionalmente, que subsidiem o desenvolvimento de insumos de controle e cura da doença tornam-se necessários e valiosos. Dessa forma, é nesse cenário que se insere a presente tese de doutorado, utilizando-se da biologia molecular como uma ferramenta para caracterizar funcionalmente os genes codificantes das proteínas fosfomanose isomerase e xilose isomerase (genes *xanB* e *xylA2*, respectivamente), as quais foram previamente identificadas por nosso grupo como potencialmente relacionadas à patogenicidade de Xcc (CARNIELLI et al., 2017; ZANDONADI et al., dados não publicados). Os genes-alvo deste trabalho podem ter relação do ponto de vista metabólico (Figura 1.1), embora tal relação ainda necessite de investigação, não sendo ainda totalmente compreendida.



**Figura 1.1** A possível relação metabólica entre PMI e XI. A via das pentoses fosfato pode ser utilizada para correlacionar as funções metabólicas da PMI e da XI, proteínas-alvo deste trabalho. GK: glicocinase. PGI: fosfoglicose isomerase. PMM/PGM: fosfomanomutase/fosfoglicomutase. XK: xiluloquinase. Fonte: próprio autor.

Para isso, foram produzidas linhagens mutante e complementante para o gene *xanB*, além de uma linhagem mutante para o gene *xylA2*. Estas linhagens foram testadas quanto à virulência em *Citrus aurantifolia*, bem como o mutante para *xanB* foi caracterizado *in vitro* quanto a fenótipos relacionados à patogenicidade, relativamente à linhagem selvagem, o que permitiu a avaliação do papel dos genes *xanB* e *xylA2* no processo infeccioso de Xcc.

Além disso, foram construídos sistemas de expressão heteróloga e de purificação das enzimas fosfomanose isomerase (PMI) e xilose isomerase (XI), os quais possibilitaram a confirmação das atividades enzimáticas preditas e, também, a produção de anticorpos contra tais moléculas. Devido aos resultados obtidos neste trabalho, a enzima recombinante PMI está sendo utilizada em ensaios com inibidores obtidos por modelagem estrutural *in silico* da PMI de *Xcc* (BARCELOS et al., 2020), para avaliação da potencialidade desses compostos como novos insumos de controle do cancro cítrico.

### **1.1 Origem, controle e incidências do cancro cítrico**

O cancro cítrico constitui uma das mais importantes doenças da citricultura e a sua origem é ainda incerta. Entretanto, o Sul da China, a Indonésia e a Índia são os locais mais prováveis para o surgimento da doença, a partir dos quais houve a disseminação para os outros continentes (DAS, 2003). Nos EUA, um dos principais países produtores de citros, o cancro cítrico foi identificado pela primeira vez em 1912, sendo o sudeste do país a porta de entrada para a doença, a partir de mudas contaminadas oriundas do Japão (DOPSON, 1964).

Embora a doença tenha sido considerada erradicada nos EUA 21 anos após a sua identificação, a doença reapareceu em meados da década de 1980 na Flórida, sendo novamente declarada erradicada no ano de 1994 (STALL; CIVEROLO, 1991). Porém, em 1995, uma área residencial de 36,3 km<sup>2</sup> próxima ao Aeroporto Internacional de Miami foi infectada pela doença e, desde 1997, novos surtos da doença vêm acometendo os EUA, afetando a produtividade do setor citricultor norte-americano (GOTTWALD; GRAHAM; SCHUBERT, 2002).

No Brasil, o berço do cancro cítrico foi o Estado de São Paulo, especificamente o município de Presidente Prudente, no ano de 1957 (BITANCOURT, 1957). Desde o seu surgimento, a doença se alastrou para diversos estados brasileiros, como Mato Grosso do Sul, Paraná, Santa Catarina e Rio Grande do Sul (NAMEKATA; ROSSI; CERAVOLO, 1996).

Em decorrência da importância econômica da citricultura para o Estado de São Paulo, a Secretaria de Agricultura e Abastecimento (SAA) atuou em campanhas de erradicação e manejo do cancro cítrico desde a década de 50 e, a partir de 1984, em



parceria com o Fundecitrus, mantendo a doença sob controle em boa parte deste período (BEHLAU; FONSECA; BELASQUE JUNIOR, 2016).

Durante muitos anos, o controle da doença foi alcançado por meio de um programa de erradicação de plantas contaminadas e, também, de plantas sem sintomas localizadas nas imediações (inicialmente 12 m e, depois, 30 m de raio), bem como pela instauração de inspeções periódicas (BEHLAU; FONSECA; BELASQUE JUNIOR, 2016). Retrato da eficiência do programa de erradicação do cancro cítrico foi a taxa de incidência mensurada no ano de 2009, com apenas 0,14% dos talhões contaminados (BELASQUE JUNIOR et al., 2010).

Em 2009, produtores de áreas mais atingidas pelo cancro cítrico pressionaram a SAA com o objetivo de modificar as diretrizes do programa de erradicação da doença, o que foi acatado. Então, de 2009 a 2013, o nível de incidência deixou de ser considerado um parâmetro do programa de erradicação, fazendo com que os produtores localizados em áreas de baixa incidência também erradicassem as plantas contaminadas e as assintomáticas localizadas em um raio de 30 m (BEHLAU; BELASQUE JUNIOR, 2014).

A mudança na legislação fez com que o Fundecitrus rescindisse o acordo com a SAA, sob a alegação de que o número de inspetores teria de ser vertiginosamente aumentado para atender o Estado todo e manter a doença sob controle. Como resultado, somente a SAA passou a regulamentar a erradicação pelos produtores das plantas contaminadas e, em 2012, o índice de talhões contaminados aumentou quase dez vezes em relação à 2009, chegando a 1,39% (BEHLAU; BELASQUE JUNIOR, 2014).

A partir de 2013, a legislação foi modificada no sentido de desobrigar esta erradicação de plantas suspeitas e, em 2017, uma nova resolução instituída pela SAA do Estado de São Paulo passou a permitir a presença de plantas sintomáticas nos pomares, exigindo apenas um Sistema de Mitigação de Riscos (SMR) para o controle da doença (SÃO PAULO, 2013, 2017).

Após o abrandamento da legislação em relação ao manejo do cancro cítrico, houve um demasiado e preocupante aumento na incidência da doença nos últimos dez anos. Em 2009, o índice de talhões contaminados no cinturão citrícola era de 0,14%, valor que permitia a classificação da doença como “controlada” (SANCHES et al., 2014). Porém, o último levantamento, realizado em 2018, mostrou que 16,43% dos talhões apresentaram sintomas da doença (FUNDECITRUS, 2018).

Atualmente, uma das principais estratégias de mitigação da incidência da doença é a pulverização dos pomares infectados com uma solução de cobre metálico 0,1%. Contudo, em elevadas concentrações, metais pesados como o cobre são tóxicos aos microrganismos e aos ecossistemas e sua utilização tem levado à seleção de linhagens de Xcc resistentes ao cobre (BEHLAU et al., 2011), nas quais foram inclusive identificados genes relacionados a essa resistência (TEIXEIRA et al., 2008). Tal evento pode requerer a utilização de concentrações crescentes de cobre, o que é fator ainda maior de preocupação ambiental.

## 1.2 Características gerais e ciclo do cancro cítrico

Após atingirem os tecidos vegetais, as bactérias causadoras do cancro cítrico penetram principalmente pelas folhas por meio de aberturas naturais, como os estômatos e os hidatódios, e também por ferimentos causados pelo minador dos citros, por espinhos, pelo contato de máquinas e equipamentos e pelo atrito provocado por grãos de areia transportados pelo vento (GRAHAM et al., 2000).

Uma vez posicionada no apoplasto, Xcc reconhece as células hospedeiras por meio do *pilus* Hrp ou do *pilus* tipo IV, componentes do Sistema de Secreção Tipo III (T3SS, *type IV secretion system*), deixam de ser planctônicas e iniciam a fase de multiplicação e agregação como consequência da formação de biofilme (BRUNINGS; GABRIEL, 2003; GRAHAM et al., 2004).

Em seguida, inicia-se o ataque às células hospedeiras, também por meio do T3SS, que resulta na translocação de fatores de virulência como a PthA, proteína que atua como reguladora da expressão do gene *upa20* da planta hospedeira, elicitando uma resposta de hipertrofia e hiperplasia (KAY et al., 2007; SCHUBERT et al., 2001). Como consequência do aumento do volume e número das células, os espaços intercelulares são diminuídos e passam a ser preenchidos pela goma xantana, polissacarídeo produzido por Xcc durante a formação de biofilme, que, por ser extremamente higroscópico, causa um dos sintomas característicos do cancro cítrico, o encharcamento do tecido vegetal (BRUNINGS; GABRIEL, 2003).

Além disso, a injeção da PthA modula a expressão de genes relacionados à síntese de auxina e giberelina em citros, além de aumentar a produção do fito-hormônio etileno (PEREIRA et al., 2014).

O cancro cítrico ocorre em uma grande diversidade de cultivares de citros e desencadeia lesões nas partes aéreas das plantas, sobretudo nas folhas e nos frutos, nos quais o tecido vegetal é induzido à hipertrofia e à hiperplasia (SWARUP et al., 1992), causando erupções de 2 a 10 mm, de aspecto corticoso e áspero (Figura 1.2) devido ao rompimento da epiderme vegetal (ROSSETTI, 2001), com margens elevadas, centros afundados e circulares (DEWDNEY; GRAHAM; ROGERS, 2014).



**Figura 1.2** Lesões características do cancro cítrico. Lesões típicas do cancro cítrico em folha, fruto e ramo infectados por Xcc, da esquerda para a direita. Fonte: Fundecitrus (<http://www.fundecitrus.com.br/ImageBank/PageFlip/pageflip.aspx?idPage=66>)

O desbalanço na produção de pigmentos em decorrência da infecção pode ocasionar o aparecimento de halos amarelos ao redor das erupções, o que está associado ao decréscimo de clorofilas a e b, caroteno e xantofila (PADMANABHAN et al., 1973). Além disso, lesões em estágios avançados podem se apresentar escurecidas em decorrência da colonização por fungos saprófitos, como *Colletotrichum* spp. (DEWDNEY; GRAHAM; ROGERS, 2014).

Ambas as superfícies foliares são acometidas pelas lesões do cancro cítrico, mas principalmente a abaxial, por conta da elevada concentração de estômatos, fazendo das folhas a região mais importante no processo de infecção (GOTTWALD; GRAHAM; SCHUBERT, 2002). Adicionalmente, a atividade do minador dos citros (*Phyllocnistis citrella*) pode facilitar a penetração do patógeno e intensificar a incidência do cancro cítrico (HALL; GOTTWALD; BOCK, 2010).

Os períodos de maior suscetibilidade ao cancro cítrico ocorrem após as brotações e, também, no início do desenvolvimento dos frutos, podendo ser agravados pela ocorrência de grandes tempestades e elevação abrupta de temperatura (SCHUBERT et al., 2001).

Entretanto, os sintomas apresentados pelas plantas contaminadas pelo cancro cítrico não constituem o principal problema da doença, uma vez que raramente tais manifestações são letais para a planta. Todavia, um dos efeitos adicionais do cancro

cítrico é o aumento na produção do fito-hormônio etileno em resposta à infecção, como mencionado anteriormente, o que provoca a abscisão prematura de folhas e frutos, tornando estes inapropriados para o comércio de frutas frescas e não comercializáveis para a indústria (CROZIER et al., 2000).

A chuva e o vento são relevantes agentes disseminadores do cancro cítrico. Ao serem molhadas pela chuva, as lesões cancrosas são exsudadas e os potenciais inóculos bacterianos têm a sua propagação facilitada pelo vento, o que as permite atingir distâncias consideravelmente longas, da ordem de alguns quilômetros em situações extremas (DAS, 2003). Frutos em estágio avançado de infecção também atuam como dispersores da doenças pois apresentam ruptura na casca, o que facilita a distribuição do patógeno (AMARAL, 2003).

A elevada taxa de sobrevivência de Xcc nas lesões nas frutas e em variadas superfícies faz do homem um outro importante agente disseminador, por meio do transporte de mudas, restos de cultura e do manuseio de equipamentos contaminados, principalmente quando as plantas estão molhadas (BOCK et al., 2006).

O manejo do cancro cítrico pode envolver medidas drásticas como a erradicação de plantas contaminadas, principalmente em áreas endêmicas, como ocorreu por um longo período de tempo no Brasil (BEHLAU; FONSECA; BELASQUE JUNIOR, 2016). Paralela e concomitantemente à erradicação, diversos citricultores desenvolveram estratégias de controle, restringindo a aquisição de mudas de citros oriundas de locais vigorosamente atingidos pelo cancro cítrico (LEITE JR; MOHAN, 1990).

A desinfecção completa de roupas, sapatos, máquinas e equipamentos utilizados em campo também tem sido uma prática de controle da disseminação do patógeno, uma vez que a capacidade de sobrevivência de Xcc em materiais de madeira, metálicos, plásticos e de couro atinge até 72 horas dependendo das condições climáticas (GOTTWALD; GRAHAM; SCHUBERT, 2002).

Dado que o vento é um dos principais agentes dispersores da doença, a implantação de quebra-ventos é uma das mais efetivas medidas de controle da dispersão de Xcc (LEITE JR; MOHAN, 1990), principalmente quando utilizada concomitantemente à pulverização de cobre nos pomares, embora este metal seja tóxico ao ambiente (BEHLAU et al., 2017) e já tenham sido identificadas linhagens resistentes (BEHLAU et al., 2011).

Outras medidas de controle do cancro cítrico incluem a aplicação de indutores de resistência (GRAHAM et al., 2000; GRAHAM; MYERS, 2009), controle do minador dos citros (HOY; NGUYEN, 1997) e, mais recentemente, o uso de produtos com atividade bactericida contra Xcc como metabólitos de *Pseudomonas* sp. (DE OLIVEIRA et al., 2011), dihidroxibenzoatos de alquila (NAZARÉ et al., 2018), galatos de alquila (SILVA et al., 2013) e o desenvolvimento de cultivares geneticamente modificadas (BARBOSA-MENDES et al., 2009; JIA et al., 2016, 2017; PENG et al., 2017; YANG et al., 2011).

Entretanto, as medidas de controle e manejo do cancro cítrico são mitigadoras de risco e paliativas, não sendo constatadas no campo medidas eficazes de controle e cura para a doença (SILVA et al., 2013).

### 1.3 O gênero *Xanthomonas*

O gênero *Xanthomonas* compreende um grupo economicamente importante de fitopatógenos com estreito espectro de hospedeiro para cada espécie, mas que na totalidade das espécies conhecidas acomete 240 gêneros de plantas distribuídas em 68 famílias diferentes, especialmente a das rutáceas (HAYWARD, 1993).

Este grupo é composto por bactérias aeróbicas restritas, Gram-negativas, portadoras de um único flagelo polar, não fastidiosas, com morfologia de bastonete e não redutoras de nitrato (GRAHAM et al., 2004; SCHAAD; JONES; CHUN, 2001).

As colônias são circulares, lisas, convexas e, por conta da produção de um aril-poliene-brominado denominado xantomonadina, apresentam coloração amarelada (SCHAAD; JONES; CHUN, 2001). Seu aspecto visual é mucoso (STALL; CIVEROLO, 1991), em decorrência da produção da goma xantana, um polissacarídeo plurivalente com aplicações nas indústrias alimentícia, farmacêutica, química e petroquímica (GOTTIG et al., 2010; LOPEZ; MORENO; RAMOS-CORMENZANA, 2001).

Dentre as espécies causadoras do cancro cítrico, destaca-se a *Xanthomonas citri* subsp. *citri* (Xcc), que infecta a maioria das espécies conhecidas de citros (DEWDNEY; GRAHAM; ROGERS, 2014). Encontrada no Uruguai, Paraguai e Argentina, *Xanthomonas fuscans* subsp. *aurantifolii* tipo B (XauB) é menos virulenta e provoca a cancrose B principalmente em *Citrus limon* (STALL; SEYMOUR, 1983). *Xanthomonas fuscans* subsp. *aurantifolii* tipo C (XauC), causadora da cancrose C, ataca um único hospedeiro, *Citrus aurantifolia* (CIVEROLO, 1984).

Apesar de haver variações a respeito da suscetibilidade de espécies de *Citrus* em relação ao cancro cítrico, a resistência à infecção artificial por Xcc não foi identificada, o que sugere que as flutuações de vulnerabilidade dos hospedeiros à bactéria decorram de componentes ambientais (BRUNINGS; GABRIEL, 2003).

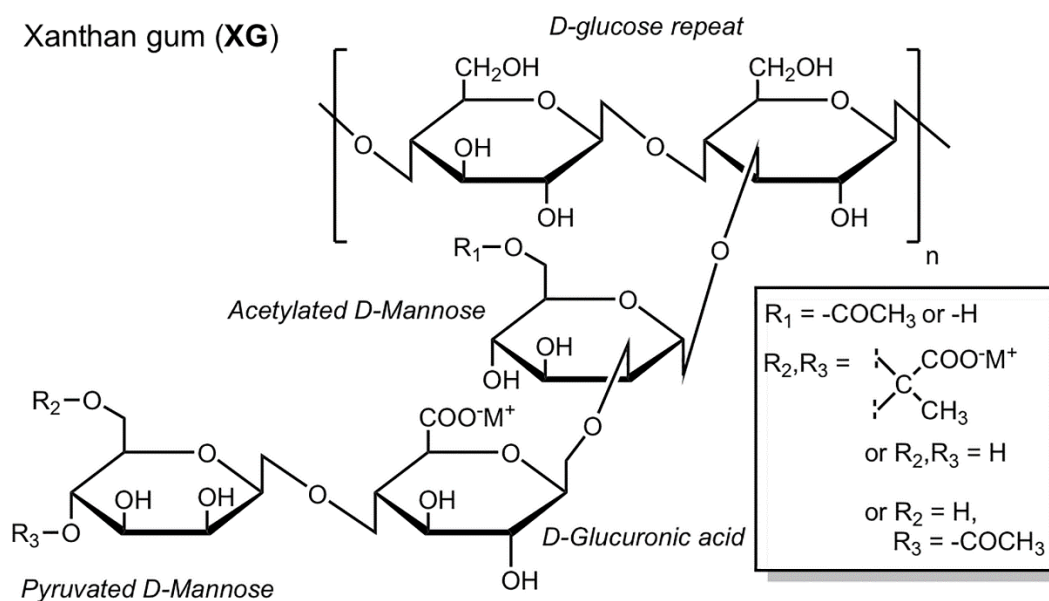
#### 1.4 Goma xantana

Dentre os fatores de virulência produzidos por *Xanthomonas* spp., a goma xantana merece especial atenção, uma vez que, além de ser um produto de interesse industrial (CARMONA et al., 2017), esse exopolissacarídeo (EPS) possui relação com a formação de biofilme e, também, com o processo infeccioso (BÜTTNER; BONAS, 2010).

A xantana é um heteropolissacarídeo constituído por um esqueleto central formado por dois resíduos de D-glicose unidos por ligações  $\beta$ -1,4, sendo que o segundo é substituído em seu oxigênio 3 (O-3) por um trissacarídeo composto por manose-( $\beta$ -1,4)-ácido glicurônico-( $\beta$ -1,2)-manose, com o primeiro resíduo de manose podendo ser acetilado e o último piruvatado (Figura 1.3).

O grau de acetilação e piruvatação dos resíduos de manose é variável entre as diversas espécies (e até mesmo entre as linhagens) de *Xanthomonas*, podendo sofrer influência adicional de fatores ambientais (CADMUS et al., 1976; KOOL et al., 2013; STANKOWSKI; MUELLER; ZELLER, 1993). Em *Xanthomonas campestris* pv. *campestris*, a piruvatação é essencial para a patogenicidade por suprimir a deposição do polissacarídeo calose, uma importante estratégia de defesa das plantas (BIANCO et al., 2016)

A via de biossíntese de xantana é bem compreendida em *Xanthomonas campestris* pv. *campestris*, sendo primariamente composta pelo *operon gumBCDEFGHIJKLM*, cujos genes codificam principalmente glicosiltransferases e proteínas transportadoras, as quais possuem atividade concentrada na membrana celular, local de síntese desse EPS (KATZEN et al., 1996).

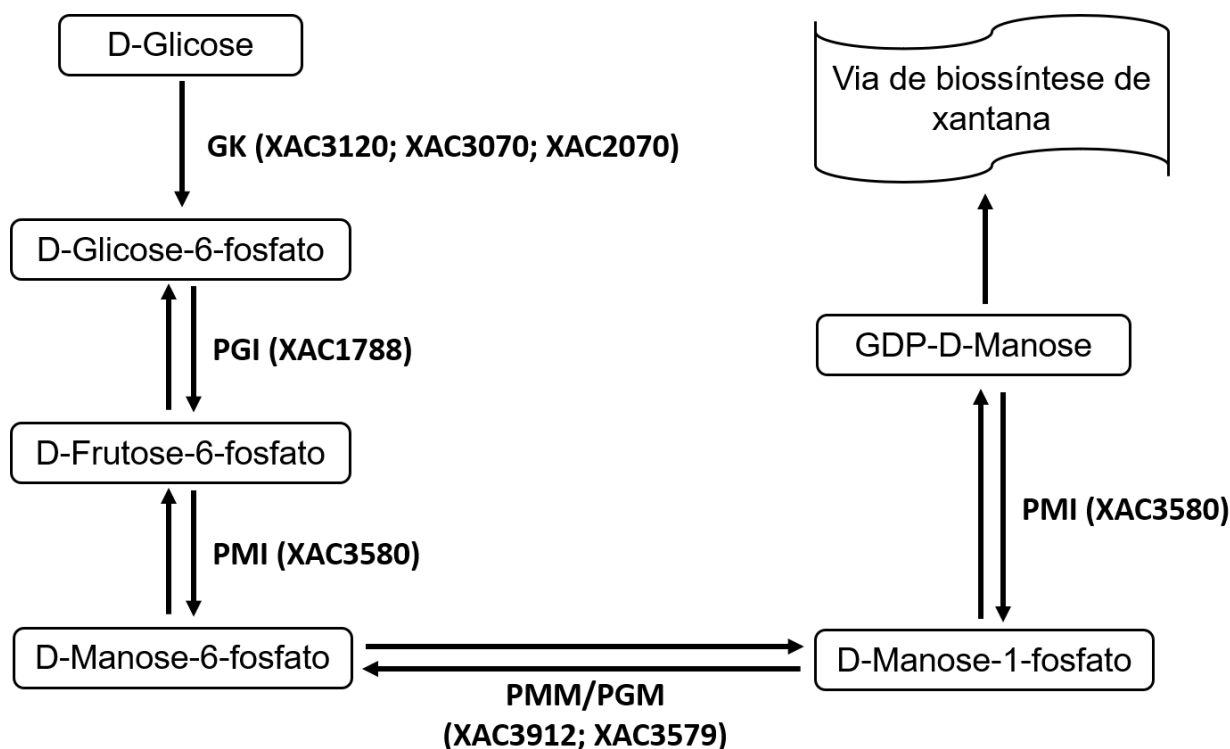


**Figura 1.3** Estrutura química da goma xantana. A figura apresenta a unidade pentassacarídica que compõe o exopolissacarídeo xantana. Fonte: HUTCHISON et al., 2019.

Os mais importantes precursores da via de biossíntese de xantana são UDP-glicose, UDP-ácido glicurônico, fosfoenolpiruvato e GDP-manose, sendo este último um composto produzido pela fosfomanose isomerase (PMI) a partir de manose-1-fosfato, que por sua vez tem como precursor a manose-6-fosfato, originada também pela ação da PMI sobre a frutose-6-fosfato (Figura 1.4) (KÖPLIN et al., 1992).

Embora não seja essencial para a patogenicidade de *Xcc* (DUNGER et al., 2007), a goma xantana está envolvida na sobrevivência epifítica por promover proteção contra a radiação ultravioleta, ao congelamento e a dessecação (MEYER; BOGDANOVE, 2009).

A relação entre a biossíntese de xantana e o processo infeccioso de *Xanthomonas* spp. tem sido estudada nas últimas décadas. Dunger e colaboradores (2007) produziram linhagens deficientes na produção de xantana por meio da deleção do gene *gumD* de *Xcc*. Apesar de não ter afetado a extensão dos sintomas provocados pela bactéria e a capacidade de induzir a resposta de hipersensibilidade, a defectividade na produção de xantana reduziu a resistência ao estresse oxidativo, além de comprometer a sobrevivência no ambiente epifítico.



**Figura 1.4** A relação entre a PMI e biossíntese de xantana. A via de biossíntese de xantana é alimentada por diversos precursores, dentre os quais se destaca a GDP-D-Manose, a qual é produzida a partir de D-manose-1-fosfato pela fosfomanose isomerase (PMI). GK: glicocquinase. PGI: fosfoglicose isomerase. PMM/PGM: fosfomanomutase/fosfoglicomutase. Fonte: próprio autor.

Em *Xanthomonas campestris*, a deleção do mesmo gene supracitado resultou em redução da virulência bacteriana por conta da alteração da via de biossíntese de xantana em experimentos utilizando repolho (KATZEN et al., 1998) e brócolis (CHOU et al., 1997) como hospedeiros.

A disrupção de um outro gene do *operon gum*, relacionado à biossíntese de xantana, *gumG*, que codifica uma acetiltransferase, levou à diminuição na produção de exopolissacarídeos e à redução severa da virulência de *Xanthomonas oryzae* pv. *oryzae* (DHARMAPURI; SONTI, 1999).

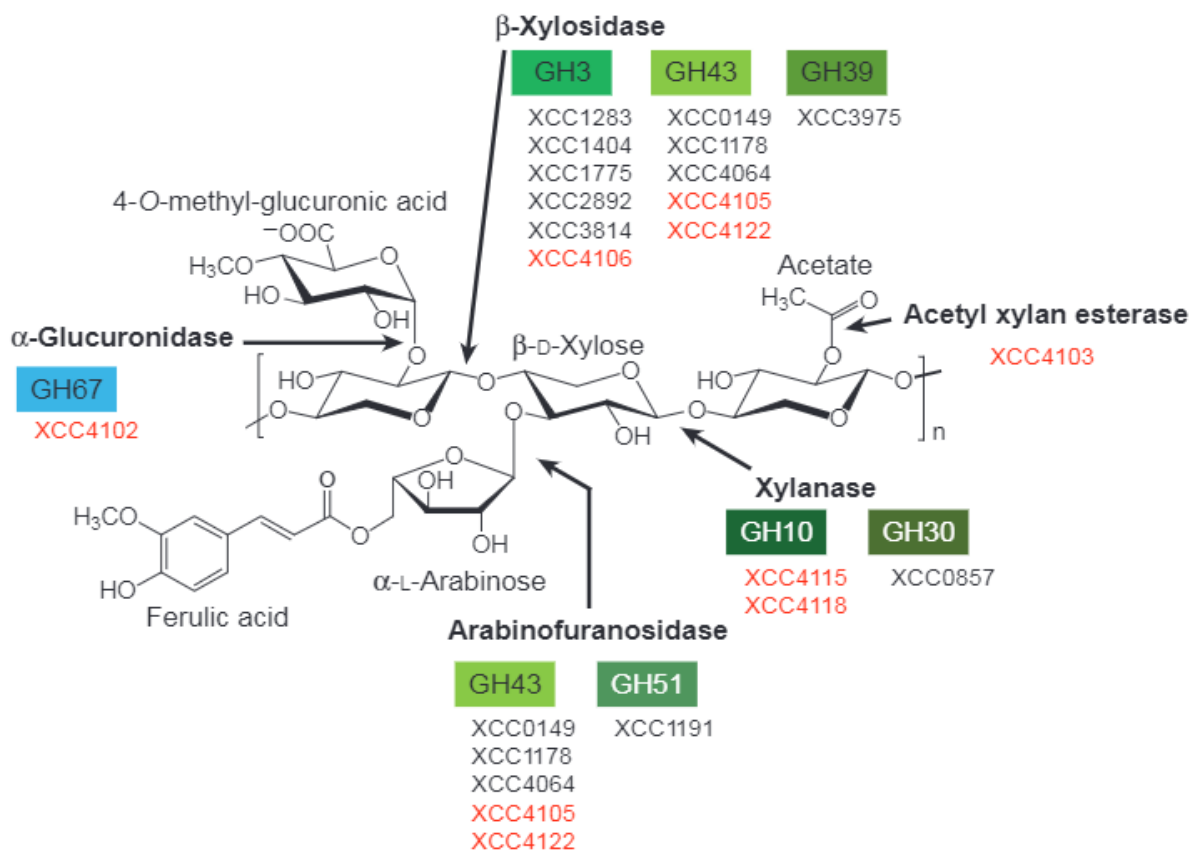
### 1.5 Xilano: estrutura e vias de degradação

Os carboidratos complexos (oligossacarídeos e polissacarídeos) são formados por uma gigantesca variedade de combinações de subunidades de resíduos de monossacarídeos unidos por ligações glicosídicas e possuem funções estruturais e de armazenamento de energia, constituindo o maior depósito de carbono passível de metabolização da biosfera (GRONDIN et al., 2017).



Unidades de D-xilose, uma aldopentose, estão unidas por ligações  $\beta$ -1,4 para formar o esqueleto central de um dos maiores componentes das hemiceluloses presentes na parede celular vegetal (Figura 1.5), o heteropolissacarídeo xilano (CHOW et al., 2015), que pode conter uma diversidade enorme de substituintes em suas ramificações (BUSSE-WICHER et al., 2016).

Sendo um heteropolissacarídeo, o xilano apresenta composição heterogênea, podendo apresentar ramificações com L-arabinofuranose ligada às posições C-3 dos resíduos de D-xilose, os quais são denominados arabinoxilanos (AX) (SCHENDEL et al., 2015). Os glicuronoxilanos (GX) possuem o ácido D-glicurônico ou 4-O-metil-glicurônico ligado às posições C-2 dos resíduos de D-xilose (TOGASHI; KATO; SHIMIZU, 2009), ao passo que os heteroxilanos que contêm uma mistura de L-arabinofuranose e ácido D-glicurônico são chamados de arabinoglicuronoxilanos (AGX) (EBRINGEROVÁ, 2005).



**Figura 1.5** Esquema da estrutura geral do xilano. Estrutura do heteropolissacarídeo xilano com a indicação das enzimas xilanolíticas, bem como de suas ORFs codificantes em *Xanthomonas campestris* pv. *campestris* (modificado a partir de DÉJEAN et al., 2013).

Sistemas enzimáticos xilanolíticos têm sido identificados em bactérias, fungos e leveduras (CHADHA et al., 2019) e são basicamente compostos por xilanases e outras glicosidases (GH, glycoside hydrolases), carboidrato esterases (CE) e mono-oxigenases líticas de polissacarídeos (LPMOs, *Lytic polysaccharide monoxygenases*), compondo vias metabólicas responsáveis pela degradação completa de xilano (MALGAS et al., 2019).

As carboidrato-esterases (CEs) são enzimas presentes em alguns microrganismos e têm como função sobrepujar uma estratégia de proteção que diversas plantas desenvolveram, baseada na esterificação parcial de polissacarídeos com ácido acético. Por exemplo, resíduos de glicosil acetilados não podem ser hidrolisados pelas glicosidases, mas após serem desacetilados pelas esterases específicas, tornam-se acessíveis às GH (BIELY, 2012).

A degradação das cadeias laterais de xilano inclui as  $\alpha$ -arabinofuranosidases (EC 3.2.1.55), que hidrolisam as ligações glicosídicas entre os resíduos de  $\alpha$ -L-arabinofuranosídeo em AXs e xilo-oligossacarídeos (XOS) (LAGAERT et al., 2014). Da mesma forma, as  $\alpha$ -glicuronidases (EC 3.2.1.139) atuam sobre os GX, liberando os substituintes de ácido D-glicurônico (SHALLOM; SHOHAM, 2003).

Já a degradação do esqueleto central xilano tem a participação principal das  $\beta$ -xilosidases (E.C.3.2.1.37) e endo-1,4- $\beta$ -xilanases (EC 3.2.1.8), enzimas que atuam na hidrólise das ligações glicosídicas  $\beta$ -1,4 entre os resíduos de xilopiranosil, liberando xilose, xilobiose (X2), xilotriose (X3), xilotetraose (X4) e até xilo-oligossacarídeos (XOs) mais longos ou também ramificados (BURLACU; CORNEA; ISRAEL-ROMING, 2016).

Importante salientar também a recente descoberta de estratégias auxiliares de hidrólise de polissacarídeos, baseadas em um mecanismo oxidativo ainda pouco compreendido, que atuam sinergicamente com as enzimas dos sistemas xilanolíticos na desconstrução dessas macromoléculas e na disponibilização de biomassa vegetal (MOREAU et al., 2019), como já descrito em *Xanthomonas* (SANTOS et al., 2014).

Diversos fitopatógenos, incluindo Xcc, têm desenvolvido eficientes vias metabólicas de assimilação da D-xilose liberada pelas enzimas xilanolíticas como fonte de energia, envolvendo o transporte dessa pentose através da membrana citoplasmática, seguida por sua isomerização a D-xilulose, catalisada pela xilose isomerase, e fosforilação a D-xilulose-5-fosfato, um componente que é finalmente metabolizado pela via das pentoses-fosfato (CHOW et al., 2015). Dessa forma, a

xilose isomerase, um dos alvos deste trabalho, está envolvida na utilização de fontes de carbono oriundas do xilano.

## 1.6 Estudos genômicos e pós-genômicos

A importância econômica da citricultura associada à falta de controle e cura eficazes do cancro cítrico motivou estudos em bactérias do gênero *Xanthomonas* utilizando diversas estratégias, dentre as quais destacam-se a genômica comparativa (FONSECA et al., 2019a, 2019b), proteômica diferencial (ARTIER et al., 2018; CARNIELLI et al., 2017; MOREIRA et al., 2017), análise de expressão gênica (ASTUA-MONGE et al., 2005), análise funcional por deleção gênica (ALEXANDRINO; GOTO; NOVO-MANSUR, 2016; UCCI et al., 2014), bem como a elucidação da estrutura e função de proteínas relacionadas ao processo infeccioso (CABREJOS et al., 2019; GOTO et al., 2016; TEIXEIRA et al., 2018).

Tais estudos foram impulsionados pela conclusão dos projetos de sequenciamento dos genomas de bactérias do gênero *Xanthomonas*, principalmente Xcc (DA SILVA et al., 2002), ampliaram o conhecimento sobre a patogênese do cancro cítrico e estão imersos em um contexto de busca por estratégias de controle para a doença.

A proteômica, considerada uma análise em larga escala da expressão proteica, é uma das áreas centrais no estudo do metabolismo, sendo amplamente utilizada em nosso grupo de pesquisa para a identificação de proteínas envolvidas no processo infeccioso de Xcc (ARTIER et al., 2018; CARNIELLI et al., 2016, 2017).

No trabalho de Carnielli e colaboradores (2017), proteínas superficiais de células intactas de Xcc foram marcadas com fluoróforos, após a exsudação das bactérias de folhas de citros infectadas (ensaio *in vivo* em *Citrus aurantifolia*) ou após cultivo em condição de não infecção (cultivo *in vitro* em meio rico em nutrientes). Após lise celular, as proteínas totais foram separadas por eletroforese bidimensional (2D-DIGE) e os *spots* diferenciais entre as duas condições foram excisados do gel, digeridos com tripsina e identificados por espectrometria de massas.

Uma das proteínas encontradas como mais abundantes na condição de indução de patogenicidade foi a fosfomanose isomerase (PMI). Embora a PMI não apresente peptídeo sinal conhecido, relacionado ao transporte à membrana, a sua identificação na superfície de Xcc desperta interesse, uma vez que tal exposição

facilita a utilização de substâncias inibidoras, caso a proteína se revele um bom alvo. Assim, a localização celular, juntamente à identificação inédita em Xcc em uma situação de indução de patogenicidade *in vivo* foram motivações para a escolha do gene codificante da PMI como alvo para este trabalho.

Em um outro trabalho desenvolvido em nosso grupo de pesquisa, Zandonadi (2012) comparou o perfil de extratos proteicos enriquecidos em periplasma de Xcc e XauB (*Xanthomonas fuscans* subsp. *aurantifolii*, tipo B), esta última causadora de uma forma mais branda da doença, a cançrose. As bactérias foram cultivadas nos meios XAM-M (indutor de patogenicidade) ou em caldo nutriente (CN), um meio rico em nutrientes e não indutor da expressão de fatores de virulência do sistema de secreção tipo III (CARVALHO, 2006). Dezenas de proteínas foram identificadas como diferencialmente expressas entre Xcc e XauB nas diferentes condições, e a xilose isomerase foi encontrada exclusivamente em Xcc quando cultivada em XAM-M. Por conta desse indício de participação da xilose isomerase no processo infeccioso de Xcc, o seu gene codificante, *xyIA2*, foi também escolhido como alvo para esta tese de doutorado.

## **1.7 Genes-alvo e seus produtos**

### **1.7.1 *xanB***

Fosfomanose isomerases (EC 5.3.1.8) são enzimas envolvidas no metabolismo de manose por meio da interconversão de D-manose-6-fosfato e D-frutose-6-fosfato, mas em alguns casos, são capazes de converter D-manose-1-fosfato em GDP-D-manose (Figura 1.4), um intermediário imprescindível para a síntese de EPS, como a xantana (JENSEN; REEVES, 1998; MILES; GUEST, 1984; PAPOUTSOPOULOU; KYRIAKIDIS, 1997; SMITH et al., 1992).

Foram identificadas três classes de PMI em bactérias e fungos, as quais possuem sequências peptídicas significativamente diferentes. As PMI tipo I possuem uma única atividade catalítica, a de interconversão de D-manose-6-fosfato e D-frutose-6-fosfato. Além da atividade apresentada pelas PMI tipo I, a conversão de D-manose-1-fosfato em GDP-D-manose é uma atividade catalítica adicional apresentada pelas PMI tipo II, também chamadas de PMI-GMP (fosfomanose isomerase-GDP-D-manose pirofosforilase). O terceiro tipo de PMI possui a mesma

atividade catalítica da PMI tipo I e ocorre somente em *Rhizobium meliloti*, sendo incluída em uma categoria única por conta de sua sequência peptídica notadamente distinta das PMI do tipo I (JENSEN; REEVES, 1998; SCHMIDT et al., 1992).

A PMI de *Xanthomonas campestris* foi categorizada na classe II por sua atividade catalítica bifuncional, sendo também identificada como uma enzima zinco-dependente, uma vez que este íon metálico atua como um cofator enzimático (PAPOUTSOPOULOU; KYRIAKIDIS, 1997).

Dado que a sequência peptídica predita a partir do gene *xanB* de Xcc compartilha uma identidade de 88,22% e semelhança de 99,78% com a sequência de *Xanthomonas campestris* (Figura 1.6), é possível inferir que a enzima de Xcc também se enquadra no tipo II.

```

citri          MSDVLP I ILSGGSGTRLWPLSRESYPKQFLPLVGDKSM LQSTWLR AAPVAGHAPIVVANE 60
campestris   MSDVLP I ILSGGSGTRLWPLSRETYPKQFLPLVGEHSM LQATWLR SAPVA AHAPIVVANE 60
                *****:*****:*****:*****:*****:*****

citri          EHRFMAAEQLQQLGVKPSAILLEPKGRNTAPAI AVAALEATR DGADP LLLVLPSDHVIQN 120
campestris   EHRFMAAEQLQQLGVKPSAILLEPKGRNTAPAI AVAALEATR NGDP LLLVLPSDHVIRD 120
                *****:*.*****:*****:*****:*****:*****

citri          EAAFQAAVT LAATAAEQ GKLVTFGIKPTAPETGYGYIKASAGAGAS AVERFVEKPD LATA 180
campestris   EAAFQAAVT VAAAAEQ GKLVTFGIKPTAPETGYGYIKAGAGTGA TAVERFVEKPD LATA 180
                *****:*.*****:*****:*****:*****:*****

citri          QSYLASGEYYWNSGMFLFRASRYLEELRKFHPAI ADACQKAWENGKRDADFTRLDKDAFA 240
campestris   QGYLASGEYYWNSGMFLFRASRYLEELRKFQPAI ADACQKAWEGGKRDADFTRLDKDAFA 240
                * .*****:*****:*****:*****:*****:*****

citri          ASPSDSIDYAVMEKTADAVVPLDAGWNDVGSWSSLLDVSNQDAQGN AHG DVIQLDCQN 300
campestris   SSPSDSIDYAVMEKTADAVVPLDAGWNDVGSWSSLLDVSEQDQG GN AHG DVIQLDCKN 300
                :*****:*****:*****:*****:*****:*****

citri          TYAYGSRLIAMVGL EDVVVET DAVLVGHRDRIQEVKDVVSQIKTAGRSEATWHRKVYR 360
campestris   TYAYGSRLIAMVGL ENVVVET DAVLVGHRDRIQEVKEVVSQIKSAGRSEATWHRKVYR 360
                *****:***** *****:*****:*****:*****:*****

citri          PWGAYDSIDMGQRHQVKRITVKPGAVLSLQMHHR AEHWIVVSGTAEVTRGEEVLLLTEN 420
campestris   PWGAYDSIDMGQR FQVKRITVKPGATLSLQMHHR AEHWIVVSGTAEVTRGDEVLLLTEN 420
                ***** .***** .*****:*****:*****:*****

citri          QSTYIPLGVTHRLRNP GKLPLELIEVQSGSYLGEDDIVRFEDTYGRA 467
campestris   QSTYIPLGVTHRLKNP GKLPLELIEVQSGSYLGEDDIVRFEDTYGRT 467
                *****:*****:*****:*****:*****:*****

```

**Figura 1.6** Alinhamento entre sequências peptídicas de PMI. As sequências peptídicas preditas a partir da PMI de Xcc e *Xanthomonas campestris* pv. *campestris* foram alinhadas utilizando o algoritmo CLUSTAL O(1.2.4), resultando em uma identidade de 88,22% e semelhança de 99,78%. O único par de resíduos de aminoácidos sem similaridades químicas está destacado em vermelho. Xcc e *Xanthomonas campestris* pv. *campestris* estão mencionadas como *citri* e *campestris*, respectivamente. Resíduos idênticos e com características principais conservadas são mostrados por (\*) e (:), respectivamente. Resíduos com características secundárias conservadas são mostrados por (.). Fonte: próprio autor.

A relação entre o gene *xanB* e a patogenicidade de *Xcc* não havia sido estabelecida, mas há trabalhos que associam a PMI de *Xanthomonas campestris* ou outras espécies a processos como a síntese de xantana (KÖPLIN et al., 1992; PAPOUTSOPOULOU; KYRIAKIDIS, 1997), adesão celular e reconhecimento (WIEDERSCHAIN, 2013), formação de biofilme (AMOS et al., 2011; SOUSA et al., 2007) e produção de lipopolissacarídeos (VORHÖLTER; NIEHAUS; PÜHLER, 2001), processos estes relacionados com a patogenicidade.

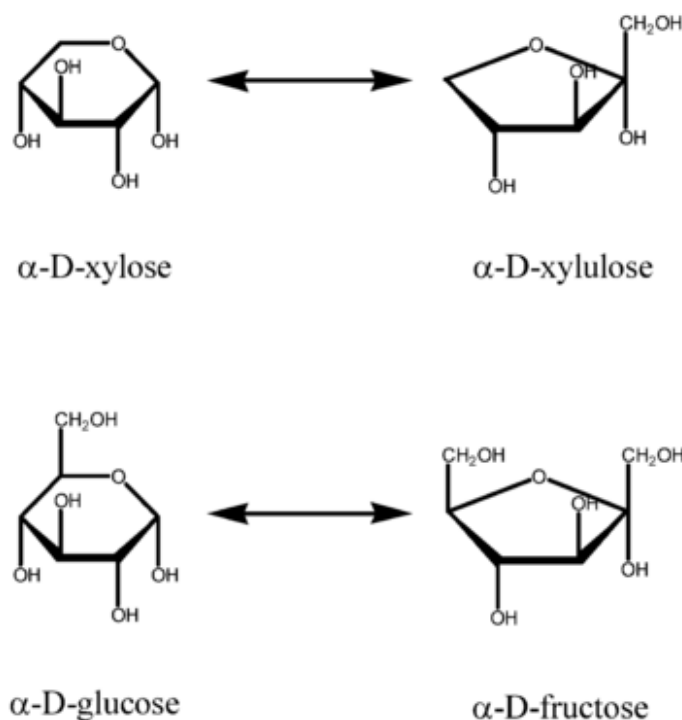
No PDB (Protein Data Bank), existem vinte estruturas cristalográficas de fosfomanose isomerases pertencentes a somente quatro espécies, as quais possuem os seguintes códigos de acesso: *Candida albicans* (1PMI e 5NW7), *Salmonella typhimurium* (3H1M, 3H1W, 3H1Y, 5ZT4, 5ZT5, 5ZT6, 5ZUW, 5ZUY, 5ZV0, 5ZVR, 5ZVU, 5ZVX e 2WFP), *Pyrobaculum aerophilum* (1X9H, 1X9I, 1TZB e 1TZC) e *Bacillus subtilis* (1QWR).

### 1.7.2 *xyIA*

A xilose isomerase (EC. 5.3.1.5), também conhecida como glicose isomerase (Figura 1.7), é uma enzima responsável por catalisar as reações de isomerização reversíveis de D-xilose em D-xilulose e D-glicose em D-frutose (STEVIS; HO, 1985), porém, sua afinidade pela xilose é 160 vezes maior do que pela glicose em *Aeromonas hydrophila* (BHOSALE; RAO; DESHPANDE, 1996).

No contexto do processo infeccioso de *Xanthomonas* spp., a fonte de carboidratos foi identificada como um componente significativo na expressão de fatores de virulência. Por exemplo, o *cluster hrp* (resposta de hipersensibilidade e patogenicidade) em *Xanthomonas oryzae* possui alguns membros com regulação positiva somente na presença de xilose (IKAWA; TSUGE, 2016).

Em *Xanthomonas* spp., os componentes desse *cluster hrp* codificam para o sistema de secreção tipo III, cuja função é injetar proteínas efetoras nas células do hospedeiro, de modo que o fitopatógeno aproveite a degradação da parede celular lignocelulósica de maneira ideal, utilizando os seus resíduos tanto como moléculas de sinalização quanto como fonte energética (BÜTTNER; BONAS, 2010).



**Figura 1.7** Esquema das reações catalisadas pela xilose isomerase. A enzima catalisa as isomerizações reversíveis entre D-xilose e D-xilulose, bem como de D-glicose e D-frutose (modificado a partir de BHOSALE; RAO; DESHPANDE, 1996).

Diferentemente das leveduras industriais, bactérias do gênero *Xanthomonas* têm a capacidade de metabolizar xilose, um dos principais componentes da massa lignocelulósica, que é negligenciada na produção de biocombustíveis (SIEVERT et al., 2017). É pela capacidade de aproveitar essa biomassa abundante e barata que o gênero *Xanthomonas* tem sido objeto de estudos voltados à produção de biocombustíveis e goma xantana (ZHANG; CHEN, 2010).

O metabolismo da xilose envolve uma série de genes frequentemente organizados em *operons* que respondem a essa aldopentose. Em *Bacillus* spp. e *Lactobacillus pentosus*, o *operon xylAB* é composto pelos genes *xylA* (codificante da xilose isomerase) e *xylB* (codificante da xiluloquinase) e é inibido pelo regulador da família LacI, chamado XylR (DAHL; SCHMIEDEL; HILLEN, 1995; LOKMAN et al., 1997).

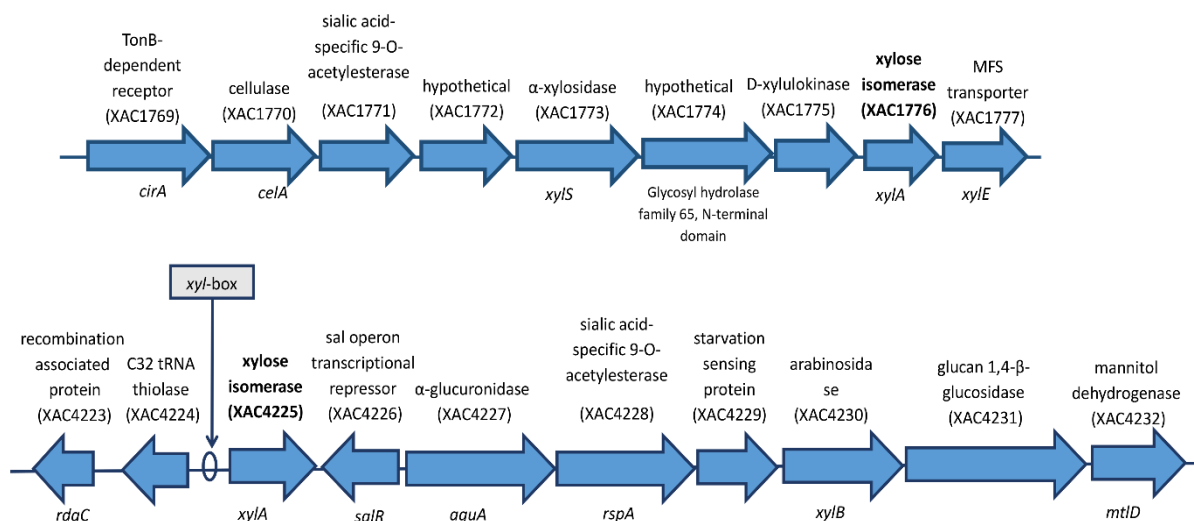
Em *Caulobacter crescentus*, XylR se liga ao operador *xyl-box* na ausência de xilose, inibindo a expressão dos genes regulados por *xyl-box*, mas se dissocia quando ligado à xilose, (STEPHENS et al., 2007). Em *Xanthomonas oryzae*, XylR também reprime a expressão de vários genes envolvidos no metabolismo da xilose que são regulados por *xyl-box*, além de estar envolvido na regulação negativa dos genes *hrp* (IKAWA et al., 2018).

A deleção do gene *xylR* em *Streptomyces lividans* aumentou quase cinco vezes a expressão do gene codificante da xilose isomerase, uma das principais enzimas do metabolismo da xilose (HEO et al., 2008).

O genoma de Xcc possui dois genes anotados como codificantes de xilose isomerase (Figura 1.8) e, para facilitar a identificação, esses genes foram aqui denominados *xylA1* (ORF XAC1776) e *xylA2* (ORF XAC4225), sendo este último o alvo do presente trabalho.

Embora as sequências peptídicas preditas a partir dos dois genes compartilhem 99,32% de identidade, os contextos genômicos são bastante distintos, uma vez que apenas *xylA2* está inserido em um operon regulado por *xyl-box*, isto é, sob provável ação de XylR ou outra proteína análoga.

Apesar do metabolismo de xilose ter sido relacionado à patogenicidade de *Xanthomonas* spp., não há estudos sobre o papel desempenhado pela xilose isomerase no processo infeccioso de Xcc, uma das enzimas centrais de via metabólica dessa aldopentose.



**Figura 1.8** Contextos genômicos dos genes *xylA* em Xcc. Os dois genes codificantes de xilose isomerase estão destacados em negrito. Apesar de compartilharem identidade de aproximadamente 99% em relação às suas sequências nucleotídicas, apenas *xylA2* (ORF XAC4225) está sob provável controle de *xyl-box* (ZANDONADI et al., dados não publicados).

No PDB (*Protein Data Bank*), existem inúmeras estruturas cristalográficas de xilose isomerases, com destaque para as espécies *Thermoanaerobacterium thermosulfurigenes* (código de acesso 1A0C), *Bacillus stearothermophilus* (1A0D), *Thermotoga neapolitana* (1A0E), *Thermus thermophilus* (1BXB), *Thermus caldophilus*



(1BXC), *Streptomyces diastaticus* (1CLK), *Streptomyces rubiginosus* (4ZB2, 4DUO), *Streptomyces olivochromogenes* (1MUW, 1S5N, 1S5M, 2GYI), *Actinoplanes missouriensis* (1BHW), *Pseudomonas aeruginosa* (2H5A) e *Arthrobacter sp* (4XIA, 5XIA).

## 1.8 Objetivos

### 1.8.1 Objetivo geral

O objetivo foi a análise funcional dos genes *xanB* e *xyIA2* de *Xcc*.

### 1.8.2 Objetivos específicos

- Expressão heteróloga das ORFs XAC3580 e XAC4225 e purificação por cromatografia de afinidade dos produtos recombinantes;
- Caracterização da atividade biológica dos produtos de expressão heteróloga dos genes *xanB* e *xyIA2* de *Xcc*.
- Obtenção de linhagens de *Xcc* mutantes deletadas nos genes *xanB* e *xyIA2* (*Xcc*Δ*xanB* e *Xcc*Δ*xyIA*, respectivamente);
- Obtenção de linhagem complementante para o gene *xanB* (*Xcc*Δ*CxanB*);
- Estudos *in vivo* dos efeitos da deleção dos genes *xanB* e *xyIA2* em relação à patogenicidade e virulência de *Xcc* em *Citrus aurantifolia*;
- Avaliação dos efeitos da deleção do gene *xanB* na motilidade, formação de biofilme e resistência à radiação ultravioleta.

## Capítulo 2 – Material e Métodos

### 2.1 Linhagens bacterianas, meios e condições de cultivo

A linhagem *Xanthomonas citri* subsp. *citri* 306 (Xcc) foi gentilmente fornecida pelo Fundo de Defesa da Citricultura (Fundecitrus, Araraquara, SP, Brasil) e armazenada a -80 °C em meio Luria-Bertani (LB, Sigma) com 10% de glicerol.

Os cultivos de Xcc, *E. coli* DH5 $\alpha$  e *E. coli* BL21(DE3) em meios sólido e líquido foram realizados em ágar-LB (Sigma) e caldo LB (Sigma), respectivamente. Para Xcc, os cultivos foram incubados a 30 °C e 200 rpm, e para as linhagens de *E. coli*, a temperatura foi de 37°C e 250 rpm.

Para os ensaios de caracterização da linhagem Xcc $\Delta$ xanB (mutante deletada no gene que codifica a fosfomanose isomerase), os cultivos foram feitos em meio indutor de patogenicidade XAM-M, que consiste no meio XAM-1 (CARVALHO, 2006), com modificações. A composição do meio XAM-M (ARTIER et al., 2018) foi a seguinte: (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 7,57 mM; KH<sub>2</sub>PO<sub>4</sub> 33,06 mM; K<sub>2</sub>HPO<sub>4</sub> 60,28 mM; citrato de sódio (C<sub>6</sub>H<sub>5</sub>Na<sub>3</sub>O<sub>7</sub>.2H<sub>2</sub>O) 1,7 mM; MgSO<sub>4</sub> 1M; casaminoácidos 0,03%; Frutose 10 mM, Sacarose 10 mM; BSA 1 mg.ml<sup>-1</sup>; pH 5,4.

### 2.2 Enzimas, kits e procedimentos em biologia molecular

Para a extração de DNA genômico de Xcc foi utilizado o *kit PureLink™ Microbiome DNA Purification Kit* (Invitrogen). DNA polimerase (*Phusion High Fidelity DNA Polimerase*), enzimas de restrição, vetor de propagação e demais *kits* de extração de DNA foram adquiridos da Fermentas-Thermo Scientific e da Cellco Biotec. Os antibióticos ampicilina, canamicina e gentamicina foram obtidos da Sigma.

Quando não detalhados, os procedimentos de biologia molecular foram conduzidos exatamente como descritos anteriormente (AUSUBEL, 2002).

### 2.3. Desenho e síntese de oligonucleotídeos

Os oligonucleotídeos utilizados na amplificação por PCR (*Polymerase Chain Reaction*, Reação em Cadeia da DNA Polimerase) das regiões codificantes da fosfomanose isomerase e da xilose isomerase de Xcc foram projetados com base nas

sequências depositadas no GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>), cujos códigos de acesso são XAC3580/AAM35493.1 (*GDP-mannose pyrophosphorylase*) e XAC4225/AAM39060.1 (*xylose isomerase*), respectivamente, de tal forma a coincidirem 15-20 bases da região 5' da fita codificante e da sua respectiva fita reversa complementar (3' do fragmento gênico).

As análises *in silico* das ORFs XAC3580 e XAC4225 utilizando o *software* SignalP 4.0 (PETERSEN et al., 2011) (<http://www.cbs.dtu.dk/services/SignalP/>) indicaram ausência de peptídeos sinais nas sequências peptídicas preditas.

Para a construção dos plasmídeos de deleção dos genes-alvo, foram projetados oligonucleotídeos designados à amplificação por PCR das regiões de 1 kb *upstream* e *downstream* às ORFs XAC3580 e XAC4225, com base na sequência do genoma completo de *Xcc*, depositada no GenBank sob o código de acesso NC\_003919.1.

Para a confirmação da deleção e complementação gênicas, foram desenhados pares de oligonucleotídeos que hibridizam em segmentos adjacentes às regiões flangeadoras *upstream* e *downstream* de 1 kb. Como tais oligonucleotídeos não hibridizam nos vetores de deleção e complementação, tal estratégia foi adotada criteriosamente no sentido de certificar que os produtos de PCR obtidos a partir das linhagens mutantes e complementante são oriundos de hibridização dos oligonucleotídeos no DNA cromossômico das linhagens.

Para a construção de um primeiro plasmídeo de complementação do gene que codifica a fosfomanose isomerase (pLAL6\_CxanB), foram desenhados oligonucleotídeos destinados à amplificação por PCR da região predita como promotora do *operon xanAB*, bem como da região codificante do gene *xanB*, com base na sequência do genoma completo de *Xcc*.

Um segundo vetor de complementação do gene *xanB* (pNPTS138\_CxanB) foi construído a partir do vetor de deleção pNPTS138\_xanB, utilizando um par de oligonucleotídeos destinados à amplificação de um fragmento gênico correspondente à região codificante do gene *xanB*.

Todos os oligonucleotídeos possuem sítios de restrição estrategicamente posicionados para a posterior clonagem dos produtos de PCR nos vetores de propagação, expressão, deleção e/ou complementação.

Foi utilizado o *software* Lasergene 7.1 ([www.dnastar.com](http://www.dnastar.com)) para a análise das sequências nucleotídicas, bem como para a escolha e posicionamento dos sítios de

restrição das endonucleases adotadas, de modo que tais sítios não estivessem presentes nas sequências nucleotídicas a serem clonadas.

Foram inseridas bases adicionais às extremidades 5' de alguns oligonucleotídeos para possibilitarem, caso necessário, clonagem direta no vetor de expressão e, também, otimização dos parâmetros termodinâmicos como temperatura de transição de fase (*melting temperature*) e temperatura de formação de superestruturas indesejáveis como *hairpins*, homo e hetero dímeros de hibridização oligo-oligo. Tais otimizações foram feitas com o *software* OligoAnalyzer 3.1 (<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer>). Os oligonucleotídeos foram sintetizados pela Exxtend (Paulínia-SP, Brasil).

A Tabela 2.1 apresenta a sequência de todos os oligonucleotídeos (e respectivos vetores utilizados) e a Tabela 2.2 mostra a relação entre os pares de oligonucleotídeos e os correspondentes produtos de PCR, especificando suas respectivas funções no presente trabalho.

**Tabela 2.1** Desenho dos oligonucleotídeos. Ao final do processo de otimização, os oligonucleotídeos resultantes estão apresentados, sendo indicados os vetores pET28 (ou 29), pLAL6 e pNPTS138 (negrito) nos quais os fragmentos amplificados com os respectivos oligonucleotídeos foram clonados. Bases em vermelho indicam regiões de sobreposição com o fragmento gênico de interesse. Em letras minúsculas estão as bases adicionadas para otimização nas regiões 5'. Sítios de restrição estão grifados em amarelo, com as respectivas enzimas de restrição identificadas acima e pontos de clivagem representados por "/". As bases que são transcritas nos códons de fim de tradução (*stop codon*) estão grifadas em azul. Os resíduos de aminoácidos das traduções conceituais previstas da região *N*-terminal de cada proteína de interesse estão mostrados em verde. As temperaturas de *melting* (TM) dos oligonucleotídeos estão indicadas. FW e RV se referem aos oligonucleotídeos direto e reverso para o fragmento indicado. KO indica os oligonucleotídeos utilizados na confirmação da deleção e complementação gênicas. pPMI denota a região promotora predita (aqui denominada *xanP*) do gene que codifica a PMI. cPMI indica a região codificante do gene *xanB*, clonada em pLAL6 ou pNPTS138 para fins de complementação gênica do mutante de deleção. 1kb\_Up e 1kb\_Down consistem nos fragmentos flanqueadores de 1 kb das respectivas ORFs a serem deletadas.

#### **pET2829\_PMI**

pET2829PMI\_FW 5' gtaa <sup>*NdeI*</sup> CA/TATGAGCGACGTCCTACCC 3' TM 65,1 °C  
M S D V L P

pET2829PMI\_RV 5' tatata <sup>*XhoI*</sup> C/TCGAGTTAGGCGCGGCCGTA 3' TM 68,8 °C

pET2829PMI\_RV\_C 5' tatata <sup>*XhoI*</sup> C/TCGAGGGCGCGGCCGTAG 3' TM 69,4 °C

#### **1kb\_Up\_PMI**

Up\_PMI\_FW 5' tatata <sup>*HindIII*</sup> A/AGCTTATACCGCAGCGCCAGGC 3' TM 68,8 °C

Up\_PMI\_RV 5' cgtata <sup>*EcoRI*</sup> G/AATTCGGGCGGGGGAAACTC 3' TM 68,7 °C

**Tabela 2.1 (Continuação)** Desenho dos oligonucleotídeos.

<b>1kb_Down_PMI</b>		
Down_PMI_FW	5' ttaaca <i>EcoRI</i> G/AATTC TCGCGCGTCTCAACGTAG 3'	TM 68,7 °C
Down_PMI_RV	5' ataatt <i>NheI</i> G/CTAGC TTGTACGACAACCTGTACCC 3'	TM 68,5 °C
<b>KO_PMI</b>		
KO_PMI_FW	5' GATCAGCTCGGACACCGG 3'	TM 63,2 °C
KO_PMI_RV	5' GGACCGACAACATCATCTTCTCG 3'	TM 63,8 °C
<b>pLAL6_pPMI</b>		
pLAL6_pPMI_FW	5' tatata <i>EcoRI</i> G/AATTC ACGAAATCATCTTGCGGATCTG 3'	TM 66,1 °C
pLAL6_pPMI_RV	5' tatata <i>XhoI</i> C/TCGAG GGGCAAATCCTGAAGT 3'	TM 66,0 °C
<b>pLAL6_cPMI</b>		
pLAL6_cPMI_FW	5' tattaa <i>XhoI</i> C/TCGAG ATGAGCGACGTCCTAC 3' M S D V L	TM 66,0 °C
pLAL6_cPMI_RV	5' tata <i>XbaI</i> T/CTAGATTAGGCGCGGCCGTAG 3'	TM 65,9 °C
<b>pNPTS_cPMI</b>		
pNPTS_cPMI_FW	5' tatata <i>EcoRI</i> G/AATTC ATGAGCGACGTCCTACCCAT 3' M S D V L P I	TM 67,3 °C
pNPTS_cPMI_RV	5' tatata <i>EcoRI</i> G/AATTC TTAGGCGCGGCCGTAGG 3'	TM 67,8 °C
<b>pET2829_XI</b>		
pET2829_XI_FW	5' tatata <i>NdeI</i> CA/TATG AGCAACACCGTGTACAT 3' M S N T V Y I	TM 65,4 °C
pET2829XI_RV	5' <i>XhoI</i> C/TCGAG TCAACGCGTCAGATACTG 3'	TM 65,3 °C
pET2829XI_RV_C	5' tatat <i>XhoI</i> C/TCGAG ACGCGTCAGATACTGAT 3'	TM 65,3 °C
<b>1kb_Up_XI</b>		
Up_XI_FW	5' tata <i>HindIII</i> A/AGCTTGGTCACGCCATGCGTC 3'	TM 68,0 °C
Up_XI_RV	5' ttaata <i>BamHI</i> G/GATCCGGGGTGAAGCTCCTG 3'	TM 68,0 °C
<b>1kb_Down_XI</b>		
Down_XI_FW	5' tatatata <i>BamHI</i> G/GATCCGCCTTGCCACTGCAC 3'	TM 68,0 °C
Down_XI_RV	5' tatata <i>SalI</i> G/TCGACGGTGGCATCGCGTAC 3'	TM 67,9 °C

**Tabela 2.1 (Continuação)** Desenho dos oligonucleotídeos.**KO\_XI**

KO_XI_FW	5' GATGGTGGCCGAGCGCGAT	TM 68,5 °C
KO_XI_RV	5' GCTGCTGGGCGTGTTCG	TM 68,1 °C

**Tabela 2.2** Pares de oligonucleotídeos e seus respectivos produtos de PCR (pPCR), funções e vetores nos quais foram clonados.

Par de oligonucleotídeos	Produto de PCR (pPCR)	Função
pET2829PMI_FW e pET2829PMI_RV	Região codificante do gene <i>xanB</i> ( <b>pPCR1</b> )	Expressão Heteróloga
pET2829PMI_FW e pET2829PMI_RV_C	Região codificante do gene <i>xanB</i> sem o stop codon nativo do gene ( <b>pPCR2</b> )	Expressão Heteróloga
pNPTS_cPMI_FW e pET2829PMI_RV	Região codificante do gene <i>xanB</i> a ser clonada no vetor pET41a ( <b>pPCR3</b> )	Expressão Heteróloga
Up_PMI_FW e Up_PMI_RV	Região de 1 kb <i>upstream</i> ao gene <i>xanB</i> ( <b>pPCR4</b> )	Deleção Gênica
Down_PMI_FW e Down_PMI_RV	Região de 1 kb <i>downstream</i> ao gene <i>xanB</i> ( <b>pPCR5</b> )	Deleção Gênica
KO_PMI_FW e KO_PMI_RV	Região de 2 kb (linhagem mutante XccΔXanB) ou 3,4 kb (linhagem selvagem Xcc)	Confirmação da deleção e complementação gênicas
pLAL6_pPMI_FW e pLAL6_pPMI_RV	Região promotora predita do gene <i>xanB</i> ( <b>pPCR6</b> )	Complementação gênica
pLAL6_cPMI_FW e pLAL6_cPMI_RV	Região codificante do gene <i>xanB</i> ( <b>pPCR7</b> )	Complementação gênica
pNPTS_cPMI_FW e pNPTS_cPMI_RV	Região codificante do gene <i>xanB</i> ( <b>pPCR8</b> )	Complementação gênica
pET2829XI_FW e pET2829XI_RV	Região codificante do gene <i>xylA</i> ( <b>pPCR9</b> )	Expressão Heteróloga
pET2829XI_FW e pET2829XI_RV_C	Região codificante do gene <i>xylA</i> sem o stop codon nativo do gene ( <b>pPCR10</b> )	Expressão Heteróloga
Up_XI_FW e Up_XI_RV	Região de 1 kb <i>upstream</i> ao gene <i>xylA</i> ( <b>pPCR11</b> )	Deleção Gênica
Down_XI_FW e Down_XI_RV	Região de 1 kb <i>downstream</i> ao gene <i>xylA</i> ( <b>pPCR12</b> )	Deleção Gênica
KO_XI_FW e KO_XI_RV	Região de 2 kb (linhagem mutante XccΔXylA) ou 3,4 kb (linhagem selvagem Xcc)	Confirmação da deleção gênica

Foram projetados dois oligonucleotídeos reversos que hibridizam na região 3' dos fragmentos gênicos codificantes da fosfomanose isomerase e da xilose isomerase, objetivando o aumento da variabilidade de produtos de expressão quanto à presença e posicionamento da(s) cauda(s) de histidinas. Além disso, um par de oligonucleotídeos permite a amplificação de um fragmento gênico que resulta na obtenção da PMI recombinante fusionada à glutationa-S-transferase (GST) quando clonado no vetor pET41a (Tabelas 2.3 e 2.4).

**Tabela 2.3** Variações de posicionamento da cauda de histidinas e fusão à glutationa-S-transferase (GST) na fosfomanose isomerase recombinante utilizando diferentes combinações de oligonucleotídeos.

Par de oligonucleotídeos	Vetor	Uso de Trombina	Cauda de Histidinas	Fusão à GST
pET2829PMI_FW e pET2829PMI_RV	pET28a	Não	N-terminal	Não
pET2829PMI_FW e pET2829PMI_RV	pET28a	Sim	Não	Não
pET2829PMI_FW e pET2829PMI_RV_C	pET28a	Não	N-terminal e C-terminal	Não
pET2829PMI_FW e pET2829PMI_RV_C	pET28a	Sim	C-terminal	Não
pET2829PMI_FW e pET2829PMI_RV	pET29b	Não	Não	Não
pET2829PMI_FW e pET2829PMI_RV_C	pET29b	Não	C-terminal	Não
pNPTS_cPMI_FW e pET2829PMI_RV	pET41a	Não	N-terminal	N-terminal

**Tabela 2.4** Variações de posicionamento da cauda de histidinas na xilose isomerase recombinante utilizando diferentes combinações de oligonucleotídeos.

Par de oligonucleotídeos	Vetor	Utilização de Trombina	Cauda de Histidinas
pET2829XI_FW e pET2829XI_RV	pET28a	Não	N-terminal
pET2829XI_FW e pET2829XI_RV	pET28a	Sim	Não
pET2829XI_FW e pET2829XI_RV_C	pET28a	Não	N-terminal e C-terminal
pET2829XI_FW e pET2829XI_RV_C	pET28a	Sim	C-terminal
pET2829XI_FW e pET2829XI_RV	pET29b	Não	Não
pET2829XI_FW e pET2829XI_RV_C	pET29b	Não	C-terminal

Sendo assim, os diferentes produtos de PCR voltados à expressão gênica de cada um dos genes-alvo puderam ser utilizados para a obtenção de distintas proteínas recombinantes (sem cauda de histidinas, com cauda N-terminal de histidinas, com cauda C-terminal de histidinas e até com as duas caudas de histidinas).

## 2.4 Extração de DNA genômico

Para a extração de DNA genômico de Xcc, uma colônia isolada foi inoculada em 10 ml de caldo LB, seguida de incubação a 30 °C e 200 rpm até que o cultivo atingisse a  $DO_{595\text{ nm}}$  igual a 0,5. As células foram coletadas por centrifugação e lavadas três vezes com PBS (Tampão 50 mM fosfato de sódio pH 7,4; 150 mM NaCl).

Em seguida, o *pellet* de bactérias foi utilizado para a extração de DNA genômico por meio do *kit PureLink™ Microbiome DNA Purification Kit* (Invitrogen), segundo instruções do fabricante. O DNA foi suspenso em água mili-Q e quantificado por espectrometria a 260 nm.

## 2.5 Amplificação e clonagem dos fragmentos gênicos de interesse

Os fragmentos gênicos de interesse descritos na Tabela 2.2 (pPCR1 a pPCR12) foram amplificados por PCR utilizando DNA genômico de Xcc como molde e os oligonucleotídeos correspondentes (Tabela 2.2).

As amplificações foram realizadas em um termociclador C1000 Touch (Bio-Rad) segundo o descrito a seguir:

---

50 ng de DNA genômico
0,25 $\mu$ M de oligonucleotídeo FW
0,25 $\mu$ M de oligonucleotídeo RV (ou RV_C)
1 U de <i>Phusion</i> DNA polimerase
1,5 mM de MgCl <sub>2</sub>
0,2 mM de dNTP
H <sub>2</sub> O MilliQ q.s.p 50 $\mu$ l

---

Os ciclos utilizados nas amplificações por PCR destinadas ao isolamento dos fragmentos gênicos de interesse estão descritos na Tabela 2.5.



**Tabela 2.5** Ciclos de PCR utilizados na amplificação dos fragmentos gênicos.

pPCR Etapa da PCR	1, 2, 3, 4, 5, 6, 7, 8 e 9		10, 11 e 12	
	Temperatura	Tempo	Temperatura	Tempo
Desnaturação inicial	98 °C	10 minutos	98 °C	10 minutos
Desnaturação	98 °C	30 segundos	98 °C	30 segundos
Hibridização	62 °C	30 segundos	64 °C	30 segundos
Extensão	72 °C	1,5 minuto	72 °C	1 minuto
Extensão final	72 °C	10 minutos	72 °C	10 minutos

35x

Após a purificação por eletroforese em gel de TAE agarose 0,8% contendo brometo de etídeo, os produtos de PCR foram utilizados em reação de ligação (com T4 DNA ligase) para clonagem no vetor pJET 1.2 (Fermentas, Anexo I), seguindo a razão molar de 3:1 (inserto:vetor).

As reações de ligação foram utilizadas na transformação de células de *E. coli* DH5 $\alpha$  competentes por tratamento com cloreto de cálcio (AUSUBEL, 2002). As células transformadas foram plaqueadas em meio ágar-LB contendo ampicilina 100  $\mu\text{g}.\text{ml}^{-1}$ .

Colônias resistentes à ampicilina foram submetidas à extração de DNA plasmidial para averiguação de clones positivos para cada um dos insertos pPCR 1-pPCR 12; Tabela 2.2) por dupla digestão do DNA plasmidial com as endonucleases específicas (Tabela 2.1), seguida de análise por eletroforese em gel de TAE agarose 0,8% contendo brometo de etídeo.

Os clones positivos (insertos com tamanho próximo ao esperado) tiveram o DNA plasmidial sequenciado pelo método do didesoxinucleotídeo marcado de Sanger (SANGER; NICKLEN; COULSON, 1977) no Centro de Pesquisas sobre o Genoma Humano e Células-tronco (CEGH-CEL), do Instituto de Biociências da USP (São Paulo-SP). As reações foram feitas utilizando o BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems™), segundo as instruções do fabricante, em um equipamento 3730 DNA Analyser (Applied Biosystems™).

## 2.6 Construção dos vetores de expressão

Os insertos referentes aos produtos de PCR 1, 2, 3 e 9 (Tabela 2.2) foram digeridos com as endonucleases específicas (Tabela 2.1), purificados por eletroforese em gel de TAE agarose 0,8% e clonados nos vetores pET28a (Novagen, Anexo II), pET29b (Novagen, Anexo III), pET41a (Novagen, Anexo IV) e pET28a, respectivamente, em reações de ligação com T4 DNA ligase (razão molar inserto:vetor de 3:1) com os vetores previamente digeridos com as enzimas de restrição específicas para cada uma das construções (*NdeI* e *XhoI* para pET28a e pET29b; *EcoRI* e *XhoI* para pET41a).

As reações de ligação foram utilizadas na transformação de células de *E. coli* DH5 $\alpha$  quimiocompetentes. As células transformantes foram plaqueadas em meio ágar-LB contendo canamicina 30  $\mu\text{g.ml}^{-1}$  e os clones resistentes foram cultivados em caldo LB adicionado de canamicina 30  $\mu\text{g.ml}^{-1}$  para fins de extração do DNA plasmidial, seguida da confirmação da presença dos insertos por eletroforese em gel de TAE agarose 0,8% após a dupla digestão com as endonucleases específicas.

Após a confirmação do sucesso das clonagens, os vetores foram denominados pET28a\_PMI, pET29b\_PMI\_C, pET41a\_GST\_PMI e pET28a\_XI, os quais possibilitam a obtenção das seguintes proteínas recombinantes: PMI fusionada a uma cauda de histidinas N-terminal, PMI fusionada a uma cauda de histidinas C-terminal, PMI fusionada à GST na região N-terminal e XI fusionada a uma cauda de histidinas N-terminal, respectivamente (ver Tabelas 2.3 e 2.4).

## 2.7 Expressão heteróloga e purificação por cromatografia de afinidade

Células quimiocompetentes de *E. coli* BL21(DE3) (Novagen) foram separadamente transformadas com os vetores de expressão obtidos e plaqueadas em ágar-LB contendo canamicina 30  $\mu\text{g.ml}^{-1}$ .

Colônias isoladas de cada uma das construções foram inoculadas em caldo LB adicionado de canamicina 30  $\mu\text{g.ml}^{-1}$  e submetidas aos ensaios de expressão induzidos por IPTG 0,1 mM, adicionado na fase *mid-log* de crescimento populacional bacteriano ( $\text{DO}_{595\text{ nm}} = 0,4$ ). Somente para os ensaios relativos à PMI, foi adicionado cloreto de zinco (Sigma) 0,2  $\mu\text{M}$  ao meio de cultura de expressão heteróloga, uma vez

que o zinco é um cofator da PMI de *Xanthomonas campestris*, proteína com elevada identidade em relação à PMI de Xcc (PAPOUTSOPOULOU; KYRIAKIDIS, 1997).

Os ensaios de expressão foram incubados por 16 horas a 18 °C em agitador orbital a 250 rpm. Ao final da indução, as células foram coletadas por centrifugação, ressuspendidas em tampão 50 mM Tris-HCl (pH 8), NaCl 100 mM e lisadas por cinco pulsos de ultrassom (Sonic Dismembrator 500, Fisher Scientific, gentilmente cedido pela Profa. Dra. Dulce Helena Ferreira de Souza) com duração de 30 segundos e amplitude de 30%, com intervalos de um minuto entre os pulsos, sob banho de gelo. Para a separação da fração solúvel, o lisado celular foi submetido à centrifugação a 4 °C, 12.000 x g por 30 minutos.

A purificação de todas as proteínas recombinantes foi realizada por cromatografia de afinidade utilizando íons metálicos imobilizados (IMAC, *Immobilized Metal Affinity Chromatography*) para as proteínas fusionadas somente à cauda de histidinas, ou glutatona para a proteína fusionada à GST.

Para IMAC, a fração solúvel do lisado celular foi aplicada em uma coluna de níquel imobilizado Ni-NTA (Novagen) e as proteínas foram eluídas da coluna por meio da aplicação de um gradiente crescente de concentrações de imidazol (de 5 mM a 500 mM). Alíquotas das frações solúvel e insolúvel, bem como alíquotas do pós-coluna (após a passagem da fração solúvel do lisado pela coluna cromatográfica de afinidade) e dos eluatos em imidazol foram submetidas à análise em SDS-PAGE (LAEMMLI, 1970) para averiguar a funcionalidade dos sistemas de expressão e purificação.

Para a PMI fusionada à GST na região N-terminal (XccGST\_PMI), o procedimento de purificação foi realizado utilizando 800 µl de uma resina de sefarose imobilizada com glutatona (Glutathione Sepharose® 4 Fast Flow, GE Healthcare), gentilmente cedida pela Profa. Dra. Dulce Helena Ferreira de Souza. A coluna foi lavada com cinco volumes de coluna de água destilada, seguida de uma etapa de equilíbrio com dez volumes de coluna com tampão 50 mM Tris-HCl (pH 8,0) 100 mM NaCl.

Em seguida, a fração solúvel do lisado celular foi aplicada à coluna por gravidade, e a eluição foi realizada com 5 ml de glutatona reduzida (Sigma) a uma concentração de 10 mM. Nesta etapa de eluição, foram coletadas alíquotas de 1 ml para averiguação da eficiência do processo de purificação, por SDS-PAGE.

Para os experimentos de atividade enzimática (item 2.8), os ensaios de expressão foram repetidos e otimizados quanto à purificação, utilizando apenas as soluções de imidazol 25 mM ou de tampão 50 mM Tris-HCl (pH 8,0), NaCl 100 mM para lavagem da coluna de níquel ou de glutatona imobilizados, respectivamente, para eliminação de contaminantes, e 500 mM de imidazol para eluição das proteínas XccPMI, XccPMI\_C e Xcc\_XI purificadas, além de 10 mM de glutatona reduzida para a eluição da proteína XccGST\_PMI.

Essas últimas frações foram dialisadas contra tampão 50 mM Tris-HCl (pH 8,0), 100 mM NaCl, e quantificadas pela medida da absorção no UV em comprimento de onda de 280 nm, em um espectrofotômetro NanoVue Plus™ (Biochrom) utilizando como base o coeficiente de extinção molar de cada uma das proteínas recombinantes obtidas.

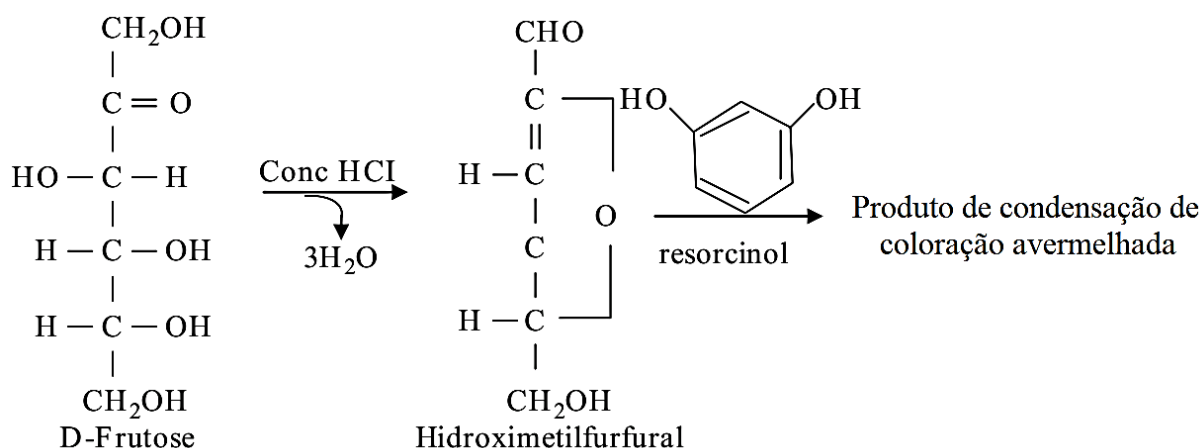
## **2.8 Ensaios de atividade enzimática**

### **2.8.1 Avaliação da atividade enzimática da PMI**

A PMI catalisa a isomerização reversível de D-manose-6-fosfato em D-frutose-6-fosfato e, portanto, o teste de Seliwanoff foi usado para detectar a formação desta cetose.

O teste de Seliwanoff (Figura 2.1) é um teste convencionalmente utilizado para a detecção de cetoses, como frutose, xilulose, ribulose e tagatose (CHAWLA, 2014), e é baseado em um reagente composto por resorcinol (benzeno-1,3-diol) 0,05% (m/v, concentração final) em ácido clorídrico (HCl) diluído em água (1:1).

A função do ácido clorídrico é desidratar a cetose e produzir um composto chamado hidroximetilfurfural, o qual reage com o resorcinol formando uma molécula de coloração avermelhada (ou esverdeada, dependendo da cetose), cuja intensidade é proporcional à quantidade de cetose presente na solução (ROE, 1934).



**Figura 2.1** Esquema das reações envolvidas no Teste de Seliwanoff (modificado a partir de KATOCH, 2011).

As reações foram constituídas por 30 µg de XccPMI\_C (cauda de histidinas na extremidade C-terminal) ou XccGST\_PMI (fusionada à GST e com cauda de histidinas na extremidade N-terminal) recombinantes, D-manose-6-fosfato (sal dissódico, Sigma) em concentração final de 0,5 mM, considerando volume final de 500 µl. Após 2 horas de reação, foram adicionados 3 ml de reagente Seliwanoff com a coloração formada sendo registrada por fotografia. Para a reação de controle negativo, o volume da enzima foi substituído pelo tampão no qual a mesma foi solubilizada (50 mM Tris-HCl (pH 8,0), 100 mM NaCl).

O controle positivo teve a mesma constituição da reação anterior, substituindo D-manose-6-fosfato por D-frutose-6-fosfato, também em concentração 0,5 mM. Os testes foram realizados em triplicata.

### 2.8.2 Avaliação da atividade enzimática da XI

A XI catalisa a isomerização da D-xilose em D-xilulose e, portanto, o teste Seliwanoff foi utilizado para detectar a formação dessa cetose em uma reação contendo XccXI recombinante (com cauda histidinas na região N-terminal), D-xilose e reagente de Seliwanoff (ROE, 1934). Para este fim, três ensaios foram preparados com as seguintes concentrações de enzima recombinante purificada: 0,05, 0,15 e 0,25 mg.ml<sup>-1</sup>.

Cada ensaio recebeu 20 µl de resorcinol a 1%, 600 µl de HCl 30%, 300 µl de D-xilose 0,5 M e seus volumes foram completados para 1,4 ml com água destilada, resultando em concentração final de 0,1 M de xilose.

Dois outros ensaios também foram preparados para servirem como branco e controle positivo. O branco recebeu 50 mM Tris-HCl (pH 8,0), 100 mM NaCl, substituindo o volume de enzima, mantendo todas as proporções dos outros compostos. O controle positivo consistiu em D-frutose 0,5 M e não recebeu enzima.

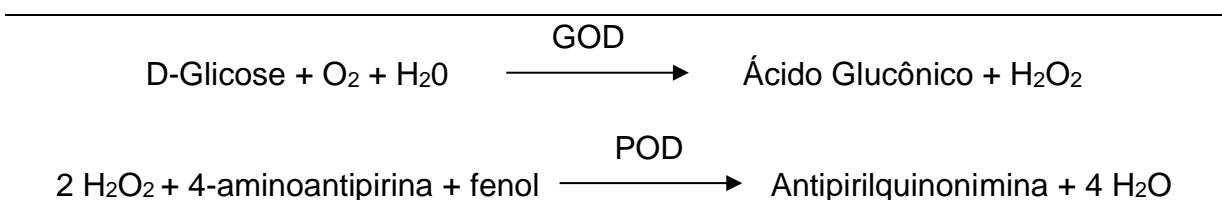
Primeiro, XccXI e seu substrato foram incubados por 2 e 3 horas. Depois disso, HCl com resorcinol e água destilada foram adicionados a todos os tubos. Em seguida, todos os tubos foram agitados vigorosamente e incubados a 80 °C por 30 minutos, com as verificações de cores realizadas a cada cinco minutos.

A segunda atividade prevista de XccXI é a interconversão de D-frutose em D-glicose. Para verificar esta atividade, um teste baseado nas reações acopladas de glicose oxidase e peroxidase foi empregado para detectar D-glicose em solução (TRINDER, 1969).

Para isso, a D-frutose foi utilizada como substrato e a detecção de D-glicose realizada com o *kit* Glucose Liquiform (Labtest), composto por tampão fosfato 30 mM, pH 7,5, fenol 1 mM, glicose oxidase 12500 U/l, peroxidase 800 U/l, 290 µM de 4-aminoantipirina, 7,5 mM de azida de sódio e surfactantes.

A D-glicose resultante da isomerização de D-frutose é oxidada pela glicose oxidase (GOD) do *kit* supracitado, formando ácido glucônico e peróxido de hidrogênio. O peróxido de hidrogênio formado participa de uma reação de oxidação envolvendo 4-aminoantipirina e fenol, catalisada pela peroxidase (POD), resultando em água e um composto avermelhado, a antipirilquinonimina, cuja intensidade de cor é diretamente proporcional à concentração de glicose liberada (Figura 2.2).

Foram utilizadas três concentrações diferentes de XccXI: 20, 60 e 100 µg.ml<sup>-1</sup>. A produção de D-glicose foi medida analisando, em triplicata, a absorbância a 490 nm das soluções após 2 e 3 horas de reação à temperatura ambiente.



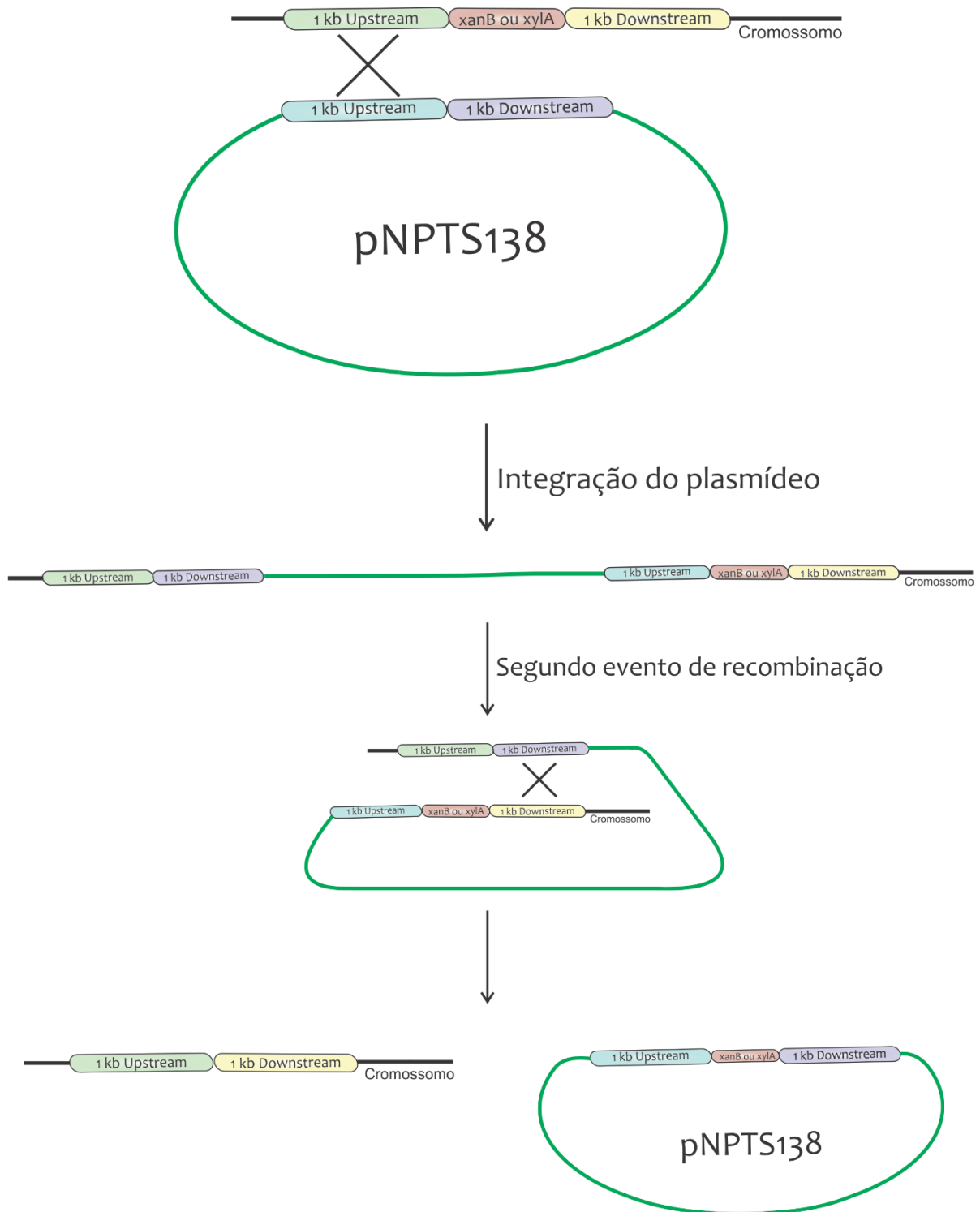
**Figura 2.2** Esquema simplificado do método de detecção de D-glicose. A D-glicose resultante da isomerização da D-frutose, catalisada pela xilose isomerase, é oxidada pela glicose oxidase (GOD), formando ácido glucônico e peróxido de hidrogênio. O peróxido de hidrogênio formado reage com 4-aminoantipirina e fenol, sob ação catalisadora da peroxidase (POD), através de uma reação oxidativa de acoplamento formando uma antipirilquinonimina vermelha, cuja intensidade de cor é proporcional à concentração de D-glicose na amostra. Adaptado do Manual de Instruções de Uso do produto Glucose Liquiform (Labtest).

## 2.9 Deleções gênicas

### 2.9.1 Construção dos vetores de deleção

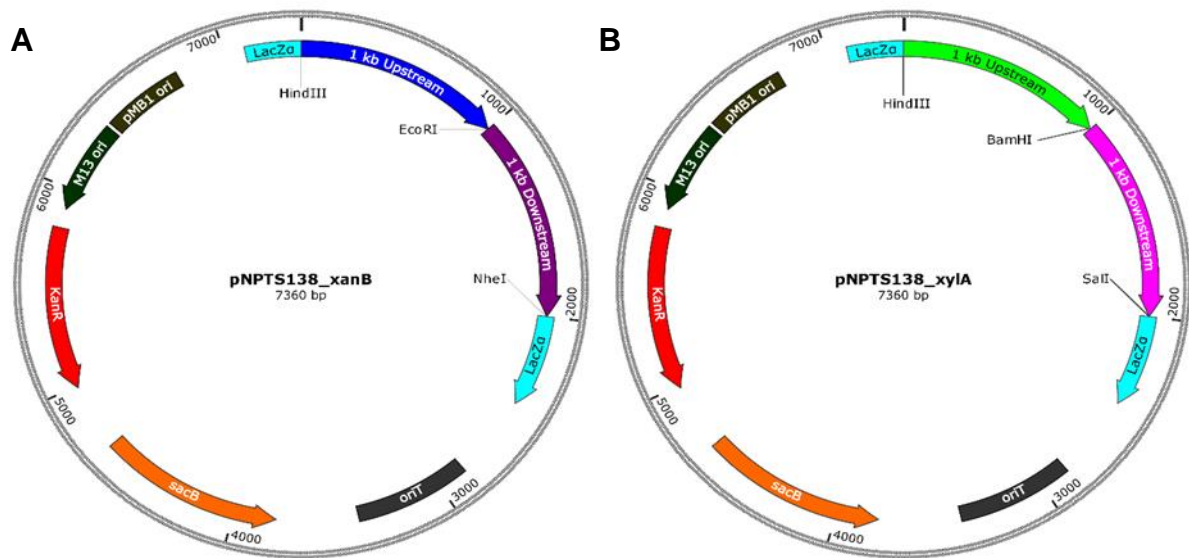
As linhagens mutantes foram obtidas usando o vetor suicida pNPTS138, gentilmente fornecido pelo Prof. Dr. Henrique Ferreira (UNESP – Rio Claro), seguindo metodologia padronizada por nosso grupo de pesquisa (ALEXANDRINO; GOTO; NOVO-MANSUR, 2016; CABREJOS et al., 2019; GOTO et al., 2016).

Os genes-alvo *xanB* e *xyIA2* foram deletados utilizando uma metodologia baseada na dupla recombinação homóloga (Figura 2.3) entre o vetor pNPTS138 (5,36 kb, Anexo V), no qual foram clonados *in tandem* dois fragmentos de aproximadamente 1 kb correspondentes, no DNA cromossomal de *Xcc*, às regiões *upstream* e *downstream* aos genes-alvo a serem deletados, resultando nos vetores de deleção do gene *xanB*, denominado pNPTS138\_*xanB*, e do gene *xyIA2*, denominado pNPTS138\_*xyIA* (Figura 2.4).



**Figura 2.3** Esquema simplificado da metodologia de deleção gênica. A metodologia de deleção gênica é baseada na dupla recombinação homóloga entre o vetor de deleção pNPTS138 e o DNA cromossomal de Xcc. O primeiro evento de recombinação ocorre com um primeiro fragmento clonado no pNPTS138 (*upstream*), e resulta na integração do vetor de deleção no DNA cromossomal de Xcc, fazendo com que as linhagens transformantes sejam resistentes à canamicina, marcador de resistência do referido vetor. O segundo evento de recombinação ocorre com um segundo fragmento clonado no pNPTS138 (*downstream*) e resulta na deleção gênica, com a transferência do gene-alvo ao vetor de deleção. As regiões *upstream* e *downstream* (de 1 kb), amplificadas e clonadas no vetor pNPTS138, são as que ladeiam, no cromossomo de Xcc, o gene a ser deletado. O vetor pNPTS138 é um vetor suicida e, portanto, não se replica nas células-filhas de Xcc. Fonte: próprio autor.





**Figura 2.4** Esquemas simplificados dos plasmídeos de deleção. O gene que confere resistência à canamicina está representado por *KanR*. *sacB* denota o gene de *Bacillus subtilis*, que codifica a enzima levansucrase, a qual converte sacarose em um composto tóxico para as células. As regiões de 1000 pb que flanqueiam no genoma de *Xcc* os genes-alvo *xanB* (A) ou *xylA* (B) estão representadas por 1 kb *downstream* e 1 kb *upstream*, com os respectivos sítios de restrição estrategicamente posicionados para clonagem dos fragmentos *in tandem*. Fonte: próprio autor.

Para isso, os insertos referentes aos produtos de PCR 4, 5, 11 e 12 foram excisados dos vetores de propagação pJET 1.2 (Fermentas) pela digestão com as endonucleases específicas para cada um deles (ver item 2.5) seguida de purificação por eletroforese em gel de TAE agarose 0,8%.

Para iniciar a construção de cada vetor de deleção, o plasmídeo pNPTS138 foi previamente digerido com as endonucleases *HindIII* e *EcoRI* para *xanB* e *HindIII* e *BamHI* para *xylA* e purificado por eletroforese em gel de TAE agarose 0,8%.

Em seguida, os fragmentos de 1 kb *upstream* purificados foram utilizados em reação de ligação com T4 DNA ligase (razão molar inserto:vetor de 3:1) juntamente com o vetor pNPTS138 digerido.

Tais reações de ligação foram utilizadas na transformação de células de *E. coli* DH5 $\alpha$  quimiocompetentes (AUSUBEL, 2002) seguida de plaqueamento em ágar-LB contendo canamicina 30  $\mu\text{g} \cdot \text{ml}^{-1}$ .

Colônias transformantes foram cultivadas em LB contendo canamicina para extração do DNA plasmidial, seguida de digestão com as endonucleases específicas para confirmação da clonagem por eletroforese em gel de TAE agarose 0,8%.

Para a clonagem dos fragmentos de 1 kb *downstream* a cada gene-alvo, seguiu-se o mesmo procedimento, utilizando os vetores já clonados com a região de 1 kb *upstream*, digerindo-os com *EcoRI* e *NheI* para *xanB* e *BamHI* e *Sall* para *xylA2*.

Novamente, os vetores digeridos e purificados compuseram reações de ligação com os fragmentos de 1 kb *downstream*, seguida da transformação de células de *E. coli* DH5 $\alpha$  quimiocompetentes (AUSUBEL, 2002) e posterior plaqueamento em ágar-LB contendo canamicina 30  $\mu\text{g}.\text{ml}^{-1}$ .

Ao final da clonagem *in tandem*, colônias transformantes de ambas as construções foram cultivadas em LB contendo canamicina 30  $\mu\text{g}.\text{ml}^{-1}$  e tiveram o DNA plasmidial extraído e digerido com as endonucleases específicas de cada vetor, permitindo a confirmação da construção dos vetores de deleção como foram planejados.

### **2.9.2 Preparação de células de Xcc eletrocompetentes**

Células de Xcc foram cultivadas em 100 ml de caldo LB, sob agitação de 200 rpm, a 30 °C até atingirem  $\text{DO}_{595 \text{ nm}}$  igual a 0,5. Em seguida, as células foram centrifugadas por 20 minutos a 10.000 x *g* a 4 °C e lavadas três vezes com água estéril gelada (AUSUBEL, 2002). As células foram concentradas 25 vezes em relação ao volume de cultivo original, alíquotadas e armazenadas a -80 °C.

### **2.9.3 Transformação de Xcc por eletroporação**

Cem nanogramas dos vetores de deleção pNPTS138\_*xanB* e pNPTS138\_*xylA* foram independentemente utilizados na eletroporação de células de Xcc eletrocompetentes. A eletroporação foi realizada com cubetas de 0,2 cm em um eletroporador *Gene Pulser Xcell Electroporation Systems* (Bio-Rad), gentilmente cedido pelo Prof. Dr. Iran Malavazi, sob as seguintes condições: 2,5 kV, 50  $\Omega$  e 50  $\mu\text{F}$ .

Após o pulso elétrico, foram adicionados 2 ml de caldo LB isento de antibióticos e o cultivo foi incubado por 2 horas sob agitação de 200 rpm a 30 °C. Em seguida, o cultivo foi plaqueado em ágar-LB contendo canamicina (30  $\mu\text{g}.\text{ml}^{-1}$ ).

#### 2.9.4 Seleção de mutantes deletados

O vetor pNPTS138 possui uma cópia do gene *sacB* (Figura 2.4), que codifica para a enzima levansucrase (SacB de *Bacillus subtilis*), que converte a sacarose em um composto tóxico para as células (BRAMUCCI; NAGARAJAN, 1996). Desse modo, o crescimento em sacarose 10% somente ocorre para as bactérias que o eliminaram (ou seja, quando há excisão do plasmídeo com o gene do cromossomo seguido de cura plasmidial).

Além disso, células resultantes de recombinação única também são eliminadas por possuírem o plasmídeo integrado no genoma. Para aumentar as chances de cura plasmidial nos mutantes que sofreram o duplo *crossing-over*, os mesmos foram cultivados repetidamente em meio de cultivo na ausência de canamicina.

Colônias mutantes de Xcc obtidas foram analisadas quanto à deleção por PCR utilizando oligonucleotídeos que hibridizam em regiões adjacentes aos fragmentos flanqueadores. Alíquotas dos cultivos bacterianos oriundos dos mutantes confirmados foram armazenadas a -80 °C.

O processo de *screening* de possíveis mutantes deletados envolve uma série de plaqueamentos e repiques, os quais estão descritos a seguir. Após a eletroporação, colônias transformantes resistentes à canamicina foram repicadas em duas condições diferentes: ágar-LB contendo canamicina a 30 µg.ml<sup>-1</sup> e sacarose 10% (Meio B) e ágar-LB contendo sacarose 10% (Meio C).

Em seguida, colônias que não tiveram crescimento nos meios B e C foram cultivadas em caldo LB isento de antibióticos (Meio D) por 12 horas, sob agitação de 200 rpm, a 30 °C, para possibilitar o primeiro evento de recombinação, que resulta na integração do vetor de deleção ao DNA cromossomal de Xcc.

Para cada processo de deleção, tal cultivo foi estriado em ágar-LB para isolamento de colônias, as quais foram repicadas em três diferentes condições:

Meio A: Ágar-LB contendo canamicina a 30 µg.ml<sup>-1</sup>;

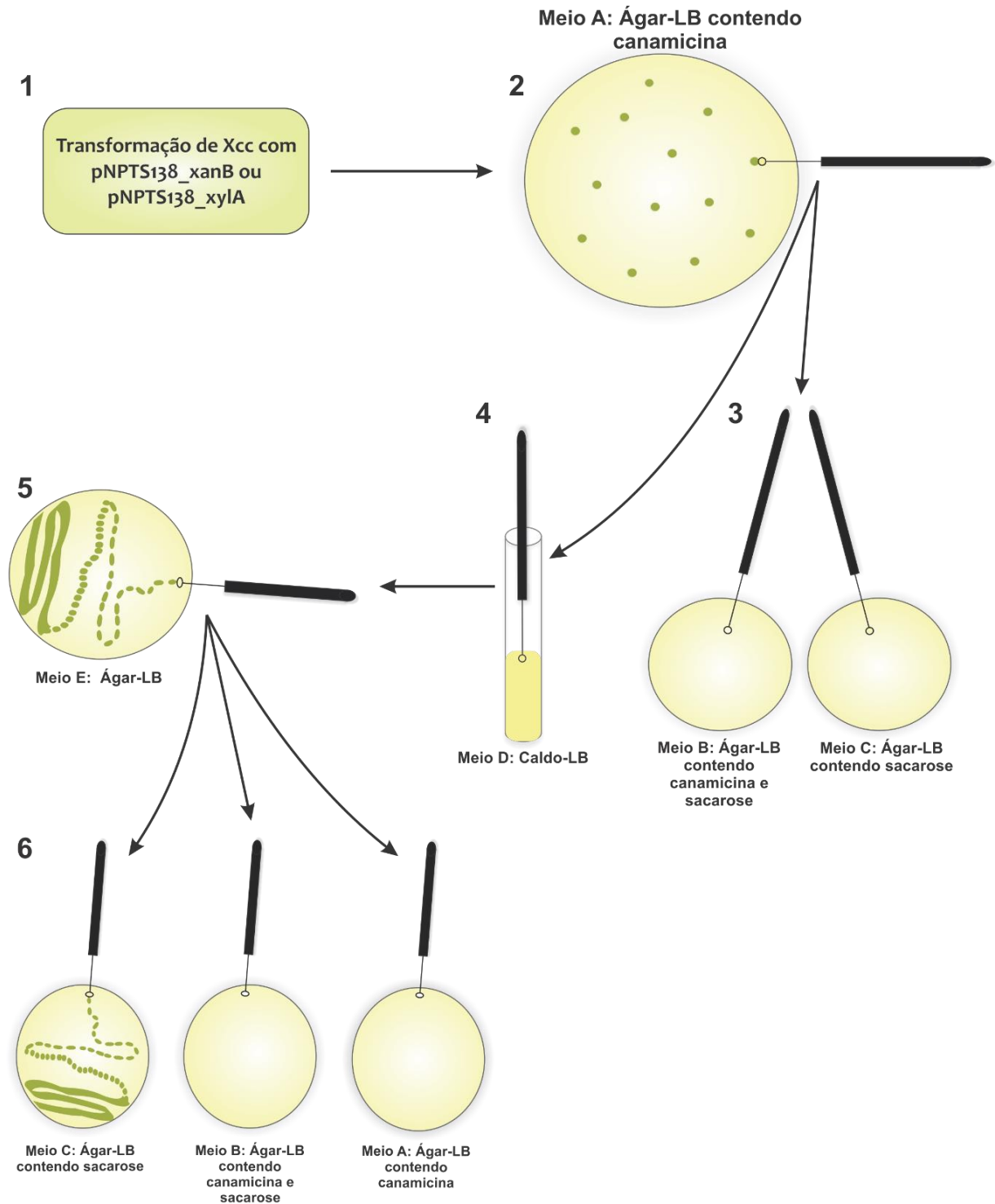
Meio B: Ágar-LB contendo canamicina a 30 µg.ml<sup>-1</sup> e sacarose 10%;

Meio C: Ágar-LB contendo sacarose 10%.

O repique de cada uma das colônias nos três meios supracitados facilita o *screening* dos possíveis mutantes deletados, uma vez que colônias resistentes à

sacarose e sensíveis à canamicina podem ser resultantes da dupla recombinação seguida da eliminação do plasmídeo de deleção.

A Figura 2.5 ilustra o procedimento adotado na seleção das possíveis colônias mutantes deletadas. Foram isoladas colônias com capacidade de crescimento no meio contendo sacarose (meio C, etapa 6, Figura 2.5) e também incapazes de crescer em meio contendo canamicina (meio A, etapa 6, Figura 2.5), sendo que tais colônias tinham se originado inicialmente de colônia capaz de crescer nesse antibiótico (meio A, etapa 2, Figura 2.5); isto denota que a célula adquire o plasmídeo mas o “elimina” posteriormente, eventos estes essenciais para a obtenção do mutante e sua confirmação, considerando que no intermédio desses dois eventos vários repiques são realizados.



**Figura 2.5** Metodologia para seleção dos possíveis mutantes deletados. Diferentes cultivos foram realizados para facilitar o *screening* de linhagens mutantes deletadas. **1-2** | Células de Xcc eletrocompetentes foram transformadas por eletroporação com cada vetor de deleção e, em seguida, plaqueadas em ágar-LB adicionado de canamicina (Meio A). **3** | Colônias transformantes foram repicadas em ágar-LB contendo canamicina e sacarose (Meio B) e ágar-LB contendo somente sacarose (Meio C). **4-5** | Colônias que não apresentaram crescimento nos meios B e C foram cultivadas em caldo-LB (Meio D) e então estriadas em ágar-LB (Meio E) para seu isolamento. **6** | Colônias isoladas foram repicadas nos meios A, B e C, para seleção dos possíveis mutantes deletados (aqueles que apresentaram crescimento apenas no meio C). Fonte: próprio autor.

### 2.9.5 Confirmação das deleções dos genes-alvo

Pares de oligonucleotídeos (KO\_PMI\_FW e KO\_PMI\_RV para o gene *xanB*; KO\_XI\_FW e KO\_XI\_RV para o gene *xyIA*, ver Tabelas 2.1 e 2.2) que hibridizam nas regiões do DNA cromossomal adjacentes às regiões flanqueadoras de 1 kb aos genes-alvo (distantes cerca de 50 pb) e não hibridizam nos respectivos vetores de deleção pNPTS138\_xanB e pNPTS138\_xyIA foram projetados para a confirmação das deleções por PCR.

A PCR utilizando esses pares de oligonucleotídeos permite diferenciar uma colônia mutante deletada de uma selvagem de acordo com os tamanhos dos produtos de PCR obtidos.

Pelo *screening* citado no item anterior, foram isoladas colônias com capacidade de crescimento no meio contendo sacarose e incapazes de crescer nos meios contendo canamicina. Cada uma dessas colônias foi ressuspensa em 49 µl de água destilada estéril e fervida por 5 minutos.

Cinco µl desse material foram utilizados como DNA molde juntamente com 5 µl de um *mix* de PCR preparado suficiente para 10 reações (descrito abaixo) de volume final de 10 µl.

---

150 pMol de oligonucleotídeo KO\_PMI\_FW ou KO\_XI\_FW  
 150 pMol de oligonucleotídeo KO\_PMI\_RV ou KO\_XI\_RV  
 0,8 mM de dNTP (0,2 mM de cada)  
 5x *Phusion Green High-Fidelity Buffer*  
 2 U de *Phusion High-Fidelity DNA* polimerase  
 H<sub>2</sub>O MiliQ q.s.p 19 µl

---

As amplificações foram realizadas em um termociclador C1000 Touch (Bio-Rad), segundo as etapas descritas na Tabela 2.6, sendo o produto da amplificação analisado por eletroforese em gel de TAE agarose 0,8%.

**Tabela 2.6** Etapas de PCR utilizadas na confirmação das deleções gênicas.

Etapa da PCR	Temperatura	Tempo
Desnaturação inicial	98 °C	10 minutos
Desnaturação	98 °C	30 segundos
Hibridização	60 °C	30 segundos
Extensão	72 °C	4 minutos
Extensão final	72 °C	10 minutos

} 45x

## 2.10 Confirmação da depleção da síntese de PMI no mutante de deleção

Complementarmente à confirmação da deleção gênica realizada por PCR, a deleção do gene *xanB* foi também avaliada por *Western blot*, por meio da constatação da depleção da síntese da PMI na linhagem mutante *XccΔxanB*.

As linhagens selvagem (*Xcc*) e mutante (*XccΔxanB*) foram cultivadas em 25 ml de meio XAM-M até que atingissem a  $DO_{595\text{ nm}}$  igual 0,4. Em seguida, os cultivos foram centrifugados a 4 °C, 12.000 x *g* por 30 minutos, ressuspendidos em 5 ml de tampão 50 mM Tris-HCl (pH 8,0), 100 mM NaCl e lisadas por cinco pulsos de ultrassom (Sonic Dismembrator 500, Fisher Scientific, gentilmente cedido pela Profa. Dra. Dulce Helena Ferreira de Souza) com duração de 30 segundos e amplitude de 30%, com intervalos de um minuto entre os pulsos, sob banho de gelo. Para a separação da fração solúvel, o lisado celular foi submetido à centrifugação a 4 °C, 12.000 x *g* por 30 minutos.

Vinte µl de cada fração solúvel do lisado celular juntamente com 20 µg da PMI recombinante (*XccPMI\_C*) foram submetidos à SDS-PAGE em duplicata de géis. Um dos géis foi corado com Coomassie Blue G-250 e o outro foi utilizado para a transferência a uma membrana de nitrocelulose (Hybond-C Extra, GE Healthcare) sob voltagem constante de 600 V por 2 horas.

Após a transferência, a membrana foi cuidadosamente retirada do sistema e com a face contendo as proteínas voltadas para cima foi corada com 0.5% Ponceau S (Sigma) em 0,1% de ácido acético para verificar o “carregamento” homogêneo entre as amostras (PEDRAS; MINIC, 2012).

A membrana foi descorada com três lavagens com água destilada e, após isso, foi adicionada uma solução de bloqueio composta por 9% (m/v) de leite em pó desnatado e liofilizado (Nestlé Molico®) em tampão TBST (20 mM Tris-HCl pH 7,4, 0,5 M NaCl, 0,05% v/v Tween-20), seguida de incubação *overnight* a 4 °C.

Após a incubação, a membrana foi lavada três vezes com TBST por 5 minutos e incubada *overnight* com o anticorpo primário IgG (anti-PMI) diluído 1:5000 em TBST. Tal anticorpo foi produzido pela empresa Proteimax (São Paulo-SP) utilizando coelhos imunizados com a proteína recombinante XccPMI obtida neste trabalho.

A membrana foi lavada três vezes com TBST por 5 minutos e incubada *overnight* com o anticorpo secundário anti-IgG de coelho conjugado com HRP, produzido em cabras (RheaBiotech, Campinas-SP), diluído 1:3000 em TBST.

Após a incubação com o anticorpo secundário, a membrana foi lavada três vezes com TBST por 5 minutos, seguida da detecção das proteínas utilizando o “*ECL Western Blotting kit*” (GE Healthcare), com as imagens sendo geradas no *software ImageLab* (Bio-Rad) acoplado ao equipamento *ChemiDOC MP Imaging System* (Bio-Rad).

## 2.11 Complementação gênica com o vetor pLAL6

### 2.11.1 Construção do vetor de complementação pLAL6\_CxanB

O pLAL6 (Anexo VI), gentilmente cedido pelo Prof. Dr. Henrique Ferreira, é um vetor de complementação integrativo, sendo construído para recombinar com o DNA cromossomal de Xcc no gene que codifica a  $\alpha$ -amilase (Xcc\_amy) (LACERDA et al., 2017).

Originalmente, o vetor apresentava um promotor induzível por arabinose (*araC*), o qual foi retirado quando da construção do vetor de complementação pLAL6\_CxanB, pois a linhagem complementante resultante seria testada *in vivo*, o que tornaria impraticável a indução da expressão do gene *xanB* por requerer inoculação de arabinose.

Como alternativa ao promotor *araC* e visando construir uma linhagem complementante contendo o promotor nativo do gene *xanB*, foi utilizada uma região de 390 pb (aqui denominada *xanP*) localizada imediatamente *upstream* ao gene *xanA* (codificante da fosfoglicomutase), identificada *in silico* como uma região promotora do



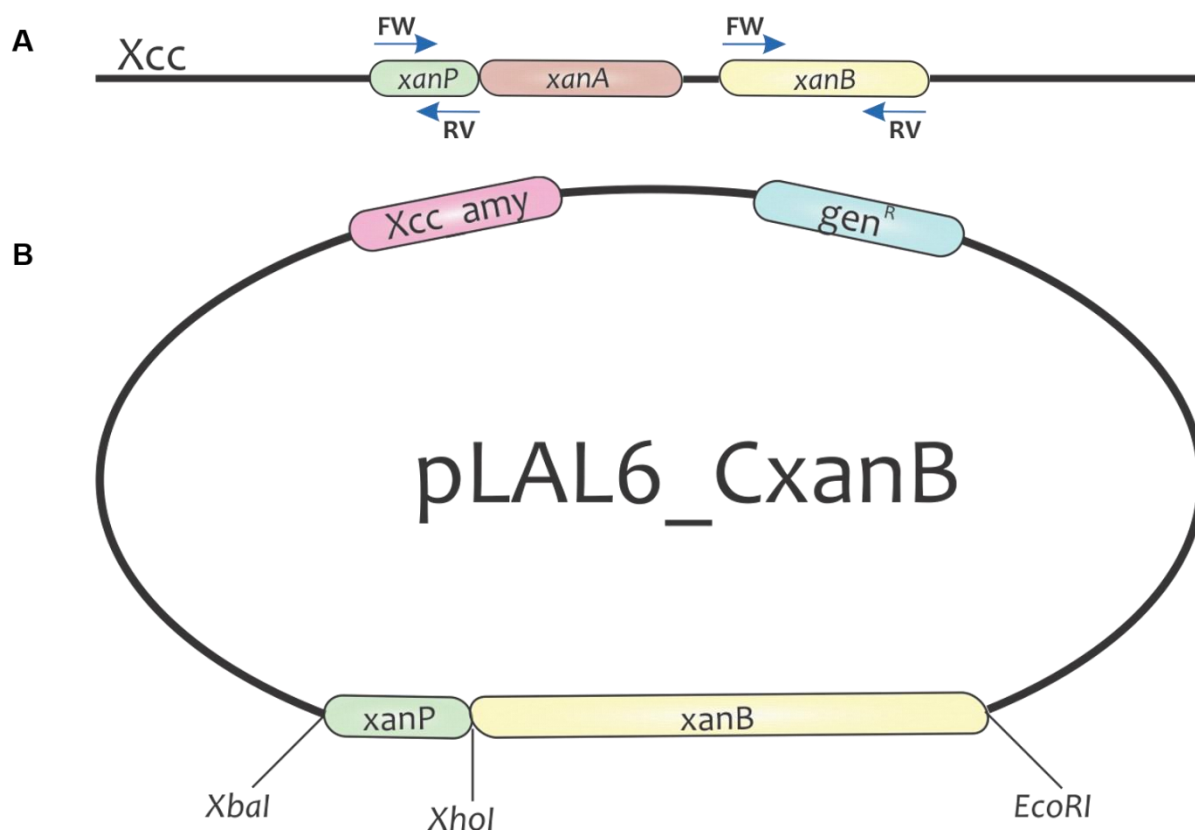
*operon xanAB* por meio do algoritmo *BPROM* (Softberry, Inc), que realiza uma predição baseada em promotores de genes regulados pelo fator  $\sigma 70$  de *E. coli*.

Assim, a região referente ao promotor predito (*xanP*) e a região codificante do gene *xanB* foram amplificadas por PCR e clonadas separadamente no vetor de propagação pJET 1.2 (Fermentas), como detalhado na seção 3.5. Em seguida, o fragmento *xanP* foi excisado do vetor de propagação pela digestão com as enzimas *XbaI* e *XhoI*, purificado por eletroforese em gel de TAE agarose 0,8% e utilizado em uma reação de ligação com T4 DNA ligase com o vetor pLAL6 previamente digerido com as mesmas enzimas supracitadas e purificado.

Essa reação de ligação foi utilizada na transformação de células quimiocompetentes de *E. coli* DH5 $\alpha$ , seguida de plaqueamento em ágar-LB suplementado com gentamicina 20  $\mu\text{g}\cdot\text{ml}^{-1}$ . Colônias transformantes foram cultivadas em caldo LB contendo gentamicina 10  $\mu\text{g}\cdot\text{ml}^{-1}$  para permitir a extração do DNA plasmidial, o qual foi digerido com as enzimas *XbaI* e *XhoI* para confirmação da clonagem por eletroforese em gel de TAE agarose 0,8%.

Após a clonagem da região *xanP*, o vetor pLAL6 recombinante foi digerido com as enzimas *XhoI* e *EcoRI*, purificado por eletroforese em gel de TAE agarose 0,8% e utilizado em reação de ligação com T4 DNA ligase juntamente com o fragmento referente à região codificante do gene *xanB*, previamente excisado do vetor de propagação pela digestão com as mesmas enzimas anteriormente citadas.

Células de *E. coli* DH5 $\alpha$  quimiocompetentes foram transformadas com a reação de ligação supracitada e plaqueadas em ágar-LB contendo gentamicina 20  $\mu\text{g}\cdot\text{ml}^{-1}$ . Do mesmo modo como foi realizado para a primeira clonagem, colônias transformantes foram cultivadas e tiveram seu DNA plasmidial digerido com as enzimas de restrição utilizadas na clonagem para confirmação por eletroforese em gel de TAE agarose 0,8%, resultando no vetor de complementação pLAL6\_C*xanB* (Figura 2.6).



**Figura 2.6** Representação do vetor de complementação pLAL6\_CxanB. **Painel A** | Esquema da organização genômica de Xcc na região do *operon xanAB* com as indicações dos pares de oligonucleotídeos (FW e RV) utilizados para isolamento e amplificação independentes das regiões xanP (região promotora predita do *operon xanAB*) e da região codificante do gene xanB. **Painel B** | Esquema simplificado da configuração do vetor de complementação pLAL6\_CxanB contendo a região Xcc\_amy, utilizada para a recombinação entre o vetor de complementação e o DNA cromossomal de Xcc, o gene que confere resistência à gentamicina (*gen<sup>R</sup>*) e as regiões xanP e xanB, clonadas *in tandem*. As enzimas de restrição utilizadas na construção do vetor de deleção estão mostradas. Fonte: próprio autor.

### 2.11.2 Seleção de complementantes

O vetor de complementação pLAL6\_CxanB foi utilizado em diferentes quantidades (1, 10, 50, 100 e 1000 ng) na eletroporação de células de XccΔxanB eletrocompetentes. A eletroporação foi realizada com cubetas de 0,2 cm de largura em um eletroporador *Gene Pulser Xcell Electroporation Systems* (Bio-Rad), gentilmente cedido para utilização pelo Prof. Dr. Iran Malavazi, sob as seguintes condições: 2,5 kV, 50 Ω e 50 μF.

Após a eletroporação, foram adicionados 2 ml de caldo LB isento de antibióticos e o cultivo foi incubado por 2 horas sob agitação de 200 rpm a 30 °C. Em seguida, o cultivo foi plaqueado em ágar-LB contendo gentamicina (20 μg.ml<sup>-1</sup>) para seleção dos transformantes.

A distinção entre colônias transformantes contendo o plasmídeo integrado no cromossomo (recombinantes) e colônias transformantes com o plasmídeo na forma episomal (não recombinantes) foi possível pelo repique de cada um dos clones resistentes à gentamicina em meio ágar-LB contendo amido 1% (m/v) seguido da sublimação de iodo (I<sub>2</sub>) nas placas, uma vez que tal molécula interage com as cadeias de amilose do amido formando um complexo de coloração azul intensa (FUWA, 1954). As colônias recombinantes teriam o vetor de complementação integrado no gene cromossomal *amy* (que codifica a  $\alpha$ -amilase), de forma a torná-lo inativo. Assim apenas as colônias recombinantes não possuem um halo de degradação do amido, estando a coloração azul presente ao redor desses transformantes, pois estas seriam incapazes de degradar o referido polissacarídeo (LACERDA et al., 2017). Analogamente, para as colônias não recombinantes são esperados halos visíveis de degradação.

## **2.12 Complementação gênica com o vetor pNPTS138**

### **2.12.1 Construção do vetor de complementação pNPTS138\_CxanB**

Como alternativa para obtenção de complementante da linhagem mutante Xcc $\Delta$ xanB, foi realizada uma tentativa inédita em Xcc de reinserção do gene *xanB* no mesmo locus cromossomal do qual foi originalmente deletado, utilizando o mutante de deleção e o vetor de deleção do gene *xanB* (pNPTS138\_xanB). A metodologia aplicada, também baseada na dupla recombinação homóloga entre o inserto no vetor (regiões flangeadoras ao gene de interesse) e o DNA cromossomal bacteriano, foi, no entanto, a inversa à utilizada para a deleção no sentido de que o gene de interesse está agora inserido no vetor pNPTS138 e ausente no cromossomo. Assim, foi realizada a construção do vetor de complementação pela clonagem do gene *xanB* no vetor recombinante pNPTS138\_xanB (utilizado na deleção de *xanB*, o qual já contém as regiões flangeadoras ao referido gene).

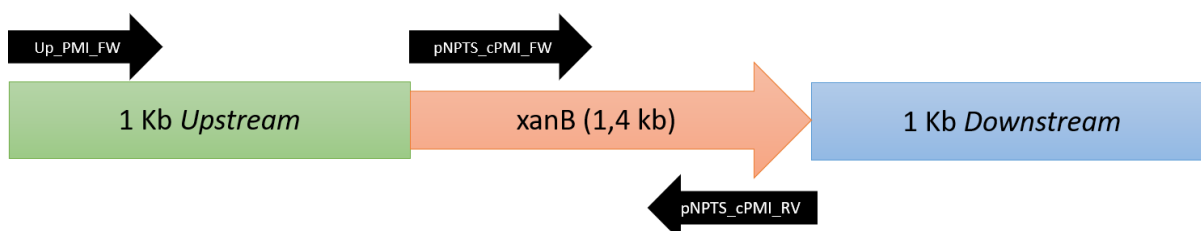
A região codificante do gene *xanB* foi amplificada por PCR a partir dos oligonucleotídeos pNPTS\_cPMI\_FW e pNPTS\_cPMI\_FW (ver Tabela 2.2), ambos com sítios de restrição para enzima *EcoRI*, sob as mesmas condições utilizadas para o fragmento obtido para a construção do vetor de expressão pET41a\_GST\_PMI.

O produto de PCR foi clonado no vetor pJET 1.2, como descrito anteriormente. Em seguida, este fragmento gênico foi clonado no vetor de deleção pNPTS138\_xanB (Figura 2.4A), previamente digerido com *EcoRI*, cujo único sítio de restrição se localiza entre os dois fragmentos flanqueadores de 1 kb, produzindo o vetor de complementação pNPTS138\_CxanB. A confirmação dessa construção foi realizada por análise de restrição por eletroforese em gel de TAE agarose 0,8%.

### 2.12.2 Confirmação da orientação de clonagem de *xanB* no pNPTS138\_CxanB

Como o fragmento gênico de *xanB* utilizado na clonagem que resultou no pNPTS138\_CxanB (item 2.12.1.) apresenta sítio de restrição para *EcoRI* em ambas as extremidades, seriam possíveis duas orientações de clonagem no vetor pNPTS138\_xanB.

Para certificar de que a clonagem ocorreu na orientação esperada, foi realizada uma PCR utilizando oligonucleotídeos que hibridizam nas seguintes regiões: codificante do gene *xanB*, 1 kb *upstream* e 1 kb *downstream* (Figura 2.7).



**Figura 2.7** Representação dos oligonucleotídeos utilizados na confirmação da orientação da clonagem no vetor pNPTS138\_CxanB. Fonte: próprio autor.

Por meio dessa estratégia, foi possível confirmar a orientação da clonagem realizada no vetor pNPTS138\_CxanB, uma vez que a PCR utilizando os oligonucleotídeos Up\_PMI\_FW e pNPTS\_cPMI\_RV resulta em um produto de PCR com 2,4 kb no caso da orientação correta da clonagem, e nenhum produto de PCR no caso de inserção de *xanB* em orientação não apropriada.

A combinação dos oligonucleotídeos Up\_PMI\_FW e pNPTS\_cPMI\_FW resulta na ausência de produtos de PCR em caso de orientação correta da clonagem, e um produto de PCR de 2,4 kb no caso da orientação inversa.

### 2.12.3 Seleção de possíveis complementantes e confirmação da complementação gênica

Após a confirmação da correta clonagem no pNPTS138\_CxanB, células de Xcc $\Delta$ xanB eletrocompetentes foram transformadas com o referido vetor de complementação, seguindo-se o mesmo procedimento utilizado na obtenção da linhagem mutante.

Após o *screening* das possíveis colônias complementantes (resistentes à canamicina, marcador de resistência do pNPTS138), foi utilizada a mesma metodologia de confirmação da deleção gênica (PCR) para atestar a obtenção da linhagem complementante pela reinserção do gene *xanB* no mutante.

## 2.13 Estudos funcionais das linhagens mutantes e complementante

### 2.13.1 Testes de patogenicidade em *Citrus aurantifolia*

As linhagens Xcc, Xcc $\Delta$ xanB, Xcc $\Delta$ CxanB e Xcc $\Delta$ xylA foram testadas quanto à patogenicidade em plantas de *Citrus aurantifolia* em estufas com temperatura e umidade controladas, gentilmente disponibilizadas pelo Fundecitrus (Araraquara-SP) e pelo Prof. Dr. Henrique Ferreira (UNESP – Rio Claro). Para o conjunto de linhagens relacionadas ao gene *xanB*, foram utilizados os métodos de aspersão e infiltração, enquanto para o gene *xylA* somente a última metodologia.

Tanto para os ensaios de aspersão quanto de infiltração, colônias isoladas foram cultivadas em 5 ml de caldo LB até atingirem OD<sub>595 nm</sub> de 0,4. Em seguida, o volume de 100  $\mu$ l dessas culturas foi centrifugado a 12.000 x g por 15 minutos a 4 °C e os *pellets* foram ressuspensos em 10 ml de solução salina a 0,9%, resultando em uma suspensão contendo aproximadamente 10<sup>6</sup> UFC.ml<sup>-1</sup>.

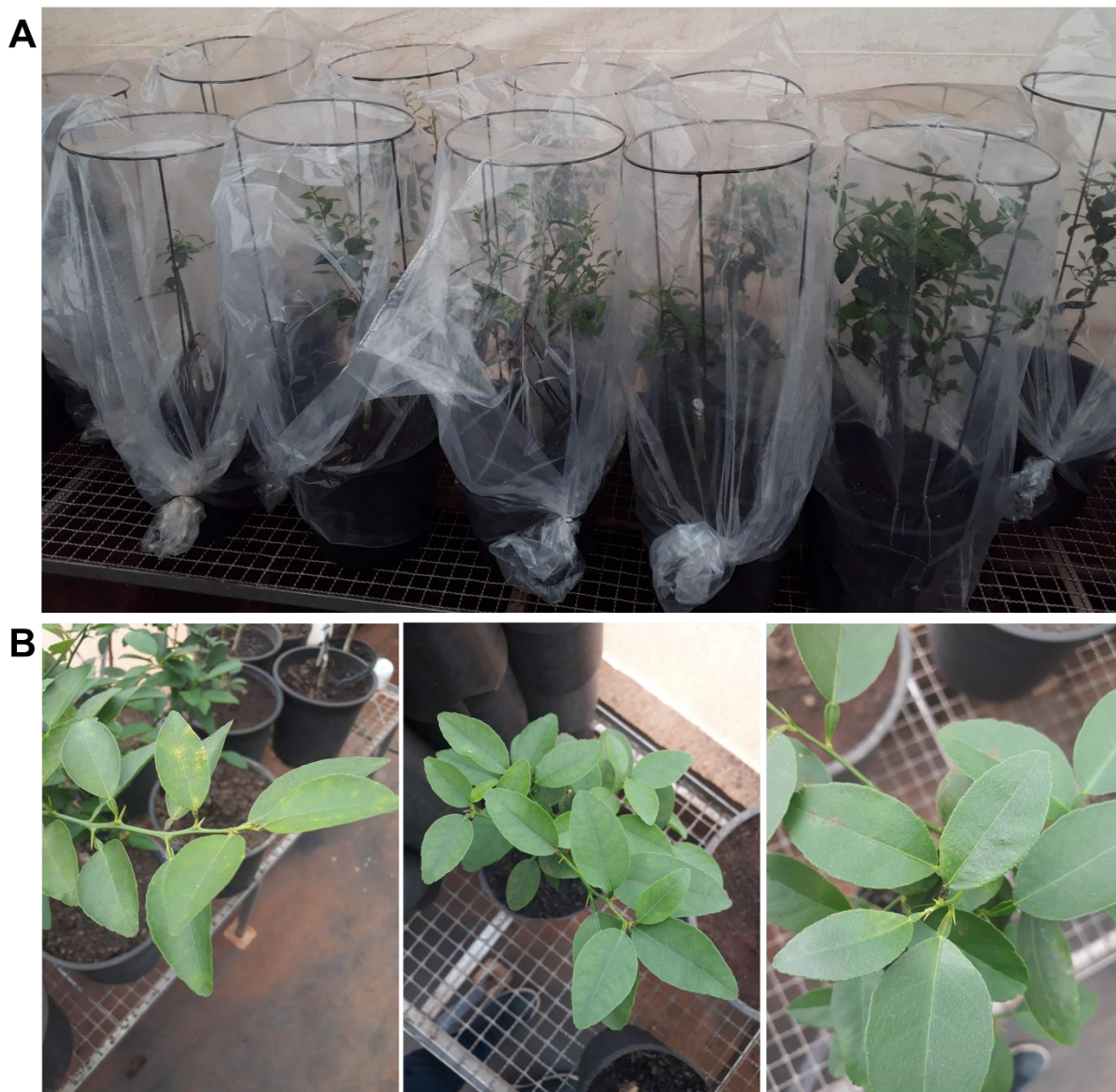
Para os testes de infiltração (Figura 2.8), 150  $\mu$ l dos cultivos celulares (ou solução salina a 0,9% para o controle negativo) foram “injetados” na face abaxial das folhas com uma seringa de 5 ml. Para cada uma das condições foram infiltradas quatro folhas de ramos independentes, atentando-se para que todas as plantas tivessem folhas infiltradas com as linhagens bacterianas e o controle negativo com solução salina.



**Figura 2.8** Exemplos de *Citrus aurantifolia*, utilizadas como hospedeiro no ensaio de patogenicidade *in vivo* por infiltração. As plantas, localizadas na casa de vegetação do Fundecitrus, na UNESP em Araraquara-SP, foram cedidas pelo Fundecitrus. Fonte: próprio autor.

Para o teste de aspersão (Figura 2.9), 10 ml das culturas bacterianas (ou solução salina a 0,9% para o controle negativo) foram pulverizados em quadruplicata de plantas, que tiveram as folhas mais suscetíveis demarcadas no início do experimento.

Os testes de infiltração e aspersão foram repetidos em dois dias diferentes, utilizando duas réplicas biológicas (culturas independentes) em cada dia e quatro repetições experimentais (quatro folhas para o teste de infiltração e quatro plantas para o teste de aspersão) para cada linhagem. As folhas foram registradas fotograficamente após vinte dias (infiltração) e 28 dias (aspersão) para comparar visualmente os sintomas do processo infeccioso. Este ensaio foi realizado na estufa do Fundecitrus em Araraquara, com a colaboração do Dr. Franklin Behlau e da Me. Tamiris Garcia.



**Figura 2.9** Ensaio de patogenicidade *in vivo* das linhagens selvagem, mutantes de deleção e complementante por aspersão. **Painel A** | Plantas de *Citrus aurantifolia* utilizadas no experimento de aspersão, após o procedimento e recobertas com plástico transparente para indução de câmara úmida. **Painel B** | Exemplos de *Citrus aurantifolia* utilizadas no experimento na casa de vegetação do Fundecitrus, em Araraquara-SP. Fonte: próprio autor.

### 2.13.2 Curvas de crescimento

As linhagens Xcc e Xcc $\Delta$ xanB foram caracterizadas quanto ao perfil de crescimento em meio XAM-M em triplicata. Para isso, as duas linhagens foram pré-inoculadas em 5 ml de caldo LB até que atingissem DO<sub>595 nm</sub> igual a 0,5. O volume total dos pré-inóculos foi transferido para 100 ml de meio XAM-M e o crescimento foi monitorado pela retirada de alíquotas de 300  $\mu$ l seguida da leitura da densidade ótica a 595 nm em um leitor de microplacas *iMark Microplate Absorbance Reader* (Bio-Rad). As leituras foram realizadas em triplicata e as médias dessas réplicas juntamente com

os desvios-padrão amostrais foram utilizados para construir uma curva de crescimento.

### 2.13.3 Ensaio de formação de biofilme

As capacidades de adesão celular e de formação de biofilme são atributos determinantes no sucesso do processo infeccioso de *Xanthomonas citri* subsp. *citri*. A mensuração da capacidade de formação de biofilme nas linhagens mutante e complementante de *xanB* produzidas, comparativamente à Xcc (selvagem) foi realizada da seguinte forma.

As linhagens Xcc, Xcc $\Delta$ xanB e XccCxanB foram cultivadas em 150  $\mu$ l de meio XAM-M dispostos em placas de 96 poços com DO<sub>595 nm</sub> inicialmente ajustada para 0,3. Os cultivos foram incubados sem agitação a 30 °C durante 24, 48 e 72 horas. Ao final de cada tempo, os poços foram lavados com NaCl 0,9% por três vezes e corados com 150  $\mu$ l de cristal violeta 0,1%.

Após trinta minutos de incubação com o corante, os poços foram lavados três vezes com água destilada. Por fim, o cristal violeta residual foi solubilizado em 150  $\mu$ l de etanol 70% e quantificado por absorbância a 595 nm (O'TOOLE; KOLTER, 1998).

### 2.13.4 Avaliação de motilidade

Para avaliar a motilidade “*swarming*”, as linhagens Xcc, Xcc $\Delta$ xanB e XccCxanB foram cultivadas em meio LB contendo 0,7% de ágar. Para este ensaio, as linhagens foram previamente cultivadas *overnight* em meio caldo LB a 30 °C sob agitação constante de 200 rpm. Em seguida, as densidades ópticas a 595 nm foram medidas e os cultivos ajustados para DO<sub>595 nm</sub> igual a 0,3.

A partir de tais cultivos, foram pipetados 3  $\mu$ l no centro das placas de Petri de 5,1 cm de diâmetro com LB 0,7%. Após isso, as placas foram incubadas a 30 °C por 48 horas sem agitação. Por fim, as placas foram registradas digitalmente e o diâmetro das colônias obtidas foi mensurado utilizando o *software* ImageJ (<https://imagej.nih.gov/ij/>) (MALAMUD et al., 2011).

A análise estatística dos dados foi realizada pelo teste de Tukey, com nível de confiança de 95% por meio do *software* Minitab (15.2).



### 2.13.5 Sobrevivência à radiação ultravioleta

A avaliação da resistência à radiação ultravioleta (UV) foi realizada a partir de um cultivo das linhagens Xcc, Xcc $\Delta$ xanB e Xcc $\Delta$ CxanB em 5 ml de LB a 37 °C por 16 horas. Em seguida, a DO<sub>595 nm</sub> foi ajustada para 0,1 (aproximadamente 3.10<sup>7</sup> UFC/ml).

Em triplicata, 100  $\mu$ l destes cultivos contidos em tubos do tipo “Eppendorf” foram expostos à radiação UV da cabine de fluxo laminar, a uma distância de 60 cm da fonte de luz, segundo descrito anteriormente (LI; WANG, 2011a). Após 15 minutos de exposição, foi feita diluição seriada e o plaqueamento em ágar-LB para contagem de UFC.

A análise estatística dos dados foi realizada pelo teste de Tukey, com nível de confiança de 95%, utilizando o *software* Minitab (15.2).

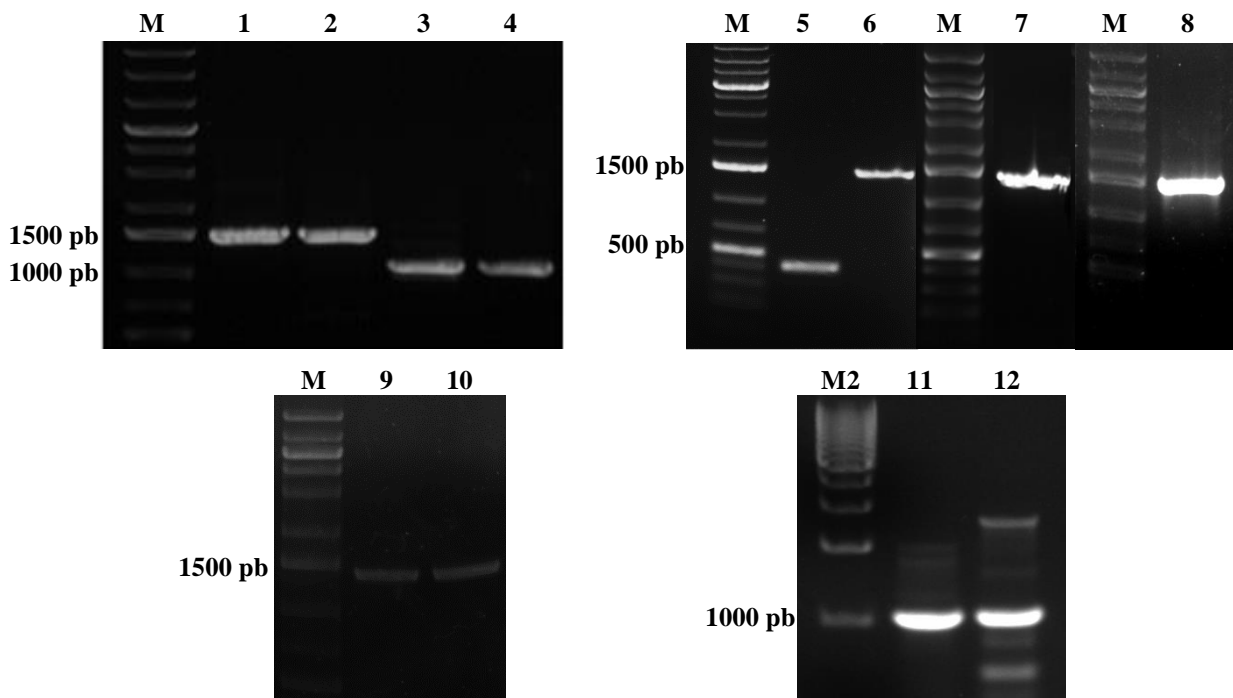
## Capítulo 3 – Resultados e Discussão

### 3.1 Amplificação por PCR e clonagem

Os pares de oligonucleotídeos desenhados foram sintetizados para permitir a amplificação dos fragmentos gênicos de interesse por PCR, seguida da clonagem em vetor de propagação pJET 1.2 (Fermentas), exceto os pares KO\_PMI\_FW-KO\_PMI\_RV e KO\_XI\_FW-KO\_XI\_RV, os quais foram utilizados apenas para confirmação da deleção e complementação gênicas.

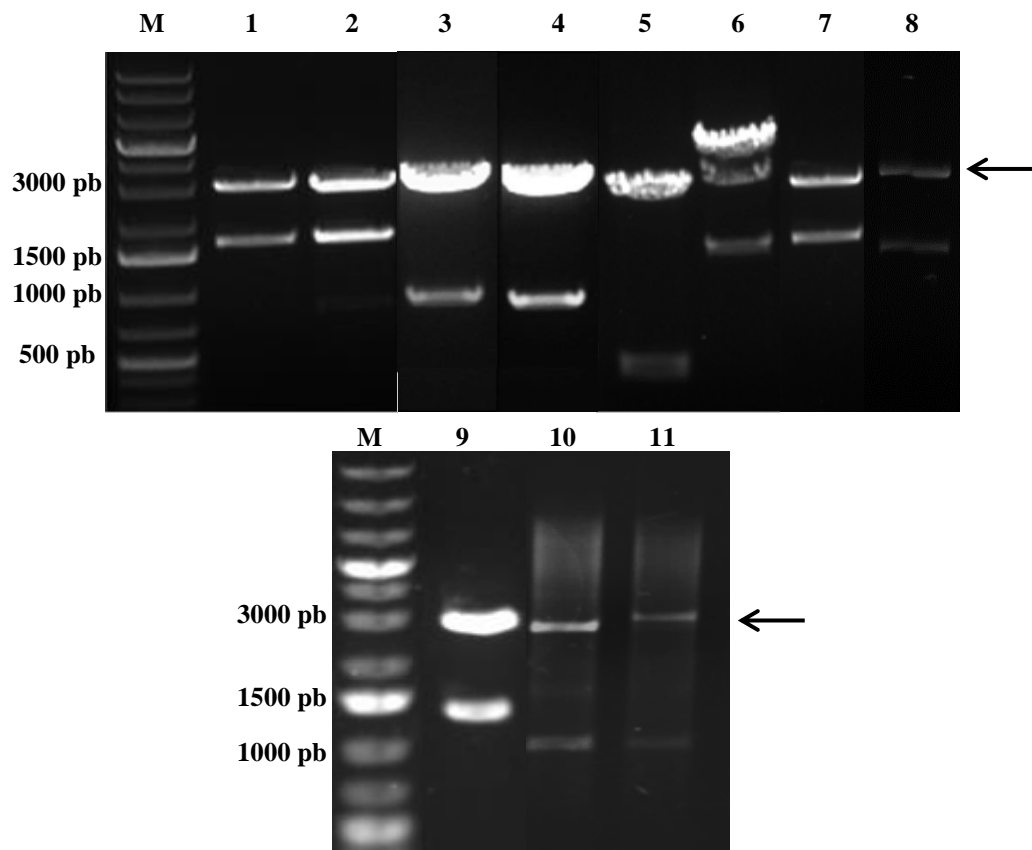
Os produtos de PCR foram analisados por eletroforese em gel de agarose 0,8% contendo brometo de etídeo (Figura 3.1), isolados a partir do gel e clonados no vetor de propagação pJET 1.2.

Os produtos das reações de ligação foram utilizados na transformação de *E. coli* DH5 $\alpha$  quimiocompetentes por tratamento com cloreto de cálcio e plaqueadas em ágar-LB adicionado de ampicilina.



**Figura 3.1** Análise dos produtos de amplificação por PCR a partir de DNA genômico de *Xcc*. Alíquotas de 3  $\mu$ L de cada reação de PCR foram analisadas em eletroforese em gel de TAE agarose 0,8%. **M** | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). **1** | Produto de PCR referente à região codificante da PMI madura (1404 pb), utilizada para clonagem no vetor pET28a. **2** | Produto de PCR referente à região codificante da PMI madura sem *stop* códon nativo (1401 pb), utilizada para clonagem no vetor pET29b. **3 e 4** | Produtos de PCR referentes aos fragmentos de 1kb *upstream* e *downstream* ao gene *xanB*, respectivamente. **5** | Produto de PCR referente à região promotora predita do gene *xanB* (390 pb). **6** | Produto de PCR referente à região codificante da PMI madura (1404 pb), utilizada na construção do vetor de complementação pLAL6\_CxanB. **7** | Produto de PCR referente à região codificante da PMI madura (1404 pb), utilizada na construção do vetor de complementação pNPTS138\_CxanB. **8** | Produto de PCR referente à região codificante da PMI madura (1404 pb), utilizada na construção do vetor de expressão pET41a\_GST\_PMI. **9** | Produto de PCR referente à região codificante da XI madura (1338 pb). **10** | Produto de PCR referente à região codificante da XI madura sem *stop* códon nativo (1335 pb). **M2** | Padrão de tamanho molecular *EZ Load™ 1 kb Molecular Ruler* (Bio-Rad). **11 e 12** | Produtos de PCR referentes aos fragmentos de 1 kb *upstream* e *downstream* ao gene *xylA*, respectivamente. Bandas do padrão mais próximas ao tamanho dos produtos estão indicadas à esquerda. Fonte: próprio autor.

Colônias isoladas foram cultivadas em caldo LB contendo ampicilina para propagação e purificação do DNA plasmidial dos transformantes. Em seguida, o DNA plasmidial referente a cada clonagem foi submetido à análise de restrição com as respectivas endonucleases para confirmação da clonagem dos insertos no vetor de propagação pJET 1.2 (Figura 3.2). Por fim, os vetores de propagação obtidos a partir das colônias positivas (tamanho dos insertos correspondente ao esperado) foram sequenciados (Centro de Pesquisas sobre o Genoma Humano e células-tronco – USP) e os insertos apresentaram 100% de identidade com as sequências dos fragmentos gênicos de interesse presentes no *GenBank*<sup>®</sup>.



**Figura 3.2** Análise de restrição dos DNAs plasmidiais para confirmação da clonagem dos fragmentos de interesse no vetor de propagação pJET 1.2. Entre 500 e 1000 ng de DNA plasmidial dos vetores de propagação foram submetidos à dupla digestão com as endonucleases específicas e analisados em eletroforese em gel de TAE agarose 0,8%. **M** | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). Cada vetor de propagação pJET 1.2 (2974 pb, Anexo I) foi digerido para excisão dos seguintes fragmentos. **1** | Região codificante do gene *xanB*. **2** | Região codificante do gene *xanB* sem o respectivo *stop* códon nativo. **3** | Região de 1 kb *upstream* ao gene *xanB*. **4** | Região de 1 kb *downstream* ao gene *xanB*. **5** | Região promotora predita (*xanP*) do operon *xanAB*. **6** | Região codificante do gene *xanB* para montagem do vetor de complementação pLAL6\_CxanB. **7** | Região codificante do gene *xanB* para montagem do vetor de complementação pNPTS\_CxanB. **8** | Região codificante do gene *xanB* para montagem do vetor de expressão pET41a\_GST\_PMI. **9** | Região codificante do gene *xyIA*. **10** | Região de 1 kb *upstream* ao gene *xyIA*. **11** | Região de 1 kb *downstream* ao gene *xyIA*. Tamanhos de bandas do padrão mais próximos aos insertos e vetor estão indicados. Uma seta indica a altura da banda do vetor de propagação pJET 1.2 (aproximadamente 3 kb), comum em todas as amostras. Fonte: próprio autor.

## 3.2 Expressão heteróloga e purificação das proteínas recombinantes

### 3.2.1 Construção dos vetores de expressão

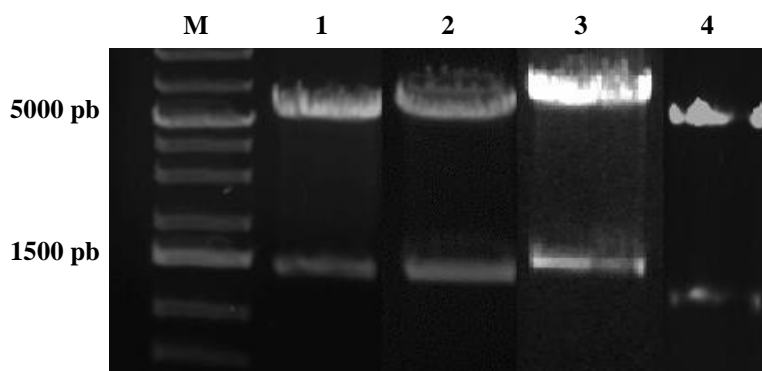
Após a dupla digestão dos vetores de propagação, os insertos referentes às regiões codificantes com *stop codon* nativo e, também, sem o referido códon (que resulta na PMI com cauda de histidinas na região C-terminal, XccPMI\_C) foram purificados a partir do gel de agarose, quantificados e utilizados em reações de ligação, utilizando T4 DNA ligase, com os vetores de expressão pET28a (PMI e XI) e

pET29b (apenas para a PMI) (Tabela 2.2), previamente digeridos com *NdeI* e *XhoI* (Tabela 2.1).

Além disso, o vetor de propagação contendo a região codificante da PMI com sítios de restrição para clonagem no pET41a foi digerido com *EcoRI* e *XhoI* (Tabelas 2.1 e 2.2) e o inserto resultante foi purificado para utilização em uma reação de ligação com o referido vetor de expressão previamente digerido com as mesmas endonucleases.

As reações de ligação foram transformadas em *E. coli* DH5 $\alpha$  quimiocompetentes por tratamento com cloreto de cálcio e os clones foram confirmados por digestão do DNA plasmidial sob eletroforese em gel de TAE agarose 0,8% (Figura 3.3).

Os clones positivos foram nomeados como pET28a\_PMI, pET29b\_PMI\_C, pET41a\_GST\_PMI e pET28a\_XI, sendo estes respectivamente os vetores de expressão da PMI com cauda de histidinas na região N-terminal (XccPMI), da PMI com cauda de histidinas na região C-terminal (XccPMI\_C), da PMI fusionada à GST na região N-terminal (XccGST\_PMI) e da XI com cauda de histidinas na região N-terminal (Xcc\_XI), respectivamente.



**Figura 3.3** Análise de restrição dos DNAs plasmidiais para confirmação da clonagem das regiões codificadoras da PMI e da XI nos vetores de expressão da série pET. Os vetores de expressão recombinante foram digeridos com as endonucleases específicas e analisados em eletroforese em gel de TAE agarose 0,8%. **M** | Padrão de tamanho molecular GeneRuler 1 kb Plus DNA Ladder (Fermentas). **1** | Vetor de expressão pET28a\_PMI digerido com *XhoI* e *NdeI*. **2** | Vetor de expressão pET29b\_PMI\_C digerido com *XhoI* e *NdeI*. **3** | Vetor de expressão pET41a\_GST\_PMI digerido com *EcoRI* e *XhoI*. **4** | Vetor de expressão pET28a\_XI digerido com *XhoI* e *NdeI*. Tamanhos de bandas do padrão mais próximos ao inserto e vetor estão indicados. Os plasmídeos pET28a e pET29b possuem cerca de 5,3 kb (Anexos II e III), já o pET41a possui 5,9 kb (Anexo IV), e os insertos referentes às regiões codificadoras da PMI e da XI possuem 1404 ou 1401 pb e 1338 pb, respectivamente. Fonte: próprio autor.

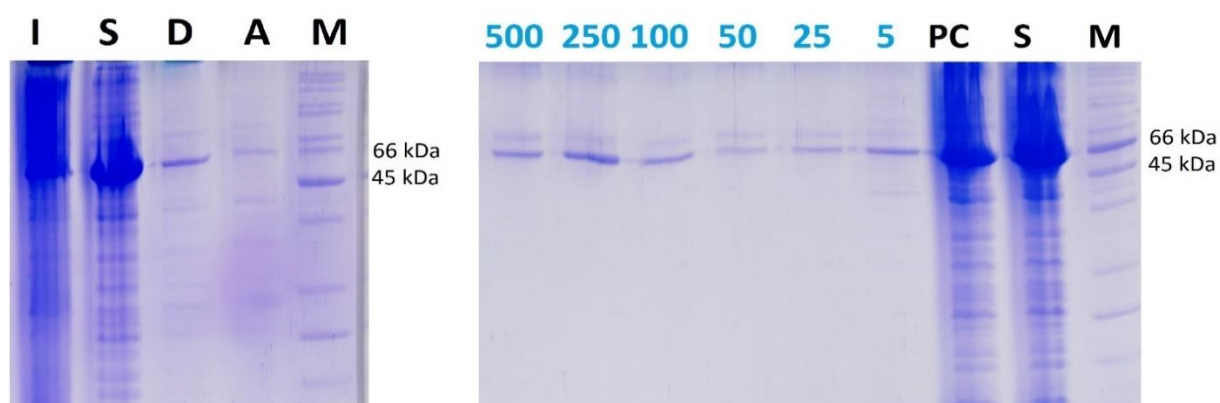
### 3.2.2 Análise da expressão heteróloga e da purificação por cromatografia de afinidade

Os vetores de expressão foram individualmente transformados em *E. coli* BL21(DE3) quimiocompetentes por tratamento com cloreto de cálcio e o plaqueamento feito em ágar-LB contendo canamicina. Uma colônia isolada de cada um dos transformantes foi utilizada para o ensaio de expressão heteróloga sob indução de IPTG. Após a indução, as células foram coletadas por centrifugação, ressuspensas em tampão e lisadas por pulsos de ultrassom, sob banho de gelo.

Para purificação dos produtos de expressão heteróloga obtidos a partir dos vetores pET28a e pET29b, a fração solúvel do lisado de cada cultivo foi aplicada em coluna de níquel imobilizado. Em seguida, a coluna foi lavada com tampão de ressuspensão (50 mM Tris-HCl (pH 8,0), 100 mM NaCl) e um gradiente crescente de concentrações de imidazol (de 5 a 500 mM) foi aplicado para eluição da proteína.

Já para a purificação da PMI fusionada à GST (pET41a\_GST\_PMI), a proteína foi eluída da coluna de glutathiona imobilizada (Glutathione Sepharose 4 Fast Flow, GE Healthcare) com 5 ml de glutathiona reduzida 10 mM.

Para todas as construções dos produtos de expressão alíquotas das frações solúveis e insolúveis, bem como alíquotas do pós-coluna e dos eluatos em imidazol e glutathiona reduzida foram submetidas à análise em SDS-PAGE (Figuras 3.4 – 3.8).



**Figura 3.4** Análise da expressão, solubilidade e purificação da XccPMI por SDS-PAGE. A figura mostra a purificação da XccPMI (aproximadamente 53 kDa), obtida a partir do vetor de expressão pET28a\_PMI. **I** | Fração insolúvel do lisado. **S** | Fração solúvel do lisado. **D** | Depois da indução por IPTG. **A** | Antes da indução por IPTG. **M** | Padrão de massa molecular *Pierce Unstained Protein MW Marker* (Thermo). **PC** | Fração solúvel após passagem pela coluna de purificação. A proteína foi eluída pela aplicação de um gradiente crescente de concentrações de imidazol (concentrações em mM, mostradas em azul e ordem decrescente de concentração). As eluições foram realizadas com volumes de imidazol referentes a cinco volumes de coluna. O ensaio de expressão teve um rendimento de 5 mg de proteína recombinante por litro de meio de cultura. Massas do padrão estão indicadas à direita. Fonte: próprio autor.

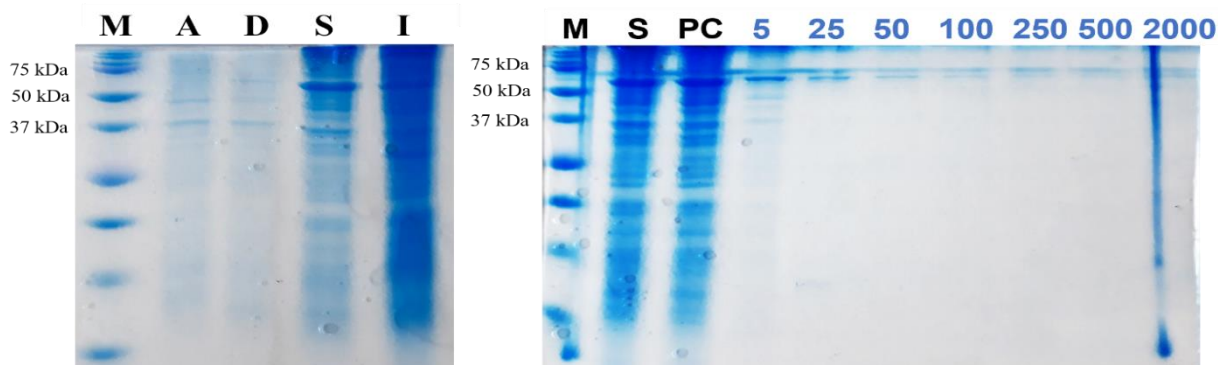
O padrão de bandas no gel da Figura 3.4 permite concluir que os sistemas de expressão e purificação funcionam a contento, uma vez que a massa molecular predita da proteína recombinante é de, aproximadamente, 53 kDa, adicionando-se a massa referente à cauda de histidinas N-terminal, pois a sequência peptídica madura da PMI consiste em aproximadamente 51 kDa.

Embora o sistema de expressão e purificação da PMI com *his-tag* N-terminal tenha funcionado satisfatoriamente, a proteína se mostrou instável em tampões TRIS, PBS, adicionados ou não de glicerol. Em todas as condições testadas, a proteína recombinante formou precipitados em diferentes temperaturas de armazenamento, incluindo -20, 4 e 25 °C.

Com o intuito de melhorar a solubilidade da proteína, foi testado o sistema de expressão e purificação da PMI fusionada à cauda de histidinas C-terminal (XccPMI\_C), utilizando o vetor pET29b\_PMI\_C (Figura 3.5).

A análise da Figura 3.5 possibilita inferir que os sistemas de expressão e purificação também funcionaram satisfatoriamente, como pode ser evidenciado pela presença de uma banda com suficiente grau de pureza na eluição com 500 mM de imidazol.

Além disso, a proteína se mostrou mais estável em temperaturas de 4 a 8 °C, permanecendo solúvel por 24 horas sob tais condições. Apesar da melhoria em relação à XccPMI, o tempo de estabilidade da XccPMI\_C ainda não foi plenamente satisfatório.



**Figura 3.5** Análise da expressão, solubilidade e purificação da XccPMI\_C sob SDS-PAGE. A figura mostra a purificação da XccPMI\_C (aproximadamente 53 kDa), sintetizada a partir do vetor de expressão pET29b\_PMI\_C. **M** | Padrão de massa molecular Precision Plus Protein™ Dual Color (Bio-Rad). **A** | Antes da indução por IPTG. **D** | Depois da indução por IPTG. **S** | Fração solúvel do lisado. **I** | Fração insolúvel do lisado. **PC** | Fração solúvel após passagem pela coluna de purificação. A proteína foi eluída pela aplicação de um gradiente crescente de concentrações de imidazol (concentrações em mM, mostradas em azul). As eluições foram realizadas com volumes de imidazol referentes a cinco volumes de coluna. O ensaio de expressão teve um rendimento de 5 mg de proteína recombinante por litro de meio de cultura. Massas do padrão estão indicadas à esquerda. Fonte: próprio autor.

Nesse sentido, foi testada a expressão, purificação, solubilidade e estabilidade da PMI fusionada à GST (XccGST\_PMI), a partir do vetor pET41a\_GST\_PMI (Figura 3.6), que resulta em uma proteína com 79 kDa, uma vez que a GST apresenta 26 kDa.

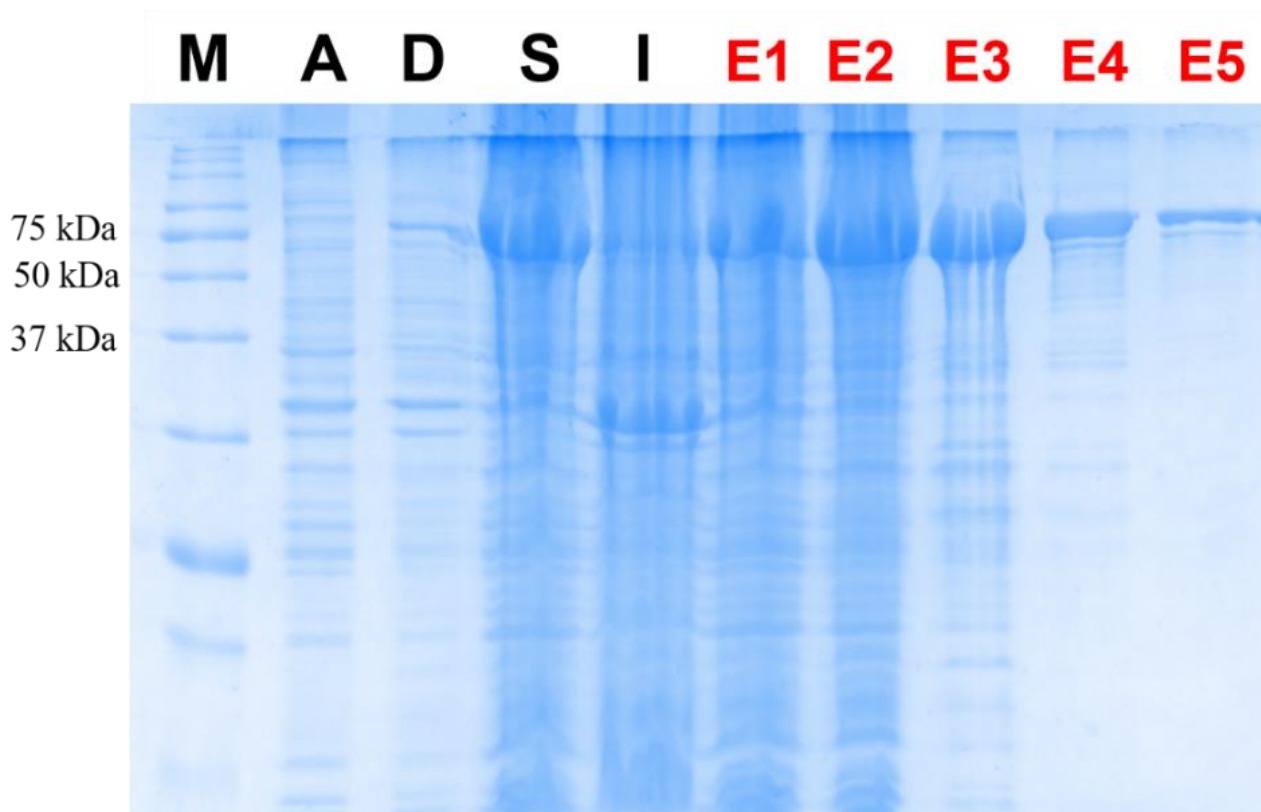
Os resultados obtidos a partir da expressão heteróloga utilizando o vetor pET41a\_GST\_PMI permitem inferir que os sistemas de expressão e purificação funcionaram de maneira adequada, fornecendo proteína relativamente pura. Ademais, a proteína purificada XccGST\_PMI apresentou-se mais estável, suportando dez dias na temperatura de 4 a 8 °C, o que tornou este sistema o mais adequado para as aplicações *downstream* da proteína recombinante.

As *tags* mais comumente utilizadas em proteínas recombinantes são 6xHis, Glutathione-S-transferase (GST), Proteína de ligação a maltose (MBP), Proteína verde fluorescente (GFP) e diversos epítomos imunogênicos. Além de permitirem a purificação, essas *tags* podem conferir vantagens à proteína-alvo durante a expressão heteróloga, aumentando a solubilidade, protegendo de proteólise, melhorando a conformação e aumentando o rendimento da produção (ARNAU et al., 2006; WOOD, 2014).

A utilização de proteínas fusionadas à GST tem sido amplamente empregada em trabalhos de expressão heteróloga de proteínas. Junior (2017) utilizou um vetor que permite a produção heteróloga de proteínas de *Trypanosoma cruzi* fusionadas à GST, resultando no correto enovelamento proteico e no aumento da solubilidade de tais proteínas. Em outro trabalho, a utilização da *tag* de GST também demonstrou ser eficiente, resultando em altos níveis de expressão do gene-alvo e, também, a otimização do processo de purificação, devido à alta afinidade da GST pela coluna de cromatografia (KAPPAUN; MARTINELLI; RITT, 2008).

Dessa forma, os aumentos de solubilidade e estabilidade atingidos com a XccGST\_PMI serão valorosos para a caracterização estrutural da proteína, cuja estrutura desse gênero ainda não foi elucidada, segundo consulta no banco de dados PDB (*Protein Data Bank*). Além disso, a obtenção da proteína recombinante pode contribuir em futuros estudos que busquem aumentar a compreensão do seu papel na interação planta-patógeno, bem como de ensaios de inibição enzimática.





**Figura 3.6** Análise da expressão, solubilidade e purificação da XccGST\_PMI por SDS-PAGE. A figura mostra a purificação da XccGST\_PMI (aproximadamente 79 kDa) em coluna de glutationa imobilizada, a partir do vetor de expressão pET41a\_GST\_PMI. **M** | Padrão de massa molecular Precision Plus Protein™ Dual Color (Bio-Rad). **A** | Antes da indução por IPTG. **D** | Depois da indução por IPTG. **S** | Fração solúvel do lisado. **I** | Fração insolúvel do lisado. A proteína foi eluída pela aplicação de 5 volumes de coluna (5 ml) de glutationa reduzida a 10 mM, sendo **E1-E5** (mostrados em vermelho) os eluatos referentes a cada um dos volumes aplicados (1 ml). Massas do padrão estão indicadas à esquerda. . O ensaio de expressão teve um rendimento de 6 mg de proteína recombinante por litro de meio de cultura. Experimento realizado durante a orientação do aluno Carlos Henrique Alves de Moraes. Fonte: próprio autor.

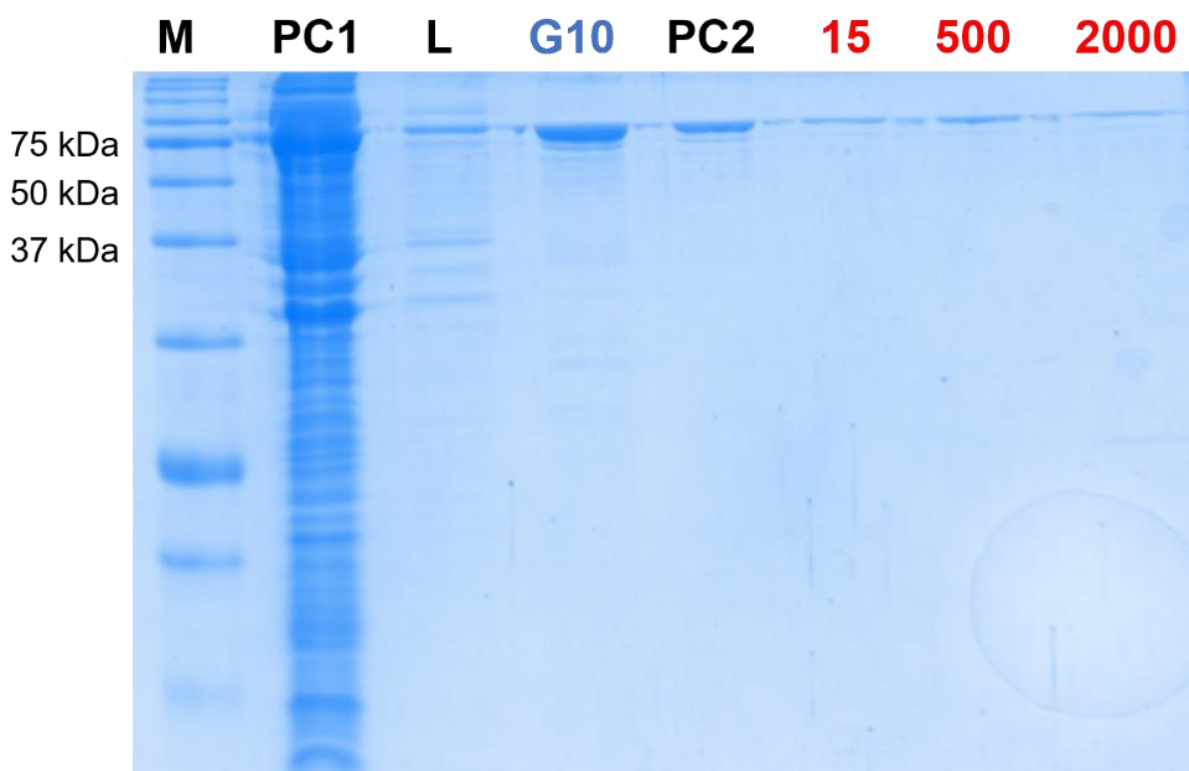
Um importante ponto a ser destacado é que o vetor pET41a\_GST\_PMI resulta em uma proteína recombinante fusionada à GST e com cauda de histidinas na região N-terminal da PMI (entre a GST e a PMI), possibilitando a purificação *in tandem* por meio da utilização de resinas com níquel e glutationa imobilizados (Figura 3.7).

A purificação *in tandem* da proteína recombinante GST\_PMI não resultou em um grau de eficiência satisfatório, uma vez que grande parte do produto recombinante atravessou a coluna de glutationa imobilizada sem adsorção pela mesma (Figura 3.7, PC1). Além disso, o mesmo ocorreu no segundo passo de purificação, quando o eluato da primeira coluna foi aplicado à coluna com níquel imobilizado (Figura 3.7, PC2).

A pouca afinidade pelas colunas poderia ser explicada pela conformação tridimensional do produto proteico, na qual tanto a cauda de histidinas quanto a GST

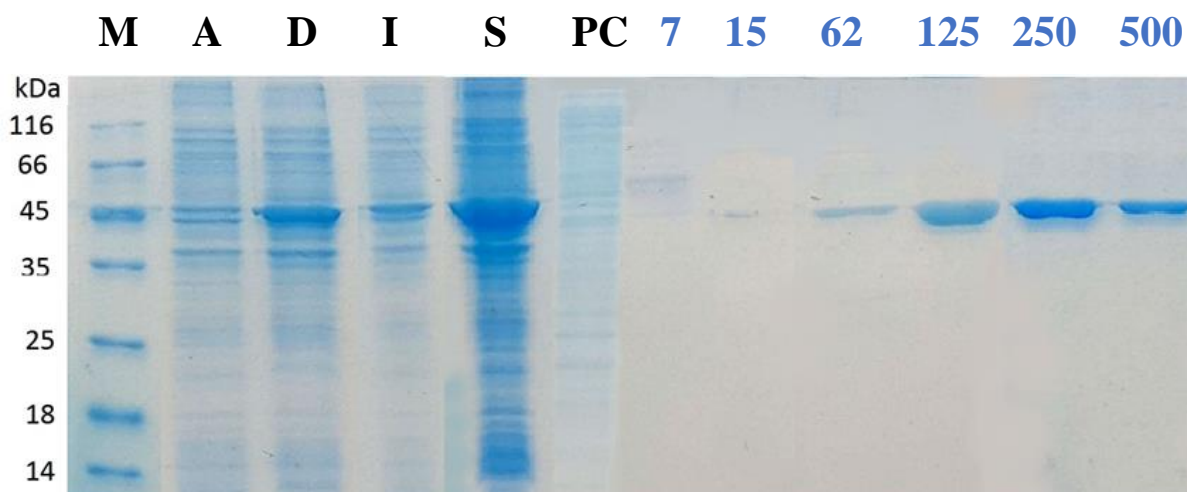
podem não estar tão expostas quanto o desejado para uma purificação por cromatografia de afinidade. Porém, mais ensaios devem ser feitos para corroborar tal hipótese.

Por outro lado, o nível de pureza obtido nos eluatos de imidazol a 500 e 2000 mM foi consideravelmente superior (Figura 3.7) se comparado ao encontrado na purificação de etapa única, utilizando a coluna com glutathiona imobilizada (Figura 3.6).



**Figura 3.7** Análise da expressão, solubilidade e purificação *in tandem* da XccGST\_PMI por SDS-PAGE. A figura mostra a purificação *in tandem* da XccGST\_PMI (aproximadamente 79 kDa), obtida a partir do vetor de expressão pET41a\_GST\_PMI. **M** | Padrão de massa molecular Precision Plus Protein™ Dual Color (Bio-Rad). **PC1** | Fração solúvel do lisado celular após passagem pela coluna de purificação com glutathiona imobilizada. **L** | Lavagem da coluna com tampão 50 mM Tris-HCl (pH 8,0), 100 mM NaCl. **G10** | Eluato a 10 mM de glutathiona reduzida. **PC2** | Eluato G10 após a passagem pela coluna de purificação com níquel imobilizado. **15** | Eluato a 15 mM de imidazol. **500** | Eluato a 500 mM de imidazol. **2000** | Eluato a 2000 mM de imidazol. Massas do padrão estão indicadas à esquerda. Fonte: próprio autor.

A partir da análise da expressão e da purificação da Xcc\_XI (Figura 3.8) foi possível averiguar que tais sistemas funcionaram de modo satisfatório, dado que a massa molecular predita da proteína recombinante é de 49,6 kDa que, somada à massa referente à cauda de histidinas N-terminal, deveria apresentar massa molecular de aproximadamente 51,6 kDa, o que é compatível com o observado.



**Figura 3.8** Análise da expressão, solubilidade e purificação da XccXI por SDS-PAGE. A figura mostra a purificação da XccXI recombinante (aproximadamente 51,6 kDa), obtida a partir do vetor de expressão pET28a\_XI. **A** | Antes da indução por IPTG. **D** | Depois da indução por IPTG. **I** | Fração insolúvel do lisado. **S** | Fração solúvel do lisado. **M** | Padrão de massa molecular Pierce Unstained Protein MW Marker (Thermo). **PC** | Fração solúvel após passagem pela coluna de purificação. A proteína foi eluída pela aplicação de um gradiente crescente de concentrações de imidazol (concentrações em mM, mostradas em azul). As eluições foram realizadas com volumes de imidazol referentes a cinco volumes de coluna. Massas do padrão estão indicadas à esquerda. O ensaio de expressão teve um rendimento de 8 mg de proteína recombinante por litro de meio de cultura. Experimento realizado durante a co-orientação da aluna Nicole Castro Silva Nicolela. Fonte: próprio autor.

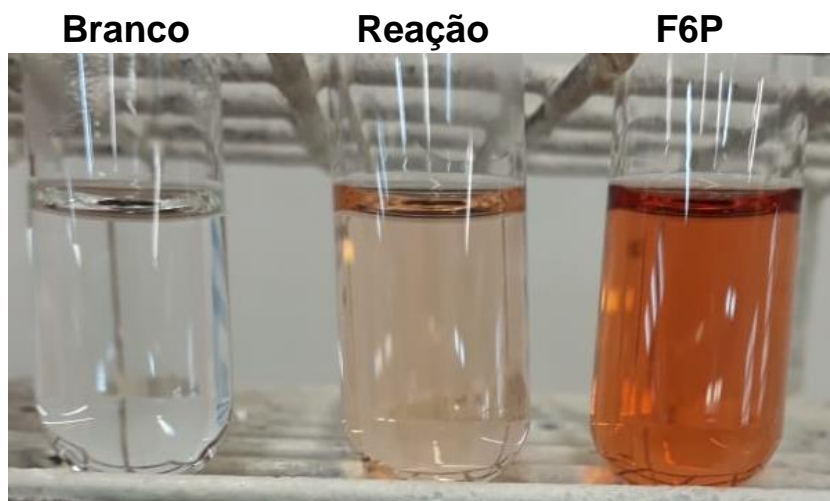
### 3.3 Avaliação das atividades enzimáticas das proteínas recombinantes

Para a análise da atividade biológica das proteínas recombinantes, os procedimentos de expressão e purificação foram repetidos, com algumas modificações para permitir a obtenção de proteínas puras e livres dos tampões de eluição que continham imidazol e glutatona.

Assim, para a XccPMI\_C, o processo de expressão foi refeito e a purificação envolveu a passagem da fração solúvel do lisado celular seguida de uma única lavagem com tampão contendo imidazol 25 mM e a eluição com tampão contendo imidazol 500 mM. Após a eluição, foi realizada diálise contra tampão 50 mM Tris-HCl (pH 8,0), 100 mM NaCl, para a retirada do imidazol do eluato.

Após diálise, a proteína recombinante purificada (0,5 ml a 0,3 mg/ml) foi utilizada em reações contendo 0,1 M de manose-6-fosfato (substrato da XccPMI\_C). Para a detecção da frutose-6-fosfato produzida pela enzima, adicionou-se 1 ml do reagente de Seliwanoff, que produz, sob altas temperaturas, uma molécula de

coloração vermelho-amarronzada na presença de cetoses, como a frutose-6-fosfato (Figura 3.9).

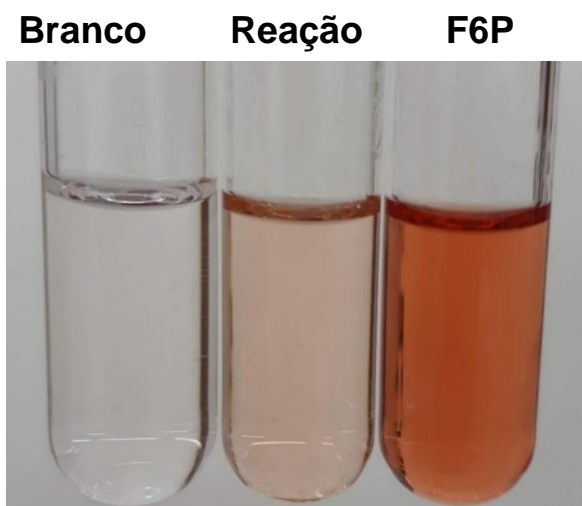


**Figura 3.9** Avaliação da atividade enzimática da XccPMI\_C. A XccPMI\_C purificada foi submetida a um teste qualitativo de atividade enzimática utilizando o reagente de Seliwanoff, o qual permite a identificação de cetoses, como a D-frutose-6-fosfato. **Branco** | Composto por 0,1 M de D-manose-6-fosfato, tampão e o reagente de Seliwanoff. **Reação** | Reação contendo 0,1 M de D-manose-6-fosfato, XccPMI\_C (0,3 mg/ml) em tampão e o reagente de Seliwanoff. **F6P** | Controle positivo contendo 0,1 M de D-frutose-6-fosfato, tampão e o reagente de Seliwanoff. A coloração vermelho-amarronzada no tubo “Reação” permitiu a conclusão de que a XccPMI\_C é capaz de desempenhar a sua atividade predita de conversão de D-manose-6-fosfato em D-frutose-6-fosfato. Todos os tampões utilizados foram compostos de 50 mM Tris-HCl (pH 8,0), 100 mM NaCl. Fonte: próprio autor.

Apesar das concentrações de manose-6-fosfato (Reação, Figura 3.9) e frutose-6-fosfato (F6P, Figura 3.9) serem as mesmas, as colorações apresentadas após a adição do reagente de Seliwanoff foram significativamente distintas, o que pode ser atribuído à não conversão total da D-manose-6-fosfato em D-frutose-6-fosfato pela XccPMI\_C, uma vez que o equilíbrio químico entre os dois monossacarídeos deve ter sido atingido.

Embora a XccPMI\_C tenha apresentado a atividade enzimática esperada, a sua instabilidade dificultou a sua utilização *downstream*, como as tentativas de cristalização. Tal proteína apresentou perda de atividade após 24 horas a 4 °C e, também, depois de um ciclo de congelamento-descongelamento.

Sob as mesmas condições testadas no ensaio anterior, a fosfomanose isomerase fusionada à GST (XccGST\_PMI) foi testada quanto à atividade enzimática de conversão de D-manose-6-fosfato em D-frutose-6-fosfato (Figura 3.10).



**Figura 3.10** Avaliação da atividade enzimática da XccGST\_PMI. A XccGST\_PMI purificada também foi submetida ao teste qualitativo de atividade enzimática utilizando o reagente de Seliwanoff. **Branco** | Composto por 0,1 M de D-manose-6-fosfato, tampão e o reagente de Seliwanoff. **Reação** | Reação contendo 0,1 M de D-manose-6-fosfato, XccGST\_PMI (0,3 mg/ml) em tampão e o reagente de Seliwanoff. **F6P** | Controle positivo contendo 0,1 M de D-frutose-6-fosfato, tampão e o reagente de Seliwanoff. A coloração vermelho-amarronzada no tubo “Reação” permitiu a conclusão de que a XccGST\_PMI também possui a atividade catalítica predita de conversão de D-manose-6-fosfato em D-frutose-6-fosfato. Todos os tampões utilizados foram constituídos de 50 mM Tris-HCl (pH 8,0), 100 mM NaCl. Fonte: próprio autor.

Assim como a XccPMI\_C (fosfomanose isomerase fusionada à cauda de histidinas na região C-terminal), a XccGST\_PMI também apresentou a atividade catalítica predita, sendo capaz de converter D-manose-6-fosfato em D-frutose-6-fosfato.

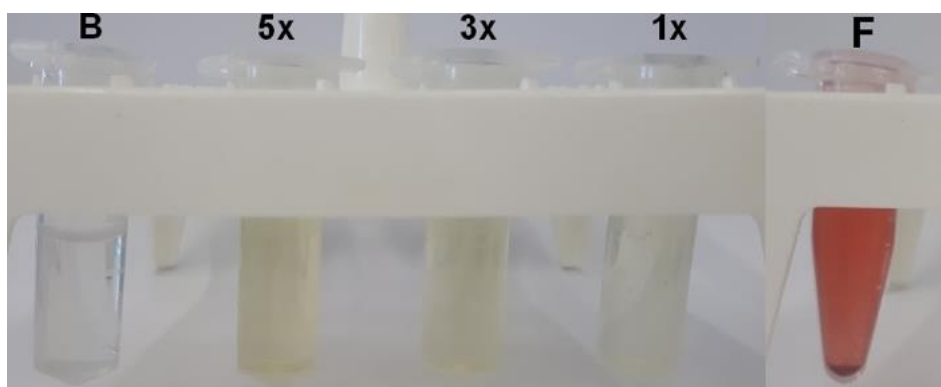
Dessa forma, tal construção foi capaz de sobrepujar a inconveniência da instabilidade apresentada pela proteína XccPMI\_C, uma vez que permaneceu estável e com atividade enzimática após dez dias de armazenamento a 4 °C e, também, após ciclos de congelamento-descongelamento.

De modo semelhante aos ensaios anteriores, a xilose isomerase recombinante (XccXI) também teve as suas atividades enzimáticas preditas avaliadas. Assim, o ensaio de expressão heteróloga foi repetido e a purificação compreendeu a passagem da fração solúvel do lisado celular seguida de uma única lavagem com tampão contendo imidazol 25 mM e a eluição com tampão contendo imidazol 500 mM. Após a eluição, foi realizada diálise contra tampão 50 mM Tris-HCl (pH 8,0), 100 mM NaCl, para a retirada do imidazol do eluato.

Primeiramente, foi testada a atividade catalítica de conversão de D-xilose em D-xilulose, utilizando o reagente de Seliwanoff. Para isso, após dialisada, a xilose isomerase recombinante purificada foi utilizada em três diferentes concentrações ( $0,25 \text{ mg.ml}^{-1}$ ,  $0,15 \text{ mg.ml}^{-1}$  e  $0,05 \text{ mg.ml}^{-1}$ ) em reações contendo  $0,1 \text{ M}$  de D-xilose.

Para a detecção da D-xilulose produzida pela enzima, adicionou-se  $1 \text{ ml}$  do reagente de Seliwanoff, sendo constatada a formação de um composto coloração amarelo-esverdeada. Como controle positivo para o reagente de Seliwanoff, foi utilizada D-frutose  $0,1 \text{ M}$  (Figura 3.11).

Apesar do teste ter sido conduzido de maneira qualitativa, a utilização de diferentes concentrações da proteína recombinante permitiu a verificação da relação diretamente proporcional (e esperada) entre a concentração da enzima e a intensidade da coloração exibida.



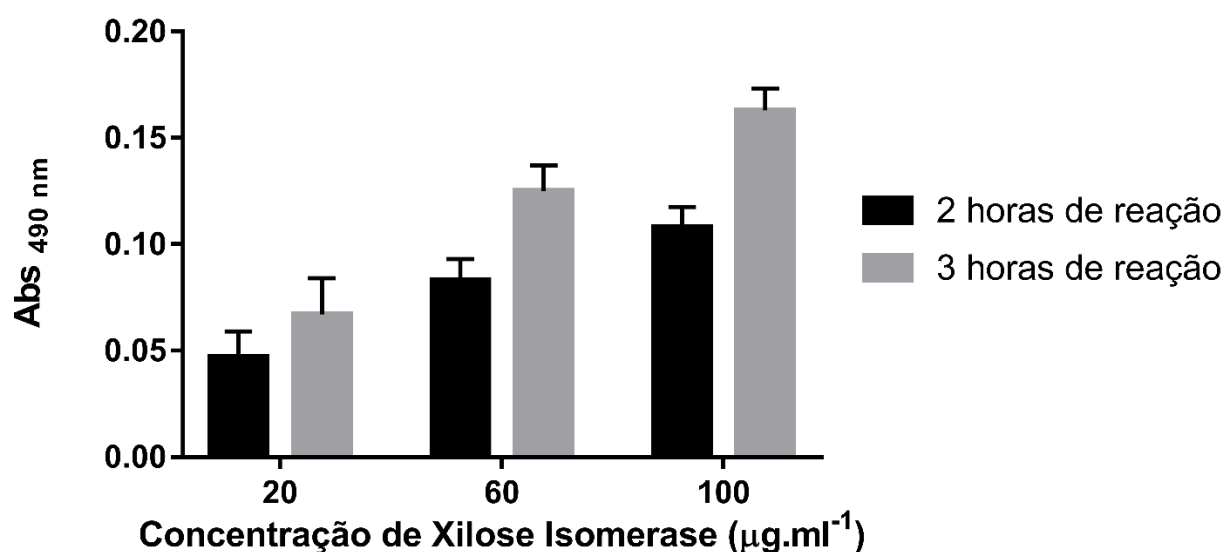
**Figura 3.11** Avaliação da atividade enzimática da XccXI. **B** | Branco contendo tampão e reagente de Seliwanoff. **5x** | Tubo contendo XI ( $0,25 \text{ mg.ml}^{-1}$ ), D-xilose  $0,1 \text{ M}$  e reagente de Seliwanoff. **3x** | Tubo contendo XI ( $0,15 \text{ mg.ml}^{-1}$ ), D-xilose  $0,1 \text{ M}$  e reagente de Seliwanoff. **1x** | Tubo contendo XI ( $0,05 \text{ mg.ml}^{-1}$ ), D-xilose  $0,1 \text{ M}$  e reagente de Seliwanoff. **F** | Tubo utilizado como controle positivo, composto de D-frutose  $0,1 \text{ M}$ , tampão e reagente de Seliwanoff. A coloração amarelo-esverdeada observada nos tubos 5x, 3x e 1x, cujas intensidades são diretamente proporcionais às quantidades de enzima utilizadas permite a inferência de que a XI recombinante apresentou sua atividade predita de conversão de D-xilose em D-xilulose. Todos os tampões utilizados eram compostos de  $50 \text{ mM}$  Tris-HCl ( $\text{pH } 8,0$ ),  $100 \text{ mM}$  NaCl. Experimento realizado durante a co-orientação da aluna Nicole Castro Silva Nicolela. Fonte: próprio autor.

A avaliação da atividade de conversão de D-frutose a D-glicose catalisada pela enzima XI foi realizada por meio da detecção de D-glicose (método de GOD-Trinder) formada após a incubação por 2 e 3 horas da enzima (em três diferentes concentrações) juntamente com o seu substrato D-frutose, a temperatura de  $25^\circ\text{C}$  (Figura 3.12).

A comparação entre as médias obtidas de cada ensaio com as diferentes concentrações de enzima nos distintos intervalos de tempo permitiu a constatação de

que, ao estender o tempo de reação, obtém-se um aumento da absorvância nas reações e, portanto, aumento da presença de D-glicose em solução, resultante da ação da XccXI sobre a D-frutose.

Além disso, foi possível verificar que quanto maior a concentração da enzima, maior a quantidade de D-glicose formada. Tais dados confirmam a existência da atividade de conversão de D-frutose em D-glicose da XccXI recombinante.



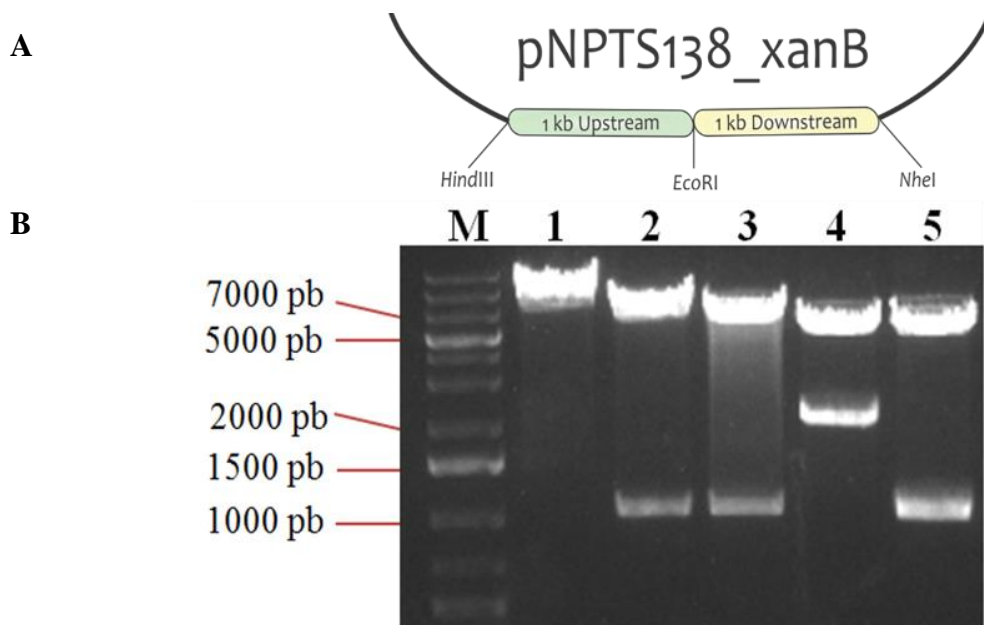
**Figura 3.12** Avaliação da atividade enzimática da XccXI. Foi avaliada a atividade enzimática da XccXI recombinante purificada na conversão de D-frutose em D-glicose. Para isso, foi utilizado um *kit* de detecção de D-glicose e três ensaios com diferentes concentrações da enzima foram realizados. As concentrações de enzima utilizadas (em µg.ml<sup>-1</sup>) foram 20, 60 e 100. Verificou-se que a absorvância em 490 nm nas reações teve um aumento em função do tempo e da concentração de XccXI adotados. O gráfico mostra as médias das medidas realizadas em triplicata, a 25 °C, após duas e três horas de reação. Experimento realizado durante a co-orientação da aluna Nicole Castro Silva Nicolela. Fonte: próprio autor.

### 3.4 Deleção gênica

#### 3.4.1 Construção dos vetores de deleção

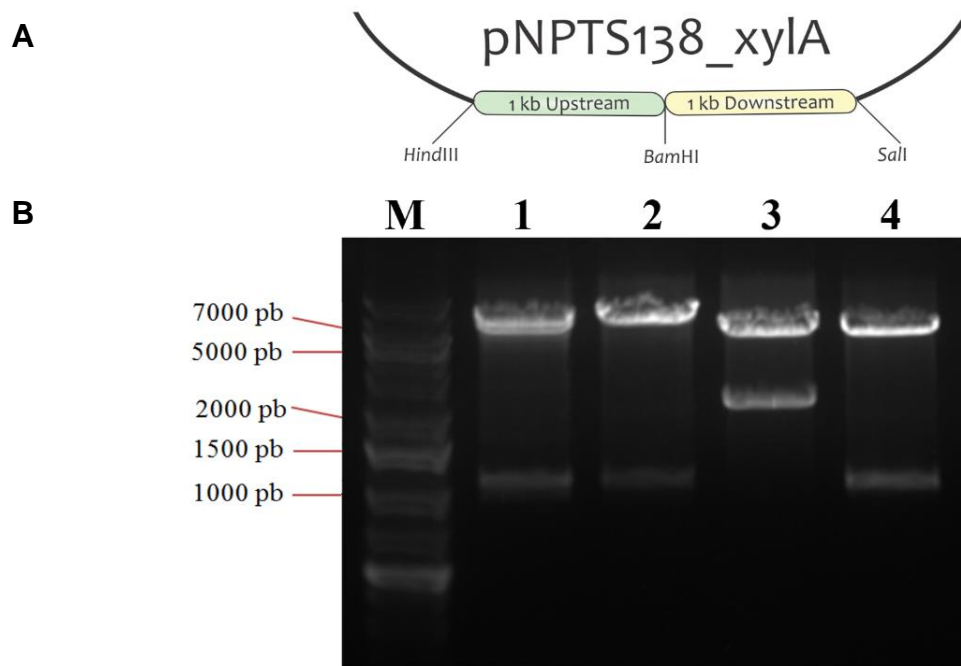
As regiões de 1 kb que flanqueiam os genes-alvo foram amplificadas por PCR usando o DNA de Xcc como molde e os pares de oligonucleotídeos mostrados nas Tabelas 2.2 e 2.3, sob as condições exibidas na Tabela 2.5. Os produtos de PCR foram clonados no vetor de propagação pJET 1.2 e sequenciados, apresentando 100% de identidade com as sequências dos fragmentos gênicos de interesse presentes no *GenBank*<sup>®</sup>.

Para cada vetor de deleção, os fragmentos gênicos referentes às regiões de 1 kb *upstream* e *downstream* aos genes-alvo foram excisados dos vetores de propagação e clonados sequencialmente *in tandem* no vetor de deleção pNPTS138, entre os sítios de restrição *Hind*III e *Nhe*I para o gene *xanB* (Figura 3.13A) e entre os sítios de restrição *Hind*III e *Sal*I para o gene *xyIA2* (Figura 3.14A). A análise de restrição dos vetores de deleção confirmou os resultados esperados (Figuras 3.13B e 3.14B).



**Figura 3.13** Análise de restrição do vetor de deleção pNPTS138\_xanB. DNA plasmidial do vetor de deleção do gene *xanB* foi submetido a diversas digestões com as endonucleases utilizadas para a sua construção e analisado em eletroforese em gel de TAE agarose 0,8%. **Painel A** | Representação gráfica de um segmento do vetor de deleção pNPTS138\_xanB. **Painel B** | Análise de restrição do vetor de deleção pNPTS138\_xanB. **M** | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). **1** | pNPTS138\_xanB digerido somente com *Eco*RI. **2** | pNPTS138\_xanB digerido com *Hind*III e *Eco*RI. **3** | pNPTS138\_xanB digerido com *Eco*RI e *Nhe*I. **4** | pNPTS138\_xanB digerido com *Hind*III e *Nhe*I. **5** | pNPTS138\_xanB digerido com *Hind*III, *Eco*RI e *Nhe*I. Tamanhos de bandas do padrão estão indicados. Fonte: próprio autor.





**Figura 3.14** Análise de restrição do vetor de deleção pNPTS138\_xylA. DNA plasmidial do vetor de deleção do gene que codifica a XI foi submetido a diversas digestões com as endonucleases utilizadas em seu desenho e analisado em eletroforese em gel de TAE agarose 0,8%. **Painel A** | Representação gráfica de um segmento do vetor de deleção pNPTS138\_xylA. **Painel B** | Análise de restrição do vetor de deleção. **M** | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). **1** | pNPTS138\_xylA digerido com *HindIII* e *BamHI*. **2** | pNPTS138\_xylA digerido com *BamHI* e *Sall*. **3** | pNPTS138\_xylA digerido com *HindIII* e *Sall*. **4** | pNPTS138\_xylA digerido com *HindIII*, *BamHI* e *Sall*. Tamanhos de bandas do padrão estão indicados. Fonte: próprio autor.

Assim, em termos de organização gênica, os vetores de deleção são semelhantes ao genoma de *Xcc*, exceto pela substituição dos genes-alvo pelo sítio de restrição de uma enzima: *EcoRI* para *xanB* e *BamHI* para *xylA2*.

### 3.4.2 Confirmação das deleções gênicas

Os vetores de deleção foram utilizados para transformação de células de *Xcc* eletrocompetentes por eletroporação, seguida de plaqueamento em ágar-LB contendo canamicina. Em seguida, colônias transformantes foram cultivadas por várias gerações na ausência de antibióticos para permitir os eventos de recombinação homóloga.

Após isso, esse cultivo foi estriado em placas contendo sacarose, o que permite a seleção de possíveis mutantes de deleção, pois o vetor pNPTS138 carrega um gene denominado *sacB*, o qual codifica para a enzima levansucrase, que converte a sacarose em um composto tóxico para a bactéria (BRAMUCCI; NAGARAJAN, 1996).

Colônias crescidas em meio de cultura contendo sacarose (o que indica cura plasmidial) foram avaliadas quanto à deleção gênica por meio de PCR de colônia, em reações utilizando os oligonucleotídeos KO\_PMI\_FW e KO\_PMI\_RV (para o gene *xanB*) e KO\_XI\_FW e KO\_XI\_RV (para o gene *xyIA*), os quais hibridizam nas regiões adjacentes às regiões flangeadoras dos genes-alvo.

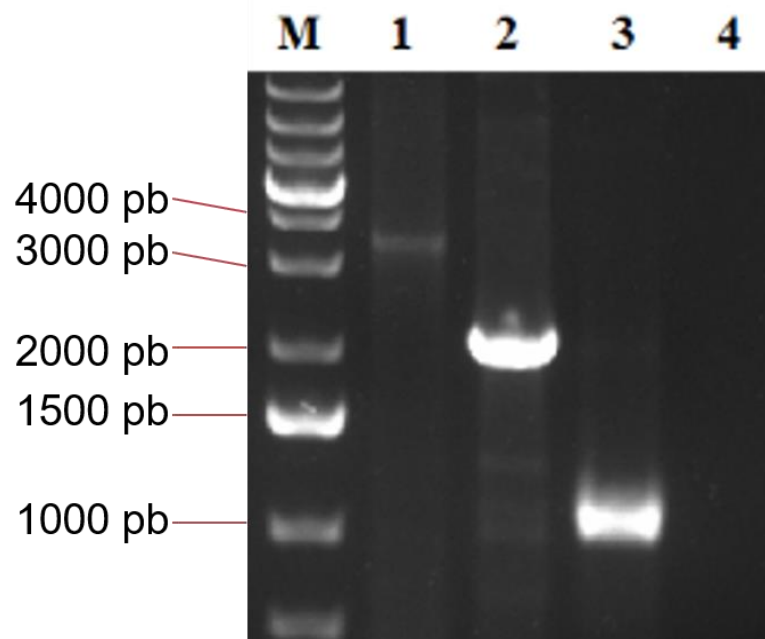
Desse modo, é possível distinguir colônias selvagens de mutantes de deleção pelo tamanho dos produtos de PCR: para o gene *xanB*, colônias selvagens resultam em produtos de PCR de 3,404 kb de tamanho (1 kb da região *upstream* + 1,404 kb da região codificante do gene *xanB* + 1 kb da região *downstream*); colônias mutantes resultam em produtos de PCR com 2 kb de tamanho (1 kb da região *upstream* + 1 kb da região *downstream*).

Já para o gene *xyIA2*, colônias selvagens resultam em produtos de PCR de 3,338 kb de tamanho (1 kb da região *upstream* + 1,338 kb da região codificante do gene *xyIA2* + 1 kb da região *downstream*); colônias mutantes resultam em produtos de PCR com 2 kb de tamanho (1 kb da região *upstream* + 1 kb da região *downstream*).

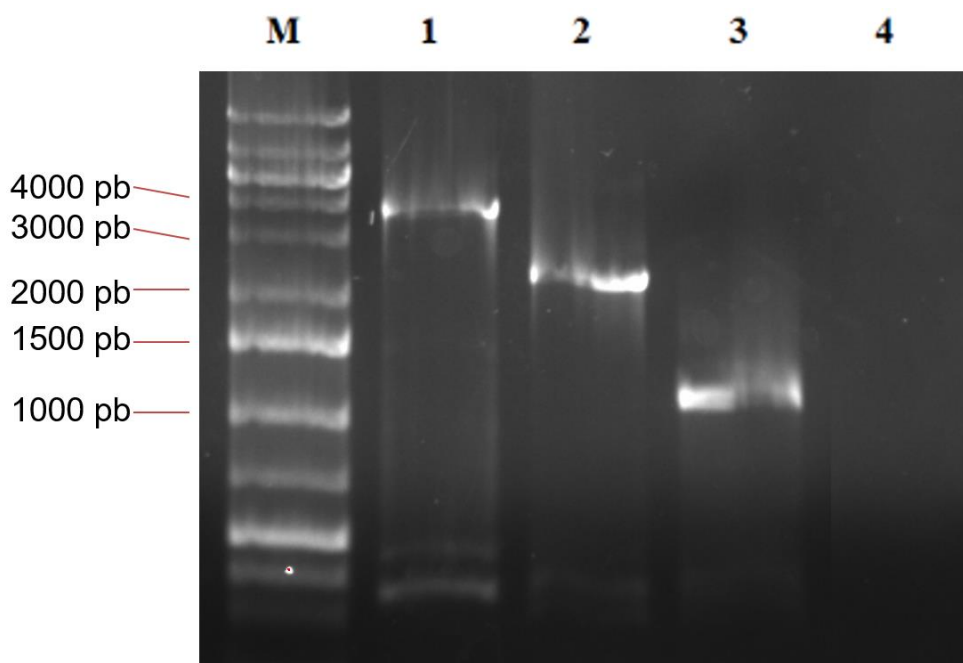
A análise dos produtos de PCR (Figuras 3.15 e 3.16) permite concluir que as deleções dos genes que codificam a PMI e a XI foram realizadas com sucesso, resultando nas linhagens denominadas *XccΔxanB* e *XccΔxyIA2*.

Objetivando uma certificação adicional de que os produtos de PCR resultantes das reações utilizando DNA genômico de *XccΔxanB* e *XccΔxyIA2* como moldes se referem às regiões cromossômicas de interesse, tais produtos foram digeridos com as enzimas de restrição cujos sítios de restrição foram inseridos após dupla recombinação homóloga (*EcoRI* para o gene *xanB* e *BamHI* para o gene *xyIA2*).

Além disso, a ausência de produtos de PCR em reações cujos moldes foram os próprios vetores de deleção confirma que os oligonucleotídeos utilizados não hibridizam nestes vetores.



**Figura 3.15** Confirmação da deleção do gene *xanB* por PCR. Análise dos produtos de PCR resultantes das reações utilizando oligonucleotídeos que hibridizam em regiões externas às regiões *upstream* e *downstream* de 1 kb ao gene *xanB*, com DNA molde de colônias de Xcc selvagem, XccΔ*xanB* e o vetor de deleção pNPTS138\_ *xanB*. **M** | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas); **1** | Produto de PCR utilizando Xcc selvagem como molde. **2** | Produto de PCR utilizando XccΔ*xanB* como molde. **3** | Produto de PCR utilizando XccΔ*xanB* como molde, digerido com a endonuclease *EcoRI*. **4** | Produto de PCR utilizando o vetor pNPTS138\_ *xanB* como molde. As bandas relevantes do padrão de tamanho molecular estão indicadas. Fonte: próprio autor.



**Figura 3.16** Confirmação da deleção do gene *xylA2* por PCR. Análise dos produtos de PCR resultantes das reações utilizando oligonucleotídeos que hibridizam em regiões externas às regiões *upstream* e *downstream* de 1 kb ao gene *xylA2*, com DNA molde de colônias de Xcc selvagem, XccΔ*xylA2* e o vetor de deleção pNPTS138\_ *xylA*. **M** | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). **1** | Produto de PCR utilizando Xcc selvagem como molde. **2** | Produto de PCR utilizando XccΔ*xylA2* como molde. **3** | Produto de PCR utilizando XccΔ*xylA2* como molde, digerido com a endonuclease *BamHI*. **4** | Produto de PCR utilizando o vetor pNPTS138\_ *xylA* como molde. As bandas relevantes do padrão de tamanho molecular estão indicadas. Fonte: próprio autor.

### 3.4.3 XccΔxanB é incapaz de sintetizar a PMI

Como uma certificação adicional da deleção do gene *xanB*, o extrato total de proteínas de XccΔxanB, após crescimento em meio indutor de patogenicidade XAM-M, foi avaliado quanto à presença da PMI por *Western blot*, utilizando como controles o extrato total de proteínas de Xcc, além da proteína recombinante XccPMI (Figura 3.17).



**Figura 3.17** Confirmação da depleção da síntese da PMI por *Western blot*. Os extratos proteicos totais das linhagens Xcc e XccΔxanB cultivadas em meio XAM-M foram avaliados acerca da presença da proteína PMI por *Western blot*. No ensaio, a proteína recombinante XccPMI foi utilizada como controle positivo. Fonte: próprio autor.

A ausência de bandas na revelação do *Western blot* obtida para a linhagem XccΔxanB corrobora a deleção do gene *xanB*. Para Xcc, foi possível observar que a proteína PMI é produzida em meio indutor de patogenicidade, XAM-M. A presença de banda fortemente marcada para a proteína recombinante assegura a funcionalidade do sistema de avaliação da expressão gênica baseado em *Western blot*.

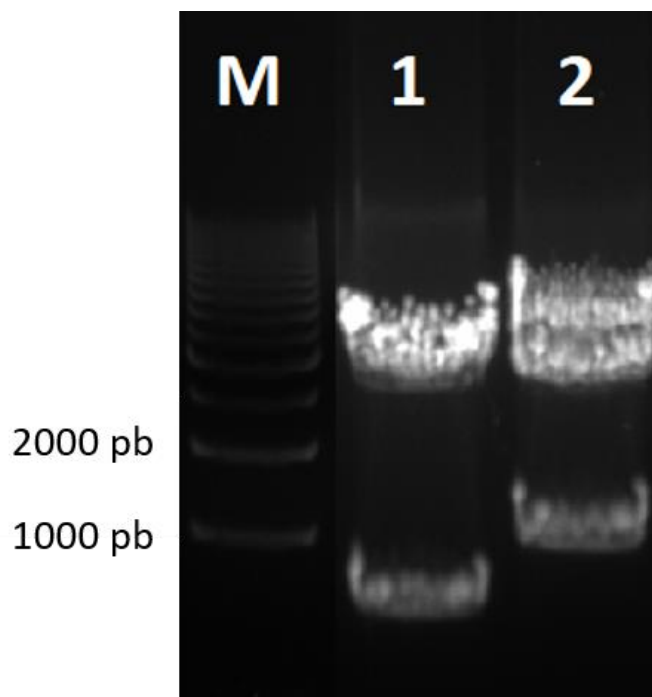
## 3.5 Complementação gênica utilizando o vetor pLAL6

### 3.5.1 Construção do vetor pLAL6\_CxanB

Os clones de propagação referentes às regiões promotora e codificante do gene *xanB* foram digeridos com suas endonucleases específicas, purificados do gel por eletroforese em gel de TAE agarose 0,8% e quantificados por espectrometria a 260 nm.

Em seguida, os dois fragmentos foram clonados *in tandem* no vetor de complementação pLAL6 (LACERDA et al., 2017), gentilmente cedido pelo Prof. Dr. Henrique Ferreira (UNESP – Rio Claro), resultando no vetor de complementação denominado pLAL6\_xanB.

A análise de restrição do vetor pLAL6\_xanB mostra que o plasmídeo foi construído a contento (Figura 3.18).



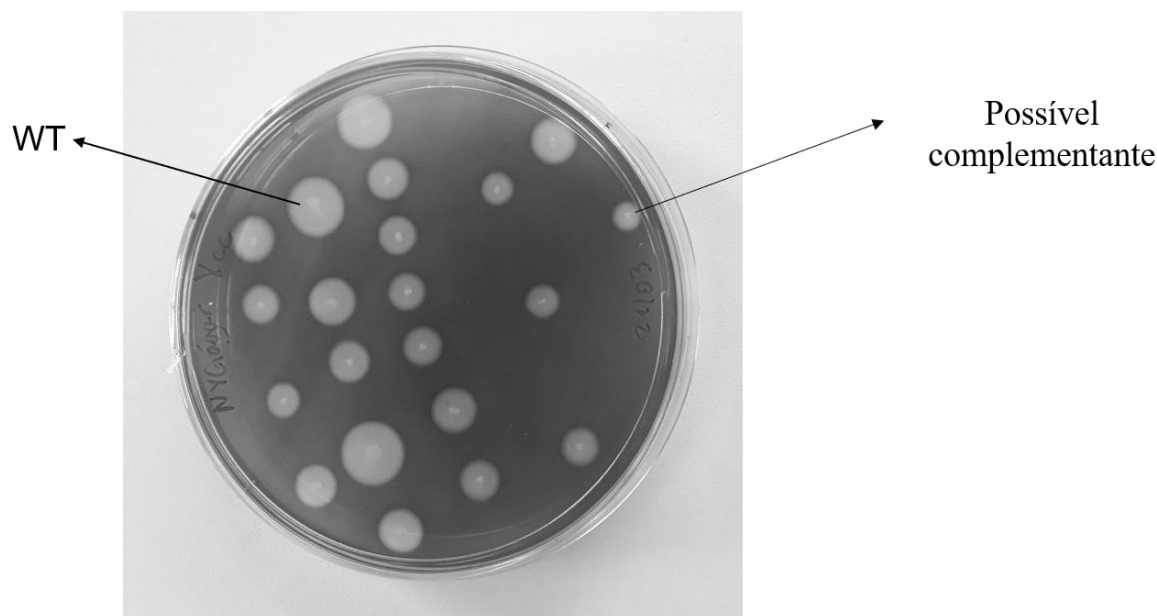
**Figura 3.18** Análise de restrição do vetor de complementação pLAL6\_xanB. DNA plasmidial do vetor de complementação pLAL6\_CxanB foi digerido com endonucleases utilizadas em sua construção e analisado por eletroforese em gel de TAE agarose 0,8%. **M** | Padrão de tamanho molecular EZ Load™ 1 kb Molecular Ruler (Bio-Rad). **1** | pLAL6\_xanB digerido com *EcoRI* e *XhoI*, excisando o fragmento referente à região predita como promotora do gene-alvo (390 pb). **2** | pLAL6\_xanB digerido com *EcoRI* e *XbaI*, excisando à construção integral, ou seja, região promotora *in tandem* com a região codificante do gene *xanB*, resultando em um fragmento de 1794 pb. Tamanhos de bandas do padrão estão indicados. Fonte: próprio autor.

### 3.5.2 Screening de possíveis complementantes

Após a confirmação da sua construção, o vetor pLAL\_xanB foi quantificado e utilizado para transformação de células de *XccΔxanB* eletrocompetentes por eletroporação, seguida do plaqueamento em ágar-LB contendo gentamicina e amido 1% (m/v).

Uma vez que o vetor foi projetado para permitir a sua integração no gene que codifica a  $\alpha$ -amilase, o sucesso da transformação implicaria na incapacidade do transformante em degradar amido, cuja avaliação foi feita pelo contato de vapor de iodo com as placas de transformantes.

Como o iodo ( $I_2$ ) possui a capacidade de interagir com as cadeias de amilose do amido formando um complexo de coloração azul intensa (FUWA, 1954), foi possível constatar que as colônias transformantes obtidas não tinham o vetor pLAL6\_xanB integrado no gene que codifica a  $\alpha$ -amilase, pois foi observado um halo de degradação de amido ao redor de tais colônias (Figura 3.19).



**Figura 3.19** *Screening* das colônias transformantes quanto à capacidade de degradar amido. Células de *XccΔxanB* eletrocompetentes foram eletrotransformadas com o vetor de complementação pLAL6\_CxanB, plaqueadas em ágar-LB contendo gentamicina, sendo repicadas individualmente em meio contendo gentamicina e amido 1% e, em seguida, realizada sublimação com iodo ( $I_2$ ). De acordo com o esperado, colônias de *Xcc* selvagem (WT) degradaram amido e apresentaram um halo resultante na ausência de iodo, ou seja, degradação de amido. Entretanto, as possíveis colônias complementantes também apresentaram o mesmo halo, indicando que a integração do vetor de complementação não ocorreu na região esperada do genoma e, conseqüentemente, a complementação não ocorreu como o planejado. Fonte: próprio autor.

É possível que a integração do vetor tenha ocorrido em outros locais do genoma bacteriano, uma vez que o pLAL6\_xanB apresenta, além da região codificante do gene *xanB*, a sua região promotora predita, a qual pode ter permitido uma recombinação homóloga com o cromossomo bacteriano, o que não foi testado.

Dessa forma, após dezenas de tentativas sem sucesso de transformação da linhagem mutante *XccΔxanB* com o pLAL6\_xanB, tal vetor foi desconsiderado para a construção da linhagem complementante.

### 3.6 Complementação gênica utilizando o vetor pNPTS138

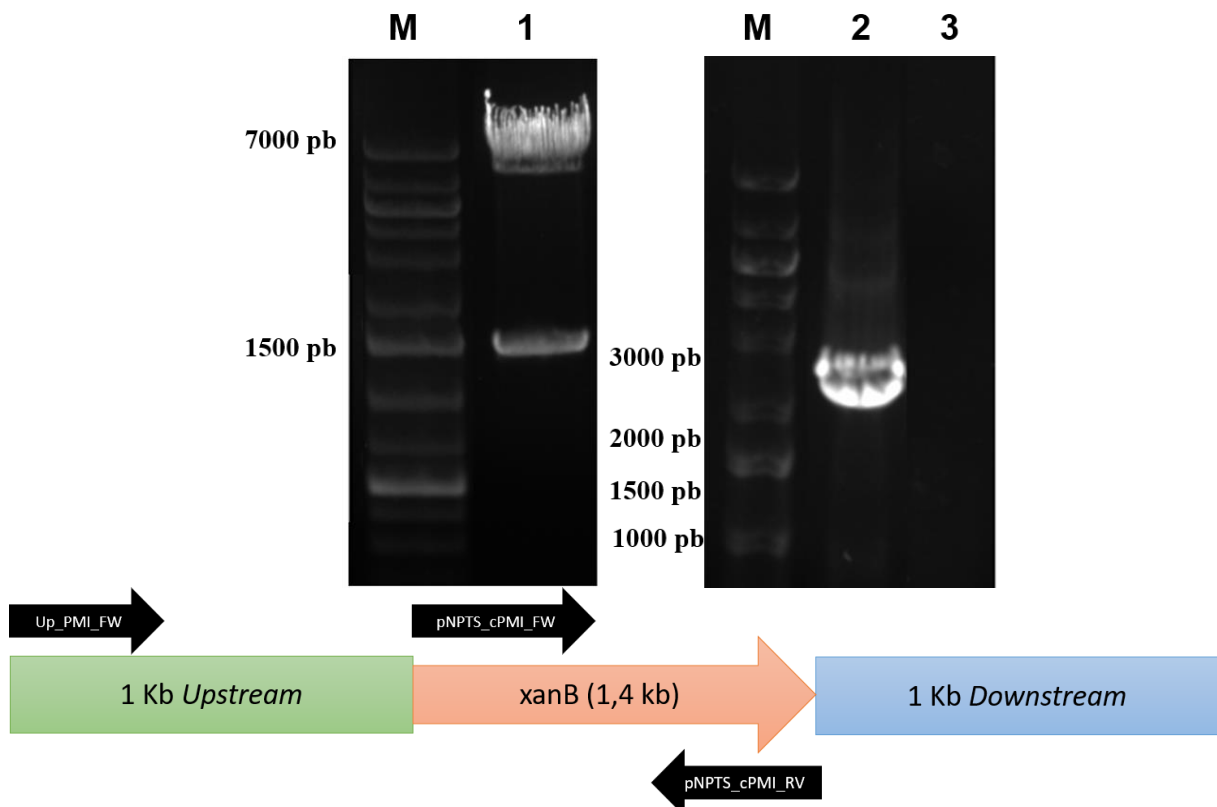
#### 3.6.1 Construção do vetor pNPTS\_CxanB

Para a construção do vetor de complementação do gene *xanB* (pNPTS138\_CxanB), a região codificante do gene *xanB* foi clonada no vetor de deleção pNPTS\_xanB no sítio de restrição para a enzima *EcoRI*. Dessa forma, a

região codificante foi inserida entre as regiões de 1 kb *upstream* e 1 kb *downstream* do vetor de deleção (Figura 3.20).

A clonagem da região codificante do gene *xanB* no vetor de deleção pNPTS\_*xanB* para construir o vetor de complementação pNPTS\_C*xanB* requereu uma confirmação da orientação de tal sequência no vetor, uma vez que o inserto apresentava o mesmo sítio de restrição em ambas as extremidades.

Para isso, foram utilizados dois pares de oligonucleotídeos Up\_PMI\_FW-pNPTS\_cPMI\_RV e Up\_PMI\_FW-pNPTS\_cPMI\_FW em PCR utilizando o vetor de complementação pNPTS\_C*xanB* como DNA molde, o que permitiu a confirmação da correta orientação da clonagem do gene *xanB* (Figura 3.20).

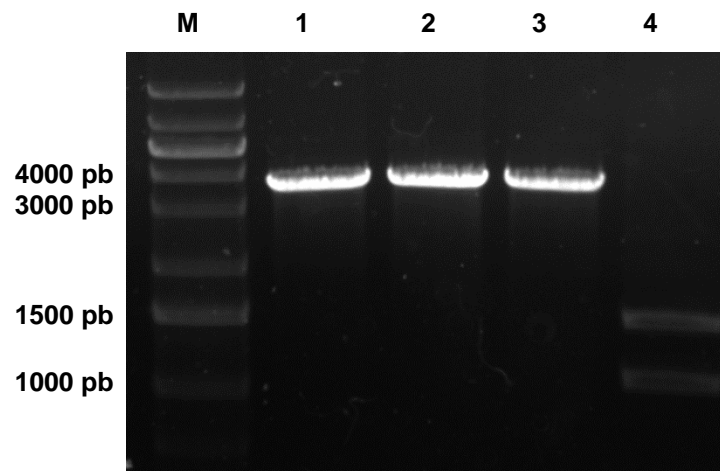


**Figura 3.20** Construção do vetor de complementação pNPTS\_C*xanB* e confirmação da orientação da clonagem do gene *xanB*. A confirmação da construção do vetor de complementação pNPTS\_C*xanB* foi realizada por digestão com endonuclease, seguida pela análise por eletroforese em gel de TAE agarose 0,8%. No painel superior: **M** | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas). **1** | pNPTS\_C*xanB* digerido com a enzima de restrição *EcoRI*. O padrão de digestão do vetor mostra que a montagem do vetor foi realizada com sucesso, uma vez que a região codificante apresenta 1404 pb, um tamanho compatível com o observado no gel. **2** | Produto de PCR utilizando o par de oligonucleotídeos Up\_PMI\_FW-pNPTS\_cPMI\_RV e o vetor pNPTS\_C*xanB* como molde. **3** | Produto de PCR utilizando o par de oligonucleotídeos Up\_PMI\_FW-pNPTS\_PMI\_FW e o vetor pNPTS\_C*xanB* como molde. A observação de uma banda na altura de 2,4 kb na coluna 2 e a inexistência de bandas na coluna 3 comprova que a orientação da clonagem da região codificante do gene *xanB* no vetor de complementação é a esperada. No painel inferior está representada a organização genômica da região do gene *xanB*, bem como os oligonucleotídeos utilizados na PCR de confirmação da orientação da clonagem. Fonte: próprio autor.

### 3.6.2 Confirmação da complementação gênica

A linhagem *Xcc*Δ*xanB* (mutante deletada no gene que codifica a PMI) foi transformada com o plasmídeo de complementação pNPTS\_C*xanB* por eletroporação e plaqueada em LB contendo canamicina. Após isso, foram seguidos os passos de seleção de duplos recombinantes entre o plasmídeo e o genoma, utilizando o mesmo método da deleção gênica.

Ao final, colônias foram testadas por PCR utilizando os oligonucleotídeos KO\_PMI\_FW e KO-PMI\_RV, os quais hibridizam nas regiões adjacentes às regiões flanqueadoras do gene-alvo para a confirmação da complementação gênica (Figura 3.21), sendo a linhagem complementante denominada *Xcc*ΔC*xanB*.



**Figura 3.21** Confirmação da complementação gênica do mutante *Xcc*Δ*xanB*. Análise dos produtos de PCR resultantes das reações utilizando oligonucleotídeos que hibridizam em regiões externas às regiões *upstream* e *downstream* de 1 kb ao gene *xanB*, com DNA genômico de *Xcc* selvagem e *Xcc*ΔC*xanB* como molde. **M** | Padrão de tamanho molecular *GeneRuler 1 kb Plus DNA Ladder* (Fermentas); **1** | Produto de PCR utilizando *Xcc* selvagem como molde. **2** | Produto de PCR utilizando *Xcc* selvagem como molde, digerido com a enzima *EcoRI*. **3** | Produto de PCR utilizando *Xcc*ΔC*xanB* como molde. **4** | Produto de PCR utilizando *Xcc*ΔC*xanB* como molde, digerido com a enzima *EcoRI*. A restrição por *EcoRI* do fragmento amplificado ocorreu para *Xcc*ΔC*xanB* e não para *Xcc*, conforme esperado, resultando em duas bandas (um fragmento de 1,4 kb e dois fragmentos de 1 kb), o que é compatível com o tamanho de *xanB* e das duas regiões flanqueadoras. As bandas relevantes do padrão de tamanho molecular estão indicadas. Fonte: próprio autor.

Esta estratégia permitiu a diferenciação das colônias selvagens das complementantes pois, apesar de os produtos de PCR de ambas as linhagens



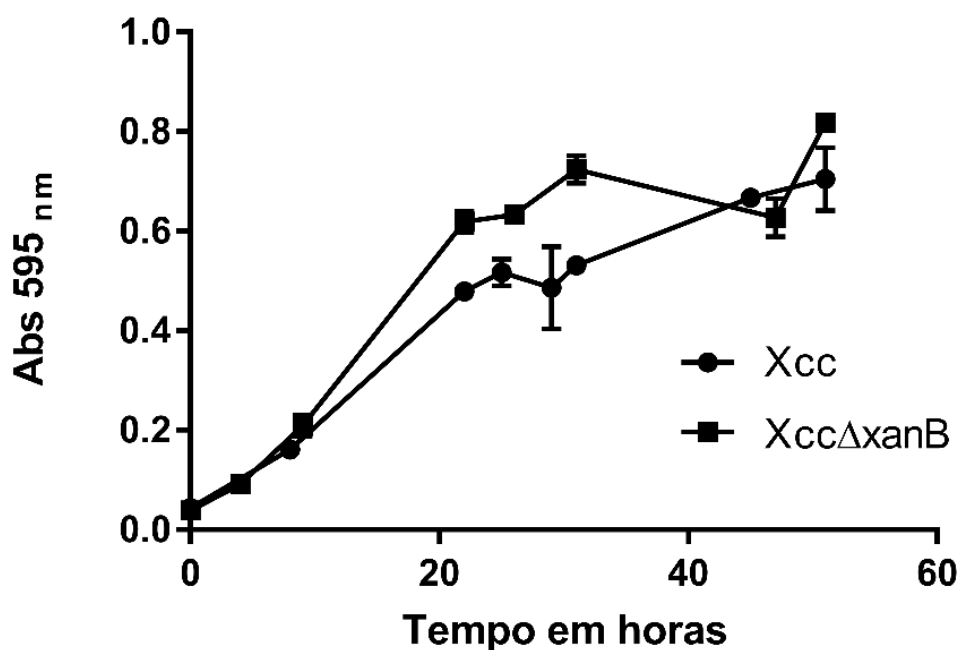
apresentarem 3,4 kb de tamanho, apenas o produto de PCR da linhagem complementante *Xcc*Δ*xanB* pode ser digerido com a enzima *EcoRI*, resultando em três fragmentos: a região codificante do gene *xanB*, cujo tamanho é 1,4 kb e as regiões flanqueadoras ao gene-alvo, as quais apresentam 1 kb de tamanho cada.

Dessa forma, foi possível distinguir as colônias selvagens das colônias complementantes, confirmando a complementação gênica de uma linhagem de *Xcc* deletada utilizando o vetor pNPTS138, uma estratégia inédita na literatura para esta espécie.

### 3.7 Caracterização fenotípica das linhagens mutantes e complementante

#### 3.7.1 Deleção de *xanB* não afeta o perfil de crescimento

As linhagens *Xcc* e *Xcc*Δ*xanB* foram cultivadas em meio XAM-M e tiveram suas densidades ópticas em 595 nm mensuradas para a comparação dos perfis de crescimento (Figura 3.22).



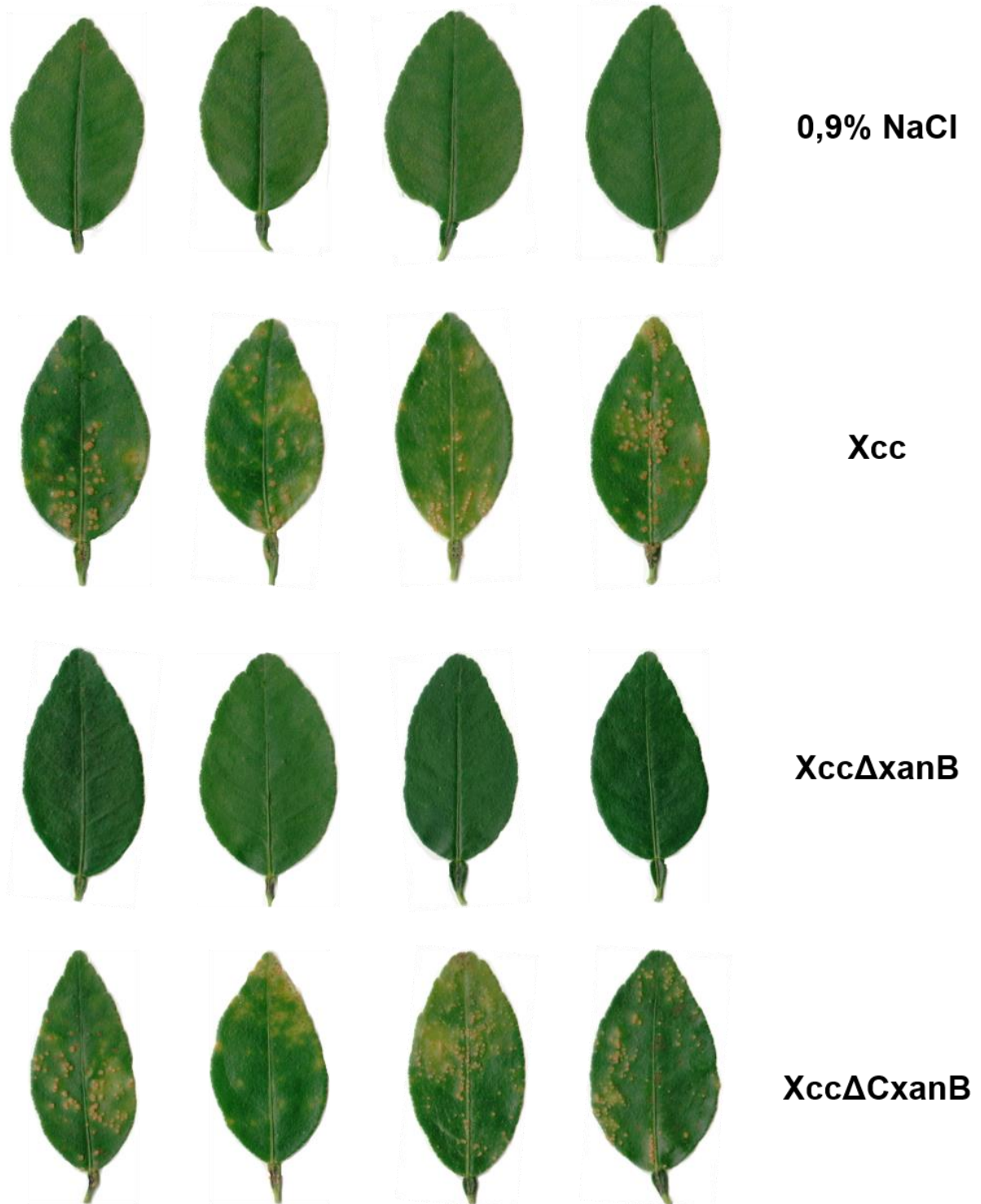
**Figura 3.22** Curva de crescimento das linhagens *Xcc* e *Xcc*Δ*xanB*. As linhagens *Xcc* e *Xcc*Δ*xanB* foram avaliadas e comparadas quanto aos perfis de suas curvas de crescimento em meio XAM-M, indutor de patogenicidade. Barras de erro indicam o desvio-padrão de triplicatas biológicas. Fonte: próprio autor.

A análise das curvas de crescimento de Xcc e Xcc $\Delta$ xanB permite concluir que a deleção do gene *xanB* não comprometeu a capacidade de crescimento da linhagem mutante e, além disso, que o padrão de crescimento foi semelhante entre as duas linhagens.

### **3.7.2 *xanB* é essencial para a patogenicidade de Xcc**

As linhagens Xcc, Xcc $\Delta$ xanB e Xcc $\Delta$ CxanB foram avaliadas quanto à patogenicidade e virulência em plantas de *Citrus aurantifolia* por meio de duas metodologias distintas – aspersão e infiltração. Para ambos os experimentos, as linhagens foram cultivadas em caldo LB e ressuspendidas em solução salina (0,9% de NaCl) de modo que a concentração final ficasse em 10<sup>6</sup> UFC.ml<sup>-1</sup>.

No ensaio de aspersão, 10 ml de cada suspensão bacteriana e de solução salina, para o controle positivo, foram pulverizados com o uso de um borrifador doméstico estéril em quadruplicata de plantas. As folhas foram registradas digitalmente após 28 dias da aspersão (Figura 3.23).



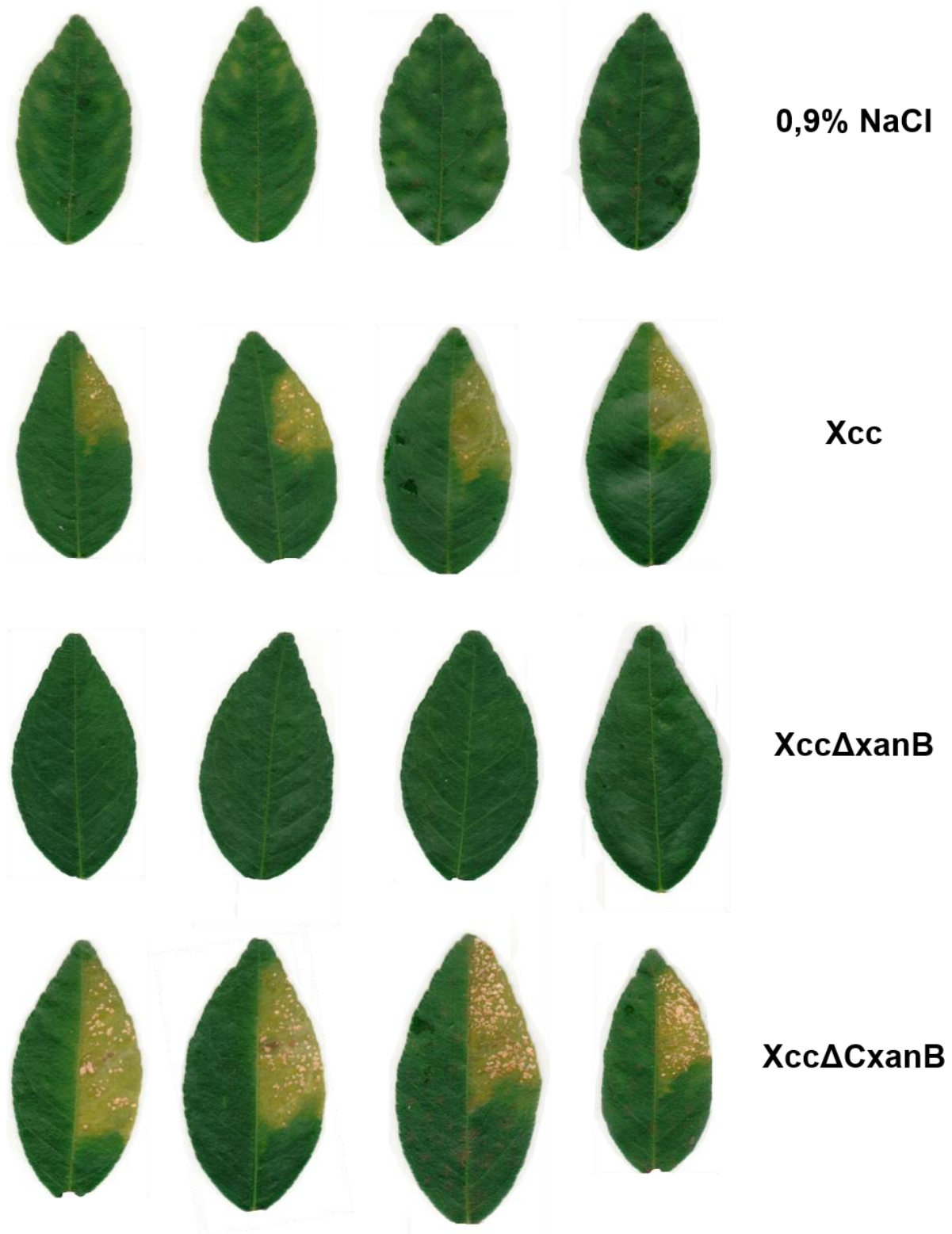
**Figura 3.23** Ensaio de patogenicidade *in vivo* por aspersão de Xcc, XccΔxanB e XccΔCxanB em *Citrus aurantifolia*. Plantas de *Citrus aurantifolia* foram utilizadas para avaliação comparativa da patogenicidade e virulência de Xcc, XccΔxanB e XccΔCxanB. Foram pulverizados 10 ml de cada cultivo bacteriano e o mesmo volume de solução salina (controle negativo) em quadruplicatas de plantas. Após serem destacadas das plantas, as folhas foram fotografadas no vigésimo oitavo dia após a aspersão. Fonte: próprio autor.

Para o ensaio de infiltração, as linhagens ou a solução salina (controle negativo) foram infiltradas na superfície abaxial de folhas de quatro plantas, de forma que todas as plantas receberam os quatro tratamentos diferentes, de forma a minimizar o efeito do hospedeiro sobre os resultados, quando da comparação entre os fenótipos apresentados. Após vinte dias das infiltrações, as folhas foram destacadas e registradas digitalmente para comparação da infecção (Figura 3.24).

Os ensaios de aspersão e infiltração apresentaram um padrão fenotípico em relação aos sintomas de cancro cítrico observados, a partir do qual foi possível constatar que a deleção do gene *xanB* fez com que a bactéria perdesse a patogenicidade. Além disso, a reversão do fenótipo observada nos ensaios com a linhagem complementante *XccΔC<sub>xanB</sub>* corrobora a hipótese de que o gene *xanB* é essencial para a patogenicidade de *Xcc*.

Comparando-se os fenótipos apresentados pelas linhagens selvagem (*Xcc*) e complementante (*XccΔC<sub>xanB</sub>*), observou-se semelhança nos graus de encharcamento e necrose do tecido vegetal, além da mesma frequência e intensidade das pústulas acastanhadas, sintomas característicos do cancro cítrico.

Embora sejam necessários estudos posteriores para a elucidação dos efeitos da deleção na expressão de fatores de virulência, a presente constatação de que o gene *xanB* é essencial para a patogenicidade de *Xcc* possibilita a futura prospecção de ferramentas biotecnológicas de controle e cura para doença, por meio da utilização de inibidores da PMI e, também, do desenvolvimento de cultivares cítricas que neutralizem a ação de tal proteína no processo infeccioso.



**Figura 3.24** Ensaio de patogenicidade *in vivo* por infiltração de Xcc, XccΔxanB e XccΔCxanB em *Citrus aurantifolia*. Plantas de *Citrus aurantifolia* foram utilizadas para avaliação comparativa da patogenicidade e virulência de Xcc, XccΔxanB e XccΔCxanB. Para cada uma das condições foram infiltradas quatro folhas de ramos independentes, atentando-se para que todas as plantas tivessem folhas infiltradas com as três linhagens bacterianas e o controle negativo com solução salina (0,9% de NaCl). As fotografias foram obtidas no vigésimo dia após a infiltração. Fonte: próprio autor.

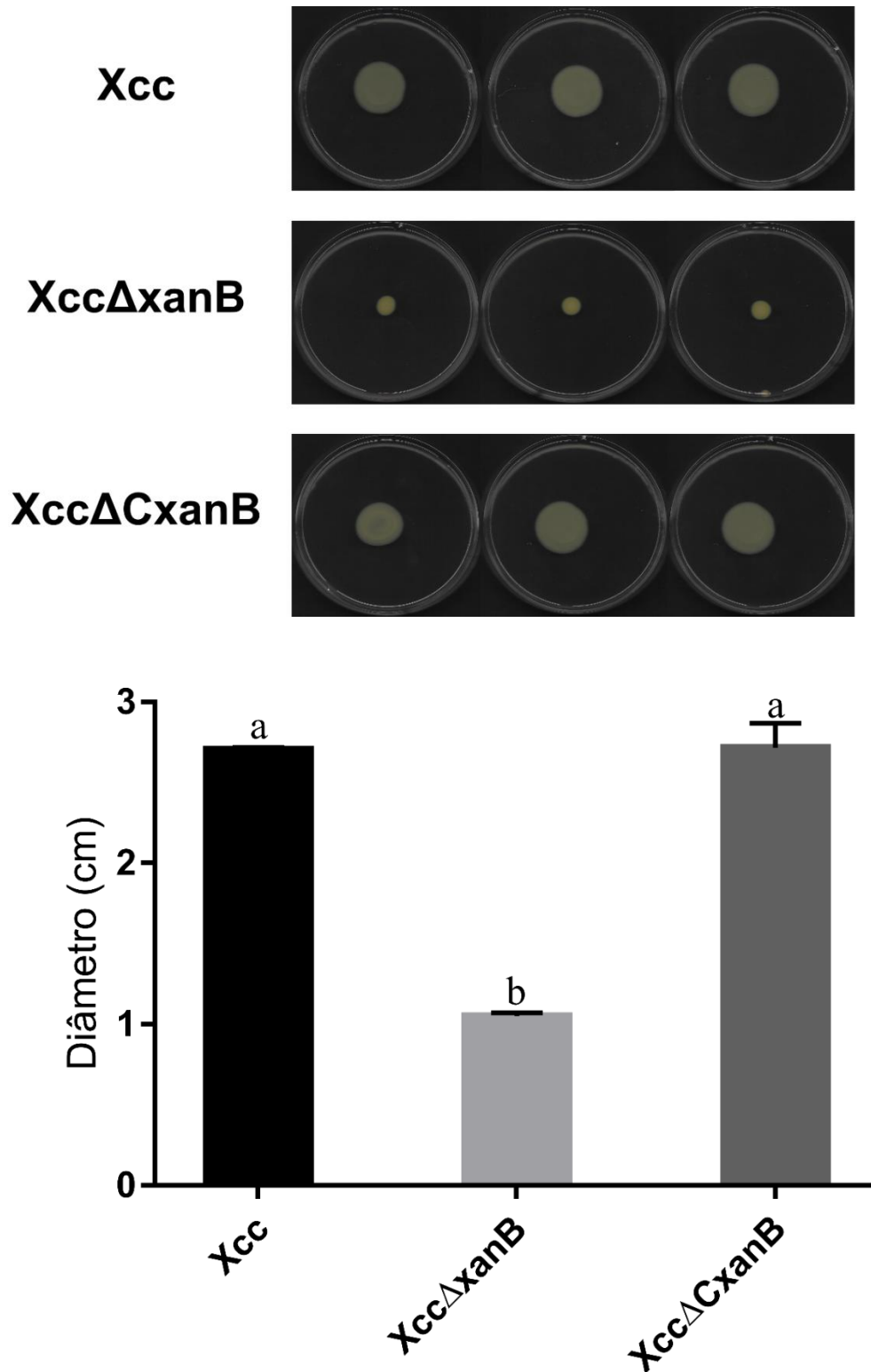
### 3.7.3 *xanB* está envolvido na motilidade, formação de biofilme e resistência à radiação UV

A capacidade de motilidade foi avaliada nas linhagens Xcc, XccΔ*xanB* e XccΔC*xanB* (Figura 3.25), segundo metodologia padronizada por (MALAMUD et al., 2011). Resumidamente, as linhagens foram cultivadas até que atingissem a DO<sub>595 nm</sub> igual a 0,3 e 3 µl destes cultivos foram pipetados no centro das placas de Petri com LB 0,7%.

O diâmetro das colônias de Xcc, XccΔ*xanB* e XccΔC*xanB* foi de  $2,703 \pm 0,016$ ,  $1,05 \pm 0,02$  e  $2,716 \pm 0,093$  cm, respectivamente. Assim, em média a linhagem XccΔ*xanB* apresentou motilidade 61,15% e 61,34% inferior às linhagem selvagem Xcc e complementante XccΔC*xanB*, respectivamente, indicando que o gene *xanB* está relacionado a esta importante característica do processo infeccioso.

Em *Stenotrophomonas maltophilia*, a mutação do gene *xanB* utilizando um *transposon* comprometeu a motilidade bacteriana, o que pode ser atribuído à constatação da ausência do flagelo polar nas células mutantes (HUANG; SOMERS; WONG, 2006).

A redução da motilidade pode ter contribuído para a perda de patogenicidade de XccΔ*xanB* (Figuras 3.23 e 3.24). Os fitopatógenos utilizam mecanismos de motilidade para atingirem superfícies ricas em nutrientes e os tecidos vegetais. Especificamente para Xcc, a motilidade está relacionada ao seu único flagelo polar, que a permite deslizar no ambiente epifítico e, também no apoplasto (GRAHAM et al., 2004; HARSHEY, 2003).



**Figura 3.25** Ensaio de motilidade. As linhagens Xcc, XccΔxanB e XccΔCxanB foram avaliadas quanto à motilidade segundo metodologia padronizada por Malamud e colaboradores (2011). As linhagens foram cultivadas até que atingissem  $DO_{595\text{ nm}}$  igual a 0,3. Três  $\mu\text{l}$  de cada cultivo foram pipetados no centro das placas de Petri de 5,1 cm de diâmetro com ágar-LB 0,7%. Após isso, as placas foram incubadas a 30 °C por 48 horas sem agitação e registradas digitalmente (Painel superior). O diâmetro das colônias obtidas foi mensurado utilizando o *software* ImageJ e a média das medidas está apresentada no painel inferior. Barras de erro indicam o desvio-padrão absoluto de cada uma das triplicatas. A análise estatística dos dados foi realizada pelo teste de Tukey, com nível de confiança de 95% por meio do *software* Minitab (15.2). Fonte: próprio autor.

O envolvimento da PMI na síntese de GDP-D-manose, um precursor nas vias de produção de lipopolissacarídeos (LPS), também pode contribuir no entendimento da diminuição da motilidade de XccΔxanB. Em Xcc, a disrupção do gene *wxacO*, participante da síntese de LPS, afetou negativamente a motilidade do fitopatógeno (LI; WANG, 2011a).

Assim, embora seja necessária a comprovação de que XccΔxanB apresenta defectividade na produção de LPS e na formação do flagelo polar, é possível sugerir que a redução de motilidade tenha relação com esses dois componentes atuantes na patogênese do cancro cítrico.

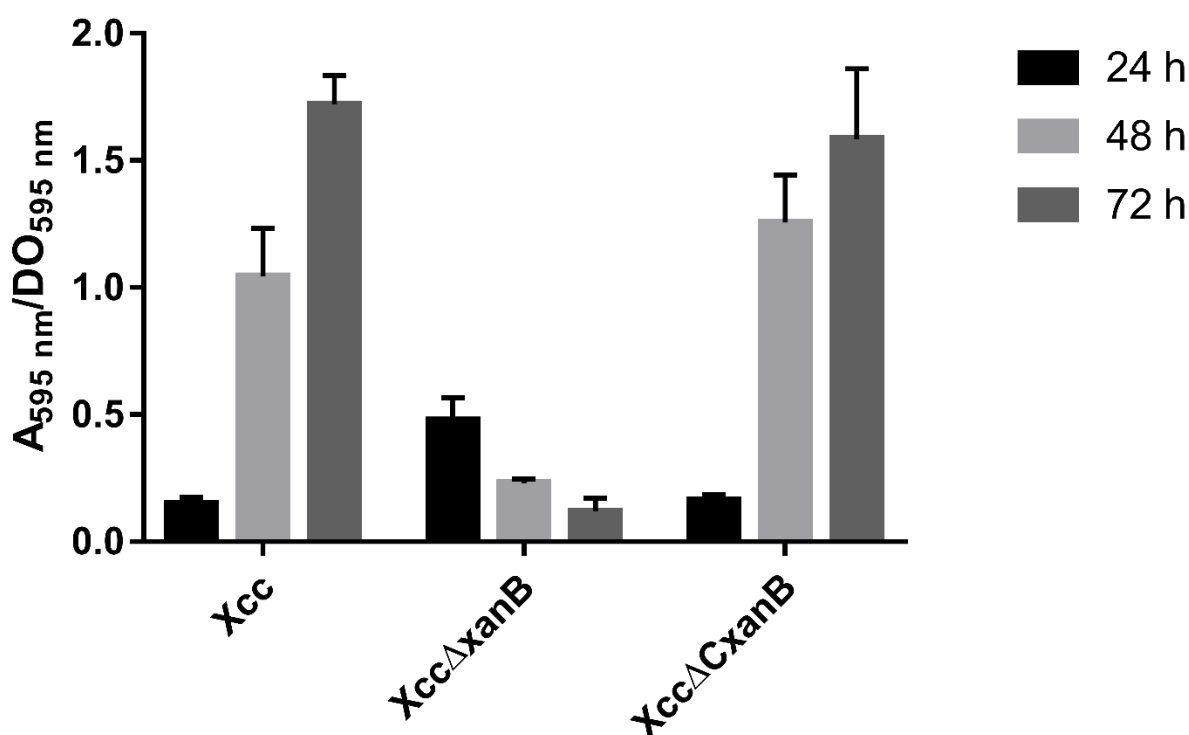
A formação de biofilme é uma capacidade que boa parte das bactérias tem de formar agregados em matrizes que permitem adesão a superfícies, possibilitando um modo de vida dinâmico às populações bacterianas (FLEMMING; WINGENDER, 2010).

As linhagens Xcc, XccΔxanB e XccΔCxanB foram avaliadas quanto à formação de biofilme ao longo de 24, 48 e 72 horas, segundo metodologia padronizada por O'Toole e Kolter (1998). Após 24 horas, a linhagem XccΔxanB apresentou formação de biofilme consideravelmente superior às linhagens selvagem Xcc e complementante XccΔCxanB. Em *Photorhabdus luminescens*, a PMI foi relacionada somente ao estágio de maturação do biofilme, e não ao estágio inicial do seu desenvolvimento (AMOS et al., 2011).

Em 48 e 72 horas houve uma inversão significativa desse padrão, com a linhagem XccΔxanB apresentando formação de biofilme 4,5 e 14,5 vezes menor (48 e 72 horas, respectivamente) em relação à linhagem selvagem Xcc (Figura 3.26).

Importante ressaltar que o comportamento da linhagem complementante XccΔCxanB foi extremamente semelhante ao de Xcc, indicando que os efeitos depletivos da formação de biofilme em XccΔxanB após 48 e 72 horas parecem ser, de fato, decorrentes da deleção do gene *xanB*.





**Figura 3.26** Ensaio de formação de biofilme. As linhagens Xcc, XccΔxanB e XccΔCxanB foram avaliadas quanto à formação de biofilme segundo metodologia padronizada por O'Toole e Kolter (1998). Os resultados estão apresentados como a média da razão entre a absorvância de cristal violeta em 595 nm e a densidade óptica de cada cultivo bacteriano imediatamente antes da mensuração da formação de biofilme ( $A_{595 \text{ nm}}/DO_{595 \text{ nm}}$ ). As medidas foram auferidas em 24, 48 e 72 horas. Barras de erro indicam o desvio-padrão absoluto de cada uma das sextuplicatas. Fonte: próprio autor.

Li e Wang (2011b) utilizaram uma abordagem de mutagênese em larga escala por meio de *transposons* e identificaram 92 genes de Xcc relacionados à formação de biofilme, dentre os quais está incluído *xanB*, o que reforça os dados aqui obtidos utilizando uma metodologia diferente para a mutação do gene-alvo.

O processo de formação e maturação dos biofilmes tem a participação de EPS e LPS, substâncias poliméricas que atuam na coesão das células e, também, na comunicação intercelular (DONLAN, 2002).

Especificamente para Xcc, estudos recentes têm demonstrado que a motilidade e a adesão são características necessárias aos estágios iniciais da formação de biofilme, enquanto maturação do biofilme depende de LPS e EPS, como a goma xantana (LI; WANG, 2011a, 2011b). Portanto, a participação do gene *xanB* nas vias de biossíntese de LPS e EPS também pode explicar, em parte, o fenótipo apresentado pela linhagem mutante deletada no referido gene.

Ademais, a formação de biofilme atua como uma estratégia para a sobrevivência epifítica de Xcc (RIGANO et al., 2007), e a redução na formação de

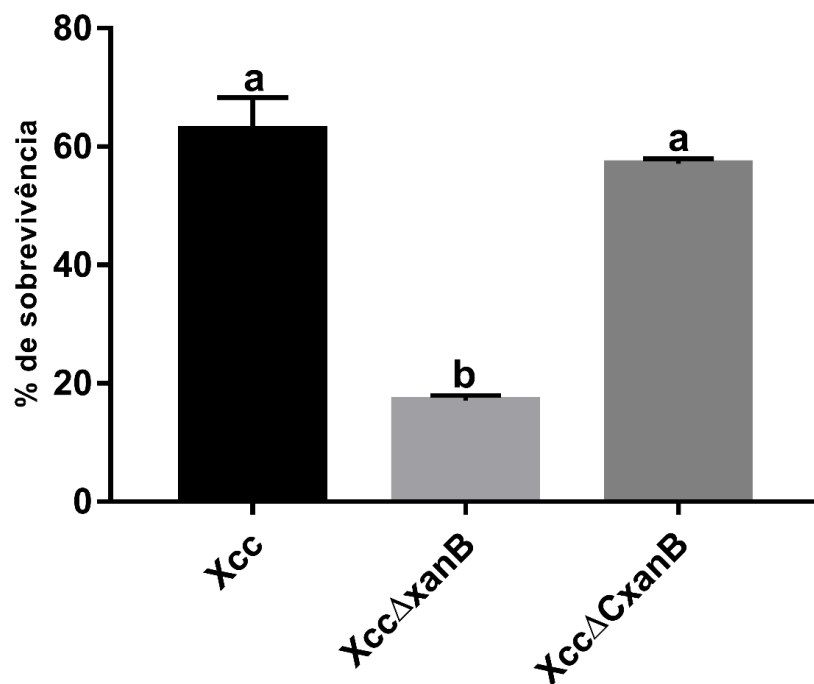
biofilme apresentada por  $Xcc\Delta xanB$  pode explicar a perda de patogenicidade desta linhagem após a deleção do gene *xanB*.

Os biofilmes atuam efetivamente na proteção contra estresses ambientais, como a resistência à radiação ultravioleta, e também na defesa contra os mecanismos responsivos do hospedeiro (KARATAN; WATNICK, 2009).

Complementarmente, as linhagens  $Xcc$ ,  $Xcc\Delta xanB$  e  $Xcc\Delta CxanB$  foram avaliadas quanto à resistência à radiação ultravioleta (UV). Para isso, as linhagens foram cultivadas até que atingissem a mesma densidade celular e foram expostas à luz ultravioleta por 15 minutos a 60 cm da fonte de luz. A porcentagem de sobrevivência de cada linhagem foi calculada em triplicata (Figura 3.27).

A linhagem  $Xcc\Delta xanB$  apresentou uma sobrevivência três vezes menor do que a linhagem selvagem  $Xcc$ , um fenótipo que foi restaurado pela complementação gênica.

A maior suscetibilidade de  $Xcc\Delta xanB$  à radiação ultravioleta pode ser explicada pela atuação da PMI na via de biossíntese de xantana, uma vez que esse EPS tem sido relacionado à resistência a diversos estresses ambientais, incluindo a exposição à radiação UV (MEYER; BOGDANOVE, 2009).



**Figura 3.27** Ensaio de resistência à radiação ultravioleta. As linhagens  $Xcc$ ,  $Xcc\Delta xanB$  e  $Xcc\Delta CxanB$  foram avaliadas quanto à sobrevivência após exposição à radiação ultravioleta por 15 minutos a uma distância de 60 cm da fonte de luz. Os resultados estão apresentados em porcentagem de sobrevivência (% de sobrevivência), a qual foi calculada a partir dos controles de cada linhagem bacteriana não exposta à radiação ultravioleta. A análise estatística dos dados foi realizada pelo teste de Tukey, com nível de confiança de 95% por meio do *software* Minitab (15.2). Fonte: próprio autor.

### 3.7.4 A deleção do gene *xyIA2* aumenta a virulência de *Xcc*

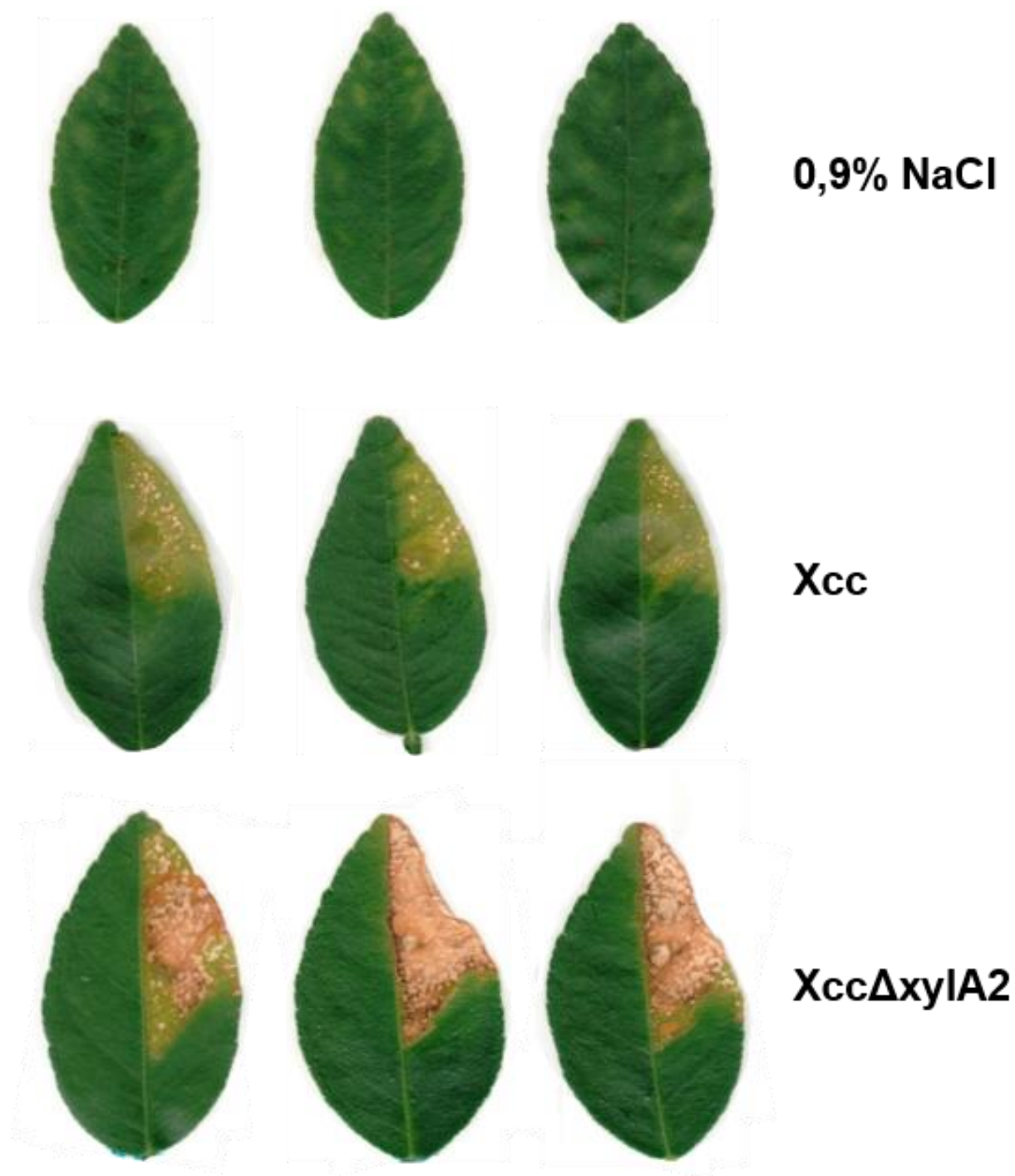
A linhagem mutante *XccΔxyIA2*, juntamente com *Xcc*, foi avaliada quanto à patogenicidade e virulência em *Citrus aurantifolia*. Resumidamente, as linhagens foram cultivadas em caldo LB e infiltradas na superfície abaxial das folhas na concentração de  $10^6$  UFC.ml<sup>-1</sup>, utilizando quatro plantas, nas quais todos os tratamentos foram testados, a fim de minimizar a variação biológica da resposta de defesa do hospedeiro. Vinte dias após a infiltração, as folhas foram destacadas das plantas e registradas digitalmente para comparação da infecção (Figura 3.28).

Para permitir uma avaliação minuciosa e integral do envolvimento da xilose isomerase no processo infeccioso de *Xcc*, foram avaliadas outras duas linhagens mutantes adicionais, uma mutante deletada na ORF XAC1776 (*XccΔxyIA1*) e outra deletada nas duas ORF codificantes de xilose isomerase (*XccΔxyIA1xyIA2*), ou seja, um duplo mutante (PRIETO, 2019, dados não publicados).

O grau de intensidade dos sintomas apresentados pelas folhas infectadas pelas diferentes linhagens testadas foi visualmente diferente. Como esperado, a infiltração com solução salina (NaCl a 0,9%) não resultou no desenvolvimento de sintomas, mostrando que as plantas utilizadas não estavam infectadas anteriormente (Figura 3.28). Para a linhagem selvagem *Xcc* (Figura 3.28), os sintomas foram característicos do cancro cítrico, com alto grau de encharcamento e clorose, presença de pústulas (principalmente na superfície abaxial) e algumas áreas de tecido necrótico (GOTTWALD; GRAHAM; SCHUBERT, 2002).

Nas folhas infiltradas com *XccΔxyIA1* (PRIETO, 2019, comunicação pessoal), houve um desenvolvimento acentuado dos sintomas, principalmente devido ao aparecimento de regiões mais necróticas quando comparadas ao tipo selvagem, sugerindo que essa deleção implica em aumento da virulência de *Xcc*.

Tanto o *XccΔxyIA2* (Figura 3.28) quanto o *XccΔxyIA1xyIA2* (duplo mutante, resultados não mostrados) apresentaram sintomas semelhantes, que foram consideravelmente mais intensos quando comparados à linhagem selvagem e também à *XccΔxyIA1*. Quando as folhas foram infiltradas com essas linhagens, as porções dos tecidos necróticos das plantas foram mais abundantes, indicando que o metabolismo da xilose está relacionado à virulência de *Xcc* e que há uma variação no envolvimento de cada gene que codifica a xilose isomerase nesse processo.



**Figura 3.28** Ensaio de patogenicidade *in vivo* por infiltração de Xcc e XccΔxylA2 em *Citrus aurantifolia*. Plantas de *Citrus aurantifolia* foram utilizadas para avaliação comparativa da patogenicidade e virulência de Xcc e XccΔxylA2. Para cada uma das condições foram infiltradas quatro folhas de ramos independentes, atentando-se para que todas as plantas tivessem folhas infiltradas com as duas linhagens bacterianas e o controle negativo com solução salina (0,9% de NaCl). As fotografias foram obtidas no vigésimo dia após a infiltração. Fonte: próprio autor.

O processo infeccioso de *Xanthomonas* spp. depende de uma conjuntura de condições que incluem adesão bacteriana (DUNGER et al., 2014), adaptação e sobrevivência ao estresse inerente à filosfera (DARSONVAL et al., 2008), colonização

do apoplasto (FATIMA; SENTHIL-KUMAR, 2015) e degradação da parede celular vegetal (CHOW et al., 2015).

Neste último processo, a bactéria utiliza enzimas de degradação da parede celular (*cell wall degradation enzymes*, CWDEs), as quais atuam na desconstrução dos componentes desta barreira física das células vegetais por meio de enzimas líticas (BÜTTNER; BONAS, 2010). Algumas dessas enzimas atuam na disrupção da estrutura do xilano, uma hemicelulose que compõe a maior parte da parede celular das plantas, usando principalmente sistemas xilanolíticos que incluem xilanases,  $\beta$ -xilosidases, arabinofuranosidases, carboidrato esterases e  $\alpha$ -glicuronidases (CHOW et al., 2015; DÉJEAN et al., 2013).

Na última década, o papel de enzimas envolvidas na degradação de xilano e na utilização de seus produtos tem sido extensamente avaliada em várias espécies do gênero *Xanthomonas*, levando a um entendimento de que tais sistemas enzimáticos estão relacionados à patogenicidade e virulência dessas bactérias (SANTOS et al., 2014).

Após a liberação dos monômeros de xilose pelos sistemas xilanolíticos, as enzimas xilose isomerase e xiluloquinase atuam na conversão de xilose em xilulose e desta em xilulose-5-fosfato, respectivamente, permitindo a integração desta via metabólica com a via das pentoses fosfato (DÉJEAN et al., 2013).

Neste trabalho, a caracterização funcional de um dos genes codificantes da xilose isomerase contribuiu para o estabelecimento da relação desta enzima na patogênese do cancro cítrico, sugerindo que a xilose pode atuar como um fator promotor da virulência de Xcc.

O genoma de Xcc possui duas ORFs preditas como codificantes de xilose isomerase, anotadas como XAC1776 (*xyIA1*) e XAC4225 (*xyIA2*). Embora as preditas sequências peptídicas resultantes de tais genes sejam 99% idênticas (alinhamento não mostrado), os contextos genômicos e regulatórios de *xyIA1* e *xyIA2* são efetivamente distintos (Figura 1.8). Enquanto *xyIA1* está contido em um *operon*, *xyIA2* apresenta-se como uma região codificante de um RNAm monocistrônico, porém com uma hipotética e regulatória região chamada *xyl-box*. Em *Xanthomonas campestris* pv. *campestris* (DÉJEAN et al., 2013), *Xanthomonas oryzae* pv. *oryzae* (IKAWA et al., 2018), *Caulobacter crescentus* CB15 (HOTTESS et al., 2004; STEPHENS et al., 2007) e *Lactobacillus pentosus* (LOKMAN et al., 1997), *xyl-box* é uma região regulatória à

qual se liga uma proteína repressora denominada XylR, codificada pela ORF XAC4226 em Xcc.

Em *X. campestris* pv. *campestris* e *X. oryzae* pv. *oryzae*, a presença de xilose inativa a proteína XylR e desreprime a transcrição dos genes regulados por *xyl-box*, a maioria dos quais é codificante de enzimas envolvidas na degradação de xilano e na utilização de xilose. Também, para *X. oryzae* pv. *oryzae*, a interação entre XylR e xilose resultou em um acúmulo de HrpX, um regulador positivo da expressão de genes *hrp* (DÉJEAN et al., 2013; IKAWA et al., 2018) e de genes do Sistema de Secreção Tipo III (T3SS) em *X. fuscans* subsp. *fuscans* (DARSONVAL et al., 2008).

Então, o mecanismo regulatório de XylR pode fornecer uma explicação para os fenótipos apresentados pela linhagem mutante aqui obtida (XccΔxylA2) e, também, para as linhagens mutantes obtidas por Prieto (2019, dados não publicados). Uma vez que não há rotas metabólicas alternativas para a degradação de xilose em Xcc, a deleção dos genes codificantes da xilose isomerase poderia resultar em aumento da concentração de xilose no ambiente intracelular, permitindo a inativação de XylR, o que levaria a um consequente aumento da expressão de genes envolvidos na degradação de xilano e, também, de genes *hrp*, estes últimos induzidos pelo possível acúmulo de HrpX ocasionado por XylR quando associada à xilose.

O ensaio *in vivo* com *Citrus aurantifolia* revelou que a linhagem XccΔxylA1 é mais virulenta do que Xcc, mas ainda mais branda que XccΔxylA2 e XccΔxylA1xylA2 (linhagem duplo mutante). Neste caso, pode-se inferir que a deleção da ORF XAC1776 (*xylA1*) restringiu o uso de xilose pela xilose isomerase expressa a partir do gene *xylA2*, cuja regulação é dependente de XylR.

Para as infecções com XccΔxylA2 e XccΔxylA1xylA2, os sintomas apresentados foram similares entre as duas linhagens e notadamente mais intensos que XccΔxylA1 e Xcc, indicando que a possível regulação por XylR, que atua apenas em *xylA2*, parece ser determinante na conversão de xilose em xilulose pela xilose isomerase, considerando que seu papel em Xcc seja análogo ao de *X. campestris* e *X. oryzae*. Assim, embora seja necessária uma confirmação adicional de que XylA2 tenha um papel metabólico predominante sobre XylA1, o aumento da virulência causada pela deleção de *xylA2* (XccΔxylA2 e XccΔxylA1xylA2) sugere um papel diferencial dos dois genes codificantes da xilose na patogênese do cancro cítrico causado por Xcc.

Assim, é possível sugerir que a alta disponibilidade de xilose em *XccΔxylA2* e *XccΔxylA1xylA2* devido ao papel principal de *XylA2* intensifica a desrepressão de *XylR*. Portanto, a modulação diferencial da virulência poderia ser explicada pelo aumento hipotético e gradual da disponibilidade de xilose nas linhagens mutantes, o que diminuiria o papel do *XylR* como repressor dos genes regulados por *xyl*-box e aumentaria sua atividade como ativador da expressão dos genes *hrp*.

## Capítulo 4 – Conclusões

- ✓ As ORFs XAC3580 e XAC4225 codificam para uma fosfomanose isomerase e uma xilose isomerase, respectivamente;
- ✓ O gene *xanB* é essencial para a patogenicidade *in vivo* de Xcc no hospedeiro *Citrus aurantifolia*, estando inclusive envolvido na formação de biofilme, motilidade e resistência à radiação ultravioleta, aspectos que podem explicar a sua relação com a patogenicidade.
- ✓ A deleção do gene *xyIA2* acentuou a virulência de Xcc, o que pode ser explicado pelo fato de xilose não atuar apenas como um substrato energético metabolizável por Xcc, mas também como uma molécula sinalizadora, cujo acúmulo poderia levar ao aumento da expressão de genes relacionados à patogenicidade, seja pela ligação desse açúcar a proteínas repressoras do operon *hrp*, ou pela regulação positiva da expressão de genes codificantes de proteínas do sistema xilanolítico.
- ✓ *xanB* codifica uma proteína candidata para o desenho de drogas destinadas ao controle do cancro cítrico, inclusive por estar possivelmente acessível na superfície bacteriana; por outro lado, *xyIA2* codifica uma proteína que não pode ser utilizada para tal finalidade.
- ✓ Embora a infecção causada por Xcc esteja relacionada a inúmeros genes e proteínas, inclusive do sistema de secreção tipo III e a mecanismos bastante complexos, a ausência de um único gene, no caso *xanB*, foi capaz de comprometer toda a capacidade infectiva de Xcc no hospedeiro cítrico, o que possivelmente pode indicar um envolvimento desse gene nas fases iniciais da infecção.



## Referências

- ALEXANDRINO, A. V.; GOTO, L. S.; NOVO-MANSUR, M. T. M. TreA codifies for a trehalase with involvement in *Xanthomonas citri* subsp. *citri* Pathogenicity. **PLoS ONE**, v. 11, n. 9, 2016.
- AMARAL, A. M. Cancro cítrico: permanente preocupação da citricultura no Brasil e no mundo. **Embrapa Recursos Genéticos e Biotecnologia-Comunicado Técnico (INFOTECA-E)**, 2003.
- AMOS, M. R. et al. Influence of the *Photobacterium luminescens* phosphomannose isomerase gene, *manA*, on mannose utilization, exopolysaccharide structure, and biofilm formation. **Appl. Environ. Microbiol.**, v. 77, n. 3, p. 776–785, 2011.
- ARNAU, J. et al. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. **Protein expression and purification**, v. 48, n. 1, p. 1–13, 2006.
- ARTIER, J. et al. Comparative proteomic analysis of *Xanthomonas citri* ssp. *citri* periplasmic proteins reveals changes in cellular envelope metabolism during in vitro pathogenicity induction. **Molecular plant pathology**, v. 19, n. 1, p. 143–157, 2018.
- ASTUA-MONGE, G. et al. Expression profiling of virulence and pathogenicity genes of *Xanthomonas axonopodis* pv. *citri*. **Journal of bacteriology**, v. 187, n. 3, p. 1201–1205, 2005.
- AUSUBEL, F. M. **Short protocols in molecular biology: a compendium of methods from Current protocols in molecular biology**. 5th ed. ed. New York: Wiley, 2002.
- BARBOSA-MENDES, J. M. et al. Genetic transformation of *Citrus sinensis* cv. Hamlin with *hrpN* gene from *Erwinia amylovora* and evaluation of the transgenic lines for resistance to citrus canker. **Scientia Horticulturae**, v. 122, n. 1, p. 109–115, 2009.
- BARCELOS, M. P. et al. Prediction of the Three-Dimensional Structure of Phosphate-6-mannose PMI Present in the Cell Membrane of *Xanthomonas citri* subsp. *citri* of Interest for the Citrus Canker Control. In: **Emerging Research in Science and Engineering Based on Advanced Experimental and Computational Strategies**. [s.l.] Springer, 2020. p. 259–276.
- BEHLAU, F. et al. Molecular characterization of copper resistance genes from *Xanthomonas citri* subsp. *citri* and *Xanthomonas alfalfae* subsp. *citrumelonis*. **Appl. Environ. Microbiol.**, v. 77, n. 12, p. 4089–4096, 2011.
- BEHLAU, F. et al. Soluble and insoluble copper formulations and metallic copper rate for control of citrus canker on sweet orange trees. **Crop protection**, v. 94, p. 185–191, 2017.
- BEHLAU, F.; BELASQUE JUNIOR, J. Cancro cítrico: a doença e seu controle. **Fundecitrus, Araraquara, Brasil**, 2014.
- BEHLAU, F.; FONSECA, A. E.; BELASQUE JUNIOR, J. A comprehensive analysis of the Asiatic citrus canker eradication programme in São Paulo state, Brazil, from 1999 to 2009. **Plant Pathology**, v. 65, n. 8, p. 1390–1399, 2016.
- BELASQUE JUNIOR, J. et al. Prováveis consequências do abrandamento da metodologia de erradicação do cancro cítrico no Estado de São Paulo. **Tropical Plant Pathology**, p. 314–317, 2010.
- BHOSALE, S. H.; RAO, M. B.; DESHPANDE, V. V. Molecular and industrial aspects of glucose isomerase. **Microbiol. Mol. Biol. Rev.**, v. 60, n. 2, p. 280–300, 1996.
- BIANCO, M. I. et al. Xanthan pyruvilation is essential for the virulence of *Xanthomonas campestris* pv. *campestris*. **Molecular Plant-Microbe Interactions**, v.

- 29, n. 9, p. 688–699, 2016.
- BIELY, P. Microbial carbohydrate esterases deacetylating plant polysaccharides. **Biotechnology advances**, v. 30, n. 6, p. 1575–1588, 2012.
- BITANCOURT, A. A. O cancro cítrico. **O biológico**, v. 23, n. 10, 1957.
- BOCK, C. H. et al. Factors affecting infection of citrus with *Xanthomonas axonopodis* pv. citri. **Phytopathology**, v. 96, n. 6, p. S14, 2006.
- BRAMUCCI, M. G.; NAGARAJAN, V. Direct selection of cloned DNA in *Bacillus subtilis* based on sucrose-induced lethality. **Applied and environmental microbiology**, v. 62, n. 11, p. 3948–3953, 1996.
- BRUNINGS, A. M.; GABRIEL, D. W. *Xanthomonas citri*: breaking the surface. **Molecular plant pathology**, v. 4, n. 3, p. 141–157, 2003.
- BURLACU, A.; CORNEA, C. P.; ISRAEL-ROMING, F. Microbial xylanase: a review. **Scientific Bulletin. Series F. Biotechnologies**, v. 20, p. 335–342, 2016.
- BUSSE-WICHER, M. et al. Evolution of xylan substitution patterns in gymnosperms and angiosperms: implications for xylan interaction with cellulose. **Plant physiology**, v. 171, n. 4, p. 2418–2431, 2016.
- BÜTTNER, D.; BONAS, U. Regulation and secretion of *Xanthomonas* virulence factors. **FEMS microbiology reviews**, v. 34, n. 2, p. 107–133, 2010.
- CABREJOS, D. A. L. et al. Structural characterization of a pathogenicity-related superoxide dismutase codified by a probably essential gene in *Xanthomonas citri* subsp. citri. **PloS one**, v. 14, n. 1, p. e0209988, 2019.
- CADMUS, M. C. et al. Colonial variation in *Xanthomonas campestris* NRRL B-1459 and characterization of the polysaccharide from a variant strain. **Canadian Journal of Microbiology**, v. 22, n. 7, p. 942–948, 1976.
- CARMONA, J. A. et al. Rheology and structural recovery kinetics of an advanced performance xanthan gum with industrial application. **Applied Rheology**, v. 27, n. 2, p. 19–27, 2017.
- CARNIELLI, C. M. et al. Interaction network and mass spectrometry data of *Xanthomonas citri* subsp. citri surface proteins from differential proteomic analysis of infectious and non-infectious cells. **Data in brief**, v. 8, p. 1400–1411, 2016.
- CARNIELLI, C. M. et al. *Xanthomonas citri* subsp. citri surface proteome by 2D-DIGE: ferric enterobactin receptor and other outer membrane proteins potentially involved in citric host interaction. **Journal of proteomics**, v. 151, p. 251–263, 2017.
- CARVALHO, F. M. DE S. **Expressão gênica em *Xanthomonas axonopodis* pv. citri controlada por promotores induzidos pela planta hospedeira** Universidade de São Paulo, , 2006.
- CHAWLA, R. **Practical clinical biochemistry: methods and interpretations**. [s.l.] JP Medical Ltd, 2014.
- CHOU, F.-L. et al. The *Xanthomonas campestris* gumD Gene Required for Synthesis of Xanthan Gum Is Involved in Normal Pigmentation and Virulence in Causing Black Rot. **Biochemical and biophysical research communications**, v. 233, n. 1, p. 265–269, 1997.
- CHOW, V. et al. Xylan utilization regulon in *Xanthomonas citri* pv. citri strain 306: gene expression and utilization of oligoxylosides. **Appl. Environ. Microbiol.**, v. 81, n. 6, p. 2163–2172, 2015.
- CIVEROLO, E. L. Bacterial canker disease of citrus [*Xanthomonas campestris*]. **Journal of the Rio Grande Valley Horticultural Society**, 1984.
- DA SILVA, A. C. R. et al. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. **Nature**, v. 417, n. 6887, p. 459, 2002.
- DAHL, M. K.; SCHMIEDEL, D.; HILLEN, W. Glucose and glucose-6-phosphate

- interaction with Xyl repressor proteins from *Bacillus* spp. may contribute to regulation of xylose utilization. **Journal of bacteriology**, v. 177, n. 19, p. 5467–5472, 1995.
- DARSONVAL, A. et al. The type III secretion system of *Xanthomonas fuscans* subsp. *fuscans* is involved in the phyllosphere colonization process and in transmission to seeds of susceptible beans. **Appl. Environ. Microbiol.**, v. 74, n. 9, p. 2669–2678, 2008.
- DAS, A. K. Citrus canker-A review. **Journal of Applied Horticulture**, v. 5, n. 1, p. 52–60, 2003.
- DE OLIVEIRA, A. G. et al. Evaluation of the antibiotic activity of extracellular compounds produced by the *Pseudomonas* strain against the *Xanthomonas citri* pv. *citri* 306 strain. **Biological Control**, v. 56, n. 2, p. 125–131, 2011.
- DÉJEAN, G. et al. The xylan utilization system of the plant pathogen *Xanthomonas campestris* pv. *campestris* controls epiphytic life and reveals common features with oligotrophic bacteria and animal gut symbionts. **New Phytologist**, v. 198, n. 3, p. 899–915, 2013.
- DEWDNEY, M. M.; GRAHAM, J. H.; ROGERS, M. E. Citrus canker. **Florida Citrus Pest Management Guide**, v. 2014, p. 93–96, 2014.
- DHARMAPURI, S.; SONTI, R. V. A transposon insertion in the gumG homologue of *Xanthomonas oryzae* pv. *oryzae* causes loss of extracellular polysaccharide production and virulence. **FEMS microbiology letters**, v. 179, n. 1, p. 53–59, 1999.
- DONLAN, R. M. Biofilms: microbial life on surfaces. **Emerging infectious diseases**, v. 8, n. 9, p. 881, 2002.
- DOPSON, R. N. The eradication of citrus canker. **Pl. Dis. Rept**, v. 48, p. 30–31, 1964.
- DUNGER, G. et al. Xanthan is not essential for pathogenicity in citrus canker but contributes to *Xanthomonas* epiphytic survival. **Archives of microbiology**, v. 188, n. 2, p. 127–135, 2007.
- DUNGER, G. et al. *Xanthomonas citri* subsp. *citri* type IV pilus is required for twitching motility, biofilm development, and adherence. **Molecular Plant-Microbe Interactions**, v. 27, n. 10, p. 1132–1147, 2014.
- FAO STAT. **Commodities by country**. Disponível em: <[http://www.fao.org/faostat/en/#rankings/commodities\\_by\\_country](http://www.fao.org/faostat/en/#rankings/commodities_by_country)>. Acesso em: 12 ago. 2019.
- FATIMA, U.; SENTHIL-KUMAR, M. Plant and pathogen nutrient acquisition strategies. **Frontiers in plant science**, v. 6, p. 750, 2015.
- FERENCE, C. 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. 1302–1318, 2018.
- FLEMMING, H.-C.; WINGENDER, J. The biofilm matrix. **Nature reviews microbiology**, v. 8, n. 9, p. 623, 2010.
- FONSECA, N. P. et al. Detection and identification of *Xanthomonas* pathotypes associated with citrus diseases using comparative genomics and multiplex PCR. **PeerJ**, v. 7, p. e7676, 2019a.
- FONSECA, N. P. et al. Analyses of seven new genomes of *Xanthomonas citri* pv. *aurantifolii* strains, causative agents of citrus canker B and C, show a reduced repertoire of pathogenicity-related genes. **Frontiers in microbiology**, v. 10, p. 2361, 2019b.
- FUNDECITRUS. Levantamento da incidência das doenças dos citros: Greening, CVC e cancro cítrico. p. 57, 2018.
- FUWA, H. A New Method For Microdetermination Of Amylase Activity By The Use Of

- Amylose As The Substrate. **The journal of Biochemistry**, v. 41, n. 5, p. 583–603, 1954.
- GOTO, L. S. et al. Structural and functional characterization of the phosphoglucosyltransferase from *Xanthomonas citri* subsp. *citri*. **Biochimica et Biophysica Acta - Proteins and Proteomics**, v. 1864, n. 12, 2016.
- GOTTIG, N. et al. Mechanisms of infection used by *Xanthomonas axonopodis* pv. *citri* in citrus canker disease. **Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology**, v. 1, n. 13, p. 196–204, 2010.
- GOTTWALD, T. R.; GRAHAM, J. H.; SCHUBERT, T. S. Citrus canker: the pathogen and its impact. **Plant Health Progress**, v. 3, n. 1, p. 15, 2002.
- GRAHAM, J. H. et al. Survival of *Xanthomonas campestris* pv. *citri* (Xcc) on various surfaces and chemical control of Asiatic Citrus Canker (ACC), Processing of the International Citrus Canker Research Workshop, June 20-22, 2000, Ft. Pierce, Florida, p. 7, 2000.
- GRAHAM, J. H. et al. *Xanthomonas axonopodis* pv. *citri*: factors affecting successful eradication of citrus canker. **Molecular plant pathology**, v. 5, n. 1, p. 1–15, 2004.
- GRAHAM, J.; MYERS, M. **Soil drenches of imidacloprid, thiamethoxam and acibenzolar-S-methyl for induction of SAR to control citrus canker in young citrus trees**. Phytopathology. **Anais...AMER PHYTOPATHOLOGICAL SOC 3340 PILOT KNOB ROAD, ST PAUL, MN 55121 USA**, 2009
- GRONDIN, J. M. et al. Polysaccharide utilization loci: fueling microbial communities. **Journal of bacteriology**, v. 199, n. 15, p. e00860-16, 2017.
- HALL, D. G.; GOTTWALD, T. R.; BOCK, C. H. Exacerbation of citrus canker by citrus leafminer *Phyllocnistis citrella* in Florida. **Florida Entomologist**, v. 93, n. 4, p. 558–566, 2010.
- HARSHEY, R. M. Bacterial motility on a surface: many ways to a common goal. **Annual Reviews in Microbiology**, v. 57, n. 1, p. 249–273, 2003.
- HAYWARD, A. C. The hosts of *Xanthomonas*. In: **Xanthomonas**. [s.l.] Springer, 1993. p. 1–119.
- HEO, G.-Y. et al. Deletion of *xylR* gene enhances expression of xylose isomerase in *Streptomyces lividans* TK24. **J. Microbiol. Biotechnol**, v. 18, n. 5, p. 837–844, 2008.
- HOTTES, A. K. et al. Transcriptional profiling of *Caulobacter crescentus* during growth on complex and minimal media. **Journal of Bacteriology**, v. 186, n. 5, p. 1448–1461, 2004.
- HOY, M. A.; NGUYEN, R. U. Classical biological control of the citrus leafminer *Phyllocnistis citrella* Stainton (Lepidoptera: Gracillariidae): Theory, practice, art and science. 1997.
- HUANG, T.-P.; SOMERS, E. B.; WONG, A. C. L. Differential biofilm formation and motility associated with lipopolysaccharide/exopolysaccharide-coupled biosynthetic genes in *Stenotrophomonas maltophilia*. **Journal of bacteriology**, v. 188, n. 8, p. 3116–3120, 2006.
- HUTCHISON, J. R. et al. Polysaccharide-based liquid storage and transport media for non-refrigerated preservation of bacterial pathogens. **PloS one**, v. 14, n. 9, 2019.
- IBGE. **Censo Agro 2017**. Disponível em: <<https://censos.ibge.gov.br/agro/2017>>.
- IKAWA, Y. et al. Concomitant regulation by a *LacI*-type transcriptional repressor *XylR* on genes involved in xylan and xylose metabolism and the type III secretion system in rice pathogen *Xanthomonas oryzae* pv. *oryzae*. **Molecular plant-microbe interactions**, v. 31, n. 6, p. 605–613, 2018.
- JENSEN, S. O.; REEVES, P. R. Domain organisation in phosphomannose

- isomerases (types I and II). **Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology**, v. 1382, n. 1, p. 5–7, 1998.
- JIA, H. et al. Modification of the PthA4 effector binding elements in Type I Cs LOB 1 promoter using Cas9/sg RNA to produce transgenic Duncan grapefruit alleviating Xcc $\Delta$ pthA4: dCs LOB 1.3 infection. **Plant biotechnology journal**, v. 14, n. 5, p. 1291–1301, 2016.
- JIA, H. et al. Genome editing of the disease susceptibility gene Cs LOB 1 in citrus confers resistance to citrus canker. **Plant biotechnology journal**, v. 15, n. 7, p. 817–823, 2017.
- JUNIOR, S. **Expressão e purificação do domínio 3XMBT-GST para uso como ferramenta no enriquecimento e identificação de sítios de metil-lisina em Trypanosoma cruzi**. [s.l.] Trabalho de Conclusão de Curso (Bacharelado em Biomedicina). Universidade Federal do Paraná, 2017.
- KAPPAUN, K.; MARTINELLI, A. H. S.; RITT, A. B. B. Expressão de urease ubíqua de soja em Escherichia coli. **Salão de Iniciação Científica (20.: 2008 out. 20-24: Porto Alegre, RS). Livro de resumos. Porto Alegre: UFRGS, 2008.**, 2008.
- KARATAN, E.; WATNICK, P. Signals, regulatory networks, and materials that build and break bacterial biofilms. **Microbiol. Mol. Biol. Rev.**, v. 73, n. 2, p. 310–347, 2009.
- KATOCH, R. **Analytical techniques in biochemistry and molecular biology**. [s.l.] Springer Science & Business Media, 2011.
- KATZEN, F. et al. Promoter analysis of the Xanthomonas campestris pv. campestris gum operon directing biosynthesis of the xanthan polysaccharide. **Journal of bacteriology**, v. 178, n. 14, p. 4313–4318, 1996.
- KATZEN, F. et al. Xanthomonas campestris pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. **Journal of Bacteriology**, v. 180, n. 7, p. 1607–1617, 1998.
- KAY, S. et al. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. **Science**, v. 318, n. 5850, p. 648–651, 2007.
- KOLLER, O. C. **Citricultura: laranja, limão e tangerina**. [s.l.] Rigel Porto Alegre, 1994.
- KOOL, M. M. et al. The influence of the primary and secondary xanthan structure on the enzymatic hydrolysis of the xanthan backbone. **Carbohydrate polymers**, v. 97, n. 2, p. 368–375, 2013.
- KÖPLIN, R. et al. Genetics of xanthan production in Xanthomonas campestris: the xanA and xanB genes are involved in UDP-glucose and GDP-mannose biosynthesis. **Journal of bacteriology**, v. 174, n. 1, p. 191–199, 1992.
- LACERDA, L. A. et al. Protein depletion using the arabinose promoter in Xanthomonas citri subsp. citri. **Plasmid**, v. 90, p. 44–52, 2017.
- LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **nature**, v. 227, n. 5259, p. 680, 1970.
- LAGAERT, S. et al.  $\beta$ -Xylosidases and  $\alpha$ -l-arabinofuranosidases: accessory enzymes for arabinoxylan degradation. **Biotechnology advances**, v. 32, n. 2, p. 316–332, 2014.
- LANZA, F. E. et al. Characteristics of citrus canker lesions associated with premature drop of sweet orange fruit. **Phytopathology**, v. 109, n. 1, p. 44–51, 2018.
- LEITE JR, R. P.; MOHAN, S. K. Integrated management of the citrus bacterial canker disease caused by Xanthomonas campestris pv. citri in the State of Paraná, Brazil. **Crop Protection**, v. 9, n. 1, p. 3–7, 1990.
- LI, J.; WANG, N. The wxacO gene of Xanthomonas citri ssp. citri encodes a protein

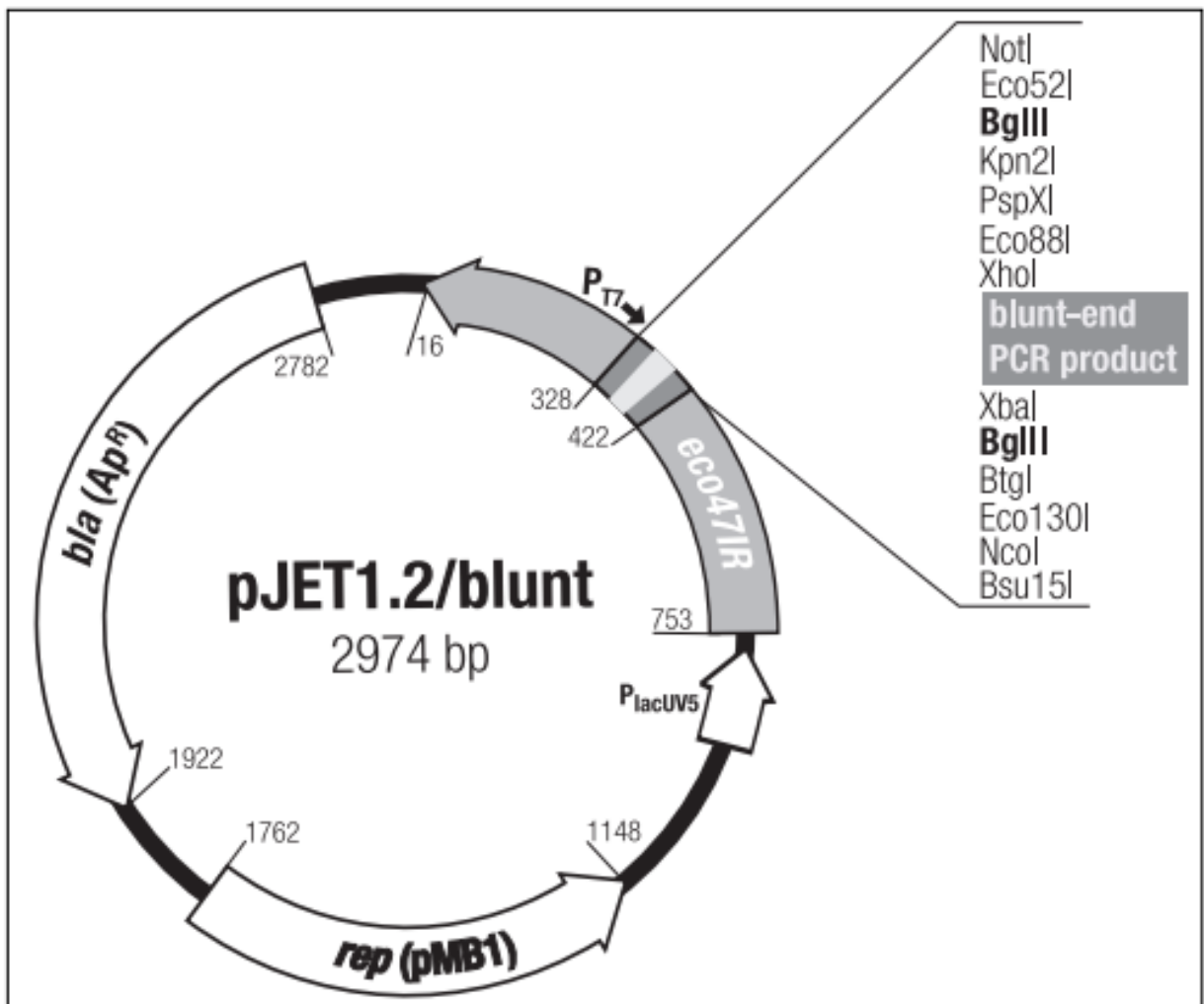
- with a role in lipopolysaccharide biosynthesis, biofilm formation, stress tolerance and virulence. **Molecular plant pathology**, v. 12, n. 4, p. 381–396, 2011a.
- LI, J.; WANG, N. Genome-wide mutagenesis of *Xanthomonas axonopodis* pv. citri reveals novel genetic determinants and regulation mechanisms of biofilm formation. **PLoS One**, v. 6, n. 7, p. e21804, 2011b.
- LOKMAN, B. C. et al. Regulation of expression of the *Lactobacillus pentosus* xylAB operon. **Journal of bacteriology**, v. 179, n. 17, p. 5391–5397, 1997.
- LOPEZ, M. J.; MORENO, J.; RAMOS-CORMENZANA, A. *Xanthomonas campestris* strain selection for xanthan production from olive mill wastewaters. **Water research**, v. 35, n. 7, p. 1828–1830, 2001.
- MALAMUD, F. et al. The *Xanthomonas axonopodis* pv. citri flagellum is required for mature biofilm and canker development. **Microbiology**, v. 157, n. 3, p. 819–829, 2011.
- MEYER, D. F.; BOGDANOVA, A. J. Genomics-driven advances in *Xanthomonas* biology. **Plant pathogenic bacteria: genomics and molecular Biology**, v. 147, p. 161, 2009.
- MILES, J. S.; GUEST, J. R. Nucleotide sequence and transcriptional start point of the phosphomannose isomerase gene (*manA*) of *Escherichia coli*. **Gene**, v. 32, n. 1–2, p. 41–48, 1984.
- MOREAU, C. et al. Lytic polysaccharide monooxygenases (LPMOs) facilitate cellulose nanofibrils production. **Biotechnology for Biofuels**, v. 12, n. 1, p. 156, 2019.
- MOREIRA, L. M. et al. Proteomics-based identification of differentially abundant proteins reveals adaptation mechanisms of *Xanthomonas citri* subsp. citri during *Citrus sinensis* infection. **BMC microbiology**, v. 17, n. 1, p. 155, 2017.
- NAMEKATA, T.; ROSSI, A. C.; CERAVOLO, L. C. Avaliação de novos métodos de erradicação de cancro cítrico. **Laranja, Cordeirópolis**, v. 17, n. 1, p. 67–78, 1996.
- NAZARÉ, A. et al. Design of Antibacterial Agents: Alkyl Dihydroxybenzoates against *Xanthomonas citri* subsp. citri. **International journal of molecular sciences**, v. 19, n. 10, p. 3050, 2018.
- NEVES, M. F.; TROMBIN, V. G. Anuário da citricultura 2017. **São Paulo: CitrusBR**, 2017.
- O'TOOLE, G. A.; KOLTER, R. Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. **Molecular microbiology**, v. 28, n. 3, p. 449–461, 1998.
- PAPOUTSOPOULOU, S. V.; KYRIAKIDIS, D. A. Phosphomannose isomerase of *Xanthomonas campestris*: a zinc activated enzyme. **Molecular and cellular biochemistry**, v. 177, n. 1–2, p. 183–191, 1997.
- PEDRAS, M. S. C.; MINIC, Z. Differential protein expression in response to the phytoalexin brassinin allows the identification of molecular targets in the phytopathogenic fungus *Alternaria brassicicola*. **Molecular plant pathology**, v. 13, n. 5, p. 483–493, 2012.
- PENG, A. et al. Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene *Cs LOB 1* promoter in citrus. **Plant biotechnology journal**, v. 15, n. 12, p. 1509–1519, 2017.
- PEREIRA, A. L. A. et al. Identification of putative TAL effector targets of the citrus canker pathogens shows functional convergence underlying disease development and defense response. **BMC genomics**, v. 15, n. 1, p. 157, 2014.
- ROE, J. H. A colorimetric method for the determination of fructose in blood and urine. **Journal of Biological Chemistry**, v. 107, n. 1, p. 15–22, 1934.

- ROSSETTI, V. V. **Manual ilustrado de doenças dos citros**. [s.l.] Fealq/Fundecitrus, 2001.
- SANCHES, A. L. R. et al. Análise econômica da prevenção e controle do cancro cítrico no estado de São Paulo. **Revista de Economia e Sociologia Rural**, v. 52, n. 3, p. 549–566, 2014.
- SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the national academy of sciences**, v. 74, n. 12, p. 5463–5467, 1977.
- SANTOS, C. R. et al. Molecular mechanisms associated with xylan degradation by *Xanthomonas* plant pathogens. **Journal of Biological Chemistry**, v. 289, n. 46, p. 32186–32200, 2014.
- SÃO PAULO. Resolução da Secretaria de Agricultura e Abastecimento (SAA) nº 147, de 31 de outubro de 2013. Diário Oficial do Estado de São Paulo, 1º de novembro de 2013, Seção 1. p. 50. . 2013.
- SÃO PAULO. Resolução da Secretaria de Agricultura e Abastecimento (SAA) nº 13, de 03 de março de 2017. Diário Oficial do Estado de São Paulo, 4 de março de 2017. Seção 1. p. 21. . 2017.
- SCHAAD, N. W.; JONES, J. B.; CHUN, W. **Laboratory guide for the identification of plant pathogenic bacteria**. [s.l.] American Phytopathological Society (APS Press), 2001.
- SCHMIDT, M. et al. The *Rhizobium meliloti* pmi gene encodes a new type of phosphomannose isomerase. **Gene**, v. 122, n. 1, p. 35–43, 1992.
- SCHUBERT, T. S. et al. Meeting the challenge of eradicating citrus canker in Florida—again. **Plant disease**, v. 85, n. 4, p. 340–356, 2001.
- SHALLOM, D.; SHOHAM, Y. Microbial hemicellulases. **Current opinion in microbiology**, v. 6, n. 3, p. 219–228, 2003.
- SIEVERT, C. et al. Experimental evolution reveals an effective avenue to release catabolite repression via mutations in *XylR*. **Proceedings of the National Academy of Sciences**, v. 114, n. 28, p. 7349–7354, 2017.
- SILVA, I. C. et al. Antibacterial activity of alkyl gallates against *Xanthomonas citri* subsp. *citri*. **Journal of bacteriology**, v. 195, n. 1, p. 85–94, 2013.
- SMITH, D. J. et al. PMI40, an intron-containing gene required for early steps in yeast mannosylation. **Molecular and cellular biology**, v. 12, n. 7, p. 2924–2930, 1992.
- SOUSA, S. A. et al. The *Burkholderia cepacia* bceA gene encodes a protein with phosphomannose isomerase and GDP-D-mannose pyrophosphorylase activities. **Biochemical and biophysical research communications**, v. 353, n. 1, p. 200–206, 2007.
- STALL, R. E.; CIVEROLO, E. L. Research relating to the recent outbreak of citrus canker in Florida. **Annual review of phytopathology**, v. 29, n. 1, p. 399–420, 1991.
- STALL, R. E.; SEYMOUR, C. P. Canker, a threat to citrus in the Gulf-Coast states. **Plant Disease**, v. 67, n. 5, p. 581–585, 1983.
- STANKOWSKI, J. D.; MUELLER, B. E.; ZELLER, S. G. Location of a second O-acetyl group in xanthan gum by the reductive-cleavage method. **Carbohydrate research**, v. 241, p. 321–326, 1993.
- STEPHENS, C. et al. Regulation of D-xylose metabolism in *Caulobacter crescentus* by a *Lacl*-type repressor. **Journal of bacteriology**, v. 189, n. 24, p. 8828–8834, 2007.
- STEVIS, P. E.; HO, N. W. Y. Overproduction of D-xylose isomerase in *Escherichia coli* by cloning the D-xylose isomerase gene. **Enzyme and microbial technology**, v. 7, n. 12, p. 592–596, 1985.

- SWARUP, S. et al. An *Xanthomonas citri* pathogenicity gene, *pthA*, pleiotropically encodes gratuitous avirulence on nonhosts. **Mol Plant Microbe Interact**, v. 5, n. 3, p. 204–213, 1992.
- TEIXEIRA, E. C. et al. The copper resistance operon *copAB* from *Xanthomonas axonopodis* pathovar *citri*: gene inactivation results in copper sensitivity. **Microbiology**, v. 154, n. 2, p. 402–412, 2008.
- TEIXEIRA, R. D. et al. A bipartite periplasmic receptor–diguanylate cyclase pair (XAC2383–XAC2382) in the bacterium *Xanthomonas citri*. **Journal of Biological Chemistry**, v. 293, n. 27, p. 10767–10781, 2018.
- TRINDER, P. Determination of glucose in blood using glucose oxidase with an alternative oxygen acceptor. **Annals of clinical Biochemistry**, v. 6, n. 1, p. 24–27, 1969.
- UCCI, A. P. et al. Asymmetric chromosome segregation in *Xanthomonas citri* ssp. *citri*. **MicrobiologyOpen**, v. 3, n. 1, p. 29–41, 2014.
- VORHÖLTER, F.-J.; NIEHAUS, K.; PÜHLER, A. Lipopolysaccharide biosynthesis in *Xanthomonas campestris* pv. *campestris*: a cluster of 15 genes is involved in the biosynthesis of the LPS O-antigen and the LPS core. **Molecular Genetics and Genomics**, v. 266, n. 1, p. 79–95, 2001.
- WIEDERSCHAIN, G. Y. Glycobiology: Progress, problems, and perspectives. **Biochemistry (Moscow)**, v. 78, n. 7, p. 679–696, 2013.
- WOOD, D. W. New trends and affinity tag designs for recombinant protein purification. **Current opinion in structural biology**, v. 26, p. 54–61, 2014.
- YANG, L. et al. Transformation of sweet orange [*Citrus sinensis* (L.) Osbeck] with *pthA*-nls for acquiring resistance to citrus canker disease. **Plant molecular biology**, v. 75, n. 1–2, p. 11–23, 2011.
- ZANDONADI, F. DA S. Análise proteômica diferencial da fração periplasmática das estirpes A, B e C de *Xanthomonas* spp. que diferem na patogenicidade e espectro de citros hospedeiros. 2012.
- ZHANG, Z.; CHEN, H. Fermentation performance and structure characteristics of xanthan produced by *Xanthomonas campestris* with a glucose/xylose mixture. **Applied biochemistry and biotechnology**, v. 160, n. 6, p. 1653–1663, 2010.

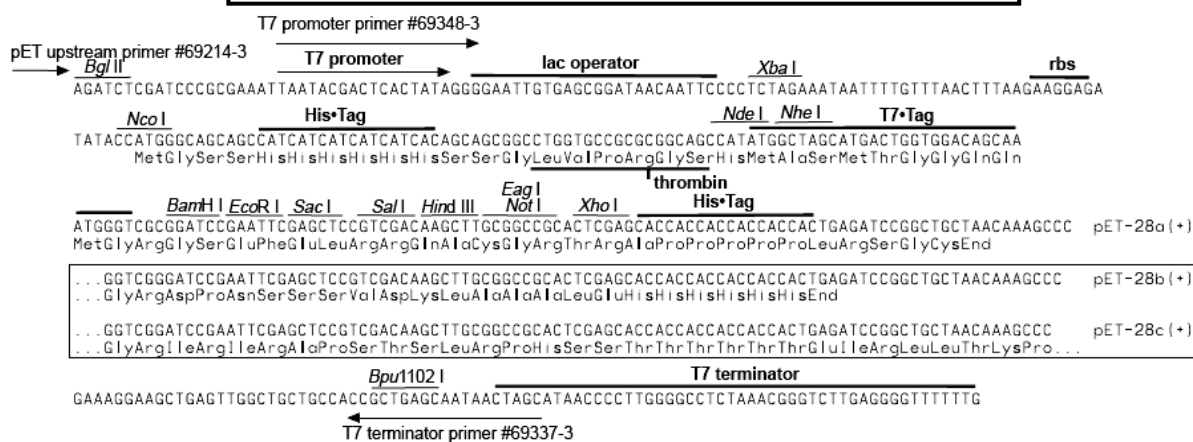
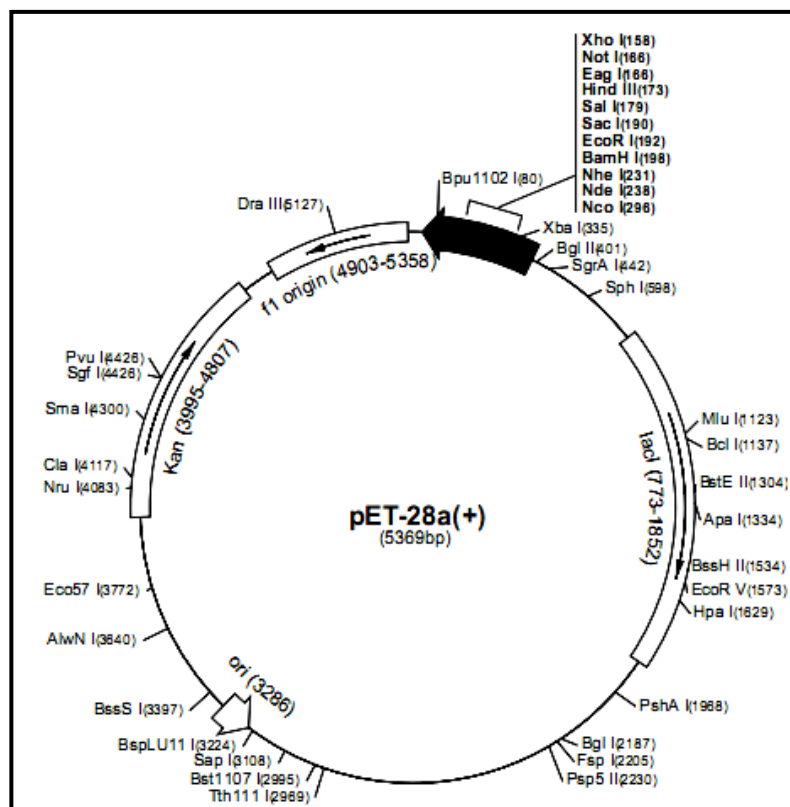


## ANEXO I



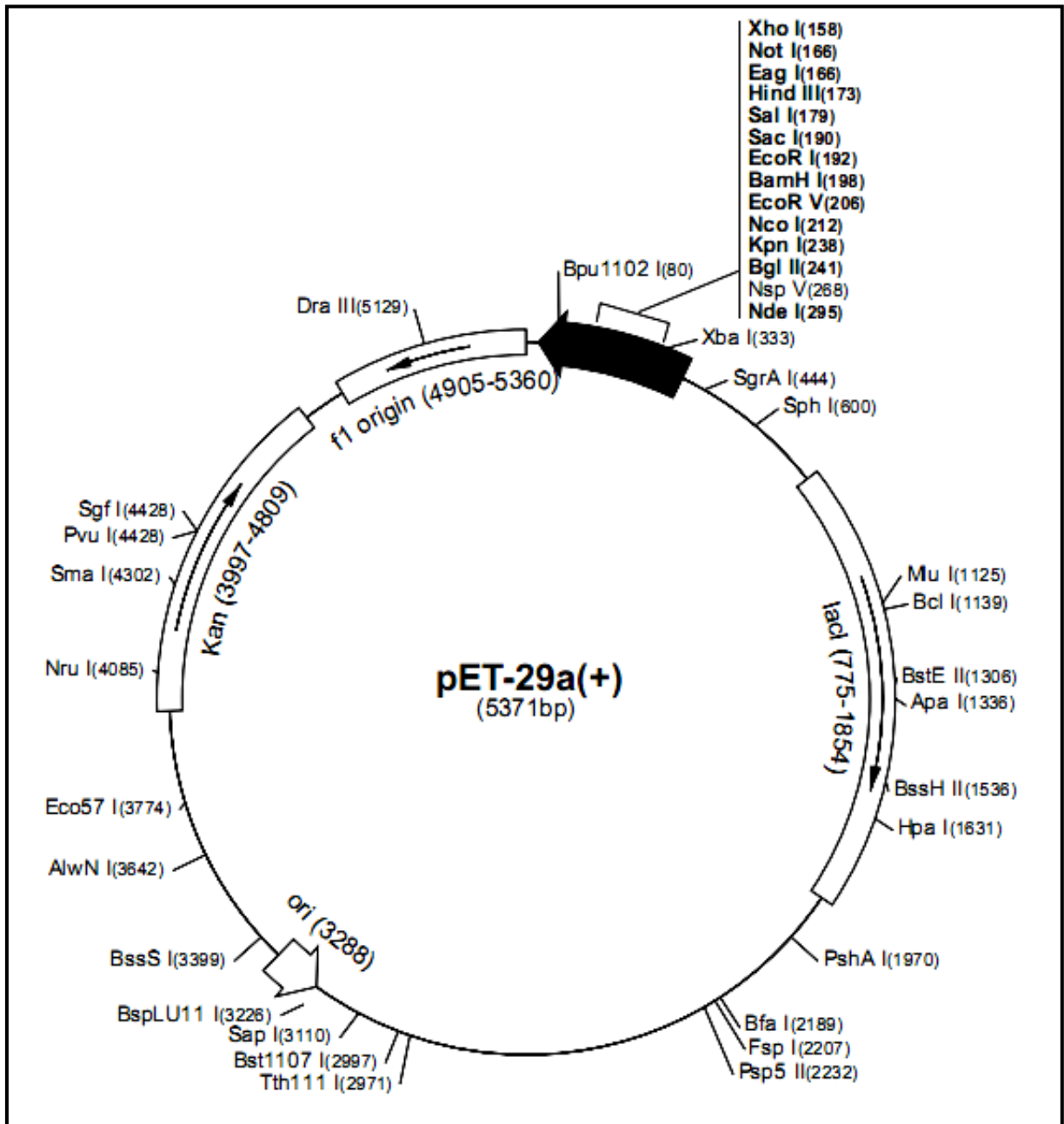
Mapa do vetor pJET 1.2 (Fermentas)

ANEXO II



Mapa do vetor pET28

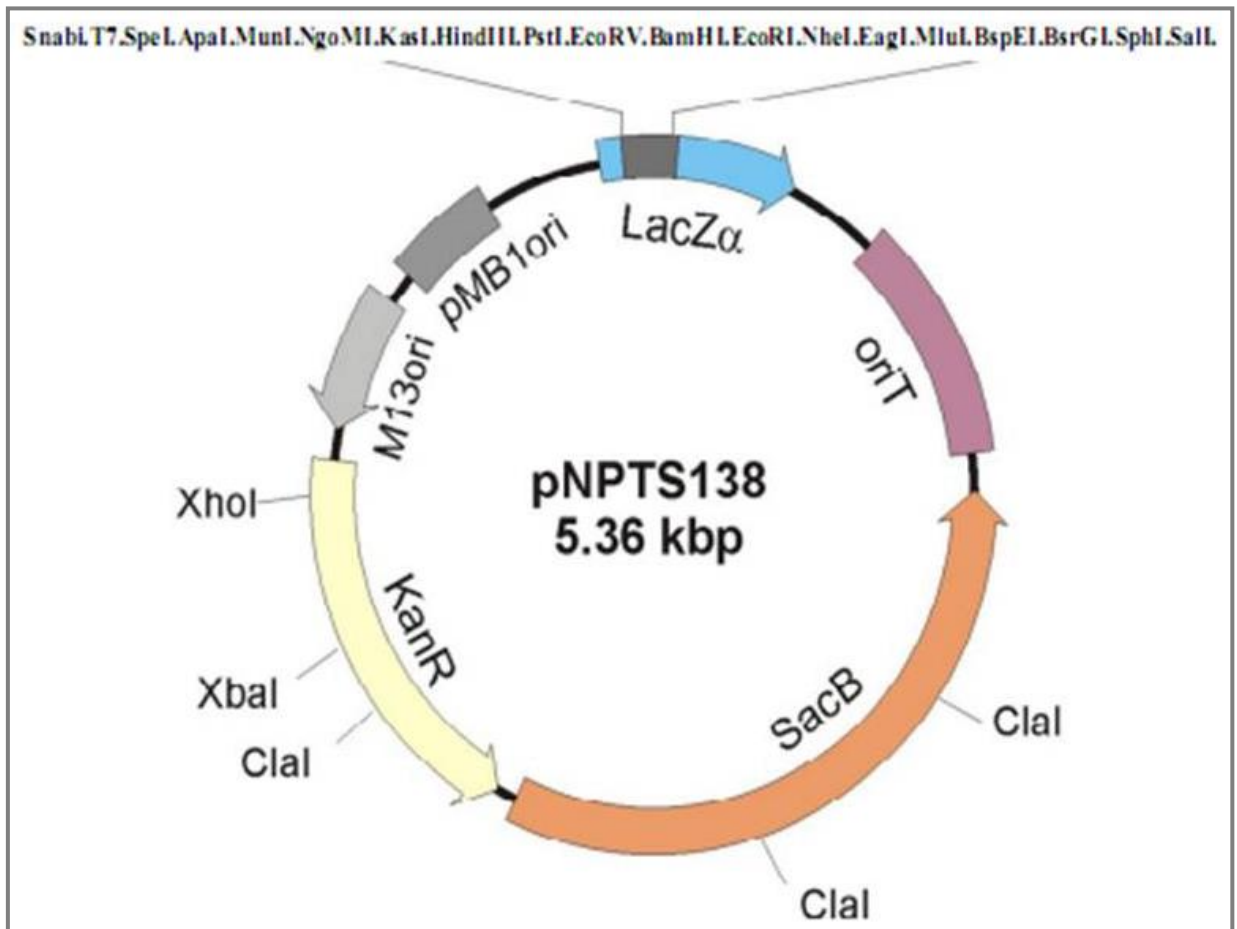
ANEXO III



Mapa do vetor pET29

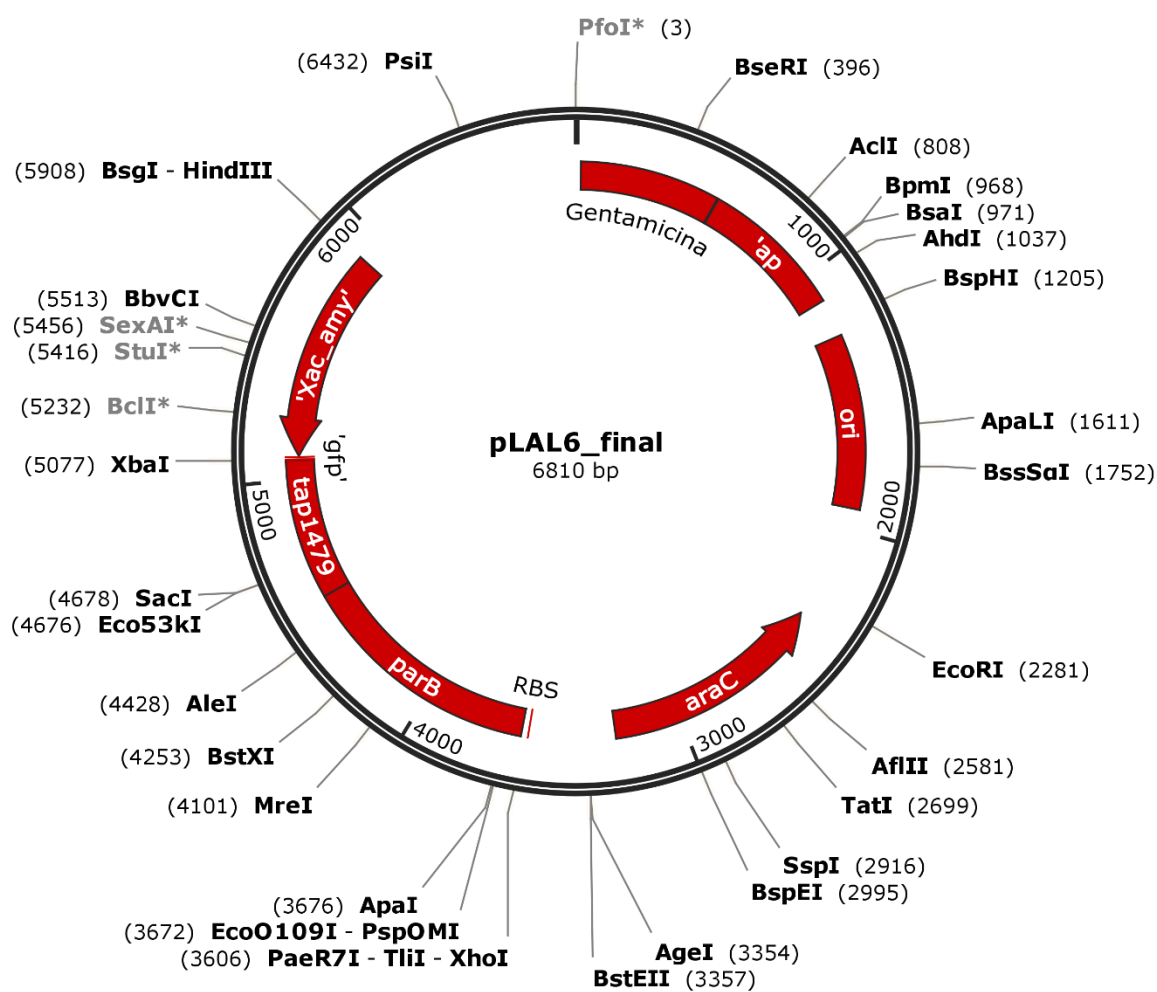


## ANEXO V



Mapa do vetor suicida pNPTS138 (Dickon Alley M. R. K., trabalho não publicado)

## ANEXO VI



Mapa do vetor pLAL6 (FERREIRA, comunicação pessoal).



**Universidade de São Paulo**  
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto  
Departamento de Química  
Programa de Pós-Graduação em Química

Planejamento computacional e avaliação *in vitro* de novos potenciais inibidores da  
fosfomanose isomerase (PMI) de interesse para o controle do cancro cítrico.

Mariana Pegrucci Barcelos

Dissertação apresentada à Faculdade de  
Filosofia, Ciências e Letras de Ribeirão Preto da  
Universidade de São Paulo, como parte das exigências para  
a obtenção do título de Mestre em Ciências, Área: **Química**

RIBEIRÃO PRETO -SP

2020





MARIANA PEGRUCCI BARCELOS

Planejamento computacional e avaliação *in vitro* de novos potenciais inibidores da  
fosfomanose isomerase (PMI) de interesse para o controle do cancro cítrico.

Versão Original

Dissertação apresentada à Faculdade de  
Filosofia, Ciências e Letras de Ribeirão Preto da  
Universidade de São Paulo, como parte das exigências para  
a obtenção do título de Mestre em Ciências, Área: **Química**

**Orientador:** Prof. Dr. Carlos Henrique Tomich de Paula da Silva  
**Co-orientador:** Dr. Leonardo Bruno Federico

RIBEIRÃO PRETO -SP

2020

Autorizo a reprodução total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Barcelos, Mariana Pegrucci

Planejamento computacional e avaliação *in vitro* de novos potenciais inibidores da fosfomanose isomerase (PMI) de interesse para o controle do cancro cítrico. Ribeirão Preto, 2020.

144 p. : il. ; 30cm

Dissertação de Mestrado, apresentada à Faculdade de Filosofia, Ciências e Letras Ribeirão Preto/USP – Área de concentração:  
Química.

Orientador: da Silva, Carlos Henrique Tomich de Paula.

1. Cancro cítrico. 2. Triagem virtual. 3. Modelagem molecular.  
4. Planejamento de inibidores.

## FOLHA DE APROVAÇÃO

Barcelos, M.P. **Planejamento computacional e avaliação *in vitro* de novos potenciais inibidores da fosfomanose isomerase (PMI) de interesse para o controle do cancro cítrico**, 2020. 86f. Dissertação (Mestrado). Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2020.

Aprovado em:

Banca Examinadora

Prof.

Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_

Prof.

Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_

Prof.

Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_

*Dedico esse trabalho à minha  
família, principalmente meus  
pais, por sempre me apoiarem  
em tudo.*

## AGRADECIMENTOS

Primeiramente, gostaria de agradecer aos meus pais, José Anselmo e Elmara, meus maiores exemplos. Mãe, seu cuidado e dedicação me deram, em alguns momentos, a esperança para seguir. Pai, sua presença significou segurança e certeza de que não estou sozinha nessa caminhada.

Não poderia deixar de agradecer ao meu avô, Paulo, que todos os dias me levava para escola e jamais permitiu que eu faltasse.

Ao meu irmão, Lucas, por todo amor e carinho. Aos meus padrinhos, tios, primos e avó que sempre estiveram presentes. Ao meu namorado, Rafael, por todo amor, carinho, paciência e compreensão que tem me dedicado e também pelas risadas nos momentos de aflição.

Ao professor Carlos Henrique que, com muita paciência e atenção, dedicou seu valioso tempo para me orientar em cada passo desse trabalho e também pela contribuição em minha vida acadêmica e influência na minha futura vida profissional.

Aos meus colegas de laboratório do LCQF (Leonardo, João Gabriel, Suzane, Guilherme, Raí, Isaque e Lorane) e ao André do LBBMA. Obrigada pelos momentos de paciência, de brincadeira e, principalmente, pela mão que se estendia quando eu precisava.

Às minhas amigas e irmãs, Leticia, Julia, Marília e Patrícia por todo apoio e cumplicidade, porque mesmo distantes estavam presentes em minha vida.

Às coordenadoras Joze e Cristiane Colégio Objetivo de Batatais por todos conselhos, apoio, torcida, risadas, companheirismo e amizade. Aos alunos, professores e funcionários do Colégio, o meu muito obrigada.

Obrigada a todos que, mesmo não citados aqui, tanto contribuíram para a conclusão desta etapa e para quem eu sou hoje.

*“A ciência compõe-se de  
erros que, por sua vez, são  
os passos até a verdade”.*  
*Júlio Verne*

## RESUMO

Barcelos, M.P. **Planejamento computacional e avaliação *in vitro* de novos potenciais inibidores da fosfomanose isomerase (PMI) de interesse para o controle do cancro cítrico**, 2020. 86 f. Dissertação (Mestrado). Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2020.

O cancro cítrico, causado pela bactéria *Xanthomonas citri* subsp. *citri* (XAC), é uma das doenças mais importantes na citricultura, e apresenta apresentando um significativo impacto na economia brasileira. Estudos de análise proteômica comparativa da superfície de células infectantes e não infectantes de XAC realizados no Laboratório de Bioquímica e Biologia Molecular Aplicada (LBBMA) do Departamento de Genética e Evolução (DGE) da Universidade Federal de São Carlos – UFSCar, apontaram para uma promissora enzima-alvo para o controle da doença, a fosfomanose isomerase (PMI). A PMI é uma enzima que catalisa a interconversão da frutose-6-fosfato (F6P) e da manose-6-fosfato (M6P) e é considerada um potencial alvo terapêutico devido a sua participação na sobrevivência e patogenicidade de diversos microrganismos como *Saccharomyces cerevisiae*, *Candida albicans*, *Mycobacterium smegmatis* e *Leishmania mexicana*. O presente projeto de pesquisa tem como meta o desenvolvimento de compostos químicos capazes de inibir a PMI de *Xanthomonas citri* subsp. *citri* encontrada na superfície bacteriana e, como consequência, investigar a contribuição para a diminuição da virulência da bactéria e/ou para o controle da infecção na planta. O estudo consistiu no desenvolvimento de um modelo tridimensional para o PMI de XAC utilizando o método de “modelagem molecular por homologia estrutural”. A metodologia utilizada para a seleção dos compostos foi a técnica de triagem virtual por forma e similaridade, além de simulações de *docking* e análises de propriedades farmacocinéticas e toxicológicas (ADME/Tox). Os modos de interações dos compostos selecionados pelos passos anteriores foram analisados através de uma inspeção visual, obtenção por fim, aqueles mais promissores. Essas moléculas foram adquiridas para serem testadas em ensaios de atividade *in vitro* utilizando o reagente de Seliwanoff.

**Palavras-chaves:** Cancro cítrico. Fosfomanose isomerase. Metodologias Computacionais





## ABSTRACT

Barcelos, M.P. **Computational plannin and *in vitro* evaluation of new phophomannose isomerase (PMI) inhibitors of interest for citrus canker control**, 2020. 86 f. Dissertation (Master's Degree). Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2020.

Citrus cancer, caused by the bacteria *Xanthomonas citri* subsp. *citri* (XAC), is one of the most important diseases in citrus, and has a significant impact on the Brazilian economy. Comparative protein analysis studies of XAC infecting and non-infecting cells performed at the Laboratory of Biochemistry and Applied Molecular Biology (LBBMA) of the Department of Genetics and Evolution (DGE) of the Federal University of São Carlos - UFSCar, pointing to a promising target enzyme for disease control, a phosphomannose isomerase (PMI). PMI is an enzyme that catalyzes the interconversion of fructose-6-phosphate (F6P) and mannose-6-phosphate (M6P), and is considered a potential therapeutic target due to its participation and pathogenicity of several microorganisms such as *Saccharomyces cerevisiae*, *Candida albicans*, *Mycobacterium smegmatis* and *Leishmania mexicana*. This research project aims to develop chemical compounds capable of inhibiting XAC PMI found on the bacterial surface and, as a consequence, to investigate a contribution to bacterial virulence and/or control. of infection in the plant. The study consisted of the development of a three-dimensional model for XAC PMI using the "structural homology molecular modeling" method. The methodology used for the selection of compounds was the virtual sorting technique by form and similarity, as well as *docking* simulations and pharmacokinetic and toxicological properties analysis (ADME/Tox). The interaction modes of the compounds selected by the previous steps were analyzed by visual inspection, finally obtaining the most promising ones. These molecules were purchased to be tested in *in vitro* activity assays using Seliwanoff reagent.

**Key-words:** Citrus canker. Phosphomannose isomerase. Computational Methodologies.

## LISTA DE FIGURAS

<b>Figura 1</b> - Sintomas do cancro cítrico presente nos frutos, folhas e caules de uma laranjeira.	20
<b>Figura 2</b> - Ciclo do cancro cítrico mostrando como ocorre a infecção e os métodos de propagação da doença. ....	21
<b>Figura 3</b> - Reação de isomerização entre manose-6-fosfato (M6P) e frutose-6-fosfato (F6P). ....	23
<b>Figura 4</b> - Alterações de proteínas na superfície de <i>Xanthomonas citri</i> subsp. <i>citri</i> por análise proteômica em células cultivadas sob condições infecciosas e não-infecciosas e suas possíveis funções. ....	26
<b>Figura 5</b> - Resultado obtido pelo <b>BLASTp</b> na procura de proteínas similares a fosfomanose isomerase de <i>Xanthomonas citri</i> subsp. <i>citri</i> (PDB ID: 2X5S, 2QH5, 2CU2 e 1H5R) .....	40
<b>Figura 6</b> - Alinhamento múltiplo final das sequências de fosfomanose isomerase e as respectivas às estruturas indicadas como homólogas pelo <b>BLASTp</b> (PDB ID 1H5R, 2CU2, 2X5S e 2QH5) obtido pelo <b>Clustal Omega</b> . ....	41
<b>Figura 7</b> - Alinhamento múltiplo final das sequências de fosfomanose isomerase e as respectivas às estruturas indicadas como homólogas pelo <b>BLASTp</b> (PDB ID 1H5R, 2CU2, 2X5S e 2QH5) após o refinamento. ....	43
<b>Figura 8</b> - Modelo gerado para a fosfomanose isomerase de <i>Xanthomonas citri</i> subsp. <i>citri</i> . utilizando o programa <b>MODELLER</b> . ....	47
<b>Figura 9</b> - Modelo final da fosfomanose isomerase de <i>Xanthomonas citri</i> subsp. <i>citri</i> , obtido por modelagem comparativa, com destaque para alguns aminoácidos presentes no sítio ativo. ....	48
<b>Figura 10</b> - Ângulos diedros de uma proteína evidenciando os ângulos Phi e Psi. ....	49
<b>Figura 11</b> - Diagrama de Ramachandram de <i>Xanthomonas citri</i> subsp. <i>citri</i> obtido pelo <b>PROCHECK</b> : 91,6% de resíduos em regiões favoráveis, 7,9% de resíduos em regiões permitidas, 0,4% de resíduos em regiões generosamente permitidas e 0% de resíduos em regiões proibidas. ....	50
<b>Figura 12</b> - Resumo da avaliação estereoquímica do modelo selecionado construído por homologia estrutural para fosfomanose isomerase de <i>Xanthomonas citri</i> subsp. <i>citri</i> ., com um G-Factor de -0.1 e distorção tetraédrica do carbono alfa de -0.1. ....	51
<b>Figura 13</b> - <b>ERRAT</b> obtido para o modelo de fosfomanose isomerase gerado de <i>Xanthomonas citri</i> subsp. <i>citri</i> com um fator de qualidade de 85.5019. ....	52
<b>Figura 14</b> - <b>Verify-3D</b> obtido para o modelo de fosfomanose isomerase gerado de <i>Xanthomonas citri</i> subsp. <i>citri</i> : 91.70% dos aminoácidos obtiveram uma pontuação igual ou superior a 0,2. ....	53
<b>Figura 15</b> - <b>PROVE</b> obtido para o modelo fosfomanose isomerase gerado de <i>Xanthomonas citri</i> subsp. <i>citri</i> . ....	54
<b>Figura 16</b> - Validação do modelo gerado para a fosfomanose isomerase <i>Xanthomonas citri</i> subsp. <i>citri</i> mostrando a posição do ligante glicose-1-fosfato e sua posição em relação do ligante cristalográfico quando utilizado os programas <b>GOLD</b> e <b>Glide</b> . ....	56
<b>Figura 17</b> - Resultado do <b>Redocking</b> da 1H5R para cálculo do RMSD, no qual foi obtido um valor de 2.5080 Å. ....	57
<b>Figura 18</b> - Curva ROC obtida para o estudo dos 30 compostos relatados como mais potentes e menos no <b>BindingDB</b> pelo molde gerado para a proteína fosfomanose isomerase de <i>Xanthomonas citri</i> subsp. <i>citri</i> . ....	60

<b>Figura 19</b> - Curva de Previsão obtida para o estudo dos 30 compostos relatados como ativos e inativos no <b>BindingDB</b> pela proteína fosfomanose isomerase de <i>Xanthomonas citri</i> subsp. <i>citri</i> . .....	61
<b>Figura 20</b> - Curva de enriquecimento com um fator de enriquecimento de 10% obtida para o estudo dos 30 compostos relatados como ativos e inativos no <b>BindingDB</b> pela proteína fosfomanose isomerase de <i>Xanthomonas citri</i> subsp. <i>citri</i> . .....	62
<b>Figura 21</b> - Alinhamento entre as sequências da fosfomanose isomerase de <i>Xanthomonas citri</i> subsp. <i>citri</i> e <i>Homo sapiens</i> feita pelo programa <b>Clustal Omega</b> . .....	64
<b>Figura 22</b> - Alinhamento entre as sequências da fosfomanose isomerase de <i>Homo sapiens</i> e as sequências selecionadas feita pelo programa <b>Clustal Omega</b> . .....	66
<b>Figura 23</b> - Alinhamento múltiplo final das sequências de fosfomanose isomerase humana (PMI) e as respectivas às estruturas indicadas como homólogas pelo <b>BLASTp</b> . (PDB IDs 5NW7 e 3H1M). .....	67
<b>Figura 24</b> - Diagrama de Ramachandram para a hipótese 3 obtido pelo <b>PROCHECK</b> : 89,2% de resíduos em regiões favoráveis, 8,9% de resíduos em regiões permitidas, 1,1% de resíduos em regiões generosamente permitidas e 0,8% de resíduos em regiões proibidas. .....	69
<b>Figura 25</b> - Resumo da avaliação estereoquímica da hipótese 3, construída por homologia estrutural para PMI humana. .....	70
<b>Figura 26</b> - <b>ERRAT</b> obtido para a hipótese 3 gerado de PMI humana. .....	71
<b>Figura 27</b> - <b>Verify-3D</b> obtido para a hipótese 3 da PMI humana: 95,51% dos aminoácidos encontram-se dentro do perfil esperado. .....	71
<b>Figura 28</b> - <b>PROVE</b> obtido para o a hipótese 3 da PMI humana. .....	72
<b>Figura 29</b> - Curva ROC obtida para o estudo dos 30 compostos relatados como ativos e inativos no <b>BindingDB</b> para a proteína PMI de <i>Homo sapiens</i> . .....	74
<b>Figura 30</b> - Curva de Previsão obtida para o estudo dos 30 compostos relatados como ativos e inativos no <b>BindingDB</b> para a proteína PMI de <i>Homo sapiens</i> . .....	75
<b>Figura 31</b> - Curva de enriquecimento com um fator de enriquecimento de 10% obtida para o estudo dos 30 compostos relatados como ativos e inativos no <b>BindingDB</b> para PMI de <i>Homo sapiens</i> . .....	76
<b>Figura 32</b> - Estrutura da G1P, utilizada como referência para os estudos de triagem virtual por similaridade nos programas <b>ROCS</b> e <b>EON</b> . .....	78
<b>Figura 33</b> - Reação química do método de Seliwanoff. .....	85
<b>Figura 34</b> - Esquema do teste empregado na avaliação da atividade enzimática. .....	85
<b>Figura 35</b> - Resultado para o ensaio de avaliação da atividade enzimática da PMI mostrando as colorações obtidas para os testes-controles e a reação entre a M6P e a enzima, após a adição do reagente de Seliwanoff e fervura por 5 minutos. .....	86
<b>Figura 36</b> - Resultado para o ensaio de avaliação da atividade enzimática da PMI mostrando as colorações obtidas para os testes-controles e a reação entre a M6P e a enzima, após a adição do reagente de Seliwanoff e fervura por 5 minutos. .....	87

## LISTA DE TABELAS

<b>Tabela 1</b> - Relação do número de compostos presentes em cada base de dados utilizada. ....	33
--------------------------------------------------------------------------------------------------	----

<b>Tabela 2</b> - Dados referentes aos ensaios in vitro realizados para a fosfomanose isomerase recombinante de <i>Xanthomonas citri</i> subsp. <i>citri</i> .....	37
<b>Tabela 3</b> - Resultados de identidade e e-value de sequências de proteínas obtidas no servidor online <b>BLAST</b> usando o algoritmo para proteínas ( <b>BLASTp</b> ) na triagem de proteínas homologadas a fosfomanose isomerase de <i>Xanthomonas citri</i> subsp. <i>citri</i> .....	40
<b>Tabela 4</b> - Apresentação dos resultados da Pontuação Geral ( <b>WHATIF</b> ), Regiões Proibidas (Ramachandran) e Fator G ( <b>PROCHECK</b> ) obtidos na etapa de validação dos modelos gerados. ....	46
<b>Tabela 5</b> - Resultado da recuperação da seleção de inibidores (15 mais potentes e 15 menos potentes) do <b>BindingDB</b> pela fosfomanose isomerase de <i>Xanthomonas citri</i> subsp. <i>citri</i> . ....	58
<b>Tabela 6</b> - Resultados de identidade e e-value de sequências de proteínas obtidas no servidor online <b>BLAST</b> usando o algoritmo de explosão de proteínas ( <b>BLASTp</b> ) para a triagem de proteínas homologadas para PMI humana. ....	65
<b>Tabela 7</b> - Apresentação dos resultados da Pontuação Geral ( <b>WHATIF</b> ), Regiões Proibidas (Ramachandran) e Fator G ( <b>PROCHECK</b> ) obtidos na etapa de validação dos modelos gerados. ....	68
<b>Tabela 8</b> - Resultado da recuperação da seleção de inibidores do <b>BindingDB</b> .....	73
<b>Tabela 9</b> - Relação entre os compostos sobreviventes das bases de dados após simulações de <i>docking</i> com esfera de 8 Å, predições ADME/Tox ( <b>DEREK</b> e <b>QikProp</b> ) e inspeção visual, utilizando a primeira estratégia. ....	79
<b>Tabela 10</b> - Relação entre os compostos sobreviventes das bases de dados após estudos triagem virtual por similaridade ( <b>ROCS</b> e <b>EON</b> ), de <i>docking</i> , predições ADME/Tox ( <b>DEREK</b> e <b>QikProp</b> ) e inspeção visual, utilizando a segunda estratégia. ....	80
<b>Tabela 11</b> - Relação entre os compostos sobreviventes das bases de dados após simulações de <i>docking</i> com esfera de 6 Å e 60 de fitness, triagem virtual por similaridade ( <b>ROCS</b> e <b>EON</b> ), predições ADME/Tox ( <b>DEREK</b> e <b>QikProp</b> ) e inspeção visual, utilizando a terceira estratégia. ....	81
<b>Tabela 12</b> - Tipo e comprimeto de interações entre os compostos comprados e os aminoácidos presentes no sítio ativo. ....	83
<b>Tabela 13</b> - Dados referentes aos ensaios in vitro realizados para a fosfomanose isomerase recombinante de <i>Xanthomonas citri</i> subsp. <i>citri</i> .....	87
<b>Tabela 14</b> - Resultados obtidos para os ensaios de inibição da enzima PMI de <i>Xanthomonas citri</i> subsp. <i>citri</i> pelo inibidor <b>DL03874928</b> . ....	88

## LISTA DE ABREVIATURAS E SIGLAS

AUC	Área sob a curva
BEDROC	Discriminação reforçada por Boltzmann do ROC

BLAST	Basic Local Aligment Search Tool
DGE	Departamento de Genética
EF	Fator de enriquecimento
EMBL-	European Molecular Biology Laboratory - European
EBI	Bioinformatics Institute
F6P	frutose-6-fosfato
FDA	Food and Drug Administration
FFP	Frações Falso-Positivas
FVP	Frações Verdadeiras Positivas
G1P	Glicose-1-fosfato
G-Factor	General Factor
LBBMA	Laboratório de Bioquímica e Biologia Molecular Aplicada
M1P	Manose-1-fosfato
M6P	manose-6-fosfato
MDCK	Madin-Darby Canine Kidney
NCBI	National Center for Biotechnology Information
ONSA	Organization for Nucleotide Sequency and Analysis
PDB	Protein Data Bank
PMI	fosfomanose isomerase
RIE	Reforço inicial robusto
RMSD	Root-Mean-Square Deviation
ROC	Receiver Operating Characteristics
Sdf	Spatial Data Format
TG	Total padronizado
UFSCar	Universidade Federal de São Carlos
XAC	<i>Xanthomonas citri</i> subsp. <i>citri</i>
XAU-B	<i>Xanthomonas fuscans</i> subsp <i>aurantifolli</i> tipo B
XAU-C	<i>Xanthomonas fuscans</i> subsp <i>aurantifolli</i> tipo C

## Sumário

1. INTRODUÇÃO .....	18
1.1 O cancro cítrico .....	18
1.2 O agente etiológico do cancro cítrico: a bactéria <i>Xanthomonas</i> sp. ....	21
1.3 Fosfomanose isomerase .....	23
1.4 Planejamento computacional de substâncias bioativas .....	26
2. OBJETIVOS .....	29
3. MATERIAIS E MÉTODOS .....	30
3.1. Construção do modelo de fosfomanose isomerase para <i>Xanthomonas citri</i> subsp. <i>citri</i> .....	30
3.1.1 Pesquisa e Alinhamento de sequências homólogas .....	30
3.1.2 Alinhamento das sequências .....	30
3.1.3 Construção e avaliação da qualidade do modelo .....	31
3.1.4. Validação do modelo de PMI gerado para simulações de <i>docking</i> .....	32
3.2 Triagem virtual de novos inibidores da fosfomanose isomerase .....	32
3.2.1. Geração de confôrmeros ( <b>OMEGA</b> ) .....	33
3.2.2. Seleção de inibidores baseados em ligantes ( <b>ROCS e EON</b> ) .....	34
3.2.3. Seleção de inibidores baseados em estrutura ( <b>GOLD</b> ).....	34
3.2.4 Predição das propriedades farmacocinéticas e toxicológicas ( <b>QikProp e DEREK</b> ) .....	34
3.3. Visualização Gráfica .....	35
3.4. Ensaio de atividade <i>in vitro</i> com a enzima PMI de <i>Xanthomonas citri</i> subsp. <i>citri</i> ... ..	35
3.4.1 Método de avaliação da atividade enzimática.....	36
3.4.2 Avaliação <i>in vitro</i> da inibição da fosfomanose isomerase pelos compostos selecionados .....	36
4. RESULTADOS E DISCUSSÃO .....	38
4.1 Construção do modelo da fosfomanose isomerase (PMI) para <i>Xanthomonas citri</i> subsp. <i>citri</i> (XAC) .....	38
4.1.1 Etapa de identificação da sequência de aprovação .....	38
4.1.2 Etapa de alinhamento das sequências .....	41
4.1.3 Etapa de construção e avaliação do modelo.....	44
4.1.4 Avaliação dos ângulos diédricos $\Psi$ versus $\phi$ através do Diagrama de Ramachandran .....	48
4.1.5 Outros parâmetros importantes para a análise (Suite <b>SAVE v5.0</b> ).....	50
4.1.6 Etapa de validação do modelo de PMI XAC. ....	54
4.2 Construção do modelo da fosfomanose isomerase (PMI) para <i>Homo sapiens</i> .....	63
4.1.1 Etapa de identificação da sequência de aprovação .....	65

4.1.2 Etapa de alinhamento .....	65
4.1.3 Etapa de construção e avaliação do modelo .....	67
4.1.4 Avaliação dos ângulos diédricos $\Psi$ versus $\varphi$ através do diagrama de Ramachandran .....	69
4.1.5 Outros parâmetros importantes para a análise dos modelos de PMI humana .....	70
4.1.6 Etapa de validação do modelo de PMI humana .....	72
4.3 Triagem virtual par seleção de inibidores da PMI de XAC. ....	77
4.4 Ensaios de atividade in vitro com a enzima PMI. ....	84
4.4.1. Método de avaliação da atividade enzimática .....	84
4.4.2. Avaliação in vitro da inibição da PMI pelo composto selecionado .....	86
5. CONCLUSÃO.....	89
6.PERSPECTIVAS FUTURAS .....	90
REFERÊNCIAS BIBLIOGRÁFICAS .....	91
APÊNDICES .....	98
APÊNDICE A - Dados relacionados a construção do modelo de PMI <i>Xanthomonas citri</i> subsp. <i>citri</i> .....	98
APÊNDICE B - Dados relacionados a construção do modelo de PMI <i>Homo sapiens</i> ....	114

## 1. INTRODUÇÃO

O Brasil é o maior produtor e exportador de laranjas do mundo, sendo a citricultura aquela com maior relevância dentro do agronegócio (CEPEA - CENTRO DE ESTUDOS AVANÇADOS EM ECONOMIA APLICADA, [s.d.]). Dois fatores contribuem para que o país ocupe essa posição: o tamanho territorial e o clima favorável. De acordo com o Departamento de Agricultura dos Estados Unidos, a produção brasileira de laranja e suco de laranja apresentou uma média de 34% e 56% da produção mundial entre as safras de 2012 e 2015. Aliando essa participação com o avanço da tecnologia empregada nos cinturões, o Brasil apresenta uma participação de 76% do mercado mundial de citros (NEVES; TROMBIN, 2017). Apesar de todos os empenhos empregados para manter a alta produtividade nos pomares brasileiros, a citricultura tem enfrentado alguns problemas como algumas doenças que atacam os citros, sendo uma das principais o cancro cítrico (GOTTWALD et al., 2001).

### 1.1 O cancro cítrico

O cancro cítrico constitui uma das principais doenças ocasionadas em pomares de citros e é causado pela bactéria de gênero *Xanthomonas sp.*, sendo o tipo mais severo ocasionado por *Xanthomonas citri* subsp. *citri* (XAC). Os sintomas são lesões circulares, salientes, corticosas, de coloração castanha, bem como eruptivas, ocorrendo em folhas, ramos e frutos (BITTANCOURT, 1957; ROSSETTI, 2011). Em casos mais severos, ocorre a queda dos frutos imaturas impactando na perda direta da produção (BEHLAU; BELASQUE JR, 2014).

O controle da doença cítrica baseia-se em medidas de erradicação e exclusão, além de medidas regulamentárias determinantes da erradicação de plantas e pomares afetados pela doença (GOTTWALD et al., 2001). Em 1999, a Defesa Agropecuária do Estado de São Paulo determinou que todas as plantas dos talhões infestados deveriam ser eliminadas obrigatoriamente quando a incidência fosse superior a 0,5% e para aquelas em que a incidência fosse igual ou menor que 0,5%, as plantas doentes e as demais contidas em um raio de trinta metros também seriam eliminadas (BEHLAU; FONSECA; BELASQUE, 2016). Esta última medida foi tomada após o aparecimento da lagarta mineradora dos citros (*Phyllocnistis citrella*) nos laranjais, cuja presença contribuiu para o aumento da incidência do cancro cítrico (SEILLE; CHRISTIANO, 2006) e, conseqüentemente, tornou ineficiente a aplicação do raio de trinta metros para a retenção da doença. (BEHLAU; BARELLI; BELASQUE, 2014; FERENGE et al., 2018).



A incidência da doença voltou a crescer nos laranjais do estado de São Paulo desde que a legislação foi amenizada, em meados de 2009, excluindo a primeira exigência de erradicação de um talhão de plantas correspondente a contaminações iguais ou maiores a 0,5%. Entre os anos de 2001 e 2009, a doença manteve baixas incidências de talhões contaminados com valores variando entre 0,08% e 0,20%, mas, a partir de 2010, a contaminação avançou e atingiu valores de 0,44% de talhões contaminados. Por fim, em 2012, atingiu 1,39% registrando a maior incidência desde seu primeiro relato em 1957 (FUNDECITRUS, 2017a).

Em 2017, uma nova medida foi tomada: caso a doença apareça em uma área onde a praga ocorre com baixa incidência, o controle é feito principalmente pelo plantio de mudas sadias e pela eliminação das plantas cítricas contaminadas e daquelas suspeitas de contaminação. Por outro lado, caso haja uma área em que o cancro cítrico já esteja difundido entre as plantas, deverá ser realizado um conjunto de medidas de manejo, como a utilização de um composto de cobre, que visam reduzir o impacto da doença na produção (FUNDECITRUS, 2016, 2017b, 2017c)

A bactéria responsável pela doença cítrica possui a capacidade de infectar folhas, frutos e ramos (Figura 1), mesmo não possuindo um mecanismo ativo de penetração através da parede celular. Como sua propagação não necessita de um vetor específico, em condições naturais de infecção a disseminação ocorre a curtas distâncias pela ação de respingos de chuvas, principalmente quando associados com ventos, e a disseminação a longas distâncias ocorre pelo transporte de material vegetal infectado (DANOS, E; BERGER, R. D.; STALL, 1984; FUNDECITRUS, 2017a; GOTTWALD, T.R., MCGUIRE, R.G. & GARRAN, 1988). Por isso, é necessária a manutenção dos equipamentos de colheita através da desinfecção, além da pulverização dos veículos antes de entrar na propriedade para impedir a difusão da doença e assim manter a propriedade livre da mesma.

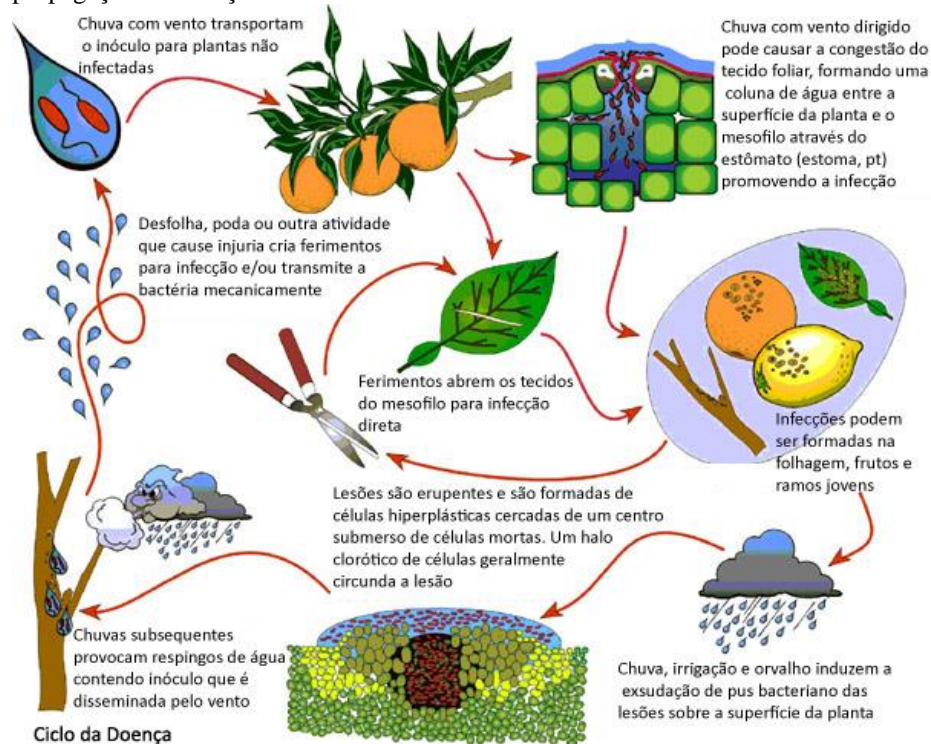
**Figura 1** - Sintomas do cancro cítrico presente nos frutos, folhas e caules de uma laranjeira.



Fonte: Schubert, 2011.

O ciclo de infecção (Figura 2) começa pelo acúmulo anormal de água nas folhas, o que acarreta a formação de uma coluna de água entre a superfície foliar e o mesófilo dos estômatos e/ou hidatódios, permitindo a entrada do patógeno. Além disso, essa água causa a mobilização de bactérias para fora das lesões foliares, as quais são células hiperplásticas em torno de um centro de células mortas colapsadas, em que um anel ou halo clorótico frequentemente rodeiam a lesão. Após a saída da bactéria pela lesão foliar, o patógeno é carregado pelo vento, disseminando a doença e inoculando plantas não infectadas. Espinhos, a poda ou outras atividades mecânicas podem causar ferimentos nas folhas criando uma porta para a infecção e/ou transmissão bacteriana (GOTTWALD, TIM R.; GRAHAM, 2005) e o ciclo recomeça.

**Figura 2** - Ciclo do cancro cítrico mostrando como ocorre a infecção e os métodos de propagação da doença.



Fonte: Cardoso, 2013.

### 1.2 O agente etiológico do cancro cítrico: a bactéria *Xanthomonas* sp..

O agente etiológico do cancro cítrico é uma bactéria pertencente ao gênero das *Xanthomonas*, com metabolismo oxidativo, Gram-negativa, aeróbica e não fixadora de nitrogênio (ARTIER, 2010). Há diferentes tipos de cancro cítrico causado pelas diferentes variantes da bactéria *Xanthomonas citri*. Como os sintomas são geralmente iguais, a sua distinção ocorre por meio do hospedeiro e pelas características fenotípicas e genotípicas, sendo assim, a doença é dividida em cancores do tipo “A”, “B”, “C”, “D” e “E” (GRAHAM et al., 2004), porém, neste estudo, será abordada apenas A, B e C, devido ao prejuízo econômico acometido pelas mesmas, pois são as com maior importância econômica.

A cancorese tipo **A** corresponde a forma mais severa da doença, sendo tipicamente da Ásia e causada por *Xanthomonas citri* subsp. *citri* (XAC), possuindo a maior gama de hospedeiros cítricos, como por exemplo, a laranja doce (*Citrus sinensis*), limão (*C. limon*) e lima ácida (*C. aurantifolia*) (STALL; SEYMOUR, 1983). É a estirpe que está mais difundida no mundo, atingindo o Japão, Argentina, Brasil, Paraguai, China, entre outros países.

A cancorese tipo **B** causada por *Xanthomonas fuscans* subsp. *aurantifolii* (Xau-B) é menos severa e de importância secundária quando comparada com a cancorese tipo **A** e atinge,

principalmente, os limoeiros verdadeiros da Argentina, Paraguai e Uruguai. A cancrose tipo **B** pode ser facilmente diferenciada do tipo **A**, mas não pela cancrose causada pelo tipo **C** (BEHLAU et al., 2007; GOTTWALD et al., 2001).

A cancrose tipo **C**, também conhecida como cancrose da limeira ácida “galego”, é causada por *Xanthomonas fuscans* subsp. *aurantifolii* tipo **C** (*Xau-C*) ocorrendo em único tipo de hospedeiro, a lima ácida. Ocorre somente no Brasil, afetando apenas alguns municípios do estado de São Paulo (BEHLAU et al., 2007; GOTTWALD et al., 2001). Como cada linhagem possui características próprias, marcadores específicos poderiam eventualmente auxiliar no diagnóstico de plantas infectadas no campo, especialmente em períodos que os sintomas não estejam presentes.

O sequenciamento completo dos genes da bactéria causadora da cancrose tipo **A** foi realizado pelos laboratórios da rede ONSA (*Organization for Nucleotide Sequencing and Analysis*) da FAPESP (DA SILVA et al., 2002a). Os resultados mostraram que o seu genoma é composto por 5.175.554 pares de bases, os quais codificam 4.313 proteínas. A comparação entre os diversos genomas das linhagens dessas bactérias poderia nos fornecer informações sobre o hospedeiro alvo e sua patogenicidade. Devido a esse fato, os projetos de sequenciamento das linhagens **B** (*XAU-B*) e **C** (*XAU-C*) também foram realizados por grupos de pesquisa da USP-SP e UNESP- Jaboticabal, com o suporte da FAPESP e do FUNDECITRUS (BODNAR et al., 2017; JALAN et al., 2011; MOREIRA et al., 2010; RICHARD et al., 2017) para que seja possível a comparação entre as sequências (ESCALON et al., 2013; GORDON et al., 2015; JALAN et al., 2013).

A *XAC* possui o sistema de secreção tipo III (SSTT) no qual a bactéria “percebe” a planta ou as condições do ambiente como favoráveis, o que ativa a produção proteica e o SSTT. Em sua grande maioria, as bactérias patogênicas gram-negativas de plantas têm dois grupos de genes modulares de suas interações com os hospedeiros, os genes *avr* (avirulência) responsáveis por determinar a especificidade do mesmo por uma interação genótipo-genótipo, e os genes *hrp* (resposta de hipersensibilidade e patogenicidade) relacionados à patogenicidade e indução de respostas de hipersensibilidade em plantas não-hospedeiras (BONAS, 1994; MOREIRA et al., 2004).

Os genes *hrp*, que se encontram ativados, produzem uma estrutura de adesão com a célula chamada de *pilus*, e através dele e da formação de um poro condutor, as proteínas são transferidas diretamente para o citosol da célula hospedeira. Uma dessas proteínas são as *avr* que, como dito anteriormente, causam uma reação de hipersensibilidade. Crê-se que algumas

dessas proteínas transportadas possam causar os sintomas da doença na planta hospedeira (BUTTNER, 2002).

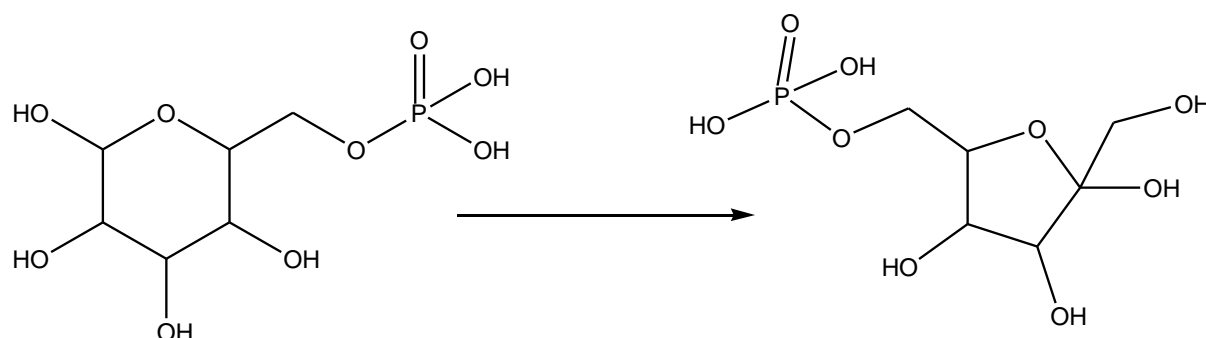
Além do Sistema de Secreção Tipo III, XAC., utilizam um mecanismo de conjugação para exportar moléculas efetoras (proteínas ou complexos DNA-proteínas) durante a infecção (YACOUBI, 2005). Estes sistemas são chamados de Sistemas de Secreção Tipo IV ou “Conjugação” Adaptada e secretam proteínas efetoras para o meio extracelular ou para o citosol das células hospedeiras, envolvendo o contato direto célula-célula (YACOUBI, 2005).

Diversos estudos sobre os mecanismos específicos das interações de XAC com seus hospedeiros têm sido realizados com a finalidade de entender a maneira pela qual a bactéria utiliza o Sistema de Secreção tipo III e o Sistema de Secreção tipo IV para interagir e modificar o metabolismo das plantas infectadas (FERENCE et al., 2018; ZHANG et al., 2015).

### 1.3 Fosfomanose isomerase

A fosfomanose isomerase (PMI), alternativamente chamada de manose-6-fosfato isomerase é uma enzima que catalisa a interconversão da frutose-6-fosfato (F6P) e da manose-6-fosfato (M6P) (Figura 3) além de permitir a síntese do difosfato de guanosina manose (GDP-manose) em organismos eucarióticos. A PMI é monomérica e uma metaloenzima, ou seja, é dependente de um íon metálico divalente para desempenhar as funções de atividade e catálise, sendo inibida pela eritose-4-fosfato, manitol-1-fosfato e, em menor grau, pelo anômero alfa da manose-6-fosfato (GAO; YU; LEARY, 2005).

**Figura 3** - Reação de isomerização entre manose-6-fosfato (M6P) e frutose-6-fosfato (F6P).



Fonte: O autor

A fosfomanose isomerase converte uma aldose (manose) em uma cetose (frutose), além de abrir e fechar os anéis para esses açúcares (STRYER, L.; BERG, J.M., TYMOCZKO, 2007). Essa enzima mostra um alto grau de seletividade para o anômero beta de M6P enquanto que o

anômero alfa não possui atividade e pode, de fato, atuar como um inibidor (ROSE, IRWIN A.; O'CONNELL, EDWARD L.; SCHRAY, 1973).

Proudfoot e colaboradores (PROUDFOOT et al., 1994) definiram três tipos de PMIs com base no alinhamento da sequência e propriedades físico-químicas e cinéticas. As enzimas dos mamíferos e da maioria das bactérias pertencem às PMIs do Tipo I, que são proteínas monofuncionais homólogas que catalisam uma única reação de isomerização (PROUDFOOT et al., 1994; PROUDFOOT; PAYTON; WELLS, 1994). Aquelas pertencentes ao Tipo II são enzimas bifuncionais e possuem similaridade de sequência limitada às enzimas do Tipo I, exceto por um motivo muito curto (9 aminoácidos) (JENSEN; REEVES, 1998). Por fim, as enzimas do Tipo III catalisam reações de isomerização reversíveis, mas compartilham pouca identidade com sequências das enzimas do Tipo I e do Tipo II (SIGDEL et al., 2015).

A enzima presente em *Xanthomonas citri* subsp. *citri* é composta por 467 resíduos e pertence à família da manose-1-fosfato guanililtransferase e manose-6-fosfato isomerase, portanto, trata-se de uma enzima bifuncional (DA SILVA et al., 2002b).

A fosfomanose isomerase apresenta várias contribuições para as vias metabólicas necessárias. Ela permite que as células transformem o M6P em F6P, que poderá ser inserido na glicólise e também permitirá que as células convertam a F6P em M6P, um identificador celular glicolítico comum para o transporte celular e a identificação da membrana celular em organismos procarióticos e eucarióticos (STRYER, L.; BERG, J.M., TYMOCZKO, 2007). Esta enzima pode ser útil no desenvolvimento de novos tratamentos antifúngicos, uma vez que a falta de atividade da PMI em células de levedura pode levar à lise celular e a enzima pode ser alvo de inibição (CLEASBY et al., 1996). Isso pode ocorrer devido ao papel da PMI na formação das paredes celulares e na biossíntese de polissacarídeos capsulares, além do mais, a M6P é uma molécula de sinalização importante, especialmente para o transporte para os lisossomos. Distúrbios que afetam a atividade da PMI podem afetar a capacidade celular para produzir rapidamente a M6P a partir da F6P abundante e, portanto, o tráfego de vesículas para os lisossomos e os endossomos pode ser alterado, impactando negativamente a célula e acarretando uma lise celular (JAEKEN; MATTHIJS, 2001).

A PMI é considerada como potencial alvo terapêutico por causa da participação na sobrevivência e patogenicidade de diversos microrganismos como *Saccharomyces cerevisiae* (PAYTON et al., 1991), *Candida albicans* (CLEASBY et al., 1996), *Mycobacterium smegmatis* (PATTERSON et al., 2003) e *Leishmania mexicana* (GARAMI; ILG, 2001). Dessa forma, a informação estrutural sobre essa enzima presente em diferentes organismos ajudará na

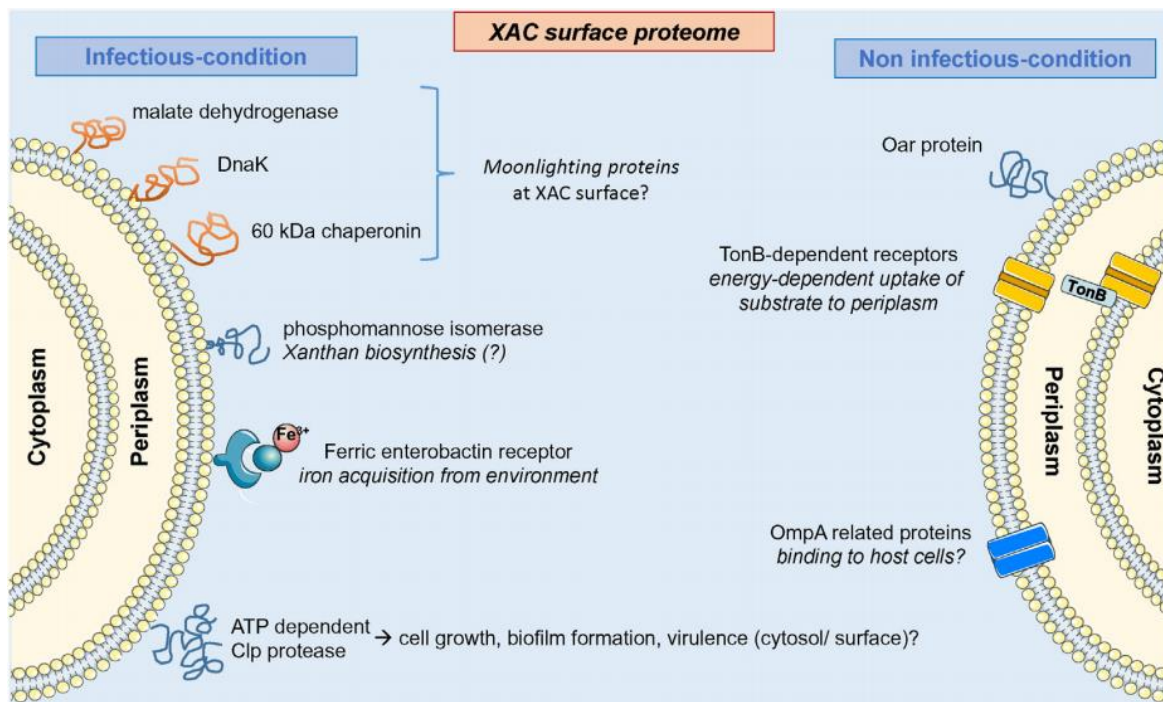
concepção e otimização de compostos que inibam seletivamente as enzimas de organismos patogênicos, de tal forma que não afetem as células do hospedeiro, mantendo-as saudáveis.

A primeira estrutura tridimensional dessa enzima foi depositada em 2004 no Protein Data Bank (PDB) (KIRCHMAIR et al., 2008) e pertence à espécie *Bacillus subtilis* (código PDB 1QWR, a 1.8 Å de resolução). Essa estrutura revelou a existência de um local ativo na superfície do domínio central com um único átomo de zinco ligado. Estudos cinéticos demonstraram a necessidade deste cofator de metal para a atividade da PMI (NOLTMANN, 1968). A *Ca*-PMI mostrou ser inibida por fosfatos de açúcares como o 4-fosfato de eritrose em concentrações milimolares (PROUDFOOT; PAYTON; WELLS, 1994). Além do zinco no local catalítico, demonstrou-se que PMI liga  $Zn^{2+}$  em dois outros locais com menor afinidade (PROUDFOOT et al., 1994). Após a resolução dessa estrutura acima descrita, mais quatro estruturas de MPI contendo  $Zn^{2+}$  foram depositadas, sendo duas com o substrato F6P (PMI de *Salmonella thyphimurium*, depositada em 2009 sob o código PDB 3HLI, a 2.08 Å; PMI de *Pyrococcus furiosus*, depositada em 2006 – código PDB 2GC3, a 2.1 Å) e duas sem (códigos PDB 3H1M de 2.5 Å e 3H1W de 1.94 Å depositadas em 2009, de *S. thyphimurium*).

O papel do cofator de zinco na catálise ainda não está claro. Uma vez que, a elucidação dos mecanismos catalíticos e regulatórios da PMI é essencial para o planejamento racional de fármacos, várias outras técnicas, como estudos de inibição, (ROUX et al., 2004; WELLS et al., 1995; WELLS; PAYTON; PROUDFOOT, 1994) rótulos específicos dos sítios (COULIN et al., 1993; PAPOUTSOPOULOU; KYRIAKIDIS, 1997; WELLS; SCULLY; MAGNENAT, 1994) e mutações direcionadas ao sítio (WELLS et al., 1995; WELLS; SCULLY; MAGNENAT, 1994) têm sido empregadas para gerar informações complementares.

Em *Xanthomonas*, a fosfomanose isomerase está relacionada com a produção de goma xantana, a qual é um mecanismo de defesa da bactéria para evitar sua desidratação em ambientes nocivos (PAPOUTSOPOULOU; KYRIAKIDIS, 1997). Em estudos realizados anteriormente no LBBMA-UFSCar por Carnielli e colaboradores (CARNIELLI et al., 2017), a fosfomanose isomerase foi encontrada na superfície da bactéria *Xanthomonas citri subsp. citri* em condições infecciosas (Figura 4), o que demonstra o caráter inédito e inovador desta pesquisa em identificar novos inibidores para essa enzima.

**Figura 4** - Alterações de proteínas na superfície de *Xanthomonas citri* subsp. *citri* por análise proteômica em células cultivadas sob condições infecciosas e não-infecciosas e suas possíveis funções.



Fonte: Carnielli et al., 2017.

#### 1.4 Planejamento computacional de substâncias bioativas

A descoberta de novas tecnologias de informação e novas ferramentas computacionais, durante os últimos anos, aumentaram a quantidade de dados em relação a moléculas, estruturas, proteínas, alvos e mecanismos de ação de fármacos (WERMUTH, 2008).

Devido ao avanço de técnicas como cristalografia e ressonância magnética nuclear (RMN), diversas estruturas estão disponíveis em bancos de dados e podem ser aplicadas em estratégias de planejamento computacional de fármacos (JORGENSEN, 2004, 2009).

Atualmente, o planejamento racional *in silico* é a metodologia mais eficiente e menos dispendiosa para o desenvolvimento de novos compostos, capaz de contribuir em todos os estágios do processo, desde a descoberta de protótipos (também conhecidos como “compostos de partida” ou “lead compounds”), sua otimização (com respeito à afinidade, especificidade, eficácia e efeitos colaterais), até a elaboração de compostos candidatos a testes clínicos (OPREA, 2005). Esta metodologia é baseada na inibição ou estimulação da atividade biológica de macromoléculas, tais como proteínas ou ácidos nucleicos (DNA ou RNA) associados a diferentes doenças. A informação estrutural do bioreceptor (e ligantes) permite a descoberta e direciona a síntese de compostos com complementaridade estérica, hidrofóbica e eletrostática ao seu sítio de ligação, os quais podem vir a se tornar potentes fármacos. Essa abordagem, em



sua essência, caracteriza o planejamento racional de fármacos por vias computacionais. O que torna ainda mais atrativo tal *approach*, quando utilizado com proteínas, é o conhecimento de que 78% dos fármacos atuais têm como alvo receptor esse tipo de biomolécula (OPREA, 2005).

Desde a concepção do alvo biológico até a descoberta de um novo fármaco, um processo que pode levar em média 11 anos ou mais, a química computacional vem se tornando, a cada dia, indispensável no planejamento racional de novos fármacos, já com inúmeros casos de sucesso envolvendo o emprego de suas metodologias (OPREA, 2005). Para Christopher Watson, da Novartis Institutes for BioMedical Research, o processo de descoberta e desenvolvimento de novos fármacos é hoje totalmente dependente das metodologias computacionais envolvidas (PEITSCH, 2004).

Em química computacional (ou “químioinformática”), genericamente citada como modelagem molecular, destaca-se o *docking* molecular como um dos métodos mais empregados. Com essa metodologia, são investigadas as possíveis orientações que determinada molécula assume no interior do sítio ligante de um bioreceptor, ou simplesmente entre duas moléculas, tal como é o caso da interação entre proteína-proteína ou proteína-DNA, caracterizando o “*docking* macromolecular”.

Os métodos de *docking*, em geral, envolvem uma função de energia contendo parâmetros eletrostáticos, de van der Waals, de ligações de hidrogênio e, algumas vezes, hidrofóbicos, baseados em modelos matemáticos que predizem as melhores orientações do ligante, segundo uma lista de escores de energia. O “Problema do *docking*” é assim centrado na geração e avaliação (quase sempre por intervenção do usuário) de plausíveis estruturas de complexos intermoleculares (CODDING, 1998). As mais recentes versões dos *programas* de maior sucesso em *docking*, **GOLD** (NISSINK et al., 2002) e **GLIDE** (FRIESNER et al., 2004), consideram a flexibilidade do ligante e, também, de algumas cadeias laterais do sítio receptor.

O método “*screening* (ou triagem) virtual” é utilizado para a seleção de novos protótipos, e a literatura vem reportando, há alguns anos, diversos casos de sucesso com o uso dessa sistemática, tal como a descoberta de isoflavonóides como inibidores não-esteroidais de 5 $\alpha$ -redutase, utilizando *constraint* farmacofórica (BRENK et al., 2003; CHEN; LIU; GILSON, 2001). Nessa era pós-genômica, o *screening* virtual complementa os conhecidos métodos experimentais de *screening* em larga escala no processo de descoberta de novos protótipos (KUNZ, 1992). O sucesso do *screening* virtual, e em geral de *docking*, depende do conhecimento de detalhes estruturais finos do sítio receptor (CARLSON; MASUKAWA; MCCAMMON, 1999).

Para o desenvolvimento de um novo fármaco também é possível hoje estimar propriedades farmacocinéticas e farmacodinâmicas, bem como propriedades “*drug-like*” ou “*lead-like*” de diferentes compostos, selecionando, durante as diversas etapas da modelagem, somente aqueles com potencial de serem fármacos. Como exemplo, a ‘Regra dos 5’ (RO5), de Lipinski, que é uma regra que os fármacos oralmente ativos e em geral seguem: peso molecular menor ou igual a 500,  $\log P$  menor ou igual a 5, número de grupos doadores de ligações de hidrogênio menor ou igual a 5 e número de grupos aceptores de ligações de hidrogênio menor ou igual a 10 (LIPINSKI; HOPKINS, 2004). Citações em *CAS SciFinder* do artigo original da RO5, de 1997, excederam 1000 somente no ano de 2004, e continuam crescendo (LIPINSKI, 2004). Uma variação dessa regra é aplicável a protótipos e, além disso, novas regras empíricas vêm sendo reportadas, tais como o número de ligações rotacionáveis em fármacos ser menor do que 8 ou, ainda, a área superficial ser menor ou igual a  $140 \text{ \AA}^2$  (OPREA, 2005).

Sendo assim, propriedades tais como absorção, distribuição, metabolismo, excreção e toxicidade (ADME/Tox) podem ser preditas a partir do *screening* em bancos de dados contendo essas informações, as quais são computadas para uma grande variedade de compostos (BAJORATH, 2004; CARLSON; MASUKAWA; MCCAMMON, 1999). Entre os *programas* mais utilizados para esses fins, destacam-se o **ADMET Predictor** (ADMET), **Meteor** (metabolismo) e **Derek** (toxicidade).

Dessa forma, utilizamos tais técnicas computacionais para o estudo de novos inibidores capazes de atuar na enzima PMI de XAC, com o objetivo de controlar o cancro cítrico.

## 2. OBJETIVOS

O presente estudo tem como objetivo o planejamento de novos potenciais inibidores de PMI utilizando metodologias computacionais, adquirir comercialmente alguns desses compostos selecionados e testá-los *in vitro*.

As metodologias computacionais incluem, especificamente, experimentos de *screening* (ou triagem) virtual em diferentes bases de dados de compostos comerciais contendo propriedades de fármacos, utilizando variadas estratégias de Química Medicinal Computacional.

Os objetivos específicos do trabalho incluem ainda:

- Desenvolvimento, validação e avaliação de um modelo proposto para a PMI de XAC através de técnicas de modelagem de proteínas por homologia estrutural;
- Desenvolvimento, validação e avaliação de um modelo proposto para a PMI humana através de técnicas de modelagem de proteínas por homologia estrutural visando sua diferença em relação ao sítio ativo com a PMI de XAC;
- Tratamento de 10 bases de dados e geração da biblioteca de confôrmeros utilizados nos estudos de triagem virtual;
- Estudos de triagem virtual baseado no ligante (similaridade por forma e eletrostática) para a investigação de compostos com propriedades inibidoras à PMI;
- Estudos do perfil farmacocinético e toxicológico dos compostos identificados, buscando selecionar os que apresentam um perfil mais adequado a novos potenciais defensivos agrícolas e que não cause interferência em humanos;
- Realizar ensaios biológicos junto aos compostos selecionados e otimizados, buscando verificar os níveis de atividade dos mesmos, tentando racionalizar os resultados por meio das relações estrutura-atividade assim observados.

### 3. MATERIAIS E MÉTODOS

Para a pesquisa de novos inibidores de PMI, foi usada a técnica de triagem virtual por forma e similaridade, além de simulações de *docking* e análises de propriedades farmacocinéticas e toxicológicas (ADME/Tox) e seus modos de interações foram analisados através de uma inspeção visual, obtendo por fim, os compostos mais promissores. Essas moléculas selecionadas foram adquiridas para serem testadas em ensaios de atividade *in vitro*.

Como não existem estruturas depositadas no PDB (*Protein Data Bank*) para a PMI de XAC, um modelo tridimensional foi desenvolvido utilizando o método de “modelagem molecular por homologia estrutural”.

#### 3.1. Construção do modelo de fosfomanose isomerase para *Xanthomonas citri* subsp. *citri*

O primeiro passo para esse estudo foi a construção de um modelo da proteína fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*, utilizando a técnica de modelagem por homologia estrutural, já que tal estrutura ainda não existe nos bancos de dados de estruturas de proteínas como o PDB. Os estudos de homologia estrutural são baseados em três passos: (1) alinhamento (a etapa crucial do processo), (2) a construção do modelo e (3) a validação do modelo.

##### 3.1.1 Pesquisa e Alinhamento de sequências homólogas

Uma busca por estruturas com sequências homólogas à de interesse (PMI de XAC), contidas no PDB, foi realizada previamente utilizando o algoritmo **BLASTp**, um algoritmo pertencente à família **BLAST** disponibilizado pelo *National Center for Biotechnology* e que realiza a comparação entre sequências proteicas (ALTSCHUL et al., 1990).

Quatro proteínas foram selecionadas para a geração do modelo, sendo uma com identidade sequencial de 29% e outras três com identidade sequencial de 38%. A estrutura de menor identidade (PDB ID: 1H5R) foi utilizada devido ao fato de possuir um ligante em seu sítio ativo e, dessa forma, ser possível observar os dobramentos ocasionados pela formação do complexo ligante receptor.

##### 3.1.2 Alinhamento das sequências

Após a obtenção das sequências do PDB, foi realizado o alinhamento dessas sequências, para o qual foi utilizado o servidor *online* **Clustal Omega** (SIEVERS et al., 2011).

Esse servidor é disponibilizado pela European Molecular Biology Laboratory – European Bioinformatics Institute (EMBL-EBO).

O **Clustal Omega** tem cinco etapas principais para gerar o alinhamento de múltiplas sequências. O primeiro passo consiste em produzir um alinhamento aos pares (*pairwise*) para, em seguida, serem agrupados (BLACKSHIELDS et al., 2010). A terceira etapa é a construção de uma árvore guia, em que os dois grupos mais próximos são combinados e repetidos até que a árvore final possa ser avaliada. Então, na etapa final, o alinhamento de múltiplas sequências é, de fato, produzido (EDDY, 2011).

Um dado importante analisado durante o alinhamento das sequências diz respeito a quais alterações de resíduos de aminoácidos essas proteínas sofreram ao longo da Evolução, estimando que a função deveria ser preservada nos trechos onde a identidade (ou mesmo a similaridade) sequencial deveria ser maior.

### 3.1.3 Construção e avaliação da qualidade do modelo

O programa **MODELLER** foi utilizado para a construção do modelo da PMI e trata-se de um algoritmo automatizado utilizado para a modelagem de estrutura proteica através de comparação com estruturas tridimensionais resolvidas. Esse programa utiliza pelo menos uma estrutura de entrada como referência (ou molde) e, a partir de outros alinhamentos fornecidos, gera um modelo tridimensional homólogo (FISER; ŠALI, 2003).

Após a construção do modelo tridimensional, o próximo passo consiste na avaliação da qualidade do modelo gerado. Para a etapa de avaliação foram utilizados os servidores *online* **WHATIF** (VRIEND, 1990) para a validação dos contatos atômicos e **SAVE v5.0**, que contém os programas **PROCHECK**, **VERIFY3D**, **ERRAT** e **PROVE**.

O servidor *online* **WHATIF** analisa contatos atômicos, entre maus contatos e contatos favoráveis (como interações intermoleculares), gerando pontuações para cada resíduo. O **PROCHECK** (LASKOWSKI et al., 1993) permite uma comparação entre os resíduos do modelo e estruturas de proteínas de alta-resolução cristalográfica. Através desse programa foram gerados: o Diagrama de Ramachandran, os gráficos de planaridade das ligações peptídicas, de distorção tetraédrica dos carbonos alfa, de energias das ligações de hidrogênio, de medidas das interações não-ligantes (van der Waals, ligações de hidrogênio, interações eletrostáticas) e General Factor (G-Factor), o qual, este último, consiste em uma média da estereoquímica geral do modelo.

O **VERIFY 3D** (LUTHY; BOWIE; EISENBERG, 1992) analisa a compatibilidade entre o modelo 3D e a sua sequência de aminoácido, baseado na localização (alfa, beta, loop) e uma sequência de ambientes químicos (polar, apolar), utilizando uma matriz proposta por Bowie e colaboradores (BOWIE; LTCY; EISENBERG, 1990). O **ERRAT** (COLOVOS; YEATES, 1993) analisa as interações não covalentes entre átomos diferentes no modelo em comparação com estruturas altamente refinadas. Por fim, o **PROVE** (PONTIUS; RICHELLE; WODAK, 1996) utiliza um algoritmo que calcula o volume dos átomos presentes nas macromoléculas, assumindo-os como esferas rígidas; e gera um desvio estatístico denominado Z-score. Esse dado estatístico provém da comparação do modelo com estruturas refinadas de alta resolução depositadas no PDB.

#### 3.1.4. Validação do modelo de PMI gerado para simulações de *docking*

Após a criação do modelo, foi realizado um experimento para a validação do mesmo, utilizando os programas **GOLD** (SUITE, 2015) e **GLIDE** (FRIESNER et al., 2004). Definiu-se, a partir do substrato da estrutura cristalográfica original (1H5R), a melhor região do sítio ativo na qual seriam realizados os cálculos. Em seguida foram feitas simulações de *docking* com o substrato da 1H5R para definir qual seria o melhor programa para os próximos estudos. Escolhido o programa, experimentos de *docking* foram realizados com a estrutura de PMI, utilizando inibidores reportados no *web servidor* **BindingDB** (CHEN; LIU; GILSON, 2001), um repositório de estruturas e dados de atividade focado na interação entre proteínas e moléculas com potenciais atividades terapêuticas.

#### *3.2 Triagem virtual de novos inibidores da fosfomanose isomerase*

Acessando o *web servidor* **BindingDB**, um total de 959 hits são descritos como inibidores de PMI, sendo que o mais potente possui  $IC_{50} = 4 \mu M$ . Sendo assim, torna-se desejável a descoberta e o desenvolvimento de novas classes de compostos e que possam inibir mais efetivamente a enzima aqui investigada. O objetivo é realizar experimentos de triagem virtual (por *docking* e similaridade) com a PMI em diferentes bases de dados de compostos disponíveis comercialmente. As bases de dados de compostos incluem: *ZINC* (subcoleções *Natural Stock*, *Drug-database* e *Drug-like*), *BindingDB* subdivisão *Drugs FDA*, ChemBridge (subcoleções *DIVERSet<sup>TM</sup>-EXPEXPRESS-Pick<sup>TM</sup> Collection(DIVERSet<sup>TM</sup>-EXP)* e *DIVERSet Core Library (DIVERSet<sup>TM</sup>-CL)*), *Maybridge* subdivisão *Screening Collection*,

*Princeton* e *IBS* (subdivisão *Natural* e *Synthetic*) e na tabela abaixo (Tabela 1) é mostrado o número de compostos presentes em cada base de dados utilizada.

**Tabela 1** - Relação do número de compostos presentes em cada base de dados utilizada.

Base de dados	Número de compostos	Referências
<i>ZINC</i> subdivisão <i>Drug-database</i>	2924	(IRWIN; SHOICHET, 2005)
<i>BindingDB</i> subdivisão <i>Drugs</i> <i>FDA</i>	6144	(CHEN; LIU; GILSON, 2001)
<i>Chembridge</i> subdivisão <i>DIVERSet-EXP</i>	50000	(CHEMBRIDGE, [s.d.])
<i>Maybridge</i> subdivisão <i>Screening Collection</i>	53000	(MAYBRIDGE, 2004)
<i>Chembridge</i> subdivisão <i>DIVERSet CL</i>	60000	(CHEMBRIDGE, [s.d.])
<i>IBS</i> subdivisão <i>Natural</i>	68497	
<i>IBS</i> subdivisão <i>Synthetic</i>	483510	
<i>ZINC</i> subdivisão <i>Natural Stock</i>	450000	(IRWIN; SHOICHET, 2005)
<i>Princeton</i>	1300000	
<i>ZINC</i> subdivisão <i>Drug-Like</i>	17900742	
<b>TOTAL</b>	20374817	

Fonte: o autor

Algumas dessas bases contêm os compostos em formato *spatial data format* (sdf), os quais já foram reescritos em formato mais apropriado, mol2, contendo cargas atômicas, tipos de átomos, ordens de ligação e estruturas tridimensionais bem definidas e mais adequadas para as simulações de triagem virtual e propósitos gerais de modelagem molecular. Também foram geradas novas bases/bibliotecas de confôrmeros, a partir dessas, utilizando o *programa OMEGA* (BOSTRÖM; GREENWOOD; GOTTFRIES, 2003).

### 3.2.1. Geração de confôrmeros (**OMEGA**)

Para a geração de confôrmeros, o programa **OMEGA** da empresa **OpenEye** foi utilizado. Esse programa realiza a geração de confôrmeros em três dimensões com baixo

desprendimento computacional e um alto efeito no alcance da conformação bioativa entre os confôrmeros gerados.

No presente estudo, os parâmetros utilizados no **OMEGA** para a geração de confôrmeros foram: 300/molécula, com uma energia de tensão de até 9 kcal/mol e um desvio médio quadrático de 0.6 Å, segundo protocolo otimizado por Carlos H.T.P. Silva (DA SILVA; TAFT, 2017).

### 3.2.2. Seleção de inibidores baseados em ligantes (ROCS e EON)

Após a preparação das bases de dados, um segundo tipo de triagem virtual foi ainda realizado, o qual é baseado em diferentes tipos de similaridade: de forma e eletrostática. Ambos os programas utilizados nessa etapa, **ROCS** e **EON**, da **OpenEye**, tem revelado compostos promissores para o tratamento de várias doenças (NAYLOR et al., 2009).

O programa **ROCS** realiza uma triagem virtual por forma, através da comparação entre os volumes e conformação de dois compostos, a molécula de referência e as moléculas presentes em bancos de dados; e o programa **EON** realiza a comparação e classificação dos compostos bases em seus mapas de superfície eletrostática. Os parâmetros adotados para ambos programas não foram alterados.

### 3.2.3. Seleção de inibidores baseados em estrutura (GOLD)

Os programas de *docking* analisam a afinidade ligante-proteína e para essa etapa do estudo foi utilizado o programa **GOLD**.

A localização do centróide utilizado para a PMI de *XAC* foi  $x = 23.977$ ,  $y = 15.659$  e  $z = 43.886$  com uma esfera de 8 Å. O tamanho dessa esfera foi modificado para algumas bases de dados devido à problemas encontrados para alguns ligantes e, nesse caso, foi utilizado uma esfera de 6 Å e 60 de fitness. Para os estudos executados na PMI de *Homo sapiens*, para a validação do modelo gerado, o centróide utilizado foi  $x = 20.21$ ,  $y = 16.63$  e  $z = 10.00$  e com uma esfera de 7 Å.

### 3.2.4 Predição das propriedades farmacocinéticas e toxicológicas (QikProp e DEREK)

Após os experimentos de triagem virtual realizados, somente as moléculas que passaram pelo crivo de propriedades adequadas a um potencial defensivo agrícola avançaram para a próxima etapa, avaliando-se, principalmente, a potencial toxicidade dos compostos frente ao



organismo humano. As moléculas foram avaliadas com respeito às suas propriedades ADME/Tox, utilizando-se os programas **QikProp** (SCHRODINGER, [s.d.]) (da empresa Schroedinger) e **DEREK** (MARCHANT; BRIGGS; LONG, 2008) (da empresa Lhasa).

Os compostos sobreviventes à predição de propriedades toxicológicas, foram submetidos à análises farmacocinéticas realizadas pelo **QikProp**, os descritores analisados foram: log CACO2, log MDCK, porcentagem de absorção oral e qualidade do modelo de absorção oral humana. Foram rejeitados todos os compostos que tiveram valores de log CACO2 e log MDCK maiores ou iguais a 500, potencial de absorção oral maior ou igual a 80% e qualidade do modelo de absorção humana igual a alta.

Na análise toxicológica realizada pelo **DEREK**, os compostos rejeitados foram aqueles que apresentaram algum alerta de toxicidade segundo *Custom Prediction* e *Lhasa Prediction*.

### 3.3. Visualização Gráfica

Os programas utilizados para a visualização gráfica foram o **Discovery Studio** (DASSAULT SYSTÈMES BIOVIA, 2015) da empresa *Biovia* bem como o **Maestro** (SCHRODINGER, 2018), da empresa *Schroedinger*.

O **Discovery Studio** foi empregado na visualização da sobreposição dos modelos obtidos pela modelagem de proteínas por homologia estrutural. Já o **Maestro** foi utilizado na análise visual dos resultados obtidos a partir dos estudos de simulações de *docking* e, conseqüentemente, na seleção das moléculas finais.

### 3.4. Ensaios de atividade *in vitro* com a enzima PMI de *Xanthomonas citri* subsp. *citri*.

A enzima recombinante de XAC foi produzida por expressão heteróloga em sistema bacteriano e foi cedida por Alexandrino *et al* (dados não publicados) (LBBMA/ UFSCar- São Carlos).

Durante o ensaio foi utilizado um sal dissódico hidratado de D-manose-6-fosfato (M6P) da Sigma (cod. M6876-10) com pureza maior ou igual a 97% e um sal dissódico hidratado de D-frutose-6-fosfato (F6P) da Sigma (cod. F3627-100) com pureza maior ou igual a 98%. Ambos sais apresentam uma massa molar (MM) de 304,10 g/mol e foram diluídos em água, de modo a obter uma concentração final de 0,3 M.

Uma solução estoque (em água) de 0,1 M do inibidor **DL03874928** (MM = 286,28 g/mol) e diluições subseqüentes a partir dessa solução-estoque foram feitas de modo a obter concentrações de 1 mM, 100 µM e 10 µM.

O reagente de Seliwanoff utilizado para a detecção colorimétrica foi feito no mesmo dia dos ensaios e consistiu em 0,025 g de resorcinol, e 50 mL de ácido clorídrico concentrado para um volume final de 50 mL.

#### 3.4.1 Método de avaliação da atividade enzimática

Como dito anteriormente, a fosfomanose isomerase é uma enzima que atua na interconversão da frutose-6-fosfato (F6P) e da manose-6-fosfato (M6P). Com isso, a atividade enzimática foi realizada por meio de um teste de detecção de frutose baseado em um método colorimétrico (SHAHIDULLAH, M; KHORASANI, 1972) com o reagente de Seliwanoff.

O procedimento consistiu na adição de 5 µL de uma solução contendo 0,3 M de M6P (0,456 mg) em 95 µL de uma solução da enzima recombinante PMI de XAC em tampão (50 mM Tris-HCl pH 8,0, 10 mM NaCl) para um volume final de 200 µL. Após incubação da mistura por 1h a 30°C foram adicionados 3 mL do reagente de Seliwanoff e a mistura foi mantida a 100°C por 5 minutos.

Os ensaios foram feitos como um teste exploratório a 25°C e a leitura de absorbância foi realizada em comprimento de onda de 415 nm no leitor de microplacas *iMark Microplate Absorbance Reader* (Bio-Rad). O experimento definitivo já está em andamento.

#### 3.4.2 Avaliação *in vitro* da inibição da fosfomanose isomerase pelos compostos selecionados

A avaliação da inibição da PMI pelo composto **DL03874928**, um dos inibidores mais promissores pois apresenta o maior número de interações com os aminoácidos do modelo gerado, foi realizada com a mesma metodologia do passo anterior.

Em cada um dos tubos foram adicionados 95 µL de enzima e 100 µL do inibidor nas concentrações de 10 µM, 100 µM e 0,1 M, de modo a obter as concentrações finais de 5 µM, 50 µM e 50 mM (Tabela 2). Essa solução foi incubada à temperatura ambiente por 10 minutos. Em seguida, foram adicionados aos tubos 5 µL da solução M6P 0,3 M; sendo que essa solução formada por inibidor, enzima e substrato foi incubada por 1h à 30°C.

Passado o tempo de incubação, cada ensaio recebeu 3 mL de Reagente de Seliwanoff e foi levado à fervura por 5 minutos. As absorbâncias das soluções foram medidas com a ajuda do leitor de microplacas *iMark Microplate Absorbance Reader* (Bio-Rad) sob comprimento de onda de 415 nm.

**Tabela 2** - Dados referentes aos ensaios *in vitro* realizados para a fosfomanose isomerase recombinante de *Xanthomonas citri* subsp. *citri*.

REAGENTES	Volume ( $\mu\text{L}$ )					
	Controle	Controle	Reação com M6P e Concentrações finais do inibidor			
	com F6P	com M6P	sem inibidor	5 $\mu\text{M}$	50 $\mu\text{M}$	50 mM
<b>M6P 0,3M</b>	0	5	5	5	5	5
<b>F6P 0,3M</b>	5	0	0	0	0	0
<b>Enzima</b>	0	0	95	95	95	95
<b>Inibidor 10 <math>\mu\text{M}</math></b>	0	0	0	100	0	0
<b>Inibidor 100 <math>\mu\text{M}</math></b>	0	0	0	0	100	0
<b>Inibidor 0,1 M</b>	0	0	0	0	0	100
<b>Tampão</b> (50 mM Tris-HCl pH 8,0, 10 mM NaCl)	195	195	100	0	0	0

## 4. RESULTADOS E DISCUSSÃO

A seguir serão apresentadas e discutidas as etapas para construção e validação dos modelos de fosfomanose isomerase para *Xanthomonas citri* subsp. *citri*, e *Homo sapiens*. Também serão apresentados os resultados da triagem virtual para os inibidores de XAC e avaliação *in vitro* para um dos compostos selecionados.

### 4.1 Construção do modelo da fosfomanose isomerase (PMI) para *Xanthomonas citri* subsp. *citri* (XAC)

A criação de um modelo tridimensional para uma proteína através da homologia estrutural consiste na comparação entre duas ou mais sequências de proteínas similares. Para que esse método seja aplicado é necessário que pelo menos três situações ocorram: (a) uma estrutura proteica seja experimentalmente resolvida; (b) compartilhar um valor de identidade sequencial acima de 25% com a proteína desejada; e (c) correto alinhamento entre as sequências envolvidas (FOMBY; CHERLIN, 2011; MARTÍ-RENOM et al., 2000; SCHWEDE et al., 2003).

A criação de um modelo é possível devido à relativa conservação que ocorre no enovelamento de uma determinada família de proteínas ao longo da evolução. Assim, caso a estrutura primária de duas proteínas sejam similares, o modelo tridimensional terá grandes chances de ser similar também. A criação de um modelo por homologia estrutural é constituída por: pesquisa por sequências homólogas, alinhamento das sequências, construção e avaliação do modelo (MARTÍ-RENOM et al., 2000; SANCHEZ; SAIL, [s.d.]).

Nos próximos tópicos são apresentadas as 4 etapas de predição de modelagem de proteínas por homologia estrutural aplicadas para construção do modelo de PMI de XAC.

#### 4.1.1 Etapa de identificação da sequência de aprovação

A etapa inicial do processo de predição de estruturas tridimensionais de proteínas é a identificação de estruturas de proteínas homologadas por sequências em busca que possuam dados cristalográficos, ou seja, sua estrutura resolvida e relatada em um banco de dados de estruturas cristalográficas, como o PDB. Para maior confiabilidade do modelo gerado, é desejável que as identidades sequenciais estejam acima de 25% em relação à sequência de proteínas em estudo (RODRIGUEZ; VRIEND, 1997).

Para a construção do modelo da PMI inicialmente obtivemos o sequenciamento de aminoácidos da mesma através do banco de dados GenBank (AAM38423.1) (“NCBI-National

Center for Biotechnology Information”, 2018). Em seguida, utilizando o algoritmo do servidor online **BLASTp** foi possível identificarmos proteínas com estruturas cristalinas disponíveis no PDB e que detêm identidades sequenciais acima de 25% compartilhadas com a PMI de interesse.

O algoritmo **BLAST** (*Basic Local Alignment Search Tool*) disponibilizado pelo NCBI (*National Center for Biotechnology*) detecta similaridades entre sequências nucleotídicas ou de proteínas. Os programas da família **BLAST** foram feitos para comparar: a) sequências de consultas de DNA; b) proteínas com bancos de dados de DNA; e c) proteínas em qualquer combinação (ALTSCHUL et al., 1990).

Uma das variantes dos algoritmos da família **BLAST** é o **BLASTp**, utilizado para comparações entre sequências de proteínas, através de uma busca por 3 aminoácidos semelhantes (*hit*) da sequência-modelo em outras sequências depositadas em diversos bancos de dados. Os alinhamentos desses *hits* recebem uma pontuação chamada de *score* e, depois, a análise é expandida nos dois lados para o restante da sequência-modelo com o intuito de se obter um maior *score*.

Uma outra pontuação analisada pelo **BLAST** é o *e-value*, com o intuito de determinar a aleatoriedade com que uma sequência homóloga foi obtida. Essa pontuação indica a probabilidade de se encontrar um alinhamento em outra sequência, aleatoriamente. Por isso, quanto menor o valor de *e-value*, mais confiável é o resultado obtido.

Por fim, um outro resultado importante fornecido pelo **BLAST** é a porcentagem de identidade sequencial, a qual é responsável por mostrar o grau de correspondência entre duas subsequências. Nesse caso, valores entre 18-25% indicam uma potencial semelhança estrutural, enquanto valores iguais ou maiores a 25% indicam adicionalmente uma potencial semelhança de função

Portanto, neste trabalho, o algoritmo **BLASTp** ajudou a detectar 3 estruturas homólogas à PMI (PDB ID 2CU2, 2X5S e 2QH5), com percentual de identidade sequencial de 36% e 38%, além de altas resoluções cristalográficas (de 1,9 a 2,35 Å), tidas como adequadas para a modelagem por homologia (ou comparativa). Foi detectada também uma quarta estrutura, sendo essa a única contendo ligante (o substrato) no sítio ativo (PDB ID 1H5R) e que apresenta 29% de identidade (Figura 5) (Tabela 3).

**Figura 5** - Resultado obtido pelo **BLASTp** na procura de proteínas similares a fosfomanose isomerase de *Xanthomonas citri* subsp. *citri* (PDB ID: 2X5S, 2QH5, 2CU2 e 1H5R)

Description	Max score	Total score	Query cover	E value	Ident	Accession
<a href="#">Chain A, Crystal Structure Of Mannose-6-phosphate Isomerase From Helicobacter Pylori</a>	207	207	97%	3e-65	40%	<a href="#">2QH5_A</a>
<a href="#">Chain A, Crystal Structure Of T. Maritima Gdp-Mannose Pyrophosphorylase In Apo State</a>	177	177	97%	1e-53	40%	<a href="#">2X5S_A</a>
<a href="#">Chain A, Crystal Structure Of Mannose-1-Phosphate Geranyltransferase From Thermus Thermophilus Hb8</a>	171	171	97%	6e-51	38%	<a href="#">2CU2_A</a>
<a href="#">Chain A, Crystal structure of glucose 1-phosphate thymidyltransferase from Aneurimibacillus thermoaerophilus</a>	41.2	41.2	67%	8e-04	29%	<a href="#">4H00_A</a>
<a href="#">Chain A, Thymidyltransferase Complexed With Tmp</a>	38.1	38.1	69%	0.007	29%	<a href="#">1H5S_A</a>
<a href="#">Chain D, Thymidyltransferase Complexed With Tmp</a>	38.1	38.1	67%	0.007	29%	<a href="#">1H5S_D</a>
<a href="#">Chain A, Crystal Structure Of Glucose-1-phosphate Thymidyltransferase RmlA Complex With Dtdp</a>	37.7	37.7	67%	0.009	25%	<a href="#">1LVW_A</a>
<a href="#">Chain A, Thymidyltransferase Complexed With Thymidylidiphosphate- Glucose</a>	37.0	37.0	67%	0.018	29%	<a href="#">1H5T_A</a>
<a href="#">Chain B, Thymidyltransferase Complexed With Thimidine And Glucose- 1-Phosphate</a>	36.6	36.6	67%	0.019	29%	<a href="#">1H5R_B</a>
<a href="#">Chain A, Crystal Structure Of Glucose-1-phosphate Thymidyltransferase From Burkholderia Vietnamensis In Complex With 2'-deoxyuridine-5'- Monophosphate And 2'-deoxy-thy</a>	35.8	35.8	11%	0.045	58%	<a href="#">5IFY_A</a>
<a href="#">Chain A, Mycobacterium tuberculosis RmlA in complex with dTDP-glucose</a>	35.4	35.4	94%	0.055	26%	<a href="#">6B5E_A</a>
<a href="#">Chain A, L89T VARIANT OF S. ENTERICA RmlA</a>	34.7	34.7	67%	0.099	28%	<a href="#">1MP3_A</a>

Fonte: (“NCBI-National Center for Biotechnology Information”, 2018)

As estruturas selecionadas (PDB ID 2QH5, 2X5S, 2CU2 e 1H5R) são as seguintes enzimas: manose-6-fosfato isomerase, GDP-manose pirofosforilase, glicose-1-fosfato timidiltransferase e uma timidililtransferase.

A manose-6-fosfato isomerase e a GDP-manose pirofosforilase realizam ambas atividades desempenhadas pela PMI de *XAC*, isto é, a isomerização entre manose-6-fosfato e frutose-6-fosfato e a transferência de manose-1-fosfato (M1P) para a guanosina trifosfato (GTP). Por fim, a glicose-1-fosfato timidiltransferase e a timidililtransferase realizam a transferência de timidila produzindo, por exemplo, a DTDP-glicose.

Devido ao fato de o modelo a ser gerado ser de interesse para os estudos de descoberta de novas moléculas baseados em estruturas, a escolha da estrutura cristalográfica para sua construção levou em consideração a presença de um ligante complexado no sítio da proteína. Assim, a escolha da proteína homóloga, usada para prever uma estrutura PMI *XAC*, foi a cadeia B da proteína com código PDB 1H5R (*Escherichia coli*, cepa K12 timidililtransferase) (Tabela 3). Como todas as estruturas apresentam uma identidade acima de 25%, elas compartilham com a estrutura de interesse uma potencial semelhança estrutural e de função.

**Tabela 3** - Resultados de identidade e *e-value* de sequências de proteínas obtidas no servidor online **BLAST** usando o algoritmo para proteínas (**BLASTp**) na triagem de proteínas homologadas a fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*.

PDB ID	Identidade	E-value
<b>2QH5_A</b>	38%	$7 \times 10^{-63}$
<b>2X5S_A</b>	38%	$5 \times 10^{-64}$
<b>2CU2_A</b>	38%	$2 \times 10^{-59}$
<b>1H5R_B</b>	29%	0,033

Fonte: o autor.



XANTHO	EKPDLATAQSYLASGEYYWNSGMFLFRASRYLEELRKFHPAIADACQKAWENGKRDADF-	231
	*: .: . : : : : . : * : :	
1H5R	-----TDINRIYLEQGRLSVAMMGRGYAWLDTGTHQSLIEASNFIATIE	244
2CU2	----ASLEEVYAGLPKISIDYG-VMEKAERVVRVVLGRFPWDDVGNWRALERVFSQDPHEN	273
2X5S	PRNFEELKKAYEKVPSISVDYA-VMEKSKKVRVVKADFEWSDLGNWSSVREIEGYTEES-	277
2QH5	KKIARLSEKSMQDLEDMSIDIA-LMQQSHKIKMVELNAKWSDLGNFNALFEEAANEPKEN	286
XANTHO	---TRLDKDAFAASPSDSIDYA-VMEKTADAVVPLDAGWNDVGSWSSLLDVSNDQAQGN	287
	.: : : : : * * * . : :	
1H5R	ER--QGL--KVSCPEEIAF-RKGFIDVEQVRKLA-----VPLIKNNYGQYLYKMTKDS	292
2CU2	VVLGEGRHVALDTFGCVVYADRGVVATLGVSGLVVAKVGDVLLVVPKDWAREVREVVKRL	333
2X5S	---DEVIL--VSDRVFVKTHNKPIAVVGLSDVIVIDTPNGILICKEEYAKVREVVKKL	332
2QH5	VSLNQTPV-----FAK-ESEEGHHHHHH-----	308
XANTHO	AHGDVIQ--LDCQNTYAY-GSRLIAMVGLDVDVVVETPDAVLVGHDRDRIQEVKDVVSQI	344
	: .	
1H5R	N-----	293
2CU2	EAQE-----	337
2X5S	FRTS-----	336
2QH5	-----	308
XANTHO	KTAGRSEATWHRKVYRPWGAYDSIDMGQRHQVKRITVKPGAVLSLQMHHHRAEHWIVVSG	404
	-----	293
	-----	337
	-----	336
	-----	308
	TAEVTRGEEVLLLTENQSTYIPLGVTHRLRNPGKLPLELIEVQSGSYLGEDDIVRFEDTY	464
1H5R	---	293
2CU2	---	337
2X5S	---	336
2QH5	---	308
XANTHO	GRA	467

A simbologia de asterisco (\*) representa a presença de um único resíduo totalmente conservado, dois pontos (:) indica a conservação entre grupos de propriedades fortemente semelhantes e um ponto (.) indica a conservação entre grupos de propriedades fracamente semelhantes.

Fonte: o autor

Essa sobreposição estrutural nos permitiu visualizar os consensos de todas as estruturas secundárias das proteínas de referência, entre hélices-alfa e fitas de folha-beta. Devido a isso, pudemos analisar se as inserções ou deleções sugeridas pelo **Clustal Omega** estavam de acordo com o esperado, estruturalmente falando. Após a análise dos alinhamentos, juntamente com as sobreposições das estruturas de referência, obteve-se o alinhamento final.

Comparando ambos alinhamentos, aquele obtido pelo **Clustal Omega** e aquele após a etapa de refinamento, pode-se observar algumas diferenças. Foram acrescentados alguns GAPS



entre as estruturas PDB ID 1H5R, 2CU2, 2X5S e 2QH5 devido as sobreposições, porém, o mais evidente ocorre na estrutura PDB ID 2QH5, em que é necessário um GAP entre os aminoácidos glicina (G11) e tirosina (Y12) como indicado na Figura 7.

**Figura 7** - Alinhamento múltiplo final das seqüências de fosfomanose isomerase e as respectivas às estruturas indicadas como homólogas pelo **BLASTp** (PDB ID 1H5R, 2CU2, 2X5S e 2QH5) após o refinamento.

1H5R	-KMRKGIILAGGSGTRLYPVTMA-VSKQLLPIYD-KPMIYYPLSTLMLAGIR--DILLIIST	57
2CU2	-MKTYALVMAGGRGERLWPLSREDRPKFPLPLFEGKTLLEATLERLAPL-VPPERTLLAVR	59
2X5S	--VMKALILAGGSGERFWPLSTPETPKQFLKLFNGKSLMRWTFERVLEE-MDPKDVIVVTH	58
2QH5	-LKIKNILLSG-----YPKQFLKLFDHKSLFELSFKRNASL-VD--ETLIVCN	57
PMI	MSDVLPIILSGGSGTRLWPLSRESYPKQFLPLVGDKSMLOSITWLRAPV-AG-HAPIVVAN	59
	: : : * * * * : : : * * * * : : :	
1H5R	PQDTPRFQQLLGDGSQWGLNL--QYKVPSPDGLAQAFIIGEE---FIG---GDDCALVLG	110
2CU2	RDQEAVAR-PYADG-----I---RLLEPLGRDTAGAVLLGVA---EALKEGAERLLVLP	108
2X5S	KDYVERTKKELPE-----LPDENIAEPMKKNTAPACFIGTKL---AD--DD-EPVLVLP	108
2QH5	EKHYFLALEEIKNEIK--NKS-VGFLLESLSKNTANAIALSALM---SD--KE-DLLIVTPS	110
PMI	EEHRFMAEQLO-QLG--VKPS-AILLEPKGRNTAPATAVAALEATRDG--AD-PLLLVLP	114
	. : . : : : : . : : : :	
1H5R	DNIFYG-HDLPKLMEEAVNK-ES--GATVFAYHVNDPE-RYGVVEFDKN-----GTAISLEE	162
2CU2	DHYVGDDEAYREALATMLEAAEEG-FVVALGLRPTREPTEYGYIRLGPREG-AWYRGEGFVE	168
2X5S	DHRIPTKFKFWKTVKKALDALEKYDGLFTFGIVPTRPETGYGYIEIGEELEEGVHKVAQFRE	170
2QH5	DHLIKDLQAYENAIIKKAIDLAQKG-FLVTFGVSIDKPNTEFGYIESP----NGLD-VKRFIE	175
PMI	DHVIQNEAAFQAAVTLAATAAEQG-KLVTFGIKPTAPETGYGYIKAS--AGAGASAVERFVE	182
	* : . : : : : : : * : * : * : . : : *	
1H5R	KP-----LEP--K-SNYAVTGLYFYDN-DVVQMAKNLKPSAR-----	195
2CU2	KPSYAEALEYIRK-GYVWNGGVFAFAPATMAELFRRHLPSSHE-ALER-LLAGAS-----	220
2X5S	KPDLETAKKFVESGRFLWNSGMFLWKAREFIEEVKVCPSIYENLKDVPDRNFEE-----	225
2QH5	KPSLDKAIETFQKSGGFYFNSGMFVFAQVFLDELKKHAPTILK-GCER-AFESLENAYFFEK	226
PMI	KPDLATAOSYLASGEYWNNSGMFLFRASRYLEELRKFHPAIAD-ACOK-AWENKRDADFT-	232
	** * .. : . . * : : : : : : : *	
1H5R	-----GELEITD-INRIYLEQGRLSVAMMGRGYAWLDTGTHQSLIEASN	238
2CU2	-----LEEVYAGLPKISIDYGVMEKA-E--RVRVVLGR--FPWDDVGNWRALERVFS	267
2X5S	-----LKKAYEKVPSISVDYAVMEKS-K--KVRVVKAD--FEWSDLGNWSSVREIEG	272
2QH5	KIARLSEKSMQDLEDMSIDIALMQOS-H--KIKMVELN--AKWSDLGNFNAL-----	273
PMI	--TRLDKDAFAASPSDSIDYAVMEKT-A--DAVVVPLD--AGWNDVGSWSSL-----	277

Os aminoácidos com coloração azul escuro pertencem a folha β enquanto que aqueles com coloração verde pertencem às α-hélices. Os aminoácidos que tiveram seus rotâmeros modificados estão realçados em azul. A simbologia de asterisco (\*) representa a presença de um único resíduo totalmente conservado, dois pontos (:) indica a conservação entre grupos de

*propriedades fortemente semelhantes e um ponto (.) indica a conservação entre grupos de propriedades fracamente semelhantes.*

Fonte: o autor.

A parte final da sequência da XAC (a partir do aminoácido 345) foi descartada devido ao não alinhamento com as outras sequências selecionadas e tal passo é possível pois essa região não está associada ao sítio ativo do modelo a ser gerado. Além disso, o alinhamento apresenta valores de 58% de identidade e 33% de similaridade no sítio ativo.

#### 4.1.3 Etapa de construção e avaliação do modelo

Após a conclusão da etapa de alinhamento, a próxima etapa a ser conduzida é a construção do próprio modelo. Tal passo foi possível através do programa **MODELLER**, que gerou 5 hipóteses de modelo.

Além da construção do modelo, **MODELLER** pode realizar tarefas auxiliares adicionais, incluindo atribuição de “voltas”, alinhamento de duas seqüências de proteínas ou seus perfis (MARTÍ-RENOM et al., 2000), alinhamento múltiplo de seqüências de proteínas e/ou estruturas (ESWAR et al., 2007), cálculo de árvores filogenéticas e modelagem de alças ou *loops* em estruturas de proteínas (FISER; ŠALI, 2003).

Para a construção de um modelo tridimensional, o **MODELLER** utiliza algumas restrições espaciais definidas entre as seqüências escolhidas e o molde. Algumas dessas restrições são: distâncias e ângulos diedros, comprimento e ângulos de ligação e restrições criadas manualmente como, por exemplo, enovelamento da estrutura secundária, ligações cruzadas, entre outros.

Primeiramente, são analisadas as distâncias existentes entre os resíduos das seqüências alinhadas, enquanto, que os comprimentos de ligações, ângulos de ligação e de diedro são provenientes do alinhamento com o modelo. Em seguida, utiliza-se uma função objetiva para determinar os comprimentos e ângulos de ligação entre átomos não-ligados. Após esses passos, surge um modelo por satisfação dessas restrições espaciais. Subsequente ao surgimento do modelo, pode ocorrer uma avaliação e modelagem das regiões de alças/*loops* através de um algoritmo, seguido pela otimização estrutural do modelo gerado (ESWAR et al., 2007; SANTOS FILHO; BICCA DE ALENCASTRO, 2003).

Em posse dos modelos gerados, o próximo passo é a etapa de avaliação para determinar o modelo mais confiável e adequado para novos estudos. Na etapa de avaliação, é aconselhável

avaliar a presença de contatos atômicos bons ou ruins e também comparar os resíduos do modelo com base nos parâmetros observados nas estruturas proteicas cristalográficas de alta resolução. Esses estudos foram desenvolvidos com o servidor on-line **WHATIF** e o programa **PROCHECK**, respectivamente, e os principais resultados são exibidos na Tabela 4.

O servidor *online* **WHATIF** analisa contatos atômicos gerando pontuações para cada resíduo e, adicionalmente, proporciona um ambiente flexível para exibir, manipular e analisar pequenas moléculas, proteínas, ácidos nucleicos e suas interações, para validação e correção de estruturas protéicas.

O **PROCHECK** (LASKOWSKI et al., 1993) realiza uma validação estereoquímica através da comparação entre os resíduos do modelo com base em parâmetros observados em estruturas de proteínas de alta-resolução cristalográfica. Uma das medidas geradas por esse programa é o General Factor (G-Factor), o qual, este último, consiste em uma média da estereoquímica geral do modelo.

Dados os resultados obtidos pelo **WHATIF**, pode-se observar que a melhor hipótese avaliada foi a Hipótese 4, que obteve uma pontuação total - 0,758, sendo que o valor de referência para este estudo é - 0,813, que corresponde ao escore geral da estrutura cristalográfica (PDB ID 1H5R) usada como referência (Tabela 4).

Outro ponto de atenção são as regiões identificadas do Gráfico de Ramachandran. O Diagrama Ramachandran é dividido em quatro regiões: altamente favorável, favoravelmente permitido, generosamente permitido e proibido. Quando um aminoácido está na região proibida, isso significa que qualquer distância interatômica entre átomos não ligados é menor que a distância de Van der Waals correspondente a esses átomos. No caso da Hipótese 4, nenhum resíduo de aminoácido é encontrado na região proibida, favorecendo sua seleção como modelo (Tabela 4).

Um ponto importante a considerar é o fator G, uma pontuação baseada nos parâmetros estereoquímicos observados de uma determinada estrutura. No caso do modelo PMI, as hipóteses apresentaram um fator G de - 0,1, um valor que está dentro da região para uma proteína com uma resolução de 1,9 Å (referente ao 1H5R) (Tabela 4).

**Tabela 4** - Apresentação dos resultados da Pontuação Geral (**WHATIF**), Regiões Proibidas (Ramachandran) e Fator G (**PROCHECK**) obtidos na etapa de validação dos modelos gerados.

<b>Modelos</b>	<b>Pontuação Geral (WHATIF)</b>	<b>Regiões Proibidas (Ramachandran Plot)</b>	<b>G-Factor</b>
<i>1H5R</i>	- 0.813	1	0
<i>Hipótese 1</i>	- 0.697	0	- 0.1
<i>Hipótese 2</i>	- 0.745	1	- 0.1
<i>Hipótese 3</i>	- 0.734	0	- 0.1
<i>Hipótese 4</i>	- 0.758	0	- 0.1
<i>Hipótese 5</i>	- 0.667	1	- 0.1

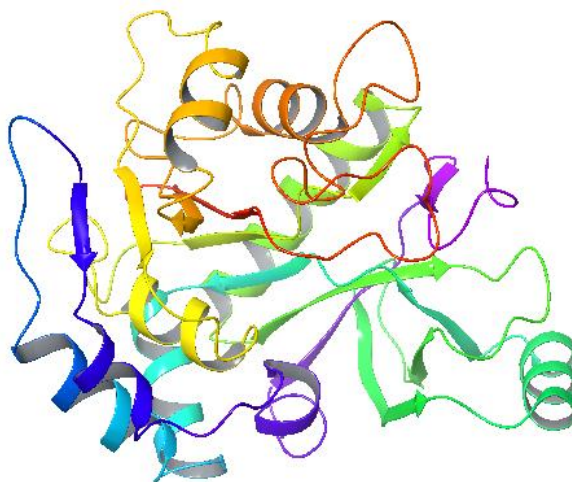
Fonte: o autor

Devem ser analisadas nesta fase do processo, as conformações dos resíduos de aminoácidos que fazem parte do local de interação em estudo. Essa inspeção visual tem como objetivo analisar e, se necessário, modificar a conformação de determinados resíduos através do uso de uma biblioteca de rotâmeros. A escolha do melhor rotâmero deve levar em consideração a direção dos resíduos da estrutura utilizada como referência para o desenvolvimento do modelo e a ausência de contato com os resíduos vizinhos.

Destacamos que, para o modelo PMI gerado (Figura 8) foi necessário modificar alguns resíduos, o que pode ser visto na Figura 7. Com essa mudança de rotâmeros, o valor de **WHATIF** para a hipótese 4 foi alterado para - 0.782, aproximando-se do valor da estrutura de referência (1H5R), além de estar entre - 1 e - 0,5 indicando que se trata de um bom modelo.

Na figura 8 é apresentado o modelo gerado para a PMI de XAC com o programa Modeller, em que é possível observar a presença das estruturas secundárias (alfa-hélices, folhas beta e alças) além de mostrar o seu enovelamento tridimensional.

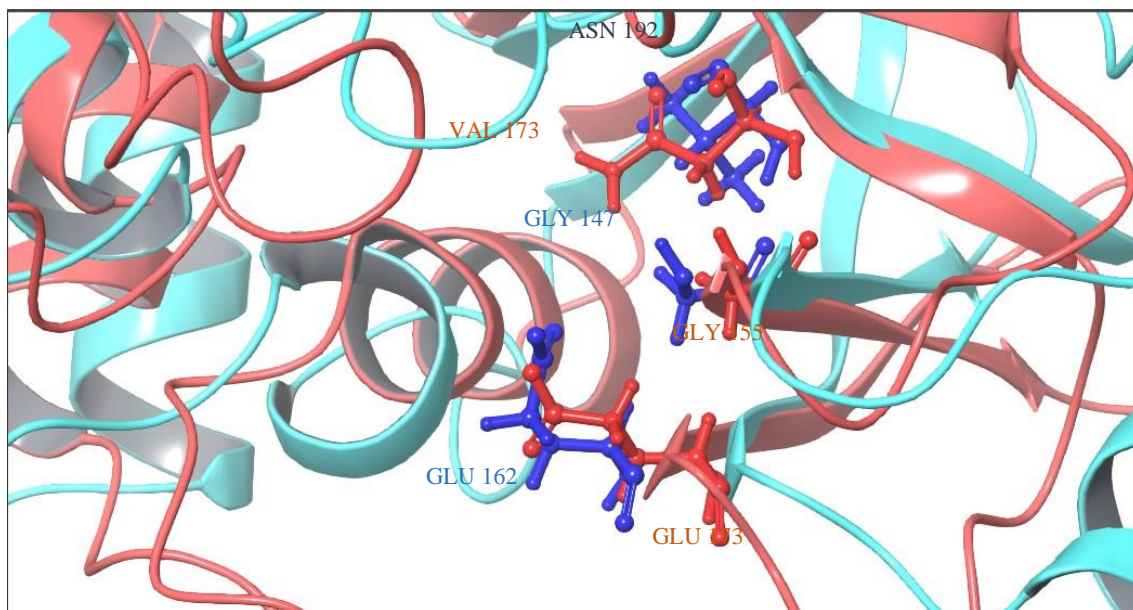
**Figura 8** - Modelo gerado para a fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*. utilizando o programa **MODELLER**.



Fonte: o autor.

A sobreposição do modelo gerado (hipótese 4) com a estrutura referência (PDB ID 1H5R) revela ainda dados importantes que aumentam a confiabilidade do modelo obtido (Figura 9). Pode ser observado, na região do sítio ativo, que os resíduos de aminoácidos descritos por Zuccotti *et al.* (2001) como responsáveis pela formação do complexo ligante-PMI, Gln147, Glu162 e Val173, permanecem alocados na mesma região, quando comparado o modelo com a estrutura de referência de PDB ID 1H5R (Figura 9). Porém, no modelo criado baseado na sequência da XAC, o aminoácido valina é substituído por uma asparagina, mas ambos (asparagina do modelo e valina da estrutura referência) apresentam-se na mesma região e sobre a mesma conformação de cadeia principal (Figura 9).

**Figura 9** - Modelo final da fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*, obtido por modelagem comparativa, com destaque para alguns aminoácidos presentes no sítio ativo.



*Em cor azul, os resíduos pertencentes ao modelo que ficaram alocados na mesma região observada para a estrutura aqui usada como referência (PDB ID 1H5R), que se encontra na cor vermelha.*

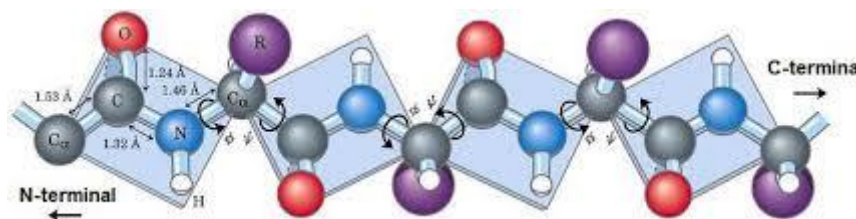
Fonte: o autor.

#### 4.1.4 Avaliação dos ângulos diédricos $\Psi$ versus $\phi$ através do Diagrama de Ramachandran

O Diagrama de Ramachandran, fornecido pelo **PROCHECK**, é uma ferramenta de análise importante pois indica erros na estrutura terciária dessas macromoléculas. O diagrama é dividido em regiões energeticamente favoráveis e desfavoráveis e mostra quais resíduos se encontram em tal regiões, permitindo uma avaliação de qualidade dos modelos teóricos ou experimentais (RAMACHANDRAN et al., 1963).

No gráfico é possível observar dois eixos contendo os ângulos da cadeia principal, Phi e Psi (Figura 10), com valores entre  $-180^\circ$  e  $+180^\circ$ , compreendendo todas as combinações energéticas/conformações possíveis que os resíduos de aminoácidos podem adquirir de acordo com esses ângulos diedros (RAMACHANDRAN et al., 1963). Entretanto há dois aminoácidos que podem ser encontrados em regiões desfavoráveis, mas não causariam dano; são eles: glicina e prolina.

**Figura 10** - Ângulos diedros de uma proteína evidenciando os ângulos Phi e Psi.

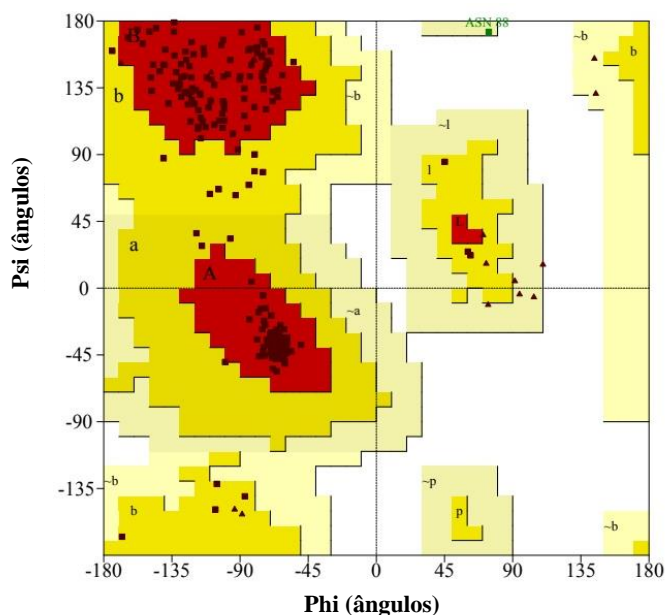


Fonte: Sancho, n.d.

A glicina possui apenas um hidrogênio como cadeia lateral, sendo menos estericamente impedido em comparação com outros aminoácidos. Este fato permite com que a cadeia polipeptídica dos resíduos de glicina assumam conformações que são ‘proibidas’ para outros resíduos. Assim, a glicina ocorre frequentemente em regiões de proteínas onde qualquer outro resíduo seria estericamente impedido. Por outro lado, a prolina é um iminoácido e possui uma cadeia lateral cíclica, tornando sua rotação restrita em torno da ligação. No caso da prolina, os valores de phi são restritos a valores em torno de  $-60^\circ$  e, portanto, trata-se do resíduo de aminoácido mais conformacionalmente restrito.

O gráfico obtido para o modelo PMI gerado (Figura 11) não indicou ângulos phi e psi desfavoráveis para os resíduos presentes no modelo, especialmente na região do local de interação. O modelo apresentou 91,6% de seus resíduos em regiões altamente favoráveis e, embora o resíduo ASN88 apareça em uma região “generosamente permitida”, seus ângulos de torção foram “herdados” das estruturas cristalográficas utilizadas como referência.

**Figura 11** - Diagrama de Ramachandram de *Xanthomonas citri* subsp. *citri* obtido pelo **PROCHECK**: 91,6% de resíduos em regiões favoráveis, 7,9% de resíduos em regiões permitidas, 0,4% de resíduos em regiões generosamente permitidas e 0% de resíduos em regiões proibidas.



Estatísticas de plotagem		
Resíduos em regiões mais favoráveis [A,B,L]	219	91.6%
Resíduos em regiões adicionais permitidas [a,b,l,p]	19	7.9%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	1	0.4%
Resíduos em regiões proibidas	0	0.0%
Número de resíduos não glicinas e não prolina	239	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	19	
Número de resíduos de prolina	17	
Número total de resíduos	277	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

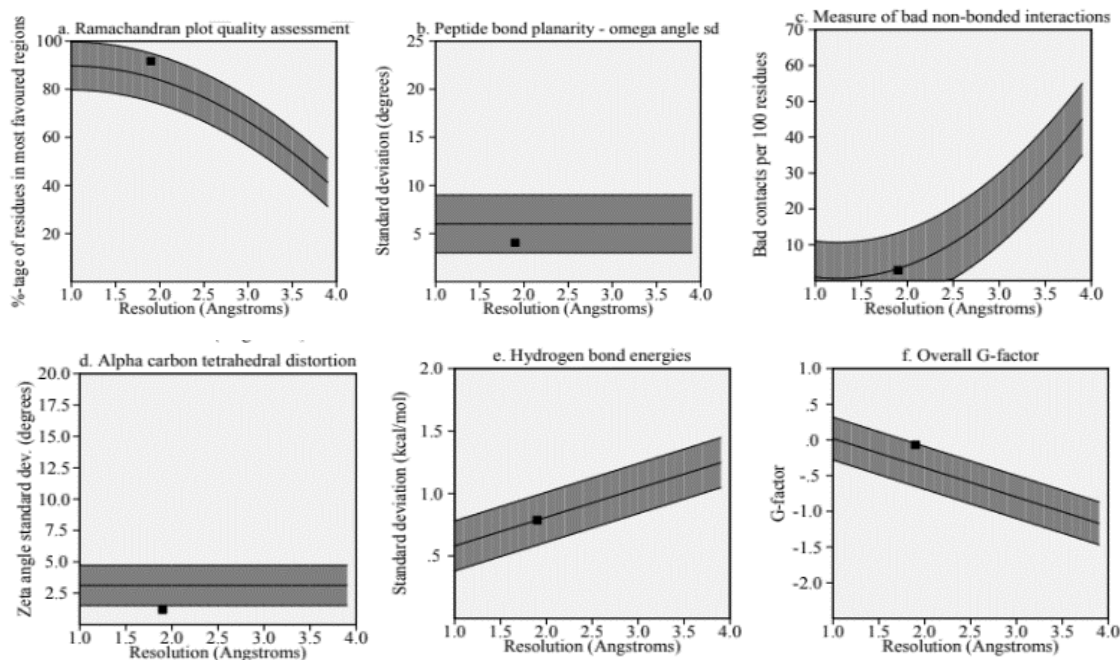
Fonte: o autor.

#### 4.1.5 Outros parâmetros importantes para a análise (Suite **SAVE v5.0**)

Outras ferramentas importantes para analisar um modelo de proteína predito podem ser encontradas no pacote **SAVE v5**. O uso do **PROCHECK** permite a análise da planaridade das ligações peptídicas, contatos atômicos fracos, distorção tetraédrica dos carbonos alfa, energia das ligações hidrogênio e fator G. Para o modelo PMI gerado, todos os valores estão dentro dos valores esperados para estruturas de proteína de resolução de 1,9 Å (referindo-se à estrutura PDB ID 1H5R, usada aqui como referência para modelagem) (Figura 12) e a distorção tetraédrica dos carbonos alfa apresentou um melhor valor que o cristalográfico.



**Figura 12** - Resumo da avaliação estereoquímica do modelo selecionado construído por homologia estrutural para fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*., com um G-Factor de -0.1 e distorção tetraédrica do carbono alfa de -0.1.



Fonte: o autor.

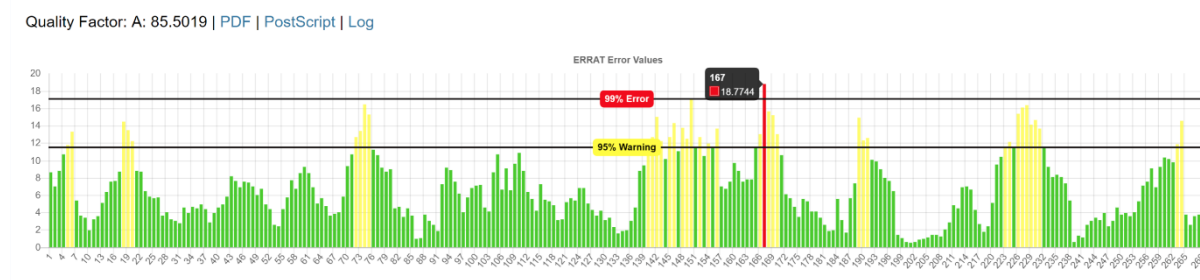
Os parâmetros apresentados juntos nas Figuras 11 e 12 ajudam a avaliar a qualidade estereoquímica geral do modelo selecionado. Os gráficos mostram que o modelo obteve valores dentro do intervalo de referência **PROCHECK** e, ao analisar os valores apresentados pelo **WHATIF** para os aminoácidos pertencentes ao local de interação do PMI, nenhum apresentou valores abaixo de -5, que, se existirem, devem ser revisados e reformados.

Além das ligações atômicas que ocorrem na cadeia peptídica, há também interações entre átomos não-ligados. A estrutura terciária de uma proteína é altamente dependente dessas interações e existem seis tipos de interações prováveis entre átomos não ligados e elas não ocorrem aleatoriamente. Assim, é possível através do programa **ERRAT** identificar regiões e fazer uma na análise estatística dessas interações.

O **ERRAT** (COLOVOS; YEATES, 1993) analisa as interações intermoleculares entre átomos diferentes no modelo em comparação com estruturas altamente refinadas, gerando um gráfico de uma função de erro por uma janela formada por nove resíduos. De um modo geral, é um método mais eficaz que a **Perfil 3-D** proposto por Bowie e colaboradores (BOWIE; LTCY; EISENBERG, 1990), porém não é melhor que o **PROCHECK**.

Observando o gráfico desenvolvido pela **ERRAT** para o modelo PMI (Figura 13), é possível identificar três regiões distintas: uma verde, uma amarela e uma vermelha. A região verde mostra que todos os aminoácidos estão dentro dos valores esperados para estruturas altamente resolvidas e refinadas. A região amarela corresponde aos aminoácidos que excedem o resultado esperado e podem ser rejeitados com 95% de confiança. Espera-se que uma boa estrutura proteica tenha 5% de sua estrutura nessa região. Finalmente, as regiões vermelhas são aquelas que podem ser rejeitadas com 99% de confiança. O modelo gerado para o PMI possui poucas regiões amarelas e apenas um único aminoácido na região vermelha (Ala167). Nenhum aminoácido presente nessas regiões faz parte do site ativo; portanto, não haverá problemas para testes futuros. Além desses dados, o fator de qualidade, valor relacionado às interações atômicas não ligadas de 85.5019, determina uma estrutura de alta qualidade (resultado para valores acima de 50).

**Figura 13 - ERRAT** obtido para o modelo de fosfomanose isomerase gerado de *Xanthomonas citri* subsp. *citri* com um fator de qualidade de 85.5019.



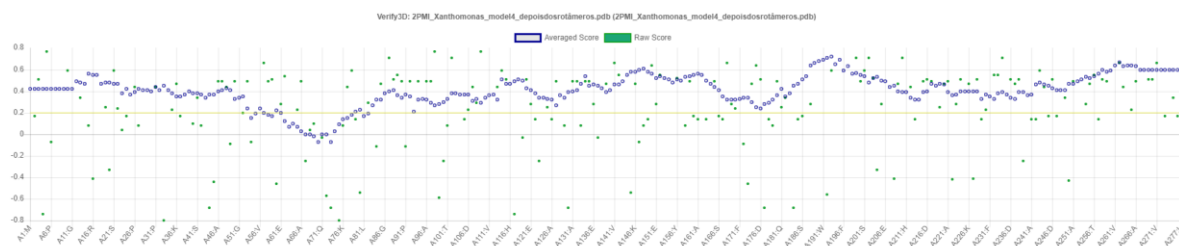
A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas. Fonte: o autor.

Outra análise interessante é verificar se um resíduo de aminoácido corresponde ao ambiente em que está, ou seja, se há compatibilidade do modelo com sua sequência de aminoácidos. Essa análise pode ser realizada usando o **Verify-3D**.

O **VERIFY 3D** (LUTHY; BOWIE; EISENBERG, 1992) analisa a compatibilidade entre o modelo 3D e a sua sequência de aminoácido, baseado na localização (alfa, beta, loop) e uma sequência de ambientes químicos (polar, apolar), utilizando uma matriz proposta por Bowie e colaboradores (BOWIE; LTCY; EISENBERG, 1990). Essa matriz permite uma análise de probabilidade de se encontrar os 20 aminoácidos em diferentes ambientes químicos, com base no tipo de aminoácido e natureza do ambiente químico. Em seguida, é realizado o alinhamento entre o modelo 3D e a sequência de aminoácidos; permitindo detectar as regiões mal resolvidas.

No estudo do PMI, pode-se observar que pelo menos 80% dos aminoácidos tiveram uma pontuação igual ou superior a 0,2 no perfil 3D-1D (Figura 14). Além disso, há uma região (Ala66 a Leu82) que caiu abaixo do limite, indicando uma provável dobra nessa região, no entanto, não é motivo de preocupação porque não está na região do site ativo do modelo.

**Figura 14 - Verify-3D** obtido para o modelo de fosfomanose isomerase gerado de *Xanthomonas citri* subsp. *citri*: 91.70% dos aminoácidos obtiveram uma pontuação igual ou superior a 0,2.

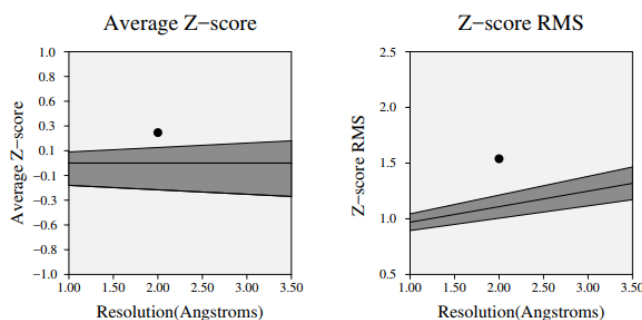


Fonte: o autor.

Finalmente, como última análise, é possível verificar o volume dos átomos. O programa **PROVE** correlaciona o volume obtido para cada aminoácido da proteína com os volumes de proteínas de alta resolução.

Como resultado, o **PROVE** fornece um valor de pontuação chamado Z-score que corresponde ao desvio estatístico programado de estruturas altamente resolvidas e refinadas. No caso do modelo PMI em estudo, tanto o Z-score quanto o Z-score RMS estavam acima da região de interesse; sendo que o Z-score possui uma média de 0,274 e o Z-score RMS com valor de 1,539 (Figura 15). Além disso, 4,6% dos resíduos analisados foram considerados outliers, indicando que mais atenção deve ser dada ao modelo. No entanto, juntamente com todas as outras análises realizadas, pode-se determinar que este é um modelo confiável para o PMI de XAC, e permite dar continuidade aos estudos.

**Figura 15** - PROVE obtido para o modelo fosfomanose isomerase gerado de *Xanthomonas citri* subsp. *citri*.



O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: o autor

#### 4.1.6 Etapa de validação do modelo de PMI XAC.

Após a criação do modelo, foi realizado uma simulação para a validação do mesmo, utilizando os programas **GOLD** e **GLIDE**. Definiu-se, a partir do substrato da estrutura cristalográfica original (1H5R), a melhor região do sítio ativo na qual seriam realizados os cálculos.

Os simulações de *docking* são cálculos que analisam a potencial afinidade entre o ligante e o sítio de ligação do receptor, propondo a pose (conformação + orientação), ou predição do modo de ligação com a maior afinidade entre ligante e receptor, estimada por uma pontuação calculada com base em uma função empírica associada ao método de mecânica molecular associado. A seguir serão descritos os métodos utilizados pelos programas **GOLD** (SUITE, 2015) e **GLIDE**.

O posicionamento dos ligantes no sítio do receptor resulta de um algoritmo adotado pelo programa escolhido para os simulações de *docking*. Os algoritmos mais usados são os genéticos, baseados em Monte Carlo ou busca sistemática. O **GOLD** utiliza o algoritmo genético enquanto o **GLIDE** utiliza um algoritmo de busca sistemática associada ao método de Monte Carlo.

No programa **GOLD**, o processo para gerar possibilidades conformacionais para a molécula ligante (e também para seletas cadeias laterais de resíduos de aminoácidos), se inicia gerando populações contendo conformações aleatórias de um dado ligante, em que cada

indivíduo dessa população é codificado e chamado de “cromossomo”. Os cromossomos possuem, então, as informações dos átomos dos ligantes como possibilidades de fazer interações. Além disso, eles também possuem informações sobre liberdades translacionais, rotacionais e conformacionais, chamadas de “genes”.

Com a ideia de gerar conformações mais “evoluídas”, os cromossomos utilizam mecanismos chamados operadores genéticos, no qual haverá uma análise de mutações e recombinações. Assim, diversas etapas como essa ocorrem até que seja obtida a população final (SUITE, 2015), evoluída de seus “ancestrais”, em processo tendenciado à maior afinidade com o sítio receptor.

Após a criação da população final, cada cromossomo/ligante é ancorado diversas vezes no sítio com diferentes orientações e cada pose é pontuada através de uma função chamada *fitness score*.

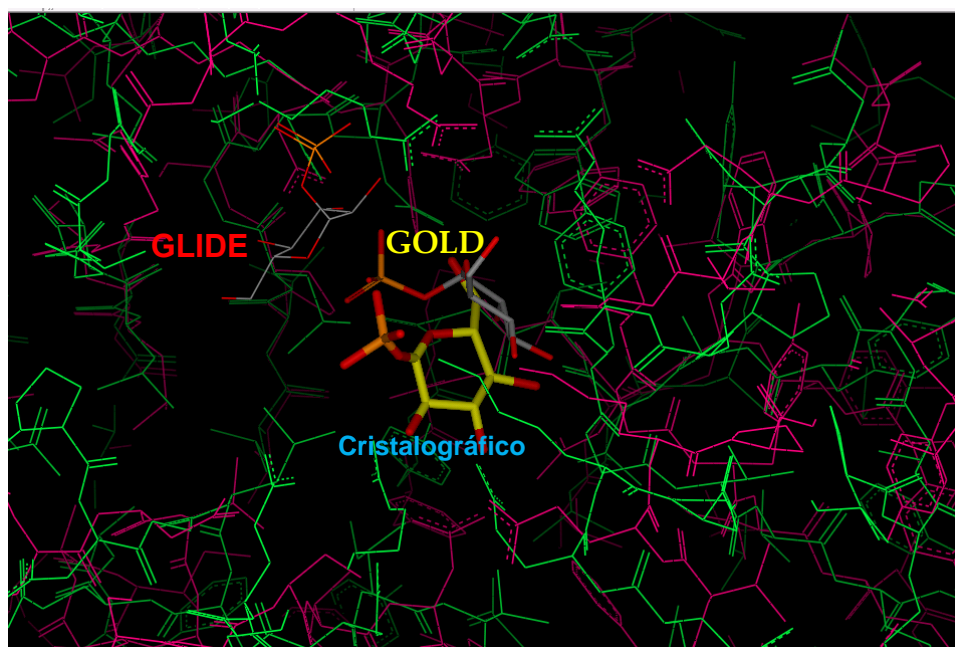
O primeiro passo do método utilizado pelo **GLIDE** consiste na construção de um *grid* no sítio de interação do receptor, determinando sua forma e propriedades. Nesse *grid* são feitas análises da presença de átomos existentes, fornecendo informações sobre os tipos de interações presentes e a pontuação para a pose do ligante. Em seguida, ocorre a geração de confôrmeros, os quais serão analisados nessa região.

A geração desses confôrmeros pelo programa **GLIDE** é baseado nas conformações de cada ligante, a qual varia conforme do número de ligações livres. A busca pelos melhores ligantes para um determinado receptor começa com a procura por regiões em que podem ocorrer interações, tanto no ligante quanto no receptor. Se ocorrer uma correspondência entre a conformação e o sítio de ancoragem, essa conformação segue para a próxima etapa.

A segunda etapa analisa os átomos do ligante, em relação à cavidade do receptor, de forma a evitar choques estéricos. Além disso, são feitas avaliação em relação ao diâmetro do ligante e suas rotações. O ligante recebe uma pontuação e caso ela seja ‘satisfatória’ de acordo com os critérios estabelecidos, ele segue em frente. Em seguida, é calculada uma nova pontuação, chamada de “gananciosa”, que irá avaliar a posição do ligante ao receptor quando esse se move em torno de 1Å nas coordenadas x, y e z. Esse movimento gera uma nova pontuação chamada de ChemScore, na qual penaliza contatos estéricos além de reconhecer diversos tipos de interações como hidrofóbica, metálicas e ligações de hidrogênio. Após isso, o ligante é novamente movimentado nas coordenadas e sua pontuação é recalculada. Essa segunda etapa ocorre a redução dos números de confôrmeros. Por fim, no terceiro passo, o algoritmo Monte Carlos explora a minimização de energia dos ligantes, pontuando as poses novamente através da função *GlideScore*.

O primeiro passo para a etapa de validação do modelo consistiu na escolha do programa para as simulações de *docking*, o **GOLD** ou o **GLIDE**. Para tanto foi feito um *redocking*, para ver qual programa conseguia posicionar melhor o ligante cristalográfico, glicose-1-fosfato (G1P), em relação à sua posição original na estrutura cristalográfica PDB ID 1H5R. Foram rodadas 10 simulações para cada programa. Como pode ser observado na Figura 15, o **GOLD** foi o programa que conseguiu um melhor posicionamento, sendo ele selecionado para futuros simulações de *docking* (Figura 16).

**Figura 16** - Validação do modelo gerado para a fosfomanose isomerase *Xanthomonas citri* subsp. *citri* mostrando a posição do ligante glicose-1-fosfato e sua posição em relação do ligante cristalográfico quando utilizado os programas **GOLD** e **Glide**.



A estrutura em verde corresponde ao modelo gerado (hipótese 4) e a em rosa corresponde a estrutura cristalográfica 1H5R.

Fonte: o autor.

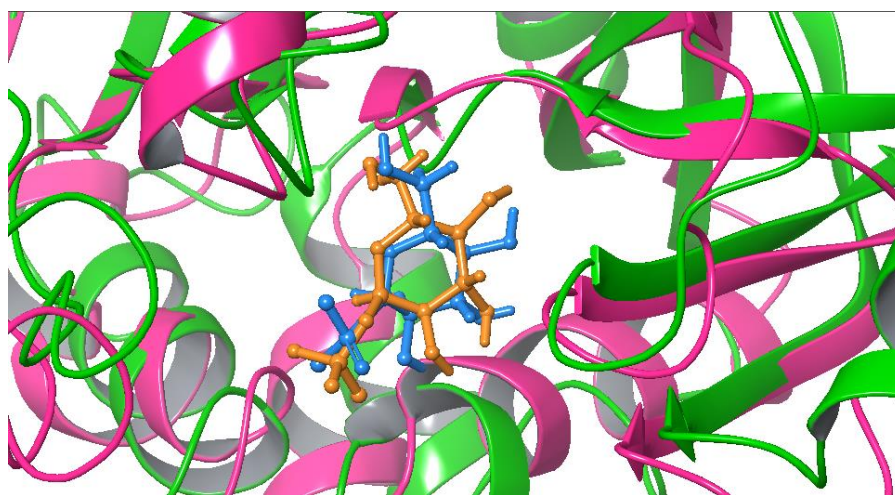
O sucesso de um programa na previsão de uma posição de ligante é geralmente medido pelo desvio médio quadrático, ou RMSD (Root Mean Square Deviation), no qual é possível determinar o grau de similaridade entre duas estruturas (COLE et al., 2005). Esse dado é calculado entre a posição do ligante observada experimentalmente e a posição do ligante prevista pelo programa, com base nas distâncias interatômicas de seus átomos, sendo o valor obtido através da seguinte expressão matemática (1):

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \|r_i^A - r_i^B\|^2} \quad (1)$$

onde  $N$  é o número total de átomos pareados entre as duas estruturas,  $r_i$  é o vetor posição do átomo  $i$ ,  $A$  é o modelo gerado e  $B$  é a estrutura molde.

A estatística mais usada é a que os modelos gerados deveriam obter um  $RMSD < 2\text{Å}$  e, embora seja comumente utilizada, é um limite arbitrário (GOHLKE; HENDLICH; KLEBE, 2000; KONTOYIANNI; MCCLELLAN; SOKOL, 2004; KRAMER; RAREY; LENGAUER, 1999). Por outro lado, alguns autores como Vieth e Bursulaya (BURSULAYA et al., 2003; VIETH et al., 1998), consideram valores de  $RMSD$  entre  $2\text{Å}$  e  $3\text{Å}$  como sucessos parciais. Para o modelo gerado da PMI de XAC foi encontrado um  $RMSD = 2.5080\text{Å}$  (Figura 17) que, de acordo com os autores citados acima, seria considerado como um sucesso parcial.

**Figura 17** - Resultado do *Redocking* da 1H5R para cálculo do  $RMSD$ , no qual foi obtido um valor de  $2.5080\text{Å}$ .



A estrutura cristalográfica PDB ID 1H5R está em cor rosa, o modelo de PMI XAC em cor verde, o ligante cristalográfico em cor azul e o ligante do redocking em cor laranja.  
Fonte: o autor.

Uma vez selecionado e refinado o modelo final, avaliamos sua eficiência em estudos de triagem virtual. Para essa avaliação, foram escolhidos 30 inibidores conhecidos da PMI, depositados e provenientes do servidor online **BindingDB**, contendo valores descritos de potência ( $IC_{50}$ ), reportados na literatura. O objetivo dessa etapa era avaliar a capacidade que o modelo possui de recuperar os ligantes sabidamente ativos e ordená-los de forma coerente em relação ao seus  $IC_{50}$ , com base em função de pontuação (**GOLD Fitness**) proveniente do programa **GOLD**.

Para que o objetivo fosse alcançado, primeiramente era necessária a definição do sítio de *docking*, a qual foi obtida através do posicionamento do ligante proveniente da PDB ID 1H5R em relação ao modelo. Com essa posição, foi possível obter-se uma esfera com as

seguintes coordenadas espaciais:  $x = 23,977$ ,  $y = 15,659$ ,  $z = 43,886$ , com uma esfera de raio de 8 Å (região onde seriam realizadas as simulações de *docking*). Os resultados dessa etapa são mostrados na Tabela 5.

**Tabela 5** - Resultado da recuperação da seleção de inibidores (15 mais potentes e 15 menos potentes) do **BindingDB** pela fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*.

Número do Inibidor	Nome PCid	Gold.PLP.Fitness	IC50
1	2998795	62.216	1.30E+03
2	1256813	60.097	1.00E+03
3	5771379	59.096	1.00E+03
4	25199560	57.191	>1,00E+5
5	4089709	56.362	707
6	25199559	55.438	>1,00E+5
7	2311082	54.774	270
8	2248499	48.79	829
9	1489253	48.566	1.23E+03
10	3093671	45.896	1.27E+03
11	4120895	44.867	>1,00E+5
12	22416235	44.529	1.25E+03
13	4917291	44.379	709
14	387196	44.267	>1,00E+5
15	25241622	43.919	>1,00E+5
16	1510389	43.654	985
17	3311395	41.868	>1,00E+5
18	25181243	41.807	1.03E+03
19	1519965	41.795	1.00E+03
20	4879298	41.788	>1,00E+5
21	3194	41.062	190
22	25181239	36.225	>1,00E+5
23	25181212	35.122	>1,00E+5
24	22765927	35.095	>1,00E+5
25	25241626	33.333	>1,00E+5
26	6112173	32.755	>1,00E+5
27	25241624	32.246	>1,00E+5
28	6033619	27.892	>1,00E+5
29	5951502	26.357	>1,00E+5
30	25199533	22.889	1.07E+03

*Os inibidores mais potentes são mostrados na cor verde e os inibidores menos potentes são mostrados na cor vermelha.*

Fonte: o autor.

Observando os dados obtidos pelo estudo de triagem virtual, o próximo passo consistiu na obtenção de curvas ROC (Receiver Operating Characteristic), de predição e de enriquecimento para melhor entendimento dos resultados. Para a criação desses gráficos, foi



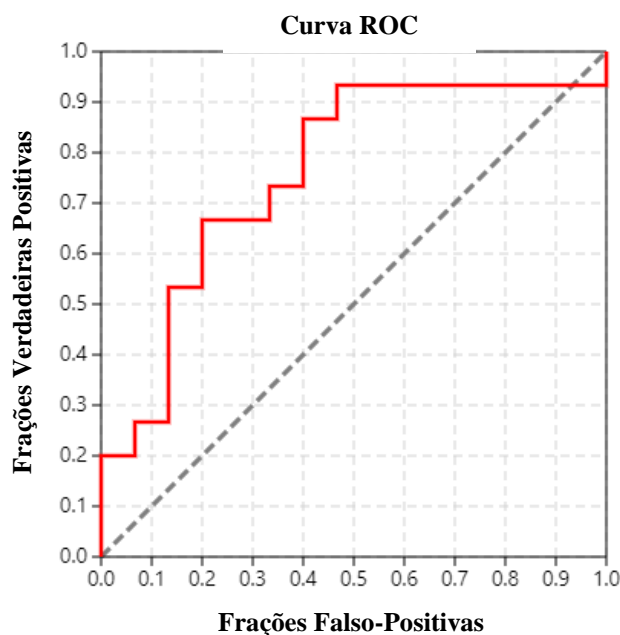
utilizado o *web servidor* Screening Explorer (EMPEREUR-MOT; ZAGURY; MONTES, 2016) com os parâmetros selecionados como padrões.

Em estudos de triagem virtual, a análise dos dados é feita com base em uma curva ROC. Trata-se de um gráfico dividido em frações positivas verdadeiras (FVP) localizadas no eixo y e frações falso-positivas (FFP) localizadas no eixo x. Um par de FVP/FFP é representado por cada ponto dentro da curva ROC e corresponde a uma fração específica de um conjunto de moléculas. As curvas ROC resumem a capacidade geral de um método para distinguir entre compostos ativos e inativos (EMPEREUR-MOT et al., 2015; EMPEREUR-MOT; ZAGURY; MONTES, 2016).

Uma função de pontuação com perfeita discriminação é aquela em que no canto superior esquerdo do gráfico, temos uma sensibilidade perfeita (FVP = 1) e uma especificidade perfeita (FFP = 0). A sensibilidade corresponde a capacidade do sistema em prever corretamente quais moléculas possuem certas características (proporção de positivos verdadeiras) ; enquanto especificidade corresponde em prever a ausência de características em algumas moléculas (proporção de falsos-positivos) (BERTRAND; TRIBALLEAU, 2005).

A linha diagonal presente no gráfico corresponde a experimentos em que a função de pontuação não seria discriminada. Uma curva entre a linha diagonal e a parte superior esquerda do gráfico é considerada boa, sendo que quanto maior a distância da linha diagonal, melhor será o sistema. De forma qualitativa, quanto mais próximo à esquerda do gráfico estiver a curva, maior será a precisão geral do estudo (BERTRAND; TRIBALLEAU, 2005; EMPEREUR-MOT et al., 2015; EMPEREUR-MOT; ZAGURY; MONTES, 2016). Abaixo, na Figura 18, é mostra o gráfico de curva ROC obtido para o estudo realizado com os 30 inibidores do BindingDB e o modelo de PMI XAC gerado. Nele é possível observar que grande parte do gráfico encontra-se distante da linha diagonal, mostrando que o estudo apresenta uma boa precisão, conseguindo diferenciar os compostos mais ativos dos menos ativos.

**Figura 18** - Curva ROC obtida para o estudo dos 30 compostos relacionados como mais potentes e menos no **BindingDB** pelo molde gerado para a proteína fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*.

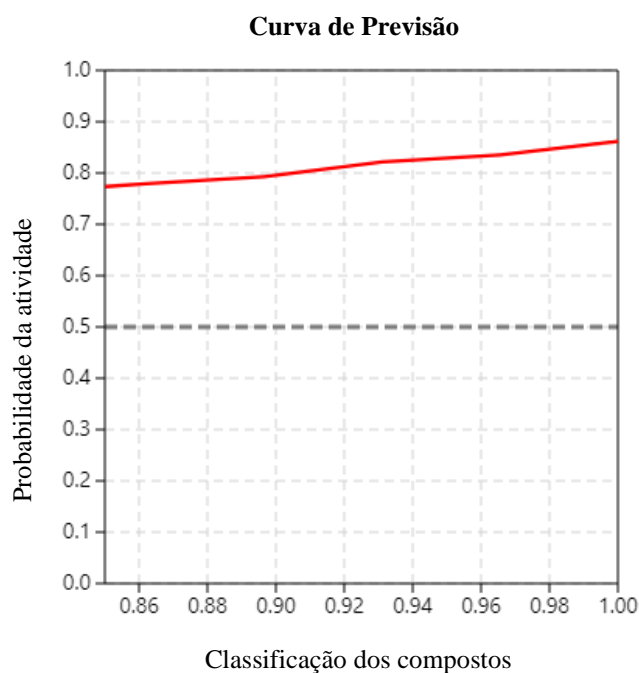


Fonte: o autor.

Caso fique difícil analisar uma curva ROC, é possível utilizar uma medida padrão chamada de área sob a curva (AUC). O AUC é um número que varia de 0 a 1 obtido por métodos de integração numérica e mostra o desempenho geral do estudo, isto é, a probabilidade de o modelo classificar um dado positivo como negativo e vice-versa. Teoricamente, quanto maior o valor obtido para AUC, melhor é o sistema utilizado (EMPEREUR-MOT; ZAGURY; MONTES, 2016). O valor de AUC ROC obtido para esse estudo foi de 0.760, um bom valor para o estudo realizado já que se encontra próximo a 1 (valor ideal).

Um outro dado obtido pelo *web server* é a curva de previsão, que quantifica a associação entre as variações de pontuação e detecção dos compostos ativos. No eixo y do gráfico é mostrado a probabilidade de atividade para cada composto presente no conjunto de dados e no eixo x, esses compostos são mostrados (EMPEREUR-MOT et al., 2015). Consequentemente, a curva de previsão permite detectar a variação emitidas por uma função de pontuação na detecção dos compostos ativos, isto é, se as variações de probabilidade de atividade são importantes o suficiente para induzir a seleção de um limiar para triagens virtuais (EMPEREUR-MOT et al., 2015). Na Figura 19 é mostrado a curva de previsão para esse estudo e ela se encontra acima da linha limite, tornando-se um bom resultado e mostrando que o modelo foi capaz de realizar uma discriminação entre os compostos utilizados.

**Figura 19** - Curva de Previsão obtida para o estudo dos 30 compostos relatados como ativos e inativos no **BindingDB** pela proteína fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*.



Fonte: o autor.

Um dado calculado a partir da curva de previsão é o ganho total padronizado (TG) (BURA; GASTWIRTH, 2001), que resume a discriminação dos compostos ativos em relação à variação das pontuações ao longo do conjunto de dados das moléculas (EMPEREUR-MOT et al., 2015).

A TG possui seu valor entre 0 e 1, sendo que sua análise deve estar associado ao valor obtido para a AUC. Valores de TG acima de 0,25 combinados com uma AUC ROC acima de 0,5, significa que as variações de pontuação são relevantes na discriminação dos compostos ativos; já valores de TG acima de 0,4 associados com uma AUC ROC acima de 0,5 indica, que o método de triagem utilizado teve um bom desempenho e seria reproduzível em condições experimentais semelhantes (EMPEREUR-MOT; ZAGURY; MONTES, 2016).

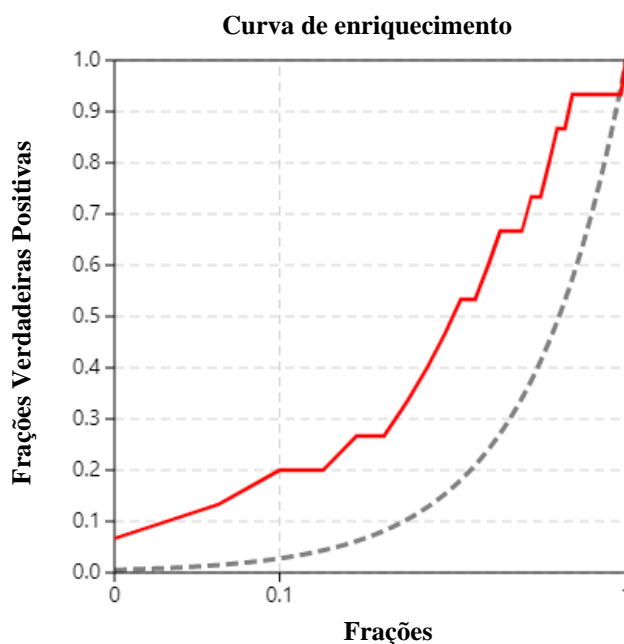
No nosso estudo, o valor de TG obtido foi de 0,312 que, associado a um AUC ROC de 0,760, mostra que o sistema foi capaz de discriminar os compostos ativos para tal modelo proteico.

Por fim, um ultimo dado obtido é a curva de enriquecimento, a qual permite avaliar a recuperação precoce de compostos ativos presentes em estudos de triagem virtual. O gráfico em escala logaritmica mostra no eixo y as FVP e no eixo x o conjunto de dados (TRUCHON; BAYLY, 2007).

Assim, como as curvas ROC, as de enriquecimento possuem uma linha diagonal que inicia no canto inferior esquerdo, subindo até o canto superior direito e que representa uma distribuição aleatória. Quanto mais distante a curva gerada estiver da linha diagonal, melhor é o estudo realizado (BERTRAND; TRIBALLEAU, 2005).

Embora as curvas de enriquecimento sejam mais fáceis de traçar, elas possuem uma grande desvantagens. Essas curvas dependem diretamente da proporção de moléculas ativas presentes no conjunto; portanto, quando a proporção desses ativos aumenta, a curva de enriquecimento ideal se aproxima da linha de classificação aleatória (BERTRAND; TRIBALLEAU, 2005). Além disso, o espaço é limitado pelas curvas ideal e aleatória, prejudicando comparações com outros dados. Na Figura 20 é mostrada a curva de enriquecimento para o estudo em questão.

**Figura 20** - Curva de enriquecimento com um fator de enriquecimento de 10% obtida para o estudo dos 30 compostos relatados como ativos e inativos no **BindingDB** pela proteína fosfomanose isomerase de *Xanthomonas citri* subsp. *citri*.



Fonte: o autor.

Um dado relacionado à curva de enriquecimento é o fator de enriquecimento (EF), que corresponde à fração de compostos ativos obtidos quando uma determinada porcentagem de compostos inativos é encontrada (TRUCHON; BAYLY, 2007) Esse dado possui uma alta variabilidade e seu valor máximo depende diretamente da proporção de moléculas ativas no conjunto (TRUCHON; BAYLY, 2007) e do valor de compostos inativos considerados

(PICCIRILLO; DO AMARAL, 2018). Nesse estudo, foi considerado a porcentagem de 10% dos compostos inativos e o valor obtido para EF foi de 2.00, mostrando que para 10 compostos inativos recuperados, 2 deles são compostos ativos.

Além dessas curvas, temos alguns outros dados para levar em consideração: discriminação reforçada por Boltzmann do ROC (BEDROC) (MUEGGE; ENYEDY, 2004) e o reforço inicial robusto (RIE) (CORNELL, 2006). Truchon & Bayly (2007) discutiram vários métodos para abordar o reconhecimento precoce, como o BEDROC e o RIE. Ambas as métricas são utilizadas para quantificar o reconhecimento precoce dos compostos ativos utilizados em um estudo de triagem virtual (EMPEREUR-MOT; ZAGURY; MONTES, 2016). O BEDROC é limitado por um valor de 0 a 1 e trata da probabilidade de um ativo ser classificado antes de um composto selecionado aleatoriamente. Por outro lado, o RIE, desenvolvido por Sheridan (SHERIDAN et al., 2001), utiliza a ponderação exponencial, atribuindo um maior peso para os ativos reconhecidos precocemente (ZHAO et al., 2009).

Uma das desvantagens do RIE é que o seu valor mínimo e máximo depende diretamente da quantidade de compostos ativos e a quantidade de compostos totais presente nos conjuntos. O valor obtido para essa métrica foi 1,832. Por outro lado, o valor obtido pra BEDROC foi 0,916 encontra-se muito próximo a 1, mostrando que o modelo é capaz de separar os compostos ativos e inativos.

Com isso, através desse resultado de validação e daqueles de avaliação mostrados anteriormente, podemos afirmar que o modelo para a PMI XAC está apto para ser utilizado nos posteriores estudos de triagem virtual.

#### 4.2 Construção do modelo da fosfomanose isomerase (PMI) para *Homo sapiens*.

Após a construção do modelo para a PMI de XAC foi decidido realizar um estudo em relação a proteína PMI de *Homo sapiens*. O primeiro passo foi realizar o alinhamento entre as sequências de aminoácidos de ambas proteínas, com o intuito de analisar as porcentagens de identidade e similaridade. Elas compartilham uma identidade de 7% e uma similaridade de 11% (Figura 21).

**Figura 21** - Alinhamento entre as sequências da fosfomanose isomerase de *Xanthomonas citri* subs. *citri* e *Homo sapiens* feita pelo programa **Clustal Omega**.

<i>Xanthomonas</i>	1	-----	0
<i>Homo sapiens</i>	1	MAAPRVFPLSCAVQOYAWGKMGNSSEVARLLASSDPLAQIAEDKPYAELW	50
<i>Xanthomonas</i>	1	-----	0
<i>Homo sapiens</i>	51	MGTHPRGDAKILDNRISQKTLQWIAENQDSLGSVKVDTFNGNLPFLFKV	100
<i>Xanthomonas</i>	1	-----	0
<i>Homo sapiens</i>	101	LSVETPLSIQAHPNKELAEKHLHQAPQHYPDANHKPEMAIALTPFQGLCG	150
<i>Xanthomonas</i>	1	-----MSDVLPIILSGSGTRLWPLSRESYPKQFLPL	32
<i>Homo sapiens</i>	151	FRPVEEIVTFLKTAAGNNMEDIF-----GELLLQL-HQQYP-----	185
<i>Xanthomonas</i>	33	VGDKSMLQSTWLRAPV-AGHAPIVVANEEHRFMAAEQLQ-----Q	72
<i>Homo sapiens</i>	186	-GDIGCFAIYFLNLLTLKPGEAMFLEANVPHAYLKGDCVECMACSDNTVR	234
<i>Xanthomonas</i>	73	LGVKPSAI----LLE-----PKGRNTAPAIAVAALAEATR DGADPLLL	110
<i>Homo sapiens</i>	235	AGLTPKFIDVPTLCEMLSYTPSSSKDRLFLP-----TRSQEDPYLS	275
<i>Xanthomonas</i>	111	V----LPSDHVIQNEAAFQAAVT----LAATAAEQGKLVTFGIKPTAPET	152
<i>Homo sapiens</i>	276	IYDPPVPDFTIMKTEV--PGSVTEYKVLALDSASILLMVQGTVIASPTTT	323
<i>Xanthomonas</i>	153	-----GYGIKASAGAGASAVERFVEKPDLATAQSYLASGEYYWNS	193
<i>Homo sapiens</i>	324	QTPIPLQGGVLFI----GANESVSLKLTEPKDL-----	353
<i>Xanthomonas</i>	194	GMFLFRASRYLEELRKFPAPAIADACQKAWENGKRDADFTRLDKDAFAASP	243
<i>Homo sapiens</i>	354	--LIFRACCLL-----	362
<i>Xanthomonas</i>	244	SDSIDYAVMEKTADAVVVPLDAGWNDVGSWSLLDVSNDQAQGNAAHGDV	293
<i>Homo sapiens</i>	363	-----	362
<i>Xanthomonas</i>	294	IQLDCQNTYAYGSRLIAMVGLEDVVVVETPDAVLVGHDRDRIQEVKDVVSQ	343
<i>Homo sapiens</i>	363	-----	362
<i>Xanthomonas</i>	344	IKTAGRSEATWHRKVYRPWGAYDSIDMQQRHQVKRITVKPGAVLSLQMH	393
<i>Homo sapiens</i>	363	-----	362
<i>Xanthomonas</i>	394	HRAEHWIVVSGTAEVTRGEEVLLLTENQSTYIPLGVTHRLRNPGLPLEL	443
<i>Homo sapiens</i>	363	-----	362
<i>Xanthomonas</i>	444	IEVQSGSYLGEDDIVRFEDTYGRA	467
<i>Homo sapiens</i>	363	-----	362

Fonte: o autor

Após essa etapa, foi criado o modelo para a PMI humana para comprovar que se trata de estruturas totalmente diferentes. O modelo também foi construído por modelagem de homologia estrutural, utilizando todos os mesmos passos que foram utilizados anteriormente, isto é, alinhamento, construção do modelo e avaliação do modelo.

#### 4.1.1 Etapa de identificação da sequência de aprovação

Para a construção do modelo da PMI, inicialmente obtivemos o sequenciamento de aminoácidos da mesma através do banco de dados GenBank (CAA53657.1) (NCBI, 2019). Em seguida, o algoritmo do servidor online **BLASTp** identificou duas estruturas (PDB ID 5NW7 e 3H1M) com identidades sequenciais acima de 25% (Tabela 6), sendo que uma possui identidade de aproximadamente 35% e a outra de 40%. As identidades dessas estruturas acima de 25% mostram que elas compartilham com a estrutura de interesse uma potencial semelhança estrutural e de função

**Tabela 6** - Resultados de identidade e *e-value* de sequências de proteínas obtidas no servidor online **BLAST** usando o algoritmo de explosão de proteínas (**BLASTp**) para a triagem de proteínas homologadas para PMI humana.

<i>PDB ID</i>	<i>Identidade</i>	<i>E-value</i>
<b>5NW7-A</b>	40.32%	$2 \times 10^{-100}$
<b>3H1M_A</b>	34.48%	$5 \times 10^{-64}$

Fonte: o autor

De posse das sequências homólogas da sequência em estudo, o próximo passo foi determinado pelo alinhamento das sequências e sua avaliação.

#### 4.1.2 Etapa de alinhamento

A partir dos resultados obtidos com o **BLASTp** foi feito o alinhamento das sequências através do servidor online **Clustal Omega** (Figura 22).

**Figura 22** - Alinhamento entre as sequências da fosfomanose isomerase de *Homo sapiens* e as sequências selecionadas feita pelo programa **Clustal Omega**.

PMI	-----MAAPRVFPLSCAVQQYAWGKMGNSSEVARLLASSDPLAQIAEDKPY	46
3H1M:A	-----ASMQKLINSVQNYAWGSKTALTELYG-----IANPQQQPM	35
5NW7:A	MRGSHHHHHHGMASMSSEKLFRIQCGYQNYDWGKIGSSSAVAQFVHNSDPSITIDETKPY	60
	: : . *:* ** . : : : : : *	
PMI	AELWMGTHPRGDAKILDNRISQKTLsq-wIAENQDSLGSVKVDTF--NGNLPFLFKVLSV	103
3H1M:A	AELWMGAHPKSSSRITTANGETVSLRDAIEKNKTAMLGEAVANRF---GELPFLFKVLCA	92
5NW7:A	AELWMGTHPSVPSKAID--LNN-QTLRDLVTAKPQEYLGESIIITKFGSSKELPFLFKVLSI	118
	*****:* * : : . : * : : * : * : * :*****.	
PMI	ETPLSIQAHPNKELAEKL-----HLQAPQHYPDANHKPEMAIALTPFQGLCGFRP	153
3H1M:A	AQPLSIQVHPNKRNSEIGFAKENAAGIPMDAAERNYKDPNHKPELVFALTPFLAMNAFRE	152
5NW7:A	EKVLSIQAHPKKLGALQ-----HAADPKNYPDDNHKPEMAIAVTDfEGFCGFKP	168
	****.*:* . . : : * * *****.:*: * * . : . *:	
PMI	VEEIVTFLKKVPEFQFLIGDEAATHLKQTM-----SHDSQAVASSLQSCFSLMKSEK	206
3H1M:A	FSDIVSLLQPVAGAHSAIA-----HFLQ-----VPNAERLSQLFASLLNMQG	194
5NW7:A	LDQLAKTLATVPELNEIIGQELVDEFISGIKLPAEVGSQDDVNNRKLLQKVFGLMNTDD	228
	..:..* * : * . : . : . : * . * . * : : :	
PMI	KVVVEQLNLLVKRISQAAAAGNNMEDI FGELLQLHQYPGDIGCFAI-YFLNLLTLKPG	265
3H1M:A	EEKSRALAVLKAAL-----NSQQGEPWQTIRVISEYYPDDSGLFSP-LLLNvVKLNPG	246
5NW7:A	DVIKQQTAKLLERTDREPQVFKDIDSRLPELIQRlnKQFPNDIGLFCGCLLNHVGLNKG	288
	. . * : : . : : : : * . * * . : * : * :	
PMI	EAMFLEANVPHAYLKGDCVECMACSDNTVRAGLTPKFIDVPTLCEMLSYTPSSSKDRLF-	324
3H1M:A	EAMFLFAETPHAYLQGVALEVMANSNDNVLRAGLTPKYIDIPELVANVKFEPKPAGELLT-	305
5NW7:A	EAMFLQAKDPHAYISGDIIECMAASDNVVRAGFTPKFKDVKNLVEMLTYSYESVEKQKMP	348
	***** *: *****.* : * ** * : *****:* * * : : . . .	
PMI	---LPTRSQEDPYLSIYDPPVPDFTIMKTEVPGSVTEYKVL-ALDSASILLMVQGTVI-A	379
3H1M:A	---APVKSGAEL---DFPIPVDDFAFSLHDLAL---QETSIGQHSAAILFCVEGEAVLR	355
5NW7:A	LQEFPRSKGDAVKSvLYDPPIAEFSVLQTI FDKSKGGKQVIEGLNGPSIVIATNGKGTIQ	408
	* . : * : * : . . : . . . : * : : * : *	
PMI	STPTTQTPIPLQRGGVLFIGANESVSLKLTEP---KDLIFRACCLL	423
3H1M:A	---KDEQRLVLPGESAFIGADESPVNASGTGRL---ARVYNKL---	393
5NW7:A	ITGDDSTKQKIDTGYVFFVAPGSSIELTADSANQDQDFTTYRAFVEA	455
	. : . * * : . . * : : . . : . .	

Fonte: o autor

A sobreposição estrutural das duas estruturas provenientes do PDB: 5NW7 e 3H1M, feita por intermédio do programa **Maestro** e permitiu visualizar os consensos de todas as estruturas secundárias das proteínas de referência, entre hélices-alfa e fitas de folha-beta e, devido a isso, pudemos analisar se as inserções ou deleções sugeridas pelo **Clustal Omega** estavam de acordo com o esperado. Após a análise dos alinhamentos juntamente com as sobreposições das estruturas de referência, obteve-se o alinhamento final indicado na Figura 23.



**Figura 23** - Alinhamento múltiplo final das sequências de fosfomanose isomerase humana (PMI) e as respectivas às estruturas indicadas como homólogos pelo **BLASTp**. (PDB IDs 5NW7 e 3H1M).

PMI H.	-----MAAPRVFPLSCAVQYAWGKMGSNSEVARLLASSDPLAQIAEDKPY	46
3H1M:A	-----ASMQKLINSVQNYAWGSK-TALTELYG-----IANPQQQPM	35
5NW7:A	MRGSHHHHHHGMASMSSEKLFRIQCQGYQNYDWGKIGSSSAVAQFVHNSDPSITIDETKPY	60
	: : . *:* ** . : : : : : : *	
PMI H.	AELWMGTHPRGDAKILDNRISQKTLISQ-WIAENQDSLGSVKVDTF--NGNLPFLFKVLSV	103
3H1M:A	AELWMAHPKSSSRITTANGETVSLRDAIEKNKTAMLGEAVANRFG---ELPFLFKVLCA	92
5NW7:A	AELWMGTHPSVPSKAID--LNNQTLRDLVTAKPQEYLGESIIITKFGSSKELPFLFKVLSI	118
	*****:** : : . : * : : : ** . : * : *****.	
PMI H.	ETPLSIQAHPNKELAEKHLQAPQ-----HYPDANKPEMAIALTPFQGLCGFRP	153
3H1M:A	AQPLSIQVHPNKRNSEIGFAKENAAGIPMDAAERNYKDPNHKPELVFALTPFLAMNAFRE	152
5NW7:A	EKVLSIQAHDPKKLGAQLHAADPK-----NYPDDNHKPEMAIAVTDTEGFCGFKP	168
	****.*:* . . : : * * *****:.*:* * . : . *	
PMI H.	VEEIVTFLKKVPEFQFLIGDEAATHLKQTM-----SHDSQAVASSLQSCFSLMKSEK	206
3H1M:A	FSDIVSLLQP---VAGA--HSAIAHFLOVP-----NAERLSQLFASLLNMQG	194
5NW7:A	LDQLAKTLATVPELNEIIGQELVDEFISGIKLPAEVGSQDDVNNRKLQKVFGLMNTDD	228
	...:.. * * : * . : : . . * . * . * : : :	
PMI H.	KVVVEQLNLLVKRISQQAAGNMMEDIFGELLQLHQYPGDIGCFAI-YFLNLLTLKPG	265
3H1M:A	EEKSRALAVLKAALNSQQ-----GEPWQTIRVISEYYPDDSGLFSP-LLLNVLNPG	246
5NW7:A	DVIKQQTAKLLERTDREPQVFKDIDSRLPELIQRLNKQFPNDIGLFCGCLLNHVGLNKG	288
	. . * : : . : : : : *:* * * . : ** : * : *	
PMI H.	EAMFLEANVPHAYLKGDVCECMACSDNTVRAGLTPKFIDVPTLCEMLSYPSSSKDRFLF	324
3H1M:A	EAMFLFAETPHAYLQGVALEVMANSDNVLRAGLTPKYIDIPELVANVKFEPKPAGELLTA	305
5NW7:A	EAMFLQAKDPHAYISGDIIECMAASDNVVRAGFTPKFKDVKNLVEMLTYSYESVEKQKMP	348
	***** * : ***** . * : * ** *** . : *** : *** : * : * : : . .	
PMI H.	PTRSQ----EDPYLSIYDPPVDFTIMKTEVPGSVTEYKVL---ALDSASILLMVQGTVI-A	379
3H1M:A	PVKSG-----AELDFPIPVDDFA--FSLHDLALQETSIGQH----SAAILFCVEGEAVLR	355
5NW7:A	LQEFPRSKGDAVKSVLYDPPIAEFVSLQTIQFDKS--KGGKQVIEGLNGPSIVVIATNGKGTIQ	408
	* . : * : : * . . : . . . : * : : : *	
PMI H.	STPTTQTPPIPLQRGGVLFIGANE-SVSLKLTEP---KDLIFRACLL	423
3H1M:A	-----KDEQRLVLKPGESAFIGADESPVNASGT-----GRLARVYNKL	393
5NW7:A	ITGDDSTKQKIDTGYVFFVAPG--SSIELTADSANQDQDFTTYRAFVEA	455
	. : . * * : . . * : : . .	

Os aminoácidos com coloração azul escuro pertencem a folha  $\beta$  enquanto que aqueles com coloração verde pertencem às  $\alpha$ -hélices. A simbologia de asterisco (\*) representa a presença de um único resíduo totalmente conservado, dois pontos (:) indica a conservação entre grupos de propriedades fortemente semelhantes e um ponto (.) indica a conservação entre grupos de propriedades fracamente semelhantes.

Fonte: o autor

#### 4.1.3 Etapa de construção e avaliação do modelo

Após a etapa de alinhamento ser concluída, a próxima etapa foi conduzida através da construção do próprio modelo. Tal passo foi possível através do programa **MODELLER**, que foi utilizado e, assim, gerou 5 hipóteses de modelo. De posse dos modelos gerados, surge a

etapa de avaliação para determinar o modelo mais confiável e adequado para novos estudos. As análises de avaliação foram desenvolvidas com o servidor *online* **WHATIF** e o programa **PROCHECK**.

Dados os resultados obtidos pelo **WHATIF**, pode-se que temos três boas hipóteses: a 3, 4 e 5. Elas obtiveram um escore total do WHAT IF de -0.675, -0.675 e -0.690, respectivamente. Esses valores são próximos entre si, porém distantes quando comparamos com os valores obtidos para as estruturas cristalográficas (Tabela 7). Outro ponto de atenção são as regiões identificadas do gráfico Ramachandran. No caso das Hipóteses 3, 4 e 5, foram encontrados 3 e 4 resíduos em regiões proibidas (Tabela 7).

Outro ponto importante a considerar é o fator G. As hipóteses apresentaram um fator G de -0,2, um valor que está dentro da região para uma proteína com uma resolução de 1,85 Å (referente ao PDB ID5NW7) (Tabela 7).

**Tabela 7** - Apresentação dos resultados da Pontuação Geral (**WHATIF**), Regiões Proibidas (Ramachandran) e Fator G (**PROCHECK**) obtidos na etapa de validação dos modelos gerados.

<b>Modelos</b>	<b>Pontuação Geral (WHATIF)</b>	<b>Regiões Proibidas (Ramachandran Plot)</b>	<b>G-Factor</b>
<i>3HIM</i>	- 0.408	2	0.0
<i>5NW7</i>	-0.231	0	-0.1
<i>Hipótese 1</i>	-0.611	7	-0.2
<i>Hipótese 2</i>	-0.707	3	-0.2
<i>Hipótese 3</i>	-0.675	3	-0.2
<i>Hipótese 4</i>	-0.675	3	-0.2
<i>Hipótese 5</i>	0.690	4	-0.1

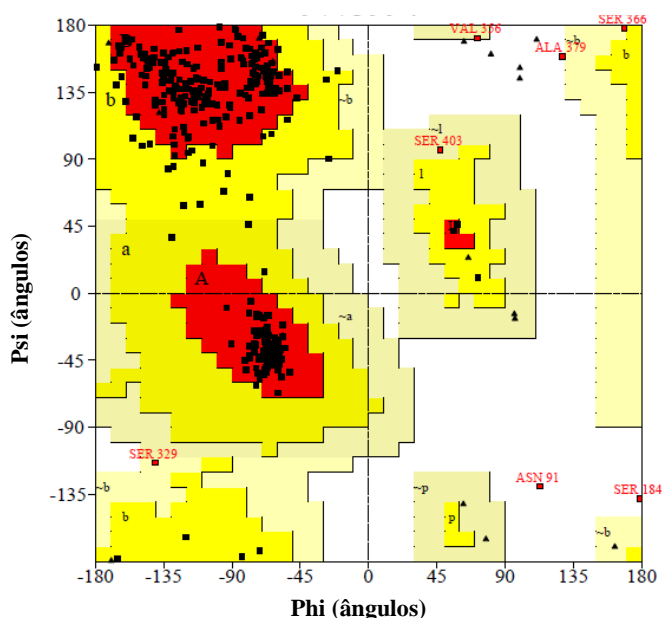
Fonte: o autor

Para que podemos realizar uma melhor seleção da hipótese escolhida, devemos analisar outros parâmetros. Em seguida foi feita a análise dos ângulos diédricos e do diagrama de Ramachandran, para poder seguir para outras análises.

#### 4.1.4 Avaliação dos ângulos diédricos $\Psi$ versus $\phi$ através do diagrama de Ramachandran

Os diagramas de **Ramachandran** obtidos para as hipóteses apresentaram alguns resíduos em regiões proibidas. As hipóteses 1 e 2 foram descartadas logo no início devido ao grande número de aminoácidos presentes em regiões proibidas, restando as hipóteses 3, 4 e 5 para análise. A hipótese escolhida foi 3, na qual apresenta 89,2% de seus resíduos em regiões altamente favoráveis e apenas 0,8% em regiões proibidas sendo os resíduos Ala (379), Asn (91) e Ser (184) (Figura 24).

**Figura 24** - Diagrama de Ramachandran para a hipótese 3 obtido pelo **PROCHECK**: 89,2% de resíduos em regiões favoráveis, 8,9% de resíduos em regiões permitidas, 1,1% de resíduos em regiões generosamente permitidas e 0,8% de resíduos em regiões proibidas.



##### Estatísticas do gráfico

Resíduos em regiões mais favoráveis [A,B,L]	330	89.2%
Resíduos em regiões adicionais permitidas [a,b,l,p]	33	8.9%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	4	1.1%
Resíduos em regiões proibidas	3	0.8%
	----	----
Número de resíduos não glicinas e não prolinas	370	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	21	
Número de resíduos de prolina	30	
	----	
Número total de resíduos	423	

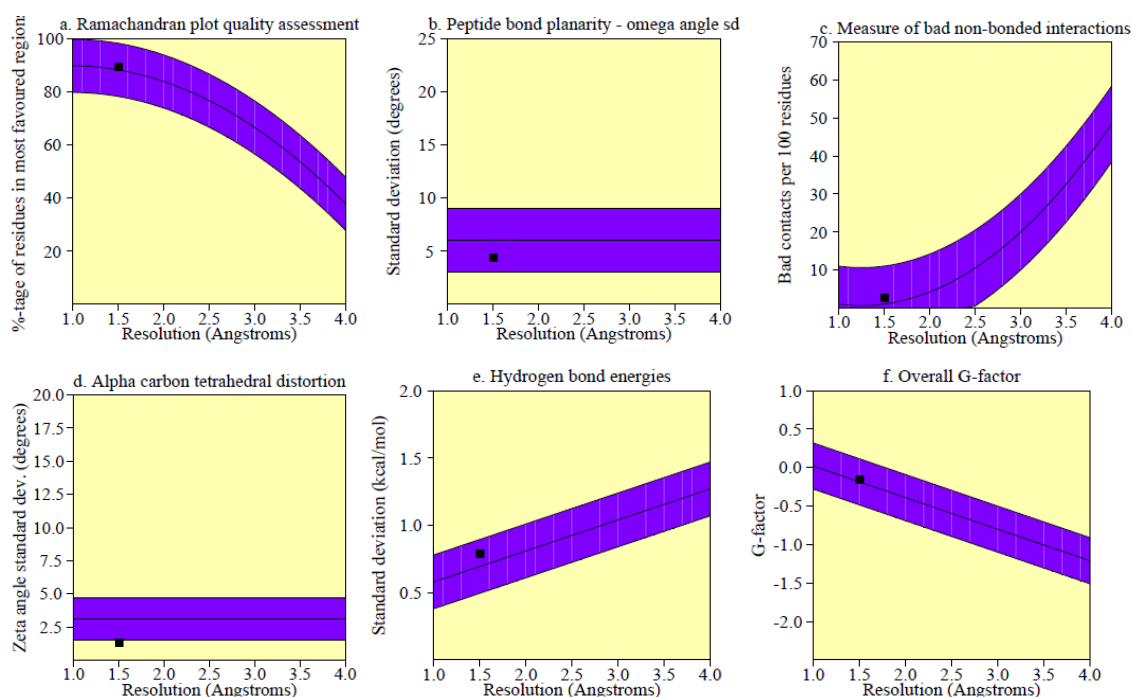
Fonte: o autor.

Embora a hipótese apresente resíduos em regiões proibidas no Ramachandran, nenhum deles se encontra na região do sítio ativo. Apenas utilizando os dados apresentados até agora, não foi possível escolher qual o melhor modelo gerado. Para que a escolha seja feita corretamente, foram utilizados outros parâmetros para a análise.

#### 4.1.5 Outros parâmetros importantes para a análise dos modelos de PMI humana

Outras ferramentas importantes para analisar um modelo de proteína predito podem ser encontradas no pacote **SAVE v5.0**. Para a hipótese selecionada, todos os valores estão dentro dos valores esperados para estruturas de proteína de resolução de 1,85 Å (referindo-se à estrutura PDB ID 5NW7) e a distorção tetraédrica dos carbonos alfa apresentou um melhor valor que o do cristalográfico (Figura 25).

**Figura 25** - Resumo da avaliação estereoquímica da hipótese 3, construída por homologia estrutural para PMI humana.

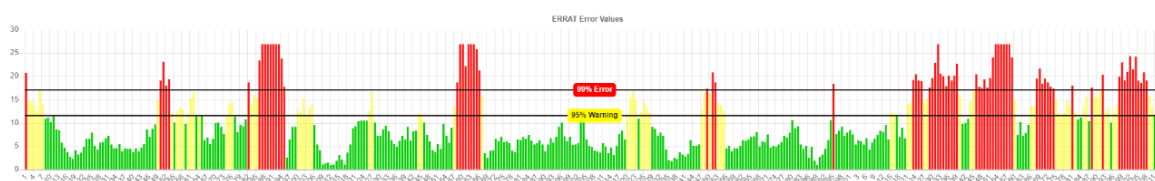


Fonte: o autor.

Além das ligações atômicas que ocorrem na cadeia peptídica, há também interações entre átomos não-ligados. A estrutura terciária de uma proteína é altamente dependente dessas interações e existem seis tipos de interações prováveis entre átomos não ligados e elas não ocorrem aleatoriamente. Assim, é possível através do programa **ERRAT** identificar regiões e fazer uma análise estatística dessas interações.

Observando o gráfico desenvolvido pela **ERRAT**, é possível identificar três regiões distintas: uma verde, uma amarela e uma vermelha. A hipótese gerada apresenta grande quantidade de aminoácidos listados como vermelho, porém esses aminoácidos não estão na região do sítio ativo. O fator de qualidade, valor relacionado às interações atômicas não ligadas, apresentados pela hipótese (61.6867) estão acima de 50, mostrando que é uma estrutura de alta qualidade (Figura 26).

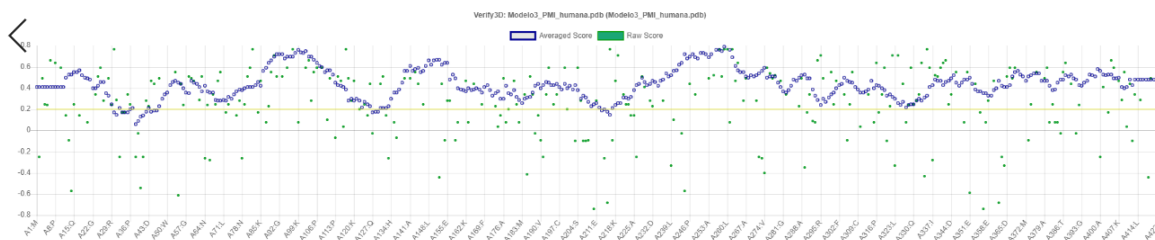
**Figura 26 - ERRAT** obtido para a hipótese 3 gerado de PMI humana.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas. Fonte: o autor.

Outra análise interessante é verificar se um resíduo de aminoácido corresponde ao ambiente em que está, ou seja, se há compatibilidade do modelo com sua sequência de aminoácidos. Essa análise pode ser realizada usando o **Verify-3D**. No estudo da hipótese 3, pode-se observar que 95,51% dos aminoácidos teve uma pontuação igual ou superior a 0,2 no perfil 3D-1D (Figura 27). Além disso, há regiões que caiu abaixo do limite, indicando uma provável dobragem nessa região; no entanto, não é motivo de preocupação porque não está na região do site ativo das hipóteses.

**Figura 27 - Verify-3D** obtido para a hipótese 3 da PMI humana: 95,51% dos aminoácidos encontram-se dentro do perfil esperado.

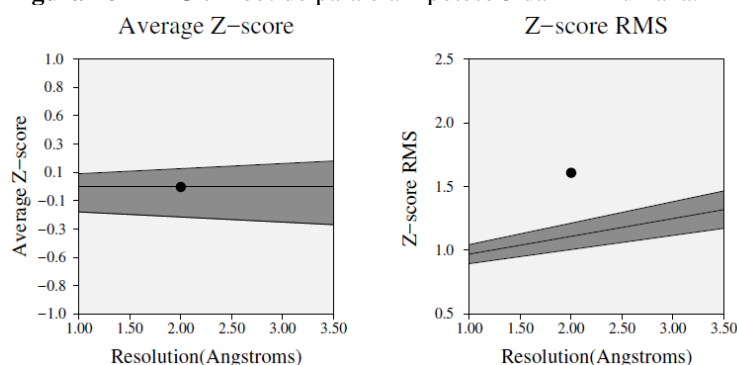


Fonte: o autor.

Finalmente, como última análise, é possível verificar o volume de átomos. O programa **PROVE** correlaciona o volume obtido para cada aminoácido da proteína com os volumes de

proteína de alta resolução. Como resultado, o **PROVE** fornece um valor de pontuação chamado escore Z que corresponde ao desvio estatístico programado de estruturas altamente resolvidas e refinadas. No caso da hipótese escolhida do modelo PMI em estudo, o Z-score encontra-se dentro da faixa esperada e o Z-score RMS acima da região de interesse (Figura 28). No entanto, juntamente com todas as outras análises realizadas, pode-se determinar que se trata da melhor hipótese de modelo para a PMI humana.

**Figura 28 - PROVE** obtido para o a hipótese 3 da PMI humana.



*O escore Z corresponde aos desvios dos valores do volume atômico em relação aos valores padrão enquanto o Z-score rms mede a média de volumes irregulares na estrutura da proteína.*

Fonte: o autor.

#### 4.1.6 Etapa de validação do modelo de PMI humana

Uma vez selecionado e refinado o modelo final, avaliamos sua eficiência em estudos de triagem virtual. Para essa avaliação, foram escolhidos 30 inibidores conhecidos da PMI, depositados e provenientes do servidor online **BindingDB**, contendo valores descritos de potência ( $IC_{50}$ ), reportados na literatura. O objetivo dessa etapa era avaliar a capacidade que o modelo possui de recuperar os ligantes sabidamente ativos (ditos “contaminantes”) e ordená-los de forma coerente em relação ao seus  $IC_{50}$ , com base em função de pontuação (**GOLD Fitness**) proveniente do programa **GOLD**.

Para que o objetivo fosse alcançado, primeiramente era necessária a definição do sítio de *docking*, uma esfera com as seguintes coordenadas espaciais:  $x = 20.21$ ,  $y = 16.63$ ,  $z = 10.00$ , com uma esfera de raio de 7 Å. Os resultados dessa etapa são mostrados na Tabela 8.

**Tabela 8 - Resultado da recuperação da seleção de inibidores do *BindingDB*.**

<b>Número de Inibidores</b>	<b>Inibidores</b>	<b>GOLD PLP Fitness</b>	<b>IC50 (nM)</b>
1	2998795	76,169	1.30E+03
2	5771379	76,061	1.00E+03
3	4089709	72,416	707
4	25199560	67,963	>1,00E+6
5	25199559	67,128	>1,00E+5
6	2248499	66,702	829
7	4917291	61,567	709
8	1256813	60,232	1.00E+03
9	3311395	58,643	>1,00E+5
10	1489253	58,532	>1,00E+5
11	1519965	58,281	1.00E+03
12	4120895	58,255	>1,00E+5
13	2311082	57,179	270
14	25181243	56,910	1.03E+03
15	1510389	54,689	985
16	25241622	54,330	>1,00E+5
17	387196	53,533	>1,00E+5
18	3093671	53,365	1.27E+03
19	22416235	52,628	1.25E+03
20	3194	51,576	190
21	4879298	46,232	>1,00E+5
22	25181212	45,228	>1,00E+5
23	25181239	44,972	>1,00E+5
24	6112173	44,039	>1,00E+5
25	22765927	43,708	>1,00E+5
26	25241626	43,102	>1,00E+5
27	25241624	42,484	>1,00E+5
28	25199533	41,919	>1,00E+5
29	6033619	40,274	>1,00E+5
30	5951502	40,144	>1,00E+5

*Os inibidores mais potentes estão em verde enquanto os menos potentes estão em vermelho.*

Fonte: o autor.

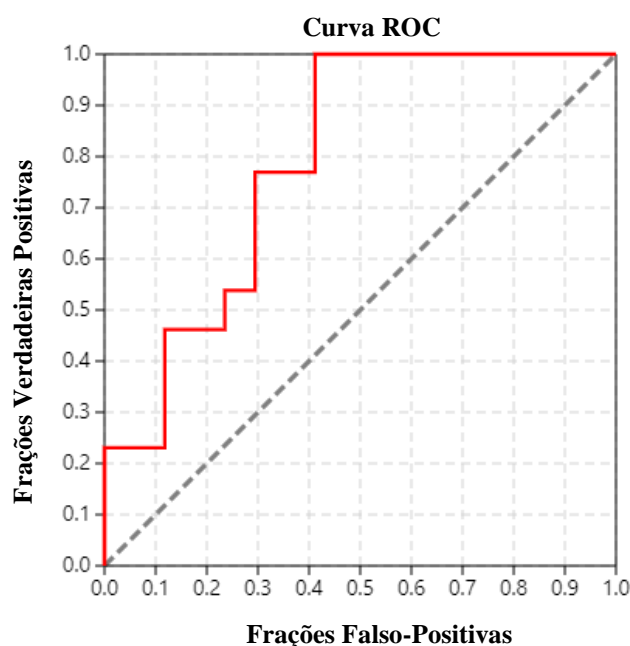
Observando os dados obtidos pelo estudo de triagem virtual, o próximo passo consistiu na obtenção de curvas ROC, de predição e de enriquecimento para melhor entendimento dos resultados. Para a criação desses gráficos, foi utilizado o *web servidor* Screening Explorer (EMPEREUR-MOT; ZAGURY; MONTES, 2016) com os parâmetros selecionados como padrões.

Em estudos de triagem virtual, a análise dos dados é feita com base em uma curva ROC (Receiver Operating Characteristic), resumem a capacidade geral de um método para

distncluir entre compostos ativos e inativos (EMPEREUR-MOT et al., 2015; EMPEREUR-MOT; ZAGURY; MONTES, 2016). Utilizando os mesmas análises feitas no item anterior, quanto maior a distância entre a linha do gráfico e a linha diagonal, melhor será a precisão geral do estudo (BERTRAND; TRIBALLEAU, 2005; EMPEREUR-MOT et al., 2015; EMPEREUR-MOT; ZAGURY; MONTES, 2016).

Abaixo, na Figura 29, é mostra o gráfico de curva ROC obtido para o estudo realizado com os 30 inibidores do BindingDB e o modelo de PMI *Homo sapiens* gerado. Nele é possível observar que grande parte do gráfico encontra-se distante da linha diagonal, mostrando que o estudo apresnta uma boa precisão, conseguindo diferenciar os compostos mais ativos dos menos ativos. Inclusive uma parte da linha pertencente ao gráfico encontra-se na parte superior, mostrando que esses inibidores encontram-se com a sensibilidade e especificidade perfeitas.

**Figura 29** - Curva ROC obtida para o estudo dos 30 compostos relatados como ativos e inativos no **BindingDB** para a proteína PMI de *Homo sapiens*.



Fonte: o autor.

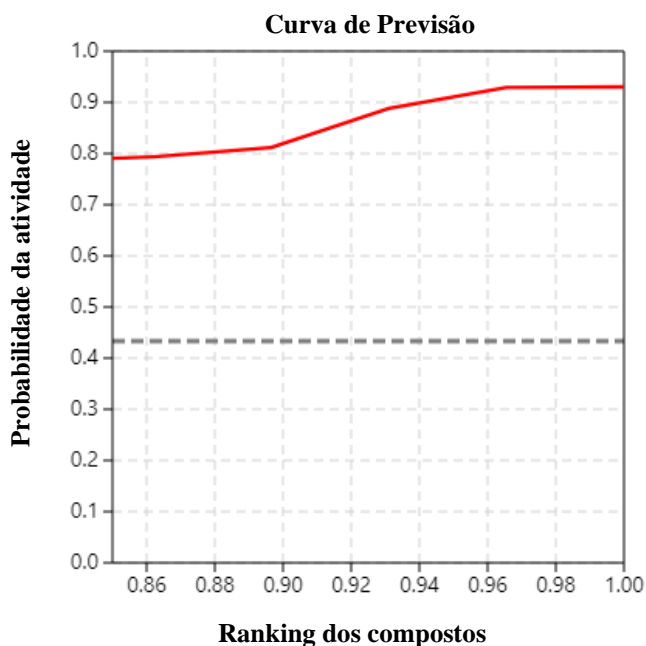
Outro dado que pode ser utilizado é a AUC, que mostra o desempenho geral do estudo. O valor de AUC ROC obtido para esse estudo foi de 0.792, um bom valor para o estudo realizado já que se encontra próximo a 1 (valor ideal).

Um outro dado obtido pelo *web server* é a curva de previsão, que quantifica se as variações de probabilidade de atividade são importantes o suficiente para induzir a seleção de um limiar para triagens virtuais (EMPEREUR-MOT et al., 2015). Na Figura 30 é mostrado a



curva de previsão para esse estudo e ela se encontra acima da linha limite, tornando-se um bom resultado e mostrando que o modelo foi capaz de realizar uma discriminação entre os compostos utilizados.

**Figura 30** - Curva de Previsão obtida para o estudo dos 30 compostos relatados como ativos e inativos no **BindingDB** para a proteína PMI de *Homo sapiens*.



Fonte: o autor.

Um dado calculado a partir da curva de previsão é o ganho total padronizado (TG) (BURA; GASTWIRTH, 2001), que resume a discriminação dos compostos ativos em relação à variação das pontuações ao longo do conjunto de dados das moléculas (EMPEREUR-MOT et al., 2015).

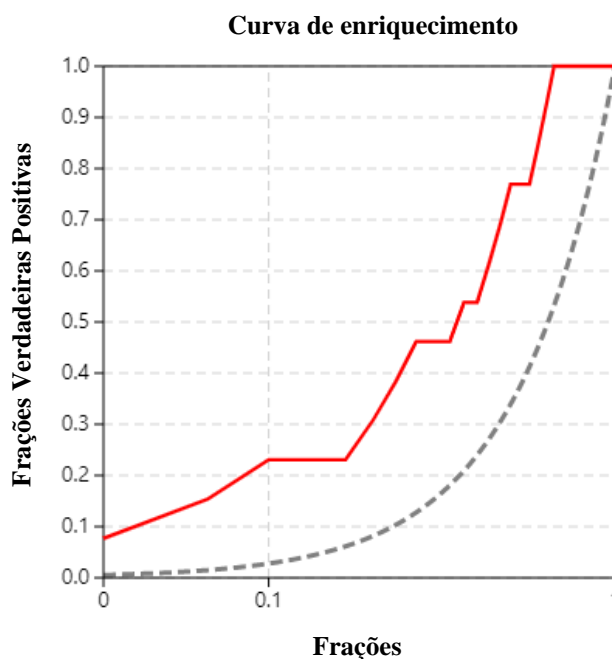
A TG possui seu valor entre 0 e 1, sendo que sua análise deve estar associado ao valor obtido para a AUC. Valores de TG acima de 0,4 associados com uma AUC ROC acima de 0,5 indica, que o método de triagem utilizado teve um bom desempenho e seria reproduzível em condições experimentais semelhantes (EMPEREUR-MOT; ZAGURY; MONTES, 2016).

No nosso estudo, o valor de TG obtido foi de 0,444 que, associado a um AUC ROC de 0,792, mostra que o método utilizado teve um bom desempenho, podendo ser reproduzível em condições experimentais semelhantes.

Por fim, um ultimo dado obtido é a curva de enriquecimento, a qual permite avaliar a recuperação precoce de compostos ativos presentes em estudos de triagem virtual. Assim,

quanto mais distante a curva gerada estiver da linha diagonal, melhor é o estudo realizado (BERTRAND; TRIBALLEAU, 2005). Na Figura 31 é mostrado a curva de enriquecimento para o estudo em questão.

**Figura 31** - Curva de enriquecimento com um fator de enriquecimento de 10% obtida para o estudo dos 30 compostos relatados como ativos e inativos no **BindingDB** para PMI de *Homo sapiens*.



Fonte: o autor.

Um dado relacionado à curva de enriquecimento é o fator de enriquecimento (EF), que corresponde à fração de compostos ativos obtidos quando uma determinada porcentagem de compostos inativos é encontrada (TRUCHON; BAYLY, 2007) Nesse estudo, foi considerado a porcentagem de 10% dos compostos inativos e o valor obtido para EF foi de 2.31.

Além dessas curvas, temos alguns outros dados para levar em consideração: discriminação reforçada por Boltzmann do ROC (BEDROC) (MUEGGE; ENYEDY, 2004) e o reforço inicial robusto (RIE) (CORNELL, 2006). Ambas as métricas são utilizadas para quantificar o reconhecimento precoce dos compostos ativos utilizados em um estudo de triagem virtual (EMPEREUR-MOT; ZAGURY; MONTES, 2016). O valor obtido para o RIE foi 2,069; enquanto que o valor obtido pra BEDROC foi 0,897 encontra-se muito próximo a 1, mostrando que o modelo é capaz de separar os compostos ativos e inativos.

Com isso, através desse resultado de validação e daqueles de avaliação mostrados anteriormente, podemos afirmar que o modelo para a PMI *Homo sapiens* provou ser capaz de diferenciar os inibidores mais potentes dos menos potentes escolhidos no **BindingDB**.

#### 4.3 Triagem virtual par seleção de inibidores da PMI de XAC.

Um dos primeiros e mais importantes passos nos estudos de planejamento e desenvolvimento de moléculas *in silico* são as conformações que os ligantes podem assumir. A posição e distância que os grupos funcionais assumem em relação aos resíduos de aminoácidos dentro do alvo molecular são fundamentais para que ocorra uma interação entre eles e, conseqüentemente, a formação de um complexo (LAPLANTE et al., 2014). Assim, avaliar e explorar o máximo de conformações que um ligante pode assumir torna-se uma atividade essencial em diversas metodologias computacionais (BAI et al., 2010).

O primeiro passo foi preparar as 6 bases de dados, descritas anteriormente, com o programa **OMEGA**. Esse passo tem o intuito de corrigir e criar diversas conformações das moléculas presentes nas bases de dados. Esse programa realiza a geração de confôrmeros em três dimensões com baixo desprendimento computacional e um alto efeito no alcance da conformação bioativa entre os confôrmeros gerados. O **OMEGA** obtém os confôrmeros em três passos. O primeiro passo, para a obtenção dos confôrmeros, consiste ao acesso à biblioteca de fragmentos, gerada através do fragmento de moléculas presente em um extenso banco de dados, para a construção da estrutura em três dimensões (3D). Os fragmentos são ligados com base em valores de distância de ligações, sendo ajustados posteriormente e minimizados por um campo de força ajustado.

Em seguida, todas as torções rotativas são numeradas de acordo com valores de um banco dados presente no programa. As rotações são ajustadas com base em ângulos e energias torcionais derivadas de um conjunto de estruturas cristalinas. Novas torções são geradas no ligante, prestando atenção a choques e duplicações, com a finalidade de gerar uma grande quantidade de confôrmeros.

Enfim, na última fase, ocorre a análise da estrutura em relação aos seus valores energéticos e geometria. Nessa etapa, a molécula pertencente ao mínimo global tem o seu RMSD comparado à outras, eliminando aquelas não pertencentes a parametrização fornecida pelo usuário (HAWKINS et al., 2010; HAWKINS; NICHOLLS, 2012).

Após a preparação das bases de dados, um tipo de triagem virtual foi realizado, o qual é baseado em diferentes tipos de similaridade: de forma e eletrostática. Essa metodologia explora o conceito de similaridade, ou seja, a triagem virtual em bases de dados de compostos

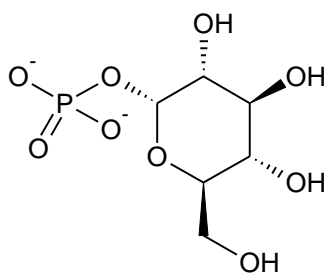
será realizada com o intuito de explorar diversidade química e selecionar novos compostos (e novas classes químicas) que sejam similares em forma e eletrostática ao substrato presente a estrutura cristalográfica PDB ID 1H5R.

O posicionamento de um ligante em um sítio ativo de uma biomolécula é um fato determinante para a atividade biológica do mesmo, já que esse sítio ativo apresenta um determinado formato para que a interação ocorra. Dessa forma, a estrutura tridimensional de um ligante é importante para que ocorra a formação do complexo ligante-proteína (HAIGH et al., 2005). Ambos os programas utilizados na etapa de triagem por similaridade e forma, **ROCS** e **EON**, da **OpenEye**, tem revelado compostos promissores para o tratamento de várias doenças (NAYLOR et al., 2009).

O programa **ROCS** realiza uma triagem virtual por forma, através da comparação entre os volumes e conformação de dois compostos, a molécula de referência (G1P) (Figura 32) e as moléculas presentes em bancos de dados. Os cálculos dos volumes são feitos através da superfície exposta de esferas de inserção colocadas na molécula referência.

Após o **ROCS**, uma outra triagem de similaridade foi feita comparando os mapas de potencial eletrostático das moléculas. O programa utilizado nessa etapa foi o **EON**, sendo que seus arquivos de entrada foram os provenientes do **ROCS**. Com esse programa se realiza a comparação e classificação dos compostos bases em seus mapas de superfície eletrostática. Os parâmetros adotados para ambos programas não foram alterados.

**Figura 32** - Estrutura da G1P, utilizada como referência para os estudos de triagem virtual por similaridade nos programas **ROCS** e **EON**.



Fonte: O autor.

Para a obtenção dos melhores inibidores possíveis para a PMI de XAC, foram utilizadas três estratégias. A primeira estratégia corresponde a apenas a etapa de *docking* com o programa **GOLD** (8Å) utilizando o modelo por homologia como receptor, seguido por predições ADME/TOX com os programas **QikProp** e **DEREK** respectivamente. A segunda estratégia inicia-se com uma triagem virtual por similaridade utilizando os programas **ROCS** e **EON**,

seguido pelo *docking* e, novamente, predições ADME/Tox. Por fim, na terceira estratégia, a triagem virtual por similaridade é realizada após o *docking*; seguindo para análises de predições ADME/Tox. Em todas as estratégias realizadas, o último passo corresponde à análise visual da interação dos compostos com o receptor.

Na primeira estratégia, as bases de dados foram submetidas ao *docking* utilizando o centroide de  $x = 23.977$ ,  $y = 15.659$  e  $z = 43.886$  e um raio de esfera de 8 Å para as bases de dados Chembridge CL, Chembridge EXP, DrugDatabase, DrugLike, *FDA Binding*, *Maybridge* e *Natural Stock*. Os arquivos finais de cada base de dados saíram com 500 compostos e foram utilizados em estudos de predições ADME/TOX. Em seguida, esses compostos passaram por uma análise visual para verificar quais possuíam uma maior interação com o modelo da PMI XAC. Abaixo (Tabela 9) são mostrados quantos compostos sobreviveram a todas as etapas.

**Tabela 9** - Relação entre os compostos sobreviventes das bases de dados após simulações de *docking* com esfera de 8 Å, predições ADME/Tox (**DEREK** e **QikProp**) e inspeção visual, utilizando a primeira estratégia.

Bases de dados	<i>Docking</i>	DEREK	QikProp	Inspeção Visual
<b><i>Chembridge CL</i></b>	500	155	5	0
<b><i>Chembridge EXP</i></b>	500	226	4	0
<b><i>DrugDatabase</i></b>	500	83	41	0
<b><i>FDA Binding</i></b>	500	92	28	5
<b><i>Maybridge</i></b>	500	5	5	0
<b><i>Natural Stock</i></b>	500	93	91	0

Fonte: o autor.

Na segunda estratégia, a primeira triagem feita foi em relação à similaridade por forma entre a molécula referência e as bases de dados. Nessa triagem foi realizada com o programa **ROCS** e foram obtidos os 2000 melhores resultados para as bases de dados *Chembridge CL*, *Chembridge EXP*, *DrugDatabase*, *DrugLike*, *FDA Binding*, *Maybridge* e *Natural Stock*. Os compostos selecionados nessa etapa seguiram para os estudos de similaridade por mapa eletrostático com o programa **EON**, no qual foram selecionados 1000 melhores resultados.

Em seguida, os arquivos obtidos na etapa anterior foram submetidos ao *docking* utilizando o centroide de  $x = 23.977$ ,  $y = 15.659$  e  $z = 43.886$  e um raio de esfera de 6 Å e 60 de fitness. O tamanho da esfera teve que ser modificado devido a problemas encontrados para alguns ligantes em relação a sua torção e aumento do tamanho das ligações. Os arquivos finais de cada base de dados saíram com 500 compostos e foram utilizados em estudos de predições ADME/Tox. Em seguida, esses compostos passaram por uma análise visual para verificar quais

possuíam uma maior interação com o modelo da PMI de XAC. Abaixo (Tabela 10) são mostrados quantos compostos sobreviveram a todas as etapas.

**Tabela 10** - Relação entre os compostos sobreviventes das bases de dados após estudos triagem virtual por similaridade (**ROCS** e **EON**), de *docking*, predições ADME/Tox (**DEREK** e **QikProp**) e inspeção visual, utilizando a segunda estratégia.

Bases de dados	ROCS	EON	<i>Docking</i>	DEREK	QikProp	Inspeção Visual
<b>Chembridge CL</b>	2000	1000	500	8	0	0
<b>Chembridge EXP</b>	2000	1000	500	1	0	0
<b>DrugDatabase</b>	2000	1000	500	13	13	0
<b>DrugLike</b>	2000	1000	500	265	83	38
<b>FDA Binding</b>	2000	1000	500	31	18	1
<b>Maybridge</b>	2000	1000	0	0	0	0
<b>Natural Stock</b>	2000	1000	500	3	3	1

Fonte: o autor.

Na terceira e final estratégia, as bases de dados *Princeton*, *IBS Natural* e *IBS Synthetic* foram submetidas ao *docking* utilizando o centroide de  $x = 23.977$ ,  $y = 15.659$  e  $z = 43.886$  e um raio de esfera de 6 Å e 60 de fitness. Nessa etapa foram obtidos 5000 compostos de cada base de dados. O próximo passo corresponde a busca de compostos com similariedade, por forma e eletrostática, utilizando o arquivo proveniente do **GOLD**. Nessa triagem foram utilizados os programas **ROCS** e **EON**, do qual foram obtidos os 3000 e 500 melhores resultados respectivamente.

Os arquivos finais de cada base de dados obtiveram como *hits* 500 compostos (provenientes do **EON**) e foram utilizados em estudos de predições ADME/Tox. Em seguida, esses compostos passaram por uma análise visual para verificar quais possuíam uma maior interação com o modelo da PMI de XAC. Abaixo (Tabela 11) são mostrados quantos compostos sobreviveram a todas as etapas.

**Tabela 11** - Relação entre os compostos sobreviventes das bases de dados após simulações de *docking* com esfera de 6 Å e 60 de fitness, triagem virtual por similaridade (**ROCS** e **EON**), predições ADME/Tox (**DEREK** e **QikProp**) e inspeção visual, utilizando a terceira estratégia.

Bases	<i>Docking</i>	ROCS	EON	DEREK	QikProp	Inspeção Visual
<i>Princeton</i>	5000	3000	500	105	9	0
<i>IBS Natural</i>	5000	3000	500	0	0	0
<i>IBS Synthetic</i>	5000	3000	500	135	4	0

Fonte: o autor.

As análises de propriedades farmacocinéticas são extremamente importantes para o desenvolvimento de novos compostos, pois o fato de algumas moléculas não apresentarem essas propriedades adequadas são um dos principais insucessos em pesquisas voltadas para essa área. A descoberta dessas propriedades antes dos ensaios biológicos, permite um menor gasto de tempo e dinheiro para substâncias biologicamente ativas (BOOBIS et al., 2002; PRENTIS; LIS; WALKER, 1988; VENKATESH; LIPPER, 2000).

O cálculo de descritores é um método de predição de atividades farmacocinéticas e um importante filtro em estudos de triagem virtual. Esses cálculos podem ser realizados através do programa **QikProp** (SCHRODINGER, [s.d.]). O programa **Qikprop**, da empresa **Schrodinger** realiza a predição de propriedades farmacocinéticas (ADME) utilizando descritores e com baixo desprendimento computacional. A estrutura da molécula é analisada como um todo e a predição é baseada em sua estrutura 3D.

Diversas importantes propriedades farmacocinéticas podem ser preditas e comparadas com um extenso banco de dados. Entre essas propriedades, destacamos aquelas utilizadas: *log CACO2*, *log MDCK*, porcentagem de absorção oral e qualidade do modelo de absorção oral humana.

O *log CACO2* é uma medida de absorção pela mucosa intestinal por difusão passiva, realizado através de estudos em células do epitélio intestinal diferenciadas, Caco-2 (QPPCaco). A linhagem celular Caco-2 é derivada de um carcinoma colo-retal humano e é fonte de para a absorção de fármacos *in-vitro* (ARTURSSON; KARLSSON, 1991). Junto com o modelo de células Caco-2, a permeabilidade das células MDCK (*Madin-Darby canine kidney*) também desempenham um importante fator a ser analisado no estudo de permeabilidade de fármacos *in-vitro* (HORSTER; STOPP, 1986; IRVINE et al., 1999), sendo um importante mimetizador da barreira sangue-cérebro.

Os valores utilizados como critério para a eliminação dos compostos nos estudos de triagem virtual foram: *log CACO2* e *log MDCK* acima ou igual a 500, porcentagem de absorção oral maior ou igual a 80% e qualidade do modelo de absorção oral humana igual a “alta”.

As predições toxicológicas foram calculadas através do programa **DEREK**, um programa em que a toxicidade em moléculas é baseado no conhecimento. A estrutura total da molécula é analisada por similaridade 2D e indica quais possíveis riscos toxicológicos que as estruturas possam ter. Além disso, também é feita uma comparação entre grupos toxicóforos, isto é, fragmentos da molécula e fragmentos presentes em bancos de dados. Dentre as predições toxicológicas mostradas por esse programa temos a mutagenicidade, carcinogenicidade, sensibilização da pele, irritação, entre outras. Essa etapa permite excluir moléculas potencialmente nocivas para futuros ensaios (COLE; BURSAVICH, 2010). Esse programa permite escolher a espécie para qual as predições serão feitas e, nesse estudo, as predições foram feitas para humanos enquanto o critério de exclusão foi baseado em qualquer alerta de toxicidade segundo Custom Prediction e Lhasa Prediction.

Após as análises farmacocinéticas e toxicológicas, todas as bases foram submetidas à uma análise visual, que possui o intuito de verificar uma maior interação com o modelo da PMI XAC através da quantidade de interações químicas. As análises dos resultados gerados serviram como critério final para a seleção de compostos, sendo um total de 45 inibidores. Ao final dos estudos, a relação dos compostos sobreviventes em cada base de dado são mostrados nas tabelas acima.

Os programas de *docking* analisam a afinidade ligante-proteína em estudos de avaliação. Essa etapa do estudo teve sua eficiência aumentada graças à evolução do processamento computacional e automação de buscas sistemáticas (JAIN, 2006; KITCHEN et al., 2004; YURIEV; AGOSTINO; RAMSLAND, 2011; YURIEV; RAMSLAND, 2013). O posicionamento dos ligantes em cavidades dos receptores é realizada por um algoritmo, sendo no caso do programa **GOLD**, o algoritmo genético.

Classificar as poses mais efetivas baseadas em uma pontuação também trata-se de uma grande dificuldade encontrada nos programas de *docking* (JAIN, 2006; KITCHEN et al., 2004; YURIEV; AGOSTINO; RAMSLAND, 2011; YURIEV; RAMSLAND, 2013). Há diferentes funções de pontuações adotadas por cada programa e, algumas delas, são baseadas no conhecimento, consenso, dados experimentais além da combinação de mecânica quântica e mecânica molecular (JAIN, 2006; KITCHEN et al., 2004; YURIEV; AGOSTINO; RAMSLAND, 2011; YURIEV; RAMSLAND, 2013).

Os resultados dos simulações de *docking* nos permitiram visualizar como os compostos se comportam no sítio ativo do modelo. Sendo assim, a seleção final se baseou prioritariamente na possibilidade dos compostos interagirem com os resíduos Arg173 Asn192, Asp249, Asp270



e Gly155; sendo as interações com os aminoácidos arginina (173) e ácido aspártico (249) as mais importantes.

O programa **GOLD** indicou que alguns compostos podem interagir de forma satisfatória no sítio ativo do modelo e, dessa forma, elegemos estes compostos como os mais promissores. Foram selecionados 1 composto da base de dados *Natural Stock*, 6 compostos da *FDA Binding* e 38 compostos da *DrugLike*.

De todos os compostos selecionados, alguns foram comprados e em seguida serão submetidos a ensaios *in vitro* e *in vivo*, para a comprovação da ação inibidora da enzima PMI de *Xanthomonas citri* subsp. *citri*. Doze inibidores foram selecionados para compra: **D05, D44, D28, D90, D02, D67, D22, D85, D39, D17, D27 e D89**; além disso também foi comprado um inibidor de PMI humana para ser testado. Esses inibidores que partiram para os ensaios biológicos foram selecionados com base nas interações com aminoácidos interessantes que compõem o sítio ativo e quantidade dessas interações; além de serem moléculas com perfis polares e não muito rígidas (Tabela 12).

**Tabela 12** - Tipo e comprimento de interações entre os compostos comprados e os aminoácidos presentes no sítio ativo.

Composto	Aminoácidos	Tipo de interação	Comprimento da interação (Å)
D05	Gly155 / Asn192	Ligações de hidrogênio	2,08 e 1,68
	Glu 173 / Asp249	Iônica	2,97 e 2,94
D44	Gly 153 / Gly155	Ligações de hidrogênio	2,00 e 2,99 / 2,10
	Asn192		2,77 e 2,91
D28	Gly155 / Glu 173	Ligações de hidrogênio	2,04 e 1,83
	Asn192 / Asp249		1,73 e 2,10 / 2,04 e 2,10
D90	His116 / Gly155	Ligações de hidrogênio	2,19 / 2,08
	Glu173 / Asp249		2,05 / 1,69
D02	Gly155 / Glu173	Ligações de hidrogênio	2,07 / 2,06
	Asn192		2,01, 2,03 e 2,10
D67	Gly155 / Asp249	Ligações de hidrogênio	1,85 / 2,60
	Asn192		1,79 e 1,96
D22	Gly155 / Glu173	Ligações de hidrogênio	1,82 / 1,72
	Asn192 / Ser193		1,69 e 2,17 / 2,18
D85	His116 / Gly155	Ligações de hidrogênio	2,10 / 2,08
	Asn192		1,73 e 2,10
D39	Gly155 / Glu173	Ligações de hidrogênio	2,07
	Asn192		1,70 e 2,00

D17	Gly153 / Gly155	Ligações de hidrogênio	2,10 e 2,11 / 2,06
	Asn192 / Asp249		1,99 e 2,03 / 2,06
D27	Gly155 / Glu173	Ligações de hidrogênio	2,02 / 2,10
	Asn192 / Asp270		1,66 e 2,07 / 2,06
D89	Gly155 / Glu173	Ligações de hidrogênio	1,88 / 1,99
	Asn192 / Asp249		1,86 e 2,10 / 2,57

Fonte: o autor

#### 4.4 Ensaio de atividade *in vitro* com a enzima PMI.

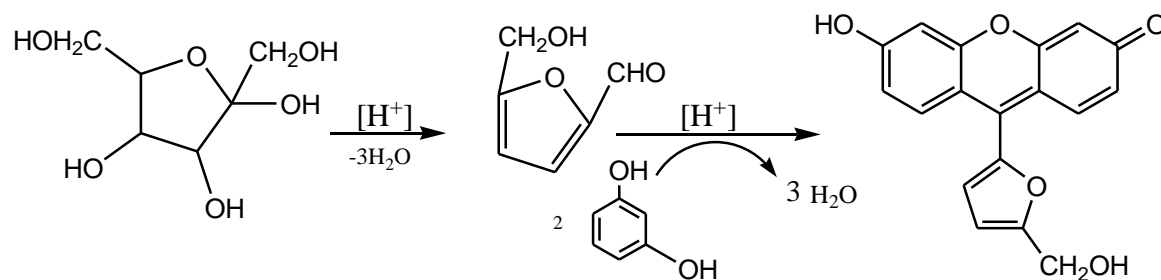
##### 4.4.1. Método de avaliação da atividade enzimática

O ensaio de atividade selecionado é o reportado por Theodor Seliwanoff em 1887, que consiste em uma reação com mudança de coloração quando utilizado frutose e resorcinol em uma solução aquosa de ácido clorídrico em fervura (SHAHIDULLAH, M; KHORASANI, 1972).

O método proposto por Seliwanoff é um teste que permite diferenciar uma aldose de uma cetose, ou seja, distinguir entre açúcares com grupamento aldeído e cetona respectivamente. Esse teste baseia-se no princípio de que, quando aquecidas, as cetoses sofrem desidratação mais rapidamente do que as aldoses. Assim, quando o reagente de Seliwanoff for utilizado em uma reação que apresenta um açúcar com grupamento cetona, a cor da solução mudará para vermelho rapidamente e quando adicionado em uma solução que tem uma aldose, a cor muda mais lentamente para o rosa (SHAHIDULLAH, M; KHORASANI, 1972).

No caso desse estudo, a frutose reagirá com o ácido clorídrico formando o furfural, o qual irá reagir com resorcinol formando um complexo de cor alaranjada pois trata-se de uma cetose (Figura 33). Já a manose, que é uma aldose, também a apresentará uma mudança de coloração, porém menos intensa do que a frutose e irá atingir uma coloração rosa.

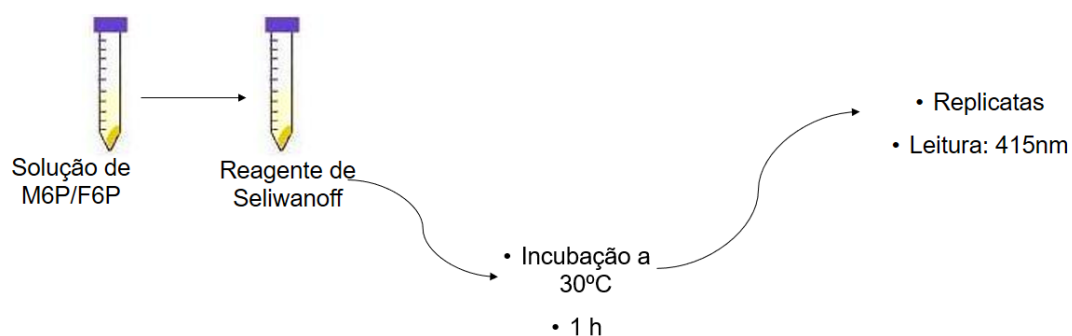
**Figura 33** – Reação química do método de Seliwanoff.



Fonte: o autor

Considerando a disponibilidade do reagente, os testes realizados para a avaliação e inibição de PMI são de caráter exploratório. O primeiro teste realizado foi para verificar se a enzima PMI estava viável para os ensaios de inibição. Nos testes-controle foram realizados com a mesma quantidade (0,456  $\mu$ g) de frutose-6-fosfato (F6P) e manose-6-fosfato (Figura 34).

**Figura 34** - Esquema do teste empregado na avaliação da atividade enzimática.



Fonte: o autor

Notou-se uma mudança rápida de coloração no controle com F6P, enquanto o controle com M6P e a reação (PMI + M6P) levaram em torno de 3 minutos para apresentarem alguma mudança em sua coloração. O primeiro tubo (aquele com a F6P) atingiu uma coloração alaranjada enquanto os outros dois apresentaram uma coloração rósea; porém, o tubo que continha a M6P e a enzima, apresentou uma coloração com intensidade maior do que a solução controle com M6P (Figura 35).

**Figura 35** - Resultado para o ensaio de avaliação da atividade enzimática da PMI mostrando as colorações obtidas para os testes-controles e a reação entre a M6P e a enzima, após a adição do reagente de Seliwanoff e fervura por 5 minutos.



Fonte: o autor

A presença da coloração mais clara no tubo de controle da reação de PMI com M6P comparada ao controle com F6P é que nem toda M6P foi convertida em F6P pois é uma reação de equilíbrio (GRACY; NOTTMANN, 1968). Com esses resultados, é possível concluir que a enzima apresentava a sua atividade de interconversão da M6P em F6P, permitindo que os testes com os inibidores pudessem seguir.

#### 4.4.2. Avaliação in vitro da inibição da PMI pelo composto selecionado

O primeiro composto selecionado para os ensaios de atividade com a enzima PMI, fornecida por Alexandrino, foi o **D28** pois ele realiza ligações entre os seus grupos hidroxilas e os aminoácidos Gly155, Glu173, Asn192 e Asp249. Esse composto foi o primeiro selecionado pois realiza ligações com os aminoácidos Asn192 e Gly155 igual o ligante cristalográfico (G1P); além de apresentar interações com os aminoácidos Glu173 e Asp249 presente no modelo PMI *Xanthomonas citri* subsp *citri* gerado.

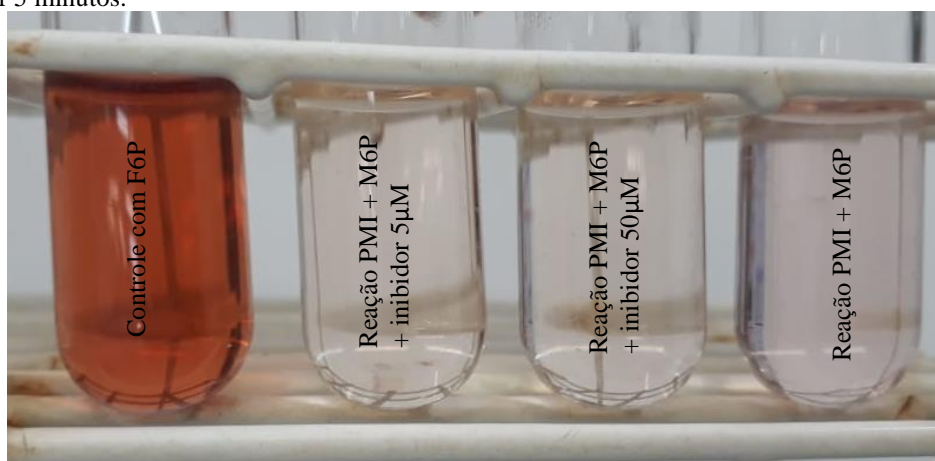
Foram testadas três concentrações finais para o inibidor **D28**: 5 $\mu$ M, 50 $\mu$ M e 50mM (Tabela 13). As concentrações de 5 $\mu$ M e 50 $\mu$ M apresentaram uma coloração rósea, porém com menor intensidade do que a solução com F6P e a solução de enzima com M6P sem inibidor (Figura 36). Para essas duas concentrações do inibidor foram medidas as absorvâncias de 0,035 e 0,028, indicando que quanto maior a concentração do composto, maior é a inibição da enzima.

**Tabela 13** - Dados referentes aos ensaios *in vitro* realizados para a fosfomanose isomerase recombinante de *Xanthomonas citri* subsp. *citri*.

REAGENTES	Volume ( $\mu\text{L}$ )					
	Controle com F6P	Controle com M6P	Reação com M6P e Concentrações finais do inibidor			
			sem inibidor	5 $\mu\text{M}$	50 $\mu\text{M}$	50 mM
<b>M6P 0,3M</b>	0	5	5	5	5	5
<b>F6P 0,3M</b>	5	0	0	0	0	0
<b>Enzima</b>	0	0	95	95	95	95
<b>Inibidor 10 <math>\mu\text{M}</math></b>	0	0	0	100	0	0
<b>Inibidor 100 <math>\mu\text{M}</math></b>	0	0	0	0	100	0
<b>Inibidor 0,1 M</b>	0	0	0	0	0	100
<b>Tampão</b>	195	195	100	0	0	0

(50 mM Tris-HCl pH 8,0, 10 mM NaCl)

**Figura 36** - Resultado para o ensaio de avaliação da atividade enzimática da PMI mostrando as colorações obtidas para os testes-controles e a reação entre a M6P e a enzima, após a adição do reagente de Seliwanoff e fervura por 5 minutos.



Fonte: o autor.

Através de cálculos simples foi possível obter a absorbância relativa e a capacidade de inibição de cada concentração utilizada (Tabela 14). A absorbância relativa foi calculada pela divisão entre absorbância obtida para o respectivo inibidor e absorbância da solução M6P +

enzima. Os resultados obtidos para 5µM e 50µM foram, respectivamente, 0,54 e 0,43 (Tabela 12). Por sua vez, considerando que a absorbância na ausência de inibidor (M6P + enzima, valor de 0,065, Tabela 12) corresponde a 100% de atividade, a porcentagem de inibição relativamente ao 100% corresponde à diferença das absorbâncias apresentadas na ausência de inibidor (0,065) e na sua presença (0,035, por exemplo). Os resultados obtidos foram 46% e 57% para as concentrações de 5µM e 50µM respectivamente (Tabela 12).

**Tabela 14** - Resultados obtidos para os ensaios de inibição da enzima PMI de *Xanthomonas citri* subsp. *citri* pelo inibidor *D28*.

<b>Concentração</b>	<b>Absorbância (415nm)</b>	<b>Absorbância relativa</b>	<b>Inibição (%)</b>
<b>F6P</b>	0.902		
<b>M6P + Enzima</b>	0.065	1	0
<b>5 µM do inibidor</b>	0.035	0,54	46
<b>50µM do inibidor</b>	0.028	0,43	57

Fonte: o autor.

Por outro lado, o ensaio com o inibidor *D28* a uma concentração de 50mM, apresentou problemas já que formou um aglomerado e não foi possível a sua detecção. Após esse fato, foi realizado um teste do inibidor com a enzima em presença do reagente de Seliwanoff e foi verificado, que em altas concentrações, o reagente utilizado no ensaio reage com o composto inibidor.

É possível notar que, embora a concentração do inibidor tenha sido aumentada em 10 vezes das concentrações de 5 µM para 50 µM, a inibição não ocorre de forma linear. Esse fato, aliado ao ensaio desse inibidor em altas concentrações, poderia indicar que o limite máximo de inibição foi atingido.

Em seguida a esse trabalho, todos os compostos comprados serão testados em relação a inibição da proteína PMI. Os resultados serão apresentados em triplicatas.

## 5. CONCLUSÃO

A estrutura tridimensional de um receptor é um dado de suma importância para químicos medicinais e computacionais que atuam na área de modelagem molecular. Explorar a região de interação do complexo ligante-receptor garante uma abordagem mais completa de estudos da área e amplia a gama de ferramentas e possibilidades a serem utilizadas.

O modelo de PMI de XAC construído se demonstrou confiável através das validações e da avaliação dos resíduos do sítio de interação. Dessa forma, tal modelo pode ser utilizado em triagens virtuais em bases de dados de compostos, na busca por novos quimiotipos ou classes químicas que poderiam interagir com o mesmo alvo, visando ao tratamento e eliminação do cancro cístico.

As estratégias de triagem virtual utilizadas foram validadas *in silico* e mostraram-se eficazes para a obtenção de diversas moléculas com potencial de inibição para a PMI. Um total de 45 compostos foram obtidos na etapa de triagem virtual, porém 12 foram comprados e passarão por ensaios biológicos para validar e quantificar suas atividades biológicas assim estimadas. Os resultados serão racionalizados do ponto de vista estrutura-atividade, com vistas a um posterior refinamento ou otimização bem como ao desenvolvimento de um futuro bactericida útil ao tratamento do cancro cístico.

## **6.PERSPECTIVAS FUTURAS**

As estratégias de triagem virtual executadas nesse estudo foram efetivas do ponto de vista teórico. Como esse Mestrado faz parte de um projeto mais amplo, novos estudos serão realizados em relação ao potencial de inibição das outras moléculas selecionadas.

As moléculas que apresentarem maior potencial de inibição, seguirão para ensaios biológicos com a bactéria *XAC* em *Citrus aurantifolia*. Após a inoculação e passados os dias necessários, as folhas serão destacadas e fotografadas para constatação e comparação do processo infeccioso da *XAC*. Após a conclusão dessa etapa, os compostos que apresentarem propriedades desfavoráveis a novos defensivos agrícolas serão otimizados de modo a apresentarem um perfil mais adequado. O objetivo é que a otimização desses compostos, após extenso estudo de relação estrutura-atividade, leve desenvolvimento de um futuro bactericida útil ao tratamento do cancro cítrico.



## REFERÊNCIAS BIBLIOGRÁFICAS

- ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403–410, 1990.
- ARTIER, J. Análise proteômica diferencial da fração periplasmática de *Xanthomonas citri* subsp. *citri*: proteínas relacionadas com a indução da patogenicidade in vitro Juliana Artier Análise proteômica diferencial da fração periplasmática de *Xanthomonas citri* subsp. 2010.
- ARTURSSON, P.; KARLSSON, J. Correlation between oral drug absorption in humans and apparent drug permeability coefficients in human intestinal epithelial (Caco2) cells. **Biochemical and Biophysical Research Communications**, v. 175, n. 3, p. 880–885, 1991.
- BAI, F. et al. Bioactive conformational generation of small molecules: A comparative analysis between force-field and multiple empirical criteria based methods. **BMC Bioinformatics**, v. 11, n. November, 2010.
- BAJORATH, J. Understanding chemoinformatics: A unifying approach. **Drug Discovery Today**, v. 9, n. 1, p. 13–14, 2004.
- BEHLAU, F. et al. Incidência e severidade de cancro cítrico em laranja “Pêra Rio” sob condições de controle químico e proteção com quebra-vento. **Fitopatologia Brasileira**, v. 32, n. 4, p. 311–317, 2007.
- BEHLAU, F. .; BELASQUE JR, J. **Cancro cítrico: A Doença e seu Controle**. 1 Ed. ed. Araraquara: Fundecitrus, 2014.
- BEHLAU, F.; BARELLI, N. L.; BELASQUE, J. Lessons from a case of successful eradication of citrus canker in a citrus-producing farm in s?o paulo state, brazil. **Journal of Plant Pathology**, v. 96, n. 3, p. 561–568, 2014.
- BEHLAU, F.; FONSECA, A. E.; BELASQUE, J. A comprehensive analysis of the Asiatic citrus canker eradication programme in São Paulo state, Brazil, from 1999 to 2009. **Plant Pathology**, 2016.
- BERTRAND, H.; TRIBALLEAU, N. Virtual Screening Workflow Development Guided by the ROC Curve Approach . Application to High-Throughput *Docking* on Metabotropic Glutamate Receptor Subtype 4 Accepting the Uncertainty ? **Receptor**, p. 2534–2547, 2005.
- BITTANCOURT, A. . O cancro cítrico. **O Biológico**, v. 23, p. 101–111, 1957.
- BLACKSHIELDS, G. et al. Sequence embedding for fast construction of guide trees for multiple sequence alignment. **Algorithms for Molecular Biology**, v. 5, n. 1, p. 1–11, 2010.
- BODNAR, A. M. et al. Complete genome sequences of three *Xanthomonas citri* strains from Texas. **Genome Announcements**, 2017.
- BONAS, U. Hrp genes of phytopathogenic bacteria. In: **Current Topics in Microbiology and Immunology**. [s.l: s.n.]. v. 192p. 79–84.
- BOOBIS, A. et al. In silico prediction of ADME and pharmacokinetics: Report of an expert meeting organised by COST B15. **European Journal of Pharmaceutical Sciences**, v. 17, n. 4–5, p. 183–193, 2002.
- BOSTRÖM, J.; GREENWOOD, J. R.; GOTTFRIES, J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. **Journal of Molecular Graphics and Modelling**, v. 21, n. 5, p. 449–462, 2003.
- BOWIE, J. U.; LTCY, R.; EISENBERG, D. A Method to Identify Protein Sequences That Fold into a Known Three-Dimensional Structure. **Science**, v. 253, n. 58, p. 164–170, 1990.
- BRENK, R. et al. Virtual screening for submicromolar leads of tRNA-guanine transglycosylase based on a new unexpected binding mode detected by crystal structure analysis. **Journal of Medicinal Chemistry**, v. 46, n. 7, p. 1133–1143, 2003.

- BURA, E.; GASTWIRTH, J. L. The Binary Regression Quantile Plot: Assessing the Importance of Predictors in Binary Regression Visually. **Biometrical Journal**, v. 43, n. 1, p. 5–21, 2001.
- BURSULAYA, B. D. et al. Comparative study of several algorithms for flexible ligand *docking*. **Journal of Computer-Aided Molecular Design**, v. 17, n. 11, p. 755–763, 2003.
- BUTTNER, D. NEW EMBO MEMBER'S REVIEW: Getting across--bacterial type III effector proteins on their way to the plant cell. **The EMBO Journal**, v. 21, n. 20, p. 5313–5322, 2002.
- CARDOSO, J. E. Cancro Cítrico. **The Plant Health Instructor**, 2013.
- CARLSON, H. A.; MASUKAWA, K. M.; MCCAMMON, J. A. Method for including the dynamic fluctuations of a protein in computer-aided drug design. **Journal of Physical Chemistry A**, v. 103, n. 49, p. 10213–10219, 1999.
- CARNIELLI, C. M. et al. Xanthomonas citri subsp. citri surface proteome by 2D-DIGE: Ferric enterobactin receptor and other outer membrane proteins potentially involved in citric host interaction. **Journal of Proteomics**, v. 151, p. 251–263, 2017.
- CEPEA - CENTRO DE ESTUDOS AVANÇADOS EM ECONOMIA APLICADA. **PIB-AGRO/CEPEA: Agronegócio Favorece Crescimento Do Pib Nacional E Ajuda No Controle Da Inflação.**
- CHEMBRIDGE. **The Glod standard in small molecule screening libraries and building blocks.**
- CHEN, X.; LIU, M.; GILSON, M. K. BindingDB: A web-accessible molecular recognition database. **J. Combi. Chem. High-Throughput Screen**, v. 4, p. 719–725, 2001.
- CLEASBY, A. et al. The x-ray crystal structure of phosphomannose isomerase from Candida albicans at 1.7 angstrom resolution. **Nature structural biology**, v. 3, n. 5, p. 470–9, 1996.
- CODDING, P. W. **Introduction to Structure-Based Drug Design**. Dordrecht: [s.n.].
- COLE, D. C.; BURSAVICH, M. G. **Aspartic Acid Proteases as Therapeutic Targets**. [s.l.] John Wiley & Sons, 2010.
- COLE, J. C. et al. Comparing protein-ligand *docking* programs is difficult. **Proteins: Structure, Function and Genetics**, v. 60, n. 3, p. 325–332, 2005.
- COLOVOS, C.; YEATES, T. O. Verification of protein structures: Patterns of nonbonded atomic interactions. **Protein Science**, v. 2, p. 1511–1519, 1993.
- CORNELL, W. D. Recent evaluations of high throughput *docking* methods for pharmaceutical lead finding. **Annual Reports in Computational Chemistry**, v. 2, p. 297–323, 2006.
- COULIN, F. et al. Identification of Cys-150 in the Active Site of Phosphomannose Isomerase from Candida albicans. **Biochemistry**, v. 32, n. 51, p. 14139–14144, 1993.
- DA SILVA, A. C. R. et al. Comparison of the genomes of two Xanthomonas pathogens with differing host specificities. **Nature**, v. 417, n. 6887, p. 459–463, 2002a.
- DA SILVA, A. C. R. et al. Comparison of the genomes of two Xanthomonas pathogens with differing host specificities. **Nature**, v. 417, n. 6887, p. 459–463, maio 2002b.
- DA SILVA, C. H. T. DE P.; TAFT, C. A. 3D descriptors calculation and conformational search to investigate potential bioactive conformations, with application in 3D-QSAR and virtual screening in drug design. **Journal of Biomolecular Structure and Dynamics**, v. 35, n. 13, p. 2966–2974, 3 out. 2017.
- DANOS, E; BERGER, R. D.; STALL, R. E. **Temporal and Spatial Spread of Citrus Canker Within Groves** *Phytopathology*, 1984.

DASSAULT SYSTÈMES BIOVIA. **Discovery Studio Modeling Environment** San Diego, CA, USA Dassault Systèmes, , 2015.

EDDY, S. R. Accelerated Profile HMM Searches. **PLoS Computational Biology**, v. 7, n. 10, p. e1002195, 20 out. 2011.

EMPEREUR-MOT, C. et al. Predictiveness curves in virtual screening. **Journal of Cheminformatics**, v. 7, n. 1, p. 1–17, 2015.

EMPEREUR-MOT, C.; ZAGURY, J. F.; MONTES, M. Screening Explorer-An Interactive Tool for the Analysis of Screening Results. **Journal of Chemical Information and Modeling**, v. 56, n. 12, p. 2281–2286, 2016.

ESCALON, A. et al. Variations in type III effector repertoires, pathological phenotypes and host range of *Xanthomonas citri* pv. *citri* pathotypes. **Molecular Plant Pathology**, 2013.

ESWAR, N. et al. Comparative protein structure modeling using MODELLER. (Supplementary). **Current protocols in protein science / editorial board, John E. Coligan ... [et al.]**, v. Chapter 2, n. November, p. Unit 2.9, 2007.

ERENCE, C. M. et al. Recent advances in the understanding of *Xanthomonas citri* ssp. *citri* pathogenesis and citrus canker disease management. **Molecular Plant Pathology**, 2018.

FISER, A.; ŠALI, A. MODELLER: Generation and Refinement of Homology-Based Protein Structure Models. **Methods in Enzymology**, v. 374, p. 461–491, 2003.

FOMBY, P.; CHERLIN, A. J. NIH Public Access. v. 72, n. 2, p. 181–204, 2011.

FRIESNER, R. A. . et al. Glide: A New Approach for Rapid, Accurate *Docking* and Scoring. Method and Assessment of *Docking* Accuracy. **Journal of Medicinal Chemistry**, v. 47, p. 1739–1749, 2004.

FUNDECITRUS. Novos Caminhos para o Cancro Citríco. **Revista Citricultor**, p. 16, 2016.

FUNDECITRUS. **Levantamento de Doenças dos Citros: HLB, CVC e Cancro Cítrico no Cinturão Citrícola de São Paulo e Triângulo/Sudoeste Mineiro**. [s.l: s.n.].

FUNDECITRUS. **Cancro Cítrico - Medidas Essenciais de Controle**. [s.l: s.n.].

FUNDECITRUS. Mitigação de Risco. **Revista Citricultor2**, p. 16, 2017c.

GAO, H.; YU, Y.; LEARY, J. A. Mechanism and kinetics of metalloenzyme phosphomannose isomerase: Measurement of dissociation constants and effect of zinc binding using ESI-FTICR mass spectrometry. **Analytical Chemistry**, v. 77, n. 17, p. 5596–5603, 2005.

GARAMI, A.; ILG, T. Disruption of mannose activation in *Leishmania mexicana*: GDP-mannose pyrophosphorylase is required for virulence, but not for viability. **EMBO Journal**, v. 20, n. 14, p. 3657–3666, 2001.

GOHLKE, H.; HENDLICH, M.; KLEBE, G. Knowledge-based scoring function to predict protein-ligand interactions. **Journal of Molecular Biology**, v. 295, n. 2, p. 337–356, 2000.

GORDON, J. L. et al. Comparative genomics of 43 strains of *Xanthomonas citri* pv. *citri* reveals the evolutionary events giving rise to pathotypes with different host ranges. **BMC Genomics**, 2015.

GOTTWALD, T.R., MCGUIRE, R.G. & GARRAN, S. Asiatic citrus canker: spatial and temporal spread in simulated new planting situations in Argentina. **Phytopathology**, v. 78, p. 739–745, 1988.

GOTTWALD, TIM R.; GRAHAM, J. . **Cancro Cítrico**. Disponível em: <<https://www.apsnet.org/edcenter/disandpath/prokaryote/pdlessons/Pages/CitrusCankerPort.aspx>>.

GOTTWALD, T. R. et al. The Citrus Canker Epidemic in Florida: The Scientific Basis of Regulatory Eradication Policy for an Invasive Species. **Phytopathology**, v. 91, n. 1, p. 30–34, 2001.

GRACY, R. W. .; NOTTMANN, E. A. Studies on Phosphomannose Isomerase - A mechanism for catalysis and for the role of zinc in the enzymatic and the nonenzymatic isomerization. **The Journal of Biological Chemistry**, v. 243, n. 20, p. 5410–5419, 1968.

GRAHAM, J. H. et al. *Xanthomonas axonopodis* pv . *citri* : factors affecting successful eradication of citrus canker. **Molecular and Plant Pathology**, v. 5, n. 1, p. 1–15, 2004.

HAIGH, J. A. et al. Small Molecule Shape-Fingerprints. n. 1, p. 673–684, 2005.

HAWKINS, P. C. D. et al. Conformer generation with OMEGA: Algorithm and validation using high quality structures from the protein databank and cambridge structural database. **Journal of Chemical Information and Modeling**, v. 50, n. 4, p. 572–584, 2010.

HAWKINS, P. C. D.; NICHOLLS, A. Conformer generation with OMEGA: Learning from the data set and the analysis of failures. **Journal of Chemical Information and Modeling**, v. 52, n. 11, p. 2919–2936, 2012.

HORSTER, M. F.; STOPP, M. Transport and metabolic functions in cultured renal tubule cells. **Kidney International**, v. 29, n. 1, p. 46–53, 1986.

IRVINE, J. D. . et al. MDCK (Madin-Darby Canine Kidney) Cells: A Tool for Membrane Permeability Screening. v. 88, n. 1, p. 28–33, 1999.

IRWIN, J. J. .; SHOICHET, B. K. ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. **Journal of Chemical Information and Modeling**, v. 45, p. 177–182, 2005.

JAEKEN, J.; MATTHIJS, G. Congenital disorders of glycosylation. **Annu Rev Genomics Hum Genet**, n. 2, p. 129–151, 2001.

JAIN, A. N. Scoring Functions for Protein-Ligand *Docking*. **Current Protein & Peptide Science**, v. 7, n. 5, p. 407–420, 2006.

JALAN, N. et al. Comparative genomic analysis of *Xanthomonas axonopodis* pv. *citrumelo* F1, which causes citrus bacterial spot disease, and related strains provides insights into virulence and host specificity. **Journal of Bacteriology**, 2011.

JALAN, N. et al. Comparative genomic and transcriptome analyses of pathotypes of *Xanthomonas citri* subsp. *citri* provide insights into mechanisms of bacterial virulence and host range. **BMC Genomics**, 2013.

JENSEN, S. O.; REEVES, P. R. Domain organisation in phosphomannose isomerases (types I and II). **Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology**, v. 1382, n. 1, p. 5–7, 1998.

JORGENSEN, W. L. **The Many Roles of Computation in Drug Discovery** Science, 2004.

JORGENSEN, W. L. Efficient drug lead discovery and optimization. **Accounts of Chemical Research**, 2009.

KIRCHMAIR, J. et al. The Protein Data Bank (PDB), its related services and software tools as key components for in silico guided drug discovery. **Journal of medicinal chemistry**, v. 51, n. 22, p. 7021–40, nov. 2008.

KITCHEN, D. B. et al. *Docking* and scoring in virtual screening for drug discovery: Methods and applications. **Nature Reviews Drug Discovery**, v. 3, n. 11, p. 935–949, 2004.

KONTOYIANNI, M.; MCCLELLAN, L. M.; SOKOL, G. S. Evaluation of *Docking* Performance: Comparative Data on *Docking* Algorithms. **Journal of Medicinal Chemistry**, v. 47, n. 3, p. 558–565, 2004.

KRAMER, B.; RAREY, M.; LENGAUER, T. Evaluation of the FlexX incremental construction algorithm for protein- ligand *docking*. **Proteins: Structure, Function and Genetics**, v. 37, n. 2, p. 228–241, 1999.

KUNZ, I. D. Structure-based strategies for drug design and discovery. **Science**, v. 257, n. 5067, p. 1078–1082,

1992.

LAPLANTE, S. R. et al. Ligand bioactive conformation plays a critical role in the design of drugs that target the hepatitis C virus NS3 protease. **Journal of Medicinal Chemistry**, v. 57, n. 5, p. 1777–1789, 2014.

LASKOWSKI, R. A. et al. Procheck Programs to check the Stereochemical Quality of Protein Structures. **Journal of Applied Crystallography**, v. 26, p. 283–291, 1993.

LIPINSKI, C. A. Lead- and drug-like compounds: The rule-of-five revolution. **Drug Discovery Today: Technologies**, v. 1, n. 4, p. 337–341, 2004.

LIPINSKI, C.; HOPKINS, A. Navigating chemical space for biology and medicine. **Nature**, v. 432, n. 7019, p. 855–861, 2004.

LUTHY, R.; BOWIE, J. U.; EISENBERG, D. Assessment of protein models with three-dimensional profiles. **Nature**, v. 83–85, 1992.

MARCHANT, C. A. ; BRIGGS, K. A. ; LONG, A. In silico tools for sharing data and knowledge on toxicity and metabolism: derek for windows, meteor, and vitic. **Toxicology Mechanism and Methods**, v. 18, n. 2–3, p. 177–187, 2008.

MARTÍ-RENO, M. A. et al. Comparative protein structure modeling of genes and genomes. v. 29, p. 291–325, 2000.

MAYBRIDGE. **The Maybridge Screening Collection**. Disponível em: <<http://www.maybridge.com>>.

MOREIRA, L. M. et al. Comparative Genomics Analyses of Citrus-Associated Bacteria. **Annual Review of Phytopathology**, v. 42, n. 1, p. 163–184, 2004.

MOREIRA, L. M. et al. Novel insights into the genomic basis of citrus canker based on the genome sequences of two strains of *Xanthomonas fuscans* subsp. *aurantifolii*. **BMC Genomics**, v. 11, n. 1, 2010.

MUEGGE, I. ; ENYEDY, I. J. Virtual screening for kinase targets. **Current Medicinal Chemistry**, v. 11, n. 6, p. 693–707, 2004.

NAYLOR, E. et al. Identification of a chemical probe for NAADP by virtual screening. **Nature Chemical Biology**, v. 5, n. 4, p. 220–226, 2009.

NCBI-National Center for Biotechnology Information. Disponível em: <<https://www.ncbi.nlm.nih.gov/protein/AAM38423.1>>.

NEVES, M. F. ; TROMBIN, V. G. **Anuário da Citricultura 2017 Citrus Br**. São Paulo: [s.n.].

NISSINK, J. W. M. et al. Proteins: Structure, Function and Genetics. **Proteins**, v. 9, p. 457–471, 2002.

NOLTMANN, A. on Phosphomannose. n. 20, 1968.

OPREA, T. I. **Introduction to Chemoinformatics in Drug Discovery**. Weinheim: Wiley-VCH, 2005.

PAPOUTSOPOULOU, S. V; KYRIAKIDIS, D. A. Phosphomannose isomerase of *Xanthomonas campestris*: a zinc activated enzyme. **Molecular and cellular biochemistry**, v. 177, n. 1, p. 183–191, 1997.

PATTERSON, J. H. et al. Mannose metabolism is required for mycobacterial growth. **Biochemical Journal**, v. 372, n. 1, p. 77–86, 2003.

PAYTON, M. A. et al. A novel *Saccharomyces cerevisiae* secretory mutant possesses a thermolabile phosphomannose isomerase. **Journal of Bacteriology**, v. 173, n. 6, p. 2006–2010, 1991.

PEITSCH, M. C. Manuel Peitsch discusses knowledge management and informatics in drug discovery. **Drug Discovery Today: BIOSILICO**, v. 2, n. 3, p. 94–96, 2004.

PICCIRILLO, E.; DO AMARAL, A. T. Virtual screening of bioactive compounds: Concepts and applications. **Química Nova**, v. 41, n. 6, p. 662–677, 2018.

PONTIUS, J.; RICHELLE, J.; WODAK, S. J. Deviations from Standard Atomic Volumes as a Quality Measure for Protein Crystal Structures. **Journal of Molecular Biology**, v. 264, n. 121–136, 1996.

PRENTIS, R.; LIS, Y.; WALKER, S. Pharmaceutical innovation by the seven UK-owned pharmaceutical companies (1964-1985). **British Journal of Clinical Pharmacology**, v. 25, n. 3, p. 387–396, 1988.

PROUDFOOT, A. E. I. et al. Purification cDNA cloning and heterologous expression of human phosphomannose isomerase. **European Journal of Biochemistry**, v. 219, n. 1–2, p. 415–423, 1994.

PROUDFOOT, A. E. I.; PAYTON, M. A.; WELLS, T. N. C. Purification and characterization of fungal and mammalian phosphomannose isomerases. **Journal of Protein Chemistry**, v. 13, n. 7, p. 619–627, 1994.

RICHARD, D. et al. Complete genome sequences of six copperresistant *Xanthomonas citri* pv. *citri* strains causing Asiatic citrus canker, obtained using long-read technology. **Genome Announcements**, 2017.

RODRIGUEZ, R.; VRIEND, G. Professional Gambling. In: **Biomolecular Structure and Dynamics**. Dordrecht: Springer Netherlands, 1997. p. 79–119.

ROSE, IRWIN A.; O'CONNELL, EDWARD L.; SCHRAY, K. J. Mannose 6-Phosphate : Anomeric Isomerase and Its LEpimerization Isomerase \* Form Used by Phosphomannose by w Phosphoglucose. **The Journal of Biological Chemistry**, v. 248, n. 6, p. 2232–2234, 1973.

ROSSETTI, V. . **Manual Ilustrado de Doenças dos Citros**. [s.l: s.n.].

ROUX, C. et al. Inhibition of Type I and Type II Phosphomannose Isomerases by the Reaction Intermediate Analogue 5-Phospho-D-Arabinonohydroxamic Acid Supports a Catalytic Role for the Metal Cofactor. **Biochemistry**, v. 43, n. 10, p. 2926–2934, 2004.

SANCHEZ, R.; SAIL, A. Advances in comparative protein-structure modelling. p. 206–214, [s.d.].

SANCHO, J. **Estructura de Macromoléculas**. Disponível em: <<http://www.bifi.es/~jsancho/estructuramacromoleculas/3peptidicoyES/3peptidicoyES.htm>>. Acesso em: 31 dez. 2019.

SANTOS FILHO, O. A.; BICCA DE ALENCASTRO, R. Modelagem de proteínas por homologia. **Química Nova**, v. 26, n. 2, p. 253–259, 2003.

SCHRODINGER. **Maestro**New York, 2018.

SCHRODINGER, L. **QikProp: Rapid ADME predictions of drug candidates**. Disponível em: <<https://www.schrodinger.com/qikprop>>.

SCHUBERT, T. **No TitleFlorida Department of Agriculture and Consumer Services**, 2011.

SCHWEDE, T. et al. SWISS-MODEL: An automated protein homology-modeling server. **Nucleic Acids Research**, v. 31, n. 13, p. 3381–3385, 2003.

SEILLE, R.; CHRISTIANO, C. Universidade de São Paulo Escola superior de Agricultura “Luiz de Queiroz” Cancro cítrico (. 2006.

SHAHIDULLAH, M; KHORASANI, S. S. M. A. The sensitivity and selectivity of the Seliwanoff test for frutose. **Analytica Chemica Acta**, v. 61, p. 317–319, 1972.

SHERIDAN, R. P. . et al. Protocols for bridging the peptide to nonpeptide gap in topological similarity searches. **Journal of Chemical Information and Computer Sciences**, v. 41, n. 5, p. 1395–1406, 2001.

SIEVERS, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal

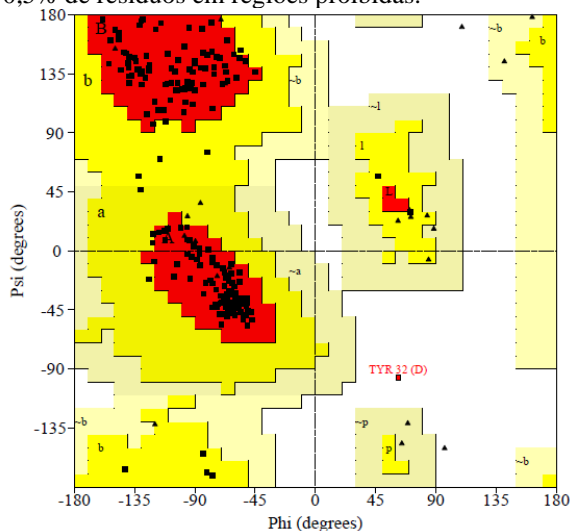
- Omega. **Molecular Systems Biology**, v. 7, n. 539, p. 1–6, 2011.
- SIGDEL, S. et al. Characterization of a Mannose-6-Phosphate Isomerase from *Bacillus amyloliquefaciens* and Its Application in Fructose-6-Phosphate Production. **PLoS ONE**, v. 10, n. 7, p. 1–15, 2015.
- STALL, R. E.; SEYMOUR, C. P. Canker, a Threat to Citrus in the Gulf-Coast States. **American Phytopathological Society**, v. 67, n. 5, p. 581–585, 1983.
- STRYER, L.; BERG, J.M., TYMOCZKO, J. **Biochemistry**. 6<sup>o</sup> ed ed. New York: [s.n.].
- SUITE, G. **GOLD User Guide A Component of the GOLD Suite**, 2015. Disponível em: <<https://www.ccdc.cam.ac.uk>>
- TRUCHON, J. F.; BAYLY, C. I. Evaluating virtual screening methods: Good and bad metrics for the “early recognition” problem. **Journal of Chemical Information and Modeling**, v. 47, n. 2, p. 488–508, 2007.
- VENKATESH, S.; LIPPER, R. A. Role of the development scientist in compound lead selection and optimization. **Journal of Pharmaceutical Sciences**, v. 89, n. 2, p. 145–154, 2000.
- VIETH, M. et al. Assessing energy functions for flexible *docking*. **Journal of Computational Chemistry**, v. 19, n. 14, p. 1612–1622, 1998.
- VRIEND, G. WHAT IF: a molecular modeling and drug design program. **Journal of Molecular Graphics**, v. 8, n. 1, p. 52–6, 1990.
- WELLS, T. N. C. et al. Mechanism of Irreversible Inactivation of Phosphomannose Isomerases by Silver Ions and Flamazine. **Biochemistry**, v. 34, n. 24, p. 7896–7903, 1995.
- WELLS, T. N. C.; PAYTON, M. A.; PROUDFOOT, A. E. I. Inhibition of Phosphomannose Isomerase by Mercury Ions. **Biochemistry**, v. 33, n. 24, p. 7641–7646, 1994.
- WELLS, T. N. C.; SCULLY, P.; MAGNENAT, E. Arginine 304 Is an Active Site Residue in Phosphomannose Isomerase from *Candida albicans*. **Biochemistry**, v. 33, n. 19, p. 5777–5782, 1994.
- WERMUTH, C. G. **The Practice of Medicinal Chemistry**. [s.l.] Elsevier, 2008.
- YACOUBI, B. EL. **BACTERIAL CITRUS CANKER: MOLECULAR ASPECTS OF A COMPATIBLE PLANT-MICROBE INTERACTION**. [s.l.: s.n.].
- YURIEV, E.; AGOSTINO, M.; RAMSLAND, P. A. Challenges and advances in computational *docking*: 2009 in review. **Journal of Molecular Recognition**, v. 24, n. 2, p. 149–164, 2011.
- YURIEV, E.; RAMSLAND, P. A. Latest developments in molecular *docking*: 2010-2011 in review. **Journal of Molecular Recognition**, v. 26, n. 5, p. 215–239, 2013.
- ZHANG, Y. et al. Positive selection is the main driving force for evolution of citrus canker-causing *Xanthomonas*. **ISME Journal**, 2015.
- ZHAO, W. et al. A statistical framework to evaluate virtual screening. **BMC Bioinformatics**, v. 10, p. 1–13, 2009.
- ZUCCOTTI, S. et al. Kinetic and Crystallographic Analyses Support a Sequential-Ordered Bi Bi Catalytic Mechanism for *Escherichia Coli* Glucose-1-Phosphate Thymidyltransferase. **J.Mol.Biol.**, v. 313, p. 831, 2001.

## APÊNDICES

### APÊNDICE A - Dados relacionados a construção do modelo de PMI *Xanthomonas citri* subsp. *citri*.

Nesse apêndice são apresentados os dados obtidos em relação as estruturas cristalográficas 1H5R, 2CU2, 2X5S e 2QH5; além dos outros 4 modelos preditos por modelagem por homologia estrutural. Os dados obtidos são os gráficos de Ramachandran, **PROCHECK**, **VERIFY3D**, **ERRAT** e **PROVE** com os respectivos valores encontrados para cada estrutura proteica.

**Figura A1** - Diagrama de Ramachandram da estrutura cristalográfica 1H5R obtido pelo **PROCHECK**: 93% de resíduos em regiões favoráveis, 6,5% de resíduos em regiões permitidas, 0% de resíduos em regiões generosamente permitidas e 0,5% de resíduos em regiões proibidas.



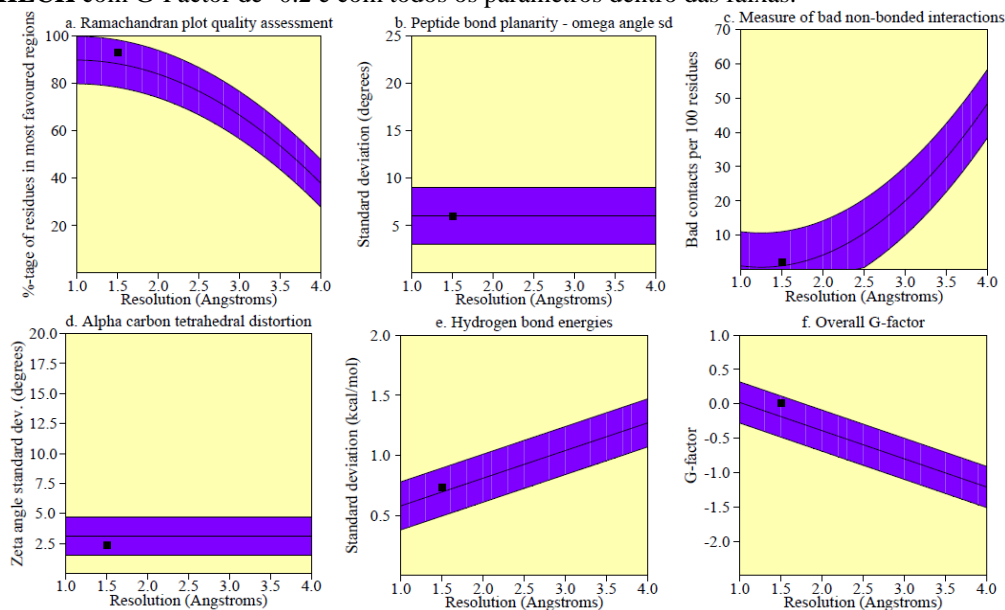
Estatísticas de plotagem		
Resíduos em regiões mais favoráveis [A,B,L]	185	93.0%
Resíduos em regiões adicionais permitidas [a,b,l,p]	13	6.5%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	0	0.0%
Resíduos em regiões proibidas	1	0.5%
	----	----
Número de resíduos não glicinas e não prolinas	199	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	23	
Número de resíduos de prolina	13	
	----	
Número total de resíduos	237	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

Fonte: O autor

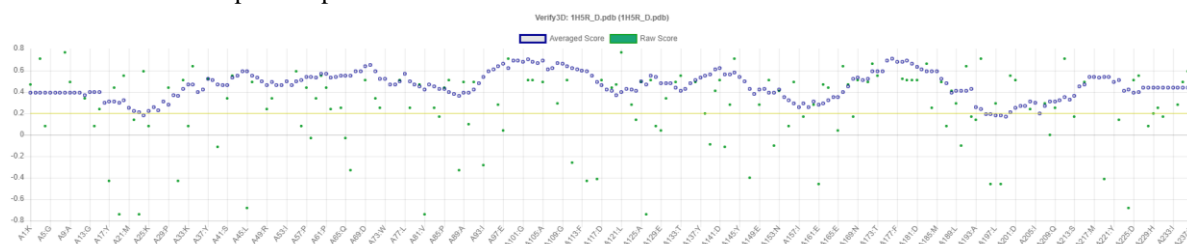


**Figura A2** - Resumo da avaliação estereoquímica da estrutura cristalográfica 1H5R realizado pelo **PROCHECK** com G-Factor de -0.2 e com todos os parâmetros dentro das faixas.



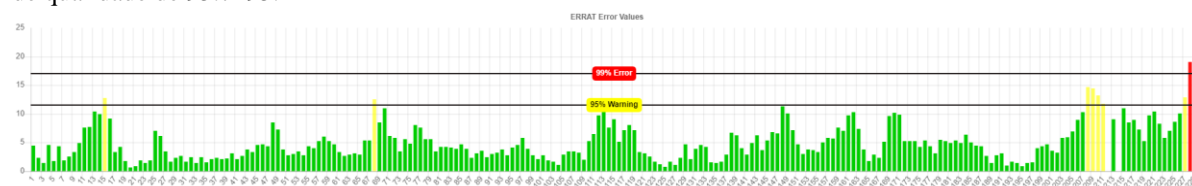
Fonte: O autor

**Figura A3** - Gráfico para a estrutura cristalográfica 1H5R obtido pelo **VERIFY3D**: 97,47% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

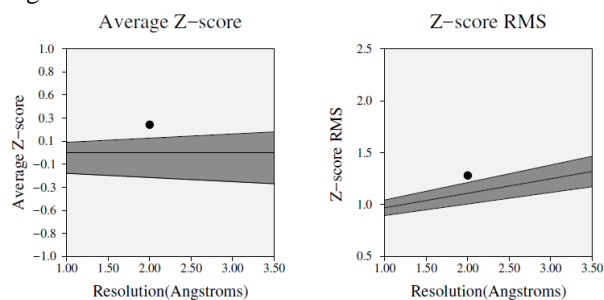
**Figura A4** - Gráfico para a estrutura cristalográfica 1H5R obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 95.7295.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

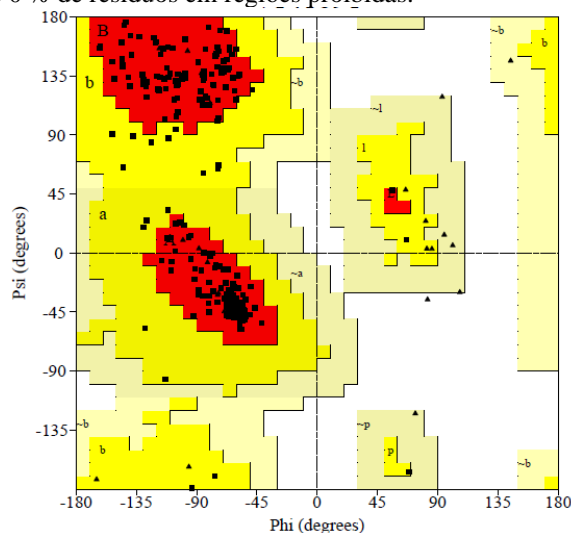
**Figura A5** - Gráfico para a estrutura cristalográfica 1H5R obtido pelo **PROVE**. Essa estrutura apresenta 2,6% dos aminoácidos foram da região de interesse.



*O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.*

Fonte: O autor

**Figura A6** - Diagrama de Ramachandram da estrutura cristalográfica 2CU2 obtido pelo **PROCHECK**: 91,9% de resíduos em regiões favoráveis, 8,1% de resíduos em regiões permitidas, 0% de resíduos em regiões generosamente permitidas e 0% de resíduos em regiões proibidas.



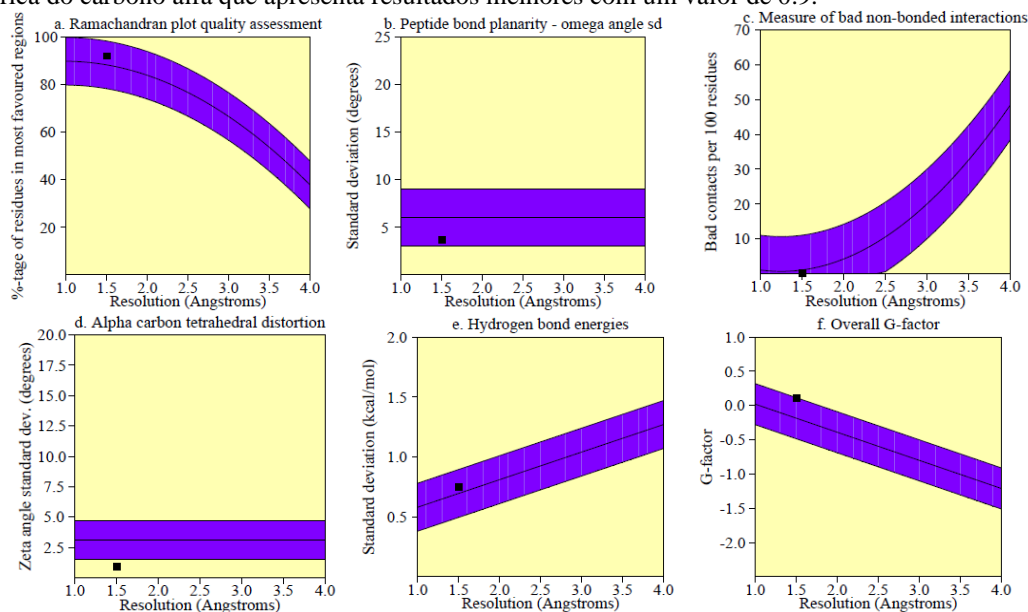
#### Estadísticas de plotagem

Resíduos em regiões mais favoráveis [A,B,L]	204	91.9%
Resíduos em regiões adicionais permitidas [a,b,l,p]	18	8.1%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	0	0.0%
Resíduos em regiões proibidas	0	0.0%
Número de resíduos não glicinas e não prolina	222	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	25	
Número de resíduos de prolina	18	
Número total de resíduos	267	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

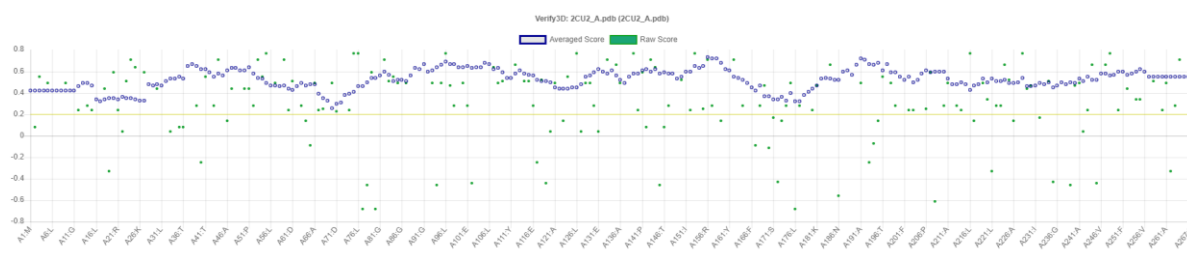
Fonte: O autor

**Figura A7** - Resumo da avaliação estereoquímica da estrutura cristalográfica 2CU2 realizado pelo **PROCHECK** com G-Factor de 0.1 e com todos os parâmetros dentro das faixas, com exceção da distorção tetraédrica do carbono alfa que apresenta resultados melhores com um valor de 0.9.



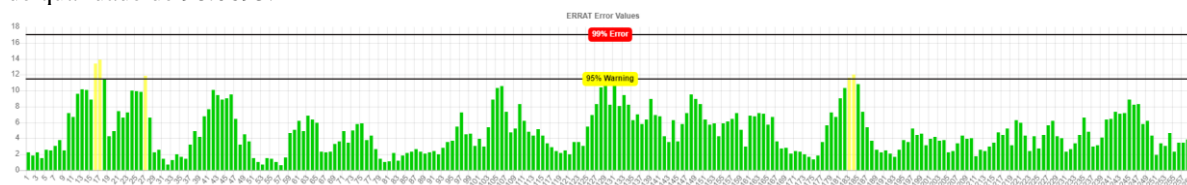
Fonte: O autor

**Figura A8** - Gráfico para a estrutura cristalográfica 2CU2 obtido pelo **VERIFY3D**: 100% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

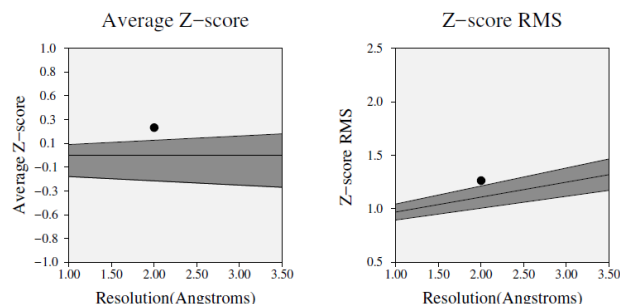
**Figura A9** - Gráfico para a estrutura cristalográfica 2CU2 obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 98.0695.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

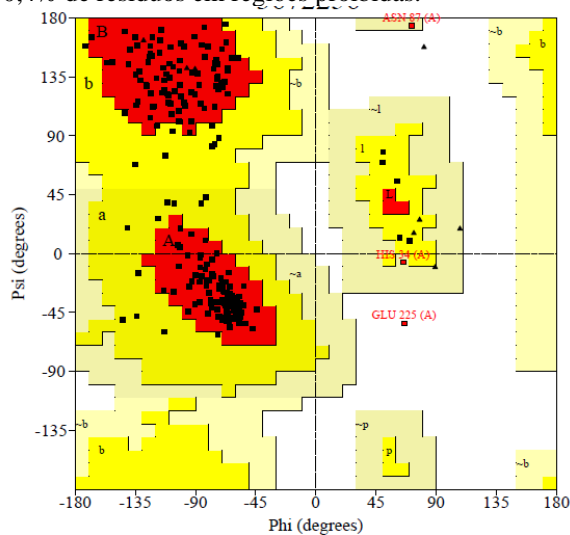
**Figura A10** - Gráfico para a estrutura cristalográfica 2CU2 obtido pelo **PROVE**. Essa estrutura apresenta 1,9% dos aminoácidos foram da região de interesse.



O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: O autor

**Figura A11** - Diagrama de Ramachandram da estrutura cristalográfica 2QH5 obtido pelo **PROCHECK**: 87,6% de resíduos em regiões favoráveis, 11,2% de resíduos em regiões permitidas, 0,9% de resíduos em regiões generosamente permitidas e 0,4% de resíduos em regiões proibidas.



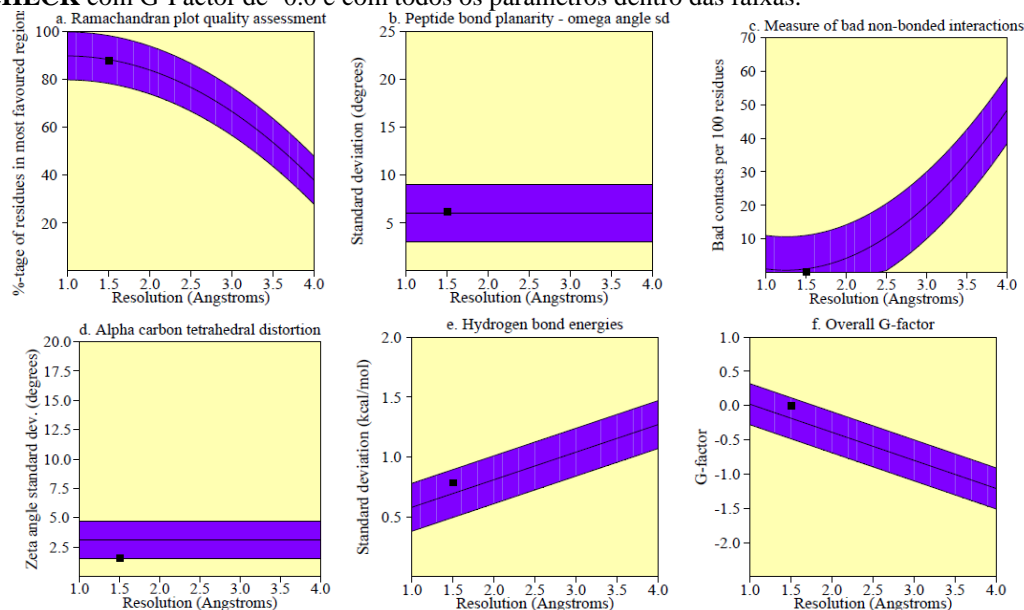
#### Estadísticas de plotagem

Resíduos em regiões mais favoráveis [A,B,L]	204	87.6%
Resíduos em regiões adicionais permitidas [a,b,l,p]	26	11.2%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	2	0.9%
Resíduos em regiões proibidas	1	0.4%
Número de resíduos não glicinas e não prolina	233	100.0%
Número de resíduos finais (exc. Glicina e prolina)	3	
Número de resíduos de glicina (mostrados como triângulos)	11	
Número de resíduos de prolina	6	
Número total de resíduos	253	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

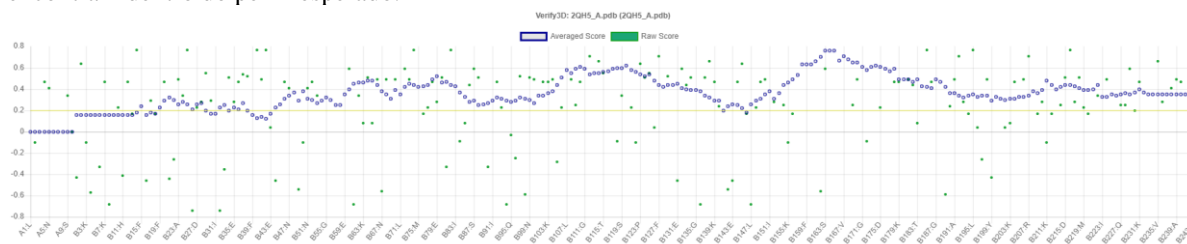
Fonte: O autor

**Figura A12** - Resumo da avaliação estereoquímica da estrutura cristalográfica 2QH5 realizado pelo PROCHECK com G-Factor de -0.0 e com todos os parâmetros dentro das faixas.



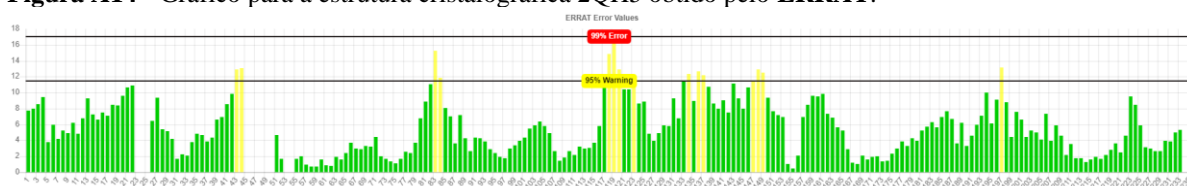
Fonte: O autor

**Figura A13** - Gráfico para a estrutura cristalográfica 2QH5 obtido pelo VERIFY3D: 86,17% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

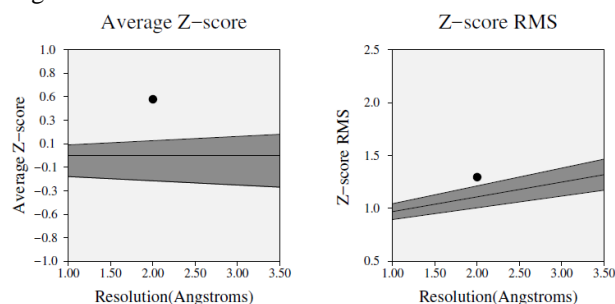
**Figura A14** - Gráfico para a estrutura cristalográfica 2QH5 obtido pelo ERRAT.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

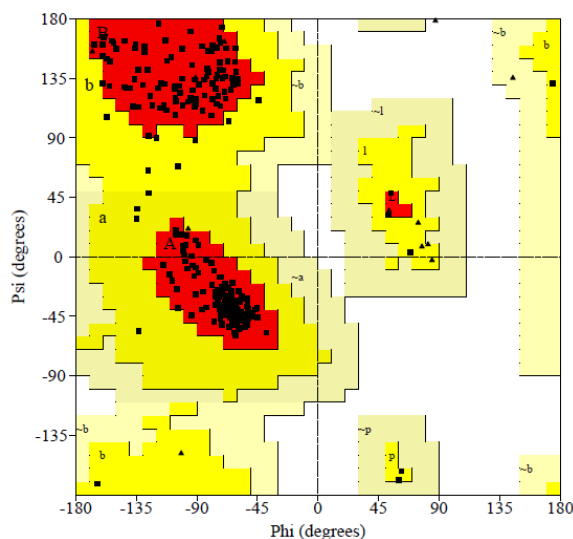
**Figura A15** - Gráfico para a estrutura cristalográfica 2QH5 obtido pelo **PROVE**. Essa estrutura apresenta 3,2% dos aminoácidos foram da região de interesse.



*O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.*

Fonte: O autor

**Figura A16** - Diagrama de Ramachandram da estrutura cristalográfica 2X5S obtido pelo **PROCHECK**: 93,7% de resíduos em regiões favoráveis, 6,3% de resíduos em regiões permitidas, 0,0% de resíduos em regiões generosamente permitidas e 0,0% de resíduos em regiões proibidas.



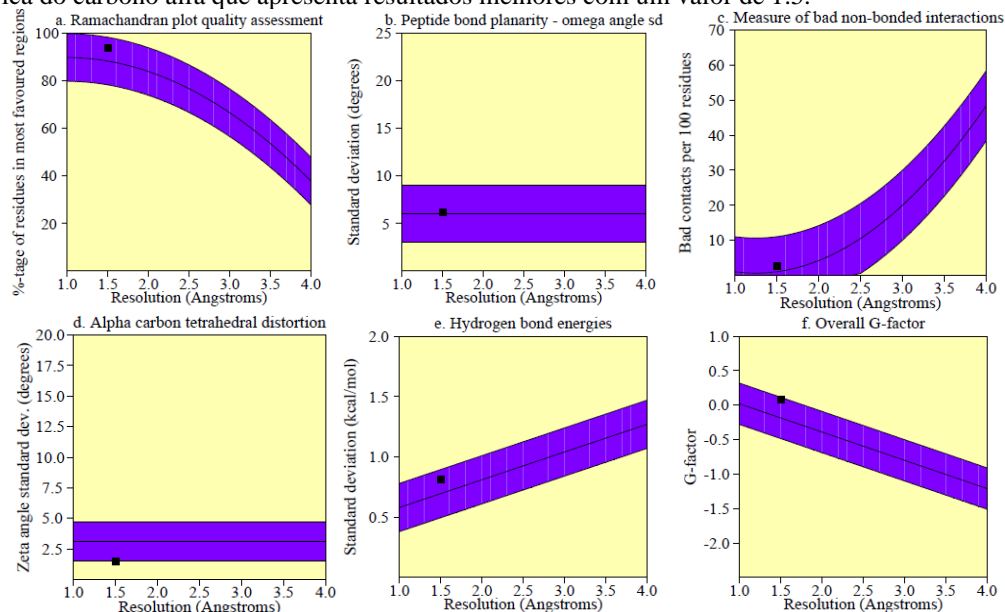
#### Estadísticas do gráfico

Resíduos em regiões mais favoráveis [A,B,L]	224	93.7%
Resíduos em regiões adicionais permitidas [a,b,l,p]	15	6.3%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	0	0.0%
Resíduos em regiões proibidas	0	0.0%
Número de resíduos não glicinas e não prolina	239	100.0%
Número de resíduos finais (exc. Glicina e prolina)	1	
Número de resíduos de glicina (mostrados como triângulos)	15	
Número de resíduos de prolina	17	
Número total de resíduos	272	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

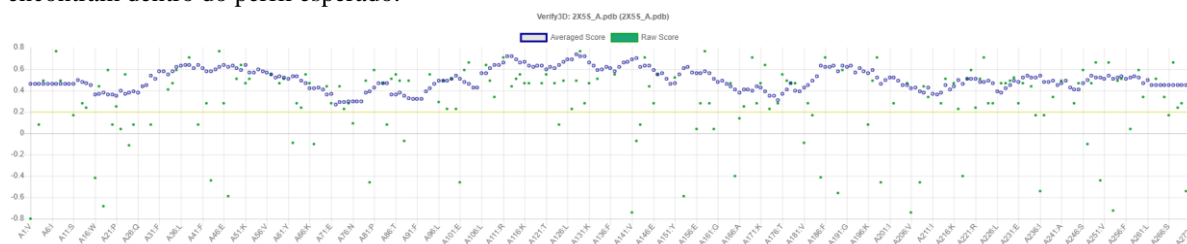
Fonte: O autor

**Figura A17** - Resumo da avaliação estereoquímica da estrutura cristalográfica 2X5S realizado pelo **PROCHECK** com G-Factor de 0.1 e com todos os parâmetros dentro das faixas, com exceção da distorção tetraédrica do carbono alfa que apresenta resultados melhores com um valor de 1.5.



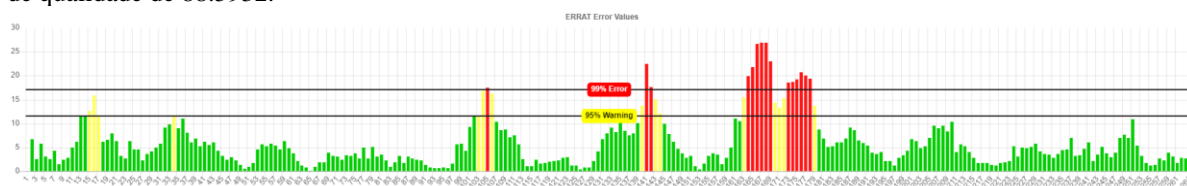
Fonte: O autor

**Figura A18** - Gráfico para a estrutura cristalográfica 2X5S obtido pelo **VERIFY3D**: 100% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

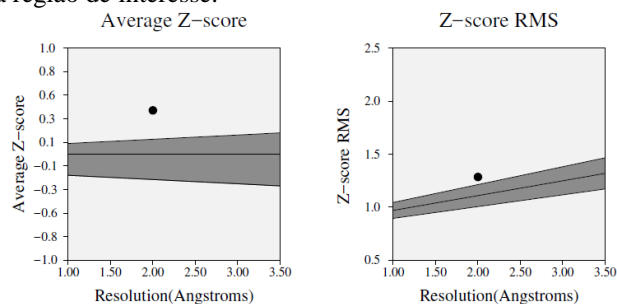
**Figura A19** - Gráfico para a estrutura cristalográfica 2X5S obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 88.5932.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

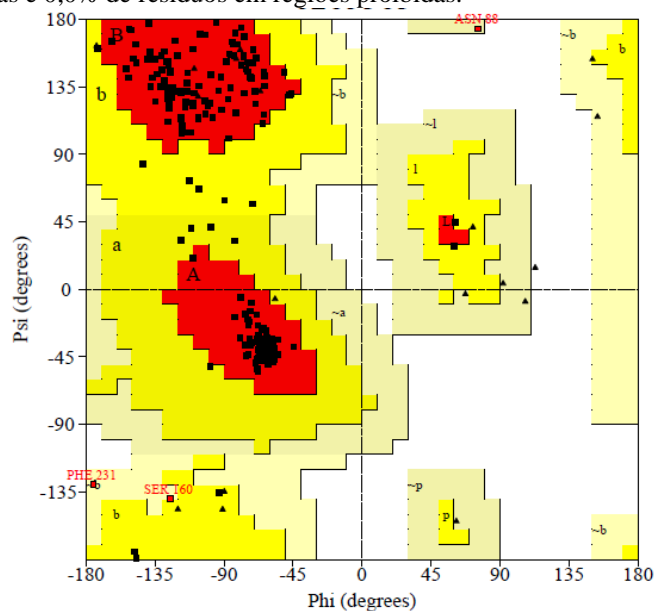
**Figura A20** - Gráfico para a estrutura cristalográfica 2X5S obtido pelo **PROVE**. Essa estrutura apresenta 2,8% dos aminoácidos foram da região de interesse.



O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: O autor

**Figura A21** - Diagrama de Ramachandram da hipótese 1 gerada para PMI XAC obtido pelo **PROCHECK**: 91,6% de resíduos em regiões favoráveis, 7,1% de resíduos em regiões permitidas, 1,3% de resíduos em regiões generosamente permitidas e 0,0% de resíduos em regiões proibidas.



Estatísticas do gráfico

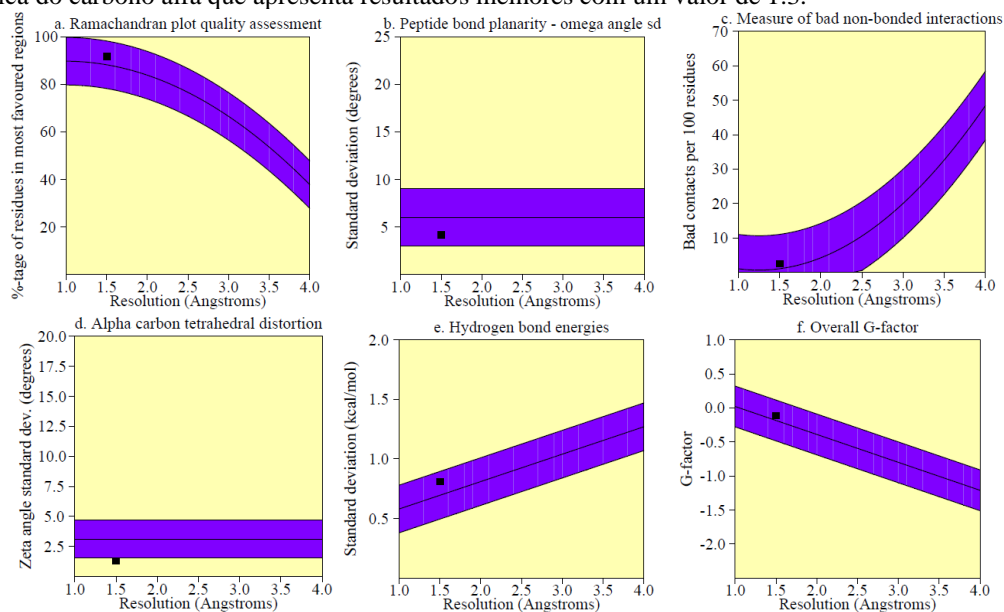
Resíduos em regiões mais favoráveis [A,B,L]	219	91.6%
Resíduos em regiões adicionais permitidas [a,b,l,p]	17	7.1%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	3	1.3%
Resíduos em regiões proibidas	0	0.0%
	----	----
Número de resíduos não glicinas e não prolina	239	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	19	
Número de resíduos de prolina	17	
	----	
Número total de resíduos	267	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

Fonte: O autor

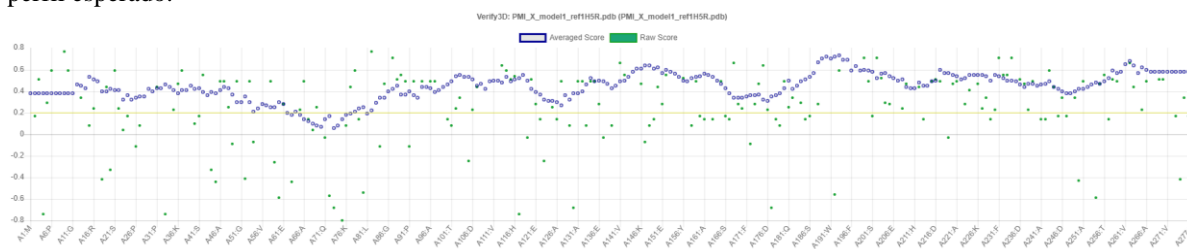


**Figura A22** - Resumo da avaliação estereoquímica da hipótese 1 gerada para PMI XAC realizado pelo **PROCHECK** com G-Factor de -0.1 e com todos os parâmetros dentro das faixas, com exceção da distorção tetraédrica do carbono alfa que apresenta resultados melhores com um valor de 1.3.



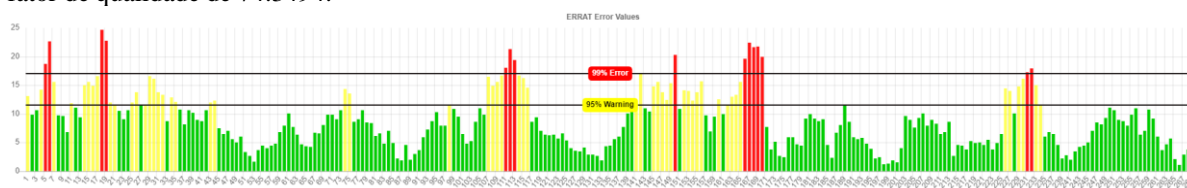
Fonte: O autor

**Figura A23** - Gráfico para a hipótese 1 obtido pelo **VERIFY3D**: 94,58% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

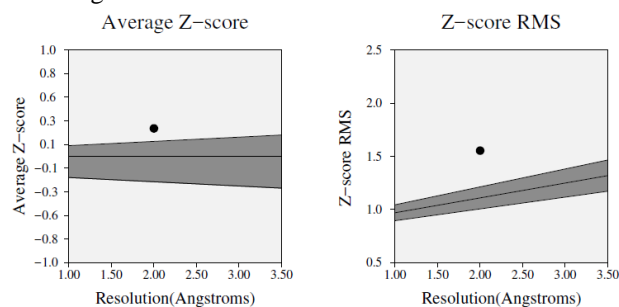
**Figura A24** - Gráfico para a hipótese 1 gerada para PMI XAC obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 74.3494.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

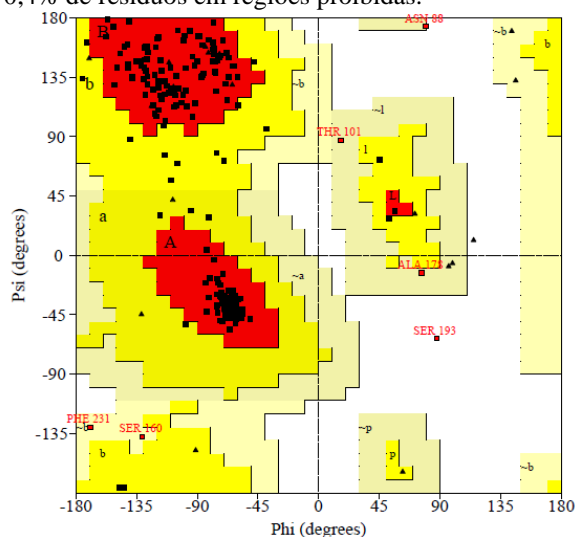
**Figura A25** - Gráfico para hipótese 1 gerada para PMI XAC obtido pelo **PROVE**. Essa estrutura apresenta 4,9% dos aminoácidos foram da região de interesse.



*O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.*

Fonte: O autor

**Figura A26** - Diagrama de Ramachandram da hipótese 2 gerada para PMI XAC obtido pelo **PROCHECK**: 89,5% de resíduos em regiões favoráveis, 7,9% de resíduos em regiões permitidas, 2,1% de resíduos em regiões generosamente permitidas e 0,4% de resíduos em regiões proibidas.



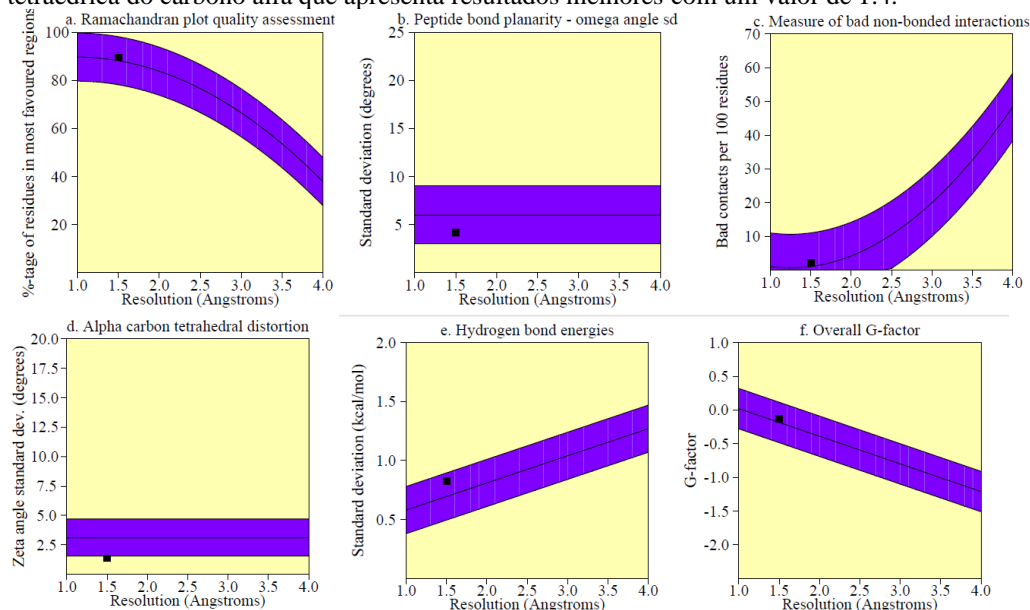
Estadísticas de gráfico

Resíduos em regiões mais favoráveis [A,B,L]	214	89.5%
Resíduos em regiões adicionais permitidas [a,b,l,p]	19	7.9%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	5	2.1%
Resíduos em regiões proibidas	1	0.4%
Número de resíduos não glicinas e não prolina	239	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	19	
Número de resíduos de prolina	17	
Número total de resíduos	277	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

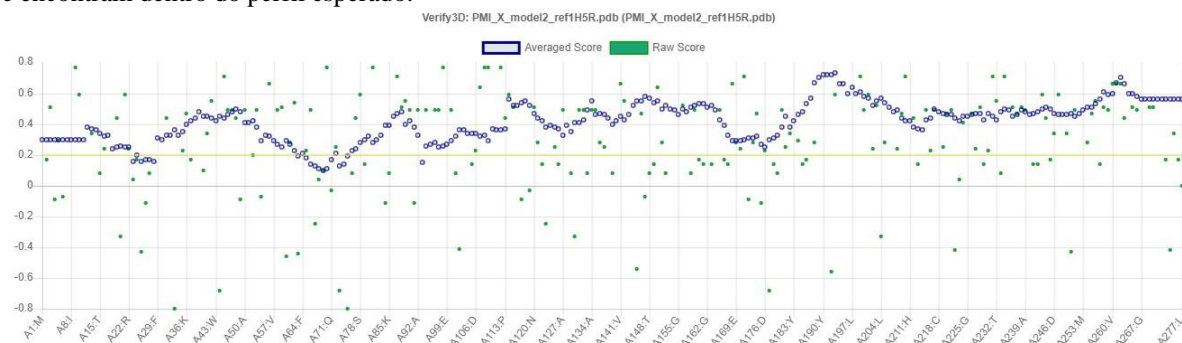
Fonte: O autor

**Figura A27** - Resumo da avaliação estereoquímica da hipótese 2 gerada para PMI XAC realizado pelo **PROCHECK** com G-Factor de -0.1 e com todos os parâmetros dentro das faixas, com exceção da distorção tetraédrica do carbono alfa que apresenta resultados melhores com um valor de 1.4.



Fonte: O autor

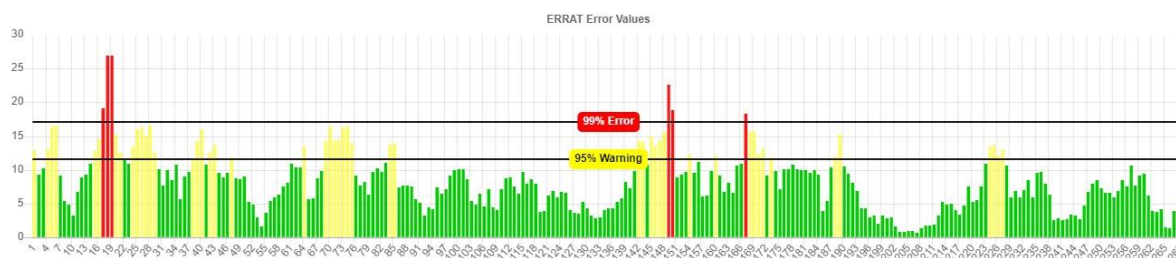
**Figura A28** - Gráfico para a hipótese 2 gerada para PMI XAC obtido pelo **VERIFY3D**: 93,88% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

**Figura A29** - Gráfico para a hipótese 2 gerada para PMI XAC obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 79.5539.

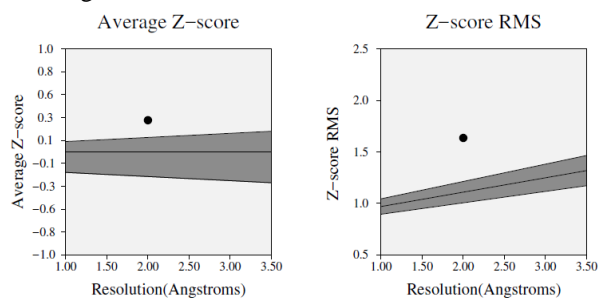
Quality Factor: A: 79.5539 | PDF | PostScript | Log



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

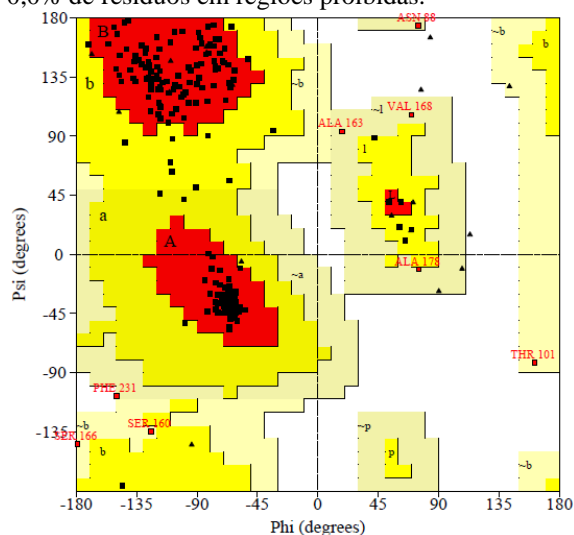
**Figura A30** - Gráfico para a hipótese 2 gerada para PMI XAC obtido pelo **PROVE**. Essa estrutura apresenta 5,6% dos aminoácidos foram da região de interesse.



O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: O autor

**Figura A31** - Diagrama de Ramachandram da hipótese 3 gerada para PMI XAC obtido pelo **PROCHECK**: 89,5% de resíduos em regiões favoráveis, 7,1% de resíduos em regiões permitidas, 3,3% de resíduos em regiões generosamente permitidas e 0,0% de resíduos em regiões proibidas.



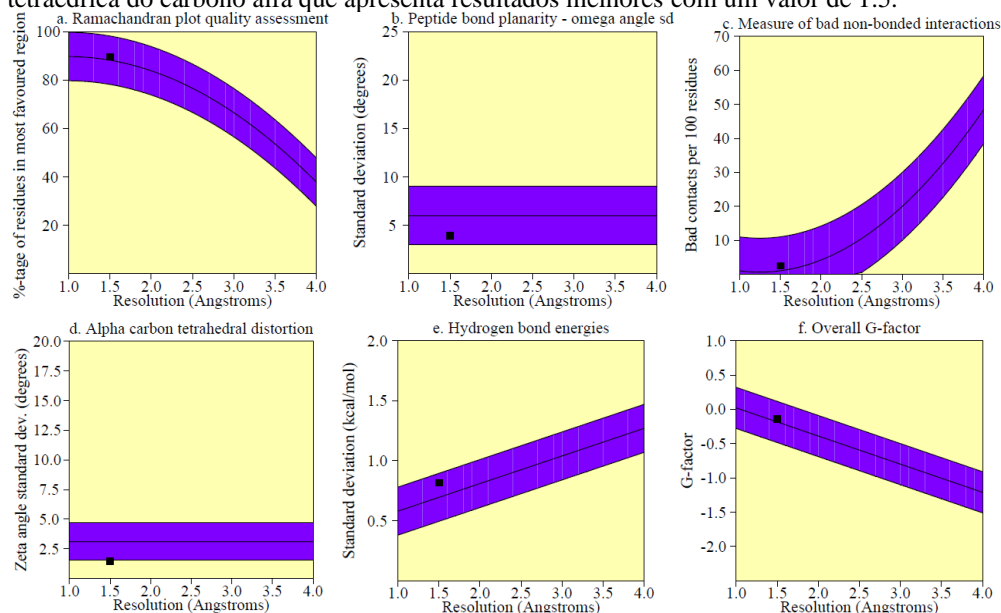
#### Estadísticas de plotagem

Resíduos em regiões mais favoráveis [A,B,L]	214	89.5%
Resíduos em regiões adicionais permitidas [a,b,l,p]	17	7.1%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	8	0.0%
Resíduos em regiões proibidas	0	0.0%
Número de resíduos não glicinas e não prolina	222	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	19	
Número de resíduos de prolina	17	
Número total de resíduos	277	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

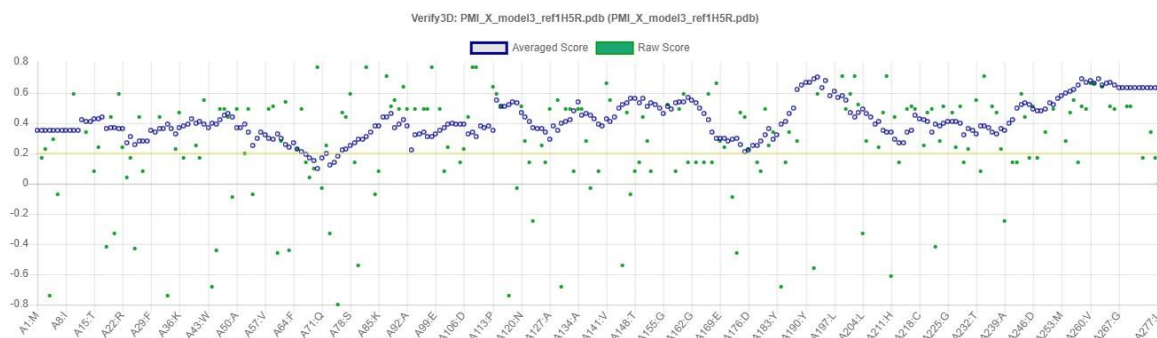
Fonte: O autor

**Figura A32** - Resumo da avaliação estereoquímica da hipótese 3 gerada para PMI XAC realizado pelo **PROCHECK** com G-Factor de -0.1 e com todos os parâmetros dentro das faixas, com exceção da distorção tetraédrica do carbono alfa que apresenta resultados melhores com um valor de 1.5.



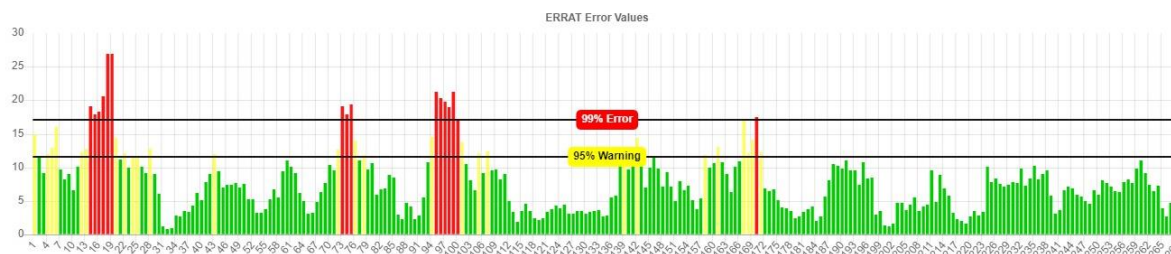
Fonte: O autor

**Figura A33** - Gráfico para a hipótese 3 obtido pelo **VERIFY3D**: 97,11% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

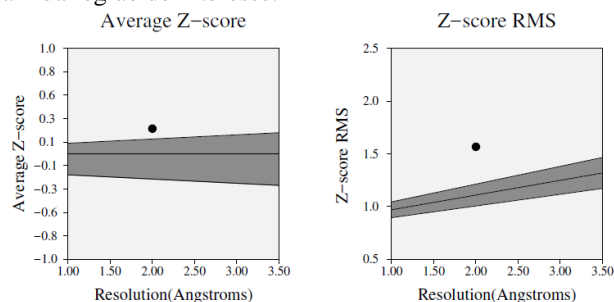
**Figura A34** - Gráfico para a hipótese 3 gerada para PMI XAC obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 84.0149.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

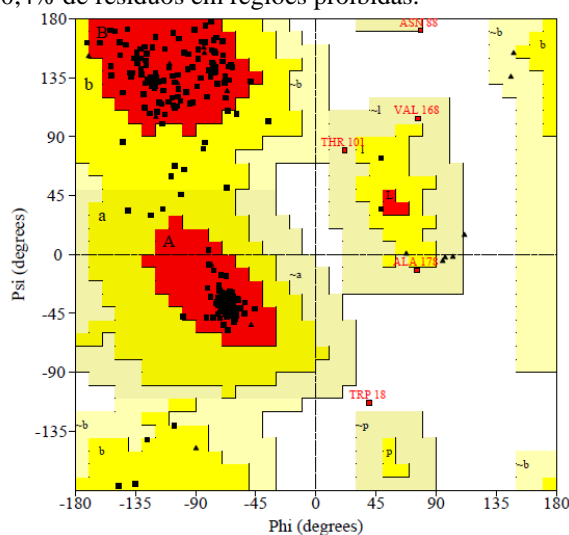
**Figura A35** - Gráfico para a hipótese 3 gerada para PMI XAC obtido pelo **PROVE**. Essa estrutura apresenta 5,7% dos aminoácidos foram da região de interesse.



O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: O autor

**Figura A36** - Diagrama de Ramachandram da hipótese 5 gerada para PMI XAC obtido pelo **PROCHECK**: 89,1% de resíduos em regiões favoráveis, 8,8% de resíduos em regiões permitidas, 1,7% de resíduos em regiões generosamente permitidas e 0,4% de resíduos em regiões proibidas.



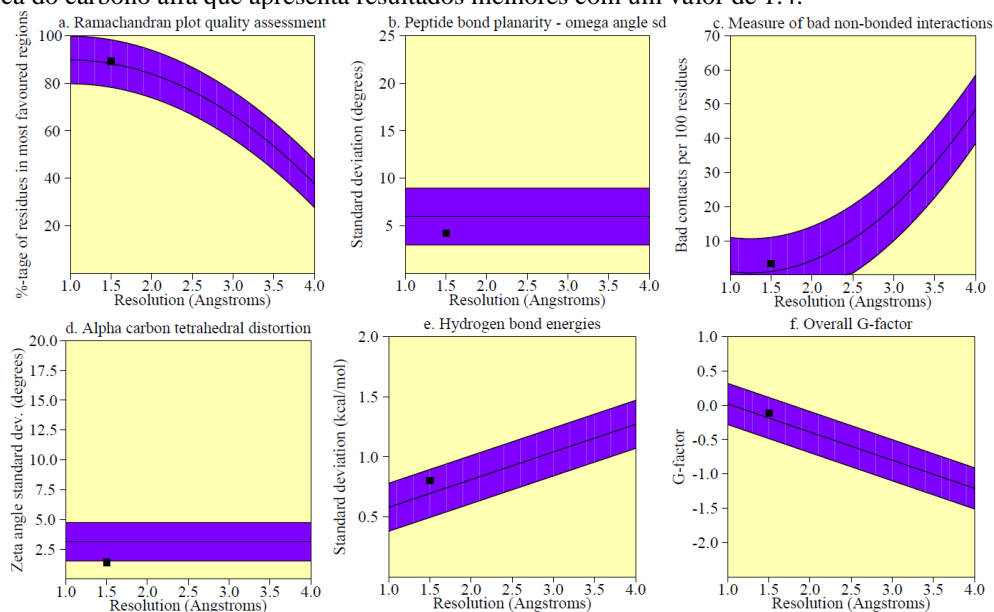
#### Estadísticas de plotagem

Resíduos em regiões mais favoráveis [A,B,L]	213	89.1%
Resíduos em regiões adicionais permitidas [a,b,l,p]	21	8.8%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	4	1.7%
Resíduos em regiões proibidas	1	0.4%
Número de resíduos não glicinas e não prolina	239	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	19	
Número de resíduos de prolina	17	
Número total de resíduos	277	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

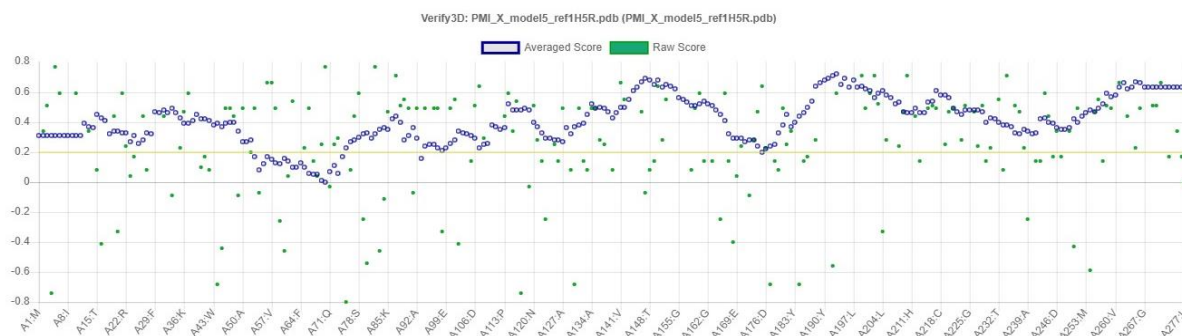
Fonte: O autor

**Figura A37** - Resumo da avaliação estereoquímica da hipótese 5 gerada para PMI XAC realizado pelo **PROCHECK** com G-Factor de -0.1 e com todos os parâmetros dentro das faixas, com exceção da distorção tetraédrica do carbono alfa que apresenta resultados melhores com um valor de 1.4.



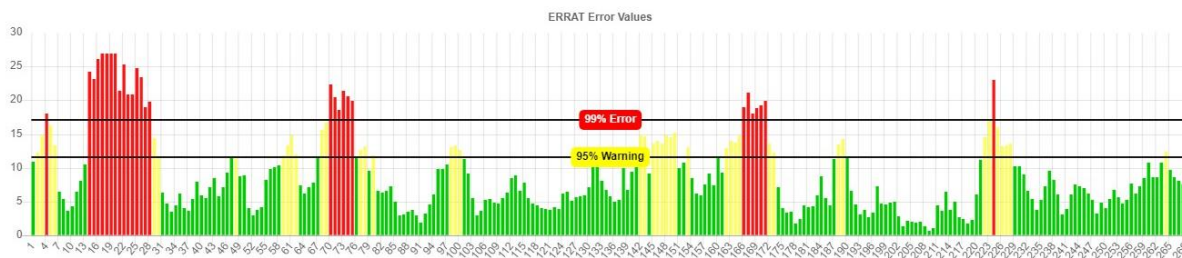
Fonte: O autor

**Figura A38** - Gráfico para a hipótese 5 gerada para PMI XAC obtido pelo **VERIFY3D**: 91,70% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

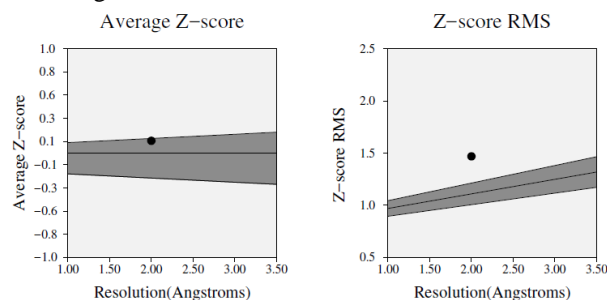
**Figura A39** - Gráfico para a hipótese 5 gerada para PMI XAC obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 73.2342.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

**Figura A40** - Gráfico para a hipótese 5 gerada para PMI XAC obtido pelo **PROVE**. Essa estrutura apresenta 4,1% dos aminoácidos foram da região de interesse.



*O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.*

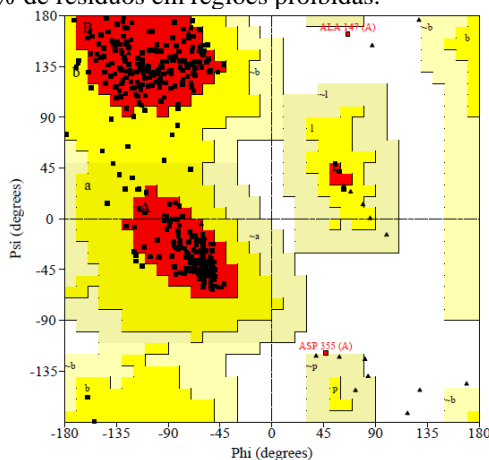
Fonte: O autor

#### *APÊNDICE B - Dados relacionados a construção do modelo de PMI Homo sapiens*

Nesse apêndice são apresentados os dados obtidos em relação as estruturas cristalográficas 3H1M e 5NW7; além dos outros 4 modelos preditos por modelagem por homologia estrutural. Os dados obtidos são os gráficos de Ramachandran, **PROCHECK**, **VERIFY3D**, **ERRAT** e **PROVE** com os respectivos valores encontrados para cada estrutura proteica.



**Figura B1** - Diagrama de Ramachandram da estrutura cristalográfica 3H1M obtido pelo **PROCHECK**: 90,6% de resíduos em regiões favoráveis, 8,8% de resíduos em regiões permitidas, 0,0% de resíduos em regiões generosamente permitidas e 0,6% de resíduos em regiões proibidas.



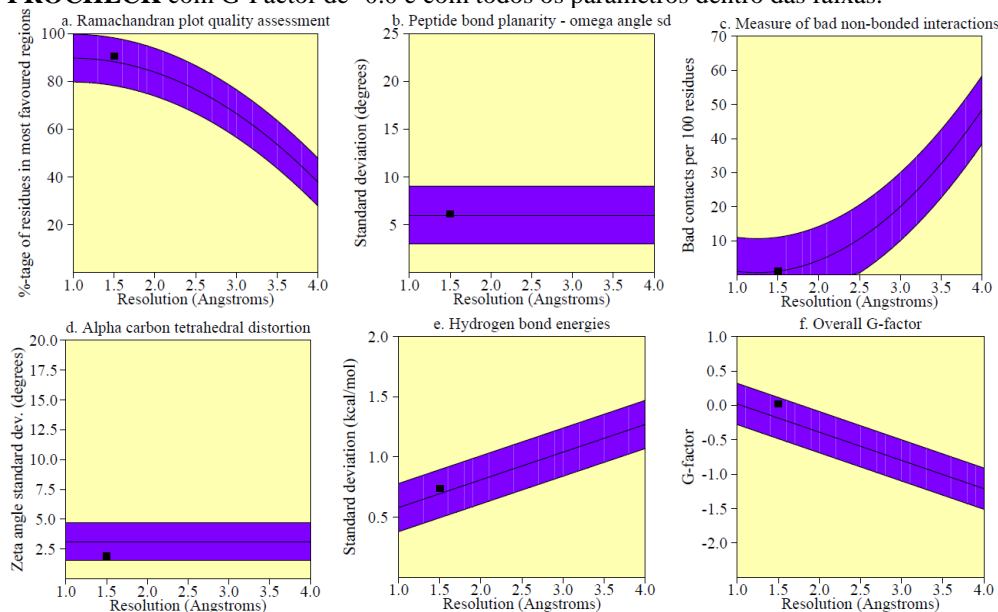
#### Estatísticas de plotagem

Resíduos em regiões mais favoráveis [A,B,L]	308	90.6%
Resíduos em regiões adicionais permitidas [a,b,l,p]	30	8.8%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	0	0.0%
Resíduos em regiões proibidas	2	0.6%
	----	----
Número de resíduos não glicinas e não prolinas	340	100.0%
Número de resíduos finais (exc. Glicina e prolina)	3	
Número de resíduos de glicina (mostrados como triângulos)	23	
Número de resíduos de prolina	26	
	----	
Número total de resíduos	392	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

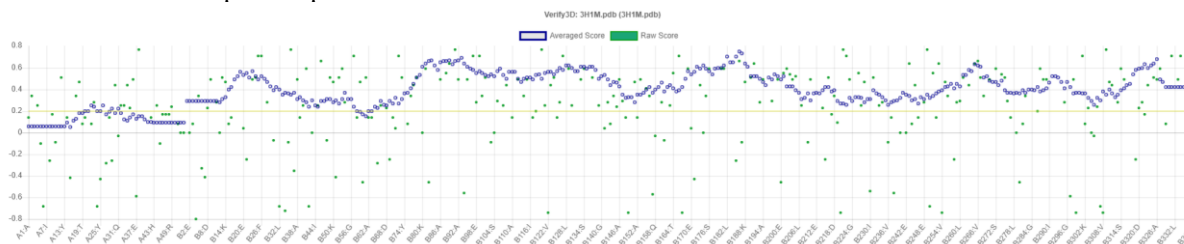
Fonte: O autor

**Figura B2** - Resumo da avaliação estereoquímica da estrutura cristalográfica 3H1M realizado pelo **PROCHECK** com G-Factor de -0.0 e com todos os parâmetros dentro das faixas.



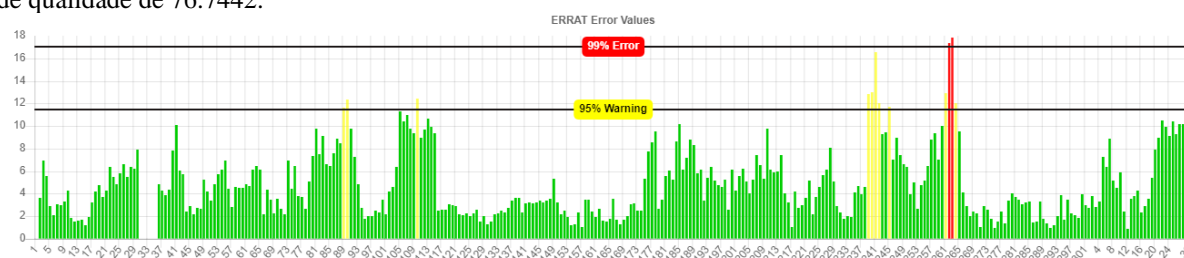
Fonte: O autor

**Figura B3** - Gráfico para a estrutura cristalográfica 3H1M obtido pelo **VERIFY3D**: 88,01% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

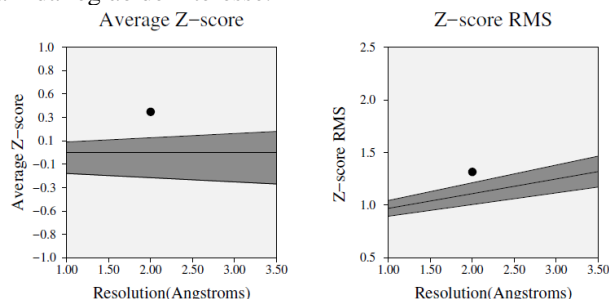
**Figura B4** - Gráfico para a estrutura cristalográfica 3H1M obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 76.7442.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

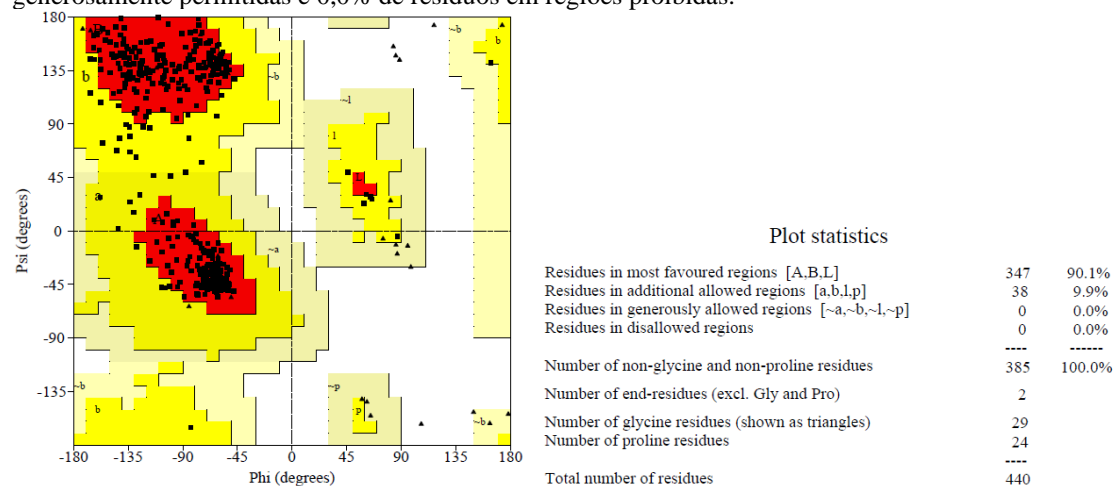
**Figura B5** - Gráfico para a estrutura cristalográfica 3H1M obtido pelo **PROVE**. Essa estrutura apresenta 3,6% dos aminoácidos foram da região de interesse.



O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: O autor

**Figura B6** - Diagrama de Ramachandram da estrutura cristalográfica 5NW7 obtido pelo **PROCHECK**: 90,1% de resíduos em regiões favoráveis, 9,9% de resíduos em regiões permitidas, 0,0% de resíduos em regiões generosamente permitidas e 0,0% de resíduos em regiões proibidas.



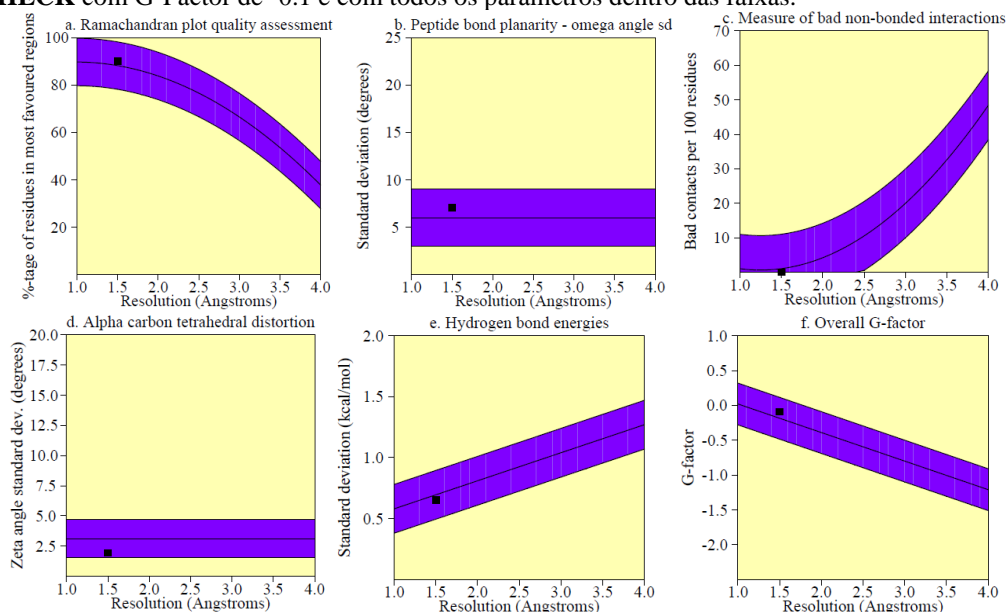
#### Estadísticas de plotagem

Resíduos em regiões mais favoráveis [A,B,L]	347	90.1%
Resíduos em regiões adicionais permitidas [a,b,l,p]	38	9.9%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	0	0.0%
Resíduos em regiões proibidas	0	0.0%
-----		
Número de resíduos não glicinas e não prolina	385	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	29	
Número de resíduos de prolina	24	
-----		
Número total de resíduos	440	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

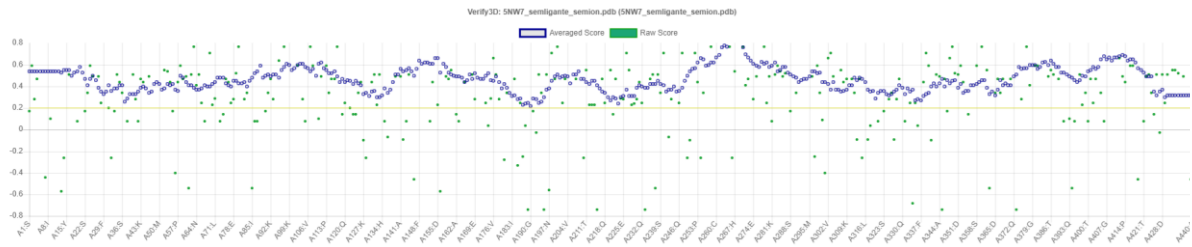
Fonte: O autor

**Figura B7** - Resumo da avaliação estereoquímica da estrutura cristalográfica 3H1M realizado pelo **PROCHECK** com G-Factor de -0.1 e com todos os parâmetros dentro das faixas.



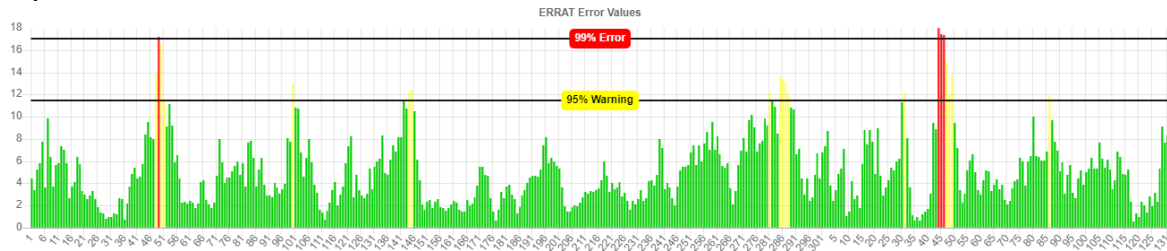
Fonte: O autor

**Figura B8** - Gráfico para a estrutura cristalográfica 3H1M obtido pelo **VERIFY3D**: 100% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

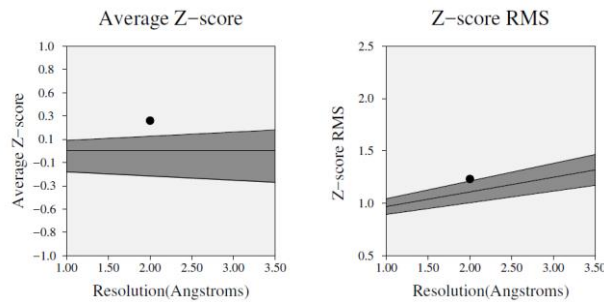
**Figura B9** - Gráfico para a estrutura cristalográfica 5NW7 obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 95.3704.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

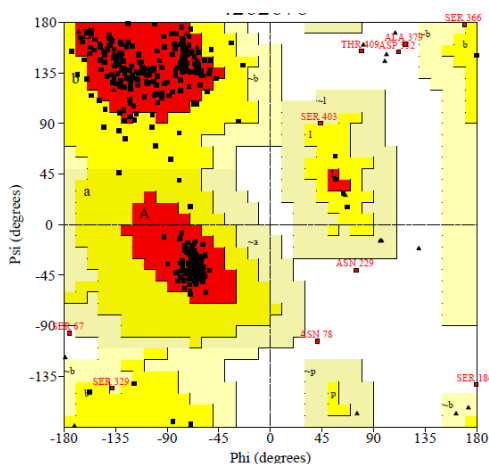
**Figura B10** - Gráfico para a estrutura cristalográfica 3H1M obtido pelo **PROVE**. Essa estrutura apresenta 2,9% dos aminoácidos foram da região de interesse.



O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: O autor

**Figura B11** - Diagrama de Ramachandram da hipótese 1 gerada para PMI *Homo sapiens* obtido pelo PROCHECK: 88,4% de resíduos em regiões favoráveis, 8,9% de resíduos em regiões permitidas, 0,8% de resíduos em regiões generosamente permitidas e 1,9% de resíduos em regiões proibidas.



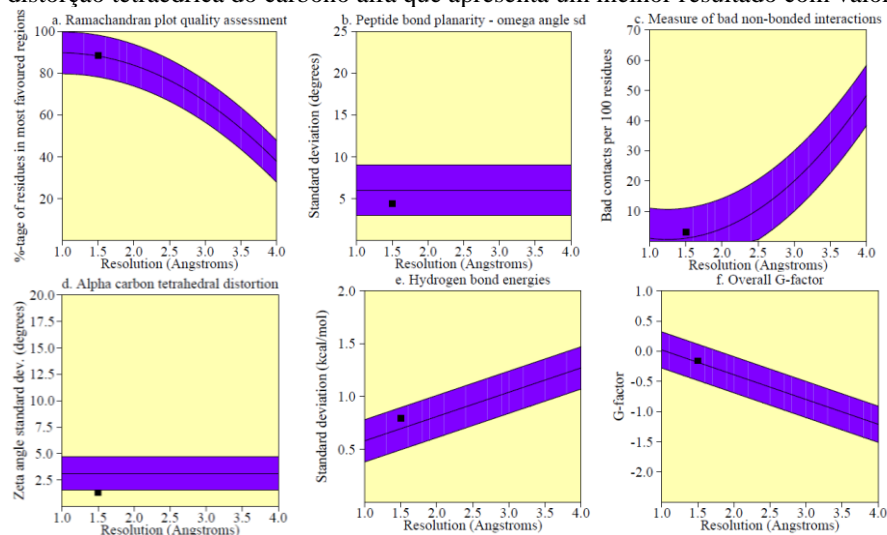
Estadísticas de plotagem

Resíduos em regiões mais favoráveis [A,B,L]	327	88.4%
Resíduos em regiões adicionais permitidas [a,b,l,p]	33	8.9%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	3	0.8%
Resíduos em regiões proibidas	7	1.9%
	----	----
Número de resíduos não glicinas e não prolina	370	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	21	
Número de resíduos de prolina	30	
	----	
Número total de resíduos	425	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

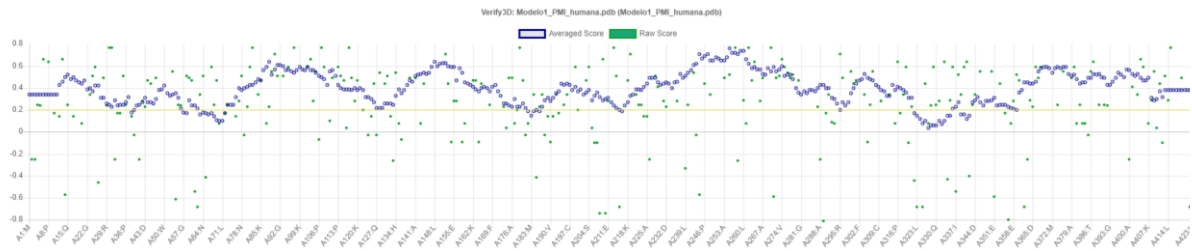
Fonte: O autor

**Figura B12** - Resumo da avaliação estereoquímica da hipótese 1 gerada para PMI *Homo sapiens* realizado pelo PROCHECK com G-Factor de -0.2 e com todos os parâmetros dentro das faixas, com exceção da distorção tetraédrica do carbono alfa que apresenta um melhor resultado com valor de 1.3.



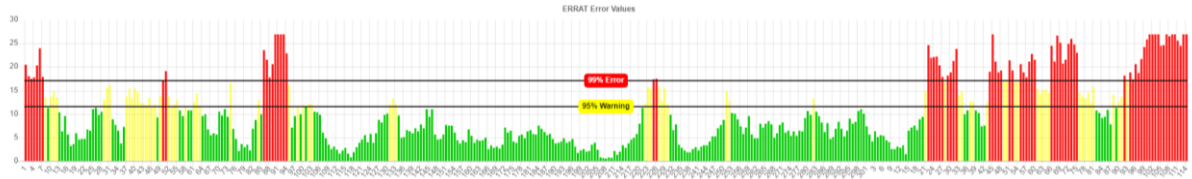
Fonte: O autor

**Figura B13** - Gráfico para a hipótese 1 gerada para PMI *Homo sapiens* obtido pelo **VERIFY3D**: 91,49% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

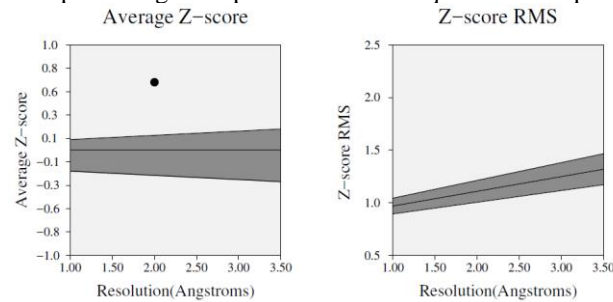
**Figura B14** - Gráfico para a hipótese 1 gerada para PMI *Homo sapiens* obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 64.3373.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

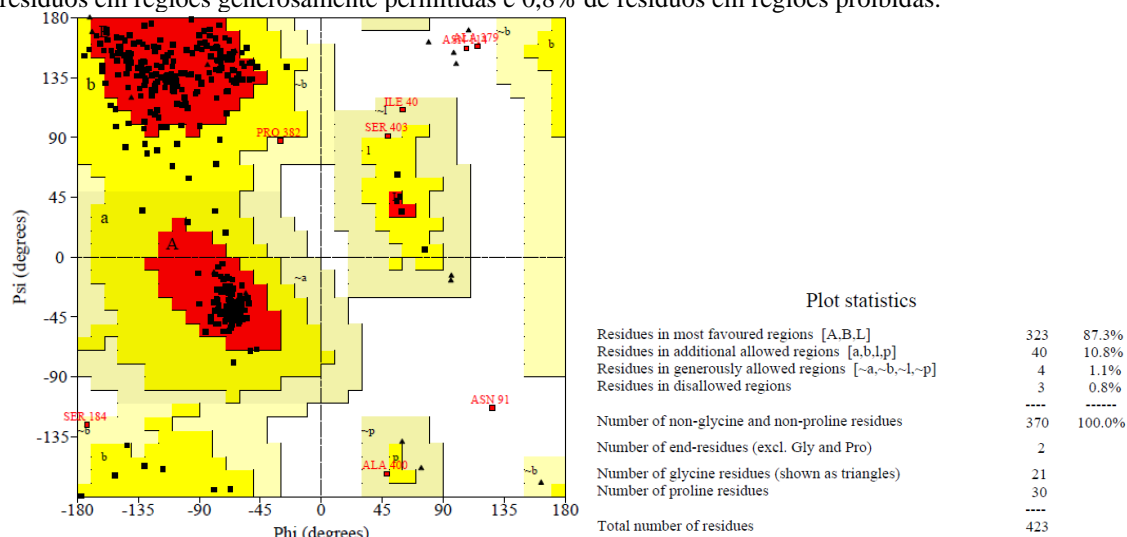
**Figura B15** - Gráfico para a hipótese 1 gerada para PMI *Homo sapiens* obtido pelo **PROVE**.



O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: O autor

**Figura B16** - Diagrama de Ramachandram da hipótese 2 gerada para PMI *Homo sapiens* obtido pelo PROCHECK: 87,3% de resíduos em regiões favoráveis, 10,8% de resíduos em regiões permitidas, 1,1% de resíduos em regiões generosamente permitidas e 0,8% de resíduos em regiões proibidas.



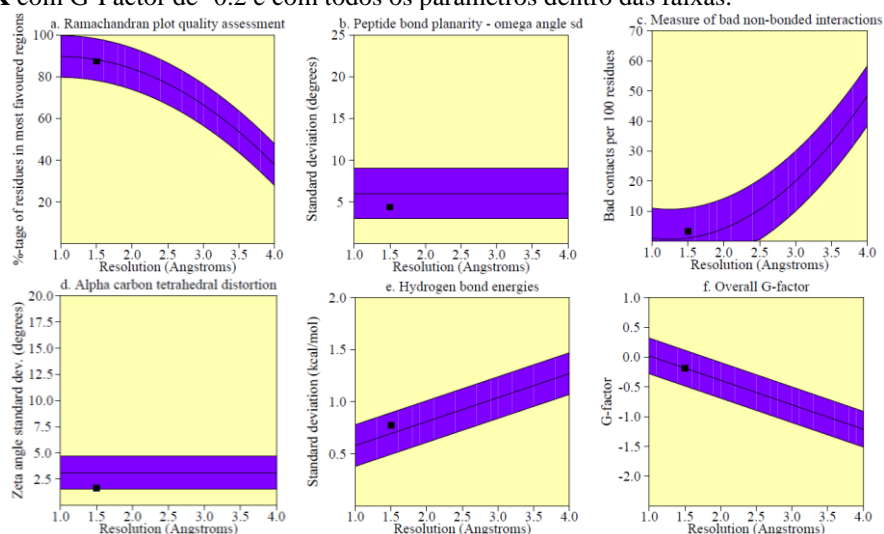
#### Estadísticas de gráfico

Resíduos em regiões mais favoráveis [A,B,L]	323	87.3%
Resíduos em regiões adicionais permitidas [a,b,l,p]	40	10.8%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	4	1.1%
Resíduos em regiões proibidas	3	0.8%
Número de resíduos não glicinas e não prolina	370	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	21	
Número de resíduos de prolina	30	
Número total de resíduos	423	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

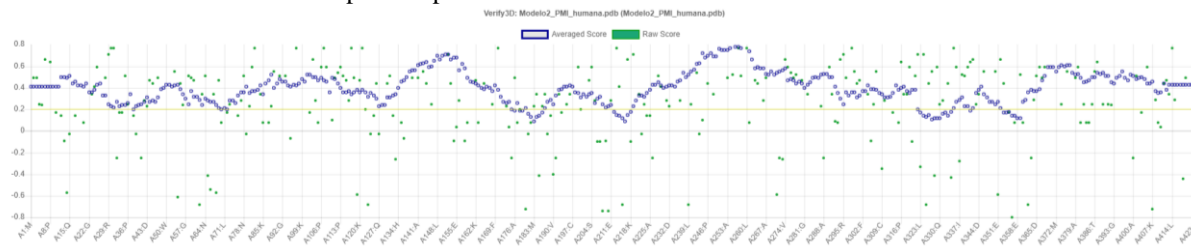
Fonte: O autor

**Figura B17** - Resumo da avaliação estereoquímica da hipótese 2 gerada para PMI *Homo sapiens* realizado pelo PROCHECK com G-Factor de -0.2 e com todos os parâmetros dentro das faixas.



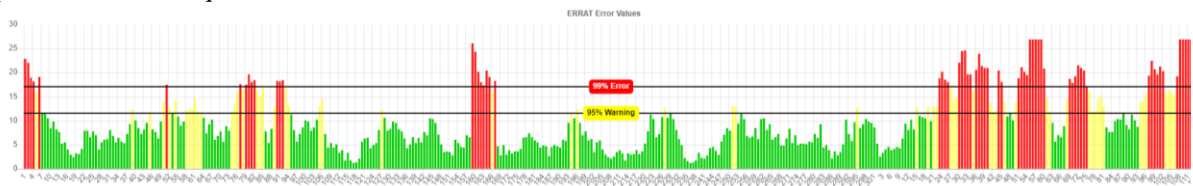
Fonte: O autor

**Figura B18** - Gráfico para a hipótese 2 gerada para PMI *Homo sapiens* obtido pelo **VERIFY3D**: 91,49% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

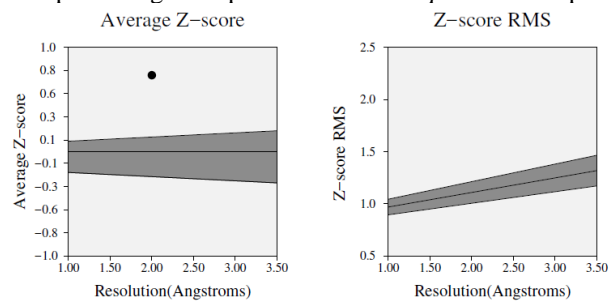
**Figura B19** - Gráfico para a hipótese 2 gerada para PMI *Homo sapiens* obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 68.6747.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

**Figura B20** - Gráfico para a hipótese 2 gerada para PMI *Homo sapiens* obtido pelo **PROVE**.

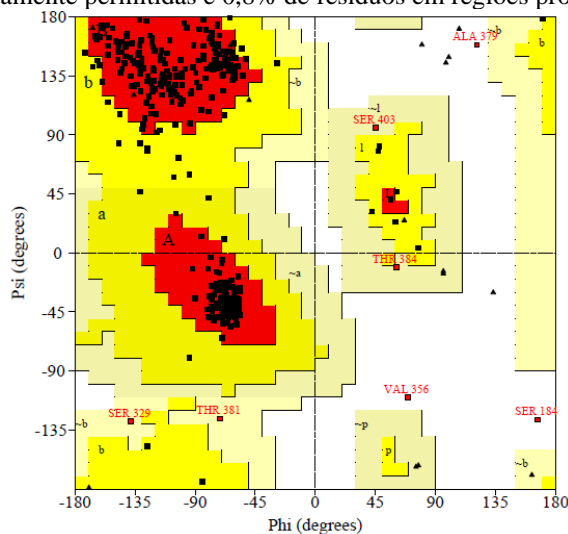


O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: O autor



**Figura B21** - Diagrama de Ramachandram da hipótese 4 gerada para PMI *Homo sapiens* obtido pelo PROCHECK: 88,9% de resíduos em regiões favoráveis, 9,2% de resíduos em regiões permitidas, 1,1% de resíduos em regiões generosamente permitidas e 0,8% de resíduos em regiões proibidas.



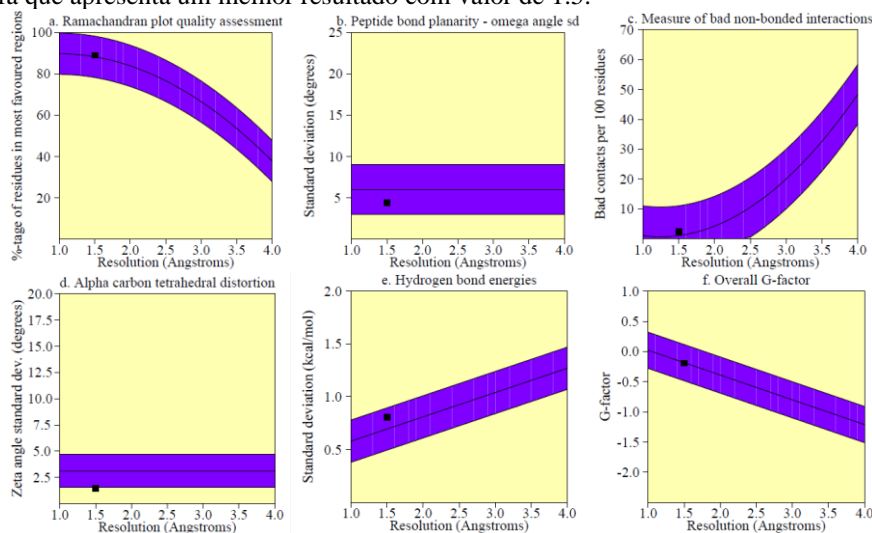
#### Estadísticas de gráfico

Resíduos em regiões mais favoráveis [A,B,L]	329	88.9%
Resíduos em regiões adicionais permitidas [a,b,l,p]	34	9.2%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	4	1.1%
Resíduos em regiões proibidas	3	0.8%
	----	----
Número de resíduos não glicinas e não prolina	370	100.0%
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	21	
Número de resíduos de prolina	30	
	----	
Número total de resíduos	423	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

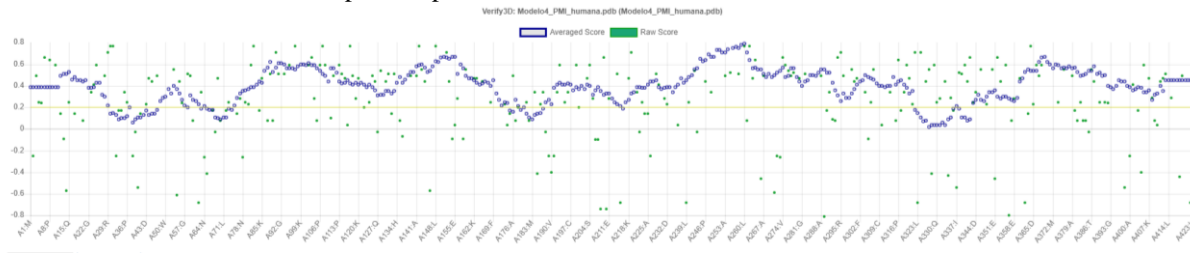
Fonte: O autor

**Figura B22** - Resumo da avaliação estereoquímica da hipótese 4 gerada para PMI *Homo sapiens* realizado pelo PROCHECK com G-Factor de -0.2 e com todos os parâmetros dentro das faixas, exceto a distorção tetraédrica do carbono alfa que apresenta um melhor resultado com valor de 1.5.



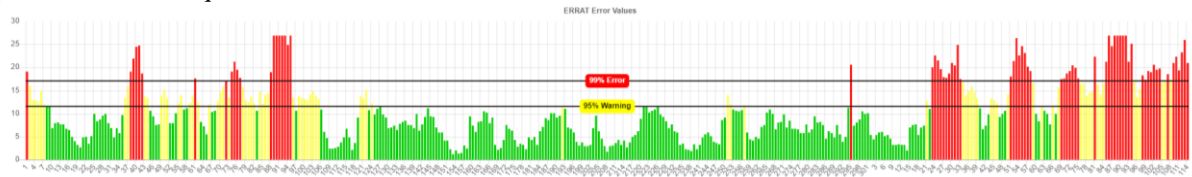
Fonte: O autor

**Figura B23** - Gráfico para a hipótese 4 gerada para PMI *Homo sapiens* obtido pelo **VERIFY3D**: 86,05% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

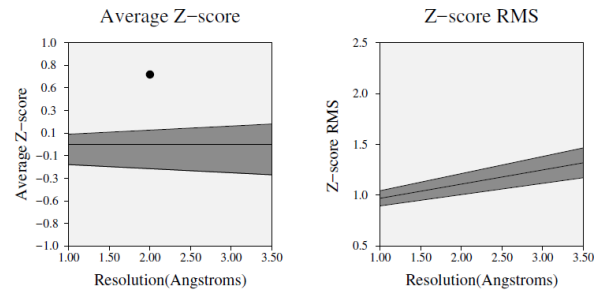
**Figura B24** - Gráfico para a hipótese 4 gerada para PMI *Homo sapiens* obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 63.8554.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

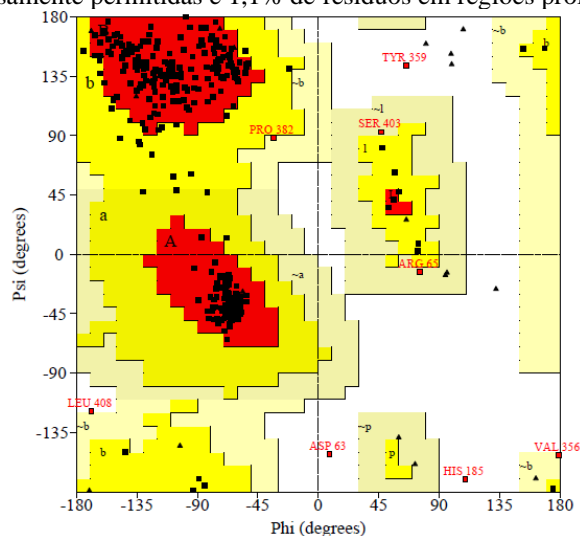
**Figura B25** - Gráfico para a hipótese 4 gerada para PMI *Homo sapiens* obtido pelo **PROVE**.



O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: O autor

**Figura B26** - Diagrama de Ramachandram da hipótese 5 gerada para PMI *Homo sapiens* obtido pelo PROCHECK: 89,2% de resíduos em regiões favoráveis, 8,9% de resíduos em regiões permitidas, 0,8% de resíduos em regiões generosamente permitidas e 1,1% de resíduos em regiões proibidas.



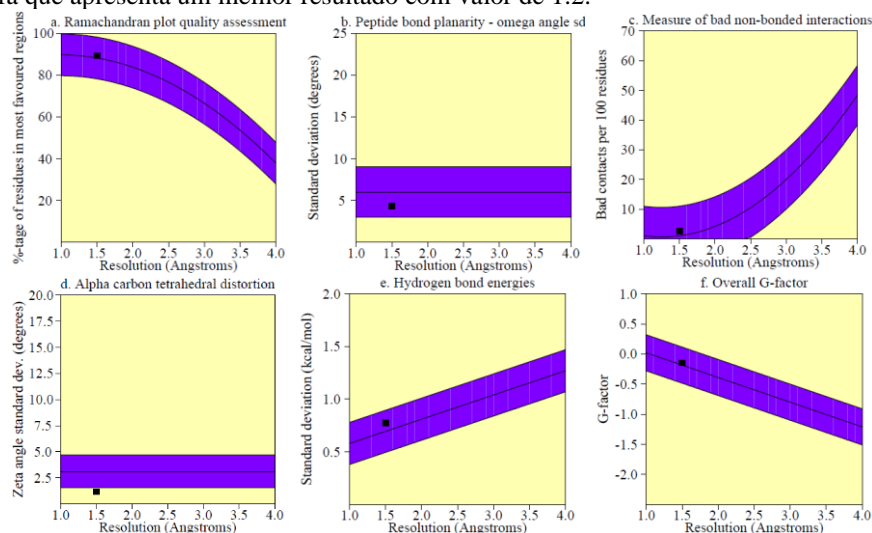
#### Estadísticas de plotagem

Resíduos em regiões mais favoráveis [A,B,L]	330	89.2%
Resíduos em regiões adicionais permitidas [a,b,l,p]	33	8.9%
Regiões em generosamente permitidas [~a, ~b, ~l, ~p]	3	0.8%
Resíduos em regiões proibidas	4	1.1%
Número de resíduos não glicinas e não prolina	---	---
Número de resíduos finais (exc. Glicina e prolina)	2	
Número de resíduos de glicina (mostrados como triângulos)	21	
Número de resíduos de prolina	30	
Número total de resíduos	423	

Com base na análise de 118 estruturas de resolução de pelo menos 2.0 Å e fator R não superior a 20%, espera-se que um modelo de boa qualidade tenha mais de 90% nas regiões mais favorecidas.

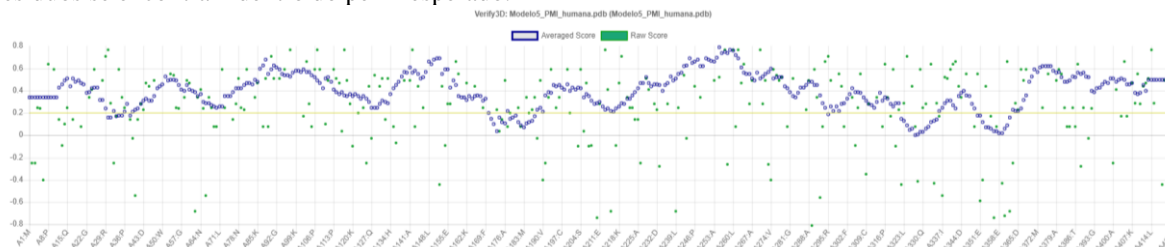
Fonte: O autor

**Figura B27** - Resumo da avaliação estereoquímica da hipótese 5 gerada para PMI *Homo sapiens* realizado pelo PROCHECK com G-Factor de -0.1 e com todos os parâmetros dentro das faixas, exceto a distorção tetraédrica do carbono alfa que apresenta um melhor resultado com valor de 1.2.



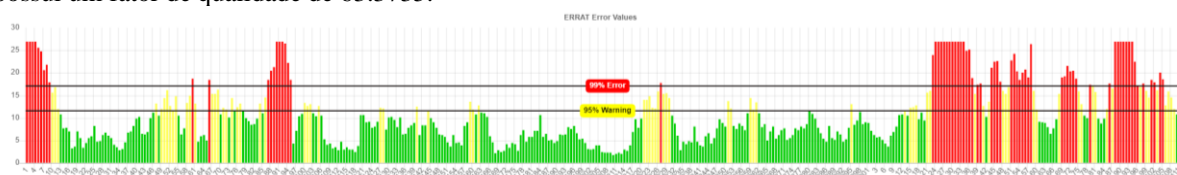
Fonte: O autor

**Figura B28** - Gráfico para a hipótese 5 gerada para PMI *Homo sapiens* obtido pelo **VERIFY3D**: 88,18% dos resíduos se encontram dentro do perfil esperado.



Fonte: O autor

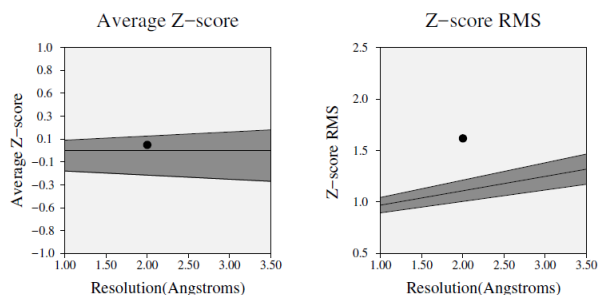
**Figura B29** - Gráfico para a hipótese 5 gerada para PMI *Homo sapiens* obtido pelo **ERRAT**. Essa estrutura possui um fator de qualidade de 63.3735.



A barra vermelha indica o aminoácido com erro acima de 99%, as barras amarelas indicam regiões com erros entre 95% e 99% e as barras verdes indicam as regiões com menor porcentagem de erro para dobrar proteínas.

Fonte: O autor

**Figura B30** - Gráfico para a hipótese 5 gerada para PMI *Homo sapiens* obtido pelo **PROVE**. Essa estrutura apresenta 6,4% dos aminoácidos foram da região de interesse.



O Z-score corresponde aos desvios dos valores do volume atômico em relação aos valores padrão e o Z-score RMS mede a média de volumes irregulares na estrutura da proteína.

Fonte: O autor