



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE DESENVOLVIMENTO SUSTENTÁVEL DO SEMIÁRIDO
CURSO DE ENGENHARIA DE BIOTECNOLOGIA E BIOPROCESSOS**

Rayane Alexandre de Abreu

**CARACTERIZAÇÃO FUNCIONAL E ESTRUTURAL COMPARATIVA
IN SILICO DE UMA PROTEÍNA PR-5 DO TIPO TAUMATINA DE
*Poncirus trifoliata***

**SUMÉ - PB
2015**

RAYANE ALEXANDRE DE ABREU

**CARACTERIZAÇÃO FUNCIONAL E ESTRUTURAL COMPARATIVA
IN SILICO DE UMA PROTEÍNA PR-5 DO TIPO TAUMATINA DE
*Poncirus trifoliata***

Monografia apresentada ao Curso de Engenharia de Biotecnologia e Bioprocessos da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Engenharia de Biotecnologia e Bioprocessos.

Orientadores: Prof. Dr. Rafael Trindade Maia

Prof. Dra. Magnólia de Araújo Campos

SUMÉ - PB

2015

A162c Abreu, Rayane Alexandre de.
Caracterização funcional e estrutural comparativa *in silico*
de uma proteína PR-5 do tipo taumatina de *Poncirus trifoliata*.
/ Rayane Alexandre de Abreu. - Sumé - PB: [s.n], 2015.

54 f.

Orientador: Prof. Dr. Rafael Trindade Maia; Prof^a. Dr^a.
Magnólia de Araújo Campos.

Monografia - Universidade Federal de Campina Grande;
Centro de Desenvolvimento Sustentável do Semiárido; Curso
de Bacharel em Engenharia de Biotecnologia de Bioprocessos.

1. Biotecnologia. 2. Proteína antifúngica. 3. Agricultura.
I. Título.

CDU: 561.28 (043.3)

RAYANE ALEXANDRE DE ABREU

**CARACTERIZAÇÃO FUNCIONAL E ESTRUTURAL COMPARATIVA
IN SILICO DE UMA PROTEÍNA PR-5 DO TIPO TAUMATINA DE
*Poncirus trifoliata***

Monografia apresentada ao Curso de Engenharia de Biotecnologia e Bioprocessos da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Engenharia de Biotecnologia e Bioprocessos.

Aprovada em: 18 de março de 2015

BANCA EXAMINADORA



Profa. Dra. Magnólia de Araújo Campos
Orientadora
UAE/CES/UFCG



Profa. Dra Thaís Gaudêncio do Rêgo
Examinadora Externa
CCGN/UFPB



Profa. Dra Cecir Barbosa de Almeida Farias
Examinadora Interna
UATEC/CDSA/UFCG

SUMÉ – PB

2015

*A Deus, primeiramente, que tem me abençoado
durante toda a minha vida;
E aos meus pais e meu irmão, pela força e incentivo,*

DEDICO.

AGRADECIMENTOS

Sobretudo a DEUS, por me guiar e me dar forças ao longo de toda a minha vida, não somente na vida universitária, mas também em todos os momentos, sei que sempre estará presente em cada detalhe de minha vida.

A Universidade Federal de Campina Grande – CDSA (Campus Sumé), pela oportunidade de realização deste curso;

Ao CNPq, pelo financiamento do projeto e pela concessão da bolsa de Iniciação científica (IC);

Ao Centro de Educação e Saúde (CES), Campus Cuité da UFCG, na pessoa da Professora Magnólia de Araújo Campos, pela oportunidade de desenvolvimento de projetos e aprendizado durante as atividades da bolsa IC e de Estágio Supervisionado no LBiotec;

Ao meu orientador Prof. Rafael Trindade Maia, por aceitar me orientar, pela contribuição ao meu crescimento acadêmico e pelo empenho e dedicação na realização desse trabalho;

A minha sempre e eterna orientadora e amiga Profa. Magnólia de Araújo Campos, por ser exemplo de profissionalismo, força, fé e por essa pessoa iluminada, que contribuiu para o meu crescimento intelectual, moral e espiritual. Mag, você me tornou uma pessoa melhor, muito obrigada por todos os cartões, por todas as risadas, conversas e por seus conselhos sábios, prometo que vou me tornar uma cientista;

Ao Dr. Marcos Antônio Machado, pesquisador do Centro APTA de Citros Sylvio Moreira/IAC, Cordeirópolis SP, por disponibilizar o acesso ao Banco de Dados dos Citros (CitEST), pela colaboração com o INCT Citros;

Aos meus queridos Professores do CDSA pela grande contribuição na minha formação acadêmica;

A coordenadora do curso de Engenharia de Biotecnologia e Bioprocessos Profa Fabiana Pimentel pelos ensinamentos, apoio e dedicação;

Aos pesquisadores da EMBRAPA Cerrados, Brasília-DF, Dra. Marília Santos Silva e Dr. Rodrigo Fragozo pela orientação do estágio como bolsista IC e contribuição na minha formação acadêmica;

Aos meus colegas de curso, pelas alegrias e chatices partilhadas;

Aos meus amigos Gércia, Emanuele, Édipo; pelas alegrias, aperreios de fim de período partilhados;

Aos meus sempre e eternos amigos Felipe Douglas e Luanna Vilar, pelos bons momentos que passamos junto; sem vocês não teria sido a mesma coisa, quero levar essa amizade construída para o resto da vida;

Aos meus primos Erivaldo e Jonailson, pelo companheirismo, alegrias e aperreios.

As minhas primas Gabriela e Morgana e as meninas que moravam com elas pelo apoio e acolhimento em Cuité;

As amigas que Cuité me deu: família LBIotec, Ana Paula, Cristiane Sousa (bundão), Gláucia Medeiros, Graciete Balbino, Ana Luiza, Larissa Cavalcante e Thamara. Meninas, muito obrigada pela amizade e companheirismo de vocês, obrigada pelas risadas, choro e loucuras que aprontamos; vocês tornaram minha estada por Cuité muito agradável;

Aos amores que passaram pela minha vida, durante esse período, pelo carinho e amizade;

As minhas queridas amigas, Girlane, Rayane, Gabriela, Valeska, Morgana e Paulinha que entenderam a minha ausência e com as quais sempre posso contar em todos os momentos da minha vida;

Aos meus avós Antônio, Maria e Carlinda, pelo exemplo de luta e sabedoria;

Aos meus tios, tias, primos e primas, em especial Tia Socorro (*in memoriam*), por ser parte muito importante da minha vida;

Ao meu irmão Raony por todo apoio e amor, obrigada por está sempre ao meu lado, pois além de ser meu irmão, ser meu amigo;

E por fim, devido a grande importância, aos meus pais Josinaldo e Rejane, pelo amor, apoio, compreensão, dedicação, investimento no meu sonho e por nunca deixarem de acreditar em mim até, quando eu não acreditava mais. Vocês são o meu exemplo de vida. E tudo o que conquistei até hoje não teria a menor importância se vocês não estivessem ao meu lado. Amo vocês mais que tudo e é tudo por vocês.

E a todos que contribuíram de alguma forma para o meu crescimento.

Muito Obrigada!

*Uma etapa importante da minha vida chega ao fim.
Descobertas, conhecimentos científicos, inovações...*

“A vida é um círculo e não um quadrado.

Tenha pressa de ser feliz” (Pe. Fábio)

É hora de virar a página.

Avante! Vamos escrever a próxima história!

“Pois, Deus me ama, não estou só.

Deus cuida de mim, quando fala pela sua voz

e me diz: Coragem!” (Adriana Arydes)

RESUMO

A busca por novos antifúngicos tem direcionado esforços para proteínas da imunidade inata de plantas, conhecidas como proteínas relacionadas à patogênese, particularmente da família PR-5. Indicar novos genes que codificam proteínas antifúngicas, dentre milhares de sequências gênicas disponíveis em bancos de dados, é uma especialidade da bioinformática de grande interesse na era pós-genômica, que requer conhecimentos sobre funções e estrutura de proteínas. Neste contexto, este trabalho descreve a caracterização funcional e estrutural comparativa *in silico* de uma proteína PR-5 do tipo taumatina de *Poncirus trifoliata*, denominada PtTLP, por uma estratégia que envolveu técnicas de modelagem e *docking* molecular. A estrutura proposta para a PtTLP apresentou um modelo estável e representativo, dentro dos padrões avaliados, o que pode ser considerado uma representação muito próxima da estrutura real da proteína. A sua estrutura obtida por modelagem comparativa evidenciou características típicas de proteínas TLPs, incluindo a existência de três domínios, oito pontes de sulfeto e uma fenda interdomínios carregada eletronegativamente, denominada sítio ativo da proteína. O modelo construído foi depositado na base de dados PMDB, disponível online. PtTLP interagiu com as moléculas de 1,3- β -glucana, 1,4 - β -glucana, 1,6 - β -glucana, N-acetilglucosamina, 1,3- β -D-glucose e 1,4- β -D-glucose, com interações comprovadas por energias negativas, pela orientação das estruturas na região que comporta o sítio ativo da proteína, ao contrário da interação com a molécula de quitina, que apresentou energia de ligação positiva e ancorou fora do sítio ativo da proteína. Estes dados dão evidências para possíveis atividades de β -glucanase e de ligação a monômeros de carboidratos solúveis para PtTLP e abrem um leque de questões a serem investigadas acerca desta proteína. Estas atividades precisam ser experimentalmente comprovadas, visto que esta proteína é forte candidata, como antifúngico e antioomiceto, para aplicações biotecnológicas na agricultura, na indústria farmacêutica e alimentícia.

Palavras-chave: TLP. Proteína antifúngica. Aplicações biotecnológicas.

ABSTRACT

The search for new antifungal agents has directed efforts to proteins of innate immunity of plants, known as pathogenesis-related proteins, particularly the PR-5 family. To indicate new genes encoding antifungal proteins, among thousands of gene sequences available in databases, is a specialty of bioinformatics of the great interest in the post-genomic era, which requires knowledge on protein structure and function. In this context, this paper describes the comparative functional and structural *in silico* characterization of a PR-5 thaumatin-like protein of *Poncirus trifoliata*, called PtTLP, by using a strategy that involved modeling techniques and molecular docking. The proposed structure for PtTLP showed a stable and representative model within the evaluated patterns, which can be considered a very close representation of the actual structure of the protein. The structure obtained by comparative modeling showed typical features of TLPs proteins, including the existence of three areas, eight disulfide bonds and interdomain cleft electronegatively charged, called active site of the protein. The constructed model was deposited into the PMDB database, and it is online available. PtTLP interacts with the molecules of β -1,3-glucan, 1,4- β -glucan, 1,6- β -glucan, N-acetylglucosamine, 1,3- β -D-glucose and 1,4- β -D-glucose, with interactions evidenced by negative energies and the orientation of the structures in the region comprising the active site of the protein, unlikely the interaction with the chitin molecule, which showed positive binding energy and anchored off the protein active site. These data provide evidence for possible β -glucanase activities and binding the monomers of soluble carbohydrates to PtTLP and open up a range of issues to be investigated on this protein. These activities need to be experimentally proven, since this protein is a strong candidate, as antifungal and antioomycetal, for biotechnological applications in agriculture, pharmaceutical and food industry.

Keywords: TLP. Antifungal protein. Biotechnological applications

LISTA DE FIGURAS

FIGURA1. Fluxograma para a predição da estrutura tridimensional de uma proteína (CAPRILES et al., 2014)	25
Figura 2. O <i>docking</i> molecular permite avaliar a energia de ligação ΔG entre duas moléculas (receptor e ligante). Ilustração obtida em http://dlab.cl/wp-content/uploads/2010/07/Docking-DeltaG.jpg?0cf197	27
Figura 3. A- Sobreposição estrutural do modelo da PR-5 de <i>Poncirus trifoliata</i> obtido através da técnica de modelagem comparativa (azul) e seu respectivo <i>template</i> (vermelho, PDB ID: 3ZS3). B- Representação em faixas (Ribbon) obtida com o programa VMD, mostrando os três domínios: azul (I), vermelho (II) e verde (III). Em <i>laranja</i> estão destacadas as oito pontes de sulfeto	33
Figura 4. Gráficos de energia local ANOLEA e QMEAN em função do teor da estrutura secundária (DSSP).....	34
Figura 5. Comparação do Qmean6 score obtido no modelo com o score de estruturas do PDB.....	35
Figura 6 Gráfico de Ramachandran com 82% dos resíduos em regiões de conformação mais favorável (vermelho), 15,3 % em regiões permitidas, 2,2% em regiões desfavoráveis (bege) e 0,5% em regiões proibidas (branco)	36
Figura 7 Gráfico do resultado do serviço ProSA para validação e confiabilidade do modelo	37
Figura 8 A- Cavidade do <i>cluster 1</i> do Ghecom, um potencial sítio de ligação. B- Representação do mapa de superfície do potencial eletrostático da PtTLP. Em <i>Azul</i> , as regiões eletropositivas, e em <i>vermelho</i> as regiões eletronegativas. Observa-se um <i>cleft</i> eletronegativo das TLPs, característica supostamente associada a atividade antifúngica desta família proteica	39
Figura 9 Gráfico mostrando as regiões da proteína que participam de cada cluster. Em <i>vermelho</i> , os resíduos que participam do melhor cluster do ranking	40
Figura 10 Melhor complexo TLP-1,3- β -D-Glucana do ranking baseado em energia de ligação. Em <i>azul</i> a 1,3- β -D-Glucana, em <i>vermelho</i> resíduos que compõem o pocket do sítio catalítico de acordo com Ghecom. Em <i>amarelo</i> , está destacado o ácido glutâmico, que apresenta um papel relevante na ação beta-glucanase das TLPs	41

Figura 11 Melhor complexo <i>TLP-1,4-β-D-Glucana</i> do ranking baseado em energia de ligação. Em <i>azul</i> a 1,4-β-D-Glucana, em <i>vermelho</i> resíduos que compõem o pocket do sítio catalítico de acordo com Ghecom. Em <i>amarelo</i> , está destacado o ácido glutâmico, que apresenta um papel relevante na ação beta-glucanase das TLPs	42
Figura 12 Melhor complexo <i>TLP-1,6-β-D-Glucana</i> do ranking baseado em energia de ligação. Em <i>azul</i> a 1,6-β-D-Glucana, em <i>vermelho</i> resíduos que compõem o pocket do sítio catalítico de acordo com Ghecom. Em <i>amarelo</i> , está destacado o ácido glutâmico, que apresenta um papel relevante na ação beta-glucanase das TLPs	44
Figura 13 Melhor complexo <i>TLP-quitina</i> do ranking baseado em energia de ligação. Em <i>azul</i> a quitina, em <i>vermelho</i> resíduos que compõem o pocket do sítio catalítico de acordo com Ghecom. Na figura, observa-se que a quitina não interage com o sítio ativo da proteína	45
Figura 14 A- Complexo <i>TLP-N-acetil-glucosamina</i> . Em <i>vermelho</i> os resíduos que fazem parte do sítio ativo apontado no <i>cluster 1</i> pelo Ghecom. Em <i>azul</i> , o N-acetil-glucosamina. Em <i>amarelo</i> , o GLU163. B- Aminoácidos presentes na interação a uma distância de 4Å	46
Figura 15 A- Complexo <i>TLP-1,3-β-D-Glucose</i> . Em <i>vermelho</i> os resíduos que fazem parte do sítio ativo apontado no <i>cluster 1</i> pelo Ghecom. Em <i>azul</i> , a 1,3-β-D-Glucose. Em <i>amarelo</i> , o GLU163. B- Aminoácidos presentes na interação a uma distância de 4Å	47
Figura 16 A- Complexo <i>TLP-1,4-β-D-Glucose</i> . Em <i>vermelho</i> os resíduos que fazem parte do sítio ativo apontado no <i>cluster 1</i> pelo Ghecom. Em <i>azul</i> , a 1,4-β-D-Glucose. Em <i>amarelo</i> , o GLU163. B- Aminoácidos presentes na interação a uma distância de 4Å	48

SUMÁRIO

1 INTRODUÇÃO.....	14
2 JUSTIFICATIVA	14
3 OBJETIVOS	16
3.1 Objetivo geral	16
3.2 Objetivos específicos	16
4 REFERENCIAL TEÓRICO	17
4.1 Importância da citricultura	17
4.2 Proteínas PR-5.....	18
4.3 Modelagem molecular.....	21
4.4 Dinâmica molecular.....	25
4.5 <i>Docking</i> molecular	26
5 METODOLOGIA.....	28
5.1 Modelagem da PtTLP de <i>Poncirus trifoliata</i>	28
5.1.1 Refinamento do modelo	29
5.1.2 Validação do modelo.....	29
5.2 <i>Docking</i> molecular	29
5.2.1 Preparação do receptor.....	29
5.2.2 Mineração e preparação dos ligantes	30
5.2.3 Simulações de <i>docking</i>	30
6 RESULTADOS E DISCUSSÃO	32
6.1 Construção e validação do modelo 3D da PtTLP:.....	32
6.2 Predição do sítio ativo da PtTLP	38
6.3 Caracterização <i>in silico</i> da atividade antifúngica por ligação a glucanas	40
6.4 Atividade de ligação a carboidratos solúveis	45
7 CONCLUSÕES	49

1 INTRODUÇÃO

Na era pós-genômica existem milhares de sequências gênicas em bancos de dados que necessitam comprovar suas funções, especialmente aquelas com potencial para aplicações biotecnológicas, tais como antifúngicos. A busca e identificação de novos antifúngicos é crescente, visando aplicação tanto na agricultura quanto na área farmacêutica, com o intuito de melhorar a qualidade de vida humana no campo e fora dele. A descoberta e comprovação da atividade antifúngica de proteínas por métodos laboratoriais são demorados e caros.

Nesse sentido, as análises computacionais contribuem e aceleram as descobertas de proteínas preditas a partir de sequências gênicas, pois, com base no dogma molecular, a informação genética para os caracteres dos seres está armazenada no DNA (ácido desoxirribonucleico), é transcrita para RNA (ácido ribonucleico) e depois decodificada ou traduzida para cadeias polipeptídicas ou proteínas funcionais. Dessa forma, através de vários programas computacionais é possível propor modelos proteicos tridimensionais (3D), determinar os sítios ativos das proteínas, demonstrar se estas interagem com outras proteínas, ou ligantes de outra origem não proteica e especular sobre o que esta interação pode causar, entre outros. Portanto, análises *in silico* apontam proteínas candidatas para uma posterior confirmação em laboratório.

2 JUSTIFICATIVA

Nas duas últimas décadas, diversos estudos têm focado em várias classes de proteínas antimicrobianas de plantas, as proteínas relacionadas à patogênese, conhecidas também como proteínas PR (de *Pathogenesis-Related*), atualmente classificadas como PR-1 a PR-17 (VAN LOON, REP e PIETERSE, 2006). Proteínas PR fazem parte do sistema imune inato de plantas e animais e, dentre aquelas com atividades antifúngicas, as da família PR-5 são de particular interesse. Para vários membros dessa família, classificados como proteínas do tipo taumatinas e osmotinas, já foi demonstrada a participação na resistência sistêmica adquirida em resposta a estresse biótico, na resistência a estresses abióticos e a ação antifúngica de morte ou inibição do crescimento de hifas e redução na germinação de esporos de fungos, por mecanismos que envolvem a permeabilização da membrana de

fungos e oomicetos, apoptose; atividade de glucanase ou a interação com receptores de patógenos (THOMPSON et al., 2006).

Dentre os bancos de dados genômicos brasileiros de grande importância agro-econômica, o banco de dados do genoma funcional e comparativo dos citros e de alguns de seus patógenos (CitEST), do Centro APTA de Citros Sylvio Moreira/IAC, Cordeirópolis SP, destaca-se por disponibilizar uma ampla diversidade de sequências expressas (ESTs, de *Expressed Sequence Tags*) em diferentes tecidos de 11 espécies de citros, sob condições contrastantes de estresses bióticos e abióticos.

Devido a importância da citricultura para o mundo e para o Brasil, como maior *commodity* brasileira de exportação, bem como os problemas causados à cultura dos citros por fatores bióticos e abióticos; análises computacionais realizadas no banco de dados CitEST vêm contribuindo para elucidar mecanismos de resposta de plantas de citros às condições estudadas. Neste contexto, em estudos anteriores, nós caracterizamos genes *PR-5* significativamente expressos em tecidos da planta *Poncirus trifoliata*, cujas sequências de aminoácidos deduzidas a partir dos unigenes compartilham alta similaridade de sequência (79 a 99%) com outras *Thaumatin-like protein* de *Citrus jambhiri*, *Vitis vinifera*, *Castanea sativa* e *Ricinus communis* (ARAÚJO et al., 2013). Devido ao fato de *P. trifoliata* apresentar tolerância a patógenos de citros, inclusive o oomiceto *Phytophthora parasitica*, o agente causal da doença gomose dos citros, os genes *PR-5* descobertos representam candidatos para validação da expressão gênica e para estudos estruturais e funcionais das proteínas codificada por eles.

Portanto, a modelagem por homologia associada ao *docking* molecular representa uma excelente estratégia para análise funcional de proteínas *PR-5* de *P. trifoliata*, deduzidas a partir da sequência gênica. As proteínas com características de interesse poderão ser utilizados em aplicações biotecnológicas para a cadeia produtiva dos citros.

3 OBJETIVOS

3.1 Objetivo geral

- Realizar a caracterização funcional e estrutural comparativa *in silico* de uma proteína PR-5 do tipo taumatina de *Poncirus trifoliata*.

3.2 Objetivos específicos

- Construir um modelo tridimensional *in silico* por comparação da sequência de aminoácidos predita a partir do gene *PtTLP* com outras funcionais disponíveis em bancos de dados;
- Utilizar *docking* molecular para prever a afinidade de ligação entre a proteína PR-5 PtTLP de *Poncirus trifoliata* e os ligantes estudados;
- Analisar possíveis funções para a PR-5 PtTLP.

4 REFERENCIAL TEÓRICO

O presente estudo busca por novas estratégias biotecnológicas para controle de patógenos de plantas e se justifica na importância da citricultura. Isso porque a cultura dos citros é atacada por diversas doenças e pragas, que reduzem a produtividade das plantas, os danos causados por esses patógenos provocam muitas perdas todos os anos. A *Phytophthora* spp. é um dos principais patógenos encontrados causando doenças em citros em todas as áreas produtoras do mundo (TEIXEIRA, 2005). Os danos causados por esse patógeno afetam a produção mundial, pois afeta o desenvolvimento dos citros e sua produção, o que gera uma perda mundial de milhões de dólares. Proteínas antifúngicas e antioomicetais, como proteínas PR-5, representam candidatas muito interessantes para controle de fungos e oomicetos. Apesar dos bancos de dados genômicos disponibilizarem uma gama de informações gênicas, a função de proteínas preditas ainda permanece uma lacuna, que pode ser minimizada pela aplicação de programas computacionais, capazes de indicar a funcionalidade das sequências de aminoácidos deduzidas com eficiência e rapidez.

4.1 Importância da citricultura

No setor da agroindústria brasileira, uma das culturas que tem um maior destaque é a citricultura, detentora de cerca de 60% da produção mundial de suco de laranja, além de ser o campeão em exportação do produto. O cultivo de laranja no Brasil, atualmente, é de cerca de 30% da safra mundial da fruta. Para manter a liderança do setor, o Ministério da Agricultura investe no apoio a adoção de sistemas mais eficientes, como a produção integrada, com medidas para reduzir os custos, aperfeiçoar e ampliar a comercialização do produto. O ministério tem, ainda, ação efetiva na fiscalização e prevenção ao aparecimento de pragas e doenças. Mais da metade da produção mundial e cerca de 80% da brasileira é destinada a produção de suco industrializado. O principal comprador da bebida brasileira é a União Europeia que aumenta significativamente o percentual de importação anualmente. Fonte: Ministério da Agricultura.

Com o avanço da citricultura no país, o Centro Avançado de Pesquisa Tecnológica do Agronegócio de Citros Sylvio Moreira (Centro APTA Citros Sylvio Moreira, ou simplesmente Centro de Citricultura Sylvio Moreira), consolidou-se na comunidade citrícola como um Centro difusor de material genético e de tecnologia, crescendo junto com a citricultura brasileira. O Centro é, atualmente, uma referência nos assuntos de pesquisa e desenvolvimento, procurando interface com todos os agentes da cadeia de produção e do conhecimento. Fonte: Centro de Citricultura Sylvio Moreira.

O projeto Citros EST (CitEST) com mais de 300.000 clones sequenciados é um projeto de sequenciamento de transcriptoma, que tem como objetivo associar conhecimentos em mapeamento genético com novos dados sobre genômica funcional e comparativa das espécies de citros e gêneros afins (REIS et al., 2007). O projeto centra-se na resposta das plantas sob diferentes estresses bióticos e abióticos. Esse projeto também incluiu a integração dos estudos genômicos funcionais de agentes patogênicos, cujos genomas de citros foram totalmente sequenciados (vírus e bactérias) no cenário de interações hospedeiro-agente patogênico.

Uma das espécies de grande interesse e estudada no genoma funcional e comparativo dos citros foi *Poncirus trifoliata* (L.). Também denominada trifoliata, essa espécie é um dos porta-enxertos mais utilizados na citricultura mundial, ocupando posição de destaque no Japão, Austrália, Nova Zelândia, Uruguai e Estados Unidos da América (PASSOS et al., 2006). Essa preferência dá-se em função de certas características induzidas por esta espécie às combinações copa/porta-enxerto, tais como porte anão, tolerância à gomose de *Phytophthora*, tolerância ao vírus da tristeza e a nematoides, e produção de frutos de excelente qualidade, embora com baixa média de produtividade e baixa tolerância à seca (CASTLE et al., 1989).

4.2 Proteínas PR-5

As plantas possuem mecanismos estruturais e bioquímicos pré e/ou pós-formados de defesa a ataque de patógenos. Os mecanismos estruturais são barreiras físicas a penetração do patógeno. Já as barreiras bioquímicas englobam substâncias que vão da capacidade de inibir o crescimento do patógeno até gerar condições adversas a sua sobrevivência no hospedeiro. O modo como esses

mecanismos vão atuar, variam de acordo com a resposta que a planta tem com o órgão/tecido afetado, com a nutrição e com as condições ambientais na qual a planta está submetida (STANGARLIN et al., 2011).

Como barreira bioquímica existe uma família de proteínas relacionadas à defesa contra patógenos, as proteínas PR, atualmente classificadas em 17 famílias (PR-1 a PR-17). Inicialmente, cinco famílias de proteínas PR (PR-1 a PR-5) foram descobertas em 1990, acumuladas intra e extracelularmente em tecidos de vegetais intactos ou em cultura de células, após o tratamento com elicitores ou ataque de patógenos (BOWLES, 1990). Dentre aquelas que possuem atividade antifúngica, a família de proteínas PR-5 merece destaque pela sua diversidade funcional.

Em plantas, proteínas PR-5 podem apresentar isoformas ácidas, que geralmente se encontram em meio extracelular, ou básicas, que são encontradas em meio intracelular e são constitutivamente expressas em altos níveis nas raízes. Uma diferença que também pode ser observada entre as isoformas ácidas e básicas é que estas últimas são induzidas em altos níveis nas folhas por etileno, enquanto que as primeiras não o são (NASSER; DE TAPIA; BURKARD, 1990). Essas proteínas são amplamente distribuídas em diversos reinos como animal, vegetal e sistemas fúngicos, que vai de fungos verdadeiros a organismos do tipo fungo; como os Oomicetos. A massa molecular de proteínas PR-5 varia entre 21 a 26 kDa, em geral têm 16 resíduos de cisteína conservados, que estão envolvidos na formação de oito ligações de pontes dissulfeto, que conferem estabilidade à proteína sob condições térmicas e variadas de pH (VELAZHAHAN; DATTA; MUTHUKRISHNAN, 1999).

As osmotinas (*Osmotin-like Proteins* ou OLPs) são isoformas básicas, de aproximadamente 26 kDa, cuja denominação foi dada devido a sua expressão ser induzida em condições de baixo teor hídrico, que chega a cerca de 15 vezes mais acumulada que em níveis hídricos normais. Essa proteína inibe o crescimento de hifas de *Bipolaris*, *Fusarium* e *Phytophthora* causando a lise do esporo, além de inibir a germinação, reduzindo assim sua viabilidade (WEBER, 2007).

As proteínas PR-5 do tipo taumatinas (*Thaumatococin-Like Proteins* ou TLPs) são isoformas básicas, que têm esse nome por sua sequência ter alta homologia com a sequência de uma proteína de sabor adocicado, denominada taumatina, isolada de *Thaumatococcus daniellii* (CORNELISSEN; HOOFT VAN HUIJSDUIJNEN; BOL, 1986). Taumatina não possui atividade antifúngica, mas as TLPs têm. As TLPs são

proteínas com atividade antifúngica *in vitro* descrita contra uma variedade de organismos patogênicos, fungos filamentosos, como *Botrytis cinerea*, *Fusarium oxysporum*, *Mycosphaerella arachidicola*, *Trichoderma viride*, entre diversos outros patógenos (GARCIA-CASADO et al., 2000).

Essas proteínas têm capacidade de romper a membrana dos fungos pela formação de esporos, além de afetar o crescimento dos fungos com atividades inibidoras β -glucanase e xilanase (SINGH et al., 2013). Além do papel previamente identificado de TLPs na resistência do patógeno, essas proteínas podem atuar com papel de destaque na tolerância ao frio, salinidade e resistência à seca, entretanto, a resistência que as TLPs conferem contra o estresse abiótico (LIU; STURROCK; EKRAMODDOULLAH, 2010). As TLPs têm diversas atividades descritas na literatura como antifúngicas, anti-congelamento, que envolvidas em processos de amadurecimento, podem atuar de diversas formas em diferentes organismos (SINGH et al., 2013).

Algumas TLPs apresentam uma atividade endo-beta-1,3- glucanase. Estudos mostram que essa atividade está relacionada a uma fenda eletronegativa na superfície dessas proteínas, mostram ainda que o posicionamento linear das (1,3) - beta-D-glucanas na fenda das proteínas permite uma interação com os resíduos de ácido glutâmico que são responsáveis pela clivagem hidrolítica do glucano (MENU-BOUAOUICHE et al., 2003). Entretanto, o posicionamento dessa fenda eletromagnética na superfície da proteína é quem irá determinar sua atividade enzimática. Através de estudos das características químicas, biológicas e estruturais podemos determinar que atividades as proteínas encontradas nessa família podem desenvolver.

A identificação ou mutação de um gene conhecido ou novo, pode nos levar a descobrir diversas coisas a respeito desse gene, como ele age, por que está expresso, em que situações está ou não expresso, o que vai acontecer caso deixe de ser expresso. São diversas questões a serem tratadas por meio do estudo desse gene e de sua sequência e os resultados que esses estudos podem levar ao desenvolvimento, por exemplo; de novos fármacos, uma planta transgênica resistente a determinadas doenças ou estresse, há diferentes aplicações biotecnológicas.

4.3 Modelagem molecular

Sequências codificantes geram estruturas tridimensionais, e com isso efeitos biológicos podem ser estudados e avaliados de forma mais precisa: o estudo estrutural de proteínas gera milhões de dados e muitas possibilidades de estudo. Os resultados gerados atuam como base para desenvolvimento de trabalhos experimentais. A predição de estruturas tridimensionais de proteínas se caracteriza por possuir aplicações práticas de grande impacto terapêutico e biotecnológico. Está diretamente relacionada a múltiplas áreas da bioinformática e modelagem molecular, tais como a interação proteína-ligante, o desenho de novas proteínas com funções específicas (nanotecnologia e engenharia de proteínas) e a própria elucidação de estruturas a partir de dados experimentais (CAPRILES et al., 2014).

A informação acerca da estrutura de uma proteína está armazenada em sua sequência codificada dos genes de um organismo, a estrutura nativa, que é uma macromolécula estável em condições fisiológicas formadas por ligações peptídicas entre aminoácidos, que apesar de ser estável não é estática; muitas vezes a função dependerá da flexibilidade da molécula (CAPRILES et al., 2014). As proteínas na sua forma nativa devem apresentar características como baixa energia livre e uma relevância biológica, para que a cadeia assuma suas funções no meio onde se encontra.

Atualmente, existem duas técnicas principais para determinação experimental da estrutura tridimensional de uma proteína, que pode ser via técnica de cristalografia por difração de raios-X e a técnica de Ressonância Magnética Nuclear (RMN). Entretanto, essas técnicas apresentam algumas limitações que variam de acordo com cada tipo de proteína, como a demora no isolamento e purificação, assim como o tempo gasta para isso que pode levar anos. Além de ser um processo muito caro, alguns desses problemas podem tornar o método experimental não confiável. De acordo com esse resultado, os métodos computacionais vêm se tornando cada vez mais utilizados para predição de proteínas, o que trará um maior embasamento para que a parte experimental seja realizada de forma mais adequada, visando um resultado previamente determinado.

Para obtenção de um modelo tridimensional existem diferentes tipos de métodos computacionais empregados, sua utilização irá variar de acordo com a complexidade da estrutura ou poucos estudos relacionados ao organismo, ou até

mesmo determinada classe da proteína. Os métodos de modelagem molecular atualmente disponíveis são *ab initio*, “threading” e modelagem por homologia ou comparativa (DORN, 2008).

A modelagem *ab initio* tenta prever a forma nativa de uma proteína, apenas a partir da sua sequência de aminoácidos. Conduz uma busca conformacional sob a orientação de energias, gerando múltiplas conformações possíveis de onde os modelos finais são selecionados (NICOLAU JUNIOR, 2013). Essa metodologia requer três elementos: uma representação da geometria da proteína, uma função potencial e uma técnica de busca do espaço da superfície da energia livre (OSGUTHORPE, 2000). São calculadas as energias envolvidas no processo de enovelamento, baseada em hipóteses termodinâmicas na qual indica a estrutura com menor energia livre global (MURZIN et al., 1995). Esse método pode ser empregado para predição de estruturas, nas quais não se sabe detalhes de sua estrutura (OSGUTHORPE, 2000).

A modelagem *Threading* é uma metodologia que consiste na entrada de uma sequência de aminoácidos, sem estruturas caracterizadas, e a partir da mesma um modelo é computado baseado em um banco de topologias existentes, por exemplo, SCOP (MURZIN et al., 1995). Esta metodologia é frequentemente usada com intuito de descobrir homologias remotas que não podem ser descobertas por alinhamento de sequências padrão (GIBAS; JAMBECK, 2001). Esse tipo de modelagem molecular *in silico* simula soluções de problemas científicos relacionados à estrutura proteica, elaborando informações relevantes a posicionamento e relacionamento dos aminoácidos para sua descrição, é um problema sofisticado que resulta em um elevado número de variáveis físico-químicas a serem manipuladas, sem contar as inúmeras variáveis conhecidas e desconhecidas envolvidas no enovelamento (MENETTI, 2006).

A modelagem por homologia é feita através da sobreposição de uma proteína com a estrutura tridimensional previamente determinada experimentalmente e depositada no banco de dados PDB (*Protein Data Bank*), que tenha uma alta similaridade, para que a modelagem *in silico* seja o mais próximo do real possível. Trata-se de uma área multidisciplinar que visa estabelecer modelos confiáveis para estruturas biológicas e fármacos (MENETTI, 2006).

A modelagem por homologia ou comparativa vem sendo mais utilizada por ser um método eficaz, que gera bons resultados e menos demorados que os métodos descritos acima. Além da qualidade dos modelos preditos, que possuem uma razoável relação evolucionária, apresentarem uma acurácia maior do que aquelas produzidas com técnicas diferentes, a confiabilidade e qualidade da estrutura predita pode ser estimada por diversos parâmetros, tais como gráficos de ANOLEA e QMEAN para verificar a qualidade local do modelo, os resultados do Dfire e QMEAN6 mostram a qualidade global do modelo e gráfico de Ramachandram revela a qualidade estereoquímica, o que torna o modelo confiável (TRAMONTANO; LESK, 2006). Para uma melhor compreensão destas técnicas, alguns conceitos devem ser levados em consideração; como proteína homóloga, similaridade e identidade entre proteínas.

A homologia refere-se à ancestralidade das proteínas, se elas apresentam ancestral comum, derivam de uma sequência em comum, a identidade refere-se à presença de um mesmo aminoácido na mesma posição em sequências diferentes, já a similaridade é a presença local de análogos em duas sequências (GIBAS; JAMBERCK, 2001). A modelagem por essa técnica é realizada utilizando uma ou mais proteínas-molde através de estruturas experimentalmente resolvidas de um banco de dados, como o PDB (*Protein Data Bank*) que tenha uma identidade acima de 25% com a proteína a ser modelada (DORN, 2008). A escolha da proteína-molde (*template*) irá depender da identidade, similaridade e da resolução com o *template*, ou seja se o mesmo e é considerado adequado para a modelagem.

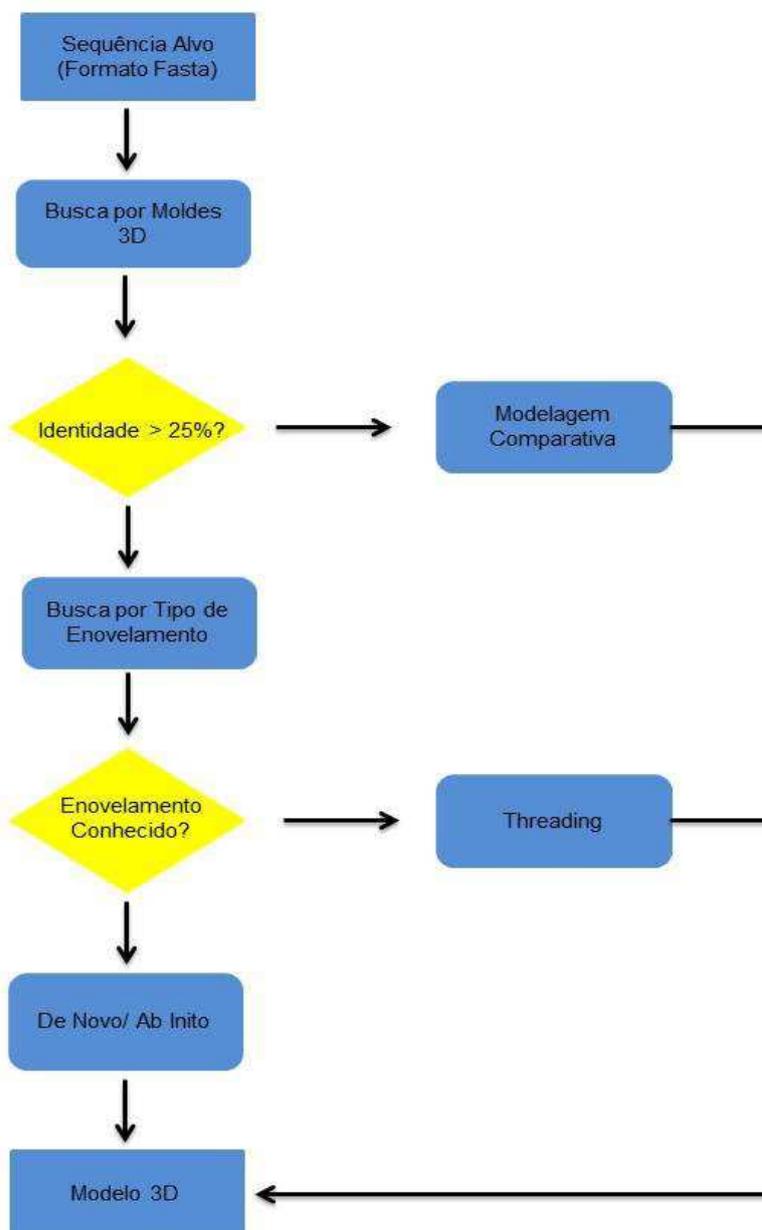
A técnica consiste basicamente em seleção da estrutura molde em banco de dados por sequências similares, o alinhamento das sequências alvo e molde, a modelagem propriamente dita, refinamento do modelo e validação através da utilização de ferramentas de bioinformática. O modelo final deve ter parâmetros físico-químicos, como qualidade local, global e estereoquímica do modelo aceitáveis (NICOLAU JUNIOR, 2013). Além de uma identidade acima de 25% entre a proteína alvo e a proteína-molde para que a modelagem possa ser considerada aceitável e de boa qualidade, e seja menos passível a erros na modelagem. A probabilidade de ocorrer erros em modelos de proteínas construídos através desse método é inversamente proporcional ao grau de identidade entre as sequências da proteína alvo e o molde utilizado. Os erros não ocorrem igualmente em toda a estrutura o que

permite a criação de um modelo satisfatório, desde que os critérios e procedimentos sejam realizados de forma correta (PACHECO, 2009).

A modelagem por homologia pode ser empregada, sempre que houver uma relação evolucionária entre a proteína a ser modelada e a proteína molde ou *template* (BUJNICKI, 2006). A relação evolucionária é uma característica crucial dessa técnica, uma proteína só pode ser modelada a partir de *templates* homólogos com estrutura tridimensional determinada experimentalmente. Essas proteínas podem ser modeladas por essa técnica porque em proteínas homólogas a estrutura destas proteínas é similar onde os aminoácidos tem características físico-químicas idênticas ocupam posições equivalentes, além dos ângulos de torção da cadeia principal também preservam um certo padrão em seus valores(DORN, 2008).

A Figura 1 mostra um fluxograma para a predição de estrutura tridimensional de uma proteína, de acordo com os três utilizados atualmente, mostrando como pode ser decidido que método utilizar para modelagem de cada proteína, dependendo das informações das respostas de análises da sequência de aminoácido da proteína.

Figura 1. Fluxograma para a predição da estrutura tridimensional de uma proteína
Fonte: (CAPRILES et al., 2014).



4.4 Dinâmica molecular

Baseada nos princípios da Mecânica Clássica e mostrando o comportamento microscópico dinâmico de átomos individuais que constituem um sistema, como uma proteína. Uma simulação por dinâmica molecular (DM) produz um conjunto de conformações da proteína em função do tempo. A partir da conformação em

equilíbrio, o valor médio de parâmetros termodinâmicos como pressão, temperatura, volume, calor específico; pode ser calculado, assim como parâmetros estruturais, incluindo o raio de giro e a estrutura média da proteína (DALL'AGNO, 2012).

De acordo com Becker (2001), entre pico (ps) e femtosegundos (fs) apenas movimentos de átomos nas cadeias laterais, vibrações de ligações e flexibilidade de ligantes são percebidos, já em nano (ns) a microssegundos (μ s) é possível a visualização e análise de mudanças em conformações de sítios ativos, movimentos de dobramento, rotação de ligação de peptídeos, rotação de ligações e formação de *loops*. Na dinâmica molecular a trajetória de um sistema molecular é causada pela solução simultânea da equação de movimento de Newton para todos os átomos do sistema molecular. Por meio da segunda lei de Newton é possível calcular a força de ação de cada átomo. Cada átomo de uma proteína possui uma energia potencial e sente conseqüentemente uma força externa igual à derivação espacial desta energia (DORN, 2008).

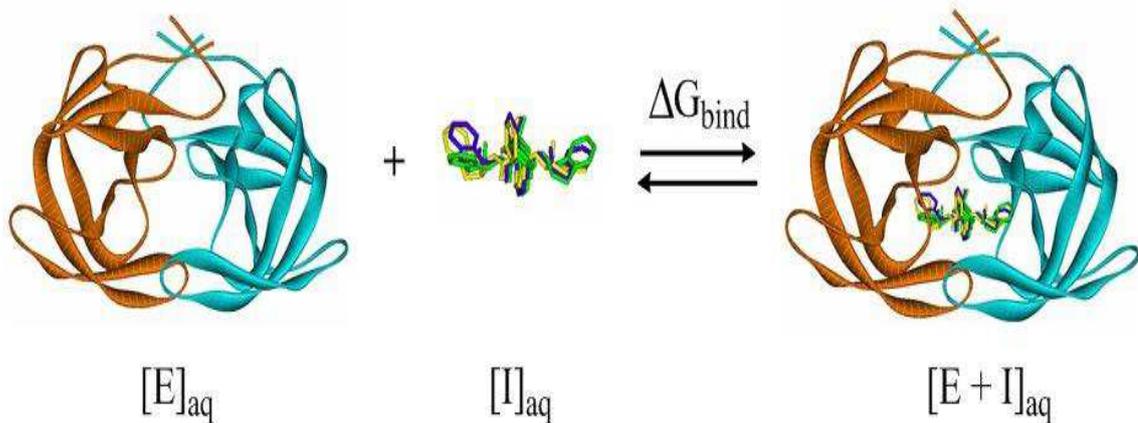
4.5 Docking molecular

Docking molecular é uma ferramenta para predição de possíveis interações entre as moléculas. Segundo Guedes, Magalhães e Dardenne (2014), para a compreensão de mecanismos e processos celulares é necessário determinar e compreender interações entre macromoléculas ou entre uma macromolécula e uma pequena molécula ligante, que pode atuar como agonista/antagonista ou substrato/inibidor em determinado processo fisiológico. Esse tipo de interação pode ocorrer entre proteínas, carboidratos ou outras moléculas.

O processo de *docking* molecular simula o curso natural de interação um ligante no sítio ativo de um receptor (macromolécula) por via de menor energia por meio da complementariedade química entre as estruturas. Através dessa interação e por ferramentas de bioinformática, é possível estimar a energia média da interação receptor-ligante (SOUSA, 2011) (Figura 2). As informações obtidas a partir do *docking* molecular, como quais as interações potenciais que ocorrem no sítio ativo, como interações hidrofóbicas, eletrostáticas, ligações de hidrogênio no complexo enzima-inibidor e como a interação ocorreu (ERICKSON et al., 2004). Esse método procura a conformação geométrica do complexo que necessite de uma menor energia para que ocorra a ligação, para que isso ocorra, o processo realizado gera

vários complexos possíveis, entretanto, só se pode especular a melhor solução com base nas informações geradas a partir do *docking* e das informações a cerca das estruturas utilizadas (SOUSA, 2011).

Figura 2. O *docking* molecular permite avaliar a energia de ligação ΔG entre duas moléculas (receptor e ligante). Fonte: <http://dlab.cl/wp-content/uploads/2010/07/Docking-DeltaG.jpg?0cf197>.



Segundo Rognan (2011), o método de *docking* é realizado em duas etapas: um algoritmo de busca conformacional que envolva todos os graus de liberdade do ligante e uma função de pontuação que ranqueie as prováveis posições espaciais para um composto. Atualmente, é a técnica dominante para identificação de novos compostos, candidatos a fármacos como a prospecção de bibliotecas de substâncias contra algum tipo de alvo biológico (NICOLAU JUNIOR, 2013). Diversos estudos acerca de substâncias depositadas em bibliotecas são feitos relacionando a melhor interação, para testes experimentais para possíveis potenciais fármacos ou desenvolvimento de organismos geneticamente modificados. Abordagens *in silico* que possam reduzir os custos e acelerar o processo de descoberta e planejamento de novos produtos para o mercado farmacêutico, alimentício e cosmético são extremamente bem vindas e necessárias.

5 METODOLOGIA

O gene *PtTLP* foi encontrado significativamente expresso em *Poncirus trifoliata*, durante interação com o oomiceto *Phytophthora parasitica* e caracterizado por Araújo et al. (2013), usando a informações do banco de dados do genoma funcional e comparativo dos citros (CitEST), gerado pelo Centro APTA dos Citros Sylvio Moreira do Instituto Agrônomo de Campinas, Cordeirópolis SP, gentilmente cedido a Profa. Magnólia de Araújo Campos (CES/UFCG) pelo Dr. Marcos Antonio Machado, dentro da colaboração no Projeto Instituto Nacional de Ciência e Tecnologia de genômica e melhoramento dos citros (INCT Citros).

5.1 Modelagem da PtTLP de *Poncirus trifoliata*

Para construção do modelo tridimensional da PR-5 PtTLP foi adotada a técnica da modelagem por homologia, que baseia-se na similaridade sequencial e estrutural entre proteínas homólogas e com isto utiliza uma ou mais estruturas de proteínas já elucidadas como referência para determinar a estrutura 3-D de uma determinada sequencia-alvo primária (GIBAS; JAMBERCK, 2001).

A sequência primária da PR-5 foi submetida a um alinhamento BLASTp (ALTSCHUL et al., 1990) contra a base de dados do PDB- Protein Data Bank na busca de uma estrutura de referência (*template*) para o alinhamento estrutural. Após a análise do *output* foi escolhido como *template* uma taumatina de maçã (*Malus domestica*; PDB ID 3ZS3) por apresentar os maiores valores de identidade (41,82%), similaridade (54%) dentre as proteínas do *set* e uma ótima resolução (1.8 Å).

O modelo foi construído com auxílio do servidor automático Swiss-model (<http://swissmodel.expasy.org/interactive>) (ARNOLD et al., 2006) utilizando a ferramenta Automatedmode. A visualização do modelo foi realizada através do VMD-Visual Molecular Dynamics (HUMPHREY, DALKE, SCHULTEN, 1996).

5.1.1 Refinamento do modelo

Para otimização da geometria e das condições termodinâmicas locais e global do modelo obtido, a estrutura foi submetida à uma minimização de energia. A minimização de energia é uma técnica na qual por meio de algoritmos e métodos computacionais busca-se a conformação com menor perturbação entrópica dentre as possíveis variações conformacionais da estrutura inicialmente submetida à minimização. Para a minimização de energia do modelo proposto foi utilizado o servidor *CHIRON* (<http://troll.med.unc.edu/chiron/login.php>) (GAIA et al., 2011). O modelo final foi então validado com a utilização de ferramentas de avaliação da qualidade estrutural.

5.1.2 Validação do modelo

A estrutura tridimensional refinada teve a qualidade local dos aminoácidos analisada pelos campos de força *ANOLEA* (MELO; FEYTMANS, 1998) e *QMEAN* (BENKERT; BIASINI; SCHWEDE, 2011). A qualidade global da proteína foi determinada pelo *Dfire* (ZHOU; ZHOU, 2002) e *QMEAN6* (BENKERT et al., 2008). A qualidade estereoquímica foi acessada pelo diagrama de Ramachandran, o qual foi obtido através do *PROCHECK* (LASKOWSKI et al., 1993). Após verificada a qualidade do modelo, sua estrutura foi depositada no *PMDB-ProteinModel Data Bank* e encontra-se disponível ao público sob o código de acesso PM0079881.

5.2 Docking molecular

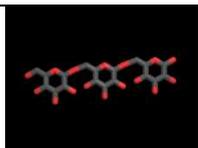
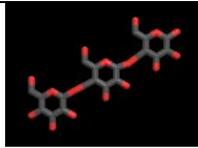
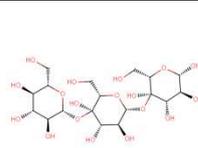
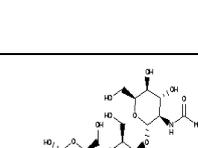
5.2.1 Preparação do receptor

O receptor (macromolécula) utilizado para as simulações de *docking* foi o modelo tridimensional da PR-5 proposto no presente trabalho. O arquivo *.pdb* foi inicialmente convertido para *.pdbqt* e parâmetros de carga *Kollman* (WEINER et al., 1984) foram assinalados, assim como hidrogênios foram adicionados à estrutura.

5.2.2 Mineração e preparação dos ligantes

Para o experimento de *docking* foram selecionados moléculas de polissacarídeos como a quitina e as glucanas e monossacarídeos como glucoses e N-acetil-glucosamina (Tabela 1). As coordenadas atômicas dos ligantes foram obtidas através de arquivos *.pdb* ou *.mol2* disponíveis nas respectivas bases de dados. Os ligantes foram assinalados com parâmetros de carga *Gasteiger* (GASTEIGER; MARSILI, 1980) e seus arquivos convertidos para *.pdbqt*.

Tabela 1. Compostos obtidos para simulação de *docking* molecular com a PR-5 de *Poncirus trifoliata*. Fonte: Dados dessa pesquisa.

Nome	Estrutura	Fórmula	Acesso	Base de dados
1,6 Beta D-Glucana		$C_{18}H_{32}O_{16}$	5409	http://molecules.gnu-darwin.org/
1,3 Beta D-Glucana		$C_{18}H_{32}O_{16}$	3775	http://molecules.gnu-darwin.org/
1,4 Beta D-Glucana		$C_{18}H_{32}O_{18}$	HMDB06944	http://www.hmdb.ca/
Quitina		$C_{24}H_{39}O_{15}N_3$	49538622	http://zinc.docking.org/

5.2.3 Simulações de *docking*

Os experimentos de *docking* foram feitos no programa Autodock1.5.6, com uma grade cúbica de energia de volume 126 x 126 x 126 Å, centrada no receptor. Foi escolhido o algoritmo genético Lamarckiano (LGA) e os parâmetros de

simulação foram de 10.000 réplicas, com análises de energia a cada 1.500.000 e 27000 gerações. O tamanho populacional foi mantido em 150 e as taxas de mutação e de crossing-over foram de 0.02 e 0.8 respectivamente. Foram geradas 10 conformações ranqueadas por energias. As conformações obtidas foram organizadas, coletadas e analisadas através da opção *Autodock tools*. A conformação de menor energia de cada complexo foi escolhida para análise no VMD.

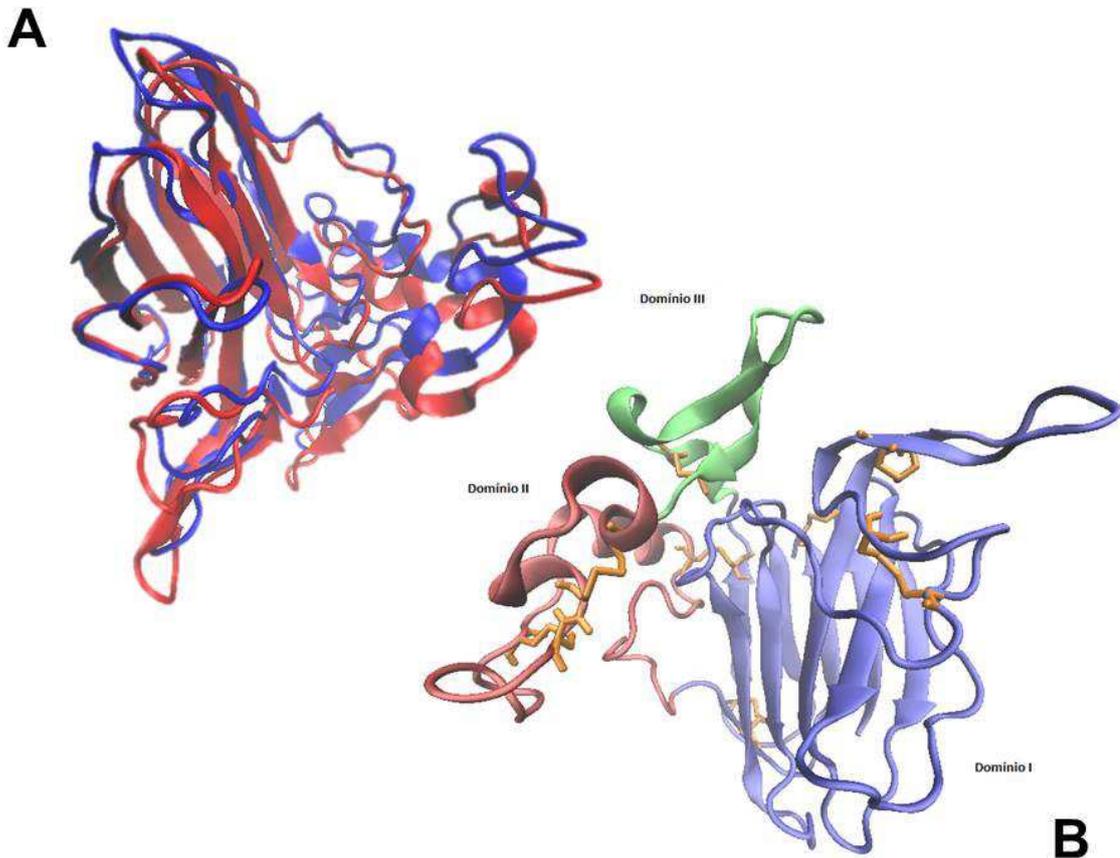
6 RESULTADOS E DISCUSSÃO

A caracterização funcional e estrutural comparativa *in silico* de uma proteína PR-5 do tipo taumatina de *Poncirus trifoliata*, denominada PtTLP, foi obtida a partir da sua sequência gênica, o gene *PtTLP*, encontrado no Banco de Dados no genoma funcional e comparativo dos Citros – CiEST (<http://www.centrodecitricultura.br/index.php>). O gene foi expresso envolvido na interação *P. trifoliata-Phytophthora parasitica*.

6.1 Construção e validação do modelo 3D da PtTLP:

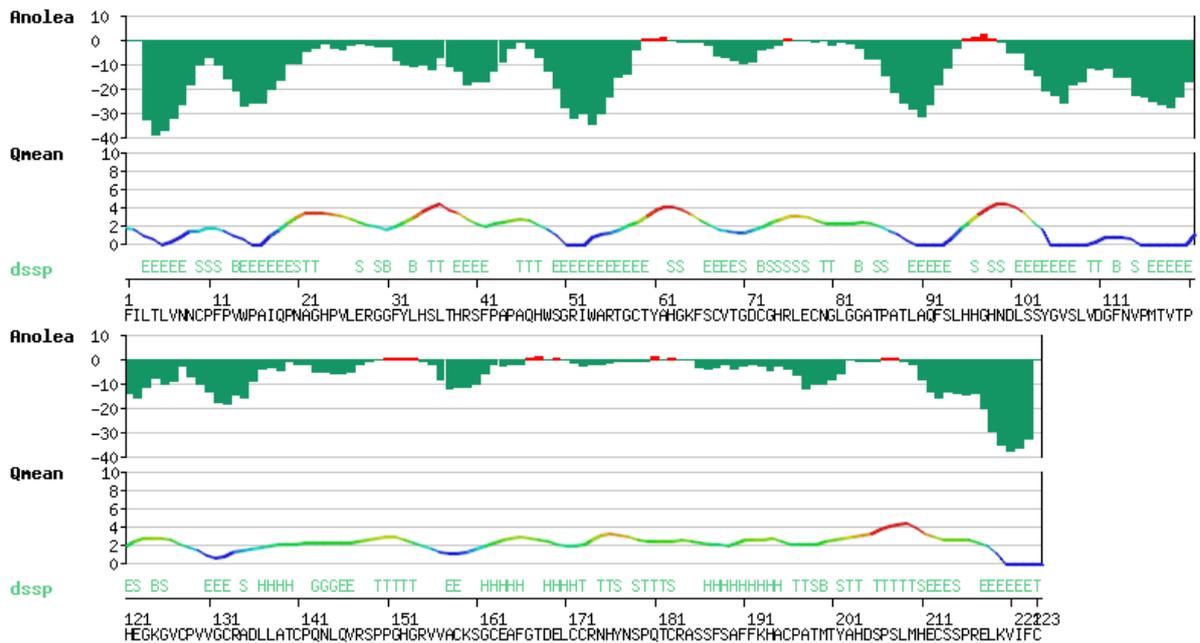
O modelo utilizado foi construído a partir do servidor Swissmodel. A estrutura 3D proposta da PtTLP evidenciou que esta proteína é constituída por três domínios (I, II e III), característica identificada entre suas proteínas homologas taumatinas (KOIWA et al., 1999). O domínio I é do tipo β -sanduiche, formado por duas folhas betas uma com quatro fitas beta e outra com cinco fitas beta, apresentando ainda quatro pontes dissulfeto e uma quinta ponte compartilhada como domínio II. O domínio II constitui-se por três alfa-hélices pequenas e algumas regiões de alça contendo ainda duas pontes dissulfeto. O domínio III apresenta duas fitas beta pequenas e duas alfa-hélices pequenas além de uma ponte dissulfeto (Figura 3B). O modelo proposto da PtTLP teve uma energia *DFire* de -295,15 (ZHOU; ZHOU, 2002) que, de acordo com Murzin et al, (1995), é considerado ótimo, pois tem uma energia total negativa. Este é um indicativo de que esta proteína necessita de pouca energia para se manter na forma ativa (Figura 3A).

Figura 3 A- Sobreposição estrutural do modelo da PR-5 de *Poncirus trifoliata* obtido através da técnica de modelagem comparativa (azul) e seu respectivo *template* (vermelho, PDB ID: 3ZS3). B- Representação em faixas (Ribbon) obtida com o programa VMD mostrando os três domínios: azul (I), vermelho (II) e verde (III), em laranja estão destacadas as oito pontes dissulfeto. Fonte: Dados dessa pesquisa.



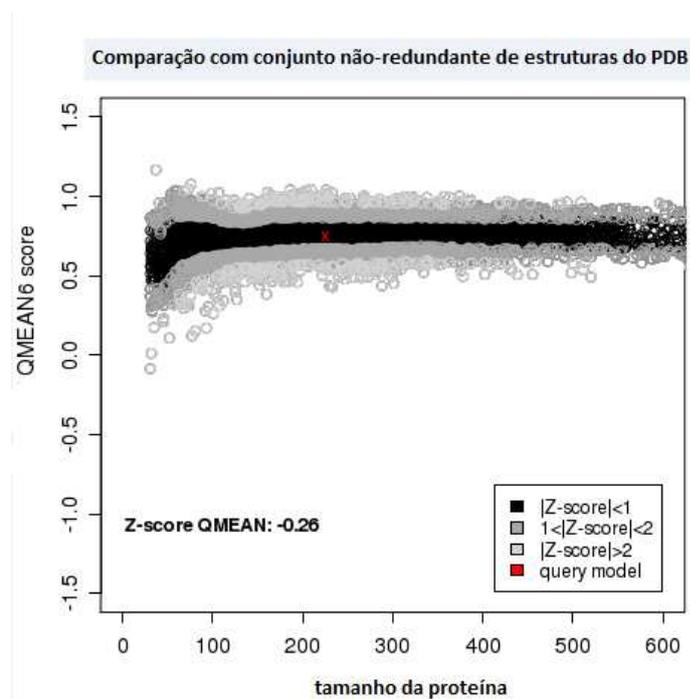
A partir dos dados obtidos do gráfico de ANOLEA (Figura 4) proveniente do modelo 3D dessa proteína revelou que as regiões onde apresentam energias positivas (desfavoráveis) são consideradas aceitáveis, por se encontrar em regiões de alça, que muitas vezes não estão diretamente relacionadas à função da proteína. O gráfico de ANOLEA mostra a qualidade do modelo de acordo com a energia gerada por cada aminoácido na proteína, energias positivas demonstram aminoácidos que estão em condições desfavoráveis de energia no sistema, que pode ser atribuído a erro ou regiões específicas de cada família de proteína (MELO; FEYTMANS, 1998).

Figura 4. Gráficos de energia local ANOLEA e QMEAN em função do teor da estrutura secundária (DSSP). Fonte: Dados dessa pesquisa



A partir dos dados do gráfico de QMEAN, obtido através das análises do modelo proposto (Figura 5) revelou que a qualidade do modelo é satisfatória, por apresentar resultados de QMEAN *score* de 0,75 e um *Z-score* -0,26, que são considerados ótimos do ponto de vista termodinâmico e de qualidade estrutural, além de apresentar valores muito próximos ou até maiores a proteínas elucidadas experimentalmente, mostrando que é um modelo confiável. Para Benkert, Schwede e Tosatto (2009), a demonstração dos dados do gráfico de QMEAN é considerado um critério importante que mostra um tipo de pontuação local para a estrutura modelada, em um *score* de 0 a 1, o modelo que obtiver QMEAN mais próximo de 1 será mais confiável, onde a estrutura terá um *Z-score* que compara as estruturas depositadas de mesmo tamanho no PDB.

Figura 5. Comparação do Qmean6 score obtido no modelo com o score de estruturas do PDB. Fonte: Dados dessa pesquisa.



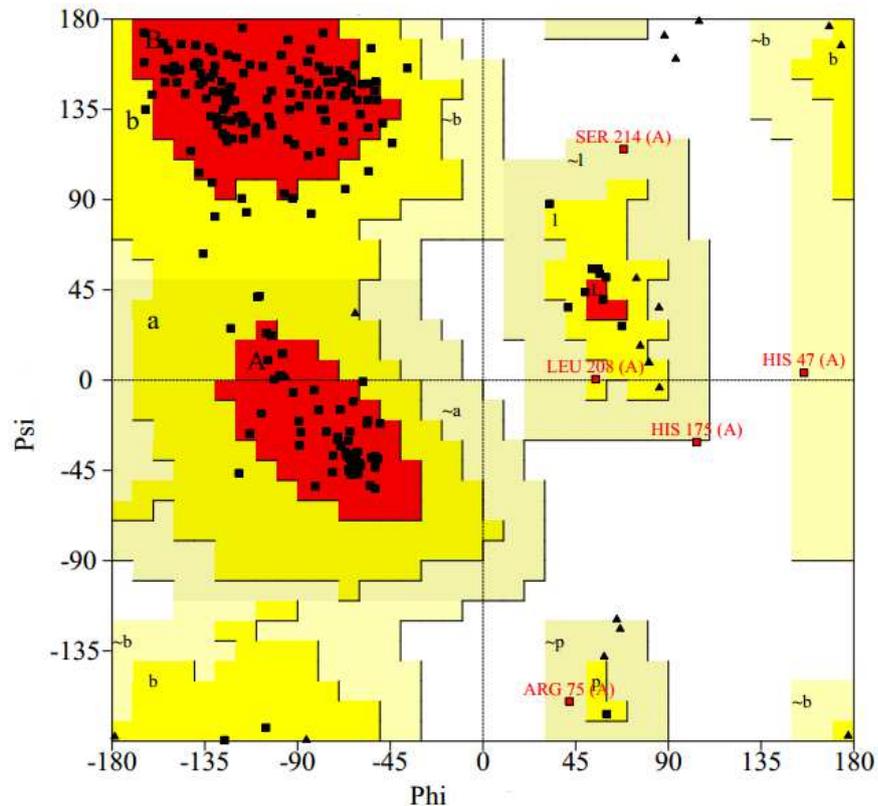
A Tabela 2 mostra os resultados do Qmean6 contendo os valores de energia para cada um dos seis parâmetros estruturais analisados, apresentando ainda um score médio. O Raw score corresponde ao score bruto calculado e o Z-score é a comparação com uma população amostral.

Tabela 2. Resultado *Qmean* para o modelo teórico. Fonte: Dados dessa pesquisa.

Score	Raw score	Z-score
Energia de interação C_beta	-91.0	-0.61
Energia de todos os pares de átomos	-4467.97	-1.31
Energia de solvatação	-5.60	-2.31
Energia das torsões angulares	-66.20	-0.19
Conformação de Estrutura secundária	79.5%	0.53
Acessibilidade ao solvente	78.1%	0.31
QMEAN6 score	0.747	0.26

Os dados gerados a partir do modelo revelaram através do gráfico de Ramachandran, que 82% dos resíduos não apresentaram impedimento na sua conformação estereoquímica (Figura 6).

Figura 6. Gráfico de Ramachandran com 82% dos resíduos em regiões de conformação mais favorável (vermelho), 15,3 % em regiões permitidas, 2,2% em regiões desfavoráveis (bege) e 0,5% em regiões proibidas (branco). Fonte: Dados dessa pesquisa.

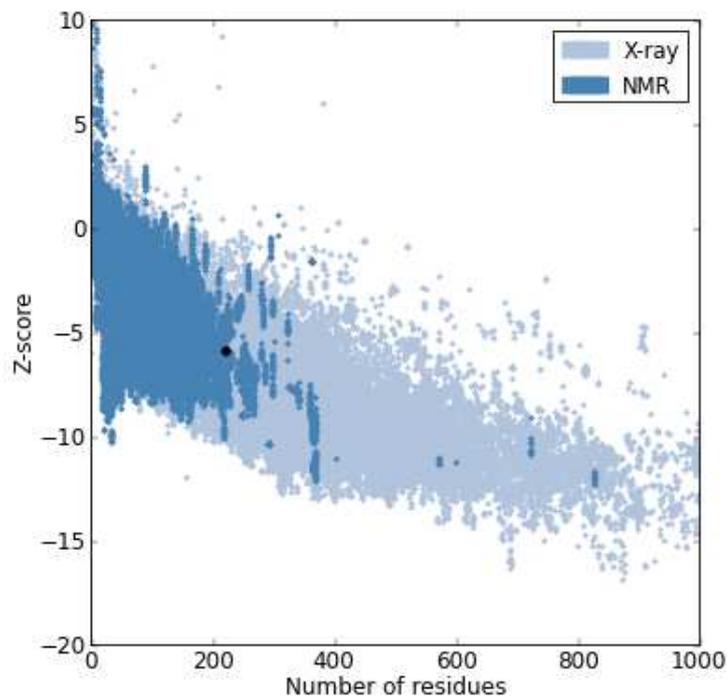


Os aminoácidos que correspondem a folhas betas e alfa-hélices estão localizados em regiões favoráveis, enquanto que alguns aminoácidos que fazem parte de alças ou *loops* estão situados em regiões menos favoráveis e permitidas, com exceção de um resíduo da histidina (His-175) que se encontra em uma região proibida. Com base nos critérios de Laskowski et al. (1993), uma pequena porcentagem de resíduos fora é considerado aceitável. Alguns resíduos de glicina também se encontram em regiões não permitidas, entretanto por ser um aminoácido de cadeia lateral simples não provoca impedimento estereoquímico que possa atrapalhar a conformação da proteína. O gráfico de Ramachandran representa a

conformação do esqueleto da cadeia polipeptídica e é descrito pela enumeração dos pares de valores Φ e Ψ para cada resíduo (LASKOWSKI et al., 1993).

A partir do gráfico com o resultado desse servidor ProSA, foi possível demonstrar que essa TLP tem padrões ótimos de aceitação comparada a modelos experimentais (Figura 7). O servidor ProSA (<https://prosa.services.came.sbg.ac.at/prosa.php>) (WIEDERSTEIN; SIPPL, 2007) calcula escores de qualidade locais (para cada aminoácido) e geral (da proteína inteira), os quais são plotados em gráficos: os valores locais fazem o gráfico de padrões de energia, e o de valor geral é mostrado comparando com outras estruturas depositadas no PDB (Z-score). Essa comparação é feita com proteínas cujas estruturas foram experimentalmente elucidadas por cristalografia de raio x e ressonância magnética.

Figura 7 Gráfico do resultado do servido ProSA para validação e confiabilidade do modelo. Fonte: Dados dessa pesquisa.

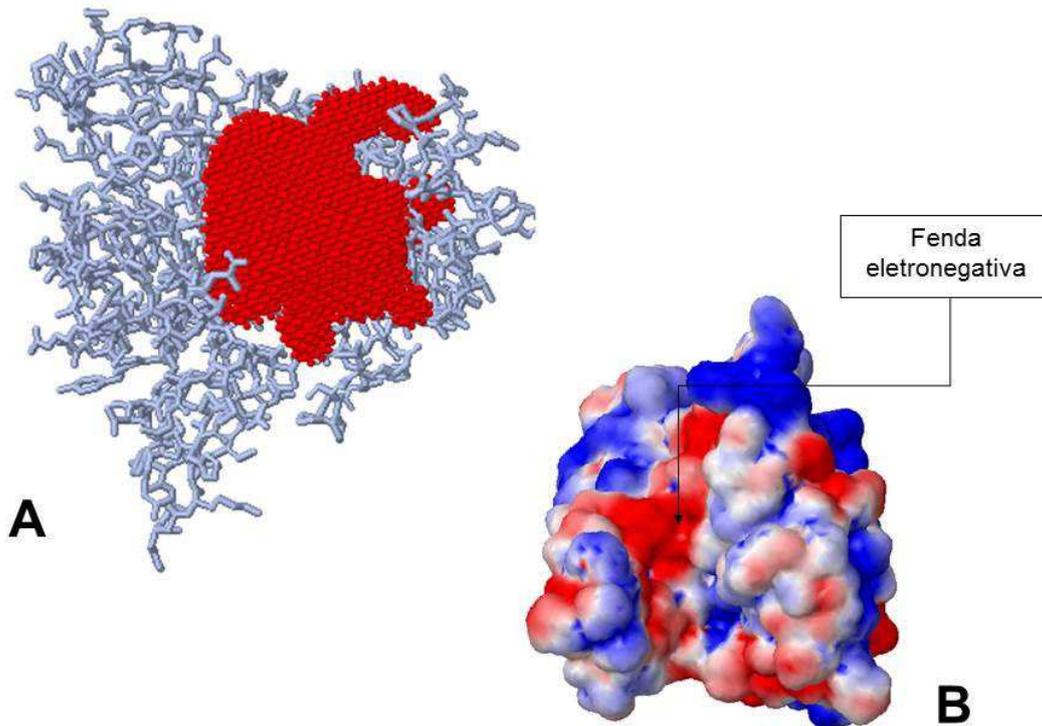


De acordo com os resultados obtidos, conclui-se que o modelo é válido e confiável, mostrando que está dentro dos padrões e pode ser considerado um modelo que tem uma representação muito próxima da estrutura real da proteína.

6.2 Predição do sítio ativo da PtTLP

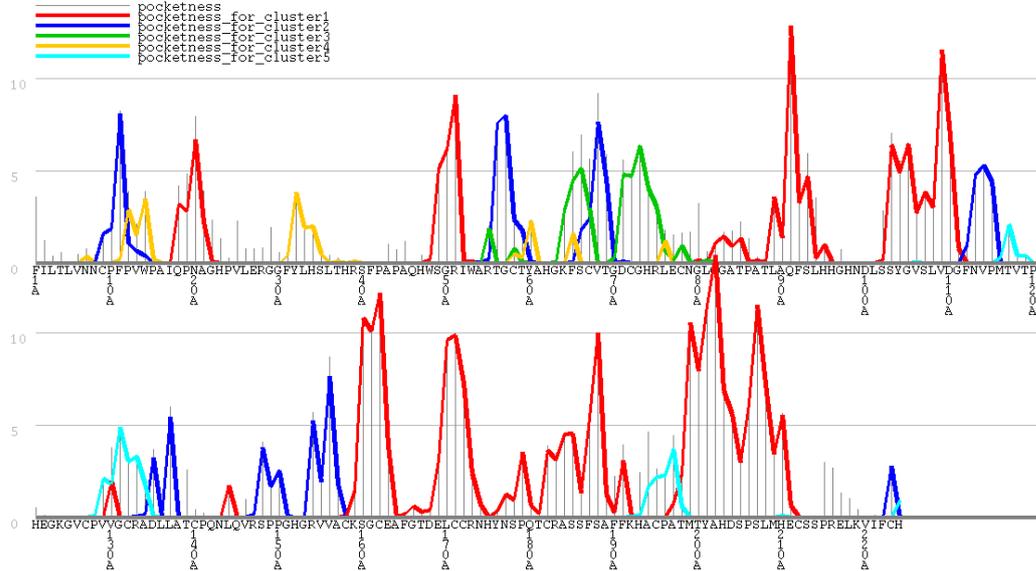
Na Figura 8A observa-se a cavidade (possível sítio de ligação) obtida pelo melhor cluster do ranking. Uma fenda interdomínio, provável candidata para ligação com glucanas proposta por Nair et al. (2010), também foi encontrada na proteína PtTLP. Esta fenda é uma das principais características da maioria das proteínas PR-5 (NAIR et al., 2010). A figura 8-B representa o mapa de superfície do potencial eletrostático da PtTLP, destaca que a proeminente fenda interdomínio possui um caráter predominantemente eletronegativo, fato esse observado em outros estudos envolvendo estruturas tridimensionais de proteínas TLPs com potencial antifúngico (KOIWA et al., 1999; GHOSH; CHAKARABARTI, 2008; THOMPSON et al., 2006). Tem sido proposto que essa fenda exerce uma importante contribuição no potencial antifúngico e na atividade de beta-glucanase para proteínas da família das TLPs (RECEVEUR-BERE'CHOT, 2006).

Figura 8. A- Cavidade do *cluster 1* do Ghecom, um potencial sítio de ligação. **B-** Representação do mapa de superfície do potencial eletrostático da PtTLP. Em *Azul*, as regiões eletropositivas, e em *vermelho* as regiões eletronegativas. Observa-se um *cleft* eletronegativo das TLPs, característica supostamente associada à atividade antifúngica desta família proteica. Fonte: Dados dessa pesquisa.



Na Figura 9, observa-se a relação de cada *cluster* com os aminoácidos da proteína. A predição do sítio ativo do modelo da PtTLP foi feita através do servidor Ghecom (<http://strcomp.protein.osaka-u.ac.jp/ghecom/>) (KAWABATA, 2010), que procura regiões na proteína com potencial para receber um ligante, analisando sua morfologia e composição.

Figura 9. Gráfico mostrando as regiões da proteína que participam de cada cluster. Em *vermelho*, os resíduos que participam do melhor cluster do ranking. Fonte: Dados dessa pesquisa.



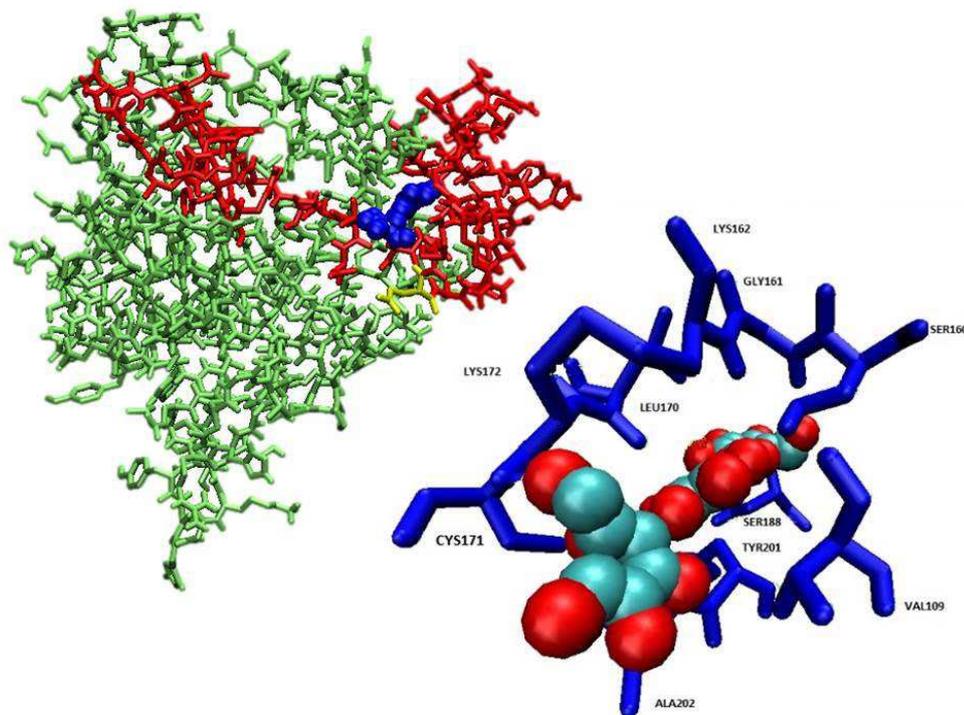
6.3 Caracterização *in silico* da atividade antifúngica por ligação a glucanas

Evidência *in silico* de atividade de Beta-glucanase para a PtTLP foi demonstrada por *docking* molecular para 3 glucanas: 1,3 beta D-glucana ($C_{18}H_{32}O_{16}$), 1,4 beta D-glucana ($C_{18}H_{32}O_{16}$) e 1,6 beta D-glucana ($C_{18}H_{32}O_{18}$), confirmada pelas energias livres de ligação -10.91, -6.28 e -11.81(Kcal/mol), respectivamente entre os compostos e a PtTLP, obtidas nos melhores complexos (Figuras 10, 11 e 12). O modelo de *docking* molecular produzido pelo programa Autodock entre a proteína PtTLP e a molécula 1,3- β -glucana apresentou uma forte interação entre as moléculas. Dez aminoácidos (Ala202, Cys171, Lys172, Leu170, Lys162, Gly161, Ser160, Val109, Ser188, Tyr201) estão envolvidos na ligação com a molécula ancorada, dentro de um raio de 4 Å de distância (Figura 10).

A interação identificada no modelo entre a 1,3 beta glucana e a fenda interdomínio da proteína é uma das principais características encontradas em proteínas com atividade antifúngica e beta-glucanase (NAIR et al., 2010), o que pode sugerir que a proteína do presente estudo desempenha as referidas funções biológicas. A energia necessária para que a ligação ocorra é baixa, o que pode evidenciar que essa interação é favorável (SOUSA, 2011). Como a 1,3 beta glucana

é um dos principais constituintes da parede celular dos oomicetos (SILVA et al., 2006), com essa interação obtida, pode-se especular o possível envolvimento dessa proteína na atividade antioomicetos.

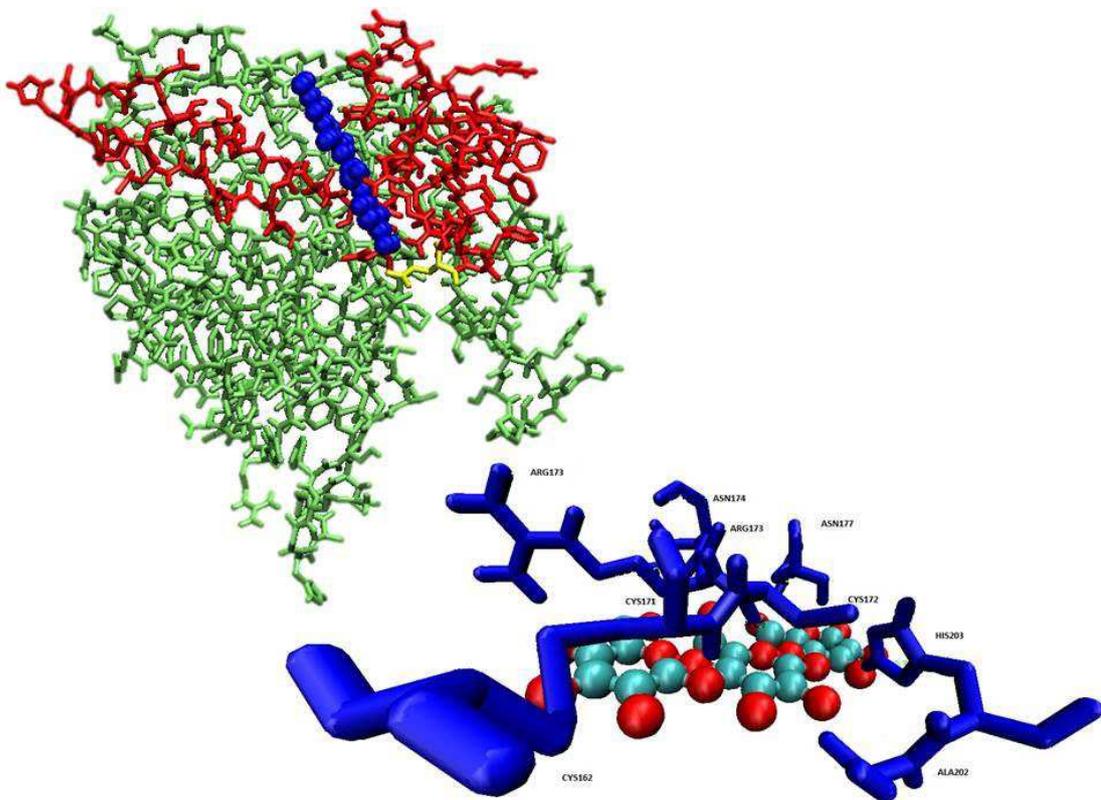
Figura 10. Melhor complexo *TLP*-1,3- β -D-Glucana do ranking baseado em energia de ligação. Em azul a 1,3- β -D-Glucana, em *vermelho* resíduos que compõem o pocket do sítio catalítico de acordo com Ghecom. Em *amarelo*, está destacado o ácido glutâmico, que apresenta um papel relevante na ação beta-glucanase das TLPs. Fonte: Dados dessa pesquisa.



A Figura 11 representa o modelo de *docking* molecular entre a proteína PtTLP e a 1,4- β -glucana. Em contraste com a ligação com a anterior, nesta, nove aminoácidos (Cys162, Cys171, Arg173, Asn174, Arg173, Asn177, Cys172, His203, Ala202) estão envolvidos na ligação com a molécula ancorada com uma distância de 4Å. Como a interação do *docking* anterior, a 1,4- β -glucana interage com a fenda eletrostática da proteína, entretanto essa interação ocorre em uma região diferente da anterior como pode ser observado nos aminoácidos descritos acima. A energia necessária de ligação para esse ancoramento foi considerada baixa entretanto maior

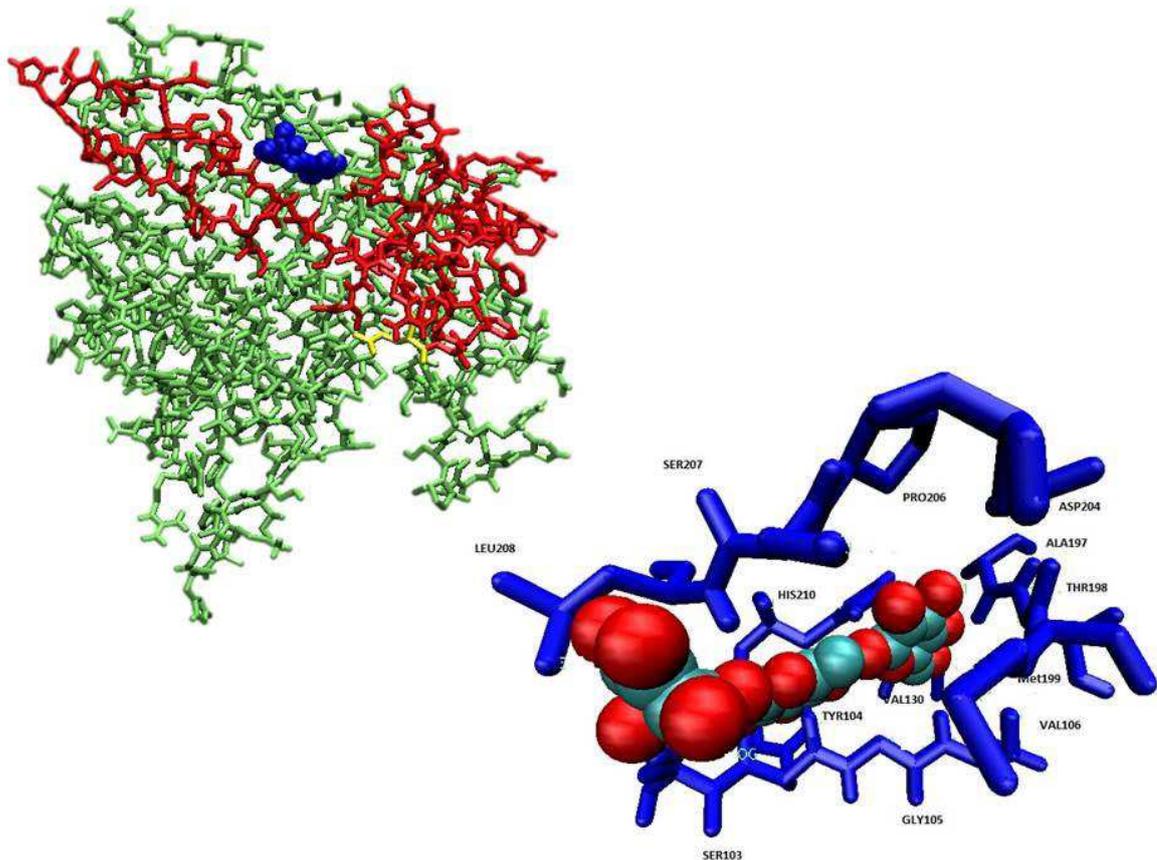
quando comparada com o complexo formado pela PtTLP e a 1,3- β -glucana , o que pode evidenciar que essa interação é favorável (SOUSA, 2011), mas não tanto quanto o complexo anterior. Essa molécula de 1,4- β -glucana está presente em poucos fungos (SILVA et al., 2006) sendo, por essa razão, pouco estudada. O motivo dessa interação não está claro ainda na literatura, necessitando, portanto, de mais estudos a cerca dessa interação, pois o resultado obtido nesse trabalho evidencia uma possível atividade antifúngica dessa proteína para fungos e/ou microrganismos mais específicos e com parede celular composta por 1,4- β -glucana .

Figura 11. Melhor complexo *TLP*-1,4- β -D-Glucana do ranking baseado em energia de ligação. Em *azul* a 1,4- β -D-Glucana, em *vermelho* resíduos que compõem o pocket do sítio catalítico de acordo com Ghecom. Em *amarelo*, está destacado o ácido glutâmico, que apresenta um papel relevante na ação beta-glucanase das TLPs. Fonte: Dados dessa pesquisa.



A representação do *docking* molecular entre a proteína PtTLP e a 1,6- β -glucana demonstrou que treze aminoácidos (Leu208, Ser207, His210, Tyr104, Ser103, Pro206, Val130, Gly105, Val106, Met199, Thr198, Ala197, Asp204) estão envolvidos nessa ligação a um raio de distância de 4Å (Figura 12). A energia de ligação mensurada para essa interação foi a mais negativa das três. Demonstrando assim, que existe uma maior interação entre a PtTLP e a 1,6- β -glucana que os demais monômeros de glucana descritos acima, pode-se observar também que as três moléculas ligadas a PtTLP interagem em locais diferentes da fenda interdomínio. A atividade glucanase é definida na literatura como requerendo uma distância de até 5Å entre os chamados dois aminoácidos catalíticos, Glu-Asp (ZECHEL; WITHERS, 2001). Como na interação acima descrita a distância entre o par catalítico foi de menos de 3 angstroms pode-se especular que a proteína PtTLP possivelmente possui atividade de 1,6- β -glucana se.

Figura 12. Melhor complexo *TLP-1,6-β-D-Glucana* do ranking baseado em energia de ligação. Em *azul* a 1,6-β-D-Glucana, em *vermelho* resíduos que compõem o pocket do sítio catalítico de acordo com Ghecom. Em *amarelo*, está destacado o ácido glutâmico, que apresenta um papel relevante na ação beta-glucanase das TLPs. Fonte: Dados dessa pesquisa.

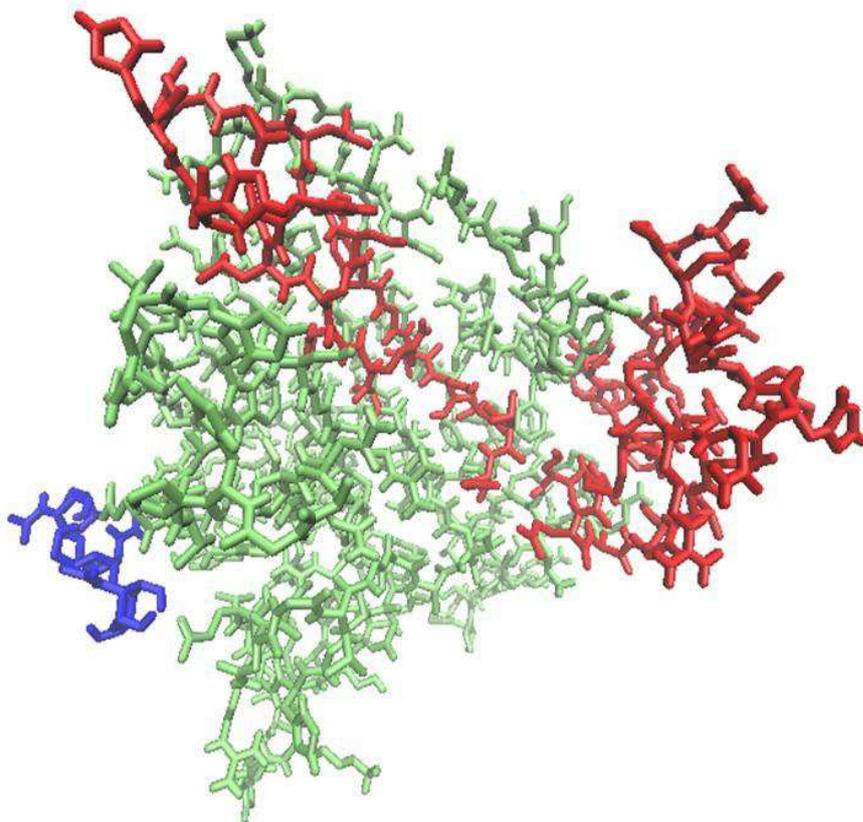


Uma vez que, tanto a localização quanto os aminoácidos envolvidos nas ligações de 1,3-β-glucana e 1,6-β-glucana são diferentes, é possível que, devido a proximidade entre eles, a PtTLP possa também clivar 1-3; 1,6-β-glucana ramificada, a qual é encontrada em *cross link* observado em parede celular de oomicetos (CLAVAUD, AIMANIANDA, LATGE, 2009) demonstrando assim sua potencial atividade antifúngica e antioomiceto. O oomiceto fitopatogênico *Phytophthora parasitica*, por exemplo, apresenta quantidades consideráveis de 1-3; 1,6-β-glucana ramificada (BLASCHEK et al., 1987), o que leva a especulações a cerca da possível atividade antioomiceto da PtTLP, atuando na interação com esta molécula específica encontrada na parede celular dessa espécie.

6.4 Atividade de ligação a carboidratos solúveis

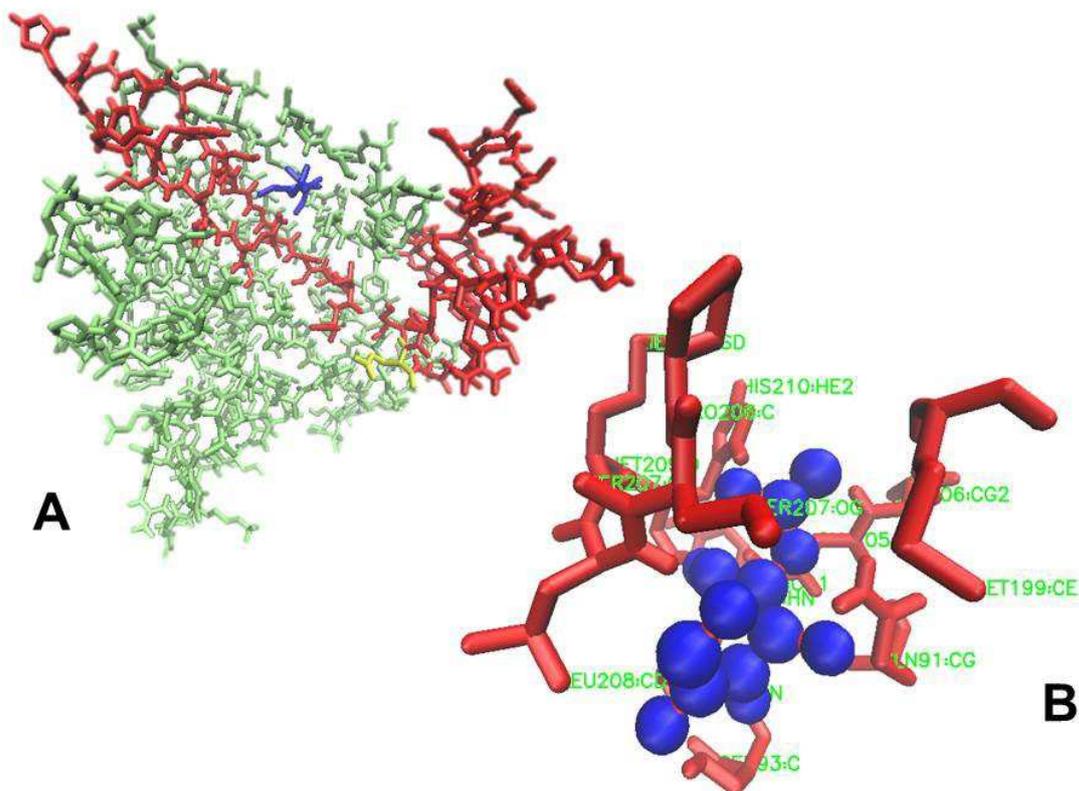
Evidência *in silico* da atividade de quitinase não foi encontrada por meio da interação molecular da PtTLP e quitina ($C_{24}H_{39}O_{15}N_3$) (Figura 13). A análise da interação revelou que a ligação a quitina é fraca, por meio de uma energia de ligação positiva, fora do sítio ativo da proteína.

Figura 13. Melhor complexo *TLP*-quitina do ranking baseado em energia de ligação. Em *azul* a quitina, em *vermelho* resíduos que compõem o pocket do sítio catalítico de acordo com Ghecom. Na figura, observa-se que a quitina não interage com o sítio ativo da proteína. Fonte: Dados dessa pesquisa.



Entretanto, curiosamente, a ligação de PtTLP com um monômero de quitina, a N-Acetil-glucosamina – NAG ($C_8H_{15}NO_6$), um monossacarídeo derivado da glicose foi estável (Figura 14) que obteve uma energia de ligação de -5,32 (Kcal/mol), este complexo demonstrou a interação com 16 aminoácidos do sítio ativo da proteína.

Figura 14 A- Complexo PtTLP-N-acetil-glucosamina. Em *vermelho* os resíduos que fazem parte do sítio ativo apontado no *cluster 1* pelo Ghecom. Em *azul*, o N-acetil-glucosamina. Em *amarelo*, o GLU163. **B-** Aminoácidos presentes na interação a uma distância de 4Å. Fonte: Dados dessa pesquisa.



Evidência *in silico* de atividade de ligação entre monossacarídeos e a PtTLP foi demonstrada por *Docking* Molecular para 2 monossacarídeos: 1,3-β-D-Glucose ($C_6H_{12}O_6$) e 1,4-β-D-Glucose ($C_6H_{12}O_6$), confirmada pelas energias livres de ligação -5,17 e -5,27 (Kcal/mol), respectivamente entre os compostos e a PtTLP, obtidas nos melhores complexos (Figuras 15 e 16). O modelo de *docking* molecular entre a proteína PtTLP e a molécula de 1,3-β-D-Glucose apresentou ligação estável, com onze aminoácidos (Arg183, Ser185, Leu170, Glu169, Lys162, Val109, Ser160,

7 CONCLUSÕES

A estrutura proposta para a PtTLP apresenta um modelo estável e representativo, dentro dos padrões avaliados, o que pode ser considerado uma representação muito próxima da estrutura real da proteína. A sua estrutura obtida por modelagem comparativa evidencia características tipicamente relatadas na literatura a proteínas TLPs, tais como a existência de três domínios, oito pontes de sulfeto e ainda a presença de uma fenda interdomínios carregada eletronegativamente, denominada sítio ativo da proteína. O modelo construído encontra-se disponível para acesso na base de dados PMDB – Protein Model Data Bank (<https://bioinformatics.cineca.it/PMDB/main.php>), sob o código de acesso PM0079881.

PtTLP interage com as moléculas de 1,3- β -glucana , 1,4- β -glucana e 1,6- β -glucana , N-acetil-glucosamina, 1,3- β -D-glucose e 1,4- β -D-glucose, com interações comprovadas por energias negativas, pela orientação das estruturas na região que comporta o sítio ativo da proteína, ao contrário da interação com a molécula de quitina, que apresenta energia de ligação positiva e ancora fora do sítio ativo da proteína.

Portanto, estes dados dão evidências para possíveis atividades de β -glucanase e de ligação a monômeros de carboidratos solúveis para a PtTLP, as quais precisam ser experimentalmente comprovadas, visto que, esta proteína é forte candidata como antifúngico e antioomiceto, para aplicações biotecnológicas na agricultura, potencialmente na citricultura, na indústria farmacêutica e alimentícia. Deste modo, os resultados aqui obtidos abrem um leque de questões a serem investigadas acerca desta proteína e suas aplicações.

REFERÊNCIAS

- ACHARYA, K., PAL, A. K., GULATI, A., KUMAR, S., SINGH, A. K., AHUJA, P. S. Overexpression of Camellia sinensis thaumatin-like protein, CsTLP in potato confers enhanced resistance to Macrophomina phaseolina and Phytophthora infestans infection. **Mol Biotechnol**, v. 54, n.2, 2013.
- ALMEIDA, J. R.; SILVA, C. H. T. P. **Estudos de modelagem molecular e relação estrutura-atividade da acetilcolinesterase e inibidores em Mal de Alzheimer**. Dissertação (Mestre em Ciências) Universidade de São Paulo, 2011.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. **J. Mol. Biol.** V.215, p.403-410, 1990.
- ARAUJO, F. N. ; FERREIRA, L. T. M. F. ; CORREIA, A. L. G. ; KISHI, L. T. ; MACHADO, M. A. ; CAMPOS, M. A. Molecular *in silico* Characterization of PR-5 genes expressed in *Poncirus trifoliata*. In: Simpósio de Biotecnologia 'Desafios, Aplicações e Inovações', 2013, João Pessoa. **Anais do I Simpósio de Biotecnologia**. João Pessoa PB, 2013.
- ARNOLD K.; BORDOLI L.; KOPP, J.; SCHWEDE, T. A SWISS-MODEL Workspace: Um ambiente baseado na web para a modelagem da estrutura de proteínas de homologia. **Bioinformática**, v.22, p.195-201, 2006.
- BATISTA, A. B.; VASCONCELOS, I. M. **Caracterização estrutural da MoCBP3, uma albumina 2S de sementes de Moringa oleífera Lamarck e seu modo de ação contra fungos fitopatogênicos**. Tese (Doutor em Bioquímica). Universidade Federal do Ceará, 2013.
- BECKER, O. **Computational Biochemistry and Biophysics**. New York, p.512, 2001.
- BENKERT, P.; BIASINI, M.; SCHWEDE, T. Para a estimativa da qualidade absoluta de modelos de estrutura de proteínas individuais. **Bioinformática**, v.27, n.3, p.343-50, 2011.
- BENKERT, P.; SCHWEDE, T.; TOSATTO, S.C.E. QMEANclust: Estimation of protein model quality by combining a composite scoring function with structural density information. **BMC Struct Biol**, v.20, p.9-35, 2009.
- BENKERT, P.; TOSATTO, S.C.E.; SCHOMBURG, D. QMEAN: A comprehensive scoring function for model quality assessment. **Proteins: Structure, Function, and Bioinformatics**, v.71, n.1, p.261-277, 2008.
- BIASINI, M.; BIENERT, S.; WATERHOUSE, A.; ARNOLD, K.; STUDER, G.; TOBIAS S.; FLORIAN K.; CASSARINO, T. G.; BERTONI, M.; BORDOLI, L.; SCHWEDE, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. **Nucleic Acids Research**, v, 42, 2014.

- BLASCHEK, W.; SCHLITZ, M.; KRAUS, J.; FRANZ, G. In vitro production of specific polysaccharides: isolation and structure of an antitumor active β -glucan from *Phytophthora parasitica*. **Food Hydrocolloids**, v.1, n.5-6, p.371-380, 1987.
- BOWLES, D.J. Defense related proteins in higher plants. **Biochemistry**, v.59, p.837-907, 1990.
- BUJNICKI, J. M. Protein structure prediction by recombination of fragments *Chembiochem*. **European Journal of Chemical Biology**, v.7, n.1, p.19–27, 2006.
- CAPRILES, P. V. S. Z.; TREVIZANI, R.; ROCHA, G. K.; DARDENNE, L. E.; CUSTÓDIO, F. L. Modelos Tridimensionais. In: VERLI, H. **Bioinformática da Biologia à Flexibilidade Molecular**. Organização de Hugo Verli. Porto Alegre, 2014.
- CARDOSO, M. B.; DE SOUZA, O. N. **Uma Proposta para a Predição Computacional da Estrutura Terciária de Polipeptídeos**. Dissertação (Mestrado em Ciência da Computação). Pontifícia Universidade Católica do Rio Grande do Sul. Porto Alegre, 2007.
- CASTLE, W. S.; TUCKER, D. P. H.; KREZDORN, A. H.; YOUTSEY, C. O. Rootstock selection: the first step to success. Rootstocks for Florida citrus. **Gainesville: University of Florida**, p.47, 1989.
- Centro de Citricultura Sylvio Moreira <<http://www.centrodecitricultura.br/index.php>> acessado em 20 de janeiro de 2015.
- CHIRON : RAMACHANDRAN, S.; KOTA, P.; DING, F.; DOKHOLYAN, N. **V.PROTEINS: Structure, Function and Bioinformatics**, v.79, p. 261-270, 2011.
- CORNELISSEN, B. J.; HOOFT VAN HUIJSDUIJNEN, R. A.; BOL, J. F. A tobacco mosaic virus-induced tobacco protein is homologous to the sweet-tasting protein thaumatin. **Nature**, v.321, p.531-532, 1986.
- DALL'AGNO, K. C. M. **Um estudo sobre a predição da estrutura 3D aproximada de proteínas utilizando o método CReF com refinamento**. Dissertação (Mestrado em Ciência da computação). Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2012.
- DIAS, S. F.; TAVEIRA, C. C. Homologia de proteínas como ferramenta na construção de novos fármacos. **Cenarium Farmacêutico**, Ano 4, n.4, 2011.
- DORN, M. **Uma proposta para a predição computacional da estrutura 3D aproximada de polipeptídeos com redução do espaço conformacional utilizando análise de intervalos**. Dissertação (Mestre em Ciência da Computação). Pontifícia Universidade Católica do Rio Grande do Sul, 2008.
- DORN, M.; SOUZA, O. N. **Uma proposta para a predição computacional da estrutura 3D aproximada de polipeptídeos com redução do espaço conformacional utilizando análises de intervalos**. Dissertação (Mestre em Ciência da Computação). Pontifícia Universidade Católica do Rio Grande do Sul, 2008.

ERICKSON, J.A.; JALAIE, M.; ROBERTSON, D.H.; LEWIS, R.A.; VIETH, M. Lessons in Molecular Recognition: The Effects of Ligand and Protein Flexibility on Molecular Docking Accuracy. **J. Med. Chem.**, v.47, p.45-55, 2004.

GAIA: KOTA, P.; DING, F.; RAMACHANDRAN, S.; DOKHOLYAN, N. V. **Bioinformatics**, 2011.

GARCIA-CASADO, G.; COLLADA, C.; ALLONA, I.; SOTO, A.; CASADO, R.; RODRIGUEZ-CEREZO, E.; GOMEZ, L.; ARAGONCILLO, C., 2000: Characterization of an apoplastic basic thaumatinlike protein from recalcitrant chestnut seeds. **Physiol. Plant.** V.110, p.172–180.

GASTEIGER, J.; MARSILI, M. Iterative partial equalization of orbital electronegativity rapid access to atomic charges. **Tetrahedron**, v. 36, p. 3219-3228, 1980.

CLAVAUD C., AIMANIANDA V., LATGE J. P. Organization of fungal, oomycete and lichen (1,3)- β -glucans. In: Chemistry, biochemistry, and biology of 1,3- β -glucans and related polysaccharides. **Academic Press**, New York, p.387-424, 2009.

GHOSH, R.; CHAKRABARTI, C. Crystal structure analysis of NP24-I: a thaumatin-like protein. **Planta**, n. 228 p.883–890, 2008.

GIBAS, C.; JAMBECK, P. *Desenvolvimento bioinformática: ferramentas de software para aplicação em biologia*. **Ed. Campus**, Rio de Janeiro, 2001.

GIBAS, C.; JAMBECK, P. *Developing Bioinformatics Computer Skills*. **O'Reilly**, 2001.

GUEDES, I. A.; MAGALHÃES, C. S.; DARDENNE, L. E. Atracamento Molecular. In: VERLI, H. **Bioinformática da Biologia à Flexibilidade Molecular**. Organização de Hugo Verli. Porto Alegre , 2014.

HUMPHREY, W., DALKE, A., SCHULTEN, K. VMD: dinâmica molecular visuais. **J Mol Graph.** v. 14, n.1, p. 33-8, 27-8, 1996

KAWABATA, T. Detection of multi-scale pockets on protein surfaces using mathematical morphology. **Proteins**, v.78, p.1195-1121, 2010.

KAWABATA, T.; GO, N. Detection of pockets on protein surfaces using small and large probe spheres to find putative ligand binding sites. **Proteins**, v.68 p.516-529, 2007.

KOIWA, H., KATO, H., NAKATSU, T., ODA, J., I., YAMADA, Y. & SATO, F. Crystal structure of tobacco PR-5d protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins. **J. Mol. Biol.**, v. 286, p.1137-1145, 1999.

KRIEGER, E.; NABUURS, S.; VRIEND, G. B. Homology Modeling. **Structural Bioinformatics**, 2003.

LASKOWSKI, R. A.; MACARTHUR M. W.; MOSS, D.; THORNTON, J. M. PROCHECK: um programa para verificar a qualidade estereoquímica de estruturas de proteínas. **J. Appl. Cryst**, v.26, p.283-291, 1993.

LIU, J. J.; STURROCK, R.; EKRAMODDOULLAH, A. K. M. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. **Plant Cell Rep.** v. 29, n.5, p.419-436, 2010.

MELO, F.; FEYTMANS, E. Assessing protein structures with a non-local atomic interaction energy. **J Mol Biol**, v.277, n.5, p.1141-1152, 1998.

MENETTI, F. S. **Modelagem da interação estrutural entre os motivos IQ da miosina Va com calmodulina.** Trabalho de Conclusão de Curso (Bacharelado em Informática Biomédica) Universidade de São Paulo, 2006.

MENU-BOUAOUICHE, L.; VRIET, C.; PEUMANS, W.J.; BARRE, A.; VAN DAMME, E.J.; ROUGÉ, P. A molecular basis for the endo-beta 1,3-glucanase activity of the thaumatin-like proteins from edible fruits, **Biochimie**.v.85, n.1-2, p.123-31, 2003.

Ministério da Agricultura <www.agricultura.gov.br/vegetal/culturas/citrus> acessado em 20 de janeiro de 2015.

MURZIN, A. G.; BRENNER, S. E.; HUBBARD, T.; CHOTHIA, C., SCOP: a structural classification of proteins database for the investigation of sequences and structures, **Journal of Molecular Biology**, v. 247, pp. 536-540, 1995.

NAIR, R. A.; KIRAN, A. G.; SIVAKUMAR, K. C.; THOMAS, G. Molecular Characterization of an Oomycete-Responsive PR-5 Protein Gene from Zingiber zerumbet, **Plant Mol Biol Rep**, v. 28 p.128–135, 2010.

NASSER, W.; DE TAPIA, M.; BURKARD, G. Maize pathogenesis-related proteins: characterization and cellular distribution of -1,3glucanases and chitinases induced by brome mosaic virus infection or mercuric chloride treatment. **Physiological and Molecular Plant Pathology**, v.36, p.1-14, 1990.

NICOLAU JUNIOR, N. **Diferenças estruturais e “docking” receptor-ligante da proteína E7 do vírus do papiloma humano (HPV) de alto e baixo risco para câncer cervical.** Tese (Doutorado em Ciências) Universidade de São Paulo, 2013.

OSGUTHORPE, D. J., Ab initio protein folding. **Current Opinion in Structural Biology**, v.10, p.146-152, 2000.

OSMOND, I. W.; HRMOVA, M.; FONTAINE, F.; IMBERTY, A.; FINCHER, G. B. Binding interactions between barley thaumatin-like proteins and (1,3)- β -D-glucans: Kinetics, specificity, structural analysis and biological implications. **Eur. J. Biochem**, v. 268 p. 4190-4199, 2001.

PACHECO, A. G. M. **Modelagem molecular comparativa e estudos de acoplamento molecular da enzima lanosterol 14 α -desmetilase do *Moniliophthora perniciosa*.** Dissertação (Mestrado em Biotecnologia). Universidade Estadual de Feira de Santana, 2009.

PASSOS, O. S.; PEIXOUTO, L. S.; SANTOS, L. C.; CALDAS, R. CO.; SOARES FILHO, W. S. Caracterização de híbridos de *Poncirus trifoliatae* de outros porta-enxertos de citros no Estado da Bahia. **Rev. Bras. Frutic**, v.28, n.3, 2006.

PATEL, S.S., VAIDYA, M.B., SHAH, D.B. Homology Modelling of Conserved rbcl Amino Acid Sequences in Leguminosae Family. **J Data Mining GenomicsProteomics**, v.5, n.2, 2014.

PETERSEN, T. N.; BRUNAK, S.; VON HEIJNE G.; NIELSEN H. Signal P 4.0: discriminar péptidos de sinal a partir de regiões transmembranas. **MétodosNatureza**, v. 8, p. 785-786, 2011.

PRASATH, D.; BALAGOPAL, A.; MAHANHESH, V.; ROSANA, O. B.; JAYASANKAR, S.; ANANDARAJ, M. Comparative study of pathogenesis-related protein 5 (PR5) of different Zingibearaceae species. **Indian Journal of Biotechnology**, v.13, p178-185, 2014.

RECEVEUR-BRE´CHOT, V.; BARRE, A.; ROUSSEL, A.; PEUMANS, W. J.; DAMME, E. J. M. V.; ROUGE, P. Crystal Structure at 1.45-Å Resolution of the Major Allergen Endo-β-1,3-Glucanase of Banana as a Molecular Basis for the Latex-Fruit Syndrome. **PROTEINS: Structure, Function, and Bioinformatics**, v.63, p.235–242, 2006.

REIS, M. S.; TAKITA, M. A.; PALMIERI, D. A.; MACHADO, M. A. Bioinformatics for the Citrus EST Project (CitEST). **Genet. Mol. Biol.** v.30, n.3, São Paulo, 2007.

ROGNAN, D. *Docking* methods for virtual screening: Principles and recent advances. **Virtual Screengin**, p.153-176, 2011.

SCHERER, N. M. **Evolução molecular darwiniana nas proteínas relacionadas à patogênese (PRs) em plantas**. Dissertação (Mestrado em Genética e Biologia Molecular). Universidade Federal do Rio Grande do Sul, 2002.

SILVA, M. L. C.; MARTINEZ, P, F.; IZELI, N. L.; SILVA, I. R.; VASCONCELOS, A. F. D.; Cardoso, S. Caracterização química de glucanas fúngicas e suas aplicações biotecnológicas. **Quim. Nova**, v. 29, n, 1, p.85-92, 2006.

SINGH, N. K.; KUMAR, K. R. R.; KUMAR, D.; SHUKLA, P.; KIRTI, P. B. (2013) Characterization of a Pathogen Induced Thaumatin-Like Protein Gene *AdTLP* from *Arachisdiogoi*, a Wild Peanut. **PLoS ONE**, v. 8, n. 12, 2013.

SIPPL, M. J. O reconhecimento de erros em estruturas tridimensionais de proteínas. **As proteínas**, v. 17, p. 355-362, 1993.

SOUSA, A. M. T. **Modelagem molecular por homologia e estudos de docking molecular de potenciais alvos terapêuticos para doença de Chagas**. Tese (Doutorado em Química), Universidade Federal de Rio de Janeiro, 2011.

SOUSA, D. T. **Análise da expressão dos genes relacionados à parede celular e à membrana plasmática de Saccharomycescerevisiae em resposta a alta pressão hidrostática**. Dissertação (Mestrado em Biotecnologia). Universidade Federal do Espírito Santo, 2012.

SOUSA, M. L. **Modelagem por homologia de estruturas protéicas do veneno de Apis melífera**. Trabalho de Conclusão de Curso (Bacharelado em Informática Biomédica). Universidade de São Paulo, 2008.

STANGARLIN, J. R.; KUHN, O. J.; TOLEDO, M. V.; PORTZ, R. L.; SCHWAN-ESTRADA, K. R. F.; PASCHOLATI, S. F. A defesa vegetal contra fitopatógenos. **Scientia Agraria Paranaensis**, v.10, n. 1, p 18-46, 2011.

TAHEREH, D.; ALI, N.; ESMAEIL, E. Identification and expression analysis of TLPs as candidate genes promoting the responses to both biotic and abiotic stresses in wheat. **Plant Omics Journal**, v. 6, n.2, p.107-115, 2013.

TEXEIRA, J. E. C. **Genes de defesa de *Citrus sunki* e *Poncirus trifoliata*: expressão constitutiva e induzida por *Phytophthora parasítica***. Dissertação (Mestrado em Genética e Melhoramento de Plantas). Universidade Federal de Lavras, 2005.

TEXEIRA, J. E. C. **Genes de defesa de *Citrus sunki* e *Poncirus trifoliata*: expressão constitutiva e induzida por *Phytophthora parasítica***. Dissertação (Mestrado em Genética e Melhoramento de Plantas). Universidade Federal de Lavras, 2005.

THOMPSON, C. E.; FERNANDES, C. L.; SOUZA, O. N.; SALZANO, F. M.; BONATTO, S. L.; FREITAS, L. B. Molecular Modeling of Pathogenesis-Related Proteins of Family 5. **Cell Biochemistry and Biophysics**, v. 44, 2006.

TRAMONTANO, A.; LESK, A. M. Protein structure prediction. **John Wiley and Sons, Inc.** Weinheim, Germany, 2006.

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

VELAZHAHAN, R., DATTA, K. S., MUTHUKRISHNAN, S. The PR-5 family: thaumatin like protein in plants. In.: DATTA SK, MUTHUKRISHNAN S, EDITORS. Pathogenesis-related proteins in plants. **CRC press Boca Raton**. p. 107-129, 1999.

WEBER, R. L. **Transformação genética de soja [*Glycinemax* (L.) MERRILL] com um gene que codifica uma osmotina de *Solanum nigrum* var. *americanum*, visando a resistência a moléstias fúngicas**. Dissertação (Mestrado em Genética e Biologia Molecular). Universidade Federal do rio Grande do Sul, 2007.

WEINER, S. J.; KOLLMAN, P. A.; NGUYEN, D. T.; CASE, D. A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. **J. Comp. Chem.** v.7 p.230-252 1986.

WIEDERSTEIN; SIPPL. Prosa-web:..service web interativo para o reconhecimento de erros em estruturas tridimensionais de proteínas. **Nucleic Acids Research**, v.35, 2007.

ZECHEL D.L, WITHERS S.G. Dissection of nucleophilic and acid– base catalysis in glycosidases. **Curr Opin Chem Biol**, v. 5, p.643–649, 2001.

ZHOU, H.; ZHOU, Y. Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. **Protein Sci**, v.11, p.2714-2726, 2002.